

UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**REDISEÑO DEL PROCESO DE ASIGNACIÓN DEL PROGRAMA
RESIDENCIA FAMILIAR ESTUDIANTIL ADMINISTRADA POR
JUNAEB**

PROYECTO DE GRADO PARA OPTAR AL GRADO DE MAGÍSTER EN
INGENIERÍA DE NEGOCIOS CON TECNOLOGÍAS DE INFORMACIÓN

NATASHA STEFANÍA FALCONÍ ROMERO

PROFESOR GUÍA:

EDUARDO CONTRERAS VILLABLANCA

MIEMBROS DE LA COMISION:

CLAUDIO JIMÉNEZ CARVAJAL

LUCIANO VILLARROEL PARRA

SANTIAGO DE CHILE

2020

RESUMEN EJECUTIVO

El presente trabajo describe una propuesta de rediseño del proceso de asignación del Programa de Residencia Familiar Estudiantil (PRFE) de la Junta Nacional de Auxilio Escolar y Becas (JUNAEB) en el contexto del Magister de Ingeniería de Negocios con Tecnologías de la Información (MBE) de la Universidad de Chile.

La propuesta busca exponer los procesos involucrados para la asignación de este beneficio, cuyo propósito es maximizar la correcta asignación planteando un modelo de focalización simplificado ajustado a los requisitos vigentes del programa.

Por lo que el proyecto propone el rediseño del proceso de asignación del PRFE cambiando el mecanismo mediante el cual se otorga el beneficio a los estudiantes que postulan o renuevan a este beneficio. Este rediseño está basado en la metodología de Ingeniería de Negocios a partir de Patrones de Proceso de Negocio que presenta desde la arquitectura de JUNAEB hasta el diseño de diagramas de procesos en notación BPMN (Business Process Management Notation) con sus respectivas lógicas de negocio. Estas lógicas de negocio están relacionadas con la utilización de técnicas de minerías de datos y de la actualización del subproceso de planificar la entrega del programa, en el que se pretende que los resultados obtenidos con la predicción de los modelos se utilice el perfil idóneo del estudiante potencial beneficiario. Igualmente, se propone la construcción de una aplicación web que pueda soportar esta solución y la aplicación de un plan detallado de gestión del cambio que permita una adecuada y posible implementación de este proyecto.

Asimismo, luego de una evaluación retrospectiva del proyecto, mediante el uso de datos históricos del programa, se logra evidenciar que la aplicación de éste arroja un 71,60% de la asignación exacta a la que se realizó en ese entonces. Además, la evaluación económica del proyecto, tal como se plantea en este trabajo, genera un ahorro en el tiempo de ejecución de las actividades del proceso rediseñado, lo que indica que el proyecto es rentable si se implementa bajo las características mencionadas en el presente trabajo.

TABLA DE CONTENIDO

RESUMEN EJECUTIVO	i
CAPÍTULO 1: INTRODUCCIÓN Y CONTEXTO.....	9
1.1 Antecedentes de la Industria.....	9
1.2 Descripción General de la Empresa	9
1.3 Problema u Oportunidad Identificada.....	10
1.4 Objetivos y Resultados Esperados del Proyecto	10
1.4.1 Objetivo General	10
1.4.2 Objetivos Específicos	11
1.4.3 Resultados Esperados	11
1.5 Alcance	11
1.6 Riesgos Potenciales.....	12
CAPÍTULO 2: MARCO TEÓRICO	13
2.1 Metodología de Ingeniería de Negocios	13
2.1.1 Enfoque por procesos con apoyo tecnológico	13
2.1.2 Ingeniería de Negocios	13
2.1.3 Posicionamiento Estratégico – Modelo Delta (HAX)	14
2.1.4 Business Process Model and Notation (BPMN).....	14
2.2 Lógica de Negocios.....	15
2.2.1 Pobreza Multidimensional	15
2.2.2 Registro Social de Hogares.....	18
2.2.3 Índice de Aislamiento Territorial	20
2.2.4 Programa de Residencia Familiar Estudiantil (PRFE)	21
2.2.5 Investigación de Operaciones y modelos de optimización.....	24
2.2.6 Programación Lineal con variables enteras	25
2.2.7 Distancia Euclidiana	26
2.2.8 Data Mining	26
2.2.9 Modelos Predictivos	27
2.2.9.1 Validación cruzada.....	28
2.2.10 Árboles de Decisión	28
2.2.10.1 Boosting	29
2.2.11 Random Forest.....	29
2.2.12 Regresión Logística.....	30
2.2.13 K nearest Neighbors.....	31
2.2.13.1 Support Vector Machine (SVM)	32
CAPÍTULO 3: PLANTEAMIENTO ESTRATÉGICO Y MODELO DE NEGOCIOS	34
3.1 Posicionamiento Estratégico.....	34
3.2 Balanced Scorecard.....	34
3.3 Modelo de Negocios	35
CAPÍTULO 4: ANÁLISIS SITUACIÓN ACTUAL.....	38
4.1 Arquitectura de Procesos.....	38

4.2	Modelamiento Detallado de Procesos	39
4.2.1	Modelamiento IDEF0.....	39
4.3	Diagnóstico de la Situación Actual.....	42
4.4	Cuantificación del Problema u Oportunidad	45
CAPÍTULO 5: PROPUESTA DE REDISEÑO DE PROCESOS		50
5.1	Direcciones de Cambio y Alcance	50
5.2	Diseño Detallado de Procesos TO BE	51
5.2.1	Asignación de Becas o Asistencialidad Estudiantil	51
5.2.1.1	Desarrollo de Modelos predictivos de Becas o Asistencialidad Estudiantil	52
5.2.1.2	Planificar la entrega de Beca o Asistencialidad Estudiantil.....	53
5.2.1.3	Mantención de Estado	56
5.3	Diseño de Lógica de Negocios	56
5.3.1	Análisis de los Datos	57
5.3.2	Preparación de los Datos	62
5.3.3	Modelamiento.....	64
5.3.3.1	Construcción de Modelos.....	64
5.3.4	Aplicando el modelo de focalización simplificado	73
5.3.5	Evaluación del Modelo	77
5.3.6	Construcción del Perfil del Beneficiario del PRFE	78
CAPÍTULO 6: PROPUESTA DE APOYO TECNOLÓGICO		80
6.1	Especificación de Requerimientos	80
6.1.1	Requerimientos Funcionales.....	80
6.1.2	Requerimientos No Funcionales	81
6.2	Arquitectura Tecnológica	81
6.3	Diseño de la Aplicación.....	82
6.3.1	Casos de Uso.....	82
6.3.2	Diagramas de Secuencia	83
6.3.3	Diagramas de Clases	85
CAPÍTULO 7: GESTIÓN DEL CAMBIO		86
7.1	Contexto de la Organización.....	86
7.2	Observación de la implementación a realizar	87
7.3	Análisis de los principios de diseño	88
7.4	Caracterización del cambio.....	89
7.4.1	Organización y estructura del cambio	90
7.4.2	Público Objetivo del cambio	91
7.4.3	Evaluación del proceso de cambio.....	91
7.5	Factores críticos de éxito	93
7.5.1	Gestión Emocional	94
7.5.2	Gestión de las Habilidades.....	94
7.5.3	Gestión del Poder.....	95
7.6	Plan de gestión del cambio	95
CAPÍTULO 8: EVALUACIÓN DEL PROYECTO		98

8.1	Análisis Retrospectivo.....	98
8.1.1	Presentación de datos.....	98
8.1.2	Resultados Obtenidos	100
8.1.2.1	Aplicación de los modelos predictivos con el Dataset PRFE 2017..	100
8.1.2.2	Comparación de los resultados de asignación oficial año 2017 versus la predicción.....	101
8.2	Definición de Beneficios Y Costos	101
8.2.1	Ahorros del Proyecto.....	101
8.2.2	Costos del Proyecto	102
8.3	Flujo de Caja.....	103
8.4	Análisis de Sensibilidad	105
CAPÍTULO 9: CONCLUSIONES.....		106
9.1	Trabajo realizado	106
9.2	Recomendaciones para trabajos futuros	108
CAPÍTULO 10: BIBLIOGRAFÍA.....		109
CAPÍTULO 11: ANEXOS.....		112
11.1	Tabla de Índice de Aislamiento por Región	112
11.2	Proceso de Postulación y Renovación al Programa Residencia Familiar Estudiantil año 2018	128
11.2.1	Requisitos para Postulantes.....	128
11.2.1.1	Documentación exigida para Postulantes.....	128
11.2.2	Requisitos para Renovantes	129
11.2.2.1	Documentación exigida para Renovantes	129
11.2.3	Causales de Término del Beneficio.....	129
11.3	Jupiter Book	130
11.3.1	Selección de variables	130
11.3.2	Separación de Datos.....	131
11.3.3	Modelo 1: Regresión Logística.....	132
11.3.4	Modelo 2: K-NN.....	135
11.3.5	Modelado 3: Árboles de Decisión.....	136
11.3.6	Modelo 4: Random Forest.....	138
11.3.7	Gradient Boosting.....	140
11.3.8	Comparación de Resultados	141
11.4	Presupuesto y Cobertura PRFE 2018.....	142
11.5	Cálculos Flujo de Caja del Proyecto	143
11.5.1	Determinación Ahorros.....	143
11.5.2	Cálculo Valor Residual	144

ÍNDICE DE TABLAS

Tabla 1: Riesgos potenciales del proyecto MBE.....	12
Tabla 2: Contribución relativa de dimensiones a la pobreza multidimensional.....	18
Tabla 3: Calificación Socioeconómica por tramo – RSH	19
Tabla 4: Modelo de Focalización PRFE 2018.....	23
Tabla 5: Cuadro de Mando Integral de Becas – JUNAE B	35
Tabla 6: Comparación Dimensiones y Variables RSH – Modelo de Focalización PRFE .	43
Tabla 7: Promedio de Notas año anterior por Becado.....	45
Tabla 8: Último curso aprobado por la madre	46
Tabla 9: Último curso aprobado por el padre.....	46
Tabla 10: Pertenencia a pueblos originarios.....	46
Tabla 11: Desplazamiento desde el domicilio particular al lugar de estudio.....	47
Tabla 12: Medio de Transporte utilizado para desplazarse	47
Tabla 13: Costo de movilización	47
Tabla 14: Chile Solidario	48
Tabla 15: SINA E	48
Tabla 16: Becados según tramo RSH.....	48
Tabla 17: Actividades del proceso de asignación	49
Tabla 18: Variables de rediseño de la propuesta.....	50
Tabla 19: Modelo de focalización PRFE simplificado	56
Tabla 20: Descripción del Dataset PRFE.....	57
Tabla 21: Variables Categóricas	63
Tabla 22: Variables Numéricas	63
Tabla 23: Matriz de confusión – Regresión Logística	64
Tabla 24: Indicadores de Rendimiento – Regresión Logística.....	65
Tabla 25: Matriz de confusión – K-NN	65
Tabla 26: Indicadores de Rendimiento – K-NN.....	66
Tabla 27: Matriz de confusión – Árboles de Decisión	67
Tabla 28: Indicadores de Rendimiento – Árboles de Decisión	67
Tabla 29: Matriz de confusión – Árboles de Decisión (cv=3).....	68
Tabla 30: Tabla de importancia de variables – Árboles de Decisión	69
Tabla 31: Matriz de confusión – Random Forest	70
Tabla 32: Indicadores de Rendimiento – Random Forest.....	70
Tabla 33: Tabla de importancia de variables – Random Forest	71
Tabla 34: Matriz de confusión – Gradient Boosting	72
Tabla 35: Indicadores de Rendimiento – Gradient Boosting.....	72
Tabla 36: Tabla de importancia de variables – Gradient Boosting	72
Tabla 37: Variables Categóricas – Modelo de Focalización simplificado	73
Tabla 38: Variables Numéricas – Modelo de Focalización simplificado	73
Tabla 39: Matriz de confusión – Árboles de Decisión (cv=3).....	74
Tabla 40: Tabla de importancia de variables – Árboles de Decisión	75
Tabla 41: Tabla de importancia de variables – Random Forest	76
Tabla 42: Tabla de importancia de variables – Gradient Boosting	76
Tabla 43: Comparación de Modelos – Exactitud de Validación.....	77

Tabla 44: Comparación de Modelos – Exactitud de Validación (con modelo de focalización simplificado)	78
Tabla 45: Variables del perfil idóneo PRFE	78
Tabla 46: Variables del perfil idóneo PRFE – Modelo de focalización simplificado	79
Tabla 47: Perfiles de usuario de la aplicación.....	80
Tabla 48: Organización y estructura del proyecto de cambio	90
Tabla 49: Prácticas del cambio	92
Tabla 50: Panel de Control del proceso de cambio	93
Tabla 51: Estructura Etapa de Preparación del Plan de Gestión del Cambio	95
Tabla 52: Estructura Etapa de Relación con las personas del Plan de Gestión del Cambio.....	96
Tabla 53: Estructura Etapa de Instalación del cambio del Plan de Gestión del Cambio ..	96
Tabla 54: Estructura Etapa de Evaluación y Cierre del Plan de Gestión del Cambio.....	97
Tabla 55: Campos del Dataset PRFE 2017	99
Tabla 56: Estado de estudiantes del PRFE en el año 2017	99
Tabla 57: Estado de estudiantes del PRFE en el año 2017	100
Tabla 58: Estado de estudiantes del PRFE en el año 2017	100
Tabla 59: Tabla comparitva resultados de asignación PRFE 2017	101
Tabla 60: Tiempo de respuesta para cada actividad si se realiza el proyecto.....	102
Tabla 61: Valorización \$/HH	102
Tabla 62: Costos estimados del proyecto	103
Tabla 63: Tasas sobre el riesgo asociado al proyecto.....	103
Tabla 64: Flujo de Caja del Proyecto	104
Tabla 65: Indicadores de Rentabilidad.....	104
Tabla 66: Variaciones VAN	105
Tabla 67: Variaciones TIR.....	105
Tabla 68: Presupuesto Asignado y Cobertura por Región Programa Residencia Familiar Estudiantil.....	142
Tabla 69: Presupuesto Asignado y Cobertura por Región Programa Residencia Familiar Estudiantil.....	143
Tabla 70: Tiempo en horas en que se realizan cada actividad del proceso de asignación con y sin proyecto	143
Tabla 71: Cálculo de Valor Presente	144

ÍNDICE DE ILUSTRACIONES

Ilustración 1: 5 Dimensiones 2017	16
Ilustración 2: Contribución relativa de dimensiones a la pobreza multidimensional	17
Ilustración 3: Incidencia de la pobreza multidimensional en los hogares por región	17
Ilustración 4: Cobertura del RSH, según región del país	20
Ilustración 5: Elementos para la determinación del componente Aislamiento Estructural	20
Ilustración 6: Elementos para la determinación del componente Aislamiento Estructural	21
Ilustración 7: Distancia Euclidiana entre dos puntos.....	26
Ilustración 8: Pasos del proceso de Descubrimiento de Conocimiento	27
Ilustración 9: Proceso de Validación cruzada	28
Ilustración 10: Ejemplo Gráfico del Árbol de Decisión de una Beca	29
Ilustración 11: Clasificación de un nuevo vector de datos (antes de aplicar el modelo KNN)	31
Ilustración 12: Ejemplo de aplicación modelo KNN seleccionando k = 5 vecinos más cercano al nuevo vector de datos.	32
Ilustración 13: Superficie obtenida aplicando el método SVM identificando el hiper-plano	33
Ilustración 14: Modelo Delta-Hax JUNAEB	34
Ilustración 15: CANVAS de Becas – JUNAEB	37
Ilustración 16: Macro 0 de la JUNAEB	38
Ilustración 17: Macro 1.1 Cadena de Valor: Entrega de Becas, Beneficios y Asistencialidad Estudiantil.....	39
Ilustración 18: Macro 1.1.2: Gestión de Asignación y Pago de Becas, Beneficios y Asistencialidad Estudiantil.....	40
Ilustración 19: Macro 1.1.2.3 – Decisión de la Asignación para pago de Becas o Asistencialidad Estudiantil.....	41
Ilustración 20: Macro 1.1.2.3.1 – Asignación de la Beca o Asistencialidad Estudiantil.....	51
Ilustración 21: Modelo BPMN – Desarrollo de modelos predictivos de Becas o Asistencialidad Estudiantil.....	53
Ilustración 22: Macro 1.1.2.3.1.2 – Planificar la entrega de Beca o Asistencialidad Estudiantil	54
Ilustración 23: Modelo BPMN – Predicción de Beneficiarios	54
Ilustración 24: Modelo BPMN – Validación de postulantes y renovante.....	55
Ilustración 25: Modelo BPMN – Planificación de operación del programa	55
Ilustración 26: Nivel educacional por estudiante	58
Ilustración 27: Pueblo originario declarado por el estudiante	58
Ilustración 28: Último grado académico cursado por la madre del estudiante.....	59
Ilustración 29: Último grado académico cursado por el padre del estudiante.....	59
Ilustración 30: Región domicilio familiar	60
Ilustración 31: Ubicación de la vivienda con respecto al lugar de estudios	60
Ilustración 32: Tramo RSH por estudiante	61
Ilustración 33: Índice de aislamiento territorial	62
Ilustración 34: Curva ROC – Regresión Logística	65
Ilustración 35: Curva ROC – K-NN	66
Ilustración 36: Curva ROC – Árboles de Decisión	68

Ilustración 37: Diagrama de Árbol de Decisión	69
Ilustración 38: Curva ROC – Random Forest	71
Ilustración 39: Arquitectura Tecnológica aplicación web	82
Ilustración 40: Caso de Uso - Asignación de potenciales beneficiarios.....	83
Ilustración 41: Diagrama de Secuencia - Asignación de potenciales beneficiarios	84
Ilustración 42: Diagrama de Clases - Asignación de potenciales beneficiarios	85
Ilustración 43: Organigrama JUNAEB.....	86
Ilustración 44: Organigrama Departamento de Becas de JUNAEB.....	87

CAPÍTULO 1: INTRODUCCIÓN Y CONTEXTO

1.1 Antecedentes de la Industria

En el factor político existe un creciente consenso en que el estado debe aumentar la cobertura de apoyo a distintos sectores sociales vulnerables en el ámbito socioeconómico. Esto implica que los organismos públicos deben asumir nuevas tareas o ampliar la oferta de servicios sociales que apuntan a al sector más vulnerable del país.

Actualmente, Chile entrega recursos estatales destinados a velar por los niños, niñas y jóvenes chilenos en condición de vulnerabilidad biopsicosocial, para que ingresen, permanezcan y tengan éxito en el Sistema Educativo. Estos recursos se transforman en Programas Sociales, los cuales en términos conceptuales son un conjunto de actividades necesarias, integradas y articuladas que proveen bienes y servicios, tendientes a lograr un propósito específico en una población objetivo. Por ejemplo, en la educación superior se creó el beneficio de la gratuidad, que actualmente cubre hasta el sexto decil de la población más vulnerable del país.

Asimismo, se ampliaron los criterios de elegibilidad a otras instituciones de educación superior (universidades privadas, centro de formación técnica e institutos profesionales).

1.2 Descripción General de la Empresa

La Junta Nacional de Auxilio Escolar y Becas (JUNAEB), es un organismo de la Administración del Estado, creado en 1964 por la Ley N.º 15.720, responsable de administrar los recursos estatales destinados a velar por los niños, niñas y jóvenes chilenos en condición de vulnerabilidad biopsicosocial, para que ingresen, permanezcan y tengan éxito en el Sistema Educativo.

Los productos y/o servicios que están a disposición de la población objetivo se materializan a través de programas sociales de las áreas de Alimentación, Salud Escolar y Becas. A partir del año 2005 se logra la implementación de nuevos programas destinados a los estudiantes de educación superior.

Dentro de los Programas de Becas y Asistencialidad estudiantil se crea el Programa de Residencia Familiar Estudiantil para responder principalmente a la necesidad de garantizar la continuidad en la educación de aquellos estudiantes que, en su lugar de residencia, preferentemente de sectores rurales, cuentan con reducida oferta académica de colegios y/o instituciones de educación superior.

El propósito del programa radica en el otorgamiento de alojamiento en el hogar de una familia tutora que brinda alimentación, apoyo pedagógico y apoyo afectivo a estudiantes en condición de vulnerabilidad, preferentemente del sector rural, que disponen de una reducida oferta de colegios educacionales cercanos a sus domicilios o cuya oferta no responde a los intereses de los estudiantes, lo cual los obliga a trasladarse a otras localidades o ciudades que cuentan con colegios o instituciones de educación superior para continuar sus estudios, ayudando a disminuir la deserción escolar.

1.3 Problema u Oportunidad Identificada

JUNAEB busca lograr dinámicas coherentes y conducentes a implementación de Programas más eficientes y dotados de calidad, pero se ha identificado, dentro del Departamento de Becas de JUNAEB que existe baja confiabilidad en el proceso de asignación y pago de las becas, principalmente en el subproceso de decisión de la asignación para el pago de becas o residencias estudiantiles. Las principales causas se presentan porque actualmente existe un modelo de focalización que permite ordenar y otorgar el beneficio a estudiantes que lo solicitan, mediante ciertas dimensiones (vulnerabilidad, aislamiento geográfico y rendimiento académico), pero éste cuenta con falencias. Estas falencias se detectan especialmente en el Programa de Residencia Familiar Estudiantil que contiene variables de carácter declarativas que son subjetivas y no pueden ser verificadas, por lo que surge la necesidad de simplificarlo y utilizar fuentes de información externas que permitan integrar información de otros servicios públicos para contar con información efectiva y confiable de los estudiantes que postulan.

Asimismo, es importante identificar claramente la población objetivo que será potencialmente beneficiaria de las becas o residencias estudiantiles, y estratificar al hogar del estudiante a través de la aplicación de instrumentos de medición que se sustente en un modelo de focalización, el cual aumente la efectividad del gasto social y asigne los recursos del Estado a las personas que presentan mayor condición de vulnerabilidad.

1.4 Objetivos y Resultados Esperados del Proyecto

1.4.1 *Objetivo General*

Rediseñar el proceso de asignación del Programa de Residencia Familiar Estudiantil, simplificando el modelo de focalización vigente con el fin de maximizar la correcta asignación del beneficio para el año 2019.

1.4.2 Objetivos Específicos

1. Analizar y levantar los procesos de asignación del Programa de Residencia Familiar Estudiantil.
2. Identificar las variables del modelo de focalización vigente con menor significancia.
3. Elaborar una propuesta de simplificación del modelo de focalización del Programa de Residencia Familiar Estudiantil con el fin de asignar el beneficio a la mayor cantidad de estudiantes que lo soliciten.
4. Implementar el rediseño del proceso de Asignación del Programa de Residencia Familiar Estudiantil.
5. Evaluar los resultados obtenidos del proyecto para el Programa de Residencia Familiar Estudiantil.

1.4.3 Resultados Esperados

El desarrollo de esta tesis espera como resultado lo siguiente:

- a. Diagramas de Flujo de Proceso en notación BPM del Programa de Residencia Familiar Estudiantil.
- b. Análisis de las variables que conforman el modelo de focalización vigente.
- c. Modelo predictivo del Programa de Residencia Familiar Estudiantil para obtener una correcta asignación del beneficio utilizando el modelo de focalización vigente y simplificado.
- d. Rediseño del proceso de Gestión de Asignación del Programa de Residencia Familiar Estudiantil.
- e. Evaluación privada con proyecto y sin proyecto para determinar el beneficio que la implementación genera.

1.5 Alcance

Se analizará el Programa de Residencia Familiar Estudiantil, centrado en la simplificación del modelo de focalización existente. Asimismo, se usarán datos del año anterior con el fin de preparar dichos datos para revisar el comportamiento de los estudiantes beneficiados. Finalmente se elaborará la variable distancia simple (mediante el cálculo de la distancia euclidiana) para utilizar un modelo predictivo que permita la correcta asignación del beneficio.

No será considerado dentro del alcance la revisión ni la modificación de la normativa vigente (de acuerdo con los manuales de procedimientos de este beneficio). Tampoco, formarán parte de este trabajo las becas dirigidas a estudiantes de nivel básica, nivel educacional media ni educación superior que entregan aportes de libre disposición a

los estudiantes que lo solicitan. Por último, no formarán parte de este trabajo las becas asociadas a los programas de Alimentación Escolar y Salud del estudiante.

1.6 Riesgos Potenciales

Los riesgos identificados que pueden comprometer la implementación exitosa de este proyecto se describen a continuación:

Tabla 1: Riesgos potenciales del proyecto MBE
Fuente: Elaboración Propia

TIPO	RIESGOS	PROBABILIDAD	IMPACTO	MITIGACIÓN DEL RIESGO
POLÍTICO	Falta de compromiso del equipo directivo	85%	ALTO	Presentar el proyecto mostrando resultados concretos de mejora en los procesos de gestión que cumplan con la misión institucional
DISEÑO	Lentitud en la toma de decisiones	75%	ALTO	Acotar los plazos en la definición del plan de ejecución de actividades y establecer reuniones periódicas para mostrar estados de avance
IMPLEMENTACIÓN	Recursos limitados para la implementación (recortes presupuestarios)	45%	MEDIO	Búsqueda fondos concursables para la implementación del proyecto
IMPLEMENTACIÓN	Baja motivación hacia el cambio	45%	MEDIO	Implementación de herramientas de gestión del cambio
GESTION DEL PROYECTO	Problemas para planificar el desarrollo del proyecto	40%	MEDIO	Realizar un cronograma de actividades
TÉCNICO	La Tecnología desarrollada no es capaz de integrarse con la aplicación existente	60%	MEDIO	Realizar un análisis de tecnologías habilitantes en forma previa para determinar esa fase

CAPÍTULO 2: MARCO TEÓRICO

El presente capítulo describe el marco teórico conceptual relevante para este trabajo de tesis, el cual se ha dividido en 2 secciones que se describen a continuación:

2.1 Metodología de Ingeniería de Negocios

2.1.1 *Enfoque por procesos con apoyo tecnológico*

Con este nuevo concepto se eliminan las barreras funcionales que existen en una organización, con el fin de que exista coordinación entre áreas que se manejan de forma independiente. Además, el enfoque por procesos elimina los roles jerárquicos existentes dentro de la estructura organizacional, ya que en las áreas funcionales que son parte de un proceso se establecen interacciones entre las personas que lo ejecutan (Barros, 2011).

Asimismo, se propone además que el uso de las tecnologías de la información en para apoyar o ejecutar procesos ayuda al trabajo en equipo, debido a que permite que intercambien y compartan información evidenciándose una comunicación horizontal en vez de jerárquica (Barros, 2011).

2.1.2 *Ingeniería de Negocios*

Es la metodología que unifica el diseño del modelo y los procesos de negocio apoyándose con tecnologías de la información, la cual se conforma de las siguientes etapas:

- **Planteamiento estratégico:** Se identifican la creencias y valores utilizadas por una organización para cumplir un objetivo. Aquí también se define el posicionamiento estratégico.
- **Definición del Modelo de Negocio:** Se definen el o los productos o servicios que entregarán valor a los clientes y por los que estarán dispuestos a pagar, relacionados con el posicionamiento estratégico identificado.
- **Diseño de la Arquitectura de procesos:** De acuerdo con el modelo de negocio definido, se generarán los macroprocesos que mediante las relaciones existentes entre ellos se ejecutan para obtener un resultado.
- **Diseño detallado de procesos del negocio:** Se detallan los macroprocesos utilizando notación de modelamiento de procesos, ejemplo BPMN.
- **Diseño de aplicación TI:** Según los procesos diseñados en la etapa anterior, se determinan los apoyos tecnológicos a implementar.

- **Construcción e Implementación:** Se construyen e implementan las aplicaciones de TI que se identificaron acompañándolas de la gestión del cambio.

2.1.3 Posicionamiento Estratégico – Modelo Delta (HAX)

El modelo Delta Hax menciona que el centro de la estrategia siempre es el cliente, por lo que todos los esfuerzos deben estar enfocados en elaborar una propuesta de valor y establecer relaciones con cada uno de ellos.

De esta misma manera, el autor plantea que para las organizaciones sin fines de lucro ya no se habla de clientes sino de los stakeholders el cual abarca a todas las personas involucradas con esta organización. Se define entonces el modelo, mediante la identificación de tres posiciones estratégicas en las que la organización puede desplazarse, las cuales se detallan a continuación (Hax, 2010):

- 1) **Mejor Producto/Servicio:** relacionado con productos commodities, aunque se encuentra en el vértice inferior derecho del triángulo de Hax, a su alrededor se identifica dos posicionamientos identificados para las organizaciones sin fines de lucro, tales como:
 - a. Eficiencia Administrativa
 - b. Diferenciación
- 2) **Solución Integral al Cliente:** desarrolla una propuesta de valor para un cliente individualizado, y si existe colaboración se puede determinar que el posicionamiento se encuentra ubicado en el vértice inferior izquierdo. Asimismo, a su alrededor se detectan estos otros posicionamientos:
 - a. Atracción y desarrollo del constituyente
 - b. Transferencia de conocimiento
 - c. Alcance horizontal
- 3) **Liderazgo de Sistema:** se propone una conexión con las empresas complementarias y los posicionamientos identificados a su alrededor son:
 - a. Canales de entrega
 - b. Sistemas de soporte
 - c. Valor Intelectual

2.1.4 Business Process Model and Notation (BPMN)

Corresponde al estándar mundial definido y administrado por la OMG (Object Management Group) para el modelamiento de procesos de negocios con el fin de mostrar el flujo de trabajo y las actividades asociadas. Cada uno de estos modelos

muestra Datos, Documentos y Reglas, para brindar apoyo a las operaciones de la empresa.

Finalmente, los objetivos principales del BPMN son proporcionar una manera fácil de utilizar la notación de modelado de procesos, accesible a los usuarios; y proporcionar facilidades para traducir los modelos a una forma ejecutable (WHITE, 2009).

2.2 Lógica de Negocios

2.2.1 Pobreza Multidimensional

En Chile la pobreza se ha medido tradicionalmente a partir de los ingresos de los hogares. Una medición de pobreza por ingresos constituye una medida indirecta del bienestar que alcanzan los individuos y sus hogares por medio del consumo de bienes y servicios (Ministerio de Desarrollo Social, Subsecretaría de Evaluación Social, 2014). Bajo este enfoque tradicional se busca simplificar la identificación de quienes se consideran bajo un estándar de vida aceptable tomando en cuenta solo un indicador.

Una medida de pobreza multidimensional, por su parte, permite medir de manera directa las condiciones de vida de la población, en tanto se mide la situación de las personas y los hogares en relación con distintas dimensiones e indicadores de bienestar que se consideran socialmente relevantes (Ministerio de Desarrollo Social, Subsecretaría de Evaluación Social, 2014). De este modo, permite visibilizar carencias que afectan a las personas y sus hogares en múltiples dimensiones, más allá de su ingreso, y analizar la importancia relativa de tales carencias en las condiciones de vida de la población en situación de pobreza.¹

Para poder realizar esta medición se utiliza la Encuesta Casen, herramienta multipropósito de responsabilidad del Ministerio de Desarrollo Social, aplicada en Chile desde 1987, la cual es el principal instrumento de medición a nivel nacional para captar información en diversas dimensiones de las condiciones de vida de la población. Adicionalmente su representatividad nacional, regional y por zona geográfica permite profundizar el análisis de las diferencias territoriales que pueda evidenciar la metodología de medición de pobreza multidimensional.

La unidad de análisis para la medición de pobreza multidimensional en Chile es el hogar, explicado por las siguientes razones:

- *Existen recursos compartidos al interior del hogar, que significan que si algún miembro del hogar tiene alguna carencia (o dispone) de estos recursos o capacidades, tal situación afecta a los demás miembros del hogar;*
- *La medición de pobreza multidimensional a nivel de hogares es consistente con la metodología aplicada tradicionalmente y la nueva metodología implementada*

¹ Fuente: Documentos Institucionales de análisis internos.

por el Ministerio de Desarrollo Social para la medición de pobreza por ingresos que se realiza a nivel de hogares (y se informada como porcentaje y número de hogares y personas en situación de pobreza por ingresos);

- La medición de pobreza multidimensional a nivel del hogar es consistente con las políticas públicas de reducción de pobreza a nivel de los hogares. (Ministerio de Desarrollo Social, Subsecretaría de Evaluación Social, 2014)

En la medición de pobreza multidimensional se contemplan las siguientes dimensiones:

- Dimensión Educación:
 - Asistencia
 - Rezago Escolar
 - Escolaridad
- Dimensión Salud
 - Malnutrición de niños
 - Adscripción a Sistema de Salud
 - Atención
- Dimensión Trabajo y Seguridad Social
 - Ocupación
 - Seguridad Social
 - Jubilaciones
- Dimensión Vivienda
 - Habitabilidad
 - Entorno
 - Servicios Básicos
- Dimensión Redes y Cohesión Social
 - Apoyo y Participación
 - Trato igualitario
 - Seguridad

En la siguiente ilustración se muestra el peso relativo que tienen cada una de las dimensiones antes mencionadas para la medición de la pobreza multidimensional.



Ilustración 1: 5 Dimensiones 2017

Fuente: Ministerio de Desarrollo Social, Encuesta Casen 2017

De acuerdo con la Encuesta Casen 2017, la contribución relativa de estas 5 dimensiones a la pobreza multidimensional se presenta en la siguiente ilustración:

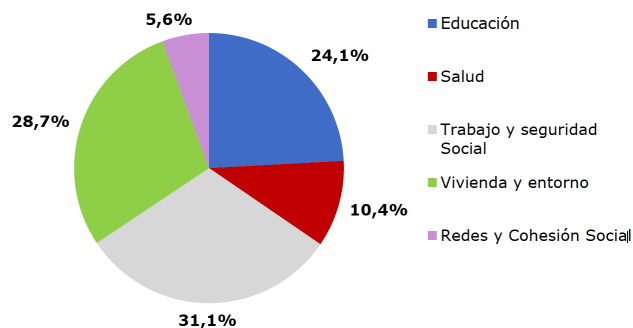
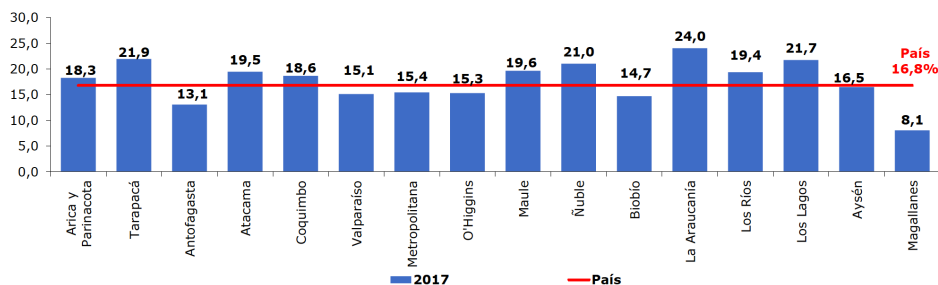


Ilustración 2: Contribución relativa de dimensiones a la pobreza multidimensional
Fuente: Ministerio de Desarrollo Social, Encuesta Casen 2017

Cabe mencionar que las dimensiones que más peso relativo tienen en esta medición son Trabajo y Seguridad Social, Vivienda y Entorno y Educación, en ese orden de importancia. Además, la Encuesta Casen 2017 mide la incidencia de la pobreza multidimensional en los hogares, por lo que la siguiente ilustración muestra esta medición por región.

(Porcentaje, hogares)



* Al 95% de confianza, las diferencias de la incidencia de pobreza multidimensional entre cada región y el total nacional NO son estadísticamente significativas para el año 2017 en las siguientes regiones: (1) Arica y Parinacota, (2) Coquimbo, (3) O'Higgins, (4) Los Ríos y (5) Aysén.
* Se excluye servicio doméstico puertas adentro y su núcleo familiar.

Ilustración 3: Incidencia de la pobreza multidimensional en los hogares por región
Fuente: Ministerio de Desarrollo Social, Encuesta Casen 2017

Finalmente, esta encuesta también realiza un análisis con respecto a las personas en situación de pobreza según sus ingresos y la compara con la pobreza multidimensional, dando los siguientes resultados:

Tabla 2: Contribución relativa de dimensiones a la pobreza multidimensional
Fuente: Ministerio de Desarrollo Social, Encuesta Casen 2017

Metodología	Situación	Total	
		Número	Porcentaje
Pobreza por ingresos	En situación de pobreza <u>no extrema</u> por ingresos	1.115.445	6,3
	En situación de <u>pobreza extrema</u> por ingresos	412.839	2,3
	En situación de pobreza por ingresos (extrema y no extrema)	1.528.284	8,6
Pobreza multidimensional	En situación de pobreza multidimensional	3.530.889	20,7

Esto quiere decir que el 20,7% de las personas en Chile se encuentran en situación de pobreza multidimensional, frente a un 8,6% en situación de pobreza por ingresos.

2.2.2 Registro Social de Hogares

A través del Ministerio de Desarrollo Social, el Estado busca constantemente fortalecer el Sistema de Protección Social, por lo que, para ello, es necesario contar también *con buenas herramientas de apoyo para los procesos de identificación y selección de quiénes requieren acceder a las prestaciones sociales del Estado* (RSH_Paper, 2018).

Esta herramienta es utilizada como el sistema de selección de usuarios para la diversidad de servicios y beneficios que ofrece el Estado denominada Registro Social de Hogares (RSH). Este es un *sistema de datos que consigna, almacena y procesa información de los atributos sociales, civiles y socioeconómicos de la población del país. Gracias a éste, se ha podido caracterizar a cerca de 5 millones de hogares y 13 millones de personas, que equivalen al 73,8% de la población nacional, un total de 27 instituciones y servicios relacionados con once Ministerios y la totalidad de las Municipalidades del país hoy cuentan con la información necesaria para verificar las condiciones de elegibilidad utilizadas en la selección y el acceso de las personas beneficiarias a los distintos programas y/o prestaciones sociales que ofrecen* (RSH_Paper, 2018).

Es un registro dinámico, ya que la información se actualiza de acuerdo con información auto reportada de personas y hogares, información proveniente de los registros administrativos disponibles en el Estado e información de los usuarios de prestaciones sociales administradas por entidades públicas y/o privadas. Además, contiene un conjunto de información multidimensional de la realidad individual/familiar disponible para el diseño e implementación de las políticas, programas, prestaciones e iniciativas legislativas más pertinentes a la realidad social, no sólo a nivel nacional, sino también a nivel regional, provincial y comunal.

El instrumento de medición es el formulario de Ingreso al Registro Social de Hogares, el cual fue diseñado por Ministerio de Desarrollo Social con el fin de consignar antecedentes del hogar.

Las preguntas del formulario se organizan temáticamente a través de los siguientes módulos:

- Localización
- Vivienda
- Personas del Hogar
- Educación
- Salud
- Ocupación
- Ingresos

Estos antecedentes serán utilizados para obtener una calificación socioeconómica, la cual estará dividida en 7 tramos. Cada uno agrupa a hogares con similares niveles ingresos y vulnerabilidad. A continuación, se muestra la agrupación por calificación socioeconómica.

Tabla 3: Calificación Socioeconómica por tramo – RSH
Fuente: Registro Social de Hogares, 2018

Tramo	Percentil	Descripción de Calificación Socioeconómica
Tramo del 40	0% – 40%	Hogares calificados en el 40% de menores ingresos o mayor vulnerabilidad
Tramo del 50	41% – 50%	Hogares calificados entre el 41% y el 50% de menores ingresos o mayor vulnerabilidad
Tramo del 60	51% – 60%	Hogares calificados entre el 51% y el 60% de menores ingresos o mayor vulnerabilidad
Tramo del 70	61% – 70%	Hogares calificados entre el 61% y el 70% de menores ingresos o mayor vulnerabilidad
Tramo del 80	71% – 80%	Hogares calificados entre el 71% y el 80% de mayores ingresos o menor vulnerabilidad
Tramo del 90	81% – 90%	Hogares calificados entre el 81% y el 90% de mayores ingresos o menor vulnerabilidad
Tramo del 100	91%-100%	Hogares calificados entre el 91% y el 100% de mayores ingresos o menor vulnerabilidad

Esta herramienta ya ha sido utilizada a nivel país, abarcando a un 73,8% de la población. La siguiente tabla muestra la cobertura del RSH por región.

