

INFORME DE TF

1ACC0057 - Machine Learning

Docente: Luis Carnaval Sanchez

Grupo: 1

Integrantes:

u202218044 - Mayhua Hinostroza, José Antonio

u202216120 - Manchay Paredes, Lucero Salome

u20191e650 - Arquiñigo Jacinto, Ibrahim Imanol Jordi

1. Situación de contexto real

En los últimos años, el mercado de criptomonedas ha experimentado una evolución significativa, tanto en volumen de inversión como en la diversidad de activos disponibles. Si bien Bitcoin y Ethereum siguen liderando como los principales criptoactivos por capitalización, existe un creciente interés por parte de inversionistas en explorar criptomonedas de baja capitalización, las cuales, aunque implican un mayor riesgo, también ofrecen un alto potencial de rentabilidad.

En abril de 2024 se llevó a cabo el cuarto halving de Bitcoin, un evento programado que reduce a la mitad la recompensa por minar nuevos bloques. Históricamente, este tipo de eventos ha sido un catalizador de ciclos alcistas en el mercado cripto, ya que disminuye la oferta de nuevos bitcoins en circulación, generando presión alcista sobre los precios (Binance Academy, 2024). Esto ha llevado a diversas firmas de inversión a anticiparse y buscar nuevas oportunidades en activos alternativos.

Una de estas firmas es Perú C-Inversiones, una administradora de capitales privados peruanos, la cual actualmente ofrece un portafolio compuesto por Bitcoin y Ethereum. Sin embargo, ante el nuevo ciclo post-halving, busca innovar mediante un producto financiero basado en criptomonedas emergentes alineadas con cuatro narrativas actuales de inversión: inteligencia artificial, videojuegos, activos del mundo real (RWA) y memecoins. Estas categorías han ganado relevancia en el ecosistema cripto por su vínculo con casos de uso reales, adopción tecnológica y fenómenos sociales virales (Investopedia, 2025).

Dada la alta volatilidad y el dinamismo de este tipo de activos, se requiere una solución basada en Machine Learning que analice grandes volúmenes de datos del mercado, identifique patrones y recomiende aquellas criptomonedas con mayor probabilidad de valorización futura. De este modo, el presente proyecto busca

responder a esa necesidad concreta mediante un modelo predictivo que apoye decisiones de inversión estratégicas.

2. Adquisición y preprocesamiento de los datos

Origen de los datos

Para el proceso de adquisición de datos, se utilizaron las funcionalidades de la API pública de CryptoCompare, la cual permite extraer información histórica diaria de diversas criptomonedas. Esta información incluye precios de apertura, cierre, máximos, mínimos y volumen de transacciones expresado en dólares estadounidenses. Además, se incorporó manualmente una clasificación por narrativa (Inteligencia Artificial, Memes, Videojuegos y RWA), la cual permite segmentar los activos según su tendencia temática en el mercado.

 Tabla 1

 Características extraídas mediante esta API

Características	Descripción
symbol	Código de la criptomoneda (ej. RNDR, SHIB)
narrativa	Categoría asignada: IA, memes, videojuegos o RWA
time	Fecha en formato UNIX (segundos desde 1970)
date	Fecha transformada en formato legible (yyyy-mm-dd)
open	Precio de apertura de la criptomoneda (a las 00:00 horas del día)
high	Precio más alto alcanzado durante el día
low	Precio más bajo durante el día

close	Precio de cierre de la criptomoneda (a las 23:59 horas del día)
volume	Volumen total operado en dólares
target	Etiqueta binaria: 1 si sube ≥15 % en los próximos 30 días, 0 si no

Nota.Elaboración propia

En la parte del preprocesamiento, se enriqueció el conjunto de datos original con nuevas variables que permiten abordar la problemática de predicción de revalorización. En particular, se creó la variable objetivo target, la cual identifica si una criptomoneda experimenta un aumento igual o superior al 15 % en los 30 días posteriores a una fecha determinada. Esta transformación convierte el problema en una tarea de clasificación binaria. Asimismo, se aseguró el correcto formato temporal de las fechas y se organizaron los registros por símbolo y secuencia cronológica.

