

## TABLA DE CONTENIDO

1. Introducción .....	1
2. Pregunta de investigación, objetivos y resultados esperados.....	4
2.1. Pregunta de investigación .....	4
2.2. Objetivo general .....	4
2.3. Objetivos específicos .....	4
2.4. Resultados esperados.....	4
3. Marco Conceptual.....	5
3.1. Concepto de Reinserción Social .....	5
3.2. Concepto de Reincidencia .....	7
3.3. ¿Cómo medir la reincidencia?.....	9
3.4. Experiencia internacional .....	11
3.4.1. Inglaterra .....	12
3.4.2. Canadá .....	12
3.4.3. Estados Unidos.....	13
3.5. Factores de riesgo asociados a la medición de reincidencia .....	14
3.6. Modelos considerados .....	15
3.6.1. Análisis prospectivo .....	15
3.6.2. Análisis Retrospectivo .....	17
3.7. Fundamentos para la elección del modelo a utilizar .....	20
3.8. Programa Volver a Empezar .....	22
3.8.1. Modelo Riesgo Necesidad y capacidad de Respuesta (RNR).....	22
3.8.2. Implementación .....	24
3.8.3. Análisis del nivel de cumplimiento de los objetivos del Programa Volver a Empezar .....	28
4. Metodología.....	33
4.1. Tratamiento de datos .....	33
4.1.1. Variables dependientes: .....	35
4.2. Regresión logística.....	36
4.3. Random Forest .....	40
4.4. XGBoost.....	41
4.5. Random Forest vs XGBoost.....	41
5. Análisis exploratorio de datos .....	43

5.1.	Valores perdidos .....	43
5.1.1.	Edad .....	43
5.1.2.	Género.....	45
5.1.3.	Profesión .....	46
5.1.4.	Nivel educacional.....	48
5.1.5.	Comuna .....	49
5.1.6.	Reincidente.....	50
5.1.7.	Grupo de referencia.....	50
5.1.8.	Cumplimiento del plan de reinserción .....	51
5.1.9.	Colocado laboralmente si lo requiere .....	52
5.2.	Distribución condicionada de la variable dependiente .....	53
5.2.1.	Reincidencia según edad .....	54
5.2.2.	Reincidencia según tamaño del grupo de referencia.....	55
5.2.3.	Reincidencia según colocación laboral.....	56
5.2.4.	Nivel educacional.....	57
5.2.5.	Reincidencia según comuna.....	58
5.2.6.	Reincidencia según género .....	59
5.2.7.	Reincidencia según variable Profesión.....	60
5.2.8.	Reincidencia según cumplimiento del programa de reinserción social 61	
6.	Resultados.....	62
6.1.	Regresión logística multivariada .....	62
6.2.	Importancias con Random Forest y XGBoost .....	64
7.	Análisis .....	68
7.1.	Regresión logística.....	68
7.2.	Análisis de importancias con Random Forest y XGBoost .....	69
7.2.1.	La variable más importante .....	69
7.2.2.	Variables de alta importancia .....	70
7.2.3.	Variables de mediana importancia.....	71
7.2.4.	Variables de menor importancia .....	71
8.	Conclusiones .....	72
9.	Bibliografía.....	76
	Anexos .....	80

Anexo A: Listado de variables.....	80
Anexo B: Resultados regresión logística multivariada .....	84
Anexo C: Importancias para XGBoost .....	85
Anexo D: Importancias para Random Forest .....	86