Región	Hogares	Distribución de Hogares en el RSH por región (%)	Personas	Distribución de Personas en el RSH por región (%)	Porcentaje respecto del total de la población
Arica y Parinacota	74.045	1,5	172.025	1,3	100,0
Tarapacá	93.367	1,9	229.422	1,8	70,2
Antofagasta	137.039	2,8	382.070	3	66,5
Atacama	89.413	1,8	233.992	1,8	83,2
Coquimbo	214.920	4,4	599.488	4,6	79,1
Valparaíso	511.508	10,4	1.344.370	10,4	73,7
Metropolitana	1.759.153	35,7	4.531.252	35	63,5
O'Higgins	301.541	6,1	763.252	5,9	83,4
Maule	347.047	7,0	901.171	7,0	87,3
Biobío	630.535	12,8	1.714.554	13,2	82,9
La Araucanía	306.157	6,2	852.706	6,6	86,7
Los Ríos	125.183	2,5	327.193	2,5	88,6
Los Lagos	258.740	5,3	692.253	5,3	80,2
Aysén	34.727	0,7	90.034	0,7	87,2
Magallanes	43.572	0,9	111.367	0,9	73,9
Total	4.926.947	100,0	12.945.149	100,0	73,8

Ilustración 4: Cobertura del RSH, según región del país

Fuente: Registro Social de Hogares, 2018

2.2.3 Índice de Aislamiento Territorial

Se define aislamiento como un espacio físico que se encuentra: lejos, apartado, desconectado, con difícil acceso, incomunicado, o en situación de isla. El aislamiento es determinado por factores físicos, de localización y demográficos y puede ser medido de forma que indique el “grado” o “nivel” de aislamiento de un territorio (Subsecretaría de Desarrollo Regional y Administrativo – SUBDERE Ministerio del Interior y Seguridad Pública, 2011). En esta misma línea, se entiende como territorios aislados a las comunas definidas por la relación existente entre los componentes de *Aislamiento Estructural* y *Grado de Integración*. El siguiente gráfico ilustra los elementos utilizados para la determinación del Aislamiento Estructural. Este ayudará a identificar a las comunas que tienen mayores dificultades de acceso según las zonas establecidas por el estudio (zona norte, centro y sur).

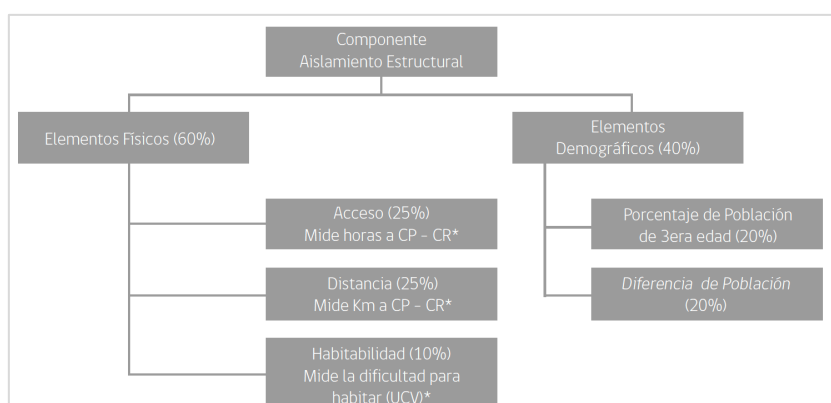


Ilustración 5: Elementos para la determinación del componente Aislamiento Estructural

Fuente: Subsecretaría de Desarrollo Regional y Administrativo – SUBDERE Ministerio del Interior y Seguridad Pública, 2011 *CP: Ciudad principal; CR: Capital regional; UCV: Universidad Católica de Valparaíso

Con respecto al componente Grado de Integración, este permite detectar el acceso a distintos elementos básicos para la población. Esto significa que, si el grado de integración de un territorio es mayor, entonces el efecto de aislamiento estructural será menor. Los elementos que componen el Grado de Integración se muestran en la siguiente ilustración.

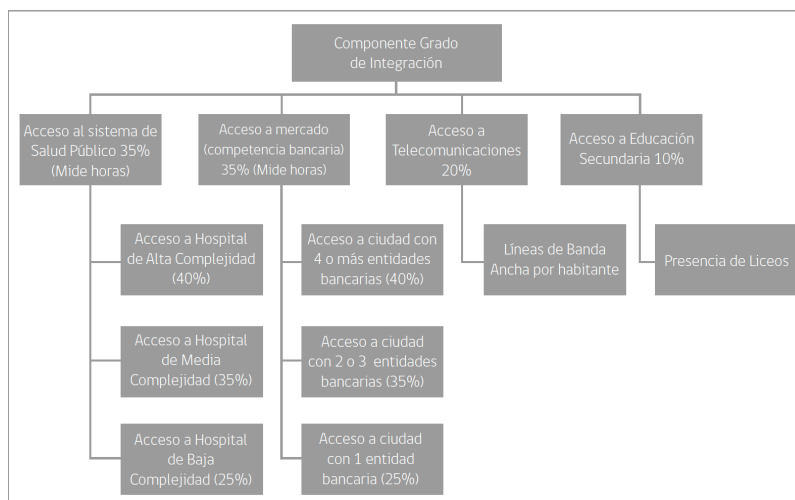


Ilustración 6: Elementos para la determinación del componente Aislamiento Estructural
Fuente: Subsecretaría de Desarrollo Regional y Administrativo – SUBDERE Ministerio del Interior y Seguridad Pública, 2011

Este estudio propone entonces, el cálculo del índice de aislamiento territorial (IAT) como:

$$IAT = (2 \times \text{Grado de Integración}) - \text{Componente Aislamiento Estructural}^2$$

Así es como el estudio presenta el IAT por comuna, cuyos datos se encontrarán en el Anexo 11.1 del presente documento.

2.2.4 Programa de Residencia Familiar Estudiantil (PRFE)

JUNAEB, actualmente es la institución responsable de la administración de programas de apoyo escolar para los estudiantes más vulnerables del país que se encuentran en riesgo de deserción. Para garantizar la continuidad en la educación de aquellos estudiantes que, en su lugar de residencia, preferentemente de sectores rurales, cuentan con reducida oferta académica de colegios y/o instituciones de educación superior para continuar sus estudios, se crea el Programa de Residencia Familiar Estudiantil. (Departamento de Becas JUNAEB, 2018).

² Fuente: Subsecretaría de Desarrollo Regional y Administrativo – SUBDERE Ministerio del Interior y Seguridad Pública, 2011

Este programa cuenta con asignación presupuestaria exclusiva para su ejecución dentro de los programas del Departamento de Becas, según la Ley de Presupuesto Anual.

Se ejecuta en las 15 regiones del país desde el año 1991 y tiene como objetivo favorecer el acceso, mantención y término en el sistema educacional de los y las estudiantes en situación de vulnerabilidad.

El programa consiste en el otorgamiento de alojamiento en el hogar de una familia tutora que brinda alimentación, apoyo pedagógico y apoyo afectivo a estudiantes en condición de vulnerabilidad, preferentemente del sector rural, que disponen de una reducida oferta de colegios educacionales cercanos a sus domicilios o cuya oferta no responde a los intereses de los estudiantes, lo cual los obliga a trasladarse a otras localidades o ciudades que cuentan con colegios o instituciones de educación superior para continuar sus estudios, ayudando a disminuir la deserción escolar.

Es administrado a través de las Direcciones Regionales de JUNAEB que materializa la entrega del beneficio antes descrito, mediante la suscripción de convenios con entidades ejecutoras que se encuentran inscritas en el registro para tal efecto, a las cuales JUNAEB transfiere los recursos para la atención de los beneficiarios del programa y administración del programa.

Para la asignación de este beneficio se crea un modelo de focalización el cual *otorga un puntaje que sirve para pesquisar y consignar la información en el Sistema Informático de Becas, así también permite asegurar una correcta medición de cada estudiante postulante o renovante, expresado en un lenguaje técnico conocido y validado en el área del trabajo social, programas y políticas sociales* (Departamento de Becas JUNAEB, 2018).

Este modelo de focalización cuenta con una base metodológica dividida en dimensiones, variables e indicadores, la cual se muestra en la siguiente tabla:

Tabla 4: Modelo de Focalización PRFE 2018
Fuente: Departamento de Becas JUNAEB, 2018

AÑO 2018		
Nombre Dimensión	Puntaje Máximo	Peso Dimensión
Vulnerabilidad	150	30%
Rendimiento Académico	125	15%
Aislamiento Geográfico	225	55%
TOTAL	500	100%
DIMENSIÓN DE VULNERABILIDAD (PUNTAJE = 150)		
SINAE		PUNTAJE MAXIMO
INDICADOR (Precargado)	Puntaje	
Primera Prioridad	40	40
Segunda Prioridad	35	
Tercera Prioridad	30	
No Vulnerable	0	
Sin cálculo prioridad	0	
CHILE SOLIDARIO		
INDICADOR (Precargado)	Puntaje	40
Familia inscrita en Chile Solidario o IEF	40	
Familia NO inscrita en Chile Solidario o IEF	0	
ULTIMO CURSO APROBADO POR LA MADRE		
INDICADOR	Puntaje	25
No tiene estudios	25	
Enseñanza Básica Incompleta	20	
Enseñanza Básica Completa	15	
Enseñanza Media Incompleta	10	
Enseñanza Media Completa	5	
Estudios Superiores Incompletos	3	
Estudios Superiores Completos	0	
ULTIMO CURSO APROBADO POR EL PADRE		
INDICADOR	Puntaje	25
No tiene estudios	25	
Enseñanza Básica Incompleta	20	
Enseñanza Básica completa	15	
Enseñanza Media Incompleta	10	
Enseñanza Media Completa	5	
Estudios Superiores Incompletos	3	
Estudios Superiores Completos	0	
PERTENENCIA A PUEBLOS ORIGINARIOS		
INDICADOR	Puntaje	20
Mapuche	20	
Aymara	20	
Yagan	20	
Kawhaskar	20	
Rapa Nui	20	
Atacameño	20	
Colla	20	
Quechua	20	
Diaguita	20	
Otro	20	
Ninguno (Sin Etnia)	0	
Dimensión Rendimiento Académico (PUNTAJE = 125)		
RENDIMIENTO ACADÉMICO		
INDICADOR (Precargado)	Puntaje	125
65 - 70	125	
60 - 64	110	
55 - 59	95	
50 - 54	80	
45 - 49	65	
40 - 44	50	
Menor que 4.0	40	
Dimensión Aislamiento Geográfico (PUNTAJE = 225)		
DESPLAZAMIENTO DESDE EL DOMICILIO PARTICULAR AL LUGAR DE ESTUDIO		
INDICADOR	Puntaje	25
Fuera de la Región	25	
Fuera de la Provincia	20	
Fuera de la Comuna	15	
Es comuna con dificultad de acceso	10	
Menos de 1 hora de traslado	5	
OFERTA EDUCACIONAL		
INDICADOR (Precargado)	Puntaje	100
Comuna Prioritaria o sin oferta	100	
Comuna con oferta casi inexistente	80	
Comuna con reducida oferta de educación	60	
Comuna con oferta educacional	40	
No prioritario	20	
MEDIO DE TRANSPORTE		
INDICADOR	Puntaje	50
Lancha, bote, barcaza	50	
Avión o avioneta	50	
A pie (más de 10 Kilómetros)	40	
Transporte municipal (en evento que no exista mic	40	
A pie y micro o bus	40	
En micro o bus (más de 10 km)	20	
En vehículo particular	20	
Otro (Caballo, bicicleta, carretera, etc)	20	
COSTO DE MOVILIZACIÓN		
INDICADOR	Puntaje	50
Es de más de 2.000 pesos diarios	50	
Es de 1.001 a 2000 pesos diarios	50	
Es de 501 a 1.000 pesos diarios	40	
Es de 301 a 500 pesos diarios	40	
Es menor igual a 300 pesos diarios	40	
No tiene costo	30	

El estudiante interesado debe realizar su postulación o renovación a través de la página web habilitada para tales efectos, ingresando sus datos personales y académicos al formulario de postulación y renovación en el Sistema Integral de Becas del programa, de acuerdo con los requisitos indicados en el mismo, y debe presentar la documentación solicitada para la revisión de antecedentes y verificación de cumplimiento de requisitos, dentro de los plazos establecidos por JUNAEB. (Ver Anexo 11.2)

El proceso de selección está conformado por una Comisión Regional, la cual determinará la continuidad de renovantes, el ingreso de nuevos becados y el análisis de los casos excepcionales.

Finalmente, hay que tener en cuenta que la asignación anual de este beneficio depende de los siguientes factores:

- Cupos regionales programados.
- Aplicación de modelo de focalización del programa que otorga puntaje
- Asignación al total de estudiantes renovantes del proceso en curso que cumplen requisitos.
- Asignación de estudiantes postulantes que cumplen requisitos y que fueron autorizados por la Comisión Regional.
- Asignación de cupos especiales.
- En caso de ser necesario, se aplican criterios de priorización considerando los puntajes del modelo de focalización.

2.2.5 Investigación de Operaciones y modelos de optimización

La Investigación de Operaciones se aplica en problemas que se refieren a la conducción y coordinación de operaciones (o actividades) dentro de una organización.

Cada vez es más frecuente que los administradores utilicen métodos cuantitativos y computadoras para llegar a la solución óptima de problemas que involucran un gran número de alternativas.

Las técnicas de la administración se aplican a las siguientes dos categorías básicas de problemas.

- 1) Problemas determinísticos**, en los que toda la información necesaria para obtener una solución se conoce con certeza.

- 2) **Problemas estocásticos**, en los que parte de la información necesaria no se conoce con certeza, sino más bien se comporta de una manera probabilística.

La optimización es una parte relevante de la investigación de operaciones, cuyas técnicas son: programación lineal (LP) y programación dinámica (DP).

Los problemas de optimización se componen de tres partes:

- 1) **Función objetivo:** es la medida cuantitativa del funcionamiento del sistema que se desea minimizar o maximizar, por ejemplo: minimizar los costos de producción o maximizar los beneficios de una empresa.
- 2) **Variables:** representan las decisiones que se pueden tomar para afectar el valor de la función objetivo.
- 3) **Restricciones:** es el conjunto de relaciones que ciertas variables están obligadas a satisfacer.

Finalmente, lo que se pretende es encontrar el valor que deben tomar las variables para hacer óptima la función objetivo satisfaciendo el conjunto de restricciones.

2.2.6 Programación Lineal con variables enteras

La Programación Lineal busca optimizar, (maximizar o minimizar) funciones lineales de variables reales con restricciones lineales (sistemas de inecuaciones lineales), optimizando una función objetivo también lineal. Si es que se le agrega una restricción adicional a las variables en las que solamente pueden tomar valores enteros, se le denomina programación lineal entera. Estos valores enteros pueden ser también binarios, es decir si sus valores son solo 0 o 1.

Un programa matemático es lineal si su modelo tiene la siguiente forma:

$$\left. \begin{array}{l} \text{Optimizar : } z = f(x_1, x_2, \dots, x_n) \\ \uparrow \\ z = c_1x_1 + c_2x_2 + \dots + c_nx_n \end{array} \right\} \text{ función objetivo}$$

(maximizar)
(o minimizar)

Sujeta a las siguientes restricciones :

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \cdot \\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{array} \right\} \begin{array}{l} \leq b_1 \\ \leq b_2 \\ \cdot \\ \cdot \\ \geq b_m \end{array} \begin{array}{l} \text{Restricciones} \\ \text{funcionales} \end{array}$$

$x_j \geq 0$ Restricciones de no negatividad (si se da el caso)

Donde:

$$\text{Parámetros del modelo} \left\{ \begin{array}{l} z = \text{valor de la medida global de efectividad} \\ x_j = \text{nivel de la actividad } j \text{ (variable de decisión)} \\ c_j = \text{incremento en } z \text{ que resulta al aumentar una unidad en el nivel de la actividad } j. \\ b_i = \text{cantidad de recurso } i \text{ disponible para asignar a las actividades} \\ a_{ij} = \text{cantidad del recurso } i \text{ consumido por cada unidad de la actividad } j. \end{array} \right.$$

2.2.7 Distancia Euclidiana

La distancia Euclidiana es la distancia entre dos puntos de un espacio euclideo, midiendo la longitud del segmento que los une (Universidad de Granada, 2014-2015). Es una consecuencia del Teorema de Pitágoras, el cual permite calcular la longitud de la hipotenusa a partir de la medida de los catetos.

Si tenemos dos puntos P y Q cuyas coordenadas en el plano son (a,b) y (c,d) respectivamente y queremos calcular la distancia euclidiana entre ellos, necesitamos obtener la longitud de los catetos del triángulo rectángulo que se definen como (c - a) y (d - b) (Bronshtein & Semendiaev, 1973), según la siguiente ilustración:

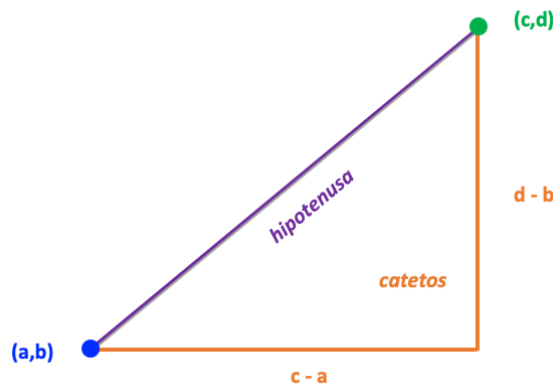


Ilustración 7: Distancia Euclidiana entre dos puntos
Fuente: Elaboración propia

Es decir, que la distancia es la longitud de la hipotenusa del triángulo rectángulo que se forma, y se determina con la siguiente fórmula:

$$d(P, Q) = \sqrt{(c - a)^2 + (d - b)^2}$$

2.2.8 Data Mining

La Minería de Datos es usada para transformar grandes volúmenes de datos en información utilizando métodos estadísticos apropiados. Este proceso de transformación busca que dicha información entregue conocimiento a los usuarios.

También es conocida como uno de los pasos esenciales para el proceso de descubrimiento de conocimiento (Han & Kamber, 2006).

La siguiente ilustración muestra las etapas del proceso de descubrimiento de conocimiento según (Han & Kamber, 2006):

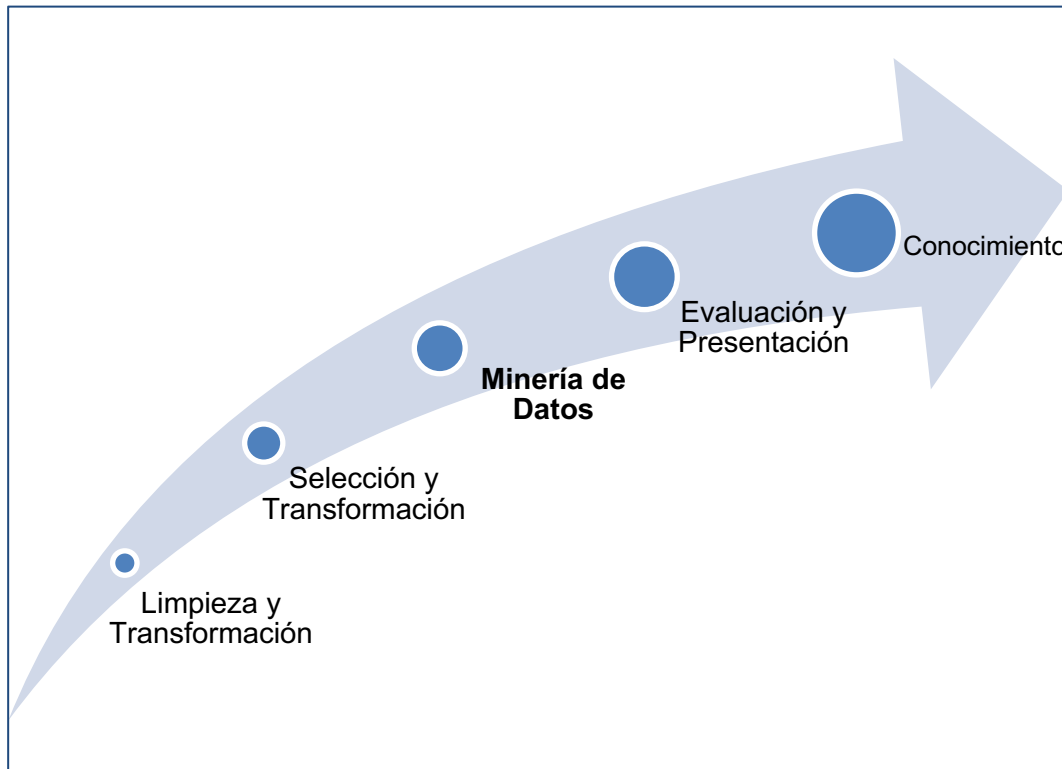


Ilustración 8: Pasos del proceso de Descubrimiento de Conocimiento

Fuente: Elaboración propia en base a (Han & Kamber, 2006)

En resumen, es un proceso esencial en el que se aplican métodos inteligentes para extraer patrones de datos. Estos métodos inteligentes pueden estar asociados con la caracterización, análisis de asociación y correlación, clasificación, predicción, análisis de conglomerados, análisis de valores atípicos y análisis de evolución de datos relevantes que fueron seleccionados en el paso anterior.

2.2.9 Modelos Predictivos

Los modelos predictivos en la Minería de Datos se componen de técnicas conocidas también como modelos supervisados. Predicen el valor de un atributo de un conjunto de datos, conocidos otros atributos. A partir de estos se induce una relación entre dicho atributo conocido y otra serie de atributos. Esas relaciones sirven para realizar la predicción en datos cuyo atributo es desconocido. Esta forma de trabajar se desarrolla en dos fases:

- Entrenamiento, en el que se construye un modelo usando un subconjunto de datos con atributo conocido.
- Prueba, en el que se prueba el modelo con el resto de los datos.

2.2.9.1 Validación cruzada

Es una técnica con la que se puede identificar la existencia de sobreajuste a los datos de entrenamiento de los modelos (Analytics Lane, 2018). Esta técnica divide en grupos de igual tamaño el conjunto de datos de entrenamiento, para poder entrenar el modelo una vez por cada uno de los grupos (iteraciones).

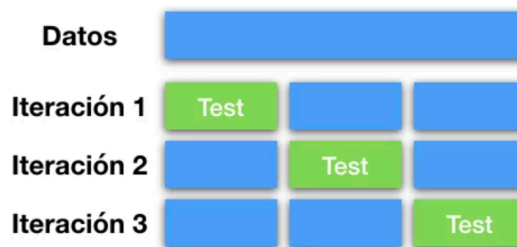


Ilustración 9: Proceso de Validación cruzada

Fuente: Sitio web - <https://www.analyticslane.com/2018/07/02/gridsearchcv/>

Algunos algoritmos de este tipo son los Árboles de Decisión, Redes Neuronales, Regresión, series temporales, entre otros.

2.2.10 Árboles de Decisión

Esta técnica de predicción permite clasificar los datos en grupos según los valores de cada una de las variables que componen este conjunto de datos. Se elige un atributo como raíz y se desarrolla el árbol de acuerdo con las variables más significativas. Esto representa literalmente un árbol cuyas divisiones describen las decisiones posibles y por lo tanto los resultados de las distintas combinaciones de decisiones y eventos. Cada evento tiene asignado probabilidades y a cada una de las ramas se le establece un resultado. Ayudan a tomar la decisión “más acertada”, desde un punto de vista probabilístico, ante una cierta cantidad de posibles decisiones.

A continuación, se muestra una representación gráfica como ejemplo que, de acuerdo con los distintos algoritmos utilizados para su construcción, muestra como resultado esta técnica.

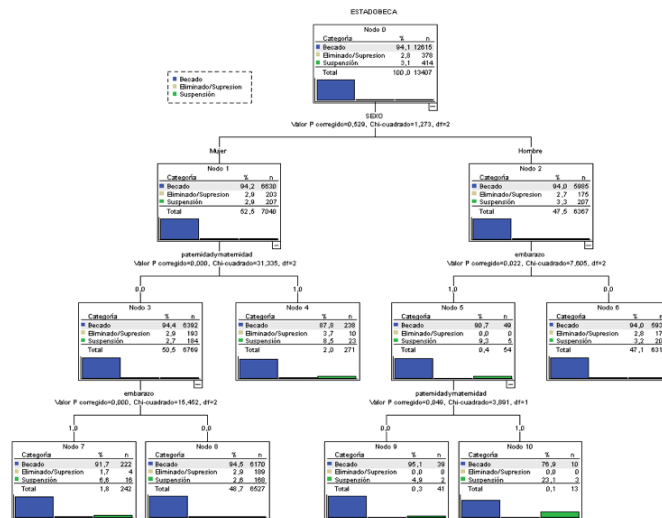


Ilustración 10: Ejemplo Gráfico del Árbol de Decisión de una Beca
Fuente: Elaboración propia

2.2.10.1 Boosting

Conjunto de algoritmos basados en crear una regla con alta precisión utilizando reglas débiles e imprecisas, es decir es un método utilizado para mejorar las predicciones del modelo de cualquier algoritmo de aprendizaje dado.

Tipos de Boosting

- *AdaBoost (Boosting Adaptativo)*: utilizado para aumentar el rendimiento de los árboles de decisión en problemas de clasificación binaria.
- *Gradient Boosting*: funciona agregando secuencialmente predictores a un conjunto, cada uno corrigiendo a su predecesor.
- *Extreme Gradient Boosting (XGBoost)*: es una implementación de árboles de decisión impulsados por gradiente diseñados para la velocidad y el rendimiento. El modelo Gradient Boosting generalmente es muy lento en su implementación debido al entrenamiento en modelos secuenciales. Por lo tanto, no son muy escalables. Por lo tanto, XGBoost se centra en la velocidad computacional y el rendimiento del modelo.

2.2.11 Random Forest

Son una combinación de predictores de árboles, de modo que cada árbol depende de los valores de un vector aleatorio muestreado de forma independiente y con la misma distribución para todos los árboles en el bosque.

Clasificador que consiste en una colección de clasificadores de árboles estructurados $\{h(\mathbf{x}, \theta_k), k = 1, \dots\}$ donde las $\{\theta_k\}$ son vectores aleatorios independientes distribuidos

idénticamente y cada árbol emite un voto unitario para la clase más popular en la entrada x .

Ventajas:

- Existen muy pocas suposiciones, por lo que la preparación de los datos es mínima.
- Maneja una gran cantidad de variables de entrada identificando las más significativas.

Desventajas:

- Modelo “caja negra”, poco control en lo que hace el modelo.
- Pérdida de interpretación
- Favorable para clasificación, pero no tanto para regresión.

2.2.12 Regresión Logística

La Regresión Logística es especialmente útil cuando solo hay dos posibles respuestas, es decir, cuando la variable de respuesta es dicotómica (éxito o fracaso). El objetivo primordial de esta técnica es modelar como influyen las variables independientes en la probabilidad de ocurrencia de un suceso en particular.

Sea Y una variable dependiente binaria (con dos posibles valores: 0 y 1). Sean un conjunto de k variables independientes, (X_1, X_2, \dots, X_k) , observadas con el fin de predecir/explicar el valor de Y .

El objetivo consiste en determinar:

$$P[Y=1/X_1, X_2, \dots, X_k] \mapsto P[Y=0/X_1, X_2, \dots, X_k] = 1 - P[Y=1/X_1, X_2, \dots, X_k]$$

Para ello, se construye el modelo

$$P[Y=1/X_1, X_2, \dots, X_k] = p(X_1, X_2, \dots, X_k; \beta)$$

donde:

$p(X_1, X_2, \dots, X_k; \beta): \mathbb{R}^k \xrightarrow{\text{función de enlace}} [0,1]$ que depende de un vector de parámetros

$$\beta = (\beta_1, \beta_2, \dots, \beta_k)$$

2.2.13 K nearest Neighbors

Método de clasificación de tipo supervisada, el cual estima una función de densidad $f(x, C_j)$, siendo x la variable independiente y C_j la clase j . Esta función determina la probabilidad a posteriori de que la variable x pertenezca a la clase j . En otras palabras, un nuevo caso se va a clasificar en la clase más frecuente a la que pertenecen sus K vecinos más cercanos.

La variable independiente x es particionada en regiones por localizaciones y etiquetas de los elementos de entrenamiento, esto quiere decir que un punto x en el espacio es asignado a la clase C , si esta es la clase más frecuente entre los k elemento más cercanos. Esta cercanía de los elementos se calcula con la fórmula de distancia euclidiana, según fue descrito en la sección 2.2.7 del presente.

Debe existir un grupo de datos de entrenamiento, los cuales sirven para almacenar los vectores característicos y las etiquetas de las clases de dichos elementos de entrenamiento.

Luego, se debe realiza la etapa de clasificación, en la cual se calcula la distancia entre los vectores almacenados y el nuevo vector y se seleccionan los k elementos más cercanos. El nuevo vector es clasificado con la clase que más se repite en los vectores seleccionados.

Las siguientes ilustraciones muestran mediante un ejemplo, cómo se clasifica un nuevo vector de datos aplicando el modelo KNN.

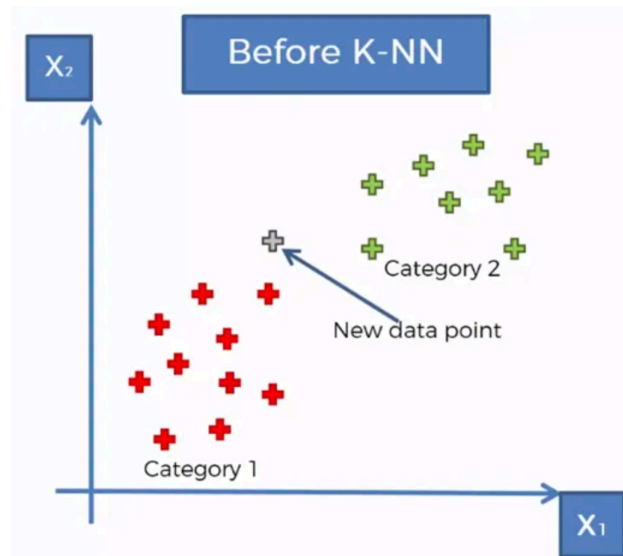


Ilustración 11: Clasificación de un nuevo vector de datos (antes de aplicar el modelo KNN)
Fuente: Sitio Web - https://www.jacobsoft.com.mx/es_mx/k-nearest-neighbors/

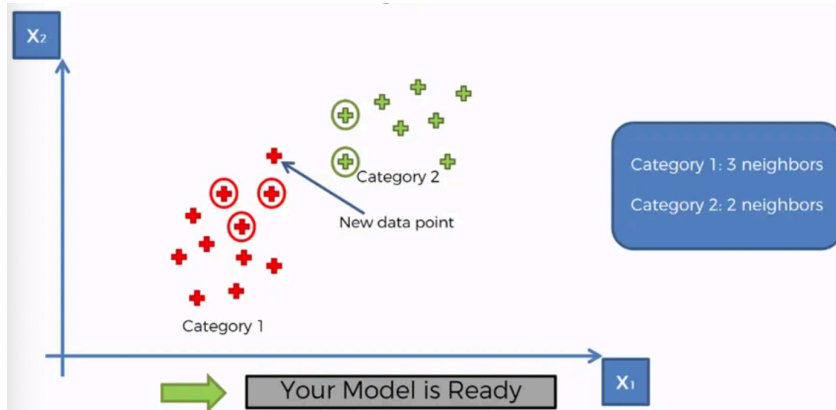


Ilustración 12: Ejemplo de aplicación modelo KNN seleccionando $k = 5$ vecinos más cercano al nuevo vector de datos.

Fuente: Sitio Web - https://www.jacobsoft.com.mx/es_mx/k-nearest-neighbors/

Finalmente, el nuevo elemento se asignó a la categoría 1, dado que para $k = 5$ existen más vecinos de dicha categoría.

2.2.13.1 Support Vector Machine (SVM)

Método de aprendizaje supervisado, en otras palabras, es un tipo de máquinas de aprendizaje, las cuales deben inicialmente entrenarse con situaciones en las que se les dice la respuesta correcta sobre muchos ejemplos y luego se convierte en una caja que devuelve la respuesta ante un nuevo caso. (¿Qué son las SVM?, 2008). El modelo está rigurosamente sustentado por las teorías estadísticas de aprendizaje propuestas por Vladimir Vapnik.

SVM sirve para predecir una clasificación binaria, multi-categoría o de regresión general. Este método logra encontrar una “superficie” que intenta separar los ejemplos negativos y positivos con el margen más grande posible a ambos lados del hiperplano (¿Qué son las SVM?, 2008).

Si es un caso bidimensional se representa por una recta. Si se utilizan 3 atributos entonces la “superficie” sería un plano, y para el caso de más dimensiones, entonces sería un hiper-plano. La siguiente ilustración muestra un ejemplo los vectores que mejor “soporten” los hiper-planos con el mayor margen de separación. (¿Qué son las SVM?, 2008)

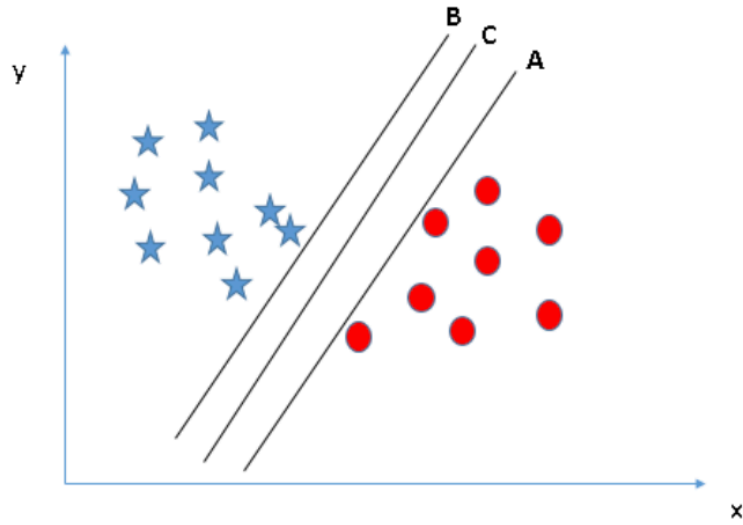


Ilustración 13: Superficie obtenida aplicando el método SVM identificando el hiper-plano
Fuente: Sitio Web - <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>

CAPÍTULO 3: PLANTEAMIENTO ESTRATÉGICO Y MODELO DE NEGOCIOS

3.1 Posicionamiento Estratégico

Para determinar el posicionamiento estratégico de la JUNAEB, se utilizará el modelo Delta de Arnoldo Hax para definir en qué enfoque se sitúa esta institución con el fin de cumplir con el propósito para lo cual fue creada.

Debido a que JUNAEB es una institución pública, y sus objetivos estratégicos se centran en la perspectiva de resultados con el fin de contribuir a facilitar la incorporación, permanencia y término en el sistema escolar de alumnos vulnerables. Esto se logra con el uso eficiente de los recursos estatales asignados durante el año y, además, ejecutando sus procesos operacionales de manera efectiva cumpliendo con los marcos normativos que la rigen. Por lo que JUNAEB se posiciona estratégicamente con un enfoque de **eficiencia administrativa**. La siguiente figura ilustra el posicionamiento estratégico de JUNAEB según el triángulo del modelo Delta y sus enfoques:

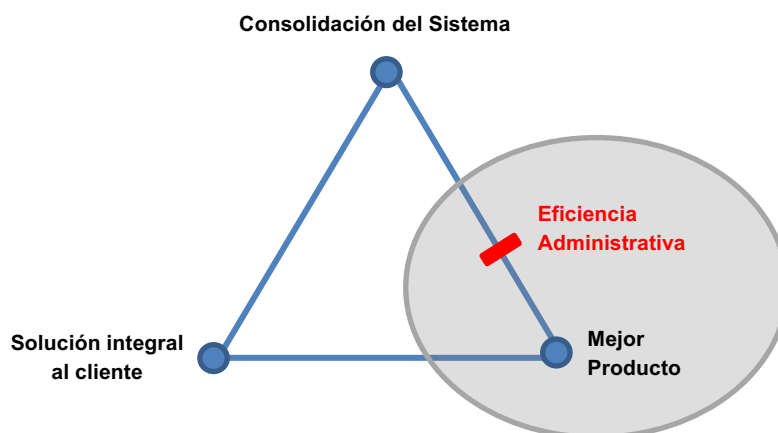


Ilustración 14: Modelo Delta-Hax JUNAEB
Fuente: Elaboración propia

3.2 Balanced Scorecard

Para dar cumplimiento con la misión de JUNAEB, BECAS ha definido los siguientes objetivos estratégicos:

- Aumentar la ejecución presupuestaria anual en base a la cobertura programada.
- Aumentar el cobro efectivo.
- Asignar los beneficios de manera oportuna.
- Mejorar los mecanismos de comunicación con el beneficiario.
- Mejorar la transparencia y calidad del proceso de pago de los beneficios.

- Monitorear avance en la gestión de formularios.
- Aumentar las supervisiones relacionadas con la correcta aplicación de la normativa.
- Fortalecer las capacidades internas asociadas a: dotación de RRHH, capacitación, sistemas de información y otros.

En relación con lo anteriormente descrito, en la siguiente tabla se muestra el cuadro de mando integral definido por Becas y que comprenden los objetivos estratégicos con sus indicadores y metas, clasificados en las 4 perspectivas establecidas por el cuadro de mando integral.

Tabla 5: Cuadro de Mando Integral de Becas – JUNAEB

Fuente: Elaboración propia

PERSPECTIVA	OBJETIVO ESTRATEGICO	INDICADOR	META
RESULTADOS	Aumentar la ejecución presupuestaria anual en base a la cobertura programada	Porcentaje de cobertura ejecutada	98,5%
		Porcentaje de Presupuesto ejecutado	98,0%
USUARIOS	Aumentar el cobro efectivo	Porcentaje de estudiantes con cuotas pagadas por cuenta rut	80,0%
	Mejorar los mecanismos de comunicación con el beneficiario	Porcentaje de estudiantes informados sobre el proceso de renovación de los beneficios	85,0%
		Nº de alumnos pagados mensualmente por modalidad masivo, que son avisados del pago a través de correo electrónico	20000
PROCESOS	Asignar los beneficios de manera oportuna	Porcentaje de estudiantes que renuevan sus beneficios	70,0%
		Porcentaje de Manuales de Procedimientos de cada beneficio elaborados oportunamente	100,0%
	Mejorar la transparencia y calidad del proceso de pago de los beneficios.	Porcentaje de disminución de alumnos con pagos pendientes NO cobrados en años anteriores	35,0%
	Monitorear avance en la gestión de formularios	Porcentaje total de formularios revisados, asociados a la revisión documental externa	100,0%
		Porcentaje de Formularios Incompletos documentalmente del total de formularios revisados	25,0%
	Aumentar las supervisiones relacionadas con la correcta aplicación de la normativa	Porcentaje de Direcciones Regionales supervisadas y con plan de acción sugerido, respecto a la aplicación del Manual de Procedimientos de los beneficios	100,0%
APRENDIZAJE	Fortalecer las capacidades internas asociadas a: dotación de RRHH, capacitación, sistemas de información y otros	Porcentaje de Direcciones Regionales capacitadas en el Sistema Informático de Becas (SIB)	100,0%

3.3 Modelo de Negocios

El modelo de negocio de BECAS cuenta con los siguientes componentes:

- En la interfaz del cliente más la propuesta de valor se encuentra, para el segmento de clientes, los beneficiarios que son los estudiantes de los distintos niveles educacionales (media, básica y superior) y de las distintas jornadas. Además, se define como cliente a las Instituciones de Educación Superior

quienes reciben el pago monetario de los beneficios de arancel que los estudiantes se han adjudicado.

Para la relación con el cliente se definió la atención personalizada que las oficinas de las Direcciones Regionales brindan a los beneficiarios.

Asimismo, para el componente de los canales se identificaron los siguientes:

- Vía Presencial (en Direcciones Regionales)
- Vía Correo Electrónico
- Vía Telefónica
- Mediante Reclamos en el sistema de atención ciudadana

Finalmente, como propuesta de valor BECAS busca contribuir a facilitar la incorporación, permanencia y término en el sistema escolar de alumnos vulnerables, ejecutando eficientemente los recursos estatales asignados anualmente.

- Para el componente de actividades clave se definen las siguientes:
 - Calcular coberturas programadas (población objetivo)
 - Asignar los beneficios
 - Informar a los beneficiarios el resultado para cobro efectivo
 - Ejecutar el presupuesto asignado

Los recursos claves están asociados a los recursos humanos (funcionarios comprometidos y con la capacidad de llevar a cabo las actividades claves antes descritas) y los sistemas informáticos que apoyan a la ejecución de estos.

De igual forma, se identifican como socios clave servicios públicos como el Ministerio de Educación, quien proporciona lineamientos e información clave para poder llevar a cabo los procesos de BECAS. El Ministerio de Desarrollo Social juega un rol activo actualmente, ya que mediante este se ha logrado disminuir la revisión documental implementado como modelo de focalización la nueva ficha del registro sociales de hogares (RSH) que principalmente entrega variables de vulnerabilidad aplicadas en estos modelos. Otro socio clave identificado es la denominada red colaboradora, compuesta por los establecimientos educacionales, Municipios y otros estamentos públicos o privados que sirven como intermediarios para llegar al beneficiario.

