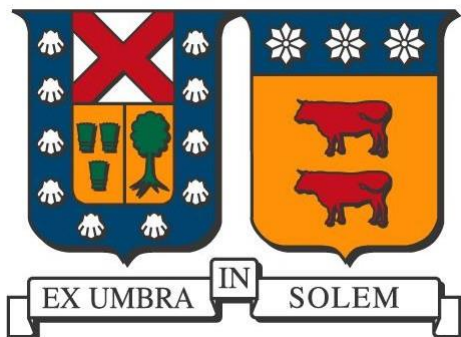


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INGENIERÍA DE MINAS, METALURGIA Y
MATERIALES



DESARROLLO DE UN INDICADOR SÍSMICO EN BASE AL
DESEMPEÑO DE HERRAMIENTAS DE MACHINE LEARNING

FELIPE ANDRÉS MUÑOZ CAMPOS

Memoria de Título para optar al título profesional de Ingeniero Civil de Minas

Profesor Guía
Gonzalo Nelis Suazo

Profesor Co-Guía
Rodrigo Estay Huidobro

2025
Santiago, Chile



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título; ☐ Tesis de Postgrado;

Título del trabajo: Desarrollo de un indicador sísmico en base al desempeño de herramientas de machine learning

Nombre del candidato(a): Felipe Andrés Muñoz Campos

Carrera / Grado: Ingeniería Civil de Minas

Campus: Santiago San Joaquín; Departamento: Ingeniería de Minas, Metalurgia y Materiales

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Gonzalo Nelis Suazo, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL

☒ El trabajo **NO** contiene información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (embargo) por:

☐ 6 meses; ☐ 12 meses; ☐ 2 años; ☐ 3 años; ☐ 5 años; ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 11-07-2025

; Firma:

Estudiante o Candidato(a):

Fecha: 11-07-2025

; Firma:

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Agradecimientos

A mis padres, Jessica y Hernán, quienes me dieron la vida y siempre han estado para apoyarme.

A mi hermana menor, Fernanda, quien tuvo que soportarme muchas madrugadas y ayudó a que el tiempo pasara volando.

También quiero agradecer a mis vecinos: Ricardo, Benjamín, Manuel, Bastián y muchos más, por todas las risas y las noches de juegos.

Agradezco a Sebastián, Reinaldo, Francisca y a todo el equipo de SONAMI, quienes me permitieron realizar mi práctica profesional y confiaron en mí para crecer como profesional.

Gracias también a todos los ayudados que tuve. Ustedes me mostraron que la docencia puede ser muy entretenida.

A mis amigos que me acompañaron durante todos estos años: Bastián, Iván, Wilson, Diego, Natalia, Benjamín, y a todos quienes, de alguna forma, aportaron su granito de arena para que hoy pueda cerrar esta etapa.

A mi mejor amigo, Javier, quien ha estado a mi lado durante estos seis años. Sin ti, nada de esto hubiera sido posible.

A mi pareja, Valentina, soy feliz de haber podido conocerte y enamorarme de ti durante estos últimos años. Tu compañía ha sido fundamental en los momentos más duros.

También una mención especial a parte del cuerpo académico del departamento, a Sergio Palma, Fabián Manríquez y Leslie, a quienes les tengo mucho cariño.

Finalmente, agradezco profundamente a mis profesores guía, Gonzalo Nelis y Rodrigo Estay. Solo tengo buenos recuerdos con ambos, y gracias a ustedes este trabajo de título fue posible. Estoy muy agradecido con ambos y espero que la vida nos permita seguir trabajando juntos.

A todos ustedes, gracias por ser parte de mi vida y por ayudarme a alcanzar este importante logro.

Resumen

La sismicidad inducida representa uno de los principales desafíos para la operación segura y eficiente de minas subterráneas, especialmente en aquellas que utilizan el método de hundimiento por bloques. En este contexto, se desarrolló una metodología para construir un indicador sísmico dinámico basado en ventanas móviles temporales, aplicado a datos reales de una mina chilena.

Se generaron indicadores mediante el agrupamiento de variables derivadas del catálogo sísmico en ventanas de 24 y 48 horas, las cuales fueron utilizadas para entrenar y evaluar 28 modelos de clasificación supervisada, incluyendo árboles de decisión, máquinas de soporte vectorial, redes neuronales y algoritmos de vecinos más cercanos.

Los resultados muestran que el algoritmo *K-Nearest Neighbors* (kNN) obtuvo el mejor desempeño general, alcanzando niveles de exactitud superiores al 95% en escenarios específicos. Además, se identificó que el uso de umbrales de magnitud más altos mejora la capacidad de los modelos para discriminar entre situaciones de mayor o menor peligro sísmico.

La metodología propuesta demostró su potencial como herramienta complementaria para la categorización temprana de condiciones que podrían estar asociadas a la ocurrencia de eventos sísmicos relevantes, contribuyendo así al fortalecimiento de las estrategias de gestión del peligro en minería subterránea.

Abstract

Induced seismicity represents one of the main challenges for the safe and efficient operation of underground mines, particularly those employing the block caving method. In this context, a methodology was developed to construct a seismic indicator based on moving time windows, applied to real data from a Chilean underground mine.

Dynamic indicators were generated using variables derived from the seismic catalog, grouped into 24- and 48-hour windows. Subsequently, 28 supervised classification models were trained and evaluated, including decision trees, support vector machines, neural networks, and K-Nearest Neighbors algorithms.

The results show that the K-Nearest Neighbors (KNN) algorithm achieved the best overall performance, with accuracy levels exceeding 95% in specific scenarios. Furthermore, it was observed that using higher magnitude thresholds improves the models' ability to discriminate between situations of higher or lower seismic hazard.

The proposed methodology demonstrated its potential as a complementary tool for the early categorization of conditions that could be associated with the occurrence of relevant seismic events, thus contributing to improved hazard management in underground mining operations.

Índice

1	Introducción	10
1.1	Hipótesis.....	11
1.2	Objetivos	11
1.2.1	Objetivo general	11
1.2.2	Objetivos específicos	11
1.3	Alcances	11
2	Antecedentes y marco teórico	12
2.1	Indicadores sísmicos y machine learning.....	12
2.2	Descripción y tratamiento de datos	15
2.3	Análisis de componentes principales (PCA).....	17
2.4	Matriz de contingencia	18
3	Metodología	21
3.1	Agrupamiento temporal	21
3.2	Reducción de dimensionalidad (PCA).....	23
3.3	Modelos de machine learning	24
3.4	Evaluación del desempeño.....	24
4	Resultados	25
4.1	Desempeño por modelos de machine learning	25
4.1.1	Desempeño árboles de decisión	25
4.1.2	Desempeño modelos discriminantes	26
4.1.3	Desempeño regresión logística	26
4.1.4	Desempeño de clasificadores bayesianos	27
4.1.5	Desempeño máquinas de vectores de soporte (support vector machine).....	27
4.1.6	Desempeño vecinos más cercanos	28
4.1.7	Desempeño métodos de ensamble	28
4.1.8	Desempeño redes neuronales	29
4.2	Desempeño por magnitudes	30
4.3	Desempeño por agrupamientos temporales	31
4.3.1	Desempeño por ventana de tiempo	31
4.3.2	Desempeño por tiempos futuros	32
4.3.3	Desempeño por tiempos de avance	32
4.4	Resultados finales	33
5	Conclusiones	34

6 Referencias..... 36

7 Anexo 38

Índice de figuras

Figura 2.1 Vista isométrica de los eventos en base a su ubicación, los colores representan la magnitud del evento (azul las magnitudes más bajas y rojo las más altas).	16
Figura 2.2 Histograma de las magnitudes antes de filtrar los eventos.	16
Figura 2.3 Ejemplo de PCA (Scholz, 2006). Inicialmente se tienen datos en 3 dimensiones (izquierda), al aplicar PCA se convierten en 2 componentes (derecha).	18
Figura 2.4 Ejemplo de curva ROC para 3 modelos distintos, modelo A representa alta precisión, modelo B es aleatorio y modelo C es un mal modelo.	20
Figura 3.1 Elección de magnitud de completitud (-0,9), utilizando el método de máxima curvatura (Wiemer & Wyss, 2000; Woessner & Wiemer, 2005).	21
Figura 3.2 Esquema del agrupamiento, Δt_i indica la duración de cada grupo, Δt_f el tiempo de búsqueda posterior, Δt_a el tiempo de avance para el siguiente grupo y $M_w > M_w^*$ la ocurrencia de un evento mayor a la magnitud relevante.	23
Figura 3.3 PCA obtenido. 4 componentes representan el 95% de la varianza explicada.	24
Figura 4.1 Gráfico ROC para indicador mediante árboles de decisión.	25
Figura 4.2 Gráfico ROC para indicador mediante modelos discriminantes.	26
Figura 4.3 Gráfico ROC para indicador mediante regresión logística.	26
Figura 4.4 Gráfico ROC para indicador mediante clasificadores bayesianos.	27
Figura 4.5 Gráfico ROC para indicador mediante SVM.	27
Figura 4.6 Gráfico ROC para indicador mediante kNN.	28
Figura 4.7 Gráfico ROC para indicador mediante modelos de ensamble.	28
Figura 4.8 Gráfico ROC para indicador mediante redes neuronales.	29
Figura 4.9 Gráfico ROC para indicador para distintas magnitudes relevantes.	30
Figura 4.10 Gráfico ROC para agrupamiento de distintas ventanas de tiempo.	31
Figura 4.11 Gráfico ROC para distintos tiempos de búsqueda futuros.	32
Figura 4.12 Gráfico ROC para distintos tiempos de avance del catálogo.	32
Figura 4.13 Diagrama ROC para el escenario con el mejor desempeño (Δt : 24 horas, Δt_f : 8 horas, Δt_a : 10 minutos).	33

Índice de tablas

Tabla 2.1 Descripción, ventajas y desventajas de los modelos de machine learning a utilizar	13
Tabla 2.2 Estadísticas básicas de parámetros a utilizar antes de filtrar eventos	17
Tabla 2.3. Matriz de contingencia (TP: Acierto positivo, FP: Falsa alarma, FN: Fallo, TN: Acierto negativo)	19

Anexos

Anexo 1 Precisiones para Δt : 24 horas, Δt_f : 8 horas, Δt_a : 10 minutos	38
Anexo 2 Precisiones para Δt : 24 horas, Δt_f : 8 horas, Δt_a : 20 minutos	39
Anexo 3 Precisiones para Δt : 24 horas, Δt_f : 12 horas, Δt_a : 10 minutos	40
Anexo 4 Precisiones para Δt : 24 horas, Δt_f : 12 horas, Δt_a : 20 minutos	41
Anexo 5 Precisiones para Δt : 48 horas, Δt_f : 8 horas, Δt_a : 10 minutos	42
Anexo 6 Precisiones para Δt : 48 horas, Δt_f : 8 horas, Δt_a : 20 minutos	43
Anexo 7 Precisiones para Δt : 48 horas, Δt_f : 12 horas, Δt_a : 10 minutos	44
Anexo 8 Precisiones para Δt : 48 horas, Δt_f : 12 horas, Δt_a : 20 minutos	45

1 Introducción

La minería en Chile es una de las principales actividades económicas del país, representando un 11,7% del Producto Interno Bruto (PIB) en el año 2024, según datos del Banco Central. Además, de acuerdo con la Sociedad Nacional de Minería (SONAMI), Chile es el mayor productor de cobre a nivel mundial, con una producción anual de 5,51 millones de toneladas de cobre fino, equivalente al 24% de la producción global durante el 2024.

No obstante, el agotamiento progresivo de los yacimientos superficiales ha obligado a la industria a migrar hacia métodos de explotación subterránea, los cuales presentan mayores desafíos técnicos, operacionales y geomecánicos. En particular, el método de hundimiento por bloques (*block caving*) se ha consolidado como una alternativa eficiente para la extracción masiva a gran escala, como lo evidencia la mina Chuquicamata Subterránea, que en 2019 inició su operación con este método tras más de un siglo de explotación a rajo abierto.

