







Universidad Nacional Rosario Castellanos

Licenciatura en Ciencia de Datos para Negocios

403

Carlos Daniel Gallegos Garcia

Juan Luis Gonzales Torrez

Salvador Rojas Salinas

Proyecto prototípico:

Desarrollo y Simulación de una Nueva Moneda Digital

en México: Modelado y Análisis

Antes de su Integración en Blockchain.









Impacto y predicción del uso de una criptomoneda como propuesta para proyecto de innovación social en la Universidad Nacional Rosario Castellanos.

Introducción

Los estudiantes universitarios enfrentan diversas limitaciones económicas que afectan su rendimiento académico, bienestar general y desarrollo profesional. La falta de acceso a recursos financieros y la escasa autonomía económica generan un entorno poco propicio para su progreso.

En este contexto, una criptomoneda universitaria surge como una alternativa innovadora y sostenible. **Rosareum**, concebida como una **moneda digital** desarrollada por y para estudiantes, representaria una solución disruptiva que integra tecnología y compromiso social para mejorar la educación financiera y el desarrollo comunitario.



• Objetivos del Proyecto

Objetivo General

Diseñar y simular una criptomoneda para uso estudiantil que contribuya a la inclusión y educación financiera, así como su desarrollo en el entorno universitario de la Universidad Nacional Rosario Castellanos.









Objetivos Específicos

- Diseñar el modelo económico y funcional de Rosareum.
- Simular su comportamiento con datos sintéticos.
- Aplicar herramientas de inteligencia artificial, estadística multivariada, ecuaciones diferenciales entre otras herramientas y disciplinas para estimar y analizar el comportamiento socioeconómico de la criptomoneda.
- Evaluar su impacto potencial y factibilidad tanto social como económica en la comunidad estudiantil.

Metodología

Fase 1: Diseño Conceptual

Definición de variables clave, reglas de emisión, circulación, y condiciones de uso de Rosareum.

Fase 2: Generación de Datos Sintéticos

Basado en encuestas hacia la comunidad de la UNRC, específicamente para la sede Santo Tomas, se generarán datos sintéticos en **Python** para poder simular el comportamiento de la moneda en la comunidad estudiantil.

Fase 3: Almacenamiento y Análisis Estadístico

Se propone insertar los registros en una base de datos **MongoDB** utilizando el formato **JSON**, el cual representará una visualización equivalente a la exportación de bloques codificados de una criptomoneda real, originalmente en un formato binario. Una vez almacenados estos datos en MongoDB, se procederá a su extracción y manipulación en el entorno estadístico **R**, con el fin de aplicar métodos analíticos y modelos estadísticos que nos permitan analizar y estimar su comportamiento respecto al tiempo.

Fase 4: Visualización y Análisis

Analizar los resultados de la simulación y predicción de la criptomoneda para determinar su viabilidad económica y aceptación social.









Modelo Económico y Funcional de Rosareum

El modelo propuesto combina oferta y demanda con una **función de utilidad social**, considerando variables como:

- Demanda estudiantil.
- Oferta disponible.
- Inflación Acumulada
- Eventos sociales y académicos.
- Precio
- Liquidez

El precio se modela mediante la siguiente fórmula:

$$P_t = \alpha \cdot \left(\frac{D_t}{S_t}\right) + \beta \cdot I_t + \gamma \cdot E_t$$

Donde:

- P_t: Precio o índice de valor de la criptomoneda en el tiempo t
- $\frac{D_t}{S_t}$: Relación entre demanda D_t y oferta S_t , reflejo directo de las presiones de mercado
- I_t: Índice de participación estudiantil, representa el grado de uso, interacción o aceptación de la criptomoneda por parte del alumnado
- E_t: Índice de sentimiento estudiantil, capturado mediante encuestas, análisis de texto, opiniones en redes o métricas indirectas de confianza o percepción
- α,β,γ: Coeficientes de ponderación que determinan la influencia relativa de cada variable sobre el precio, calibrados empíricamente o por simulación









