



## Informe Escrito

### Introducción a las Redes neuronales y Deep Learning

«*Predicción de Abandono Estudiantil mediante implementación de FFNN en PyTorch*»

Bruno Morici - Ing. Civil Informática: 202373555-8

Juan Pablo Fuenzalida Fredes - Ing. Civil Informática: 202373102-1

Segundo Semestre 2025

## Índice

<b>1. Problema</b>	<b>2</b>
<b>2. Solución</b>	<b>3</b>
<b>3. Análisis de Resultados</b>	<b>7</b>
<b>4. Análisis de Sesgos e Interpretabilidad del Modelo FFNN</b>	<b>9</b>
<b>5. Conclusiones y aplicaciones futuras</b>	<b>12</b>
<b>6. Bibliografía y Anexos (Repositorio y Presentación)</b>	<b>13</b>



## 1. Problema

Según el último **Informe de Retención Estudiantil de Educación Superior** del *Sistema de Información de la Educación Superior (SIES) de Chile (2024)*, la tasa de **retención de alumnos de primer año** en Universidades, Institutos Profesionales y Centros de Formación Técnica alcanzó un **77%**. Esto implica que el **23 % de los estudiantes desertan** tan solo en el primer año. Además, informes históricos revelan que, a nivel global, aproximadamente solo el **53 % de los alumnos matriculados** logra alcanzar el quinto año de su plan de estudios. Esta persistente y **elevada tasa de deserción universitaria** constituye un **problema latente** con implicaciones significativas, afectando negativamente tanto el desarrollo profesional individual de los estudiantes como el desarrollo del país, al limitar el flujo de *profesionales calificados y certificados*.

En respuesta a esta **problemática observada** y dentro del contexto académico universitario, este proyecto tiene como objetivo desarrollar una **herramienta predictiva** para *mitigar las cifras de deserción*. Buscamos **identificar de manera proactiva y oportuna** a los alumnos que presenten un **alto riesgo de abandono** para implementar **estrategias de apoyo focalizadas** ya sean de índole *socioeconómica, académica o demográfica*, generando una mejora tangible en el ambiente universitario y aumentando las *tasas de graduación*.

Para alcanzar este objetivo, se procederá a la construcción y entrenamiento de una **Red Neuronal Artificial** capaz de *analizar patrones* en las características de los alumnos y **predecir su tendencia** a desertar o a graduarse. La Red Neuronal será diseñada y entrenada utilizando el **dataset público** denominado "*Predict Students' Dropout and Academic Success*". Este conjunto de datos abarca registros de estudiantes **graduados, matriculados y desertores**, incluyendo una *amplia gama de variables y características* de tipo académico, socioeconómico y demográfico, las cuales servirán como **inputs fundamentales** para el modelo, permitiendo el *aprendizaje de los patrones subyacentes* que correlacionan dichas variables con el **éxito o el abandono académico**.



## 2. Solución

### 2.1. Metodología y Obtención del Modelo de Clasificación

El proyecto tuvo como objetivo principal la predicción del abandono estudiantil mediante un modelo de clasificación basado en una Red Neuronal *Feed-Forward* (FFNN), implementado en **PyTorch**. El procedimiento de trabajo se estructuró en seis fases clave: elección y limpieza del *dataset*, análisis exploratorio y preparación, creación y configuración del modelo, entrenamiento, exportación para una aplicación demostrativa y, finalmente, un análisis de sesgos en base a características propias de los alumnos.

#### 2.1.1. Características del Dataset

El conjunto de datos inicial comprende **4424** instancias y **36** variables o *features*. Estas variables se agrupan en tres categorías principales que influyen en la deserción, he aquí algunas:

- **Registro Académico:** Incluye rendimiento, unidades curriculares aprobadas e inscritas, y notas promedio por semestre (*e.g.*, `curricular_units_2nd_sem_grade`).
- **Características Demográficas:** Información como género (*e.g.*, `gender_0`) y edad al momento de la inscripción (*e.g.*, `age_at_enrollment`).
- **Factores Socio-Económicos:** Consideraciones como tener o no beca (*e.g.*, `scholarship_holder_1`) y estado de las tasas de matrícula.

Los datos corresponden a alumnos registrados hasta su segundo año de estudio, independiente de si se tituló o no.

