

## TABLA DE CONTENIDO

CAPÍTULO 1: INTRODUCCIÓN .....	1
1.1 ANTECEDENTES GENERALES DE LA INDUSTRIA MINERA .....	1
1.2 ANTECEDENTES GENERALES DE LA EMPRESA .....	3
1.2.1 DECLARACIONES ESTRATÉGICAS .....	4
1.2.2 SECTOR INDUSTRIAL .....	4
CAPÍTULO 2: DESCRIPCIÓN DEL ESTUDIO .....	7
2.1 PROBLEMA.....	7
2.1.1 CONTEXTO .....	7
2.1.2 JUSTIFICACIÓN .....	8
2.2 HIPÓTESIS Y ALTERNATIVAS DE SOLUCIÓN .....	8
2.2.1 HIPÓTESIS .....	8
2.2.2 ALTERNATIVAS DE SOLUCIÓN .....	8
CAPÍTULO 3: OBJETIVOS.....	10
3.1 OBJETIVO GENERAL.....	10
3.2 OBJETIVOS ESPECÍFICOS.....	10
CAPÍTULO 4: MARCO TEÓRICO.....	11
4.1 ACCIDENTABILIDAD EN MINERÍA .....	11
4.2 SEGURIDAD EN MINERÍA .....	14
4.2.1 DIFERENCIA ENTRE PELIGRO Y RIESGO.....	14
4.2.2 NORMA OHSAS 18001 .....	14
4.2.3 INDICADORES DE ACCIDENTABILIDAD .....	15
4.2.4 GESTIÓN DE CONTROLES CRÍTICOS .....	16
4.2.5 OBSERVACIONES CONDUCTUALES EN PREVENCIÓN.....	17
4.3 PROCESAMIENTO DE TEXTO .....	18
4.3.1 TOKENIZATION.....	18
4.3.2 LIMPIEZA Y ESTANDARIZACIÓN.....	19
4.3.3 STOPWORDS .....	19
4.3.4 TF-IDF .....	20
4.4 MACHINE LEARNING.....	21
4.4.1 LOGIT.....	21
4.4.2 SUPPORT VECTOR MACHINE.....	22

4.4.3 REDES NEURONALES ARTIFICIALES .....	23
4.5 MÉTRICAS DE EVALUACIÓN .....	25
4.5.1 LOGARITHMIC LOSS .....	25
4.5.2 MATRIZ DE CONFUSIÓN .....	25
4.5.3 AREA BAJO LA CURVA.....	27
4.5.4 F1 SCORE .....	28
4.6 TRABAJOS SIMILARES .....	28
CAPÍTULO 5: METODOLOGÍA .....	31
5.1 ANALISIS DE LA SITUACIÓN ACTUAL.....	31
5.2 PREPARACIÓN DE BASES DE DATOS NECESARIAS .....	31
5.3 PROCESAMIENTO Y TRANSFORMACIÓN DE LA DATA.....	31
5.4 MODELAMIENTO.....	32
5.5 EVALUACIÓN .....	32
CAPÍTULO 6: ALCANCES Y RESULTADOS ESPERADOS .....	33
6.1 ALCANCES.....	33
6.2 RESULTADOS ESPERADOS.....	33
CAPÍTULO 7: IMPLEMENTACIÓN DE METODOLOGÍA.....	34
7.1 LEVANTAMIENTO DE LA SITUACIÓN ACTUAL .....	34
7.2 CONSOLIDACIÓN DE LA DATA .....	34
7.3 PROCESAMIENTO Y TRANSFORMACIÓN DE LA DATA.....	35
7.3.1 ANALISIS EXPLORATORIO .....	35
7.3.2 TRANSFORMACIÓN DE LA DATA.....	40
7.4 MODELAMIENTO.....	40
7.4.1 LOGIT.....	41
7.4.2 SUPPORT VECTOR MACHINE.....	43
7.4.3 RED NEURONAL .....	45
7.5 MODELO FINAL .....	48
7.5.1 DISCUSIÓN .....	52
CAPÍTULO 8: CONCLUSIONES .....	54
8.1 SOBRE LA CALIDAD DE LOS DATOS.....	55
8.2 TECNICAS DE MACHINE LEARNING.....	55
8.3 HIPOTESIS .....	57
8.4 SOBRE LAS CATEGORIAS .....	57

8.5 APLICACIÓN .....	58
CAPÍTULO 9: RECOMENDACIONES Y TRABAJOS FUTUROS .....	60
CAPÍTULO 10: BIBLIOGRAFÍA .....	61
CAPÍTULO 11: ANEXOS .....	65
Anexo A: Clasificación de OPS según 30 categorías .....	65
Anexo B: Ranking de las 10 palabras más usadas por categoría.....	66
Anexo C: Curvas ROC modelo Logit.....	67
Anexo D: Métricas de evaluación Logit.....	69
Anexo E: Curvas ROC modelo SVM.....	70
Anexo F: Métricas de evaluación SVM .....	72
Anexo G: Estructura red neuronal en Python usando Keras .....	73
Anexo H: Loss Function red neuronal.....	74
Anexo I: Curvas ROC usando Redes Neuronales.....	75
Anexo J: Métricas de evaluación Red Neuronal.....	77

## ÍNDICE DE TABLAS

Tabla 1: Producción de cobre según tamaño de empresa .....	2
Tabla 2: Actos inseguros en accidentes fatales .....	13
Tabla 3: Condiciones inseguras en accidentes fatales .....	13
Tabla 4: Stopwords NLTK .....	20
Tabla 5: Matriz de confusión .....	25
Tabla 6: Categorías de clasificación de OPS .....	34
Tabla 7: Número de OPS por faena según clasificación .....	39
Tabla 8: Selección de las 10 palabras más usadas que distinguen una categoría .....	39
Tabla 9: Resultados AUC por modelo .....	49
Tabla 10: Resultados Log Loss por modelo.....	49
Tabla 11: Comparación de modelos (ratio verdaderos positivos, ratio verdaderos negativos y accuracy).....	50

## ÍNDICE DE GRAFICOS

Gráfico 1: Porcentaje que representa el cobre del PIB de Chile .....	2
Gráfico 2: Número de trabajadores en faenas mineras en Chile .....	3
Gráfico 3: Ventas promedio por país .....	5
Gráfico 4: Número de empleados promedio por país .....	6
Gráfico 5: Cantidad de fallecidos en Chile por año .....	12
Gráfico 6: Número de OPS según clasificación .....	37
Gráfico 7: Número de OPS mensuales .....	38
Gráfico 8: Histograma largo de OPS .....	38

## ÍNDICE DE ILUSTRACIONES

Ilustración 1: Logo de la empresa.....	3
Ilustración 2: Proceso de gestión de controles críticos .....	17
Ilustración 3: Output regresión lineal vs output regresión Logística .....	22
Ilustración 4: SVM .....	23
Ilustración 5: Red Neuronal Artificial .....	24
Ilustración 6: Flujo de proceso para entrenar y evaluar clasificadores para identificar reportes por tipo y severidad.....	29
Ilustración 7: Esquema de clasificación .....	30