Uno de los principales desafíos asociados a este tipo de minería es la sismicidad inducida, fenómeno que corresponde a la liberación repentina de energía acumulada en el macizo rocoso como consecuencia de la redistribución de esfuerzos. Esta actividad sísmica puede generar eventos de alta magnitud, con capacidad de dañar infraestructuras, interrumpir operaciones y poner en riesgo la seguridad del personal. El entendimiento de los fenómenos sísmicos debido a la actividad minera en procesos de *caving*, es limitado por la complejidad de los problemas que se suscitan (Ortega, 2014)

El desarrollo e implementación de sistemas de monitoreo sísmico ha permitido registrar en tiempo real el comportamiento del macizo rocoso, proporcionando datos esenciales para estimar la sismicidad. A partir de esta información, se han desarrollado diversos indicadores sísmicos, los cuales buscan sintetizar la actividad del macizo e identificar condiciones críticas para activar protocolos de alerta. En este contexto, el uso de técnicas de machine learning (ML) ha emergido como una herramienta prometedora, al permitir detectar patrones en grandes volúmenes de datos sísmicos y construir modelos que mejoran la anticipación de eventos de alto riesgo.

El presente trabajo propone una evaluación del desempeño de indicadores sísmicos construidos mediante herramientas de ML, utilizando una base de datos real proveniente de una mina chilena. Mediante un enfoque temporal, se generan conjuntos de datos utilizados para entrenar modelos de clasificación. Estos modelos buscan anticipar la ocurrencia de eventos sísmicos relevantes dentro

de horizontes de predicción. El rendimiento fue evaluado con métricas de desempeño, bajo distintos umbrales de magnitud.

1.1 Hipótesis

Es posible construir indicadores sísmicos mediante técnicas de machine learning, y a partir de los resultados obtenidos evaluar su desempeño, logrando así mejorar la capacidad predictiva frente a eventos sísmicos de alta magnitud.

1.2 Objetivos

1.2.1 Objetivo general

Evaluar el desempeño de un indicador de peligro sísmico construido en base al desempeño de herramientas de machine learning.

1.2.2 Objetivos específicos

- Identificar las variables más representativas de un catálogo sísmico para la evaluación de indicadores sísmicos, aplicando técnicas de reducción de dimensionalidad (PCA).
- Entrenar distintos modelos de clasificación supervisada y evaluar su desempeño utilizando métricas como matriz de contingencia y curva ROC.

1.3 Alcances

Este estudio se enmarca en el análisis de sismicidad inducida en minería subterránea, utilizando un conjunto de datos reales registrados en una faena nacional que opera bajo el método de hundimiento por bloques.

El análisis se enfoca en la evaluación temporal de indicadores sísmicos, contruidos a partir de agrupaciones de eventos en ventanas móviles, considerando magnitudes superiores a ciertos umbrales y horizontes de predicción de corto plazo. La evaluación del desempeño se realiza mediante métricas de desempeño, sin incorporar variables geomecánicas ni geológicas adicionales.

Por lo tanto, los resultados y conclusiones presentados en esta memoria se circunscriben a las condiciones particulares del conjunto de datos utilizado. Para su generalización o aplicación en otras minas, se requerirá un proceso de validación complementario adaptado a las características específicas de cada operación.

2 Antecedentes y marco teórico

La sismicidad inducida es un fenómeno que ocurre como resultado de la redistribución de tensiones al interior del macizo rocoso producto de las excavaciones subterráneas. En el contexto de la minería por hundimiento, este fenómeno se intensifica debido a la magnitud del colapso planificado del macizo, lo cual genera eventos sísmicos de diversa intensidad. Estos eventos pueden afectar la estabilidad de la infraestructura minera, interrumpir la continuidad operacional e incluso- representar un riesgo crítico para la integridad de los trabajadores.

Para monitorear este comportamiento, las minas suelen instalar sistemas sísmicos compuestos por geófonos y acelerómetros, los cuales permiten detectar el momento, ubicación y magnitud de los eventos. Esta información es procesada y organizada en catálogos sísmicos que contienen parámetros fundamentales para el análisis operacional y geomecánico.

De acuerdo con Vallejos & McKinnon (2013) y González et al. (2021), la sismicidad inducida representa uno de los principales desafíos operacionales en minas subterráneas de gran escala, donde la profundidad y el volumen involucrado incrementan la probabilidad de eventos de alta energía. Frente a este escenario, resulta esencial contar con herramientas capaces de anticipar eventos sísmicos significativos, permitiendo aplicar protocolos preventivos o correctivos en las labores afectadas.

2.1 Indicadores sísmicos y machine learning

Los indicadores sísmicos sirven para monitorear la respuesta en el tiempo del macizo rocoso a la minería aplicada, utilizando diversos parámetros obtenidos a partir de los registros de sismicidad inducida mediante geófonos o acelerómetros instalados en la mina (Estay, 2014). Su propósito es advertir sobre posibles situaciones de riesgo al detectar niveles inusuales que puedan señalar un comportamiento inestable del macizo.

El aprendizaje automático es una rama de la inteligencia artificial que permite que los computadores aprendan patrones a partir de datos históricos sin necesidad de ser programados explícitamente para cada tarea. En contextos como la minería subterránea, el ML puede ser utilizado para construir modelos capaces de anticipar eventos sísmicos de alta magnitud, a partir de parámetros derivados del monitoreo sísmico continuo. La Tabla 2.1 presenta los modelos de ML que se utilizarán.

Tabla 2.1 Descripción, ventajas y desventajas de los modelos de machine learning a utilizar.

Método	Breve descripción	Ventajas	Desventajas	Referencia
Árbol de decisión (fino/medio/grueso)	Divide el espacio de predicción mediante umbrales en los atributos. La profundidad controla complejidad y sobreajuste.	Interpretables, rápidos, capturan relaciones no lineales.	Sensibles al sobreajuste si son muy profundos.	(Hastie et al., 2009)
Análisis Discriminante Lineal (LDA)	Proyecta datos maximizando la separación lineal entre clases, asumiendo covarianzas iguales y distribución gaussiana.	Simple, eficiente en datos linealmente separables.	No relaciones lineales.	(Hastie et al., 2009)
Análisis Discriminante Cuadrático (QDA)	Similar al LDA, pero permite covarianzas diferentes entre clases, adaptándose a fronteras no lineales.	Mayor flexibilidad, adecuado para relaciones complejas.	Requiere más datos, puede sobreajustar.	(Hastie et al., 2009)
Regresión logística	Modelo lineal que estima la probabilidad de clase usando la función logística.	Probabilístico, interpretable, robusto en datos simples.	Limitado en relaciones lineales.	(Hosmer et al., 2013)
Naive Bayes Gaussiano	Clasificador probabilístico bajo supuestos de independencia y distribución normal de las variables.	Eficiente, funciona bien si las suposiciones se cumplen.	Sensible a violaciones de la independencia o no normalidad.	(Hastie et al., 2009)
Naive Bayes con kernel	Similar al anterior, pero estima distribuciones sin suponer forma paramétrica.	Flexible, adecuado para datos gaussianos.	Puede requerir mayor tiempo de cómputo.	(Hastie et al., 2009)
SVM lineal	Clasifica mediante un hiperplano, útil para datos linealmente separables.	Rápido, eficaz en alta dimensionalidad.	No relaciones lineales.	(Cortes y Vapnik, 1995)
SVM polinómico (cuadrático/cúbico)	Usa kernels polinómicos para modelar fronteras curvas.	Mayor flexibilidad frente a relaciones complejas.	Puede sobreajustar, requiere sintonización de hiperparámetros.	(Cortes y Vapnik, 1995)

SVM Gaussiano (fino/medio/grueso)	Kernel RBF que permite fronteras altamente no lineales; el parámetro de suavidad regula complejidad.	Alta capacidad de modelar relaciones complejas.	Riesgo de sobreajuste si no se regula adecuadamente.	(Cortes y Vapnik, 1995)
kNN (fino/medio/grueso)	Clasifica según la mayoría de los k vecinos más cercanos; la cantidad de vecinos regula sensibilidad y robustez.	Intuitivo, no paramétrico.	Sensible al ruido y escala de los datos.	(Cover y Hart, 1967)
kNN con distancia del coseno	Usa similitud angular en lugar de distancia euclidiana, útil en datos direccionales.	Robusto en contextos donde importa más la dirección que la magnitud.	Requiere normalización adecuada de los datos.	(Cover y Hart, 1967)
kNN cúbico	Emplea distancia cúbica, modificando el peso de las distancias.	Ajusta el criterio de vecindad según contexto.	Más sensible a la configuración de la métrica.	(Cover y Hart, 1967)
kNN ponderado	Da mayor peso a vecinos cercanos, mejorando la sensibilidad a patrones locales.	Captura relaciones locales importantes.	Puede ser sensible al ruido en los datos.	(Cover y Hart, 1967)
Árboles agregados	Combina múltiples árboles sobre subconjuntos de datos para reducir la varianza.	Mejora precisión y estabilidad.	Menos interpretable, requiere más recursos.	(Breiman, 1996)
Árboles potenciados	Entrena árboles secuencialmente, corrigiendo los errores del anterior.	Alta precisión, capta relaciones complejas.	Riesgo de sobreajuste, sensible al ruido.	(Freund y Schapire, 1997)
Árboles potenciados con submuestreo	Boosting con submuestreo aleatorio de la clase mayoritaria, útil para datos desbalanceados.	Mejora rendimiento en clases de desequilibrio.	Puede complicar la interpretación y requiere ajuste fino.	(Freund y Schapire, 1997)
Red neuronal angosta	Una capa oculta con pocas neuronas, baja complejidad.	Bajo riesgo de sobreajuste, rápido.	Capacidad limitada para relaciones complejas.	(Goodfellow et al., 2016)
Red neuronal media	Una capa oculta, cantidad intermedia de neuronas.	Balance entre capacidad de representación y generalización.	Requiere más datos que la red angosta.	(Goodfellow et al., 2016)

Red neuronal ancha	Una capa oculta con muchas neuronas, alta capacidad de modelado.	Captura relaciones complejas.	Requiere muchos datos y regularización.	(Goodfellow et al., 2016)
Red neuronal de dos capas ocultas	Dos capas ocultas, permite representaciones jerárquicas.	Modela estructuras complejas en los datos.	Mayor riesgo de sobreajuste, mayor complejidad computacional.	(Goodfellow et al., 2016)
Red neuronal de tres capas ocultas	Aumenta la profundidad y capacidad del modelo.	Capacidad de aprender patrones altamente complejos.	Requiere gran cantidad de datos y técnicas de regularización.	(Goodfellow et al., 2016)

Para la construcción de estos modelos se divide el conjunto de datos en dos partes: una destinada al entrenamiento, donde el algoritmo aprende los patrones del comportamiento sísmico; y otra reservada para la validación, en la que se evalúa la capacidad del modelo para realizar predicciones sobre datos no vistos previamente.

Estas técnicas han demostrado ser especialmente útiles para resolver problemas de clasificación, detección de anomalías y predicción en entornos con alta complejidad y no linealidad. Esto permite aplicar distintos modelos de clasificación supervisada para predecir la ocurrencia de eventos sísmicos relevantes, construyendo así un indicador de peligro que puede ser evaluado objetivamente en función de su desempeño.

Distintos autores como Reyes et al. (2013), Araneda et al. (2023) y Cabrera et al. (2024) han demostrado la efectividad de algoritmos como redes neuronales y árboles de decisión en la predicción de eventos. Además, en el contexto de la minería de hundimiento, se han propuesto indicadores de peligro sísmico aplicando técnicas de ML sobre datos reales (Estay, 2014 y Estay & Pavez-Orrego, 2025).

2.2 Descripción y tratamiento de datos

El caso de estudio se basa en información sísmica registrada en una mina subterránea chilena durante el período comprendido entre enero y marzo del año 2016. Durante este intervalo, se registraron un total de 45.143 eventos sísmicos, los cuales fueron detectados a través del sistema de monitoreo instalado en la faena. Los eventos se distribuyen en un área de $7,1 \times 10^9 \text{ m}^3$ (2098 x 2147 x 1572 metros), cuya distribución de magnitudes va desde -2,3 a 1,8 (Figura 2.1). La

distribución en frecuencia de los datos se muestra en la Figura 2.2 y los parámetros y estadísticas de cada uno de ellos se presentan en la Tabla 2.2.