- Para el componente de estructura de costos se identifican los más importantes, los cuales están relacionados a la dotación de personal, implementación de proyectos informáticos (nuevo sistema de información integral de becas) y el presupuesto asignado para la elaboración de licitaciones buscan contratar servicios externos para la revisión documental (formularios y documentos físicos necesarios para la asignación del beneficio), supervisiones en

programas de residencia estudiantil (con el fin de determinar el cumplimiento de la normativa de los programas) y estudios de satisfacción usuario a varios programas que sirven para conocer las necesidades del beneficiario y estas aplicarlas en rediseños de modelos de focalización para la asignación del beneficio.

- Por último, para el componente de ingresos este engloba principalmente el presupuesto asignado en el año aprobado por la Dirección de Presupuesto y que se basa en las proyecciones y cálculo realizados para cubrir con la demanda de estudiantes que requieren hacer uso de los beneficios entregados por BECAS. Cada beca/programa, cuenta con una línea presupuestaria aprobada anualmente.

A continuación, se presenta en la siguiente figura, el modelo de negocio de BECAS:

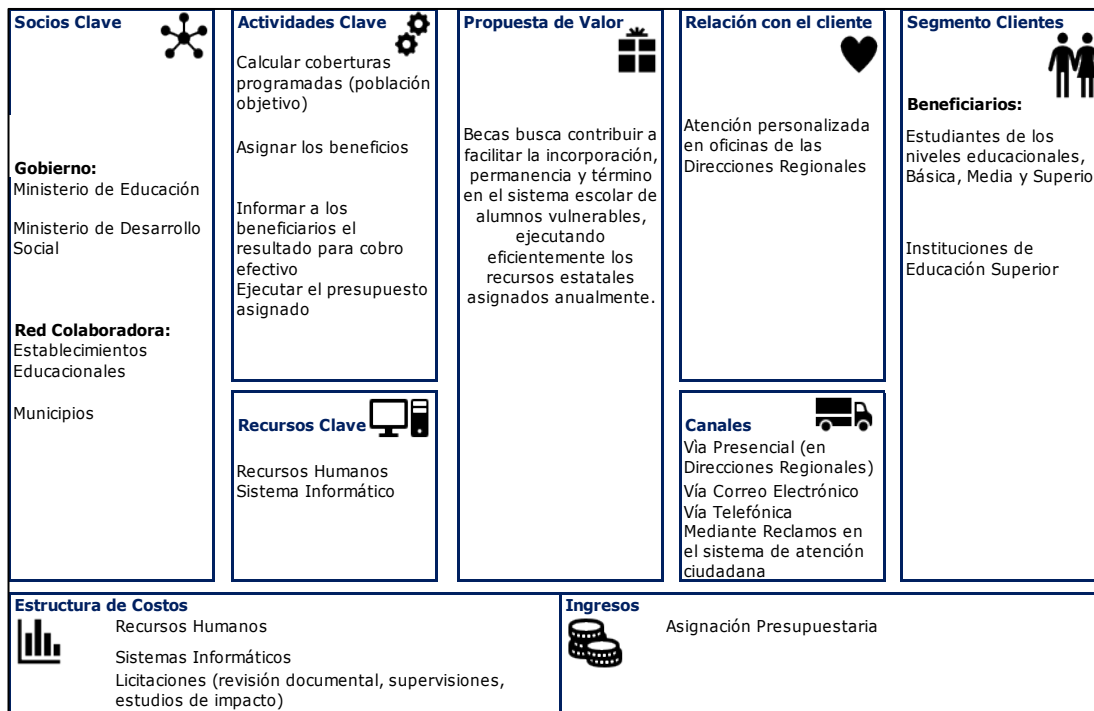


Ilustración 15: CANVAS de Becas – JUNAEB

Fuente: Elaboración propia

Según la posición estratégica definida anteriormente comprobamos mediante este modelo de negocio, que BECAS se ubica también en la estrategia de eficiencia administrativa debido a que su operación se centra en los resultados. Pero, podemos notar también que se está enfocando cada vez más al beneficiario implementando actividades relacionadas con los mismos, para poder entregar el beneficio en el tiempo requerido.

CAPÍTULO 4: ANÁLISIS SITUACIÓN ACTUAL

En este capítulo se describirá y analizará la situación actual del proceso de gestión de asignación del Programa de Residencia Familiar Estudiantil (PRFE).

4.1 Arquitectura de Procesos

El siguiente diagrama elaborado en IDEF0 muestra la arquitectura de los procesos de JUNAEB en la actualidad, en el cual se identifican las siguientes macros:

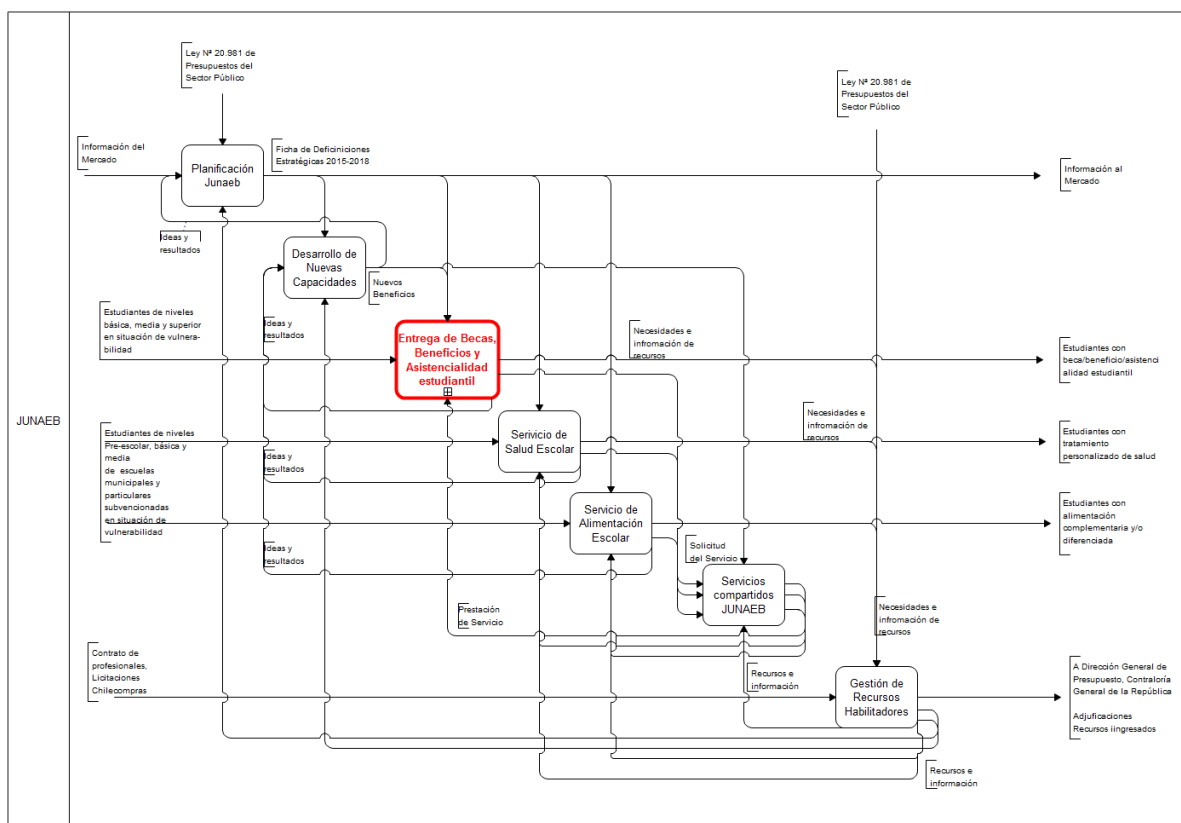


Ilustración 16: Macro 0 de la JUNAEB

Fuente: Elaboración propia

Cadena de Valor: La JUNAEB cuenta con varias líneas de negocio que componen su cadena de valor las cuales se centran en:

- Entrega de Becas, Beneficios y Asistencialidad Estudiantil
- Servicio de Salud Escolar
- Servicio de Alimentación Escolar

Asimismo, cuenta con Servicios compartidos en los cuales destaca el servicio de planificación y estudios, los cuales son los encargados de establecer las relaciones entre los servicios externos para la obtención de información estratégica para la operación de cada una de las cadenas de valor.

2. Desarrollo de Nuevas Capacidades
3. Planificación JUNAEB
4. Gestión de Recursos Habilitadores

Analizando esta arquitectura se puede establecer que JUNAEB mantiene sus procesos de negocio con baja estandarización y alta integración, en la cual se integran los procesos y se comparte la información entre estos de manera que la institución es vista por los usuarios como única.

4.2 Modelamiento Detallado de Procesos

4.2.1 Modelamiento IDEF0

Desagregando la Macro 1.1 de nombre “Entrega de Becas, Beneficios y Asistencialidad Estudiantil” se mostrará el detalle de los procesos involucrados, el cual se presenta en el siguiente diagrama:

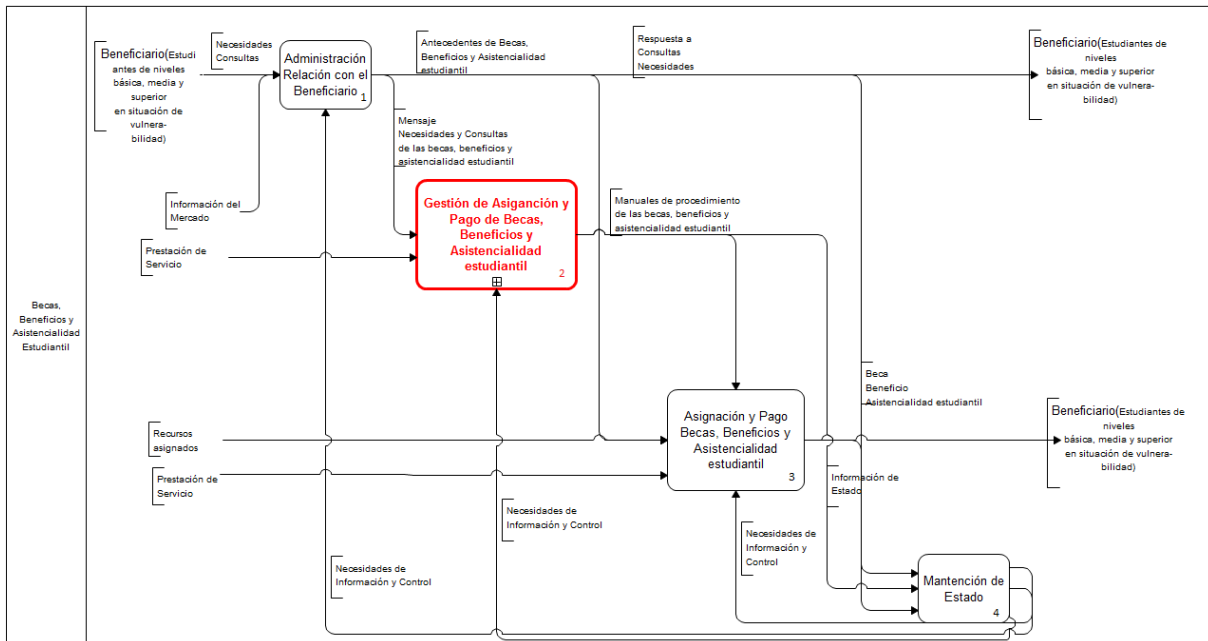


Ilustración 17: Macro 1.1 Cadena de Valor: Entrega de Becas, Beneficios y Asistencialidad Estudiantil

Fuente: Elaboración propia

Los principales procesos involucrados en esta macro son:

1. Administración de la relación con el beneficiario, el cual recibe como entrada información de los beneficiarios actuales.
2. **Gestión de Asignación y Pago de Becas, Beneficios y Asistencialidad Estudiantil**, cuya entrada se relaciona con la información de salida que arroja el proceso anterior, tales como las necesidades de beneficiarios y las consultas que surgen por cada beca.

3. Supervisión y Control de Becas, Beneficios y Asistencialidad Estudiantil, en el cual los insumos de supervisión control se determinan de acuerdo con los manuales de procedimientos de las becas que son el resultado del proceso anterior.
4. Mantenimiento de Estado, encargada de recibir y proporcionar información relevante para la toma de decisiones para cada uno de los procesos precedentes.

Si ingresamos al proceso de Gestión de Asignación y Pago de Becas, Beneficios y Asistencialidad Estudiantil, nos vamos a encontrar con varios subprocesos involucrados, necesarios para su ejecución:

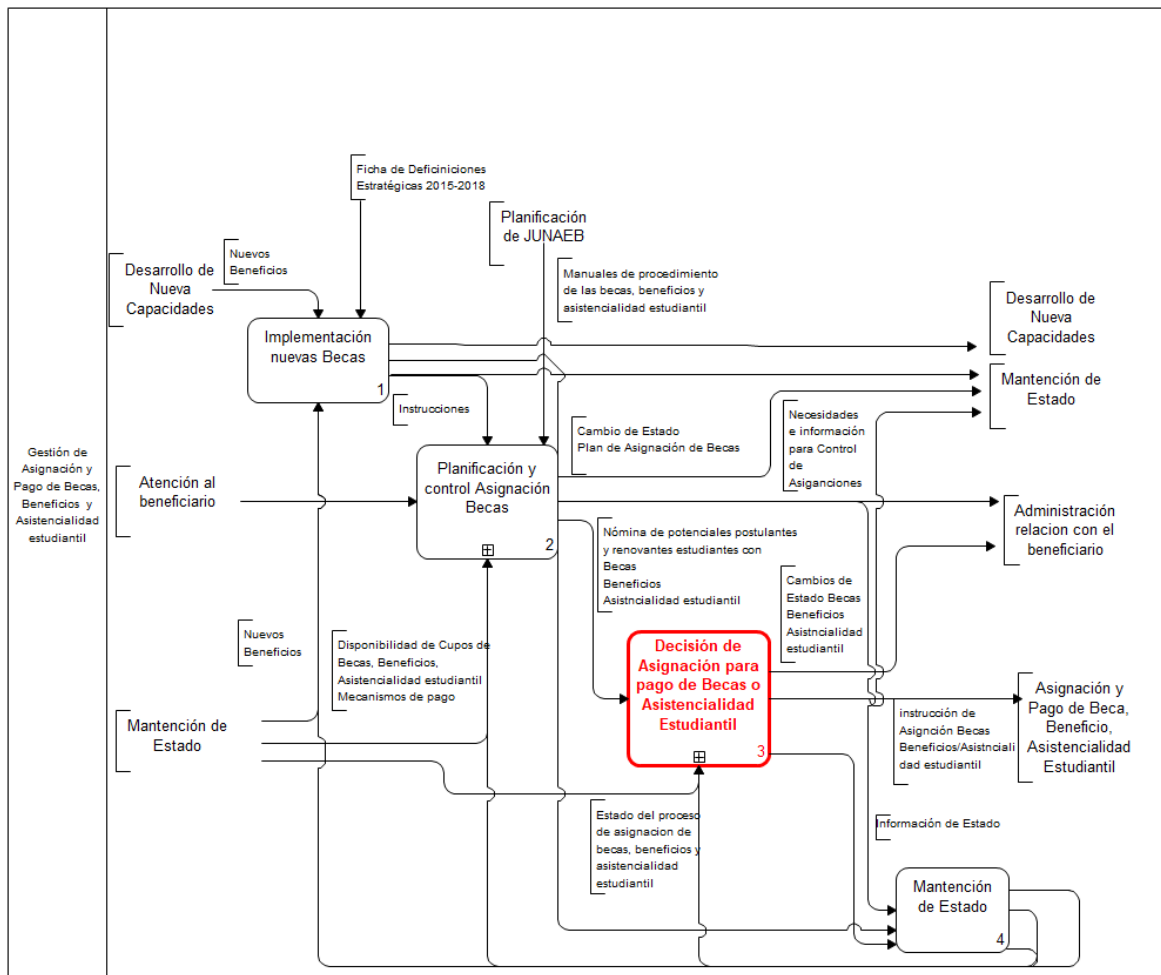


Ilustración 18: Macro 1.1.2: Gestión de Asignación y Pago de Becas, Beneficios y Asistencialidad Estudiantil

Fuente: Elaboración propia

Ahora se puede notar el detalle de los procesos involucrados en la Macro 1.1.2, que cuenta con tres subprocesos:

1. Implementación de nuevas Becas, el cual se encuentra relacionado con la creación de nuevos beneficios de acuerdo con la Ficha de Definiciones Estratégicas 2015-2018.
2. Planificación y Control Asignación de Becas, es el proceso que, utilizando el resultado del proceso anterior (instrucciones), ejecuta sus subprocesos para poder determinar las coberturas de cada beca en el año y así poder planificar el presupuesto anual y establecer los planes de control necesario para ejecutar eficientemente los recursos.
3. **Decisión de la Asignación para el pago de Becas o Asistencialidad Estudiantil**, proceso que se ejecuta una vez al año y en el que, mediante las normativas vigentes de cada programa y la información de los postulantes y renovantes, se realiza la asignación, es decir, se determina cuantos estudiantes se adjudicarán o no cada uno de los beneficios.
4. Mantenimiento de Estado, encargada recibir y proporcionar información relevante para la toma de decisiones para cada uno de los procesos precedentes.

Si siguiendo con este análisis, si ingresamos al proceso de Decisión de la Asignación para el pago de Becas o Asistencialidad Estudiantil, se mostrarán los siguientes subprocesos que figuran a continuación:

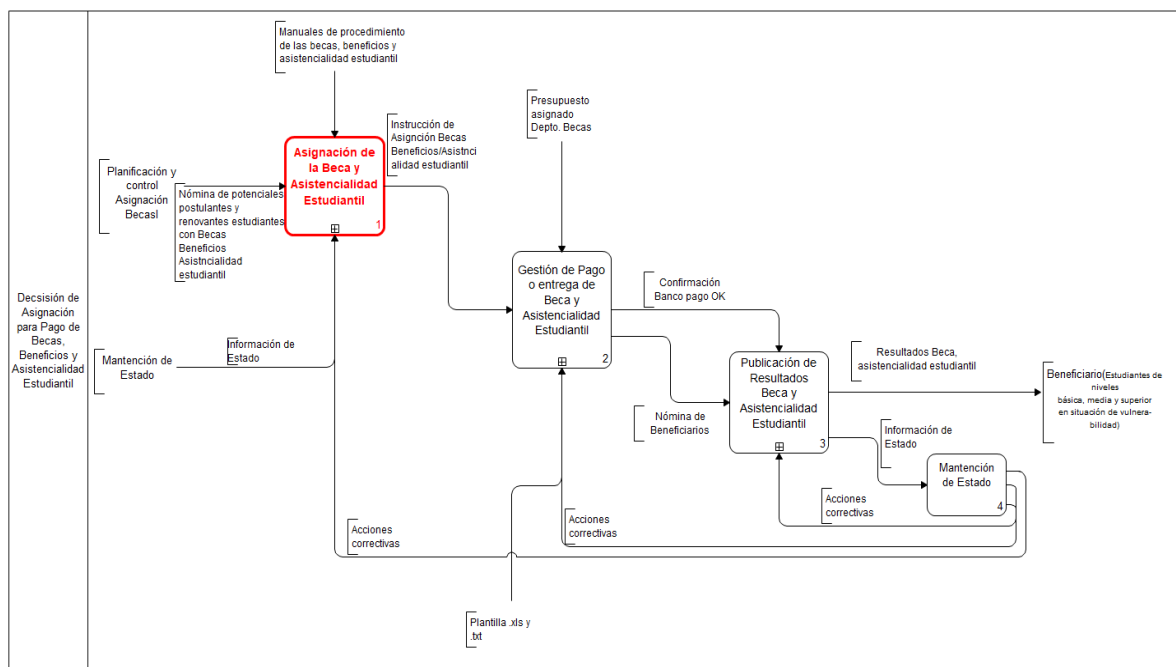


Ilustración 19: Macro 1.1.2.3 – Decisión de la Asignación para pago de Becas o Asistencialidad Estudiantil

Fuente: Elaboración propia

Los procesos involucrados en esta Macro 1.1.2.3 son:

1. **Asignación de la Beca y Asistencialidad Estudiantil:** es el proceso que se propone rediseñar. Este contiene subprocesos asociados con la consolidación de los postulantes y renovantes. También incluye la planificación de la entrega de la beca mediante la validación de los requisitos aplicando el modelo de focalización como una lógica de negocio (para el caso de PRFE), en el que como producto se confecciona la nómina de becados.
2. Gestión de Pago o Entrega de Becas o Asistencialidad Estudiantil, este proceso contiene todas las tareas asociadas con el pago del beneficio, si es dinero de libre disposición, o entrega de la beca o asistencialidad estudiantil para el caso del PRFE. Este proceso requiere como insumo, la nómina final de beneficiarios que se desprende del proceso anterior.
3. Publicación de Resultados de Becas o Asistencialidad Estudiantil, proceso donde se realizan las actividades relacionadas con la publicación de los resultados de las postulaciones y renovaciones de las becas y asistencialidad estudiantil.
4. Mantención de Estado, es el proceso que recibe y proporciona información relevante para la toma de decisiones para cada uno de los procesos precedentes.

4.3 Diagnóstico de la Situación Actual

Actualmente, el 85% de los programas de becas y asistencialidad estudiantil (residencias) utilizan la información socioeconómica que entrega el Registro Social de Hogares (RSH) para medir la vulnerabilidad. Esto se determina gracias a la calificación que asocian al grupo familiar agrupados en tramos, que son definidos por este instrumento creado y administrado por el Ministerio de Desarrollo Social (MDS). Pero, el Programa de Residencia Familiar Estudiantil (PRFE) no lo utiliza. El programa cuenta con su propio modelo de focalización que permite discriminar a nivel regional a los estudiantes que cumplen con los requisitos y que gracias a un puntaje pueden ser beneficiados con este programa.

Actualmente, este modelo contiene 3 dimensiones denominadas Dimensión de Vulnerabilidad, Dimensión Académica y Dimensión de Aislamiento Geográfico. Estas 3 dimensiones están compuestas por variables que, de acuerdo con las categorías asociadas entregan un puntaje, cuyo peso relativo es más significativo para la dimensión de Aislamiento Geográfico, según la naturaleza de este programa. El mayor inconveniente que existe es que el modelo no ha sido actualizado hace más de 10 años, por lo que no se ha analizado ni comparado con la medición de vulnerabilidad vigente como es el instrumento RSH. Si comparamos el modelo actual del RSH versus

el modelo de focalización vigente del PRFE, quedará en evidencia las variables que no son contempladas para la medición de la vulnerabilidad, por lo que no sería una medición integral.

Tabla 6: Comparación Dimensiones y Variables RSH – Modelo de Focalización PRFE

Fuente: Elaboración propia

Registro Social de Hogares		Focalización PRFE	
DIMENSIONES	VARIABLES	DIMENSIONES	VARIABLES
I.- Educación	Asiste a algún establecimiento educacional	I.- Educación	Último curso aprobado por el padre
	Por qué no asiste		Último curso aprobado por la madre
	Hasta qué curso llego		Rendimiento Académico
II.- Salud	Condiciones de forma permanente	II.- Salud	Declaración de salud a los postulantes
	Puede hacer compras o ir al médico solo		
	Puede salir solo		
	Puede bañarse, lavarse los dientes solo		
	Puede moverse solo dentro de la casa		
	Puede controlar esfínteres		
III.- Ocupación	Trabaja	III.- Económica	SINAE CHILE SOLIDARIO
	Ocupación actual		
	Rama de actividad económica		
	Temporalidad / permanencia		
	Contrato de trabajo		
	Relación contractual		
	Horas de trabajo a la semana		
	Buscó trabajo en los últimos 2 meses		
IV.- Vivienda	Tipo de vivienda	IV.- Vivienda	NO APLICA
	Situación del sitio		
	Situación de la vivienda		
	Fuente de agua		
	Distribución de agua		
	Sistema de baño		
	N° de habitaciones de uso dormitorio		
	Materialidad muros		
	Materialidad pisos		
	Materialidad techos		
V.- Sociocultural	NO APLICA	V.- Sociocultural	Pertenencia a pueblos originarios
VI.- Territorial	NO APLICA	VI.- Territorial	Desplazamiento desde el domicilio particular al lugar de estudio
			Oferta Educacional
			Medio de transporte
			Costo movilización

Este cuadro comparativo muestra que la dimensión “Vivienda” no está presente en el modelo de focalización de PRFE, la cual es necesaria para el cálculo de la calificación socioeconómica según el RSH. También se visualiza que las dimensiones “Sociocultural” y “Territorial” sí son contempladas en el modelo de focalización. Principalmente la dimensión “Territorial” es utilizada por la naturaleza del programa. Por otra parte, para la dimensión “Económica” el PRFE aplica información de del Sistema Chile Solidario, *sistema público de protección social dirigido a las familias y sus integrantes en situación de extrema pobreza*.³ Simplemente, otorga puntaje a los postulantes que se encuentren dentro del registro de Chile Solidario.

Además, emplea la variable Sistema Nacional de Asignación con Equidad (SINAE)⁴, medición construida por JUNAEB para la condición de vulnerabilidad, la que identifica principalmente prioridades de atención las desigualdades en que se encuentra la población escolar.

De igual forma, la mayoría de las variables contempladas en el modelo de focalización son declarativas, es decir, los postulantes y renovantes ingresan la información solicitada y esta no es corroborada con fuentes externas, por lo que no son confiables al momento de otorgar el puntaje.

Otro problema claramente identificable es que existe mucho trabajo manual, principalmente relacionado con la verificación documental, que como ya fue mencionado, la información que declaran los estudiantes no es corroborada por fuentes secundarias (por ejemplo, otros servicios, ministerios, entre otros). Estas tareas manuales hacen que el proceso de la gestión de la asignación dure mucho tiempo, por lo que la prolijidad en la decisión podría verse afectada, ya que las fechas de cierre del proceso están definidas anteriormente por la institución y deben cumplirse.

Existen tareas que son llevadas a cabo por estos varios actores en el proceso de asignación del PRFE, que, aunque se conocen como prácticas que ya están instauradas, no se encuentran formalizadas como un proceso estandarizado. Esto genera atrasos y duplicidad de funciones que se resuelven al paso, puesto que se deben cumplir con los plazos establecidos por la institución.

³ Fuente: <https://www.ips.gob.cl/servlet/internet/content/1421810829144/chile-solidario>

⁴ JUNAEB (2005). Obtenido de Sistema Nacional de Asignación con Equidad para Becas JUNAEB: https://www.junaeb.cl/wp-content/uploads/2013/02/libro_junaeb.pdf

4.4 Cuantificación del Problema u Oportunidad

En primer lugar, se conoce que cada programa o beca cuenta con un presupuesto programado anual, así como una cantidad de cupos máximos que deben ser asignados (Ver Anexo 11.3). Esto es importante de mencionar, ya que hay que tener en cuenta que no todos los estudiantes que postulan al beneficio serán asignados.

Para el año 2018 se programó un total de 7.436 cupos para asignar a nivel nacional, de los cuales fueron realmente asignados 7.308 estudiantes, lo que equivale al 98,3% de la cobertura programada. Cabe indicar que no se llega al 100% debido a que existen estudiantes que egresan del programa por las causales descritas en el Anexo 11.1.3.

Para evidenciar el comportamiento de los beneficiarios en el año 2018, se individualizarán todas las variables de cada una de las dimensiones del modelo de focalización con la cantidad de estudiantes en cada categoría.

Se desplegarán inicialmente las variables declarativas ordenadas por dimensión, ya que estas al ser ingresadas por el estudiante no se contrastan ni con ninguna fuente de datos externas que permita confirmar esta información. Solamente se comprueba mediante la revisión documental que se hace para los casos que se considera necesario.

Inicialmente, se mostrará el número de estudiantes becados asociados al promedio de notas del año anterior declarada en la dimensión “Rendimiento Académico”:

Tabla 7: Promedio de Notas año anterior por Becado
Fuente: Elaboración propia

NOTA	Nº BECADOS	PESO RELATIVO
6.5 – 7.0	603	8%
6.0 – 6.4	1657	23%
5.5 – 5.9	2116	29%
5.0 – 5.4	2011	28%
4.5 – 4.9	686	9%
4.0 – 4.4	150	2%
Menor que 4.0	43	1%
Sin nota	42	1%
TOTAL	7308	100%

Cabe mencionar que las notas más altas son las que entregan mayor puntaje.

Asimismo, se mostrará el número de estudiantes becados asociados al último curso aprobado por la madre y el padre, y la pertenencia a pueblos originarios declarada en la dimensión “Vulnerabilidad”.

Hay que destacar que solo los estudiantes que pertenecen a pueblos originarios, se les otorga puntaje sino su puntaje será cero.

Tabla 8: Último curso aprobado por la madre
Fuente: Elaboración propia

ESCOLARIDAD MADRE	BECADOS	PESO RELATIVO
No tiene estudios	114	2%
Enseñanza Básica Incompleta	2406	33%
Enseñanza Básica Completa	1858	25%
Enseñanza Media Incompleta	828	11%
Enseñanza Media Completa	1734	24%
Estudios Superiores Incompletos	60	1%
Estudios Superiores Completos	308	4%
TOTAL	7308	100%

Tabla 9: Último curso aprobado por el padre
Fuente: Elaboración propia

ESCOLARIDAD PADRE	BECADOS	PESO RELATIVO
No tiene estudios	382	5%
Enseñanza Básica Incompleta	2231	31%
Enseñanza Básica completa	2215	30%
Enseñanza Media Incompleta	780	11%
Enseñanza Media Completa	1348	18%
Estudios Superiores Incompletos	44	1%
Estudios Superiores Completos	308	4%
TOTAL	7308	100%

La tabla se encuentra ordenada de forma descendente según el puntaje asociado. Por ejemplo, las variables escolaridad de la madre y del padre, la categoría “No tiene estudios” otorga un mayor puntaje.

Tabla 10: Pertenencia a pueblos originarios
Fuente: Elaboración propia

ETNIA	BECADOS	TOTAL	PESO RELATIVO
	2681	7308	37%

Ahora, se mostrarán las variables denominadas “desplazamiento desde el domicilio particular al lugar de estudio”, “medio de transporte” y “costo de movilización” de la dimensión de “Aislamiento Geográfico”. Como se evidencian en las tablas a

continuación, mientras más haya que desplazarse o el costo sea más elevado el puntaje asociado será mayor.

Tabla 11: Desplazamiento desde el domicilio particular al lugar de estudio

Fuente: Elaboración propia

DESPLAZAMIENTO AL LUGAR DE ESTUDIO	BECADOS	PESO RELATIVO
Fuera de la Región	279	4%
Fuera de la Provincia	1032	14%
Fuera de la Comuna	2557	35%
Es comuna con dificultad de acceso	3141	43%
Menos de 1 hora de traslado	223	3%
Sin Información	76	1%
TOTAL	7308	100%

Tabla 12: Medio de Transporte utilizado para desplazarse

Fuente: Elaboración propia

MEDIO DE TRANSPORTE	BECADOS	PESO RELATIVO
Avión o avioneta (financiada por terceros)	39	1%
Lancha, bote, barcaza	550	8%
A pie (más de 10 Kilómetros)	141	2%
A pie y micro o bus	3204	44%
Transporte municipal	203	3%
En micro o bus (más de 10 km)	2881	39%
En vehículo particular	191	3%
Otro (Caballo, bicicleta, carretera, etc)	22	0,3%
Sin Información	77	1%
TOTAL	7308	100%

Tabla 13: Costo de movilización

Fuente: Elaboración propia

COSTO DE MOVILIZACIÓN	BECADOS	PESO RELATIVO
Es de más de 2.000 pesos diarios	4339	59%
Es de 1.001 a 2000 pesos diarios	1467	20%
Es menor igual a 300 pesos diarios	46	1%
Es de 501 a 1.000 pesos diarios	878	12%
Es de 301 a 500 pesos diarios	189	3%
No tiene costo	313	4%
Sin Información	76	1%
TOTAL	7308	100%

Las únicas variables que se cruzan con fuentes secundarias son “Chile Solidario” y “SINAE” de la dimensión de “Vulnerabilidad”. Las siguientes tablas muestran la información de los becados según sus categorías.

Tabla 14: Chile Solidario
Fuente: Elaboración propia

CHILE SOLIDARIO	BECADOS	PESO RELATIVO
Familia inscrita en Chile Solidario o IEF	2981	41%
Familia NO inscrita en Chile Solidario o IEF	4327	59%
TOTAL	7308	100%

Tabla 15: SINAIE
Fuente: Elaboración propia

SINAIE	BECADOS	PESO RELATIVO
Primera Prioridad	4680	64%
Segunda Prioridad	366	5%
Tercera Prioridad	719	10%
Mal rendimiento en Ed.Media	480	7%
Sin cálculo de prioridad	3	0,04%
Sin Información	1060	15%
TOTAL	7308	100%

Esta información muestra que la mayoría de los estudiantes becados en el PRFE son estudiantes vulnerables según esas variables. Pero es importante dimensionar la magnitud de estudiantes becados que cuentan con el Registro Social de Hogares y se encuentran en los tramos más vulnerables. A continuación, se muestra una tabla resumen de los becados en el 2018 individualizados por tramo socioeconómico del RSH.

Tabla 16: Becados según tramo RSH
Fuente: Elaboración propia

TRAMO	BECADOS	PESO RELATIVO
Sin RSH	169	3%
tramo 40%	5861	80%
tramo 50%	561	8%
tramo 60%	302	4%
tramo 70%	192	3%
tramo 80%	111	2%
tramo 90%	98	1%
tramo 100%	14	0%
Total general	7308	100%

Al revisar los resultados, se demuestra que un 80% de los becados 2018 se encuentran en el tramo del 40% que indica que su calificación socioeconómica se posiciona dentro de ese tramo por ende son los más vulnerables. Por otra parte, es relevante indicar que 423 estudiantes quedaron en estado ingresado, es decir no fueron asignados debido a que no obtuvieron el puntaje necesario para la obtención de este beneficio.

Otra información importante por resaltar está asociada con el tiempo que se dedica para la ejecución de las actividades del proceso de gestión de asignación del PRFE. Estas actividades están asociadas a un calendario que elabora la encargada nacional del programa. La siguiente tabla especifica estas tareas junto con el tiempo en días que son necesarios para su ejecución.

Tabla 17: Actividades del proceso de asignación

Fuente: Elaboración propia

NO.	ACTIVIDADES	DIAS	RESPONSABLES
1	Consolidar información, Recepción y revisión documentación	57	Entidad Ejecutora/Hogar
2	Validación de formularios de postulantes y renovantes	62	Unidad de Becas Regional
3	Constitución y sesión Comisión Regional de Selección	27	Integrantes de Comisión
4	Publicación de Resultados	2	Departamento de Becas

La actividad que toma mayor tiempo en realizarse es la Validación de formularios de postulantes y renovantes, pero hay que tener en cuenta que las actividades uno y dos se ejecutan en paralelo.

CAPÍTULO 5: PROPUESTA DE REDISEÑO DE PROCESOS

Este capítulo detallará la propuesta de rediseño del proceso de asignación del Programa de Residencia Familiar Estudiantil (PRFE), indicando el alcance de esta propuesta, cómo debe ser el nuevo diseño del proceso y finalmente, la lógica de negocio a implementar.

5.1 Direcciones de Cambio y Alcance

La propuesta del rediseño del proceso de asignación de becas o asistencialidad estudiantil implica el uso de la minería de datos empleando modelos predictivos que ayuden a determinar los potenciales beneficiarios del Programa de Residencia Familiar Estudiantil (PRFE). Se pretende determinar un “perfil de estudiante idóneo” que cumpla con las características normativas de este programa, con el fin de optimizar las tareas relacionadas con la decisión final asignación. Se ajustará, además, el modelo de focalización en el que se propone el uso del índice de aislamiento territorial descrito en la sección 2.2.3 del presente documento. Las variables de rediseño necesarias para llevar a cabo a esta propuesta se detallan en la siguiente tabla.

Tabla 18: Variables de rediseño de la propuesta
Fuente: Elaboración propia

VARIABLE	AS-IS	TO-BE	LÓGICA DE NEGOCIO
Anticipación	No existe proceso claramente definido al desagregar el Macroproceso: <u>1.1.2.3. Decisión de la Asignación para el pago de Becas o Asistencialidad Estudiantil</u>	- Modelar proceso de asignación del programa con el modelo predictivo de los potenciales beneficiarios. - Crear un perfil de beneficiario idóneo con las características del programa.	Aplicación de la minería de Datos para predecir a los potenciales beneficiarios obteniendo las variables con mayor grado de importancia para formular un perfil de beneficiario idóneo para el programa.
Coordinación	- No existe proceso claramente definido al desagregar el Macroproceso: <u>1.1.2.3. Decisión de la Asignación para el pago de Becas o Asistencialidad Estudiantil</u> , ya que solamente se usa el manual de procedimientos para la asignación. - No existe un proceso formal de colaboración para la confirmación con fuentes secundarias de las variables declarativas.	Modelo de Focalización ajustado según la importancia de las variables determinado por el modelo predictivo y sujeto al cumplimiento de requisitos del programa.	Variables de tipo: - <i>Regla</i> : reglas asociadas al modelo de focalización ajustado para el cumplimiento de requisitos del programa. - <i>Colaboración</i> : proceso formal de colaboración para la obtención de información de las fuentes secundarias para la correcta asignación del programa. En este caso, se utiliza información de la SUBDERE relacionadas con el índice de aislamiento territorial.

Mantenición consolidada de estado	Existe asimetría de información, no existen una fuente única que entregue datos confiables. Hay que buscarlas en distintas fuentes y la comunicación entre macroprocesos no es constante.	Se agrega una tarea de almacenamiento al nuevo modelo el cual tendrá un impacto tecnológico para que se pueda proveer de información a todos los actores de este proceso.	Entrega de información mediante reportes. Se agregarán reglas de negocio de tipo <i>pauta</i> .
Prácticas de trabajo	Las tareas del proceso de decisión de la asignación no se encuentran estandarizadas ni semiautomatizadas.	Se establece un nuevo modelo de focalización estandarizado para todos los programas de residencias.	Modelación en notación BPMN del proceso asignación incluyendo el nuevo modelo de focalización.

5.2 Diseño Detallado de Procesos TO BE

El proceso de la Macro 1.1.2.3, muestra el subproceso de asignación de Becas o Asistencialidad Estudiantil, el cual será rediseñado según la siguiente propuesta.

5.2.1 Asignación de Becas o Asistencialidad Estudiantil

Rediseñar este proceso implica que los subprocesos *Consolidar postulantes y renovantes la Beca o Asistencialidad Estudiantil* se convierta en el subproceso de nombre **Desarrollo de modelos predictivos de becados**, el cual contempla un cambio en la lógica del modelo de focalización del PRFE. Además, el subproceso *Planificar la entrega de Beca o Asistencialidad Estudiantil* tendrá algunas modificaciones en el desarrollo de las actividades de este. Al desagregarlo entonces al cuarto nivel se visualizará la propuesta de rediseño y los procesos involucrados que se destacarán en color amarillo y se describirán a continuación:

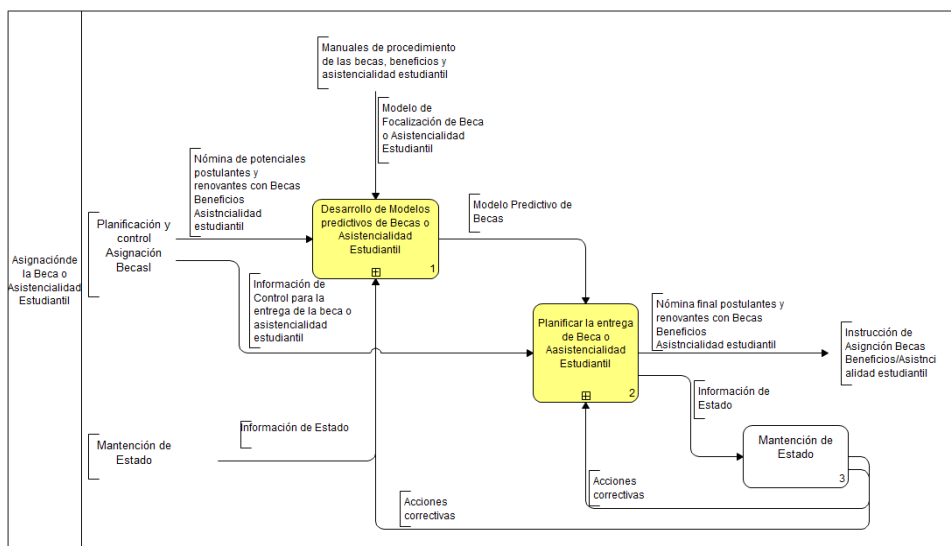


Ilustración 20: Macro 1.1.2.3.1 – Asignación de la Beca o Asistencialidad Estudiantil
Fuente: Elaboración propia

5.2.1.1 Desarrollo de Modelos predictivos de Becas o Asistencialidad Estudiantil

Es el proceso en el que se desarrollarán los modelos predictivos de becas o asistencialidad estudiantil, iniciando por el modelo del PRFE. Este servirá como insumo para la planificación de la entrega de este beneficio.