#### 2.1.2. Proceso de Limpieza y Preparación

La calidad de los datos es fundamental para el rendimiento del modelo, por lo que se realizaron las siguientes operaciones:

- **Eliminación de Nulos y Estandarización:** Se confirmó la ausencia de valores nulos y se estandarizaron los nombres de las variables.
- **Normalización Estándar:** Se aplicó la normalización para escalar las *features* numéricas, asegurando que ninguna variable dominara el proceso de entrenamiento debido a su magnitud.
- **Eliminación de Outliers:** Se utilizó el método de la puntuación Z (*Z-Score*) para la detección y eliminación de valores atípicos, mejorando la robustez del modelo.

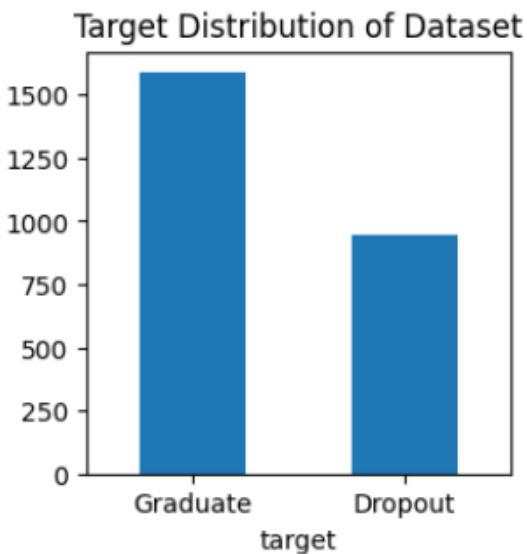


Figura 1: Distribucion del Dataset

- **Reducción a Clasificación Binaria:** La variable objetivo (`target`) se redujo de tres clases (`Graduate`, `Dropout`, `Enrolled`) a dos, **eliminando** la clase `Enrolled` (alumnos activos). Esto centró la predicción en el resultado final: graduación o deserción.
- **One-Hot Encoding:** Las variables categóricas o binarias se codificaron para ser interpretadas correctamente por la red neuronal.

### 2.1.3. Balanceo de Clases

La distribución de la variable objetivo mostró un desbalance de clases (más `Graduate` que `Dropout` al inicio). Tras eliminar `Enrolled`, el desbalance persistió, por lo que se aplicó un Balanceo Artificial mediante la Modificación de la Función Objetivo (*Weighted Cross-Entropy Loss*). Esta técnica asigna un mayor peso a los errores de clasificación de la clase minoritaria (`Dropout`) durante el entrenamiento, obligando al modelo a prestarles más atención. Se reservaron los datos de la clase `Enrolled` para el testing posterior.

### 2.1.4. Selección de Características y PCA

Tras un análisis de correlación con la variable `target`, se seleccionaron las **16 features** más relevantes, tanto positivas como negativamente correlacionadas. Posteriormente, se aplicó el Análisis de Componentes Principales (PCA). Esta técnica se utiliza para reducir la dimensionalidad de los datos sin perder información significativa, simplificando el modelo y acelerando el entrenamiento.

