



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

DISEÑO Y CONSTRUCCIÓN DE MODELO DE CLASIFICACIÓN DE INCIDENTES
DE SEGURIDAD USANDO NLP EN LOS REGISTROS DE TEXTO ESCRITO PARA
AUTOMATIZAR ETIQUETACIÓN

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

PAMELA ANDREA MUÑOZ CÁCERES

PROFESORA GUÍA:
ASTRID CONTRERAS FUENTES

MIEMBROS DE LA COMISIÓN:
IGNACIO CALISTO LEIVA
ENRIQUE JOFRE ROJAS

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: Ingeniera Civil Industrial
POR: Pamela Andrea Muñoz Cáceres
PROFESORA GUÍA: Astrid Contreras
Fuentes
FECHA: 14/05/2020

DISEÑO Y CONSTRUCCIÓN DE MODELO DE CLASIFICACIÓN DE INCIDENTES DE SEGURIDAD USANDO NLP EN LOS REGISTROS DE TEXTO ESCRITO PARA AUTOMATIZAR ETIQUETACIÓN

La seguridad en minería es un tema de vital importancia, los accidentes son muchas veces fatales, lo que trae consecuencias económicas a las empresas y, más importante aún, afecta a todo el entorno familiar del involucrado. Por lo antes mencionado se toman diversas medidas de control para evitar cualquier tipo de accidente, entre ellas el registro de OPS, tema central del informe.

OPS es la abreviación de Observación Preventiva de Seguridad, son observaciones realizadas por trabajadores con respecto al comportamiento de otros en relación a la seguridad. La mayor parte de este registro corresponde a data no estructurada, específicamente, texto.

Actualmente cada vez que un trabajador registra una OPS, la clasifica en una categoría específica en función del comportamiento riesgoso encontrado. Dado que esta clasificación es manual, hace el trabajo más tedioso y quita tiempo al trabajador, por lo que se propone un sistema de clasificación automática para las OPS.

Para clasificar automáticamente se entrenan tres algoritmos: Logit, Support Vector Machine y una red neuronal. Los resultados muestran un buen desempeño de Logit en la mayoría de las categorías en comparación con los otros dos enfoques, aunque la red neuronal, tiene un rendimiento muy similar.

Los modelos demuestran que es posible clasificar automáticamente los textos OPS, siempre priorizando la identificación de verdaderos positivos, con el fin de identificar de manera correcta las conductas riesgosas, sin afectar de manera significativa el desempeño final, logrando clasificar correctamente en la categoría correcta en más del 80% de los casos.

A mis abuelos.

AGRADECIMIENTOS

Quiero agradecer a mi familia por haber tenido paciencia infinita a lo largo de toda mi carrera universitaria.

A mis amigas que me han acompañado durante todos estos años, dejando la vara alta en muchos aspectos, sin ustedes no sonreiría tantas veces al día.

A Tesseracto que me han acogido y brindado distintas oportunidades para desarrollarme tanto profesional como personalmente. Gracias por confiar y creer en mí.

A Natalia Fuentes, por darme ánimos durante todo este proceso y haberme ayudado a corregir este trabajo.

A Ingrid Cáceres, la Tita, por cumplir todos mis caprichos y corregir hasta el último error gramatical de mi memoria.

A mi compañero de la vida, Nicolas, eres el mejor y queda mucho por recorrer juntos.

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

CAPÍTULO 1: INTRODUCCIÓN

1.1 ANTECEDENTES GENERALES DE LA INDUSTRIA MINERA

La industria minera en Chile representa uno de los sectores más importantes para la economía. En el año 2018 la participación en el PIB nacional alcanzó un 9,8%, siendo la participación específica del cobre de un 8,9%, lo que en producción mundial de cobre corresponde a un 28,3%.

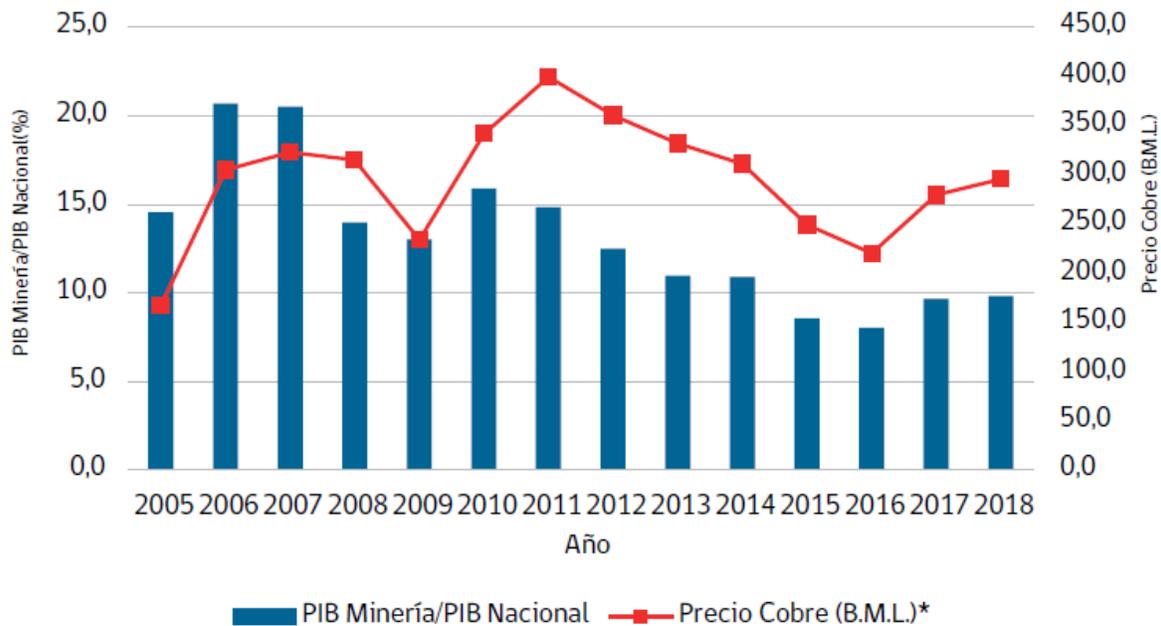
En Chile se explotan siete minerales metálicos, se considera un mineral metálico el contenido en la corteza terrestre que posea una concentración natural de un compuesto metálico con potencial de ser explotado. En el 2018 Chile ocupó el primer lugar mundial en la explotación de cobre con 5.872.027 toneladas métricas y el segundo lugar en la producción de molibdeno con 60.248 toneladas métricas. [1]

El cobre es un excelente conductor de electricidad y calor, además de resistir los microbios y la corrosión. Por estas y otras razones es una de las materias primas con mayor uso industrial en el mundo, considerándose sus variaciones de precio un indicador de la economía mundial.

El molibdeno, por otra parte, es utilizado en la fabricación de vehículos, electrodomésticos e infraestructura, este mineral no se encuentra en la naturaleza en estado natural, sin embargo, se puede extraer como un subproducto de otros minerales.

En el grafico que se encuentra a continuación se aprecia la proporción que alcanza la producción de cobre en el producto interno bruto nacional, cuyo punto más alto se registró en 2006, año en que dicha proporción superó el 20%.

Gráfico 1: Porcentaje que representa el cobre del PIB de Chile



Fuente 1: Anuario de la Minería de Chile

Las empresas mineras se categorizan como grandes, medianas o pequeñas, en función de las horas hombre que se registran en una faena minera. En el año 2018, las empresas grandes produjeron más de 5 millones de toneladas métricas finas de cobre, una cifra ampliamente superior a las 169 mil de empresas medianas y 49 mil de pequeñas empresas (Tabla 1).

Tabla 1: Producción de cobre según tamaño de empresa

TAMAÑO	Cu (tmf)
TOTAL	5.872.027
GRANDES	5.653.698
MEDIANAS	169.380
PEQUEÑAS	48.949

Fuente 2: Anuario de la Minería de Chile

Es importante también mencionar la magnitud de personas que trabajan en la industria minera de Chile, alcanzando en el año 2018 un total de 228.340, lo que representó un aumento del 5% con respecto al año anterior.

Gráfico 2: Número de trabajadores en faenas mineras en Chile



Fuente 3: Anuario de la Minería de Chile

1.2 ANTECEDENTES GENERALES DE LA EMPRESA

Tesseracto es una empresa dedicada a diseñar, desarrollar e implementar soluciones tecnológicas de advanced analytics e inteligencia artificial. El nombre proviene de un término matemático que hace referencia a un cubo en cuatro dimensiones (ilustración 1), siendo la meta final de la empresa encontrar una cuarta dimensión en mejoras de productividad y costos. [2]

Ilustración 1: Logo de la empresa



Fuente 4: Sitio web de la empresa

La empresa inició actividades en septiembre de 2018, actualmente se ubica en la comuna de Las Condes y solo en su primer año registró ingresos por 100.000 USD

La empresa, cuenta con un equipo multidisciplinario de 10 personas, siendo el Gerente General de la misma el señor José Tomas Cumsille. Con respecto al crecimiento futuro se proyectan ventas por 500.000 USD el primer año, 700.000 USD el segundo año y 1.000.000 USD el tercer año.

Tesseracto posee capacidades e infraestructura en 3 diferentes áreas necesarias para diseñar, construir e implementar tecnología de inteligencia artificial. La primera corresponde a las matemáticas, para construir algoritmos y herramientas de procesamiento estadístico avanzado. La segunda es la informática, para integrar sus soluciones a distintas arquitecturas de software-hardware según lo requiera el cliente. Finalmente, la experiencia en interfaz de usuario y gestión de cambio para que sus productos funcionen de forma sustentable y escalable generando los beneficios esperados.

1.2.1 DECLARACIONES ESTRATÉGICAS

Misión

Creamos tecnología de advanced analytics y machine learning para ver la dimensión oculta de los datos disponibles.

Visión

Ser líderes mundiales en el diseño e implementación de tecnologías que ayuden a las personas y organizaciones en la toma de decisiones considerando dimensiones más allá de conocidas en el cuidado de su salud y sostenibilidad.

1.2.2 SECTOR INDUSTRIAL

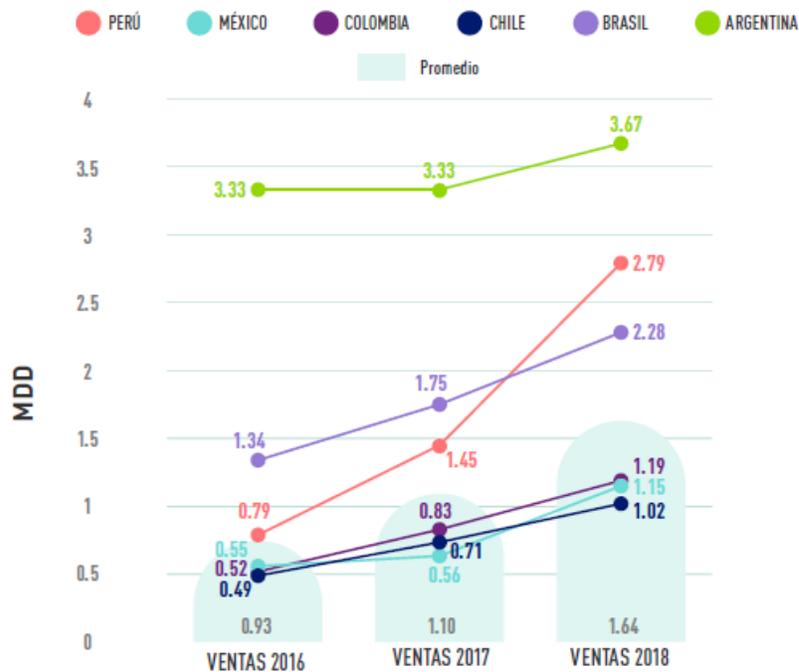
La empresa se desenvuelve en el sector de tecnologías de información en Chile, más específicamente en el de desarrollo de informática y software, enfocado principalmente en la inteligencia artificial y el uso de herramientas de machine learning.

Según un estudio publicado en el año 2018, dentro del campo de la inteligencia artificial en América latina el sector que más usa este tipo de

herramientas es el de desarrollo de software y herramientas para empresas (43%), seguido por el sector salud (17%). [3]

Según estadísticas de Endeavor y Everis sobre emprendimientos en América latina, las ventas se han incrementado en los últimos tres años. En el caso de Chile, estas se han aumentado de 0.49 millones de dólares en el año 2016, a 1.02 millones de dólares en el año 2018 (Gráfico 3).

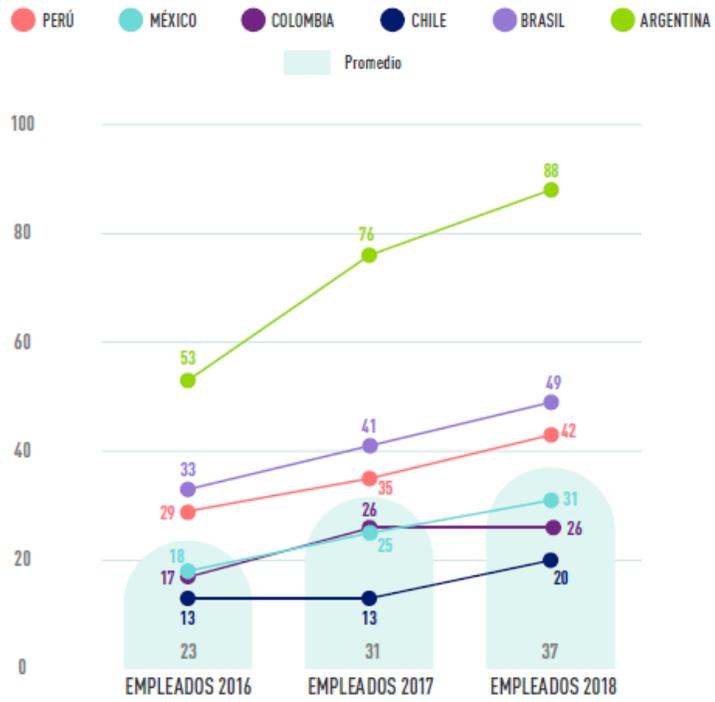
Gráfico 3: Ventas promedio por país



Fuente 5: El impacto de la IA en América Latina, Everis y Endeavor

Al igual que las ventas, el número de empleados promedio en emprendimientos enfocados en inteligencia artificial ha aumentado en Chile y en el resto de América latina. Para Chile este era 13 en el 2016 y luego 20 en el 2018 (Gráfico 4).

Gráfico 4: Número de empleados promedio por país



Fuente 6: El impacto de la IA en América Latina, Everis y Endeavor

CAPÍTULO 2: DESCRIPCIÓN DEL ESTUDIO

2.1 PROBLEMA

2.1.1 CONTEXTO

El proyecto se enmarca en la utilidad subyacente de los registros de texto escrito que se realizan en los controles de seguridad en minería.

Para gestionar los riesgos en minería se realizan controles con el objetivo de comprobar las medidas de seguridad que se deben cumplir. Actualmente hay tres tipos de registros escritos de texto que indican si se cumplen las condiciones de seguridad: las observaciones de verificación de control crítico (OVCC), las observaciones preventivas de seguridad (OPS) y las observaciones planeadas de las tareas (OPT).

Las OVCC son observaciones a condiciones de seguridad, las OPS se centran en las conductas de los trabajadores, mientras que las OPT buscan establecer si se cumple con el procedimiento estipulado. Todos estos registros son principalmente texto en donde un observador describe la situación de riesgo observada.

En particular, la magnitud de OPS, para una de las empresas mineras más grande en nuestro país puede alcanzar a ser de 15.000 registros mensuales. En estos momentos no existe ningún mecanismo capaz de analizar el texto que contiene cada una de estas, y todos los análisis entorno a estos datos se basa en la cantidad de registros por lugar, pero no en su contenido en particular.

El proyecto consiste en utilizar los registros de OPS y generar un modelo de clasificación de estos que permita a los tomadores de decisión entender de forma más inmediata el contenido general de estos a lo largo del tiempo.

2.1.2 JUSTIFICACIÓN

Actualmente las OPS se clasifican de manera manual por cada trabajador al momento de ingresarlas al sistema, un sistema de clasificación automático permitiría una estandarización en el proceso y además un ahorro de tiempo para cada trabajador.

Por otra parte, la empresa minera estudiada no es la única empresa minera que registra posibles incidentes en seguridad mediante texto escrito, por ende, crear un algoritmo de clasificación automático permitiría ponerlo en práctica en otra empresa con características similares de manera rápida y sencilla.

Por último, un algoritmo de clasificación sería capaz de entregar un posible clasificador para cualquier texto ingresado al sistema, por lo que los tomadores de decisión contarían con información acerca de todos los registros ingresados.

2.2 HIPÓTESIS Y ALTERNATIVAS DE SOLUCIÓN

2.2.1 HIPÓTESIS

(H1) Es posible encontrar patrones similares entre los registros de observaciones preventivas de seguridad (OPS)

(H2) Es posible encontrar un modelo que automatice la clasificación de observaciones preventivas de seguridad

2.2.2 ALTERNATIVAS DE SOLUCIÓN

- Se propone usar procesamiento de lenguaje natural para lograr transformar los textos de OPS a un lenguaje que pueda ser entendido por las computadoras.

- Se confeccionarán algoritmos de clasificación, como Logit, SVM y redes neuronales, usando los registros obtenidos del paso anterior.
- En base al desempeño de los algoritmos se entregará una recomendación acerca del sistema más efectivo para clasificar OPS automáticamente.

CAPÍTULO 3: OBJETIVOS

3.1 OBJETIVO GENERAL

“Diseñar y construir un modelo de clasificación de posibles accidentes para una gran empresa minera usando algoritmos de Natural Language Processing”

3.2 OBJETIVOS ESPECÍFICOS

- Establecer las formas actuales de clasificación accidentes en la industria.
- Establecer los algoritmos de Natural Language Processing que permitan construir un modelo predictivo.
- Construir un modelo capaz de clasificar los incidentes de seguridad, conocidos como Observaciones preventivas de seguridad, de manera automática.

CAPÍTULO 4: MARCO TEÓRICO

4.1 ACCIDENTABILIDAD EN MINERÍA

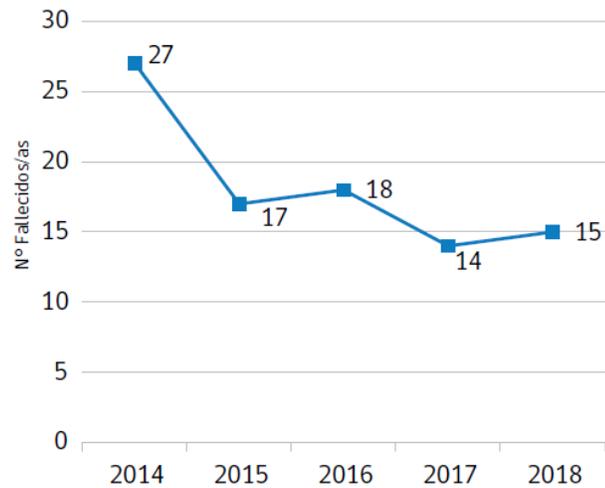
Según cifras del Anuario de la Minería en Chile en el total de faenas mineras en Chile para el año 2017 se desempeñaban 217.404 trabajadores. Por otra parte, dentro del mismo año se registraron 13 accidentes fatales, 47 accidentes graves y 712 accidentes que resultaron en tiempo perdido para el trabajador o fueron incapacitantes.