Tabla 2Variables derivadas durante el preprocesamiento de datos

Variable	Descripción
target	Variable binaria que toma el valor 1 si el precio de cierre de una criptomoneda aumenta ≥15 % en los 30 días posteriores a una fecha dada; en caso contrario, toma el valor 0
date	Fecha convertida a formato legible (yyyy-mm-dd) a partir del valor UNIX original
narrativa	Etiqueta categórica que agrupa a cada criptomoneda según su tendencia temática: Inteligencia Artificial, Memes, Videojuegos o RWA.

Nota. Elaboración propia. Estas variables fueron construidas con base en los datos históricos extraídos desde la API de CryptoCompare.

Link del repositorio:

https://github.com/YuhiTTo/TF Mayhua Manchay Arquinigo.git

Análisis Exploratorio de los Datos (EDA)

El Análisis Exploratorio de Datos (EDA) es un paso fundamental en el proceso de Machine Learning, ya que permite comprender la estructura de los datos, identificar patrones, detectar posibles inconsistencias y preparar el conjunto de datos para el modelado predictivo. A continuación, se detalla el procedimiento seguido en el presente trabajo:

Carga e inspección inicial de los datos

Para este estudio, los datos fueron recolectados a través de la API de CryptoCompare, obteniendo información histórica diaria de 8 criptomonedas pertenecientes a diferentes narrativas: Inteligencia Artificial, Memes, Videojuegos y RWA. El dataset final utilizado contiene los siguientes campos:

- symbol: Símbolo identificador de la criptomoneda.
- narrativa: Categoría o narrativa de la criptomoneda.
- date: Fecha de la observación (formato yyyy-mm-dd).
- open, high, low, close: Precios de apertura, máximo, mínimo y cierre diario (en USD).
- volume: Volumen de transacciones en USD.
- target: Variable binaria que indica si el precio de la criptomoneda se valorizó al menos un 15% en los 30 días posteriores.

El dataset final contiene aproximadamente 6,500 registros, organizados cronológicamente para cada criptomoneda. En la primera inspección se confirmó que

todas las variables están correctamente tipadas y no se observaron problemas graves de inconsistencia.

Análisis de valores faltantes

Se realizó un análisis exhaustivo de valores faltantes e inconsistencias en las principales variables numéricas. El resultado evidenció que:

- No existen registros nulos en las columnas de precios (open, high, low, close)
 ni en el volumen de transacciones.
- No se encontraron valores infinitos ni errores numéricos en las variables críticas.

Por lo tanto, el conjunto de datos no requirió procesos adicionales de imputación o eliminación de registros, lo que garantiza la integridad y confiabilidad de los análisis posteriores.

Creación de variables adicionales

Con el objetivo de enriquecer el análisis y facilitar la detección de patrones relevantes para el modelado, se generaron nuevas variables derivadas, específicamente:

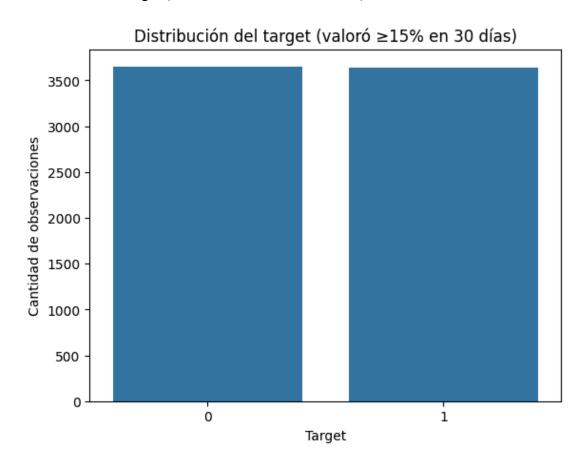
- return_pct: Retorno diario porcentual de la criptomoneda, calculado como la variación porcentual del precio de cierre respecto al día anterior.
- volatility: Volatilidad diaria, medida como la diferencia entre el precio máximo y el precio mínimo de cada jornada.
- ma_7: Media móvil simple de 7 días aplicada al precio de cierre, que permite identificar tendencias de corto plazo y suavizar las fluctuaciones diarias.

