



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**ESTUDIO DE FACTIBILIDAD DE UN MODELO DE CLASIFICACIÓN DE
ESTUDIANTES UNIVERSITARIOS SEGÚN EL ESTADO CRÍTICO DE
ESTRÉS, ANSIEDAD Y SINTOMATOLOGÍA DEPRESIVA**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

JAVIERA VALENTINA OVALLE TOLEDO

PROFESOR GUÍA:
ÁNGEL JIMÉNEZ MOLINA

MIEMBROS DE LA COMISIÓN:
DANIEL SCHWARTZ PERLROTH
PATRICIA MUÑOZ SILVA
VANIA MARTÍNEZ NAHUEL

Este trabajo ha sido parcialmente financiado por:
PROYECTO U-REDES-C 2018 07 & BEAUCHEF PROYECTA

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL
POR: **JAVIERA VALENTINA OVALLE TOLEDO**
FECHA: 2020
PROF. GUÍA: ÁNGEL JIMÉNEZ MOLINA

ESTUDIO DE FACTIBILIDAD DE UN MODELO DE CLASIFICACIÓN DE ESTUDIANTES UNIVERSITARIOS SEGÚN EL ESTADO CRÍTICO DE ESTRÉS, ANSIEDAD Y SINTOMATOLOGÍA DEPRESIVA

La salud mental ha cobrado visibilidad en el último tiempo y es cada vez más relevante levantar información al respecto para tomar medidas tanto preventivas, como correctivas en todas las etapas de vida de las personas. En específico, es importante estudiar la época universitaria debido a que estudios indican que el 30% de los estudiantes sufre de estrés, ansiedad y depresión de manera simultánea, cifra alarmante debido a los efectos sobre la calidad de vida, rendimiento académico, vida personal y profesional.

El presente trabajo de título está enfocado en los estudiantes de la Facultad de Ingeniería de la Universidad de Chile, donde se presentan las patologías antes mencionadas, y solo es posible abarcar el tema una vez que el estudiante lo identifica. Frente a esto, se plantea un estudio de factibilidad en el cual se busca clasificar a los estudiantes por medio de modelos matemáticos, según su estado de depresión, ansiedad y estrés, en pro de un futuro Ensayo Clínico Aleatorizado (RCT).

Siguiendo los estudios de otros investigadores, se evaluaron las actitudes y los comportamientos normalizados en época universitaria que agravan las patologías, utilizando metodología EMA con cuestionarios para determinar depresión, ansiedad, estrés, higiene del sueño, caracterización y contextos. Bajo la misma línea, para el levantamiento de datos se utilizó la aplicación móvil Ánimi, debido al éxito del uso de tecnologías móviles en medicina.

Los resultados muestran que la aplicación presentó una retención del 76% de los usuarios activos, decayendo su uso a partir de la ejecución número 16 de un total de 42. Los valores se encuentran dentro de lo esperado, pero no permiten concluir sobre la cantidad óptima de ejecuciones debido a su alta desviación estándar. En cuanto a los variables de estudio, se destaca que tanto la higiene del sueño como el contexto personal, situacional y académico son factores relevantes, resultando imperativo su estudio y la incorporación de más aspectos relacionados en trabajos futuros. De la misma manera, se obtuvo que los algoritmos de clasificación Bosques Aleatorios, Regresión Logística y Máquinas de Vectores de Soporte predicen correctamente las variables estudiadas, cumpliendo así el objetivo propuesto.

A modo de conclusión, todo lo anterior no solo demuestra la factibilidad de lo estudiado y el éxito de la aplicación móvil, sino que también permite determinar las directrices para el posterior desarrollo de un RCT y entrega valor social, ya que fomenta el auto conocimiento y colabora potencialmente con las autoridades universitarias para la implementación de medidas preventivas y correctivas asociadas a la temática.

*A mi familia.
que gracias a ellos soy todo,
sin ellos soy nada*

Con amor

Agradecimientos

Quiero agradecer a mis amigos y amigas Camila, Javiera, Dennis, Gabriel, Rubén y Pablo por apoyarme durante este proceso, estar junto a mi y regalarme tantos buenos momentos. También a mis amigos y amigas de la UCH, USM, del colegio y de la vida, ya que han hecho de mi camino algo inolvidable.

También le agradezco a Gonzalo, Franco, J. Guillermo y Sebastián por hacer más entretenidas las 3 horas del ramo, por las juntas, las largas horas de trabajo, la buena onda y los carretes. Gracias a Guillermo P. por ser un buen amigo y por darse el tiempo de leer mi memoria y gracias a Rodrigo por caminar a mi lado estos últimos meses y por ayudarme a escoger el nombre de la aplicación. Ánimi tiene múltiples interpretaciones, como todo lo que pasa por nuestra mente.

Gracias a Vania Martínez, Álvaro Jiménez y toda la red de CEMERA e ImHay por proponer esta memoria, darme recomendaciones, dejarme participar en reuniones con expertos en el tema y trabajar arduamente por la salud mental juvenil. A Ángel Jiménez y Daniel Schwartz por elegirme para llevar a cabo este estudio y ser mis profesores guía y co-guía, respectivamente.

Gracias a mis profesores, compañeros y compañeras de la Sección 8 de Trabajo de Título. Gracias a Beauchef Proyecta y el Proyecto U-Redes por darme financiamiento para el desarrollo de la aplicación móvil. También a todo el equipo de Marketing Digital de CrediTú por acogerme bajo mis condiciones y permitir que me desarrollara profesionalmente en período de escritura de título.

Gracias a los equipos docentes de TIC's, Organización Industrial y Diseño de Procesos de Negocios por permitirme poner en práctica mi amor por la docencia, los proyectos, los negocios y la tecnología.

Infinitas gracias a los y las estudiantes de la FCFM que decidieron participar voluntariamente en la toma de datos de mi trabajo. Gracias por creer que la salud mental es tan importante como la salud física.

Por último, gracias a quienes han sido mi pilar en la vida: mis padres, hermanos, cuñadas, sobrinos y sobrinas, por apoyarme incondicionalmente, aceptar mis cambios de humor, alegrarme y estar presentes en todo momento. Gracias Cuky por ser la perrita más atenta, cariñosa y regalona que conozco, y gracias mama Pilar por haberme enseñado de valentía y resiliencia. Un beso y un abrazo al cielo.

Tabla de Contenido

1. Introducción	1
1.1. Organización	1
1.1.1. Características generales	1
1.1.2. Área oferente	1
1.1.3. Áreas implicadas	2
2. Descripción General	4
2.1. Justificación	4
2.1.1. Contexto y problema	4
2.1.2. Enfoque del problema	8
2.1.3. Efectos del problema	8
2.2. Población y grupo de estudio	9
2.3. Hipótesis principal y secundaria	10
2.3.1. Hipótesis principal	10
2.3.2. Hipótesis secundaria	11
2.4. Impacto y diferenciación del proyecto	11
2.5. Objetivos	12
2.5.1. Objetivo general	12
2.5.2. Objetivos específicos	12
2.6. Alcances	12
2.7. Apoyos institucionales	14
2.7.1. Financiero	14
2.7.2. Académico	14
3. Marco Conceptual	15
3.1. Conceptos y cuestionarios de salud y psicología	15
3.1.1. Conceptos	15
3.1.2. Cuestionarios y escalas	16
3.1.3. Estudios de factibilidad y RCT	16
3.1.4. Evaluación Ecológica Momentánea (EMA)	17
3.1.5. Tratamiento de datos	17
3.1.6. Modelos de clasificación	19
3.2. Estado del arte	20
3.2.1. Estrés e higiene del sueño en estudiantes universitarios	20
3.2.2. Depresión y ansiedad en estudiantes universitarios	21
3.2.3. Aplicaciones móviles en salud mental	23

4. Metodología	26
5. Recolección de datos	29
5.1. Cuestionarios	29
5.1.1. Caracterización general	29
5.1.2. Higiene de sueño	31
5.1.3. Trayectoria del estado anímico	31
5.1.4. Contexto situacional	33
5.1.5. Preguntas finales	33
5.2. Aplicación móvil	35
5.2.1. Descripción general	35
5.2.2. Funcionalidades	36
5.2.3. Ilustraciones de la aplicación móvil	37
5.2.4. Manejo de la aplicación móvil	45
6. Resultados	50
6.1. Uso de la aplicación móvil Ánimi	50
6.2. Pre-procesamiento y transformación de datos	51
6.2.1. Análisis exploratorio de la muestra	52
6.2.2. Análisis exploratorio de los datos	55
6.2.2.1. Depresión, Ansiedad y Estrés	55
6.2.2.2. Selección de variables relevantes	57
6.3. Minería de datos	61
6.3.1. Trabajo y modelos de Machine Learning	61
6.3.1.1. Bases de datos	61
6.3.1.2. Balanceo de datos	61
6.3.1.3. Programación de modelos	63
6.3.1.4. Comparación de modelos	63
6.3.1.5. Resultados con matriz de correlación	66
6.3.1.6. Resultados con análisis de componentes principales	67
6.3.1.7. Resultados con bosques aleatorios	68
6.4. Comparaciones	69
6.4.1. Preguntas iniciales y finales	69
7. Discusión	71
8. Conclusión	76
9. Trabajo Futuro	78
Bibliografía	80
Anexo A. Matrices de correlación	85
Anexo B. Resultados modelos con MC	86
Anexo C. Resultados modelos con PCA	101
Anexo D. Resultados modelos con RF	113

Índice de Tablas

5.1.	Preguntas de Contexto General	30
5.2.	Preguntas de Realidad Nacional	30
5.3.	Preguntas de Percepción Personal	31
5.4.	Preguntas de Higiene del Sueño	31
5.5.	Preguntas de Depresión y Ansiedad	32
5.6.	Preguntas de Estrés	32
5.7.	Preguntas de Contexto Situacional	33
5.8.	Preguntas Finales - DASS21	34
5.9.	Preguntas Finales - Comparativas	35
6.1.	Variables de interés para Depresión	58
6.2.	Variables de interés para Ansiedad	59
6.3.	Variables de interés para Estrés	59
6.4.	Importancia de variables encontradas con RF	60
6.5.	Proporciones y frecuencias sin balancear	62
6.6.	Proporciones y frecuencias balanceadas	62
6.7.	Proporciones y frecuencias desbalanceadas PCA	62
6.8.	Proporciones y frecuencias balanceadas PCA	63
6.9.	Métricas por modelo con MC	66
6.10.	Métricas por modelo con PCA	67
6.11.	Métricas por modelo con BA	68
B.1.	Depresión día con Regresión Logística - MC	86
B.2.	Depresión tarde con Regresión Logística - MC	87
B.3.	Ansiedad día con Regresión Logística - MC	91
B.4.	Ansiedad tarde con Regresión Logística - MC	92
B.5.	Estrés día con Regresión Logística - MC	96
B.6.	Estrés tarde con Regresión Logística - MC	97
C.1.	Depresión día con Regresión Logística - PCA	101
C.2.	Depresión tarde con Regresión Logística - PCA	101
C.3.	Ansiedad día con Regresión Logística - PCA	105
C.4.	Ansiedad tarde con Regresión Logística - PCA	105
C.5.	Estrés día con Regresión Logística - PCA	109
C.6.	Estrés tarde con Regresión Logística - PCA	109
D.1.	Depresión tarde con Regresión Logística - RF	114
D.2.	Depresión día con Regresión Logística - RF	115
D.3.	Ansiedad día con Regresión Logística - RF	119
D.4.	Ansiedad tarde con Regresión Logística - RF	120
D.5.	Estrés día con Regresión Logística - RF	124

D.6. Estrés tarde con Regresión Logística - RF	125
--	-----

Índice de Ilustraciones

2.1.	Árbol de Causas	7
3.1.	Ejemplo de aplicación de SMOTE	18
3.2.	Ejemplo de gráfico de proporciones de PCA	18
3.3.	Ejemplo de matriz de correlación	19
4.1.	Etapas de Metodología KDD	26
5.1.	Funcionamiento general aplicación Ánimi	35
5.2.	Vista inicial	38
5.3.	Caracterización	39
5.4.	Cuestionarios - parte 1	40
5.5.	Cuestionarios - parte 2	41
5.6.	Cuestionarios - parte 3	42
5.7.	Cuestionario final y fragmento de DASS-21	43
5.8.	Mensajes según puntaje de corte	44
5.9.	Retroalimentación automática DASS-21	45
5.10.	Base de datos Animi	46
5.11.	Ejemplo tabla desbloques de pantalla	47
5.12.	Ejemplo tabla de registro de llamadas	47
5.13.	Ejemplo tabla de respuestas diarias	47
5.14.	Ejemplo tabla de geolocalización	47
5.15.	Ejemplo tabla datos pasivos recopilados cada 15 min.	47
5.16.	Firebase vista número de ejecuciones	48
5.17.	Firebase vista mensajes	48
5.18.	Firebase vista preguntas diarias	49
5.19.	Firebase vista uso de sensores	49
6.1.	Frecuencia de cuestionarios respondidos	50
6.2.	Reducción de variables y unificación de bases de datos	52
6.3.	Distribución según edad, género, región de procedencia y semestre	53
6.4.	Percepción de estado anímico en relación con otras variables	54
6.5.	Percepción de estado anímico en relación con otras variables	55
6.6.	Distribución de variables de Depresión, Ansiedad y Estrés	56
6.7.	Depresión, Ansiedad y Estrés observados	57
6.8.	Ejemplo matriz de confusión	64
6.9.	Comparación de percepción de carga académica pre y post cuestionarios	69
6.10.	DASS 21: Depresión, Ansiedad y Estrés observados	70
9.1.	Visualización general del proceso propuesto	79
A.1.	Matriz de correlación por jornada	85
B.1.	Depresión día con Bosques Aleatorios - MC	88

B.2.	Importancia de variables para Depresión día con Bosques Aleatorios - MC . .	88
B.3.	Depresión tarde con Bosques Aleatorios - MC	89
B.4.	Importancia de variables para Depresión tarde con Bosques Aleatorios - MC .	89
B.5.	Depresión día con Árboles de Clasificación - MC	90
B.6.	Depresión tarde con Árboles de Clasificación - MC	90
B.7.	Ansiedad día con Bosques Aleatorios - MC	93
B.8.	Importancia de variables para Ansiedad día con Bosques Aleatorios	93
B.9.	Ansiedad tarde con Bosques Aleatorios - MC	94
B.10.	Importancia de variables para Ansiedad tarde con Bosques Aleatorios - MC .	94
B.11.	Ansiedad día con Árboles de Clasificación - MC	95
B.12.	Ansiedad tarde con Árboles de Clasificación - MC	95
B.13.	Estrés día con Bosques Aleatorios - MC	98
B.14.	Importancia de variables para Estrés día con Bosques Aleatorios - MC	98
B.15.	Estrés tarde con Bosques Aleatorios - MC	99
B.16.	Importancia de variables para Estrés tarde con Bosques Aleatorios - MC . . .	99
B.17.	Estrés día con Árboles de Clasificación - MC	100
B.18.	Estrés tarde con Árboles de Clasificación - MC	100
C.1.	Depresión día con Bosques Aleatorios - PCA	102
C.2.	Importancia de variables para Depresión día con Bosques Aleatorios - PCA .	102
C.3.	Depresión tarde con Bosques Aleatorios - PCA	103
C.4.	Importancia de variables para Depresión tarde con Bosques Aleatorios - PCA	103
C.5.	Depresión día con Árboles de Clasificación - PCA	104
C.6.	Depresión tarde con Árboles de Clasificación - PCA	104
C.7.	Ansiedad día con Bosques Aleatorios - PCA	106
C.8.	Importancia de variables para Ansiedad día con Bosques Aleatorios - PCA . .	106
C.9.	Ansiedad tarde con Bosques Aleatorios - PCA	107
C.10.	Importancia de variables para Ansiedad tarde con Bosques Aleatorios - PCA .	107
C.11.	Ansiedad día con Árboles de Clasificación - PCA	108
C.12.	Ansiedad tarde con Árboles de Clasificación - PCA	108
C.13.	Estrés día con Bosques Aleatorios - PCA	110
C.14.	Importancia de variables para Estrés día con Bosques Aleatorios - PCA . . .	110
C.15.	Estrés tarde con Bosques Aleatorios - PCA	111
C.16.	Importancia de variables para Estrés tarde con Bosques Aleatorios - PCA . .	111
C.17.	Estrés día con Árboles de Clasificación - PCA	112
C.18.	Estrés tarde con Árboles de Clasificación - PCA	112
D.1.	Depresión día con Bosques Aleatorios - RF	113
D.2.	Importancia de variables para Depresión día con Bosques Aleatorios - RF . .	116
D.3.	Depresión tarde con Bosques Aleatorios - RF	116
D.4.	Importancia de variables para Depresión tarde con Bosques Aleatorios - RF .	117
D.5.	Depresión día con Árboles de Clasificación - RF	117
D.6.	Depresión tarde con Árboles de Clasificación - RF	118
D.7.	Ansiedad tarde con Bosques Aleatorios - RF	121
D.8.	Importancia de variables para Ansiedad día con Bosques Aleatorios - RF . . .	121
D.9.	Ansiedad tarde con Bosques Aleatorios - RF	122
D.10.	Importancia de variables para Ansiedad tarde con Bosques Aleatorios - RF . .	122
D.11.	Ansiedad día con Árboles de Clasificación - RF	123
D.12.	Ansiedad tarde con Árboles de Clasificación - RF	123

D.13.	Estrés tarde con Bosques Aleatorios - RF	126
D.14.	Importancia de variables para Estrés día con Bosques Aleatorios - RF	126
D.15.	Estrés tarde con Bosques Aleatorios - RF	127
D.16.	Importancia de variables para Estrés tarde con Bosques Aleatorios - RF	127
D.17.	Estrés día con Árboles de Clasificación - RF	128
D.18.	Estrés tarde con Árboles de Clasificación - RF	128

Capítulo 1

Introducción

1.1. Organización

1.1.1. Características generales

La presente memoria de título se llevó a cabo en la Universidad de Chile, casa de estudios de carácter nacional y público, que está orientada a la contribución a la sociedad por medio de la generación de conocimiento y formación de profesionales de manera íntegra, a través de la docencia, creación y extensión, con especial énfasis en la investigación y postgrado.

Actualmente ocupa el primer lugar en el “Academic Ranking of World Universities 2019” [59] y el segundo en el “Quacquarelli Symonds (QS) Ranking 2020” [17], los cuales evalúan el nivel y las estadísticas de citación de artículos, colaboraciones nacionales e internacionales, calidad de los estudiantes, intercambios estudiantiles, premios a académicos, calidad de trabajo de personal y equipo académico.

Específicamente, se trabajó en la **Facultad de Ciencias Físicas y Matemáticas (FCFM)** que fue fundada en 1842 y actualmente se especializa en el desarrollo de las ciencias básicas, las ciencias de la ingeniería y las aplicaciones tecnológicas. Con sus 5.700 estudiantes justifica el 29 % de las matrículas, ocupando el primer puesto en ingresos a nivel universidad. Su oferta alcanza 13 carreras, donde un tercio de los estudiantes proviene de regiones y un 33 % es de género femenino [27].

1.1.2. Área oferente

El estudio fue solicitado y coordinado por el Psicólogo y Doctor en Sociología Dr. Álvaro Jiménez, y la Psiquiatra y Doctora en Psicoterapia Dra. Vania Martínez, siendo integrante y directora del Núcleo Milenio ImHay, respectivamente.

Como contraparte se encuentran los profesores Dr. Ángel Jiménez y Dr. Daniel Schwartz, ambos del departamento de Ingeniería Industrial de la FCFM, quienes cumplen el rol de guía y co-guía, según corresponde.

Ahora bien, dentro de la Facultad fue **Beauchef Proyecta** quien presentó el tema como una memoria de título interdisciplinaria, siendo un trabajo colaborativo entre la Facultad de Ciencias Físicas y Matemáticas y la Facultad de Medicina, ambas de la Universidad de Chile, enmarcándose dentro del Proyecto “Una nueva Ingeniería para el 2030.”

1.1.3. Áreas implicadas

A continuación, se describen las áreas implicadas y la relación que tienen con los resultados de la memoria:

- **Unidad de Calidad de Vida:** Funciona desde el año 2008 con el fin de prestar apoyo a los estudiantes que se encuentran viviendo alguna situación que atente contra su estabilidad emocional y/o psicológica, facilitando tanto la experiencia universitaria como el desarrollo personal y profesional. Sus líneas de acción corresponden a la prevención, detección y derivaciones tempranas, seguimiento; y orientación y acompañamiento en la desvinculación [36].

La implicancia es indirecta, ya que un alcance del proyecto es potenciar el autoconocimiento, lo que llevaría, en el mejor de los casos, a la solicitud de ayuda. Es aquí donde la Unidad se ve referenciada, ya que tentativamente esos estudiantes podrían requerir horas de atención. Ahora bien, dado que se desconoce la identidad de los participantes, para evaluar este punto los psicólogos asociados deberían preguntar a los pacientes por el uso de la aplicación aquí desarrollada y qué los impulsó a la consulta. Dicho aumento de pacientes puede generar problemas de oferta de profesionales.

- **Unidad de Orientación Vocacional:** Esta unidad cumple el rol de orientar a los y las estudiantes de la facultad en torno a problemática académicas y vocacionales, apoyar en situaciones complejas de la vida universitaria y derivar a otras unidades en caso de encontrarlo pertinente, todo esto a través de una conversación confidencial, empática y diligente [38].

La implicación es indirecta, dado que el resultado esperado contempla el apoyo al autoconocimiento del estudiante, lo cual, junto a la retroalimentación personalizada de la aplicación móvil a ejecutar, incitarán a solicitar ayuda. Este último punto es el más relevante para la unidad, ya que su diagnóstico está centrado en que la baja demanda se debe a que no pueden diagnosticar a los alumnos sin que estos se acerquen a sus oficinas.

- **Unidad de Bienestar Estudiantil:** Servicio dependiente de la Subdirección de escuela, que tiene como misión intervenir y contribuir sobre variables socioeconómicas y personales que pueden afectar el rendimiento académico de los y las estudiantes [37].

El beneficio es indirecto, dado que los estudiantes presentan licencias médicas por causas de salud mental, por lo que al aumentar el auto conocimiento y en un futuro poder predecir estados de salud mental se podrán tomar medidas preventivas.

- **Estudiantes de la FCFM:** Se ven directamente implicados, ya que el trabajo está orientado a ellos, los datos son solicitados directamente y se entregarán resultados personalizados. Tanto la implementación de los cuestionarios como la retroalimentación personalizada permitirán tener mayor consciencia del estado anímico y las condiciones que lo gatillan.

Se espera que en un futuro se puedan tomar medidas preventivas a modo de evitar crisis de salud mental asociadas a las variables aquí estudiadas.

- **Núcleo Milenio ImHay:** ImHay está conformado por un grupo de investigadores provenientes de las ciencias de la salud, ciencias sociales y neurociencias. Su trabajo científico incorpora una mirada comprensiva, empíricamente fundada, sistémica, multinivel e intersectorial, que considera el curso de vida y la pertinencia cultural y territorial; y su objetivo principal es generar evidencia sobre las intervenciones para mejorar la salud mental de adolescentes y jóvenes [44].

Esta organización se ve implicada de manera directa porque, tal como indica su objetivo, a través de la memoria se está validando la factibilidad del uso de un modelo de clasificación de estudiantes, lo cual puede significar tanto educación como sensibilización y tratamiento precoz de la temática. Además, este trabajo da pie para una predicción de estados de salud mental a partir de modelos matemáticos y tecnologías.

Capítulo 2

Descripción General

2.1. Justificación

2.1.1. Contexto y problema

Durante el año 2019, por medio de las protestas asociadas a los temas de salud mental que realizaron los estudiantes de diversas casas de estudio se hizo evidente cuán deficiente es el sistema en cuanto al bienestar de los estudiantes respecta. De hecho, en los últimos años ha aumentado la prevalencia de enfermedades de salud mental en los universitarios, alcanzando el 20 % a nivel mundial [46].

En la misma línea, según la Primera Encuesta de Salud Mental Universitaria en Chile, las estadísticas indican que el 45.5 % de los alumnos sufre de ansiedad, 33.5 % de estrés, 46 % de depresión, y que el 29.7 % posee las tres patologías simultáneamente [15]. Sin embargo, solo el 11 % de ellos está en tratamiento psicológico, dejando ver una situación alarmante y una falta de atención que puede desencadenar crisis futuras con consecuencias a largo plazo.

Analizando la literatura, Baader et Al. [14] realizaron un estudio de diagnósticos de prevalencias de trastornos mentales a través de la aplicación de cuestionarios PHQ-9, HBS, MDQ, DEP-ADO, AUDIT, EDDS y RE, y factores de riesgo emocionales, demostrando que tanto la transición a la vida universitaria como la adquisición de más responsabilidades, las expectativas propias, los estresores psicosociales, el género y el perfil sociodemográfico afectan en mayor medida la prevalencia de trastornos. Lo anterior se traduce, entre otras cosas, en un comportamiento disruptivo, ansioso y afectivo en los estudiantes universitarios chilenos.

Frente a esto, cabe evaluar a fondo qué es lo causa o agrava el estado de salud mental, encontrando que existen múltiples causas que van desde la misma persona, su entorno y forma de vida, hasta el contexto en el cual se encuentra. En específico, una de las razones recae en el ámbito académico, donde Figueiredo et Al. [39] estudiaron el desgaste psíquico y los problemas de salud asociados a través de la aplicación de cuestionarios de desencanto, desilusión y percepción de carga académica. De este estudio se obtuvo una relación positiva entre 1) el desgaste psíquico por agotamiento con desencanto y desilusión, y 2) el desencanto psíquico por agotamiento con la sobrecarga académica percibida. Si bien este artículo se realizó sobre estudiantes de psicología, los resultados se extrapolan a estudiantes de carreras demandantes como medicina e ingeniería, donde hay alta competitividad y exigencia académica.

Tal como en el caso anterior, Pérez-Villalobos et Al. [56] estudiaron los problemas de salud mental de estudiantes chilenos de primer año universitario a los cuales aplicaron los cuestionarios BDI-II, BAI y AUDIT para medir depresión, ansiedad y consumo problemático de alcohol, respectivamente. También evaluaron características demográficas y culturales, encontrando que depresión y ansiedad se presentan con mayor frecuencia en género femenino, que el consumo de alcohol agrava la situación y tiene mayor efecto en el género masculino y que ni religión ni colegio de procedencia tienen gran incidencia sobre los resultados.

Así mismo, la Organización Mundial de la Salud (OMS) [12] también ha estudiado el tema e indica que en la época universitaria se ve reflejado el comportamiento seguido durante la adolescencia (10 a 19 años) de los jóvenes, siendo este último aquel donde se desarrollan los hábitos sociales y emocionales. En específico, explican que estos hábitos se ven afectados por la calidad de vida hogareña, las relaciones con los pares, el desarrollo personal, las condiciones de vida, las enfermedades crónicas, el trastorno de autismo, discapacidad intelectual, entre otras, que generalmente no son diagnosticadas a tiempo y que empeoran sin tratamiento, desencadenando así enfermedades mentales a futuro. En efecto, según estudios realizados sobre el 10 % de los estudiantes de cada facultad de la Universidad de Chile [1], se identificó que predomina la necesidad de afecto por parte de la familia por sobre los problemas económicos, teniendo la primera una incidencia mayor cuando al bienestar emocional respecta.

Incorporando más estadísticas relacionadas al tema recién presentado, según la información presentada por el Consejo Universitario de la Universidad de Chile, a partir del Formulario de Caracterización Estudiantil (FOCES) se obtuvo que el 46 % de los estudiantes destina al menos dos horas diarias a transporte entre el hogar y la facultad, lo cual tiene un impacto en las horas destinadas a sueño, descanso, ocio y actividades académicas. El 7 % trabaja y estudia, lo cual si bien no siempre provoca mal desempeño, si afecta en cuanto a estrés, distribución del tiempo, prioridades, entre otros. El 25 % no cuenta con un sector óptimo en su hogar para estudiar y el 37 % no tiene computador o comparte uno [7], condiciones que no son óptimas para una buena adaptación y uso de recursos fuera de la facultad. Además, el 46 % de ellos pertenece a la primera generación de su familia en acceder a la educación superior, agregando la presión y expectativas desde el mismo estudiante y desde su entorno familiar.

Cabe destacar que la Institución es consciente de que hay un problema y año a año modifica su programa para brindar apoyo a la comunidad estudiantil en aspectos psicológicos, económicos, emocionales, académicos y sociales. Ejemplos son los Programas de Acompañamiento Académico, Formación Docente, Herramientas de Coordinación y Gestión Docente, Centros de Enseñanza/Aprendizaje, Programas Financieros de Bienestar Estudiantil, Programas de Autocuidado y Desarrollo Cultural, Programas y Actividades de Salud Mental y Unidades de Bienestar. Sin embargo, en base a la entrevista realizada a Julio Lira, Orientador Vocacional de la FCFM, el principal problema para todas las unidades relacionadas es que el acceso a los programas puede ser activado una vez que el estudiante se acerca y solicita ayuda, lo cual sucede con menos frecuencia de la esperada, por lo que hay áreas subutilizadas y otras sobrecargadas.

De hecho, tal como se muestra en el “Informe sobre Programas de Acompañamiento Académico y Psicológico” [26] entre los años 2015 y 2018, en promedio ocurrieron 4.787 atenciones psicológicas y 1.018 atenciones psiquiátricas en el Servicio Médico y Dental (SEMDA) a lo largo de los cinco campus de la Universidad de Chile, llevadas a cabo por los 9 psicólogos y 2 psiquiatras pertenecientes al centro; es decir, cada profesional realiza aproximadamente 600 atenciones promedio al año, distribuidas como sigue a continuación:

- SEMDA Ingeniería: 55 horas semanales en psicología.
- SEMDA Juan Gómez Millas: 33 horas semanales en psicología.
- SEMDA Campus Sur: 33 horas semanales en psicología.
- SEMDA Central: 128 horas semanales en psicología y 55 en psiquiatría.

En el caso de la Unidad de Calidad de Vida, según el “Informe de Actividades 2014 - 2016” [3], en el año 2016 se brindó apoyo efectivo a 247 estudiantes distribuidos en 521 citas efectuadas de un total de 682 citas solicitadas. Este centro abarca tanto el ámbito clínico de primer acceso como el ámbito académico, por lo que es preocupante que se abarque a menos del 10 % del universo de la FCFM, demostrando así que actualmente la oferta es insatisfecha con la demanda; y en el caso de que fuera satisfecha, se requeriría un aumento de profesionales de la salud mental.

También, según la información presentada por las unidades internas de la FCFM, existen tres causales con alta frecuencia de atención: 1) actitudes normalizadas, como lo son pocas horas de sueño, largas estadías en la facultad y competitividad, 2) Alta carga académica, entendiéndose como alta cantidad de cursos por semestre y el calendario de actividades asociadas y 3) patologías propias que poseen los pacientes debido a situaciones no tratadas en su adolescencia o niñez y a la propia personalidad.

En consecuencia, a partir de lo expuesto anteriormente, bajo la creencia de que actualmente los usuarios utilizan los servicios proporcionados por la casa de estudios o acuden a profesionales externos cuando son capaces de reconocer que su estado anímico no es el que desearían tener y agregando los factores externos que inciden en la vida diaria, se presenta el siguiente árbol de causas:

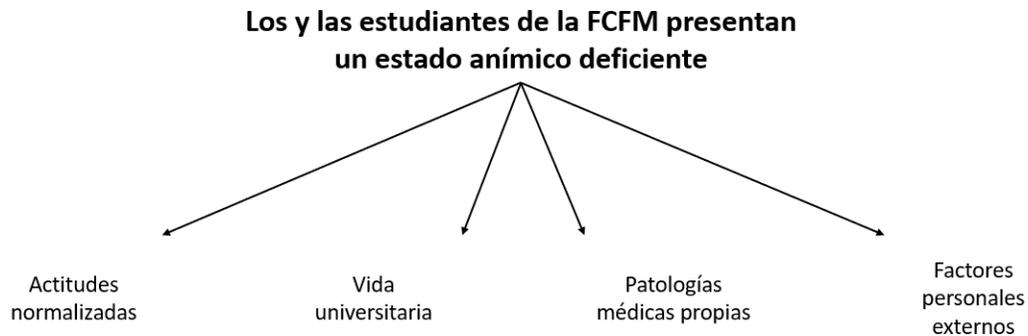


Figura 2.1: Árbol de Causas

Como se muestra en la Figura 2.1, se entiende que:

- **Actitudes normalizadas:** Se destacan situaciones como dormir pocas horas, faltar a clases para dormir o estudiar sobre ejercicios o controles, no dormir por más de un día para alcanzar a cumplir con las responsabilidades, llorar antes de una evaluación, no tener tiempo para realizar actividades que antes eran parte de su vida, no ver a la familia o amigos en semanas, entre otras.
- **Vida universitaria:** Mirada desde un punto de vista académico, trata de la cantidad de créditos cursados por semestre, la distribución horaria de los cursos, tareas, actividades en grupo y controles asociados, aprobación o reprobación de ramos, hábitos y métodos de estudio, entre otras.
- **Patologías médicas propias:** Consta de aquellas patologías o conductas reiteradas a lo largo de la vida del estudiante que se ven potenciadas o vuelven a aparecer luego de verse sometidos a situaciones de estrés. Cabe notar que cada persona reacciona distinto a los estímulos, es decir, una persona puede estar bien con carga académica elevada y trabajar bajo presión, mientras que otra persona presenta cuadros ansiosos o crisis de pánico por lo mismo.
- **Factores personales externos:** Corresponden a todas las situaciones o condiciones de vida a las cuales se ve expuesto el estudiante y que no dependen totalmente de sí mismo. Por ejemplo, problemas familiares, problemas económicos, discriminaciones, creencias religiosas, crisis dentro de una relación amorosa, discusiones con amistades, entre otras.

2.1.2. Enfoque del problema

Las variables expuestas apuntan a ser relevantes y potenciadoras de patologías futuras, lo cual deriva en consecuencias a corto plazo (en la universidad) y largo plazo (ambiente laboral y personal adulto). En específico, en la FCFM se cumple que la mayoría proviene de regiones, hay diversidad de pensamientos, religiones, procedencias, existen problemas económicos, emocionales, alta competitividad y alta exigencia, dándose así todas las condiciones para desestabilizar el estado anímico.

Sin embargo, a modo de realizar un trabajo más preciso y debido a que el foco del mismo está orientado a analizar la conducta anímica en base a la carga académica, se abarcará aquellos grupos de causas que están directamente relacionadas con esta última, dejando propuesto evaluar los factores intrínsecos y externos en un trabajo futuro. De esta manera se evitará sobrecargar a los participantes con cuestionarios muy largos, se mantendrá el *attrition* de la metodología EMA que será utilizada en el caso y se podrán enfocar esfuerzos según las conclusiones que se obtengan.

2.1.3. Efectos del problema

Sobre la base de un enfoque de los estados anímicos a partir de la carga académica, dentro de los efectos que provoca el problema descrito en el segmento anterior se encuentra un deterioro en la calidad de vida de los estudiantes, la cual puede definirse como la percepción que tiene el sujeto de sí mismo con respecto a las distintas dimensiones de la vida en un momento en específico. De aquí, y a partir del estudio realizado por Molina et Al. [33] donde se aplicó a el Test de Kidscreen-52 para medir la calidad de vida relacionada con la salud a estudiantes de primer año, se infiere que un deficiente estado de salud mental impacta de manera negativa en la calidad de vida, dado que se ven alteradas las metas, expectativas, preocupaciones y estándares de cada individuo.

Ahora bien, dado que este trabajo de título busca analizar la presencia de Depresión, Ansiedad y Estrés, a continuación se presentan los efectos del subdiagnóstico de cada uno, orientado a los estudios revisados anteriormente:

Depresión

La presencia de depresión a lo largo de la vida universitaria provoca efectos como sentimientos constantes de tristeza, vacío y desesperanza. Así mismo, provoca irritabilidad, frustración desmedida, pérdida de interés o placer en pasatiempos y actividades favoritos de los afectados, alteraciones del sueño, falta de energía, cambios en el apetito, en el desempeño académico, dolores musculares inexplicados, ansiedad, sentimientos de culpa, dificultad en la concentración y pensamientos negativos recurrentes.

Ansiedad

En el caso de la ansiedad también se genera una situación de nerviosismo, pánico y peligro inminente. Además, se presenta un aumento del ritmo cardíaco, hiperventilación, temblores, sensación de debilidad, problemas de concentración, problemas de sueño, problemas gastro-

intestinales, dificultad para controlar las preocupaciones y necesidad de evitar situaciones que generan ansiedad.

Estrés

Cuando se trata de estrés es posible que se den cuadros de diarrea o estreñimiento, mala memoria, dolores físicos y de cabeza frecuentes, falta de energía y concentración, problemas sexuales, cansancio extremo y problemas para dormir, malestares estomacales, uso de alcohol y drogas como relajantes y problemas de peso.

Finalmente, según los profesionales del Servicio Médico y Dental, para los estudiantes de la Universidad de Chile todos estos efectos no tratados de manera oportuna se transforman en patologías, siendo los diagnósticos más reiterados: Episodios depresivos, trastorno afectivo bipolar, trastorno de ansiedad generalizada o de ansiedad y depresión, trastorno específico de la personalidad y trastorno de pánico, derivando en situaciones como la solicitud de reducción de carga académica o entrega de facilidades, internación o reposo.

2.2. Población y grupo de estudio

La población corresponde a los y las estudiantes de la Facultad de Ciencias Físicas y Matemáticas que ingresaron a la carrera de Ingeniería y Ciencias, Plan Común entre los años 2014 a 2019, donde:

- Estudiantes de ingreso 2014: Corresponde a los estudiantes de último año de las carreras de ingeniería. Actualmente se encuentran en período de entrega de memoria de título o en proceso de continuidad de estudios hacia post grado.
- Estudiantes de ingreso 2015: Actualmente se encuentran cursando quinto año, terminando sus carreras en caso de haber seleccionado alguna licenciatura o en su penúltimo año en caso de escoger una ingeniería.
- Estudiantes de ingreso 2016: Actualmente se encuentran cursando cuarto año, el cual corresponde al segundo año del de la carrera escogida por cada estudiante.
- Estudiantes de ingreso 2017: Actualmente se encuentran cursando el segundo semestre del primer año de sus carreras de estudio. Nuevos compañeros y profesores, métodos de enseñanza, aprendizaje y exigencias. Recientemente se vieron enfrentados a situaciones estresantes como la elección de carrera, distanciamiento con amistades de Plan Común debido a la bifurcación de caminos y adaptación a los departamentos académicos en cuestión.
- Estudiantes de ingreso 2018: Corresponde a los estudiantes que actualmente cursan IV Semestre. En condiciones normales y sin reprobaciones cuentan con 6 cursos con un 88 % de aprobación promedio (según información disponible en Catálogo de Cursos FCFM [5]). Adicionalmente, están prontos a escoger un área de estudio dentro de las 13

carreras que ofrece la Facultad, lo que conlleva cambios a nivel emocional, contextual y académico.

- Estudiantes de ingreso 2019: Corresponde a estudiantes de Segundo Semestre de primer año de estudio. Cursan 6 ramos que, en promedio, cuentan con un 82 % de aprobación (según información disponible en Catálogo de Cursos FCFM [5]). Además, este grupo se encuentra en transición a la vida universitaria, adquiriendo nuevas responsabilidades, formas de vida y, en muchas ocasiones, trasladarse a Santiago para estudiar.

Se definió un grupo de estudio perteneciente a la población abarcando todos los años de ingreso ya presentados y catalogando a los participantes en tres grandes segmentos: el primero corresponde a los estudiantes de ingreso 2018 y 2019 los cuales se encuentran cursando Plan Común, el segundo a los estudiantes de ingreso 2015 a 2017 quienes se encuentran cursando la especialidad, y el tercero a quienes actualmente cursan su sexto y último año de estudio.

De aquí, luego de publicitar la memoria y la aplicación móvil creada para el levantamiento de datos, se obtuvo una participación voluntaria de 62 usuarios cuya distribución se explica más adelante en el informe.

2.3. Hipótesis principal y secundaria

De manera previa a la definición de las hipótesis es importante recalcar que realizar una clasificación de usuarios permite encontrar diferencias y similitudes entre grupos, así como identificar patrones y causas en común sobre alguna variable de interés, los cuales pueden ser analizados o tratados de manera individual y conjunta en el futuro.

2.3.1. Hipótesis principal

Se tiene como hipótesis principal que **es factible clasificar a los estudiantes según su estado anímico y sintomatología depresiva a partir de un modelo matemático que contempla:**

- Información personal: Caracterización del usuario según edad, región de procedencia, convivientes habitacionales y percepciones sobre sí mismo.
- Información académica: Carrera, semestre en curso, cantidad de créditos inscritos y percepción de carga académica.
- Cuestionarios auto reportados: Datos respectivos a sintomatología de ansiedad, depresión, estrés, percepciones y contextos, recopilados a través de una aplicación móvil.

El detalle de los cuestionarios y fuentes de datos nombradas se encuentran más adelante.

2.3.2. Hipótesis secundaria

Es factible predecir las variables de Depresión, Ansiedad y Estrés por un modelo de clasificación de estados anímicos y sintomatología orientada a caracterización del usuario, contexto situacional, higiene del sueño, información académica y percepción sobre sí mismo.

Para validar o no esta hipótesis es necesario estudiar cada respuesta obtenida por los usuarios, a modo de evaluar la posibilidad de segmentar a la población sin la necesidad de aplicar un cuestionario de estados anímicos de forma reiterada.

2.4. Impacto y diferenciación del proyecto

En la medida que se implemente la herramienta propuesta más adelante, se permitirá que los estudiantes sean más conscientes de las fluctuaciones de su estado anímico y de las condiciones que las provocan, pudiendo así evitar los riesgos de crisis, es decir, se enfrenta la principal preocupación expresada por las partes implicadas: el conocimiento, ya que entregará información relevante sobre los patrones y posibles situaciones de crisis a las autoridades correspondientes, y permitirá identificar comportamientos reiterados de los estudiantes, fomentando así la solicitud de atención psicológica y el posterior diagnóstico de patologías.

Por otro lado, se entregará un valor social al disminuir la cantidad de estudiantes con problemas asociados a su estado anímico, incitando a que estos acudan a centros de atención profesional ya sean dentro o fuera de la universidad, mejorando así la calidad de vida y el rendimiento académico de cada uno. Ahora bien, en caso de que utilicen el servicio psicológico que ofrece la casa de estudio se requerirá aumentar el personal destinado a salud mental para poder satisfacer la demanda. Dicha situación conlleva la potencial disminución de la cantidad de camas usadas por urgencias e internaciones de salud mental, debido al diagnóstico temprano y las estrategias de prevención.

Desde el punto de vista de las áreas involucradas, los resultados aquí encontrados permiten conformar una base de datos de una población universitaria, la cual puede ser utilizada para desarrollar proyectos con fines similares. De la misma manera, a través de este estudio piloto del modelo de clasificación se sustenta el estudio futuro de un modelo de predicción de estado de salud mental, que reemplace el formulario DASS-21.

Finalmente, el resultado de esta memoria puede ser extendido a todos los estudiantes universitarios, lo que requeriría un convenio con las casas de estudio para proveer de datos académicos en caso de que éstos sean significativos dentro del modelo de clasificación.

2.5. Objetivos

2.5.1. Objetivo general

En consecuencia, sobre el problema de que el estado anímico estudiantil está subdiagnosticado en la FCFM, en miras de la hipótesis principal se propone el siguiente objetivo general:

“Determinar la factibilidad del uso de un modelo de clasificación y de alerta del estado anímico y sintomatología depresiva de los estudiantes universitarios basado en tecnologías móviles”

2.5.2. Objetivos específicos

Para satisfacer lo anterior, se definen los siguientes objetivos específicos:

- Dar contexto general con respecto a la salud mental universitaria, y obtener variables de interés preliminares por medio de la recopilación y análisis de información primaria y secundaria respectiva al tema.
- Determinar la viabilidad de procedimientos asociados a estados anímicos, así como los participantes, el acceso a ellos, barreras de participación y disponibilidad de asignación al estudio de manera aleatoria.
- Recopilar información respectiva a la situación emocional, contextual/situacional y carga académica del estudiante por medio del desarrollo de una aplicación móvil como herramienta de monitoreo y evaluación remota.
- Generar un modelo de clasificación y definir la incidencia de las variables relevantes a partir de la determinación de estas.
- Determinar barreras de acceso, adhesión de los usuarios, idoneidad del grupo objetivo y riesgos de contaminación de los datos y modelos.

2.6. Alcances

La memoria de título aquí presente corresponde a un estudio para determinar la factibilidad de clasificar a los estudiantes según su estado de depresión, ansiedad y estrés por medio de un modelo matemático de clasificación. Esta contempla, además, el uso de un piloto de aplicación móvil para la recopilación de datos.

Para determinar los alcances se realizaron entrevistas y lecturas de investigaciones asociadas al tema. Además, se levantó información a través de la unidad de Orientación Vocacional

y de los informes respectivos de Calidad de Vida y Bienestar Estudiantil, de donde se identificaron tanto las variables potenciales por las cuales los estudiantes pueden sufrir las patologías y las principales causas tratadas actualmente, siendo estas:

- Carga académica en créditos.
- Percepción de carga académica.
- Caracterización sociodemográfica.
- Autoexigencia.
- Aislación social.
- Propósito de vida.
- Apreciación personal.
- Higiene del sueño.
- Ansiedad.
- Depresión.
- Estrés.
- Contexto situacional.

