



UNIVERSIDAD PERUANA UNIÓN

Facultad de Ingeniería y Arquitectura

# Predicción de la Deserción Académica mediante Redes Neuronales

Proyecto de Inteligencia Artificial

**Autores:** Javier Tello, Sebastian Chinchay, Verónica Vergara, Pamela Vallejos

**Docente:** Guillermo Mamani Apaza

**Curso:** Inteligencia Artificial

**Fecha:** 5 de noviembre de 2025

Lima, Perú — 2025

# Capítulo 1

## Etapa 1: Inteligencia

### 1.1. Contexto del Caso

En la Universidad Peruana Unión (UPeU), se identificaron variaciones considerables en el rendimiento académico y en la continuidad de los estudiantes. Algunos logran un desempeño satisfactorio, mientras que otros enfrentan dificultades que los llevan a la deserción. Para abordar esta problemática, se diseñó un modelo predictivo que permita identificar el riesgo de deserción, con base en variables representativas del bienestar integral del estudiante.

### 1.2. Objetivo

El objetivo general es **predecir el riesgo de deserción universitaria** mediante redes neuronales artificiales, integrando factores académicos, personales, de salud y económicos.

### 1.3. Variables Consideradas

- **Hábitos y salud:** Sueño\_horas, Actividad\_física, Alimentación, Estilo\_de\_vida.
- **Personales y emocionales:** Estrés\_académico, Apoyo\_familiar, Bienestar.
- **Académicas:** Asistencia, Horas\_estudio, Interés\_académico, Rendimiento\_académico, Promedio\_acumulado.
- **Socioeconómicas:** Carga\_laboral, Beca, Deudor.
- **Variable objetivo:** Riesgo\_deserción.

# Capítulo 2

## Etapa 2: Proceso de Negocio

### 2.1. Obtención de Datos

Los datos fueron generados a partir de un modelo probabilístico inicial basado en redes bayesianas, posteriormente transformados en un conjunto sintético de 2000 registros mediante.

### 2.2. Preprocesamiento

Las variables categóricas fueron convertidas a formato numérico mediante One-Hot Encoding, y la variable objetivo fue codificada con LabelEncoder. El dataset fue dividido en un 80 % para entrenamiento y un 20 % para prueba, asegurando la representatividad de las clases de riesgo.

### 2.3. Modelo de Negocio

El modelo predice el nivel de riesgo de deserción para cada estudiante: *Sin riesgo*, *Riesgo leve*, *Riesgo moderado*, *Riesgo alto o Riesgo crítico*, permitiendo priorizar intervenciones preventivas en el entorno institucional.

# Capítulo 3

## Etapa 3: Tecnología

### 3.1. Diseño del Modelo Neuronal

Se implementó un modelo **Sequential** con la biblioteca **TensorFlow Keras**, con la siguiente arquitectura:

- Capa de entrada: 128 neuronas, activación ReLU.
- Capa oculta: 64 neuronas, activación ReLU.
- Capa de salida: activación Softmax (clasificación multiclas).

### 3.2. Entrenamiento

El modelo fue entrenado con 40 épocas, tamaño de lote de 32 y una validación interna del 20 %. La función de pérdida empleada fue *categorical crossentropy* y el optimizador *Adam*.

### 3.3. Resultados

El modelo alcanzó una precisión global del 99.25 %. La Tabla 1 resume las métricas obtenidas.

Clase	Precision	Recall	F1-Score	Support
Riesgo Alto	1.00	1.00	1.00	94
Riesgo Crítico	0.00	0.00	0.00	1
Riesgo Leve	0.99	1.00	0.99	84
Riesgo Moderado	1.00	1.00	1.00	220
Sin Riesgo	0.00	0.00	0.00	1
<b>Accuracy</b>			0.995	
<b>Macro Average</b>	0.60	0.60	0.60	400
<b>Weighted Average</b>	0.99	0.99	0.99	400

Tabla 1: Reporte de clasificación del modelo de predicción de deserción.

Las curvas de pérdida y precisión mostraron una convergencia estable, indicando que el modelo generaliza adecuadamente sin sobreajuste.