Para efectos de ejemplificar el comportamiento de este proceso y, como el alcance de este rediseño afecta directamente al Programa de Residencia Familiar Estudiantil (PRFE), se mencionarán los subprocesos, tareas, insumos y resultados relacionados con este beneficio.

Este proceso utilizará como insumo la nómina de estudiantes que postularon o renovaron al PRFE, para obtener una base de datos que servirá para entrenar y probar el modelo predictivo desarrollado.

Como regla de negocio, deberá contemplar el modelo de focalización del PRFE descrito anteriormente en este documento.

A continuación, se ilustra con un diagrama BPMN y se describen los subprocesos involucrados:

- **Preparar Datos** contiene las tareas de extracción y verificación de los datos que serán utilizados para la construcción de los modelos predictivos, las cuales deben consolidar todos los datos relacionados con los postulantes y renovantes, de manera que estén disponibles como DataSet, en el formato requerido para los modelos.
- **Desarrollo de modelos predictivos** actividades relacionadas con la creación de los modelos predictivos para la asignación de los beneficios. Este proceso se ejecutará cada vez que se requiera un modelo nuevo y para cada una de las becas o asistencialidad estudiantil vigentes. Contiene scripts que corren algoritmos con modelos predictivos.
- **Evaluar modelos predictivos**, es el que evaluará los modelos obtenidos y que deberá realizarse periódicamente para ajustar cualquier componente o cambios que se presenten con los datos. Este contiene una tarea de tipo script, la cual busca mostrar el desempeño de cada uno de los modelos ejecutados en el subproceso anterior.
- **Seleccionar mejor modelo predictivo**, es el que contiene tareas asociadas con la selección y obtención de resultados del modelo predictivo que cumpla con las características del PRFE.

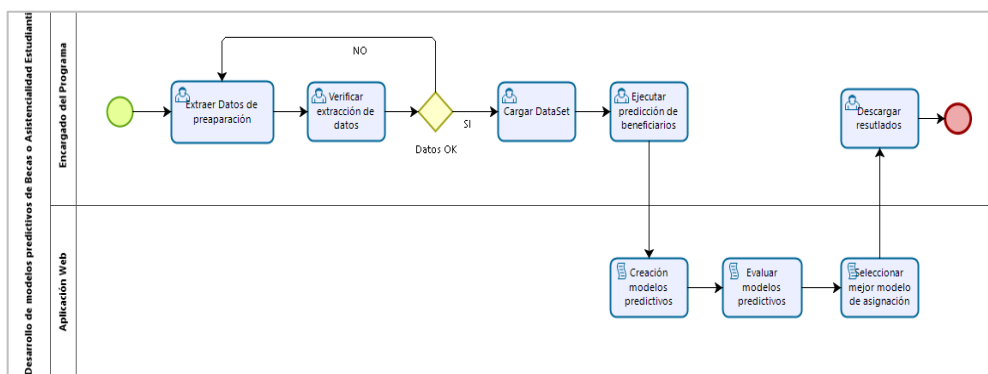


Ilustración 21: Modelo BPMN – Desarrollo de modelos predictivos de Becas o Asistencialidad Estudiantil

Fuente: Elaboración propia

5.2.1.2 Planificar la entrega de Beca o Asistencialidad Estudiantil

En este proceso se desarrollarán todas las actividades imprescindibles para generar la instrucción final de asignación de las becas o asistencialidad estudiantil, utilizando el insumo “Modelo Predictivo” que resulta del proceso anterior.

Inicia con la tarea de predicción de los beneficiarios, según el modelo predictivo desarrollado en el proceso anterior. Como producto, se obtendrá una combinación de predictores y variables, que resultan del modelo seleccionado, para identificar al grupo de postulantes y renovantes que efectivamente serán beneficiarios del PRFE. El nombre de este producto se define como perfil del beneficiario, el cual servirá de insumo para realizar la actividad de validación de los postulantes y renovantes.

Luego de esto, se genera una nómina final de beneficiarios postulantes y renovantes al beneficio, el cual permitirá que los responsables del proceso planifiquen eficientemente la entrega del beneficio. Para este caso, involucra la convocatoria y aceptación de las familias tutoras que acogerán a los estudiantes beneficiarios del programa durante el año. También, abarca la planificación operativa del programa, desde el pago de los aportes a las familias tutoras y Entidades Ejecutoras hasta los planes de supervisión y control necesarios para dar correcto cumplimiento a la normativa vigente del beneficio.

Es así como el encargado del programa construirá finalmente una instrucción de asignación, que contendrá la nómina definitiva de beneficiarios (quienes cumplan con el perfil predictor y hayan sido validados anteriormente) y el instrumento administrativo que la institución designe para estos efectos.

Desagregando este proceso, se evidencia la arquitectura del macroproceso 1.1.2.3.1.2 ilustrada a continuación:

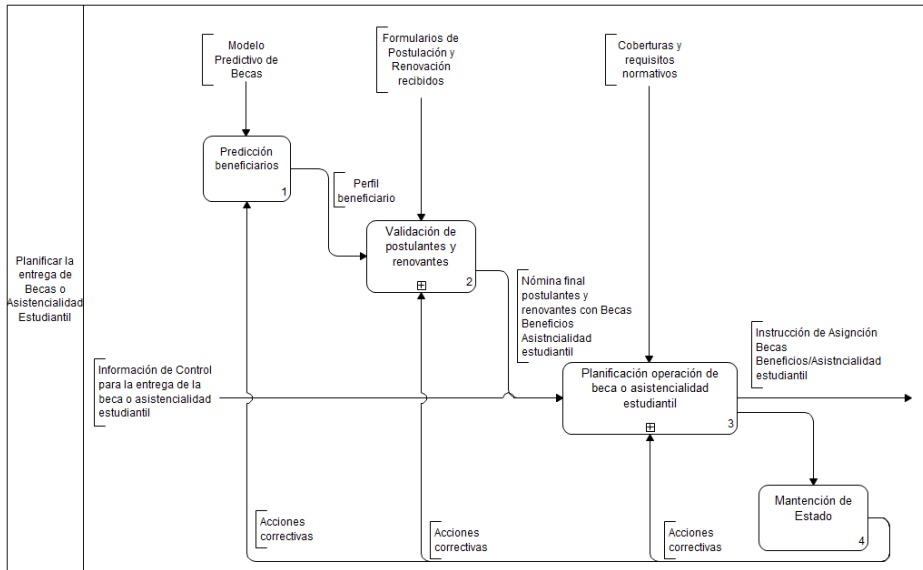


Ilustración 22: Macro 1.1.2.3.1.2 – Planificar la entrega de Beca o Asistencialidad Estudiantil
Fuente: Elaboración propia

Para poder explicar los subprocesos involucrados, se visualizarán en diagramas BPMN los que se indican a continuación:

- Predicción beneficiarios**, es el proceso que utiliza los modelos predictivos confeccionados, cuyo resultado debe arrojar un perfil idóneo de estudiante que se podría convertir en beneficiario, aplicando el modelo de focalización vigente. Esto se realiza al inicio del año con información histórica de beneficiarios. Entonces, el proceso inicia cuando se carga la nómina de potenciales postulantes y renovantes (por ejemplo, los estudiantes que obtuvieron el beneficio el año anterior). Con esta base de datos, se corren los modelos predictivos escogidos por su mejor rendimiento y se revisan los resultados obtenidos. Si todo está correcto, se extrae el perfil de los potenciales beneficiarios para continuar con los siguientes subprocesos.

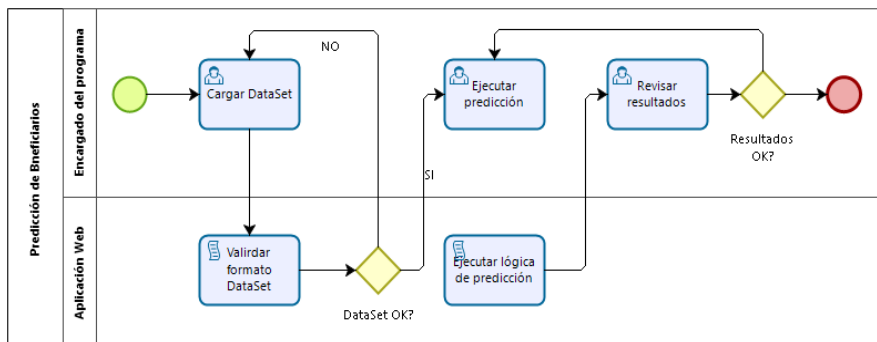


Ilustración 23: Modelo BPMN – Predicción de Beneficiarios
Fuente: Elaboración propia

- Validación de postulantes y renovantes**, proceso en el que se constituye una comisión de asignación regional, la cual se encarga de validar los formularios que fueron ingresados dentro del plazo establecidos por el programa y que, con el perfil del beneficiario ya definido en el proceso precedente, se construye la nómina final de postulantes y renovantes al beneficio mediante el cambio de estado en el sistema informático de becas de la institución.

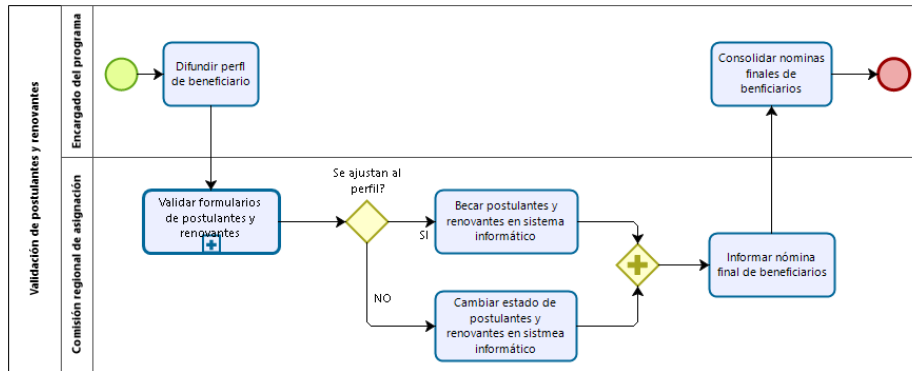


Ilustración 24: Modelo BPMN – Validación de postulantes y renovante
Fuente: Elaboración propia

- Planificación de operación del programa**, es el proceso en el que se planifica la operación del programa, es decir, se realizan los planes de convocatoria y aceptación de las entidades ejecutoras, quienes, a su vez, convocan a las familias tutoras que acogerán a los estudiantes beneficiarios del programa. Asimismo, se programan los aportes monetarios que se entregan a las mismas entidades y se elaboran los planes de supervisión a las casas de las familias tutoras. Finalmente, el producto resultante de este proceso debe ser la instrucción final de asignación del beneficio rigiéndose bajo la normativa vigente del programa.

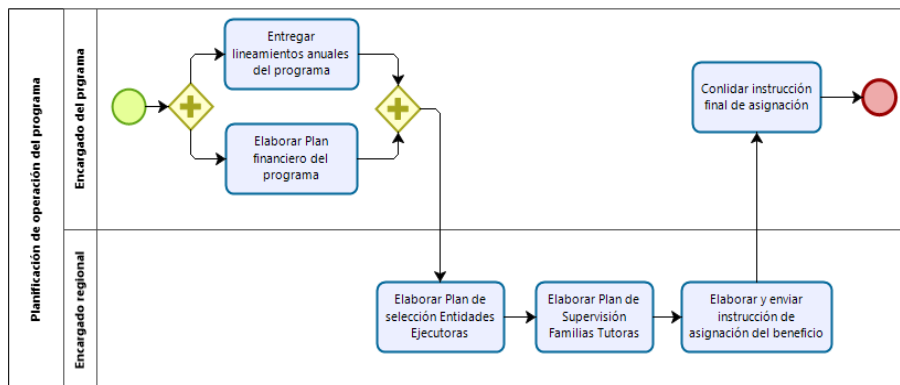


Ilustración 25: Modelo BPMN – Planificación de operación del programa
Fuente: Elaboración propia

5.2.1.3 Mantención de Estado

Es el proceso que recibe y proporciona información relevante para las mejoras que se requieran aplicar a los modelos predictivos o en su defecto, a los planes de entrega de Becas o Asistencialidad estudiantil.

5.3 Diseño de Lógica de Negocios

En esta sección se describirá en detalle la lógica de negocios propuesta para el rediseño del proceso asignación de Becas o Asistencialidad Estudiantil, explicando los modelos predictivos utilizados, los resultados de rendimiento de cada uno de ellos y la selección de el o los mejores modelos predictivos que se ajusten a la solución de esta problemática. Se propone entonces, utilizar la herramienta Jupiter book, para el desarrollo de los modelos con el lenguaje de programación python. Al mismo tiempo, se propone simplificar el modelo de focalización debido a que se conoce que las fuentes de información secundarias⁵ que se pueden utilizar contienen información socioeconómica y de vulnerabilidad que pueden ser utilizadas directamente sin necesidad de solicitar al estudiante que la declare. Por lo anterior, la propuesta de un modelo de focalización se materializaría según se muestra en la siguiente tabla:

Tabla 19: Modelo de focalización PRFE simplificado
Fuente: Elaboración propia

DIMENSIÓN DE VULNERABILIDAD	
TRAMO RSH	
INDICADOR (Precargado)	Puntaje
0% - 40%	300
41% - 50%	250
51% - 60%	200
61% - 70%	150
71% - 80%	100
81% - 90%	50
91% - 100%	25

PERTENENCIA A PUEBLOS ORIGINARIOS	
INDICADOR (Precargado)	Puntaje
Mapuche	100
Aymara	100
Yagan	100
Kawhaskar	100
Rapa Nui	100
Atacameño	100
Colla	100
Quechua	100
Diaguita	100
Otro	100
Ninguno (Sin Etnia)	50

DIMENSIÓN RENDIMIENTO ACADÉMICO	
RENDIMIENTO ACADÉMICO	
INDICADOR (Precargado)	Puntaje
6.5 - 7.0	125
6.0 - 6.4	110
5.5 - 5.9	95
5.0 - 5.4	80
4.5 - 4.9	65
4.0 - 4.4	50
Menor que 4.0	40

DIMENSIÓN DE AISLAMIENTO GEOGRÁFICO	
INDICADOR (Precargado según Índice de Aislamiento Territorial)	Puntaje
Comuna Aislada	600
Comuna no Aislada	0

⁵ Fuentes de información secundarias: Registro Social de Hogares (RSH) del Ministerio de Desarrollo Social e Índice de Aislamiento Territorial de la Subsecretaría de Desarrollo Regional y Administración (SUBDERE)

5.3.1 Análisis de los Datos

Para poder aplicar el modelo predictivo para la asignación de la beca o Programa, se debe contar con datos limpios. Los datos necesarios son los asociados a las variables de la focalización PRFE con el puntaje correspondiente, así como la dirección de su domicilio familiar y de estudios. Cada registro constituye a un estudiante que en el año 2018 postuló o renovó a este beneficio y el resultado (becado o no becado), es un dato más de la nómina. La extracción de estos datos se realizó desde el sistema informático disponible en la institución y de fuentes de información secundarias. Estos datos conforman una nómina de estudiantes que obtuvieron el beneficio en el año 2018, la cual cuenta con 13.562 registros. Entonces, para obtener el Dataset inicial se toman los siguientes campos que se describen en la tabla a continuación:

Tabla 20: Descripción del Dataset PRFE
Fuente: Elaboración propia

NOMBRE CAMPO	DESCRIPCIÓN CAMPO	TIPO
COD_FORMULARIO	Código único identificador	Integer
DESC_ESTADO	Estado de Beneficio: Becado: que obtuvo el beneficio. No Becado: no se adjudicó el beneficio. Suspensión: el beneficio se detiene pero se reserva el cupo para el siguiente año, esto es solicitado por el beneficiario. Titulado/Egresado: el beneficiario estuvo becado, pero solicita retiro por ende se debe egresar del programa.	String
DESC_ETNIA	Pueblo originario seleccionado por el estudiante	String
NOT_DESCRIPCION	Nota del año anterior declarada por el estudiante	Float
CURSO_MADRE	Último grado académico cursado por la madre declarada por el estudiante.	String
CURSO_PADRE	Último grado académico cursado por el padre declarada por el estudiante	String
UBICORLUGEST_DESC	Ubicación de la vivienda con respecto al lugar de estudios declarado por el estudiante.	String
MEDIOTRANSPORTE_DESC	Medio de transporte utilizado para la movilización hacia el lugar de estudios declarado por el estudiante.	String
COSTOMEDIOTRANSPORTE_DESC	Costo en el que incurre el estudiante para poder movilizarse hacia el lugar de estudios declarado por el estudiante.	String
PUNTAJE	Suma del puntaje por cada variable del modelo de focalización del programa.	Float
COMPOSICION_GRUPO_FAMILIAR	Composición del Hogar: Parentesco con el Jefe (a) de Hogar	String
PADRE_EN_HOGAR	Padre en el hogar: (0 = NO; 1 = SI)	Boolean
GRUPFAMI_NUMEROINT	Número de integrantes en el grupo familiar	Integer
ESTUDIANTE_EN_RND	Estudiante se encuentra en el registro nacional de discapacidad: (0 = NO; 1 = SI)	Boolean
HOGAR_MONOPARENTAL	Hogar formado por un progenitor (madre o padre) y uno o varios hijos: (0 = NO; 1 = SI)	Boolean
UNIPERSONAL	Hogar compuesto por una sola persona: (0 = NO; 1 = SI)	Boolean
DS_TRAMO_CSE	Tramo del 40 (0% – 40%): Hogares calificados en el 40% de menores ingresos o mayor vulnerabilidad Tramo del 50(41% – 50%): Hogares calificados entre el 41% y el 50% de menores ingresos o mayor vulnerabilidad Tramo del 60(51% – 60%): Hogares calificados entre el 51% y el 60% de menores ingresos o mayor vulnerabilidad Tramo del 70(61% – 70%): Hogares calificados entre el 61% y el 70% de menores ingresos o mayor vulnerabilidad Tramo del 80(71% – 80%): Hogares calificados entre el 71% y el 80% de menores ingresos o mayor vulnerabilidad Tramo del 90(81% – 90%): Hogares calificados entre el 81% y el 90% de menores ingresos o mayor vulnerabilidad Tramo del 100(91%-100%): Hogares calificados en el 91% de menores ingresos o mayor vulnerabilidad	String
ID_TRAMO_CSE	Identificador del tramo de calificación socioeconómica del hogar	Integer
AGRUPACION_HABITACIONAL_LOC	Localidad vivienda	String
NIVEL	Nivel de estudios del estudiante	String
ESTABLECIMIENTO_EDUCACIONAL	Nombre del Establecimiento Educativo en el que se encuentra cursando estudios regulares el estudiante	String
REGION_EST_EDU	Región del Establecimiento Educativo	Integer
COMUNA_EST_EDU	Comuna Establecimiento Educativo	String
DOMICILIO_EST_EDU	Dirección Establecimiento Educativo	String
REGION_DOM_FAMILIAR	Región domicilio familiar	Integer
COMUNA_FAMILIAR	Comuna Domicilio Familiar	String
INDICE AISLAMIENTO TERRITORIAL	Índice de aislamiento Territorial entre -1.0 y 0.9 (excluyendo el 0), se considera como "Comuna Aislada". Índice de Aislamiento Territorial mayor a 0.9 o igual a 0, se considera como "Comuna no aislada".	String

Cabe mencionar que la variable objetivo será DESC_ESTADO la cual identifica si el estudiante estuvo o no becado en el programa.

Continuando con la revisión de los datos, se puede evidenciar variables relacionadas con la composición familiar del estudiante, su domicilio familiar, su información académica y su información económica. A continuación, se mostrarán gráficamente las variables más importantes y las cantidades que están representando con respecto al total de los registros del dataset.

- **Nivel**, variable de tipo académica que indica si el estudiante pertenece al nivel educacional básica o media o cursa estudios superiores. De acuerdo con el siguiente gráfico, en el 2018, el 61,25% estudiaron en el nivel básica o media, mientras que el 38,75% estudiaron en el nivel superior, tal como se muestra en la siguiente ilustración:

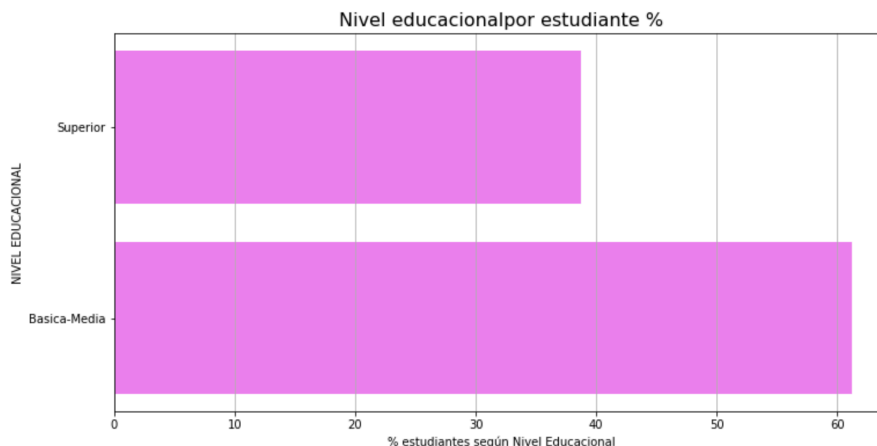


Ilustración 26: Nivel educacional por estudiante
Fuente: Elaboración propia

- **Etnia**, variable de tipo cultural, la cual muestra a qué pueblo originario pertenece el estudiante, según la información declarada por este. En el caso de no pertenecer a ningún pueblo originario, el estudiante escoge la opción “No aplica”. Esta última opción, es la que cuenta con el mayor peso relativo con el 58,68%, seguido de la etnia “Mapuche”, que representa un 32,07% de los registros, según se visualiza en el siguiente gráfico:

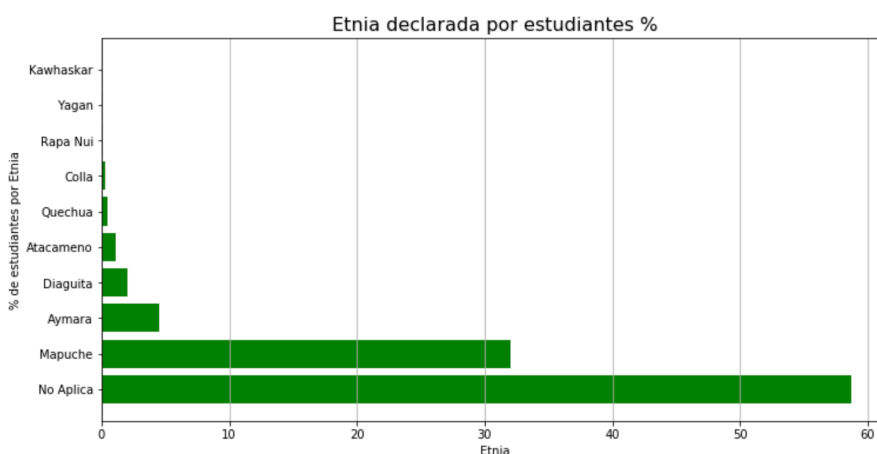


Ilustración 27: Pueblo originario declarado por el estudiante
Fuente: Elaboración propia

- **Curso Madre**, variable de tipo socioeconómica que indica que el 30,05% cursó como último grado académico, la enseñanza media completa. Esta información es señalada por el estudiante y en el caso de no tener información también se

debe declarar. La siguiente ilustración muestra los pesos relativos de cada una de las alternativas de esta variable.

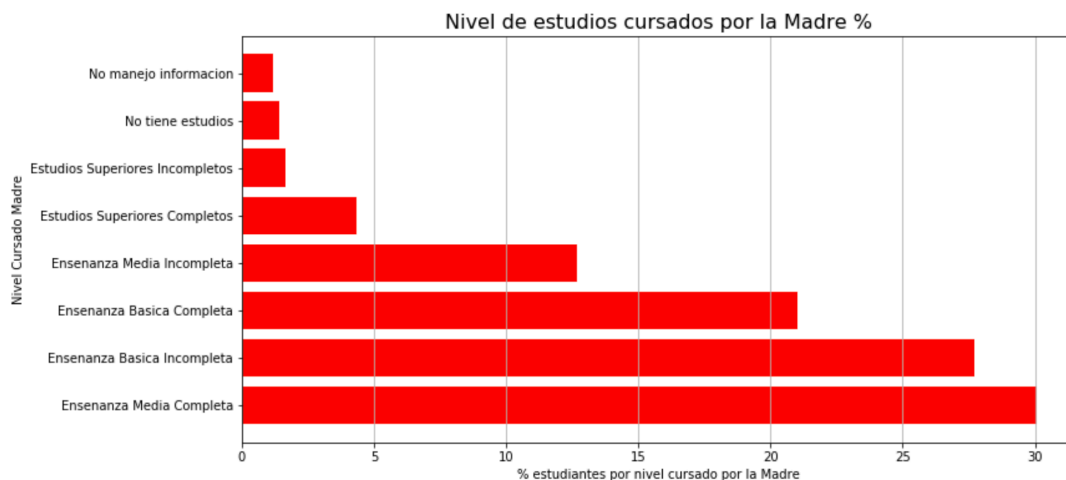


Ilustración 28: Último grado académico cursado por la madre del estudiante
Fuente: Elaboración propia

- **Curso Padre**, variable de tipo socioeconómica que indica que el 26,58% cursó como último grado académico, la enseñanza básica pero incompleta, tal como lo declara el estudiante al completar el formulario de postulación o renovación del programa. La siguiente ilustración muestra los pesos relativos de cada una de las alternativas de esta variable.

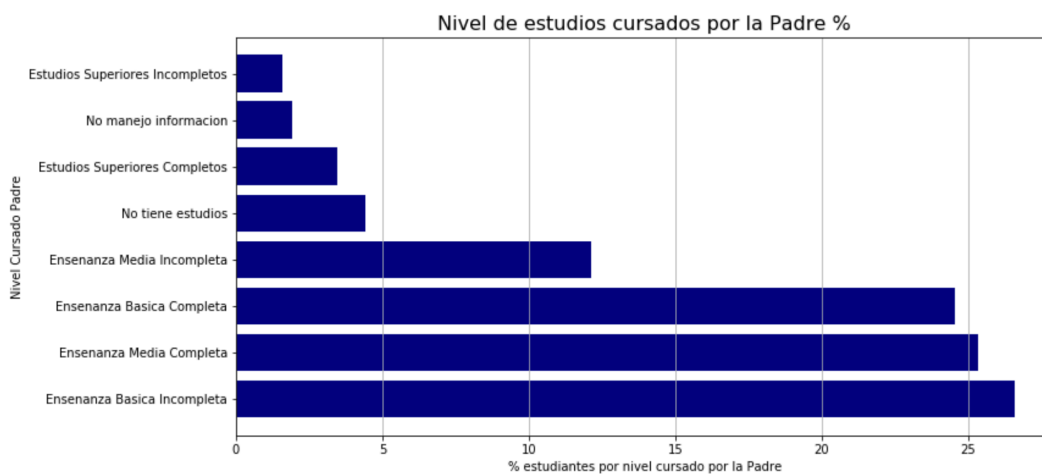


Ilustración 29: Último grado académico cursado por el padre del estudiante
Fuente: Elaboración propia

- **Región domicilio familiar**, variable de tipo territorial en la que se identifica la región de origen del estudiante. Según la ilustración a continuación, se muestra que el 21,05% provienen de la región del BíoBío (8), mientras que solamente el

1,28% de los estudiantes que postularon o renovaron al beneficio en el 2018, pertenecen a la región de Magallanes (12), tal como lo muestra la siguiente ilustración:

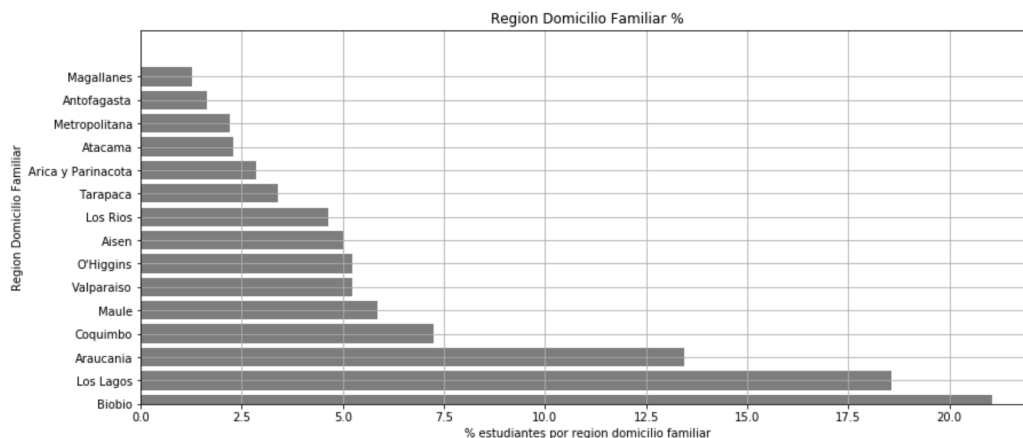


Ilustración 30: Región domicilio familiar
Fuente: Elaboración propia

- **Ubicación lugar de estudios**, variable de tipo territorial, la cual es declarativa y señala la ubicación del domicilio familiar del estudiante con respecto al lugar de estudios. Según los datos, el 32,56% de los postulantes o renovantes al programa corresponde a la opción “Es una comuna con dificultad de acceso”. El siguiente gráfico ilustra la proporción de las opciones correspondientes a esta variable.

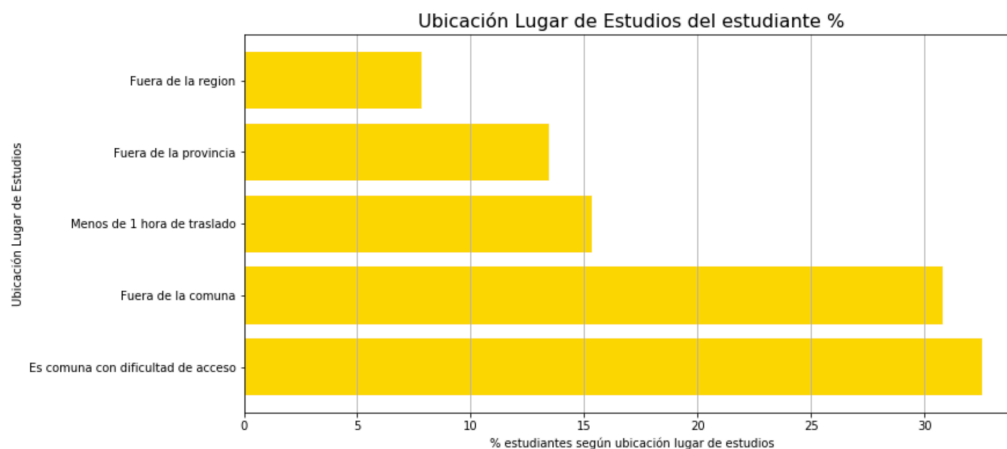


Ilustración 31: Ubicación de la vivienda con respecto al lugar de estudios
Fuente: Elaboración propia

- **Tramo RSH**, variable de tipo económica que proviene del registro social de hogares, que ubica a la población en tramos de vulnerabilidad, los datos del programa PRFE indican que el 75,99% de los postulantes o renovantes

pertencen al 40% de menores ingresos o mayor vulnerabilidad, tal como se evidencia en la siguiente ilustración.

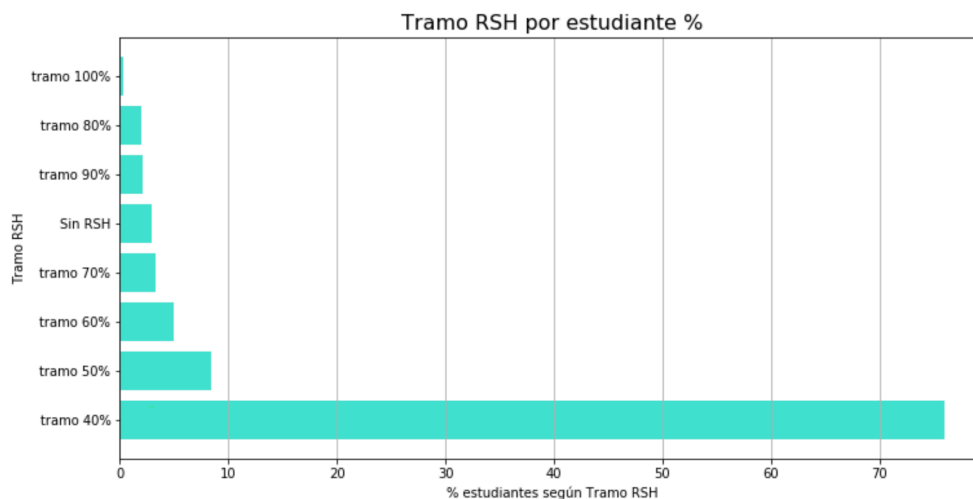
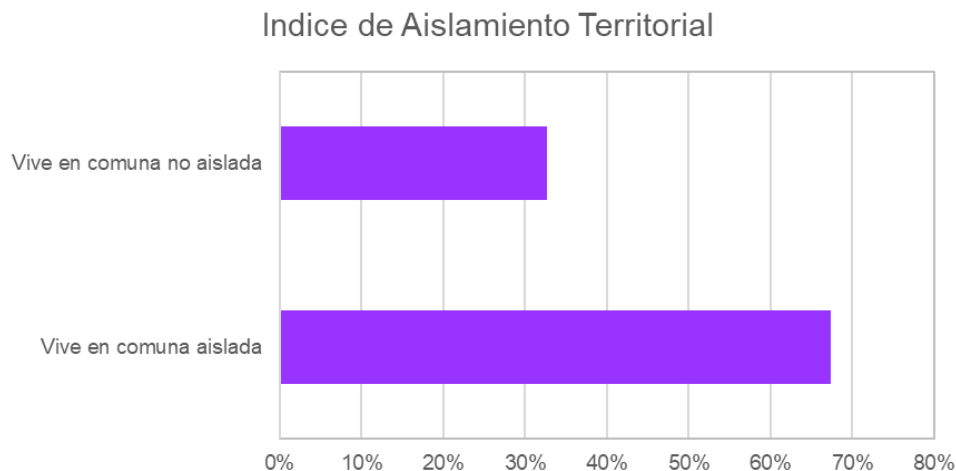


Ilustración 32: Tramo RSH por estudiante
Fuente: Elaboración propia

- **Índice de aislamiento territorial**, variable de tipo territorial proporcionada por la SUBDERE y se asocia según la comuna del domicilio familiar del estudiante, el cual identifica si el estudiante vive en una comuna aislada o no. Esto fue clasificado según la siguiente condición:
 - Índice de aislamiento Territorial entre -1.0 y 0.9 (excluyendo el 0), se considera como “Comuna Aislada”.
 - Índice de Aislamiento Territorial mayor a 0.9 o igual a 0, se considera como “Comuna no aislada”.

El siguiente gráfico muestra la proporción resultante de los postulantes y renovantes 2018 al programa.



Hay que destacar que se logró cruzar datos con información de fuentes secundarias como el Registro Social de Hogares (RSH) y de la SUBDERE para obtener datos más fidedignos, lo que provocó que se descartaran los que no eran comprobables.

5.3.2 Preparación de los Datos

Con el dataset ya constituido y con la comprensión de los campos más relevantes, el siguiente paso es preparar los datos para que se ajusten al modelo y predicción que se requieren para esta propuesta.

Para comenzar, se realiza una limpieza de los datos, en el que se eliminan los registros que cuenten con filas vacías, con caracteres especiales y que no pueden ser reemplazados por otro formato propio del lenguaje de programación Python. Como estos datos fueron extraídos de distintas fuentes y algunos campos son digitados por un usuario, es probable que se presenten este tipo de errores de datos. Hay que mencionar que solamente el 0,5% de los datos presentó este tipo de errores y fueron eliminados del dataset.

Inmediatamente, se procede a transformar los datos que consideramos relevantes como en caso de la variable dependiente DESC_ESTADO, la cual primero se convirtió en solo 2 opciones, ya que los otros valores (“Egresado” y “Suspensión”) se unificaron en el estado “No Becado”. Como siguiente paso, se transforma DESC_ESTADO a una variable de tipo boolean, en la que el valor “Becado” es 1 y el valor “No Becado” es 0. Asimismo, se transforma la variable INDICE DE AISLAMIENTO TERRITORIAL en 1 para los valores “Comuna aislada” y 0 para los valores “Comuna no aislada”.

Para finalizar con la preparación de los datos, debemos dejar seleccionadas las variables categóricas y las variables numéricas que serán utilizadas para la construcción de los modelos predictivos. Esto significa que no todas las variables categóricas y numéricas serán escogidas, ya que se busca predecir el estado del estudiante con la información utilizada para la asignación. La única variación es que para la dimensión territorial del modelo de focalización del programa, se agregan los datos del índice de aislamiento territorial.

En resumen, las variables que se utilizarán para los modelos predictivos son:

- **Variables Categóricas:** variables descriptivas y que son relevantes para la aplicación del modelo. Para el PRFE, son importantes las variables que

proporcione información socioeconómica, territorial y de pertenencia, según las dimensiones del modelo de focalización del proceso de asignación del beneficio. Las variables seleccionadas son:

Tabla 21: Variables Categóricas
Fuente: Elaboración propia

NOMBRE CAMPO	DESCRIPCION CAMPO	DIMENSIÓN
DESC_ETNIA	Pueblo originario seleccionado por el estudiante	Pertenencia
CURSO_MADRE	Último grado académico cursado por la madre declarada por el estudiante.	Socioeconómica
CURSO_PADRE	Último grado académico cursado por el padre declarada por el estudiante	Socioeconómica
UBICORLUGEST_DESC	Ubicación de la vivienda con respecto al lugar de estudios declarado por el estudiante.	Territorial
MEDIOTRANSPORTE_DESC	Medio de transporte utilizado para la movilización hacia el lugar de estudios declarado por el estudiante.	Territorial
COSTOMEDIOTRANSPORTE_DESC	Costo en el que incurre el estudiante para poder movilizarse hacia el lugar de estudios declarado por el estudiante.	Territorial
DS_TRAMO_CSE	Tramo del 40 (0% – 40%): Hogares calificados en el 40% de menores ingresos o mayor vulnerabilidad Tramo del 50(41% – 50%): Hogares calificados entre el 41% y el 50% de menores ingresos o mayor vulnerabilidad Tramo del 60(51% – 60%): Hogares calificados entre el 51% y el 60% de menores ingresos o mayor vulnerabilidad Tramo del 70(61% – 70%): Hogares calificados entre el 61% y el 70% de menores ingresos o mayor vulnerabilidad Tramo del 80(71% – 80%): Hogares calificados entre el 21% y el 30% de mayores ingresos o menor vulnerabilidad Tramo del 90(81% – 90%): Hogares calificados entre el 11% y el 20% de mayores ingresos o menor vulnerabilidad Tramo del 100(91%-100%): Hogares calificados en el 10% de mayores ingresos o menor vulnerabilidad	Socioeconómica
INDICE AISLAMIENTO TERRITORIAL	Índice de aislamiento Territorial entre -1.0 y 0.9 (excluyendo el 0), se considera como "Comuna Aislada". Índice de Aislamiento Territorial mayor a 0.9 o igual a 0, se considera como "Comuna no aislada".	Territorial
COMUNA_EST_EDU	Comuna Establecimiento Educacional	Territorial

- **Variables Numéricas:**

Son las variables numéricas relevantes para la aplicación del modelo predictivo. Para el PRFE, las variables NOT_DESCRIPCION, ESTUDIANTE_EN_RND y HOGAR_MONOPARENTAL, son las que se seleccionan para la ejecución del modelo.

Tabla 22: Variables Numéricas
Fuente: Elaboración propia

NOMBRE CAMPO	DESCRIPCION CAMPO	DIMENSIÓN
NOT_DESCRIPCION	Nota del año anterior declarada por el estudiante.	Académica
ESTUDIANTE_EN_RND	Estudiante se encuentra en el registro nacional de discapacidad: (0 = NO; 1 = SI)	Socioeconómica
HOGAR_MONOPARENTAL	Hogar formado por un progenitor (madre o padre) y uno o varios hijos: (0 = NO; 1 = SI)	Socioeconómica

- **Variable dependiente:**

Es la variable que describe el estado final del postulante o renovante al beneficio: "Becado" (1) y "No Becado" (0) que ya fueron transformadas anteriormente.