Por otro lado, las variables a no evaluar en este estudio son:

- Desórdenes alimenticios, debido a que no se registra en el historial de causales para la solicitud de horas a la Unidad de Calidad de Vida u Orientación Vocacional.
- Alto consumo de sustancias (drogas y alcohol), debido a que no hay evidencia dentro de la FCFM de que sea una causal para la solicitud de horas a la Unidad de Calidad de Vida u Orientación Vocacional.
- Suicidio, debido a que es un tema sensible y se puede gatillar tanto la aceptación de la situación para el usuario como la acción en sí misma, y la estudiante memorista no posee las capacidades ni facultades legales para tratar el tema.
- Factores intrínsecos, entendiéndose estos como aquellos comportamientos que tiene la persona y que no dependen de las circunstancias, dado que no estarían siendo consecuencia de la carga académica ni la relación con la universidad.
- Factores externos, siendo éstos los problemas familiares, de pareja, creencias, entre otras, debido que escapan del enfoque propuesto.
- Rendimiento y calendario académico, el cual difiere por cada estudiante debido a las consecuencias del estallido social que sufrió Chile en el mes de octubre del 2019. Dicha situación implicó un receso académico, nuevas condiciones evaluativas y cambios tanto en las actividades a realizar, como en evaluaciones y exigencia, viéndose enfrentados a un fin de semestre extraordinario. Sin embargo, se agregaron preguntas relacionadas debido a que están ligadas con el enfoque propuesto.

Finalmente, debido a atrasos en la programación y entrega de la aplicación móvil, solo se utilizarán las funciones de levantamiento de datos por medio de interacción con el usuario, quedando el uso de sensores del dispositivo telefónico disponible para trabajos posteriores.

2.7. Apoyos institucionales

2.7.1. Financiero

Esta memoria fue parcialmente financiada por el **Proyecto U-Redes-C 2018 07** alojado en el Hospital Clínico de la Universidad de Chile, y liderado por la Dr. Graciela Rojas, directora del mismo centro.

Por otro lado, también se contó con el apoyo financiero de **Bauchef Proyecta**, quienes destinan presupuesto para el desarrollo de las memorias interdisciplinarias que ofrecen a los y las estudiantes.

Cabe destacar que ambos apoyos financieros fueron destinados al pago por concepto de la aplicación móvil para Android que permitirá realizar la toma de datos auto reportados y adquiridos de manera pasiva.

Finalmente, **Fondecyt - Proyecto 1181809** otorgó un porcentaje de su prepuesto para subsidiar a la estudiante durante el desarrollo de esta memoria.

2.7.2. Académico

Desde el lado académico, se trabajó directamente con los profesores de la carrera de Ingeniería Civil Industrial de la Universidad de Chile, Dr. Ángel Jiménez y Dr. Daniel Schwartz, quienes corresponden a los profesores guía y co-guía, respectivamente. Además, se recibió el apoyo constante de la Red de Salud Mental de la Universidad de Chile, específicamente de la Dra. Vania Martínez, directora del Núcleo Milenio ImHay; y el Dr. Álvaro Jiménez, Investigador del Núcleo Milenio ImHay. Este equipo multidisciplinario proporciona conocimiento en áreas de tecnología, estadística y salud mental.

También se mantuvo relación presencial con la Unidad de Orientación Vocacional, específicamente con el profesor Julio Lira, quien está encargado de recibir a los estudiantes que poseen bajo rendimiento académico, problemas personales y/o que son derivados por diversas temáticas.

Por último, está contemplada la colaboración con diversos estudiantes de doctorado, máster y pregrado que se encuentran trabajando en proyectos de características similares, ya sea a través del uso compartido de información o aplicaciones.

Capítulo 3

Marco Conceptual

3.1. Conceptos y cuestionarios de salud y psicología

3.1.1. Conceptos

- **Salud:** “Un estado de completo bienestar físico, mental y social, y no solamente la ausencia de afecciones o enfermedades” [28].
- **Salud mental:** Bienestar emocional, psicológico y social, afectando la forma de pensar, sentir y actuar de las personas [54].
- **Estudiante universitario:** Toda persona que curse asignaturas en alguna institución de educación superior, y que en promedio se sitúe entre los 18 y 29 años [35].
- **Trastorno del sueño:** Este trastorno afecta seriamente la calidad de vida y tiene como síntomas la somnolencia diurna o sensación de no haber logrado descanso reparador, sensaciones desagradables en las piernas, falta de concentración y dificultad para realizar las tareas cotidianas [22].
- **Depresión:** Trastorno del estado de ánimo en donde los sentimientos de tristeza, melancolía o infelicidad interfieren con la vida diaria. Puede ser temporal, permanente y se clasifica en términos de leve, moderada o severa [61].
- **Trastorno de ansiedad generalizada:** Patrón de preocupación y ansiedad frecuente y constante acerca de muchos acontecimientos o actividades diferentes. Puede comenzar en cualquier momento de la vida y es ligeramente más frecuente en mujeres que en hombres. Sus síntomas además pueden abarcar dificultad para concentrarse, fatiga, irritabilidad, problemas para conciliar el sueño y permanecer dormido, nerviosismo, entre otros [62].

- **Estrés:** Sentimiento de tensión física o emocional que puede provenir de cualquier situación o pensamiento que genere frustración, furia o nerviosismo. Puede ser positivo si sucede en períodos cortos o dañino para la salud en caso contrario [53].

3.1.2. Cuestionarios y escalas

- **Patient Health Questionnaire (PHQ-4):** Cuestionario simplificado de auto reporte que permite detectar sintomatología depresiva leve, moderada o severa. Ha demostrado ser un instrumento de diagnóstico con una sensibilidad del 92% y especificidad del 89% en la detección de pacientes depresivos. Además, presenta validez de constructo y predictiva recurrente con los criterios de CIE-10 para depresión; y su versión en español está validada en Chile [48].
- **Escala de Depresión, Ansiedad y Estrés (DASS-21):** Instrumento estandarizado que permite detectar las patologías de ansiedad, estrés y depresión de manera oportuna y auto reportada. Su validez en versión abreviada en español para los estudiantes universitarios chilenos está dada por Alpha de Cronbach, y su validez de concurrente y divergente fue verificada con el BAI, BDI – II y SCL - 90 – R [63].
- **Escala de Estrés Percibido (EEP-4):** Es un instrumento simplificado que permite medir el grado en el que las personas evalúan las situaciones cotidianas a las que se ven enfrentados. Estas pueden ser catalogadas como impredecibles, incontrolables o sobrecargadas. Tiene una consistencia interna entre 0.6 y 0.8, y estructura interna de factor único [34].

3.1.3. Estudios de factibilidad y RCT

Un estudio de factibilidad busca establecer la posibilidad de realizar un estudio completo (factibilidad) y si las componentes necesarias para el mismo funcionarán de manera correcta en su conjunto (piloto)[8] [18]. Es utilizado para planificar preliminarmente un Ensayo Clínico Aleatorizado (Randomized Clinical Trial o RCT por sus siglas en inglés), pudiendo realizar medidas de los resultados y modificaciones previas al trabajo completo. Así mismo, permite determinar las condiciones que se deben cumplir en términos de diseño del experimento, acceso a los participantes y profesionales pertinentes, acceso a los datos y cantidad de casos a evaluar si es que no existen recopilaciones previas.

Los resultados deben ser descriptivos, incorporando el reclutamiento, capacidad de administrar los tratamientos evaluados, aceptabilidad del proyecto, fidelidad, adherencia, retención, integridad y adecuación de resultados, debido a que el contexto en el cual fueron obtenidos puede cambiar los resultados y, por ende, causar errores de interpretación[18]. Cabe destacar que debido al uso de pequeñas muestras, los valores obtenidos pueden ser engañosos, por lo que los estudios estadísticos deben ser tratados con extrema precaución y no ser utilizados para tomar conclusiones.

3.1.4. Evaluación Ecológica Momentánea (EMA)

Como describen Shiffman et Al.[42] *Ecological Momentary Assessment* por sus siglas en inglés corresponde a un conjunto de métodos de investigación que permite realizar un seguimiento del estado del sujeto en cuestión, debido a que utiliza una colección repetida de datos que son recolectadas de manera auto reportada y en tiempo real (M: Momentary) en un ambiente no intervenido (E: Ecological) por medio de una serie de tareas que se ejecutan aleatoriamente durante el día, por un período de tiempo definido (A: Assessment).

Además, como muestra el estudio de literatura de Colombo et Al. [24], se identificó que el método se ha aplicado en diversos campos: desregulación de cortisol, patrones de sueño, actividad física, ruminación y afecto y reacción emocional; y que a través de éste es posible reducir el sesgo por retrospección, mejorar la comprensión de procesos que dinámicos a lo largo del tiempo, realizar un análisis de comportamiento y señales psicológicas día a día, contemplar el contexto, obtener una colección de datos representativos y entregar una retroalimentación personalizada en tiempo real. Así mismo, el estudio indica que el uso de este instrumento presenta un promedio de 73 % de retención y bastante variedad en formato, pudiendo presentarse entre 1 a 20 veces al día por períodos cortos de 3 días (para evaluar patrones de sueño, por ejemplo) o más largos de 1 o 2 meses (para evaluar la asociación de actividad física con depresión, entre otros).

Históricamente se ha implementado EMA vía papel, pero hace un tiempo se incorporó el uso de tecnologías. Aquí tanto Colombo et Al. como Larry Chan et Al. [23] declaran que el uso de tecnologías puede potenciar de mejor manera los beneficios antes nombrados, ya que permite tomar datos reportados por el usuario y adquiridos a través de los sensores del dispositivo. Algunos ejemplos son Mobylyze!, Help4Mood y Medlink.

Estos últimos autores realizaron una serie de “focus groups” para analizar cómo afecta la frecuencia, extensión y diseño de preguntas en las respuestas de los estudiantes, encontrando que en promedio los usuarios respondían entre 15 a 19 preguntas repartidas entre 10 ejecuciones al día, entre las 9:00 y 22:00. Además, los participantes indicaron que: 1) Estarían dispuestos a responder con mayor frecuencia sobre su estado anímico y estrés, 2) Es más sencillo responder frente a una pregunta certera con escalas que frente a imágenes, y 3) A través del estudio fueron siendo más conscientes de su estado anímico.

3.1.5. Tratamiento de datos

- Balanceo de datos: Es bastante común que los datos se encuentren desbalanceados, es decir, que haya una mayor recurrencia de una clase por sobre la otra, lo cual conlleva problemas en modelos como árboles aleatorios y otros que priorizan minimizar el error [2]. Por ello el balanceo de datos corresponde a un paso necesario para obtener resultados más certeros. El método más conocido es el remuestreo, que genera datos ficticios a partir de los conocidos y/o elimina datos de la clase mayoritaria. El más usado es SMOTE (en inglés, *Synthetic Minority Over-sampling Technique*) que permite balancear dos clases al generar datos ficticios en la clase minoritaria a través de combinaciones lineales con los datos cercanos a la vecindad [4]. En la Figura 3.1 es posible ver la clase minoritaria en rojo oscuro y los valores ficticios creados en rojo

claro, obteniendo así clases balanceadas.

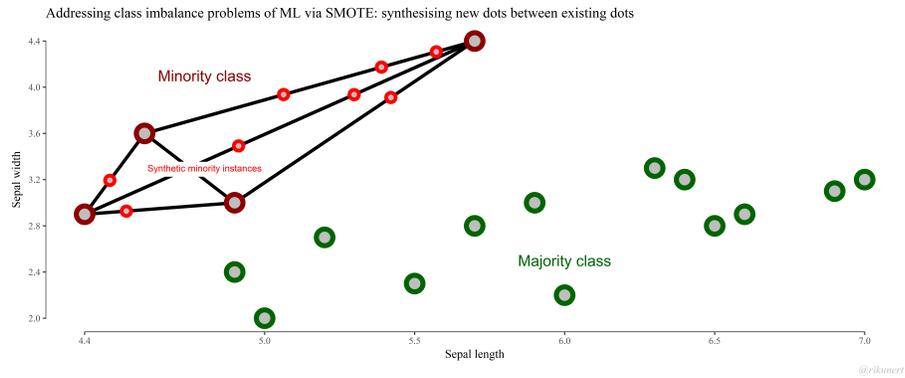


Figura 3.1: Ejemplo de aplicación de SMOTE

- **Análisis de Componentes Principales:** El PCA (*Principal Component Analysis* por sus siglas en inglés) es una técnica que permite reducir la dimensionalidad de grandes conjuntos de datos ya sea numéricos o mixtos, a modo de aumentar la capacidad de interpretación minimizando la pérdida de información. Para esto crea nuevas variables no correlacionadas que maximizan la varianza a partir de combinaciones lineales de las variables existentes y luego selecciona aquellas que presentan una mayor varianza acumulada con respecto a un porcentaje de corte seleccionado [47]. De manera gráfica, en la siguiente Figura se puede revisar la proporción de la varianza que explica cada componente:

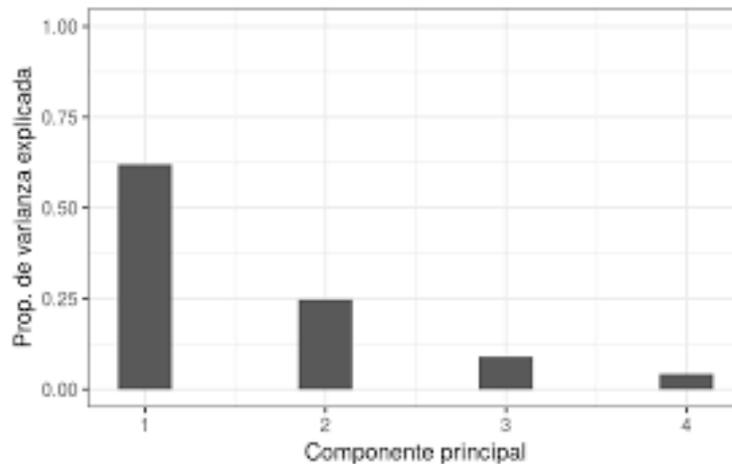


Figura 3.2: Ejemplo de gráfico de proporciones de PCA

- **Matriz de Correlación:** Antes de trabajar con modelos es necesario identificar aquellas variables que son de interés para obtener un resultado que se ajuste de mejor manera a los datos reales. Para esto existe la matriz de correlación, que permite identificar qué tan relacionadas se encuentran las variables entregando el coeficiente y el tipo de correlación (positiva o negativa). Se puede representar de manera gráfica y numérica, mostrándose la primera opción en la Figura 3.3 donde el color azul representa

correlación positiva disminuyendo según la gradiente del color hasta convertirse en rojo representando correlación negativa [19].

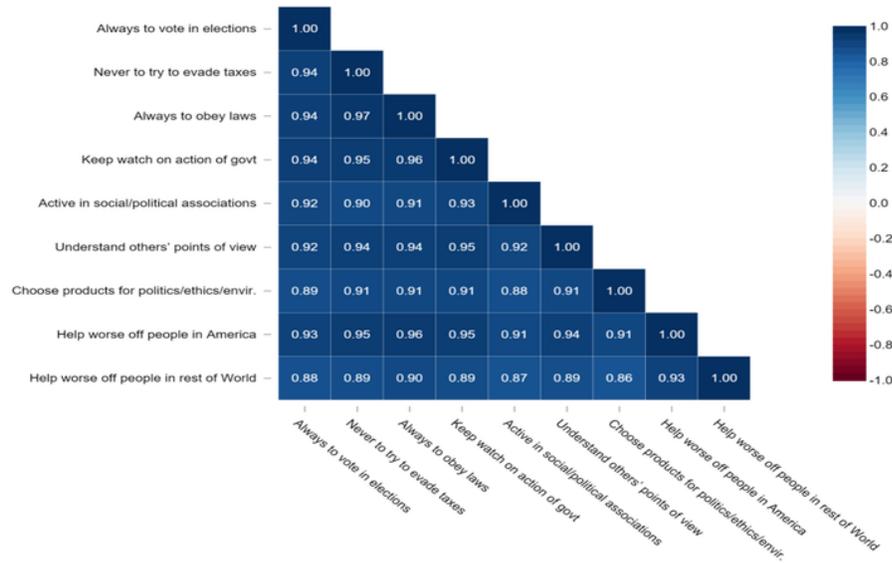


Figura 3.3: Ejemplo de matriz de correlación

3.1.6. Modelos de clasificación

Según Entezari-Maleki et Al. [32] los modelos de clasificación son en base a aprendizaje supervisado y buscan predecir el comportamiento de la variable dependiente a partir de comportamiento conocido y la influencia de las variables independientes asociadas. En esta memoria se considerarán:

- Árboles de decisión:** Esta herramienta sigue el aprendizaje inductivo y son similares a un diagrama de flujo, donde cada nodo significa una prueba de un atributo y cada rama representa el resultado de la prueba. De aquí, el nodo superior es el nodo raíz, mientras que los nodos hijos son cada posible respuesta a los cuestionamientos sobre la variable en estudio. Luego, los nodos finales representan una decisión coincidente con una variable de clase.

El proceso consta de dos etapas, donde la primera corresponde a la inducción del árbol: sobre el conjunto de datos se escoge un atributo de prueba y el conjunto de entrenamiento restante se divide en subconjuntos. Cuando en un nodo se tienen objetos de más de una clase se genera un nodo interno; de lo contrario, se forma una hoja, la cual lleva una etiqueta de clase. Por otro lado, la segunda etapa consta de clasificar cada objeto nuevo, el que sigue un camino basado en las decisiones que se tomen en cada nodo interno, hasta llegar a una hoja o nodo final [57].

- Bosques aleatorios:** Corresponde a una extensión del modelo de árbol, siendo una implementación más avanzada de un algoritmo de agregación. Aquí, cada árbol se genera

a partir de una muestra trazada por sustitución a partir del conjunto de entrenamiento. Luego, se obtiene el resultado de cada árbol y la clasificación predominante obtenida corresponde al resultado del bosque [21].

- **Máquinas de vectores de soporte (*Support Vector Machine*):** Método de clasificación para datos lineales y no lineales, en el cual se utiliza un mapeo no lineal para transformar los datos originales de prueba a una dimensión mayor. Luego, en dicha dimensión, se busca el hiperplano de separación lineal óptimo usando vectores de soporte y márgenes [45].
- **Regresión logística (*LOGIT*):** Corresponde a un modelo de regresión que mediante el uso de la función logit establece la probabilidad de que la variable dependiente adquiera uno de los dos valores posibles, a partir de la estimación de los parámetros beta asociados a cada variable explicativa. De aquí, cuando se tiene una variable dependiente binaria con valores posibles 1 y 0, y sucede para un caso i , se tiene que [25]:

$$\frac{1}{1 + e^{-\beta^T X_i}} \quad (3.1)$$

Con:

k : número de variables explicativas

X_i : $(X_{(1,i)}, \dots, X_{(k,i)})$ vector de variables explicativas del caso i

β : $(\beta_1, \dots, \beta_k)$ vector de parámetros estimados

p_i : Probabilidad de que la variable dependiente del caso i sea igual a 1

3.2. Estado del arte

3.2.1. Estrés e higiene del sueño en estudiantes universitarios

Gotelli et Al. [13] realizaron un estudio con el objetivo de estimar la prevalencia de los principales desórdenes no psicóticos en estudiantes universitarios de en promedio 21 años. A ellos se les aplicaron los cuestionarios: Salud del Paciente (PHQ-9), Escala de Desesperanza de Beck (HBS), Trastorno del Estado de Ánimo (MDQ), Detección de Consumo Problemático de Alcohol y Drogas en Adolescentes (DEP-ADO), Cuestionario AUDIT, Escala de Diagnóstico del Trastorno Alimentario (EDDS) y Escala de Restricción (RE). Los resultados indican que el 13,7% cumple con criterios de depresión mayor, el 27% posee algún tipo de depresión y el 10% posee un trastorno bipolar (ambos más frecuentes en mujeres). Además, el 5% presenta alta puntuaciones en desesperanza (sin mayores variaciones según el sexo del participante), el 24% presenta un consumo riesgoso de alcohol, y el 30% ha consumido marihuana en el último año (más frecuentes en hombres). Estas patologías coinciden con las estudiadas por Barrera [15], haciendo notar la relevancia y prevalencia de las mismas dentro

de las actuales generaciones universitarias.

Llevando lo encontrado al árbol de causas de la Figura 2.1, los artículos nombrados en el párrafo anterior, Antúnez & Vinet [10] y Lin et Al. [50] basan estos resultados en la etapa evolutiva por la que pasan los estudiantes al egresar del colegio y entrar a la universidad, período en el cual comienzan a manejarse a sí mismos, asumen otras responsabilidades, deben responder a expectativas, vivir alejados de la familia, sustentarse económicamente, adaptarse a un nuevo contexto y enfrentarse a nuevas exigencias académicas. En sí, estas situaciones se centran en la adaptación al cambio académico y personal derivado, y son factores que gatillan estrés, el cual puede ser bueno para ciertas circunstancias, pero que tiene impacto negativo sobre la salud y, en caso de no ser tratado, se puede convertir en una patología.

También se identifican problemas en la higiene del sueño, donde Amaral et Al. [9] analizaron estudiantes a través cuestionarios de percepción de estrés y pensamiento de perseverancia, encontrando una relación directa con el estrés, el afecto negativo, el pensamiento negativo, autculpabilidad y otras estrategias cognitivas emocionales autorreferentes.

A modo de conclusión, el estrés es causado por seis grandes grupos de causas y subcausas:

1. Notas: Sentimiento de no estudiar lo suficiente, cansancio debido a pocas horas de sueño para continuar estudiando.
2. Resultados: Miedo a reprobado ramos, diferencia en el rendimiento del colegio en comparación a la universidad, no poder mejorar el rendimiento.
3. Padres: Sentimiento de no cumplir las expectativas o que crean que no se están tomando la universidad en serio.
4. Profesores: Exceso de carga académica, métodos de enseñanza poco efectivos, mucho material asociado y en otros idiomas.
5. Compañeros/as de curso: Trabajos en equipo, inseguridades al presentarse frente a otros, competencia en desempeño, ambientes y grupos de estudio, sobrecarga y distribución de actividades.

A esto se agrega uno de los factores más importantes: la persona en sí misma, ya que se reconocen variables de estrés como la autoexigencia con respecto al rendimiento académico, compararse con el resto de los compañeros en diversos ámbitos y participación en actividades sociales o extracurriculares y manejo del tiempo.

Todo lo anterior será preguntado a través de cuestionarios de percepción de carga académica, higiene del sueño, bienestar emocional y estrés percibido, utilizando para este último la Escala de Estrés Percibido en su versión simplificada de cuatro preguntas.

3.2.2. Depresión y ansiedad en estudiantes universitarios

Depresión y ansiedad son dos patologías que generalmente se presentan en conjunto, pero que no necesariamente están ligadas. Hettema [41] indica en su estudio que más de la mitad de los pacientes que padecen depresión desarrollan trastorno de ansiedad a lo largo de

su vida, lo cual se debe, tentativamente, a la relación de origen encontrada en la genética del fenotipo, desde donde también se desarrollarían condiciones como ataques de pánico y fobias. De aquí, si bien no se busca una relación causa consecuencia, es pertinente tratar ansiedad y depresión de manera separada, e interesante estudiar su correlación en los individuos.

Por otro lado, históricamente se ha pensado que las mujeres presentan una mayor prevalencia de depresión y ansiedad que los hombres, pero esto puede deberse simplemente al hecho de que las mujeres tienden a expresar más sus sentimientos y solicitar ayuda profesional. Ibrahim et Al. [43], Antúnez & Vinet [10] y Wiseman et Al. [65] indican que aunque la diferencia no es demasiado significativa, si existe. En particular, indican que los hombres presentan una mayor prevalencia de soledad y falta de expresión que las mujeres, mientras que estas tienen mayor prevalencia de depresión y ansiedad. Además, destacan que ambas condiciones están correlacionadas de manera moderada, lo que puede deberse a que comparten factores de origen, los que, como se mostró en el párrafo anterior, se ven potenciados por la genética de las personas estudiadas.

Para analizar estos descubrimientos, dentro del cuestionario de caracterización de este trabajo se solicita a los estudiantes su género y percepción de soledad en la etapa universitaria (con quién vive, región de procedencia, si tiene un contacto con quien conversar cuando se siente solo/a, etc.) esperando concluir lo mismo que los autores; y se evalúan ambas patologías de manera conjunta a través del cuestionario PHQ-4, donde las dos primeras preguntas están orientadas a depresión y las dos últimas a ansiedad.

Los artículos expuestos concluyen que la depresión se asocia significativamente con una mayor tendencia a alimentación anormal, consumo nocivo de alcohol, comportamiento auto-destructivo, disturbio del sueño, falta de autocuidado, ansiedad, falta de interés en actividades cotidianas y falta de concentración [43] [65] [52]. Estas actitudes ocurren con mayor frecuencia de la deseada dentro de los y las estudiantes de ingeniería y no siempre son reconocidas como potenciadoras o síntomas de patologías, sino que entendidas como parte de la vida universitaria.

Además, Oliver & Burkham [52] estudiaron la relación entre depresión - edad, año de estudio y calendario académico, variables que se esperan sean determinantes en la aparición de depresión, ansiedad y estrés. En específico, se espera una relación positiva con la carga académica (entiéndase entrega de tareas y rendición de pruebas) ya que tienden a ser situaciones de alta exigencia e intensidad dentro del semestre, y negativa con el año cursado, entendiendo que los primeros años son más difíciles en términos de adaptación y toma de decisiones. Los resultados que encontraron indican que, en efecto, existe una relación negativa entre año cursado y depresión, no así con la edad ni el calendario académico, donde no hay diferencias significativas y se presenta un patrón luego de tres semanas de aplicar cuestionarios, respectivamente. Sin embargo, estas variables serán estudiadas debido a que son reiteradas causas en conversaciones con el estudiantado y permitirán obtener conclusiones propias a la carrera y grupo de estudio.

Desde el lado de la ansiedad, al igual que el resto de las patologías, Antúnez y Vinet [10] afirman que la etapa de transición colegio universidad y la adolescencia juegan un rol fundamental al momento de la aparición de trastornos de salud mental y ansiedad debido

al cambio en la responsabilidad del estudiante, nuevas condiciones de vida, exigencias académicas y contextos sociales. En su estudio contextualizaron a los participantes según año cursado, lugar de procedencia, género y nivel socioeconómico, y aplicaron los cuestionarios DASS-21, Inventario de Síntomas de Derogatis, Inventario de Ansiedad de Beck e Inventario de Depresión de Beck en modo auto-reportado. De ellos, identificaron una mayor declaración de sintomatología y prevalencia de ansiedad y psicoticismo en mujeres, mientras que depresión, paranoia y estrés en hombres.

En conclusión, depresión y ansiedad corresponden a patologías que se presentan con alta frecuencia y que en su mayoría son mal diagnosticadas. En este estudio no se pretende diagnosticarlas, sino que identificar los patrones de comportamiento que poseen los y las estudiantes universitarias que son comunes dentro de la población clínica revisada por todos los autores presentados. Ahora bien, como el grupo de estudio seleccionado no es clínico, las preguntas a realizar serán de carácter sencillo y contemplarán variables como el género, nivel de estudios, región de procedencia, ansiedad, depresión, estrés, higiene del sueño y contexto situacional; todo esto utilizando los cuestionarios PHQ-4, que corresponde a una versión simplificada del PHQ-9 utilizado en los artículos mencionados, y DASS-21.

3.2.3. Aplicaciones móviles en salud mental

Torous et Al. [60] realizaron una encuesta sobre el interés en el uso de aplicaciones móviles para monitorear la salud mental a través de EMA en pacientes que presentaban, en su mayoría, trastornos de ansiedad y de depresión. De ellos, dado que la población objetivo de esta memoria de título son estudiantes universitarios de entre 18 y 27 años, se consideran relevantes los resultados del grupo menor a 30 años, donde el 57 % declaró querer recibir un mensaje de texto por parte del centro asistencial, un 82 % quería acceder a información de salud mental a través de sus teléfonos móviles, un 78 % quería usar una aplicación para registrar los cambios en la condición de salud mental, un 71 % descargaría una aplicación relacionada y un 72 % la usaría de manera diaria. Estas cifras no son completamente extrapolables debido a que los encuestados contaban con tratamiento psiquiátrico, a diferencia de los sujetos de este estudio. Sin embargo, corresponden a un buen indicador de tasas de descarga de la aplicación e interés en utilizar la misma.

En línea con lo anterior, el Instituto Nacional de Salud Mental [51] indica que, además de facilitar procedimientos, el uso de aplicaciones de seguimiento permite detectar cambios en el comportamiento o estado de salud a través de sensores, pudiendo alertar a terceros cuando el usuario presenta riesgo de crisis. También, destaca que hay una conveniencia en tiempo y lugar, permiten mantener la anonimidad, funcionan como un punto de partida para comenzar a cuidar la salud mental, son de bajo costo y gran alcance, facilitan o entregan monitoreo las 24 horas y complementan tratamientos presenciales. Sin embargo, dada la poca indagación en estos temas, se duda de la efectividad que pueden tener en comparación a los métodos tradicionales y, además, requiere una garantía tanto a los datos recopilados, como a los tipos de preguntas realizadas y el manejo de la información durante y después de ejecutados los cuestionarios.

En lo que a aplicaciones existentes se refiere, es posible encontrar MindShift™ CBT [20], lanzada en Canadá para ayudar a los usuarios a relajarse y estar atentos, desarrollar maneras

más efectivas de pensar y usar pasos activos para hacerse cargo de su ansiedad. Desde su relanzamiento en febrero de 2019, hoy cuenta con más de 75.000 usuarios activos. En Estados Unidos es posible encontrar aplicaciones como Moodpath [55] que permite llevar un registro y monitoreo de sintomatología depresiva. Además, facilita la comunicación con centros de asistencia entregando reportes cada dos semanas. Posee una evaluación promedio de 4.6 sobre 5 y supera las 500.000 descargas. Como estas, existen muchas otras con distintos enfoques a lo largo del mundo y en sí los argumentos del uso de las aplicaciones para monitorear a la persona demuestran la capacidad que posee el paciente de responder las actividades requeridas por la aplicación, y la efectividad del auto reporte a través de esta. Además, se sustentan en la conclusión de que el auto reporte diario en formato EMA puede ayudar a predecir pensamientos de daño auto provocado.

Ahora bien, Becker et Al. [16] estudiaron cómo predecir el estado anímico a partir de datos recopilados a través del teléfono, y cómo influye este método con el desempeño de las predicciones de una serie de modelos estadísticos. El proyecto consistió en implementar una serie de cuestionarios vía la aplicación móvil iYouVU con 5 ejecuciones al día, por 6 semanas, datos que fueron previamente tratados y luego llevados a los modelos: *Mean Model*, *Linear Regression*, *Support Vector Machine*, *Lasso Regression* y *Bayesian Hierarchical Linear Regression*, cuyos resultados más significativos fueron: Mean Model obtuvo un mayor Root Mean Square Error (RMSE) por sobre el resto de los modelos, que solo los conceptos psicológicos medidos influyen significativamente en el estado de ánimo, y que se obtienen mejores resultados cuando se trata a los sujetos de manera individual.

Además, destaca otros estudios similares como la aplicación móvil que ayuda a individuos depresivos a través de EMA realizada por Burns, que infiere el estado mental del usuario utilizando información de localización, actividad, entorno social y uso del teléfono; y se demostró que los patrones de cambio anímico obtenidos de los registros de GPS ayudan a determinar síntomas depresivos. Así mismo existen aplicaciones como MoodMiner y MoodScope que siguen la misma metodología y obtuvieron resultados similares. Sin embargo, no es posible asegurar que los datos tomados de manera pasiva contribuyan significativamente a la predicción.

En línea descrito en los párrafos anteriores, el uso de la tecnología ha sido extendido a campos médicos. En específico, se han implementado tratamientos de salud mental, permitiendo así que más pacientes sean cubiertos, se mejore el autoconocimiento y se puedan tomar mejores decisiones conforme la fidelidad de los datos obtenidos. De hecho, se destaca la facilidad de uso en tiempo y lugar, el bajo costo y masivo acceso; y el monitoreo de 24 horas que permite realizar con cada uno de los involucrados. Sin embargo, también se indica que se debe garantizar la privacidad de los datos, analizar a cada persona por separado y ser responsables tanto con el tratamiento de la información como la presentación de esta.

Frente a lo anterior, resulta pertinente implementar la toma de datos a través de aplicaciones móviles. Para esto, el formato EMA presenta una alta tasa de respuesta y fidelidad de los registros, dado que las preguntas se realizan de manera aleatoria y reiterada, captando información en tiempo real y evitando el sesgo de respuesta por retrospcción o situacional. Cabe destacar que esta será proporcionada de manera gratuita a los usuarios con una interfaz amigable y de fácil uso; contemplará formularios EMA con recordatorios de llenado y los

datos recopilados serán guardados de manera anónima bajo altos estándares de seguridad y utilizados únicamente para fines académicos. Además, no se entregará un diagnóstico a los participantes, sino que solo se clasificarán respecto a los indicadores obtenidos y se entregará una recomendación personalizada de dónde acudir para obtener ayuda profesional en caso de ser requerida.

Capítulo 4

Metodología

Se adoptó la metodología de Descubrimiento de Conocimiento e Información desde Bases de Datos o *Knowledge Discovery in Databases* (KDD), la cual corresponde a una de las metodologías de manejo de datos más utilizada cuando se trata de grandes volúmenes [6], y consta de todo el proceso de descubrir conocimiento útil a partir de los datos, incorporando teoría, algoritmos y métodos de *machine learning*, estadística, computación de alto desempeño, entre otros.

Sus etapas corresponden a la selección, preprocesamiento, transformación, minería de datos e interpretación de los datos, tal como se puede ver en la Figura 4.1:

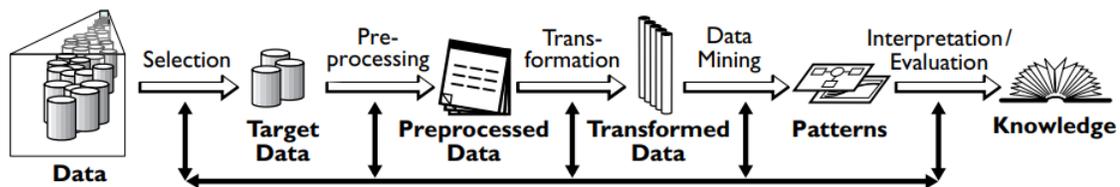


Figura 4.1: Etapas de Metodología KDD

1. **Selección:** Selección de los datos o el enfoque del trabajo en un conjunto de variables o segmentos donde se pueden obtener descubrimientos relevantes (*Target Dataset*).
2. **Pre-procesamiento:** Tratamiento de los datos seleccionados realizando operaciones básicas como la remoción de datos atípicos y modificación de datos faltantes o incompletos, con la finalidad de obtener un conjunto de datos limpios para trabajar (*Preprocessed data*).
3. **Transformación:** Encontrar las herramientas relevantes para representar los datos según el objetivo que se quiera lograr a partir de la misma. Aquí se utilizan métodos de reducción y/o transformación, obteniendo así un número reducido de variables o datos representativos (*Transformed data*).

4. **Minería de datos:** Desarrollo de modelos analíticos orientados al cumplimiento de los objetivos del proceso, los cuales permitirán identificar patrones desconocidos por medio del aprendizaje obtenido de los datos transformados (*Patterns*). Estos modelos pueden ser de segmentación o clasificación.
5. **Interpretación y análisis:** Interpretación de los resultados obtenidos, así como un posible retorno a los pasos previos para obtener nuevos patrones, los cuales son analizados y traducidos en conocimiento (*Knowledge*).

En línea con el objetivo general planteado en la sección 2.5, a continuación, se muestra un resumen de la metodología a seguir para lograr el cumplimiento de los objetivos específicos. Todo esto con la finalidad de aplicar la metodología KDD de obtención de conocimiento a partir de los datos, el que será utilizado posteriormente para validar o no las hipótesis presentadas.

Objetivo específico 1: Dar contexto general con respecto a la salud mental universitaria, y obtener variables de interés preliminares por medio de la recopilación y análisis de información primaria y secundaria respectiva al tema.

- Levantamiento y análisis de artículos relacionados con factores que explican la salud mental de la población universitaria, referidos principalmente a depresión, ansiedad, estrés y factores incidentes relevantes como: higiene del sueño, carga académica y contexto situacional.
- Análisis de estadísticas nacionales e internacionales respecto a la prevalencia de enfermedades de salud mental en estudiantes universitarios.
- Obtención de variables preliminares a través la solicitud y el análisis de datos de caracterización del estudiante: edad, región de procedencia, año de ingreso, cantidad de créditos, vivienda; además de entrevistas y lecturas de temáticas asociadas.

Objetivo específico 2: Determinar la viabilidad de procedimientos asociados a estados anímicos, así como los participantes, el acceso a ellos, barreras de participación y disponibilidad de asignación al estudio de manera aleatoria.

- Análisis de los distintos grupos de estudiantes dentro de la facultad, así como el interés por temas asociados a estados anímicos, métodos de acceso a ellos y su comportamiento respectivo al uso de aplicaciones.
- Revisión de literatura asociada a aplicaciones móviles de seguimiento de estados anímicos y salud mental en estudiantes universitarios.

Objetivo específico 3: Recopilar información respectiva a la situación emocional, contextual/situacional y carga académica del estudiante por medio del desarrollo de una aplicación móvil como herramienta de monitoreo y evaluación remota.

- Recopilación de datos auto reportados a través del desarrollo de una aplicación móvil: cuestionarios de caracterización, higiene del sueño, depresión, ansiedad, estrés y contexto.
- Análisis preliminar de los datos recopilados por medio de gráficos y matrices a modo de evaluar el comportamiento y las tendencias de los mismos.

Objetivo específico 4: Generar un modelo de clasificación y definir la incidencia de las variables relevantes a partir de la determinación de estas.

- Desarrollo de modelos matemáticos para depresión, ansiedad y estrés, contemplando variables de contexto situacional, respuestas diarias de cuestionarios e información académica (créditos y percepción de carga académica).
- Contraste del resultado de los modelos con las categorizaciones proporcionadas por la aplicación del cuestionario DASS-21.

Objetivo específico 5: Determinar barreras de acceso, adhesión de los usuarios, idoneidad del grupo objetivo y riesgos de contaminación de los datos y modelos.

- Análisis y discusión respecto a los resultados obtenidos y su comparación con estudios asociados al tema.
- Obtención de métricas respectivas a la participación, uso de la aplicación y levantamiento de dificultades y riesgos potenciales.

Capítulo 5

Recolección de datos

La recolección de datos se realizó a través de la aplicación Ánimi diseñada exclusivamente para este trabajo de memoria, la cual se compone de una sección que consta de cuestionarios diarios respectivos a temas de ansiedad, depresión y estrés, y otra sección que por problemas técnicos de programación queda disponible para trabajos futuros, pudiendo así recopilar información no solo a través de cuestionarios, sino también a través de un conjunto de sensores de los dispositivos móviles.

5.1. Cuestionarios

En esta sección se presentan los cuestionarios a aplicar una o dos veces al día durante el período del estudio, con sus respectivas alternativas de respuesta:

5.1.1. Caracterización general

Las preguntas mostradas en la Tabla 5.1 buscan obtener una caracterización general del estudiante, abarcando edad, género, lugar de procedencia, situación académica y bienestar emocional. Ahora bien, las interrogantes 6, 7, 8 y 9 apuntan a caracterizar académicamente al estudiante al consultar por su carrera, semestre en curso, créditos y percepción de carga académica asociada.

Finalmente, semanas previas a la ejecución de toma de datos para la memoria de título el pueblo chileno comenzó a protestar contra los constantes abusos y violaciones de derechos que se han ejecutado contra la sociedad por parte de privados y del Gobierno, abarcando temas como salud, educación, transporte, sueldo mínimo, pensiones, privatización de servicios básicos, entre otros [58]. Esta situación puede sesgar los resultados obtenidos debido a que es de alta tensión y esfuerzo tanto físico como mental para todos los involucrados. Frente a ello, se agregan las preguntas mostradas en la Tabla 5.2, las cuales permitirán determinar el impacto de la situación y la influencia sobre el comportamiento de la persona en una situación cotidiana.

Tabla 5.1: Preguntas de Contexto General

Pregunta - Alternativas
1. Edad Ingresar número
2. Género F - M - No binario - Prefiero no decirlo
3. Región de procedencia Listado de regiones chilenas
4. ¿Con quién vives en período de clases? Sol@ - familiares – Pareja - Compañeros/as de departamento/casa – Amigos/as
5. Año de ingreso Listado de años desde el 2014 al 2019
6. Carrera Listado de carreras
7. ¿De qué semestre son la mayoría de tus ramos? Listado desde semestre I a semestre XII+
8. ¿Cuántos créditos tienes inscritos este semestre? Ingresar número
9. Según tu percepción, ¿cómo evalúas la carga académica que tienes este semestre? baja - normal - alta

Tabla 5.2: Preguntas de Realidad Nacional

Pregunta - Alternativas
1. Durante las últimas dos semanas, ¿cuánto crees que han afectado las movilizaciones en tu día a día? Nada - Poco - Algo - Bastante - Mucho
2. Durante las últimas dos semanas, ¿cuánto crees que ha afectado la situación actual del país en tu estado de ánimo? Nada - Poco - Algo - Bastante - Mucho

Además, se presentan las preguntas mostradas en la Tabla 5.3 de propósito de vida (pregunta 1) y apreciación personal (pregunta 2) dado que se ha demostrado que se relacionan con la sintomatología estudiada[30]. Específicamente, indican que cuando se tiene un propósito se experimentan menores niveles de depresión, ansiedad y estrés [29]. Cabe destacar que se utilizaron las preguntas 8 y 23 de la Escala de Bienestar RyFF [64], respectivamente.

Tabla 5.3: Preguntas de Percepción Personal

Pregunta - Alternativas
10. Tengo claro donde va mi vida muy en desacuerdo - en desacuerdo – indiferente – de acuerdo – muy de acuerdo
11. Según tu apreciación y modo general, ¿cómo evalúas tu estado anímico? muy malo - malo – indiferente – bueno – muy bueno

5.1.2. Higiene de sueño

Según estudios [31] [49] donde se analizaron artículos relacionados a la higiene del sueño y su efecto sobre depresión y desempeño, existe una correlación significativa que vuelve a esta una variable importante a estudiar. Sin embargo, no existe claridad respecto a si una deficiente higiene del sueño provoca las sintomatologías nombradas, o al revés.

En la Tabla 5.4 las primeras dos preguntas permiten calcular las horas efectivas de sueño, mientras que la tercera pregunta evalúa la calidad de sueño percibida.

Tabla 5.4: Preguntas de Higiene del Sueño

Pregunta - Alternativas
1. ¿A qué hora te dormiste anoche? Listado de horas, opciones cada 30 min.
2. ¿A qué hora te despertaste por la mañana? Listado de horas, opciones cada 30 min.
3. En una escala de 1 a 7, como las notas de la universidad ¿Qué tan reparador fue tu sueño de anoche? Seleccionar nota

5.1.3. Trayectoria del estado anímico

Estas preguntas buscan identificar la sintomatología de las patologías estudiadas: depresión, ansiedad y estrés. Para ello se utiliza el Cuestionario de Salud del Paciente PHQ-4 que permite evaluar depresión (preguntas 1 y 2) y ansiedad (preguntas 3 y 4), y la Escala de Estrés Percibido EEP-4 (PSS-4 en inglés), para evaluar estrés.

Ambos son versiones simplificadas de los cuestionarios PHQ y EEP explicados en la sección 3.1.2, donde luego de un testeo con estudiantes se requirió realizar una adaptación de las respuestas de ambos cuestionarios. Esto porque referenciaban a la frecuencia de las ocurrencias en las últimas dos semanas y en este caso se requiere la asertividad de la afirmación en el momento.

En las Tablas 5.5 y 5.6 se encuentran las respectivas preguntas y respuestas posibles.

Tabla 5.5: Preguntas de Depresión y Ansiedad

Pregunta - Alternativas
1. ¿Tienes poco interés o placer al hacer las cosas? 0: Muy en desacuerdo 1: En desacuerdo 2: De acuerdo 3: Muy de acuerdo
2. ¿Te sientes bajoneado/a, deprimido/a, irritable o desesperanzado/a? 0: Muy en desacuerdo 1: En desacuerdo 2: De acuerdo 3: Muy de acuerdo
3. ¿Te sientes muy nervioso/a, angustiado/a o con los nervios de punta? 0: Muy en desacuerdo 1: En desacuerdo 2: De acuerdo 3: Muy de acuerdo
4. ¿No has podido dejar de preocuparte o controlar la preocupación? 0: Muy en desacuerdo 1: En desacuerdo 2: De acuerdo 3: Muy de acuerdo

Tabla 5.6: Preguntas de Estrés

Pregunta - Alternativas
1. ¿Te sientes incapaz de controlar las cosas importantes de tu vida? 0: Muy en desacuerdo 1: En desacuerdo 2: Ni de acuerdo ni en desacuerdo 3: De acuerdo 4: Muy de acuerdo
2. ¿Estás seguro/a sobre tu capacidad de manejar tus problemas personales? 0: Muy en desacuerdo 1: En desacuerdo 2: Ni de acuerdo ni en desacuerdo 3: De acuerdo 4: Muy de acuerdo
3. ¿Sientes que las cosas están resultando como tú quieres? 0: Muy en desacuerdo 1: En desacuerdo 2: Ni de acuerdo ni en desacuerdo 3: De acuerdo 4: Muy de acuerdo
4. ¿Sientes que no puedes resolver todas las situaciones que tienes que enfrentar? 0: Muy en desacuerdo 1: En desacuerdo 2: Ni de acuerdo ni en desacuerdo 3: De acuerdo 4: Muy de acuerdo

Para el caso de estrés la puntuación de las respuestas es contraria para las preguntas positivas (2 y 3), donde a la respuesta “Muy en desacuerdo” se le asigna un puntaje de 4, mientras que a “Muy de acuerdo” se le asigna un puntaje de 0. En las preguntas negativas (1 y 4) la puntuación es normal.

