



U N I V E R S I D A D
Panamericana

Introducción al Aprendizaje Automático: Clasificación Binaria de Estrellas y Galaxias

Juan Pablo de Alba Tamayo

June 2025

Índice

1. Introducción	4
2. Definición del Proyecto	5
2.1. Objetivo	5
2.2. Expectativas	5
3. Exploración de Datos	6
3.1. Descripción de Datos	6
3.2. Distribución de Datos	6
3.3. Correlación de Datos	6
4. Preparación de Datos	7
4.1. Limpieza	7
4.2. Selección de Variables	7
4.3. Preprocesado (Logaritmo, RobustScaler)	7
5. Modelos de Machine Learning	8
5.1. Random Forest	8
5.2. Logistic Regression	8
5.3. SVM	8
5.4. KNN	8
6. Resultados	9
6.1. Comparación de Modelos	9
6.2. RandomizedSearch	9
6.3. Modelo Final	9
6.4. Probar con Dataset de Test	9
7. Conclusión	10

1. Introducción

2. Definición del Proyecto

2.1. Objetivo

2.2. Expectativas

3. Exploración de Datos

3.1. Descripción de Datos

3.2. Distribución de Datos

3.3. Correlación de Datos

4. Preparación de Datos

4.1. Limpieza

4.2. Selección de Variables

4.3. Preprocesado (Logaritmo, RobustScaler)

5. Modelos de Machine Learning

5.1. Random Forest

5.2. Logistic Regression

5.3. SVM

5.4. KNN

6. Resultados

6.1. Comparación de Modelos

6.2. RandomizedSearch

6.3. Modelo Final

6.4. Probar con Dataset de Test

7. Conclusión

Referencias