Cada uno de estos accidentes genera un impacto no solo para el trabajador involucrado, sino que, para todo su entorno familiar, causando en muchas ocasiones, daños importantes en la estructura familiar de los integrantes, ya sea en términos económicos, de salud, sociales, psicológicos, entre otros.

Para las empresas mineras, además de querer asegurar la seguridad de sus trabajadores, existen consecuencias económicas a raíz de los accidentes. Por cada accidente se debe detener la producción por algunos días y realizar una investigación con el fin de identificar las causas y futuras recomendaciones.

Por lo último es que se han modificado e instaurado nuevas medidas de seguridad con el fin de que se reduzca la tasa de accidentes en la industria, la que sin lugar a duda ha ido disminuyendo como muestra la figura a continuación.

Gráfico 5: Cantidad de fallecidos en Chile por año



Fuente 7: Anuario de la Minería de Chile

El análisis de accidentes en la industria actualmente conlleva el acto inseguro que realizó el trabajador en el accidente, y la condición insegura que había.

Con el primer análisis se desprende que las conductas de riesgo que han llevado a una mayor cantidad de accidentes en los últimos años son las de "No asegurar ni advertir el peligro" y "Colocarse en posición o postura peligrosa".

Tabla 2: Actos inseguros en accidentes fatales

ACTOS INSEGUROS O SUB ESTÁNDAR	TOTAL	2014	2015	2016	2017	2018
TOTAL	180	60	28	40	26	26
No asegurar ni advertir el peligro	59	19	11	11	9	9
Colocarse en posición o postura peligrosa	40	13	8	8	4	7
Actuar sin órdenes o desobedecer a éstas	26	12	4	6	1	3
Operar o trabajar a velocidades inseguras	10	4	1	2	3	0
No usar equipo de protección disponible	8	4	0	1	0	3
Neutralizar la operación de dispositivos de seguridad	7	2	0	0	5	0
Usar equipo inseguro	7	1	1	5	0	0
No se determinó acto inseguro	6	1	1	2	0	2
Falta de atención a superficies de apoyo y alrededores	5	1	0	1	3	0
Realizar trabajos para los cuales no está debidamente autorizado	3	1	0	1	0	1
No asegurar bloqueo del equipo	3	1	1	1	0	0
Error en la conducción	3	1	0	1	0	1
Reparar el equipo en movimiento	2	0	1	1	0	0
Usar herramienta insegura	1	0	0	0	1	0

Fuente 8: Anuario de la Minería de Chile

Mientras que las condiciones inseguras que más destacan en la lista de accidentes fatales son las de “métodos o procedimientos peligrosos” y “falta de resguardo o defensa inadecuada”.

Tabla 3: Condiciones inseguras en accidentes fatales

CONDICIONES INSEGUROS O SUBESTANDAR	TOTAL	2014	2015	2016	2017	2018
TOTAL	203	58	39	52	26	28
Método o procedimientos peligrosos	42	12	8	11	3	8
Falta de resguardo o defensa inadecuada	40	9	8	9	8	6
Defecto del terreno	20	3	3	6	5	3
Sistemas de advertencia inexistentes	18	7	1	5	4	1
Falta acuanadura, falta o fortificación inadecuada	16	9	1	3	1	2
Defecto de equipo	13	3	4	5	1	0
Defecto de materiales	8	1	3	4	0	0
Iluminación deficiente	7	3	1	2	0	1
Falta o insuficiencia de entrenamiento	7	1	3	2	0	1
Espacio reducido para realizar la tarea	6	1	1	0	1	3
Defecto de herramientas	6	2	0	2	1	1
Desarrollo incorrecto del proyecto de explotación	5	3	2	0	0	0
Defecto de camino	3	0	2	1	0	0
Falta de equipo de protección personal que era necesario	3	0	2	0	0	1
Limpieza y orden deficiente	2	2	0	0	0	0
Visibilidad deficiente	2	1	0	1	0	0

Fuente 9: Anuario de la Minería de Chile

4.2 SEGURIDAD EN MINERÍA

4.2.1 DIFERENCIA ENTRE PELIGRO Y RIESGO

Peligro

Se considera peligroso cualquier objeto o acción que tenga la potencialidad de causar daños en un individuo. Por ejemplo, el agua es un peligro porque tiene la potencialidad de que alguien pudiese ahogarse en ella. [4]

Riesgo

El riesgo es un indicador de la probabilidad de que algo peligroso pueda ocasionar un daño, en conjunto con las consecuencias que tendrían aquellos daños. Para calcular el riesgo se debe tener en cuenta las características ambientales, así como también el cómo y qué tanto una persona está expuesta al peligro.

4.2.2 NORMA OHSAS 18001

La norma OHSAS 18001, del inglés Occupational Health and Safety Management Systems, es un marco para un sistema de gestión de seguridad y salud desarrollado en el Reino Unido para mejorar los estándares de salud y seguridad que una empresa ofrece a sus empleados. [5]

Los beneficios que plantea la aplicación de esta norma son: la disminución de accidentes laborales, el logro de una percepción de entorno más seguro para los trabajadores, el ahorro de costos debido a bajas laborales y el cumplimiento con las regulaciones del país en que ésta se aplique.

Según la norma, para decidir si un riesgo es aceptable se debe realizar un adecuado análisis, con el fin de estimar su probabilidad y magnitud, lo que busca este análisis es entender las causas, posibles consecuencias, determinar una posible frecuencia y definir los procedimientos y medidas de prevención. [6]

Para la identificación de riesgos existen métodos cuantitativos y cualitativos. Los primeros buscan entregar una valoración numérica con respecto a la frecuencia de determinado suceso, mientras que los métodos cualitativos no usan cálculos numéricos y se pueden subdividir en dos: métodos comparativos y métodos generalizados.

Los métodos comparativos se basan en técnicas obtenidas a través de experiencias anteriores, ejemplos de estos son los manuales técnicos que indiquen el funcionamiento de los equipos, las listas de comprobación en donde se determina la adecuación a los equipos y procedimientos establecidos y, también, el análisis de los accidentes histórico en situaciones similares.

Por último, los métodos generalizados, llevan consigo un análisis más profundo y lógico de la deducción de errores y fallos en equipos, procedimientos y operaciones

4.2.3 INDICADORES DE ACCIDENTABILIDAD

Para poder comparar, ya sea a lo largo del tiempo o entre la industria, la accidentabilidad se usan los siguientes indicadores:

LTIFR (Lost Time Injury Frequency Rate)

Este indicador se calcula como el número de LTI (Lost Time Injury) multiplicado por 1 millón y dividido por el número de horas hombre en el periodo reportado. El termino LTI se refiere a cualquier lesión o enfermedad que haya significado al menos un día de trabajo perdido. [7]

TRIFR (Total Recordable Injury Frequency Rate)

A diferencia de la métrica anterior este indicador incluye todas las lesiones o enfermedades que requieren tratamiento de un profesional. El indicador se calcula como el número de fatalidades sumado a la cantidad de accidentes, multiplicado por un millón y dividido por la cantidad de horas hombre en el periodo. [8]

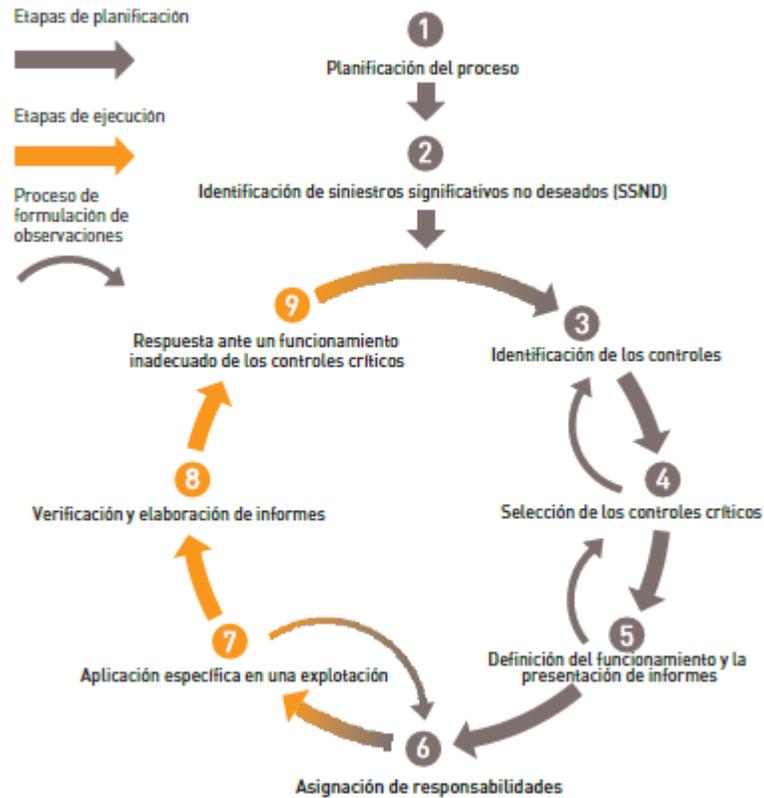
4.2.4 GESTIÓN DE CONTROLES CRÍTICOS

Se considera control un acto, objeto o sistema cuya finalidad sea prevenir un incidente no deseado. Los controles críticos son aquellos que resultan esenciales para prevenir un accidente o al menos mitigar las consecuencias de uno, usualmente un control que impida más de un siniestro o mitigue más de una consecuencia se considera un control crítico. [9]

El proceso de gestión de controles críticos consta de 9 etapas, las primeras buscan definir el proceso a ejecutar, identificar los siniestros no deseados y los controles adecuados. Con posterioridad se define el funcionamiento, los reportes a usar y la asignación de responsabilidades. Finalmente, se verifican los controles, se crean informes y se genera una respuesta a los funcionamientos inadecuados. [10]

El paso de verificación de controles críticos considera informes acerca del estado de éstos, pero también incluye un proceso de supervisión activa, el que puede llevarse a cabo mediante la verificación de registros, el acompañamiento a operarios en inspecciones rutinarias o la reevaluación periódica de requisitos de desempeño.

Ilustración 2: Proceso de gestión de controles críticos



Fuente 10: Guía para la ejecución de la gestión de los controles críticos

4.2.5 OBSERVACIONES CONDUCTUALES EN PREVENCIÓN

Las observaciones conductuales consisten en realizar observaciones en terreno del trabajo que diariamente realizan los trabajadores para entregar una recomendación en el lugar que contenga propuestas de mejora para el trabajador. Se espera que de ser bien implementadas sea posible aminorar los accidentes, ya que el trabajador puede entender de mejor forma la tarea a realizar y cómo realizarla.

Por otra parte, también se cree que este tipo de observaciones logran aumentar el grado de compromiso de los trabajadores con la seguridad, ya que las observaciones las realiza gran parte del personal permitiendo así hacer partícipes a cada trabajador de la seguridad de sus compañeros. [11]

OBSERVACIONES PREVENTIVAS DE SEGURIDAD

Las OPS son un tipo de observación conductual, estas se realizan en una de las más grandes empresas mineras en Chile, con el fin de detectar malas conductas y entregar un feedback a sus trabajadores.

Para esta empresa el registro debe contar con la siguiente estructura:

- 1- Observación: Enfocada en el actuar de los trabajadores siempre con el permiso de la persona y sin causar distracciones
- 2- Detención de actividad: Se solicita a la o las personas involucradas detener la actividad para que éstos expliquen el procedimiento y los riesgos involucrados
- 3- Comentarios positivos y negativos: Se realizan comentarios positivos y negativos respecto a la conducta observada.
- 4- Compromiso: Se obtiene el compromiso de la persona observada a trabajar de forma segura
- 5- Agradecimientos: Se agradece la participación del trabajador.

4.3 PROCESAMIENTO DE TEXTO

4.3.1 TOKENIZATION

Proceso en el que los textos se dividen en tokens, típicamente constituidos por palabras individuales. El objetivo es dividir el texto en pequeños fragmentos significativos.

Por ejemplo, en la oración "*Se observa trabajador no usando dispositivos de seguridad*" al aplicar este procedimiento quedaría como:

['Se', 'observa', 'trabajador', 'no', 'usando', 'dispositivos', 'de', 'seguridad']

4.3.2 LIMPIEZA Y ESTANDARIZACIÓN

Al trabajar con palabras en programación se consideran palabras diferentes aquellas que contengan mayúsculas en vez de minúsculas, por lo que es necesario estandarizar las palabras para finalmente solo usar palabras en minúsculas.

En esta parte del proceso también se remueven signos de puntuación, ya que éstos no aportan valor a las conclusiones futuras.

Continuando con el ejemplo anterior:

['se', 'observa', 'trabajador', 'no', 'usando', 'dispositivos', 'de', 'seguridad']

4.3.3 STOPWORDS

No todas las palabras en un texto proveen de un significado como tal, las palabras que no poseen significado se denominan stopwords.

A continuación, se presenta la lista de stopwords en el idioma español según la librería de procesamiento de lenguaje natural NLTK

Tabla 4: Stopwords NLTK

de, la, que, el, en, y, a, los, del, se, las, por, un, para, con, no, una, su, al, lo, como, más, pero, sus, le, ya, o, este, sí, porque, esta, entre, cuando, muy, sin, sobre, también, me, hasta, hay, donde, quien, desde, todo, nos, durante, todos, uno, les, ni, contra, otros, ese, eso, ante, ellos, e, esto, mí, antes, algunos, qué, unos, yo, otro, otras, otra, él, tanto, esa, estos, mucho, quienes, nada, muchos, cual, poco, ella, estar, estas, algunas, algo, nosotros, mi, mis, tú, te, ti, tu, tus, ellas, nosotras, vosotros, vosotras, os, mío, mía, míos, mías, tuyo, tuya, tuyos, tuyas, suyo, suya, suyos, suyas, nuestro, nuestra, nuestros, nuestras, vuestro, vuestra, vuestros, vuestras, esos, esas, estoy, estás, está, estamos, estáis, están, esté, estés, estemos, estéis, estén, estaré, estarás, estará, estaremos, estaréis, estarán, estaría, estarías, estaríamos, estaríais, estarían, estaba, estabas, estábamos, estabais, estaban, estuve, estuviste, estuvo, estuvimos, estuvisteis, estuvieron, estuviera, estuvieras, estuviéramos, estuviérais, estuvieran, estuviese, estuvieses, estuviésemos, estuvieseis, estuviesen, estando, estado, estada, estados, estadas, estad, he, has, ha, hemos, habéis, han, haya, hayas, hayamos, hayáis, hayan, habré, habrás, habrá, habremos, habréis, habrán, habría, habrías, habríamos, habrías, habrían, había, habías, habíamos, habíais, habían, hube, hubiste, hubo, hubimos, hubisteis, hubieron, hubiera, hubieras, hubiéramos, hubierais, hubieran, hubiese, hubieses, hubiésemos, hubieseis, hubiesen, habiendo, habido, habida, habidos, habidas, soy, eres, es, somos, sois, son, sea, seas, seamos, seáis, sean, seré, serás, será, seremos, seréis, serán, sería, serías, seríamos, seríais, serían, era, eras, éramos, erais, eran, fui, fuiste, fue, fuimos, fuisteis, fueron, fuera, fueras, fuéramos, fuerais, fueran, fuese, fueses, fuésemos, fueseis, fuesen, sintiendo, sentido, sentida, sentidos, sentidas, siente, sentid, tengo, tienes, tiene, tenemos, tenéis, tienen, tenga, tengas, tengamos, tengáis, tengan, tendré, tendrás, tendrá, tendremos, tendréis, tendrán, tendría, tendrías, tendríamos, tendríais, tendrían, tenía, tenías, teníamos, teníais, tenían, tuve, tuviste, tuvo, tuvimos, tuvisteis, tuvieron, tuviera, tuvieras, tuviéramos, tuvierais, tuvieran, tuviese, tuvieses, tuviésemos, tuvieseis, tuviesen, teniendo, tenido, tenida, tenidos, tenidas, tene

Fuente 11: Elaboración propia

Finalmente aplicando este procedimiento al texto de ejemplo:

['observar', 'trabajador', 'usar', 'dispositivo', 'seguridad']

4.3.4 TF-IDF

Algoritmo que es el resultado de la multiplicación de dos otros algoritmos. Estos algoritmos son Term Frequency (TF) y Inverse Document Frequency (IDF).

El primer algoritmo indica la frecuencia de una palabra en un documento con respecto al largo de este mismo, es decir, la cantidad de veces que aparece una palabra en un documento dividido por el total de palabras en este.

$$tf(t, d) = \frac{\text{Número de ocurrencias de un término en el documento}}{\text{Número de palabras totales en el documento}}$$

El segundo algoritmo (IDF) evalúa que tan única es una palabra. Para esto se utilizan múltiples documentos y se calcula el cociente entre el número total de documentos y el número de documentos que tienen el término en cuestión.

$$idf(t, d) = \log \left(\frac{\text{Número de ocurrencias de un término en el documento}}{\text{Número de documentos que contienen el término}} \right)$$

4.4 MACHINE LEARNING

Existen diversos algoritmos que permiten crear modelos de clasificación de texto, entre ellos, Logit, SVM, Random Forest, XGboost y redes neuronales [12]. En este trabajo se utilizan tres de ellos: Logit, SVM y redes neuronales, la descripción de cada uno de ellos se presenta a continuación.

4.4.1 LOGIT

Algoritmo de clasificación supervisada que busca predecir el resultado de una variable categoría en función de múltiples parámetros.

Logit representa uno de los algoritmos más simples para poder clasificar problemas que se compongan de dos grupos distintos. [13]

Si se considera una regresión típica de la forma:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Donde y corresponde a la variable dependiente.

Aplicando la función Sigmoid, que es:

$$p = \frac{1}{1 + e^{-y}}$$

Tenemos que

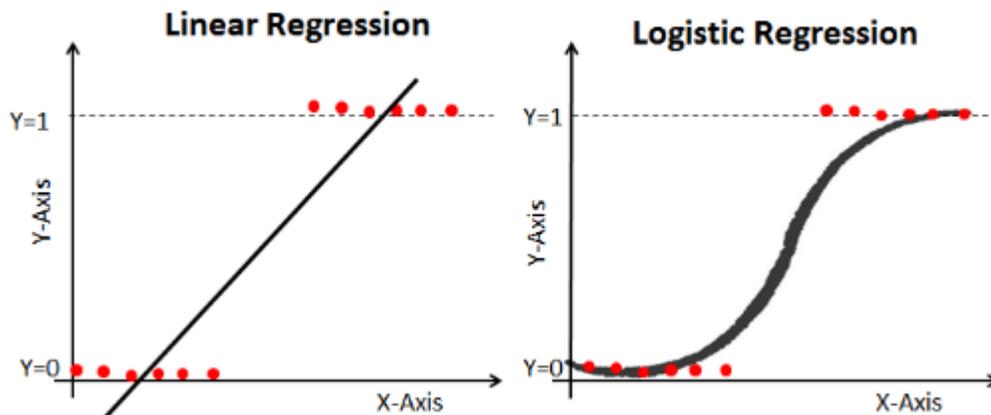
$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}}$$

Las probabilidades que entrega el algoritmo se modelan de la siguiente forma:

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i}$$

La principal diferencia con un modelo de regresión lineal es que la regresión logística entrega un resultado constante, mientras que la regresión lineal un resultado continuo, esto porque la primera utiliza el método de máxima verosimilitud y la segunda el de mínimos cuadrados ordinarios.