5.1.4. Contexto situacional

Se agregan las preguntas mostradas en la Tabla 5.7 a modo de identificar gatillantes que puedan afectar a la sintomatología de las variables estudiadas.

Principalmente, se busca identificar la actividad que se encontraba realizando la persona al momento de responder (pregunta 1), si es que se encontraba acompañado/a socialmente (pregunta 2) y cuál es su percepción de su día en general (pregunta 3). En sí, estas preguntas están relacionadas con las mostradas en la Tabla 5.3 y pueden sugerir patrones de comportamiento una vez se analicen los datos.

Tabla 5.7: Preguntas de Contexto Situacional

Pregunta - Alternativas
1. ¿Qué estabas haciendo antes de contestar estas preguntas? a. Trabajando b. Estudiando c. Descansando/Ocio d. Desplazándote de un punto a otro (en trayecto) e. Otro (especificar)
2. ¿Con quién estabas antes de contestar estas preguntas? a. Solo/a b. Compañeros/as de universidad c. Compañeros/as de trabajo d. Amigos/as e. Pareja amorosa (pololo/polola, andante, etc.) f. Desconocidos g. Familia h. Otro
3. En una escala de 1 a 7, como las notas de la universidad ¿Con qué nota calificarías todo este día? Seleccionar nota.

5.1.5. Preguntas finales

En la Tabla 5.8 se indican las 21 preguntas del Cuestionario de Depresión, Ansiedad y Estrés DASS-21, el cual, como se presentó anteriormente, permite realizar una clasificación auto reportada de la situación de la persona sobre la base del estado anímico durante la semana previa a la ejecución.

Cabe destacar que de estas preguntas 7 corresponden a depresión, 7 a ansiedad y 7 a estrés. De aquí, bajo condiciones normales (llámese cotidianas y sin situaciones que escapan de control) los puntajes de corte son mayor o igual a 6 para depresión y estrés, mientras que para ansiedad es mayor o igual a 5. Sin embargo, debido a la situación excepcional del país en la cual el estrés y la ansiedad se han visto incrementados, bajo recomendación y respaldo de los profesionales correspondientes, se considera pertinente ajustar a los siguientes puntos de corte para identificar a aquellos individuos que, efectivamente, padecen una patología duradera y esta no está siendo causada por los acontecimientos sociales.

En consecuencia, los nuevos puntajes de corte y las preguntas asociadas son las siguientes:

- Depresión: $P5 + P10 + P21 + P17 + P16 + P3 + P13 \geq 11$

- Ansiedad: $P19 + P4 + P20 + P7 + P9 + P15 + P2 \geq 8$
- Estrés: $P18 + P6 + P11 + P1 + P12 + P8 + P14 \geq 13$

Tabla 5.8: Preguntas Finales - DASS21

Preguntas DASS-21
1. Me costó mucho calmarme.
2. Me di cuenta que tenía la boca seca.
3. No podía sentir nada positivo.
4. Tuve problemas para respirar (ej. respirar muy rápido, o perder el aliento sin haber hecho un esfuerzo físico).
5. Se me hizo difícil motivarme para hacer cosas.
6. Me descontrolé en ciertas situaciones.
7. Me sentí tembloroso/a (ej. manos temblorosas).
8. Sentí que estaba muy nervioso/a.
9. Estuve preocupado/a por situaciones en las cuales podría sentir pánico y hacer el ridículo.
10. Sentí que no tenía nada por qué vivir.
11. Noté que me estaba poniendo intranquilo/a.
12. Se me hizo difícil relajarme.
13. Me sentí triste y deprimido/a.
14. No soporté que algo me apartara de lo que estaba haciendo.
15. Estuve a punto de tener un ataque de pánico.
16. Fui incapaz de entusiasmarme con algo.
17. Sentí que valía muy poco como persona.
18. Sentí que estaba muy irritable.
19. Sentí los latidos de mi corazón a pesar de que no había hecho ningún esfuerzo físico.
20. Sentí miedo sin saber por qué.
21. Sentí que la vida no tenía ningún sentido.

Las respuestas posibles para cada caso son:

0: No describe **nada** de lo que me pasó o sentí en la semana.

1: Sí, esto me pasó o lo sentí en **alguna medida o en algún momento**.

2: Sí, esto me pasó **bastante** o lo sentí muchas veces.

3: Si, esto me pasó **mucho, o casi siempre**.

Finalmente, en la Tabla 5.9 se agregan preguntas para contrastar resultados. La primera busca identificar el cambio en la percepción de la carga académica del estudiante conforme avanza el tiempo, y la segunda, analiza tanto la toma de consciencia, como la identificación respecto al estado anímico propio.

Tabla 5.9: Preguntas Finales - Comparativas

Preguntas Finales
1. Según tu percepción, ¿cómo evalúas la carga académica que tienes este semestre? baja - adecuada - alta
2. Según tu apreciación y modo general, ¿cómo evalúas tu estado anímico? muy malo - malo - indiferente - bueno - muy bueno

5.2. Aplicación móvil

5.2.1. Descripción general

Ánimi es la aplicación de salud mental creada para levantar datos primarios de los estudiantes pertenecientes al grupo de estudio de este trabajo. Su funcionamiento está representado en la Figura siguiente, donde se comienza con el uso del celular a través de un inicio de sesión verificado a través de Firebase. Aquí también se realiza la primera consulta de la aplicación para mostrar el mensaje de inicio a los usuarios, las preguntas y la cantidad de ejecuciones.

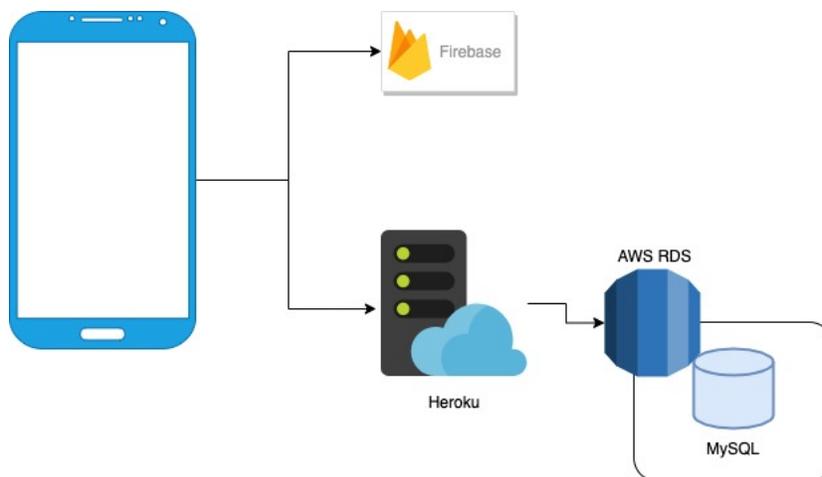


Figura 5.1: Funcionamiento general aplicación Ánimi

De manera paralela, se interactúa con la base de datos a través de Heroku Web Service, donde se encuentran las API con el cuestionario inicial, los cuestionarios diarios y la captura de datos a través de sensores del teléfono, actualizando, además, la cantidad de respuestas faltantes para cada participante.

Una vez que el contador de cada teléfono llega a su fin se activa un temporizador de una semana para el lanzamiento del cuestionario final, el cual se encuentra en una página web que muestra una serie de preguntas y recupera tanto el usuario como las respuestas. La aplicación envía una notificación a Heroku (que a su vez realiza consultas a la base de datos) para así enviar un correo a los participantes con el enlace de respuesta y, luego de ejecutar los algoritmos correspondientes, mostrar una retroalimentación personalizada automática.

Con respecto a los datos pasivos, los sensores recopilan información cada 15 minutos y la guardan en un arreglo. Finalmente, cuando el usuario termina de responder un cuestionario se envía tanto el conjunto de datos pasivos como las respuestas obtenidas a Heroku y se guardan en la base de datos “Ánimo” que se encuentra en Amazon Web Service.

5.2.2. Funcionalidades

La aplicación se ejecuta dos veces al día: la primera a las 10:00 ya que el grueso de los estudiantes comienza su jornada escolar entre las 8:30 y 10:15, presentándose 15 minutos de descanso entre cambio de clase a las 10:00. La segunda ejecución se realiza a las 21:00, ya que las clases terminan a las 17:45, 19:30 o 21:00, por lo que en dicho momento se espera que para todos y todas las estudiantes se haya acabado el día.

Dentro de las funcionalidades se contempla:

- Consentimiento informado, el cual explica el motivo, duración, uso de los datos y responsables del estudio.
- Los cuestionarios detallados en la sección anterior que permiten tomar datos de manera periódica sobre una serie de preguntas relacionadas a estrés, ansiedad, depresión, higiene del sueño y situaciones de contexto, las que requerirán respuesta 42 veces (2 veces al día, por 21 días) en escala Likert o en preguntas de selección, según corresponda.
- Sistema de recordatorio de llenado del formulario y contador de ejecuciones restantes.
- Envío de correo con acceso al formulario DASS-21 que permite obtener una clasificación preliminar auto reportada y válida del estado anímico del individuo.
- Envío de correo con acceso al cuestionario final que contempla dos preguntas para contrastar percepciones al inicio y final del estudio.
- Toma de datos de manera pasiva durante el período ejecución, es decir, utilizando los sensores del dispositivo móvil, siendo estos:
 - Geolocalización: Latitud, longitud y altura.
 - Luminosidad: Valor de luminosidad de la pantalla.

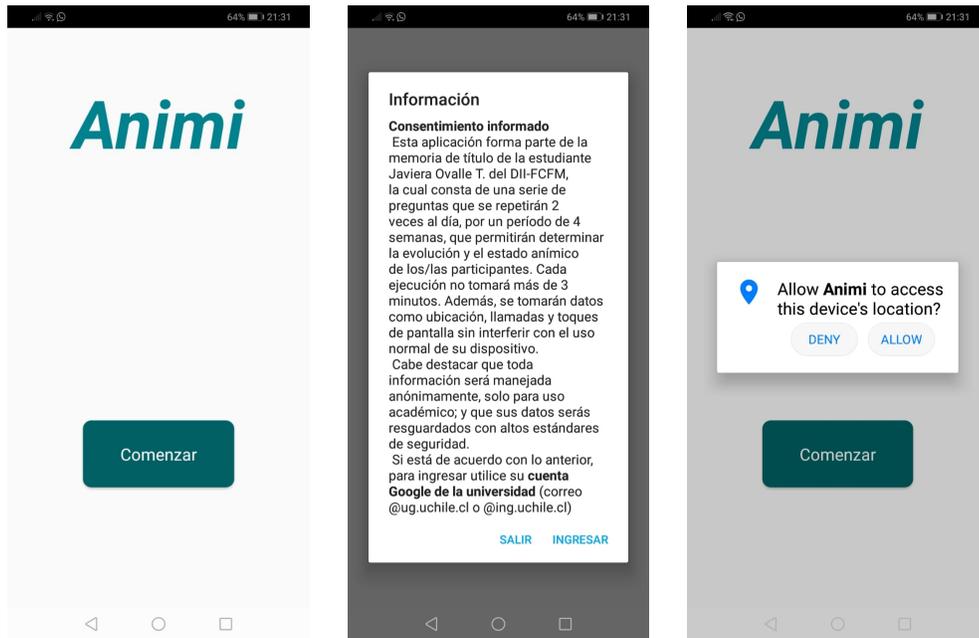
- Proximidad: 1 o 0 según el sensor del teléfono.
- Giroscopio: Angulares de cada eje y posiciones de cada eje (X,Y,Z).
- Acelerómetro: Angulares de cada eje (X,Y,Z).
- Temperatura del dispositivo: En caso de que sea soportado por el mismo. Si no, toma valor -1.
- Temperatura ambiente: En caso de que sea soportado por el mismo. Si no, toma valor -1.
- Presión Atmosférica: En caso de que sea soportado por el mismo. Si no, toma valor -1.
- Registro de llamadas: Inicio de llamada, tipo, duración y número telefónico encriptado.
- Decibeles: Registrados a través del micrófono.

Cabe destacar que si bien la aplicación está programada para recopilar todo el listado de datos pasivos, el acceso a ellos dependerá de la versión de cada dispositivo y del permiso que otorgue el usuario.

5.2.3. Ilustraciones de la aplicación móvil

La imagen mostrada en la Figura 5.2a corresponde a la primera vista que tendrá el o la estudiante al descargar la aplicación. Luego, como se muestra en la Figura 5.2b, se presenta el consentimiento informado y la solicitud de acceso a los sensores del dispositivo móvil en la Figura 5.2c .

En caso de aceptar el consentimiento informado, se abre una sección para ingresar con cuenta de la Universidad de Chile, la cual funciona como filtro para los estudiantes dentro de la tienda de aplicaciones de Android. Si el usuario rechaza el consentimiento, se cierra la aplicación.



(a) Vista inicial.

(b) Vista consentimiento informado.

(c) Vista solicitud uso de geolocalizador.

Figura 5.2: Vista inicial

Posteriormente, tal como se puede ver en la Figura 5.3, al momento de la creación de la cuenta se activa el cuestionario inicial que se divide en seis pasos: edad - género, región de procedencia - con quién vives, año de ingreso - carrera - semestre en curso, créditos actuales - percepción de carga académica, claridad de vida - evaluación de estado anímico y finalmente, preguntas sobre los efectos de la realidad nacional en el estado anímico de la persona.

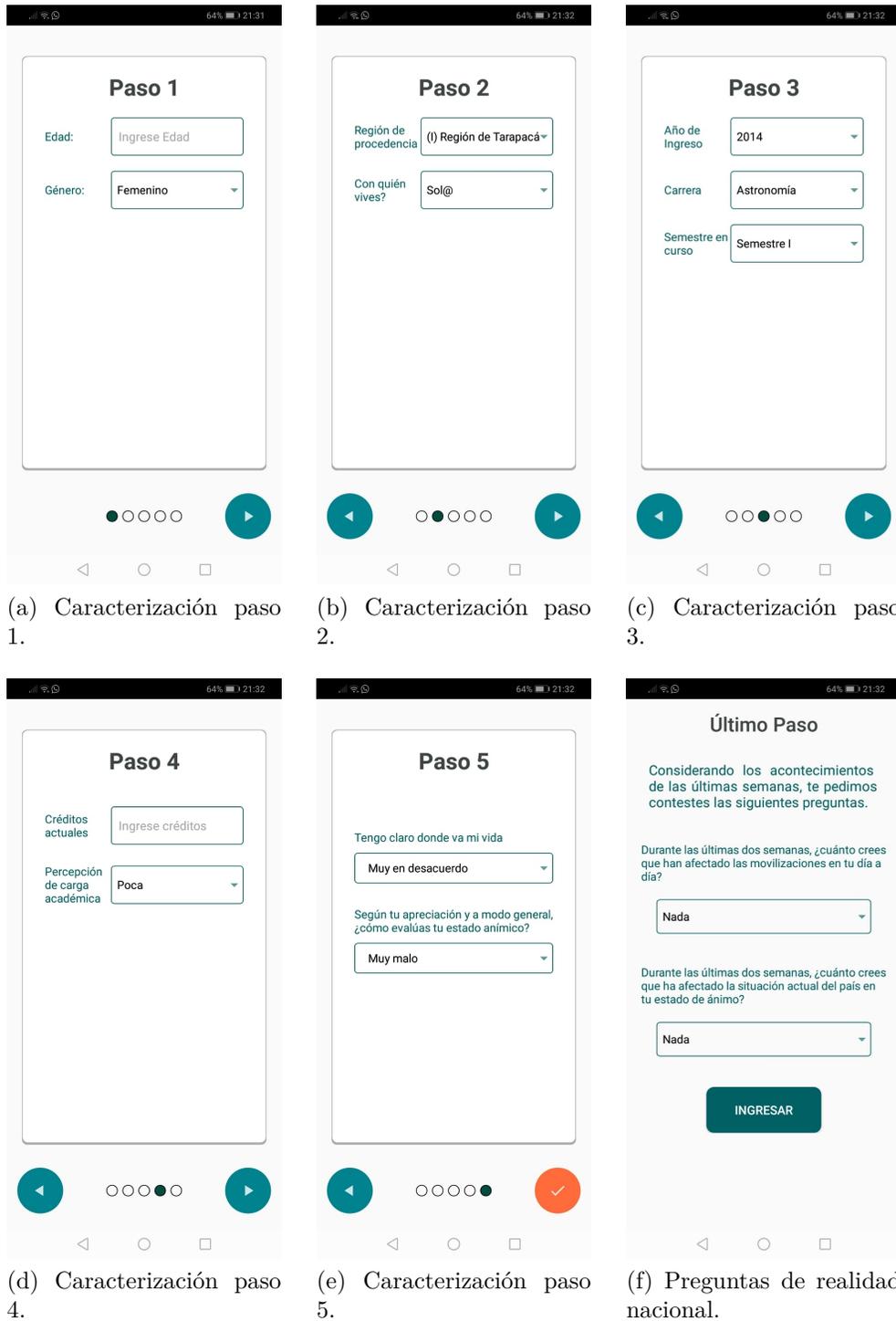
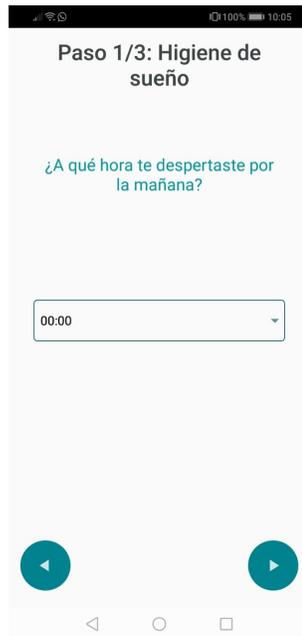


Figura 5.3: Caracterización

Luego, se activa una notificación a las 10:00 del día siguiente al registro, momento en que se solicitará responder el primer cuestionario que abarca higiene del sueño, estado anímico y contexto situacional, tal como se muestra en las Figuras 5.4 a 5.6:



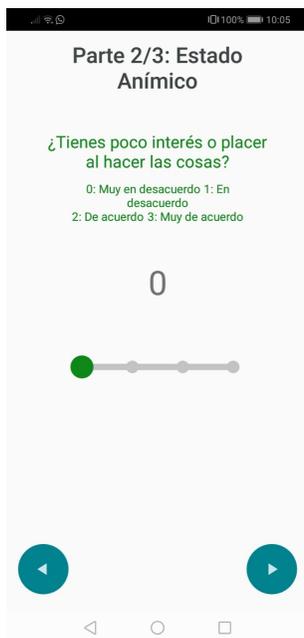
(a) Hora de dormir.



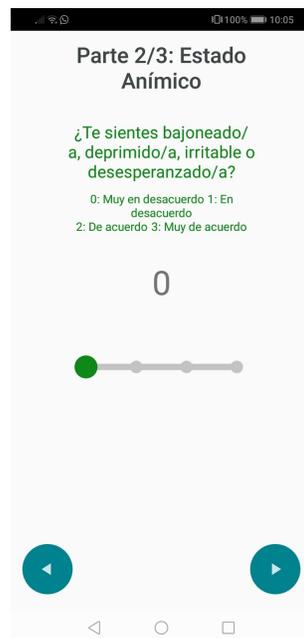
(b) Hora de levantarse.



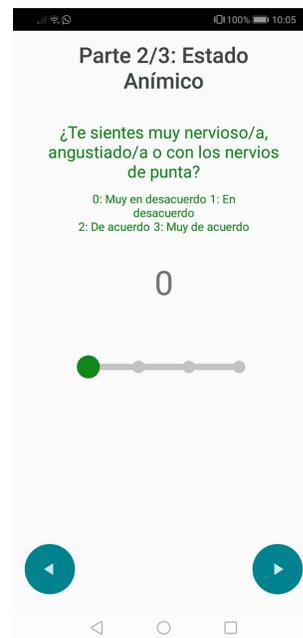
(c) Calificación del dormir.



(d) Depresión 1.



(e) Depresión 2.



(f) Ansiedad 1.

Figura 5.4: Cuestionarios - parte 1

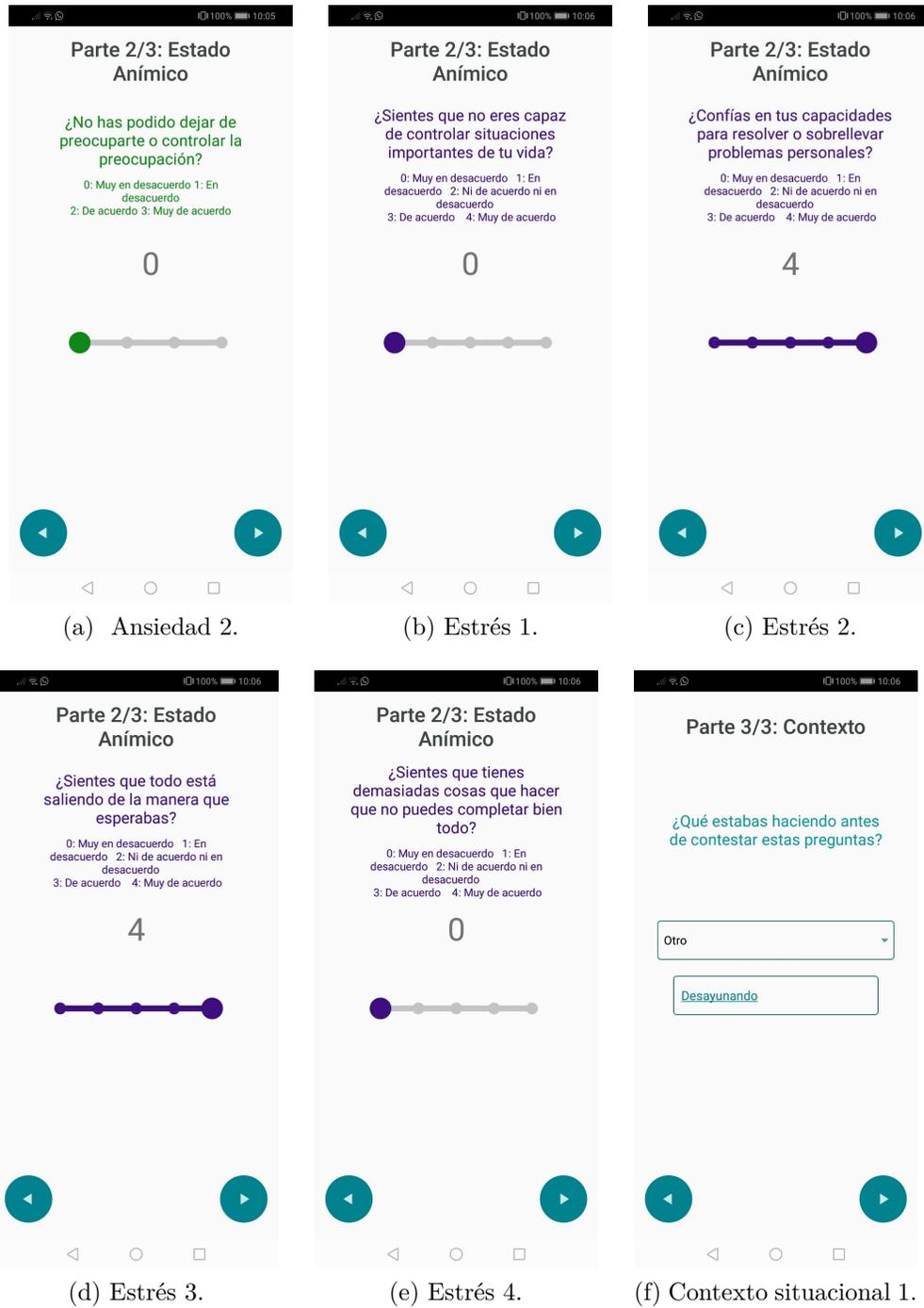


Figura 5.5: Cuestionarios - parte 2

Cabe destacar que durante la mañana se realizan las preguntas de higiene del sueño. No así durante la noche, caso en que se agrega una pregunta asociada a la calificación general del día según la percepción del usuario:



Figura 5.6: Cuestionarios - parte 3

Posteriormente, luego de una semana de terminados los cuestionarios diarios, la aplicación funciona como gatillante del envío de un correo con el enlace para responder el cuestionario final y DASS-21 a cada usuario. De aquí, una vez que este ingresa, las respuestas quedan asociadas a su correo y son enviadas a la base de datos. El cuestionario se ve como sigue en la Figura 5.7:

Cuestionario Final

A continuación, verás 2 preguntas sobre tu percepción de carga académica y del efecto de este proyecto en ti.

Según tu percepción, ¿Cómo evalúas la carga académica que tienes este semestre?

Baja Adecuada Alta

Según tu apreciación y a modo general, ¿cómo evalúas tu estado anímico?

Muy Malo Malo Indiferente Bueno Muy Bueno

A continuación, verás una serie de preguntas asociadas a tu **estado de ánimo**, donde deberás responder qué tan de acuerdo estás con las afirmaciones de lo ocurrido **durante la última semana**.

(0) Nada
(1) En alguna medida
(2) Muchas Veces
(3) Casi siempre

1. Me costó mucho calmarme.

0 1 2 3

2. Me di cuenta que tenía la boca seca.

Figura 5.7: Cuestionario final y fragmento de DASS-21

Es pertinente nombrar que cada respuesta seleccionada se guarda de manera numérica en escala de 0 a 3 en el caso de estar planteada positivamente y en escala de 3 a 0 en caso contrario.

Finalmente, en caso de que se cumplan los puntos de corte para depresión, ansiedad o estrés, la retroalimentación automática diseñada y validada por el Psicólogo Dr. Álvaro Jiménez y la Psicoterapeuta Dr. Vania Martínez es la que sigue:

¡Hola! Muchas gracias por haber participado del estudio Ánima. De acuerdo con el cuestionario que respondiste, existe la probabilidad de que estés presentando algunas dificultades emocionales. Un cuestionario no puede establecer un diagnóstico, por lo que te recomendamos que consultes a un servicio de atención en salud mental. Si tienes alguna duda, puedes escribirnos al correo saludmental@imhay.org

El Núcleo Milenio para mejorar la salud mental de adolescentes y jóvenes (Imhay) te invita a seguir los siguientes consejos:

+ Consejos para el manejo del estrés: <https://www.instagram.com/p/B5lBa6kHTut/>

+ Consejos para cuidar el ánimo: <https://vimeo.com/177616822>

Equipo Ánima

(a) Sobre puntos de corte.

¡Hola! Muchas gracias por haber participado del estudio Ánima. De acuerdo con el cuestionario que respondiste, tu salud mental se encuentra en buen estado.

Sin embargo, es importante realizar cotidianamente algunas actividades para prevenir la emergencia de dificultades emocionales en el futuro. El Núcleo Milenio para mejorar la salud mental de adolescentes y jóvenes (Imhay) te invita a seguir los siguientes consejos:

+ Consejos para el manejo del estrés: <https://www.instagram.com/p/B5lBa6kHTut/>

+ Consejos para cuidar el ánimo: <https://vimeo.com/177616822>

Si tienes alguna duda, puedes escribirnos al correo saludmental@imhay.org

Equipo Ánima

(b) Bajo puntos de corte.

Figura 5.8: Mensajes según puntaje de corte

Finalmente, tal como se muestra en la Figura 5.9, el usuario ve el mensaje inmediatamente después de enviar sus respuestas:



Figura 5.9: Retroalimentación automática DASS-21

5.2.4. Manejo de la aplicación móvil

Como se explicó en los párrafos anteriores, la aplicación Ánimi fue diseñada e implementada para la recolección de datos primarios de los participantes de este estudio. Para ello, se creó un repositorio de datos en MySQL que se encuentra alojado en la nube de Amazon Web Service.

En la Figura 5.10 es posible ver el modelo de datos relacional asociado a la aplicación móvil con entidades para cada pregunta de caracterización e identificación del usuario - dispositivo móvil y para los registros de datos pasivos (ubicación, desbloques de pantalla, registro de llamadas, acelerómetro, entre otras).

De aquí, tres entidades están relacionadas directamente a la información de usuarios y de los cuestionarios de auto reporte que responden diariamente:

- AD_USUARIOS: Esta tabla contiene tanto el número de usuario, como la conexión a Google ID, fecha de ingreso, si está habilitado para responder, y las preguntas del cuestionario inicial.
- P_DIARIAS: La cual recupera el usuario, sesión, tipo de jornada (1 para jornada mañana, 2 para tarde), y las respuestas a las preguntas de los cuestionarios diarios. Si una pregunta no corresponde a una jornada, se le asigna valor -1.

- P_DASS21: Contiene las preguntas asociadas al cuestionario final DASS21 y el usuario que respondió, la sesión respectiva y la fecha de respuesta.

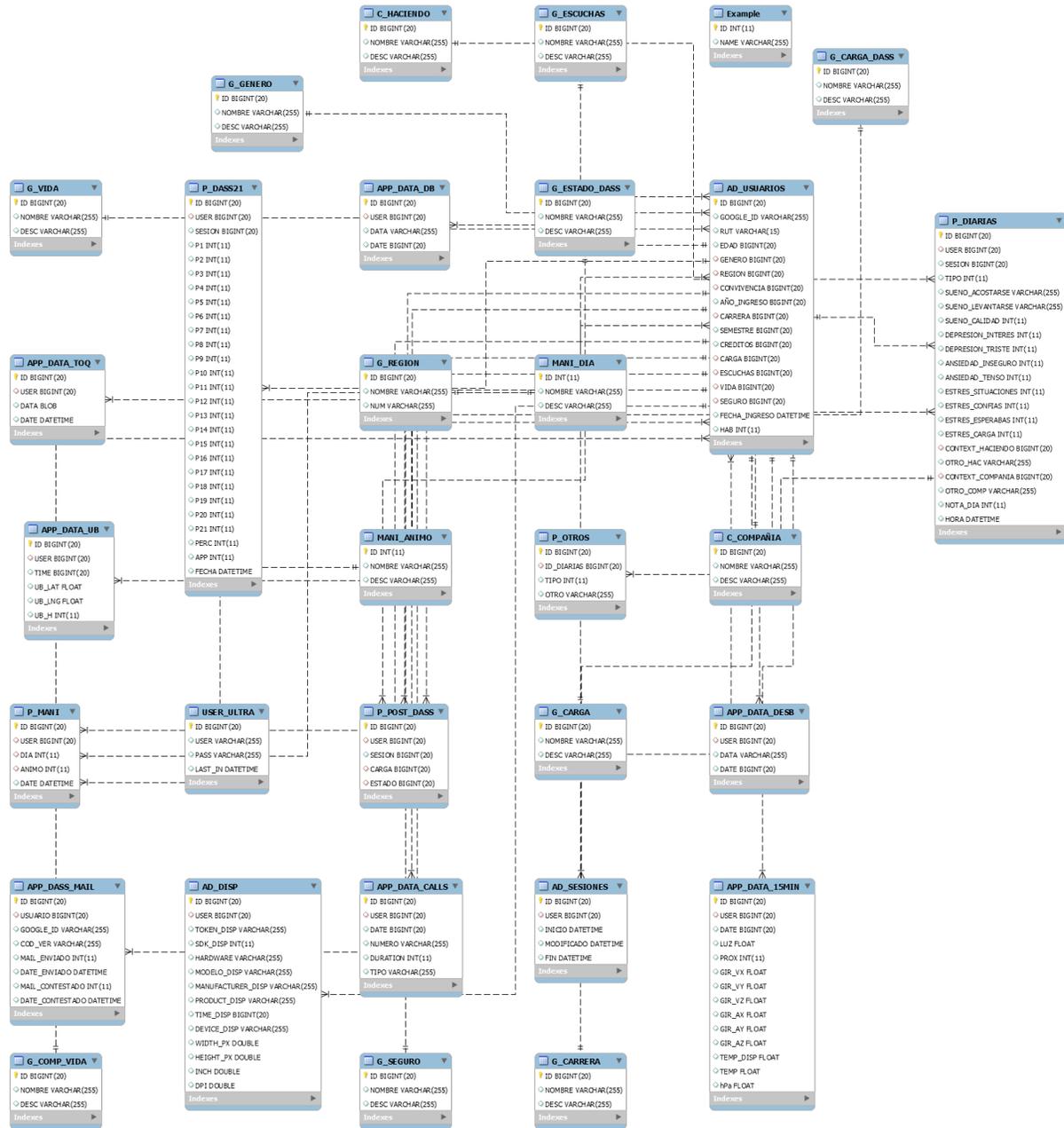


Figura 5.10: Base de datos Animi

Ahora bien, los datos asociados a la base se pueden tratar y/o consultar utilizando diversas herramientas de manejo de bases de datos más el punto de salida que proporciona AWS. A continuación, se muestra una selección de consultas de lectura realizadas a través de MySQL Workbench:

ID	USER	DATA	DATE
3	64	SCREEN_UNLOCK	1585802107791
4	65	SCREEN_UNLOCK	1585831698623
5	66	SCREEN_UNLOCK	1585839007835
6	69	SCREEN_OFF	1585848962492
7	69	SCREEN_ON	1585849145617
8	69	SCREEN_UNLOCK	1585849146814
9	69	SCREEN_UNLOCK	1585849146929
10	69	SCREEN_OFF	1585849698651
11	69	SCREEN_ON	1585849709783

Figura 5.11: Ejemplo tabla desbloques de pantalla

ID	USER	DATE	NUMERO	DURATION	TIPO
5	74	1585856051626	KzU2OTgxODIyNDI3	144	INCOMING
6	84	1586186520435	KzU2OTY2NzgwOTg1	0	MISSED

Figura 5.12: Ejemplo tabla de registro de llamadas

ID	USER	TIME	UB_LAT	UB_LNG	UB_H
222	77	1585858743530	-33.4468	-70.636	661
223	81	1585860003508	-33.4469	-70.6361	623
224	81	1585860063510	-33.4469	-70.6361	627
225	81	1585860159509	-33.4465	-70.6357	562
226	82	1585860402517	-33.447	-70.6361	619
227	82	1585860562084	-33.4468	-70.6359	613
228	82	1585860625012	-33.4468	-70.6359	608
229	82	1585860684023	-33.447	-70.6359	622
230	82	1585860745027	-33.4469	-70.6358	612

Figura 5.13: Ejemplo tabla de respuestas diarias

ID	USER	SESION	TIPO	SUENO_ACOSTARSE	SUENO_LEVANTARSE	SUENO_CALIDAD	DEPRESION_INTERES	DEPRESION_TRISTE	ANSIEDAD_INSEGURO
41	85	85	1	01:30	08:00	7	0	0	0
42	84	84	2			-1	0	1	0
43	86	86	2			-1	0	1	1
44	84	84	1	03:00	09:30	5	1	2	1
45	84	84	2			-1	0	1	0
46	84	84	1	01:30	09:00	5	1	0	0
47	84	84	2			-1	0	1	0
48	84	84	1	01:30	08:00	5	1	0	1

Figura 5.14: Ejemplo tabla de geolocalización

ID	USER	DATE	LUZ	PROX	GIR_VX	GIR_VY	GIR_VZ	GIR_AX	GIR_AY	GIR_AZ	TEMP_DISP	TEMP	hPa
68	77	1585858501148	73	0	-0.00129808	-0.00014508	0.0000820924	35.908	88.9586	89.7966	-1	-1	944.001
69	77	1585858561843	71	0	-0.00041233	0.0000992658	-0.000162255	34.2452	88.9855	90.3581	-1	-1	944
70	77	1585858622499	76	0	0.000328347	-0.0000152716	0.000137448	36.772	88.9778	91.0266	-1	-1	944.003
71	77	1585858683320	74	0	0.00122937	-0.000150807	0.000250073	35.4995	88.9782	90.1083	-1	-1	944.022
72	77	1585858744149	74	0	0.000061091	0.0000553597	-0.000082083	32.9211	88.9707	90.4558	-1	-1	944.004

Figura 5.15: Ejemplo tabla datos pasivos recopilados cada 15 min.

Cabe destacar que la fecha es guardada en formato EPOCH o UNIX, y que las preguntas categóricas tienen respuestas numéricas asociadas, las cuales deben ser cambiadas a los respectivos valores de texto para estudios cualitativos.

Finalmente, desde Firebase es factible modificar la cantidad de ejecuciones, mensajes inicial y final, texto de las preguntas diarias y utilización de los sensores del móvil, tal como se puede ver en las imágenes respectivas:

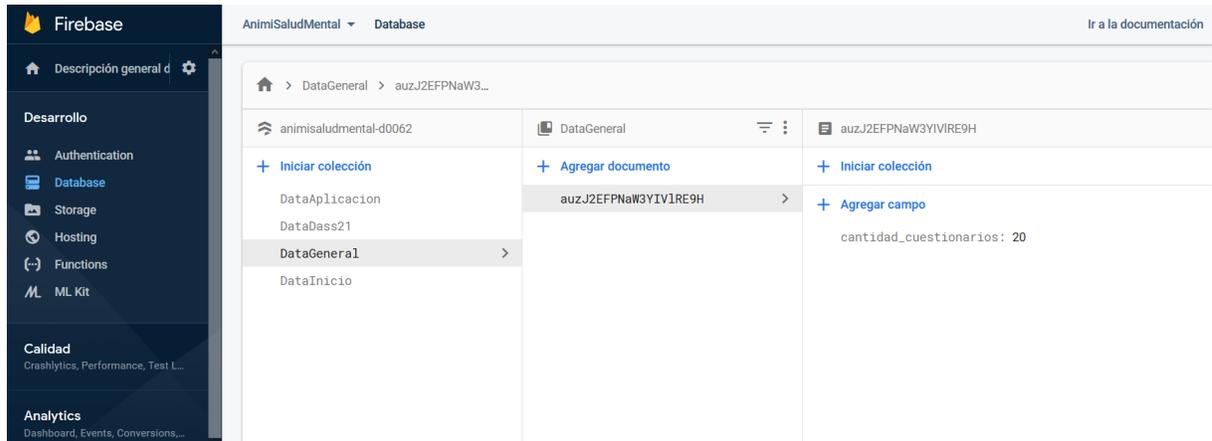


Figura 5.16: Firebase vista número de ejecuciones

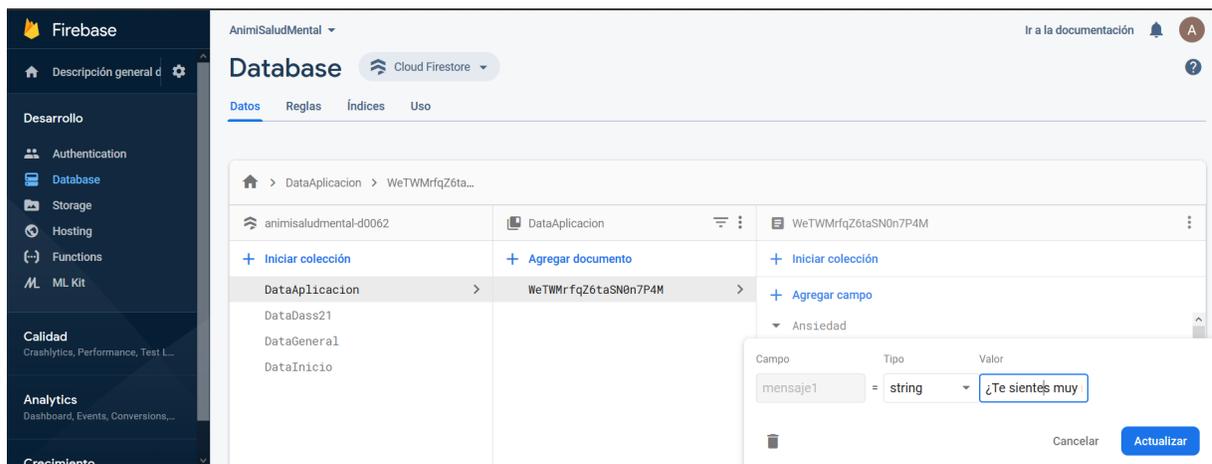


Figura 5.17: Firebase vista mensajes

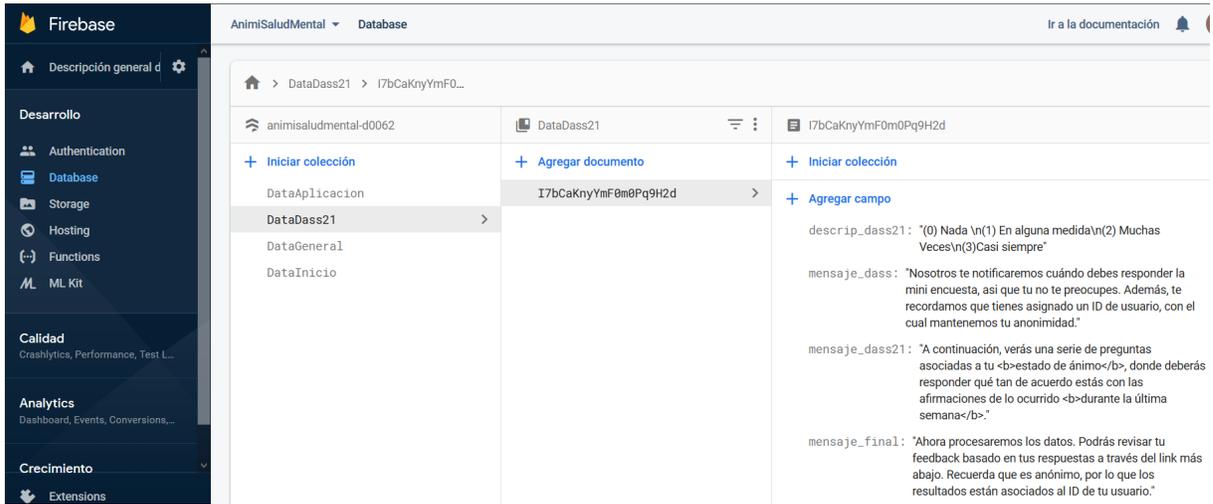


Figura 5.18: Firebase vista preguntas diarias

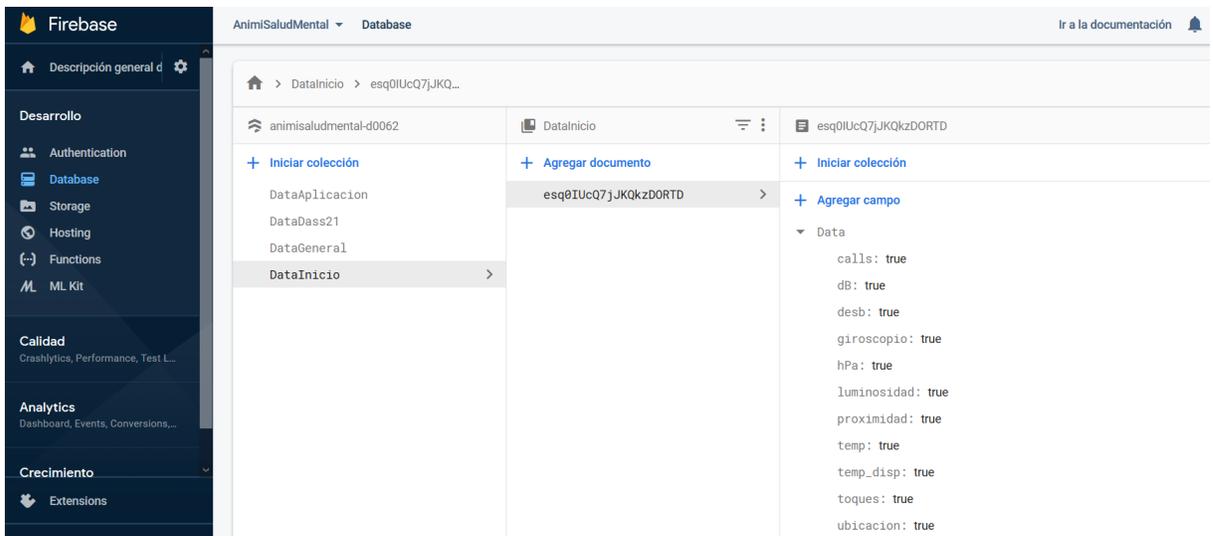


Figura 5.19: Firebase vista uso de sensores

Capítulo 6

Resultados

6.1. Uso de la aplicación móvil Ánimi

La aplicación alcanzó un total de 123 descargas. De estas, 62 personas (50.4%) iniciaron sesión y completaron el cuestionario inicial, mientras que 47 (38.2%) respondieron al menos un cuestionario diario, lo que se traduce en una retención del 76% de los usuarios activos.

Ahora bien, como se muestra en la Figura 6.1, se tiene que 17 (36%) participantes respondieron al menos la mitad de los cuestionarios, destacando una moda de 5, mediana de 14, desviación estándar de 13 y promedio de 16 cuestionarios respondidos. Cabe notar que si bien a cada estudiante se le asignaron 42 ejecuciones a partir del día siguiente de completado el cuestionario inicial, el máximo respondido fueron 40 (2 casos).