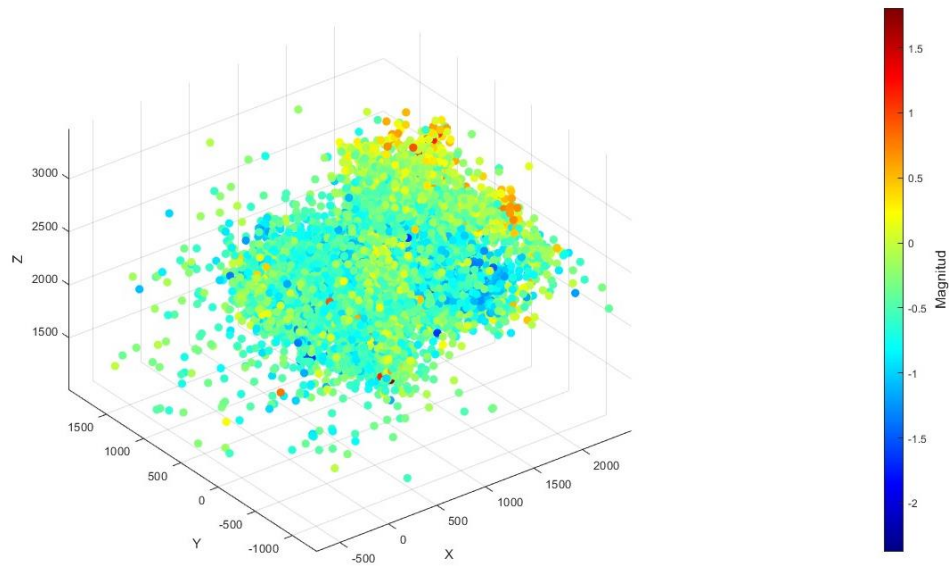


Figura 2.1 Vista isométrica de los eventos en base a su ubicación, los colores representan la magnitud del evento (azul las magnitudes más bajas y rojo las más altas).

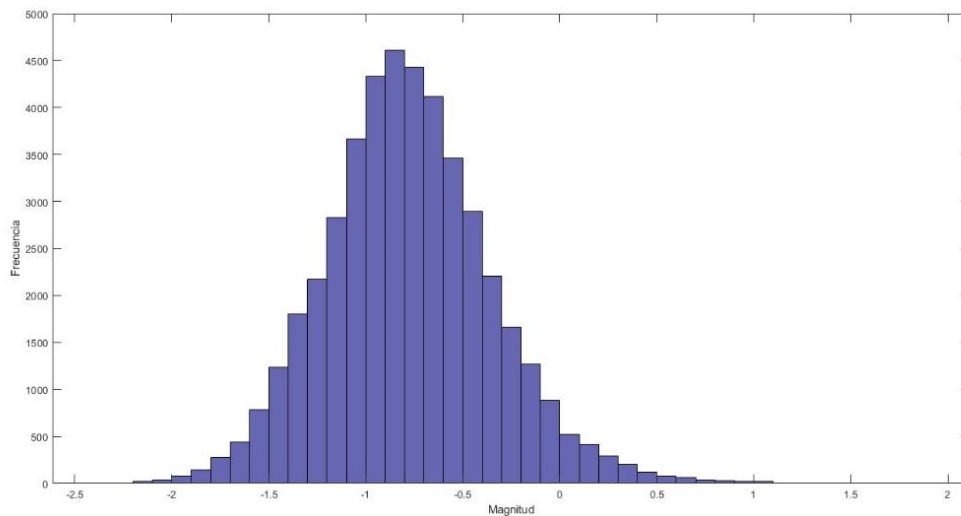


Figura 2.2 Histograma de las magnitudes antes de filtrar los eventos.

Tabla 2.2 Estadísticas básicas de parámetros a utilizar antes de filtrar eventos.

	Mínimo	Máximo	Media	Varianza	Unidad
Momento	$2,7 \times 10^5$	$5,1 \times 10^{11}$	$3,3 \times 10^8$	$2,0 \times 10^{19}$	Nm
Energía	0,0	$1,9 \times 10^6$	164,7	$1,5 \times 10^8$	J
Momento P	$5,9 \times 10^4$	$8,8 \times 10^{11}$	$5,0 \times 10^8$	$4,6 \times 10^{19}$	Nm
Energía P	0,0	$1,9 \times 10^5$	28,6	$2,6 \times 10^6$	J
Momento S	$1,3 \times 10^5$	$5,2 \times 10^{11}$	$2,5 \times 10^8$	$1,4 \times 10^{19}$	Nm
Energía S	0,0	$1,7 \times 10^6$	136,2	$1,2 \times 10^8$	J
Frecuencia de esquina	10,5	2520,0	126,3	$1,6 \times 10^4$	Hz
Caída de esfuerzo estática	4,7	$7,9 \times 10^6$	$2,7 \times 10^4$	$2,0 \times 10^{10}$	Pa
Caída de esfuerzo dinámica	82,4	$3,0 \times 10^7$	$6,1 \times 10^4$	$1,4 \times 10^{11}$	Pa
Desviación del momento	0,0	$3,4 \times 10^{11}$	$2,2 \times 10^8$	$1,0 \times 10^{19}$	Nm
Desviación de la energía	0,0	$9,5 \times 10^5$	103,8	$4,8 \times 10^7$	J
Frecuencia de esquina S	10,5	2520,0	126,3	$1,6 \times 10^4$	Hz
Magnitud	-2,4	1,8	-0,8	0,2	--

Para asegurar la confiabilidad de los datos, es necesario verificar que estos sean confiables, para ello, se consideran los siguientes filtros.

- Error de localización menor a 50 metros.
- Eventos detectados por al menos 3 geófonos.
- Filtro por magnitud de completitud (Wiemer & Wyss, 2000; Woessner & Wiemer, 2005).

2.3 Análisis de componentes principales (PCA)

El análisis de componentes principales es una técnica estadística utilizada para reducir la dimensionalidad de un conjunto de datos multivariado, manteniendo la mayor parte de su variabilidad. Su objetivo es transformar un conjunto de variables posiblemente correlacionadas en un nuevo conjunto de variables no correlacionadas, denominadas componentes principales. Estas se obtienen como combinaciones lineales de las variables originales que maximizan la varianza explicada, y se construyen a partir de los autovalores y autovectores de la matriz de covarianzas o de correlaciones del conjunto de datos (Maćkiewicz y Ratajczak, 1993). El primer componente captura la mayor proporción de la varianza total, el segundo la mayor parte de la varianza remanente, y así sucesivamente, de forma que los componentes están ordenados según su importancia.

Por ejemplo, si se analizan varios parámetros que caracterizan eventos sísmicos—como la energía liberada, la magnitud, el momento sísmico y la duración del evento—es probable que muchos de ellos estén correlacionados entre sí. Aplicando PCA, es posible combinar estos indicadores en un

número reducido de componentes que resuman la dinámica general del proceso sísmico, facilitando así la interpretación y el análisis de tendencias.

En esta oportunidad, el PCA se aplicó a un conjunto de indicadores sísmicos agrupados por ventanas temporales. El propósito fue reducir el número de variables manteniendo al menos un 95% de la varianza total, lo cual permitió disminuir la complejidad del modelo sin comprometer su capacidad descriptiva ni predictiva. Además, al eliminar redundancias y colinealidades entre variables, se facilita la visualización de los datos y la interpretación de patrones relevantes en la evolución de la actividad sísmica.

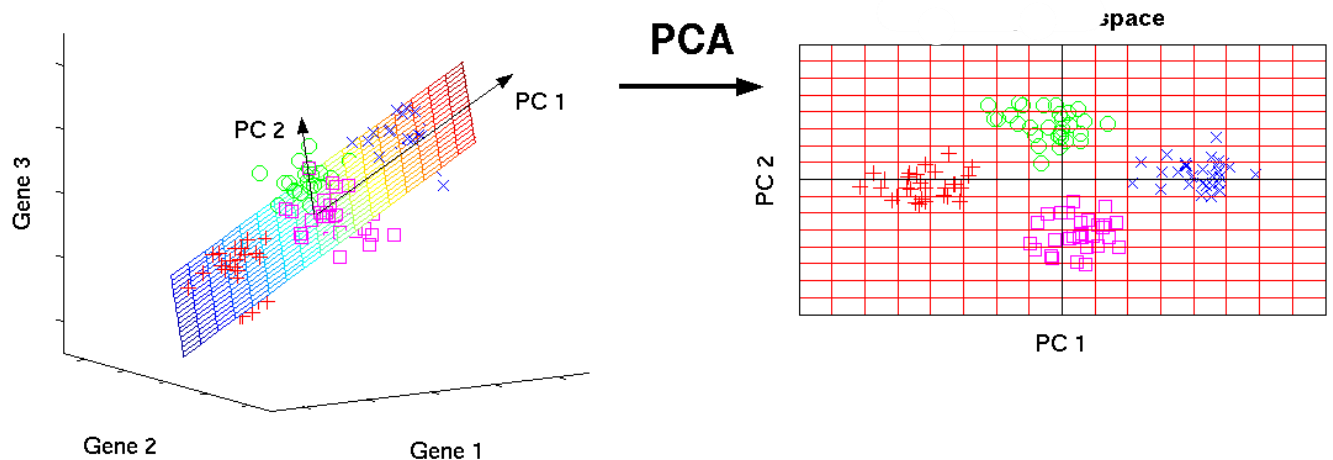


Figura 2.3 Ejemplo de PCA (Scholz, 2006). Inicialmente se tienen datos en 3 dimensiones (izquierda), al aplicar PCA se convierten en 2 componentes (derecha).

2.4 Matriz de contingencia

Una de las maneras de evaluar la calidad de los modelos de clasificación implementados, es a través de la matriz de contingencia (Fawcett, 2005). En ella se va cuantificando la capacidad del modelo de predecir o clasificar de manera correcta comparando las predicciones del modelo con los resultados reales observados. En la construcción de indicadores sísmicos, la matriz de contingencia permite evaluar el rendimiento al identificar si fue capaz de anticipar correctamente un evento de magnitud relevante.

La matriz de contingencia se compone de los siguientes escenarios:

- Verdaderos positivos (TP): Casos en los que el indicador predijo correctamente la ocurrencia de un evento sísmico significativo (aciertos).

- Falsos positivos (FP): Casos en que el indicador predijo un evento, pero no ocurrió ninguno (falsas alarmas).
- Verdaderos negativos (TN): Casos en que el modelo predijo correctamente que no ocurriría ningún evento.
- Falsos negativos (FN): Casos en que el modelo no anticipó un evento sísmico, pero este sí ocurrió (fallos).

Tabla 2.3. Matriz de contingencia (TP: Acierto positivo, FP: Falsa alarma, FN: Fallo, TN: Acierto negativo).

		Valor real	
		Ocorre	No ocurre
Valor predicho	Ocorre	TP	FP
	No ocurre	FN	TN

A partir de esta matriz se derivan métricas que permiten cuantificar la capacidad del modelo para predecir y categorizar los eventos de forma precisa dentro del horizonte de predicción. Entre las métricas de desempeño se encuentran la tasa de verdaderos positivos (TPR, ecuación 1), la tasa de falsos positivos (FPR, ecuación 2), la exactitud (ACC, ecuación 3) y la tasa de falsas alarmas (PSS, ecuación 4)

$$TPR = \frac{TP}{TP + FN} \quad (1)$$

$$FPR = \frac{FP}{FP + TN} \quad (2)$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

$$PSS = TPR - FPR \quad (4)$$

Estas métricas de desempeño se pueden graficar utilizando la curva ROC (características operativas del receptor). En esta curva se grafica el TPR vs el FPR. El área bajo la curva (AUC) se utiliza como medida cuantitativa del desempeño del modelo, donde un valor cercano a 1 indica alta capacidad predictiva, mientras que valores cercanos a 0,5 indican aleatoriedad equivalente al

lanzamiento de una moneda. En la Figura 2.4 se tiene de ejemplo una curva ROC, donde la curva A es la más precisa, la curva B representa aleatoriedad y la curva C es un mal indicador.