Ilustración 3: Output regresión lineal vs output regresión Logística



Fuente 12: Understanding Logistic Regression in Python

El modelo Logit es extensible a la clasificación de múltiples clases a la vez, en esta variante cada categoría se compara de a una contra el resto de las categorías, es decir estima de manera separada

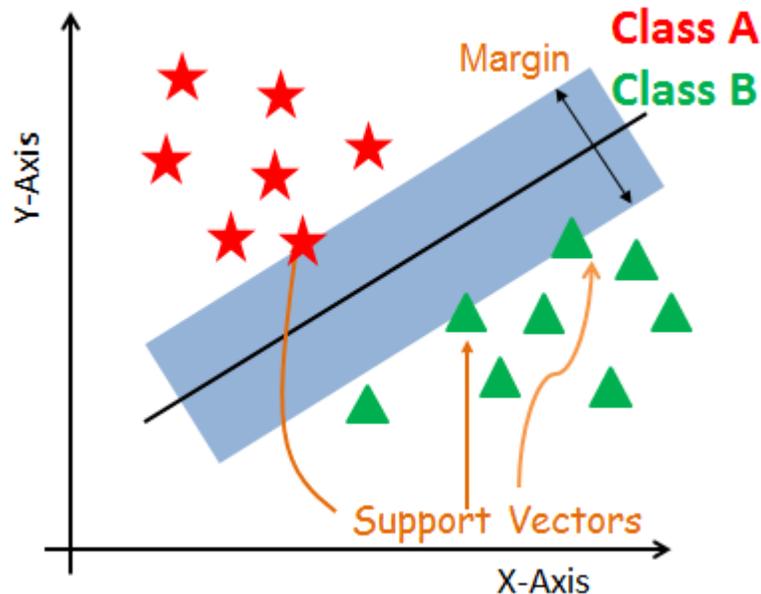
4.4.2 SUPPORT VECTOR MACHINE

Support vector machine (SVM), es un algoritmo de clasificación en donde las clases se dividen mediante a un hiperplano multidimensional. El objetivo es encontrar el hiperplano que mejor divide la data en distintas clases.

Se distinguen tres elementos de este algoritmo: el hiperplano, los vectores de soporte y el margen. El hiperplano es el límite de decisión entre clases, por ende, separa datos en función de la clase a la que corresponden.

Los vectores de soporte son aquellos datos que están cerca del hiperplano o que están mal clasificados. Por último, el margen es un espacio formado por dos hiperplanos paralelos al del límite de decisión y se calcula como la distancia perpendicular entre el límite de decisión y el punto más cercano.

Ilustración 4: SVM



Fuente 13: Support Vector Machines with Scikit-learn

Para lograr el objetivo el algoritmo genera los hiperplanos que logran segregar clases de mejor forma. Luego se maximiza el margen del plano usando los vectores de soporte donde, finalmente, el hiperplano definitivo será aquel que logre el mayor margen. [14]

4.4.3 REDES NEURONALES ARTIFICIALES

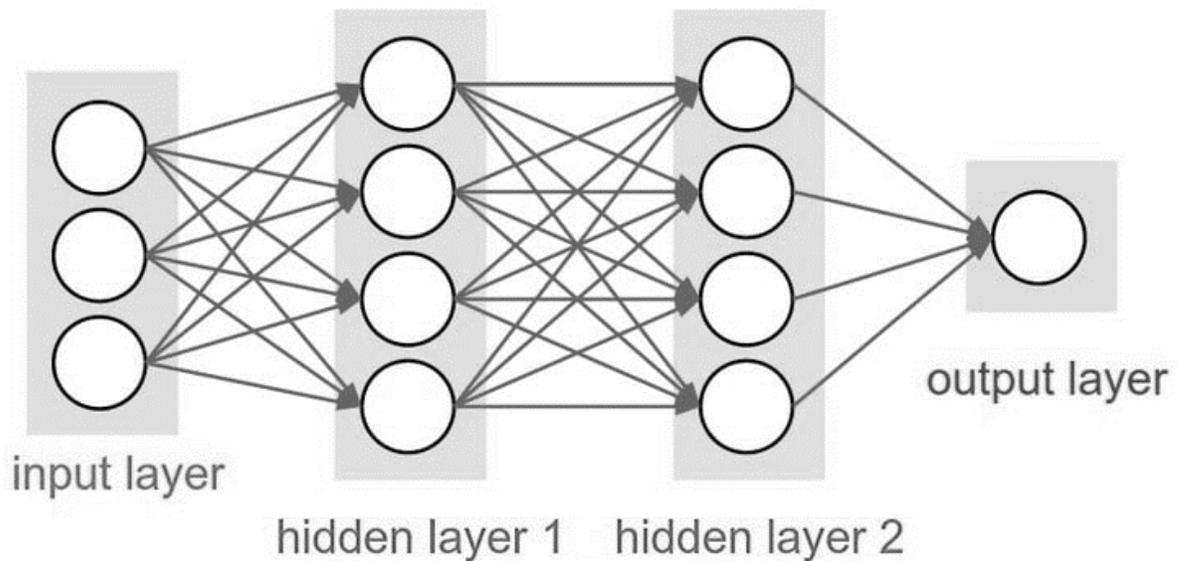
Los algoritmos conocidos como redes neuronales están inspirados en cómo funciona el cerebro, específicamente en la forma en que el cerebro humano aprende. [15]

Las redes neuronales funcionan como cajas negras en donde cada fragmento de información es analizado por pequeñas unidades (neuronas), divididas en distintas capas que se conectan entre sí. Las conexiones se

conocen como vectores y estos son ajustados en cada iteración del algoritmo. [16]

Toda red neuronal se compone de una capa de entrada que consiste en la data que caracteriza los registros a entrenar, luego, se encuentran la capa oculta (hidden layer), que podría ser más de una. Por último, al final de la estructura, se encuentra la capa de salida, en ésta, para los problemas de clasificación, existe un nodo por cada clase a analizar.

Ilustración 5: Red Neuronal Artificial



Fuente 14: What is an artificial neural network? Here's everything you need to know

Una arquitectura de red se puede describir en función de su ancho y profundidad. Se conoce como el ancho de una capa a la cantidad de nodos en ésta, mientras que el ancho de la red se considera el mayor de estos. Por otra parte, el largo de la red se refiere al número de capas que ésta contiene. Estos valores son conocidos como hiperparámetros y son valores que se pueden escoger manualmente cuando se diseña una red.

A diferencia de otros modelos, las redes neuronales conllevan un componente de no linealidad en su algoritmo, esto se logra gracias a las funciones de activación, existiendo distintas funciones de activación y siendo algunas de las más utilizadas ReLu, softmax y sigmoid.

El proceso de entrenamiento de una red conlleva múltiples iteraciones hasta alcanzar el resultado óptimo. Cada iteración lleva dos etapas: Forward Propagation y Backpropagation.

El proceso de Forward Propagation corresponde al paso de la data de entrada a través de la red para calcular el resultado final. En el proceso de Backpropagation se calcula el gradiente, esto consiste en determinar la contribución de cada parámetro al error obtenido y actualizar los vectores para un mejor funcionamiento. [17]

4.5 MÉTRICAS DE EVALUACIÓN

4.5.1 LOGARITHMIC LOSS

Logarithmic Loss (Log Loss) es una métrica de comparación de modelos que está basada en la función de verosimilitud. Los valores que entrega esta métrica no pueden interpretarse por sí solos, aunque siempre un menor valor implica un mejor modelo. [18]

4.5.2 MATRIZ DE CONFUSIÓN

La matriz de confusión es una tabla que se usa para poder medir el desempeño de un modelo de clasificación. Esta tabla, se divide en filas según la etiqueta real y en columnas en función del valor predicho. [19]

A continuación, se presenta un ejemplo de matriz de confusión donde se considera un total de 165 registros.

Tabla 5: Matriz de confusión

		Valor Predicho	
		NO	SI
Valor real	NO	50	10
	SI	5	100

Fuente 15: Elaboración propia

Los elementos dentro de una matriz de confusión son los siguientes:

- 1- Verdaderos positivos (VP): Número de casos en los que el modelo predijo que pertenecía a la clase y así era.
- 2- Falsos positivos (FP): Número de casos en que se predice que pertenece a la clase, pero realmente no era así.
- 3- Verdaderos negativos (VN): Número de casos en que no pertenece a la clase y el modelo efectivamente predice eso.
- 4- Falsos negativos (FN): Número de casos en que se predice que no pertenece a la clase, cuando si pertenecía.

Algunas métricas que derivan de la matriz de confusión son:

- 1- Accuracy: Mide con qué frecuencia un clasificador está correcto.

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN}$$

- 2- Tasa de error: Mide con qué frecuencia un clasificador esta incorrecto.

$$Tasa\ de\ error = \frac{FP + FN}{VP + VN + FP + FN}$$

- 3- Ratio de verdaderos positivos: Mide cuantas de las veces en que un registro pertenecía a la clase el modelo predijo exactamente eso.

$$Ratio\ de\ verdaderos\ positivos = \frac{VP}{VP + FN}$$

- 4- Ratio de verdaderos negativos: Mide cuantas de las veces en que un registro no pertenecía a la clase el modelo predijo exactamente eso.

$$\text{Ratio de verdaderos negativos} = \frac{VN}{VN + FP}$$

5- Ratio de falsos positivos: Mide cuantas de las veces en donde no pertenece a la clase, el modelo predice que sí.

$$\text{Ratio de falsos positivos} = \frac{FP}{FP + VN}$$

6- Precision: Mide cuantas veces de las que el modelo predice que si, efectivamente era así.

$$\text{Precision} = \frac{VP}{VP + FP}$$

7- Prevalence: Mide que tanto la condición de pertenecer a la clase ocurre en la muestra.

$$\text{Prevalence} = \frac{VP + FN}{VP + VN + FP + FN}$$

4.5.3 AREA BAJO LA CURVA

La curva ROC es una forma de mostrar el desempeño de un clasificador binario de manera visual.

Los algoritmos de clasificación como la regresión logística no solo son capaces de discernir la clase a la que pertenece el registro, sino que también pueden entregar la probabilidad con la que esto pasa.

Se le denomina "threshold" a la probabilidad mínima con la que un registro será admitido en la clase, comúnmente se utiliza un threshold de 0.5 lo que significa que todos aquellos registros que tengan una probabilidad igual o mayor a 0.5 serán clasificados como parte de la clase.

La curva ROC, representa la tasa de verdaderos positivos en función de la tasa de falsos positivos para cada valor posible de threshold. Mientras mayor sea el área bajo la curva, mejor será la performance del clasificador. [20]

4.5.4 F1 SCORE

La métrica F1 score es una función del ratio de verdaderos positivos (recall) y la precisión (precisión) del clasificador, su uso se recomienda cuando se tiene un problema de clasificación con clases desbalanceadas. [21]

La fórmula es la siguiente:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

4.6 TRABAJOS SIMILARES

La clasificación de texto es un procedimiento que puede ser útil en múltiples escenarios dado la gran cantidad de data no estructurada que se genera día a día.

Uno de los primeros problemas que se intentó resolver con esta tecnología fue la clasificación de spam en el correo electrónico, que incluso hoy continúa siendo revisado y mejorado por múltiples autores. [22]

Desde entonces han surgido otros trabajos que proyectan poder etiquetar texto en más de dos categorías, ejemplos de esto son los algoritmos que intenta predecir el sentimiento de las personas detrás de un texto [23] [24], e incluso algunos que buscan mejorar el uso de texto en los reportes médicos. [25] [26]

Dado que en este trabajo se busca categorizar texto en múltiples categorías, siendo posible que estos pertenezcan a mes de una a la vez, los trabajos que mas se asemejan son los que se detallan a continuación:

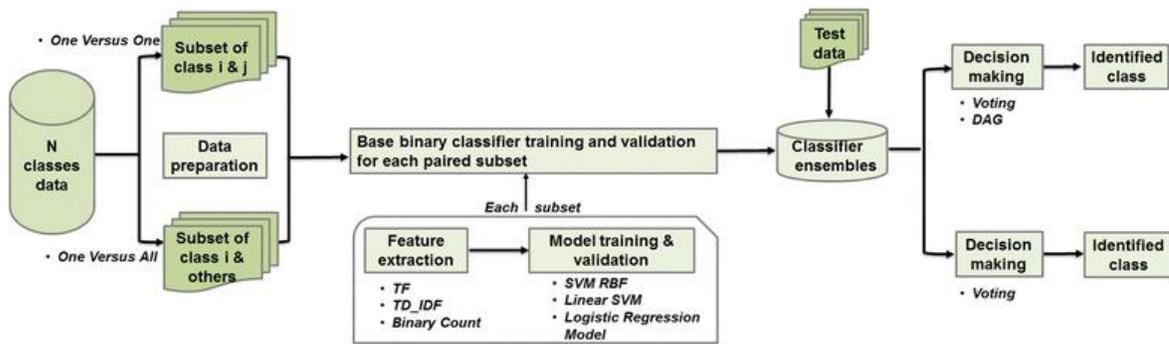
CLASIFICACIÓN DE INCIDENTES DE SEGURIDAD

En este trabajo, los autores buscan realizar una clasificación de incidentes de seguridad para todas las admisiones a un hospital que corresponden a eventos adversos, en donde el paciente resultó o pudo haber resultado lesionado. [27]

El objetivo de esta clasificación es reconocer el problema de seguridad subyacente y para ello es necesario entender el tipo de incidente.

Los autores utilizan algoritmos de clasificación basados en Naive Bayes, como son Logit y SVM, para clasificar los incidentes en 10 clases distintas. Para evaluar la performance de los clasificadores utilizan la medida F-score.

Ilustración 6: Flujo de proceso para entrenar y evaluar clasificadores para identificar reportes por tipo y severidad



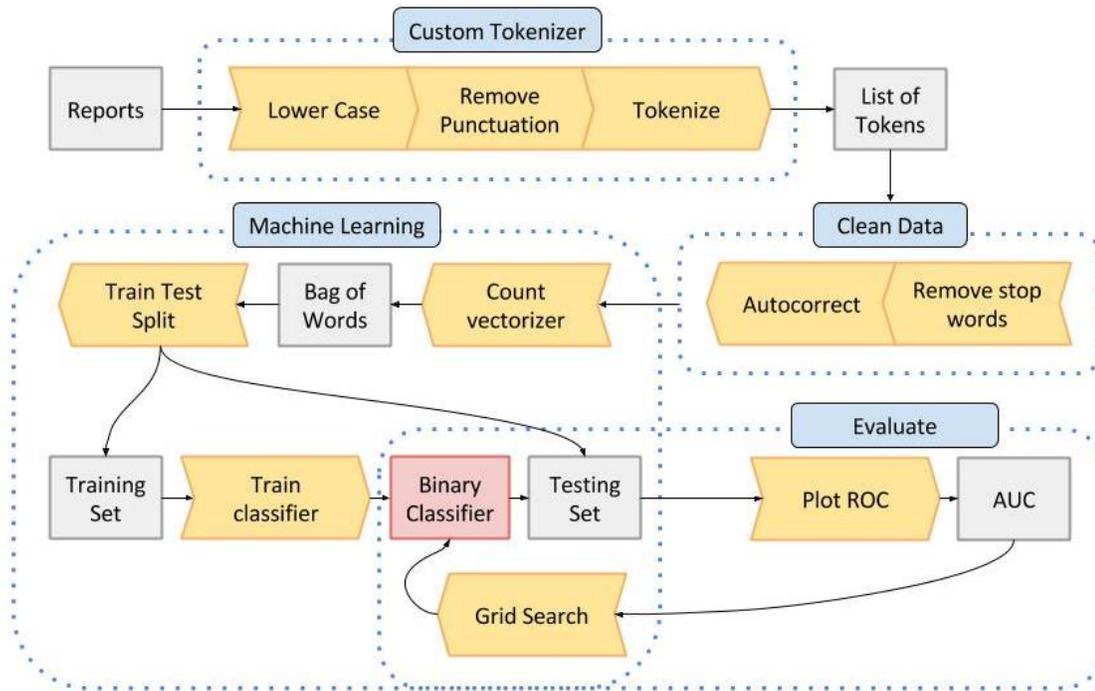
Fuente 16: Using multiclass classification to automate the identification of patient safety incident reports by type and severity

CLASIFICACIÓN DE REPORTES DE INCIDENTES EN RADIOLOGÍA

En este trabajo se pretende clasificar los reportes de radiología, que están compuestos de texto, según la etapa en la que se encuentra el paciente. [28]

Para clasificar estos reportes de incidentes se utilizan clasificadores bayesianos y luego se evalúa el desempeño de estos mediante curvas ROC.

Ilustración 7: Esquema de clasificación



Fuente 17: Natural language processing applied to incident reports in radiation oncology—determination of completeness and classification

CAPÍTULO 5: METODOLOGÍA

5.1 ANALISIS DE LA SITUACIÓN ACTUAL

En secciones anteriores se define la forma de clasificación de incidentes de seguridad actual, para tener un mejor entendimiento sobre la clasificación final de cada OPS será necesaria una evaluación como insumo para el diseño experimental.

5.2 PREPARACIÓN DE BASES DE DATOS NECESARIAS

Los datos a utilizar son los registros de OPS desde julio de 2018 a febrero de 2019 para cuatro operaciones mineras en nuestro país.

Dichos registros se encuentran en bases de datos distintas, las cuales deben unirse en una única fuente de datos, en este procedimiento se debe constatar que los registros sean semejantes para que no existan problemas de formato.

Los registros de OPS suelen incluir el nombre y el lugar de trabajo de quien hizo la observación, el lugar donde se hizo la observación, la fecha de observación, el texto asociado y la respectiva clasificación.

Se deben eliminar todos aquellos registros que no contengan texto, o no se le haya determinado una clasificación, tampoco serán analizados los textos que se encuentren en un idioma distinto al español y, por último, aquellos que no contengan suficientes palabras.

5.3 PROCESAMIENTO Y TRANSFORMACIÓN DE LA DATA

La base resultante debe someterse a una limpieza, con el fin de trabajar con bases aptas para el análisis y la vectorización de texto. En específico, será necesario eliminar símbolos de puntuación, transformar toda la data a texto en minúscula y eliminar aquellos textos que no representen más de 10

palabras, ya que se subentiende que no explica de manera correcta a situación, ni cumple con los requisitos de una OPS.

Luego del proceso de limpieza se procede a vectorizar los textos en función de dos algoritmos, TF_IDF y Bag of words, se espera poder concluir cuál es el más apto para el modelo de clasificación.

5.4 MODELAMIENTO

Dado que una OPS puede pertenecer a más de una categoría, no es posible utilizar un enfoque de modelamiento en donde se evalúen todas las clases a la vez, sino que es necesario crear un modelo de clasificación para cada una de las seis clases a evaluar, en donde se indica si el texto respectivo pertenece o no la clase respectiva.

Lo anterior se aplica para Logit y SVM, sin embargo, las redes neuronales sí permiten crear un modelo que pueda clasificar un texto para seis clases al mismo tiempo.

5.5 EVALUACIÓN

Para evaluar el desempeño de los distintos algoritmos se utilizarán las métricas descritas en la sección anterior.