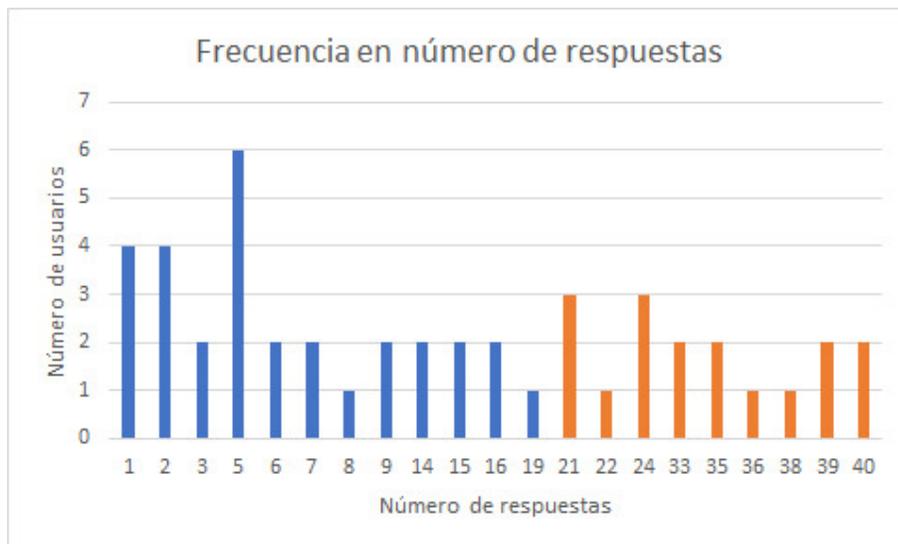


Figura 6.1: Frecuencia de cuestionarios respondidos

6.2. Pre-procesamiento y transformación de datos

Las respuestas se almacenaron en la base de datos mostrada en la Figura 5.10, de donde se obtuvieron cuatro archivos exportables en formato CSV para ser trabajados en Microsoft Excel 365 ProPlus y en el software R Studio, versión 1.2.5019: base de usuarios registrados, base de respuestas diarias, base de respuestas manifestaciones y base de respuestas DASS-21.

Sobre la primera base se excluyeron variables asociadas a la identificación del usuario y la fecha de registro (Figura 6.2a); lo mismo se realizó en la segunda base de datos eliminando, además, el tipo de usuario, la hora del registro y las respuestas asociadas a categorías “otro” en las preguntas de contexto (Figura 6.2b).

Con respecto a la tercera y cuarta base de datos, dado que se enviaron a todos los usuarios registrados, en ambos casos se seleccionaron únicamente a aquellos usuarios que efectivamente respondieron los cuestionarios diarios. En el caso de las respuestas de DASS-21 se agregaron las respuestas iniciales sobre percepción de carga académica y percepción de estado anímico de la base de usuarios para realizar comparaciones.

Finalmente, se unificaron las bases de datos de usuarios y respuestas a modo de obtener una única fuente de información que permita hacer análisis de información cruzada, y se agruparon las variables respectivas a higiene del sueño, depresión, ansiedad y estrés según:

- Higiene del sueño: Variable numérica construida sobre la base de las variables hora de acostarse y hora de levantarse, convirtiendo éstas en texto y luego realizando la resta de los valores, a modo de obtener las horas de sueño utilizando la función `HORA()` de Excel.
- Depresión: Variable binaria construida a partir de las variables Depresión Interés y Depresión Triste según los criterios explicados más adelante.
- Ansiedad: Variable binaria construida a partir de las variables Ansiedad Inseguro y Ansiedad Tenso según los criterios explicados más adelante.
- Estrés: Variable binaria construida a partir de las variables Estrés Situaciones, Estrés Confías, Estrés Esperabas y Estrés Carga según los criterios explicados más adelante.

Con esto, la base de datos consolidada queda tal como se puede ver en la Figura 6.2c:



(a) Base de usuarios.

(b) Base de respuestas.

BASE RESPUESTAS DIARIAS

Edad	Depresion
Genero	Ansiedad
Region	Estres
Convivencia	Contexto_hacienda
Anolngreso	Contexto_compania
Carrera	Nota_dia
Semestre	Estado_Animico
Creditos	Tipo
Carga	Horas_sueno
Vida	Sueno_calidad

(c) Base de datos unificada.

Figura 6.2: Reducción de variables y unificación de bases de datos

6.2.1. Análisis exploratorio de la muestra

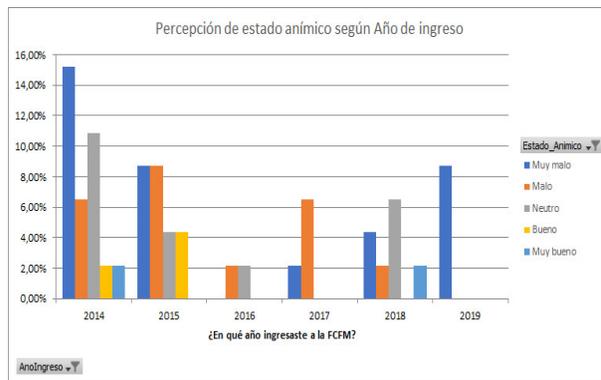
A partir de la muestra de 62 participantes pertenecientes al grupo de estudio se obtuvo una retención del 76 % (47 usuarios) a lo largo de la aplicación de los cuestionarios. De ellos, tal como se muestra en la figura siguiente, se identifica una distribución equitativa en términos de género, donde un 48 % se encuentra entre los 23 y 24 años, 74 % vive en la RM y que el 26 %, 20 % y 13 % cursa su segundo, décimo y onceavo semestre, respectivamente. Además, los participantes presentan una tendencia a tener claro hacia donde va su vida (41 % de acuerdo o muy de acuerdo) y un 78 % vive con familiares durante la etapa universitaria.



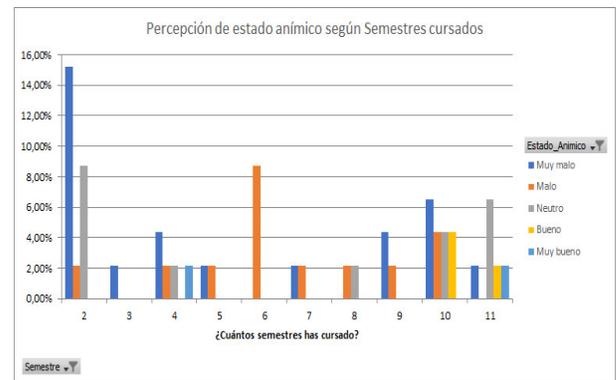
Figura 6.3: Distribución según edad, género, región de procedencia y semestre

Por otro lado, se consultó a los estudiantes por la percepción de carga académica, encontrando que el 60 % considera que tiene una carga académica normal, mientras que un 31 % cree que es alta. Además, al realizar un análisis relacionado al estado anímico se obtuvo que cuando se cuenta con una alta carga académica destaca un mal estado anímico. En caso contrario, destaca un estado anímico neutro o indiferente. Sin embargo, es posible notar diversos estados anímicos cuando la carga es normal, lo que indica que este factor no es determinante con respecto a la variable estudiada (Figura 6.4c).

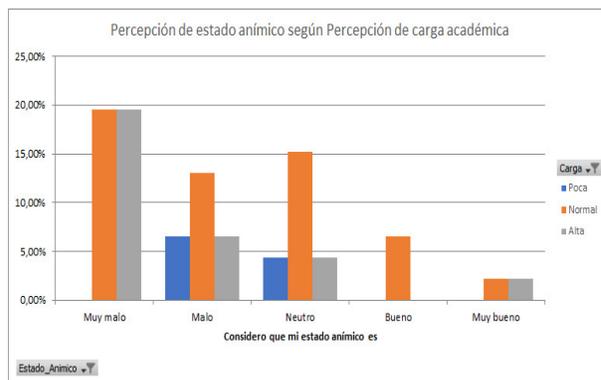
También se realizó un análisis respecto al estado anímico y la claridad en propósito de vida que tiene la persona, donde se obtuvo una predominancia de mal estado anímico cuando no hay claridad en la dirección de vida; no así en el caso contrario, donde predomina un estado anímico neutro o bueno (Figura 6.4f).



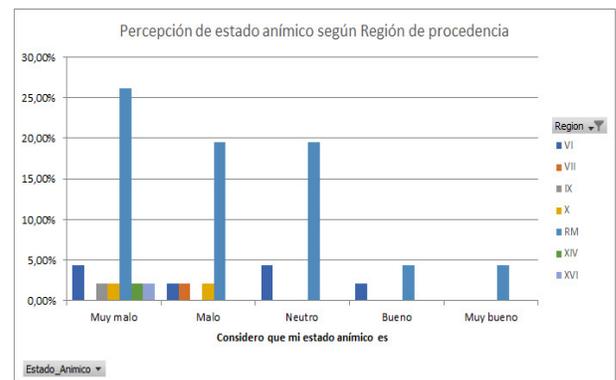
(a) Estado anímico y Año de ingreso



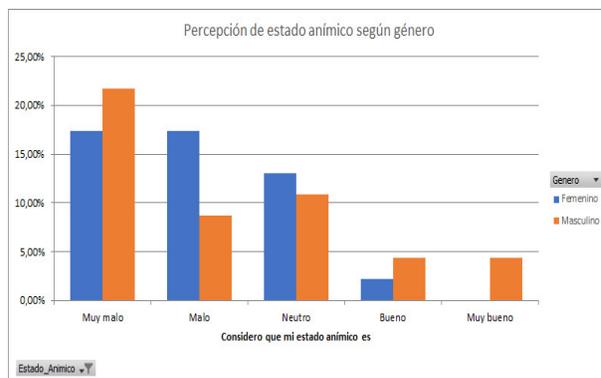
(b) Estado anímico y Semestres cursados



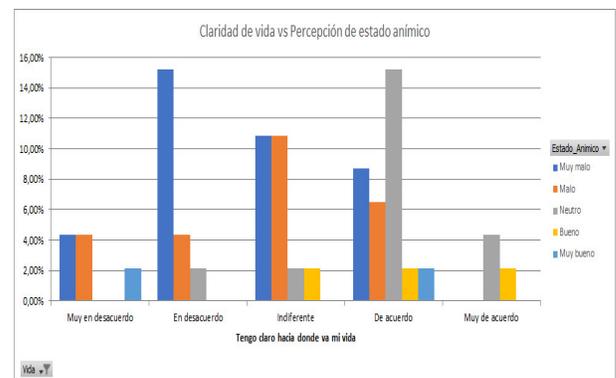
(c) Estado anímico y Percepción de carga académica



(d) Estado anímico y Región de procedencia



(e) Estado anímico y Género



(f) Estado anímico y Claridad de vida

Figura 6.4: Percepción de estado anímico en relación con otras variables

Ahora bien, tanto los estudiantes de primer semestre son aquellos que declaran un muy mal estado anímico. Lo mismo ocurre en los últimos semestres y en los usuarios que provienen de una región distinta a la Región Metropolitana, donde si bien se presentan las demás opciones, predomina un mal sentir (Figura 6.4d).

Al mismo tiempo, como se muestra en la Figura 6.4e, se tiene que quienes peor evalúan su estado anímico son los hombres, mientras que las mujeres indican valorar en su mayoría el estado de ánimo como neutro o malo, destacando este último.

Adicionalmente, en la Figura 6.5 se analizó la base de respuestas asociadas a las problemáticas sociales que aquejan el país, encontrando que los usuarios declaran que las manifestaciones afectaron bastante su día a día (53 %) y que vieron afectado de alguna manera su estado anímico (88 %). De este último punto, el 27 % se vio muy afectado.

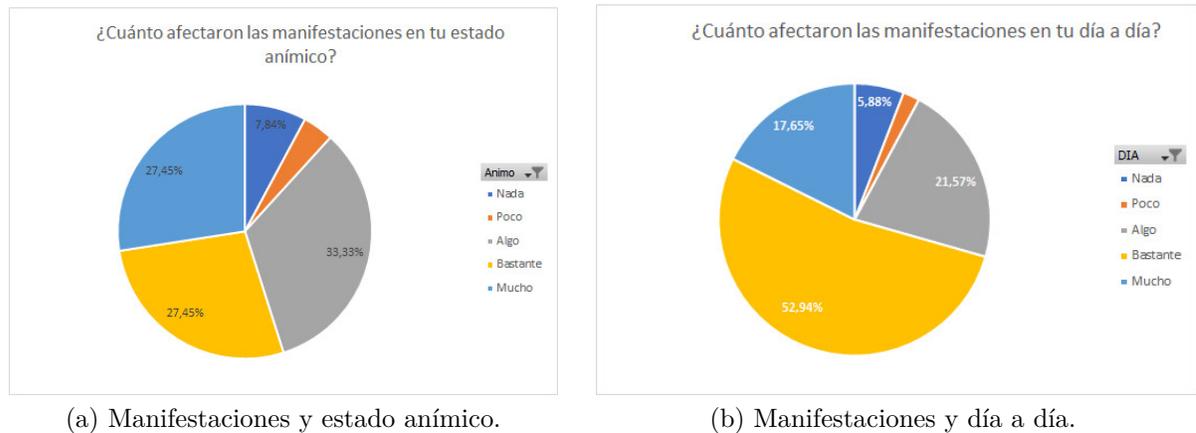


Figura 6.5: Percepción de estado anímico en relación con otras variables

Como se explicó anteriormente, estas preguntas se agregaron para poder identificar qué tan influenciadas pueden estar las respuestas por lo que acontece a nivel nacional, es decir, puede que el porcentaje de la muestra que presenta alguna patología sea alto debido al efecto que declaran las personas y no por las condiciones que se estudian en este trabajo. Este punto se debe tener presente al momento de tomar conclusiones.

6.2.2. Análisis exploratorio de los datos

6.2.2.1. Depresión, Ansiedad y Estrés

A modo de disminuir la cantidad de variables se decidió categorizar las respuestas obtenidas de las preguntas de depresión, ansiedad y estrés en tres variables binarias que indiquen o no presencia de la patología, las cuales, según lo expuesto en la sección 5.1.4 adquieren valor de 1 cuando:

- Depresión: La suma de preguntas 1 y 2 de la Tabla 5.5 es valor mayor o igual a 3.
- Ansiedad: La suma de preguntas 3 y 4 de la Tabla 5.5 es valor mayor o igual a 3.
- Estrés: La suma de preguntas 1 a 4 de la Tabla 5.6 es valor mayor o igual a 8.

De aquí, al realizar un análisis de la Figura 6.6 que muestra la distribución de las respuestas diarias según el momento en el que fueron consultadas, se identificó una distribución equitativa entre la jornada AM (Tipo 1) y PM (Tipo 2) para el caso de ansiedad y estrés. No así en el caso de depresión, donde durante la mañana se presentaron más indicios de la patología que por las tardes, lo que puede justificarse por la relación entre higiene del sueño y la variable.

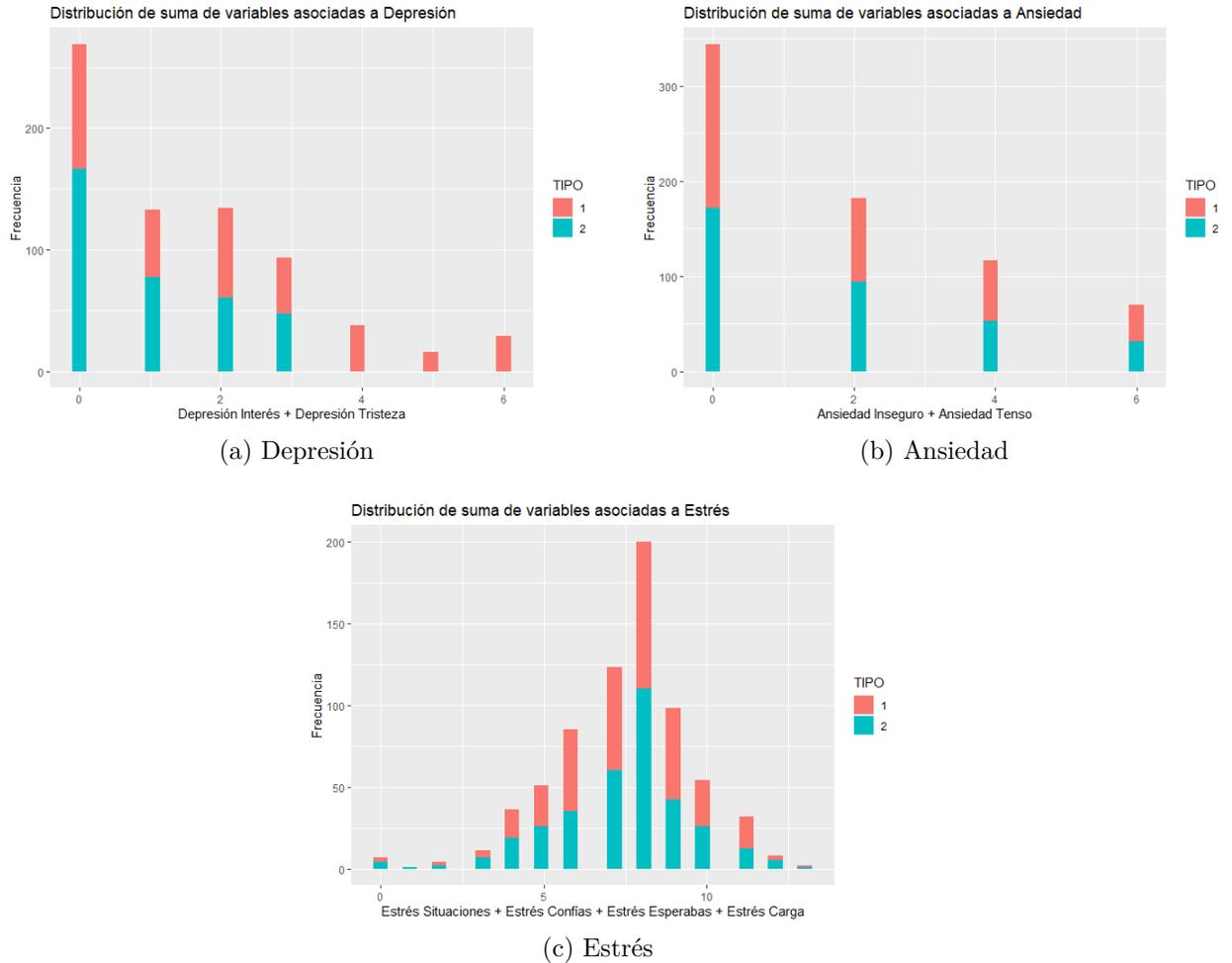


Figura 6.6: Distribución de variables de Depresión, Ansiedad y Estrés

Ahora bien, a partir de la información recopilada, en la Figura 6.7 se puede evidenciar que el 52 % de la muestra presenta ansiedad, mientras que el 55 % padece de estrés y el 24 % de depresión. Sin embargo, hay que considerar que la toma de datos se realizó durante una crisis social y que la muestra declaró que esta le afectó en gran medida en aspectos personales y académicos - situaciones que fomentan el estrés y la ansiedad - por lo que los valores aquí expuestos podrían variar en condiciones normales. Frente a esto, condicional a la recomendación de los especialistas, se propone penalizar los puntos de corte en contextos sociales anormales, a modo de evitar los falsos positivos.

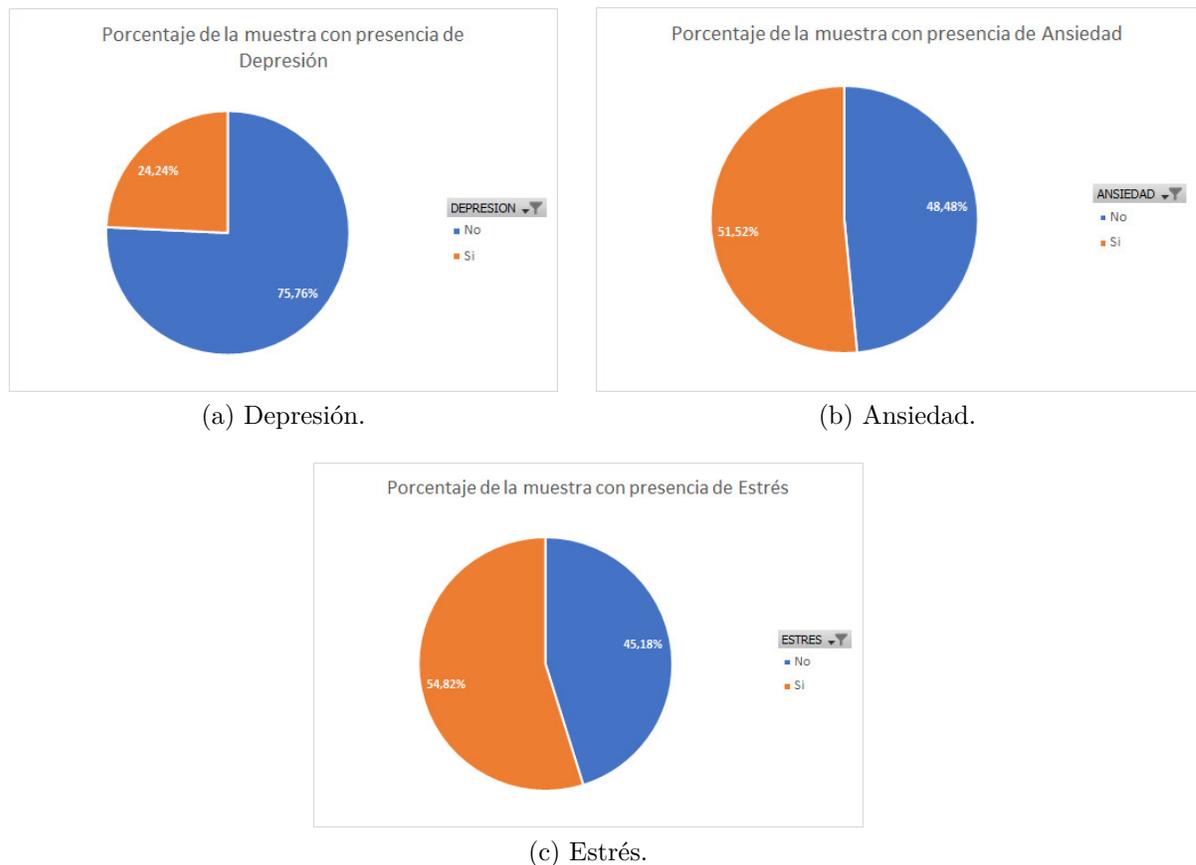


Figura 6.7: Depresión, Ansiedad y Estrés observados

6.2.2.2. Selección de variables relevantes

En consecuencia, a partir de todo lo anterior se seleccionaron 18 factores preliminares correspondientes a: edad, género, región, convivencia actual, carrera, semestre en curso, cantidad de créditos actuales, percepción de carga académica, claridad de vida, percepción de estado anímico, horas de sueño, nota de calidad del sueño, depresión, ansiedad, estrés, contexto de actividad al momento de responder y nota del día.

De la misma manera, se descartaron las variables Año de Ingreso y Contexto Compañía de la base de datos unificada (Figura 6.2c) ya que la primera es completamente explicada por el número de semestres cursados y puede no representar el avance académico real, mientras

que la segunda no presenta diferencias significativas al ser evaluada sobre depresión, ansiedad y estrés.

Sin embargo, aún existen muchas variables que pueden estar explicando el comportamiento de las patologías, por lo que es pertinente identificar y eliminar aquellos factores que son irrelevantes. De esta forma se obtendrá un resultado más limpio, menor complejidad y rendimiento computacional eficiente.

A continuación, se presentan tres métodos de selección de variables relevantes que posteriormente permitirán ejecutar los modelos y realizar comparaciones en pro de obtener conclusiones certeras: Matriz de Correlación, Análisis de Componentes Principales y Bosques aleatorios, a partir de los cuales se diseñaron modelos específicos para el caso de depresión, ansiedad y estrés, siendo estos expuestos en la sección siguiente “Minería de datos”.

1. Matriz de Correlación

En primera instancia se analizó la relación de las variables nombradas anteriormente por medio de la matriz de correlación, pudiendo así identificar factores independientes que explican depresión, ansiedad y estrés. Para esto, se procedió a calcular y graficar la matriz de correlación para las variables numéricas (Anexo A): Edad, Semestres cursados, Créditos, Carga, Claridad de vida, Estado anímico, Calidad del sueño, Depresión, Ansiedad, Estrés, Nota del día y Horas de sueño.

Cabe destacar que los 46 casos que presentaron errores de registro en Horas de sueño (presente únicamente en la jornada de la mañana) fueron eliminados de la base de datos debido a que el factor se encuentra presuntamente relacionado con las variables de estudio, por lo cual realizar una imputación de los datos por el promedio o la mediana sesgaría los resultados.

En la Tablas 6.1, 6.2 y 6.3 se muestran aquellas variables altamente correlacionadas con depresión, ansiedad y estrés, con su respectivo coeficiente y jornada del cuestionario, debido a que durante la mañana se consultó respecto a higiene de sueño (317 filas) y en la tarde, con respecto a la nota final del día (362 filas).

Tabla 6.1: Variables de interés para Depresión

	Depresión	
	Correlación Positiva	Correlación Negativa
Día	Ansiedad (0,47)	Sueño calidad (-0,38)
	Carga (0,32)	Estado anímico (-0,47)
		Semestres cursados (-0,15)
		Claridad de Vida (-0,33)
Tarde	Ansiedad (0,32)	Nota del día (-0,56)
	Carga (0,24)	Semestres cursados (-0,13)
		Estado Anímico (-0,39)
		Claridad de vida (-0,27)

Tabla 6.2: Variables de interés para Ansiedad

	Ansiedad	
	Correlación Positiva	Correlación Negativa
Día	Depresión (0,46)	Sueño calidad (-0,41)
	Carga (0,22)	Estado anímico (-0,37)
		Semestres cursados (-0,19)
		Claridad de Vida (-0,3)
Tarde	Depresión (0,32)	Nota del día (-0,41)
	Carga (0,16)	Semestres cursados (-0,16)
		Estado Anímico (-0,31)
		Claridad de vida (-0,31)

Tabla 6.3: Variables de interés para Estrés

	Estrés	
	Correlación Positiva	Correlación Negativa
Día	Ansiedad (0,1)	Sueño calidad (-0,12)
	Claridad de vida (0,06)	Créditos (-0,15)
Tarde	Claridad de vida (0,1)	Edad (-0,1)
	Nota del día (0,07)	Semestres cursados (-0,08)
		Créditos (-0,09)

2. Análisis de Componentes Principales

Por otro lado, se trabajó con el Análisis de Componentes Principales para datos mixtos, utilizando el paquete “pcamix” de la librería “PCAmixdata” de R Studio.

Para esto, se separó la base de datos según jornada y tipo de variable, obteniendo:

- Variables cuantitativas jornada día: Edad, Semestre, Creditos, HorasSueno, SuenoCalidad.
- Variables catagóricas jornada día: Genero, RM, Convivencia, AnoIngreso, Carrera, Depresion, Ansiedad, Estres, Carga, Vida, EstadoAnimico, ContextoHaciendo, ContextoCompania.
- Variables cuantitativas jornada tarde: Edad, Semestre, Creditos, NotaDia.
- Variables categóricas jornada tarde: Genero, RM, Convivencia, AnoIngreso, Carrera, Depresión, Ansiedad, Estrés, Carga, Vida, EstadoAnimico, ContextoHaciendo, ContextoCompania.

De aquí, se obtuvieron 5 dimensiones que explican la varianza de la variable dependiente en cuestión, siendo esta depresión, ansiedad o estrés, según corresponde.

Posteriormente, estas dimensiones fueron transformadas en una base de datos, la cual se segmentó en dos grupos: entrenamiento y testeo, sobre los cuales se aplicaron los modelos que se explican en la sección siguiente.

3. Bosques Aleatorios

Finalmente, se aplicó el método de selección de variables a través de Bosques Aleatorios, el cual permite trabajar con datos tanto categóricos como numéricos y proporciona la importancia de las variables.

Se trabajó con el paquete randomForest de la librería que lleva el mismo nombre en R, determinando 100 árboles. Los resultados se presentan a continuación:

Tabla 6.4: Importancia de variables encontradas con RF

Variable	Depresión		Ansiedad		Estrés	
	Día	Tarde	Día	Tarde	Día	Tarde
AnoIngreso	6,8315	2,45833	3,6716	5,6929	5,3572	5,6631
Ansiedad	12,2451	2,85237	-	-	2,9428	3,7013
Carga	4,3907	0,65976	4,0275	3,6321	2,892	3,3885
Carrera	21,4676	6,80241	7,7934	11,3836	12,3567	19,4615
Compañía	5,4812	2,30327	4,1961	8,7146	6,8801	8,5535
Haciendo	8,5318	2,77849	7,2424	10,034	10,6967	13,7654
Convivencia	3,1547	1,42766	2,0565	2,8332	1,6226	3,1132
Creditos	9,1988	1,73887	8,0322	8,2564	8,0143	9,1891
Depresion	-	-	12,0553	5,0739	2,3635	1,3741
Edad	8,8261	2,58108	5,2843	5,2715	5,893	8,6328
Estado_Animico	12,3104	12,11143	5,453	5,8533	3,0937	5,1793
Estres	1,9926	0,51933	2,4242	4,002	-	-
Genero	1,2514	0,29007	1,2614	1,9849	2,2685	1,64
HorasSueno	8,375	-	7,3737	-	8,6231	-
RM	2,582	0,3963	1,2573	2,4069	1,1295	1,6916
Semestre	10,9823	5,4704	3,8955	7,2827	3,1331	6,7821
SuenoCalidad	22,8684	-	15,4445	-	9,0154	-
Vida	6,4845	1,801	6,7057	7,2607	8,1337	9,424
NotaDia	-	17,7889	-	18,2878	-	11,0623

En este caso se determinó como punto de corte el 20 % del valor máximo de cada variable, se procediendo como sigue:

- Depresión: En este caso se evaluó el modelo con aquellas que presentaron importancia mayor a 5 en la jornada de la mañana y mayor a 1 en la jornada de la tarde, obteniendo mayor AUC con esta selección en comparación con todas haber utilizado todas las variables en cada caso.
- Ansiedad: Se evaluó el modelo con aquellas que presentaron importancia mayor a 3 en la jornada de la mañana y mayor a 5 en la jornada de la tarde, obteniendo mayor AUC

con esta selección en comparación con todas haber utilizado todas las variables en cada caso.

- Estrés: El modelo fue ejecutado con las variables que presentaron importancia mayor a 5 en la jornada de la mañana y mayor a 3 en la jornada de la tarde, obteniendo mayor AUC con esta selección en comparación con todas haber utilizado todas las variables en cada caso.

6.3. Minería de datos

6.3.1. Trabajo y modelos de Machine Learning

6.3.1.1. Bases de datos

En primera instancia se segmentó la base de datos de respuestas diarias en dos archivos, uno que contiene las respuestas de la jornada de día y otro con la jornada tarde, contemplando únicamente las variables respectivas a los cuestionarios implementados en cada caso. Luego, se generaron las variables auxiliares binarias para las componentes categóricas Género, Convivencia, Carrera, AñoIngreso y Contexto haciendo: ¿qué hacías antes de responder el cuestionario?, utilizando la función “dummy_cols” de la librería *fastDummies* de R, removiendo la primera *dummy*. Posteriormente se llevó cada variable dependiente a factor y se realizó la división del conjunto de datos en dos grupos: uno de entrenamiento y otro de testeo a proporción 70 % - 30 %, respectivamente.

6.3.1.2. Balanceo de datos

Tal como se explicó anteriormente, para obtener un resultado certero es necesario que las clases a evaluar se encuentren balanceadas. Esto, principalmente porque algoritmos como Bosques aleatorios buscan minimizar la tasa de error global y son muy sensibles a las proporciones, tendiendo a favorecer la clase con mayor proporción, provocando sesgo en los resultados.

En consecuencia, al analizar la base de datos se identificó que para el conjunto de entrenamiento las clases se encontraban desbalanceadas para depresión en ambas jornadas, por lo cual se procedió a balancear las mismas por medio de la función “SMOTE” de la librería *DMwR* de R, indicando un sobremuestreo de 160 y submuestreo de 200. En el caso de estrés y ansiedad la diferencia proporcional de las clases fue mínima, por lo que no fue necesario balancear. Lo anterior fue aplicado tanto para selección a través de matriz de correlación como para el caso de Bosques Aleatorios.

En las Tablas 6.5 y 6.6 se pueden ver las proporciones antes y después de realizado el balanceo para seleccionar el grupo de entrenamiento inicial:

Tabla 6.5: Proporciones y frecuencias sin balancear

		Depresión		Ansiedad		Estrés		
		Clase	Día	Noche	Día	Noche	Día	Noche
Proporción	0		0,65	0,87	0,47	0,45	0,46	0,45
	1		0,35	0,13	0,53	0,51	0,55	0,55
Frecuencia	0		164	220	121	125	116	113
	1		90	33	134	128	139	140

Tabla 6.6: Proporciones y frecuencias balanceadas

		Depresión		Ansiedad		Estrés		
		Clase	Día	Noche	Día	Noche	Día	Noche
Proporción	0		0,5	0,5	0,47	0,45	0,46	0,45
	1		0,5	0,5	0,53	0,51	0,55	0,55
Frecuencia	0		180	66	121	125	116	104
	1		180	66	134	128	139	118

De la misma manera, antes de realizar el Análisis de Componentes Principales se procedió con un balanceo de clases utilizando (100, 200) y (300, 150) para sobremuestreo y submuestreo en el caso de depresión día y tarde, respectivamente. Lo mismo sucedió para ansiedad: (100, 200) y (200, 200), y estrés: (100, 200) y (100, 200).

Los valores previos y posteriores se evidencian en las Tablas siguientes:

Tabla 6.7: Proporciones y frecuencias desbalanceadas PCA

		Depresión		Ansiedad		Estrés		
		Clase	Día	Noche	Día	Noche	Día	Noche
Proporción	0		0,65	0,87	0,67	0,89	0,61	0,89
	1		0,34	0,13	0,33	0,11	0,39	0,1
Frecuencia	0		43	65	45	67	41	67
	1		23	10	22	8	26	8

Tabla 6.8: Proporciones y frecuencias balanceadas PCA

		Depresión		Ansiedad		Estrés	
		Clase	Día	Noche	Día	Noche	Día
Proporción	0	0,5	0,53	0,57	0,57	0,5	0,5
	1	0,5	0,47	0,43	0,43	0,5	0,5
Frecuencia	0	46	45	44	32	52	16
	1	46	40	44	24	52	16

6.3.1.3. Programación de modelos

A partir de la selección de variables explicada en la sección correspondiente se programaron los modelos respectivos a depresión, ansiedad y estrés para cada metodología de selección y algoritmo de *machine learning*, utilizando las siguientes funciones sobre cada fila de la base de datos:

- Regresión Logística: Las funciones “glm” y “predict.glm” con familia binomial.
- Bosques Aleatorios: Las funciones “RandomForest” con 100 árboles y “predict”.
- Máquinas de Vectores de Soporte (SVM): Las funciones “svm” de tipo C-Classification, Kernel lineal, y “predict”.
- Árboles de Clasificación: Las funciones “rpart” de método Class y “predict”.

De aquí, por cada variable dependiente se ejecutaron dos modelos respectivos a Regresión Logística, dos para Bosques Aleatorios, dos para Máquinas de Vectores de Soporte y dos para Árboles de Clasificación, representando la jornada de día y de tarde para cada una.

Cabe destacar que se utilizó la base de entrenamiento para, como su nombre lo indica, entrenar los modelos, y luego la base de testeo para predecir.

Los resultados obtenidos en cada caso pueden ser consultados en los Anexos B, C y D para el caso de reducción de variables con matriz de correlación, análisis de componentes principales y bosques aleatorios, respectivamente.

6.3.1.4. Comparación de modelos

A modo de evaluar el desempeño de los modelos se definió una serie de métricas aplicables a cada algoritmo y metodología de selección, para lo cual se procedió calculando la matriz de confusión, la cual permite realizar una comparación de los datos obtenidos a través de la predicción con los datos reales, mostrando los falsos positivos, falsos negativos, verdaderos positivos y verdaderos negativos. En la Figura 6.8 se muestra un ejemplo del resultado visual y la interpretación de sus resultados[11]:

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

- **VP** es la cantidad de *positivos* que fueron *clasificados correctamente* como positivos por el modelo.
- **VN** es la cantidad de *negativos* que fueron *clasificados correctamente* como negativos por el modelo.
- **FN** es la cantidad de *positivos* que fueron *clasificados incorrectamente* como negativos. **Error tipo 2** (Falsos Negativos)
- **FP** es la cantidad de *negativos* que fueron *clasificados incorrectamente* como positivos. **Error tipo 1** (Falsos positivos)

Figura 6.8: Ejemplo matriz de confusión

Con esto se identifican las siguientes métricas relevantes:

Accuracy o exactitud: La cual se refiere a lo cerca que está el resultado de una medición del valor verdadero. Se calcula según la fórmula:

$$\frac{VP + VN}{VP + VN + FP + FN} \quad (6.1)$$

Precision o precisión: Se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud. Cuanto menor es la dispersión mayor la precisión. Para calcular su valor se utiliza la fórmula:

$$\frac{VP}{VP + FP} \quad (6.2)$$

Recall o sensibilidad: Es la proporción de casos positivos que fueron correctamente identificadas por el algoritmo. Se calcula según la fórmula:

$$\frac{VP}{VP + FN} \quad (6.3)$$

Error: El cual se refiere a lo alejado que está el resultado de una medición del valor verdadero. Se calcula según la fórmula:

$$\frac{FP + FN}{VP + VN + FP + FN} \quad (6.4)$$

F1 - Score: Combina precisión y sensibilidad en una sola métrica. Se calcula según la ecuación:

$$\frac{2 * (Recall * Precision)}{(Recall + Precision)} \quad (6.5)$$

Existen cuatro casos posibles para cada clase según las métricas nombradas:

- Alta precision y alto recall: El modelo maneja perfectamente esa clase.
- Alta precision y bajo recall: El modelo no detecta la clase muy bien, pero cuando lo hace es altamente confiable.
- Baja precisión y alto recall: La clase detecta bien la clase pero también incluye muestras de otras clases.
- Baja precisión y bajo recall: El modelo no logra clasificar la clase correctamente.

Además, para cada modelo se calculó el valor **AUC** o área bajo la curva ROC, la cual representa la probabilidad de que el modelo clasifique un ejemplo positivo aleatorio más alto que un ejemplo negativo aleatorio [40]. Para ello se utilizó la función “multiclass.roc” de la librería “pROC” de R. Según el valor AUC una prueba es considerada regular [0.6, 0.75), buena [0.75, 0.9), muy buena [0.9, 0.97) o excelente [0.97, 1): Test excelente.

Cabe destacar que estos modelos fueron aplicados sobre la base de datos original en el caso de la selección de variables con matriz de correlación, y sobre la base de datos generada por el método de selección de variables PCA o Bosques Aleatorios, para el caso correspondiente. Así mismo, dado que dos variables de interés se presentaron solo en la jornada AM y otra en la jornada PM, los resultados se muestran según el tipo de jornada.

A continuación, se presentan las métricas e indicadores obtenidos por cada algoritmo utilizando las variables de interés elegidas a través de matriz de correlación, análisis de componentes principales y bosques aleatorios, respectivamente:

6.3.1.5. Resultados con matriz de correlación

Tabla 6.9: Métricas por modelo con MC

		Modelo	Accu.	Recall	Precision	Error	F1	AUC
Dep.	D	Logística	0,67	0,64	0,53	0,33	0,58	0,78
		Random F	0,76	0,86	0,59	0,24	0,70	0,72
		SVM	0,69	0,72	0,64	0,31	0,68	0,68
		Árboles Clas.	0,71	0,54	0,62	0,29	0,58	0,68
	T	Logística	0,76	0,93	0,34	0,24	0,50	0,83
		Random F	0,91	0,93	0,79	0,09	0,85	0,86
		SVM	0,9	0,91	0,86	0,10	0,88	0,88
		Árboles Clas.	0,91	0,79	0,61	0,09	0,69	0,86
Ans.	D	Logística	0,83	0,81	0,85	0,17	0,83	0,86
		Random F	0,85	0,87	0,84	0,15	0,85	0,85
		SVM	0,83	0,83	0,83	0,17	0,83	0,83
		Árboles Clas.	0,8	0,72	0,87	0,20	0,79	0,8
	T	Logística	0,72	0,73	0,73	0,28	0,73	0,81
		Random F	0,78	0,78	0,78	0,22	0,78	0,78
		SVM	0,72	0,7	0,73	0,28	0,71	0,72
		Árboles Clas.	0,8	0,73	0,82	0,20	0,77	0,78
Est.	D	Logística	0,6	0,54	0,65	0,40	0,59	0,62
		Random F	0,6	0,6	0,59	0,40	0,59	0,6
		SVM	0,62	0,68	0,58	0,38	0,63	0,63
		Árboles Clas.	0,65	0,69	0,67	0,35	0,68	0,65
	T	Logística	0,68	0,75	0,69	0,32	0,72	0,68
		Random F	0,67	0,67	0,67	0,33	0,67	0,67
		SVM	0,63	0,61	0,65	0,37	0,63	0,63
		Árboles Clas.	0,69	0,77	0,65	0,31	0,70	0,68

Para el caso de depresión se tiene que los algoritmos Bosques Aleatorios, Árboles de Clasificación y SVM presentan mayor exactitud (en orden decreciente) para ambas jornadas del día, siendo Bosques Aleatorios aquel que cumple de mejor manera al utilizar un criterio combinado de Accuracy, Recall, Precision y AUC.

Por el lado de ansiedad, en la primera jornada del día los algoritmos más exactos fueron Bosques Aleatorios, Regresión Logística y SVM, mientras que para la tarde destacan Árboles de Clasificación y Bosques Aleatorios, en el orden respectivo, siendo Bosques Aleatorios el algoritmo que se desempeñó de manera superior al evaluar las métricas en su conjunto.

Finalmente, en el caso de estrés las pruebas tuvieron resultados regulares (AUC menor a 0.75) y los valores encontrados fueron notoriamente más bajos en todos los indicadores, destacando en exactitud Árboles de Clasificación y SVM en la jornada de la mañana, y tanto Árboles de Clasificación como Regresión Logística y Bosques Aleatorios para la jornada de la tarde. Aquí, la combinación óptima de métricas se encuentra en Árboles de Clasificación.

6.3.1.6. Resultados con análisis de componentes principales

Tabla 6.10: Métricas por modelo con PCA

		Modelo	Accu.	Recall	Precision	Error	F1	AUC
Dep.	D	Logística	0,90	0,8	0,88	0,1	0,84	0,94
		Random F	0,93	0,95	0,9	0,07	0,92	0,92
		SVM	0,86	0,95	0,7	0,14	0,81	0,83
		Árboles Clas.	0,83	0,8	0,73	0,17	0,76	0,82
	T	Logística	0,83	0,8	0,57	0,12	0,67	0,87
		Random F	0,76	0,79	0,6	0,24	0,68	0,71
		SVM	0,76	0,75	0,8	0,24	0,77	0,78
		Árboles Clas.	0,79	0,6	0,38	0,21	0,46	0,71
Ans.	D	Logística	0,79	0,44	0,8	0,21	0,57	0,91
		Random F	0,82	0,95	0,56	0,18	0,70	0,76
		SVM	0,82	1	0,44	0,18	0,61	0,72
		Árboles Clas.	0,79	0,77	0,64	0,21	0,7	0,78
	T	Logística	0,97	1	0,8	0,03	0,88	0,85
		Random F	0,94	0,96	0,75	0,06	0,84	0,86
		SVM	0,97	0,96	0,75	0,03	0,84	0,88
		Árboles Clas.	0,9	0,75	0,6	0,1	0,66	0,84
Est.	D	Logística	0,82	0,91	0,71	0,18	0,8	0,89
		Random F	0,86	0,77	0,92	0,14	0,84	0,88
		SVM	0,79	0,77	0,82	0,21	0,79	0,79
		Árboles Clas.	0,64	0,85	0,52	0,36	0,65	0,67
	T	Logística	0,82	0,94	0,4	0,21	0,57	0,95
		Random F	0,79	0,79	0,75	0,21	0,77	0,77
		SVM	0,76	0,76	0,75	0,24	0,75	0,76
		Árboles Clas.	0,67	0,93	0,27	0,33	0,42	0,77

Al analizar la exactitud obtenida en depresión se identifican los mayores valores para los algoritmos Bosques Aleatorios (primera jornada), Regresión Logística (ambas jornadas) y Árboles de Clasificación (segunda jornada); destacando Bosques Aleatorios y Regresión Logística con los mejores criterios combinados para las mañana y tarde, respectivamente.

En el caso de la variable dependiente ansiedad, en la primera jornada del día los algoritmos más exactos fueron Bosques Aleatorios, SVM, mientras que para la tarde destacan SVM y Regresión Logística, siendo Bosques Aleatorios el algoritmo que se desempeñó de manera superior al evaluar las métricas en su conjunto en la primera jornada y SVM en la segunda.

De manera similar, al evaluar estrés se obtuvieron mejores métricas con respecto a la selección de variables relevantes a través de matriz de correlación y los algoritmos con mayor exactitud fueron Bosques Aleatorios y Regresión Logística, siendo esta última la que se desempeñó mejor al usar criterio de métricas combinadas, para ambas jornadas.