Estas variables son esenciales para entender la dinámica del mercado y serán utilizadas en etapas posteriores para la construcción de modelos predictivos.

Visualización y exploración de los datos

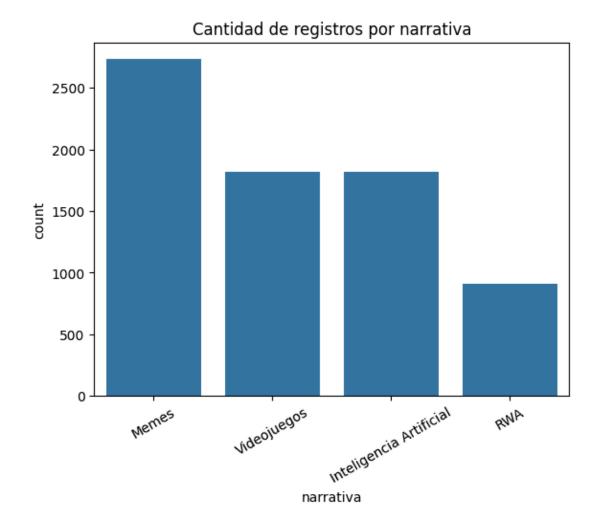
Se generaron diversas visualizaciones para explorar de manera gráfica las características principales del dataset:

Figura 1Distribución del target (valorización ≥15% en 30 días)



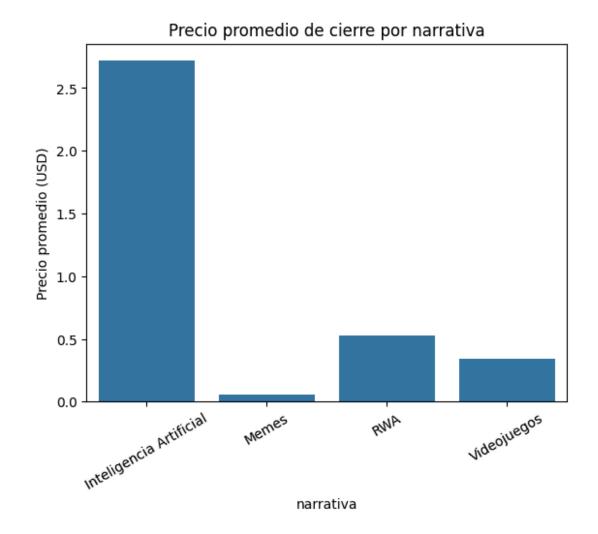
Nota. Gráfico de barras que muestra la proporción de registros donde el precio de la criptomoneda logró una valoración significativa. Se observa un leve desbalance entre las clases, típico en problemas de predicción de eventos de alta volatilidad en mercados financieros. Elaboración propia

Figura 2Cantidad de registros por narrativa



Nota. Distribución de registros según la narrativa a la que pertenece cada token. Se evidencia una mayor cantidad de datos en las narrativas de Memes e Inteligencia Artificial, lo cual debe considerarse para evitar sesgos en el análisis. Elaboración propia

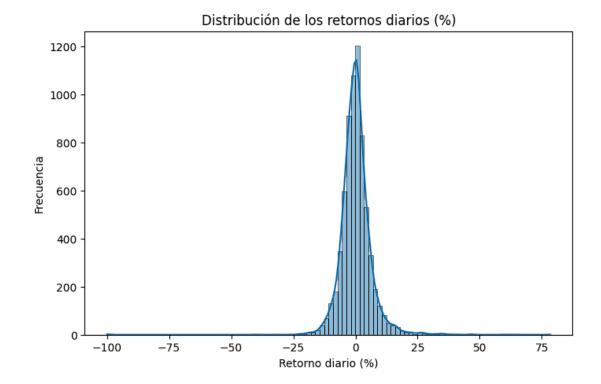
Figura 3Precio promedio de cierre por narrativa