## INDICE DE TABLAS

Tabla 1 - Ciclos de Capacitación VAE (elaboración externa) .....	25
Tabla 2 - Reincidencia según criterios (elaboración externa).....	29
Tabla 3 - Continuidad del programa VAE (elaboración externa) .....	31
Tabla 4 - Valores perdidos por variable (elaboración propia).....	35
Tabla 5 - Medidas de tendencia central, variable Edad (elaboración propia) .....	43
Tabla 6 - Distribución de la variable edad (elaboración propia).....	45
Tabla 7 - Distribución de la variable comuna, antes y después de la limpieza de datos (elaboración propia).....	49
Tabla 8 - Medidas de tendencia central para la variable comuna, antes y después de la limpieza de datos (elaboración propia) .....	50
Tabla 9 - Reincidencia según tramo etario (elaboración propia) .....	54
Tabla 10 - Reincidencia según tamaño del grupo de referencia (elaboración propia) .....	56
Tabla 11 - Reincidencia según variable colocado laboralmente (elaboración propia) .....	57
Tabla 12 - Reincidencia según nivel educacional (elaboración propia) .....	57
Tabla 13 - Reincidencia según comuna (elaboración propia) .....	59
Tabla 14 - Reincidencia según cumplimiento del programa de reinserción (elaboración propia) .....	61
Tabla 15 - Resultados regresión logística (elaboración propia).....	63
Tabla 16 - Estadísticos Regresión logística.....	64
Tabla 17 - Importancias RF con función de pérdidas MSE (elaboración propia) ..	64
Tabla 18 - Importancias RF con función de pérdidas NodePurity (elaboración propia) .....	65
Tabla 19 - Importancias XGB (elaboración propia) .....	65
Tabla 20 - Importancias finales normalizadas (elaboración propia) .....	66
Tabla 21 - Resumen variables significativas (elaboración propia).....	69

## INDICE DE FIGURAS

Figura 1 - Flujograma de atenciones (Elaboración externa).....	27
--	----

## INDICE DE GRÁFICOS

Gráfico 1 - Encuesta utilidad del programa (elaboración externa).....	30
---	----

Gráfico 2 - Encuesta evaluación del programa (elaboración externa).....	31
Gráfico 3 - Matriz de correlación (elaboración propia).....	35
Gráfico 4 - Comparación Random Forest vs XGBoost (elaboración propia) .....	42
Gráfico 5 - Histograma variable edad, previo a la limpieza de datos (elaboración propia) .....	44
Gráfico 6 - Histograma variable edad, posterior a la limpieza de datos (elaboración propia) .....	44
Gráfico 7 - Distribución de la variable género, previo a la limpieza de datos (elaboración propia) .....	45
Gráfico 8 - Distribución de la variable género, posterior a la limpieza de datos (elaboración propia) .....	46
Gráfico 9 - Distribución de la variable profesión, previo a la limpieza de datos (elaboración propia) .....	47
Gráfico 10 - Distribución de la variable profesión, posterior a la limpieza de datos (elaboración propia) .....	47
Gráfico 11 - Distribución de la variable nivel educacional, antes y después de la limpieza de datos (elaboración propia).....	48
Gráfico 12 - Distribución de la variable grupo de referencia, antes y después de la limpieza de datos (elaboración propia).....	51
Gráfico 13 - Distribución de la variable cumplimiento del plan de reinserción, previo a la limpieza de datos (elaboración propia).....	52
Gráfico 14 - Distribución de la variable cumplimiento del plan de reinserción, posterior a la limpieza de datos (elaboración propia) .....	52
Gráfico 15 - Distribución de la variable colocado laboralmente, antes y después de la limpieza de datos (elaboración propia).....	53
Gráfico 16 - Distribución de la variable dependiente (reincidencia) (elaboración propia) .....	54
Gráfico 17 - Reincidencia según tramo etario (elaboración propia).....	55
Gráfico 18 - Reincidencia según tamaño del grupo de referencia (elaboración propia) .....	56
Gráfico 19 - Reincidencia según nivel educacional (elaboración propia) .....	58
Gráfico 20 - - Reincidencia según comuna (elaboración propia).....	59
Gráfico 21 - Reincidencia según género (elaboración propia).....	60
Gráfico 22 - Reincidencia según profesión (elaboración propia) .....	60

## INDICE DE CÓDIGOS

Código 1 - Regresión logística .....	63
Código 2 - Random Forest y XGBoost .....	67