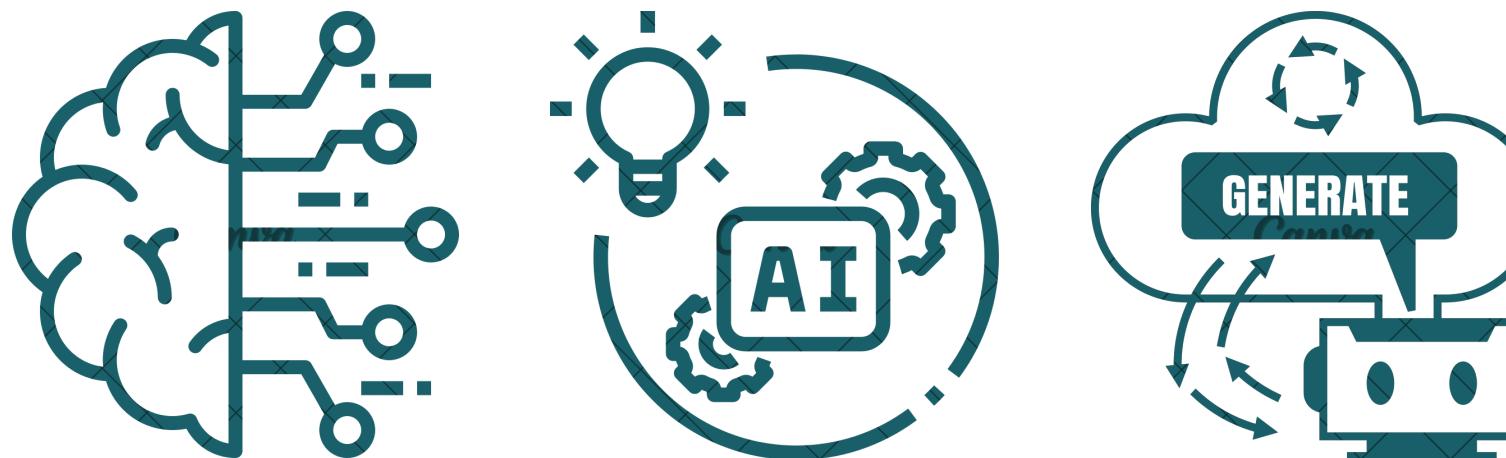


OPTIMIZACIÓN PARA MACHINE LEARNING_A
Docente: Oswaldo Velásquez Castañon

Aplicación y optimización de modelos de *Machine Learning* para evaluar el rendimiento académico de estudiantes de secundaria