Fundamento del modelo

Este modelo surge como una herramienta de simulación y predicción del precio de una criptomoneda en un entorno universitario. Dado que el valor de este activo no está regido por mercados financieros tradicionales, sino por el **uso social, la participación activa y la confianza de la comunidad universitaria**, es necesario incluir:

- 1. Factores económicos tradicionales, como la oferta y la demanda.
- 2. **Factores sociales y psicológicos**, como el compromiso (índice de participación) y el sentimiento (confianza o entusiasmo).
- 3. **Un enfoque flexible** que permita simular escenarios bajo diferentes condiciones sociales.

Mecanismo de Conversión y Alcance Social

Dado que la universidad no puede administrar formalmente una moneda, **Rosareum** se plantea como una criptomoneda descentralizada, administrada por los propios estudiantes, quienes podrían utilizarla para:

- Intercambiar bienes y servicios dentro del campus.
- Recibir becas y ayudas estudiantiles.
- Participar en actividades académicas y extracurriculares.

El ecosistema fue simulado utilizando tecnologías como **MongoDB**, **Python** y **R**, que permitieron la generación, almacenamiento y análisis de datos sintéticos basados en variables socioeconómicas relevantes. Entre estas variables se incluyeron aspectos como la **confianza en la moneda**, el **nivel de adopción por parte de los estudiantes**, la **frecuencia de eventos estudiantiles** dentro de la universidad, entre otras variables sociales seleccionadas estratégicamente para reflejar el entorno universitario.









Como muestra estadística se utilizó la sede **Santo Tomás**, donde se aplicó una **encuesta diagnóstica** con el propósito de establecer los pesos iniciales de las variables mencionadas. El cuestionario incluyó preguntas como:

- ¿Qué noción tienes de las criptomonedas?
- ¿Qué riesgo consideras más importante al utilizar una criptomoneda?
- ¿Consideras que una criptomoneda podría tener relevancia en nuestro entorno estudiantil?

Estas preguntas estaban diseñadas con opciones de respuesta cerradas, lo que permitió cuantificar la frecuencia de cada elección. Con base en estas frecuencias, se asignaron los **pesos iniciales** a cada variable, los cuales fueron utilizados en la generación y calibración de los datos sintéticos. Este enfoque permitió modelar un ecosistema digital fundamentado en percepciones reales de la comunidad estudiantil.

Generación y Almacenamiento de Datos

Los datos fueron simulados en un notebook de Python y almacenados en una base de datos MongoDB, clasificados según su naturaleza: **económica**, **social** y **temporal**. Además, se organizaron cronológicamente para representar un ciclo académico completo de **dos años**.

Se definió el ciclo académico considerando los **dos semestres universitarios por año**, así como los periodos vacacionales. A cada periodo se le asignó un valor categórico: **1** para los semestres activos, **0** para los periodos de transición y **-1** para las vacaciones. Esta codificación permitió incorporar el calendario académico al modelo de simulación de datos.

Posteriormente, utilizando nuestro **modelo de precio** y los **pesos previamente asignados a cada variable**, se generó la simulación de datos a lo largo de los dos años. Para garantizar la reproducibilidad de los resultados, se estableció una **semilla aleatoria fija**, asegurando así que cada ejecución del modelo produzca los mismos datos simulados.









Enlace a la encuesta y el notebook de generación de datos sintéticos: Contiene el proceso completo de simulación utilizando Python, incluyendo la definición de variables socioeconómicas, establecimiento del ciclo académico, asignación de pesos derivados de la encuesta.