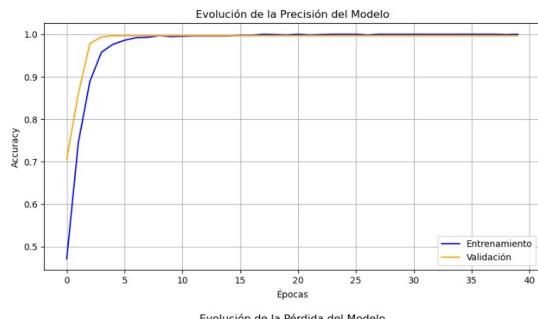


Figura 3.1: Evolución de la precisión del modelo.



Figura 3.2: Evolución de la pérdida del modelo.

### 3.4. Arquitectura de Red Neuronal

Tabla 2: Resumen de la arquitectura de la red neuronal secuencial.

Capa (tipo)	Forma de salida (Output Shape)	Parámetros
Dense (dense)	(None, 128)	6,144
Dropout (dropout)	(None, 128)	0
Dense (dense_1)	(None, 64)	8,256
Dense (dense_2)	(None, 5)	325
<b>Total de parámetros</b>		44,177 (172.57 KB)
<b>Parámetros entrenables</b>		14,725 (57.52 KB)
<b>Parámetros no entrenables</b>		0 (0.00 B)
<b>Parámetros del optimizador</b>		29,452 (115.05 KB)

Arquitectura de la Red Neuronal para Predicción de Deserción

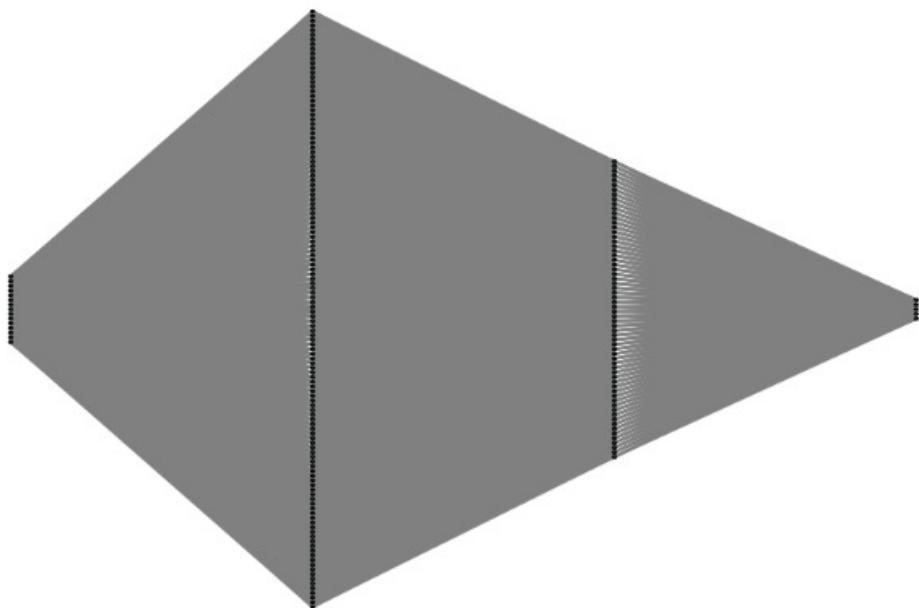


Figura 3.3: Arquitectura de la red neuronal utilizada para la predicción de deserción académica.

# Capítulo 4

## Etapa 4: Despliegue

### 4.1. Guardado del Modelo

El modelo final fue persistido en formato .h5 junto con los objetos de codificación.

```
model.save("modelo_riesgo_desercion.h5")
joblib.dump(le, "label_encoder.pkl")
joblib.dump(X_encoded.columns.tolist(), "columnas_X.pkl")
```

### 4.2. Predicción con Nuevos Casos

Se desarrolló una función de inferencia que permite ingresar las características de un nuevo estudiante y obtener su clasificación de riesgo.

```
pred = model.predict(nuevo_encoded)
clase_predicha = le.inverse_transform([np.argmax(pred)])
print("Riesgo predicho:", clase_predicha[0])
```

### 4.3. Impacto Institucional

El despliegue del modelo permitirá establecer un sistema de alerta temprana institucional, capaz de detectar y atender de forma oportuna los casos con alto riesgo de deserción.