También es importante evaluar que el resultado de la data de entrenamiento sea similar al de la data de prueba, ya que esto comprobará que el modelo es capaz de reconocer los patrones subyacentes sin sobre ajustar el resultado.

CAPÍTULO 6: ALCANCES Y RESULTADOS ESPERADOS

6.1 ALCANCES

Como se indica anteriormente el registro de texto que se usará en esta memoria es el de OPS, dejando fuera del análisis los registros de OPT y OVCC.

El modelo de clasificación no logrará incorporar futuros cambios, por lo que cualquier análisis de registros futuros contemplará la clasificación según los clusters del modelo y no podrá generar nuevas clasificaciones.

También quedarán fuera de todo análisis aquellos registros que no estén escritos en el idioma español, dejando mayoritariamente fuera registros de OPS en inglés.

Por otra parte, el desarrollo del modelo no incluye una implementación o integración con la empresa, por lo que cualquier uso posterior depende exclusivamente de decisiones de la empresa.

6.2 RESULTADOS ESPERADOS

Los resultados esperados para el desarrollo de esta memoria son los siguientes

- 1- **Modelo de clasificación de texto para cada categoría:** Se requiere realizar un modelo de clasificación que permita determinar si un registro en particular pertenece a una categoría específica.
- 2- **Modelo de clasificación de texto:** En función de lo anterior se pretende generar un modelo capaz de predecir a qué categoría pertenece un texto en particular.

CAPÍTULO 7: IMPLEMENTACIÓN DE METODOLOGÍA

7.1 LEVANTAMIENTO DE LA SITUACIÓN ACTUAL

Actualmente la clasificación de Observaciones Preventivas de Seguridad la realiza cada trabajador al momento de crear el registro en el sistema.

El trabajador solo clasificará la OPS si es que ésta representa una situación riesgosa, en caso contrario solo ingresara el texto que describe la situación.

Una OPS puede ser clasificada en 30 categorías distintas (Ver Anexo A), aunque estas categorías se pueden agrupar en 6 de mayor envergadura, las que se alistan a continuación.

Tabla 6: Categorías de clasificación de OPS

EPP
Línea de fuego/Posición de persona
Área/Entorno de trabajo
Manejo manual
Procedimiento/Evaluación de Riesgo
Herramientas y Equipos

Fuente 18: Elaboración propia

No es necesario que una OPS pertenezca solo una categoría, ya que una misma observación podría evaluar distintos aspectos de seguridad en el trabajo.

7.2 CONSOLIDACIÓN DE LA DATA

Para consolidar los datos en una base de datos única, se concatena todas las bases de dato en función de sus columnas en común, una vez hecho eso, se procede a revisar el formato de los textos en ella.

En primera instancia, la base de datos resultante cuenta con 111.673 registros de OPS, donde 1.006 registros estaban en un idioma distinto al español, siendo éste mayoritariamente inglés (840 casos). Para detectar el idioma de cada OPS se utilizó la librería de Python Langdetetc, que analiza textos y devuelve el idioma asociado. [29]

De los 110.667 registros restantes, 44.510 se asocian a una conducta riesgosa, por ende, tienen una clasificación respecto a ésta y pueden ser analizados.

7.3 PROCESAMIENTO Y TRANSFORMACIÓN DE LA DATA

7.3.1 ANALISIS EXPLORATORIO

Tal como se menciona antes, las OPS pueden ser clasificadas en 30 categorías distintas, las que a su vez se enmarcan en 6 grupos diferentes, las que, para efectos de este trabajo, serán las 6 categorías a analizar, éstas son las siguientes:

1- EPP (Elementos de protección personal)

En esta categoría se abordan los problemas relacionados con elementos de protección personal. Se incluye, la ausencia de estos, el mal uso o la mala elección por parte del trabajador, ejemplos son la falta de protección auditiva, cascos o lentes, la mala elección de lentes apropiados a la tarea y el mal posicionamiento de mascarillas respiratorias.

Las subcategorías especifican cuál es la parte del cuerpo que se ve sin la debida protección (cabeza, cara, oídos, cuerpo, pies).

2- Línea de fuego/Posición de persona

Esta categoría hace patente cuando los trabajadores se ubican en una posición riesgosa habiendo una alternativa apropiada para la realización

de la tarea, algunos ejemplos son el movimiento por zonas no permitidas y el traspaso de barreras establecidas

Las subcategorías describen la consecuencia que puede tener la persona (golpeado, atrapado, caída de altura) y la falta de atención a la tarea o el entorno.

3- Área o entorno de trabajo

Esta categoría alerta sobre malas condiciones en el entorno de trabajo en que se desempeñan los trabajadores. Se detallan problemas de suciedad, desorden, mala ubicación de los equipos, falta de iluminación, mala calidad del aire y falta de segregación, ejemplos son la existencia de materiales o equipos desperdigados, y la falta de iluminación al momento del trabajo.

4- Manejo Manual

Esta categoría refleja malas prácticas en el manejo de una tarea. En especial, se destaca el levantamiento de carga, posturas inadecuadas y problemas de espacio al momento de realizar labores.

5- Procedimiento/Evaluación de riesgo

Esta categoría resalta problemas en el procedimiento a realizar y en el conocimiento o evaluación de los riesgos asociados. Los textos que se clasifican de esta manera indican situaciones en donde: no se sigue el procedimiento (ya sea porque no se conoce o no se quiere), el trabajador no entiende el riesgo al que se expone y no se cumplen las condiciones necesarias para realizar la labor correspondiente.

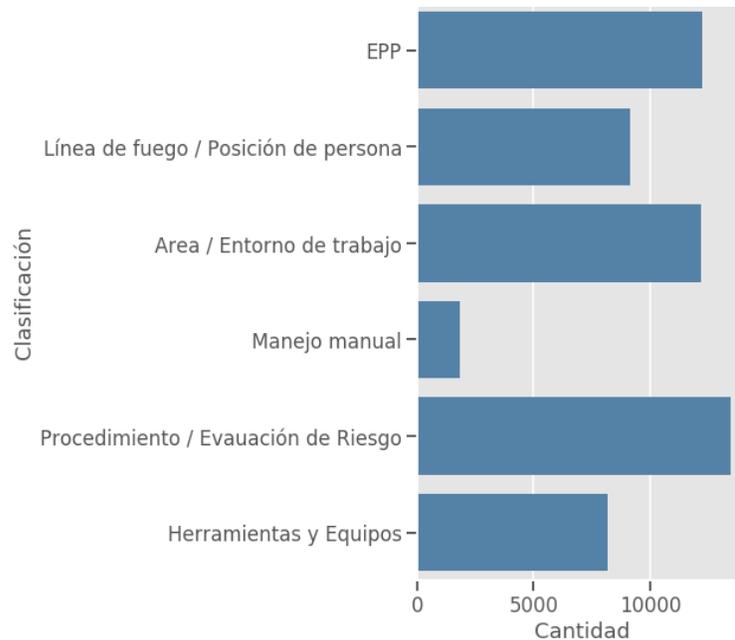
6- Herramientas y equipos

En esta categoría se describen aquellas situaciones en que el trabajador desempeña una tarea con el equipo o herramienta inadecuada, esto

puede ser, porque no tenía la herramienta, porque la usaba mal, o porque esta no se encontraba en la condición apropiada.

El gráfico 6 muestra la cantidad de OPS por cada una de estas categorías, la mayoría de estas tiene más de 7500 registros, salvo la de Manejo Manual que representa tan solo 1824 de los registros.

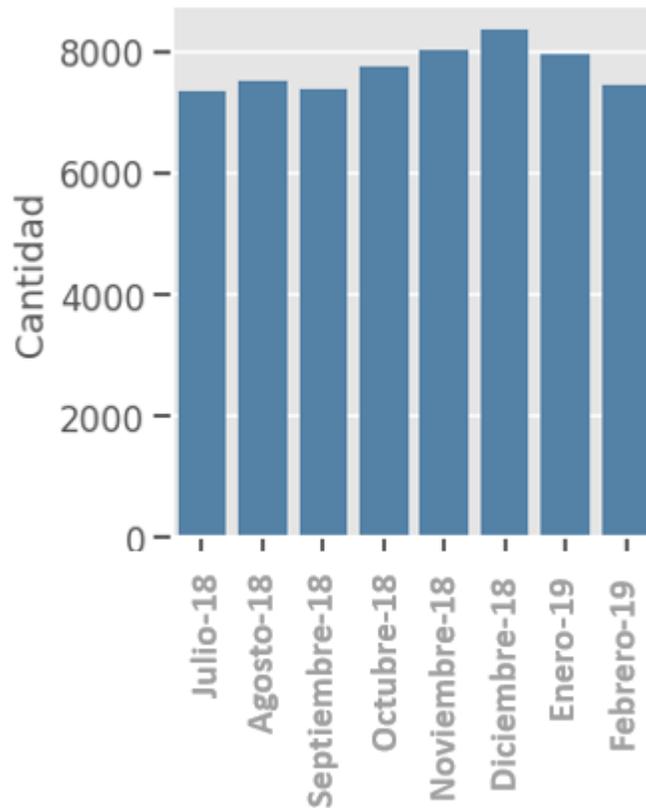
Gráfico 6: Número de OPS según clasificación



Fuente 19: Elaboración propia

En el gráfico 7 se puede observar el número de OPS por cada mes existente en los registros a analizar, siendo diciembre el mes en que se registraron más OPS y entregando un promedio mensual de 7670 registros.

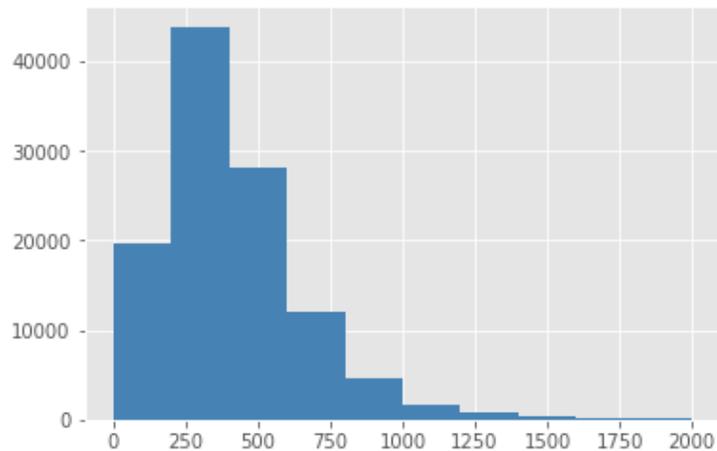
Gráfico 7: Número de OPS mensuales



Fuente 20:Elaboración propia

Por otra parte, el largo de una OPS, medido en las palabras que contiene, se puede observar en el grafico 8, este muestra que la mayoría de los registros tienen más de 250 palabras, siendo el largo promedio de 415 palabras.

Gráfico 8: Histograma largo de OPS



Fuente 21: Elaboración propia

Dado que las OPS corresponden a tres faenas mineras distintas, es importante entender la distribución de estos registros dentro de las faenas, la tabla a continuación muestra dicha distribución.

Tabla 7: Número de OPS por faena según clasificación

Faena	EPP	Línea de fuego/Posición de persona	Área/Entorno de trabajo	Manejo manual	Procedimiento/Evaluación de Riesgo	Herramientas y Equipos	TOTAL
1	10633	8034	10809	1571	12258	7196	50501
2	328	300	539	59	375	330	1931
3	1202	704	738	180	754	566	4144
TOTAL	12163	9038	12086	1810	13387	8092	56576

Fuente 22: Elaboración propia

De la tabla anterior se desprende que la mayor parte de las OPS clasificadas corresponden a la faena 1 y que la clasificación más recurrente es la asociada a errores en los procedimientos y evaluación de riesgo, sin embargo, no es la categoría más relevante para todas las faenas, ya que en el caso de la faena 2 y 3 las categorías más nombradas son las de área y uso de equipos de protección personal respectivamente, esto probablemente se deba a que las condiciones de trabajo no son exactamente las mismas y los riesgos varían en función de éstas.

Lo que distingue a una categoría de otra son los patrones subyacentes en los textos, una forma de ver estas diferencias es analizar las palabras más usadas por cada clase (Ver Anexo B).

El análisis anterior muestra que las palabras más usadas en la categoría tienen relación con lo que busca identificar la clase, como muestra la tabla a continuación.

Tabla 8: Selección de las 10 palabras más usadas que distinguen una categoría

CATEGORÍA	PALABRAS
EPP	"Guante", "EPP"
POSICIÓN DE PERSONA	"Sector", "Riesgo"
ÁREA DE TRABAJO	"Área", "Sector"
MANEJO MANUAL	"Postura"
PROCEDIMIENTO	"Actividad"
HERRAMIENTAS Y EQUIPOS	"Herramienta", "sin"

Fuente 23: Elaboración propia

7.3.2 TRANSFORMACIÓN DE LA DATA

Para poder aplicar modelos de Machine Learning es necesario transformar la data a un formato adecuado para estos, este proceso se denomina vectorización.

En primer lugar, como los textos contienen palabras que no conllevan ningún significado de por sí (stopwords), estas deben ser eliminadas. También, se eliminan números y signos de puntuación ya que tampoco agregan valor al texto en cuestión.

Luego, es necesario que cada palabra a lo largo de todos los registros tenga todas sus letras en minúsculas, por lo que se realiza esa transformación en conjunto con la eliminación de tildes. Este proceso es útil ya que una palabra escrita en mayúscula es considerada distinta a su símil en minúscula.

Finalmente, se separa cada texto en un conjunto de tokens, estos se definieron como n-gramas de 1x2, este formato quiere decir que los tokens pueden componerse de una palabra unitaria o de un par de palabras. Se definió de esta forma debido a que existe una diferencia significativa entre el token "usar" y el token "no, usar".

En paralelo al proceso de tokenización se realiza el proceso de lematización en donde cada palabra dentro del texto es transformada a su raíz semántica.

Finalmente, se utiliza el algoritmo TF-IDF para llevar los textos a su respectiva representación vectorial.

7.4 MODELAMIENTO

Para automatizar la clasificación de OPS se utilizarán tres modelos distintos: Logit, Support Vector Machine y redes neuronales. Dado que una OPS puede pertenecer a más de una categoría, tanto para Logit como para SVM se genera un modelo por cada categoría, este enfoque se conoce como

One vs All, y se separa el problema de clasificación múltiple en distintos problemas de clasificación binaria. [30]

Para la correcta evaluación de los modelos se decide separar la data en un segmento para el entrenamiento de los algoritmos y otro para la evaluación de éstos, la proporción es de 80% y 20%. Esta segmentación es necesaria para descartar el sobreajuste de los modelos.

Para los modelamientos usados el resultado del algoritmo indica la probabilidad con que un registro en particular pertenezca a la clase o no, típicamente se considera que aquellos registros en donde la probabilidad es mayor a 0.5 son parte de la clase analizada. Sin embargo, dado los requerimientos del cliente, es preferible identificar correctamente aquellos registros que pertenecen a la clase y para incorporar este requerimiento es que el valor de la probabilidad necesaria será seleccionado dependiendo del modelo en cuestión.

La incorrecta predicción de los registros que no pertenecen a la clase se denomina error tipo 1, donde la hipótesis de pertenencia a la clase es aceptada siendo que el registro no pertenecía a ésta, al contrario, el error tipo 2 es predecir que el registro no tenía la clasificación cuando si era así. [31]

La probabilidad con la que un registro se define como parte de la clase analizada se denomina threshold [32]. En los modelamientos posteriores se define este valor en función de los resultados de la curva ROC, ajustándolo hasta que se alcance un 75% de eficacia en predecir los registros que pertenecen a la clase en cuestión. De esta manera se limita la pérdida de ajuste producto del cambio, es decir, se maneja el error tipo 2 sin perturbar de manera drástica el tipo 1.

7.4.1 LOGIT

El modelamiento usando Logit se establece con un vector de palabras de largo 3000, eliminando aquellas palabras que aparecen en más del 50% de los textos, debido a que una palabra que aparece en más del 50% de las OPS no pertenece a ninguna clasificación como tal.

Las curvas ROC generadas a partir de esta categoría se encuentran en el Anexo C. Estas curvas permiten una comparación visual entre la data de entrenamiento y prueba y, además, especifican qué tan bueno es el modelo para ordenar la data entre menor probabilidad de ser un caso positivo a mayor probabilidad de ser un caso positivo.

En base a las curvas Logit podemos observar que no parece haber un sobreajuste significativo para ninguna de las categorías, siendo las que más difieren: "Herramientas y equipos", "Manejo Manual" y "Procedimiento", con tan solo 6 centésimas de diferencia.

La categoría "EPP" es la que mejor comportamiento tiene en base a la curva ROC, su AUC (área bajo la curva) llega a un valor de 0.95 para la data de prueba, lo que significa que su capacidad de predicción es bastante buena ya que está a solo 5 centésimas del modelo perfecto.

La categoría "Procedimiento" es la que peor se desempeña con respecto a las demás, consiguiendo un AUC de 0.78 para la data de prueba. El promedio de AUC de todas las categorías es de 0.865, demostrando una buena capacidad de predicción para el modelo Logit.

En el ANEXO D se muestran los resultados generales para cada una de las categorías usando el algoritmo Logit. El threshold fue modificado para lograr predecir al menos un 75% de los resultados positivos.

Para la categoría "EPP" se logra la mejor accuracy, alcanzando una cifra de 92% para la data de prueba, logrando predecir correctamente el 82% de los casos que pertenecen a la categoría, mientras que identifica correctamente el 95% de los textos que no pertenecen a la categoría.

La categoría "Línea de fuego/Posición de persona", alcanza un accuracy de 80% para toda la data, con un ratio de verdaderos positivos de 0.75 y una precisión para encontrar e identificar aquellos datos que no pertenecen a la clase de 76%.

Por otra parte, la categoría de "Área/Entorno de trabajo" logra predecir correctamente un 85% de la data de prueba, alcanzando a predecir el 75% que pertenece a la categoría gracias a la modificación de threshold, eso

inevitablemente lleva consigo una menor ratio de verdaderos negativos, logrando un ajuste de 88% en este valor.

La categoría "Manejo Manual", es la que menor cantidad de datos presenta para poder entrenar. En esta categoría, el valor de accuracy alcanza un 89% para la data de evaluación, donde nuevamente producto de un cambio en el valor del threshold se alcanza a predecir correctamente el 75% de los datos que pertenecen a ella, llevando consigo una ratio de verdaderos negativos de 0.89. La precisión es el valor que más se ve afectado en esta categoría, ya que, dado que son muy pocos los casos donde es positiva, al tratar de predecir de mejor manera éstos, aumentan significativamente los casos de falsos positivos.

La categoría "Procedimiento" representa para la data de prueba un ajuste total de 69%, siendo la categoría que peor ajuste tiene para el algoritmo Logit. El ratio de verdaderos positivos es de 0.75, gracias a una modificación del umbral y el de verdaderos negativos es de 0.67, lo que significa que el algoritmo en este caso es mucho mejor prediciendo aquellos datos que pertenecen a la clase que los que no.

7.4.2 SUPPORT VECTOR MACHINE

El modelamiento usando SVM, a diferencia del hecho con Logit, no considera una vectorización de largo 3000, sino que la vectorización se realizó usando un vector de largo 1200, en este modelamiento también se incluye la restricción para no incluir aquellos tokens que pertenecen a más del 50% de las OPS.

