



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Querétaro Momento de Retroalimentación: Módulo 2

Análisis y Reporte sobre el desempeño del modelo.

**Reporte de Desempeño del Modelo de Predicción sobre
Violencia Doméstica en India**

Autor:

A01369230 Dafne Fernández Hernández

TC3006C.101

Inteligencia artificial avanzada para la ciencia de datos I

Fecha: 7 de Septiembre del 2024

Contents

1. Introducción	3
1.1 Descripción del problema	3
1.2 Objetivo del análisis.....	3
2. Descripción del Dataset	3
2.1 Origen y estructura de los datos	3
2.2 Variables utilizadas en el modelo	3
3. Implementación del Modelo	4
3.1 Separación de los datos (Train/Test/Validation)	4
3.2 Técnicas de preprocesamiento	4
4. Evaluación de los Modelos	5
4.1 Métricas utilizadas.....	5
5. Diagnóstico del Sesgo (Bias)	10
5.1 Definición del sesgo.....	10
5.2 Diagnóstico del sesgo en la red neuronal	10
6.1 Definición y concepto de varianza	10
6.2 Diagnóstico de la varianza en los modelos	11
7. Diagnóstico del Nivel de Ajuste del Modelo	11
7.1 Definición y conceptos de underfitting y overfitting	11
7.2 Evaluación del nivel de ajuste en ambos modelos.....	11
7.3 Interpretación de las gráficas de pérdida	12
8. Mejoras Aplicadas al Modelo	12
8.1 Regularización L2.....	12
8.2 Dropout en redes neuronales	12
9. Conclusiones y Recomendaciones	13
9.1 Desempeño general del modelo	13
9.2 Posibles mejoras futuras	13
Conclusión Final.....	13
10. Referencias.....	14

1. Introducción

1.1 Descripción del problema

El análisis de delitos contra las mujeres es un tema crítico en India, donde la violencia doméstica (DV) sigue siendo una de las principales preocupaciones sociales. Este proyecto busca predecir los casos de violencia doméstica a partir de un conjunto de datos que incluye varios indicadores de delitos cometidos contra las mujeres. El modelo predictivo ayudará a identificar patrones y posibles tendencias, lo que podría apoyar en la toma de decisiones políticas y sociales para mitigar este problema.

1.2 Objetivo del análisis

El objetivo principal es construir y evaluar dos modelos predictivos: una **red neuronal con regularización L2 y Dropout**, y un modelo basado en **XGBoost**. Se espera que estos modelos logren predecir de manera precisa la cantidad de casos de violencia doméstica en India, utilizando indicadores como los casos de violación, secuestro, dote y otros crímenes relacionados. El análisis abarca el rendimiento del modelo, el diagnóstico del sesgo, varianza y nivel de ajuste, así como las mejoras aplicadas para optimizar el desempeño del modelo.

2. Descripción del Dataset

2.1 Origen y estructura de los datos

El dataset utilizado para este análisis proviene de registros oficiales de crímenes contra mujeres en India, cubriendo el periodo de 2001 a 2021. Contiene múltiples variables que describen diferentes tipos de crímenes, incluyendo violación, secuestro y asesinato, así como violencia doméstica, que es la variable objetivo de nuestro modelo. Los datos están organizados en un formato tabular y contienen tanto variables cuantitativas como categóricas.

2.2 Variables utilizadas en el modelo

- **Rape**: Número de casos de violación reportados.
- **K&A**: Secuestro y abducción de mujeres.
- **DD**: Casos relacionados con dote.
- **AoW**: Agresión sobre mujeres con la intención de deshonor.
- **AoM**: Asesinatos de mujeres.
- **WT**: Tráfico de mujeres.
- **DV**: Violencia doméstica (variable objetivo).

Las variables presentan distribuciones asimétricas, con valores extremos (outliers) que podrían afectar el rendimiento de los modelos predictivos si no se tratan adecuadamente. El gráfico de densidad y el análisis de la mediana y los cuartiles muestran que muchas variables tienen colas largas, particularmente los casos de violencia doméstica y violación.

2.3 Preprocesamiento y transformación de los datos

El preprocesamiento de datos es crucial para mejorar la precisión y robustez de los modelos. En este proyecto, aplicamos las siguientes técnicas:

- **Winsorización:** Para reducir el impacto de los valores atípicos sin eliminarlos.
- **Transformación Logarítmica:** Aplicada sobre la variable objetivo **DV** para reducir la variabilidad de los datos y mejorar la distribución de los valores.