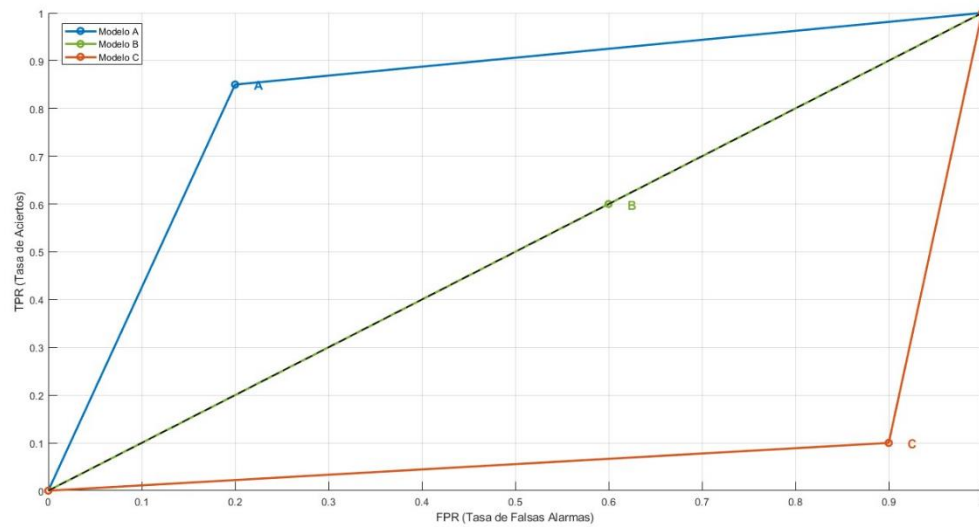


Figura 2.4 Ejemplo de curva ROC para 3 modelos distintos, modelo A representa alta precisión, modelo B es aleatorio y modelo C es un mal modelo.

3 Metodología

Para garantizar la calidad y confiabilidad del conjunto de datos, se aplicaron diversos criterios de filtrado al catálogo sísmico original. Se eliminaron todos los eventos con magnitud inferior a la magnitud de completitud ($M_c = -0,9$), determinada mediante el método de máxima curvatura propuesto por Wiemer & Wyss (2000). Este valor corresponde al bin con la frecuencia más alta dentro de la distribución frecuencia–magnitud (Figura 3.1). Adicionalmente, se descartaron los eventos detectados por menos de tres geófonos o que tuvieran un error de localización mayor a 50 metros, con el objetivo de asegurar una localización confiable.

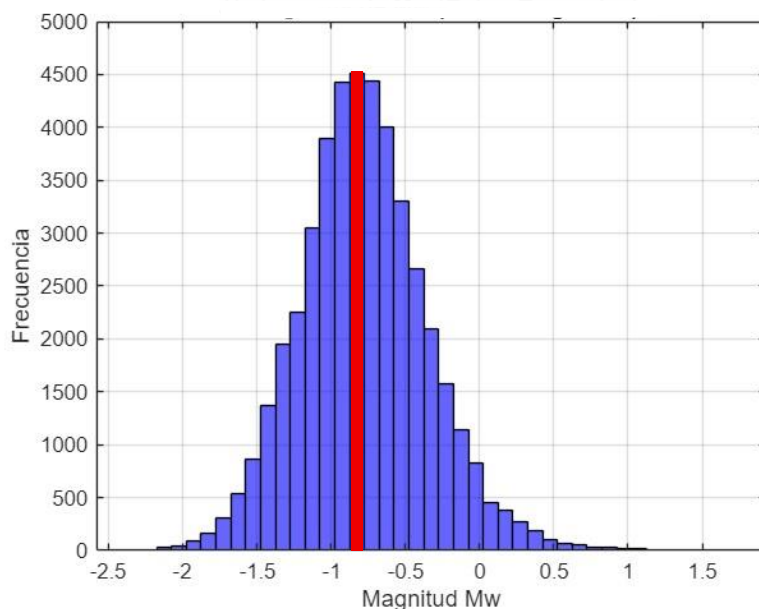


Figura 3.1 Elección de magnitud de completitud (-0,9), utilizando el método de máxima curvatura (Wiemer & Wyss, 2000; Woessner & Wiemer, 2005).

Posteriormente, se estandarizaron todas las variables a media cero y desviación estándar uno, permitiendo comparaciones homogéneas entre atributos de distinta magnitud y escala.

3.1 Agrupamiento temporal

Se implementó un enfoque de ventanas móviles para agrupar y caracterizar la evolución temporal de la actividad sísmica. Este enfoque permite construir grupos con la información del catálogo dentro de un intervalo temporal definido.

Además, para cada grupo es necesario evaluar si ocurre algún evento sísmico que pueda representar un riesgo para la operación minera. Para ello, se define una magnitud relevante, que actúa como

umbral a partir del cual un evento se considera potencialmente peligroso. Este umbral permite identificar si dentro de cada grupo ocurre al menos un evento sísmico que supere dicho valor.

Las variables utilizadas dentro de cada grupo corresponden a las descritas en la Sección 2.2. A estas se añadieron estadísticas descriptivas asociadas a la magnitud de los eventos, tales como:

- Magnitud máxima, mínima y promedio.
- Varianza de la magnitud.
- Medidas de forma de la distribución: curtosis y asimetría.

Para cada grupo, se calcularon los valores promedio de las variables consideradas. Es decir, cada parámetro se resume mediante un valor medio representativo, calculado a partir de todos los eventos contenidos en el grupo, independientemente de cuántos eventos haya en cada caso.

La Figura 3.2 muestra un ejemplo visual de cómo se crean los grupos. El agrupamiento se hizo considerando las siguientes ventanas temporales:

- **Ventana de tiempo (Δt_i):** Duración del agrupamiento. Se consideran 2 escenarios (24 y 48 horas).
- **Tiempo de búsqueda (Δt_f):** Intervalo posterior al terminar la ventana de tiempo, en el cual se evalúa si ocurre un evento sísmico con magnitud relevante. Se analizaron dos escenarios (8 y 12 horas).
- **Tiempo de avance (Δt_a):** Paso temporal con el que se desplaza la ventana respecto al inicio del grupo anterior. Se consideraron dos escenarios (10 y 20 minutos).

Para cada evento se definió un valor del tipo indicador (binario) en base a una magnitud relevante (M_w^*). Esta magnitud corresponde al umbral a partir del cual un evento se considera con potencial riesgo de generar daño. Se consideraron 3 escenarios: 0,7; 0,8; 0,9. Este criterio de magnitud relevante es el que se usará para evaluar el desempeño del indicador.

Hora	x	y	z	Parámetro	$M_w \geq M_w^*$
00:10	-200	40	100	...	0
01:23	400	0	100	...	0
02:20	0	25	110	...	1
03:50	30	50	120	...	0
04:00	25	30	100	...	0
05:57	70	20	130	...	1
08:40	80	50	120	...	1
10:00	60	70	100	...	0
12:20	80	40	120	...	0
15:00	100	20	80	...	1

Figura 3.2 Esquema del agrupamiento, Δt_i indica la duración de cada grupo, Δt_f el tiempo de búsqueda posterior, Δt_a el tiempo de avance para el siguiente grupo y $M_w > M_w^*$ la ocurrencia de un evento mayor a la magnitud relevante.

Esta configuración dio lugar a 24 combinaciones de escenarios, al combinar las distintas duraciones de agrupamiento, ventanas futuras, pasos de avance y umbrales de magnitudes. Esta estrategia permitió realizar un análisis robusto del desempeño de los modelos bajo diferentes condiciones operativas.

3.2 Reducción de dimensionalidad (PCA)

Inicialmente, cada grupo quedó representado por 18 variables. Dado el elevado número de variables y la posible colinealidad entre ellas, se aplicó la técnica de análisis de componentes principales (PCA) con el objetivo de reducir la dimensionalidad sin pérdida significativa de información.

Donde se obtiene que las variables más relevantes son:

- Curtosis
- Varianza
- Máxima magnitud
- Mínima magnitud

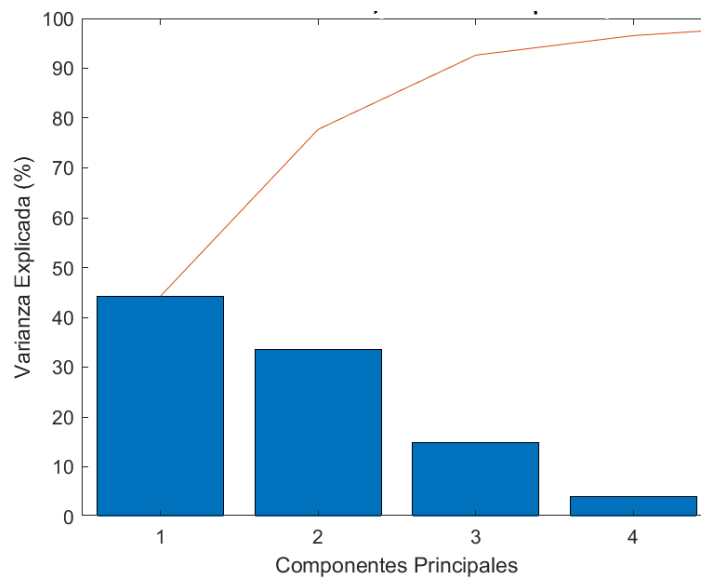


Figura 3.3 PCA obtenido. 4 componentes representan el 95% de la varianza explicada.

3.3 Modelos de machine learning

Se implementaron múltiples modelos de clasificación con el objetivo de evaluar su capacidad de categorizar adecuadamente la ocurrencia de eventos sísmicos de magnitud relevante en la ventana futura, considerando los parámetros definidos en el PCA. Para el entrenamiento de los modelos, se utilizó el 70% de los datos generados a partir del agrupamiento temporal, reservando el 30% restante para la etapa de validación. Estos datos de entrada corresponden a los obtenidos tras aplicar el PCA.

La selección de modelos incluyó las diversas familias de algoritmos descritas previamente en la sección 2.1, cubriendo desde enfoques simples hasta técnicas avanzadas de aprendizaje automático.

3.4 Evaluación del desempeño

El desempeño de cada modelo fue evaluado utilizando las métricas descritas en la sección 2.4. El análisis de estas métricas permitió comparar el rendimiento de los distintos modelos bajo condiciones operativas variadas, identificando aquellos con mayor capacidad de anticipar eventos sísmicos de alta magnitud.

4 Resultados

En esta sección se presentan los resultados obtenidos a partir de la aplicación de los modelos de clasificación supervisada para la predicción de eventos sísmicos de alta magnitud, utilizando los indicadores construidos mediante agrupamiento temporal de datos y reducción de dimensionalidad.

Para cada uno de los modelos que se presentarán a continuación, cada punto graficado corresponde a una de las 24 combinaciones posibles descritas con anterioridad.

4.1 Desempeño por modelos de machine learning

A continuación, se presenta el desempeño de los modelos de machine learning, de acuerdo con lo señalado en la sección 2.4. Además, las precisiones obtenidas a partir de las matrices de contingencia para cada modelo, considerando los distintos escenarios propuestos y las magnitudes relevantes, se incluyen en la sección Anexo 7.1.

4.1.1 Desempeño árboles de decisión

La Figura 4.1 muestra el desempeño de los tres modelos de árboles de decisión. Se observa que los árboles finos presentan los mejores resultados, seguidos por los árboles medianos, mientras que los árboles gruesos registran el menor desempeño.

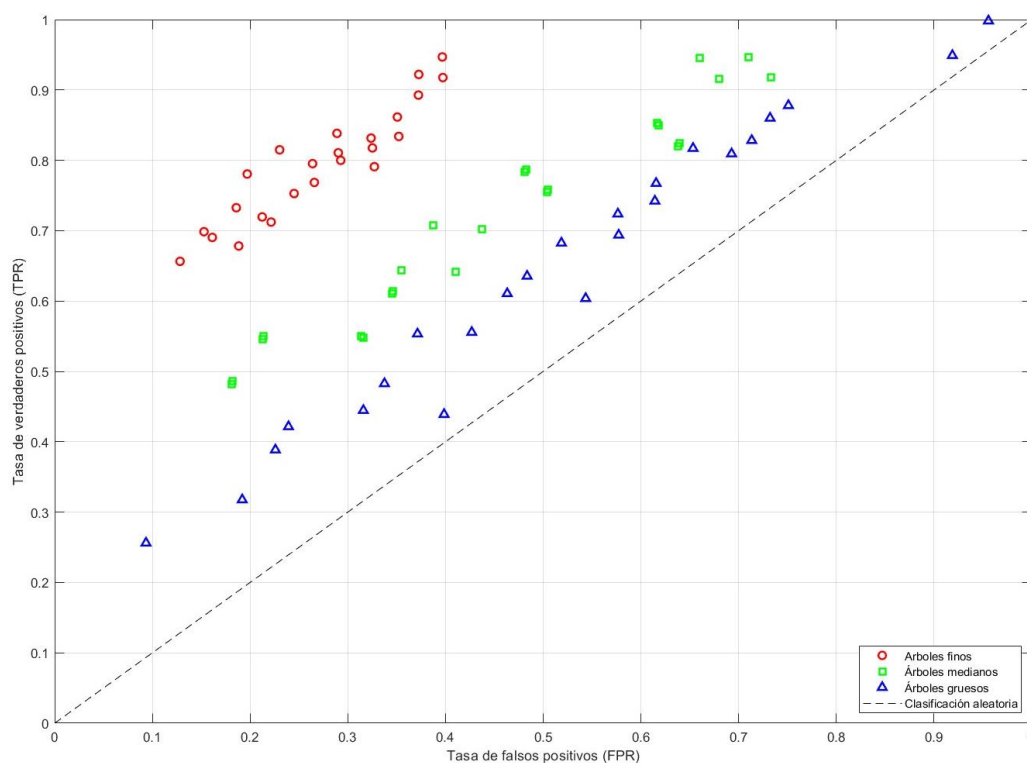


Figura 4.1 Gráfico ROC para indicador mediante árboles de decisión.