El cambio en la magnitud del vector se debe a dos razones, en primer lugar, el procesamiento de datos usando SVM es mucho más lento que el hecho con Logit (siendo de aproximadamente 45 minutos versus 1 minuto respectivamente) y, en segundo lugar, porque los modelos preliminares mostraban una tendencia al sobreajuste y delimitar el largo del vector impide que el algoritmo pueda memorizar la data de entrenamiento.

Las curvas ROC (Anexo E) muestran sobreajuste para ninguna de las categorías usando SVM, la mayor diferencia es de 3 centésimas y se da en "Procedimiento" y "Manejo Manual". La categoría que mejor AUC tiene es nuevamente "EPP" con un 0.94, mientras que la categoría con peor capacidad

de ordenar los datos de manera correcta es "Procedimiento" con un AUC de 0.76 en la data de prueba.

En el Anexo F se encuentran los resultados para cada una de las categorías a analizar usando SVM. Nuevamente, el umbral de clasificación fue modificado para alcanzar a tener al menos un 75% de precisión en clasificar aquellos textos que sí pertenecen a una categoría en particular.

Al igual que en el modelamiento usando Logit, la categoría "EPP", es la que mejor ajuste total tiene alcanzando un 91%, su ratio de verdaderos positivos es de 78% y el de verdaderos negativos de 97%, logrando así predecir en casi la totalidad de los casos cuando un texto no pertenece a la categoría.

Por otra parte, la categoría de "Área/Entorno de trabajo", alcanza un ajuste total de 84% tanto para la data de prueba como la de entrenamiento, el ratio de verdaderos positivos es de 75%, mientras que el ratio de verdaderos negativos es de 87%.

La categoría de "Línea de fuego/Posición de persona", logra una precisión en la predicción total de 73%, junto con un ratio de verdaderos positivos de 75% y de verdaderos negativos de 74%, lo que significa que predice levemente mejor cuando las OPS sí pertenecen a la clase en cuestión.

En la categoría de "Procedimiento" es donde nuevamente encuentra el peor porcentaje de accuracy, siendo este de tan solo el 65% para la data de evaluación. El ratio de verdaderos positivos es del 75% gracias a la modificación del umbral, mientras que el de verdaderos negativos es del 60%, representando el peor ratio de verdaderos negativos para SVM y, por ende, prediciendo mucho mejor cuando los textos no pertenecen a la clase en cuestión.

La categoría de "Manejo Manual", requirió de una modificación de umbral significativa al ser tan pocos los casos en donde se presenta (4% de los datos). En función de lo anterior, su ratio de verdaderos positivos alcanza el 75% deseado y el de verdaderos negativos alcanza un 85%.

Finalmente, la categoría de "Herramientas y equipos" logra un accuracy de 68%, siendo principalmente producto de la gran cantidad de falsos positivos, lo que lleva consigo un ratio de verdaderos negativos igual a 0.67, mientras que el porcentaje de verdaderos positivos se logra llevar al 75% modificando el umbral.

7.4.3 RED NEURONAL

La red neuronal difiere de los otros modelos al no usar una vectorización de palabras mediante el algoritmo TF-IDF, si bien es posible de implementar, cuando se trata de modelos de clasificación de texto en redes neuronales, se suelen utilizar los llamados embedding layers.

Como los algoritmos no son capaces de entender las palabras en los textos se busca crear una representación vectorial de estas, para esto se utiliza un procedimiento denominado One Hot Encoding, con esto, se retornará una versión de los textos codificada en números enteros, en donde cada número representará una palabra dentro del texto.

Al igual que en los modelos anteriores se delimita el máximo de palabras que el vocabulario final puede guardar, de esta manera se conservan aquellas que más importancia tienen, es decir, aquellas que mejor describen los patrones subyacentes de cada clase. Este número se fijó en 3000, para mantener el estándar antes fijado en el modelo Logit.

El embedding layer usado en la red neuronal, permite que la vectorización a través del algoritmo One Hot Encoding se haga al mismo tiempo que la red es entrenada. [33]

Con respecto a la arquitectura de la red, ésta consta de siete capas (layers), la primera que es el embedding layer, una capa convolucional, una capa de transformación, dos capas de dropout, una capa densa y la capa final. (Ver Anexo G)

Las redes convolucionales suelen utilizarse mayoritariamente en el procesamiento de imágenes [34], sin embargo, así como las imágenes pueden entenderse como una secuencia de píxeles, los textos se pueden considerar como una secuencia específica de palabras. La gran ventaja de este enfoque

es que permite a la red aprender patrones en el espacio de data en lugar de a través del tiempo, como lo haría una red recurrente. [35]

Las capas de dropout son capas que permiten a la red no considerar ciertos nodos en algunas etapas del entrenamiento con el fin de no sobreajustar el resultado. Al ignorar algunos de los nodos, la capa anterior al dropout debe forzar a otros nodos a tomar más o menos responsabilidad por el resultado deseado. [36]

La capa densa (Dense layer) se diferencia de la convolucional en que cada nodo está relacionado con todos los demás nodos, ésta se ubica justo antes de la capa final y también tiene una capa de dropout para prevenir el sobreajuste.

La capa final de toda la arquitectura consta de tan solo seis nodos, cada uno representa una de las clases en cuestión. En esta capa se delimita que los nodos solo pueden tomar valores entre 0 y 1 ya que estos representan la probabilidad con la que el texto pertenece a la clase en cuestión. Esta última capa indicará al mismo tiempo si un registro pertenece o no a cada una de las clases.

ENTRENAMIENTO Y EVALUACIÓN

El entrenamiento de la red neuronal se realiza con el 80% de la data, conservando lo ya estipulado en los otros dos modelamientos, la única diferencia es que el 10% de la data de entrenamiento se utiliza como data de validación durante el entrenamiento.

En las redes neuronales artificiales es una buena práctica separar la data en tres partes, data de entrenamiento, data de validación y data de prueba. Con la data de entrenamiento el algoritmo aprende, es decir, modifica los vectores que conectan los nodos para alcanzar el resultado deseado y, luego de cada una de estas modificaciones, se utiliza la data de validación para evaluar la performance del modelo.

Es importante conocer cómo se desempeña el modelo durante el entrenamiento ya que de esta forma se evita el sobreajuste de éste. Un modelo se comienza a sobreajustar cuando el ajuste de la data de entrenamiento

aumenta, pero el ajuste a la data de validación disminuye. Para evitar que esto suceda se agrega al algoritmo un Early Stopping (Parada temprana), este hará que cuando la mejora solo beneficie a la data de entrenamiento y no a la de validación el algoritmo se detenga y finalice su entrenamiento. [37]

Tal como se revisó en secciones anteriores, los registros de OPS se encuentran desbalanceados con respecto a las distintas clases (véase gráfico 6), sin embargo, balancear la data no es una buena opción pues las redes neuronales son algoritmos no lineares que suelen necesitar más registros para funcionar de manera apropiada y al balancearla se disminuye considerablemente el número de registros para el entrenamiento. [38]

La técnica usada para tratar el problema de la data desbalanceada fue imponer una función de costo (loss function) que penaliza con mayor fuerza los errores tipo II, es decir, predecir que no pertenece a la clase cuando si era así (ver Anexo H).

Usando todo lo detallado anteriormente el modelo de redes neuronales demora alrededor de dos horas en ser ejecutado, lo que es significativamente mayor al algoritmo Logit, aunque, por otra parte, considerando que logra ajustar las 6 categorías a la vez, supera con creces la velocidad de SVM.

En el Anexo I se encuentran las curvas ROC generadas a partir del modelamiento con redes neuronales. Nuevamente la categoría "EPP" es la que ordena de mejor forma las probabilidades, mientras que la de peor desempeño es "Procedimiento" con un AUC de 0.77 en la data de evaluación. Con respecto al sobreajuste de la data, la mayor diferencia es de 7 centésimas y se da en la categoría de "Herramientas y Equipos", siendo muy similar a lo obtenido en los modelos anteriores.

Los resultados obtenidos al evaluar el desempeño del modelo de clasificación usando redes neuronales se encuentran en el Anexo J. El umbral de clasificación fue modificado en cada una de las categorías para alcanzar a identificar al menos un 75% de los resultados positivos.

En primer lugar, la categoría "EPP" es la que demuestra el mejor desempeño, tal cual se dio en los modelos anteriores. Esta categoría logra un accuracy de 0.93 para la data de evaluación, logrando predecir, sin

modificación de threshold, el 83% de los verdaderos positivos y el 96% de los verdaderos negativos.

En segundo lugar, por desempeño final, se encuentra la categoría "Manejo Manual" la que logra un accuracy de 87%. Mientras que su tasa de verdaderos positivos es del 75% y la de verdaderos negativos alcanza un 87%.

La categoría "Área/Entorno de trabajo", es la que posee la tercera mejor métrica de accuracy, con un total de 83% de predicciones correctas. La tasa de verdaderos positivos es de 81%, sin modificación de umbral, y la de verdaderos negativos de 83%.

En cuarto lugar, se encuentra la categoría "Línea de fuego/Posición de persona" con un 76% de lo resultado predichos de manera correcta. Logrando al mismo tiempo un ratio de verdaderos positivos igual a 0.75 y de verdaderos negativos de 0.76.

La siguiente categoría es "Herramientas y Equipos" con un 69% de accuracy, alcanzando el 75% de verdaderos positivos y un 68% de verdaderos negativos, lo que hace más propenso a fallar en predecir cuando los textos no pertenecen a la clase.

Finalmente, la categoría que peor desempeño tiene es "Procedimiento", logrando un accuracy final de 0.65, gran parte del error se debe a la predicción de falsos positivos, alcanzando un ratio de verdaderos negativos de apenas un 0.59, mientras que el ratio de verdaderos positivos es fijado en un 75%.

7.5 MODELO FINAL

En el apartado anterior se analizaron las métricas de ajuste para cada una de las clases usando Logit, SVM y redes neuronales. En esta sección se pretende escoger un modelo final para todas las clases disponibles.

Los modelos anteriores están hechos para diferenciar si un determinado texto se puede clasificar dentro de una clase en particular o no, pero el objetivo final de este trabajo de título es crear un algoritmo de clasificación capaz de

identificar a qué categorías pertenece un determinado texto, por ende, es necesaria la elección de uno de los tres algoritmos.

Una de las primeras métricas analizadas fue el área bajo la curva ROC, dado que es la capacidad del algoritmo de ordenar la data, es necesario analizar cual modelo logra mejor esta tarea, en la tabla 8 se pueden ver los resultados.

Tabla 9: Resultados AUC por modelo

Categoría	Logit	SVM	Redes neuronales
EPP	0.95	0.94	0.95
Línea de fuego/Posición de persona	0.84	0.82	0.84
Área/Entorno de trabajo	0.91	0.89	0.90
Manejo manual	0.91	0.88	0.90
Herramientas y Equipos	0.80	0.79	0.79
Procedimiento/Evaluación de Riesgo	0.78	0.76	0.77
Promedio	0.865	0.847	0.858

Fuente 24: Elaboración propia

En la tabla se aprecia que para cada una de las categorías SVM es el que muestra un peor desempeño alcanzando un AUC promedio de 0.847, por otra parte, el modelo que mejor se desempeña es Logit, aunque las redes neuronales logran igualar el valor de este en dos categorías.

Otra de las métricas relevantes para evaluar el desempeño de un modelo es Log Loss, si bien no tiene una interpretación directa, es útil para comparar distintos modelos de clasificación asumiendo siempre que, a menor valor, mejor predice el modelo.

Tabla 10: Resultados Log Loss por modelo

Categoría	Logit	SVM	Redes neuronales
EPP	0.249	0.251	0.259
Línea de fuego/Posición de persona	0.402	0.392	0.447
Área/Entorno de trabajo	0.361	0.359	0.405

Manejo manual	0.109	0.119	0.121
Herramientas y Equipos	0.408	0.391	0.510
Procedimiento/Evaluación de Riesgo	0.537	0.512	0.640

Fuente 25: Elaboración propia

La tabla 9, muestra los resultados de Log Loss para los modelos, a diferencia de la métrica anterior, esta tabla define a SVM como el mejor modelo para cuatro de las seis categorías analizadas, mientras que Logit solo es mejor en las categorías "EPP" y "Manejo Manual". Por otra parte, las redes neuronales presentan el peor desempeño en esta métrica, presentando valores muy por encima a Logit y SVM.

También, como se fijó un porcentaje mínimo necesario de verdaderos positivos, podemos comprobar cómo afecta este en la tasa de verdaderos negativos y en el accuracy del modelo. La tabla 10 muestra los resultados.

Tabla 11: Comparación de modelos (ratio verdaderos positivos, ratio verdaderos negativos y accuracy)

		Logit	SVM	Red Neuronal
EPP	Ratio verdaderos positivos	0.8241	0.7813	0.8366
	Ratio verdaderos negativos	0.9573	0.9710	0.9649
	Accuracy	0.9207	0.9189	0.9299
Línea de fuego /Posición de persona	Ratio verdaderos positivos	0.7517	0.7561	0.7514
	Ratio verdaderos negativos	0.7690	0.7408	0.7671
	Accuracy	0.7655	0.7439	0.7639
Área/Entorno de trabajo	Ratio verdaderos positivos	0.7596	0.7526	0.8191
	Ratio verdaderos negativos	0.8848	0.8742	0.8380
	Accuracy	0.8506	0.8410	0.8327
Manejo manual	Ratio verdaderos positivos	0.7541	0.7569	0.7530
	Ratio verdaderos negativos	0.8965	0.8597	0.8746
	Accuracy	0.8907	0.8554	0.8700
Herramientas y Equipos	Ratio verdaderos positivos	0.7505	0.7579	0.7547
	Ratio verdaderos negativos	0.6991	0.6702	0.6823
	Accuracy	0.7085	0.6863	0.6958
Procedimiento /Evaluación de Riesgo	Ratio verdaderos positivos	0.7517	0.7599	0.7868
	Ratio verdaderos negativos	0.6703	0.6096	0.5926
	Accuracy	0.6950	0.6551	0.6522

Fuente 26: Elaboración propia

En la categoría "EPP" las redes neuronales tienen el mejor desempeño, tanto para el ratio de verdaderos positivos, como para el accuracy total. El modelo usando Logit logra resultados bastante similares, pero con un menor desempeño, mientras que SVM logra un mejor ratio de verdaderos negativos, pero con un ratio de verdaderos positivos mucho menor a los demás modelos.

Para la categoría "Línea de fuego/Posición de persona", el modelamiento usando Logit logra un mejor desempeño que las redes neuronales en las tres métricas analizadas. Mientras que SVM tiene una mejor predicción de verdaderos positivos, pero disminuye bastante su capacidad de predecir correctamente aquellos textos que no pertenecen a la clase.

La categoría "Área/Entorno de trabajo" muestra un mejor desempeño en identificar verdaderos positivos usando redes neuronales, sin embargo, esto significa una disminución del accuracy y el ratio de verdaderos negativos. Usando Logit, el modelo logra mejores resultados que SVM y un ratio de verdaderos negativos mayor que al usar redes neuronales.

Por otra parte, la categoría "Manejo Manual" tiene un ratio de verdaderos positivos prácticamente igual en cada uno de los modelos, y como el enfoque usando Logit logra una mejor accuracy y al mismo tiempo identificar de mejor manera los resultados que no pertenecen a la clase, entonces se puede concluir que es el mejor modelo para identificar los textos de la clase.

La categoría "Herramientas y equipos", al igual que la anterior, tiene un desempeño similar en la identificación de verdaderos positivos en todos los modelos y es el modelo Logit el que logra los mejores resultados identificando los resultados que no pertenecen a la clase y en el accuracy final.

Por último, la categoría "Procedimiento", logra identificar los verdaderos positivos un 3% mejor usando redes neuronales, aunque eso significa una disminución significativa en las otras dos métricas. Usando Logit se logra un mejor resultado en el desempeño total y en la identificación de verdaderos negativos en comparación con las redes neuronales y SVM.

7.5.1 DISCUSIÓN

Elegir el mejor enfoque para clasificar OPS depende de la métrica a analizar, una primera aproximación es comparar el accuracy, pero esto no es del todo significativo para clases desbalanceadas, como es el caso en todas las clasificaciones analizadas.

Además, como los textos indican condiciones de seguridad, muchas veces es mejor identificar mal cuando no hay un problema que cuando lo hay, por eso la importancia de fijar el umbral de clasificación para que al menos logrará clasificar el 75% de los resultados que si pertenecían a la clase.

Por otra parte, no se debe perder el foco de predecir, lo que significa que no con el fin de lograr identificar la mayor cantidad de valores en la clase no se debe tomar a la ligera la disminución en el ratio de verdaderos negativos.

También es importante considerar siempre que la modificación de umbral de clasificación solo es posible cuando el modelo logra ordenar los textos de mayor a menor probabilidad, la forma de comprobar esto es mediante el valor del AUC. Mientras menor el valor del AUC, mayor será el número de falsos positivos que entregará el modelo al variar el umbral de clasificación.

Otro factor importante es el reentrenamiento del modelo, en este caso, dejar fijo un modelo en particular sin una revisión pertinente cada cierto periodo de tiempo puede significar no identificar nuevos factores de riesgo. Por lo anterior, al momento de volver a entrenar el modelo es útil uno que sea sencillo y no requiera demasiado tiempo para los tomadores de decisión.

El modelo con un rendimiento más estable en todas las categorías es Logit, logrando un accuracy final de 80.8%. Sin embargo, la red neuronal propuesta logra superar a este modelo en algunas de las métricas, es posible que esto se dé por la cantidad de datos con las que se entrena el modelo. Las redes neuronales son modelos complejos que necesitan grandes cantidades de datos para entrenarse de manera correcta, de modo que, en el futuro, a medida que se acumule la data, este enfoque podría superar a los demás ampliamente.

Clasificar OPS usando SVM lograba métricas menores a los otros enfoques y una mayor tendencia al sobreajuste de la data. De todas formas, es importante recalcar que este modelo se entrenaba con un menor número de variables que los demás, lo que significa que al aumentar el largo del vector de palabras se podrían obtener mejores resultados, pero eso a la vez permitiría al modelo sobreajustar la data.

Asimismo, existe un caso que queda fuera de los modelos de clasificación entrenados, pero que es parte de la realidad. Para todos los modelos de clasificación entrenados se aplicaron distintos filtros a los textos con el fin de lograr una data que permitiera capturar patrones en ellos. Sin embargo, en la realidad, no solo es posible, sino que existen textos OPS que no tienen un texto lógico, ya sea porque está incompleto, en blanco o porque no tiene sentido. En estos casos los modelos no encontrarán una categoría en la que asignar el texto, lo que significa que es necesario separarlos de la muestra e identificar porque se producen, para tratarlos como una categoría completamente distinta.

Finalmente, una de las grandes dificultades de la data a trabajar es que fue manipulada y creada por distintos usuarios. Cada uno bajo su comprensión de lo que significaba el acto de riesgo clasificó el comportamiento en la categoría que creyó pertinente. Si bien las diferencias entre personas pueden ser bastantes, con la suficiente data y el suficiente entrenamiento un modelo de clasificación puede ser capaz de encontrar el patrón de riesgo subyacente para casos en que los propios operadores no fueron capaces de clasificar, siempre priorizando, por supuesto la obtención de verdaderos positivos.