Google Forms (Encuesta):

https://docs.google.com/forms/d/e/1FAIpQLSc14xmBZ0gcVXk6VOARgCOdm_LvB dJxqcbAsyeohWI65y3JGQ/viewform?usp=sharing&ouid=11761917423386451886 3

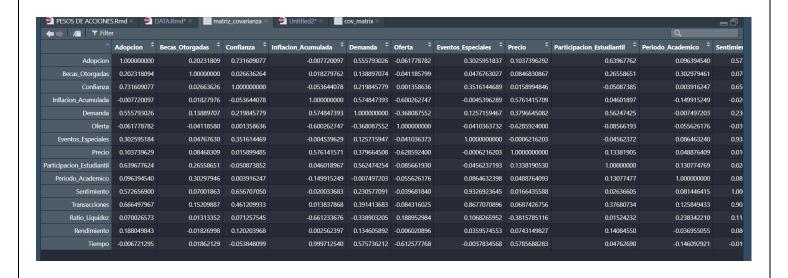
Google Colab (Notebook Python):

https://colab.research.google.com/drive/1GGbblXnULpAk-ODleuGVoFhgBTUEwqGN?usp=sharing

Análisis de Datos

Normalización y Matriz de Varianza-Covarianza

A partir del dataframe generado en R, se procedió a normalizar los datos con el objetivo de calcular los rendimientos diarios del precio. Posteriormente, se construyó una matriz de varianza-covarianza, la cual permitió analizar las relaciones entre variables, identificando patrones de comportamiento y posibles correlaciones significativas dentro del ecosistema simulado.











Correlaciones Positivas Más Relevantes

- Eventos Especiales y Sentimiento (0.932): Los eventos mejoran notablemente el sentimiento estudiantil.
- Sentimiento y Transacciones (0.903): Un buen ánimo colectivo incrementa el uso de la moneda.
- Eventos Especiales y Transacciones (0.867): Los eventos también motivan mayor circulación.
- Confianza y Sentimiento (0.657): Mayor confianza se traduce en un mejor sentimiento hacia la moneda.
- Participación y Adopción (0.640): La participación activa incrementa la adopción de la moneda.

Correlaciones Negativas Significativas

- Inflación Acumulada y Oferta (-0.600): Mayor inflación reduce la oferta para evitar devaluación.
- **Precio y Oferta (-0.629)**: A mayor precio, menor oferta, posiblemente por especulación o retención.
- Liquidez e Inflación Acumulada (-0.661): Alta inflación reduce la seguridad en el intercambio.
- **Liquidez y Precio (-0.382)**: Estudiantes prefieren retener la moneda si creen que aumentará de valor.
- **Demanda y Liquidez (-0.339)**: Alta demanda puede coincidir con retención en lugar de circulación.

Las correlaciones indican una fuerte relación entre variables sociales y el uso de la moneda, como el sentimiento, la confianza y la participación. También se observan efectos económicos importantes, donde la inflación, el precio y la liquidez influyen en la circulación y la oferta de la moneda.







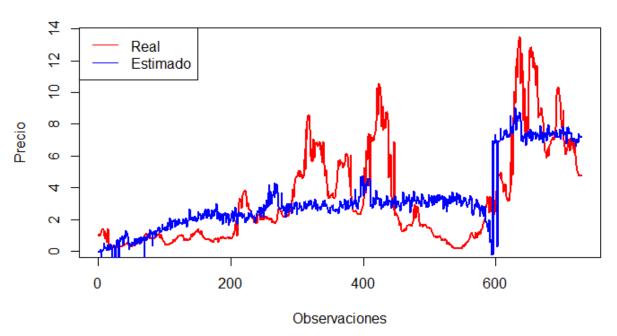


Modelos Predictivos

1. Regresión Lineal

Se aplicó un modelo lineal para predecir el precio en función de las variables clave. El modelo resultó limitado por la naturaleza no lineal de los datos, produciendo predicciones suavizadas en la estimación del precio.

Precio Estimado vs Real



2. Random Forest (Bosques Aleatorios)

Para mejorar la **precisión en la estimación del precio** de la criptomoneda a partir de los **datos sintéticos generados**, se implementó un modelo de **aprendizaje automático supervisado** utilizando el algoritmo **Random Forest**. Este modelo es especialmente eficaz cuando se trabaja con datos con múltiples variables y relaciones no lineales, como en este caso.