5.3.3 Modelamiento

Tal como se ha mencionado en el presente documento, se propone construir modelos predictivos para la definición final del mejor perfil de estudiante que debe ser beneficiado con el PRFE y ajustado con los requisitos del programa. Por lo anterior, se tiene distintos Dataset de entrenamiento y validación, ya que se modelará tanto aplicando el modelo de focalización vigente y el modelo de focalización simplificado propuesto.

Para este proyecto se generan 5 algoritmos de aprendizaje supervisado, los cuales serán descritos en las siguientes secciones.

5.3.3.1 Construcción de Modelos

Para iniciar se divide el Dataset en 2 grupos. Uno corresponde a los datos de entrenamiento y el otro grupo a los datos de validación. Para no sobreajustar el modelo se separan los datos en 2 oportunidades. La primera utilizando el 10% de los datos como validación y, a continuación, se utiliza un 20% de esos datos como de validación. Se desarrollan modelos de tipo supervisados, ya que como se ha explicado anteriormente, se requiere predecir a los estudiantes “Becados” y “No Becados”, utilizando clasificadores o en su defecto predictores.

Como primer modelo predictivo, se construye un modelo basado en *Regresión Logística*, el que muestra la siguiente una matriz de confusión:

Tabla 23: Matriz de confusión – Regresión Logística
Fuente: Elaboración propia

Observación	RL	Predicción	
		Becados	No Becados
Becados		1036	206
No Becados		343	856

Esta indica que 1.036 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 856 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 206 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 343 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Por otro lado, se obtienen los indicadores de rendimiento del modelo según la siguiente tabla:

Tabla 24: Indicadores de Rendimiento – Regresión Logística
Fuente: Elaboración propia

Rendimiento	Resultado
Exactitud (Accuracy)	77,51%
Precisión	75,13%

Este modelo entonces clasifica correctamente al 77,51% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 75,13% de los estudiantes.

Por último, la curva ROC que arroja el modelo muestra un área bajo la curva del 77,4% que indica que, en esa proporción, el modelo podrá distinguir las clases dadas.

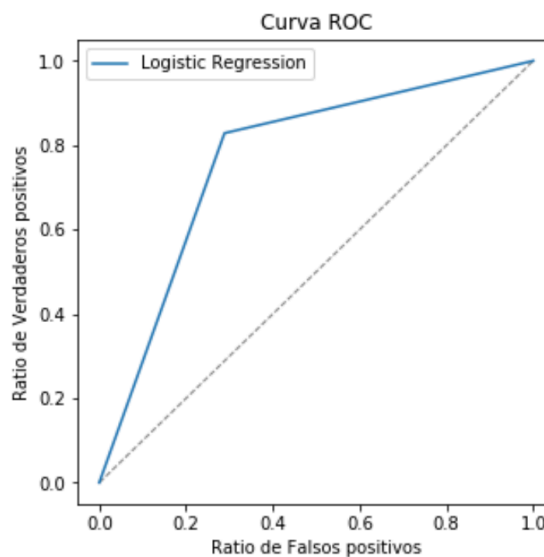


Ilustración 34: Curva ROC – Regresión Logística
Fuente: Elaboración propia

Como segunda opción, se construye el modelo de *K-Nearest Neighbour*, en el que se utiliza como parámetro la selección de 5 objetos ($k = 5$), cuyos resultados se explican con la matriz de confusión que se muestra a continuación:

Tabla 25: Matriz de confusión – K-NN
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
	Becados	970	206
	No Becados	423	833

Se evidencia que 970 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 833

estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 206 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 423 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

También, el resultado del rendimiento indica que este modelo clasifica correctamente al 74% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 71,01% de los estudiantes, según se muestra en la siguiente tabla:

Tabla 26: Indicadores de Rendimiento – K-NN
Fuente: Elaboración propia

Rendimiento	Resultado
Exactitud (Accuracy)	74,00%
Precisión	71,01%

Por otro lado, la curva ROC que arroja el modelo muestra un área bajo la curva del 77% que indica que, en esa proporción, el modelo podrá distinguir a los *Becados* y *No Becados*.

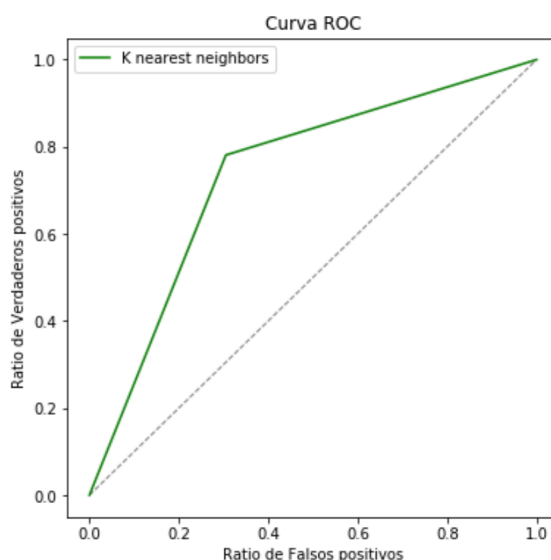


Ilustración 35: Curva ROC – K-NN
Fuente: Elaboración propia

El tercer modelo que se usa para la predicción de la asignación del PRFE es *Árboles de Decisión*. Con ese modelo se busca clasificar a los estudiantes *Becados* y *No Becados* mediante la determinación de los atributos y las decisiones que podrían resultar de esta predicción según las características entregadas (las variables categóricas y numéricas mencionadas anteriormente).

Es así como, los resultados obtenidos con respecto a la correcta predicción del atributo *Estado* se presentan en la siguiente matriz de confusión:

Tabla 27: Matriz de confusión – Árboles de Decisión
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
	Becados	1183	59
	No Becados	725	474

Revisando estos resultados, se indica que 1.183 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 474 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 59 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 725 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Para medir el desempeño de este modelo se calculan 2 indicadores que entregan el siguiente resultado:

Tabla 28: Indicadores de Rendimiento – Árboles de Decisión
Fuente: Elaboración propia

Rendimiento	Resultado
Exactitud (Accuracy)	67,88%
Precisión	62,00%

Este modelo clasifica correctamente al 67,88% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 62% de los estudiantes.

También, la curva ROC que arroja el modelo muestra un área bajo la curva del 67% que indica que, en esa proporción, el modelo podrá distinguir a los *Becados* y *No Becados*.

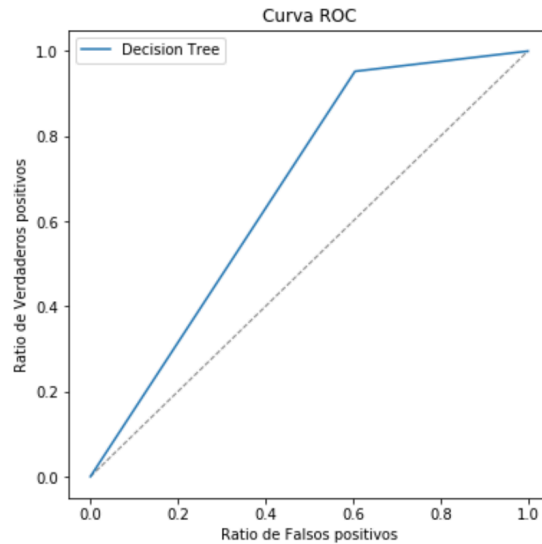


Ilustración 36: Curva ROC – Árboles de Decisión
Fuente: Elaboración propia

Para poder mejorar el desempeño de este modelo se utiliza la validación cruzada entregando como parámetro $cv = 3$, es decir, se divide el conjunto de datos de entrenamiento en 3 para después validar el modelo en 3 iteraciones. Para este caso, los resultados obtenidos son:

- La matriz de confusión mejorada indica que 1.000 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 670 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 242 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 529 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Tabla 29: Matriz de confusión – Árboles de Decisión ($cv=3$)
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
Becados		1000	242
No Becados		529	670

Este ajuste hace que el modelo clasifique correctamente al 68,4% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 65,4% de los estudiantes. Asimismo, la curva ROC que arroja el modelo muestra un área bajo la curva del

68,2% que indica que, en esa proporción, el modelo podrá distinguir a los *Becados* y *No Becados*.

Otra respuesta que se puede obtener son las variables más prioritarias según las clasifica este modelo. En la siguiente tabla se resumen las más importantes según el peso relativo entregado por el mismo modelo.

Tabla 30: Tabla de importancia de variables – Árboles de Decisión
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
UBICORLUGEST_DESC	"Menos de 1 hora de traslado"	50,77%
UBICORLUGEST_DESC	"Fuera de la región"	13,68%
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"	7,39%
HOGAR_MONOPARENTAL	1	7,07%
UBICORLUGEST_DESC	"Es comuna con dificultad de acceso"	3,25%
DS_TRAMO_CSE	"Sin RSH"	2,73%
COSTOMEDIOTRANSPORTE_DESC	"Es menor igual a 300 pesos"	2,39%
COMUNA_EST_EDU	"Temuco"	2,15%

Esto significa que un estudiante que haya declarado que la ubicación de su domicilio familiar con respecto a su lugar de estudios demora menos de 1 hora o es vive en una comuna con dificultad de acceso o vive fuera de la región de su lugar de estudios, que vive efectivamente en una comuna aislada (según el índice de aislamiento territorial), cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, que no haya inscrito su información en el registro social de hogares, que el costo del medio de transporte que utiliza para trasladarse a su lugar de estudios es menos a 300 pesos chilenos y que estudie en la comuna de Temuco, es asignado correctamente por el modelo como *Becado*.

Finalmente, el diagrama de árbol resultante se muestra en la siguiente ilustración:

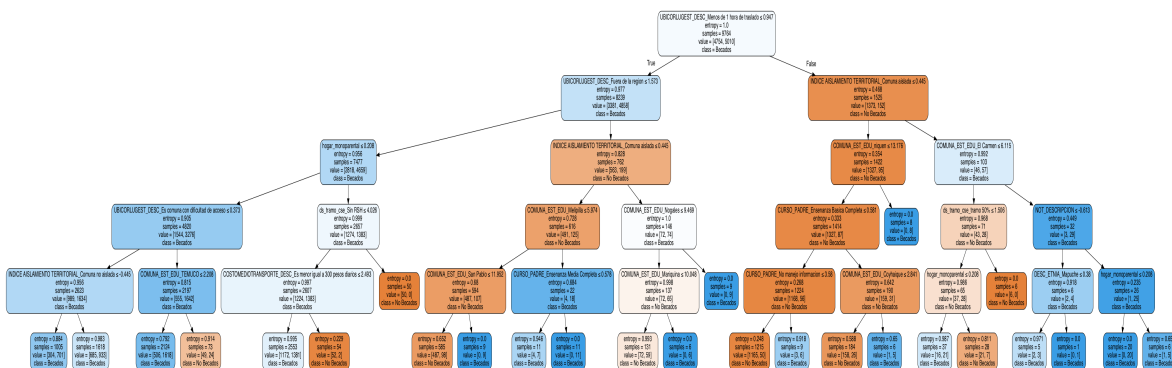


Ilustración 37: Diagrama de Árbol de Decisión

Fuente: Elaboración propia

Como cuarto modelo se construye un Random Forest (siguiendo la línea de la clasificación mediante árboles de decisión). En este caso el modelo genera los siguientes resultados:

- Número de árboles: 13
- Mejor Criterio de Evaluación: gini
- Mejor tamaño: 21
- Mejor tamaño hojas:

Además, presenta una matriz de confusión como se indica:

Tabla 31: Matriz de confusión – Random Forest
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
Becados		1052	190
No Becados		459	740

1.052 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 740 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 190 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 459 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*. También, para medir el desempeño de este modelo se calculan 2 indicadores que entregan el siguiente resultado:

Tabla 32: Indicadores de Rendimiento – Random Forest
Fuente: Elaboración propia

Rendimiento	Resultado
Exactitud (Accuracy)	73,41%
Precisión	69,62%

Este modelo clasifica correctamente al 73,41% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 69,62% de los estudiantes. Igualmente, la curva ROC que arroja el modelo muestra un área bajo la curva del 73,2% que indica que, en esa proporción, el modelo podrá distinguir a los *Becados* y *No Becados*.

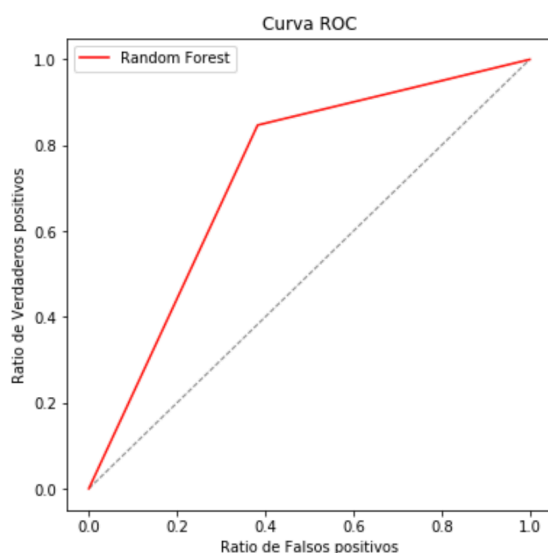


Ilustración 38: Curva ROC – Random Forest
Fuente: Elaboración propia

Otro indicador que se puede obtener son las variables más prioritarias según las clasifica este modelo. En la siguiente tabla se resumen las más importantes según el peso relativo entregado por el mismo modelo.

Tabla 33: Tabla de importancia de variables – Random Forest
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
UBICORLUGEST_DESC	"Menos de 1 hora de traslado"	7,99%
NOT_DESCRIPCIÓN		6,06%
INDICE AISLAMIENTO TERRITORIAL	"Comuna no aislada"	5,97%
COSTOMEDIOTRANSPORTE_DESC	"Es más de 2.000 pesos"	5,12%
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"	3,69%
HOGAR_MONOPARENTAL	1	3,40%
COSTOMEDIOTRANSPORTE_DESC	"Es menor igual a 300 pesos"	2,66%
UBICORLUGEST_DESC	"Fuera de la comuna"	2,34%
CURSO_MADRE	"Enseñanza media completa"	2,73%
CURSO_PADRE	"Enseñanza media completa"	2,15%

Es decir que, un estudiante que haya declarado que la ubicación de su domicilio familiar con respecto a su lugar de estudios demora menos de 1 hora o vive fuera de la comuna de su lugar de estudios, que vive efectivamente en una comuna aislada (según el índice de aislamiento territorial) o no vive en una comuna aislada, que el costo del medio de transporte que utiliza para trasladarse a su lugar de estudios es menos de 300 pesos chilenos o más de 2.000 pesos, que hayan declarado que último grado académico cursado por la madre es la enseñanza media completa y que el último grado académico cursado por el padre es también la enseñanza media completa, es asignado correctamente por el modelo como *Becado*.

El último modelo desarrollado es Gradient Boosting el que clasifica los datos según la información que muestra la matriz de confusión:

Tabla 34: Matriz de confusión – Gradient Boosting
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
	Becados	952	290
No Becados	464	735	

Esta indica que 952 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 735 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 290 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 464 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Por otro lado, el modelo presenta un desempeño según se indica:

Tabla 35: Indicadores de Rendimiento – Gradient Boosting
Fuente: Elaboración propia

Rendimiento	Resultado
Exactitud (Accuracy)	69,10%
Precisión	67,23%

Este modelo clasifica correctamente al 69,10% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 67,23% de los estudiantes. Además, otro indicador que se puede obtener son las variables más prioritarias según las clasifica este modelo. En la siguiente tabla se resumen las más importantes según el peso relativo entregado por el mismo modelo.

Tabla 36: Tabla de importancia de variables – Gradient Boosting
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
UBICORLUGEST_DESC	"Menos de 1 hora de traslado"	27,75%
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"	16,67%
CURSO_MADRE	"Enseñanza básica incompleta"	6,97%
CURSO_PADRE	"Enseñanza media completa"	6,89%
COSTOMEDIOTRANSPORTE_DESC	"Entre 301 a 500 pesos"	5,46%
COMUNA_EST_EDU	"Ancud"	3,21%
COSTOMEDIOTRANSPORTE_DESC	"Es menor igual a 300 pesos"	3,11%
COSTOMEDIOTRANSPORTE_DESC	"Entre 501 a 1000 pesos"	2,37%
CURSO_PADRE	"Enseñanza media completa"	2,08%
COMUNA_EST_EDU	"Concepción"	1,96%

Esto significa que un estudiante que haya declarado que la ubicación de su domicilio familiar con respecto a su lugar de estudios demora menos de 1 hora, que vive efectivamente en una comuna aislada (según el índice de aislamiento territorial), que el costo del medio de transporte que utiliza para trasladarse a su lugar de estudios llega hasta los 1.000 pesos chilenos, que hayan declarado que último grado académico cursado por la madre es la enseñanza básica incompleta, que el último grado académico cursado por el padre es enseñanza media completa y que su comuna de estudio es Concepción, es asignado correctamente por el modelo como *Becado*.

5.3.4 Aplicando el modelo de focalización simplificado

En este apartado se describen los resultados obtenidos aplicando los modelos predictivos construidos y aplicados en la sección 5.3.2 de este documento. La variante que se aplica es la selección de las variables categóricas y numéricas según las dimensiones del modelo de focalización simplificado. Por lo anterior, las variables categóricas seleccionadas son:

Tabla 37: Variables Categóricas – Modelo de Focalización simplificado
Fuente: Elaboración propia

NOMBRE CAMPO	DESCRIPCION CAMPO	DIMENSIÓN
DESC_ETNIA	Pueblo originario seleccionado por el estudiante	Pertenencia
DS_TRAMO_CSE	Tramo del 40 (0% – 40%): Hogares calificados en el 40% de menores ingresos o mayor vulnerabilidad Tramo del 50(41% – 50%): Hogares calificados entre el 41% y el 50% de menores ingresos o mayor vulnerabilidad Tramo del 60(51% – 60%): Hogares calificados entre el 51% y el 60% de menores ingresos o mayor vulnerabilidad Tramo del 70(61% – 70%): Hogares calificados entre el 61% y el 70% de menores ingresos o mayor vulnerabilidad Tramo del 80(71% – 80%): Hogares calificados entre el 21% y el 30% de mayores ingresos o menor vulnerabilidad Tramo del 90(81% – 90%): Hogares calificados entre el 11% y el 20% de mayores ingresos o menor vulnerabilidad Tramo del 100(91%-100%): Hogares calificados en el 10% de mayores ingresos o menor vulnerabilidad	Socioeconómica
INDICE AISLAMIENTO TERRITORIAL	Índice de aislamiento Territorial entre -1.0 y 0.9 (excluyendo el 0), se considera como "Comuna Aislada". Índice de Aislamiento Territorial mayor a 0.9 o igual a 0, se considera como "Comuna no aislada".	Territorial
COMUNA_EST_EDU	Comuna Establecimiento Educacional	Territorial

Mientras que las variables numéricas seleccionadas son:

Tabla 38: Variables Numéricas – Modelo de Focalización simplificado
Fuente: Elaboración propia

NOMBRE CAMPO	DESCRIPCION CAMPO	DIMENSIÓN
ESTUDIANTE_EN_RND	Estudiante se encuentra en el registro nacional de discapacidad: (0 = NO; 1 = SI)	Socioeconómica
HOGAR_MONOPARENTAL	Hogar formado por un progenitor (madre o padre) y uno o varios hijos: (0 = NO; 1 = SI)	Socioeconómica

Como resultados, obtenemos que los siguientes indicadores de desempeños de los modelos:

- **Regresión Logística**

Este modelo indica que 850 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* y 950 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 349 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 292 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Como medidas de desempeño este modelo clasifica correctamente al 73,74% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 73,13% de los estudiantes.

Por último, el modelo determina un área bajo la curva del 73,69% que indica que, en esa proporción, el modelo podrá distinguir las clases dadas.

- **Árboles de Decisión (cv=3)**

La matriz de confusión indica que 700 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 929 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 499 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 313 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Tabla 39: Matriz de confusión – Árboles de Decisión (cv=3)
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
Becados		700	499
No Becados		313	929

Este ajuste hace que el modelo clasifique correctamente al 66,70% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 65,06% de los estudiantes. También, la curva ROC que arroja el modelo muestra un área bajo la curva del 66,59% que indica que, en esa proporción, el modelo podrá distinguir a los *Becados* y *No Becados*.

Otra respuesta que se puede obtener son las variables más prioritarias según las clasifica este modelo. En la siguiente tabla se resumen las más importantes según el peso relativo entregado por el mismo modelo.

Tabla 40: Tabla de importancia de variables – Árboles de Decisión
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"	37,33%
HOGAR_MONOPARENTAL	1	22,30%
DS_TRAMO_CSE	"Sin RSH"	10,41%
COMUNA_EST_EDU	"Temuco"	6,10%
COMUNA_EST_EDU	"Coyhaique"	5,61%
DS_TRAMO_CSE	"Tramo 40%"	5,51%

Esto significa que un estudiante que provenga de una comuna aislada, cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, que pertenezca al 40% de la población más vulnerable o que no haya inscrito su información en el registro social de hogares y que estudie en las comunas de Temuco o Coyhaique, será clasificado correctamente por el modelo como *Becado*.

- **Random Forest**

La matriz de confusión indica que 949 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 788 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 411 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 293 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

El modelo entonces clasifica correctamente al 71,16% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 69,78% de los estudiantes. También, la curva ROC que arroja el modelo muestra un área bajo la curva del 69,78% que indica que, en esa proporción, el modelo podrá distinguir a los *Becados* y *No Becados*.

Otra respuesta que se puede obtener son las variables más prioritarias según las clasifica este modelo. En la siguiente tabla se resumen las más importantes según el peso relativo entregado por el mismo modelo.

Tabla 41: Tabla de importancia de variables – Random Forest
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
HOGAR_MONOPARENTAL	1	13,36%
INDICE AISLAMIENTO TERRITORIAL	"Comuna no aislada"	12,52%
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"	11,37%
DS_TRAMO_CSE	"Tramo 40%"	3,81%
DS_TRAMO_CSE	"Sin RSH"	3,60%
COMUNA_EST_EDU	"Temuco"	3,53%

Esto significa que un estudiante que provenga de una comuna aislada o no, cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, que pertenezca al 40% de la población más vulnerable o que no haya inscrito su información en el registro social de hogares y que estudie en la comuna de Temuco, será clasificado correctamente por el modelo como *Becado*.

- **Gradient Boosting**

El modelo muestra que 748 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 902 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 494 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 297 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

Este modelo clasifica correctamente al 67,6% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 71,58% de los estudiantes.

Otro indicador que se puede obtener son las variables más prioritarias según las clasifica este modelo. En la siguiente tabla se resumen las más importantes según el peso relativo entregado por el mismo modelo.

Tabla 42: Tabla de importancia de variables – Gradient Boosting
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"	31,11%
HOGAR_MONOPARENTAL	1	13,07%
DS_TRAMO_CSE	"Sin RSH"	7,21%
COMUNA_EST_EDU	"Ancud"	5,87%

Esto significa que un estudiante que provenga de una comuna aislada, cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, que

no haya inscrito su información en el registro social de hogares y que estudie en la comuna de Ancud, será clasificado correctamente por el modelo como *Becado*.

- **K-NN**

Resulta que 852 estudiantes con estado *Becado* del conjunto de datos de validación fueron clasificados correctamente como *Becados* por el modelo y 900 estudiantes con estado *No Becado* fueron clasificados correctamente como *No Becados*. Asimismo, 347 estudiantes con estado *Becado* fueron clasificados incorrectamente como *No Becados* por el modelo y 342 estudiantes *No Becados* fueron clasificados incorrectamente como *Becados*.

También, el resultado del rendimiento de este modelo indica que clasifica correctamente al 72% de los estudiantes en el estado correspondiente de asignación. Para el caso del estado *Becado*, el modelo clasifica correctamente al 72,17% de los estudiantes. Asimismo, el modelo determina un área bajo la curva del 71,76% que indica que, en esa proporción, el modelo podrá distinguir las clases dadas.

5.3.5 Evaluación del Modelo

Con el fin de determinar cuál es el mejor modelo que se ajusta al Programa de Residencia Familiar Estudiantil, se realizará una comparación del rendimiento y desempeño de cada uno de los modelos utilizados. Esta comparación se divide también por el modelo de focalización empleado.

- **Modelo de Focalización vigente**

En la siguiente tabla se muestra los indicadores resultantes de cada modelo:

Tabla 43: Comparación de Modelos – Exactitud de Validación
Fuente: Elaboración propia

Modelo	Accuracy de Validación
Regresión Logística	77,51%
K-NN(5)	74,00%
Random Forest	73,41%
Gradient Boosting	70,30%
Árboles de Decisión (cv=3)	68,40%

De acuerdo con los resultados de exactitud de los modelos, se evidencia que el mejor modelo predictivo para el programa PRFE es Regresión Logística, que

clasifica correctamente al 77,51% de los estudiantes en el estado correspondiente de asignación.

- **Modelo de Focalización simplificado**

Para el caso de los modelos construidos aplicando el modelo de focalización simplificado, la tabla comparativa entre modelos se muestra a continuación:

Tabla 44: Comparación de Modelos – Exactitud de Validación (con modelo de focalización simplificado)

Fuente: Elaboración propia

Modelo	Accuracy de Validación
Regresión Logística	73,74%
K-NN(5)	72,00%
Random Forest	71,16%
Gradient Boosting	67,20%
Árboles de Decisión (cv=3)	66,70%

Según se indica en la tabla, el mejor modelo predictivo para el programa PRFE es Regresión Logística, que clasifica correctamente al 73,74% de los estudiantes en el estado correspondiente de asignación.

5.3.6 Construcción del Perfil del Beneficiario del PRFE

De acuerdo con la evaluación de los modelos elaborada anteriormente, se puede obtener el perfil idóneo de beneficiarios PRFE según la importancia de las variables que los mejores modelos mostraron como predictores. Esto deja en evidencia que la información socioeconómica y territorial tienen gran relevancia dentro del programa, por lo que mientras más información fidedigna se obtenga del estudiante sobre estas dos dimensiones, más correcta será la asignación de este beneficio. Por lo tanto, el perfil idóneo de un estudiante que postula o renueva al beneficio debe contemplar la siguiente información mínima:

- **Modelo de Focalización vigente**

Tabla 45: Variables del perfil idóneo PRFE

Fuente: Elaboración propia

VARIABLE	VALOR
UBICORLUGEST_DESC	"Menos de 1 hora de traslado"
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"
CURSO_MADRE	"Enseñanza básica incompleta"
CURSO_PADRE	"Enseñanza media completa"
COSTOMEDIOTRANSPORTE_DESC	"Es más de 2.000 pesos"
HOGAR_MONOPARENTAL	1

Un estudiante que contenga como mínimo la información de la ubicación desde su vivienda hasta el lugar de estudios, que viva efectivamente en una comuna aislada, que su madre haya cursado como último grado la enseñanza básica incompleta, que su padre haya cursado como último grado la enseñanza básica completa, que el costo del medio de transporte que utilizan para trasladarse al lugar de estudios es más de 2.000 pesos chilenos y cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, tiene una alta probabilidad de obtener el beneficio del PRFE.

- **Modelo de Focalización simplificado**

Tabla 46: Variables del perfil idóneo PRFE – Modelo de focalización simplificado
Fuente: Elaboración propia

VARIABLE	VALOR
INDICE AISLAMIENTO TERRITORIAL	"Comuna aislada"
HOGAR_MONOPARENTAL	1
DS_TRAMO_CSE	"Sin RSH"
DS_TRAMO_CSE	"Tramo 40%"

Un estudiante que viva efectivamente en una comuna aislada, cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, que pertenezca al 40% de la población más vulnerable o que no haya inscrito su información en el registro social de hogares, tiene una alta probabilidad de obtener el beneficio del PRFE.

CAPÍTULO 6: PROPUESTA DE APOYO TECNOLÓGICO

6.1 Especificación de Requerimientos

Se propone desarrollar una aplicación web en la que se pueda cargar la nómina de los estudiantes que han solicitado el beneficio; esta aplicación ejecutará el modelo de minería de datos propuesto calculando la variable más importante del modelo de focalización de este programa y mostrará los resultados obtenidos. Permitirá finalmente, identificar a los potenciales beneficiarios que cumplen con los requisitos para ser asignados.

6.1.1 *Requerimientos Funcionales*

La aplicación web de gestión de la asignación del PRFE permitirá a los usuarios obtener un reporte de los potenciales beneficiarios, de acuerdo con los resultados arrojados por el modelo predictivo, según la propuesta descrita en el capítulo anterior. Los perfiles de usuarios de la aplicación propuesta se describen mediante la tabla a continuación:

Tabla 47: Perfiles de usuario de la aplicación
Fuente: Elaboración propia

No.	Perfil	Descripción
1	Administrador de Aplicación	Realiza la configuración del modelo predictivo, administra los usuarios y los permisos de acceso a las funcionalidades de la aplicación.
2	Encargado Nacional de Beca/Programa	Tiene todas las funcionalidades de la aplicación, desde la carga de nóminas hasta la exportación de reportes con los resultados. Puede realizar ajustes a los parámetros del programa relacionados con el modelo de focalización y su puntaje. Opera las funciones de cargar nóminas, ejecutar proceso de obtención de potenciales beneficiarios, extraer reportes del programa que administra a nivel nacional.
3	Encargado Regional Beca/Programa	Puede extraer reportes de asignación según la región a la que pertenece.

En relación con los requerimientos funcionales, esta aplicación debe realizar las siguientes acciones mínimas:

- Debe permitir el ingreso por medio de la identificación de un usuario y contraseña asignados.
- Debe permitir la creación de al menos 3 roles (administrador, encargado nacional y encargado regional) con permisos asociados.
- Debe permitir la visualización y descarga del reporte (en formato .xls o .csv) de potenciales beneficiarios según el rol.

- Debe permitir, al Encargado Nacional modificar parámetros de acuerdo con el modelo de focalización de la beca o programa.
- Debe permitir al administrador habilitar, crear o eliminar usuarios y cambiar permisos.

6.1.2 **Requerimientos No Funcionales**

Los requerimientos no funcionales están asociados con la arquitectura tecnológica que se sugiere utilizar y que se describirá a profundidad en el punto 6.2 del presente capítulo. Estos están asociados a las siguientes categorías que toda aplicación web debe contemplar:

- Calidad: Asociada a la funcionalidad, usabilidad, eficiencia, confiabilidad, y portabilidad. Por ejemplo, la aplicación debe ser fácil de usar y de comprender. Además, debe funcionar en cualquier navegador disponible en la actualidad.
- Interfaz: Las interfaces de usuario tienen que ser atractivas a la vista y fáciles de usar, es decir, los campos a completar no deberán tener ninguna dificultad para ser llenados.
- Organizacional: Al ser una institución pública, la aplicación deberá desarrollarse bajo las políticas establecidas por el Departamento de Informática. De todas maneras, la aplicación se desarrollará bajo una arquitectura de software libre, es decir, sin necesidad de adquirir alguna licencia.

6.2 **Arquitectura Tecnológica**

Como el modelo predictivo se desarrolló en el lenguaje de programación Python, se propone utilizar el framework⁶ Django para la aplicación web descrita. Este framework mantiene una lógica Model-View-Controller (MVC), aunque sus creadores lo denominan Model-Template-View (MTV) por la estructura que utiliza para el desarrollo de este tipo de aplicación.

La «vista» es la función de “callback”⁷ en Python para una URL en particular, porque esta función callback describe cuál dato es presentado (Django Software, 2019). Ese dato se presenta a través de una «plantilla» y el «modelo» contiene la comunicación con la base de datos de la aplicación. (Ver Ilustración 39)

⁶ Entorno de trabajo, que nos ofrece una forma estándar para escribir código en un lenguaje concreto. Fuente: <https://lenguajesdeprogramacion.net/diccionario/que-es-un-framework/>

⁷ Palabra en inglés utilizada en el lenguaje de programación Python para representar la frase: “volver a llamar”.

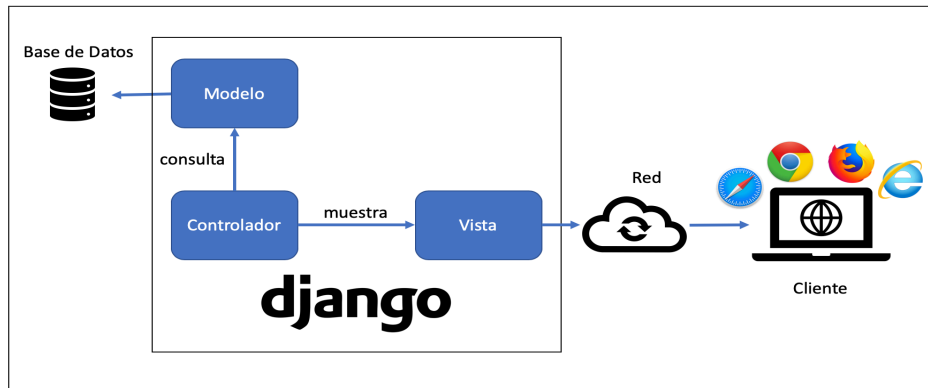


Ilustración 39: Arquitectura Tecnológica aplicación web

Fuente: Elaboración propia basado en Sitio Web - <https://lenguajesdeprogramacion.net/diccionario/que-es-un-framework/>

Los beneficios de utilizar este tipo de arquitectura son, por ejemplo, que el diseño modular permite a los diseñadores y a los desarrolladores trabajar conjuntamente. También, hace que la aplicación sea muy ligera, debido a que usa menos código por el lenguaje de programación Python que es dinámico.

6.3 Diseño de la Aplicación

En esta sección se presenta el diseño de la aplicación a implementar, identificando el detalle de los módulos más importantes describiendo sus funcionalidades representados en los siguientes diagramas UML⁸.

6.3.1 Casos de Uso

El siguiente diagrama de Caso de Uso (ver Ilustración 15) describe cómo los diferentes usuarios involucrados en el proceso de gestión de la asignación de la beca/programa, ejecutan y el modelo predictivo para obtener el resultado de los potenciales beneficiarios.

⁸ Lenguaje Unificado de Modelado para la documentación de los desarrollos de software.

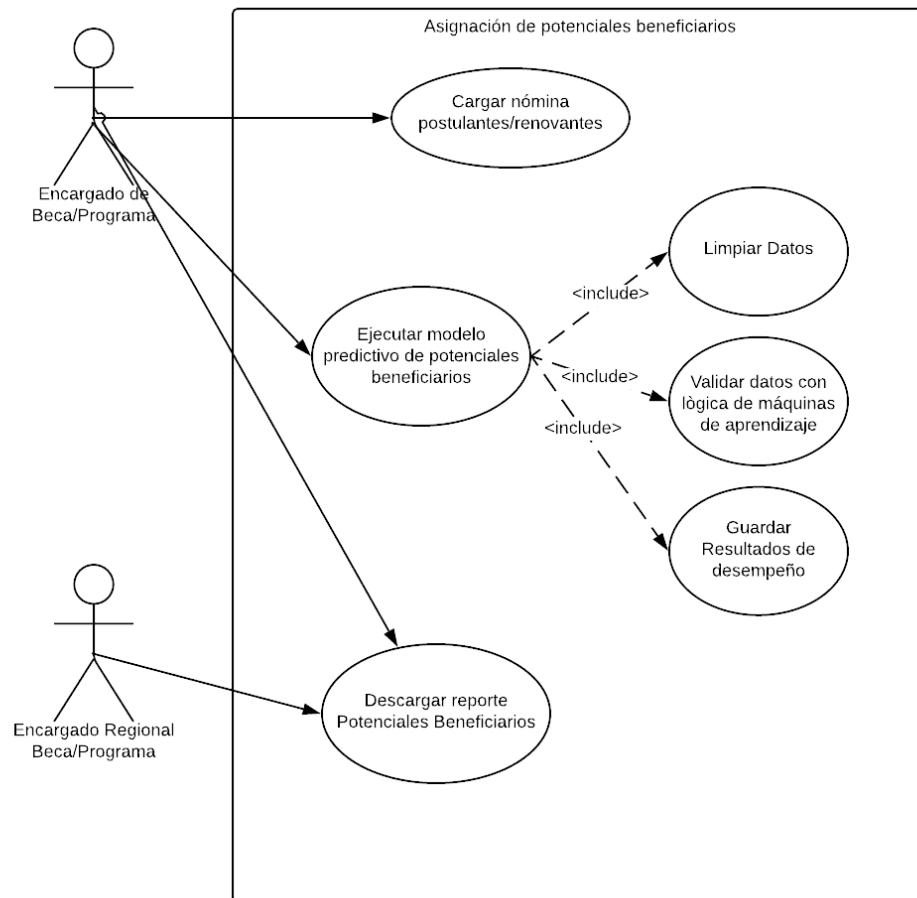


Ilustración 40: Caso de Uso - Asignación de potenciales beneficiarios
Fuente: Elaboración propia

6.3.2 Diagramas de Secuencia

La siguiente ilustración muestra la interacción de los distintos objetos de la aplicación web propuesta.

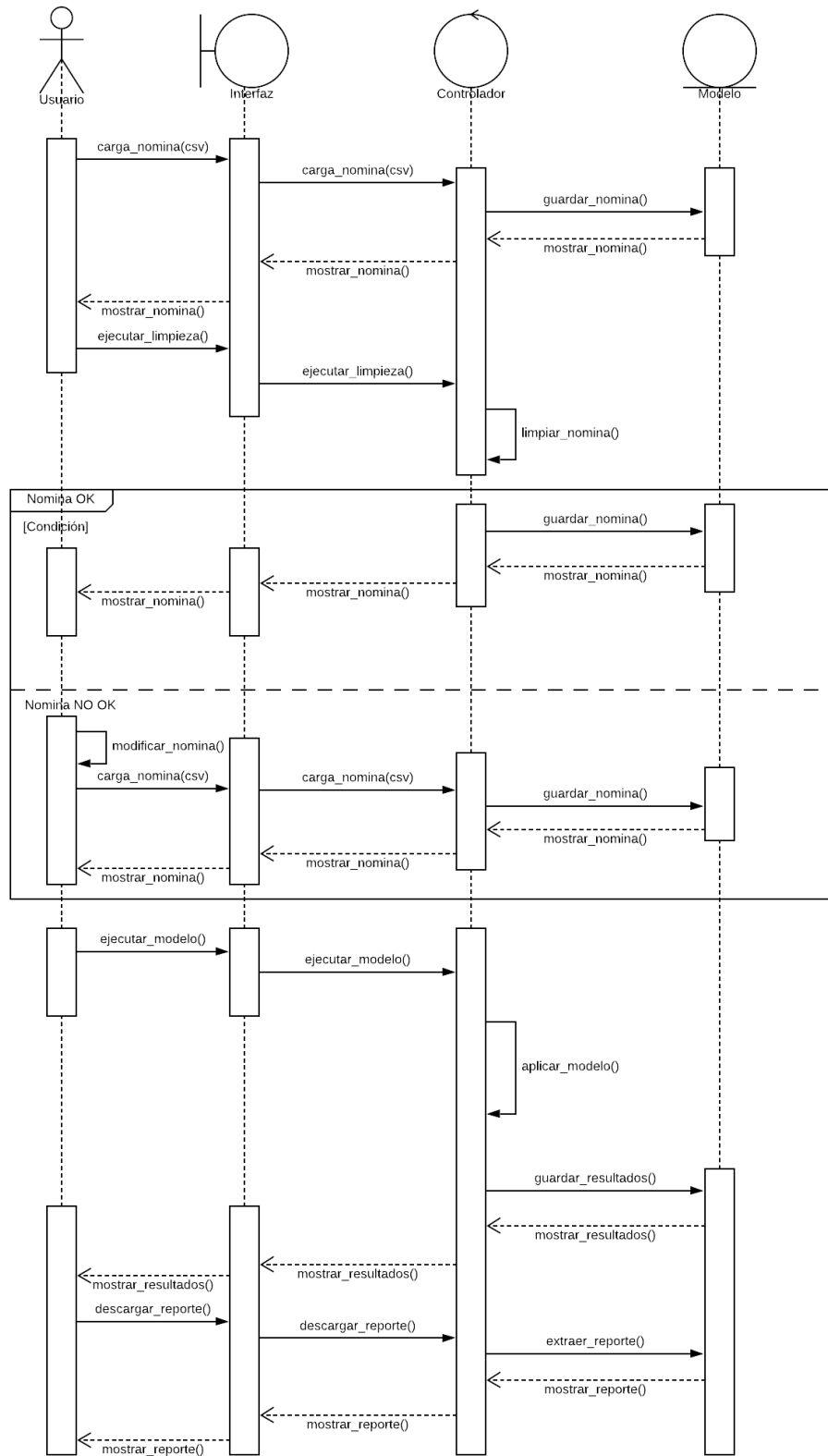


Ilustración 41: Diagrama de Secuencia - Asignación de potenciales beneficiarios
Fuente: Elaboración propia

6.3.3 Diagramas de Clases

La siguiente ilustración muestra las clases, sus atributos y operaciones de la aplicación web antes descrita.