CAPÍTULO 8: CONCLUSIONES

En primer lugar, se establecen las formas actuales de clasificación de accidentes en Chile para la industria minera. Los accidentes se pueden clasificar en base a la causa que los originó, las consecuencias que generan, la frecuencia con la que ocurren, la probabilidad de ocurrencia, las condiciones inseguras que existían y el acto inseguro subyacente.

La importancia de hacer una clasificación de accidentes no es solo generar una estadística al respecto, sino que utilizar estos números para decidir dónde poner foco en la prevención, de ahí la importancia de entender cuál era el acto inseguro de un trabajador, eje central de una OPS.

En segundo lugar, para lograr la clasificación de texto usando NLP existen variadas técnicas de Machine Learning a utilizar, la vectorización del texto, por ejemplo, puede lograrse mediante Bag of Words, el algoritmo TF-IDF y One Hot Encoding, entre otros. Con respecto al modelo de clasificación los enfoques son igual de variados, es posible usar algoritmos como Logit, SVM, Random Forest, Xgboost y redes neuronales.

En este trabajo se utilizaron Logit, SVM y redes neuronales, ya que los dos primeros son los enfoques tradicionales para el análisis de texto, son ampliamente usados y muestran un buen desempeño. Mientras que las redes neuronales son buenos algoritmos cuando se tiene gran cantidad de data, y han logrado ganar terreno en el análisis de datos en el último tiempo.

Usando los tres algoritmos mencionados para la clasificación se obtienen modelos que permiten clasificar acciones inseguras de trabajo de manera automática. Estos modelos presentan un buen desempeño final ya que logran predecir de mejor forma que un modelo aleatorio la clase de cada texto. Además, la modificación del umbral de clasificación consigue poner foco en identificar el riesgo en la mayoría de los casos para que se puedan tomar acciones en prevenir accidentes, alcanzándose con esto el objetivo general de este trabajo de título.

8.1 SOBRE LA CALIDAD DE LOS DATOS

Con respecto a los registros y su respectiva clasificación, existen dos desventajas que son destacables a la hora de realizar los modelos de clasificación.

En primer lugar, toda clasificación la hace la persona que registra la OPS, por ende, son más de 2000 personas distintas quienes clasifican, no siempre considerando los mismos criterios para esto, por lo que es posible que una persona en particular considere que un registro pertenece a una categoría y otra persona no.

La segunda desventaja es el desbalance entre las etiquetas para las distintas categorías, siendo la categoría de Manejo Manual la más afectada. Esta situación hace más difícil la capacidad de aprendizaje de algoritmo ya que tiene menos casos positivos de los que aprender.

8.2 TECNICAS DE MACHINE LEARNING

En relación a los modelos de clasificación usados, cada uno tuvo un comportamiento distinto, ya que la metodología es distinta y los parámetros iniciales varían, las principales conclusiones para cada modelo se detallan a continuación:

Logit: Un método rápido para entrenar, que logra superar en varias de las métricas a los demás enfoques, sobre todo en términos del ratio de verdaderos positivos sin que esto afecte el accuracy final. Destaca mayoritariamente por no tener una tendencia marcada al sobreajuste.

SVM: Los resultados obtenidos son inferiores a los demás enfoques en la mayoría de las métricas, en especial en la cantidad de falsos positivos que produce afectando la capacidad de predecir aquellos datos que no pertenecen a la clase en cuestión. Sin embargo, esto se puede deber a la menor cantidad de variables con las que fue entrenado, propuesta por su notoria tendencia al sobreajuste y el gran tiempo de entrenamiento que requiere.

Red Neuronal: El algoritmo se destacó sobre los demás en algunas de las categorías, sin embargo, en otras mostró un rendimiento bastante menor. La gran ventaja de este enfoque es que permite la clasificación de todas las categorías a la vez.

De los modelos planteados Logit podría ser una buena estrategia a corto plazo, ya que puede capturar los patrones en las categorías, logrando un rendimiento destacable en varias de las métricas. Sin embargo, si en el futuro se requiere evaluar las 30 categorías originales, un modelamiento usando Logit requeriría de al menos 30 modelos distintos, si se sigue la modalidad planteada en este informe.

Por otra parte, el modelamiento a través de la red neuronal logra buenos resultados en las categorías de mayor presencia, pero es inferior en capturar los patrones de las que no se encuentran demasiado en la data. En el futuro, con la acumulación de data este problema es factible de ser solucionado, lo que dotaría a este algoritmo de mejor capacidad de predicción. Además, a diferencia del anterior, un solo modelo sería capaz de predecir las 30 clases a la vez.

Con respecto a los indicadores utilizados para evaluar los modelos cabe destacar ciertas ventajas y desventajas.

La métrica de accuracy no parece ser el mejor indicador para clases desbalanceadas, ya que en algunas de las categorías tan solo predecir un mismo resultado para toda la data podría lograr un accuracy mayor al 90%.

El ratio de verdaderos positivos fue fijado a través del umbral de clasificación en la mayoría de los modelamientos, por ende no es directamente comparable entre algoritmos.

Con lo anterior adquiere relevancia el ratio de verdaderos negativos del modelo, ya que se puede identificar aquel modelo en el que la modificación de umbral para obtener mayor cantidad de verdaderos positivos afecta en menor medida el ratio de verdaderos negativos.

El AUC permitió demostrar que existen patrones en la data que permiten ordenar los textos de mayor a menor probabilidad de pertenecer a una

categoría en particular, esta conclusión es relevante ya que permite la modificación de umbral, sin un AUC suficientemente bueno, cambiar el umbral de clasificación no lograría mejorar la identificación de registros y podría disminuir el desempeño final del modelo de manera significativa.

Usar Log Loss como métrica de evaluación permitió comparar modelos entre sí, sin embargo, esta métrica no es del todo relevante luego de la modificación del umbral de clasificación.

8.3 HIPOTESIS

Con respecto a las hipótesis planteadas al inicio queda demostrado que existen patrones en la data que representan cada categoría y que éstos pueden ser modelados en un algoritmo de clasificación ya que todos los modelos logran un ajuste general superior al 77%, es decir, predicen la categoría correcta en la mayoría de los casos.

Además, si se toma el modelo Logit como ejemplo, dada la modificación del umbral, se puede identificar qué texto pertenece a una categoría en particular en un 80% de los casos, con un accuracy final de más del 80%.

8.4 SOBRE LAS CATEGORIAS

El procedimiento aplicado a cada una de las categorías analizada es esencialmente el mismo a lo largo del trabajo realizado, sin embargo, dado los resultados obtenidos es evidente que los patrones inherentes a cada categoría difieren en la facilidad de ser encontrados.

La categoría "EPP" muestra la mejor performance para cada uno de los modelos, esto se debe esencialmente a que describe una condición de uso no uso de equipos, fácil de identificar para los modelos cuando un equipo en particular se encuentra en el texto. Una situación similar ocurre con las categorías "Área/Entorno de trabajo" y "Línea de fuego/Posición de persona".

Otras categorías, como "Procedimiento" y "Herramientas y equipos" son más difíciles de identificar, esencialmente porque tienen una mayor diversidad

de posibilidades en las que pueden aparecer. Por lo anterior, a pesar de tener gran cantidad de datos, no logran un desempeño al nivel de las demás.

Por otra parte, la categoría "Manejo Manual" presenta un gran desafío por ser la categoría que menos se presenta en la data, lo que en todos los modelamientos significó disminuir considerablemente el umbral de clasificación y con ello aceptar un aumento de falsos positivos. Se espera que en el futuro mayor cantidad de datos permitan mejorar el desempeño de esta categoría y lograr así aumentar la precisión de esta.

Finalmente, una de las mayores dificultades es la diferenciación entre las categorías. Dentro de la data original existen registros OPS en donde muchas o incluso todas las categorías han sido seleccionadas a la vez, si bien es posible que una OPS describa un comportamiento de riesgo para cada clasificación, esto podría ser un indicio de cierta dificultad por parte de los trabajadores en identificar la categoría más pertinente.

8.5 APLICACIÓN

Para que la automatización de la clasificación de registros entregue el valor esperado, es necesario que sea acompañada de un sistema y control acorde.

En primer lugar, debe tener la capacidad de alertar tanto a supervisores como altos cargos de la operación, con el debido nivel de agregación, sobre el desempeño en medidas de seguridad en la faena.

Por otra parte, el sistema podría ser capaz de entregar una calificación para cada área respecto a que tan bien se adecuan a las medidas correspondientes y a cuanto estas han mejorado con respecto a tiempo.

En la misma línea anterior, toda la información debe poder ser consultada de manera rápida y fácil en un dashboard. Este puede incluir el número de OPS, los lugares con más registros, el aumento en el tiempo, las zonas más críticas por cada tipo, entre otros.

Por lo demás, asumiendo que el sistema en el estado actual logra un ajuste del 77%, el reentrenamiento y la revisión por parte del personal

adecuado es indispensable si esto se quiere mejorar. Es importante que se puedan detectar con mayor eficacia aquellas conductas que representan la mayor cantidad de accidentes a lo largo de la historia y por ello, el cruce de información entre prevención y resultados de seguridad se hace vital.

Por último, dado que el sistema será capaz de realizar análisis a tiempo real, estos pueden incluso sugerir los riesgos más preponderantes y aquellos que representan consecuencias más importantes con el fin de enlistar las capacitaciones y controles preventivos en esas áreas en específico.

CAPÍTULO 9: RECOMENDACIONES Y TRABAJOS FUTUROS

Una forma de mejorar los resultados de los modelos de clasificación es generar una mayor estandarización de la clasificación que se les da a cada OPS, para lograr esto se podría capacitar a los trabajadores con respecto a las distintas categorías o se puede encargar a una persona experta en seguridad la clasificación de un gran número de OPS de tal manera de entrenar el algoritmo con datos basados en el experto.

Otra posibilidad de obtener mejores resultados en los algoritmos es entrenar los modelos con la data histórica o una mayor porción de lo que se utilizó en este trabajo, en específico eso permitiría mejorar aquellas categorías que tenían pocos registros con respecto al resto.

Si un modelo de clasificación como este llegase a implementarse es necesario revisar la performance de éste a medida que pasan los años, ya que las condiciones de la mina podrían cambiar o incluso podría ser necesario agregar una nueva categoría de incidentes.

Por otra parte, es una buena práctica seleccionar ciertas OPS ya clasificadas por el algoritmo y comprobar que estén en las categorías correctas, con estos ejemplos se podría volver a correr el algoritmo y mejorar las clasificaciones futuras.

Considerando que se logró una clasificación eficiente para los registros OPS, si se aplica el modelo en una empresa similar a la analizada, el paso siguiente es expandir estos algoritmos a las 30 categorías reales de la data, con esto los encargados de tomar las decisiones tendrán incluso un detalle más específico acerca de los comportamientos que pueden ocasionar incidentes de seguridad.

En último término, esta data puede emplearse para intentar predecir futuros accidentes, este trabajo prueba que existen ciertos patrones subyacentes entre registros, por lo que muy probablemente estos (u otros) puedan ser usados para predecir accidentes cruzando estos registros con la base histórica de accidentes en cada faena respectiva.

CAPÍTULO 10: BIBLIOGRAFÍA

[1] Servicio Nacional de Geología y Minería, «Anuario de la Minería de Chile,» Santiago, 2019.

[2] Tesseracto, «¿Qué es Tesseracto?,» [En línea]. Available: <http://tesseracto.com/index>. [Último acceso: 7 Agosto 2019].

[3] Endeavor, «El Impacto de la Inteligencia Artificial en el Emprendimiento,» New York, 2018.

[4] Campaign for Accuracy in Public Health Research, «Risk vs. Hazard,» [En línea]. Available: <https://campaignforaccuracyinpublichealthresearch.com/risk-vs-hazard/>. [Último acceso: 18 Agosto 2019].

[5] Escuela Europea de Excelencia, «Qué es la OHSAS 18001. Definición y origen,» Nueva ISO 45001:2018, 16 Octubre 2015. [En línea]. Available: <https://www.nueva-iso-45001.com/2015/10/que-es-la-ohsas-18001/>. [Último acceso: 2 Noviembre 2019].

[6] ISO Tools, «La norma OHSAS 18001».

[7] G. Robotham, «Lost Time Injury Frequency Rate,» Safety Risk, 2 Octubre 2013. [En línea]. Available: <https://safetyrisk.net/lost-time-injury-frequency-rate/>. [Último acceso: 27 Octubre 2019].

[8] Sitemate, «TRIFR Calculation: How to calculate total recordable injury frequency rate,» [En línea]. Available: <https://sitemate.com/resources/articles/safety/trifr-calculation/>. [Último acceso: 27 Octubre 2019].

[9] International Council on Mining & Metals, «Gestión de controles críticos para la salud y la seguridad,» Reino Unido, 2015.

[10] International Council on Mining & Metals, «Guía para la ejecución de la gestión de los controles críticos,» Reino Unido, 2015.

[11] Paritarios, «El sentido de las observaciones conductuales en prevención,» [En línea]. Available: http://www.parityarios.cl/especial_el_sentido_de_las_observaciones_conductuales_en_prevencion.html. [Último acceso: 18 Agosto 2019].

[12] S. Bansal, «A Comprehensive Guide to Understand and Implement Text Classification in Python,» Analytics Vidhya, 23 Abril 2018. [En línea].

Available: <https://www.analyticsvidhya.com/blog/2018/04/a-comprehensive-guide-to-understand-and-implement-text-classification-in-python/>. [Último acceso: 6 Diciembre 2019].

[13] A. Navlani, «Understanding Logistic Regression in Python,» DataCamp, 7 Septiembre 2018. [En línea]. Available: <https://www.datacamp.com/community/tutorials/understanding-logistic-regression-python>. [Último acceso: 20 Agosto 2019].

[14] A. Navlani, «Support Vector Machines with Scikit-learn,» Datacamp, 12 Julio 2018. [En línea]. Available: <https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python>. [Último acceso: 21 Agosto 2019].

[15] L. Dormehl, «What is an artificial neural network? Here's everything you need to know,» Digital Trends, 5 Enero 2019. [En línea]. Available: <https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/>. [Último acceso: 3 Noviembre 2019].

[16] A. Merchant, «Neural Networks Explained,» Medium, 29 Diciembre 2018. [En línea]. Available: <https://medium.com/datadriveninvestor/neural-networks-explained-6e21c70d7818>. [Último acceso: 3 Noviembre 2019].

[17] 365 DataScience, Machine Learning Course Notes - Section 6, 2019.

[18] D. Becker, «What is Log Loss?,» Kaggle, 2018. [En línea]. Available: <https://www.kaggle.com/dansbecker/what-is-log-loss>. [Último acceso: 26 Septiembre 2019].

[19] Data school, «Simple guide to confusion matrix terminology,» Data school, 25 Marzo 2014. [En línea]. Available: <https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/>. [Último acceso: 1 Septiembre 2019].

[20] Data school, «ROC curves and Area Under the Curve explained (video),» Data school, 19 Noviembre 2014. [En línea]. Available: <https://www.dataschool.io/roc-curves-and-auc-explained/>. [Último acceso: 1 Septiembre 2019].

[21] R. Joshi, «Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures,» EXSILO SOLUTIONS, 6 Septiembre 2016. [En línea]. Available: <https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/>. [Último acceso: 1 Septiembre 2019].

- [22] A. Malge, «An Efficient Framework for Spam Mail Detection in Attachments using NLP,» International Journal of Science and Research, 2015.
- [23] C.-H. Cheng, «Sentimental text mining based on an additional features method for text classification,» PLOS ONE, 2019.
- [24] R. C. Balabantaray, M. Mohammad y N. Sharma, «Multi-Class Twitter Emotion Classification: A New Approach,» International Journal of Applied Information Systems (IJ AIS) –, New York, 2012.
- [25] Y. Kim, E. Riloff y S. M. Meystre, «Improving classification of medical assertions in clinical notes,» 2011.
- [26] Y. Sasaki, B. Rea y S. Ananiadou, «Multi-Topic Aspects in Clinical Text Classification,» International Conference on Bioinformatics and Biomedicine, Manchester, 2007.
- [27] Y. Wang, E. Coiera, W. Runciman y F. Magrabi, «Using multiclass classification to automate the identification of patient safety incident reports by type and severity,» 2017.
- [28] H. Wang y J. Kildea, «Natural language processing applied to incident reports in radiation oncology–determination of completeness and classification,» 2017.
- [29] Python Software Foundation, «langdetect 1.0.7,» PyPI, [En línea]. Available: <https://pypi.org/project/langdetect/>. [Último acceso: 29 Agosto 2019].
- [30] M. Aly, «Survey on Multiclass Classification Methods,» Pasadena, 2005.
- [31] Soporte de Minitab, «¿Qué son los errores de tipo I y tipo II?,» Minitab, [En línea]. Available: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/basics/type-i-and-type-ii-error/>. [Último acceso: 17 Octubre 2019].
- [32] Google Developers, «Clasificación: Umbral,» [En línea]. Available: <https://developers.google.com/machine-learning/crash-course/classification/thresholding>. [Último acceso: 17 Octubre 2019].
- [33] J. Brownlee, «What Are Word Embeddings for Text?,» Machine Learning Mastery, 11 Octubre 2017. [En línea]. Available: <https://machinelearningmastery.com/what-are-word-embeddings/>. [Último acceso: 10 Noviembre 2019].

[34] C. T. B. Hons, «An introduction to Convolutional Neural Networks,» Towards Data Science, 27 Mayo 2019. [En línea]. Available: <https://towardsdatascience.com/an-introduction-to-convolutional-neural-networks-eb0b60b58fd7>. [Último acceso: 10 Noviembre 2019].

[35] S. Ghelano, «Text Classification — RNN's or CNN's?,» Towards Data Science, 1 Junio 2019. [En línea]. Available: <https://towardsdatascience.com/text-classification-rnns-or-cnn-s-98c86a0dd361>. [Último acceso: 10 Noviembre 2019].

[36] J. Brownlee, «A Gentle Introduction to Dropout for Regularizing Deep Neural Networks,» Machine Learning Mastery, 3 Diciembre 2018. [En línea]. Available: <https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks/>. [Último acceso: 10 Noviembre 2019].

[37] Elite Data Science, «Overfitting in Machine Learning: What It Is and How to Prevent It,» Elite Data Science, [En línea]. Available: <https://elitedatascience.com/overfitting-in-machine-learning>. [Último acceso: 11 Noviembre 2019].

[38] J. Brownlee, «How Much Training Data is Required for Machine Learning?,» Machine Learning Mastery, 24 Julio 2017. [En línea]. Available: <https://machinelearningmastery.com/much-training-data-required-machine-learning/>. [Último acceso: 14 Noviembre 2019].