Integrantes:
• Fabrizio Montalvo
• Francesco Carlevarino

Contenido

1. Introducción

2. Objetivos

3. Marco Teórico

4. Metodología

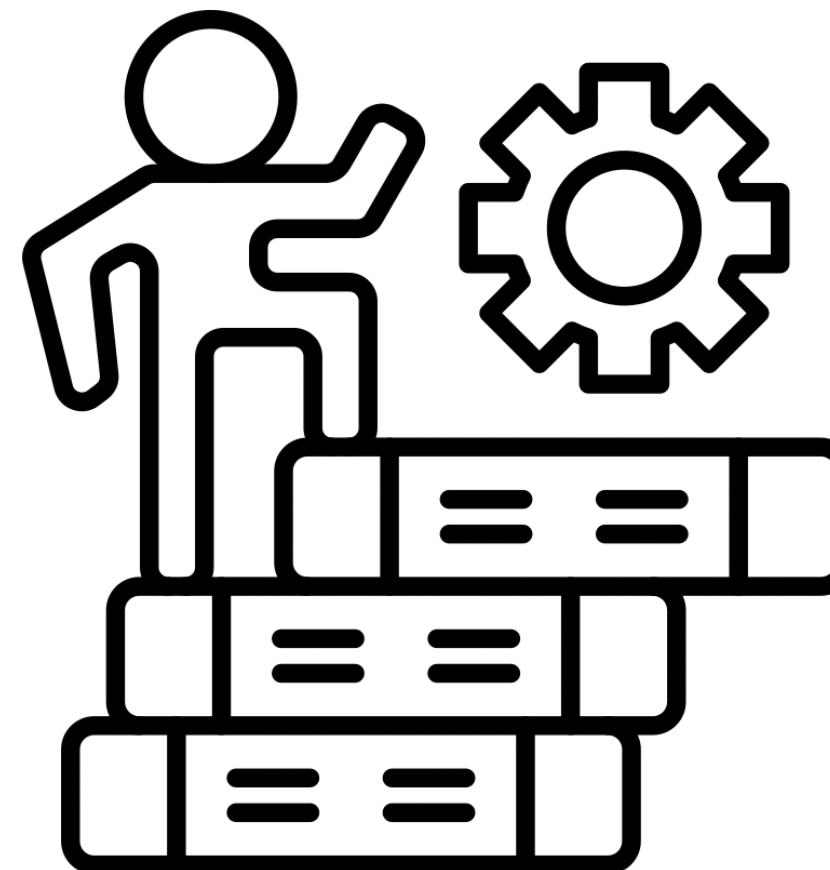
5. Resultados

6. Conclusiones



Introducción

Rendimiento Académico



Indicador clave para el desarrollo educativo y profesional de los estudiantes (Tomasevic et al., 2020).



Se mide mediante parámetros como calificaciones generales, tasas de deserción escolar y pruebas estandarizadas, como el SAT en Estados Unidos o PISA a nivel internacional.

Introducción

Factores que influyen el rendimiento académico

- Factores internos: motivación, habilidades cognitivas
- Factores externos: socioeconómicos, educativos y familiares



En EE. UU., estudiantes de bajos ingresos tienen un menor desempeño en matemáticas y lectura (NCES, 2023).

En Europa, PISA reporta brechas de más del 30% en comprensión lectora y matemáticas entre niveles socioeconómicos (OCDE, 2019).



Problema

Los estudiantes de secundaria enfrentan una etapa crucial en su desarrollo, influida por factores como el entorno familiar, condiciones socioeconómicas y prácticas de estudio (Díaz, 2003).

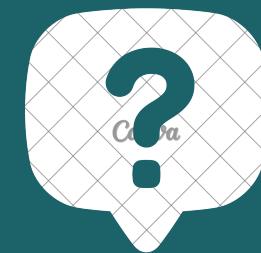


Los algoritmos de *machine learning* analizan múltiples variables y detectan patrones complejos en el rendimiento académico (Muros, 2024).

Muchos estudios excluyen variables como hábitos de salud (consumo de alcohol, horas de sueño) y otros factores familiares (nivel educativo de los padres), lo que podría reducir la precisión de los modelos predictivos.



Pregunta de investigación



¿Cómo impacta la inclusión de variables como el consumo de licor o el nivel socioeconómico en el rendimiento de los algoritmos de *machine learning* para predecir el rendimiento académico de los estudiantes de secundaria?

Contenido

1. Introducción

2. Objetivos

3. Marco Teórico

4. Metodología

5. Resultados

6. Conclusiones



Objetivos

Objetivo general

Optimizar los modelos de aprendizaje automático utilizados para predecir el rendimiento académico de estudiantes de secundaria, mejorando la precisión predictiva mediante técnicas avanzadas de optimización de la función de costos.



Objetivos específicos

- 1 Implementar optimizadores avanzados para la mejora de la función de costos en los modelos de machine learning.
- 2 Evaluar el impacto de diferentes estrategias de optimización en la precisión y eficiencia de los modelos.
- 3 Comparar el rendimiento de los modelos optimizados con los modelos iniciales para determinar las mejoras en la predicción del rendimiento académico.

Relevancia del estudio



Instituciones educativas

Facilita el diseño de programas de apoyo específicos para estudiantes en riesgo.

Promueve la retención escolar y mejora el desempeño académico al crear entornos inclusivos adaptados a las necesidades individuales.

Contribución Social

Ayuda a construir un sistema educativo más equitativo y de mayor calidad.

Permite reducir brechas educativas mediante estrategias justas, promoviendo cohesión social y acceso igualitario a oportunidades de aprendizaje.

Aporte a la Comunidad Científica

Amplía el uso del aprendizaje automático en la predicción del rendimiento académico.

Mejora la precisión de los modelos predictivos y fomenta el desarrollo de metodologías avanzadas e innovadoras para la inclusión educativa.