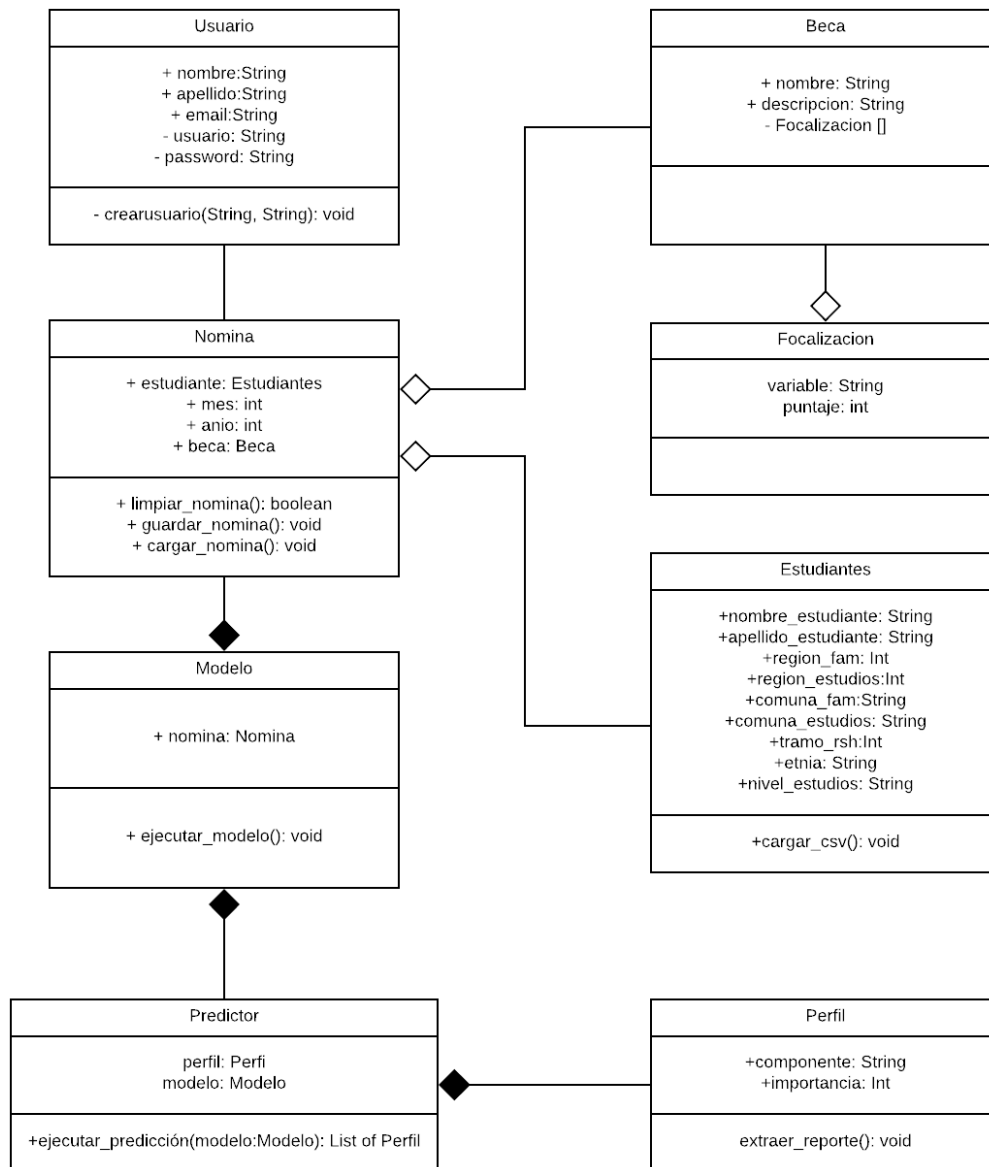


Ilustración 42: Diagrama de Clases - Asignación de potenciales beneficiarios
Fuente: Elaboración propia

Las clases mostradas en este diagrama evidencian la interacción que tendrán las tablas en la base de datos de la aplicación web que se propone desarrollar. Es importante destacar que todos los resultados obtenidos por el modelo predictivo serán registrados y, en su defecto, se podrán descargar para que estén disponibles para el usuario.

CAPÍTULO 7: GESTIÓN DEL CAMBIO

7.1 Contexto de la Organización

Como ya ha sido mencionado anteriormente, JUNAEB es un organismo administrado por el Estado que depende de los recursos públicos que anualmente se le asignan.

Tiene 55 años sirviendo a la comunidad chilena, en especial a los niños, niñas y jóvenes en condición de vulnerabilidad biopsicosocial.

Cuenta con más de 120 funcionarios de planta, más de 800 profesionales contratados y más de 100 funcionarios a honorarios, a lo largo de las 16 direcciones regionales, 19 direcciones provinciales y la dirección nacional.⁹ Esto indica una dotación aproximada total de 1.020 funcionarios en todo Chile.

Su estructura organizacional es jerárquica distribuida verticalmente, lo que implica un alto grado de burocracia para cada proceso que se desarrolla internamente. A continuación, se muestra el organigrama de JUNAEB

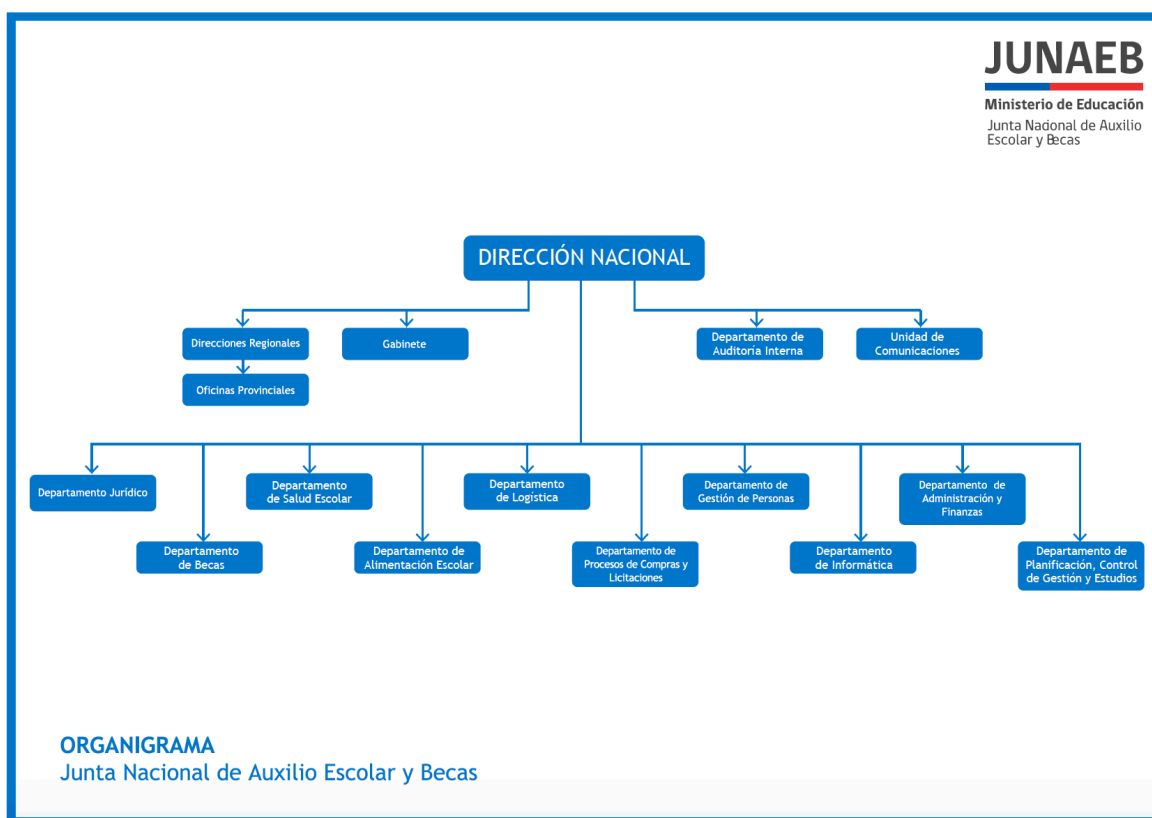


Ilustración 43: Organigrama JUNAEB
Fuente: Sitio Web - <https://www.junaeb.cl/organigrama/>

⁹ Fuente: http://transparencia.junaeb.cl/iris_junaeb/gdesk/album.jsp

De modo que el proyecto propuesto se aspira implementar en el Departamento de Becas JUNAEB, se analizará la estructura, cultura organizacional y el comportamiento de los actores clave para el proyecto dentro de este departamento.

7.2 Observación de la implementación a realizar

El Departamento de Becas de JUNAEB cuenta con la siguiente estructura jerárquica:

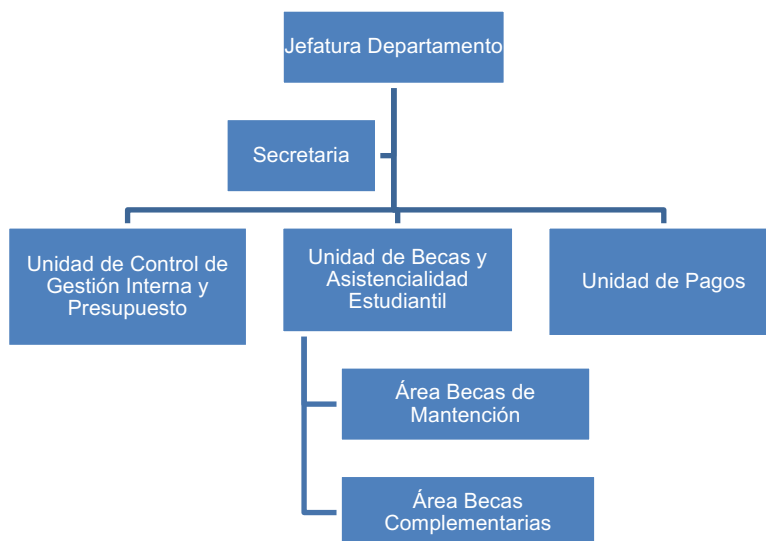


Ilustración 44: Organigrama Departamento de Becas de JUNAEB

Fuente: Departamento de Becas de JUNAEB, 2018

Esta jerarquía indica a groso modo como se lleva a cabo el proceso de toma de decisiones interno. Esto hace que el proceso sea lento y burocrático, teniendo que pasar por las jefaturas de área o de unidad para que la jefatura de departamento conozca el problema, se trabaje con éste y se entregue alguna directriz para su solución. Igualmente, no se debe dejar de lado el innumerable cumplimiento de normas que, en algunos casos, puede ser innecesario viniendo de la aplicación discrecional del poder de quien la ejecuta.

Si se alude a lo que respecta con la cultura organizacional cuenta con un sinfín de prácticas, creencias, expectativas, etc., de el “cómo se hacen las cosas”, lo que se traduce en prácticas concretas y cotidianas, negativas o positivas. (Waissbluth, 2008)

Hay que observar también las restricciones que esto cambios traen a cualquier institución pública, tales como, presupuesto, políticas internas, comunicacionales, múltiples actores claves que participan en el proyecto, entre otras.

7.3 Análisis de los principios de diseño

Para poder diseñar un buen Plan de Gestión de Cambio y que el proyecto pueda implementarse de la mejor manera, se debe analizar algunos principios que serán relevantes tener en cuenta al momento de llevar a cabo este plan para poder cumplir con los resultados esperados.

Primero, se debe establecer una estrategia de cambio, es decir, debe existir un propósito claro del porque es importante realizar este cambio. El foco debe estar centrado en lo que se desea cambiar y en las personas que se transformarán en los actores clave de este proceso de cambio. Esto significa que se le debe dar un sentido este proceso de manera tal que las personas se sientan “seducidas” y finalmente exista un compromiso por parte de ellos.

Tal como fue mencionado anteriormente, al realizar un proceso de cambio dentro de un departamento perteneciente a un servicio público, se tiene que tomar en cuenta que no “todo” se puede cambiar, existen prácticas instauradas que se rigen por políticas ya definidas que no dependen de los actores claves ni de los líderes de este proceso. Por lo tanto, se tiene que identificar que no se puede cambiar, que se va a conservar y declararlo para que desde el inicio quede establecido su alcance.

Otro principio relevante es la identificación de los actores clave, el equipo de trabajo y sus roles. Estos equipos deben estar integrados y comunicados entre sí y deben tener claro cuáles son las preocupaciones que deben cuidar con el fin de que el proyecto finalice exitosamente.

También, debe existir una gestión emocional, lo que implica conocer los estados de ánimo de los líderes, los equipos de trabajo, entre otros de tal manera de poder diseñar estrategias que permita intervenir en ellos y generar cambios internos que fortalezcan el compromiso y sentido que busca obtener como resultado este proyecto.

Ligado con lo anterior, las comunicaciones entre los actores clave es muy relevante. Se deben crear instancias de comunicación ayudando a generar los contextos para que existan conversaciones importantes para el proyecto, donde todos puedan participar.

Una vez identificados los actores clave del proyecto, es necesario determinar las habilidades técnicas y relacionales existentes, así como las que son precisas cambiar. Aquí es obligatoria diseñar e instalar capacidades en la institución que sean importantes para el proceso de cambio, lo que implica implementar actividades de capacitación técnica para que se reduzca la brecha histórica de estas habilidades. Hay que mencionar que para los servicios públicos muchas habilidades instauradas se convierten en prácticas que llevan mucho tiempo llevándose a cabo por el simple hecho de que “siempre se han hecho así las cosas”.

Un principio clave para tener en cuenta es la gestión del poder, y más en este caso cuando se identifica una cultura organizacional cerrada y completamente establecida y el piso político que pudieran tener los directivos de este servicio. Por lo tanto, hay que identificar los espacios de poder posibles y necesarios con el fin de instaurar la relevancia e importancia que este cambio va a generar.

En último lugar, el proceso de cambio debe establecer un sistema de seguimiento y control, lo que derivará a reuniones de evaluación que ayudarán a mejorar el diseño inicial con el que fue planteado este cambio. Esto debe transformarse en la instauración de hitos y ritos que contemplen un inicio y un fin y que involucre a los distintos actores y equipos del proceso.

7.4 Caracterización del cambio

Este proyecto requerirá la implementación de un sistema web como apoyo tecnológico, lo que implica el uso de una nueva herramienta informática que no ha sido utilizada anteriormente.

Hay que tener muy en cuenta la cultura organizacional actual en la que se tiene interiorizado que “las directrices de gabinete se deben hacer y no se puede hacer nada si no estamos de acuerdo” o “esto es urgente como para ayer”. Estas acciones se repiten diariamente, por ende, la Jefatura del Departamento no se da el tiempo de instalar alguna nueva creencia tanto en ella como para el equipo que tiene a cargo. En consecuencia, esta será una de las cosas que este proyecto busca cambiar y no conservar.

7.4.1 Organización y estructura del cambio

Los actores clave serán parte importante de la organización del proyecto, no obstante, la composición estará establecida a partir de un equipo de cambio, en el que el Coordinador de Área Becas Complementarias tendrá la dirección del proyecto.

Los actores clave de este proyecto se identifican a continuación:

1. Jefatura de Unidad: 1 persona
2. Coordinador de Área Becas Complementarias: 1 persona
3. Integrantes de Área Becas Complementarias: 3 personas + estudiante MBE
4. Coordinador de Área Becas de Mantenimiento: 1 persona
5. Coordinador de Proyectos TI Becas: 1 persona
6. Direcciones regionales: 16 personas
7. Entidades Ejecutoras: 40 personas

La estructura de este proyecto estará compuesta por 3 áreas denominadas:

- Dirección del Proyecto
- Liderazgo del Proyecto
- Equipos de Trabajo

La siguiente tabla muestra las funciones y los integrantes de cada una según los actores clave anteriormente identificados:

Tabla 48: Organización y estructura del proyecto de cambio
Fuente: Elaboración propia

Dirección del Proyecto	Liderazgo del Proyecto	Equipos de Trabajo
<p>Funciones esperadas:</p> <ul style="list-style-type: none"> • Definir lineamientos y estrategias del proyecto (coordinar acciones). • Obtener y aplicar recursos en pos de un resultado. • Gestionar el proyecto (definir nombre, alcance, cambiarlo si es necesario y declarar las ideas fuerza). <p>Integrantes:</p> <ul style="list-style-type: none"> • Coordinador Área Becas Complementarias 	<p>Funciones esperadas:</p> <ul style="list-style-type: none"> • Generar aprendizaje para la organización. • Generar espacios de reflexión. • Establecer diversas instancias que permitan escuchar, identificar e intervenir en los estados de ánimo de la organización. <p>Integrantes:</p> <ul style="list-style-type: none"> • Encargada Nacional PRFE • Estudiante MBE 	<p>Funciones esperadas:</p> <ul style="list-style-type: none"> • Establecer las tareas específicas para cada actor. • Establecer espacio de acciones y prácticas para el proyecto. • Conducir reuniones de trabajo efectivas. <p>Integrantes:</p> <ul style="list-style-type: none"> • Jefatura Unidad • Integrantes de Área Becas Complementarias • Coordinador de Área Becas de Mantenimiento • Coordinador de Proyectos TI Becas

La comunicación dentro de esta organización debe ser fluida y eficaz, lo que implica establecer una estrategia comunicacional robusta, en la que se escuchen las

resistencias y obstáculos que se van presentando en el proyecto, de tal forma de poder lidiar con ellas. Esta comunicación será de manera horizontal, es decir que toda la información relacionada con el proyecto será transmitida de manera transversal para que todas las áreas e integrantes la manejen y se puedan tomar decisiones en función del éxito de éste.

El equipo de la gestión del proyecto en conjunto con el equipo de gestión del cambio deberá realizar una comunicación efectiva basada en la confianza, es decir, mediante datos duros se deberá comprobar que el cambio en el modelo de focalización propuesto es necesario hacerlo para poder beneficiar a más estudiantes.

Asimismo, se debe tener una comunicación directa con las direcciones regionales, ya que para este proyecto son una parte importante al momento de la implementación de este. Se propone entonces, realizar videoconferencias para comunicar el avance de este proyecto, con el fin de que estén en conocimiento.

El liderazgo del proyecto debe declarar el inicio, etapas y término del proceso y proyectos de cambio. Así como observar la evolución del proyecto y generación de los hitos y ritos relevantes. Además de diseñar y realizar el proceso de cierre y evaluación final del proyecto. Cabe indicar que la gestión del poder juega un rol muy importante en el proceso de cambio, y es un proceso continuo a lo largo de todo el proyecto. Es responsabilidad del liderazgo del proyecto hacerse cargo de construir poder de manera continua para que el proyecto pueda avanzar, de lo contrario puede que el proyecto fracase.

7.4.2 Público Objetivo del cambio

Este proyecto está dirigido a los estudiantes vulnerables de educación media y superior que por no tener oferta educacional cerca de su domicilio familiar deben trasladarse y por ende solicitan el beneficio de residencia familiar estudiantil a través de las entidades ejecutoras y gestionado por las direcciones regionales de JUNAEB.

Estos dos últimos, son el público objetivo de cambio del proyecto, ya que son los que administran este beneficio y son los que atienden presencialmente a los estudiantes postulantes y renovantes.

7.4.3 Evaluación del proceso de cambio

El proceso de cambio incorporará prácticas nuevas y las ya existentes se ajustarán de acuerdo con las necesidades del proyecto y de la organización. Estas prácticas deben estar claramente detalladas, ya que serán uno de los ítems a controlar dentro de la

evaluación y seguimiento del cambio. La siguiente tabla muestra las prácticas que se conservarán y las que se agregarán en el proceso de cambio de este proyecto:

Tabla 49: Prácticas del cambio
Fuente: Elaboración propia

PRÁCTICAS	CONSERVAR	INCORPORAR
Nombre	Conversar las ideas	Reunión equipo técnico
Propósito	Incorporar las ideas del equipo para que se realice al menos una	Realizar seguimiento de los acuerdos que impactan directamente a las becas en el ámbito técnico (por ejemplo: modelo de focalización)
Descripción	Los jueves a la hora del almuerzo, salimos los 5 y conversamos que de las cosas que han pasado en la semana y cómo podemos mejorar nuestras tareas.	Semanalmente, el equipo se reunirá formalmente (en una sala que será reservada para estos fines) con el fin de revisar los avances que el proyecto tenga y se validen el/los modelos implementados para el rediseño del proceso de asignación.
Hito/Tiempo	Jueves 13:30	Viernes 11:00 – 12:00

La práctica que se decide conservar ha dado muchos frutos hasta el momento ya que se han identificado las oportunidades de mejora que han aportado a este proyecto para la búsqueda de la mejor solución.

Lo que se espera de esta práctica es que el equipo reporte a la Dirección del Proyecto los avances que se están llevando a cabo con el fin de tomar acciones preventivas o correctivas dependiendo de lo que se observe.

Es así como realizar estas prácticas e instaurar la observación del cambio son fundamentales para el seguimiento y la evaluación de éste. Al respecto, se ha creado un panel de control que contemplan todos los dominios del modelo ontológico de la gestión del cambio que han sido descrito en el presente capítulo. (Ver tabla 50)

Tabla 50: Panel de Control del proceso de cambio

Fuente: Elaboración propia basado en los 10 elementos del Modelo Ontológico de la Gestión del Cambio

DOMINIO	OBJETIVO	DESCRIPCIÓN	INDICADOR	MEDIDA	RESPONSABLES	META		
						ÓPTIMO	TOLERABLE	DEFICIENTE
Liderazgo y Gestión del Proyecto	Definir alcance, estructura, etapas y subproyectos según lo solicitado por "el cliente".	Debe existir coordinación entre los actores clave, para manejar el proyecto con las narrativas que se presenten. Si se articulan correctamente todos estos elementos el proyecto será exitoso.	Grado de articulación del proyecto	%	1. Dirección del Proyecto 2. Líder del Proyecto	100%	[99 - 70] %	> 70%
Estrategia y Sentido	Construir sensibilización a medida que se desarrolla el proyecto combinando los relatos identificados de los	El sentido debe contruirse y desconstruirse a medida que se desarrolla el proyecto. Es dinámico por lo que debe estar siempre presente.	Personas que conocen el sentido del proyecto en relación a los actores clave del cambio	%	1. Dirección del Proyecto 2. Líder del Proyecto	100%	[99 - 60] %	> 60%
Cambio y conservación	Identificar procesos de cambio y conservación del proyecto.	El proyecto tiene actividades, procesos, etc., que son difíciles de cambiar, pero también que son importantes conservar, por lo tanto es importante identificarlas, declararlas y realizar acciones para que se noten los cambios y conversiones del proyecto.	Cantidad de procesos que se cambian y se conservan en el proyecto	%	1. Dirección del Proyecto 2. Líder del Proyecto	100%	[99 - 70] %	> 70%
Organización y Estructura	Definir los equipos, integrantes y sus roles	El proyecto debe establecer la forma en que los equipos se integrarán identificando a los participantes y definiendo sus roles. Se deben realizar instancias de participación entre todos los equipos durante el desarrollo del proyecto.	Grado de integración entre los equipos	%	Dirección del Proyecto	100%	[99 - 80] %	> 80%
Hitos, Ritos y Etapas	Diseñar la planificación del proyecto que contengan las etapas, hitos y ritos del cambio.	El proyecto debe contemplar el diseño de una planificación que declare el inicio y cierre de las etapas definidas, la realización de ritos entre otras actividades que apoyen a la gestión del proyecto.	Cantidad de actividades realizadas según planificación	%	Dirección del Proyecto	100%	[99 - 60] %	> 60%
Gestión Emocional	Gestionar las emociones y estados de ánimo de los actores clave del cambio mediante el escuchar y observar.	El proyecto debe diseñar estrategias para escuchar estados de ánimo y establecer procesos de intervención para que exista una movilización de los estados de ánimo que se presentarán durante el desarrollo del mismo.	Cantidad de procesos de intervención para la movilización exitosa de estados de ánimo	%	Líder del Proyecto	100%	[99 - 75] %	> 75%
Gestión de las Comunicaciones	Generar instancias de comunicación entre los actores clave del proyecto.	La comunicación para el cambio debe ser bidireccional, es por esto que el Líder del Proyecto debe identificar las conversaciones relevantes, generando contextos para que ocurran.	Cantidad de instancias de comunicación realizadas	%	Líder del Proyecto	100%	[99 - 60] %	> 60%
Gestión de las Habilidades	Capacitar a los actores del proceso de cambio en las nuevas habilidades que el proyecto requiere.	Se deben desarrollar habilidades técnicas y relacionales para el proceso de cambio con el fin de instalar estas capacidades en la organización mediante la realización de capacitaciones para todos los actores.	Cantidad de personas capacitadas según el plan de capacitación establecido para el proyecto	%	Equipo de Trabajo	100%	[99 - 75] %	> 75%
Gestión del Poder	Identificar los espacios de poder necesarios para el desarrollo del proyecto.	Se deben acumular capacidades para ejercer el poder dentro del proyecto, ya que es necesario que las cosas funcionen para que este sea exitoso.	Cantidad de espacios de poder identificados que se han acumulado para el proyecto	%	Dirección del Proyecto	100%	[99 - 80] %	> 80%
Observación y Seguimiento	Establecer un sistema de seguimiento y control del proyecto.	Se debe observar e identificar qué está ocurriendo en el proyecto para poder dar seguimiento y tomar acciones correctivas.	Nivel de cumplimiento de los objetivos definidos por los equipos de trabajo	%	1. Dirección del Proyecto 2. Líder del Proyecto	100%	[99 - 80] %	> 80%

7.5 Factores críticos de éxito

Más allá de la planificación que se establezca para este proyecto, del cumplimiento de los plazos, indicadores, etc., es importante recalcar que los tres dominios críticos para el éxito son la gestión emocional, la gestión de las habilidades y la gestión del poder. A continuación, se describirán cada uno destacando las acciones clave que deben realizarse para obtener el resultado esperado del proyecto.

7.5.1 Gestión Emocional

Este dominio es fundamental para que este proyecto se lleve a cabo, ya que no existe dentro de la cultura de la institución por ende sería algo nuevo a instaurar. De hecho, es fácil detectar las emociones que se encuentran presentes en la institución, independientemente de la implementación del proyecto.

Estas emociones están relacionadas con los miedos que las personas tienen por hacer mal un trabajo. Al ser una institución pública, los funcionarios están muy expuestos a las críticas positivas o negativas. Otra emoción que se presenta es el descontento debido a la falta de apreciación por los logros que el equipo tiene, ya sea por parte de la jefatura directa o por el propio Director Nacional. Esto hace que las personas muestren desinterés en su trabajo y que lo ejecuten por cumplir sin algún esfuerzo adicional para aportar o proponer acciones nuevas.

La relevancia de identificar estas emociones se presenta cuando también se generan actividades para poder conversarlas y ayudar a crear oportunidad de mejora para no perder el sentido que el proyecto tiene.

En resumen, corresponde al liderazgo y dirección del proyecto, detectar cuáles son las narrativas que los actores están diciendo, con el fin de entender que les ocurre para poder hacerse cargo de los posibles quiebres. Hay que tener en cuenta que siempre hay variaciones en las emociones de cada persona y obviamente en el proyecto, pero lo relevante es tener la capacidad de escucharlas y observarlas para asegurar una adecuada implementación del proyecto.

7.5.2 Gestión de las Habilidades

Para que el proyecto se desarrolle adecuadamente en la institución, es vital identificar las habilidades que existen en los actores involucrados en el proceso de cambio. Asimismo, hay que detectar las brechas técnicas que se pudieran presentar para así crear instancias de capacitación con el fin de reforzar n las habilidades existentes y formar a los actores en las nuevas que se requieran.

Lo relevante de esto es lograr la instalación de capacidades en la institución para este y futuros procesos internos de cambio. Una herramienta que puede ayudar a la gestión de éstas es la creación de mapas de habilidades, los cuales pueden ser creados para cada equipo de trabajo del proyecto. Esto fortalecerá las habilidades históricas y generará aprendizaje para el proceso.

7.5.3 Gestión del Poder

A medida que el proyecto vaya avanzando en su implementación, hay que gestionar el poder para que todo lo planificado y requerido suceda. En efecto, se puede detectar claramente que se requiere tener el poder de autoridad, ya que como la propuesta fue presentada por un profesional de un área del Departamento de Becas, es necesario que exista un actor que le dé más peso al proyecto, es probable que con los cambios políticos que se generan en la institución, puede pasar inadvertido o que no se le dé la relevancia que se necesita.

Hay que tener claro que el poder político juega un papel fundamental por lo anteriormente mencionado, pero si se tiene el poder pragmático, es decir el “know-how” del proyecto y existe una articulación fuerte con los actores clave, se puede acumular el capital necesario para que lo planeado funcione.

En resumen, los factores críticos de éxito de este proyecto consisten en definir a los actores, escucharlos, anticipar resistencias y estados de ánimo e involucrarlos en el proceso de cambio mediante actividades que los motiven y generen sentido sobre el éxito de éste.

7.6 Plan de gestión del cambio

Después de conocer los dominios ontológicos de la gestión del cambio y como se identifican para este proyecto, se detallará a continuación un plan general de la gestión del cambio aplicado a esta propuesta.

Esto será mostrado en la siguiente tabla que contendrá los puntos clave a considerar:

Tabla 51: Estructura Etapa de Preparación del Plan de Gestión del Cambio
Fuente: Elaboración propia

I. ETAPA DE PREPARACIÓN	
Nombre del Proyecto:	
Descripción del proyecto:	
Objetivos:	
Resultados Esperados:	
Posibles riesgos y acciones de mitigación:	
<u>Riesgo</u>	<u>Acción de mitigación</u>

En la primera etapa denominada “Etapa de Preparación” corresponde identificar los resultados esperados del proyecto. Además, hay que detectar los riesgos que se pueden presentar junto con las acciones que permitan mitigar ese riesgo. La clave es declararlo en el plan para que todos los actores estén en conocimiento de ello.

Por ejemplo, en este proyecto un riesgo claramente identificado es la baja motivación hacia el cambio por parte de los integrantes de toda la institución, por lo que el desafío

se presenta en poder sensibilizar a las personas para que incorporen el cambio como un aprendizaje dentro de sus labores y que no se conformen con lo existente por el simple hecho de ser lo “más cómodo” o lo “conocido” de cómo se hacen las cosas.

Tabla 52: Estructura Etapa de Relación con las personas del Plan de Gestión del Cambio
Fuente: Elaboración propia

II. ETAPA DE RELACION CON LAS PERSONAS		
Roles, Funciones y Responsables:		
Organización del proyecto:		
Estados de ánimo al inicio del proyecto:		
Actores Clave	Estado de Ánimo	Habilidades para el Cambio

La etapa de “Relación con las personas” es la segunda dentro de este plan, la cual contempla la integración y articulación que debe estar entre los actores clave de este proyecto. Un ejemplo concreto se puede evidenciar en el punto 7.4.1 del presente capítulo.

Tabla 53: Estructura Etapa de Instalación del cambio del Plan de Gestión del Cambio
Fuente: Elaboración propia

III. ETAPA DE INSTALACIÓN DEL CAMBIO	
Prácticas instaladas:	
Nombre:	
Propósito:	
Descripción:	
Hito/Tiempo:	
Red de Apoyo:	
Capacitación:	
Habilidad:	
Lugar y Fecha:	
Participantes:	
Comunicación del proyecto:	
Identificando el Poder del proyecto:	

Esta etapa de instalación del cambio contiene la descripción de cada una de las prácticas que se instalan en el proceso, de las actividades de capacitación de las nuevas habilidades, del plan de comunicación que debe estar declarado para una óptima implementación del proyecto y la declaración del poder que se va captando para el éxito de este.

En este capítulo, en los puntos precedentes, se detallan la aplicación de esta etapa en el proyecto propuesto.

Tabla 54: Estructura Etapa de Evaluación y Cierre del Plan de Gestión del Cambio
Fuente: Elaboración propia

IV. ETAPA DE EVULUACIÓN Y CIERRE
Panel de Control:
Reuniones de Evaluación:
Diseño de celebración de Logros:
Diseño de Cierre del proyecto:

Por último, la etapa denominada “Evaluación y Cierre”, contiene la herramienta de seguimiento que será utilizada en este proceso como es el Panel de Control que fue mencionado en el apartado 7.4.3 del presente capítulo. Asimismo, es relevante tener un registro de las reuniones de evaluación, los logros que cada hito, así como el diseño del cierre de este proyecto.

CAPÍTULO 8: EVALUACIÓN DEL PROYECTO

En esta sección se presenta la factibilidad del proyecto implementado en la institución, debido a que los beneficios de éste están relacionados con mejorar los tiempos de respuesta en las actividades internas del programa y en los tiempos de respuesta a los usuarios. Se realizará una evaluación privada del proyecto contemplando el escenario actual y proyectando el comportamiento del programa si el proyecto es implementado según lo propuesto.

Debido a que no fue factible realizar una prueba piloto con información de nuevos beneficiarios del programa, se realiza en este capítulo un análisis retrospectivo utilizando información histórica para tales efectos.

8.1 Análisis Retrospectivo

Luego de obtener el perfil idóneo del estudiante que es potencial beneficiario del PRFE, se comparará con la nómina final de beneficiarios del año 2017 para determinar el grado de acierto y de error relacionado con el estado final del estudiante, es decir si estuvo o no becado en el programa. Esto implica que se corran los modelos predictivos utilizados para el proyecto con los datos del año 2017, para analizar los resultados y comprender su comportamiento con respecto a la asignación oficial del programa.

8.1.1 Presentación de datos

La base de datos utilizada para este análisis contiene los campos que se utilizaron en el año 2017 para poder asignar el beneficio. En ese entonces, para medir la vulnerabilidad del estudiante se solicitaba la información al departamento de planificación y estudios de la institución sobre el puntaje SINAЕ¹⁰ y si el estudiante pertenecía al Sistema Chile Solidario¹¹.

Una variable nueva que no se analizó en la propuesta para el año 2018, es la denominada OFERTA EDUCACIONAL. Esta era la variable relevante para otorgar el mayor puntaje en la dimensión de aislamiento territorial, ya que guarda relación con la existencia de oferta educación en la comuna del domicilio familiar. La siguiente tabla resume las variables de la base de datos PRFE 2017 con su respectiva descripción.

¹⁰ Fuente: Sistema Nacional de Asignación con Equidad para Becas, Junta Nacional de Auxilio Escolar y Becas

¹¹ Fuente: <https://www.ips.gob.cl/servlet/internet/content/1421810829144/chile-solidario>

Tabla 55: Campos del Dataset PRFE 2017

Fuente: Elaboración propia

NOMBRE CAMPO	DESCRIPCION CAMPO	TIPO
COD_FORMULARIO	Código único identificador	Integer
DESC_ESTADO	Estado de Beneficio: Becado: que obtuvo el beneficio. No Becado: no se adjudicó el beneficio.	String
DESC_ETNIA	Pueblo originario seleccionado por el estudiante	String
PUNTAJE	Suma del puntaje por cada variable del modelo de focalización del programa.	Float
ds_tramo_cse	Tramo del 40 (0% – 40%): Hogares calificados en el 40% de menores ingresos o mayor vulnerabilidad Tramo del 50(41% – 50%): Hogares calificados entre el 41% y el 50% de menores ingresos o mayor vulnerabilidad Tramo del 60(51% – 60%): Hogares calificados entre el 51% y el 60% de menores ingresos o mayor vulnerabilidad Tramo del 70(61% – 70%): Hogares calificados entre el 61% y el 70% de menores ingresos o mayor vulnerabilidad Tramo del 80(71% – 80%): Hogares calificados entre el 71% y el 80% de mayores ingresos o menor vulnerabilidad Tramo del 90(81% – 90%): Hogares calificados entre el 81% y el 90% de mayores ingresos o menor vulnerabilidad Tramo del 100(91%-100%): Hogares calificados en el 91% de mayores ingresos o menor vulnerabilidad	String
ID_TRAMO_CSE	Identificador del tramo de calificación socioeconómica del hogar	Integer
NIVEL	Nivel de estudios del estudiante	String
COMUNA_EST_EDU	Comuna Establecimiento Educacional	String
REGION_DOM_FAMILIAR	Región domicilio familiar	Integer
COMUNA_FAMILIAR	Comuna Domicilio Familiar	String
OFERTA EDUCACIONAL	Existencia en la comuna de domicilio familiar de oferta de educacional para estudiantes	String
INDICE AISLAMIENTO TERRITORIAL	Índice de aislamiento territorial calculado por la SUBDERE asociado por la comuna de domicilio familiar y clasificada como: Comuna aislada o Comuna no aislada	String
NOT_DESCRIPCION	Nota del año anterior declarada por el estudiante	Float
CURSO MADRE	Último grado académico cursado por la madre declarada por el estudiante.	String
CURSO PADRE	Último grado académico cursado por el padre declarada por el estudiante	String
UBICORLUGEST_DESC	Ubicación de la vivienda con respecto al lugar de estudios declarado por el estudiante.	String
MEDIOTRANSPORTE_DESC	Medio de transporte utilizado para la movilización hacia el lugar de estudios declarado por el estudiante.	String
COSTOMEDIOTRANSPORTE_DESC	Costo en el que incurre el estudiante para poder movilizarse hacia el lugar de estudios declarado por el estudiante.	String
CHISOL	Sistema Chile Solidario que promueve la incorporación de familias y personas en situación de extrema pobreza a	String
DESC_SINAE	Clasificación de prioridades según las siguientes categorías: Pobreza Rural Problemas de Rendimiento y deserción en Ed.Media Deserción Escolar en Ed.Media Mal rendimiento en Ed.Media Sin calculo de prioridad	String

Tenemos entonces que, en el 2017, 5.423 estudiantes fueron beneficiados con el PRFE según se muestra en la siguiente tabla:

Tabla 56: Estado de estudiantes del PRFE en el año 2017

Fuente: Elaboración propia

DESC_ESTADO	N° DE ESTUDIANTES
Becado	5423
No Becado	4299
Total general	9722

De los estudiantes becados, se desprende la siguiente información:

- 55,25% no pertenecen a ningún pueblo originario, mientras que el 26,98% pertenece al pueblo originario Mapuche.
- El 93,64% no tiene RSH y el 5,07% pertenece al tramo del 40% más vulnerable o menores ingresos. Esto se explica, ya que el instrumento de medición del RSH nace recién a partir del año 2016, y para el PRFE no fue requisito sino hasta el 2018, tener vigente el RSH.
- El 58,40% vive en una comuna aislada, según la clasificación que se le otorga a partir del índice de aislamiento territorial agregado para efectos de este análisis.

- Con respecto a la variable OFERTA EDUCACIONAL, se desprende que el 48,88% de los estudiantes pertenecen a una comuna de origen sin oferta o poca oferta educacional.
- El 33,72% de los estudiantes sí pertenecen al Sistema Chile Solidario (CHISOL).
- La clasificación de primera prioridad del SINAIE denominada *Pobreza Rural* contiene al 47,62% de los estudiantes.