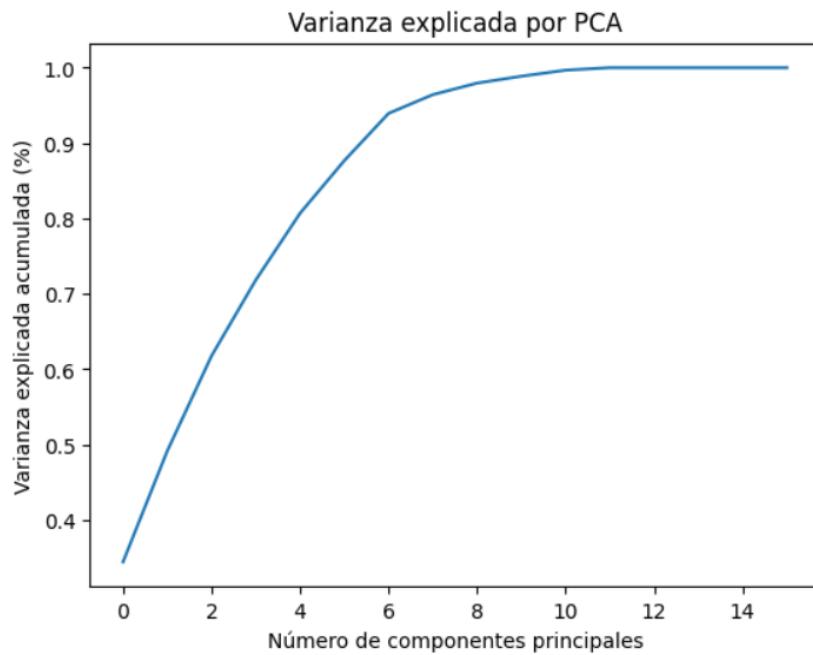


Figura 2: Varianza del PCA

Es fundamental tener esto en claro, debido a que la Red Neuronal no se entrega directamente con las características de los alumnos, si no más bien con un PCA o combinación lineal de las mismas. Por lo que luego cuando se realice el análisis de sesgo, es necesario recomponer estas combinaciones lineales para extraer la relevancia de cada característica en el modelo.

## 2.2. Arquitectura y Entrenamiento de la Red Neuronal

### 2.2.1. Modelo *Feed-Forward*

Se diseñó un modelo de Red Neuronal *Feed-Forward* con una arquitectura de 4 capas ocultas, caracterizada por un diseño cónico (disminución de neuronas) y el uso de técnicas de regularización avanzadas.

- **Capa de Entrada:** 16 *features*.
- **Capas Ocultas:** Con tamaños [128, 64, 32, 16].
- **Función de Activación:** *ReLU* (Rectified Linear Unit), elegida por su eficiencia computacional y capacidad para introducir no linealidad en el modelo.
- **Regularización (*Dropout*):** Se aplicó un *Dropout* progresivo de [0.4, 0.3, 0.2, 0.1]. Esta técnica



desactiva neuronas aleatoriamente durante el entrenamiento, previniendo el sobreajuste (*overfitting*).

- **Capa de Salida:** 2 neuronas (para Dropout y Graduate) con función *Softmax* para generar probabilidades de pertenencia a cada clase.

### 2.2.2. Hiperparámetros y Estrategia de Entrenamiento

El modelo fue entrenado utilizando los siguientes hiperparámetros y estrategias:

- **Optimizador:** Adam, un algoritmo popular para el descenso de gradiente.
- **Tasa de Aprendizaje (*Learning Rate*):** 0,0005.
- **Batch Size:** 32.
- **Weight Decay (Regularización L2):** 0,0005.
- **Learning Rate Scheduler:** *ReduceLROnPlateau*, que adapta dinámicamente el *learning rate* si la pérdida de validación deja de mejorar, optimizando la convergencia.
- **Early Stopping:** Mecanismo para detener el entrenamiento cuando la pérdida de validación no mejora tras un número predefinido de épocas, lo que previene el sobreajuste y guarda el mejor modelo encontrado.