4.1.2 Desempeño modelos discriminantes

La Figura 4.2 presenta el desempeño de los dos modelos discriminantes. Ambos modelos muestran resultados similares, con un rendimiento cercano al 50%. No obstante, el discriminante cuadrático presenta un desempeño levemente superior.

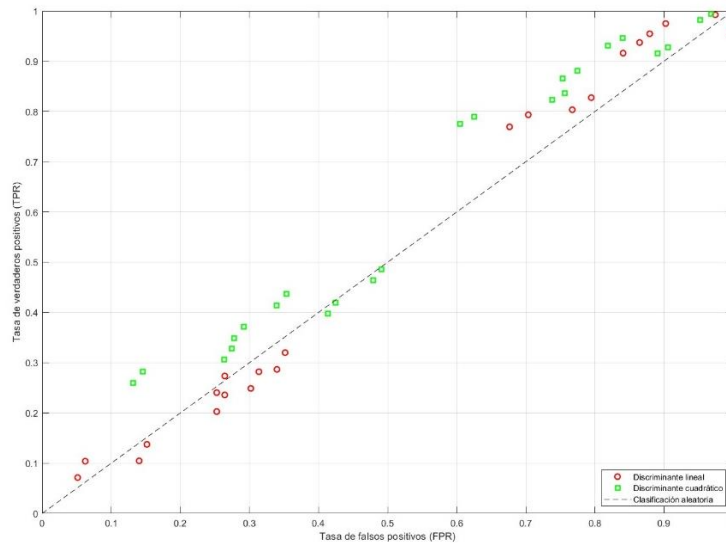


Figura 4.2 Gráfico ROC para indicador mediante modelos discriminantes.

4.1.3 Desempeño regresión logística

La Figura 4.3 presenta los resultados para el modelo de regresión logística. En este caso, se observa un rendimiento cercano al 50%, lo que indica que su capacidad predictiva es prácticamente aleatoria.

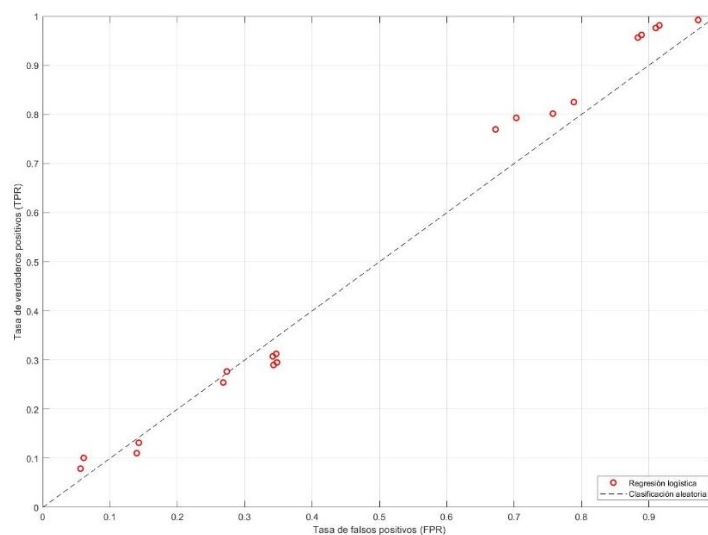


Figura 4.3 Gráfico ROC para indicador mediante regresión logística.

4.1.4 Desempeño de clasificadores bayesianos

La Figura 4.4 muestra el desempeño de los modelos bayesianos. Se aprecia que el modelo gaussiano obtiene resultados levemente superiores en comparación al modelo de kernel, cuyo rendimiento es cercano al 50%.

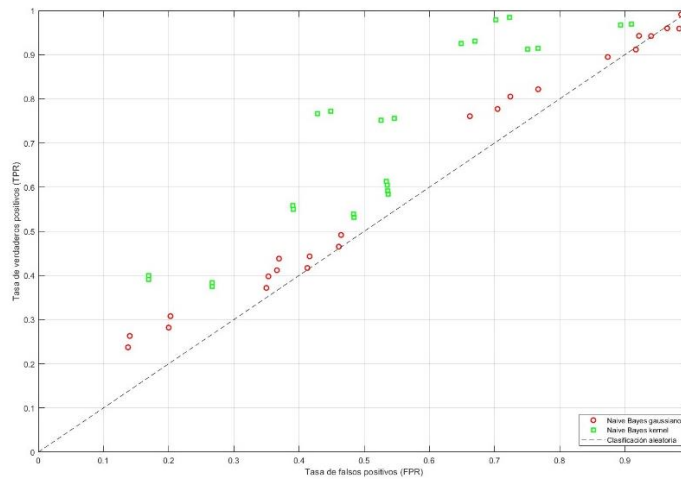


Figura 4.4 Gráfico ROC para indicador mediante clasificadores bayesianos.

4.1.5 Desempeño máquinas de vectores de soporte (support vector machine)

La Figura 4.5 presenta los resultados para los modelos de máquinas de vectores de soporte. Se observa que el modelo gaussiano fino obtiene el mejor desempeño, mientras que el resto de los modelos se comporta de forma aleatoria.

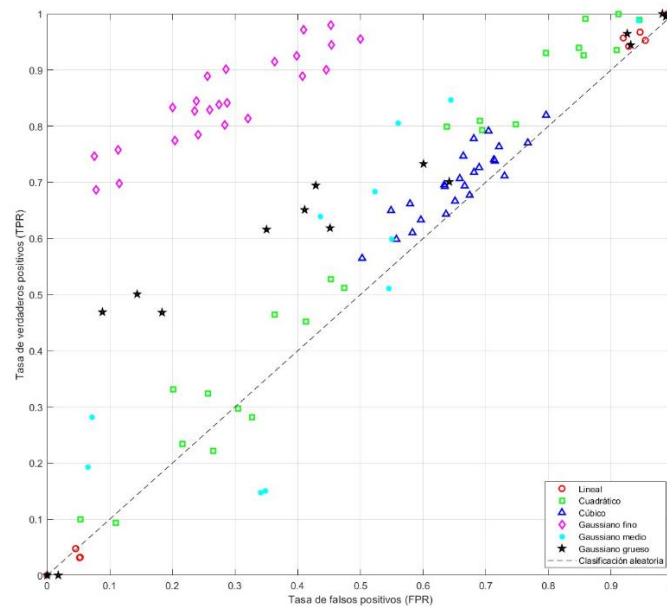


Figura 4.5 Gráfico ROC para indicador mediante SVM.

4.1.6 Desempeño vecinos más cercanos

La Figura 4.6 muestra el desempeño de los modelos kNN. El kNN fino entrega los mejores resultados, muy similares al KNN ponderado, seguidos por los modelos cúbico, medio y distancia del coseno.

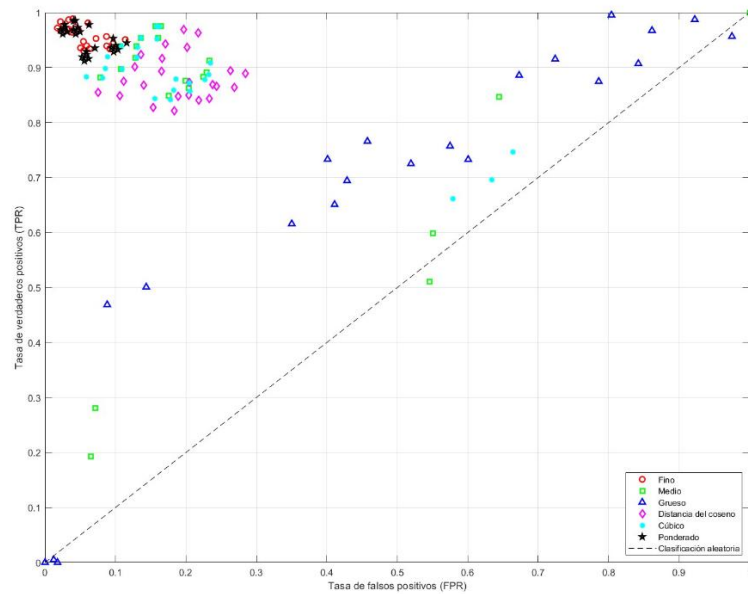


Figura 4.6 Gráfico ROC para indicador mediante kNN.

4.1.7 Desempeño métodos de ensamble

La Figura 4.7 presenta el desempeño de los modelos de ensamble. Se destaca que los árboles potenciados entregan los mejores resultados dentro de esta categoría.

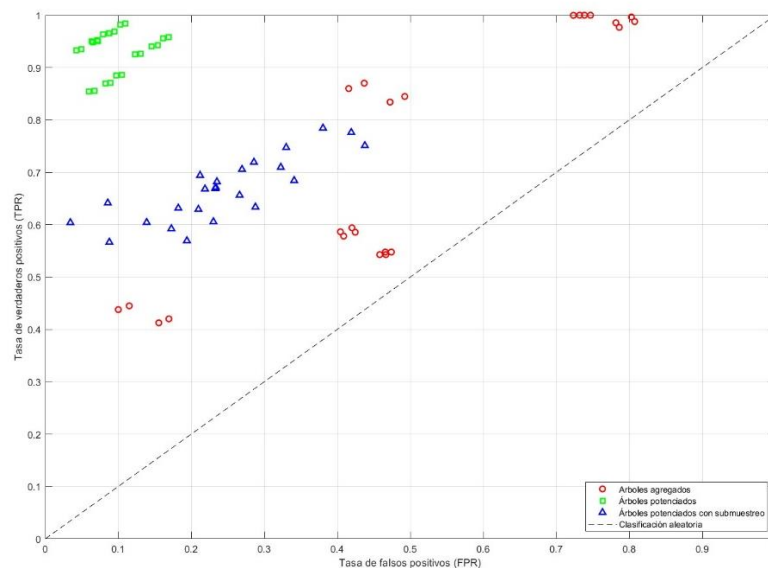


Figura 4.7 Gráfico ROC para indicador mediante modelos de ensamble.

4.1.8 Desempeño redes neuronales

La Figura 4.8 presenta el desempeño de los modelos de redes neuronales. Se aprecia que el mejor resultado corresponde al modelo de redes neuronales anchas, mientras que el modelo de redes neuronales angostas presenta el menor desempeño. El resto de los modelos muestra resultados similares entre sí.

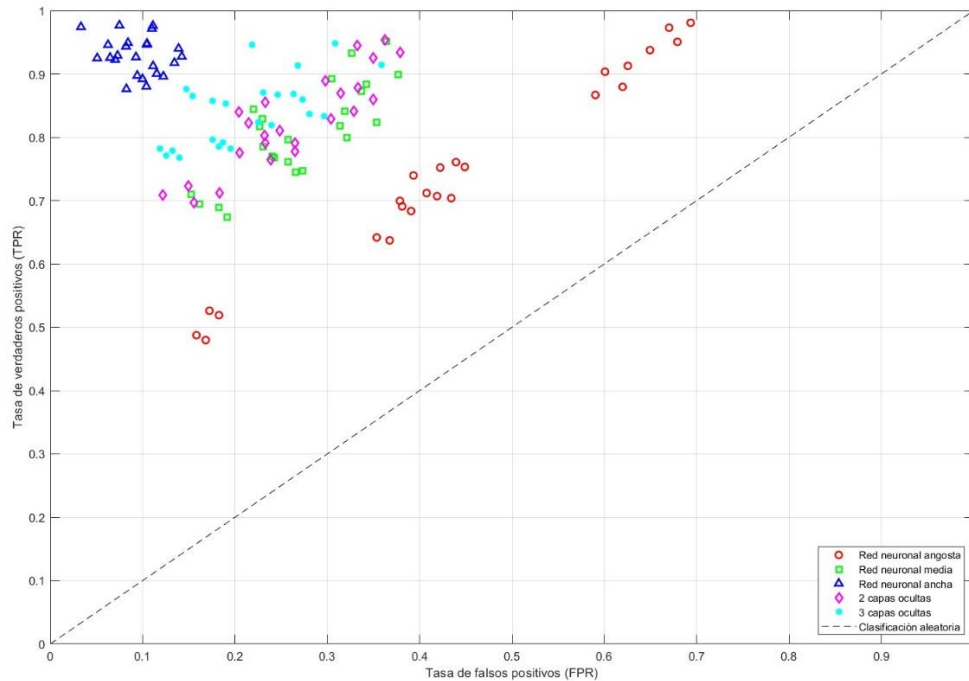


Figura 4.8 Gráfico ROC para indicador mediante redes neuronales.