8.1.2 Resultados Obtenidos

8.1.2.1 Aplicación de los modelos predictivos con el Dataset PRFE 2017

Los resultados del rendimiento de los modelos arrojan que el mejor modelo para la predicción es el Gradient Boosting con una exactitud de entrenamiento del 73.18% y con una exactitud validación del 71,60%. Seguido de la Regresión Logística con una exactitud de entrenamiento del 73,70% y una exactitud de validación del 71,31%. La siguiente tabla resume los resultados obtenido con respecto a la exactitud de validación, indicador que sirve para selección del mejor modelo según su desempeño:

Tabla 57: Estado de estudiantes del PRFE en el año 2017
Fuente: Elaboración propia

Modelo	Accuracy de Validación
Gradient Boosting	71,60%
Regresión Logística	71,31%
Árboles de Decisión (cv=3)	70,86%
Random Forest	70,63%
K-NN(5)	69,00%

Finalmente, al resultar mejor modelo el Gradient Boosting, se puede determinar un perfil idóneo de estudiante beneficiario con las variables seleccionadas por el modelo con mayor importancia. En la tabla que se muestra a continuación, se indican las variables más importantes señaladas por el modelo:

Tabla 58: Estado de estudiantes del PRFE en el año 2017
Fuente: Elaboración propia

VARIABLE	VALOR	IMPORTANCIA
UBICORLUGEST_DESC	"Menos de 1 hora de traslado"	23,41%
COSTOMEDIOTRANSPORTE_DESC	"Es menor igual a 300 pesos"	10,73%
OFERTA EDUCACIONAL	"Comuna no prioritaria"	9,08%
UBICORLUGEST_DESC	"Fuera de la región"	7,17%
CHISOL	"No pertenece a Chisol"	5,58%
OFERTA EDUCACIONAL	"Con oferta casi inexistente"	5,51%
CHISOL	"Sin información"	5,10%
DESC_SINAE	"Pobreza rural"	4,09%

Un estudiante con información de la ubicación desde su vivienda hasta el lugar de estudios sea menos de 1 hora de traslado o fuera de la región, que el costo del medio de transporte que utilizan para trasladarse al lugar de estudios es menor o igual a 300 pesos chilenos, que viva en una comuna prioritaria según la nula o casi inexistente oferta de establecimientos educacionales, que no pertenezca al Sistema Chile Solidarios y que su clasificación según SINAIE sea *pobreza rural*, tiene una alta probabilidad de obtener el beneficio del PRFE.

8.1.2.2 Comparación de los resultados de asignación oficial año 2017 versus la predicción

Realizando una comparación de los resultados de la asignación del PRFE del año 2017, y del perfil idóneo del beneficiario que arroja el modelo predictivo aplicado con estos datos, se obtienen los siguientes resultados.

- Modelo predictivo: Gradient Boosting
- Muestra: 1750 estudiantes
- Resultados:

Tabla 59: Tabla comparativa resultados de asignación PRFE 2017
Fuente: Elaboración propia

Observación	AD	Predicción	
		Becados	No Becados
	Becados	805	169
	No Becados	328	448

Como se puede observar, el modelo predictor indica que 805 estudiantes con estado *Becados* (según asignación oficial PRFE 2017), se clasifican correctamente como *Becados*. Mientras que 448 estudiantes con estado oficial *No Becado*, el modelo los clasifica correctamente como *No Becados*.

8.2 Definición de Beneficios Y Costos

En este apartado se definirán los beneficios del proyecto como los ahorros que este genera para la institución si es implementado, así como los costos en los que se incurren.

8.2.1 Ahorros del Proyecto

Este proyecto como está enfocado en optimizar la asignación del Programa de Residencia Familiar Estudiantil, rediseñando el proceso de gestión de asignación, no

genera ingresos monetarios para la institución, pero si genera ahorro en el tiempo de ejecución de las actividades de este proceso.

Para calcular los ahorros que este proyecto genera, se determina el tiempo en minutos aproximados que tomaría realizar cada actividad del proceso de asignación si este proyecto se lleva a cabo. La siguiente tabla detalla este tiempo:

Tabla 60: Tiempo de respuesta para cada actividad si se realiza el proyecto
Fuente: Elaboración propia

Actividades	Tiempo en minutos con proyecto
Consolidar información, recepción y revisión documentación	180
Validación de formularios de postulantes y renovantes	2
Constitución y sesión Comisión Regional de Selección	1
Publicación de Resultados	40
Total minutos	223
Total horas	3,72

Asimismo, para valorizar este ahorro de tiempo se utilizó el valor hora hombre del profesional encargado del programa según la siguiente información:

Tabla 61: Valorización \$/HH
Fuente: Elaboración propia

Horas al mes	176
Honorarios x hora	\$5.398
Sueldo promedio mensual encargado del Programa	\$950.000

Cabe indicar que anualmente el sueldo del encargado del programa a nivel nacional incrementa de acuerdo con el valor porcentual del IPC, por lo que se considera para este proyecto un aumento del 5% cada año.

8.2.2 Costos del Proyecto

Con respecto a los costos del proyecto será necesario para el desarrollo y puesta en marcha se debe tener en cuenta lo siguiente:

- Para poder desarrollar la aplicación web anteriormente mencionada, es necesario la contratación de un Desarrollador, el cual estará durante 4 meses para las tareas relacionadas al desarrollo y pruebas de esta.
- Para la implementación de este proyecto, es importante realizar actividades de capacitación. La principal estará enmarcada en la elaboración de una Jornada en la que participarán las contrapartes de este beneficio de todas las regiones del país.

La valorización de estos 2 ítems se muestra en la siguiente tabla:

Tabla 62: Costos estimados del proyecto
Fuente: Elaboración propia

Costo desarrollador por proyecto	\$6.000.000 (x 6 meses)
Jornada de capacitación (implementación del proyecto)	\$ 2.000.000

8.3 Flujo de Caja

Para la elaboración de este flujo de caja, se deben tener en cuenta las siguientes consideraciones:

- Los cupos disponibles para este programa están decreciendo debido a la naturaleza de este, esto significa que la brecha de acceso a la educación está disminuyendo y por tanto la demanda tiende a la baja. Asimismo, cada año y de acuerdo con las políticas del gobierno actual, el presupuesto asignado para este programa solo incrementa de acuerdo con IPC lo que no permite un aumento en la cantidad anual de beneficios para asignar. Dado lo anterior, y según el análisis del comportamiento de los cupos realizado, existe una disminución promedio del 8% anual.
- El horizonte de evaluación de este proyecto será de 5 años, ya que se estima que la mayor cantidad de beneficiarios serán asignados durante esos años. Cabe indicar que, a partir del año 6, el decrecimiento del número de beneficiarios es mínima, por lo que el Flujo de Capitales será constante. Es importante mencionar que existe una alta probabilidad que para el año 8 este programa se fusione con otro programa social impulsado por el gobierno actual, lo que implica que éste ya no mantenga las mismas condiciones y características vigentes o que en su defecto desaparezca de la oferta de programas.
- Se considerará una evaluación privada del proyecto, por lo que, para determinar la tasa de descuento más adecuada, se utilizará la fórmula del CAPM con los siguientes datos sobre el riesgo asociado:

Tabla 63: Tasas sobre el riesgo asociado al proyecto
Fuente: Elaboración propia

CAPM	14%
R _{lr}	5,18%
R _m	14,40%
B (educación)	1,05

Para este cálculo se consideró como la tasa libre de riesgo el 5.18% anual nominal, según el Banco Central en los últimos 10 años. Para el riesgo del mercado se utilizó el 14.40% nominal anual según IPSA en los últimos 25 años.

Finalmente, para el β o el coeficiente de riesgo sistemático se utilizará el factor 1.05¹² asociado al sector Educación, puesto que se considera oportuno asumir este coeficiente ya que el PRFE está enfocado en asegurar que los beneficiados ingresen, permanezcan y terminen sus estudios secundarios y superiores.

Con todo lo anteriormente descrito, el flujo de caja del proyecto se muestra en la siguiente tabla:

Tabla 64: Flujo de Caja del Proyecto
Fuente: Elaboración propia

		0	1	2	3	4	5
Ahorros	\$ asignación		\$ 2.395.455	\$ 2.028.363	\$ 1.866.676	\$ 1.717.924	\$ 1.581.072
Costos Fijos			\$ -	\$ -	\$ -	\$ -	\$ -
Flujo Operacional			\$ 2.395.455	\$ 2.028.363	\$ 1.866.676	\$ 1.717.924	\$ 1.581.072
Inversión Inicial							
	Capacitación	\$ 2.000.000					
	RRHH (Desarrollador)	\$ 6.000.000					
Flujo de Capitales		-\$ 8.000.000	\$ 2.395.455	\$ 2.028.363	\$ 1.866.676	\$ 1.717.924	\$ 1.581.072
Valor Residual							\$3.615.938
Flujo de Caja Total		-\$ 8.000.000	\$ 2.395.455	\$ 2.028.363	\$ 1.866.676	\$ 1.717.924	\$ 5.197.010

Es importante mencionar que al final del año 5 se presenta un Valor Residual de \$3.615.938, porque de acuerdo con las proyecciones del comportamiento de los beneficiarios con respecto al uso del beneficio, se determina que este generará flujos de efectivo por 3 años más.

Para medir la rentabilidad de este proyecto se calcula la tasa interna de retorno (TIR) y el Valor Actual Neto (VAN) cuyos resultados mostrados a continuación indican que el proyecto es rentable bajo las condiciones anteriormente descritas.

Tabla 65: Indicadores de Rentabilidad
Fuente: Elaboración propia

TIR	17%
VAN	\$384.204

¹² Fuente: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/Betas.html

8.4 Análisis de Sensibilidad

Para el análisis de sensibilidad se utilizan como parámetros el Valor Actual Neto (VAN) y la Tasa Interna de Retorno (TIR) para determinar hasta que valores con respecto al porcentaje de disminución de cobertura y el porcentaje de incremento del costo en HH (sueldo de profesional de becas) puede presentar este proyecto para que siga siendo rentable. Las siguientes tablas muestran las variaciones de acuerdo con el VAN y la TIR:

Tabla 66: Variaciones VAN

Fuente: Elaboración propia

VAN	Disminución % cobertura				
	\$384.204	8%	10%	12%	14%
Incremento % costo HH	5,0%	\$384.204	-\$48.090	-\$451.512	-\$827.638
	5,5%	\$419.200	-\$15.152	-\$420.496	-\$798.412
	6,0%	\$454.196	\$17.785	-\$389.479	-\$769.187
	6,5%	\$489.192	\$50.722	-\$358.463	-\$739.962
	7,0%	\$524.188	\$83.660	-\$327.447	-\$710.736

Tabla 67: Variaciones TIR

Fuente: Elaboración propia

TIR	Disminución % cobertura				
	17%	8%	10%	12%	14%
Incremento % costo HH	5,0%	16,91%	14,60%	12,28%	9,97%
	5,5%	17,10%	14,78%	12,46%	10,15%
	6,0%	17,28%	14,96%	12,64%	10,32%
	6,5%	17,46%	15,14%	12,82%	10,50%
	7,0%	17,65%	15,32%	13,00%	10,67%

CAPÍTULO 9: CONCLUSIONES

Este capítulo contiene la información de cierre del trabajo realizado, en el que se resumen los resultados obtenidos del proyecto, los principales hallazgos y conclusiones relacionadas, factores fuera de alcance que son importantes mencionar y las recomendaciones para trabajos futuros por parte de la institución, en el caso de implementar el proyecto y querer mejorarlo.

9.1 Trabajo realizado

Este trabajo propone inicialmente rediseñar el proceso de gestión de asignación del Programa de Residencia Familiar Estudiantil, simplificando el modelo de focalización vigente con el fin de maximizar la correcta asignación del beneficio, ya que se evidencian varias problemáticas relacionadas con el subproceso de asignación del beneficio, el trabajo manual en las revisiones documentales de la información proporcionada por los estudiantes y los tiempo en los que se toman los responsables para ejecutar correctamente este proceso.

Se efectúan distintas actividades que ayudan a cumplir con el propósito planteado este trabajo, en el que se analiza y levantan los procesos involucrados para la asignación final del PRFE, logrando visibilizar las inconsistencias del subproceso de asignación del beneficio, ya que no se encuentra estandarizado y toma mucho tiempo realizar la validación de formularios de postulantes y renovantes. Esto último, constituye el subproceso más importante para asignar correctamente el beneficio.

Otra acción realizada es la identificación las variables del modelo de focalización vigente con menor significancia, lo que ayuda a proponer uno simplificado que cuente con información comprobable y no declarativa, tal y como se obtenía con el modelo anterior. Para esto se plantea el uso del índice de asilamiento territorial demanda de SUBDERE, el cual nos entrega información relevante para la dimensión territorial del modelo de focalización y sobre la comuna de domicilio familiar del estudiante. Asimismo, se incorpora la información proporcionada por el Ministerio de Desarrollo Social, sobre la calificación socioeconómica del estudiante según el tramo de vulnerabilidad.

Asimismo, se implementa el rediseño del proceso de Asignación del Programa de Residencia Familiar Estudiantil. Por lo anterior, se efectúa la desagregación de las tareas necesarias para cumplir con este subproceso y se expone cada una de las tareas involucradas modelándolos en BPMN evidenciando así los actores involucrados y el flujo que estos siguen. Gracias a esto, se logra proponer una forma en la que el PRFE puede entregar el beneficio mediante una planificación previa que involucra la realización de los planes de convocatoria y aceptación de las entidades ejecutoras,

quienes, a su vez, convocan a las familias tutoras que acogerán a los estudiantes beneficiarios del programa. Además, se programan los aportes monetarios que se entregan a las mismas entidades y se elaboran los planes de supervisión a las casas de las familias tutoras. Esto entrega un trabajo sistematizado de las tareas que cada actor debe realizar durante el año del beneficio.

Se consigue confeccionar varios modelos predictivos, de los cuales se comparan sus indicadores de desempeño para poder escoger el mejor de ellos. Para el caso de PRFE se realizan 2 análisis para revisar el comportamiento de los datos según el modelo de focalización utilizado (el vigente o el simplificado). Resulta que, para el modelo de focalización vigente, el mejor modelo es Regresión Logística, el cual indica una exactitud de validación del 77,51%. Este puntaje se considera óptimo, lo que permite utilizar el modelo para predecir si un estudiante con las características relacionadas con su información académica, cultural, socioeconómica y territorial será o no becado en el programa.

Se obtiene también, un perfil idóneo de un potencial beneficiario PRFE según la importancia de las variables predictoras en el que se demuestra que si un estudiante contiene como mínimo la información de la ubicación desde su vivienda hasta el lugar de estudios, que viva efectivamente en una comuna aislada, que su madre haya cursado como último grado la enseñanza básica incompleta, que su padre haya cursado como último grado la enseñanza básica completa, que el costo del medio de transporte que utilizan para trasladarse al lugar de estudios es más de 2.000 pesos chilenos y cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos.

Para el caso del modelo de focalización simplificado, se presenta una exactitud del 73,74% para el modelo Regresión Logística, lo que presume que éste es el mejor modelo para aplicar en el Programa de Residencia Familiar Estudiantil (PRFE). Además, la propuesta de perfil idóneo de un potencial beneficiario depende de que un estudiante que viva efectivamente en una comuna aislada, cuyo hogar esté formado por un progenitor (madre o padre) y uno o varios hijos, que pertenezca al 40% de la población más vulnerable o que no haya inscrito su información en el registro social de hogares, tiene una alta probabilidad de obtener el beneficio del PRFE.

En conclusión, se evalúan los resultados realizando un análisis retrospectivo utilizando datos de la asignación del programa correspondiente al año 2017, con lo que se puede concluir que de una muestra de 1.750 estudiantes 1.253 se clasifican correctamente en la asignación, dando así una gran proporción de resultados fidedignos que comprueban que el uso del modelo predictivo es una herramienta confiable para la correcta asignación.

Con respecto a la evaluación económica los resultados obtenidos son que el proyecto se implementa según la propuesta descrita en este documento, se evidencia que se generaría un ahorro en cuanto al costo de hora hombre identificado, ya que se reduce el tiempo de ejecución de las actividades asociadas al proceso de asignación gracias a la utilización de modelos predictivos y automatización de otras tareas. Además, genera una tasa interna de retorno del 19% considerándose así, un proyecto rentable para la institución.

Para finalizar, es importante recalcar que el proyecto no se ha implementado dentro de la institución, por lo que todo el análisis realizado se ha enfocado en la revisión de datos de años anteriores para evaluar el comportamiento del programa frente al rediseño y mejoras propuestas en este documento.

9.2 Recomendaciones para trabajos futuros

Refiriéndose a los próximos trabajos que pueden realizarse con el proyecto implementado se indican algunas recomendaciones:

- Para lograr obtener datos confiables de los estudiantes relacionados con su información socioeconómica, deben generarse acuerdos o convenios de interoperatividad con otras instituciones del estado que contengan este tipo de información. Si esto se realiza, se podrían descargar base datos que sirvan de insumo para los modelos predictivos y así obtener una mejor exactitud que la ya presentada.
- La evaluación de un modelo de focalización simplificado generaría mayor oportunidad al momento de la postulación y renovación del beneficio, por lo que, en conjunto con una revisión normativa, se podría considerar la disminución de los requisitos para acceder al programa.
- La gestión del cambio debe ser considerada como fundamental dentro de cualquier institución pública, por lo que debe ser contempladas dentro de la planificación estratégica institucional, ya que sin compromiso de los actores claves, no se puede implementar ningún proyecto. Es conocido que todos los proyectos de impliquen mejoras, requieren de una planificación integral para su puesta en marcha.
- La implementación de una plataforma web que permita obtener información sobre la predicción de la asignación es fundamental, por ende, realizar una integración con el sistema de becas actual sería ideal para que el encargado del programa pueda planificar la gestión de su beneficio de la mejor manera.

CAPÍTULO 10: BIBLIOGRAFÍA

- Consultoría Estratégica. (s.f.). Obtenido de <http://cirpq.tripod.com/consultoriaestrategica/modelodelta.pdf>
- Dirección de Presupuesto. (s.f.). Obtenido de http://www.dipres.gob.cl/595/articulos-157551_doc_pdf.pdf
- Gobierno Transparente. (s.f.). Obtenido de https://www.agci.cl/gobiernotransparente/solicitud_informacion.html
- JUNAEB. (s.f.). Obtenido de <http://www.junaeb.cl>
- Junta Nacional de Auxilio Escolar y Becas (JUNAEB). (2005). INFORME DE CUMPLIMIENTOS DE COMPROMISOS AL 30-06-2016. INFORME DE CUMPLIMIENTO, DIRECCIÓN DE PRESUPUESTOS.
- Junta Nacional de Auxilio Escolar y Becas (JUNAEB). (19 de abril de 2017). Quienes Somos: JUNAEB. Recuperado el 19 de abril de 2017, de Sitio Web de JUNAEB: <https://www.junaeb.cl/quienes-somos>
- Hax, A; Ugarte, Juan José; (2014). Hacia la Gran Universidad Chilena, Un modelo de transformación estratégica. Santiago, Chile: Ediciones Universidad Católica de Chile.
- Becas, J. N. (21 de 02 de 2017). Gobierno Transparente. Obtenido de <http://anfitrion.cl/GobiernoTransparente/junaeb/2017/otras/382.pdf>
- ¿Qué es un FRAMEWORK? (s.f.). *Lenguajes de Programación*. Obtenido de <https://lenguajesdeprogramacion.net/diccionario/que-es-un-framework/>
- ¿Qué son las SVM? (23 de Marzo de 2008). Obtenido de Optimization & Machine Learning: <http://optimachine.blogspot.com/2008/03/qu-son-las-svm.html>
- *Analytics Lane*. (02 de Julio de 2018). Obtenido de <https://www.analyticslane.com/2018/07/02/gridsearchcv/>
- Barros, O. (2011). DISEÑO INTEGRADO DE NEGOCIOS, PROCESOS Y APLICACIONES TI. En O. Barros, *INGENIERÍA DE NEGOCIOS*. Universidad de Chile.

- Breiman, L. (2001). Random Forests. *Machine Learning*(45), 5–32.
- Bronshtein, I., & Semendiaev, K. (1973). *Manual de Matemáticas para Ingenieros y Estudiantes*. Moscú: MIR Moscú.
- Brownlee, J. (12 de Agosto de 2019). Recuperado el Abril de 2020, de Machine Learning Mastery: <https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/>
- Departamento de Becas JUNAEB. (2018). *Manual de Procedimientos Programa Residencia Familiar Estudiantil*. Santiago.
- División de Políticas y Estudios Departamento de Estudios y Evaluación. (2012). *ESTUDIO IDENTIFICACIÓN DE LOCALIDADES EN CONDICIONES DE AISLAMIENTO 2012*. Santiago.
- Django Software, F. (2019). *Django Versión 2.2*. Obtenido de <https://djangoproject.com>
- D'Souza, J. (21 de Marzo de 2018). *A Quick Guide to Boosting in ML*. Obtenido de Medium: <https://medium.com/greyatom/a-quick-guide-to-boosting-in-ml-acf7c1585cb5>
- Fernando Cartes, E. C. (1999). <http://repositorio.uchile.cl>. Obtenido de http://repositorio.uchile.cl/bitstream/handle/2250/125319/Cartes_Fernando.pdf?sequence=1
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. San Francisco.
- Hax, A. C. (2010). *The Delta Model, Reinventing Your Business Strategy*. Massachusetts: Springer.
- Instituto de Previsión Social. (2019). *Chile Solidario*. Obtenido de <https://www.ips.gob.cl/servlet/internet/content/1421810829144/chile-solidario>
- Junta Nacional de Auxilio Escolar y Becas. (2005). *Sistema Nacional de Asignación con Equidad para Becas JUNAEB*. Santiago.
- K-NN, C. (s.f.). Obtenido de Intelligent Systems Group: <http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t9knn.pdf>
- Ministerio de Desarrollo Social, Subsecretaría de Evaluación Social. (2014). *Observatorio*. Obtenido de Ministerio de Desarrollo Social:

http://observatorio.ministeriodesarrollosocial.gob.cl/documentos/Pobreza_Multi_dimensional_Chile_heidi_Berner.pdf

- RSH_Paper. (2018). *Registro Social de Hogares de Chile*. Obtenido de Ministerio de Desarrollo Social, Subsecretaría de Evaluación Social: http://www.registrosocial.gob.cl/wp-content/uploads/2015/09/RSH_paper_2.pdf
- Subsecretaría de Desarrollo Regional y Administrativo – SUBDERE Ministerio del Interior y Seguridad Pública. (2011). *Estudio Identificación de Territorios Aislados 2011*. Santiago, Chile: Departamento de Estudios y Evaluación, Unidad de Análisis Territorial.
- Support Vector Machines for Classification. (2015). En M. A. Mariette Awad, *Efficient Learning Machines* (págs. 39-66).
- Universidad de Granada. (2014-2015). Obtenido de El espacio Euclideo Lección 1: <https://www.ugr.es/~rpaya/documentos/AnalisisI/2014-15/Euclideo.pdf>
- Waissbluth, M. (2008). Obtenido de Logotec: http://www.logotec.cl/descargas/gestion_del_cambio.pdf
- WHITE, S. A. (2009). *Guía de Referencia y Modelado BPMN. Comprendiendo y utilizando BPMN*. Lighthouse Point, Florida, Estados Unidos: Future Strategies Inc.

CAPÍTULO 11: ANEXOS

11.1 Tabla de Índice de Aislamiento por Región

REGION	NOMBRE REGION	COMUNA	PROVINCIA	AISLAMIENTO ESTRUCTURAL	GRADO DE INTEGRACIÓN	ÍNDICE DE AISLAMIENTO
15	Arica y Parinacota	General Lagos	Parinacota	0,4302	0,1294	-0,1714
15	Arica y Parinacota	Putre	Parinacota	0,4341	0,47	0,506
15	Arica y Parinacota	Camarones	Arica	0,5154	0,5176	0,5198
15	Arica y Parinacota	Arica	Arica	0,4047	0,9502	1,4957
1	Tarapacá	Colchane	Tamarugal	0,4075	0,1014	-0,2047
1	Tarapacá	Camiña	Tamarugal	0,4344	0,3066	0,1788
1	Tarapacá	Pica	Tamarugal	0,5596	0,5226	0,4856
1	Tarapacá	Huara	Tamarugal	0,5088	0,6089	0,7091
1	Tarapacá	Pozo Almonte	Tamarugal	0,4406	0,6934	0,9462
1	Tarapacá	Alto Hospicio	Iquique	0,3448	0,8313	1,3177
1	Tarapacá	Iquique	Iquique	0,4011	1	1,5989
2	Antofagasta	Ollagüe	El Loa	0,4173	0,1245	-0,1682
2	Antofagasta	Taltal	Antofagasta	0,405	0,4407	0,4763
2	Antofagasta	San Pedro de Atacama	El Loa	0,4902	0,4952	0,5003
2	Antofagasta	Atacama	El Loa	0,4902	0,4952	0,5003

2	Antofagasta	Sierra Gorda	Antofagasta	0,4129	0,5027	0,5925
2	Antofagasta	Mejillones	Antofagasta	0,4008	0,7291	1,0573
2	Antofagasta	Tocopilla	Tocopilla	0,4261	0,8322	1,2384
2	Antofagasta	María Elena	Tocopilla	0,3029	0,7757	1,2484
2	Antofagasta	Antofagasta	Antofagasta	0,4143	0,9832	1,5521
2	Antofagasta	Calama	El Loa	0,3748	0,9909	1,6069
3	Atacama	Alto del Carmen	Huasco	0,5386	0,5953	0,652
3	Atacama	Freirina	Huasco	0,469	0,6653	0,8616
3	Atacama	Chañaral	Chañaral	0,4085	0,6483	0,8881
3	Atacama	Huasco	Huasco	0,4673	0,6835	0,8998
3	Atacama	Caldera	Copiapó	0,4212	0,7481	1,075
3	Atacama	Diego de Almagro	Chañaral	0,3338	0,7161	1,0984
3	Atacama	Vallenar	Huasco	0,4607	0,8037	1,1468
3	Atacama	Tierra Amarilla	Copiapó	0,3527	0,765	1,1774
3	Atacama	Copiapó	Copiapó	0,3905	0,9299	1,4692
4	Coquimbo	Canela	Choapa	0,492	0,5698	0,6475
4	Coquimbo	Combarbalá	Limarí	0,5211	0,5979	0,6747
4	Coquimbo	Paihuano	Elqui	0,4436	0,5595	0,6753
4	Coquimbo	La Higuera	Elqui	0,4451	0,6041	0,763

4	Coquimbo	Los Vilos	Choapa	0,4747	0,6549	0,8351
4	Coquimbo	Río Hurtado	Limarí	0,5148	0,679	0,8431
4	Coquimbo	Andacollo	Elqui	0,448	0,6536	0,8591
4	Coquimbo	Salamanca	Choapa	0,4095	0,6458	0,8821
4	Coquimbo	Vicuña	Elqui	0,4139	0,6667	0,9195
4	Coquimbo	Punitaqui	Limarí	0,4938	0,7221	0,9505
4	Coquimbo	Monte Patria	Limarí	0,4047	0,6995	0,9943
4	Coquimbo	Illapel	Choapa	0,4699	0,741	1,0121
4	Coquimbo	Ovalle	Limarí	0,4164	0,8475	1,2786
4	Coquimbo	Coquimbo	Elqui	0,4293	0,9127	1,3961
4	Coquimbo	La Serena	Elqui	0,4445	0,9719	1,4993
5	Valparaíso	Petorca	Petorca	0,5004	0,4954	0,4904
5	Valparaíso	Papudo	Petorca	0,4102	0,4902	0,5703
5	Valparaíso	Zapallar	Petorca	0,3956	0,4887	0,5817
5	Valparaíso	Cabildo	Petorca	0,4223	0,5825	0,7427
5	Valparaíso	El Quisco	San Antonio	0,4965	0,6441	0,7916
5	Valparaíso	Algarrobo	San Antonio	0,4232	0,6243	0,8254
5	Valparaíso	Rinconada	Los Andes	0,3804	0,6067	0,8329
5	Valparaíso	Catemu	San Felipe	0,3188	0,5826	0,8463
5	Valparaíso	Santa María	San Felipe	0,3795	0,6194	0,8592

5	Valparaíso	La Ligua	Petorca	0,4296	0,6676	0,9056
5	Valparaíso	Llay Llay	San Felipe	0,3083	0,6112	0,9141
5	Valparaíso	Puchuncaví	Valparaíso	0,3088	0,6119	0,915
5	Valparaíso	El Tabo	San Antonio	0,4857	0,7019	0,9181
5	Valparaíso	Panquehue	San Felipe	0,2969	0,6378	0,9786
5	Valparaíso	Quintero	Valparaíso	0,296	0,6404	0,9848
5	Valparaíso	Olmué	Marga Marga	0,261	0,6414	1,0218
5	Valparaíso	Hijuelas	Quillota	0,2676	0,6549	1,0421
5	Valparaíso	Cartagena	San Antonio	0,4083	0,7345	1,0607
5	Valparaíso	Calle Larga	Los Andes	0,3118	0,7126	1,1134
5	Valparaíso	Putendo	San Felipe	0,4127	0,7748	1,1369
5	Valparaíso	Santo Domingo	San Antonio	0,3148	0,7381	1,1615
5	Valparaíso	La Cruz	Quillota	0,3015	0,7398	1,178
5	Valparaíso	San Esteban	Los Andes	0,3281	0,7534	1,1786
5	Valparaíso	Nogales	Quillota	0,2681	0,7361	1,2041
5	Valparaíso	Limache	Marga Marga	0,2424	0,7897	1,3371
5	Valparaíso	Casablanca	Valparaíso	0,335	0,8424	1,3498
5	Valparaíso	Calera	Quillota	0,2346	0,7988	1,363
5	Valparaíso	San Felipe	San Felipe	0,3297	0,8763	1,423

5	Valparaíso	San Antonio	San Antonio	0,2718	0,8581	1,4443
5	Valparaíso	Los Andes	Los Andes	0,3297	0,892	1,4544
5	Valparaíso	Quillota	Quillota	0,244	0,887	1,5301
13	Metropolitana	Alhué	Melipilla	0,3802	0,4699	0,5597
13	Metropolitana	San Pedro	Melipilla	0,3946	0,5184	0,6423
13	Metropolitana	Tiltil	Chacabuco	0,3129	0,6635	1,0141
13	Metropolitana	Curacaví	Melipilla	0,2775	0,6812	1,0849
13	Metropolitana	Calera de Tango	Maipo	0,2636	0,6972	1,1308
13	Metropolitana	Paine	Maipo	0,2433	0,7061	1,169
13	Metropolitana	El Monte	Talagante	0,2076	0,694	1,1804
13	Metropolitana	María Pinto	Melipilla	0,2631	0,7263	1,1895
13	Metropolitana	Padre Hurtado	Talagante	0,2195	0,7134	1,2073
13	Metropolitana	San José de Maipo	Cordillera	0,3779	0,7982	1,2185
13	Metropolitana	Pirque	Cordillera	0,2585	0,7528	1,2471
13	Metropolitana	Isla de Maipo	Talagante	0,2619	0,7568	1,2516
13	Metropolitana	Colina	Chacabuco	0,2255	0,7481	1,2707
13	Metropolitana	Lampa	Chacabuco	0,2585	0,8035	1,3484
13	Metropolitana	Buin	Maipo	0,2033	0,816	1,4287
13	Metropolitana	Talagante	Talagante	0,2056	0,8219	1,4382

13	Metropolitana	Melipilla	Melipilla	0,2205	0,8414	1,4624
13	Metropolitana	Peñaflor	Talagante	0,207	0,8541	1,5011
6	O'Higgins	Navidad	Cardenal Caro	0,5784	0,4069	0,2354
6	O'Higgins	La Estrella	Cardenal Caro	0,4354	0,3768	0,3181
6	O'Higgins	Paredones	Cardenal Caro	0,5845	0,4647	0,3449
6	O'Higgins	Pumanque	Colchagua	0,4536	0,415	0,3764
6	O'Higgins	Pichilemu	Cardenal Caro	0,6252	0,5936	0,562
6	O'Higgins	Litueche	Cardenal Caro	0,4266	0,5057	0,5848
6	O'Higgins	Lolol	Colchagua	0,4671	0,5445	0,6218
6	O'Higgins	Marchihue	Cardenal Caro	0,4531	0,5819	0,7106
6	O'Higgins	Las Cabras	Cachapoal	0,3544	0,6402	0,926
6	O'Higgins	Peralillo	Colchagua	0,3897	0,6636	0,9375
6	O'Higgins	Pichidegua	Cachapoal	0,3205	0,6379	0,9552
6	O'Higgins	Coínco	Cachapoal	0,3049	0,6416	0,9782
6	O'Higgins	Coltauco	Cachapoal	0,2693	0,6456	1,022
6	O'Higgins	Chépica	Colchagua	0,3094	0,6669	1,0245
6	O'Higgins	Palmilla	Colchagua	0,3137	0,685	1,0562
6	O'Higgins	Peumo	Cachapoal	0,2815	0,704	1,1266

6	O'Higgins	Nancagua	Colchagua	0,2894	0,708	1,1267
6	O'Higgins	Mostazal	Cachapoal	0,2156	0,6857	1,1558
6	O'Higgins	Doñihue	Cachapoal	0,2392	0,6995	1,1598
6	O'Higgins	Chimbarongo	Colchagua	0,276	0,7228	1,1697
6	O'Higgins	Codegua	Cachapoal	0,2094	0,6909	1,1725
6	O'Higgins	Placilla	Colchagua	0,2562	0,7242	1,1922
6	O'Higgins	Quinta De Tilcoco	Cachapoal	0,2306	0,713	1,1953
6	O'Higgins	Santa Cruz	Colchagua	0,3631	0,7874	1,2117
6	O'Higgins	El Olivar	Cachapoal	0,1625	0,6885	1,2144
6	O'Higgins	Malloa	Cachapoal	0,2376	0,7358	1,2341
6	O'Higgins	Machalí	Cachapoal	0,1883	0,7174	1,2465
6	O'Higgins	San Vicente	Cachapoal	0,2834	0,7666	1,2499
6	O'Higgins	Graneros	Cachapoal	0,1823	0,7392	1,2961
6	O'Higgins	Requínoa	Cachapoal	0,1796	0,7626	1,3457
6	O'Higgins	Rengo	Cachapoal	0,2171	0,8008	1,3844
6	O'Higgins	San Fernando	Colchagua	0,2157	0,8536	1,4914
6	O'Higgins	Rancagua	Cachapoal	0,1615	0,9116	1,6617
7	Maule	Vichuquén	Curicó	0,4599	0,3938	0,3278
7	Maule	Pelluhue	Cauquenes	0,4917	0,4858	0,4799

7	Maule	Licantén	Curicó	0,3926	0,466	0,5395
7	Maule	Hualañé	Curicó	0,3781	0,4735	0,5689
7	Maule	Chanco	Cauquenes	0,4072	0,5006	0,594
7	Maule	Curepto	Talca	0,3707	0,4942	0,6178
7	Maule	Empedrado	Talca	0,289	0,5216	0,7541
7	Maule	San Rafael	Talca	0,2475	0,5575	0,8674
7	Maule	Río Claro	Talca	0,2546	0,5798	0,9051
7	Maule	Colbún	Linares	0,2767	0,6514	1,026
7	Maule	Sagrada Familia	Curicó	0,2374	0,6507	1,0639
7	Maule	Teno	Curicó	0,2946	0,6943	1,094
7	Maule	Rauco	Curicó	0,2817	0,7014	1,121
7	Maule	Romeral	Curicó	0,2782	0,7118	1,1454
7	Maule	Pelarco	Talca	0,2209	0,6866	1,1523
7	Maule	Longaví	Linares	0,2656	0,7129	1,1603
7	Maule	Cauquenes	Cauquenes	0,2955	0,7309	1,1663
7	Maule	Retiro	Linares	0,2548	0,7123	1,1698
7	Maule	Constitución	Talca	0,2726	0,7252	1,1777
7	Maule	Pencahue	Talca	0,2316	0,7108	1,1901
7	Maule	Villa Alegre	Linares	0,2672	0,7288	1,1905

7	Maule	San Clemente	Talca	0,2303	0,7119	1,1936
7	Maule	Molina	Curicó	0,2314	0,7127	1,1939
7	Maule	Yerbas Buenas	Linares	0,2247	0,7239	1,223
7	Maule	Maule	Talca	0,2452	0,7362	1,2272
7	Maule	San Javier	Linares	0,2575	0,7731	1,2888
7	Maule	Parral	Linares	0,2639	0,7962	1,3286
7	Maule	Linares	Linares	0,2301	0,8579	1,4857
7	Maule	Curicó	Curicó	0,2496	0,8701	1,4906
7	Maule	Talca	Talca	0,1825	0,8727	1,5629
8	Biobío	Tirúa	Arauco	0,6113	0,2415	-0,1282
8	Biobío	Alto Biobío	Biobío	0,5982	0,3003	0,0024
8	Biobío	Antuco	Biobío	0,583	0,4407	0,2984
8	Biobío	Cobquecura	Ñuble	0,4746	0,3967	0,3187
8	Biobío	Contulmo	Arauco	0,4548	0,4945	0,5342
8	Biobío	San Fabián	Ñuble	0,5479	0,5436	0,5394
8	Biobío	Yungay	Ñuble	0,4466	0,5243	0,602
8	Biobío	Tucapel	Biobío	0,4545	0,5462	0,6379
8	Biobío	Florida	Concepción	0,3963	0,5176	0,6388
8	Biobío	Lebu	Arauco	0,3528	0,5064	0,66

8	Biobío	Quillón	Ñuble	0,4622	0,5611	0,66
8	Biobío	Pemuco	Ñuble	0,4073	0,5342	0,6611
8	Biobío	Quirihue	Ñuble	0,3948	0,5295	0,6641
8	Biobío	Trehuaco	Ñuble	0,3285	0,516	0,7036
8	Biobío	Ránquil	Ñuble	0,435	0,5716	0,7083
8	Biobío	Quilleco	Biobío	0,4286	0,5713	0,714
8	Biobío	El Carmen	Ñuble	0,3803	0,5527	0,7251
8	Biobío	Ninhue	Ñuble	0,3703	0,5503	0,7304
8	Biobío	Quilaco	Biobío	0,4172	0,577	0,7368
8	Biobío	Santa Bárbara	Biobío	0,4326	0,5855	0,7384
8	Biobío	Santa Juana	Concepción	0,3142	0,5378	0,7613
8	Biobío	San Ignacio	Ñuble	0,37	0,5703	0,7706
8	Biobío	Portezuelo	Ñuble	0,3367	0,5775	0,8183
8	Biobío	Yumbel	Biobío	0,3798	0,6143	0,8488
8	Biobío	Cañete	Arauco	0,3903	0,6281	0,8659
8	Biobío	Coelemu	Ñuble	0,282	0,5759	0,8698
8	Biobío	Los Álamos	Arauco	0,3302	0,6001	0,87
8	Biobío	San Rosendo	Biobío	0,4198	0,663	0,9063
8	Biobío	Coihueco	Ñuble	0,3844	0,6459	0,9075

8	Biobío	Pinto	Ñuble	0,4124	0,6643	0,9162
8	Biobío	San Nicolás	Ñuble	0,3549	0,636	0,9171
8	Biobío	Laja	Biobío	0,3945	0,6606	0,9267
8	Biobío	Negrete	Biobío	0,3152	0,6285	0,9419
8	Biobío	Cabrero	Biobío	0,3755	0,6622	0,9489
8	Biobío	Nacimiento	Biobío	0,3305	0,6433	0,9561
8	Biobío	Mulchén	Biobío	0,3259	0,6474	0,9688
8	Biobío	Ñiquén	Ñuble	0,3823	0,6828	0,9833
8	Biobío	Bulnes	Ñuble	0,3259	0,6854	1,0449
8	Biobío	Chillán Viejo	Ñuble	0,3491	0,7012	1,0532
8	Biobío	Hualqui	Concepción	0,2366	0,678	1,1194
8	Biobío	Arauco	Arauco	0,2964	0,7195	1,1426
8	Biobío	Curanilahue	Arauco	0,1958	0,6826	1,1694
8	Biobío	San Carlos	Ñuble	0,323	0,8366	1,3503
8	Biobío	Los Ángeles	Biobío	0,2989	0,8597	1,4205
8	Biobío	Chillán	Ñuble	0,2902	0,8879	1,4857
8	Biobío	Tomé	Concepción	0,1992	0,8547	1,5101
8	Biobío	Lota	Concepción	0,1509	0,8528	1,5548
8	Biobío	Coronel	Concepción	0,1736	0,8696	1,5655
9	Araucanía	Melipeuco	Cautín	0,626	0,1122	-0,4017

9	Araucanía	Lonquimay	Malleco	0,6045	0,2933	-0,0178
9	Araucanía	Curarrehue	Cautín	0,5406	0,334	0,1273
9	Araucanía	Toltén	Cautín	0,5201	0,3972	0,2744
9	Araucanía	Saavedra	Cautín	0,5279	0,4669	0,4059
9	Araucanía	Cunco	Cautín	0,4053	0,4726	0,5398
9	Araucanía	Teodoro Schmidt	Cautín	0,3939	0,4957	0,5975
9	Araucanía	Loncoche	Cautín	0,3688	0,4905	0,6121
9	Araucanía	Pucón	Cautín	0,4676	0,5575	0,6473
9	Araucanía	Lumaco	Malleco	0,441	0,5793	0,7175
9	Araucanía	Curacautín	Malleco	0,3804	0,5518	0,7232
9	Araucanía	Carahue	Cautín	0,4297	0,5772	0,7246
9	Araucanía	Los Sauces	Malleco	0,4018	0,566	0,7302
9	Araucanía	Galvarino	Cautín	0,3842	0,5626	0,741
9	Araucanía	Purén	Malleco	0,3709	0,5585	0,7462
9	Araucanía	Vilcún	Cautín	0,3178	0,5429	0,7681
9	Araucanía	Renaico	Malleco	0,4556	0,6175	0,7794
9	Araucanía	Gorbea	Cautín	0,3638	0,5904	0,817
9	Araucanía	Chol-chol	Cautín	0,3594	0,6262	0,893
9	Araucanía	Villarrica	Cautín	0,3137	0,6096	0,9055

