

Executive Summary

- En este proyecto, aplicamos ciencia de datos para analizar y predecir el éxito de los aterrizajes del Falcon 9 de SpaceX.
- Utilizamos datos reales para entender los factores determinantes, desarrollamos visualizaciones interactivas, modelos predictivos y dashboards.
- Incluye: limpieza, análisis, visualización, ML y presentación de resultados.

Introduction

- Falcon 9 es un lanzador reutilizable. El objetivo fue analizar datos históricos para predecir la probabilidad de éxito de cada aterrizaje.
- Preguntas clave:
 - - ¿Qué factores afectan el éxito?
 - - ¿Podemos predecir el resultado?

Data Collection & Wrangling

- - Datos desde SpaceX API, Wikipedia, CSVs.
- - Limpieza de nulos y duplicados.
- - Conversión de tipos.
- - Creación de columna Class (0=failure, 1=success).
- - One-hot encoding de variables categóricas.

EDA & Visual Analytics - Metodología

- - Uso de seaborn, matplotlib y plotly para visualización.
- - Folium para análisis espacial.
- - Dash para construir dashboard interactivo.

EDA con SQL - Resultados

- - Número de lanzamientos por sitio.
- - Payload máximo y mínimo.
- - Ranking de outcomes por fecha.
- - Uso de subqueries y agrupaciones.

Mapa Interactivo con Folium

- - Sitios de lanzamiento marcados con círculos.
- - Líneas a ciudades/carreteras.
- - Cálculo de distancias.
- - Uso de `MousePosition` y `MarkerCluster`.

Dash Dashboard Interactivo

- - Drop-down para Launch Site.
- - Pie chart de tasa de éxito.
- - Range slider para payload.
- - Scatter plot por selección.

Modelado Predictivo - Metodología

- - División train/test.
- - Estandarización con StandardScaler.
- - Modelos usados: KNN, SVM, Árbol de Decisión, Regresión Logística.
- - Validación cruzada y GridSearchCV.

Resultados de Clasificación

- - KNN: 94.44% accuracy (mejor rendimiento).
- - SVM: kernel 'sigmoid' mejor resultado (~83%).
- - Árbol de Decisión: ~88.89%.
- - Evaluación con matriz de confusión.

Conclusión

- Se logró predecir con alta precisión los aterrizajes exitosos del Falcon 9.
- Las visualizaciones geoespaciales y dashboards mejoraron la comprensión.
- El proyecto demuestra una aplicación completa de ciencia de datos.