Random Forest funciona construyendo un conjunto de árboles de decisión a partir de diferentes subconjuntos del conjunto de datos y de las variables predictoras. Cada árbol realiza una predicción independiente, y la modelo final promedia o vota los resultados individuales de todos los árboles, lo que mejora la robustez y precisión del modelo frente a posibles ruidos o sobreajustes.









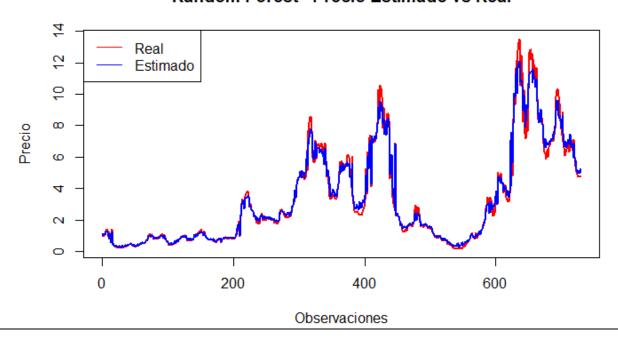
En este proyecto, el modelo fue entrenado con variables socioeconómicas sintetizadas, como:

- Nivel de confianza en la criptomoneda.
- Adopción estudiantil a lo largo del ciclo académico.
- Eventos universitarios y su frecuencia.
- Factores económicos simulados, como inflación o incentivos.
- Participación social y sentimiento colectivo.
- Oferta de la moneda, es decir, la cantidad total disponible en circulación.
- Demanda de la moneda, reflejada por el interés y uso real dentro del ecosistema estudiantil.
- **Liquidez**, entendida como la facilidad con la que puede intercambiarse la moneda sin afectar drásticamente su precio.

El objetivo fue predecir el **precio estimado de la criptomoneda** en función de estas variables, evaluando la influencia relativa de cada una sobre el valor final. El modelo permitió identificar cuáles factores tienen un mayor peso en la variación del precio y, a su vez, ofreció una herramienta predictiva confiable para simular diferentes escenarios dentro del entorno estudiantil.

Además, la capacidad de Random Forest para manejar interacciones complejas entre variables y su resistencia al sobreajuste lo convierten en una elección adecuada para este tipo de análisis basado en datos sintéticos.

Random Forest - Precio Estimado vs Real











Resultados del Modelo:

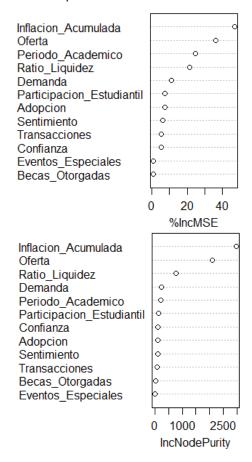
MAE (Error Absoluto Medio): 0.192

RMSE (Raíz del Error Cuadrático Medio): 0.327

Importancia de Variables (Random Forest):

Variable	%IncMSE	IncNodePurity
Inflación Acumulada	46.91	2979.25
Oferta	36.15	2108.32
Ratio de Liquidez	21.55	771.16
Periodo Académico	24.53	215.93
Demanda	11.28	228.88

Importancia de Variables



Estas variables explican significativamente las fluctuaciones en el precio de la criptomoneda, sugiriendo que los factores económicos y cíclicos (como el calendario escolar) tienen un peso relevante en su comportamiento.

 Análisis de Componentes Principales (PCA) del Ámbito Social Y Creación del índice de popularidad social (IPS)

Con el objetivo de sintetizar la información contenida en múltiples variables sociales del conjunto de datos, se aplicó un Análisis de Componentes Principales (PCA, por sus siglas en inglés). Este método estadístico permite reducir la dimensionalidad del dataset, conservando la mayor parte de la varianza y eliminando la redundancia entre variables correlacionadas.