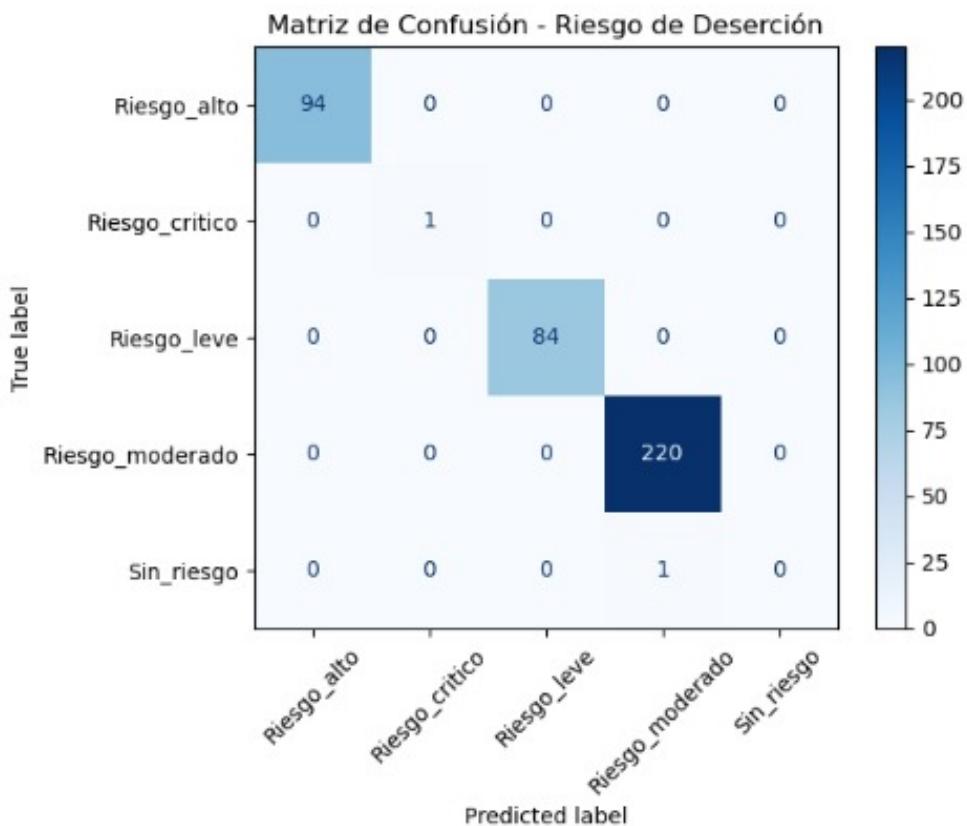


Figura 4.1: Matriz de confusión neuronal para la predicción de deserción académica.

#### 4.4. Toma de decisiones basada en el modelo

El sistema de predicción de deserción proporciona una clasificación automática del riesgo estudiantil, permitiendo definir estrategias de intervención según el nivel identificado. La Tabla ?? resume las recomendaciones institucionales por nivel de riesgo.

<b>Nivel de riesgo</b>	<b>Interpretación</b>	<b>Acciones sugeridas</b>
Sin riesgo	Estudiante estable	Seguimiento regular y refuerzo positivo.
Riesgo leve	Señales tempranas	Tutoría preventiva y monitoreo de asistencia.
Riesgo moderado	Factores combinados	Consejería académica y apoyo emocional.
Riesgo alto	Alta probabilidad de abandono	Intervención conjunta con bienestar.
Riesgo crítico	Riesgo inminente	Activación de protocolo de retención urgente o apoyo personalizado.

Tabla 3: Estrategias de intervención según nivel de riesgo.

# Conclusiones

El uso de redes neuronales artificiales demostró ser una herramienta eficaz para predecir el riesgo de deserción estudiantil. El modelo desarrollado combina datos de diversas dimensiones —académica, personal y socioeconómica— obteniendo un alto nivel de precisión y capacidad de generalización.

La implementación futura en entornos reales permitirá a la UPeU desarrollar estrategias de retención basadas en evidencia y aprendizaje automático.

# Bibliografía

- [1] W. Ha, L. Ma, Y. Cao, Q. Feng, y S. Bu, “The effects of class attendance on academic performance: Evidence from synchronous courses during Covid-19 at a Chinese research university,” *International Journal of Educational Development*, vol. 104, p. 102952, 2024.
- [2] F. Chollet, “Keras: The Python Deep Learning library,” *Github repository*, 2015. [Online]. Disponible en: <https://keras.io/>
- [3] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [4] Universidad Peruana Unión, “Plan Estratégico Institucional 2025–2030,” UPeU, 2024.