Nota. Comparativa del precio promedio de cierre en USD para cada narrativa. Se observa que los tokens de la narrativa RWA y algunos de Inteligencia Artificial tienden a tener precios más elevados que los tokens de tipo Meme o Videojuegos. Elaboración propia

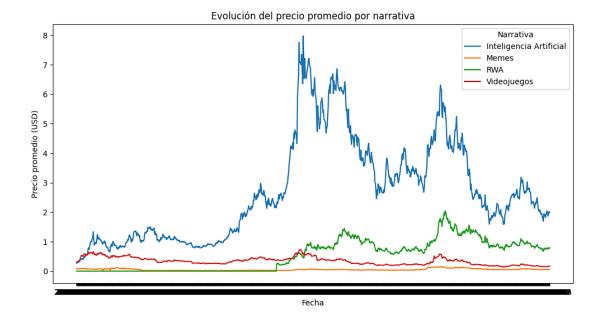
Figura 4

Distribución de los retornos diarios (%) de las criptomonedas



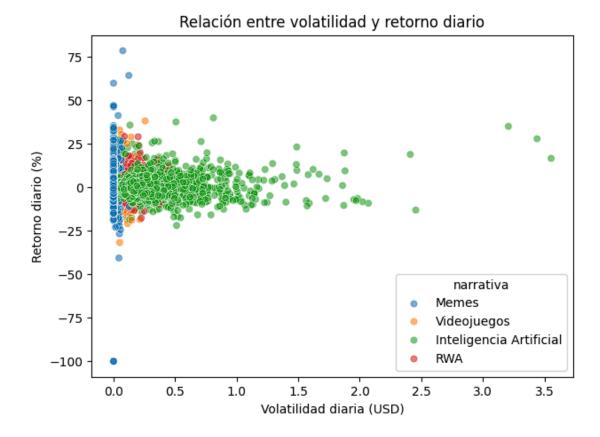
Nota. Histograma con la distribución de los retornos porcentuales diarios, mostrando una concentración alrededor de valores cercanos a 0%, pero con colas más pesadas que lo esperado bajo un supuesto de normalidad, lo cual es característico de mercados altamente volátiles como el de las criptomonedas. Elaboración propia

Figura 5Evolución del precio promedio por narrativa en el tiempo



Nota. Gráfico de líneas que presenta la evolución temporal de los precios promedio por narrativa, permitiendo observar tendencias generales y periodos de alta volatilidad diferenciados por categorías. Elaboración propia

Figura 6Relación entre volatilidad diaria y retorno diario



Nota. Gráfico de dispersión que revela una relación débil entre la volatilidad diaria y el retorno, aunque se identifican algunos patrones diferenciados según la narrativa del token. Elaboración propia

Figura 7

Mapa de calor - Correlación entre variables numéricas



Nota. Mapa de calor que ilustra la matriz de correlación entre las variables numéricas del dataset. Elaboración propia.

3. Propuesta

Para abordar la problemática relacionada con la identificación de criptomonedas emergentes con alto potencial de valorización en el corto plazo, se adoptará la metodología CRISP-DM (Cross Industry Standard Process for Data Mining) como marco estructurado para el desarrollo del proyecto. Esta metodología, ampliamente utilizada en proyectos de ciencia de datos, permite organizar de forma sistemática las fases necesarias para transformar los datos disponibles en conocimiento útil para la toma de decisiones. A continuación, se describen las seis fases que componen esta metodología:

3.1. Comprensión del negocio

Esta fase se enfoca en comprender los objetivos del proyecto desde una perspectiva estratégica y de valor para el usuario. En el presente caso, el objetivo principal es identificar criptomonedas emergentes con alto potencial de revalorización en el corto plazo (30 días), con base en sus características históricas y su narrativa de mercado (como inteligencia artificial, videojuegos, activos del mundo real o memecoins). La problemática se plantea como una tarea de clasificación binaria que permita apoyar la toma de decisiones de inversión en un entorno financiero altamente volátil.