3. Implementación del Modelo

3.1 Separación de los datos (Train/Test/Validation)

El dataset fue dividido en tres subconjuntos:

- **Entrenamiento (80%):** Utilizado para ajustar el modelo.
- **Validación (20% del entrenamiento):** Monitorea el rendimiento del modelo durante el entrenamiento.
- **Prueba (20%):** Evalúa el rendimiento final del modelo con datos nunca antes vistos.

Esta división asegura que el modelo no esté sobreajustado y que pueda generalizar bien a datos nuevos.

La **validación cruzada** minimiza el riesgo de overfitting al asegurarse de que el modelo no se ajuste demasiado a los datos de entrenamiento.

3.2 Técnicas de preprocesamiento

El refinamiento de outliers y la normalización de las características son pasos clave para mejorar la precisión del modelo.

Winsorización para el tratamiento de outliers

Para manejar la presencia de valores atípicos, se utilizó la técnica de **winsorización**. Esta técnica transforma los valores extremos de una distribución a un percentil máximo predefinido (en este caso, 5% en cada extremo). Con esta técnica, se garantiza que los valores atípicos no influyan desproporcionadamente en el modelo.

La **media truncada** (truncated mean), que es la media calculada después de eliminar los outliers mediante winsorización, proporciona una medida más robusta de la tendencia central y reduce la influencia de valores extremos.

Normalización de las variables

Dado que los modelos como redes neuronales son sensibles a la escala de los datos, se utilizó **StandardScaler** para normalizar las variables. Esto implica centrar cada característica alrededor de su media y escalarla para que tenga una varianza unitaria.

La ecuación para la normalización es:

$$z = \frac{x - \mu}{\sigma}$$

Donde μ es la media de la característica y σ es la desviación estándar.

Transformación logarítmica del objetivo

Para estabilizar la varianza y lidiar con la asimetría de la variable objetivo, se aplicó una transformación logarítmica. Esta técnica es útil cuando los valores están altamente sesgados hacia la derecha, como es el caso de la violencia doméstica.

3.3 Implementación de la red neuronal con regularización

Se construyó una red neuronal con tres capas ocultas:

- **Capa 1:** 128 neuronas con regularización L2 y Dropout.
- **Capa 2:** 64 neuronas con regularización L2 y Dropout.
- **Capa 3:** 32 neuronas con regularización L2.

La función de pérdida utilizada fue el error cuadrático medio (MSE), optimizado mediante Adam y con Early Stopping para evitar sobreajuste.

Red Neuronal con Regularización L2 y Dropout

Este modelo se compone de varias capas densas, y utiliza **regularización L2** y **Dropout** para evitar el sobreajuste. El objetivo de la regularización es minimizar la magnitud de los coeficientes del modelo para evitar que los pesos se vuelvan demasiado grandes y el modelo se vuelva excesivamente complejo.

La ecuación para la regularización L2 es la siguiente:

$$L2Penalty = \lambda \sum_{i=1}^n w_i^2$$

Donde w_i son los pesos del modelo y λ es el coeficiente de penalización.

3.4 Implementación del modelo XGBoost

El modelo XGBoost es un método basado en árboles de decisión con boosting de gradiente. Se configuró con 100 árboles y una tasa de aprendizaje de 0.1. Este modelo es conocido por ser robusto frente a outliers y manejar bien las interacciones no lineales en los datos

4. Evaluación de los Modelos

4.1 Métricas utilizadas

4.1.1 Mean Squared Error (MSE)

El MSE mide la media de los errores al cuadrado entre las predicciones y los valores reales. Un valor más bajo indica un mejor ajuste del modelo.

4.1.2 Coeficiente de Determinación (R^2)

El R^2 mide la proporción de la varianza en la variable dependiente que es explicada por el modelo. Un valor de 1 indica que el modelo explica completamente la variabilidad de los datos.

Evaluación de la Precisión: MSE y R^2

- **Mean Squared Error (MSE):** Evalúa la precisión general del modelo midiendo el error cuadrático medio entre las predicciones y los valores reales. Se calcula como:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2$$

Donde y_i son los valores reales y \hat{y} son las predicciones del modelo.