Contenido

1. Introducción

2. Objetivos

3. Marco Teórico

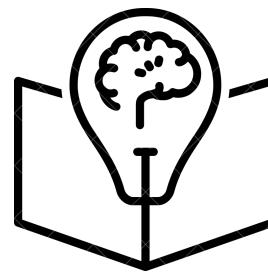
4. Metodología

5. Resultados

6. Conclusiones



Marco Teórico



ADAM

ADAM es un optimizador que ajusta de manera adaptativa la tasa de aprendizaje para cada parámetro.

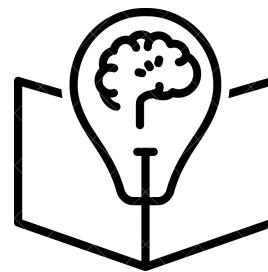
Utiliza promedios móviles del gradiente y de su varianza para acelerar la convergencia y mejorar la estabilidad del entrenamiento. Es especialmente útil en problemas con datos ruidosos o gradientes dispersos.

ADAMW

ADAMW es una variante de ADAM que corrige el manejo del decaimiento de pesos separándolo explícitamente del cálculo de los gradientes.

Evita que el regularizador interfiera con las actualizaciones adaptativas, logrando entrenamientos más estables y una mejor generalización del modelo.

Marco Teórico



BFGS

BFGS es un método de optimización cuasi-Newton que aproxima la matriz Hessiana para guiar las actualizaciones de los parámetros de forma más precisa que los métodos basados solo en gradientes.

Permite converger rápidamente hacia óptimos locales en problemas suaves y bien condicionados.

SGD

SGD actualiza los parámetros del modelo utilizando pequeños lotes de datos, lo que acelera el entrenamiento y permite escapar de mínimos locales.

Este comportamiento favorece una mejor exploración del espacio de parámetros. Es un optimizador simple, eficiente y ampliamente utilizado en modelos de aprendizaje profundo.