3.2. Comprensión de los datos

En esta etapa se recopila y revisa la información disponible para familiarizarse con su estructura, calidad y relevancia. Se utilizaron datos históricos extraídos desde la API pública de CryptoCompare, que provee información diaria de diversas criptomonedas. Las variables recolectadas incluyen precios de

apertura, cierre, máximos y mínimos, volumen de transacciones y fechas. Asimismo, se añadió manualmente una variable categórica llamada "narrativa" que agrupa a las criptomonedas según su tendencia temática. Durante esta fase se realiza también un análisis exploratorio preliminar para detectar patrones, valores atípicos y posibles inconsistencias.

3.3. Preparación de los datos

Esta fase implica el procesamiento y transformación del conjunto de datos para adecuarlo al análisis predictivo. Incluye tareas como la conversión de fechas, la limpieza de registros incompletos, la normalización de variables y la asignación de categorías narrativas. Adicionalmente, se construye la variable objetivo (target), la cual indica si una criptomoneda experimentó un incremento igual o superior al 15 % en los 30 días posteriores a una fecha determinada. Este proceso convierte el problema en una tarea de clasificación binaria. El resultado de esta fase es un dataset estructurado y listo para ser utilizado en el modelado.

3.4. Modelado

Para abordar la problemática planteada por Perú C-Inversiones, se requiere la implementación de un modelo de Machine Learning capaz de identificar criptomonedas emergentes que presentan alta probabilidad de revalorización en el corto plazo. Específicamente, se busca predecir si un activo tendrá un incremento igual o superior al 15% en los siguientes 30 días, considerando tanto sus características históricas como su narrativa temática en el mercado. Dado que el problema se plantea como una tarea de clasificación binaria supervisada, se seleccionaron dos modelos robustos y complementarios para

ser comparados, considerando su idoneidad técnica, su aplicabilidad en entornos financieros y los algoritmos trabajados durante el presente proyecto:

1. Random Forest Classifier

El modelo de Random Forest es un algoritmo basado en ensambles de árboles de decisión, ampliamente utilizado por su robustez y su capacidad para manejar tanto variables numéricas como categóricas. Entre sus principales ventajas destacan:

- Reducción del sobreajuste mediante la combinación de múltiples árboles.
- Manejo eficiente de conjuntos de datos con relaciones no lineales y alta dimensionalidad.
- Interpretabilidad, al permitir conocer la importancia relativa de cada variable en las predicciones.

Este modelo es particularmente adecuado para contextos financieros como el presente, donde los datos presentan cierta complejidad y se requiere un balance entre precisión y capacidad de generalización.

2. Support Vector Machine (SVM)

El algoritmo de Support Vector Machine es una técnica de clasificación binaria que busca identificar el hiperplano óptimo que separa las clases. Utiliza los llamados "vectores de soporte" para definir esta frontera de decisión. Las razones para su selección son:

- Alto rendimiento en problemas de clasificación binaria, incluso con conjuntos de datos medianos.
- Capacidad de capturar relaciones complejas mediante el uso de kernels no lineales, como el kernel radial (RBF).

 Buen desempeño en espacios de alta dimensionalidad, como los que pueden derivarse de variables financieras enriquecidas.

Previa a la implementación del SVM, se realizarán los procesos necesarios de escalado y normalización de las variables, dado que este modelo es sensible a las diferencias de magnitud entre características.

3.5. Evaluación

Una vez entrenados los modelos, se procede a su evaluación crítica para determinar si cumplen con los objetivos definidos en la fase de comprensión del negocio. Se utilizarán métricas de evaluación como la matriz de confusión, precisión (accuracy), sensibilidad (recall) y la medida F1 para valorar el desempeño predictivo. Además, se analizará la importancia de las variables predictoras y se compararán los resultados obtenidos por cada algoritmo.

3.6. Despliegue

Aunque no es el foco principal de este proyecto académico, se contempla como fase futura el despliegue del modelo seleccionado en un entorno interactivo, como un dashboard o una aplicación web, que permita a los usuarios ingresar datos recientes y recibir recomendaciones automatizadas de inversión. Este paso garantiza la aplicabilidad práctica del conocimiento generado.