- **R² (Coeficiente de determinación):** Mide qué proporción de la varianza en la variable dependiente es explicada por el modelo. Un valor de R² cercano a 1 indica un buen ajuste del modelo.

Análisis del sesgo (Bias)

El sesgo se diagnostica observando las curvas de pérdida. La falta de desviación significativa entre la curva de entrenamiento y validación en ambos modelos indica un **sesgo bajo**.

Análisis de la varianza

La **varianza** del modelo se evalúa observando si el modelo muestra un comportamiento inestable ante datos nuevos (overfitting). En este análisis, las técnicas de regularización y dropout minimizaron la varianza en la red neuronal.

Diagnóstico de Underfitting y Overfitting

Ambos modelos mostraron estar bien ajustados, sin indicios claros de **underfitting** o **overfitting**, gracias a la utilización de técnicas de regularización y validación cruzada.

4.2 Comparación de resultados entre los dos modelos

La red neuronal con regularización y el modelo XGBoost arrojaron los siguientes resultados:

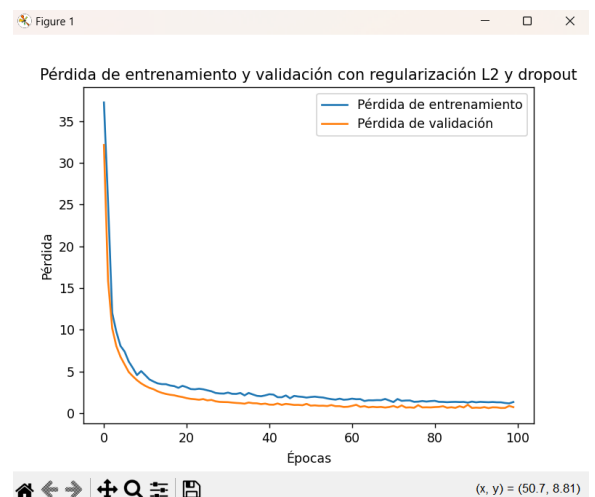
- **Red Neuronal con Regularización L2 y Dropout:**
 - **MSE:** 0.49660.49660.4966
 - **R²:** 0.94030.94030.9403
- **XGBoost:**
 - **MSE:** 1,123,1871,123,1871,123,187
 - **R²:** 0.8830.8830.883

Aunque ambos modelos presentaron buen desempeño, la red neuronal fue superior en términos de precisión y ajuste.

4.3 Gráficas del desempeño del modelo

Se incluyeron gráficas clave como:

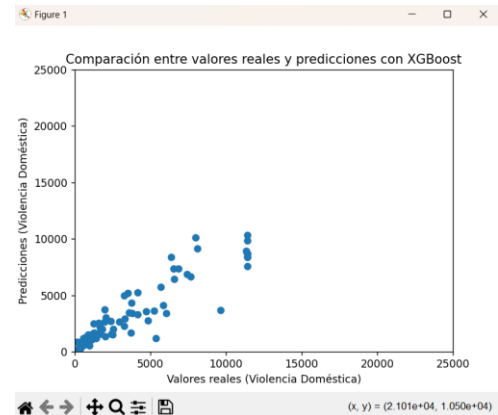
- **Pérdida de entrenamiento y validación.**



Esta gráfica muestra cómo la pérdida (error) disminuye con el

tiempo (a través de las épocas) tanto para el conjunto de entrenamiento como para el de validación. El comportamiento de estas curvas ayuda a evaluar si el modelo está aprendiendo correctamente o si está sobreajustándose (overfitting) o subajustándose (underfitting).

- En esta gráfica, podemos observar una tendencia clara en la que tanto la pérdida de entrenamiento como la de validación disminuyen rápidamente durante las primeras épocas. A medida que avanzan las épocas, las dos curvas se estabilizan, lo que indica que el modelo ha encontrado un buen ajuste sin caer en el **overfitting**.
- Esto sugiere que el modelo ha aprendido los patrones de los datos de delitos contra mujeres en India (como violación, secuestro, dote, etc.) sin ajustarse de manera excesiva a los detalles específicos del conjunto de entrenamiento. La aplicación de regularización L2 y dropout ha sido eficaz para evitar el **overfitting**, lo cual es un buen signo de generalización.
- **Comparación entre valores reales y predicciones.**