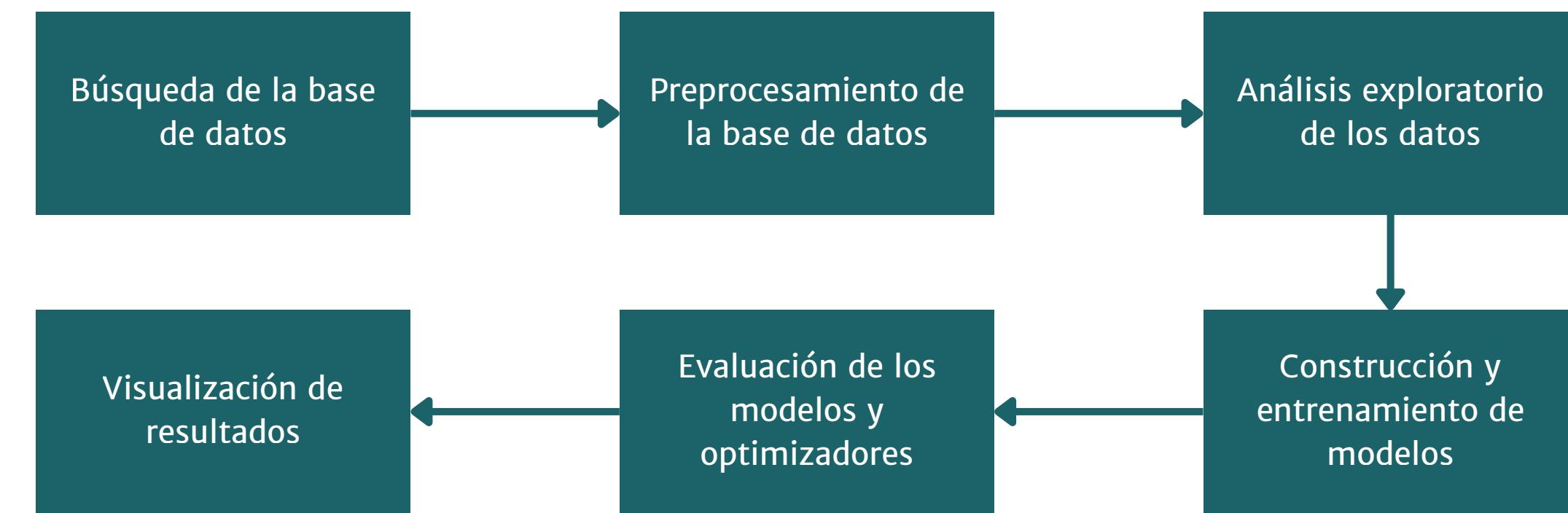
Contenido

1. Introducción
2. Objetivos
3. Marco Teórico
- 4. Metodología**
5. Resultados
6. Conclusiones



Metodología

Flujograma



Metodología

Búsqueda de la base de datos

- Se exploraron fuentes como la Plataforma Nacional de Datos Abiertos, GitHub y Kaggle.
- Se seleccionó una base de datos con información sobre el rendimiento académico de aproximadamente 400 estudiantes de secundaria.
- Los datos corresponden a dos escuelas de Portugal.

Preprocesamiento de la base de datos

- Se transformaron las variables categóricas en numéricas utilizando la librería LabelEncoder.

Metodología

Análisis exploratorio de los datos

- Se utilizaron herramientas de visualización, como:
 - Histogramas.
 - Diagramas de dispersión.
 - Matrices de correlación.
- Estas herramientas ayudaron a identificar patrones en las variables.
- Se analizaron las relaciones entre:
 - Variables predictoras: nivel socioeconómico, hábitos de estudio, entre otras.
 - Variable objetivo: rendimiento académico.

Metodología

Construcción y entrenamiento de modelos

- Se construyó un modelo de red neuronal utilizando la librería Keras. El modelo contó con dos capas ocultas y una capa de salida con activación sigmoide para la predicción de la variable de rendimiento académico.
- El modelo se compiló utilizando diferentes optimizadores y funciones de pérdida.
 - Para los optimizadores Adam, AdamW y SGD, se utilizó la función de pérdida mean_squared_error propia de Keras.
 - Para el optimizador BFGS, se implementó una clase personalizada desde cero, que utilizó una función de pérdida basada en `scipy.optimize` para realizar la optimización.
- Durante el entrenamiento, se utilizó el conjunto de entrenamiento para ajustar los pesos del modelo, y se validó el modelo en el conjunto de validación para prevenir el sobreajuste.

Metodología

Evaluación de modelos

- Una vez entrenado el modelo, se evaluó su rendimiento utilizando el conjunto de prueba. Se compararon las pérdidas (loss) del entrenamiento y la validación, así como las métricas de rendimiento en el conjunto de prueba. Las métricas clave fueron la pérdida de entrenamiento, la pérdida de validación y la pérdida en el conjunto de prueba, que se utilizaron para comparar la efectividad de los diferentes optimizadores y funciones de pérdida.