4.2 Desempeño por magnitudes

La Figura 4.9 muestra el desempeño de los modelos considerando los distintos umbrales de magnitud. Se observa que los tres casos de estudio presentan una distribución similar, sin tendencias marcadas que permitan identificar diferencias significativas entre los umbrales evaluados.

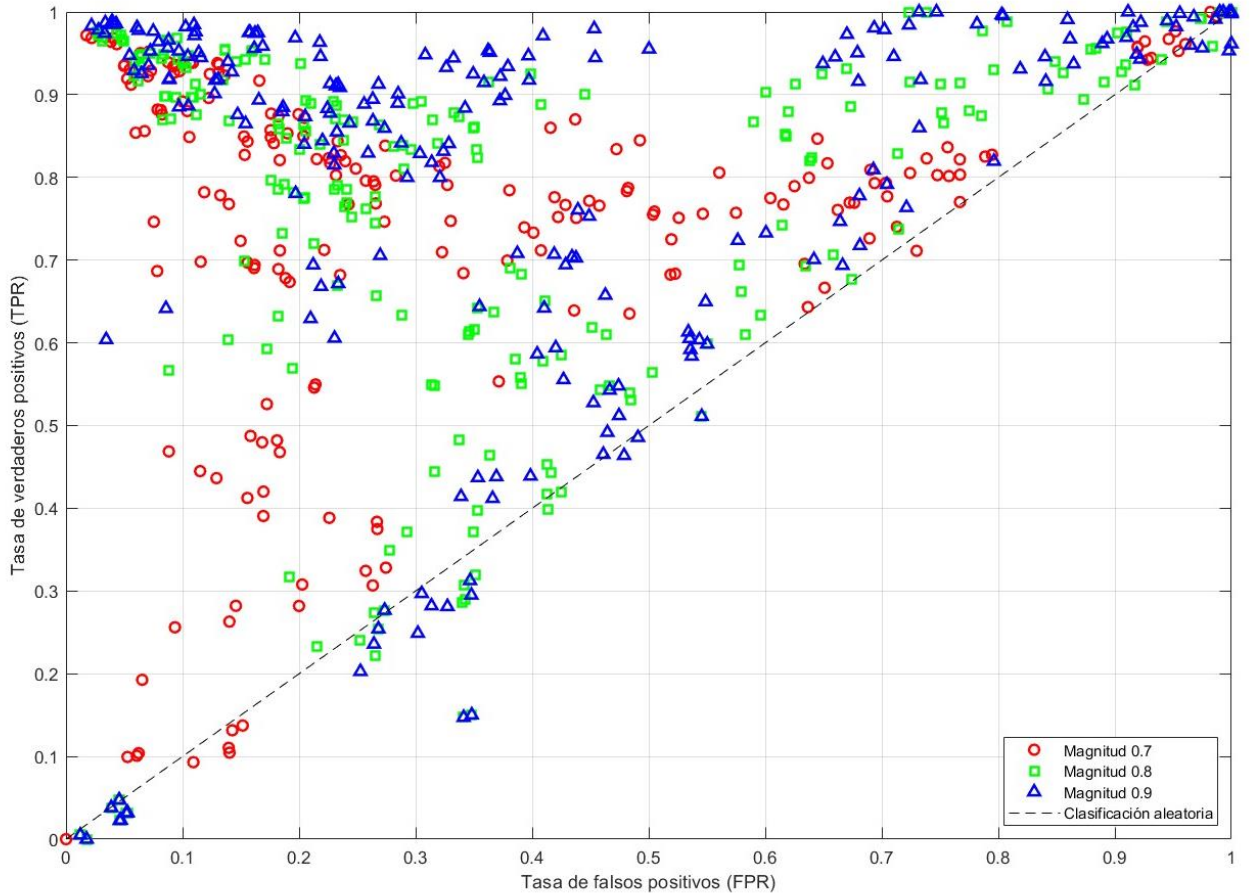


Figura 4.9 Gráfico ROC para indicador para distintas magnitudes relevantes.

A pesar de que el rendimiento general es comparable entre los umbrales, se concluye que la mejor alternativa corresponde al umbral $M_w^* > 0,7$, ya que representa una magnitud capaz de generar daños estructurales y, además, concentra un mayor número de eventos en el catálogo. Esto último favorece la construcción de modelos más robustos y con mejor capacidad de aprendizaje.

4.3 Desempeño por agrupamientos temporales

Para evaluar el efecto de los parámetros temporales en el desempeño de los modelos, se analizaron los escenarios por separado según ventana de tiempo, tiempo futuro y tiempo de avance.

4.3.1 Desempeño por ventana de tiempo

La Figura 4.10 presenta el desempeño de los modelos considerando ventanas de tiempo de 24 y 48 horas. Se observa que las ventanas de 24 horas tienden a ubicarse más arriba, pero desplazadas a la derecha en el gráfico ROC, mientras que las de 48 horas se posicionan más abajo y hacia la izquierda. Si bien no se evidencian diferencias notorias, los agrupamientos de 24 horas muestran, en general, un leve mejor desempeño.

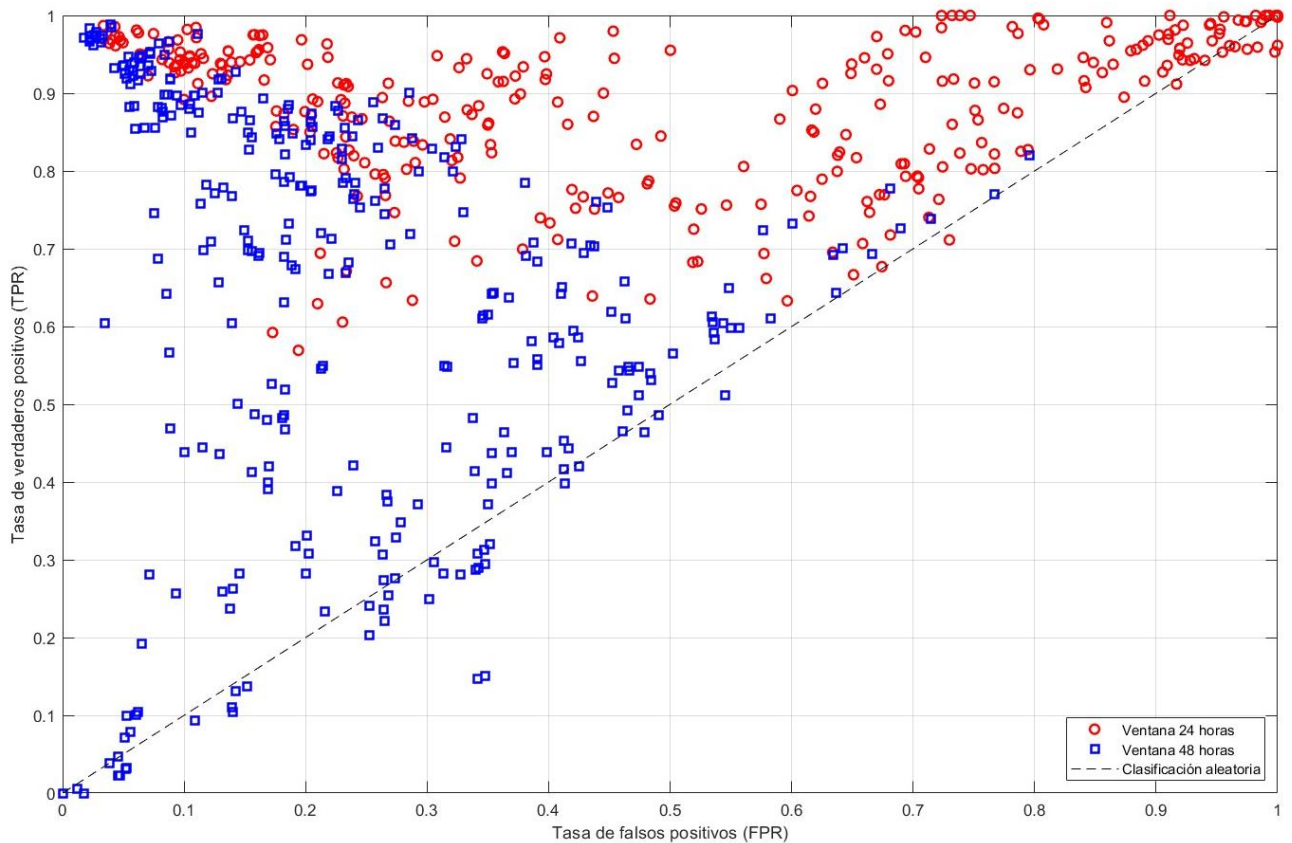


Figura 4.10 Gráfico ROC para agrupamiento de distintas ventanas de tiempo.

4.3.2 Desempeño por tiempos futuros

La Figura 4.11 muestra los resultados asociados al tiempo de predicción futura, considerando ventanas de 8 y 12 horas. Se aprecia un comportamiento similar al observado en las ventanas de agrupamiento, donde las predicciones a 8 horas se ubican más arriba pero más a la derecha, y las de 12 horas se concentran en la parte inferior e izquierda. No obstante, de manera general, los tiempos de predicción de 8 horas presentan mejores resultados.

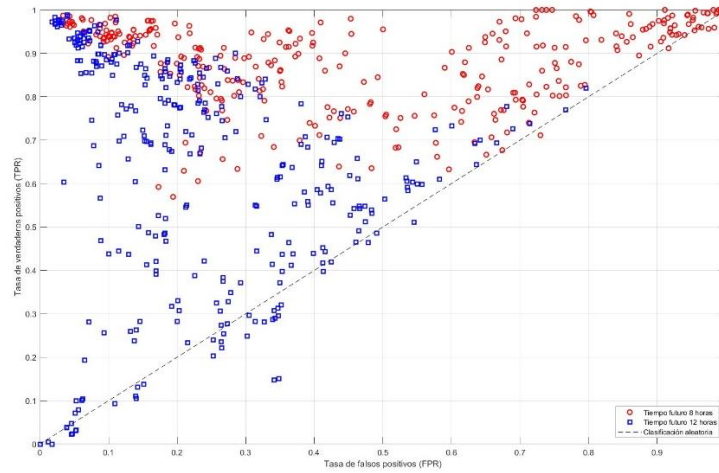


Figura 4.11 Gráfico ROC para distintos tiempos de búsqueda futuros.

4.3.3 Desempeño por tiempos de avance

En cuanto al tiempo de avance, se observa en Figura 4.12 como los datos se distribuyen de manera similar entre los escenarios evaluados. Sin embargo, se observa una mayor concentración de casos con altos valores de AUC en aquellos que utilizan un paso de avance de 10 minutos, lo que se atribuye a una mayor resolución temporal en el seguimiento de los indicadores.

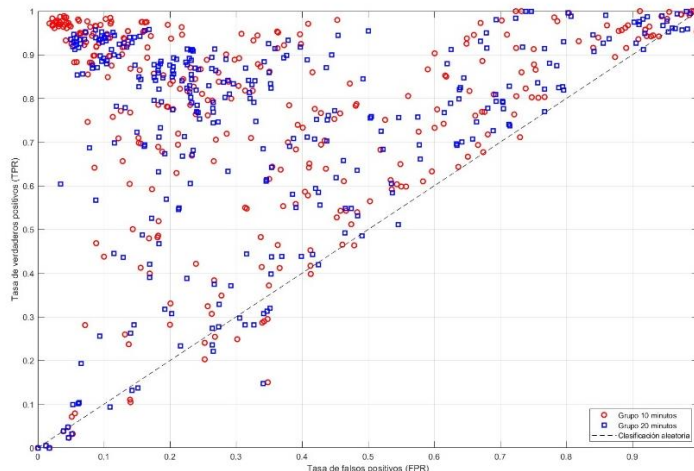


Figura 4.12 Gráfico ROC para distintos tiempos de avance del catálogo.