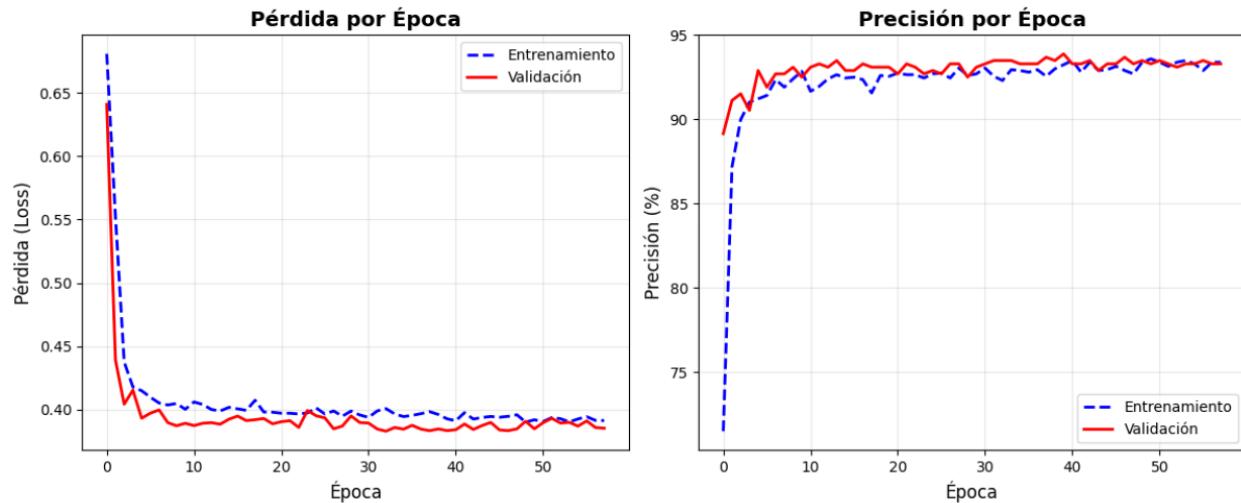


Figura 3: Pérdida y Precisión por Época

### 3. Análisis de Resultados

#### 3.1. Resultados y Métricas

El entrenamiento se detuvo en la época 72 por la activación del *Early Stopping*, siendo el mejor modelo guardado en la época 47. Los gráficos de pérdida y precisión por época mostraron una convergencia rápida y estable, sin signos de sobreajuste significativo.

##### 3.1.1. Métricas de Rendimiento

La evaluación final sobre el conjunto de prueba arrojó un alto rendimiento en todas las métricas:

- **Precisión (*Accuracy*): 93,89 %.**
- ***Precision* Promedio Ponderada: 93,93 %.**
- ***Recall* (Predicciones Correctas por Clase): 93,89 %.**
- **F1-Score: 93,84 %**

##### 3.1.2. Análisis de la Matriz de Confusión

La matriz de confusión proporciona una visualización detallada del desempeño por clase.

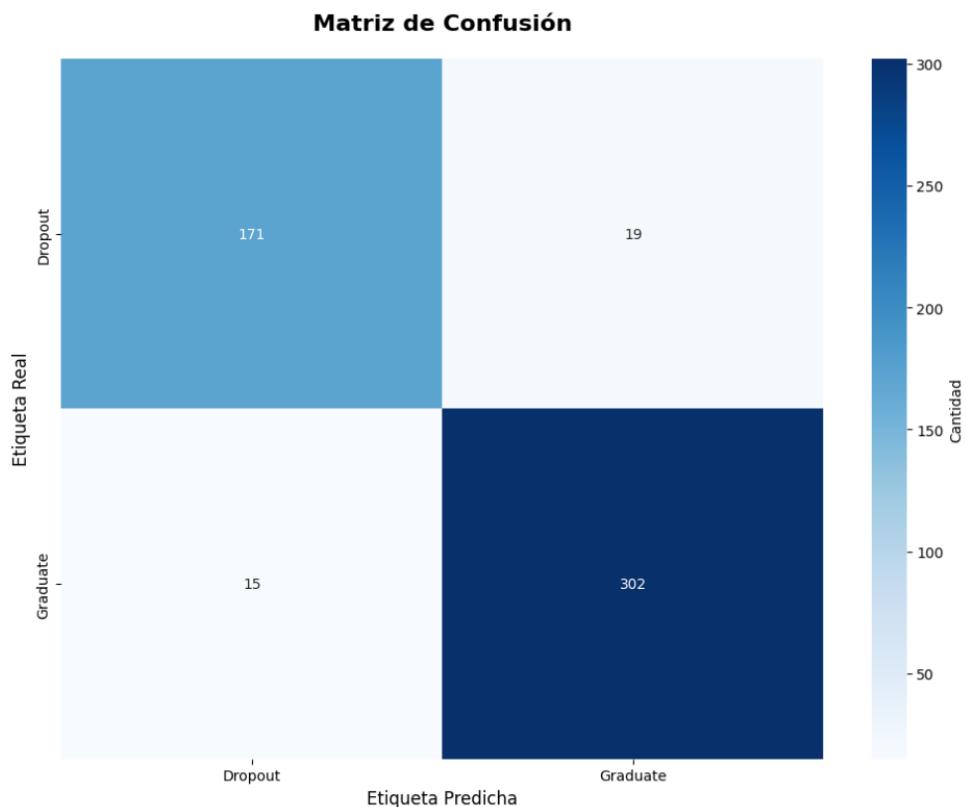


Figura 4: Matriz de Confusión

- **168** alumnos Desertores fueron clasificados correctamente, y **22** fueron falsamente clasificados como Graduados.
- **308** alumnos Graduados fueron clasificados correctamente, y **9** fueron clasificados falsamente como Desertores (falsos positivos).