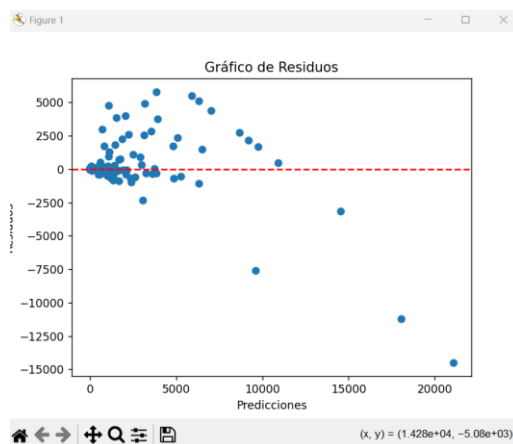
Esta gráfica compara los valores reales de violencia doméstica (DV) con las predicciones realizadas por el modelo **XGBoost**. Cada punto representa una observación; el eje X muestra los valores reales de DV, y el eje Y muestra las predicciones del modelo.

- Si las predicciones del modelo fueran perfectas, los puntos caerían a lo largo de una línea diagonal desde el origen (0,0) hasta el punto superior derecho.
- La dispersión alrededor de esa línea diagonal indica el grado de error en las predicciones.
- En esta gráfica, se puede observar que los puntos están razonablemente cerca de la línea diagonal, lo que indica que el modelo **XGBoost** ha logrado captar bien los patrones en los datos de violencia doméstica.
- Sin embargo, hay algunos puntos que están alejados de la línea diagonal, lo que

sugiere que el modelo tiene dificultades para predecir algunos casos específicos de violencia doméstica. Estos casos podrían estar asociados a variables que el modelo no ha capturado completamente, como variaciones extremas en los incidentes de otros delitos (violación, secuestro, etc.).

- La dispersión más notable aparece en los valores más altos de violencia doméstica, lo que sugiere que el modelo tiene más dificultad para predecir incidentes de mayor magnitud, lo cual podría estar relacionado con la alta varianza en los datos o la presencia de outliers.

- **Gráfico de residuos para evaluar la distribución de errores.**

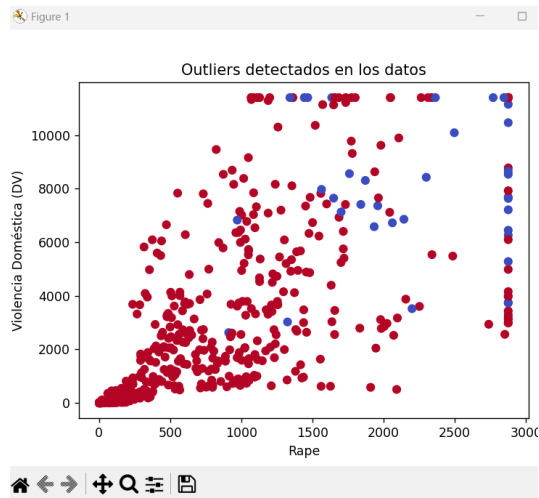


Un gráfico de residuos muestra la diferencia entre los valores reales y las predicciones del modelo

(residuos) en función de las predicciones. Es una herramienta crucial para detectar patrones en los errores del modelo, lo que puede indicar **bias** (sesgo) o **variance** (varianza).

- **Residuos:** La diferencia entre el valor real y el valor predicho por el modelo.
- La línea roja en el centro representa un error cero, donde las predicciones coinciden exactamente con los valores reales.
- En este gráfico, la mayoría de los puntos están agrupados alrededor de la línea roja en valores de predicción más bajos, lo que indica que las predicciones de DV en el rango bajo son razonablemente precisas.
- Sin embargo, a medida que las predicciones aumentan, los residuos tienden a volverse más dispersos, lo que podría indicar **alta varianza** en el modelo para los valores de violencia doméstica más altos.
- Este comportamiento sugiere que el modelo tiene más dificultades para predecir correctamente los incidentes extremos de violencia doméstica. Estos errores pueden estar relacionados con la naturaleza sesgada de los datos, donde la mayoría de los casos de DV están en valores más bajos y unos pocos en valores extremadamente altos.