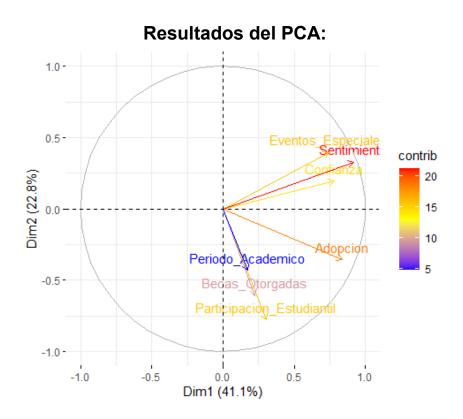


El PCA se aplicó sobre las siguientes variables sociales clave:

- Confianza en la criptomoneda.
- Adopción de la moneda por parte del estudiantado.
- Participación estudiantil en actividades económicas y sociales.
- Eventos especiales organizados dentro de la comunidad universitaria.
- Becas otorgadas, como estímulo económico institucional.
- Periodo académico, clasificado como activo, transitorio o vacacional.

Este análisis transformó el conjunto de variables originales —potencialmente correlacionadas, según la matriz de varianza-covarianza— en un nuevo conjunto de componentes principales no correlacionados. Cada componente es una combinación lineal de las variables originales y está ordenado de acuerdo con la varianza explicada, permitiendo reducir la complejidad sin perder información relevante.

El uso de PCA facilitó la interpretación de los datos y mejoró la eficiencia de los modelos predictivos posteriores (como el modelo de precios con Random Forest), al trabajar con un conjunto reducido de variables que capturan los patrones subyacentes del comportamiento social dentro del ecosistema universitario.









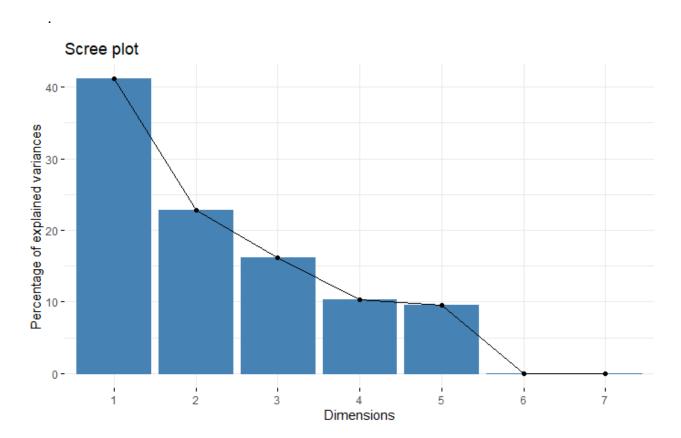


Se observa que las variables sentimiento, adopción, participación estudiantil y eventos académicos presentan una alta carga en los primeros componentes principales, lo que indica su fuerte influencia dentro del ámbito social. En contraste, la variable periodo académico, que fue determinante en el análisis del precio, muestra una baja contribución al componente social.

Construcción del Índice de Popularidad Social

Con base en los resultados del Análisis de Componentes Principales (PCA), se procedió a construir un **Índice de Popularidad Social**, diseñado para resumir y representar de manera integral el comportamiento colectivo dentro del ecosistema universitario. Este índice fue generado mediante una **combinación ponderada de los tres primeros componentes principales**, los cuales explican la mayor parte de la **varianza acumulada** en las variables sociales analizadas.

Las **ponderaciones asignadas** a cada componente se basaron en el **porcentaje de varianza explicada**, permitiendo que aquellos componentes con mayor capacidad descriptiva tuvieran un peso proporcionalmente más alto en el índice final. De esta manera, el índice captura de forma eficiente la dinámica social a lo largo del ciclo académico.