• Cronograma de trabajo del proyecto (basado en CRISP-DM)

Semana	Fase del proyecto	Tareas principales
14	Comprensión del negocio	Definición del problema, objetivos, alcance y

		formulación del enfoque predictivo
14	Comprensión de los datos	Recolección desde la API de CryptoCompare, revisión de variables y análisis inicial
14	Preparación de los datos	Limpieza, transformación, asignación de narrativas y creación de variable target
14	Análisis exploratorio (EDA)	Visualización de distribuciones, correlaciones, balance de clases y narrativa
14	Modelado	Entrenamiento de modelos (Random Forest y Support Vector Machine (SVM))
14	Evaluación	Evaluación de métricas, análisis de importancia de variables, comparación de modelos
14	Redacción del informe final	Consolidación de resultados, interpretación y presentación de hallazgos
14	(Opcional) Despliegue	Diseño de una interfaz demo o propuesta de dashboard predictivo

4. Ingeniería de Características

Luego del análisis exploratorio, se procedió a diseñar y seleccionar un conjunto final de características que permita mejorar la capacidad predictiva del modelo de clasificación binaria propuesto. Las características se construyeron a partir de los datos históricos extraídos mediante la API pública de CryptoCompare y se complementaron con variables derivadas diseñadas durante el preprocesamiento.

En esta etapa se priorizó el uso de variables numéricas calculadas por cada criptomoneda a lo largo del periodo de análisis (2023–2025), considerando la narrativa asignada como información contextual. El objetivo fue capturar aspectos como la variabilidad del precio, la magnitud del volumen de operaciones y la dinámica de crecimiento, que pudieran estar asociadas al comportamiento futuro del activo.

Tabla 3Características utilizadas para la clasificación binaria

Características	Descripción	
symbol	Identificador único de la criptomoneda	
narrativa	Categoría temática de la criptomoneda (IA, Memes, Videojuegos, RWA)	
volatility	Medida de volatilidad diaria: diferencia entre el valor máximo y mínimo del día	
return_pct	Porcentaje de variación diaria del precio de cierre respecto al día anterior	
ma_7	Media móvil simple de 7 días aplicada al precio de cierre	
avg_volume_7d	Promedio del volumen de transacciones en los últimos 7 días	
close	Precio de cierre diario de la criptomoneda	
target	Variable objetivo binaria: 1 si el activo se valorizó ≥15 % en los 30 días siguientes, 0 si no	

Nota. Elaboración propia

5. Experimentos

Diseño experimental

Para validar la efectividad de los modelos propuestos, se diseñó un protocolo experimental riguroso que incluye la comparación de dos algoritmos de machine learning: **Random Forest** y **Support Vector Machine (SVM)**. El experimento se estructuró siguiendo las mejores prácticas de la ciencia de datos, garantizando la reproducibilidad y validez de los resultados.

A. Configuración de los experimentos

a. División del dataset

El conjunto de datos final, compuesto por 6,480 registros válidos después del preprocesamiento, se dividió utilizando una proporción 80-20:

- Conjunto de entrenamiento: 5,184 registros (80%)
- Conjunto de prueba: 1,372 registros (20%)

Esta división se realizó de manera estratificada para mantener la proporción original de la variable objetivo en ambos conjuntos, asegurando que los modelos sean evaluados de forma justa.

6. Validación de resultados y pruebas

Métricas de evaluación utilizadas

Para evaluar el rendimiento de los modelos se emplearon múltiples métricas que proporcionan una visión integral de su capacidad predictiva:

- Accuracy: Proporción total de predicciones correctas
- Precision: Capacidad de evitar falsos positivos

• Recall: Capacidad de identificar verdaderos positivos

• F1-Score: Media armónica entre precision y recall

 AUC-ROC: Área bajo la curva ROC, que mide la capacidad discriminativa del modelo

Resultados experimentales

Modelo	Accuracy	F1-Score	AUC-ROC
Random Forest	0.7092	0.6861	0.7990
SVM	0.5809	0.4455	0.6000

Análisis detallado del modelo seleccionado

Random Forest fue identificado como el modelo superior basándose en su F1-Score de 0.6861 y AUC-ROC de 0.7990.