6.3.1.7. Resultados con bosques aleatorios

Tabla 6.11: Métricas por modelo con BA

		Modelo	Accu.	Recall	Precision	Error	F1	AUC
Dep.	D	Logística	0,74	0,7	0,61	0,26	0,65	0,8
		Random F	0,76	0,81	0,67	0,24	0,75	0,74
		SVM	0,72	0,74	0,67	0,28	0,7	0,71
		Árboles Clas.	0,69	0,49	0,57	0,31	0,52	0,65
	T	Logística	0,96	0,97	0,97	0,04	0,96	0,95
		Random F	0,99	0,97	0,98	0,01	0,97	0,92
		SVM	0,97	0,95	0,99	0,03	0,91	0,97
		Árboles Clas.	0,92	0,9	0,93	0,08	0,92	0,92
Ans.	D	Logística	0,77	0,74	0,8	0,23	0,77	0,86
		Random F	0,8	0,82	0,78	0,2	0,79	0,8
		SVM	0,72	0,76	0,68	0,28	0,71	0,72
		Árboles Clas.	0,74	0,64	0,82	0,26	0,72	0,74
	T	Logística	0,65	0,59	0,68	0,34	0,63	0,74
		Random F	0,8	0,83	0,76	0,2	0,79	0,8
		SVM	0,68	0,78	0,57	0,32	0,66	0,68
		Árboles Clas.	0,68	0,63	0,69	0,32	0,66	0,68
Est.	D	Logística	0,79	0,87	0,77	0,21	0,82	0,85
		Random F	0,94	0,92	0,06	0,14	0,94	0,95
		SVM	0,76	0,63	0,88	0,24	0,73	0,76
		Árboles Clas.	0,79	0,81	0,81	0,21	0,81	0,79
	T	Logística	0,75	0,66	0,63	0,27	0,64	0,58
		Random F	0,6	0,56	0,69	0,4	0,57	0,58
		SVM	0,58	0,47	0,66	0,42	0,52	0,56
		Árboles Clas.	0,75	0,61	0,64	0,27	0,63	0,59

Utilizando los mismos criterios anteriores se identificó que para el caso de depresión los algoritmos que entregaron mayor exactitud fueron Regresión Logística y Bosques Aleatorios en el caso de la mañana y tanto SVM como Bosques Aleatorios en el caso de la tarde. Ahora bien, al analizar las métricas en su conjunto, el algoritmo con mejor desempeño fue Bosques Aleatorios para ambos casos.

De la misma manera, para ansiedad tanto Regresión Logística (jornada mañana), Bosques Aleatorios (ambas jornadas) y SVM (jornada tarde) presentaron mayor exactitud al ejecutarse los modelos, destacando Bosques Aleatorios en ambas jornadas al evaluar todas las métricas.

Por último, para el caso de estrés los valores AUC obtenidos fueron los más bajos de los tres métodos de selección en la tarde (menor a 0.6) y levemente menores a los comparativos en el caso de la mañana. Tanto Bosques Aleatorios como Regresión Logística y Árboles de clasificación presentaron alta exactitud en ambas jornadas. Sin embargo, Bosques Aleatorios

se desempeña de mejor manera en la mañana al analizar el conjunto de métricas, mientras que en la tarde lo hace Regresión Logística.

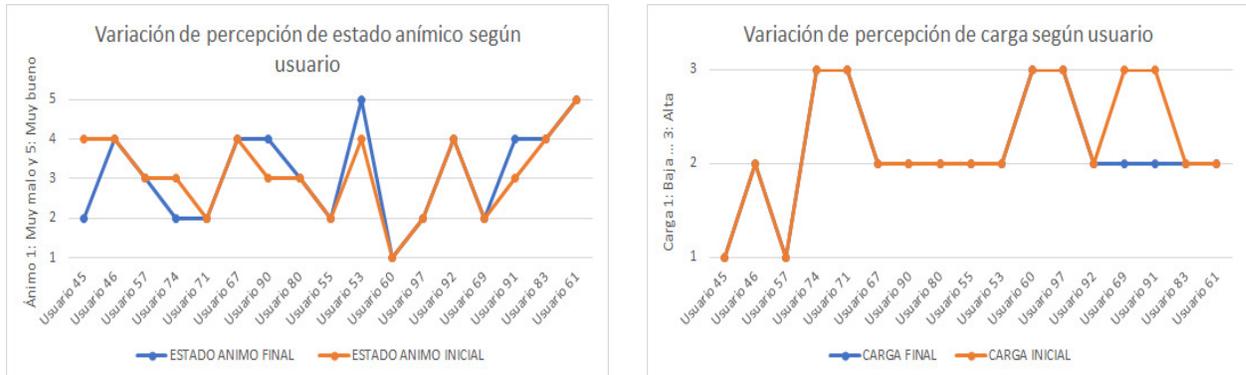
6.4. Comparaciones

6.4.1. Preguntas iniciales y finales

Una vez finalizados los cuestionarios diarios se aplicaron dos preguntas para realizar contrastes y analizar cómo cambia la percepción de la persona cuando es más consciente de su estado y sus acciones.

Destacando que a la fecha de entrega de esta memoria de título solo 17 usuarios respondieron el formulario final, en la Figura 6.9a se evidencia un leve cambio en la percepción de estado anímico de los participantes, empeorando en 2 casos y mejorando en 3. Esto permite concluir de manera preliminar que la percepción de estado anímico si varía conforme avanza el tiempo y esta variación se puede captar a través de un seguimiento constante.

Por otro lado, al analizar el gráfico 6.9b se evidencia que solo existe una variación en la percepción de la carga académica una vez finalizado el cuestionario, permitiendo concluir preliminarmente que el valor de esta variable no cambia con el paso del tiempo ni con la consciencia del estado anímico. Sin embargo, se requieren más datos para obtener conclusiones certeras.



(a) Percepción anímica.

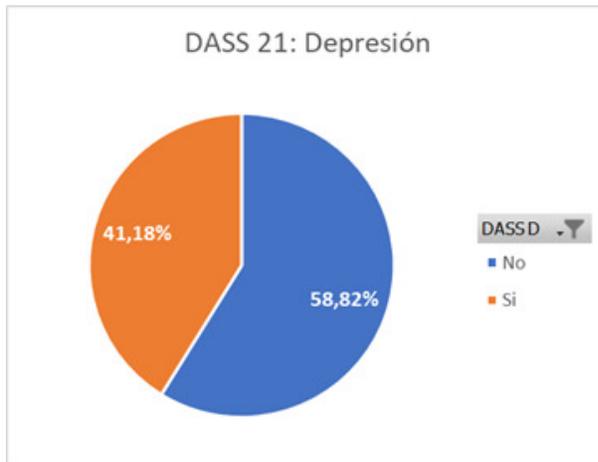
(b) Percepción de carga académica.

Figura 6.9: Comparación de percepción de carga académica pre y post cuestionarios

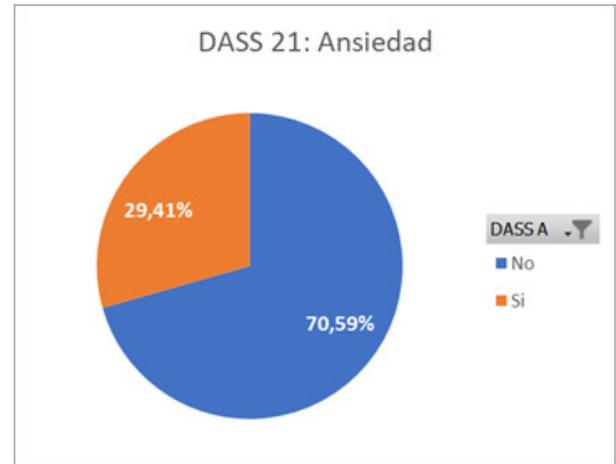
Finalmente, a partir de los resultados del formulario DASS-21 mostrados en la Figura 6.10, que entrega una categorización automática y validada a partir de las respuestas de los participantes, se comprende que la submuestra se encuentra en buen estado anímico, presentando baja prevalencia de depresión, ansiedad y estrés.

En contraste con las tendencias de prevalencia encontrados en las Figuras 6.7a, 6.7b y 6.7c, destaca la mayor presencia de depresión y menor presencia de ansiedad y estrés, lo cual puede explicarse en primera instancia por la cantidad de respuestas obtenidas al momento de

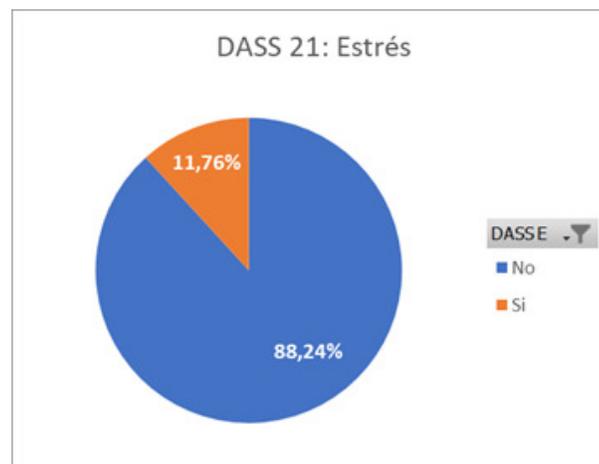
la entrega de esta memoria (submuestra no es representativa) y/o por la disminución de las protestas a nivel nacional (que tuvieron efecto significativo sobre las personas), finalización de cursos y comienzo de vacaciones.



(a) DASS 21: Depresión.



(b) DASS 21: Ansiedad.



(c) DASS 21: Estrés.

Figura 6.10: DASS 21: Depresión, Ansiedad y Estrés observados

Capítulo 7

Discusión

El presente trabajo de memoria analiza tanto la depresión, como la ansiedad y el estrés en estudiantes universitarios. En primera instancia se identificaron los factores relevantes para cada variable dependiente en miras de satisfacer el objetivo general y la hipótesis. En segundo lugar, se levantaron los datos a través del diseño e implementación de la aplicación *Ánima*, y en tercera instancia se obtuvieron los resultados asociados a los modelos de clasificación sobre cada fila de datos obtenidos.

1) Aplicación Móvil

Una vez definidos los factores incidentes para depresión, ansiedad y estrés sobre la base de la literatura y entrevistas a actores relevantes, se trabajó en la aplicación para realizar un levantamiento de datos que fuera robusto, seguro y a la vez sencillo de entender y utilizar. De aquí cabe notar lo siguiente:

- Se debe considerar que la aplicación fue promocionada y puesta a disposición de los usuarios en medio de un período excepcional de alteración de la rutina por causas del estallido social, recopilando datos en el período de noviembre de 2019 a enero de 2020. Los resultados expuestos consideran este alcance.
- 62 personas (50.4 % del total de descargas) iniciaron sesión y completaron el cuestionario inicial, mientras que 47 (38.2 % del total de descargas) respondieron al menos un cuestionario diario, lo que se traduce en una retención del 76 % de los usuarios activos, valor acorde a los márgenes de uso evidenciados en investigaciones similares. Lo anterior puede haber sido causado porque los participantes decidieron no aceptar el consentimiento informado, por desinterés o por requisitos de uso, dado que solo permite el ingreso con correo institucional, el cual no es ampliamente utilizado, pero al que toda la población tiene acceso y permite llevar un control de uso desde la plataforma Google Play donde se encontraba disponible. También puede deberse al requisito de acceso a internet, el cual no está asegurado fuera de la universidad.
- Con respecto a las respuesta obtenidas, 17 participantes (36 % de los usuarios activos) respondieron al menos la mitad de los cuestionarios. La moda fue de 5 cuestionarios,

mediana de 14, desviación estándar de 13 y un promedio de 16 cuestionarios respondidos. Cabe destacar que a cada estudiante se le asignaron 42 ejecuciones posteriores al cuestionario inicial.

- Se debe tener en cuenta que la aplicación contempla el uso de internet al momento de recibir y enviar el cuestionario. También requiere los permisos específicos del usuario para el acceso a los sensores del celular, y posee una ventana de tiempo en el cual la respuesta es válida, no permitiendo responder la ejecución de manera atrasada.
- Es importante considerar que existen nuevas versiones de *Android* que no permiten la utilización de todos los sensores pasivos del celular, por lo que pueden existir funcionalidades de la aplicación que no sean válidas en todos los dispositivos.

2) Variables dependientes e independientes

Al analizar las variables dependientes, se encontró que depresión y ansiedad están altamente correlacionadas, siendo explicadas entre sí y, además, por la carga académica, calidad del sueño, claridad de vida y estado anímico. Lo anterior coincide con la literatura, donde se destacaba que ambos trastornos tenían factores en común y, si bien no se presentaban necesariamente en conjunto, podrían estar correlacionadas, lo cual era un resultado esperado por parte de la investigadora. Sin embargo, es importante destacar que en el método de reducción de variables por matriz de correlación se encontraron exactamente las mismas variables independientes para depresión y ansiedad, lo cual hace que no tenga sentido evaluar ambas por separado. Sin embargo, los otros dos métodos agregan variabilidad y/o nuevas variables ortogonales que coinciden con lo evidenciado en los artículos expuestos en el estado del arte.

En el caso del factor género, a pesar de que se evidenció una leve diferencia para depresión y ansiedad, no fue una variable significativa en ningún caso, lo que contradice algunos estudios anteriores que indican que hay mayor presencia en mujeres. Lo mismo ocurrió con el caso de la región de procedencia, convivencia habitacional y compañía situacional, factores que se esperaban fueran relevantes dada su alta relación con cambios de hábito al momento de comenzar la universidad, pero que tampoco alcanzaron valores significativos. Como razones se puede nombrar que la conducta se presenta en personas que recién ingresan a la educación superior y existió una baja participación de su parte al momento del levantamiento de los datos. Además, que un gran porcentaje de la muestra vive con familiares y que la gran mayoría proviene de la Región Metropolitana. Sin embargo, se propone no descartar estas variables en estudios futuros con población distribuida de manera homogénea.

Luego, al momento de evaluar estrés, se identificó que las horas de sueño fueron menos significativas que la calidad de sueño percibida por los estudiantes, presentando esta última una correlación negativa con la variable de estudio, otorgando una mirada cualitativa al efecto de la higiene del sueño expuesta por otros investigadores. Sobre la misma base, se esperaba que fueran relevantes factores situacionales, personales y académicos, pero no se encontró una relación determinante para los dos primeros casos; también se esperaba algún tipo de correlación con depresión y ansiedad, pero los valores obtenidos no mostraron mayor

significancia, situación que fue avalada por el psicólogo asociado, explicando que no hay evidencia de que las variables estén relacionadas.

Finalmente, se recomienda trabajar con los semestres cursados y semestre actual del estudiante, debido a que al considerar el año de ingreso se estima que cumple la malla curricular a tiempo, situación que no se ajusta a la realidad y que puede sesgar los resultados.

3) Tipo de reducción de variables y modelos de clasificación

Evaluando el desempeño de distintos algoritmos y métodos de agrupación, se identificó que estos difieren según la jornada del día en la cual se obtuvieron los datos y la variable dependiente estudiada, situación esperada debido a la diferencia de variables de interés en cada caso.

Con respecto a la selección y reducción de variables relevantes, es importante destacar que las tres metodologías requieren un correcto balanceo de clases, estandarización de niveles y selección de puntajes de corte para las elecciones. De hecho, Bosques Aleatorios permite trabajar con grandes cantidades de datos y consta de ejecutar múltiples árboles de decisión explorando así diversas alternativas y eliminando relaciones lineales o no esperadas entre las variables iniciales, entregando la importancia de estas al investigador. En cambio, el método de Matriz de Correlación también trabaja con las variables iniciales, entregando el valor de la relación entre las mismas, pero se vuelve contraproducente al trabajar con datos numéricos y categóricos al mismo tiempo debido al peso que le entrega a cada uno.

Por otro lado, en el método de Análisis de Componentes Principales se crean nuevas variables que resultan de combinaciones lineales de las variables iniciales para minimizar errores, lo cual permite obtener resultados más cercanos a la realidad. Lo anterior se puede evidenciar en los porcentajes de exactitud y de probabilidad de éxito tanto de depresión, como de ansiedad y estrés, siendo estos más altos que con los otros métodos.

Ahora bien, al analizar los algoritmos de clasificación se debe tener en cuenta que todos ellos tienen ventajas, desventajas y generalmente resultados similares, por lo que es pertinente considerar cuál es el tipo de variable que se estudia, cuáles son las variables independientes y cómo se quieren interpretar los resultados al momento de elegir con cual trabajar. Aquí, Regresión Logística es altamente utilizado en campos médicos por su fácil interpretación, pero depende bastante de las características que se le dan al modelo. Situación similar sucede con Bosques Aleatorios, donde, tal como se expresó anteriormente, tiende a sobreajustar y parcializar hacia variables con más categorías. Sin embargo, a diferencia de Regresión Logística y Árboles de Clasificación, este último es preferible cuando el número de variables independientes es alto.

Siguiendo la línea de lo recién explicado, se trabajó con estos métodos y algoritmos debido a que cumplían con los criterios necesarios tanto para variables dependientes como independientes. Sin embargo, al momento de escoger el mejor modelo se decidió evaluar bajo la combinación de las métricas *Accuracy*, *Recall*, *Precision* y *AUC*, debido a que de esta manera se abarcan todos los aspectos en los cuales se puede estar realizando sobreajuste o parcialización de resultados, seleccionando así aquellos con mayores valores bajo los estándar

res expresados para cada caso.

Finalmente, en caso de que los usuarios completen exitosamente el levantamiento de datos, sería interesante aplicar los modelos de manera agrupada sobre cada uno de ellos. Durante este trabajo no se cumplió una cantidad base para los participantes, teniendo que realizar los análisis a nivel de ejecución recopilada y no de usuarios.

4) Avance a Ensayo Clínico Aleatorizado (RCT)

Por medio de los resultados y alcances obtenidos a través de este estudio de factibilidad se evidencia que para conducir un RCT es necesario aumentar la aceptabilidad de los usuarios universitarios a aplicaciones de salud mental. Esto conlleva la presentación de la aplicación en conjunto de estudios que avalen la normalidad de temas asociados a deficiencias en el estado anímico y por sobre todo, a solicitar ayuda profesional cuando se requiere.

Así mismo, destaca la necesidad de asignar aleatoriamente a los participantes, lo cual puede ser realizado a través de una asignación aleatoria por año de ingreso o por semestre en curso. Si la población corresponde a los estudiantes de la FCFM, se deben abarcar 362 estudiantes de un total de 6014 activos en la base de datos a diciembre 2019; todo esto para alcanzar un 95 % de confianza. Ahora bien, en caso de extender el grupo objetivo a otras universidades y/o facultades, se requieren modificaciones en la aplicación que permitan inicio de sesión con correos externos al institucional de la Universidad de Chile y, además, la modificación de preguntas respectivas a carrera de estudio, año de ingreso y semestre en curso. Sin embargo, la metodología de asignación aleatoria y llegada a través de correo institucional es la misma.

A partir de las tasas de uso de la aplicación, el período de levantamiento de datos de 21 días - 42 ejecuciones parece ser elevado debido a que ningún participante terminó de manera exitosa. Sin embargo, debido a la alta desviación estándar presente en la cantidad de ejecuciones realizadas por los usuarios, no es posible recomendar un número exacto de cuestionarios, pero independiente a este número, es pertinente seguir respetando una semana de diferencia con el cuestionario final DASS-21 para no alterar la validez.

Adicionalmente, dado que en la literatura se indica que los períodos de transición son aquellos que desencadenan o potencian los factores de depresión, ansiedad y estrés, y considerando que los resultados aquí encontrados muestran que los estudiantes que cursan sus primeros semestres o que están atrasados con respecto a la malla académica por varios semestres son aquellos que declaran tener mal o muy mal estado anímico, se recomienda realizar un análisis enfocado al avance académico y un seguimiento a lo largo de la carrera.

Por otro lado, se propone incorporar el trabajo con los datos que se recuperan a través de los sensores del dispositivo móvil, a modo de realizar análisis más completos donde se requiera una menor interacción con el usuario y se de más peso a su contexto y comportamiento en el momento de la respuesta.

De manera adicional, una vez conducido el RCT se propone estudiar las siguientes hipótesis secundarias:

1. Es posible predecir las variables de depresión, ansiedad y estrés con un modelo de clasificación de estados anímicos y sintomatología orientada a caracterización del usuario, contexto situacional, higiene del sueño, rendimiento académico, interacción con plataformas de estudio, información recopilada por sensores del celular (datos pasivos), y percepción sobre sí mismo.
2. Es posible clasificar a los estudiantes según su estado anímico a partir de información estudiantil proporcionada por la FCFM y datos pasivos recopilados a través de una aplicación móvil.
3. Es posible predecir el resultado de las variables de depresión, ansiedad y estrés del formulario DASS-21 por un modelo de clasificación de estados anímicos y sintomatología orientada a caracterización del usuario, contexto situacional, higiene del sueño, rendimiento académico, interacción con plataformas de estudio, información recopilada por sensores del celular (datos pasivos), y percepción sobre sí mismo.

Capítulo 8

Conclusión

A través de este trabajo se identificaron los principales factores que influyen en los estudiantes universitarios y que generan o potencian depresión, ansiedad y estrés. Así mismo, por medio del diseño e implementación de la aplicación móvil *Ánimi*, se levantaron datos asociados a contexto personal, habitacional, situacional y se relativos a las variables de estudio. Todo lo anterior permitió la selección de variables relevantes, el diseño de modelos de clasificación que entregaron resultados favorables para los efectos propuestos, y la determinación de los indicadores para el desarrollo de un Ensayo Clínico Aleatorio.

Con respecto a los factores asociados a depresión, ansiedad y estrés es imperativo evaluar higiene del sueño, carga académica, avance curricular, contexto situacional al momento de responder, evaluación final del día y percepción propia del estado anímico, debido a que resultaron ser aquellos con mayor incidencia sobre lo estudiado. Además, no se debe dejar fuera el género del participante y su carrera, debido a que hay evidencia que indica que las patologías se presentan mayoritariamente en mujeres y en carreras demandantes y competitivas.

Luego de realizado el análisis a los datos, se determinó segmentar la base en jornada *AM* y *PM*, debido a la existencia de preguntas específicas que aportan valor al estudio. Con esto, y gracias a la comparación de métodos de reducción de variables y algoritmos de clasificación, se redujeron las variables independientes por medio del Análisis de Componentes Principales, y se obtuvo mejores métricas para clasificar con Bosques Aleatorios, Regresión Logística y Máquinas de Vectores de Soporte para el caso de depresión, ansiedad y estrés, respectivamente. Cabe destacar que por resultados similares y ventajas asociadas al manejo de grandes volúmenes de datos, se propone trabajar con Bosques Aleatorios para seleccionar variables en una investigación completa.

A partir de la información obtenida, la aplicación móvil diseñada exclusivamente para el levantamiento de datos permitió tanto llevar el control y seguimiento de los usuarios, como un manejo simple, sencillo y eficaz de la información recopilada. Su tasa de retención alcanzó el 76 %, posee una conexión a la nube AWS de Amazon y conexión con Heroku. Lo anterior y su fácil manejo a través de Firebase hacen de *Ánimi* la aplicación idónea para los objetivos planteados.

Ahora bien, según los resultados y el análisis asociado a ellos, la cantidad de ejecuciones promedio alcanzó los 16 cuestionarios, pero no se puede concluir que este sea el número óptimo debido a la alta desviación estándar encontrada en la muestra.

De esta manera, todo lo expuesto anteriormente permite concluir que es factible clasificar a los estudiantes según su estado anímico con las variables expuestas y por medio de la utilización de una aplicación móvil, validando así la hipótesis presentada.

Por otro lado, para obtener representatividad a un 95% de confianza en la población de la FCFM el grupo objetivo debe ser de 362 estudiantes. Así mismo, se requiere de la colaboración de docentes, autoridades, unidades de Calidad de Vida, Bienestar Estudiantil, Orientación Vocacional y SEMDA para efectos de promoción y seguimiento, y resulta importante agregar factores académicos para conducir un Ensayo Clínico Aleatorizado (RCT).

Finalmente, la presente memoria de título, además de determinar la factibilidad de lo estudiado, agrega valor social a la comunidad, ya que 1) fomenta tanto el autoconocimiento, como la identificación y normalización de estados anímicos, y 2) una vez se desarrolle el RCT, permitirá a las entidades pertinentes poder determinar qué estudiante necesita apoyo profesional para sobrellevar su etapa universitaria, cubriendo una de las principales problemáticas expuestas en torno al tema.

Capítulo 9

Trabajo Futuro

En línea con lo expuesto a lo largo de la memoria, y dado que los análisis indican que la carga académica, la carrera y el semestre en curso son relevantes cuando a depresión, ansiedad y estrés respecta, como trabajo futuro se propone incorporar tanto información relativa al desempeño académico, la cual se encuentra en U-Campus (caracterización, avance curricular, historial de notas, cantidad de créditos inscritos), y U-Cursos (calendario académico), como caracterización general del estudiante, a obtenerse preliminarmente de Bienestar Estudiantil con el FOCES.

Por otro lado, a modo de disminuir la interacción directa con el usuario y hacer uso de las tecnologías, se propone incorporar al estudio los datos obtenidos a través de los sensores del teléfono. De hecho, como se explica en la sección 5.2, la aplicación Ánimi contempla las funcionalidades de levantamiento de datos de manera interactiva con el usuario y la recopilación de datos entregada por los sensores del dispositivo móvil. Además, almacena la información en la nube AWS con altos estándares de seguridad y permite la descarga de la misma en archivos de formato CSV. Cabe destacar que la aplicación se encuentra disponible para distribución en formato APK.

También, es importante realizar una vinculación con autoridades de la universidad, entiéndase esto como Decanato, Docencia, Orientación Vocacional, Bienestar Estudiantil y Calidad de Vida, a modo de realizar validación, promoción y seguimiento. Específicamente, como la Unidad de Calidad de Vida posee programas de prevención y acompañamiento, se propone utilizar los resultados obtenidos a través del uso de la aplicación Ánimi, determinando puntos de corte internos para tomar la acción óptima según cada caso. De aquí los casos extremos deberían ser derivados a SEMDA para tratamientos a largo plazo, mientras que aquellos de nivel medio o bajo ser atendidos por la misma unidad en conjunto con Bienestar Estudiantil y Orientación Vocacional según su naturaleza. El proceso general se muestra en la Figura siguiente:

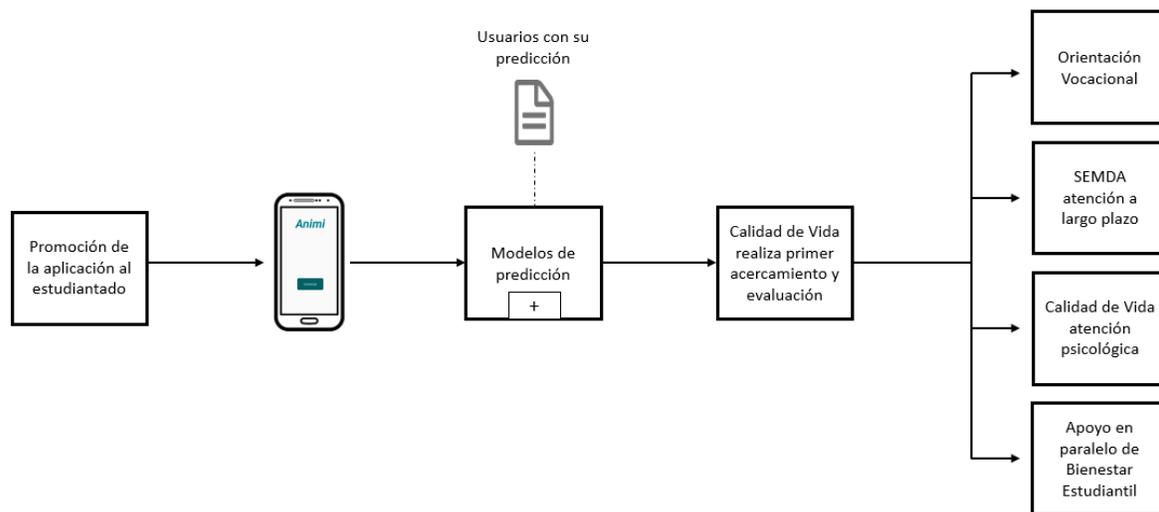


Figura 9.1: Visualización general del proceso propuesto

A partir de los resultados expuestos y lo recién propuesto, para llevar a cabo el Ensayo Clínico Aleatorizado ya sea dentro de la FCFM y/o fuera de ella, se debe:

- Realizar las modificaciones correspondientes a la aplicación en términos de número de ejecuciones y formas de acceso (actualmente permite correo electrónico con dominio “@ug.uchile.cl” y “@uchile.cl”).
- Seleccionar la población y el grupo objetivo según el nivel de confianza escogido.
- Llevar a cabo el análisis de datos, realizar la reducción de variables, implementar el modelo de clasificación y obtener conclusiones.

Finalmente, se propone realizar seguimiento a los estudiantes, a modo de acompañarlos a lo largo de su vida universitaria, normalizar estados anímicos, fomentar la solicitud de ayuda profesional y detectar potenciales casos de manera temprana.

Bibliografía

- [1] Los conflictos familiares afectan la salud mental de los universitarios, 2006. URL <https://www.emol.com/noticias/Tendencias/2006/01/05/727549/los-conflictos-familiares-afectan-la-salud-mental-de-los-universitarios.html>.
- [2] ¿es necesario balancear los datos?, 2007. URL <http://powerhousedm.blogspot.com/2007/10/es-necesario-balancear-los-datos.html>.
- [3] Informe de actividades 2014-2016 fcfm, 2016. URL <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=7&ved=2ahUKEwiS1Yy-5KTPAhUZJrkGHffiANIQFjAGegQIBBAB&url=http%3A%2F%2Fescuela.ingenieria.uchile.cl%2Fdam%2Fjcr%3Aa0347b54-62dc-4c3b-a901-5e345428b642%2Finforme-escuela2014-1016version-webfinal.pdf&usq=AOvVaw0WW50ZEoKIUv6gEJu20UC7>.
- [4] Smote explained for noobs - synthetic minority over-sampling technique line by line, 2017. URL http://rikunert.com/SMOTE_explained.
- [5] Catálogo de cursos fcfm, Diciembre 2019. URL https://ucampus.uchile.cl/m/fcfm_catalogo/.
- [6] ¿qué es kdd y minería de datos?, Diciembre 2019. URL <http://fcojlanda.me/es/ciencia-de-los-datos/kdd-y-mineria-de-datos-espanol/>.
- [7] Salud mental en los estudiantes: el tema que remece a la universidad de chile, 2019. URL <https://www.theclinic.cl/2019/07/04/salud-mental-en-los-estudiantes-el-tema-que-remece-a-la-universidad-de-chile/>.
- [8] J. Haxby Abbott. The distinction between randomized clinical trials (rcts) and preliminary feasibility and pilot studies: What they are and are not. *Journal of Orthopaedic & Sports Physical Therapy*, 44(8):555–558, 2014. doi: 10.2519/jospt.2014.0110. URL <https://doi.org/10.2519/jospt.2014.0110>. PMID: 25082389.
- [9] Ana Amaral, Maria João Soares, Ana Margarida Pinto, Ana Pereira, Nuno Madeira, Sandra Bos, Mariana Marques, Carolina Roque, and António Macedo. Sleep difficulties in college students: The role of stress, affect and cognitive processes. *Psychiatry Research*, 260, 11 2017. doi: 10.1016/j.psychres.2017.11.072.
- [10] Zayra Antúnez and Eugenia Vinet. Problemas de salud mental en estudiantes de una universidad regional chilena. *Revista médica de Chile*, 141:209–216, 02 2013. doi: 10.4067/S0034-98872013000200010.
- [11] Juan Ignacio Barrios Arce. La matriz de confusión y sus métricas, 2019. URL <https://www.juanbarrios.com/matriz-de-confusion-y-sus-metricas/>.

- [12] Varios Autores. Salud mental del adolescente, 2018. URL who.int/es/news-room/fact-sheets/detail/adolescent-mental-health.
- [13] Tomas Baader, Carmen C, José Molina, Marcelo V, Catalina Alamo, Carlos F, Silvia B, and Paula B. Diagnóstico de la prevalencia de trastornos de la salud mental en estudiantes universitarios y los factores de riesgo emocionales asociados. *Revista chilena de neuro-psiquiatría*, 52:167–176, 09 2014. doi: 10.4067/S0717-92272014000300004.
- [14] Tomas Baader M, Carmen Rojas C, José Luis Molina F, Marcelo Gotelli V, Catalina Alamo P, Carlos Fierro F, Silvia Venezian B, and Paula Dittus B. Diagnóstico de la prevalencia de trastornos de la salud mental en estudiantes universitarios y los factores de riesgo emocionales asociados. *Revista chilena de neuro-psiquiatría*, 52:167 – 176, 09 2014. ISSN 0717-9227. URL https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0717-92272014000300004&nrm=iso.
- [15] Ana Barrera. Los universitarios no piden ayuda psicológica porque se asume como un signo de debilidad, 2019. URL <https://n9.cl/jmoz>.
- [16] Dennis Becker, Vincent Bremer, Burkhardt Funk, Joost Asselbergs, Heleen Riper, and Jeroen Ruwaard. How to predict mood? delving into features of smartphone-based data. 08 2016. URL https://www.researchgate.net/publication/310598732_How_to_Predict_Mood_Delving_into_Features_of_Smartphone-Based_Data.
- [17] Radio Bio Bio. Ranking qs: Universidades chilenas rompen récord pero aún están en deuda en investigación científica, 2020. URL <https://www.biobiochile.cl/noticias/nacional/chile/2019/06/19/ranking-qs-universidades-chilenas-rompieron-record-pero-aun-se-evidencia-brecha.shtml>.
- [18] Amanda Blatch-Jones, Wei Pek, Emma Kirkpatrick, and Martin Ashton-Key. Role of feasibility and pilot studies in randomised controlled trials: a cross-sectional study. *BMJ Open*, 8:e022233, 09 2018. doi: 10.1136/bmjopen-2018-022233.
- [19] Tim Bock. What is a correlation matrix? URL <https://www.displayr.com/what-is-a-correlation-matrix/>.
- [20] Anxiety Canada. New mindshift™ cbt app gives Canadians free anxiety relief, 2019. URL <https://anxietycanada.com/articles/new-mindshift-cbt-app-gives-canadians-free-anxiety-relief/>.
- [21] IBM Knowledge Center. Nodo bosque aleatorio, 2019. URL https://www.ibm.com/support/knowledgecenter/es/SS3RA7_sub/modeler_mainhelp_client_ddita/clementine/python_nodes_rf.html.
- [22] Universidad Católica Centro del Sueño. Trastorno del sueño, 2019. URL <http://redsalud.uc.cl/ucchristus/centro-del-sueno/trastornos-del-sueno/>.
- [23] Larry Chan, Vedant Das Swain, Christina Kelley, Kaya Barbaro, Gregory Abowd, and Lauren Wilcox. Students’ experiences with ecological momentary assessment tools to report on emotional well-being. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2:1–20, 03 2018. doi: 10.1145/3191735.
- [24] Desirée Colombo, Javier Fernández-Álvarez, Andrea Patané, Michelle Semonella, Marta Kwiatkowska, Azucena García-Palacios, Pietro Cipresso, Giuseppe Riva, and Cristina

- Botella. Current state and future directions of technology-based ecological momentary assessment and intervention for major depressive disorder: A systematic review. *Journal of Clinical Medicine*, 8(4), 2019. ISSN 2077-0383. doi: 10.3390/jcm8040465. URL <https://www.mdpi.com/2077-0383/8/4/465>.
- [25] Andrew Cucchiniara. Applied logistic regression. *Technometrics*, 34:358–359, 03 2012. doi: 10.1080/00401706.1992.10485291.
- [26] Vicerrectoría de Asuntos Académicos. Informe sobre programas de acompañamiento académico y psicológico. Technical report, Universidad de Chile, 2019.
- [27] Universidad de Chile. Hechos y cifras, 2019. URL <http://escuela.ingenieria.uchile.cl/la-escuela/hechos-y-cifras>.
- [28] Organización Mundial de la Salud. ¿qué es la salud?, 2019. URL <https://www.who.int/es/about/who-we-are/frequently-asked-questions>.
- [29] Organización Mundial de la Salud. Salud mental: fortalecer nuestra respuesta, 2019.
- [30] Naelys Diaz, E. Gail Horton, and Tammy Malloy. Attachment style, spirituality, and depressive symptoms among individuals in substance abuse treatment. *Journal of Social Service Research*, 40(3):313–324, 2014. doi: 10.1080/01488376.2014.896851. URL <https://doi.org/10.1080/01488376.2014.896851>.
- [31] Bragança M. Dinis, J. Quality of sleep and depression in college students: A systematic review. *Sleep science (Sao Paulo, Brazil)*, 11(4):290–301, 2018. doi: 10.5935/1984-0063.20180045. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6361309/>.
- [32] Reza Entezari-Maleki, Arash Rezaei, and Behrouz Minaei. Comparison of classification methods based on the type of attributes and sample size. *JCIT*, 4:94–102, 09 2009. doi: 10.4156/jcit.vol4.issue3.14.
- [33] Molina et al. Calidad de vida relacionada con la salud en estudiantes universitarios de primer año, 2016. URL <https://n9.cl/13ei>.
- [34] Salgado Madrid et al. Escalas para estudiar percepción de estrés psicológico en el climaterio. *Revista Ciencias Biomédicas*, page 323, 7 2013. URL https://www.academia.edu/30510654/ESCALAS_PARA_ESTUDIAR_PERCEPCION_DE_ESTRES_PSICOLOGICO_EN_EL_CLIMATERIO.
- [35] UCH. Facultad de Economía y Negocios. Estudiantes se titularon a los 28,8 años, la edad promedio más alta desde 2007, 2015. URL <http://www.economiaynegocios.cl/noticias/noticias.asp?id=193890>.
- [36] FCFM. Unidad de calidad de vida, 2019. URL <http://ingenieria.uchile.cl/admision/vida-en-beauchef/89970/unidad-de-calidad-de-vida-estudiantil>.
- [37] FCFM. Unidad de bienestar estudiantil, 2019. URL <http://ingenieria.uchile.cl/facultad/estructura/97871/unidad-de-bienestar-estudiantil>.
- [38] FCFM. Unidad de orientación vocacional, 2019. URL <http://escuela.ingenieria.uchile.cl/noticias/fcfm-cuenta-con-orientador-vocacional-para-estudiantes-de-pregrado>.
- [39] Hugo Figueiredo-Ferraz, Sara Cardona, and Pedro Gil-Monte. Desgaste psíquico y problemas de salud en estudiantes de psicología. *Psicología em Estudo*, 14:349 – 353, 06

2009. ISSN 1413-7372. URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-73722009000200016&nrm=iso.
- [40] Google. Clasificación: Roc y auc. URL <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=es>.
- [41] John Hettema. What is the genetic relationship between anxiety and depression? *American journal of medical genetics. Part C, Seminars in medical genetics*, 148C:140–6, 05 2008. doi: 10.1002/ajmg.c.30171.
- [42] Saul Shiffman; Arthur A. Stone; Michael R. Hufford. Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 4 2008. URL <https://ssrn.com/abstract=1141344>.
- [43] Ahmed Ibrahim, Shona Kelly, Clive Adams, and Cris Glazebrook. A systematic review of studies of depression prevalence in university students. *Journal of psychiatric research*, 47, 12 2012. doi: 10.1016/j.jpsychires.2012.11.015.
- [44] ImHay. Núcleo milenio imhay, 2019. URL <https://www.imhay.org/>.
- [45] Jian Pei Jiawei Han, Micheline Kamber. *Data Mining, Concepts and Techniques*, volume 3. 2011. URL <https://n9.cl/7u4m>.
- [46] Graciela; Martínez Vania Jiménez Molina, Álvaro; Rojas. Problemas de salud mental en estudiantes universitarios (i): ¿consecuencias de la (sobre) carga académica?, 2019. URL <https://n9.cl/13ow>.
- [47] Ian Jolliffe and Jorge Cadima. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374:20150202, 04 2016. doi: 10.1098/rsta.2015.0202.
- [48] Kurt Kroenke, Robert Spitzer, Janet Williams, and Bernd Löwe. An ultra-brief screening scale for anxiety and depression: The phq-4. *Psychosomatics*, 50:613–21, 11 2009. doi: 10.1176/appi.psy.50.6.613.
- [49] Harvard Mental Health Letter. Sleep and mental health, 2009. URL https://www.health.harvard.edu/newsletter_article/sleep-and-mental-health.
- [50] Y.M. Lin, S.C. Lin, M.Y. Wang, and F.S. Chen. What causes the academic stress suffered by students at universities and colleges of technology? 7:77–81, 01 2009.
- [51] National Institute of Mental Health. Technology and the future of mental health treatment, 2019. URL <https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment/index.shtml>.
- [52] Burkham R. Oliver, J. M. Depression in university students: Duration, relation to calendar time, prevalence, and demographic correlates. *Journal of Abnormal Psychology*, 88:667–670, 1979. URL <http://dx.doi.org/10.1037/0021-843X.88.6.667>.
- [53] Medline Plus. Estrés y salud, 2019. URL <https://medlineplus.gov/spanish/ency/article/003211.htm>.
- [54] Medline Plus. Salud mental, 2019. URL <https://medlineplus.gov/spanish/mentalhealth.html>.
- [55] Psyberguide. Moodpath app – depression, burnout & anxiety test, 2019. URL <https://>

[//psyberguide.org/apps/moodpath-app-depression-burnout-anxiety-test/](https://psyberguide.org/apps/moodpath-app-depression-burnout-anxiety-test/).

- [56] Cristhian Pérez-Villalobos, Carmen Bonnefoy-Dibarrat, Adherys Cabrera-Flores, Shila Peine-Grandón, Katerina-Macaya Abarca, Marjorie Baqueano-Rodríguez, and Jorge Jiménez-Espinoza. Mental health problems of freshmen in concepción, chile. *Anales de Psicología / Annals of Psychology*, 28(3):797–804, 1 1. doi: 10.6018/analesps.28.3.156071. URL <https://revistas.um.es/analesps/article/view/analesps.28.3.156071>.
- [57] Héctor Gabriel Acosta Mesa Ivonne Rabatte Suárez María del Carmen Gogeoascoechea Trejo Patricia Pavón León Sobeida L. Blázquez Morales Rocío Erandi Barrientos Martínez, Nicandro Cruz Ramírez. Árboles de decisión como herramienta en el diagnóstico médico. *Revista Médica de la Universidad Veracruzana*, 2:19–24, 09 2009. URL <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=27872>.
- [58] Ana María Roura. Protestas en chile: las grietas del modelo económico chileno que las manifestaciones dejaron al descubierto, Octubre 2019. URL <https://www.bbc.com/mundo/noticias-america-latina-50208871>.
- [59] ShanghaiRanking. Academic ranking of world universities 2019, 2019. URL <http://www.shanghai ranking.com/arwu2019.html>.
- [60] John Torous, Rohn Friedman, and Matcheri Keshavan. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth uHealth*, 2(1):e2, Jan 2014. ISSN 2291-5222. doi: 10.2196/mhealth.2994. URL <http://mhealth.jmir.org/2014/1/e2/>.
- [61] Red Salud UC. Depresión, 2019. URL <http://redsalud.uc.cl/ucchristus/VidaSaludable/Glosario/D/depresion.act>.
- [62] Red Salud UC. Trastorno de ansiedad generalizada, 2019. URL http://redsalud.uc.cl/ucchristus/VidaSaludable/glosario/A/trastorno_de_ansiedad_generalizada.act.
- [63] Antúñez; Vinet. Escalas de depresión, ansiedad y estrés (dass - 21): Validación de la versión abreviada en estudiantes universitarios chilenos, 2012. URL https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-48082012000300005.
- [64] Véliz. Validación escala de bienestar psicológico de ryff, 2008. URL <http://celappa.com/wp-content/uploads/2015/11/Validación-Escala-de-Bienestar-Psicológico-de-Ryff-véliz-2008.pdf>.
- [65] Hadas Wiseman, Daniel G. Guttfreund, and Itamar Lurie. Gender differences in loneliness and depression of university students seeking counselling. *British Journal of Guidance & Counselling*, 23(2):231–243, 1995. doi: 10.1080/03069889508253008. URL <https://doi.org/10.1080/03069889508253008>.