4.4 Resultados finales

Luego de analizar todos los escenarios por separado, se obtiene que el escenario con desempeños más altos es el de ventanas de tiempo de 24 horas, tiempos de búsqueda de 8 horas, avances en el catálogo de 10 minutos y una magnitud relevante de 0,7. El diagrama ROC para este escenario es presentado en la Figura 4.13.

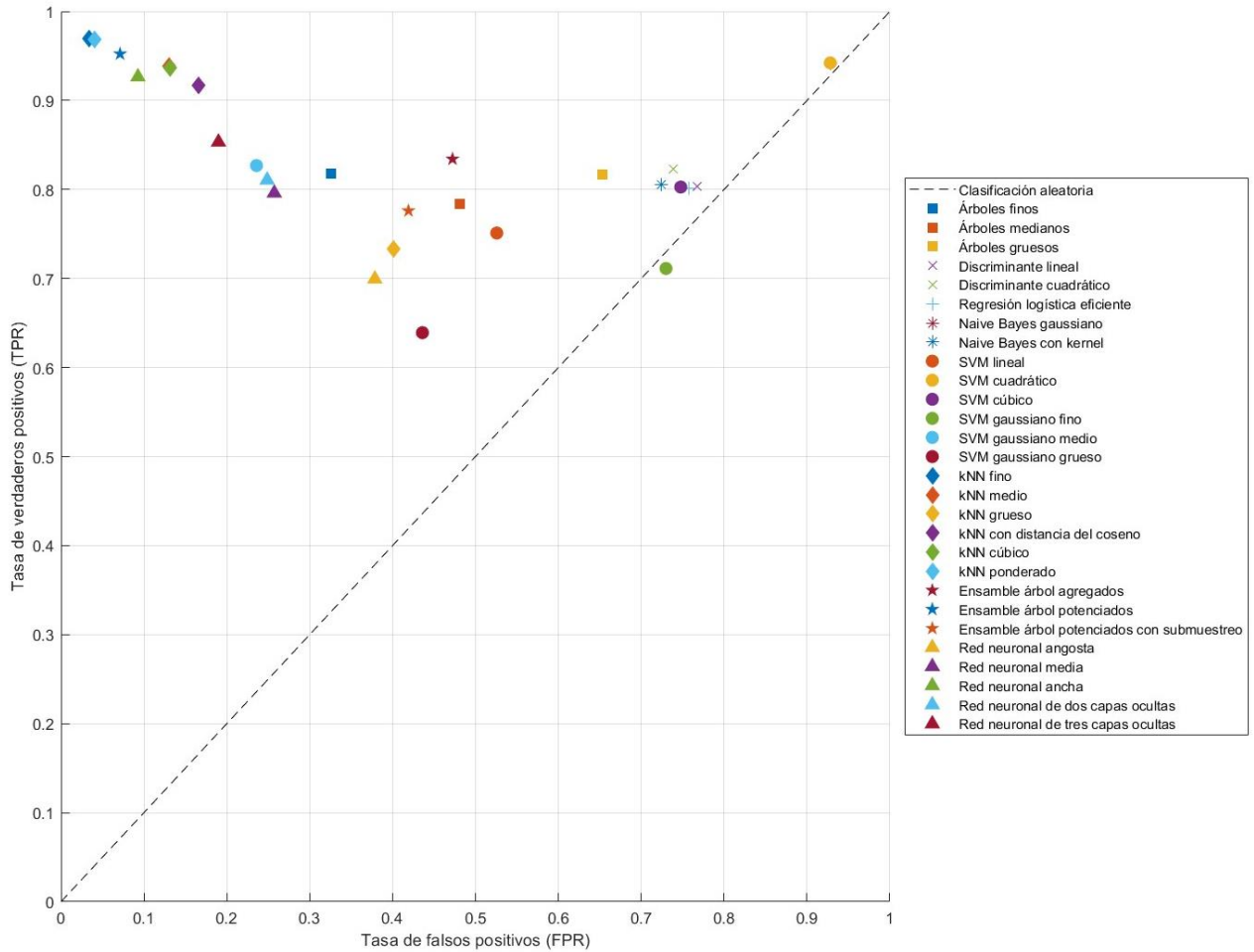


Figura 4.13 Diagrama ROC para el escenario con el mejor desempeño (Δt : 24 horas, Δt_f : 8 horas, Δt_a : 10 minutos).

De esta figura, es posible apreciar como el mejor modelo corresponde a kNN fino, seguido de kNN ponderado, además también presentan muy buenos resultados los modelos de árboles potenciados, redes neuronales anchas y otros modelos de kNN.

5 Conclusiones

En este trabajo se desarrolló y evaluó una metodología para caracterizar el comportamiento de parámetros sísmicos, utilizando ventanas móviles temporales y técnicas de reducción de dimensionalidad, aplicada a datos reales de una mina subterránea chilena que opera mediante el método de hundimiento por bloques.

La metodología demostró ser eficaz para describir la evolución temporal de la actividad sísmica y construir vectores de características que permiten categorizar condiciones potencialmente asociadas a eventos de mayor magnitud.

Los resultados evidencian que el escenario con mejor desempeño corresponde al uso de ventanas de 24 horas, tiempos de búsqueda de 8 horas, avances de 10 minutos y un umbral de magnitud $M_w^* > 0,7$, configuración que logra una adecuada resolución temporal y un equilibrio entre la cantidad de eventos y la relevancia operacional de los mismos.

Entre los 28 modelos de machine learning evaluados, el algoritmo *K-Nearest Neighbors* (kNN) en su configuración fina presentó el mejor desempeño general, con valores de AUC superiores al 95% en los escenarios óptimos, lo que refleja su capacidad para identificar patrones en los datos y discriminar entre situaciones de mayor o menor peligro sísmico.

Modelos como los árboles potenciados y las redes neuronales anchas también destacaron, lo que sugiere que enfoques tanto basados en proximidad como en aprendizaje no lineal tienen potencial para estas aplicaciones. En contraste, se observó que modelos como los discriminantes o la regresión logística presentan desempeños cercanos al azar bajo esta configuración de datos.

La reducción de dimensionalidad mediante análisis de componentes principales (PCA) permitió simplificar la representación de los datos, manteniendo el 95% de la varianza explicada con solo cuatro variables, lo que favorece la interpretabilidad y la eficiencia de los modelos.

A pesar de las limitaciones del estudio, como la ausencia de variables geomecánicas, geográficas o de contexto operacional, los resultados obtenidos demuestran el potencial de los indicadores basados en machine learning para la categorización temprana de situaciones sísmicas que podrían implicar mayor riesgo, lo que aporta a la construcción de herramientas de apoyo en la gestión del peligro sísmico en minería subterránea.

En conclusión, la metodología propuesta representa un avance hacia la incorporación de herramientas de análisis de datos avanzadas en la gestión de la sismicidad inducida, ofreciendo oportunidades para fortalecer la seguridad y la continuidad operacional en faenas subterráneas de gran escala.

6 Referencias

- Araneda, M., Valencia, K., Sepúlveda, A., & Muñoz, F. (2023). Forecasting mine seismicity with hybrid LSTM and attention-based deep learning models. *Journal of Rock Mechanics and Geotechnical Engineering*, 15(10), 2217–2232. <https://doi.org/10.1016/j.jrmge.2023.03.009>
- Cabrera, L., A. Ardid, I. Melchor, S. Ruiz, B. Symmes Lopetegui, J. Carlos Báez, F. Delgado, P. Martinez-Yáñez, D. Dempsey, and S. Cronin (2024). Eruption Forecasting Model for Copahue Volcano (Southern Andes) Using Seismic Data and Machine Learning: A Joint Interpretation with Geodetic Data (GNSS and InSAR), *Seismol. Res. Lett.* XX,1–16, doi: 10.1785/0220240022
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27
- Derras, B., Bard, P. Y., & Cotton, F. (2014). Towards fully data driven ground-motion prediction models for Europe. *Bulletin of Earthquake Engineering*, 12(3), 495–516. <https://doi.org/10.1007/s10518-013-9481-0>
- Estay, R. 2014. Metodología para la evaluación del desempeño de indicadores sísmicos en sismicidad inducida por la minería, Msc. Thesis, Universidad de Chile.
- Estay, R. & Pavez-Orrego, C. (2025). An induced seismicity indicator using accumulated microearthquakes' frictional's energy. *Mining* 5(2), 27. <https://doi.org/10.3390/mining5020027>
- Faradonbeh, R. S., Shakeri, J., Ghaderi, Z., Mikula, P. A., Jang, H., & Taheri, A. (2024). Harnessing machine learning for seismic event discrimination in deep underground mining: A case study from Western Australia. In P. Andrieux & D. Cumming-Potvin (Eds.), *Deep Mining 2024: Proceedings of the 10th International Conference on Deep and High Stress Mining* (pp. 831–852). Australian Centre for Geomechanics. https://doi.org/10.36487/ACG_repo/2465_52
- Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874.
- Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1), 119–139.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

- Gonzalez, F. J., Vallejos, J. A., Rojas, E., & Landeros, P. (2021). Evaluation of the seismic rock mass response to mining and the impact of preconditioning using an epidemic-type aftershock model. *International Journal of Rock Mechanics and Mining Sciences*, 141, 104975. <https://doi.org/10.1016/j.ijrmms.2021.104975>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.
- Małkiewicz, A., & Ratajczak, W. (1992). *Principal Components Analysis (PCA)*.
- Ortega, V. 2014. *Análisis de sismicidad inducida por tronaduras de incorporación de área en mina el teniente, Codelco, Chile, Universidad de Chile*.
- Scholz, M. (2006). *Approaches to Analyse and Interpret Biological Profile Data* (Doctoral dissertation). Universität Potsdam
- SONAMI. 2025. ¿Cómo superar el estancamiento y potenciar la industria?
- Vallejos Massa, J. A., & McKinnon, S. D. (2013). Logistic regression and neural network classification of seismic records. *International Journal of Rock Mechanics and Mining Sciences*, 62, 86–95. <https://doi.org/10.1016/j.ijrmms.2013.04.005>

7 Anexo

7.1 Precisiones para los distintos escenarios

Anexo 1 Precisiones para Δt : 24 horas, Δt_f : 8 horas, Δt_a : 10 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	75,20%	78,68%	86,34%
Arboles medianos	66,18%	68,53%	79,84%
Arboles gruesos	60,08%	65,66%	76,64%
Discriminante lineal	54,11%	65,22%	75,53%
Discriminante cuadrático	56,49%	63,19%	75,50%
Regresión logística	54,46%	65,25%	75,53%
Naive Bayes gaussiano	54,21%	64,81%	75,70%
Naive Bayes con kernel	56,19%	63,17%	75,36%
SVM lineal	62,39%	67,95%	75,82%
SVM cuadrático	54,21%	64,81%	75,70%
SVM cúbico	54,97%	63,96%	75,70%
SVM gaussiano fino	50,85%	55,34%	64,57%
SVM gaussiano medio	79,81%	81,73%	87,87%
SVM gaussiano grueso	60,46%	65,99%	75,73%
kNN fino	96,79%	96,76%	98,20%
kNN medio	90,69%	91,37%	94,10%
kNN grueso	67,15%	68,92%	76,59%
kNN con distancia del coseno	87,89%	88,84%	91,93%
kNN cúbico	90,54%	91,30%	94,10%
kNN ponderado	96,45%	96,59%	97,86%
Árboles agregados	69,33%	70,83%	79,91%
Árboles potenciados	94,15%	94,62%	96,16%
Árboles potenciados con submuestreo	68,63%	68,36%	71,69%
Red neuronal angosta	66,36%	70,59%	79,49%
Red neuronal media	77,15%	79,90%	86,99%
Red neuronal ancha	91,77%	92,94%	96,41%
Red neuronal de dos capas ocultas	78,34%	80,41%	87,70%
Red neuronal de tres capas ocultas	83,34%	82,72%	90,62%