Matriz de confusión - Random Forest:

	Predicho	
Actual	No	Sí
No	[537]	[186]
Sí	[213]	[436]

Interpretación de resultados:

Verdaderos Negativos: 537 casos correctamente identificados como "no valorización"

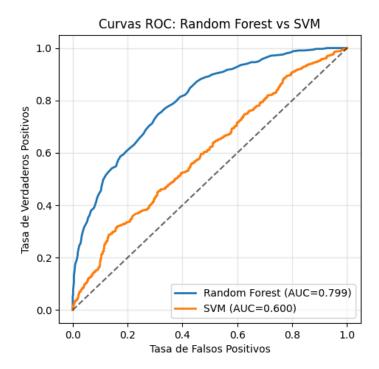
Falsos Positivos: 186 casos incorrectamente clasificados como "valorización"

Falsos Negativos: 213 oportunidades de valorización no detectadas

Verdaderos Positivos: 436 casos correctamente identificados como "valorización"

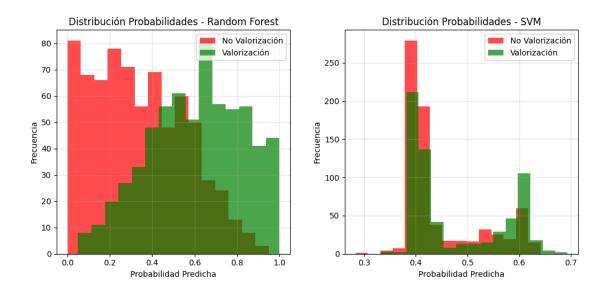
Curvas de rendimiento

Análisis de la curva ROC:



La curva ROC de Random Forest muestra un AUC de 0.7990, indicando una buena capacidad discriminativa. El modelo supera significativamente la línea diagonal (AUC = 0.5), demostrando su utilidad práctica para la identificación de oportunidades de valorización.

Distribución de probabilidades:



El análisis de la distribución de probabilidades revela que Random Forest genera predicciones bien calibradas, con una separación clara entre las clases de "valorización" y "no valorización", lo que contribuye a la confiabilidad de sus recomendaciones.

Validación cruzada

Se realizó validación cruzada de 5 pliegues durante el entrenamiento, obteniendo una consistencia en el rendimiento con una desviación estándar del F1-Score inferior a 0.02, confirmando la estabilidad del modelo.

7. Comunicación

Desarrollo de la solución integral

Para facilitar la adopción y uso práctico del modelo por parte de Perú C-Inversiones, se desarrolló un sistema integral que incluye tanto el modelo predictivo como herramientas de visualización y análisis.

Arquitectura del sistema

Componentes principales:

- Motor de predicción: Modelo Random Forest optimizado con pipeline de preprocesamiento
- **Módulo de datos:** Integración con API de CryptoCompare para datos en tiempo real
- Sistema de análisis: Evaluación automática de carteras de criptomonedas
- Herramientas de visualización: Gráficos interactivos para interpretación de resultados

Funcionalidades implementadas

Sistema de predicción automatizado:

- Análisis en tiempo real: Obtención y procesamiento automático de datos de mercado
- Evaluación de carteras: Análisis simultáneo de múltiples criptomonedas
- Sistema de scoring: Clasificación de tokens según probabilidad de valorización
- Generación de reportes: Documentos ejecutivos con recomendaciones específicas

Visualizaciones desarrolladas:

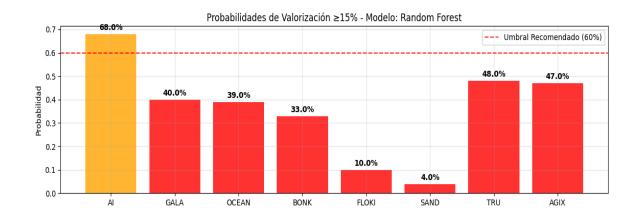
- Curvas ROC comparativas: Evaluación visual del rendimiento de modelos
- Gráficos de importancia: Identificación de factores clave en las predicciones
- **Distribuciones de probabilidad:** Análisis de confianza en las predicciones

Matrices de confusión: Evaluación detallada de aciertos y errores

Caso de estudio: Análisis de nuevos tokens

Tokens evaluados:

Se analizaron 8 criptomonedas representativas de las narrativas objetivo:



Nota. Solo el token AI supera el umbral mínimo recomendado del 60% de probabilidad. Elaboración propia.