9	Araucanía	Perquenco	Cautín	0,319	0,6156	0,9123
9	Araucanía	Collipulli	Malleco	0,3346	0,6532	0,9718
9	Araucanía	Freire	Cautín	0,2924	0,6466	1,0009
9	Araucanía	Pitrufquén	Cautín	0,3306	0,666	1,0014
9	Araucanía	Ercilla	Malleco	0,3146	0,6595	1,0045
9	Araucanía	Nueva Imperial	Cautín	0,3654	0,7096	1,0537
9	Araucanía	Traiguén	Malleco	0,4097	0,7864	1,1631
9	Araucanía	Lautaro	Cautín	0,2895	0,7341	1,1788
9	Araucanía	Padre Las Casas	Cautín	0,2035	0,7174	1,2313
9	Araucanía	Angol	Malleco	0,3924	0,849	1,3057
9	Araucanía	Victoria	Malleco	0,2469	0,8418	1,4367
9	Araucanía	Temuco	Cautín	0,1914	0,9155	1,6397
14	Los Ríos	Corral	Valdivia	0,4027	0,3653	0,3278
14	Los Ríos	Futrono	Ranco	0,4185	0,388	0,3576
14	Los Ríos	Lago Ranco	Ranco	0,3746	0,419	0,4635
14	Los Ríos	Lanco	Valdivia	0,4104	0,5159	0,6215
14	Los Ríos	Panguipulli	Valdivia	0,4145	0,5235	0,6324
14	Los Ríos	Los Lagos	Valdivia	0,3416	0,5087	0,6759
14	Los Ríos	Mariquina	Valdivia	0,3263	0,5694	0,8125

14	Los Ríos	Máfil	Valdivia	0,2995	0,588	0,8765
14	Los Ríos	Paillaco	Valdivia	0,3137	0,6115	0,9093
14	Los Ríos	Río Bueno	Ranco	0,2591	0,6545	1,0498
14	Los Ríos	La Unión	Ranco	0,2385	0,6576	1,0766
14	Los Ríos	Valdivia	Valdivia	0,1807	0,9188	1,6569
10	Los Lagos	Palena	Palena	0,5207	0,4927	0,4647
10	Los Lagos	Futaleufú	Palena	0,5167	0,6118	0,7069
10	Los Lagos	Chaitén	Palena	0,4853	0,6046	0,7238
10	Los Lagos	Puqueldón	Chiloé	0,5117	0,6246	0,7375
10	Los Lagos	Hualaihue	Palena	0,5182	0,6411	0,7641
10	Los Lagos	Cochamó	Llanquihue	0,5379	0,6539	0,7699
10	Los Lagos	San Juan de la Costa	Osorno	0,518	0,7569	0,9957
10	Los Lagos	Quemchi	Chiloé	0,5333	0,7678	1,0024
10	Los Lagos	Curaco de Vélez	Chiloé	0,5615	0,7826	1,0038
10	Los Lagos	Fresia	Llanquihue	0,5013	0,7584	1,0155
10	Los Lagos	Queilén	Chiloé	0,514	0,7857	1,0573
10	Los Lagos	Puyehue	Osorno	0,5304	0,7966	1,0629
10	Los Lagos	San Pablo	Osorno	0,5343	0,8011	1,0679
10	Los Lagos	Mauñín	Llanquihue	0,4809	0,7762	1,0715

10	Los Lagos	Quinchao	Chiloé	0,4904	0,7822	1,074
10	Los Lagos	Calbuco	Llanquihue	0,5091	0,7925	1,0759
10	Los Lagos	Los Muermos	Llanquihue	0,4966	0,7897	1,0829
10	Los Lagos	Puerto Octay	Osorno	0,4898	0,7934	1,097
10	Los Lagos	Dalcahue	Chiloé	0,5516	0,8335	1,1155
10	Los Lagos	Quellón	Chiloé	0,5216	0,8229	1,1241
10	Los Lagos	Chonchi	Chiloé	0,5231	0,8332	1,1433
10	Los Lagos	Purranque	Osorno	0,5223	0,8388	1,1553
10	Los Lagos	Río Negro	Osorno	0,5042	0,833	1,1617
10	Los Lagos	Frutillar	Llanquihue	0,4991	0,8408	1,1825
10	Los Lagos	Llanquihue	Llanquihue	0,4968	0,8529	1,2089
10	Los Lagos	Ancud	Chiloé	0,5501	0,8862	1,2223
10	Los Lagos	Castro	Chiloé	0,5243	0,9552	1,3861
10	Los Lagos	Puerto Varas	Llanquihue	0,5049	0,9575	1,41
10	Los Lagos	Osorno	Osorno	0,5059	0,9865	1,4672
10	Los Lagos	Puerto Montt	Llanquihue	0,4817	0,992	1,5023
11	Aisén	O'Higgins	Capitán Prat	0,5479	0,1647	-0,2186
11	Aisén	Guaitecas	Aisén	0,5026	0,2762	0,0497
11	Aisén	Tortel	Capitán Prat	0,5046	0,3477	0,1907
11	Aisén	Lago Verde	Coihaique	0,4932	0,496	0,4988

11	Aisén	Chile Chico	General Carrera	0,5739	0,6033	0,6326
11	Aisén	Río Ibáñez	General Carrera	0,5478	0,6067	0,6657
11	Aisén	Cochrane	Capitán Prat	0,5147	0,5953	0,676
11	Aisén	Cisnes	Aisén	0,4615	0,708	0,9545
11	Aisén	Aysén	Aisén	0,4834	0,8684	1,2533
11	Aisén	Coyhaique	Coyhaique	0,4684	0,9549	1,4414
12	Magallanes	Cabo de Hornos	Antártica Chilena	0,3818	0,1007	-0,1804
12	Magallanes	Timaukel	Tierra del Fuego	0,5893	0,4209	0,2524
12	Magallanes	Primavera	Tierra del Fuego	0,4067	0,4939	0,5811
12	Magallanes	Torres del Paine	Última Esperanza	0,6249	0,6207	0,6166
12	Magallanes	San Juan de la Costa	Magallanes	0,4961	0,5718	0,6475
12	Magallanes	Río Verde	Magallanes	0,4938	0,6044	0,7149
12	Magallanes	Laguna Blanca	Magallanes	0,4462	0,6117	0,7771
12	Magallanes	Natales	Última Esperanza	0,6049	0,8803	1,1556
12	Magallanes	Porvenir	Tierra del Fuego	0,4839	0,8223	1,1606
12	Magallanes	Punta Arenas	Magallanes	0,4645	1	1,5355

11.2 Proceso de Postulación y Renovación al Programa Residencia Familiar Estudiantil año 2018

11.2.1 Requisitos para Postulantes

Para postular al programa los estudiantes deberán cumplir con los siguientes requisitos:

- Vivir en sectores rurales o urbanos con reducida oferta de colegios (con reducida oferta educativa).
- Ser estudiante de 7° y 8° año de Enseñanza Básica; Educación Media Científico Humanista o Técnico Profesional; de colegios municipales, particulares subvencionados y administración delegada (diurnos); primer año de enseñanza superior, siempre que provengan del programa (no se excluye a los extranjeros si cuentan con residencia definitiva en el país).
- La necesidad de trasladarse a otra localidad para continuar estudios ya sea porque no existen colegios en su lugar de residencia o porque la oferta educacional es reducida y no responde a los intereses de los estudiantes.
- Tener una edad mínima de 12 años cumplidos al momento de la postulación. En situaciones de excepción se permitirá el ingreso de estudiantes menores a la edad señalada, debiendo ser resuelto por la Comisión Regional de selección, con todos los antecedentes que den cuenta de la necesidad de acceder al programa.
- Preferentemente en condición de vulnerabilidad socioeconómica, lo cual implica en caso de ser necesario aplicar criterios de priorización, se deberán considerar los puntajes del modelo de focalización descrito con posterioridad en el presente manual.
- Estar inscrito el Registro Social de Hogares

11.2.1.1 Documentación exigida para Postulantes

Los postulantes deben presentar, ante la entidad ejecutora correspondiente, los siguientes documentos:

- Fotocopia de cédula de identidad.
- Certificado de alumno regular y/o de matrícula.
- Declaración de salud firmada por el beneficiario y su madre o padre, donde declara que no presenta impedimentos de salud o indicación de tratamiento farmacológico en caso de que corresponda
- Certificado que acredite la residencia, otorgado por la Dirección de Desarrollo Comunitario del Municipio correspondiente o declaración jurada otorgada ante

notario o certificado otorgado por Junta de Vecinos, autorizada formalmente por medio de acto administrativo para emitir certificados de residencia. Para el caso de los estudiantes con descendencia indígena se considerará válido el certificado de residencia en comunidad indígena.

- En caso de estudiantes de un curso menor a 7° básico o 12 años que requieran acceder al beneficio, deben contar y presentar a la entidad ejecutora un informe de un/a psicólogo/a que dé cuenta del grado de madurez y capacidad de adaptación que se requiere para ser beneficiario del programa.

11.2.2 Requisitos para Renovantes

- Mantener la residencia en sectores rurales y/o aislados o con reducida oferta educacional, por lo que requieren trasladarse desde su localidad de origen para continuar estudios.
- Mantener la condición de alumno regular.
- Estar inscrito en el Registro Social de Hogares.

11.2.2.1 Documentación exigida para Renovantes

Los renovantes deben presentar, ante la entidad ejecutora correspondiente, los siguientes documentos:

- Certificado de Alumno Regular y/o de matrícula.
- Certificado que acredite la residencia, otorgado por la Dirección de Desarrollo Comunitario del Municipio correspondiente o declaración jurada otorgada ante notario o certificado otorgado por Junta de Vecinos, autorizada formalmente por medio de acto administrativo para emitir certificados de residencia. Para el caso de los estudiantes con descendencia indígena se considerará válido el certificado de residencia en comunidad indígena.

11.2.3 Causales de Término del Beneficio

- Incumplimiento por parte del beneficiario y/o de la familia de origen, de alguna de las obligaciones establecidas en el presente manual respecto a las obligaciones de los usuarios.
- Extinción del período máximo de goce permitido para los estudiantes de Educación Superior en relación con la duración normal de la carrera de acuerdo con lo dispuesto en el presente manual en el punto 6.2 “Extensión de la duración del beneficio para estudiantes de Educación Superior”
- Beneficiario incurre en conductas que pudieran revestir caracteres de delito o se consideraran contrarios al objeto del programa.
- Titulación de Educación Superior en modalidad: CFT-IP o Universidades.
- Fallecimiento del beneficiario/a.

- Haberse comprobado la falsedad de la información consignada en el formulario de postulación o renovación.
- Beneficiario no se adapta al programa, entendiéndose por ello que durante un año académico el estudiante es ubicado, por la entidad ejecutora, con tres familias tutoras distintas y persiste la problemática. Es decir, que durante el período de permanencia en el programa se ha reubicado en más de tres viviendas.
- En el evento que el beneficiario no sea promovido de curso durante dos años continuos.
- Pérdida de la calidad de alumno/a regular del usuario/a.
- En caso de que el beneficiario deserte del sistema educacional.
- Cuando un estudiante tenga ingreso de asistencia 0 durante 2 meses continuos, será causal de egreso en el Sistema Informático durante los 5 primeros días del mes siguiente.
- Ser objeto de una medida de protección, o formar parte de algún procedimiento, como consecuencia de la vulneración de derechos de los beneficiarios.

El embarazo adolescente en ningún caso será causal de expulsión del programa. La familia de origen deberá solicitar que el colegio mantenga a la estudiante hasta que su condición física y emocional lo permita, apoyando y potenciando el auto cuidado propio del estado prenatal, con el apoyo y acuerdo de la familia de origen. No es objeto de este programa el apoyo en la crianza y cuidado de los hijos de las/os beneficiarias/os del programa, no constituye una obligación del programa ni de las familias tutoras.

11.3 Jupiter Book

11.3.1 Selección de variables

In [42]:

```
dummie_cols = ['CURSO_PADRE', 'DESC_ETNIA', 'ds_tramo_cse', 'UBICORLUGEST_DESC',
'MEDIOTRANSPORTE_DESC',
'INDICE AISLAMIENTO TERRITORIAL', 'COSTOMEDIOTRANSPORTE_DESC', 'C
URSO_MADRE', 'COMUNA_EST_EDU']
```

In [43]:

```
num_cols = ['NOT_DESCRIPCION', 'estudiante_en_rnd', 'hogar_monoparental']
```

In [44]:

```
df_sample = pd.concat([dfna[num_cols], pd.get_dummies(dfna[dummie_cols]), dfna[
'DESC_ESTADO']], axis = 1)
```

11.3.2 Separación de Datos

In [52]:

```
df_sample.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
Int64Index: 13562 entries, 0 to 13561  
Columns: 364 entries, NOT_DESCRIPCION to DESC_ESTADO  
dtypes: float64(1), int64(3), uint8(360)  
memory usage: 5.2 MB
```

In [53]:

```
df_sample_x = df_sample[df_sample.columns[:-1]] #todas las columnas menos la última.  
df_sample_y = df_sample[df_sample.columns[-1]] #todas las columnas.
```

In [54]:

```
df_sample_y = df_sample[df_sample.columns[-1]] #todas las columnas.
```

In [55]:

```
x_train_smp, x_test_smp, y_train_smp, y_test_smp = train_test_split(df_sample_x,  
                                                                    df_sample_y,  
                                                                    test_size =0.1,  
                                                                    random_state=  
43)
```

Para no sobreajustar el modelo se vuelve a separar los datos para obtener el conjunto de datos de validación

In [56]:

```
x_train2_smp, x_valid_smp, y_train2_smp, y_valid_smp = train_test_split(x_train_smp,  
                                                                    y_train_smp,  
                                                                    test_size =0.2,  
                                                                    random_state=  
43)
```


11.3.3 Modelo 1: Regresión Logística

In [61]:

```
from sklearn import linear_model# primero importamos el modulo de modelos lineal
es
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import precision_score
from sklearn import metrics
from sklearn.model_selection import cross_validate
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import cross_val_predict
from sklearn.metrics import roc_curve, auc, roc_auc_score
import lime
import lime.lime_tabular
```

In [62]:

```
##C= es el inverso de lambda, si lo hacemos muy grande, sacamos el efecto de la
regularizacion
logreg = linear_model.LogisticRegression(C=1e90, solver='lbfgs', multi_class='mu
ltinomial') # Creamos un objeto con los parametros que nosotros definamos
# Se pueden ajustar posteriormente mediante validación cruzada
```

In [63]:

```
logreg.fit(x_train2_smp, y_train2_smp) # Entrenamos con la data de entrenamiento
```

Out[63]:

```
LogisticRegression(C=1e+90, class_weight=None, dual=False, fit_intercept=True,
                   intercept_scaling=1, l1_ratio=None, max_iter=100,
                   multi_class='multinomial', n_jobs=None, penalty
='l2',
                   random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
                   warm_start=False)
```

In [64]:

```
e_training_rl = str(round(logreg.score(x_train2_smp, y_train2_smp),4)*100)
print ("Error de entrenamiento = "+ e_training_rl+"%")
```

Error de entrenamiento = 78.68%

In [65]:

```
e_testing_rl = str(round(logreg.score(x_test, y_valid_smp),4)*100)
print ("Error de validación (Accuracy) = "+ e_testing_rl+"%") # Para comparar con
entrenamiento
```

Error de validación (Accuracy) = 77.51%

In [66]:

```
y_pred_lr = logreg.predict(x_test)
Y_true = y_valid_smp
y_pred_lr.mean()
```

Out[66]:

0.5649324047521508

In [67]:

```
#logreg.predict(x_pred_2017)
```

In [68]:

```
feature_list_lr = list(x_train_smp.columns)
```

In [70]:

```
precision = (round(precision_score(Y_true, y_pred_lr), 4)*100)
print('Precisión del modelo:')
print(precision,"%")
```

Precisión del modelo:
75.13 %

In [71]:

```
accuracy = (round(metrics.accuracy_score(Y_true, y_pred_lr), 4)*100)
print('Exactitud:')
print(accuracy,"%")
```

Exactitud:
77.51 %

In [72]:

```
print ("% de estudiantes Becados: ", round(y_pred_lr.mean(), 4)*100,"%")
```

% de estudiantes Becados: 56.489999999999995 %

Matriz de Confusión

In [73]:

```
cm = confusion_matrix(Y_true, y_pred_lr)
print("Matriz de Confusión \n" , cm)
```

Matriz de Confusión
[[856 343]
 [206 1036]]

In [74]:

```
print ('Reporte:')
print(classification_report(Y_true, y_pred_lr))
```

Reporte:

	precision	recall	f1-score	support
0	0.81	0.71	0.76	1199
1	0.75	0.83	0.79	1242
accuracy			0.78	2441
macro avg	0.78	0.77	0.77	2441
weighted avg	0.78	0.78	0.77	2441

In [75]:

```
print ("AUC : ",round(metrics.roc_auc_score(Y_true, y_pred_lr), 4)*100,"%")
```

AUC : 77.4 %

11.3.4 Modelo 2: K-NN

In [77]:

```
from sklearn import neighbors
```

In [78]:

```
knn = neighbors.KNeighborsClassifier(n_neighbors = 10)
```

In [79]:

```
knn.fit(x_train2_smp, y_train2_smp)
```

Out[79]:

```
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkows  
ki',  
                    metric_params=None, n_jobs=None, n_neighbors=1  
0, p=2,  
                    weights='uniform')
```

In [80]:

```
e_training_knn = str(round(knn.score(x_train2_smp, y_train2_smp),4)*100)  
print ("Error de entrenamiento =" + e_training_knn)
```

Error de entrenamiento =78.75999999999999

In [81]:

```
e_testing_knn = str(round(knn.score(x_test, Y_true),2)*100)  
print ("Error de validación (Accuracy) =" + e_testing_knn+"%")
```

Error de validación (Accuracy) =74.0%

In [82]:

```
y_pred_knn = knn.predict(x_test)
```

In [83]:

```
precision = (round(precision_score(Y_true, y_pred_knn),4)*100)  
print('Precisión del modelo:')  
print(precision,"%")
```

Precisión del modelo:
72.6 %

In [84]:

```
cm = confusion_matrix(Y_true, y_pred_knn)  
print("Matriz de Confusión \n" , cm)
```

Matriz de Confusión
[[833 366]
 [272 970]]

In [85]:

```
print ('Reporte:')  
print(classification_report(Y_true, y_pred_knn))
```

```
Reporte:  
              precision    recall  f1-score   support  
  
   0       0.75      0.69      0.72      1199  
   1       0.73      0.78      0.75      1242  
  
 accuracy          0.74  
 macro avg         0.74      0.74      0.74      2441  
 weighted avg     0.74      0.74      0.74      2441
```

In [86]:

```
print ("AUC : ",metrics.roc_auc_score(Y_true, y_pred_knn))
```

AUC : 0.7378720055225838

11.3.5 Modelado 3: Árboles de Decisión

In [88]:

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import classification_report
```

In [89]:

```
dt = DecisionTreeClassifier(criterion = 'entropy', max_features= None, random_state = 0, max_depth=5)
#por defecto toma el "criterion" = gini el que minimiza la probabilidad de clasificación errónea.
```

In [90]:

```
dt.fit(x_train2_smp, y_train2_smp)
```

Out[90]:

```
DecisionTreeClassifier(class_weight=None, criterion='entropy', max_depth=5,
                       max_features=None, max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, presort=False,
                       random_state=0, splitter='best')
```

In [91]:

```
e_training_dt = str(round(dt.score(x_train2_smp, y_train2_smp),2)*100)
print ("Error de entrenamiento =" + e_training_dt+"%")
```

Error de entrenamiento =70.0%

In [92]:

```
e_testing_dt = str(round(dt.score(x_test, y_valid_smp),4)*100)
print ("Error de validación =" + e_testing_dt+"%")
```

Error de validación =67.88%

In [93]:

```
y_pred_dt = dt.predict(x_test)
```

In [94]:

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(Y_true, y_pred_dt)
print("Matriz de Confusión \n" , cm)
```

```
Matriz de Confusión
[[ 474  725]
 [  59 1183]]
```

In [95]:

```
#Calculo la precisión del modelo
from sklearn.metrics import precision_score
precision = (round(precision_score(Y_true, y_pred_dt),4)*100)
print('Precisión del modelo:')
print(precision, "%")
```

```
Precisión del modelo:
62.0 %
```

In [96]:

```
print("Accuracy:", (round(metrics.accuracy_score(Y_true, y_pred_dt),4)*100),"%")
```

Accuracy: 67.88 %

Valicación Cruzada (cv=3)

In [106]:

```
crits = ['entropy', 'gini']
min_smps_leaf = np.arange(70,130)
max_d = np.arange(1,10)
```

In [107]:

```
dt_gcv = model_selection.GridSearchCV(estimator = dt, param_grid = {'criterion':
crits, 'max_depth': max_d,
                                                                    'min_samples
_leaf': min_smps_leaf}, cv = 3)
```

In [108]:

```
dt_gcv.fit(x_train2_smp, y_train2_smp)
```

Out[108]:

```
GridSearchCV(cv=3, error_score='raise-deprecating',
             estimator=DecisionTreeClassifier(class_weight=None,
                                             criterion='entropy', m
ax_depth=5,
                                             max_features=None,
                                             max_leaf_nodes=None,
                                             min_impurity_decrease=
0.0,
                                             min_impurity_split=None,
                                             min_samples_leaf=1,
                                             min_samples_split=2,
                                             min_weight_fraction_le
af=0.0,
                                             presort=False, random_
state=0,
                                             splitter='best'),
             iid='warn', n_jobs...
             'max_depth': array([1, 2, 3, 4, 5, 6, 7, 8,
9]),
             'min_samples_leaf': array([ 70, 71, 72,
73, 74, 75, 76, 77, 78, 79, 80, 81, 82,
83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94,
95,
96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 1
08,
109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 1
21,
122, 123, 124, 125, 126, 127, 128, 129])}),
             pre_dispatch='2*n_jobs', refit=True, return_train_score
=False,
             scoring=None, verbose=0)
```

In [111]:

```
print('Mejor Criterio de Evaluación:', dt_gcv.best_estimator_.criterion)
print('Mejor tamaño:', dt_gcv.best_estimator_.max_depth)
print('Mejor tamaño hojas:', dt_gcv.best_estimator_.min_samples_leaf)
```

```
Mejor Criterio de Evaluación: gini
Mejor tamaño: 7
Mejor tamaño hojas: 120
```

In [112]:

```
y_pred_dt_gcv = dt_gcv.predict(x_test)
```

In [113]:

```
cm = confusion_matrix(Y_true, y_pred_dt_gcv)
print("Matriz de Confusión \n" , cm)
```

```
Matriz de Confusión
[[ 670  529]
 [ 242 1000]]
```

In [114]:

```
#Calculo la precisión del modelo
precision = (round(precision_score(Y_true, y_pred_dt_gcv),4)*100)
print('Precisión del modelo:')
print(precision, "%")
```

```
Precisión del modelo:
65.4 %
```

In [115]:

```
e_training_dt_gcv = str(round(dt_gcv.score(x_train2_smp, y_train2_smp),4)*100)
print ("Error de entrenamiento =" + e_training_dt_gcv+"%")
```

```
Error de entrenamiento =70.53%
```

In [116]:

```
e_testing_dt_gcv = str(round(dt_gcv.score(x_test, Y_true),3)*100)
print ("Error de validación =" + e_testing_dt_gcv+"%")
```

```
Error de validación =68.4%
```

In [117]:

```
print ("AUC : ",(round(metrics.roc_auc_score(Y_true, y_pred_dt_gcv),4)*100),"%")
```

```
AUC : 68.2 %
```

11.3.6 Modelo 4: Random Forest

In [129]:

```
from sklearn import model_selection
from sklearn import ensemble
```

In [130]:

```
rf = ensemble.RandomForestClassifier()
```

In [131]:

```
n_est = [10,11,12,13,14,15,16,17,18,19,20]
min_smps_leaf = np.arange(1,50)
max_d = np.arange(1,50)
crit = ['entropy','gini']
```

In [132]:

```
rf_gcv = model_selection.RandomizedSearchCV(estimator = rf, param_distributions=
{'n_estimators': n_est,
'criterion': crit, 'max_depth': max_d,
'min_samples_leaf': min_smps_leaf}, random_state = 42)
```

Out[133]:

```
RandomizedSearchCV(cv='warn', error_score='raise-deprecating',
                   estimator=RandomForestClassifier(bootstrap=True,
                                                    class_weight=None,
                                                    criterion='gini',
                                                    max_depth=None,
                                                    max_features='auto',
                                                    max_leaf_nodes=None,
                                                    min_impurity_decrease=0.0,
                                                    min_impurity_split=None,
                                                    min_samples_leaf=1,
                                                    min_samples_split=2,
                                                    min_weight_fraction_leaf=0.0,
                                                    n_estimators='warn',
                                                    n_jobs=None,
                                                    oob_score=True,
                                                    random_state=None,
                                                    verbose=0),
                   estimator_params=(),
                   estimator__min_samples_leaf=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49]),
                   estimator__n_estimators=[10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20],
                   pre_dispatch='2*n_jobs', random_state=42, refit=True,
                   return_train_score=False, scoring=None, verbose=0)
```

In [134]:

```
best_score = str(rf_gcv.best_score_)
print ("Best Score = " + best_score)
```

```
Best Score = 0.745391233101188
```

In [135]:

```
print('Numero de arboles:', rf_gcv.best_estimator_.n_estimators)
print('Mejor Criterio de Evaluacion:', rf_gcv.best_estimator_.criterion)
print('Mejor tamaño:', rf_gcv.best_estimator_.max_depth)
print('Mejor tamaño hojas:', rf_gcv.best_estimator_.min_samples_leaf)
```

```
Numero de arboles: 13
Mejor Criterio de Evaluacion: gini
Mejor tamaño: 21
Mejor tamaño hojas: 1
```

In [136]:

```
e_training_rf = str(round(rf_gcv.score(x_train2_smp, y_train2_smp),3)*100)
print ("Error de entrenamiento =" + e_training_rf+"%")
```

Error de entrenamiento =86.5%

In [137]:

```
e_testing_rf = str(round(rf_gcv.score(x_test, Y_true),4)*100)
print ("Error de validación =" + e_testing_rf+"%")
```

Error de validación =73.41%

In [138]:

```
results_rf_gcv = pd.DataFrame({'variables': x_train_smp.columns,
                              'importancia': rf_gcv.best_estimator_.feature_imp
                              ortances_})
```

In [140]:

```
y_pred_rf = rf_gcv.predict(x_test)
print(classification_report(Y_true, y_pred_rf))
```

	precision	recall	f1-score	support
0	0.80	0.62	0.70	1199
1	0.70	0.85	0.76	1242
accuracy			0.73	2441
macro avg	0.75	0.73	0.73	2441
weighted avg	0.75	0.73	0.73	2441

In [141]:

```
#Calculo la precisión del modelo
precision = round(precision_score(Y_true, y_pred_rf),4)*100
print('Precisión del modelo:')
print(precision, "%")
```

Precisión del modelo:
69.62 %

In [142]:

```
print ("AUC : ",(round(metrics.roc_auc_score(Y_true, y_pred_rf),3)*100),"%")
```

AUC : 73.2 %

Matriz de Confusión

In [143]:

```
print(pd.crosstab(Y_true, y_pred_rf, rownames=['REAL'], colnames=['PREDICCION'
]))
```

PREDICCION	0	1
REAL		
0	740	459
1	190	1052

11.3.7 Gradient Boosting

In [145]:

```
from sklearn.ensemble import GradientBoostingClassifier
```

In [146]:

```
gb_clf = GradientBoostingClassifier(n_estimators=40, learning_rate=0.9, max_features=2, max_depth=2, random_state=0)
```

In [147]:

```
gb_clf.fit(x_train2_smp, y_train2_smp)
```

Out[147]:

```
GradientBoostingClassifier(criterion='friedman_mse', init=None,
                           learning_rate=0.9, loss='deviance', max_depth=2,
                           max_features=2, max_leaf_nodes=None,
                           min_impurity_decrease=0.0, min_impurity_split=None,
                           min_samples_leaf=1, min_samples_split=2,
                           min_weight_fraction_leaf=0.0, n_estimators=40,
                           n_iter_no_change=None, presort='auto',
                           random_state=0, subsample=1.0, tol=0.0001,
                           validation_fraction=0.1, verbose=0,
                           warm_start=False)
```

In [148]:

```
print("Accuracy score (Error de entrenamiento): {0:.3f}".format(gb_clf.score(x_train2_smp, y_train2_smp)))
print("Accuracy score (Error de validación): {0:.3f}".format(gb_clf.score(x_test, Y_true)))
```

```
Accuracy score (Error de entrenamiento): 0.703
Accuracy score (Error de validación): 0.691
```

In [149]:

```
e_testing_gb = str(round(gb_clf.score(x_train2_smp, y_train2_smp),3)*100)
```

In [150]:

```
e_training_gb = str(round(gb_clf.score(x_test, Y_true),3)*100)
```

Métricas

In [153]:

```
y_pred_bo = gb_clf.predict(x_test)
print(classification_report(Y_true, y_pred_bo))
```

	precision	recall	f1-score	support
0	0.72	0.61	0.66	1199
1	0.67	0.77	0.72	1242
accuracy			0.69	2441
macro avg	0.69	0.69	0.69	2441
weighted avg	0.69	0.69	0.69	2441

In [154]:

```
#Calculo la precisión del modelo
precision = round(precision_score(Y_true, y_pred_bo),4)*100
print('Precisión del modelo:')
print(precision, "%")
```

Precisión del modelo:
67.23 %

In [155]:

```
print ("AUC : ",(round(metrics.roc_auc_score(Y_true, y_pred_bo),4)*100),"%")
```

AUC : 68.97999999999999 %

Mátriz de Confusión

In [156]:

```
print(pd.crosstab(Y_true, y_pred_bo, rownames=['REAL'], colnames=['PREDICCION']
))
```

```
PREDICCION    0    1
REAL
0             735  464
1             290  952
```

11.3.8 Comparación de Resultados

In [168]:

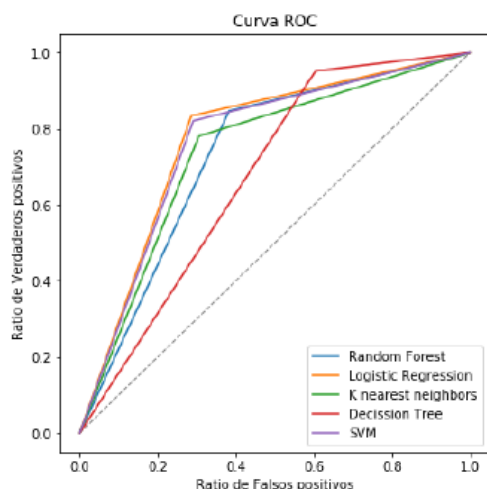
```
resultados_finales = pd.DataFrame({'Modelo': ['R. Logistica', 'K-NN (5)', 'Árboles de Decisión', 'A.D-CV=3', 'A.D-CV=10', 'Random Forest', 'Gradient Boosting', 'SVM'],
'E_Validación': [e_testing_rl, e_testing_knn, e_testing_dt, e_testing_dt_gcv, e_testing_dt_gcv2, e_testing_rf, e_testing_gb, e_testing_svm],
'E_Entrenamiento': [e_training_rl, e_training_knn, e_training_dt, e_training_dt_gcv, e_training_dt_gcv2, e_training_rf, e_training_gb, e_training_svm]})
```

In [169]:

```
resultados_finales[['Modelo', 'E_Entrenamiento', 'E_Validación']].sort_values(by='E_Validación', ascending = False)
```

Out[169]:

	Modelo	E_Entrenamiento	E_Validación
0	R. Logistica	78.68	77.51
7	SVM	80.0	76.61
1	K-NN (5)	78.75999999999999	74.0
5	Random Forest	86.5	73.41
6	Gradient Boosting	69.1	70.3
3	A.D-CV=3	70.53	68.4
4	A.D-CV=10	70.93	68.17
2	Árboles de Decisión	70.0	67.88



11.4 Presupuesto y Cobertura PRFE 2018

De acuerdo con la Ley de Presupuestos del Sector Público N° 21.053 del año 2018, publicada en el Diario Oficial del 27 de diciembre de 2017, para el Programa Residencia Familiar estuvo fue aprobado un total de CLP\$ 8.169.806.000, por lo que se determina la cobertura por región del programa a partir del presupuesto asignado según consta en las siguientes tablas:

Tabla 68: Presupuesto Asignado y Cobertura por Región Programa Residencia Familiar Estudiantil
Fuente: Elaboración del Departamento de Becas de JUNAEB basado en la Ley de Presupuestos del Sector Público N° 21.053 del año 2018

PROGRAMA RESIDENCIA FAMILIAR - INVERSIÓN											
Región	Ed. Básica y Media				Ed. Superior				Total Cobertura	Total Inversión (\$)	Total Inversión (M\$)
	Cobertura	Familias Tutoras	Entidades Ejecutoras	Total Inversión E.B y E.M. (\$)	Cobertura	Familias Tutoras	Entidades Ejecutoras	Total Inversión E.S. (\$)			
Tarapacá	112	83.070.400	12.135.200	95.205.600	50	54.005.000	8.156.500	62.161.500	162	157.367.100	157.367
Antofagasta	44	32.634.800	4.767.400	37.402.200	25	27.002.500	4.078.250	31.080.750	69	68.482.950	68.483
Atacama	101	74.911.700	10.943.350	85.855.050	28	30.242.800	4.567.640	34.810.440	129	120.665.490	120.665
Coquimbo	300	222.510.000	32.505.000	255.015.000	155	167.415.500	25.285.150	192.700.650	455	447.715.650	447.716
Valparaíso	250	172.425.000	20.707.500	193.132.500	187	189.805.000	21.927.620	211.732.620	437	404.865.120	404.865
O'Higgins	174	120.007.800	14.412.420	134.420.220	112	113.680.000	13.133.120	126.813.120	286	261.233.340	261.233
Maule	342	235.877.400	28.327.860	264.205.260	104	105.560.000	12.195.040	117.755.040	446	381.960.300	381.960
Biobío	1.102	760.049.400	91.278.660	851.328.060	501	508.515.000	58.747.260	567.262.260	1.603	1.418.590.320	1.418.590
La Araucanía	935	644.869.500	77.446.050	722.315.550	198	200.970.000	23.217.480	224.187.480	1.172	986.399.640	986.400
Lonquimay	39	35.525.100	4.371.510	39.896.610		0	0	0			
Los Lagos	1.030	710.391.000	85.314.900	795.705.900	313	317.695.000	36.702.380	354.397.380	1.506	1.316.850.650	1.316.851
Provincia Palena	163	148.476.700	18.270.670	166.747.370		0	0	0			
Aysén	383	766.000.000	90.284.590	856.284.590	64	128.000.000	15.086.720	143.086.720	447	999.371.310	999.371
Magallanes	26	23.683.400	3.809.520	27.492.920	18	24.359.400	3.833.280	28.192.680	44	55.685.600	55.686
Metropolitana	61	42.071.700	5.052.630	47.124.330	24	24.360.000	2.814.240	27.174.240	85	74.298.570	74.299
Los Ríos	354	244.153.800	29.321.820	273.475.620	45	45.675.000	5.276.700	50.951.700	399	324.427.320	324.427
Arica y Parinacota	141	104.579.700	15.277.350	119.857.050	55	59.405.500	8.972.150	68.377.650	196	188.234.700	188.235
Dirección Nacional		0	0	0		0	910.321.940	910.321.940		910.321.940	910.322
Total		4.421.237.400	544.226.430	4.965.463.830		1.996.690.700	1.154.315.470	3.151.006.170	7.436	8.116.470.000	8.116.470
										Prácticas Estivales	53.336
										TOTAL M\$	8.169.806
										Ley Ppto 2018 (M\$)	8.169.806

11.5 Cálculos Flujo de Caja del Proyecto

11.5.1 Determinación Ahorros

En el año 2018 se considera una cobertura para el Programa de Residencia Familiar Estudiantil el siguiente número de cupos disponibles según la modalidad:

- Postulantes: 5.205
- Renovantes: 2.231

También, se debe tener en cuenta el tiempo en minutos que se utilizan para desarrollar las actividades principales del proceso de asignación de este programa, el cual se detalla en la tabla a continuación:

Tabla 69: Presupuesto Asignado y Cobertura por Región Programa Residencia Familiar Estudiantil
Fuente: Elaboración propia

ACTIVIDADES PROCESO DE ASIGNACIÓN	NO. BENEFICIARIOS	TIEMPO/MINUTOS
CONSOLIDAR INFORMACION POR REGION SOBRE POSTULANTES Y RENOVANTES		180
REVISAR CUMPLIMIENTO DE REQUISITOS DE POSTULANTES	5.205	20.821
REVISAR CUMPLIMIENTO DE REQUISITOS DE RENOVANTES	2.231	4.462
ELABORAR ARHCIVO DE PUBLICACIÓN DE RESULTADOS		120
TOTAL	7.436	25.582

Como se puede observar, el tiempo total en minutos que se necesitan para poder ejecutar el proceso de asignación es 25.502, los que al transformarlos en horas resulta un valor de 426.

Para determinar el ahorro real que representa implementar el proyecto propuesto se realiza una comparación de los tiempos que se necesitan para ejecutar el proceso si se realiza o no el proyecto. El resultado de este análisis se muestra en la siguiente tabla:

Tabla 70: Tiempo en horas en que se realizan cada actividad del proceso de asignación con y sin proyecto

Fuente: Elaboración propia

ACTIVIDADES PROCESO DE ASIGNACIÓN	SIN PROYECTO (SP)	CON PROYECTO (CP)	SP - CP
CONSOLIDAR INFORMACION POR REGION SOBRE POSTULANTES Y RENOVANTES	180	180	→ 0
REVISAR CUMPLIMIENTO DE REQUISITOS DE POSTULANTES	20.821	2	↓ -20.819
REVISAR CUMPLIMIENTO DE REQUISITOS DE RENOVANTES	4.462	1	↓ -4.461
ELABORAR ARHCIVO DE PUBLICACIÓN DE RESULTADOS	120	40	↓ -80
TOTAL MINUTOS	25.582	223	↓ -25.359
TOTAL HORAS	426	4	↓ -423

Se aprecia que cuatro de las cinco actividades disminuyen notoriamente su tiempo de ejecución, lo que representa un ahorro de 423 horas en total del tiempo necesario para desarrollar la asignación de este programa.

Con estos datos, se valoriza el ahorro calculado según la siguiente información conocida:

- A. Tiempo ahorrado en horas del proceso de asignación: 423
- B. Sueldo mensual profesional encargada del programa: CLP\$ 950.000
- C. Horas trabajadas al mes: 176
- D. Valor hora/hombre: CLP\$ 5.397
- E. Incremento anual hora/hombre: 5%
- F. Disminución anual cobertura (a partir del año 2): 8%

La fórmula utilizada para el cálculo del ahorro en el año 1 con proyecto, de acuerdo con la información presentada anteriormente:

$$D \cdot A / (1 + E)$$

En relación con el año 2, para el cálculo del ahorro, la formula tiene una pequeña variante, debido a la disminución de la cobertura del 8%.

11.5.2 Cálculo Valor Residual

Es importante mencionar que al final del año 5 se presentan CLP\$ 3.615.938, lo cual es calculado mediante el método del Valor Residual. Para esto, se conoce que para los años 6, 7 y 8 el monto del Flujo Operacional es constante, es decir, equivalente a CLP\$ 1.581.072. Se debe calcular el valor presente de cada uno de estos 3 valores, cuyo año base es el año 5. Para esto se utiliza la siguiente formula:

$$VF / (1 + t)^n$$

Tal que,

- $VF = 1.581.072$
- $t = \text{tasa de descuento} = 14,9\%$
- $n = \text{número de periodos}$

El resultado para cada año se muestra en la siguiente tabla:

Tabla 71: Cálculo de Valor Presente
Fuente: Elaboración propia

VALOR PRESENTE (CLP\$)	6	7	8
	1.376.042	1.197.599	1.042.297

Finalmente, como los 3 valores se encuentran en el mismo periodo, se suman. Este resultado es entonces el Valor Residual del proyecto en el año 5.