Anexo A

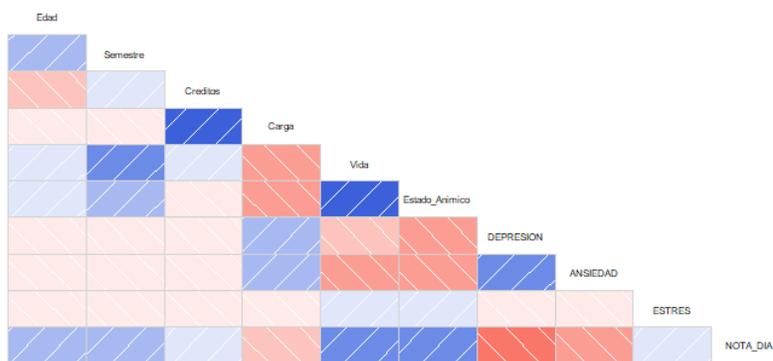
Matrices de correlación

Matriz de Correlación jornada día



(a) Jornada Día

Matriz de Correlación jornada tarde



(b) Jornada Tarde

Figura A.1: Matriz de correlación por jornada

Anexo B

Resultados modelos con MC

Tabla B.1: Depresión día con Regresión Logística - MC

DepresionD_MC_Logit	Estimate	Std.	Error	z	
(Intercept)	7.08e+00	2.07e+00	3.43	0.00061	***
Genero_2	-4.41e-01	5.30e-01	-0.83	0.40526	
RM	-5.17e+00	1.09e+00	-4.73	2.3e-06	***
Convivencia_1	-2.38e+00	1.30e+00	-1.83	0.06699	.
Convivencia_2	-3.87e+00	1.27e+00	-3.05	0.00227	**
CONTEXTO_HACIENDO_1	-1.09e+00	1.01e+00	-1.08	0.28037	
CONTEXTO_HACIENDO_2	-1.59e+00	8.18e-01	-1.95	0.05110	.
CONTEXTO_HACIENDO_3	-2.88e-01	5.95e-01	-0.48	0.62872	
CONTEXTO_HACIENDO_5	-1.28e-01	8.91e-01	-0.14	0.88541	
CONTEXTO_COMPANIA_2	-2.97e+00	2.48e+00	-1.20	0.22964	
CONTEXTO_COMPANIA_3	-1.74e+00	2.51e+00	-0.69	0.48891	
CONTEXTO_COMPANIA_4	-1.26e+00	9.29e-01	-1.36	0.17394	
CONTEXTO_COMPANIA_5	-1.30e+00	1.13e+00	-1.15	0.25016	
CONTEXTO_COMPANIA_6	-4.54e-01	3.33e+00	-0.14	0.89159	
CONTEXTO_COMPANIA_7	5.62e-02	1.03e+00	0.05	0.95633	
CONTEXTO_COMPANIA_8	3.86e+01	8.08e+03	0.00	0.99619	
AnoIngreso_2015	-4.07e-01	7.02e-01	-0.58	0.56164	
AnoIngreso_2016	1.33e+02	1.21e+04	0.01	0.99125	
AnoIngreso_2018	3.59e+00	8.20e-01	4.37	1.2e-05	***
AnoIngreso_2017	2.72e+00	8.38e-01	3.25	0.00115	**
ANSIEDAD	2.67e+00	5.51e-01	4.84	1.3e-06	***
Carga	2.08e-01	4.62e-01	0.45	0.65306	
SUENO_CALIDAD	-6.76e-01	1.93e-01	-3.50	0.00046	***
Estado_Animico	-1.59e+00	4.06e-01	-3.91	9.3e-05	***
Semestre	4.34e-01	1.17e-01	3.72	0.00020	***
Vida	-6.11e-02	2.61e-01	-0.23	0.81501	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla B.2: Depresión tarde con Regresión Logística - MC

DepresionT_MC_Logit	Estimate	Std.	Error	z value
(Intercept)	2.44e+02	5.83e+06	0	1
Genero_1	-7.86e+00	8.86e+05	0	1
RM	-3.55e+01	9.97e+05	0	1
Convivencia_1	-1.37e+01	1.57e+06	0	1
Convivencia_4	-1.09e+02	4.42e+05	0	1
Convivencia_2	2.08e+02	3.76e+06	0	1
Carrera_9	1.61e+02	4.55e+06	0	1
Carrera_7	1.23e+02	4.45e+06	0	1
Carrera_5	9.31e+01	4.06e+06	0	1
Carrera_14	6.49e+01	4.26e+06	0	1
Carrera_4	-1.54e+02	3.93e+06	0	1
Carrera_11	9.88e+01	5.03e+06	0	1
Carrera_13	1.27e+02	4.26e+06	0	1
Carrera_3	1.20e+02	5.17e+06	0	1
CONTEXTO_HACIENDO_1	-1.02e+00	1.76e+05	0	1
CONTEXTO_HACIENDO_2	2.39e+01	5.53e+05	0	1
CONTEXTO_HACIENDO_5	-6.50e+01	1.83e+05	0	1
CONTEXTO_HACIENDO_4	1.47e+01	1.50e+06	0	1
CONTEXTO_COMPANIA_2	1.29e+02	6.14e+05	0	1
CONTEXTO_COMPANIA_3	NA	NA	NA	NA
CONTEXTO_COMPANIA_4	-1.45e+00	3.47e+05	0	1
CONTEXTO_COMPANIA_5	-6.99e+01	1.06e+06	0	1
CONTEXTO_COMPANIA_6	-1.84e+02	3.91e+05	0	1
CONTEXTO_COMPANIA_7	2.22e+01	3.26e+05	0	1
CONTEXTO_COMPANIA_8	1.13e+02	3.81e+05	0	1
AnoIngreso_2015	-2.81e+01	3.26e+05	0	1
AnoIngreso_2016	6.07e+01	2.82e+06	0	1
AnoIngreso_2018	-1.26e+00	2.52e+05	0	1
AnoIngreso_2017	-1.04e+02	5.05e+05	0	1
ANSIEDAD	9.04e-01	1.40e+05	0	1
Carga	-3.13e+01	4.62e+05	0	1
NOTA_DIA	-4.05e+01	2.12e+04	0	1
Estado_Animico	-4.10e+01	4.32e+05	0	1
Semestre	4.04e+00	1.10e+05	0	1
Vida	9.39e+00	1.03e+05	0	1

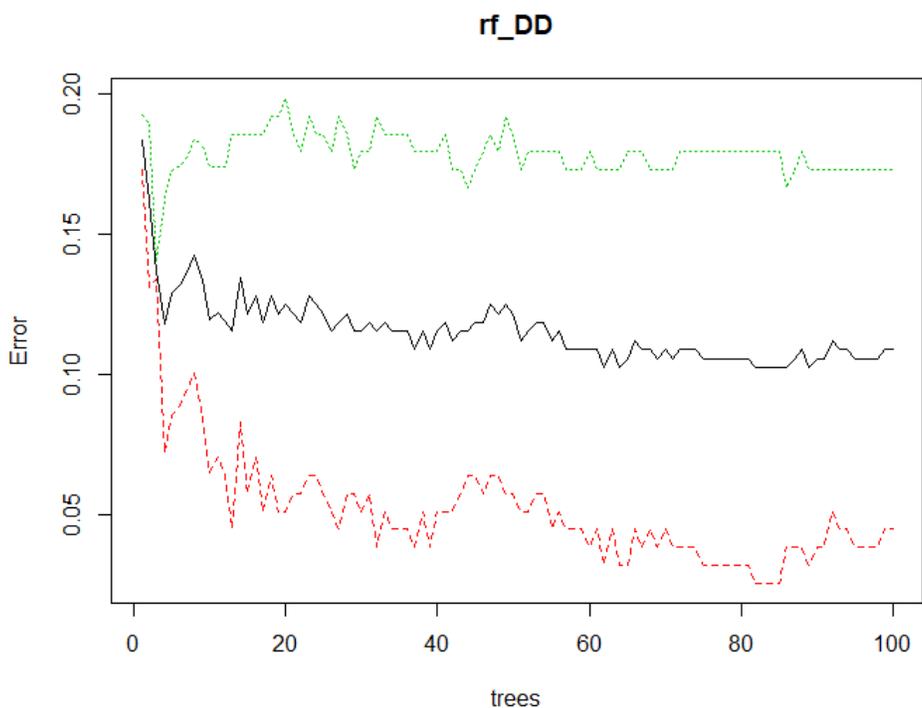


Figura B.1: Depresión día con Bosques Aleatorios - MC

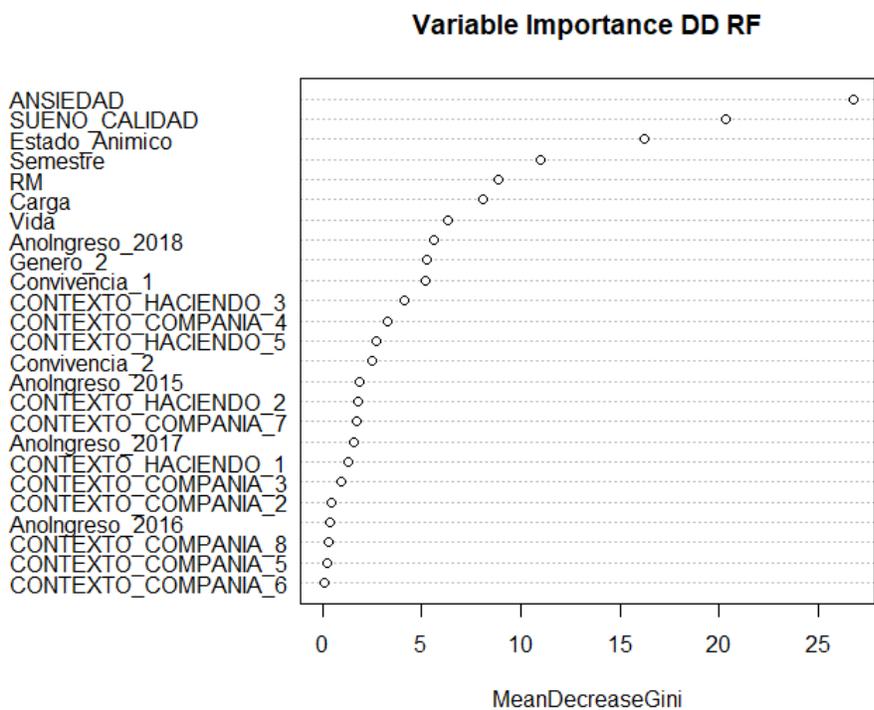


Figura B.2: Importancia de variables para Depresión día con Bosques Aleatorios - MC

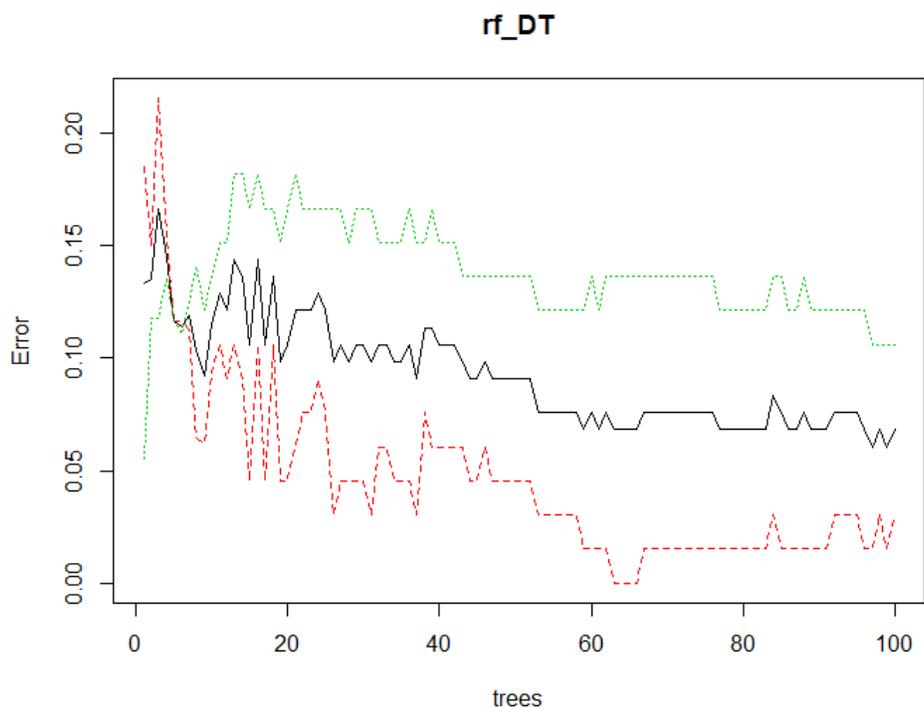


Figura B.3: Depresión tarde con Bosques Aleatorios - MC

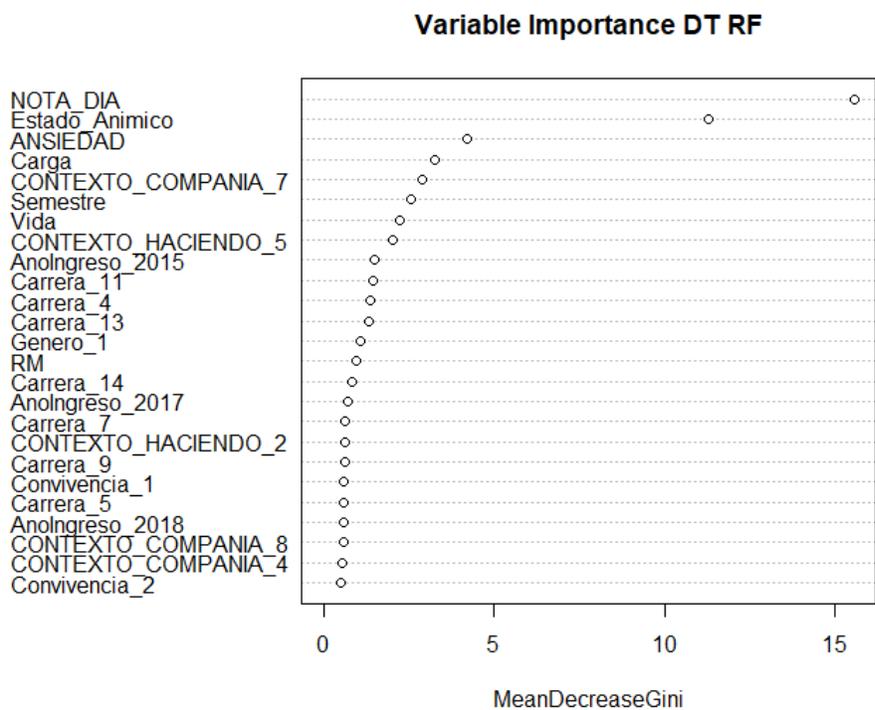


Figura B.4: Importancia de variables para Depresión tarde con Bosques Aleatorios - MC

Árbol Clasificación DD

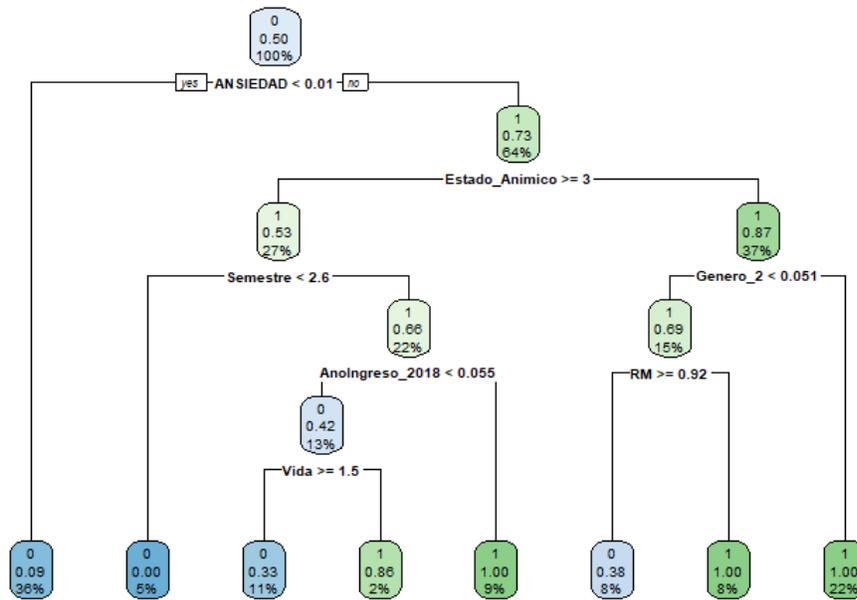


Figura B.5: Depresión día con Árboles de Clasificación - MC

Árbol Clasificación DT

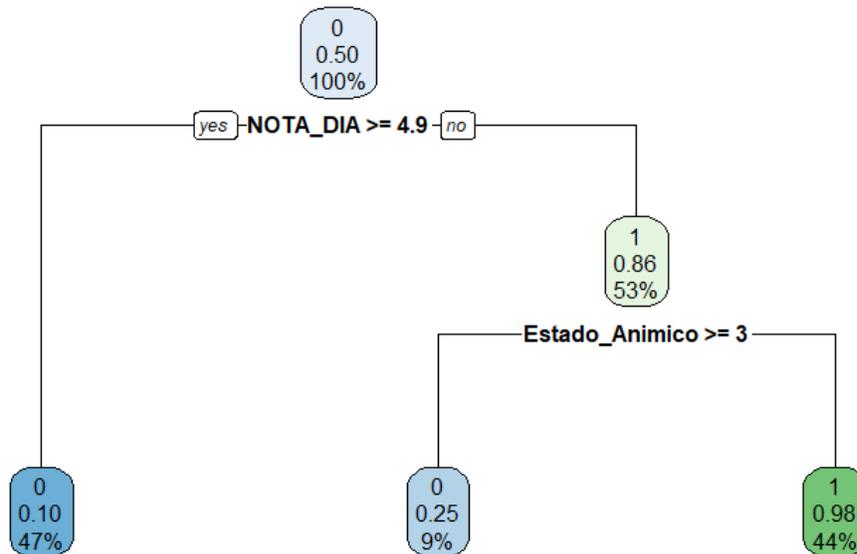


Figura B.6: Depresión tarde con Árboles de Clasificación - MC

Tabla B.3: Ansiedad día con Regresión Logística - MC

AnsiedadD_MC_Logit	Estimate	Std.	Error	z	
(Intercept)	14.373	23.094	0.62	0.53369	
Genero_2	12.885	0.7833	1.65	0.09996	.
RM	28.112	0.9277	3.03	0.00244	**
Convivencia_1	62.982	17.843	3.53	0.00042	***
Convivencia_2	65.113	23.027	2.83	0.00469	**
Carrera_8	-10.158	13.527	-0.75	0.45268	
Carrera_7	-34.303	0.9976	-3.44	0.00058	***
Carrera_5	-20.645	12.718	-1.62	0.10454	
Carrera_14	-30.517	11.308	-2.70	0.00696	**
Carrera_4	-42.170	20.813	-2.03	0.04275	*
Carrera_9	-11.130	0.8347	-1.33	0.18240	
Carrera_13	-28.508	12.084	-2.36	0.01832	*
Carrera_3	-0.7604	13.139	-0.58	0.56278	
CONTEXTO_HACIENDO_1	-0.8509	0.7199	-1.18	0.23725	
CONTEXTO_HACIENDO_2	-0.6175	0.7873	-0.78	0.43282	
CONTEXTO_HACIENDO_3	-0.8221	0.5174	-1.59	0.11208	
CONTEXTO_HACIENDO_5	13.159	0.7414	1.77	0.07591	.
DEPRESION	16.931	0.5375	3.15	0.00163	**
Carga	0.7771	0.5324	1.46	0.14438	
SUENO_CALIDAD	-0.3648	0.1780	-2.05	0.04045	*
Estado_Animico	-0.1294	0.3756	-0.34	0.73039	
Semestre	-0.1118	0.0934	-1.20	0.23138	
Vida	-0.5069	0.2791	-1.82	0.06931	.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla B.4: Ansiedad tarde con Regresión Logística - MC

AnsiedadT_MC_Logit					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	34.981	13.292	2.63	0.0085	**
Genero_1	0.3035	0.4737	0.64	0.5217	
RM	0.9865	0.6088	1.62	0.1052	
Convivencia_1	0.7524	0.8085	0.93	0.3521	
Convivencia_4	NA	NA	NA	NA	
Convivencia_2	24.125	15.198	1.59	0.1124	
Carrera_9	-0.7362	10.074	-0.73	0.4649	
Carrera_7	-14.711	0.8738	-1.68	0.0923	.
Carrera_5	-0.5177	0.9940	-0.52	0.6025	
Carrera_14	-0.4561	0.7535	-0.61	0.5450	
Carrera_4	0.2517	10.988	0.23	0.8188	
Carrera_11	-0.4338	0.9982	-0.43	0.6638	
Carrera_13	-0.4820	11.377	-0.42	0.6718	
Carrera_3	-169.279	8.850.304	-0.02	0.9847	
CONTEXTO_HACIENDO_1	0.7673	0.7410	1.04	0.3004	
CONTEXTO_HACIENDO_2	0.5849	0.5778	1.01	0.3114	
CONTEXTO_HACIENDO_5	-0.3628	0.4585	-0.79	0.4287	
CONTEXTO_HACIENDO_4	0.7238	0.5652	1.28	0.2004	
DEPRESION	0.9669	0.7316	1.32	0.1863	
Carga	0.1926	0.4248	0.45	0.6503	
NOTA_DIA	-0.4792	0.1479	-3.24	0.0012	**
Estado_Animico	-0.4710	0.2600	-1.81	0.0701	.
Semestre	-0.0226	0.0790	-0.29	0.7743	
Vida	-0.1392	0.2155	-0.65	0.5184	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

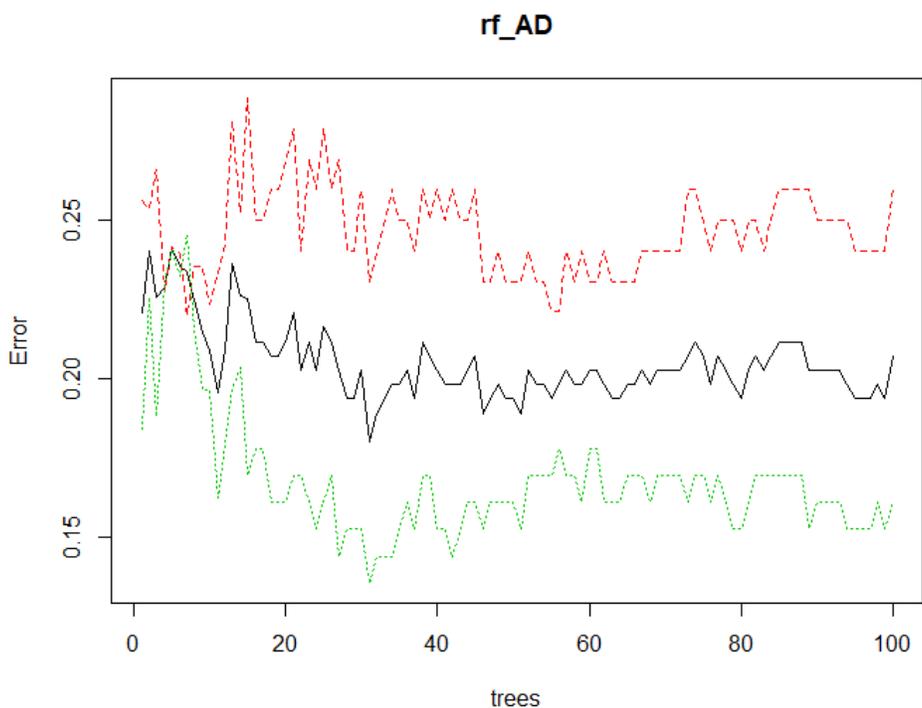


Figura B.7: Ansiedad día con Bosques Aleatorios - MC

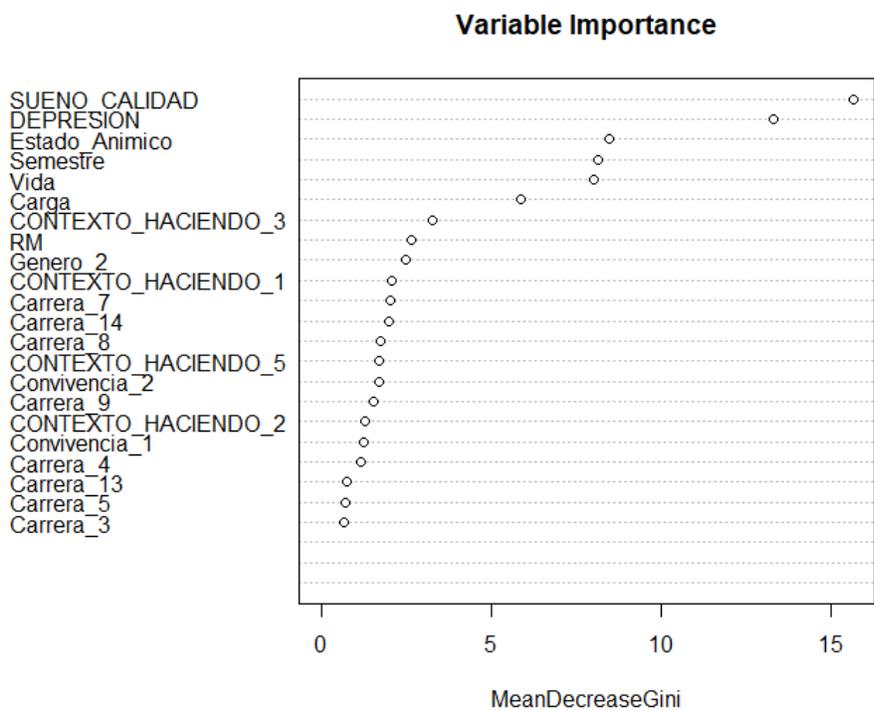


Figura B.8: Importancia de variables para Ansiedad día con Bosques Aleatorios

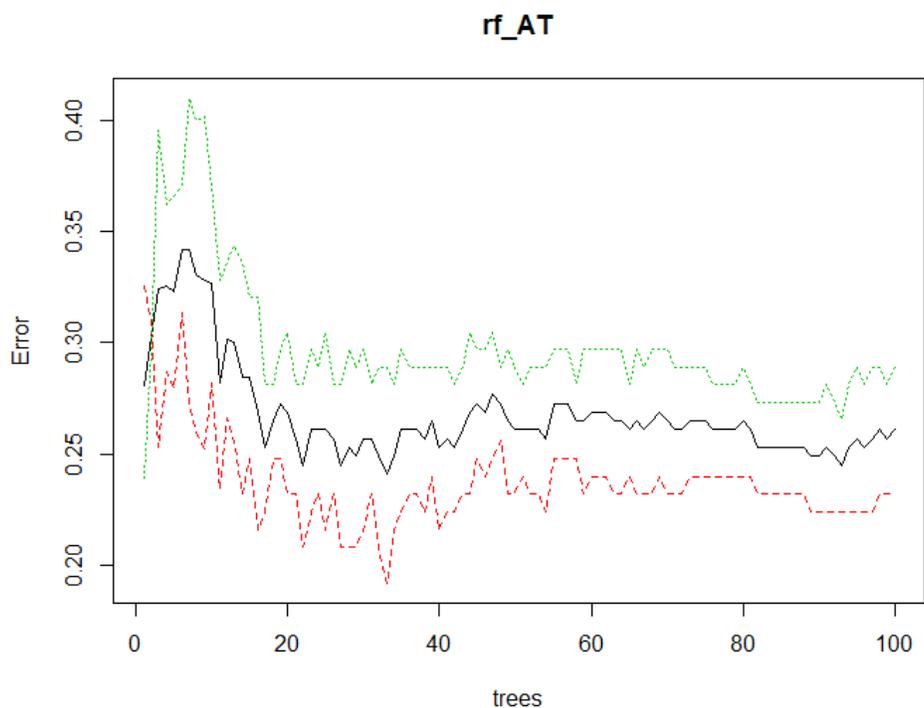


Figura B.9: Ansiedad tarde con Bosques Aleatorios - MC

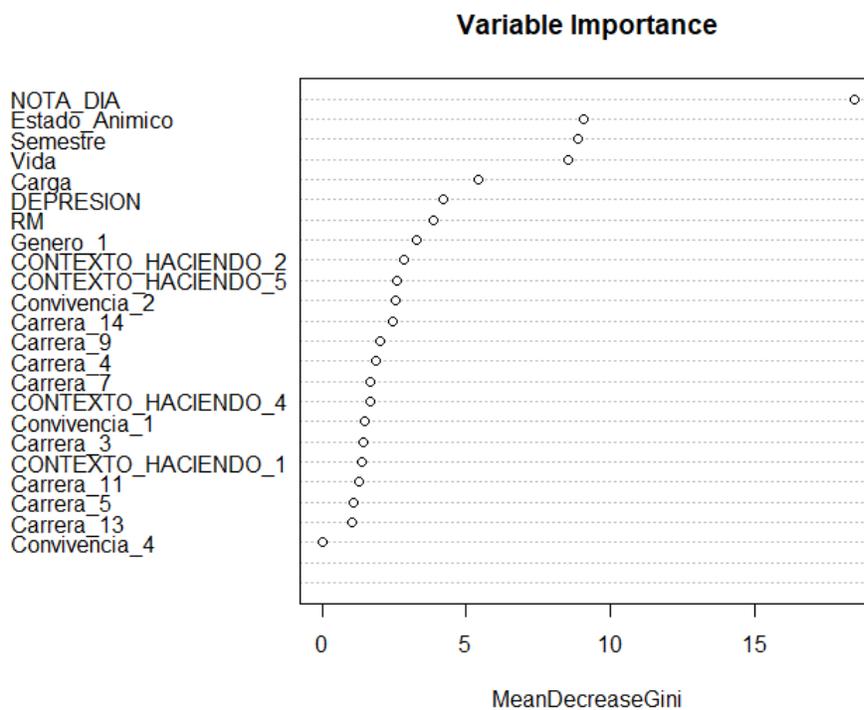


Figura B.10: Importancia de variables para Ansiedad tarde con Bosques Aleatorios - MC

Árbol Clasificación AD

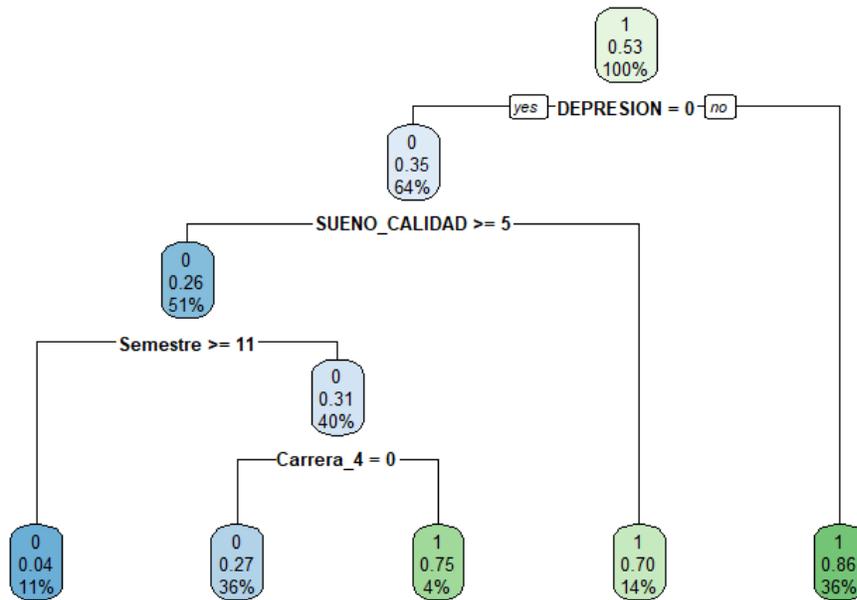


Figura B.11: Ansiedad día con Árboles de Clasificación - MC

Árbol Clasificación AT

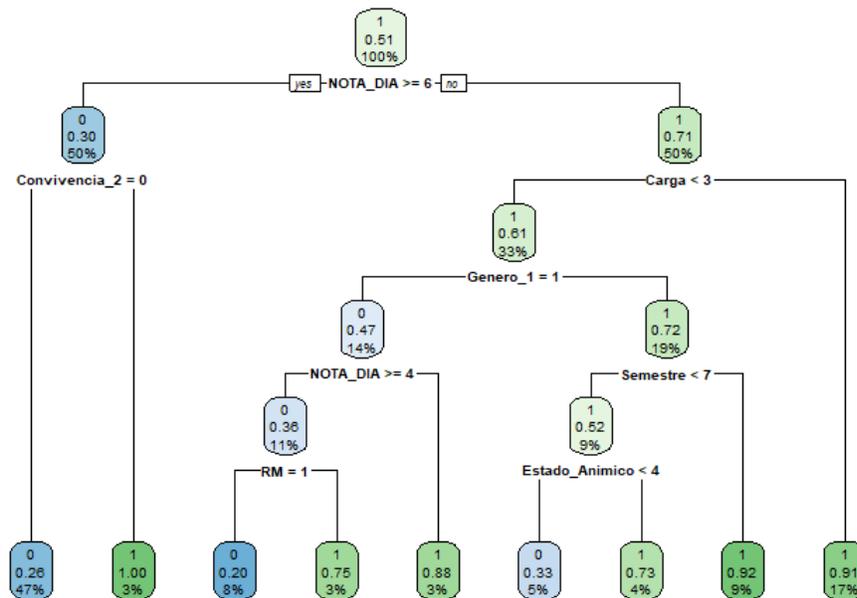


Figura B.12: Ansiedad tarde con Árboles de Clasificación - MC

Tabla B.5: Estrés día con Regresión Logística - MC

EstrésD_MC_Logit					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	24.477	11.834	2.07	0.03860	*
Genero_2	-0.3972	0.5099	-0.78	0.43604	
RM	-0.0973	0.5837	-0.17	0.86764	
Convivencia_1	0.2831	10.384	0.27	0.78513	
Convivencia_2	0.7858	11.828	0.66	0.50649	
Carrera_8	0.9902	0.8270	1.20	0.23115	
Carrera_7	-0.6025	0.7877	-0.76	0.44434	
Carrera_5	0.6280	0.9407	0.67	0.50438	
Carrera_14	-0.0534	0.7464	-0.07	0.94299	
Carrera_4	0.1780	0.9917	0.18	0.85753	
Carrera_9	17.131	0.6463	2.65	0.00803	**
Carrera_13	16.449	0.8869	1.85	0.06365	.
Carrera_3	-0.2143	0.8620	-0.25	0.80367	
CONTEXTO_HACIENDO_1	-0.7953	0.5622	-1.41	0.15715	
CONTEXTO_HACIENDO_2	-0.1282	0.6591	-0.19	0.84578	
CONTEXTO_HACIENDO_3	-0.5640	0.4206	-1.34	0.17990	
CONTEXTO_HACIENDO_5	0.1719	0.5378	0.32	0.74933	
ANSIEDAD	-0.1201	0.3630	-0.33	0.74071	
Vida	0.3203	0.1901	1.69	0.09193	.
SUENO_CALIDAD	-0.4718	0.1403	-3.36	0.00077	***
Creditos	-0.0271	0.0163	-1.67	0.09573	.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla B.6: Estrés tarde con Regresión Logística - MC

EstrésT_MC_Logit					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	59.622	25.824	2.31	0.02095	*
Genero_1	-0.4505	0.4670	-0.96	0.33470	
RM	-0.3466	0.5576	-0.62	0.53426	
Convivencia_1	-12.218	0.7910	-1.54	0.12241	
Convivencia_4	-156.509	10.256.911	-0.02	0.98783	
Convivencia_2	14.355	12.452	1.15	0.24899	
Carrera_9	-14.845	0.8575	-1.73	0.08340	.
Carrera_7	-29.786	0.8502	-3.50	0.00046	***
Carrera_5	-32.063	0.9557	-3.35	0.00079	***
Carrera_14	-16.017	0.8018	-2.00	0.04577	*
Carrera_4	-26.435	11.338	-2.33	0.01972	*
Carrera_11	-20.175	0.9171	-2.20	0.02781	*
Carrera_13	0.8167	11.189	0.73	0.46547	
Carrera_3	-25.297	10.927	-2.32	0.02061	*
CONTEXTO_HACIENDO_1	-0.2889	0.7602	-0.38	0.70390	
CONTEXTO_HACIENDO_2	10.543	0.5393	1.95	0.05059	.
CONTEXTO_HACIENDO_5	0.1452	0.4070	0.36	0.72124	
CONTEXTO_HACIENDO_4	0.6939	0.5575	1.24	0.21325	
Vida	0.2891	0.1894	1.53	0.12689	
NOTA_DIA	0.0100	0.1097	0.09	0.92725	
Edad	-0.1966	0.1042	-1.89	0.05920	.
Semestre	0.1149	0.0764	1.50	0.13272	
Creditos	-0.0260	0.0143	-1.82	0.06804	.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

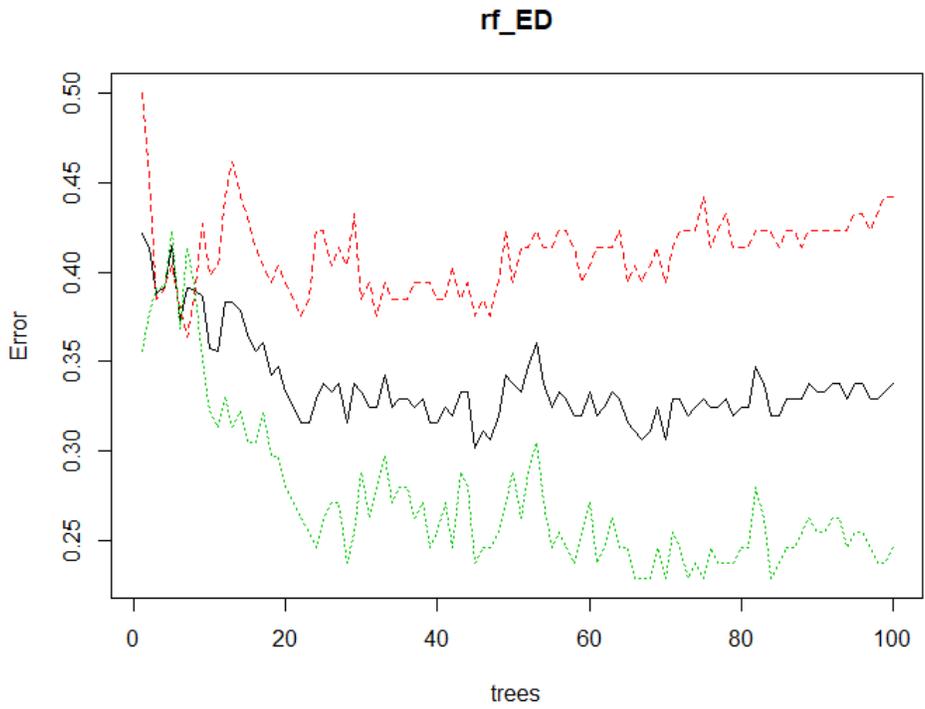


Figura B.13: Estrés día con Bosques Aleatorios - MC

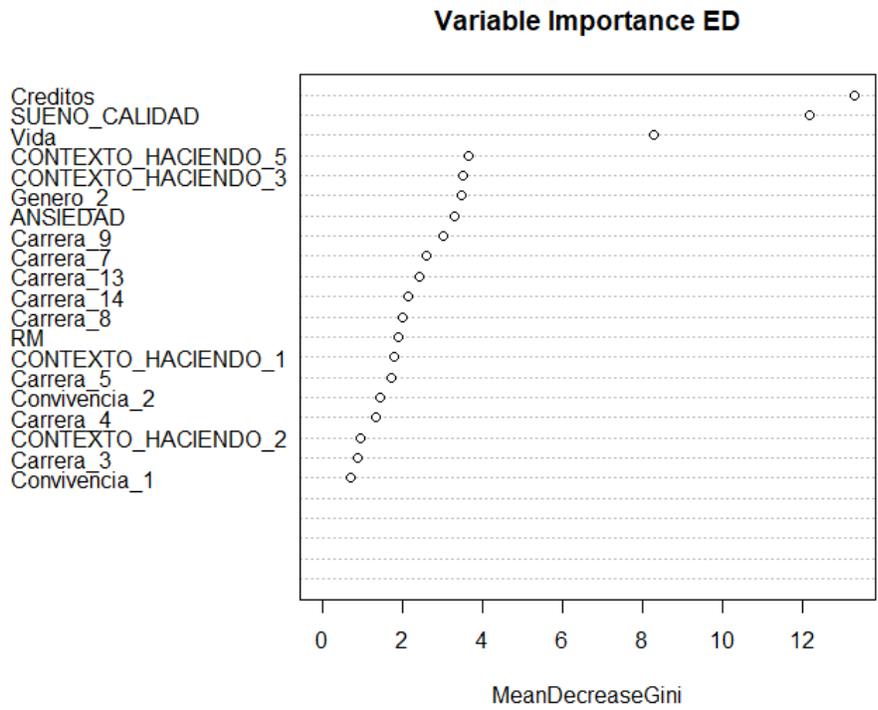


Figura B.14: Importancia de variables para Estrés día con Bosques Aleatorios - MC

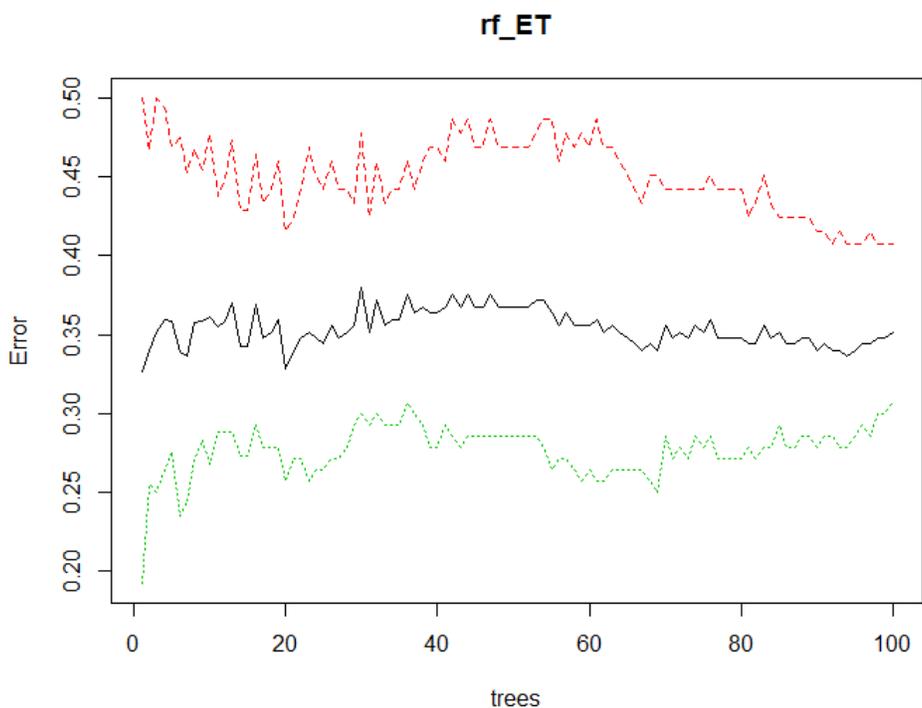


Figura B.15: Estrés tarde con Bosques Aleatorios - MC

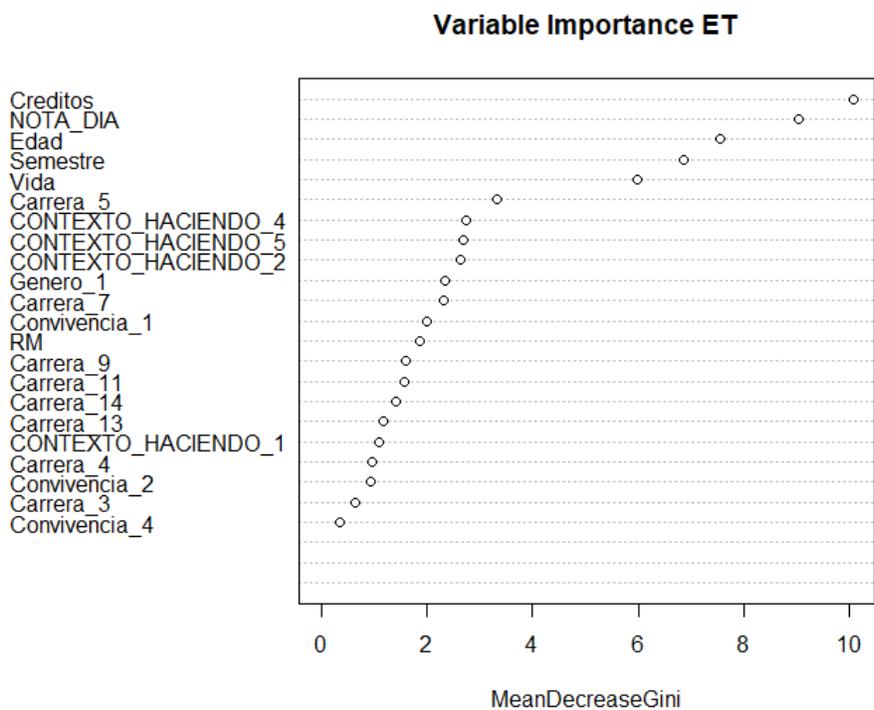


Figura B.16: Importancia de variables para Estrés tarde con Bosques Aleatorios - MC