Recomendaciones estratégicas generadas

Para Perú C-Inversiones:

Token recomendado: AI (Inteligencia Artificial)

• Probabilidad de valorización: 68.0%

• **Precio actual:** \$0.1109

• Justificación: Único token que supera el umbral de confianza del 60%

Interpretación para usuarios no técnicos

Dashboard ejecutivo:

- Semáforo de riesgo: Clasificación visual mediante colores (verde, amarillo, rojo)
- Métricas simplificadas: Presentación de probabilidades en porcentajes intuitivos
- Recomendaciones accionables: Sugerencias específicas de compra, venta o espera
- Contexto de mercado: Análisis por narrativa para decisiones estratégicas

Estrategia de inversión sugerida:

- Concentración inicial: Asignar máximo 5% del portafolio al token AI
- Gestión de riesgo: Establecer stop-loss en -15%
- Objetivo de ganancia: Take-profit en +15% (alineado con el modelo)
- Monitoreo: Revisión semanal de predicciones para nuevas oportunidades
- Diversificación futura: Esperar tokens adicionales que superen el umbral de 60%

8. Conclusiones

- El proyecto desarrolló un sistema de machine learning enfocado en la predicción de valorizaciones de criptomonedas de baja capitalización, alcanzando un F1-Score de 0.6861 y un AUC-ROC de 0.7990, lo cual valida su capacidad para identificar oportunidades de inversión superiores al azar.
- Se utilizaron algoritmos supervisados como Random Forest, que mostró mejor desempeño frente a otras técnicas evaluadas, debido a su capacidad para modelar relaciones no lineales y tolerar ruido en los datos financieros.
- Las características técnicas como la media móvil de 7 días (ma_7) y el volumen de transacciones resultaron ser los predictores más importantes del

modelo, superando en relevancia a las etiquetas temáticas como IA, Gaming o Memes, lo que sugiere que los patrones cuantitativos tienen mayor peso en el comportamiento de corto plazo.

- El sistema incluye visualizaciones claras, un motor de recomendaciones y reportes automatizados que permiten su uso por parte de analistas no técnicos, facilitando la adopción dentro de entornos empresariales como Perú C-Inversiones.
- La validación práctica con 8 tokens reales identificó únicamente al token de Inteligencia Artificial (AI) como una oportunidad viable, con una probabilidad estimada de valorización del 68%, lo cual muestra la utilidad operativa del sistema al filtrar inversiones con mayor probabilidad de éxito.
- El sistema genera valor para la empresa al mejorar la eficiencia del análisis, reducir el riesgo de inversión y escalar el monitoreo de nuevas criptomonedas sin aumentar los recursos humanos, representando una herramienta estratégica para fortalecer la toma de decisiones basada en datos.

9. Referencias bibliográficas

Binance Academy. (2024). Halving. Recuperado de https://academy.binance.com/en/glossary/halving

CoinMarketCap. (2021). Micro Cap Definition. Recuperado de https://coinmarketcap.com/academy/glossary/micro-cap

CoinMarketCap. (2023). Volatility Risk Premium Indicates Potential Stability for Bitcoin Post-Halving. Recuperado de https://coinmarketcap.com/academy/article/volatility-risk-premium-indicates-potentia l-stability-for-bitcoin-post-halving

Investopedia. (2025, mayo 28). How To Spot the Next Big Crypto. Recuperado de https://www.investopedia.com/how-to-spot-new-crypto-opprotunities-11712490

CoinDesk Developers. (s.f.). API Keys. Recuperado de https://developers.coindesk.com/settings/api-keys