Contenido

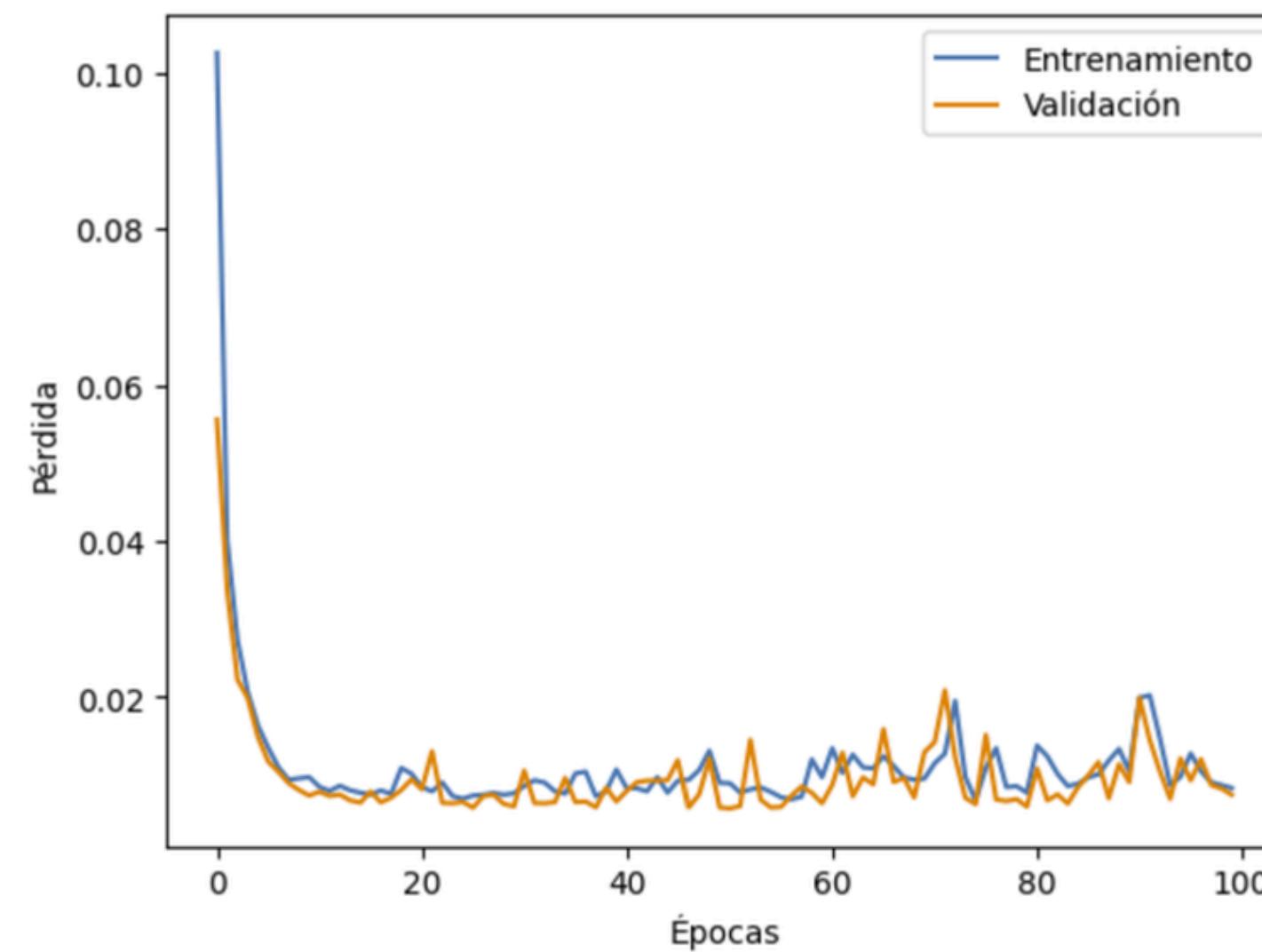
1. Introducción
2. Objetivos
3. Marco Teórico
4. Metodología
- 5. Resultados**
6. Conclusiones