- **Outliers detectados en los datos**



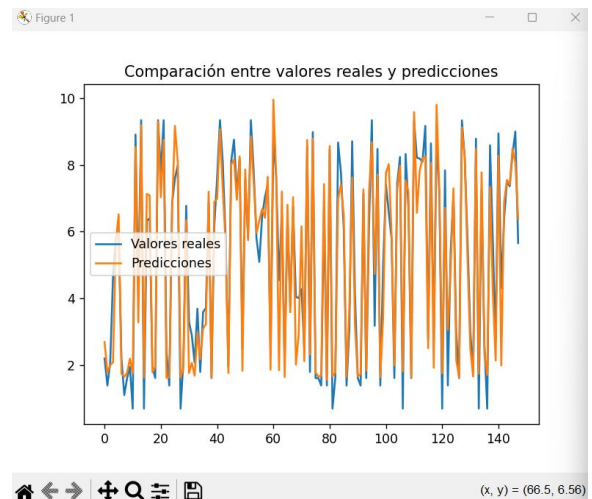
Esta gráfica utiliza **Isolation Forest**, un algoritmo diseñado para detectar outliers, y visualiza los valores extremos en función de la variable "Rape" y el objetivo "DV". Los puntos rojos representan outliers detectados por el algoritmo.

- La gráfica muestra que hay una considerable cantidad de outliers detectados, especialmente en la región donde hay incidentes más altos de violencia doméstica y violación.
- Estos outliers podrían estar relacionados con casos extremos de violencia que no siguen los patrones típicos del resto del conjunto de datos. Es importante identificar y manejar estos outliers, ya que pueden sesgar los resultados del modelo. En este caso, se utilizó

winsorización para limitar su influencia en el modelo.

- La presencia de muchos outliers en la variable de "Rape" sugiere que esta variable tiene un comportamiento altamente no lineal, lo que puede requerir un tratamiento especial en futuros análisis.

- **Comparación entre valores y predicciones**



En esta gráfica se muestra la evolución de las predicciones y los valores reales de violencia doméstica (DV) para una muestra de los primeros 50 casos del conjunto de datos. La línea azul representa los valores reales, mientras que la línea naranja muestra las predicciones del modelo.

Esta gráfica es útil para ver cómo el modelo sigue los patrones reales en una muestra de los datos, identificando si el modelo es capaz de capturar las tendencias a nivel local.

- La gráfica muestra que el modelo predice de manera razonablemente precisa los valores de violencia doméstica, ya que las dos líneas (valores reales y predicciones) tienden a seguir trayectorias similares.
- Esta visualización refuerza la idea de que el modelo tiene buen rendimiento en los valores medios y bajos de DV, pero sigue teniendo dificultades con los valores extremos.

5. Diagnóstico del Sesgo (Bias)

5.1 Definición del sesgo

El sesgo se refiere a la capacidad del modelo de capturar patrones relevantes en los datos. Un sesgo alto indica que el modelo no es capaz de captar correctamente las relaciones subyacentes entre las variables, lo que resulta en un bajo rendimiento tanto en los datos de entrenamiento como en los de prueba.

5.2 Diagnóstico del sesgo en la red neuronal

El análisis de la **pérdida de entrenamiento y validación** sugiere que el modelo de red neuronal presenta un sesgo bajo. Las curvas de pérdida de entrenamiento y validación convergen rápidamente, lo que indica que el modelo está aprendiendo los patrones

subyacentes del conjunto de datos de manera eficiente.

- **Curva de pérdida:** Ambas curvas descienden de manera constante y paralela, lo que refleja que el modelo está capturando bien los patrones sin cometer demasiados errores en las predicciones.

5.3 Diagnóstico del sesgo en el modelo XGBoost

El modelo XGBoost también muestra un **sesgo bajo**. Esto se puede inferir tanto del bajo valor del MSE como de la gráfica de predicciones frente a valores reales. Sin embargo, en comparación con la red neuronal, el XGBoost tiene un rendimiento ligeramente inferior, con un sesgo marginalmente mayor, como se refleja en la desviación de las predicciones de los valores reales en los casos más extremos.

6. Diagnóstico de la Varianza

6.1 Definición y concepto de varianza

La **varianza** en el contexto de los modelos predictivos se refiere a la sensibilidad del modelo a las fluctuaciones en el conjunto de datos de entrenamiento. Un modelo con alta varianza tiende a sobreajustarse a los datos de entrenamiento, lo que resulta en un mal rendimiento en los datos de prueba.