El alto rendimiento del modelo, con una precisión general cercana al 94 %, confirma su utilidad para identificar a estudiantes en riesgo de deserción (Clase Dropout) y a aquellos con alta probabilidad de éxito (Graduate).



## 4. Análisis de Sesgos e Interpretabilidad del Modelo FFNN

Para asegurar la confiabilidad y detectar posibles sesgos en los modelos de predicción de abandono, se implementó una metodología de **Inteligencia Artificial Explicable (XAI)** centrada en la librería SHAP (SHapley Additive exPlanations).

### 4.1. Metodología de Creación y Evaluación

Se construyó una Red Neuronal de Propagación hacia Adelante (**FFNN**) con cuatro capas ocultas. La metodología de análisis se basó en dos modelos secuenciales con diferentes conjuntos de *features*.

#### 4.1.1. Categorización de Features

Las variables de entrada fueron agrupadas para entender el origen de la influencia:

- **Factores Académicos:** Notas, unidades aprobadas/matriculadas (`curricular_units,admission_grade`)
- **Factores Financieros/Deuda:** Estado de matrícula, condición de deudor (`debtor`), becas.
- **Factores Sociodemográficos/Contextuales:** Género (`gender`), edad, estado civil, modo de aplicación.

#### 4.1.2. El Rol Crítico del PCA y sus Implicaciones

Antes de ingresar a la FFNN, se aplicó el **Análisis de Componentes Principales (PCA)** para reducir la dimensionalidad de las *features* codificadas (que eran  $> 22$ ) a **16 Componentes Principales** ( $PC_1$  a  $PC_{16}$ ).

- **Base del Modelo:** La red neuronal no ve las *features* originales, sino una **combinación lineal** de ellas.
- **Implicación:** La importancia se asigna a estas combinaciones ( $PC_i$ ), y no directamente a las variables originales.

#### 4.1.3. Estabilización del Análisis con SHAP

Debido a la naturaleza de la FFNN profunda, las transformaciones no lineales (ReLU) y las capas de Dropout dificultaron el cálculo de los gradientes internos que requiere el `shap.DeepExplainer`.

\* **Problema:** La incompatibilidad con las capas internas provocó errores de dimensionalidad (`RuntimeError`).



\* **Solución:** Se implementó un modelo temporal minimalista (`MinimalLogitModel`) que **omitió las capas de Dropout y Softmax** durante el cálculo de los valores SHAP. Esto estabilizó la retropropagación de gradientes y permitió la explicación del modelo.

#### 4.1.4. Explicación Simple de SHAP

SHAP (SHapley Additive exPlanations) es una técnica basada en la teoría de juegos que asigna un valor de importancia justo a cada *feature* o componente.

- **Valor SHAP:** Representa la contribución de un  $PC_i$  a la diferencia entre la predicción y la predicción base.
- **Regla de Interpretación:** Un  $PC_i$  alto en valor absoluto es crucial. Su signo (positivo o negativo) determina si ese componente **fomenta la predicción de Abandono (Clase 1)** o de Graduación.

### 4.2. Resultados Finales y Conclusiones sobre Sesgos

Se realizó un análisis de composición para los  $PC$  con mayor valor SHAP para determinar su vínculo con las *features* originales.

#### 4.2.1. Características Liderando la Predicción

El análisis SHAP/PCA reveló que la predicción está dominada por solo dos factores, consistentes en ambos modelos:

Cuadro 1: Factores Dominantes en la Predicción de Abandono		
Factor Principal	Componente	Peso Dominante
Factor Académico	$PC_1$	$\approx 0,32$ ( <code>curricular_units, grade</code> )
Factor Financiero	$PC_2$	$\approx 0,49$ ( <code>debtor, tuition_fees_up_to_date</code> )

- **Riesgo Financiero ( $PC_2$ ):** Un alto valor en este PC (impulsado por ser deudor y tener la matrícula no al día) **aumenta la probabilidad de abandono** (contribución SHAP positiva).
- **Rendimiento Académico ( $PC_1$ ):** Un alto valor en este PC (impulsado por buenas notas y muchas unidades aprobadas) **disminuye la probabilidad de abandono** (contribución SHAP negativa).