Modelado de la Evolución del IPS con Ecuaciones Diferenciales

Para modelar la **evolución temporal** del Índice de Popularidad Social, se empleó una **ecuación diferencial de crecimiento exponencial** de la forma:

$$\frac{dI}{dT} = r \cdot I$$

Donde:

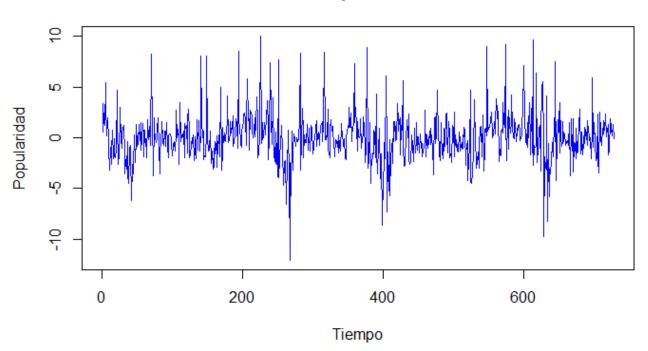
I representa el **índice de popularidad social**.

T es el **tiempo**, medido en unidades correspondientes al ciclo académico.

r es la **tasa de crecimiento**, asumida constante, estimada mediante regresión lineal del logaritmo del índice, lo cual permite linealizar la ecuación para su análisis.

 $\frac{dI}{dT}$ es la **derivada del índice con respecto al tiempo**, es decir, su tasa de cambio instantánea.

Índice de Popularidad Social











Este modelo supone que el crecimiento del índice es **proporcional a su valor actual**, lo que lo convierte en una representación adecuada para **fenómenos acumulativos**, como la **adopción social**, la **difusión de ideas** o el aumento progresivo de la **popularidad** dentro de una comunidad.

La solución de esta ecuación proporciona una curva de crecimiento continuo del tipo:

$$I(T) = C_0 \cdot e^{rT}$$

Donde C_0 es el valor inicial del índice en el momento T=0

Valor inicial del índice cuando:

$$T = 0 I = 0.04940699$$

$$T = 1 I = 2.204332$$

Obteniendo el siguiente resultado para la modelación con la ecuación diferencial exponencial:

$$I(T) = 0.4940699 \cdot e^{3.7973T}$$

Este modelo nos permite analizar la evolución del comportamiento colectivo a lo largo del tiempo, proporcionando una herramienta útil para proyectar escenarios futuros. Al ajustar el valor del parámetro r, es posible simular diferentes tasas de adopción y evaluar cómo cambios en las condiciones sociales podrían influir en la aceptación de la criptomoneda. De esta manera, el modelo sirve como base para estimar el **impacto potencial** de su implementación en la comunidad universitaria.



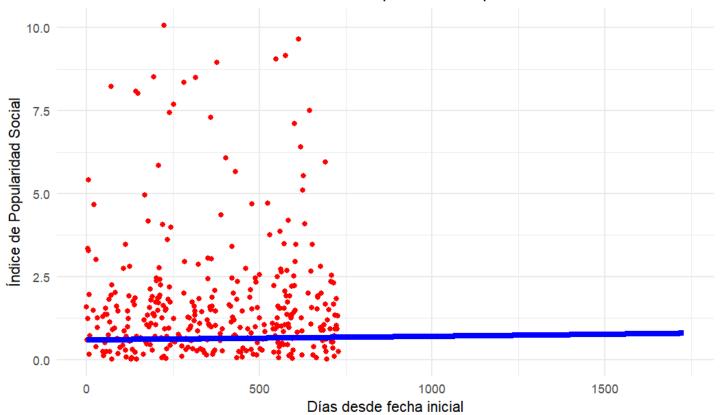






Resultados del modelado:





• Análisis Interpretativo del Modelo

La implementación del modelo de crecimiento exponencial permitió proyectar de forma eficaz la evolución temporal del Índice de Popularidad Social asociado a la criptomoneda universitaria. Los resultados indican un comportamiento acumulativo y sostenido del índice, que es coherente con fenómenos de adopción colectiva impulsados por el efecto de red, el boca a boca y la influencia de eventos institucionales.