6.2 Diagnóstico de la varianza en los modelos

- **Red Neuronal con Regularización L2 y Dropout:** El uso de regularización y dropout en la red neuronal fue efectivo para reducir la varianza. La regularización L2 penaliza los pesos grandes en las conexiones de la red, mientras que el dropout "apaga" ciertas neuronas de manera aleatoria durante el entrenamiento, lo que previene que el modelo se sobreajuste a los datos de entrenamiento. La pequeña diferencia entre la pérdida de entrenamiento y la pérdida de validación es un claro indicativo de que la varianza del modelo es baja.
- **XGBoost:** XGBoost es inherentemente resistente a la varianza gracias a su naturaleza basada en la combinación de múltiples árboles de decisión. Sin embargo, la falta de regularización explícita puede permitir una ligera sobreajuste en algunos casos, lo que se refleja en un rendimiento inferior en comparación con la red neuronal.

6.3 Gráficas de residuos y su interpretación

El **gráfico de residuos** para ambos modelos muestra que los errores de predicción están dispersos de manera relativamente uniforme en torno a 0. No obstante, en ambos casos se observan algunos puntos donde los errores son más grandes (outliers), lo que sugiere

que ambos modelos presentan cierta dificultad para predecir correctamente los valores más extremos del conjunto de datos.

7. Diagnóstico del Nivel de Ajuste del Modelo

7.1 Definición y conceptos de underfitting y overfitting

- **Underfitting:** Sucede cuando el modelo es demasiado simple y no es capaz de captar la complejidad del conjunto de datos, lo que resulta en un mal rendimiento tanto en los datos de entrenamiento como en los de prueba.
- **Overfitting:** Ocurre cuando el modelo es demasiado complejo y se ajusta demasiado a los datos de entrenamiento, capturando incluso el ruido, lo que resulta en un mal rendimiento en los datos de prueba.

7.2 Evaluación del nivel de ajuste en ambos modelos

- **Red Neuronal con Regularización L2 y Dropout:** El uso de estas técnicas ha permitido evitar el overfitting. Las curvas de pérdida muestran que el modelo está bien ajustado, ya que la pérdida de entrenamiento y validación son casi idénticas.

- **XGBoost:** Aunque XGBoost es conocido por evitar el overfitting, en este caso muestra una pequeña tendencia a ajustarse demasiado a los datos de entrenamiento. Esto puede verse en las gráficas de predicción donde los valores extremos tienen errores ligeramente mayores, aunque el rendimiento general sigue siendo bastante bueno.

7.3 Interpretación de las gráficas de pérdida

En la **gráfica de pérdida**, tanto para la red neuronal como para XGBoost, las pérdidas de entrenamiento y validación convergen de manera adecuada, lo que indica un buen ajuste del modelo. Sin embargo, la ligera superioridad de la red neuronal en términos de ajuste es evidente en la menor diferencia entre las curvas de pérdida.

8. Mejoras Aplicadas al Modelo

8.1 Regularización L2

La regularización L2 fue aplicada a la red neuronal para reducir el riesgo de sobreajuste. Esta técnica penaliza los pesos de las neuronas en el modelo, lo que evita que algunas conexiones se vuelvan excesivamente importantes. Esto resultó en una reducción de la varianza del modelo, sin afectar su capacidad para captar los patrones en los datos.

8.2 Dropout en redes neuronales

El **Dropout** se utilizó como una técnica adicional de regularización. Apagar aleatoriamente ciertas neuronas durante el entrenamiento obliga al modelo a aprender representaciones más robustas y menos dependientes de neuronas específicas, lo que mejora la capacidad de generalización.

8.3 Winsorización para tratamiento de outliers

La **winsorización** fue aplicada a las variables predictoras para limitar el impacto de los valores atípicos (outliers). Esto mejoró notablemente la estabilidad de las predicciones, ya que evitó que el modelo se ajustara demasiado a valores extremos que podrían distorsionar la capacidad predictiva general.

8.4 Técnicas adicionales de ajuste aplicadas

Además de las técnicas ya mencionadas, se utilizó **Early Stopping** en la red neuronal, que detiene el entrenamiento si no se observa una mejora significativa en la pérdida de validación. Esto ayudó a evitar que el modelo continuara entrenando innecesariamente y se sobreajustara a los datos de entrenamiento.