Árbol Clasificación ED

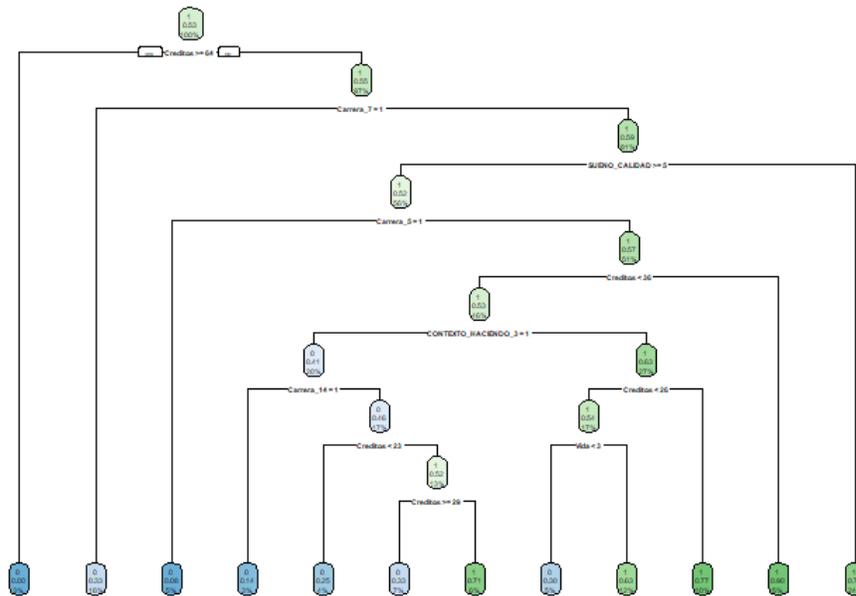


Figura B.17: Estrés día con Árboles de Clasificación - MC

Árbol Clasificación ET

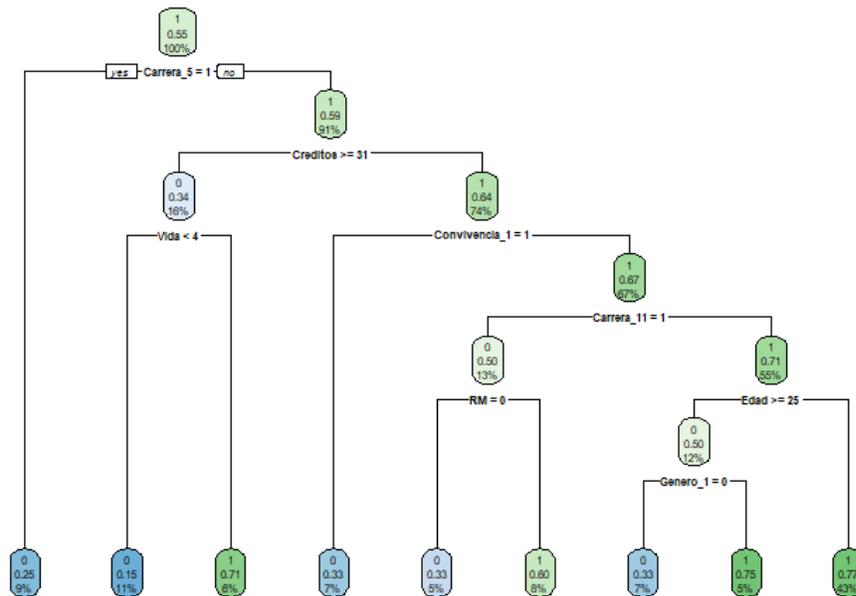


Figura B.18: Estrés tarde con Árboles de Clasificación - MC

Anexo C

Resultados modelos con PCA

Tabla C.1: Depresión día con Regresión Logística - PCA

DepresiónD_PCA_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	-0.495	0.312	-1.59	0.11190	
dim1	-1.154	0.330	-3.50	0.00047	***
dim2	0.483	0.199	2.42	0.01541	*
dim3	-0.311	0.166	-1.88	0.06029	.
dim4	0.086	0.166	0.52	0.60534	
dim5	-0.561	0.216	-2.59	0.00962	**
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla C.2: Depresión tarde con Regresión Logística - PCA

DepresiónT_PCA_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	-5.064	2.140	-2.37	0.018	*
dim1	-6.774	2.720	-2.49	0.013	*
dim2	-2.505	1.151	-2.18	0.029	*
dim3	-0.363	0.646	-0.56	0.574	
dim4	-1.192	0.577	-2.06	0.039	*
dim5	-0.918	0.839	-1.09	0.274	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

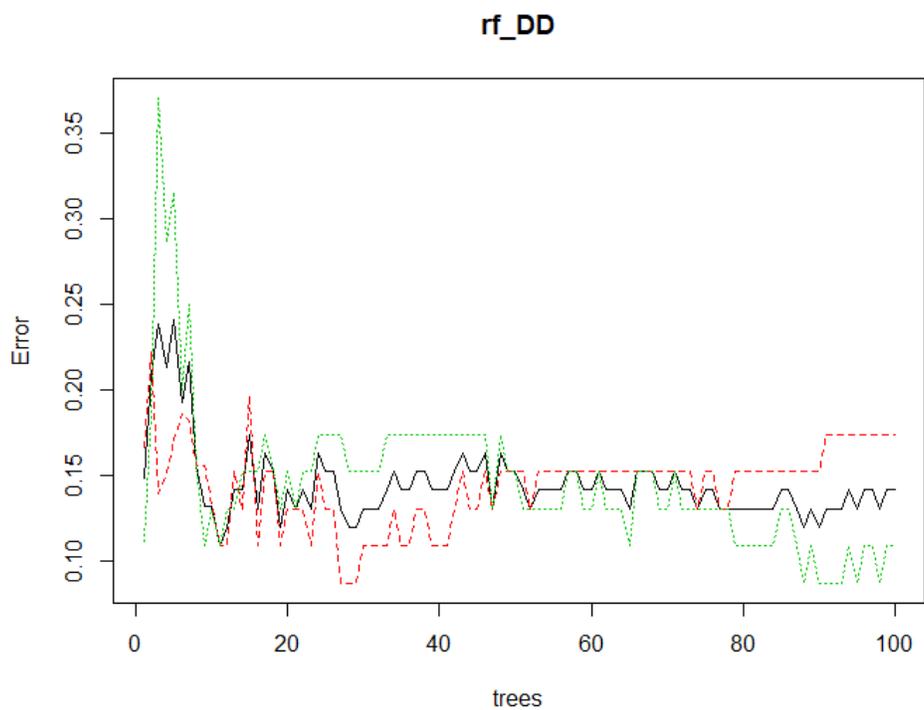


Figura C.1: Depresión día con Bosques Aleatorios - PCA

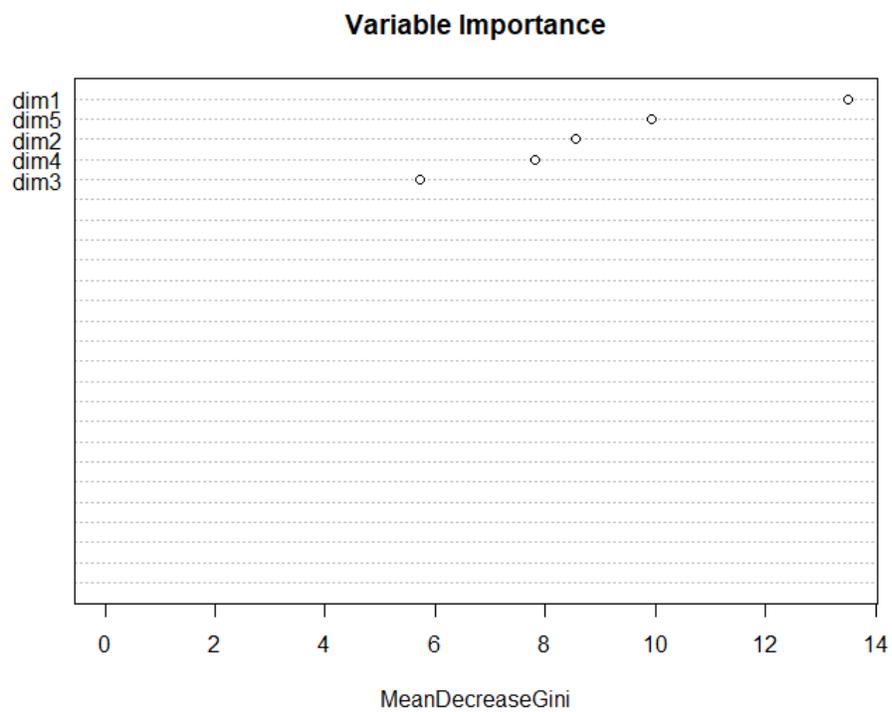


Figura C.2: Importancia de variables para Depresión día con Bosques Aleatorios - PCA

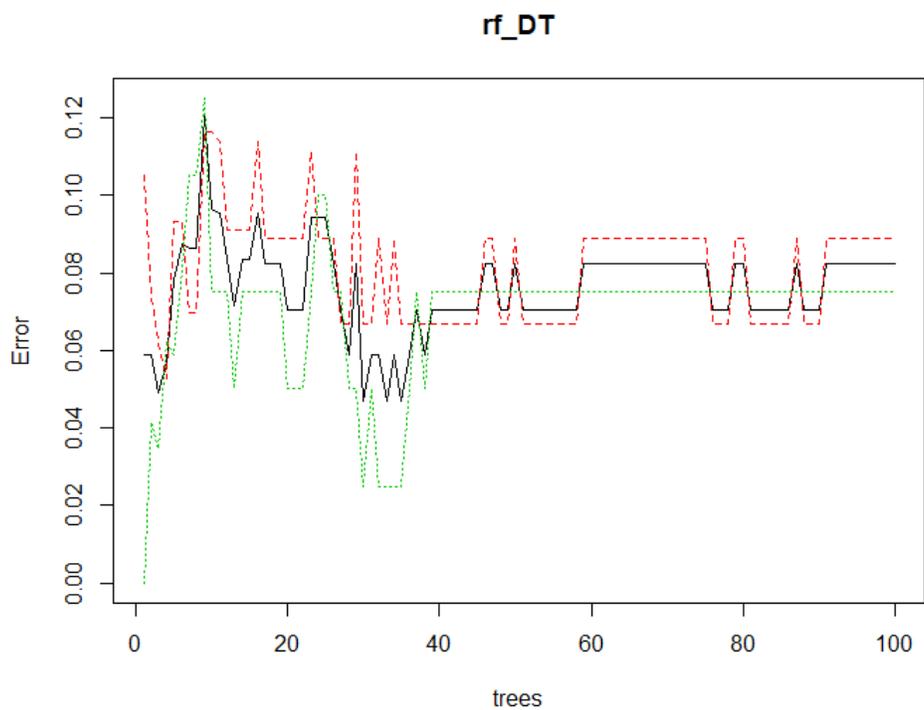


Figura C.3: Depresión tarde con Bosques Aleatorios - PCA

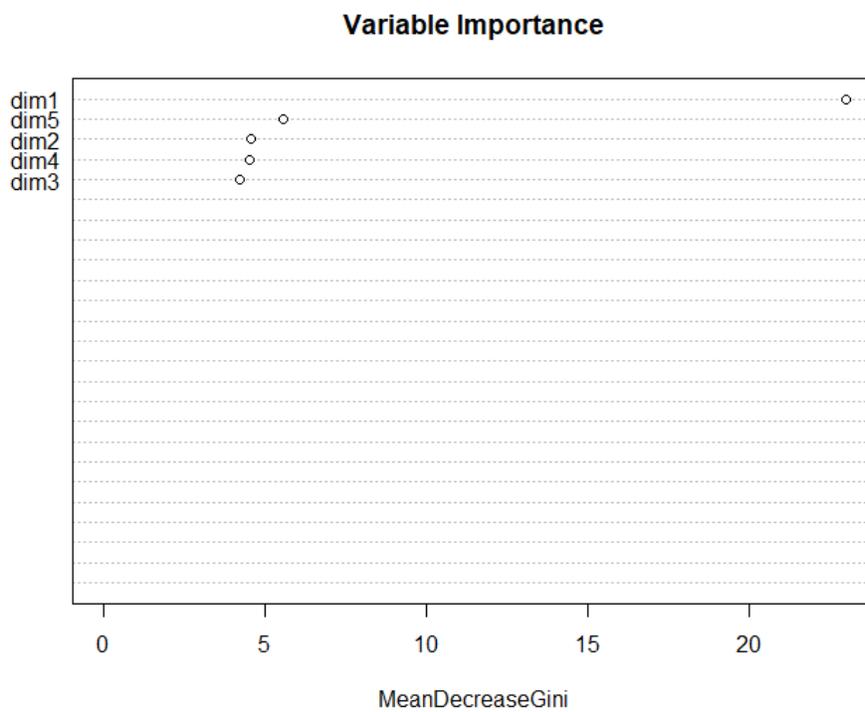


Figura C.4: Importancia de variables para Depresión tarde con Bosques Aleatorios - PCA

Árbol Clasificación DD

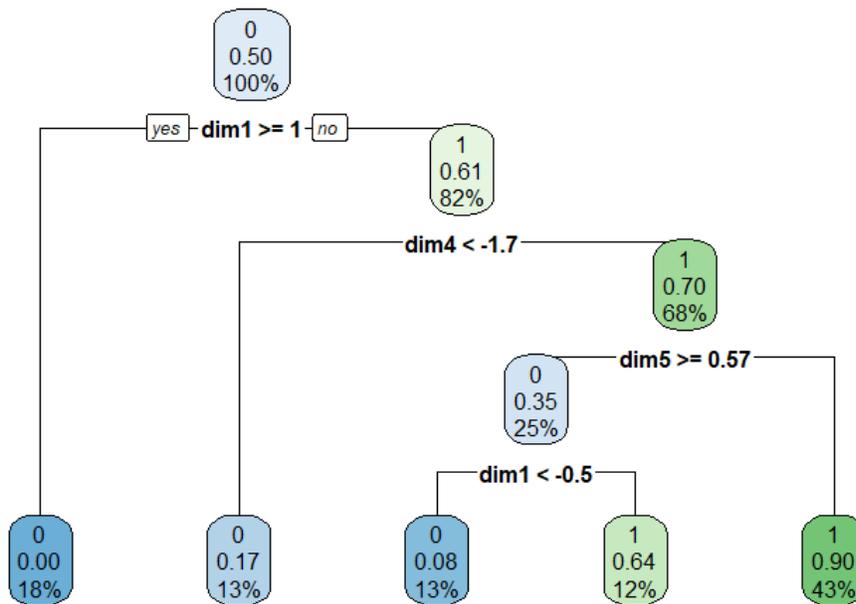


Figura C.5: Depresión día con Árboles de Clasificación - PCA

Árbol Clasificación DT

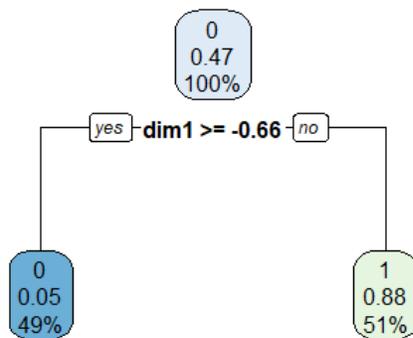


Figura C.6: Depresión tarde con Árboles de Clasificación - PCA

Tabla C.3: Ansiedad día con Regresión Logística - PCA

AnsiedadD_PCA_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	-1.378	0.547	-2.52	0.01182	*
dim1	-1.770	0.503	-3.52	0.00044	***
dim2	-0.943	0.371	-2.54	0.01095	*
dim3	0.583	0.355	1.64	0.10116	
dim4	0.082	0.355	0.23	0.81740	
dim5	-0.494	0.275	-1.80	0.07234	.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla C.4: Ansiedad tarde con Regresión Logística - PCA

AnsiedadT_PCA_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	-168.36	96741.85	0	1	
dim1	-64.08	34506.47	0	1	
dim2	-13.59	17971.31	0	1	
dim3	34.56	22602.51	0	1	
dim4	18.64	22161.28	0	1	
dim5	2.77	37989.86	0	1	

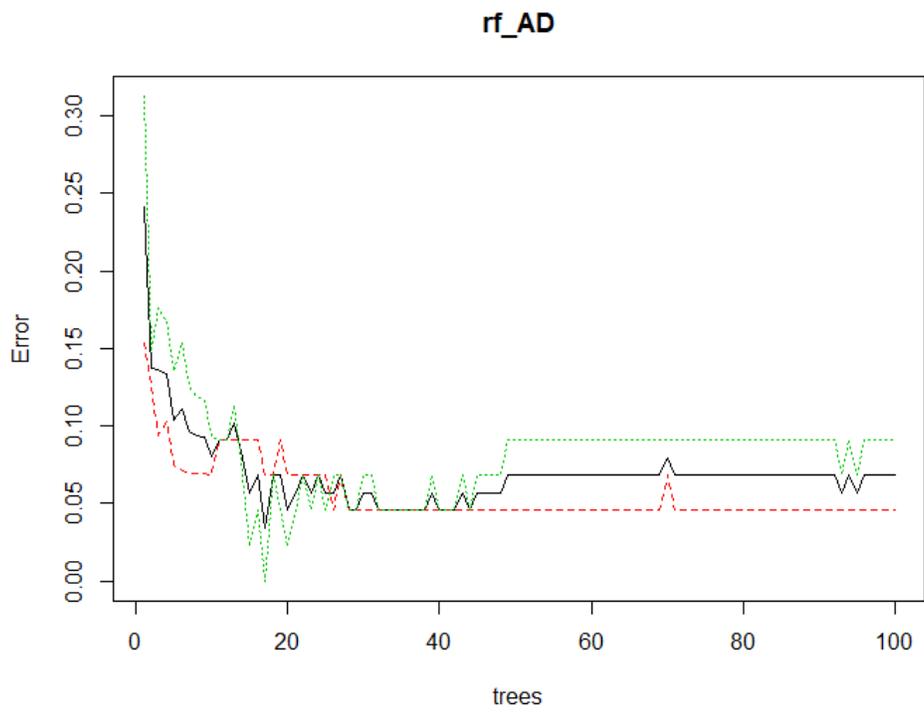


Figura C.7: Ansiedad día con Bosques Aleatorios - PCA

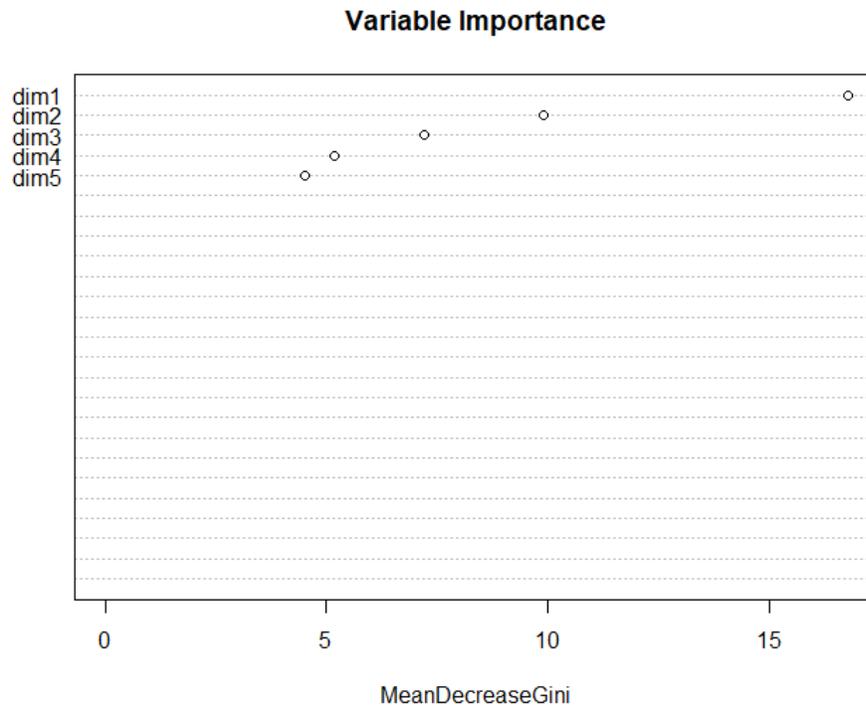


Figura C.8: Importancia de variables para Ansiedad día con Bosques Aleatorios - PCA

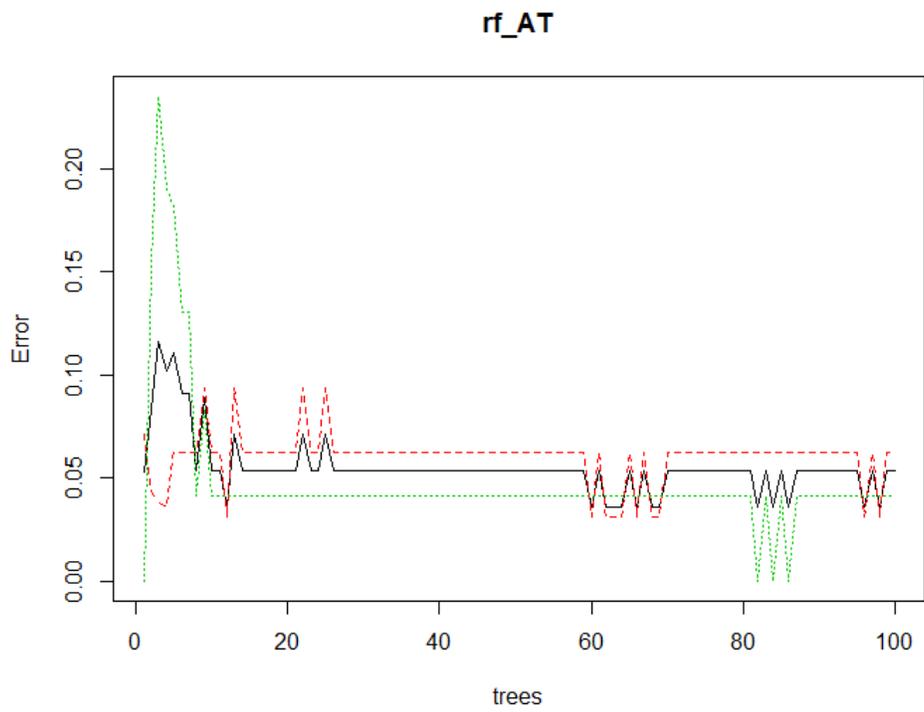


Figura C.9: Ansiedad tarde con Bosques Aleatorios - PCA

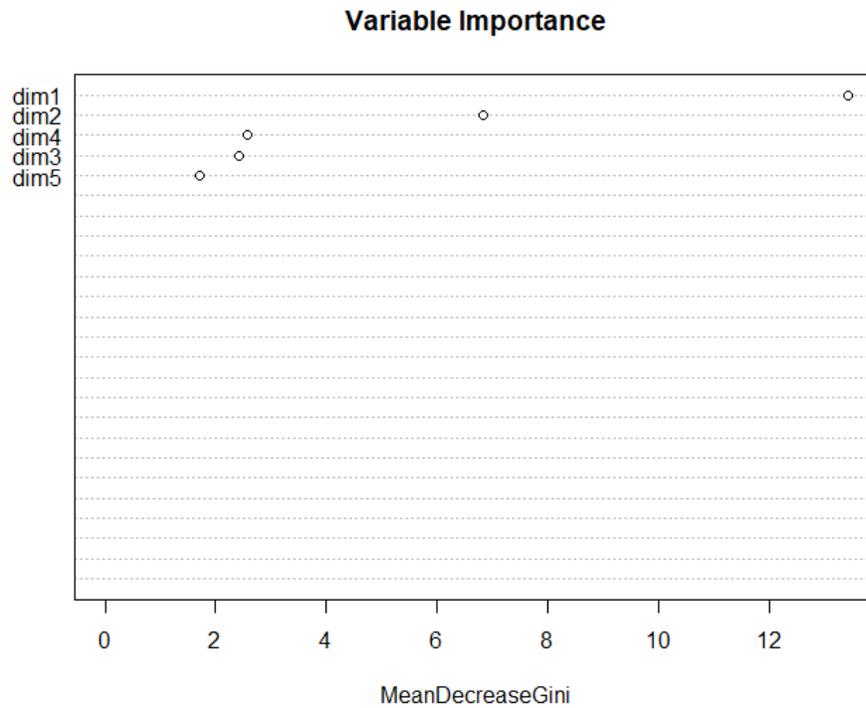


Figura C.10: Importancia de variables para Ansiedad tarde con Bosques Aleatorios - PCA

Árbol Clasificación AD

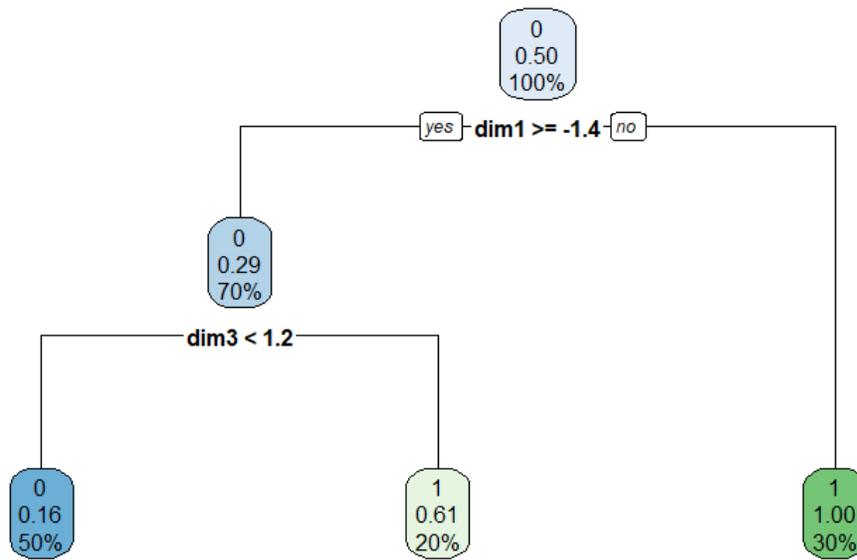


Figura C.11: Ansiedad día con Árboles de Clasificación - PCA

Árbol Clasificación AT

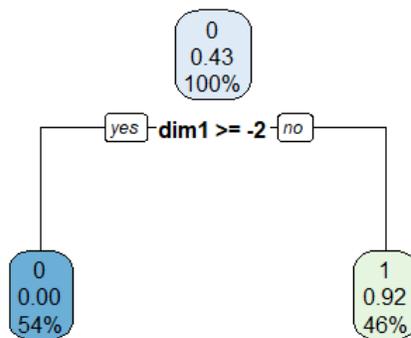


Figura C.12: Ansiedad tarde con Árboles de Clasificación - PCA

Tabla C.5: Estrés día con Regresión Logística - PCA

EstrésD_PCA_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	-0.520	0.408	-1.27	0.2025	
dim1	-1.878	0.451	-4.16	3.2e-05	***
dim2	0.233	0.215	1.08	0.2788	
dim3	-0.165	0.225	-0.73	0.4627	
dim4	-0.561	0.190	-2.95	0.0032	**
dim5	0.117	0.279	0.42	0.6756	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla C.6: Estrés tarde con Regresión Logística - PCA

EstrésT_PCA_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	-503	84.840	-0.01	1	
dim1	-306	51.676	-0.01	1	
dim2	-223	40273	-0.01	1	
dim3	205	35417	0.01	1	
dim4	221	37678	0.01	1	
dim5	-112	19847	-0.01	1	

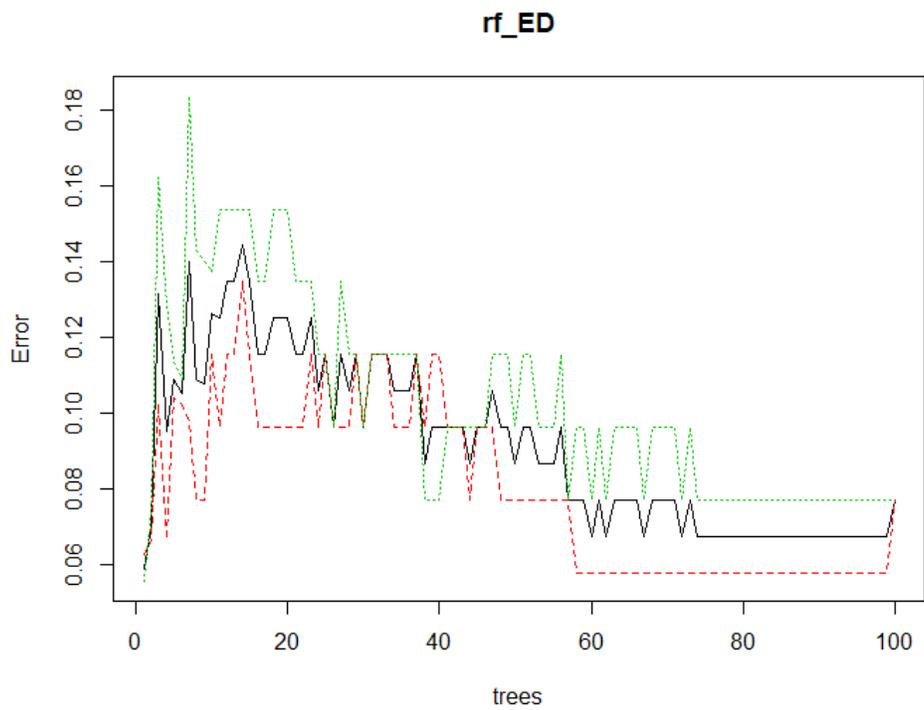


Figura C.13: Estrés día con Bosques Aleatorios - PCA

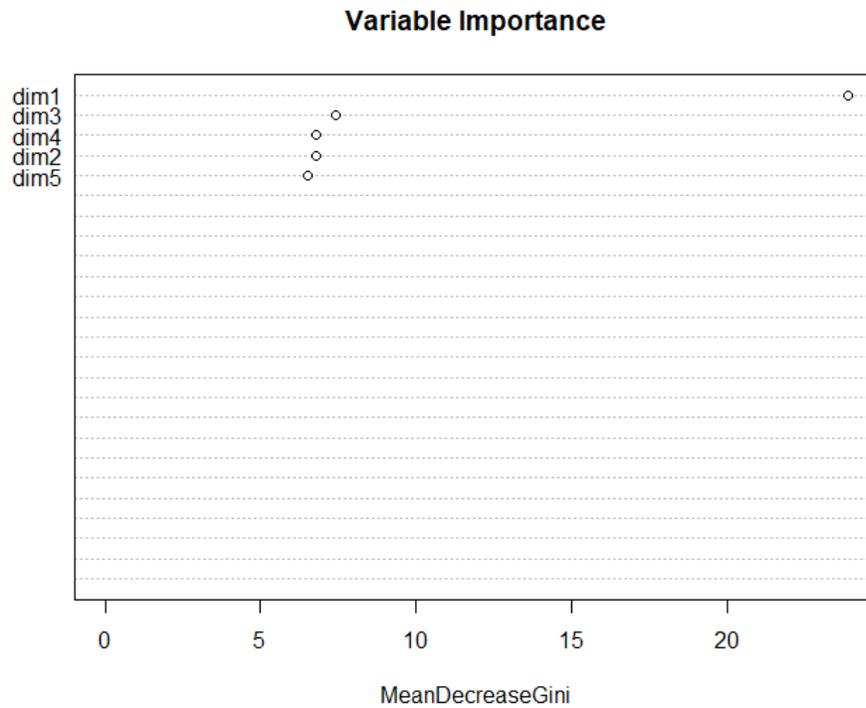


Figura C.14: Importancia de variables para Estrés día con Bosques Aleatorios - PCA

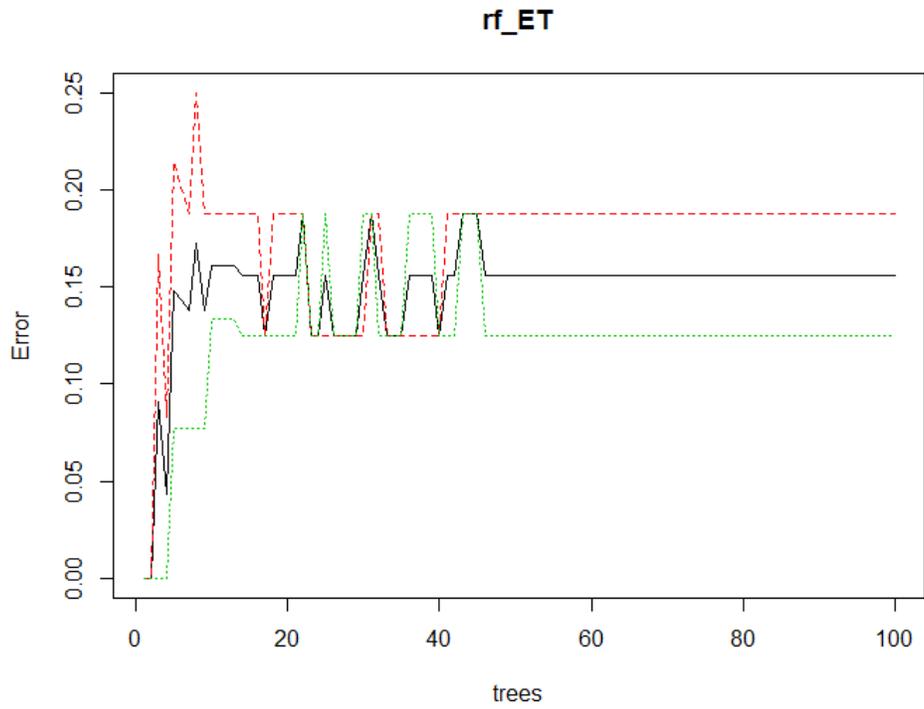


Figura C.15: Estrés tarde con Bosques Aleatorios - PCA

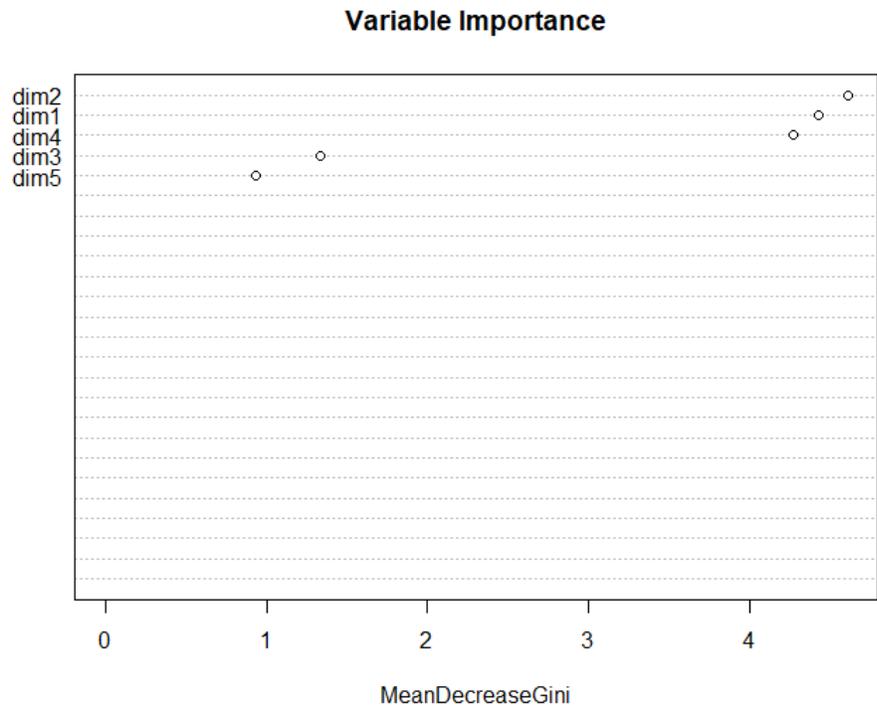


Figura C.16: Importancia de variables para Estrés tarde con Bosques Aleatorios - PCA

Árbol Clasificación ED

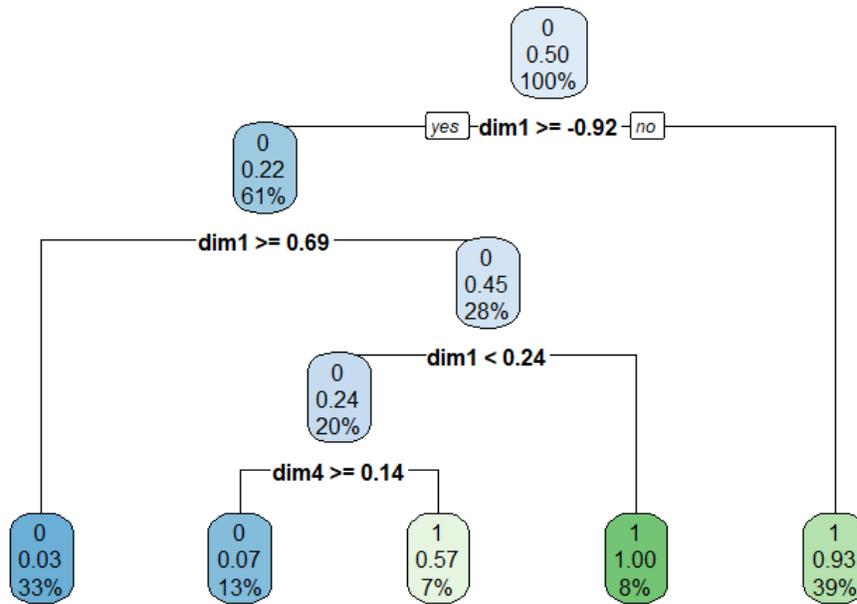


Figura C.17: Estrés día con Árboles de Clasificación - PCA

Árbol Clasificación ET

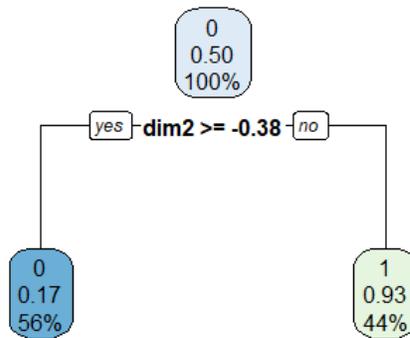


Figura C.18: Estrés tarde con Árboles de Clasificación - PCA

Anexo D

Resultados modelos con RF

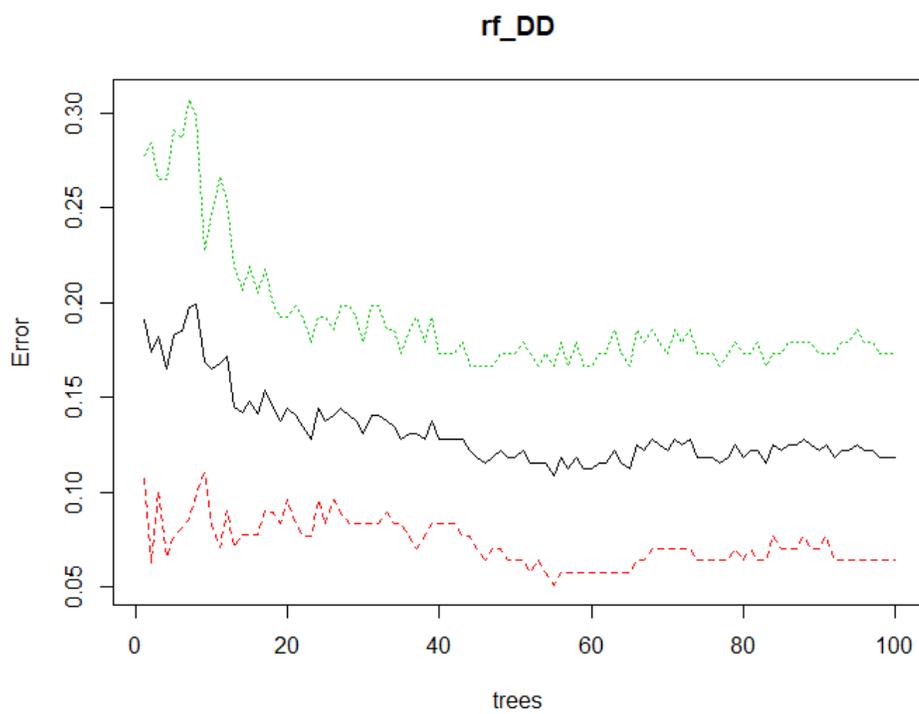


Figura D.1: Depresión día con Bosques Aleatorios - RF

Tabla D.1: Depresión tarde con Regresión Logística - RF

Coefficients:	Estimate	Std.	Error	z
(Intercept)	2.18e+01	5.78e+03	0.00	1.00
Creditos	-7.56e-02	1.62e+00	-0.05	0.96
Vida2	4.90e+01	5.78e+03	0.01	0.99
Vida3	4.47e+01	5.78e+03	0.01	0.99
Vida4	4.22e+01	5.78e+03	0.01	0.99
Vida5	3.60e+01	7.59e+04	0.00	1.00
CONTEXTTO_COMPANIA2	-3.42e+00	9.42e+04	0.00	1.00
CONTEXTTO_COMPANIA3	2.34e+01	7.97e+04	0.00	1.00
CONTEXTTO_COMPANIA4	-1.74e+01	1.84e+04	0.00	1.00
CONTEXTTO_COMPANIA5	-1.60e+01	3.55e+04	0.00	1.00
CONTEXTTO_COMPANIA6	-1.88e+01	7.95e+04	0.00	1.00
CONTEXTTO_COMPANIA7	-1.73e+00	4.16e+00	-0.42	0.68
CONTEXTTO_COMPANIA8	1.91e+01	3.20e+04	0.00	1.00
AnoIngreso2015	9.51e+00	1.88e+01	0.51	0.61
AnoIngreso2017	-1.42e+01	2.96e+01	-0.48	0.63
AnoIngreso2018	1.79e+01	3.80e+04	0.00	1.00
AnoIngreso2019	1.54e+01	2.13e+04	0.00	1.00
Edad	-1.05e+00	3.73e+00	-0.28	0.78
CONTEXTTO_HACIENDO2	-4.99e+00	1.40e+01	-0.36	0.72
CONTEXTTO_HACIENDO3	-4.23e+00	1.30e+01	-0.33	0.75
CONTEXTTO_HACIENDO4	-1.70e+00	1.31e+01	-0.13	0.90
CONTEXTTO_HACIENDO5	-6.20e+00	1.37e+01	-0.45	0.65
ANSIEDAD1	-1.04e+00	3.00e+00	-0.35	0.73
Semestre	-1.76e+00	3.87e+00	-0.45	0.65
Carrera4	-2.26e+00	3.28e+01	-0.07	0.94
Carrera5	1.50e+01	1.52e+04	0.00	1.00
Carrera7	1.16e+00	2.31e+01	0.05	0.96
Carrera8	-3.89e+00	2.57e+01	-0.15	0.88
Carrera9	4.72e+00	2.89e+01	0.16	0.87
Carrera11	5.22e+01	5.78e+03	0.01	0.99
Carrera13	-1.04e+01	3.23e+01	-0.32	0.75
Carrera14	2.05e+01	3.75e+04	0.00	1.00
Estado_Animico3	-4.91e+00	9.42e+00	-0.52	0.60
Estado_Animico4	-4.77e+01	5.78e+03	-0.01	0.99
Estado_Animico5	-1.53e+01	5.62e+04	0.00	1.00
NOTA_DIA	-4.49e+00	2.76e+00	-1.63	0.10

Tabla D.2: Depresión día con Regresión Logística - RF

DepresiónD_RF_RL				
Coefficients:	Estimate	Error	z	
(Intercept)	140.852	0.00	0.9980	
SUENO_CALIDAD	-0.7377	-3.98	6.8e-05	***
Carrera4	148.226	0.00	0.9963	
Carrera5	190.624	0.01	0.9952	
Carrera7	169.177	0.01	0.9958	
Carrera8	170.481	0.01	0.9957	
Carrera9	182.739	0.01	0.9954	
Carrera11	191.018	0.01	0.9952	
Carrera13	157.122	0.00	0.9961	
Carrera14	180.923	0.01	0.9955	
Estado_Animico2	-192.407	0.00	0.9966	
Estado_Animico3	-199.551	0.00	0.9965	
Estado_Animico4	-219.877	0.00	0.9961	
Estado_Animico5	-377.086	-0.01	0.9938	
ANSIEDAD1	14.682	3.21	0.0013	**
Semestre	0.1266	1.03	0.3015	
Creditos	0.0402	1.39	0.1634	
Edad	-0.4449	-2.64	0.0084	**
CONTEXTO_COMPANIA2	-23.388	-0.51	0.6133	
CONTEXTO_COMPANIA3	12.019	0.86	0.3884	
CONTEXTO_COMPANIA4	-0.0215	-0.03	0.9721	
CONTEXTO_COMPANIA5	-24.665	-2.16	0.0305	*
CONTEXTO_COMPANIA6	-0.7085	-0.63	0.5306	
CONTEXTO_COMPANIA7	-0.4608	-0.67	0.5010	
CONTEXTO_COMPANIA8	180.397	0.00	0.9967	
CONTEXTO_HACIENDO2	0.8712	0.94	0.3475	
CONTEXTO_HACIENDO3	22.268	2.88	0.0040	**
CONTEXTO_HACIENDO4	18.546	2.22	0.0267	*
CONTEXTO_HACIENDO5	12.369	1.33	0.1830	
HORAS_SUENO	-0.3205	-1.82	0.0688	.
AnoIngreso2015	-14.129	-1.52	0.1279	
AnoIngreso2016	179.269	0.01	0.9936	
AnoIngreso2017	-22.228	-2.21	0.0274	*
AnoIngreso2018	12.922	1.37	0.1717	
AnoIngreso2019	160.131	0.01	0.9953	
Vida2	16.501	1.07	0.2848	
Vida3	0.8164	0.52	0.6022	
Vida4	0.6724	0.47	0.6417	
Vida5	-145.592	-0.01	0.9910	
Vida4	0.6724	0.47	0.6417	
Vida5	-145.592	-0.01	0.9910	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