Aunque se trata de un modelo matemáticamente simple, logra capturar adecuadamente las tendencias fundamentales del comportamiento social en torno a la moneda. Además, su estructura puede ser ampliada en el futuro mediante modelos más complejos, como el modelo logístico o los modelos tipo SIR, que permiten incorporar dinámicas de saturación o interacción entre individuos. La curva obtenida al ajustar el modelo a los datos simulados mostró una trayectoria suavemente creciente y de forma exponencial, reflejando el aumento progresivo del interés y aceptación de nuestra criptomoneda.

• Estimación del Índice de Popularidad Social mediante Redes Neuronales con Keras.

Ahora utilizaremos una red neuronal para determinar la estimación del índice de popularidad social con respecto al tiempo. Hay que tomar en cuenta que **no es lo mismo el modelado de la evolución exponencial del índice** (que se puede hacer, por ejemplo, mediante ecuaciones diferenciales), que **la estimación del índice en función de variables sociales en un periodo dado**, para lo cual empleamos una red neuronal implementada con la librería **Keras** en R.

La red neuronal fue diseñada como un modelo secuencial que recibe como entrada una matriz de variables sociales normalizadas (como *Participación Estudiantil*, *Confianza*, *Becas Otorgadas*, entre otras), y tiene como salida el **índice de popularidad** generado previamente con Análisis de Componentes Principales (PCA).

La arquitectura del modelo contiene dos capas densas ocultas: la primera con 16 neuronas y la segunda con 8, ambas con activación ReLU, y una capa de salida con una sola neurona para predecir el valor continuo del índice. Se utilizó *dropout* en la primera capa para evitar el sobreajuste, y el modelo fue compilado con la función de pérdida **mean squared error (MSE)** y el optimizador **Adam**.

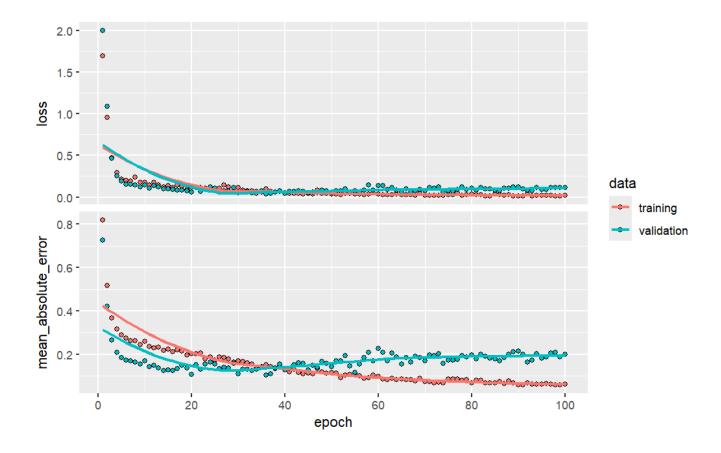
El entrenamiento se realizó durante 100 épocas, con un tamaño de lote (*batch size*) de 8 y un 20% de los datos reservado para validación. Finalmente, se visualizaron las métricas de entrenamiento, permitiendo observar la evolución del error y la estabilidad del modelo durante el aprendizaje.











La gráfica muestra cómo fue el desempeño de la red neuronal durante 100 épocas de entrenamiento. Hay dos métricas principales para los datos de entrenamiento y validación:

Loss: Aquí vemos el error cuadrático medio (MSE) entre las predicciones y los valores reales del índice de popularidad. El error baja rápido en las primeras 20 épocas y luego se mantiene más o menos estable. El error en validación es un poco más alto que en entrenamiento, como es normal, pero no hay señales claras de sobreajuste.

Mean Absolute Error: Esta muestra el error absoluto medio y sigue una tendencia parecida al loss: baja al principio y luego se estabiliza. A partir de la época 50 se ve un pequeño aumento en el error de validación, que podría ser el comienzo de un leve sobreajuste, pero sigue siendo algo aceptable.