#### 4.2.2. Evaluación del Sesgo de Género

El modelo fue reentrenado incluyendo las variables de género (`gender_0`, `gender_1`) para evaluar el sesgo directo.

- **Peso Máximo del Género:** El análisis de composición del *PC* más importante ( $PC_1$ ) mostró que el peso de la variable de género más influyente fue de **0.1727**.
- **Prueba de Paridad:** La prueba de predicción directa (cambiando solo el género de un estudiante con perfil fijo) mostró que el cambio **no alteró la predicción final** (`Graduate` o `Dropout`) y tuvo un impacto mínimo ( $\leq 5$  puntos porcentuales) en las probabilidades predichas.



## 5. Conclusiones y aplicaciones futuras

El proyecto cumplió exitosamente su objetivo principal: construir un modelo predictivo (**FFNN** con  $\approx 93\%$  de precisión) y someterlo a un riguroso análisis de interpretabilidad (**XAI**) mediante SHAP. Logrando descomponer las predicciones complejas del modelo PCA y FFNN, determinando la jerarquía de los factores que realmente impulsan el riesgo de abandono. Además se validó la integración del modelo entrenado (.pt) en una aplicación web simple (Django), demostrando la viabilidad de un **evaluador de riesgo** en sistemas de gestión universitaria o aulas virtuales.

El análisis SHAP/PCA proporcionó una jerarquía clara y accionable de los factores de riesgo y posterior toma de decisiones en base a estos, siendo el Riesgo Financiero el factor **más influyente**. El modelo lo utiliza para identificar un alto riesgo cuando el estudiante es **deudor** y **no tiene la matrícula al día**. El Rendimiento Académico es el segundo factor más fuerte. Un alto valor en este (buenas notas y aprobaciones) actúa como un poderoso **factor protector** contra el abandono. El modelo facilita un enfoque **proactivo**, permitiendo a la universidad priorizar recursos e intervenciones focalizadas (ej., asistencia financiera, tutorías obligatorias) en lugar de un enfoque reactivo.

El análisis de sesgo directo, centrado en la variable `gender`, demostró la equidad del modelo, el análisis SHAP reveló que la contribución del **género** al modelo fue mínima ( $\approx 0,17$ ), quedando muy por debajo de los factores de rendimiento y financieros. Esto valida que las listas de “alumnos de riesgo” generadas por el modelo se basan en **factores objetivos** y neutros al género. Esto es fundamental para la **confianza** en el sistema y para evitar la discriminación en la asignación de ayuda.

Finalmente el modelo se puede integrar directamente en los Sistemas de Información Estudiantil para generar puntuaciones de riesgo en tiempo real, utilizando el output de SHAP para generar explicaciones automáticas, transformando el modelo en una herramienta transparente, que permite que la intervención se base en una comprensión clara de las causas subyacentes de su riesgo académico.



## 6. Bibliografía y Anexos (Repositorio y Presentación)

- **Predict Students' Dropout and Academic Success Dataset** - <https://archive.ics.uci.edu/dataset/697/predict+students+dropout+and+academic+success>
- **Investigación revela alto nivel de abandono de estudios superiores** - <https://www.latercera.com/diario-impreso/investigacion-revela-alto-nivel-de-abandono-de-estudios-superiores-50-se-re>
- **RETENCIÓN DE 1er AÑO DE PREGRADO** [https://mifuturo.cl/wp-content/uploads/2025/08/Informe\\_Retencion\\_2025\\_SIES.pdf](https://mifuturo.cl/wp-content/uploads/2025/08/Informe_Retencion_2025_SIES.pdf)
- **Predict Students' Dropout and Academic Success** — Repositorio GitHub: <https://github.com/alecnco/Predict-students-dropout-and-academic-success>
- **Repositorio con código fuente del proyecto** — Repositorio GitHub: [https://github.com/moriciii3/proyecto\\_redes\\_neuronales](https://github.com/moriciii3/proyecto_redes_neuronales)
- **Enlace a video de presentación del proyecto** — Enlace a Drive: <https://drive.google.com/file/d/11N-FEaKh48GAUbOOK6PAUZ77ajCHGzsE/view?usp=sharing>