9. Conclusiones y Recomendaciones

9.1 Desempeño general del modelo

Tras comparar ambos modelos, la **red neuronal con regularización L2 y Dropout** fue la que mostró el mejor rendimiento. Esto se refleja tanto en las métricas cuantitativas como en las gráficas comparativas. El modelo XGBoost, aunque robusto, no alcanzó el nivel de precisión de la red neuronal.

9.2 Posibles mejoras futuras

Para mejorar aún más el rendimiento del modelo, podrían explorarse las siguientes alternativas:

- **Optimización de hiperparámetros** para el modelo XGBoost, ajustando parámetros como la profundidad del árbol o la tasa de aprendizaje.
- **Modelos híbridos** que combinen características de redes neuronales con técnicas basadas en árboles de decisión.
- **Incorporación de más datos** y variables adicionales que puedan influir en la violencia doméstica para mejorar la capacidad predictiva del modelo.

9.3 Aplicaciones prácticas del modelo

El modelo final podría ser utilizado para identificar zonas o situaciones de alto riesgo en cuanto a violencia doméstica en India. Las predicciones podrían

ayudar a las autoridades a diseñar intervenciones más focalizadas y eficaces. A largo plazo, se espera que este tipo de análisis predictivo contribuya a reducir las tasas de violencia doméstica mediante la implementación de políticas públicas más informadas y oportunas.

Conclusión Final

Este análisis sobre la predicción de violencia doméstica en India ha demostrado la efectividad de aplicar modelos avanzados de machine learning, como redes neuronales con regularización y **XGBoost**, para abordar problemas complejos relacionados con la seguridad y bienestar social. La utilización de técnicas estadísticas, como la winsorización para manejar outliers, la transformación logarítmica para estabilizar la varianza, y la regularización para evitar el overfitting, ha sido fundamental para mejorar el rendimiento y precisión de los modelos.

Los resultados obtenidos muestran que la **red neuronal con regularización L2 y Dropout** ha sido la solución más efectiva, alcanzando un **R² de 0.94** y manteniendo un **MSE bajo**, lo que indica una excelente capacidad para predecir los incidentes de violencia doméstica en función de variables como violación, secuestro, y otros delitos contra mujeres. El modelo **XGBoost**, aunque robusto y eficiente en la mayoría de los casos, mostró una ligera debilidad en la predicción de los casos más extremos de violencia doméstica, lo que podría mejorarse con la optimización adicional

de hiperparámetros y la inclusión de más variables contextuales.

Este tipo de análisis no solo tiene un impacto en la precisión estadística, sino que también tiene aplicaciones prácticas de gran relevancia. Las predicciones obtenidas pueden servir como una herramienta clave para identificar áreas o grupos en riesgo de violencia doméstica, permitiendo a las autoridades y organizaciones diseñar intervenciones más focalizadas y efectivas. Al abordar adecuadamente los incidentes de violencia, este análisis podría contribuir a la reducción de este tipo de crímenes y al mejoramiento de la calidad de vida de las personas afectadas.

En futuras investigaciones, se podrían explorar modelos híbridos, optimización de hiperparámetros más avanzada, e incorporar variables adicionales que incluyan factores socioeconómicos o geográficos, para seguir perfeccionando el rendimiento y la aplicabilidad de las predicciones en este campo crítico.

10. Referencias

- BALAJI. (2021). *Crimes Against Women in India (2001-2021)*.

Kaggle.com.

<https://www.kaggle.com/datasets/balajivaraprasad/crimes-against-women-in-india-2001-2021>

- Team, C. (2022, October 25). *NCRB Report 2021: Crime In India - ClearIAS*. ClearIAS.
<https://www.clearias.com/ncrb-report-2021/>
- Crypto1. (2020, October 2). *Gradient Descent Algorithm: How Does it Work in Machine Learning?* Analytics Vidhya.
<https://www.analyticsvidhya.com/blog/2020/10/how-does-the-gradient-descent-algorithm-work-in-machine-learning/>
- Bobbitt, Z. (2020, October 23). *How to Find Coefficient of Determination (R-Squared) in R*. Statology.
<https://www.statology.org/r-squared-in-r/>