Contenido



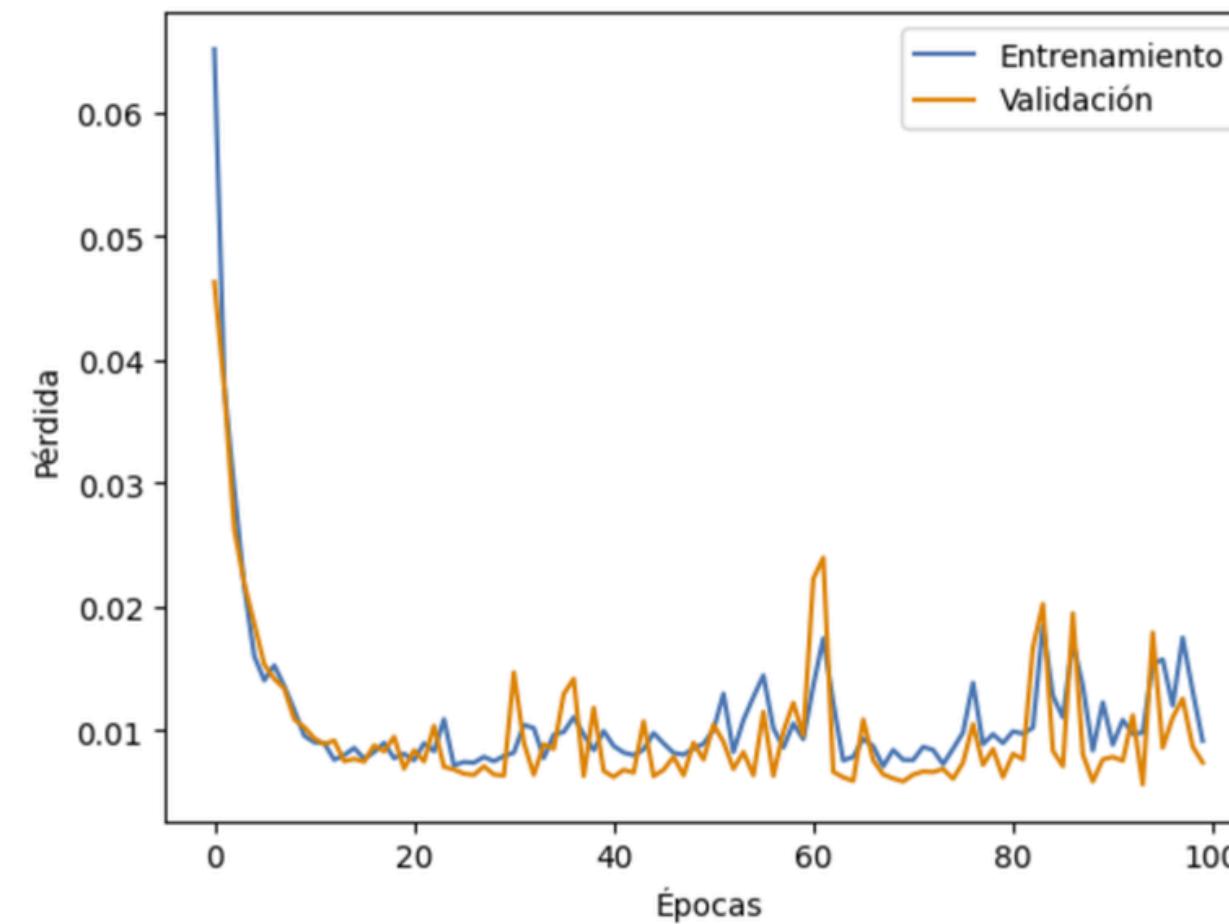
Resultados

Gráfico 1. Curvas de pérdida Adam



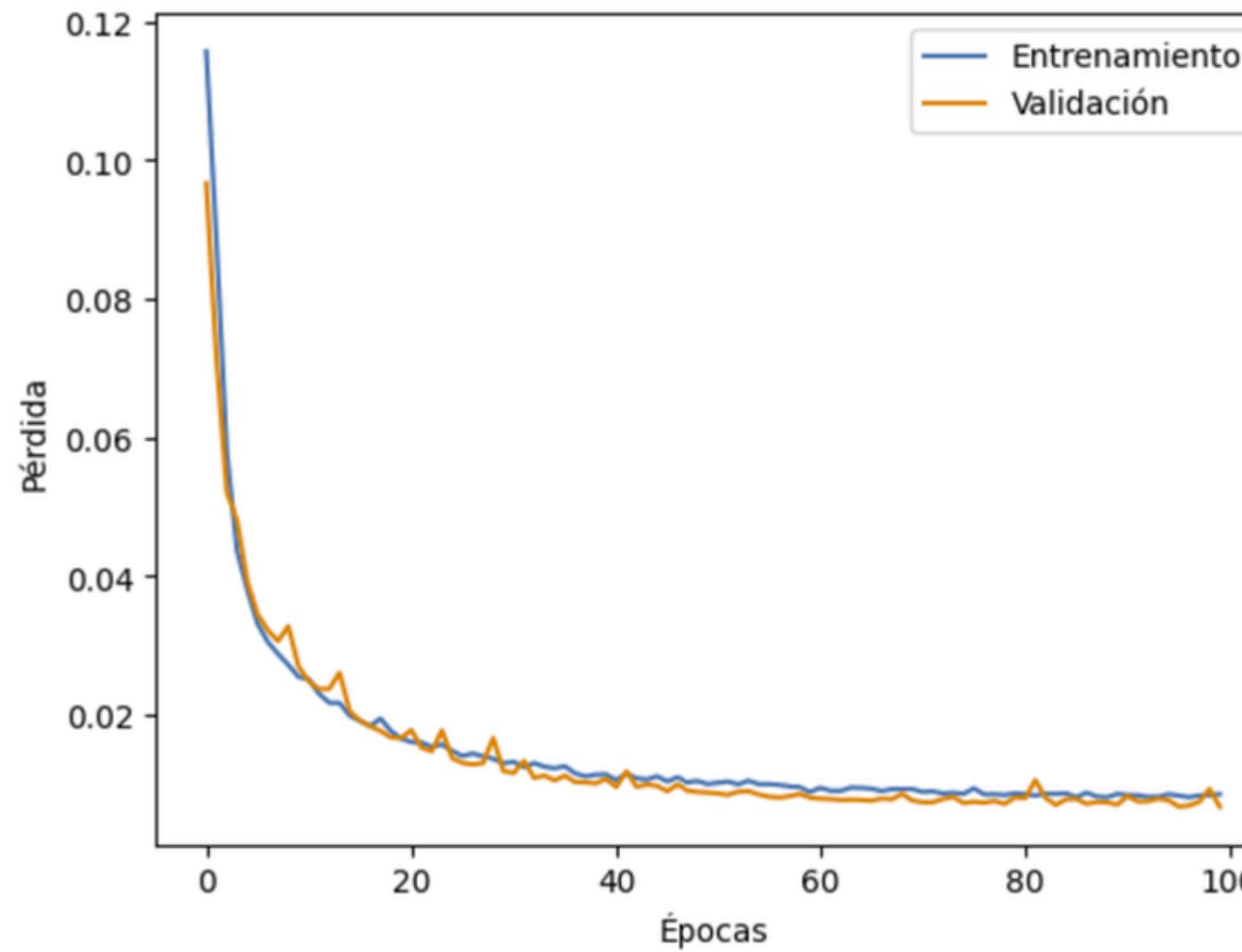
Resultados

Gráfico 2. Curvas de pérdida AdamW



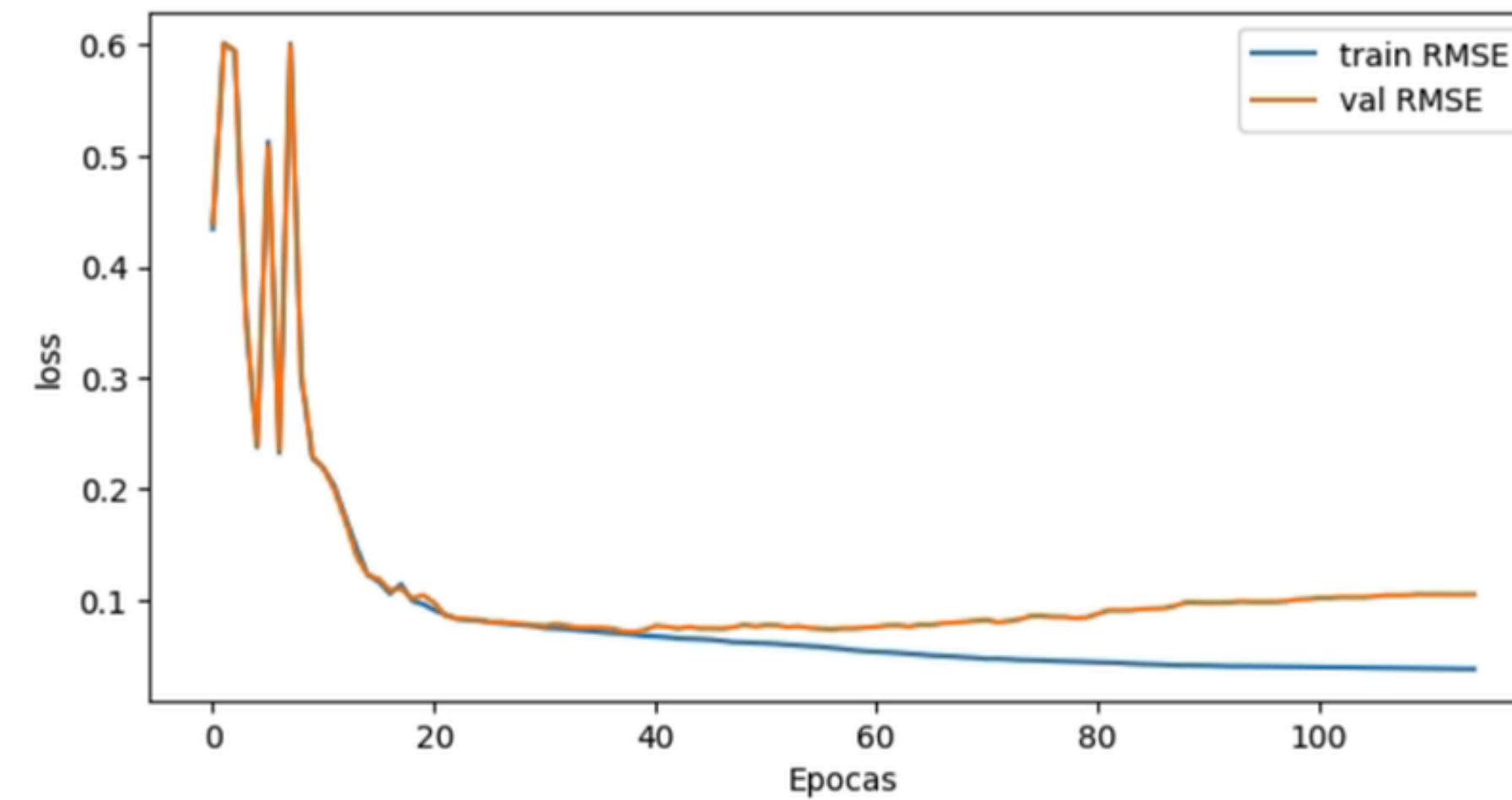
Resultados

Gráfico 3. Curvas de pérdida SGD



Resultados

Gráfico 4. Curvas de pérdida BFGS



Resultados

Tabla 1. Resultados obtenidos

Modelo	train_RMSE	val_RMSE	test_RMSE
Adam	1.732	1.733	2.062
AdamW	1.974	1.720	2.412
SGD	1.763	1.646	2.283
BFGS	8.410	8.691	9.028

Contenido

1. Introducción
2. Objetivos
3. Marco Teórico
4. Metodología
5. Resultados
6. Conclusiones



Conclusiones

El proyecto logró cumplir el objetivo general de optimizar los modelos de aprendizaje automático para predecir el rendimiento académico, demostrando que la elección del optimizador tiene un impacto directo en la precisión final del modelo. Entre las alternativas evaluadas, Adam obtuvo el mejor desempeño, alcanzando los valores más bajos de RMSE en los conjuntos de entrenamiento, validación y prueba, lo que confirma su eficacia para problemas con múltiples variables y datos limitados.