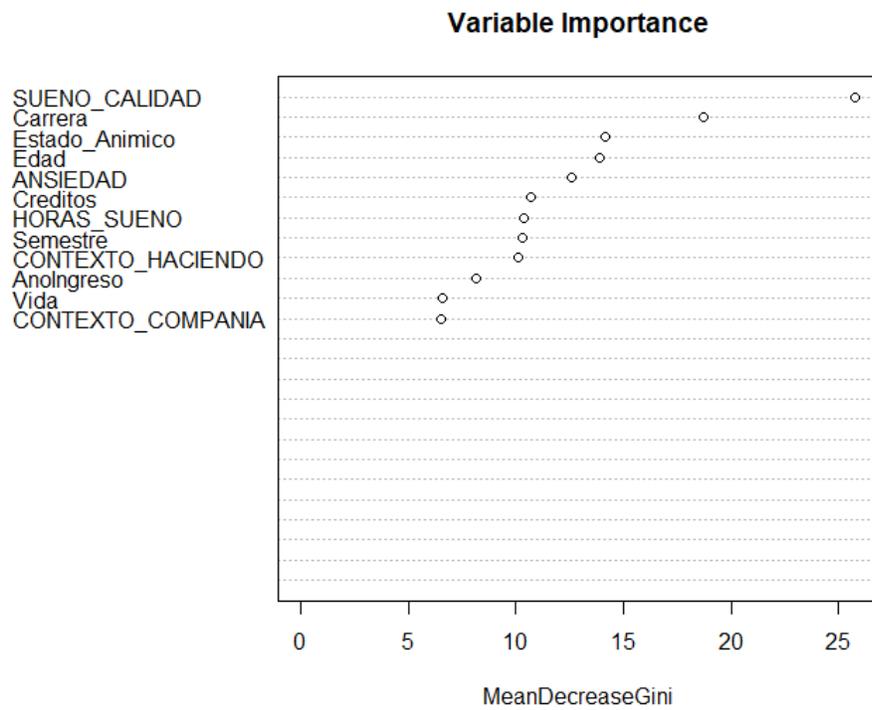


Figura D.2: Importancia de variables para Depresión día con Bosques Aleatorios - RF

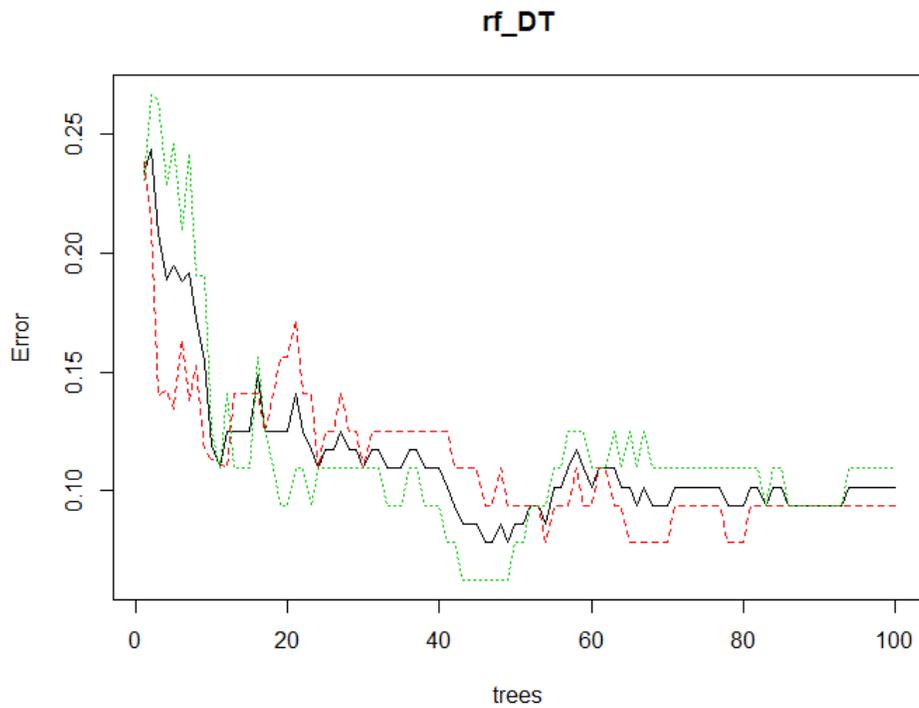


Figura D.3: Depresión tarde con Bosques Aleatorios - RF

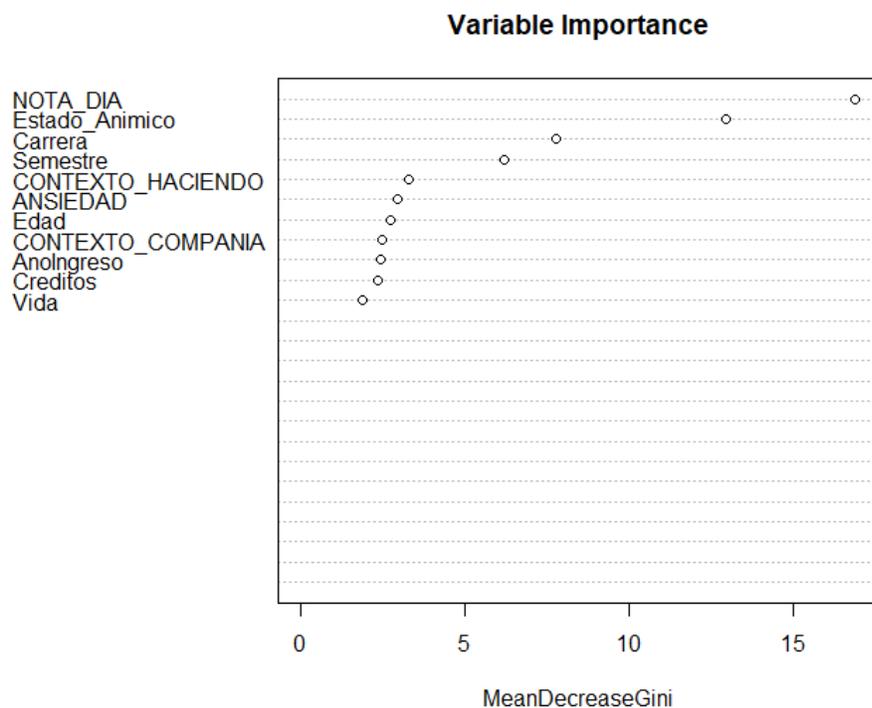


Figura D.4: Importancia de variables para Depresión tarde con Bosques Aleatorios - RF

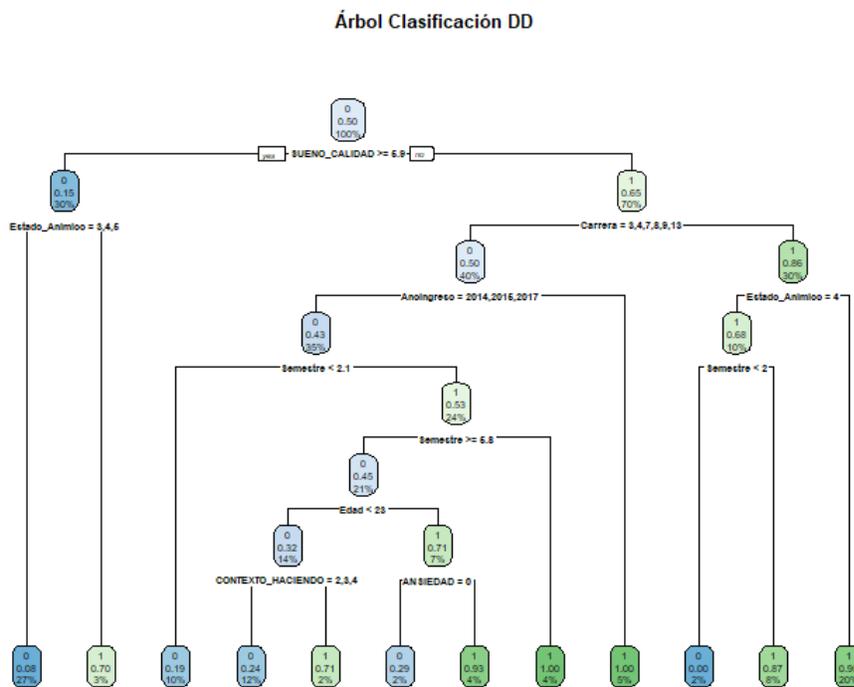


Figura D.5: Depresión día con Árboles de Clasificación - RF

Árbol Clasificación DT

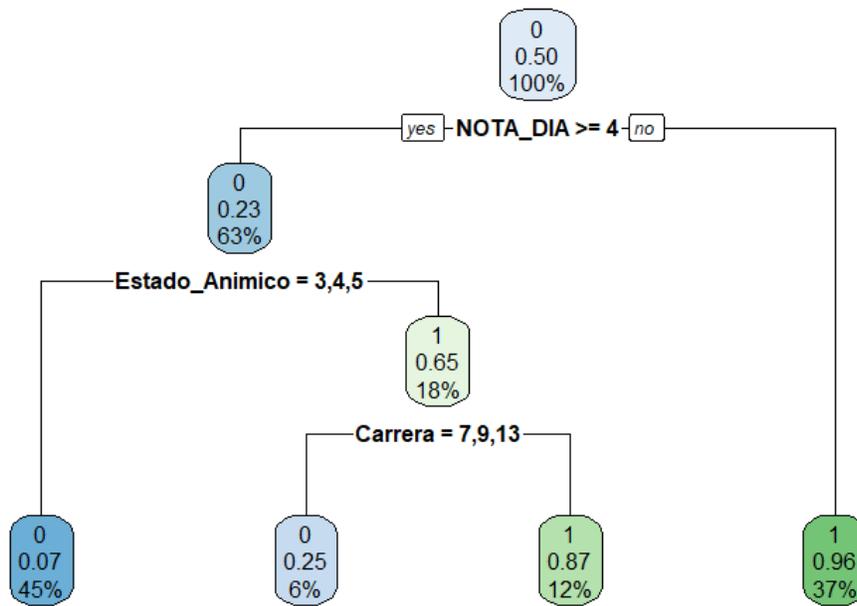


Figura D.6: Depresión tarde con Árboles de Clasificación - RF

Tabla D.3: Ansiedad día con Regresión Logística - RF

AnsiedadD_RF_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	2.58e+01	6.52e+03	0.00	0.99685	
AnoIngreso2015	5.00e+00	1.39e+00	3.60	0.00032	***
AnoIngreso2016	2.07e+01	4.40e+03	0.00	0.99624	
AnoIngreso2017	2.68e+00	2.69e+00	0.99	0.32041	
AnoIngreso2018	-4.80e-01	4.37e+00	-0.11	0.91259	
AnoIngreso2019	-3.43e+00	4.56e+00	-0.75	0.45204	
Semestre	-3.58e-01	3.71e-01	-0.97	0.33442	
Carga2	-8.72e-01	1.28e+00	-0.68	0.49572	
Carga3	-1.60e+00	1.97e+00	-0.82	0.41488	
CONTEXTO_COMPANIA2	-1.31e+00	4.35e+00	-0.30	0.76263	
CONTEXTO_COMPANIA3	-7.96e-01	1.23e+00	-0.65	0.51667	
CONTEXTO_COMPANIA4	-1.01e+00	9.81e-01	-1.03	0.30348	
CONTEXTO_COMPANIA5	1.45e-01	1.34e+00	0.11	0.91396	
CONTEXTO_COMPANIA6	-1.62e+00	1.27e+00	-1.27	0.20242	
CONTEXTO_COMPANIA7	8.57e-01	9.69e-01	0.88	0.37650	
CONTEXTO_COMPANIA8	1.64e+01	4.14e+03	0.00	0.99685	
Edad	-1.44e-01	2.61e-01	-0.55	0.58181	
Estado_Animico2	-1.80e+01	6.52e+03	0.00	0.99780	
Estado_Animico3	-2.07e+01	6.52e+03	0.00	0.99746	
Estado_Animico4	-1.71e+01	6.52e+03	0.00	0.99791	
Estado_Animico5	-1.82e+01	6.52e+03	0.00	0.99777	
Vida2	2.30e+00	3.59e+00	0.64	0.52121	
Vida3	3.92e+00	2.84e+00	1.38	0.16819	
Vida4	1.46e+00	2.07e+00	0.71	0.48058	
Vida5	-1.87e+01	1.69e+03	-0.01	0.99117	
CONTEXTO_HACIENDO2	-9.31e-01	1.18e+00	-0.79	0.42889	
CONTEXTO_HACIENDO3	-1.87e+00	9.45e-01	-1.98	0.04826	*
CONTEXTO_HACIENDO4	-7.09e-01	9.90e-01	-0.72	0.47434	
CONTEXTO_HACIENDO5	-1.14e+00	1.20e+00	-0.95	0.34059	
HORAS_SUENO	-1.67e-01	2.29e-01	-0.73	0.46470	
Carrera4	6.05e+00	1.86e+00	3.25	0.00116	**
Carrera5	3.00e+00	2.90e+00	1.03	0.30119	
Carrera7	-9.14e-01	2.50e+00	-0.36	0.71524	
Carrera8	-2.82e+00	2.98e+00	-0.94	0.34477	
Carrera9	5.97e-01	2.39e+00	0.25	0.80298	
Carrera11	1.42e+00	2.72e+00	0.52	0.60122	
Carrera13	1.94e+00	3.14e+00	0.62	0.53678	
Carrera14	NA	NA	NA	NA	
Creditos	-9.08e-03	5.70e-02	-0.16	0.87347	
DEPRESION1	2.71e+00	8.12e-01	3.34	0.00085	***
SUENO_CALIDAD	-5.07e-01	2.22e-01	-2.29	0.02213	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tabla D.4: Ansiedad tarde con Regresión Logística - RF

AnsiedadD_RF_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	402.932	23.995.532	0.02	0.9866	
DEPRESION1	16.867	0.9088	1.86	0.0635	.
Edad	-0.2800	0.2043	-1.37	0.1705	
AnoIngreso2015	22.413	0.9590	2.34	0.0194	*
AnoIngreso2016	-16.753	20.864	-0.80	0.4220	
AnoIngreso2017	-40.111	17.648	-2.27	0.0230	*
AnoIngreso2018	-65.774	28.638	-2.30	0.0216	*
AnoIngreso2019	-71.881	27.950	-2.57	0.0101	*
Estado_Animico2	-230.440	23.995.456	-0.01	0.9923	
Estado_Animico3	-233.722	23.995.458	-0.01	0.9922	
Estado_Animico4	-215.587	23.995.455	-0.01	0.9928	
Estado_Animico5	-199.852	23.995.458	-0.01	0.9934	
Vida2	-21.124	16.803	-1.26	0.2087	
Vida3	-0.0437	16.174	-0.03	0.9784	
Vida4	-17.020	13.682	-1.24	0.2135	
Vida5	-95.885	34.019	-2.82	0.0048	**
Semestre	-0.8797	0.2257	-3.90	9.7e-05	***
Creditos	0.0323	0.0254	1.27	0.2031	
CONTEXTO_COMPANIA2	0.2048	14.747	0.14	0.8895	
CONTEXTO_COMPANIA3	17.940	15.670	1.14	0.2523	
CONTEXTO_COMPANIA4	0.5989	0.9094	0.66	0.5102	
CONTEXTO_COMPANIA5	24.485	11.307	2.17	0.0304	*
CONTEXTO_COMPANIA6	141.131	23.995.450	0.01	0.9953	
CONTEXTO_COMPANIA7	-0.3262	0.4434	-0.74	0.4620	
CONTEXTO_COMPANIA8	193.805	9.072.725	0.02	0.9830	
CONTEXTO_HACIENDO2	-0.0753	0.9783	-0.08	0.9386	
CONTEXTO_HACIENDO3	-14.436	0.8101	-1.78	0.0748	.
CONTEXTO_HACIENDO4	-13.217	10.668	-1.24	0.2154	
CONTEXTO_HACIENDO5	-15.082	0.8885	-1.70	0.0896	.
Carrera4	29.242	16.253	1.80	0.0720	.
Carrera5	12.620	18.979	0.66	0.5061	
Carrera7	-0.4327	18.292	-0.24	0.8130	
Carrera8	-66.556	24.373	-2.73	0.0063	**
Carrera9	-0.7189	18.194	-0.40	0.6928	
Carrera11	-0.0959	18.958	-0.05	0.9596	
Carrera13	-25.154	28.112	-0.89	0.3709	
Carrera14	NA	NA	NA	NA	
NOTA_DIA	-0.4197	0.1671	-2.51	0.0120	*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

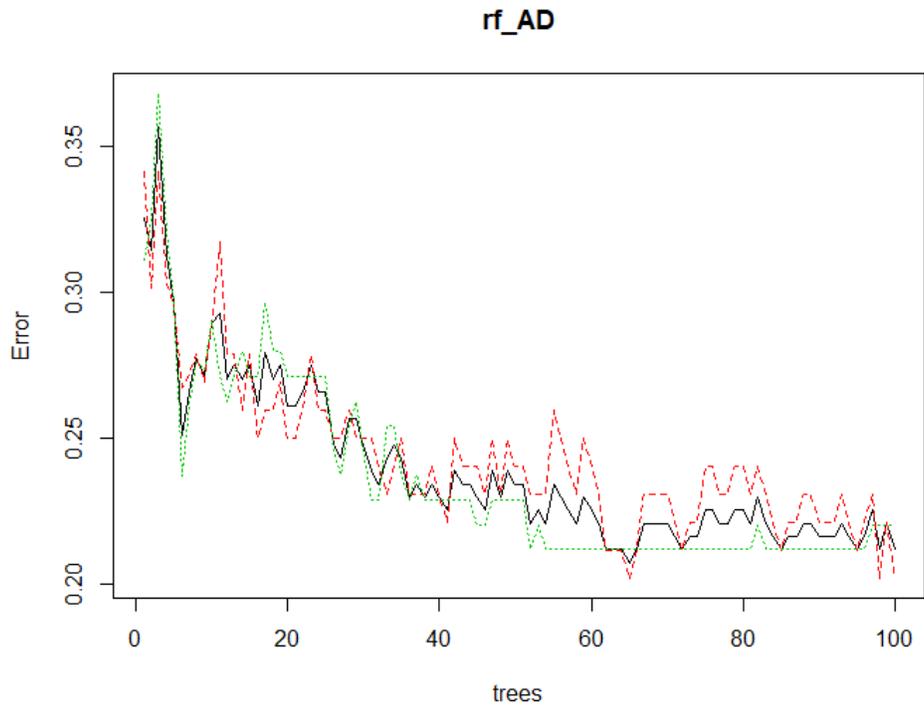


Figura D.7: Ansiedad tarde con Bosques Aleatorios - RF

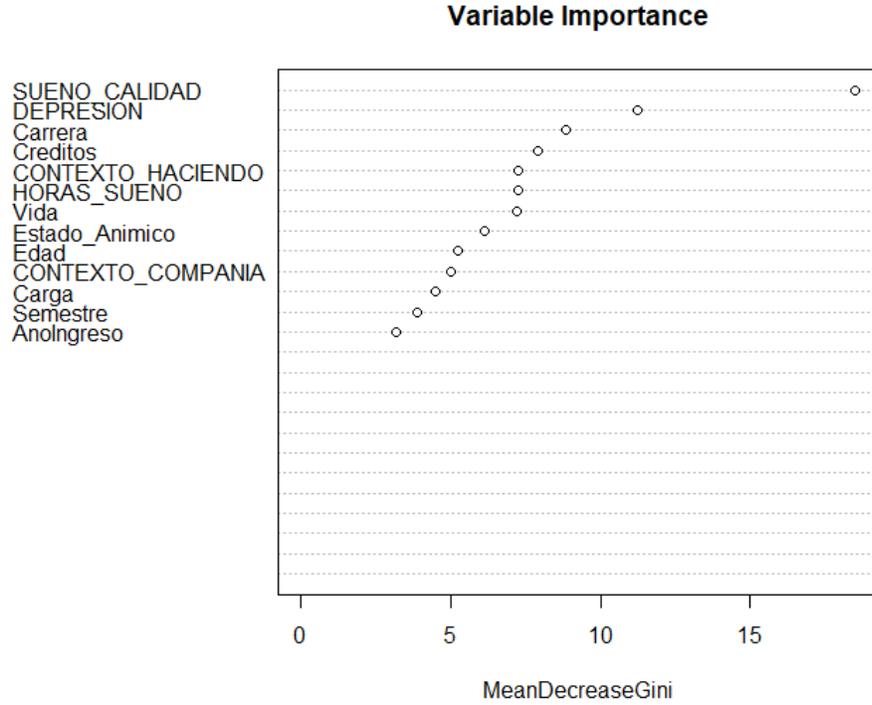


Figura D.8: Importancia de variables para Ansiedad día con Bosques Aleatorios - RF

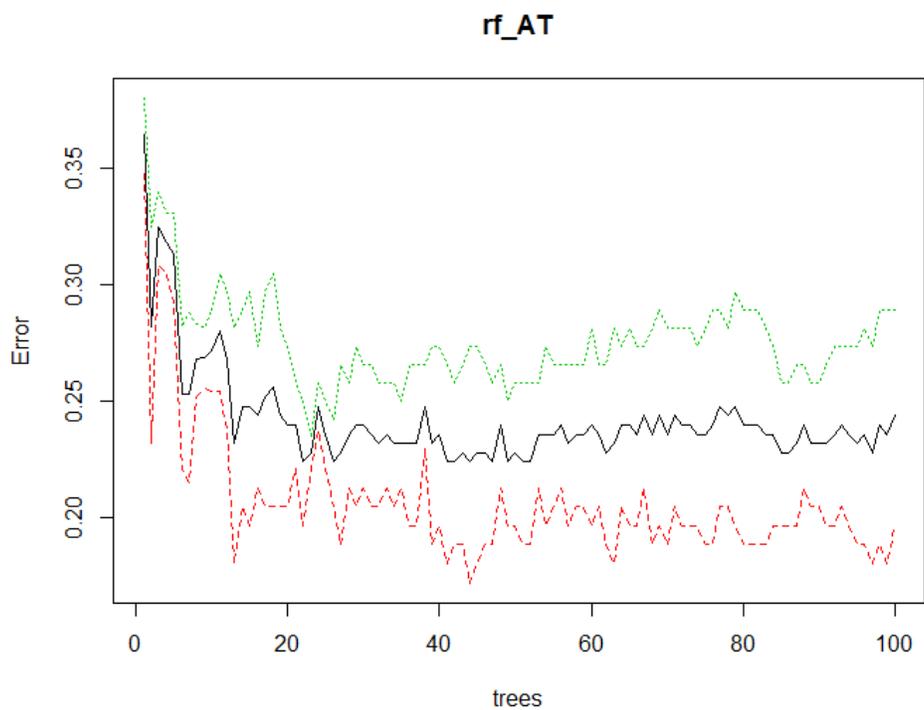


Figura D.9: Ansiedad tarde con Bosques Aleatorios - RF

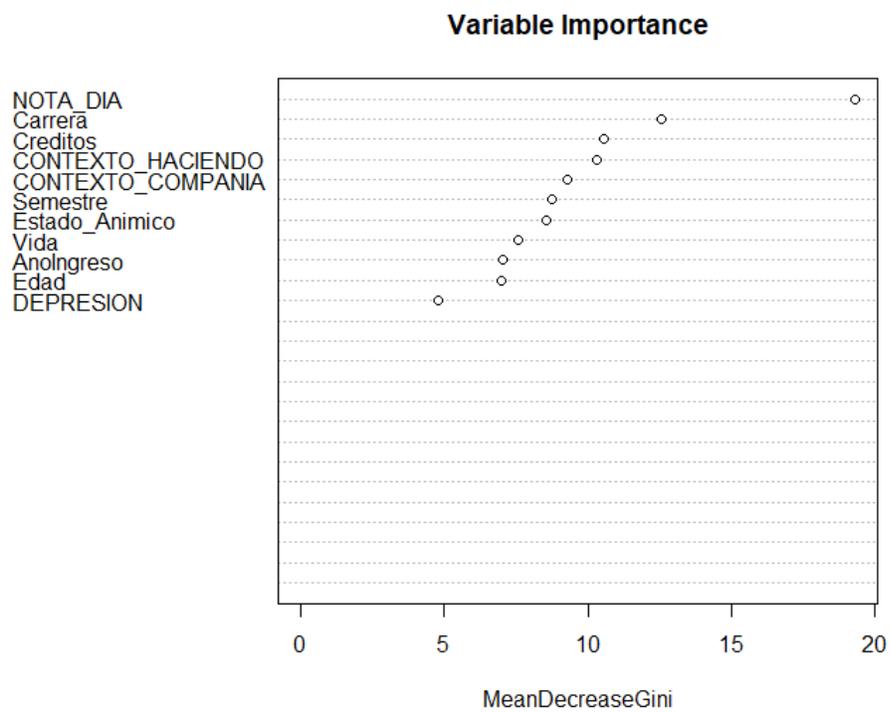


Figura D.10: Importancia de variables para Ansiedad tarde con Bosques Aleatorios - RF

Árbol Clasificación AD

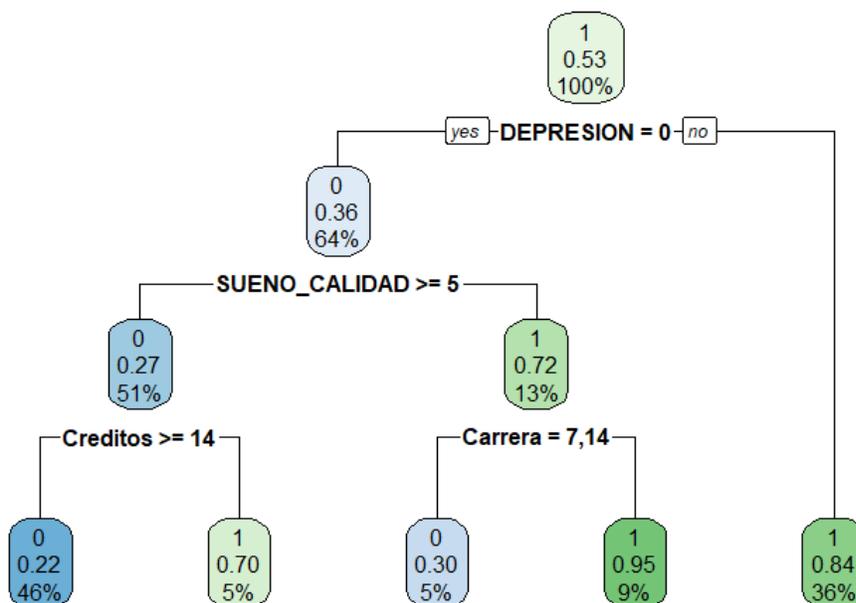


Figura D.11: Ansiedad día con Árboles de Clasificación - RF

Árbol Clasificación AT

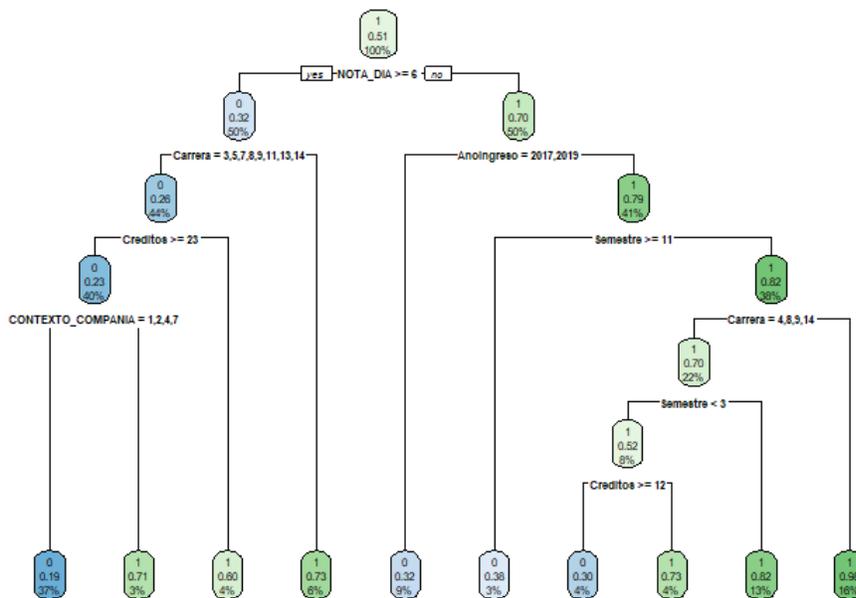


Figura D.12: Ansiedad tarde con Árboles de Clasificación - RF

Tabla D.5: Estrés día con Regresión Logística - RF

EstrésD_RF_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	75.867	67.244	1.13	0.25922	
AnoIngreso2015	17.159	10.769	1.59	0.11107	
AnoIngreso2016	124.633	16.546.393	0.01	0.99399	
AnoIngreso2017	-18.348	48.944	-0.37	0.70775	
AnoIngreso2018	61.647	26.820	2.30	0.02153	*
AnoIngreso2019	47.148	52.615	0.90	0.37021	
Semestre	0.4607	0.2460	1.87	0.06113	.
Carga2	0.8856	10.654	0.83	0.40584	
Carga3	-11.510	47.942	-0.24	0.81026	
CONTEXTO_COMPANIA2	-165.202	10.798.612	-0.02	0.98779	
CONTEXTO_COMPANIA3	0.4762	11.798	0.40	0.68651	
CONTEXTO_COMPANIA4	-0.2596	0.7833	-0.33	0.74030	
CONTEXTO_COMPANIA5	-0.7182	12.480	-0.58	0.56496	
CONTEXTO_COMPANIA6	-0.9646	10.363	-0.93	0.35194	
CONTEXTO_COMPANIA7	0.6971	0.6385	1.09	0.27497	
CONTEXTO_COMPANIA8	0.6306	15.980	0.39	0.69314	
Edad	-0.4739	0.2259	-2.10	0.03593	*
Estado_Animico3	-0.7571	48.758	-0.16	0.87660	
Estado_Animico4	-19.753	47.020	-0.42	0.67441	
Estado_Animico5	-70.590	55.646	-1.27	0.20460	
Vida2	20.453	23.067	0.89	0.37526	
Vida3	-0.7181	13.523	-0.53	0.59541	
Vida4	10.570	13.710	0.77	0.44072	
Vida5	17.486	21.167	0.83	0.40876	
CONTEXTO_HACIENDO2	0.2857	0.9064	0.32	0.75258	
CONTEXTO_HACIENDO3	-0.0696	0.7618	-0.09	0.92718	
CONTEXTO_HACIENDO4	-0.0773	0.7877	-0.10	0.92181	
CONTEXTO_HACIENDO5	0.1931	0.8737	0.22	0.82511	
HORAS_SUENO	-0.1471	0.1629	-0.90	0.36670	
Carrera4	31.278	15.581	2.01	0.04470	*
Carrera5	24.095	19.912	1.21	0.22625	
Carrera7	28.019	48.330	0.58	0.56210	
Carrera8	95.691	30.581	3.13	0.00175	**
Carrera9	96.829	33.954	2.85	0.00435	**
Carrera11	40.640	17.348	2.34	0.01915	*
Carrera13	83.651	22.178	3.77	0.00016	***
Carrera14	NA	NA	NA	NA	
Creditos	-0.0573	0.0274	-2.09	0.03687	*
DEPRESION1	-0.4958	0.4702	-1.05	0.29163	
SUENO_CALIDAD	-0.3127	0.1697	-1.84	0.06539	.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Tabla D.6: Estrés tarde con Regresión Logística - RF

EstrésT_RF_RL					
Coefficients:	Estimate	Std.	Error	z	
(Intercept)	92.706	23.995.506	0.00	0.99692	
Convivencia2	79.991	19.646	4.07	4.7e-05	***
Convivencia3	16.976	14.432	1.18	0.23947	
Convivencia4	-161.907	13.367.771	-0.01	0.99034	
Carga2	-28.041	16.212	-1.73	0.08369	.
Carga3	-32.683	17.769	-1.84	0.06587	.
ANSIEDAD	-0.0360	0.3939	-0.09	0.92721	
Estado_Animico2	-113.569	23.995.452	0.00	0.99622	
Estado_Animico3	-134.075	23.995.453	-0.01	0.99554	
Estado_Animico4	-132.649	23.995.451	-0.01	0.99559	
Estado_Animico5	-147.562	23.995.452	-0.01	0.99509	
AnoIngreso2015	0.8286	0.9837	0.84	0.39959	
AnoIngreso2016	12.288	20.228	0.61	0.54354	
AnoIngreso2017	17.474	16.173	1.08	0.27995	
AnoIngreso2018	46.732	26.920	1.74	0.08257	.
AnoIngreso2019	26.208	28.605	0.92	0.35956	
Semestre	0.0849	0.1444	0.59	0.55651	
CONTEXTO_COMPANIA2	-0.4062	11.573	-0.35	0.72559	
CONTEXTO_COMPANIA3	-11.005	15.874	-0.69	0.48816	
CONTEXTO_COMPANIA4	-0.0313	0.6877	-0.05	0.96365	
CONTEXTO_COMPANIA5	-0.2730	0.6897	-0.40	0.69220	
CONTEXTO_COMPANIA6	-164.901	16.773.098	-0.01	0.99216	
CONTEXTO_COMPANIA7	-0.1580	0.4027	-0.39	0.69472	
CONTEXTO_COMPANIA8	145.377	10.754.556	0.01	0.98921	
Edad	-0.0704	0.2020	-0.35	0.72741	
Creditos	-0.0131	0.0188	-0.70	0.48507	
Vida2	-16.583	15.277	-1.09	0.27768	
Vida3	-0.9623	13.526	-0.71	0.47678	
Vida4	14.012	11.702	1.20	0.23117	
Vida5	17.088	19.319	0.88	0.37642	
NOTA_DIA	-0.0473	0.1304	-0.36	0.71659	
CONTEXTO_HACIENDO2	20.740	0.8485	2.44	0.01451	*
CONTEXTO_HACIENDO3	15.284	0.7173	2.13	0.03311	*
CONTEXTO_HACIENDO4	16.329	0.8565	1.91	0.05660	.
CONTEXTO_HACIENDO5	22.992	0.8155	2.82	0.00481	**
Carrera4	17.806	22.987	0.77	0.43856	
Carrera5	29.849	23.024	1.30	0.19483	
Carrera7	38.230	21.465	1.78	0.07490	.
Carrera8	55.540	16.788	3.31	0.00094	***
Carrera9	56.626	21.481	2.64	0.00839	**
Carrera11	41.598	22.133	1.88	0.06018	.
Carrera13	70.769	25.432	2.78	0.00539	**
Carrera14	NA	NA	NA	NA	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

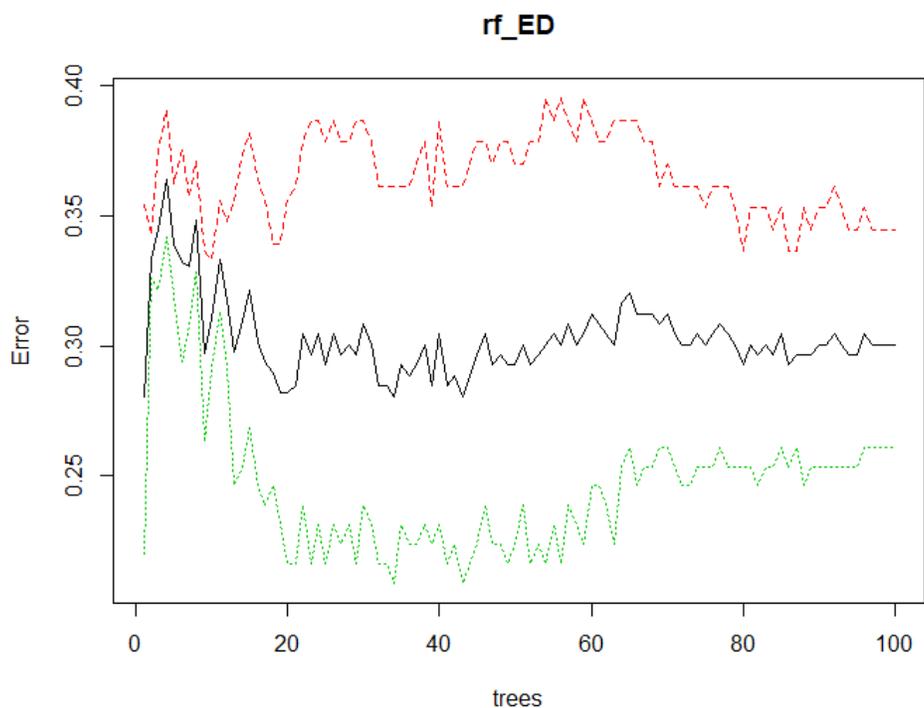


Figura D.13: Estrés tarde con Bosques Aleatorios - RF

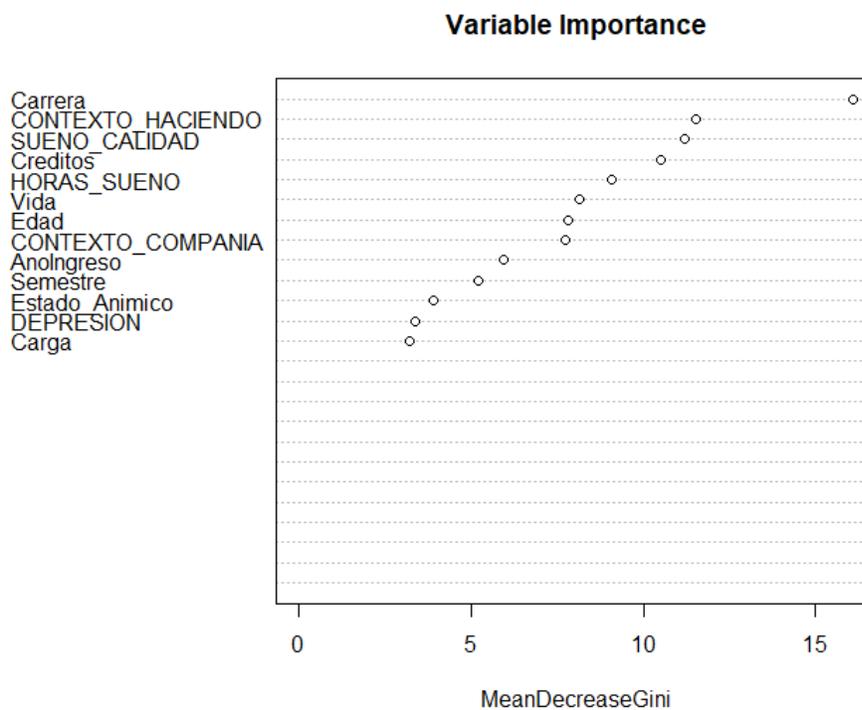


Figura D.14: Importancia de variables para Estrés día con Bosques Aleatorios - RF

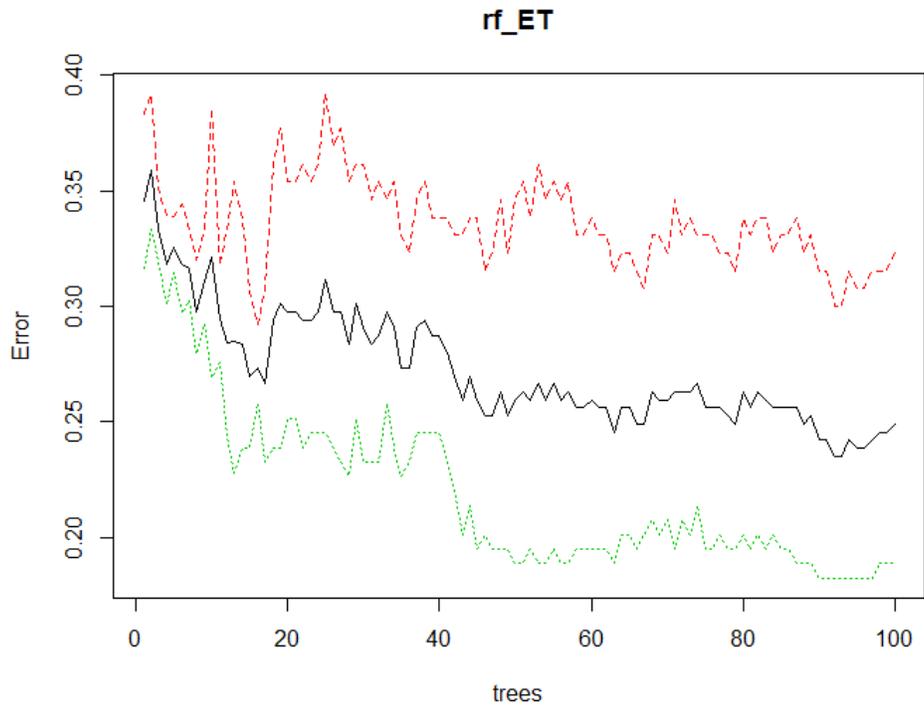


Figura D.15: Estrés tarde con Bosques Aleatorios - RF

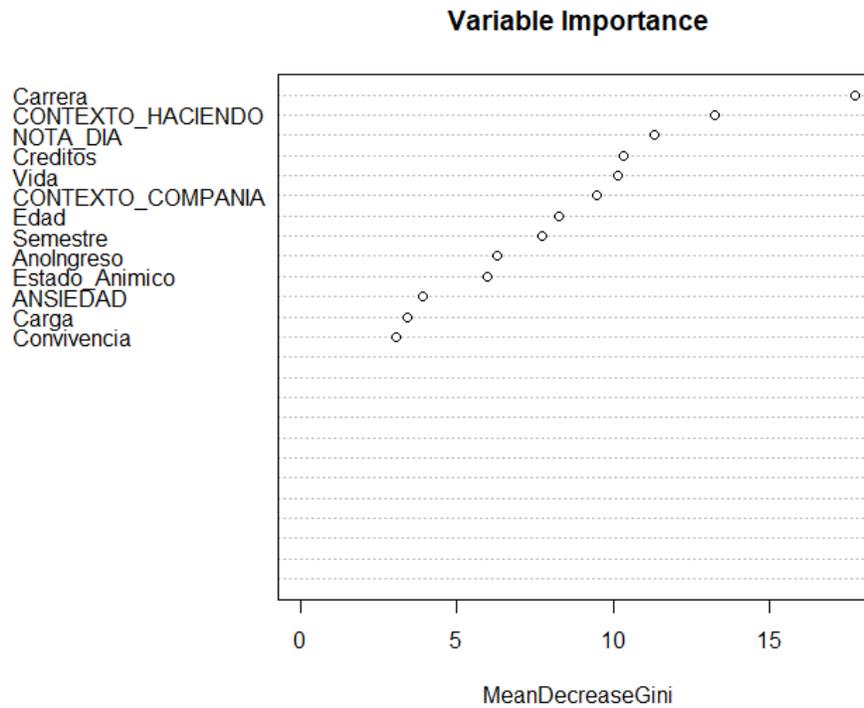


Figura D.16: Importancia de variables para Estrés tarde con Bosques Aleatorios - RF

Árbol Clasificación ED

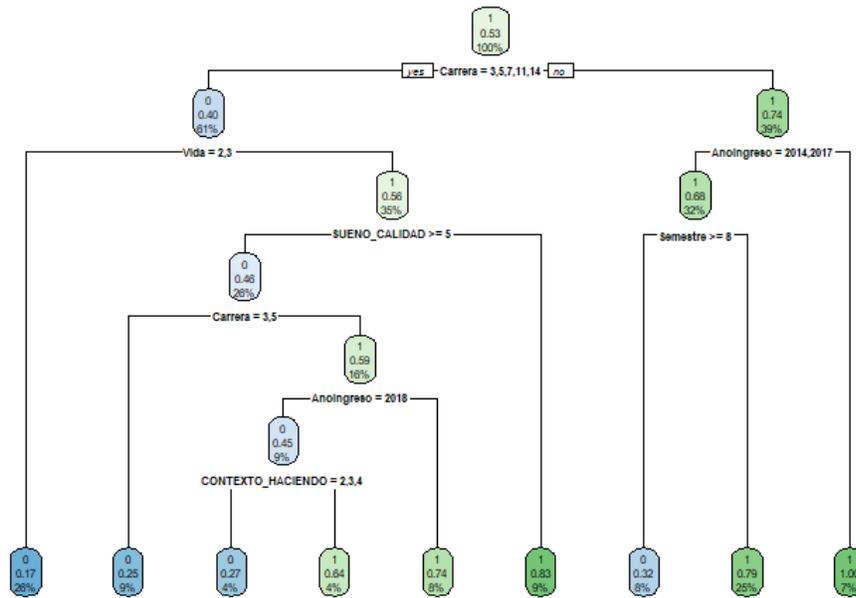


Figura D.17: Estrés día con Árboles de Clasificación - RF

Árbol Clasificación ET

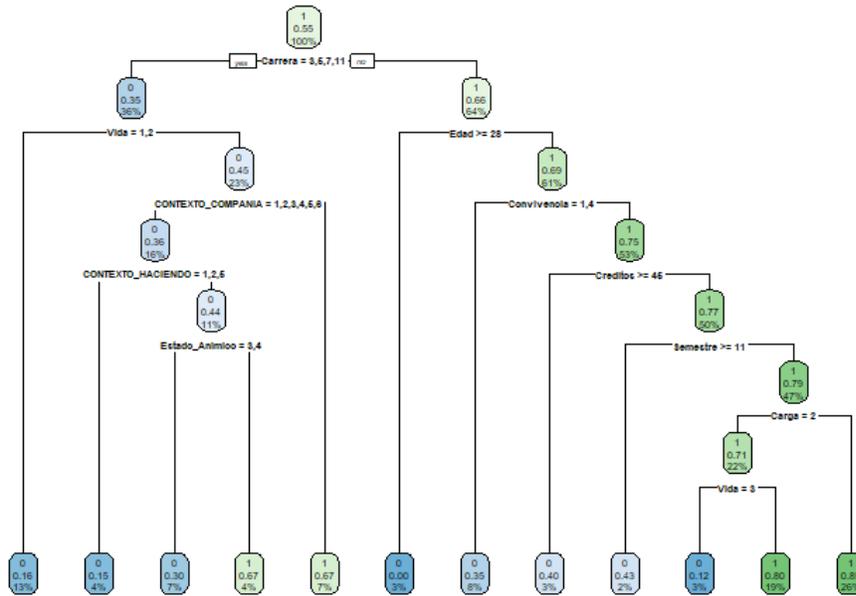


Figura D.18: Estrés tarde con Árboles de Clasificación - RF