CAPÍTULO 11: ANEXOS

Anexo A: Clasificación de OPS según 30 categorías

EPP	EPP - Oídos
	EPP - Ojos y Cara
	EPP - Cabeza
	EPP - Manos y Brazos
	EPP - Pie y Piernas
	EPP - Respiratorio
	EPP - Cuerpo
Línea de fuego/Posición de persona	Línea de fuego/Posición de persona - Golpeado contra
	Línea de fuego/Posición de persona - Atrapado entre
	Línea de fuego/Posición de persona - Golpeado por
	Línea de fuego/Posición de persona - Ojos (atención) en la tarea
	Línea de fuego/Posición de persona - Ojos (atención) en el camino
Área/Entorno de trabajo	Línea de fuego/Posición de persona - Potencial de caer
	Área/Entorno de trabajo - Pobre orden y mantención de lugar
	Área/Entorno de trabajo - Almacenamiento inadecuado
	Área/Entorno de trabajo - Área de trabajo sucia
	Área/Entorno de trabajo - Pobre iluminación
	Área/Entorno de trabajo - Clima extremo
	Área/Entorno de trabajo - Exposición a polvo /humos
Manejo manual	Área/Entorno de trabajo - No segregado
	Manejo manual - Izar y doblar
	Manejo manual - Sobre extensión
	Manejo manual - Estrecho
Procedimiento/Evaluación de Riesgo	Manejo manual - Postura inadecuada
	Procedimiento/Evaluación de Riesgo - Inadecuado para el trabajo
	Procedimiento/Evaluación de Riesgo - No conocido o no entendido
Herramientas y Equipos	Procedimiento/Evaluación de Riesgo - No seguido
	Herramientas y Equipos - Incorrecto para el trabajo
	Herramientas y Equipos - Usado incorrectamente
	Herramientas y Equipos - En condición insegura

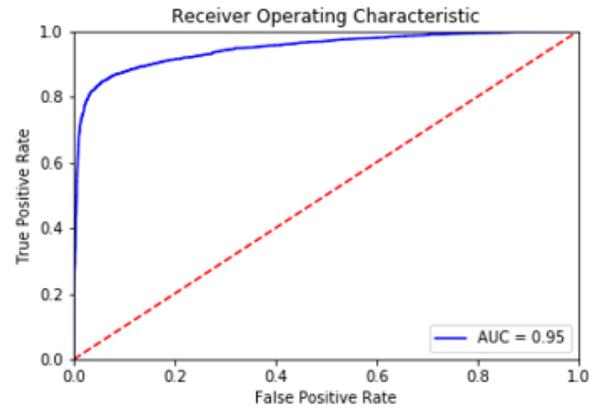
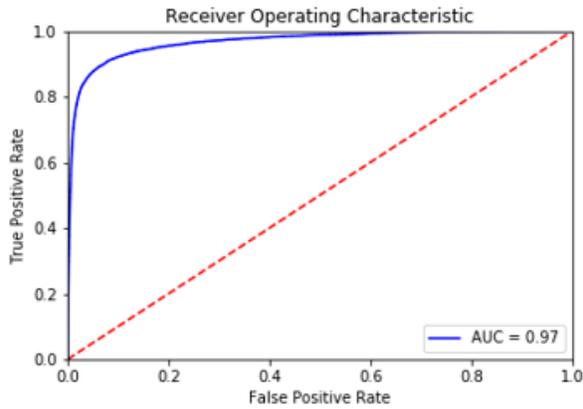
Anexo B: Ranking de las 10 palabras más usadas por categoría

EPP		POSICIÓN DE PERSONA		ÁREA DE TRABAJO	
observar	7560	trabajador	5192	área	5405
usar	6923	este	4715	personal	5202
este	5862	trabajar	4284	encontrar	4462
trabajar	5516	riesgo	3787	sector	4320
epp	4999	personal	3443	trabajador	4205
utilizar	4854	detener	2630	riesgo	3753
personal	4364	tarea	2612	actividad	3675
sin	4155	sector	2465	tarea	3555
guante	4085	converso	2369	estar	3068
tarea	4014	actividad	2353	trabajo	3049

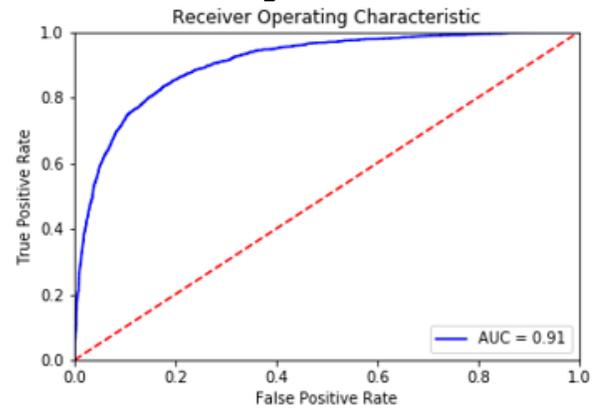
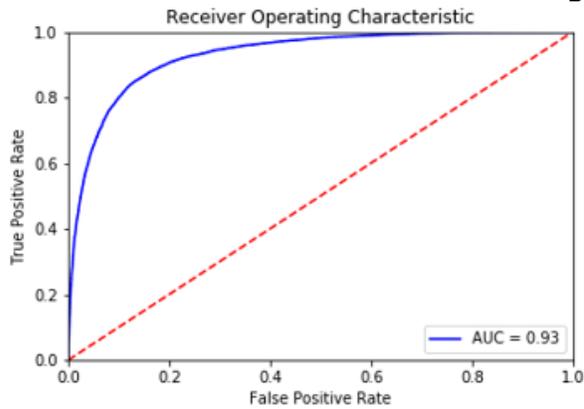
MANEJO MANUAL		PROCEDIMIENTO		HERRAMIENTAS Y EQUIPOS	
actividad	782	trabajar	6345	trabajador	3481
tarea	751	personal	5420	personal	2950
postura	721	riesgo	4500	estar	2720
personal	695	tarea	4236	encontrar	2611
detener	653	actividad	4060	herramienta	2350
riesgo	577	estar	3988	riesgo	2154
converso	490	encontrar	3819	sin	2054
bueno	425	converso	3604	detener	1948
formar	400	sin	3555	converso	1925
trabajo	399	detener	3473	tarea	1913

Anexo C: Curvas ROC modelo Logit

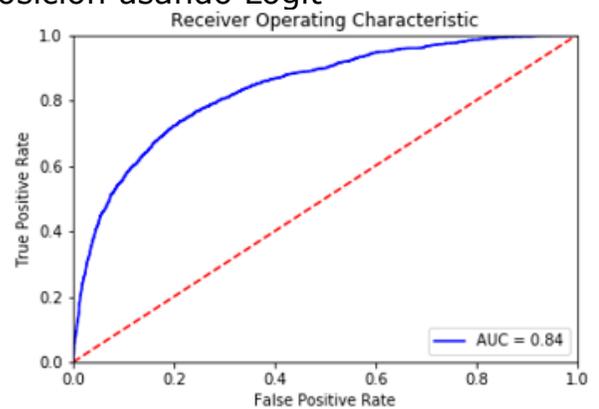
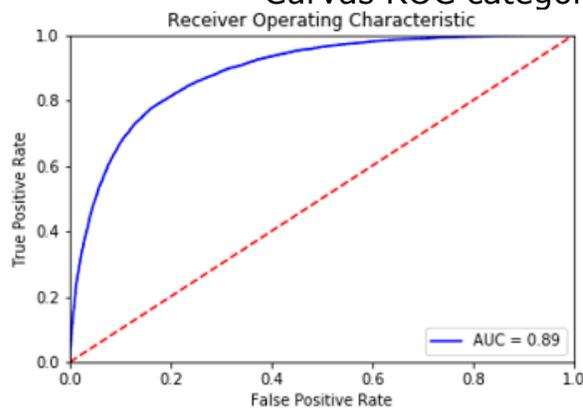
Curvas ROC categoría EPP usando Logit



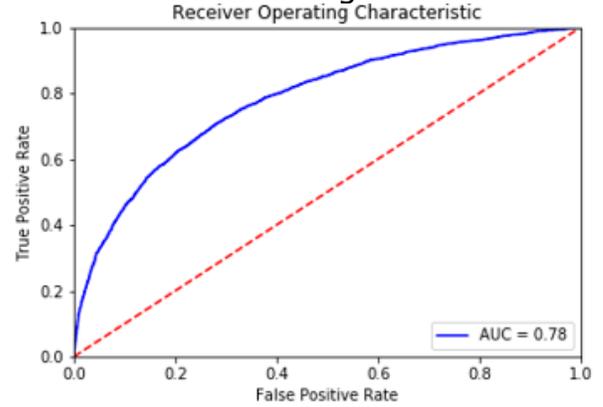
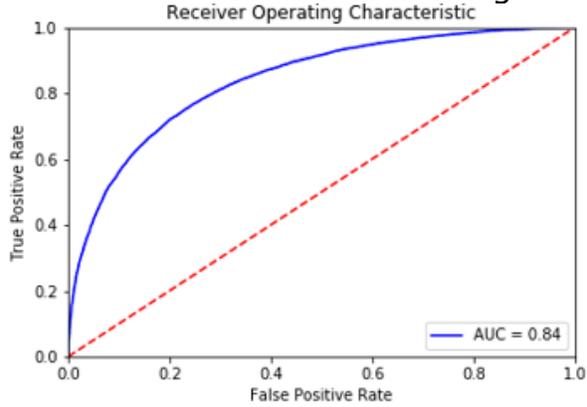
Curvas ROC categoría Área usando Logit



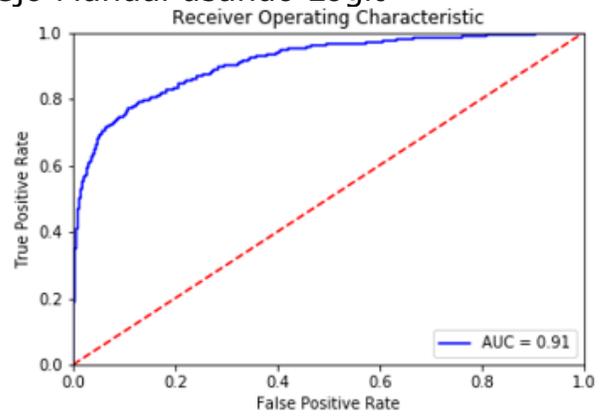
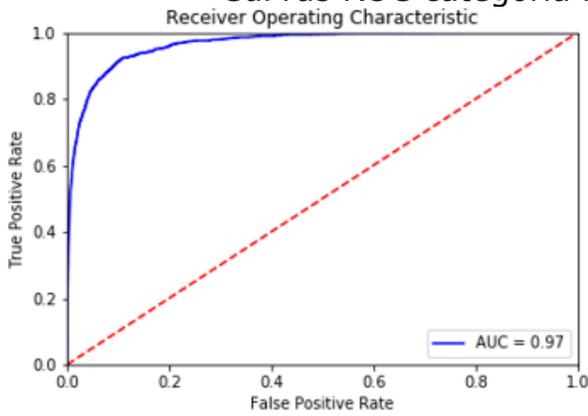
Curvas ROC categoría Posición usando Logit



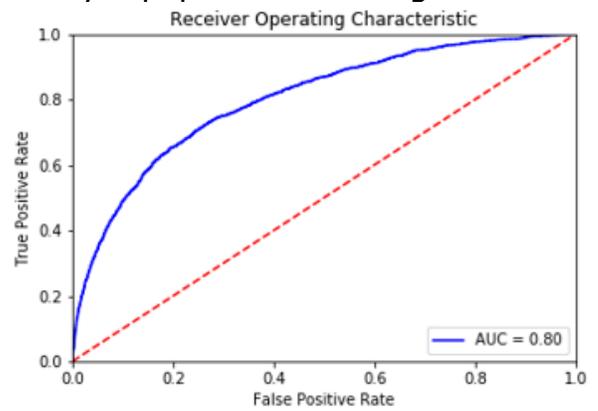
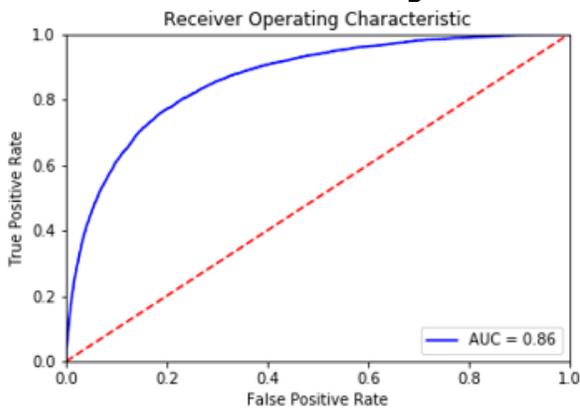
Curvas ROC categoría Procedimiento usando Logit



Curvas ROC categoría Manejo Manual usando Logit



Curvas ROC categoría Herramientas y equipos usando Logit

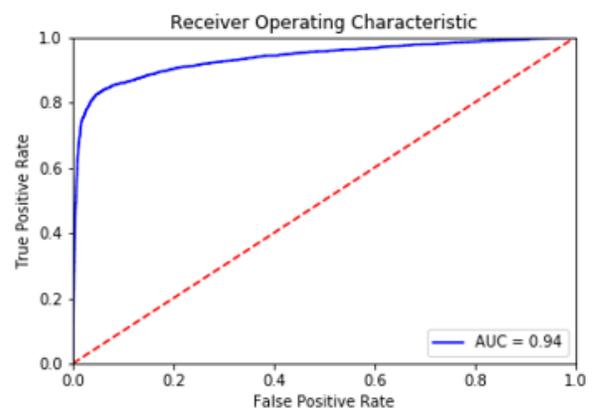
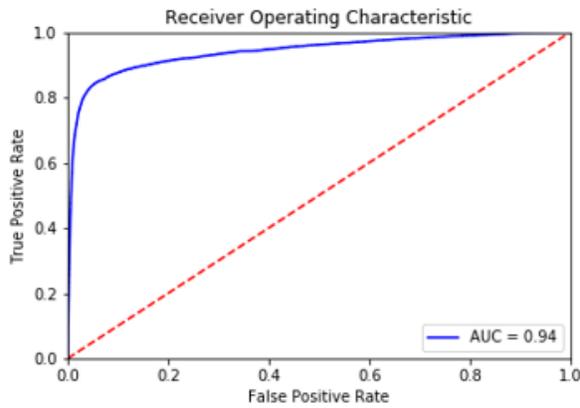


Anexo D: Métricas de evaluación Logit

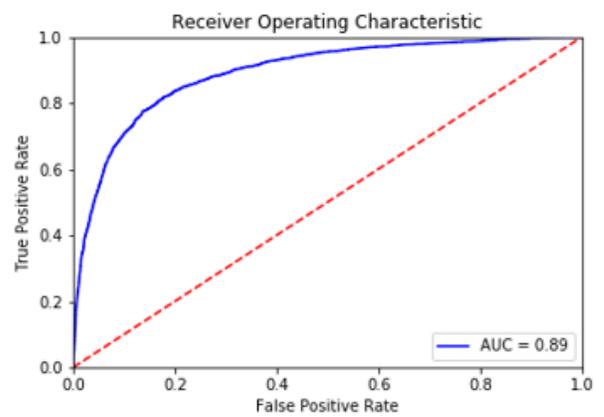
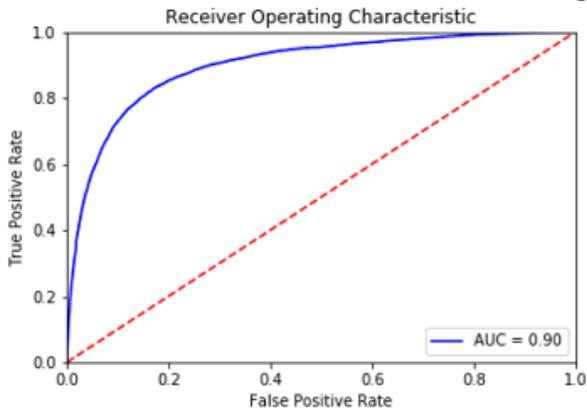
Categoría	Log Loss	Accuracy	Tasa de error	Ratio verdaderos positivos	Ratio Verdaderos Negativos	Ratio Falsos positivos	Precisión	Prevalence	F1
EPP (Train)	0.221	0.932	0.068	0.849	0.964	0.036	0.900	0.275	0.874
EPP	0.249	0.921	0.079	0.824	0.957	0.043	0.880	0.275	0.851
Línea de fuego/Posición de persona (Train)	0.358	0.800	0.200	0.815	0.796	0.204	0.507	0.204	0.625
Línea de fuego/Posición de persona	0.402	0.765	0.235	0.752	0.769	0.231	0.455	0.204	0.567
Área/Entorno de trabajo (Train)	0.328	0.871	0.129	0.812	0.893	0.107	0.740	0.273	0.774
Área/Entorno de trabajo	0.361	0.851	0.149	0.760	0.885	0.115	0.712	0.273	0.735
Manejo manual (Train)	0.089	0.909	0.091	0.897	0.910	0.090	0.298	0.041	0.447
Manejo manual	0.109	0.891	0.109	0.754	0.897	0.103	0.237	0.041	0.361
Herramientas y Equipos (Train)	0.363	0.747	0.253	0.835	0.727	0.273	0.407	0.183	0.547
Herramientas y Equipos	0.408	0.709	0.291	0.750	0.699	0.301	0.358	0.183	0.485
Procedimiento/Evaluación de Riesgo (Train)	0.482	0.736	0.264	0.808	0.704	0.296	0.542	0.303	0.649
Procedimiento/Evaluación de Riesgo	0.537	0.695	0.305	0.752	0.670	0.330	0.497	0.303	0.599

Anexo E: Curvas ROC modelo SVM

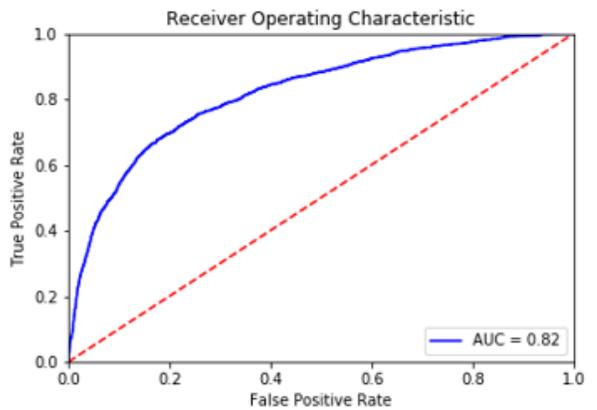
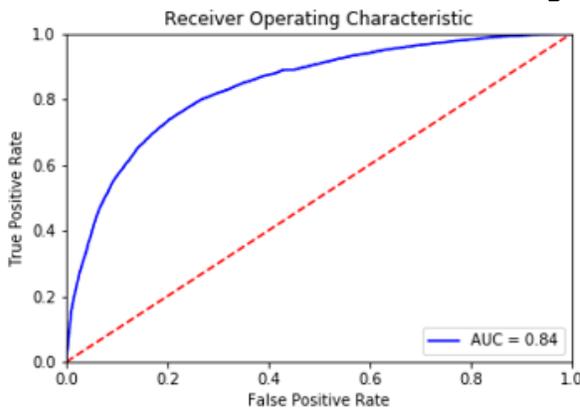
Curvas ROC categoría EPP usando SVM