- Adam fue el optimizador con mejor precisión (RMSE más bajo) y la mejor capacidad de generalización.
- SGD obtuvo resultados competitivos, aunque con convergencia más lenta.
- AdamW aportó estabilidad pero no superó a Adam en precisión.
- BFGS tuvo el peor rendimiento, mostrando que no es adecuado para este tipo de red ni para este flujo de entrenamiento.
- En conjunto, los optimizadores adaptativos (Adam/AdamW) son los más efectivos para este problema.

Referencias bibliográficas

- Agyemang, E. F., et al. (2024). Predicting students' academic performance via machine learning algorithms: An empirical review and practical application. *Computer Engineering and Intelligent Systems*. <https://doi.org/10.7176/CEIS/15-1-09>
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2), 281–305. <https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf>
- Castrillón, O. D., Sarache, W., & Ruiz-Herrera, S. (2020). Predicción del rendimiento académico por medio de técnicas de inteligencia artificial. *Formación Universitaria*, 13(1), 93–102.
- Chen, S., & Ding, Y. (2023). A machine learning approach to predicting academic performance in Pennsylvania's schools. *Social Sciences*, 12(3), 118.
- Díaz, A. L. (2003). Factores personales, familiares y académicos que afectan al fracaso escolar en la Educación Secundaria. *Electronic Journal of Research in Educational Psychology*, 1(1), 43–66.
- Espezúa, S (s.f) Ingeniería de Atributos – Data Mining
- Falát, L., & Piscová, T. (2022). Predicting GPA of university students with supervised regression machine learning models. *Applied Sciences*, 12(17), 8403.
- Li, I. W., & Dockery, A. (2015). Does school socio-economic status influence university outcomes?. *Australian Journal of Labour Economics*, 18(1), 75–94.
- Muros Anguita, J. G. (2024). Utilización de algoritmos de machine learning/deep learning (ML/DL) para la predicción de parámetros de gestión del transporte aéreo.
- NCES. (2023). Report on the Condition of Education 2023. National Center for Education Statistics.
- OCDE. (2019). PISA 2018 Results: What students know and can do (Volume I). OCDE Publishing.
- Pedregosa, F et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- Tomasevic, N., et al. (2020). Academic achievement prediction in higher education through interpretable modeling. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0309838>.