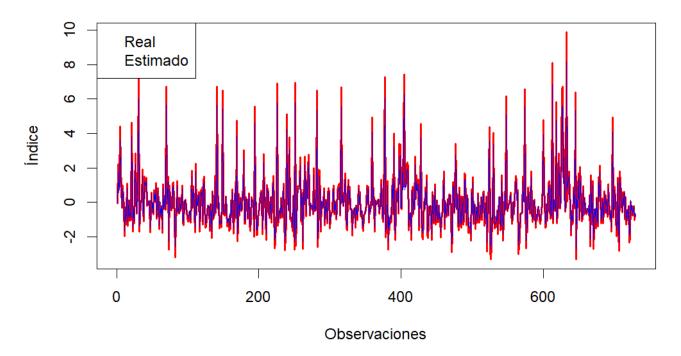






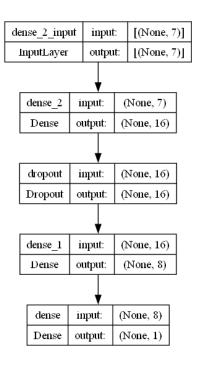
Resultados de la red:

Índice Popularidad Real vs Estimado



Podemos observar una alta concordancia entre el índice de popularidad real y el estimado por el modelo. A lo largo de las observaciones, ambas series siguen patrones similares, lo que sugiere que nuestro modelo es capaz de capturar adecuadamente la dinámica del fenómeno social representado. Aunque hay ciertos picos donde el valor real presenta variaciones más bruscas que el estimado, estas discrepancias no son constantes ni sistemáticas, lo cual indica un buen ajuste general.

Esta alineación entre ambos índices valida la capacidad predictiva del modelo y refuerza su utilidad para simular escenarios futuros o analizar tendencias en contextos universitarios. Además, la estabilidad en las diferencias sugiere que el modelo no está sobreajustado y mantiene una buena generalización sobre los datos.











Enlace al notebook en R con el análisis estadístico completo: Incluye el procesamiento de datos, normalización, construcción del índice de popularidad social mediante PCA, cálculo de la matriz de varianza-covarianza, modelado con ecuaciones diferenciales.

Notebook R:

https://tubular-custard-49d994.netlify.app/

Análisis integral de los resultados, su transformación con sentido social y potencial de impacto

Los resultados del proyecto muestran que la implementación de una criptomoneda universitaria tiene viabilidad no solo como un pequeño sistema financiero interno, sino también como herramienta de cohesión social. La moneda, al estar diseñada con base en variables sociales y económicas contextualizadas, demostró potencial para fortalecer el sentido de pertenencia a la universidad, incentivar la participación estudiantil y facilitar el acceso equitativo a recursos que fomenten la educación financiera.

El análisis estadístico y el modelado dinámico demostraron que la participación estudiantil crece de forma exponencial cuando la confianza es sólida y los beneficios son tangibles. En este contexto, factores sociales como la adopción, los eventos y la participación resultaron más determinantes que las variables académicas, reafirmando que el valor simbólico y comunitario de la moneda supera con creces su función monetaria.

Además, se identificó una relación directa entre confianza y adopción, lo que indica que el éxito de la moneda depende de políticas claras, seguras y enfocadas en la percepción positiva del estudiantado

Finalmente, este modelo puede replicarse en otras sedes de la universidad, adaptándose a los contextos específicos de cada una, lo que permitirá evaluar de manera precisa su desempeño individual y, a su vez, calcular una viabilidad ponderada y general del sistema a nivel nacional. Rosareum actúa como un catalizador para el desarrollo integral del estudiantado, al incentivar el aprendizaje de la educación financiera y promoviendo la toma de decisiones colectivas desde un enfoque práctico e innovador para el desarrollo de la comunidad estudiantil.