Anexo 2 Precisiones para Δt : 24 horas, Δt_f : 8 horas, Δt_a : 20 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	73,65%	77,06%	84,57%
Arboles medianos	63,58%	65,88%	76,97%
Arboles gruesos	59,10%	64,70%	75,71%
Discriminante lineal	54,09%	65,23%	75,57%
Discriminante cuadrático	56,34%	62,91%	74,97%
Regresión logística	54,24%	65,06%	75,37%
Naive Bayes gaussiano	53,67%	64,06%	74,73%
Naive Bayes con kernel	55,06%	62,06%	74,29%
SVM lineal	61,67%	67,12%	74,83%
SVM cuadrático	53,49%	64,12%	75,04%
SVM cúbico	56,88%	65,90%	77,67%
SVM gaussiano fino	53,15%	57,55%	66,60%
SVM gaussiano medio	76,29%	78,17%	84,19%
SVM gaussiano grueso	58,85%	64,40%	74,17%
kNN fino	92,69%	92,66%	94,07%
kNN medio	84,14%	84,81%	87,48%
kNN grueso	61,25%	62,98%	70,50%
kNN con distancia del coseno	80,84%	81,79%	84,89%
kNN cúbico	83,66%	84,40%	87,15%
kNN ponderado	92,21%	92,35%	93,59%
Árboles agregados	68,95%	70,45%	79,56%
Árboles potenciados	90,02%	90,48%	91,99%
Árboles potenciados con submuestreo	66,40%	66,14%	69,40%
Red neuronal angosta	66,04%	70,28%	79,21%
Red neuronal media	73,74%	76,50%	83,61%
Red neuronal ancha	88,75%	89,92%	93,40%
Red neuronal de dos capas ocultas	76,50%	78,58%	85,89%
Red neuronal de tres capas ocultas	79,23%	78,61%	86,53%

Anexo 3 Precisiones para Δt : 24 horas, Δt_f : 12 horas, Δt_a : 10 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	78,61%	81,95%	89,31%
Arboles medianos	68,59%	70,85%	81,96%
Arboles gruesos	62,58%	68,19%	78,99%
Discriminante lineal	59,99%	70,91%	81,04%
Discriminante cuadrático	62,63%	69,06%	80,89%
Regresión logística	60,01%	70,62%	80,72%
Naive Bayes gaussiano	60,25%	70,43%	80,89%
Naive Bayes con kernel	61,39%	68,41%	80,39%
SVM lineal	65,92%	71,26%	78,82%
SVM cuadrático	60,25%	70,67%	81,62%
SVM cúbico	61,29%	70,33%	82,14%
SVM gaussiano fino	50,45%	54,76%	63,83%
SVM gaussiano medio	83,54%	85,38%	91,42%
SVM gaussiano grueso	67,12%	72,68%	82,48%
kNN fino	97,33%	97,30%	98,72%
kNN medio	92,03%	92,70%	95,38%
kNN grueso	71,50%	73,24%	80,78%
kNN con distancia del coseno	89,68%	90,64%	93,74%
kNN cúbico	91,81%	92,54%	95,29%
kNN ponderado	96,94%	97,07%	98,32%
Árboles agregados	71,62%	73,09%	82,23%
Árboles potenciados	94,78%	95,24%	96,76%
Árboles potenciados con submuestreo	71,62%	71,35%	74,63%
Red neuronal angosta	69,88%	74,04%	82,99%
Red neuronal media	79,29%	82,06%	89,19%
Red neuronal ancha	93,95%	95,10%	98,59%
Red neuronal de dos capas ocultas	81,06%	83,09%	90,43%
Red neuronal de tres capas ocultas	83,35%	82,74%	90,69%

Anexo 4 Precisiones para Δt : 24 horas, Δt_f : 12 horas, Δt_a : 20 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	77,99 %	81,34%	88,72%
Arboles medianos	69,22 %	71,48%	82,63%
Arboles gruesos	64,86 %	70,49%	81,32%
Discriminante lineal	60,70 %	71,65%	81,82%
Discriminante cuadrático	62,73 %	69,18%	81,04%
Regresión logística	60,67 %	71,31%	81,44%
Naive Bayes gaussiano	60,34 %	70,55%	81,04%
Naive Bayes con kernel	62,31 %	69,35%	81,37%
SVM lineal	65,62 %	70,98%	78,56%
SVM cuadrático	60,34 %	70,79%	81,78%
SVM cúbico	61,10 %	70,17%	82,02%
SVM gaussiano fino	47,54 %	51,87%	60,96%
SVM gaussiano medio	81,05 %	82,90%	88,95%
SVM gaussiano grueso	64,04 %	69,62%	79,45%
kNN fino	93,94 %	93,91%	95,33%
kNN medio	84,81 %	85,48%	88,17%
kNN grueso	67,83 %	69,58%	77,14%
kNN con distancia del coseno	83,90 %	84,86%	87,98%
kNN cúbico	84,38 %	85,11%	87,87%
kNN ponderado	93,15 %	93,28%	94,54%
Árboles agregados	71,04 %	72,52%	81,68%
Árboles potenciados	90,60 %	91,06%	91,06%
Árboles potenciados con submuestreo	73,20 %	72,93%	72,93%
Red neuronal angosta	70,80 %	74,97%	74,97%
Red neuronal media	78,11 %	80,88%	80,88%
Red neuronal ancha	90,12 %	91,27%	91,27%
Red neuronal de dos capas ocultas	79,99 %	82,03%	82,03%
Red neuronal de tres capas ocultas	84,20 %	83,59%	83,59%

Anexo 5 Precisiones para Δt : 48 horas, Δt_f : 8 horas, Δt_a : 10 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	76,77%	80,18%	87,69%
Arboles medianos	66,14%	68,50%	79,58%
Arboles gruesos	58,86%	64,46%	75,22%
Discriminante lineal	56,14%	67,03%	77,13%
Discriminante cuadrático	59,55%	66,27%	78,33%
Regresión logística	56,30%	66,87%	76,95%
Naive Bayes gaussiano	54,84%	65,47%	76,14%
Naive Bayes con kernel	56,34%	63,18%	75,13%
SVM lineal	67,56%	73,14%	80,85%
SVM cuadrático	54,80%	65,43%	76,10%
SVM cúbico	57,61%	66,42%	77,93%
SVM gaussiano fino	51,84%	56,24%	65,29%
SVM gaussiano medio	78,60%	80,53%	86,54%
SVM gaussiano grueso	63,48%	68,90%	78,44%
kNN fino	96,27%	96,24%	97,65%
kNN medio	91,78%	92,45%	95,12%
kNN grueso	66,16%	67,94%	75,45%
kNN con distancia del coseno	89,56%	90,49%	93,52%
kNN cúbico	91,64%	92,38%	95,13%
kNN ponderado	95,89%	96,03%	97,28%
Árboles agregados	73,13%	74,60%	83,50%
Árboles potenciados	94,38%	94,84%	96,35%
Árboles potenciados con submuestreo	69,45%	69,18%	72,44%
Red neuronal angosta	67,76%	72,00%	80,72%
Red neuronal media	79,66%	82,42%	89,37%
Red neuronal ancha	92,58%	93,75%	97,15%
Red neuronal de dos capas ocultas	80,48%	82,56%	89,70%
Red neuronal de tres capas ocultas	84,19%	83,57%	91,31%

Anexo 6 Precisiones para Δt : 48 horas, Δt_f : 8 horas, Δt_a : 20 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	75,22%	78,57%	86,10%
Arboles medianos	63,54%	65,90%	76,77%
Arboles gruesos	57,90%	63,39%	73,93%
Discriminante lineal	56,13%	66,80%	76,70%
Discriminante cuadrático	59,34%	65,92%	77,75%
Regresión logística	56,09%	66,45%	76,32%
Naive Bayes gaussiano	54,29%	64,95%	75,41%
Naive Bayes con kernel	55,21%	61,91%	73,62%
SVM lineal	66,86%	72,32%	80,06%
SVM cuadrático	54,07%	64,73%	75,44%
SVM cúbico	59,47%	68,10%	79,38%
SVM gaussiano fino	54,14%	58,46%	67,32%
SVM gaussiano medio	75,10%	77,04%	83,07%
SVM gaussiano grueso	61,81%	67,25%	76,82%
kNN fino	92,18%	92,15%	93,57%
kNN medio	85,23%	85,90%	88,52%
kNN grueso	60,28%	62,06%	69,60%
kNN con distancia del coseno	82,48%	83,39%	86,36%
kNN cúbico	84,76%	85,49%	88,25%
kNN ponderado	91,66%	91,80%	93,05%
Árboles agregados	72,67%	74,15%	82,87%
Árboles potenciados	90,25%	90,71%	92,19%
Árboles potenciados con submuestreo	67,22%	66,96%	70,23%
Red neuronal angosta	67,44%	71,70%	80,25%
Red neuronal media	76,26%	78,96%	85,77%
Red neuronal ancha	89,56%	90,71%	94,05%
Red neuronal de dos capas ocultas	78,65%	80,68%	87,68%
Red neuronal de tres capas ocultas	80,06%	79,45%	87,22%

Anexo 7 Precisiones para Δt : 48 horas, Δt_f : 12 horas, Δt_a : 10 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	80.15%	83.42%	90.98%
Arboles medianos	68.55%	70.92%	81.82%
Arboles gruesos	61.38%	66.89%	77.46%
Discriminante lineal	61.98%	72.44%	82.37%
Discriminante cuadrático	65.63%	72.23%	84.09%
Regresión logística	61.81%	71.97%	81.64%
Naive Bayes gaussiano	60.87%	71.56%	81.81%
Naive Bayes con kernel	61.54%	68.11%	79.58%
SVM lineal	71.11%	76.59%	84.35%
SVM cuadrático	60.83%	71.52%	82.26%
SVM cúbico	63.88%	72.34%	83.39%
SVM gaussiano fino	51.44%	55.67%	64.56%
SVM gaussiano medio	82.33%	84.26%	90.18%
SVM gaussiano grueso	70.08%	75.41%	84.79%
kNN fino	96.82%	96.79%	98.18%
kNN medio	93.12%	93.78%	96.35%
kNN grueso	70.53%	72.32%	79.70%
kNN con distancia del coseno	91.36%	92.25%	95.16%
kNN cúbico	92.91%	93.63%	96.39%
kNN ponderado	96.39%	96.53%	97.78%
Árboles agregados	75.34%	76.79%	85.33%
Árboles potenciados	95.01%	95.46%	96.91%
Árboles potenciados con submuestreo	72.42%	72.16%	75.44%
Red neuronal angosta	71.25%	75.52%	84.09%
Red neuronal media	81.75%	84.46%	91.13%
Red neuronal ancha	94.74%	95.90%	99.24%
Red neuronal de dos capas ocultas	83.21%	85.25%	92.11%
Red neuronal de tres capas ocultas	84.18%	83.57%	91.36%

Anexo 8 Precisiones para Δt : 48 horas, Δt_f : 12 horas, Δt_a : 20 minutos

Modelo/Magnitud	0,7	0,8	0,9
Arboles finos	79.57%	82.85%	90.26%
Arboles medianos	69.18%	71.50%	82.18%
Arboles gruesos	63.66%	69.05%	79.42%
Discriminante lineal	62.74%	72.99%	82.72%
Discriminante cuadrático	65.67%	72.14%	84.04%
Regresión logística	62.48%	72.43%	81.91%
Naive Bayes gaussiano	60.95%	71.43%	81.47%
Naive Bayes con kernel	62.46%	68.90%	80.41%
SVM lineal	70.82%	76.32%	83.92%
SVM cuadrático	60.91%	71.63%	82.16%
SVM cúbico	63.69%	71.99%	83.07%
SVM gaussiano fino	48.51%	52.65%	61.57%
SVM gaussiano medio	79.89%	81.83%	87.76%
SVM gaussiano grueso	66.94%	72.28%	81.69%
kNN fino	93.44%	93.41%	94.80%
kNN medio	85.88%	86.52%	89.04%
kNN grueso	66.88%	68.67%	75.91%
kNN con distancia del coseno	85.54%	86.42%	89.27%
kNN cúbico	85.49%	86.20%	88.97%
kNN ponderado	92.61%	92.75%	94.01%
Árboles agregados	74.69%	76.11%	84.48%
Árboles potenciados	90.83%	91.28%	92.73%
Árboles potenciados con submuestreo	74.02%	73.76%	76.98%
Red neuronal angosta	72.18%	76.36%	84.76%
Red neuronal media	80.64%	83.29%	89.83%
Red neuronal ancha	90.92%	92.05%	95.32%
Red neuronal de dos capas ocultas	82.14%	84.14%	90.87%
Red neuronal de tres capas ocultas	85.02%	84.40%	92.04%