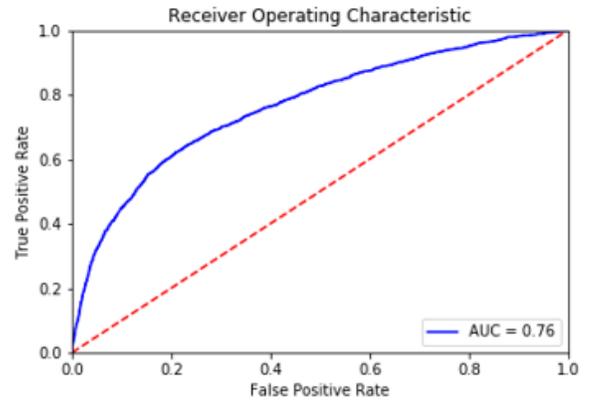
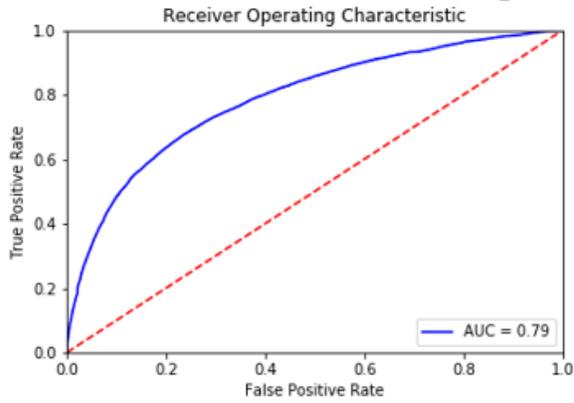
Curvas ROC categoría Área usando SVM



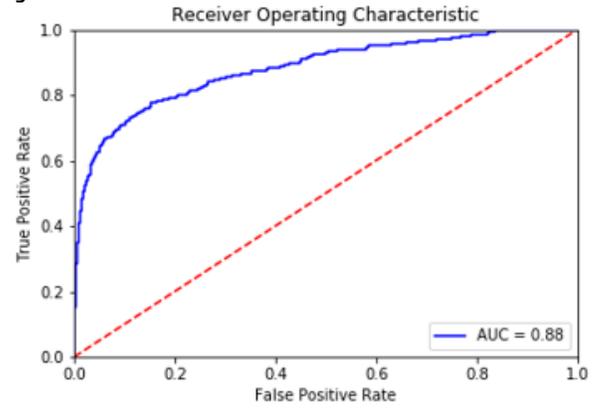
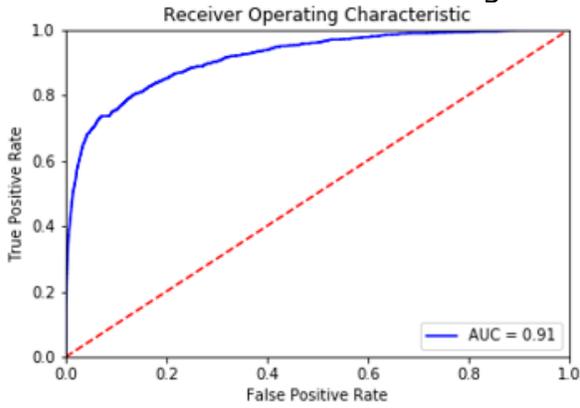
Curvas ROC categoría Posición usando SVM



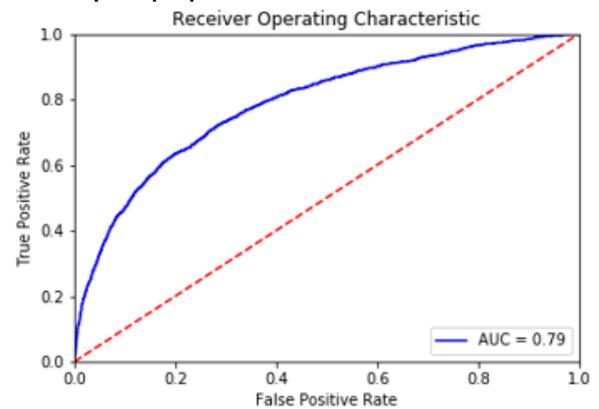
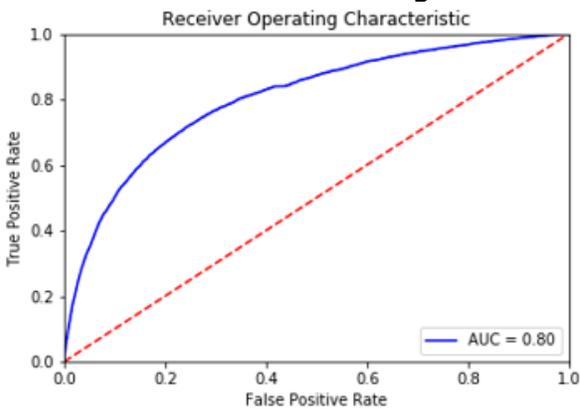
Curvas ROC categoría Procedimiento usando SVM



Curvas ROC categoría Manejo Manual usando SVM



Curvas ROC categoría Herramientas y equipos usando SVM



Anexo F: Métricas de evaluación SVM

Categoría	Log Loss	Accuracy	Tasa de error	Ratio verdaderos positivos	Ratio Verdaderos Negativos	Ratio Falsos positivos	Precisión	Prevalence	F1
EPP (Train)	0.247	0.921	0.079	0.784	0.972	0.028	0.915	0.275	0.844
EPP	0.251	0.919	0.081	0.781	0.971	0.029	0.911	0.275	0.841
Línea de fuego/Posición de persona (Train)	0.380	0.757	0.243	0.782	0.751	0.249	0.446	0.204	0.568
Línea de fuego/Posición de persona	0.392	0.744	0.256	0.756	0.741	0.259	0.428	0.204	0.547
Área/Entorno de trabajo (Train)	0.348	0.847	0.153	0.773	0.874	0.126	0.698	0.273	0.734
Área/Entorno de trabajo	0.359	0.841	0.159	0.753	0.874	0.126	0.692	0.273	0.721
Manejo manual (Train)	0.109	0.858	0.142	0.805	0.860	0.140	0.197	0.041	0.316
Manejo manual	0.119	0.855	0.145	0.757	0.860	0.140	0.187	0.041	0.300
Herramientas y Equipos (Train)	0.382	0.709	0.291	0.770	0.696	0.304	0.362	0.183	0.492
Herramientas y Equipos	0.391	0.686	0.314	0.758	0.670	0.330	0.340	0.183	0.469
Procedimiento/Evaluación de Riesgo (Train)	0.494	0.675	0.325	0.785	0.627	0.373	0.477	0.303	0.594
Procedimiento/Evaluación de Riesgo	0.512	0.655	0.345	0.760	0.610	0.390	0.458	0.303	0.571

Anexo G: Estructura red neuronal en Python usando Keras

```
model = Sequential()
model.add(layers.Embedding(vocab_size, embedding_dim,
input_length=maxlen))
model.add(layers.Conv1D(500, 5, activation='relu'))
model.add(layers.GlobalMaxPooling1D())
model.add(layers.Dropout(0.5))
model.add(layers.Dense(50, activation='relu'))
model.add(layers.Dropout(0.4))
model.add(layers.Dense(6, activation='sigmoid'))
model.compile(optimizer='adam',
              loss=weighted_binary_crossentropy,
              metrics=['binary_accuracy'])
```

Anexo H: Loss Function red neuronal

```
import tensorflow as tf
import keras.backend.tensorflow_backend as tfb

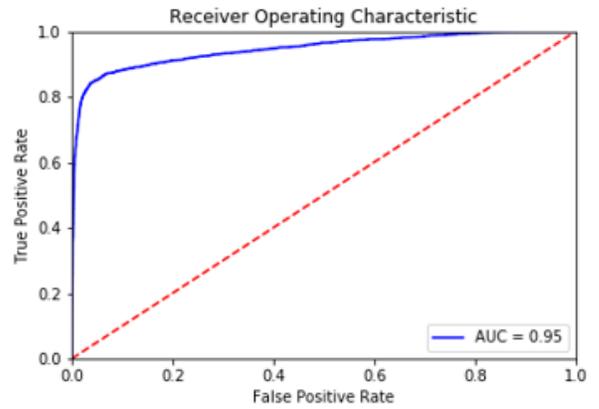
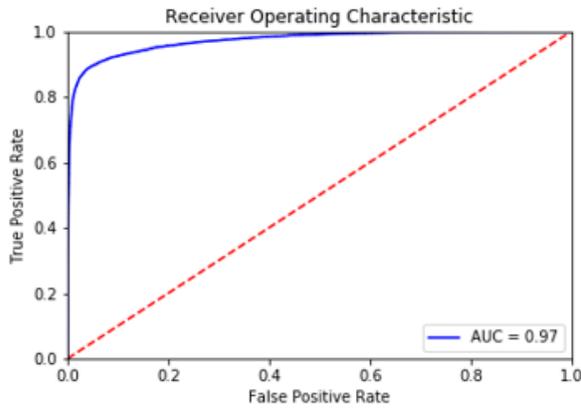
POS_WEIGHT = 3

def weighted_binary_crossentropy(target, output):
    """
    Weighted binary crossentropy between an output tensor
    and a target tensor. POS_WEIGHT is used as a multiplier
    for the positive targets.

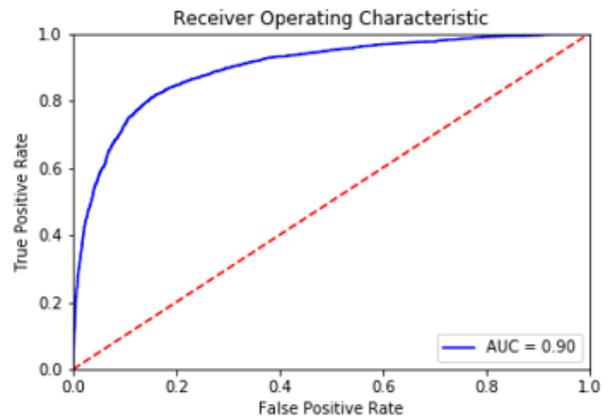
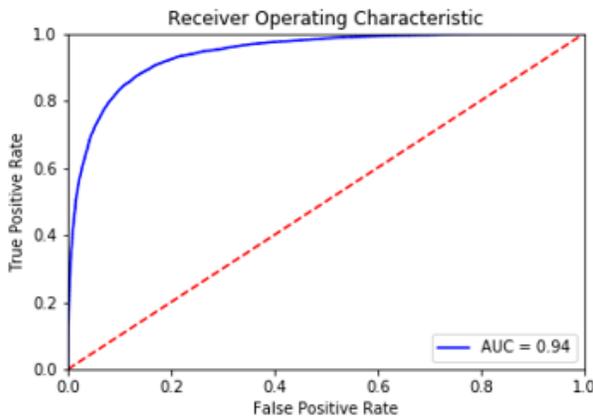
    Combination of the following functions:
    * keras.losses.binary_crossentropy
    * keras.backend.tensorflow_backend.binary_crossentropy
    * tf.nn.weighted_cross_entropy_with_logits
    """
    # transform back to logits
    _epsilon = tfb._to_tensor(tfb.epsilon(), output.dtype.base_dtype)
    output = tf.clip_by_value(output, _epsilon, 1 - _epsilon)
    output = tf.log(output / (1 - output))
    # compute weighted loss
    loss = tf.nn.weighted_cross_entropy_with_logits(targets=target,
                                                    logits=output,
                                                    pos_weight=POS_WEIGHT)
    return tf.reduce_mean(loss, axis=-1)
```

Anexo I: Curvas ROC usando Redes Neuronales

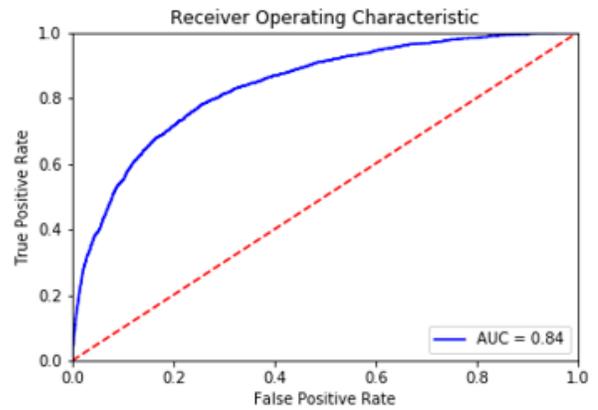
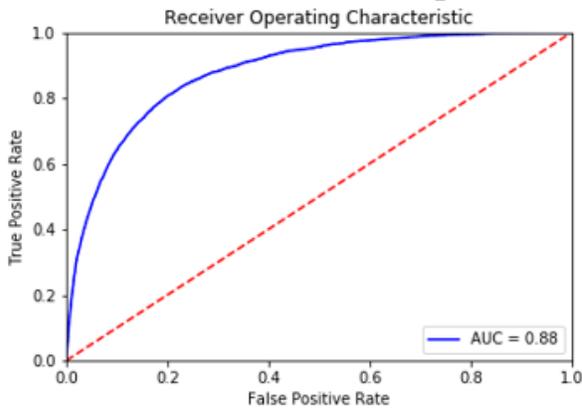
Curvas ROC categoría EPP usando Redes Neuronales



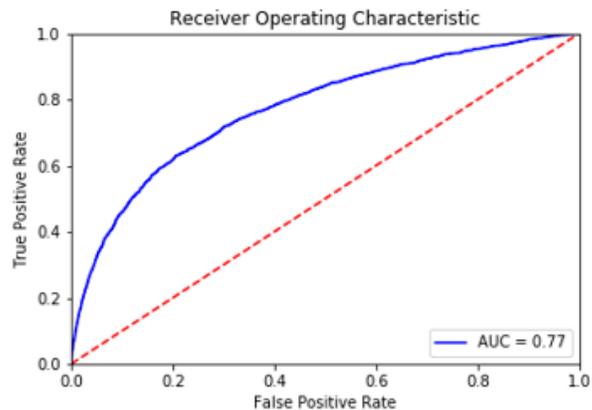
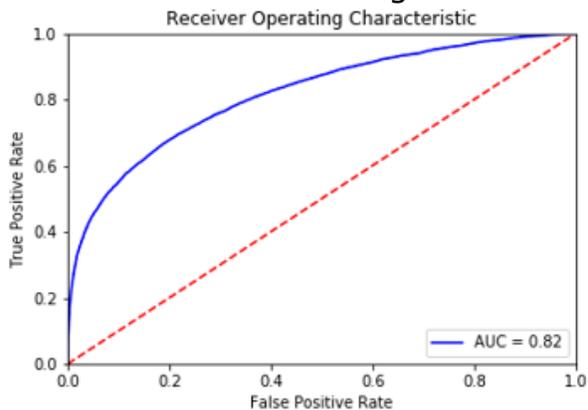
Curvas ROC categoría Área usando Redes Neuronales



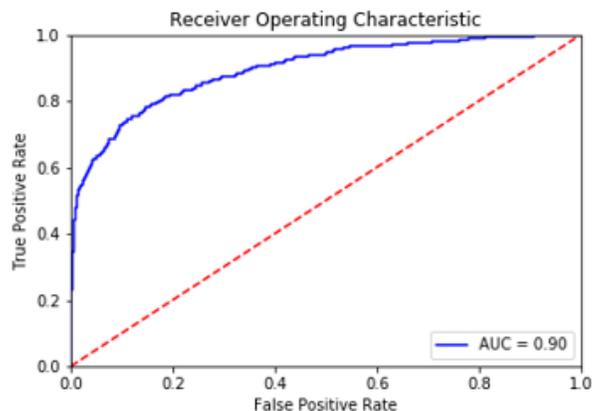
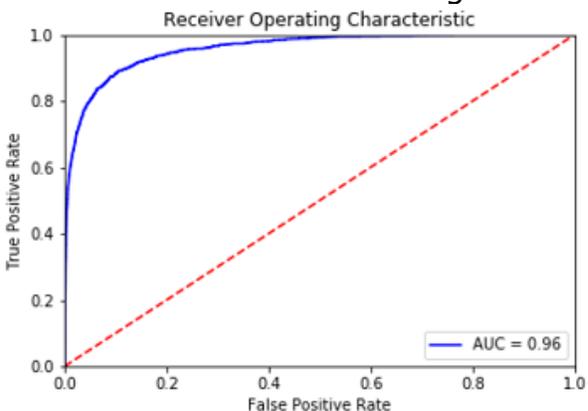
Curvas ROC categoría Posición usando Redes Neuronales



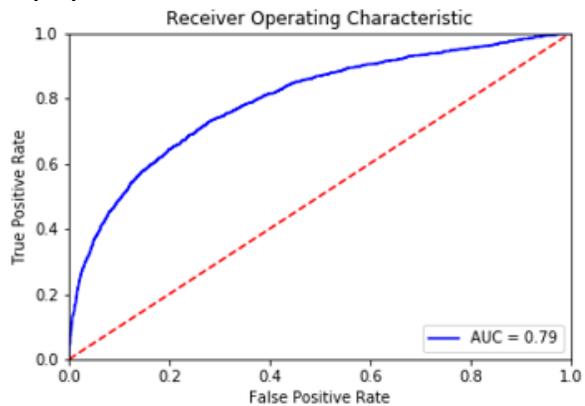
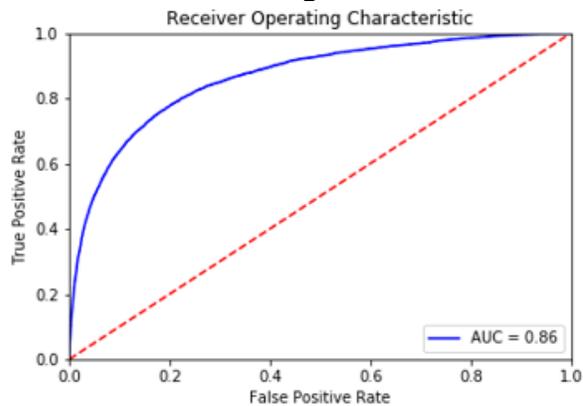
Curvas ROC categoría Procedimiento usando Redes Neuronales



Curvas ROC categoría Manejo Manual Redes Neuronales



Curvas ROC categoría Herramientas y equipos usando Redes Neuronales



Anexo J: Métricas de evaluación Red Neuronal

Categoría	Log Loss	Accuracy	Tasa de error	Ratio verdaderos positivos	Ratio Verdaderos Negativos	Ratio Falsos positivos	Precisión	Prevalence	F1
EPP (Train)	0.219	0.943	0.057	0.864	0.973	0.027	0.923	0.276	0.893
EPP	0.259	0.930	0.070	0.837	0.965	0.035	0.899	0.272	0.867
Línea de fuego/Posición de persona (Train)	0.413	0.794	0.206	0.815	0.789	0.211	0.498	0.204	0.619
Línea de fuego/Posición de persona	0.447	0.764	0.236	0.751	0.767	0.233	0.453	0.204	0.565
Área/Entorno de trabajo (Train)	0.340	0.866	0.134	0.878	0.861	0.139	0.702	0.272	0.780
Área/Entorno de trabajo	0.405	0.833	0.167	0.819	0.838	0.162	0.662	0.279	0.732
Manejo manual (Train)	0.099	0.891	0.109	0.891	0.890	0.110	0.261	0.042	0.404
Manejo manual	0.121	0.870	0.130	0.753	0.875	0.125	0.192	0.038	0.306
Herramientas y Equipos (Train)	0.456	0.747	0.253	0.837	0.727	0.273	0.406	0.182	0.546
Herramientas y Equipos	0.510	0.696	0.304	0.755	0.682	0.318	0.351	0.185	0.479
Procedimiento/Evaluación de Riesgo (Train)	0.592	0.683	0.317	0.809	0.629	0.371	0.485	0.302	0.606
Procedimiento/Evaluación de Riesgo	0.640	0.652	0.348	0.787	0.593	0.407	0.461	0.307	0.581