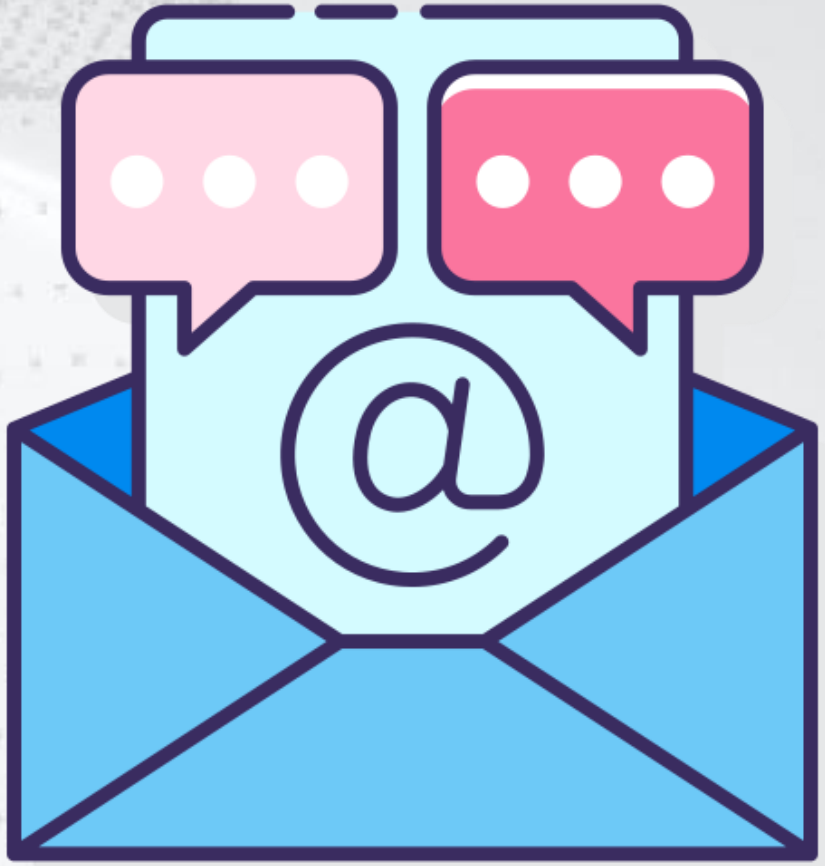
KIBERNUM
IT Academy

Módulo 5
Clase 11A

Introducción a la Ciencia de Datos

Tuning Modelo Email Spam



KIBERNUM

Caso Email Spam

- El objetivo de esta actividad aplicar las técnicas para la optimización del modelo de clasificación.

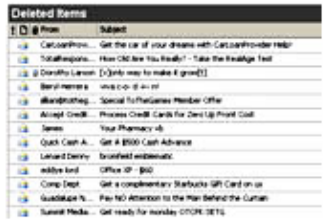
UCI Machine Learning Repository

Center for Machine Learning and Intelligent Systems

Spambase Data Set

Download: Data Folder, Data Set Description

Abstract: Classifying Email as Spam or Non-Spam



Data Set Characteristics:	Multivariate	**Number of Instances:**	4601	**Area:**	Computer
Attribute Characteristics:	Integer, Real	**Number of Attributes:**	57	**Date Donated**	1999-07-01
Associated Tasks:	Classification	**Missing Values?**	Yes	**Number of Web Hits:**	514344

https://archive.ics.uci.edu/ml/datasets/spambase

Instrucciones

- Construya un notebook ordenado, reproducible y documentado
- Aplique técnicas de balanceo de datos
- Aplique técnicas de optimización de hiperparámetros
- Utilice parámetros de regularización
- Compare los resultados del modelo antes de la optimización y después de la optimización
- Elabore conclusiones

KIBERNUM

¡Gracias!