

TFM

PYTHON



JosÉ cabezas

Agradecimientos

Contenido

[Abstract 1](#_Toc138151959)

[Introducción 2](#_Toc138151960)

[Objetivos generales del trabajo 2](#_Toc138151961)

[Justifica la elección del tema. 2](#_Toc138151962)

[Resumen metodología 3](#_Toc138151963)

[Estructura 4](#_Toc138151964)

[Planteamiento del problema 4](#_Toc138151965)

[Objetivos del trabajo 5](#_Toc138151966)

[Metodología 6](#_Toc138151967)

[Evaluación de los resultados (¿o quiere decir que evaluemos que variables hacen que existan mayor riesgo de accidentes? , se podría usar la matriz de coorelación usada cuando se generó el modelo) 6](#_Toc138151968)

[Conclusiones 7](#_Toc138151969)

[Referencias 8](#_Toc138151970)

[Anexos TFM I](#_Toc138151971)

# Abstract

Our Master's thesis focuses on developing a Python-based system for predicting airplane accidents. To achieve this, we utilized the Streamlit library and worked with a preprocessed dataset of airplane accidents. Additionally, we employed the PyCaret library for the prediction process.

Firstly, we dedicated time to preprocess the airplane accident dataset. This involved tasks such as data cleaning, outlier removal, and variable transformation. Through this process, we obtained a refined dataset ready for use in our prediction model.

Next, we utilized the PyCaret library to develop a machine learning model capable of predicting the occurrence of airplane accidents. PyCaret allowed us to explore various classification algorithms and identify the most suitable one for our dataset. Furthermore, it greatly simplified the model training and evaluation process.

Once the prediction model was generated, we utilized the Streamlit library to create an interactive interface. We designed a form where users could input relevant data for making predictions, including the date, airplane type, and airline, among others. Streamlit facilitated the creation of this user interface by enabling quick design and deployment of interactive web applications.

In summary, our group's Master's thesis focused on developing a Python-based airplane accident prediction system. We employed Streamlit to create a user input form and PyCaret for training and evaluating the prediction model. Our work involved preprocessing the airplane accident dataset, selecting the most suitable classification algorithm, and creating an interactive interface for users. With this system, we aim to contribute to enhancing safety in the aviation industry and preventing potential accidents.

# Introducción

## Objetivos generales del trabajo

El objetivo general de nuestro trabajo de Fin de Máster (TFM) ha sido desarrollar un sistema de predicción de accidentes de aviones utilizando programación en Python. Para lograr este objetivo, hemos utilizado las bibliotecas Streamlit y PyCaret, junto con un conjunto de datos de accidentes de aviones previamente procesado.

Nuestro principal propósito ha sido contribuir a la mejora de la seguridad en la industria de la aviación, brindando herramientas que permitan predecir la ocurrencia de accidentes de aviones. La capacidad de prever posibles accidentes puede facilitar la toma de decisiones y la implementación de medidas preventivas por parte de las autoridades, las compañías aéreas y otros actores involucrados en la industria.

Además, nos hemos planteado el objetivo de aprovechar las ventajas de la programación en Python y las bibliotecas Streamlit y PyCaret para desarrollar un sistema eficiente y fácil de usar. Queremos ofrecer una interfaz intuitiva que permita a los usuarios ingresar los datos relevantes y obtener predicciones precisas sobre la probabilidad de que ocurra un accidente de avión.

En resumen, nuestro trabajo se ha centrado en el desarrollo de un sistema de predicción de accidentes de aviones con el fin de mejorar la seguridad en la industria de la aviación. Mediante el uso de programación en Python, las bibliotecas Streamlit y PyCaret, y un conjunto de datos preprocesado, hemos buscado ofrecer una herramienta eficiente y accesible para predecir la probabilidad de ocurrencia de accidentes de aviones.

## Justifica la elección del tema.

La elección del tema para nuestro trabajo de Fin de Máster (TFM) se basa en diversas justificaciones fundamentales. La seguridad en la industria de la aviación es un tema de vital importancia, tanto para la protección de vidas humanas como para la economía global. Los accidentes de aviones tienen consecuencias devastadoras en términos de pérdida de vidas, daños materiales y repercusiones legales y financieras. Por lo tanto, desarrollar un sistema de predicción de accidentes de aviones se presenta como una medida crucial para prevenir y reducir los riesgos asociados, mejorando así la seguridad en este sector crítico.

Además, los avances tecnológicos en el procesamiento de datos y las técnicas de aprendizaje automático han brindado nuevas oportunidades para abordar problemas complejos como la predicción de accidentes de aviones. Al aprovechar el potencial de la programación en Python y las bibliotecas especializadas como Streamlit y PyCaret, podemos desarrollar un sistema eficiente y preciso.

La elección de utilizar un conjunto de datos de accidentes de aviones previamente recopilados también se justifica por la disponibilidad de información valiosa y la reducción de costos y tiempo asociados con la recopilación de nuevos datos. Al analizar patrones y factores de riesgo en estos datos históricos, podemos generar modelos predictivos confiables que contribuyan a una mejor comprensión de los accidentes de aviones y a la implementación de medidas preventivas más efectivas.

Por último, la creación de un sistema de predicción de accidentes de aviones ofrece una herramienta valiosa para los responsables de la toma de decisiones en la industria de la aviación, como las autoridades de regulación, las compañías aéreas y los investigadores de seguridad. Al proporcionar información precisa sobre la probabilidad de ocurrencia de accidentes, se pueden implementar estrategias de seguridad más efectivas y tomar medidas preventivas proactivas.

En resumen, la elección del tema de nuestro TFM se fundamenta en la importancia de mejorar la seguridad en la industria de la aviación, los avances tecnológicos en el aprendizaje automático, la utilización de datos existentes y la contribución a la toma de decisiones en este sector crítico.

## Resumen metodología

En nuestro trabajo de Fin de Máster (TFM), realizado en equipo por tres personas, hemos utilizado una metodología colaborativa para desarrollar un sistema de predicción de accidentes de aviones. Cada miembro del equipo se ha enfocado en construir un modelo de predicción utilizando diferentes agrupaciones de columnas de datos. Uno de nosotros se ha encargado de los datos de la tripulación, otro de los datos de la aeronave y el tercero de los datos climatológicos.

En primer lugar, cada miembro del equipo ha trabajado de manera individual en la construcción de su propio modelo de predicción. Utilizando las agrupaciones de columnas asignadas, hemos aplicado algoritmos y técnicas de aprendizaje automático para entrenar modelos capaces de predecir la probabilidad de que ocurra un accidente de avión.

A continuación, hemos utilizado la biblioteca Streamlit para crear páginas individuales donde cada miembro del equipo ha desarrollado un formulario para introducir los datos específicos de su modelo. Estas páginas permiten a los usuarios ingresar información relacionada con la tripulación, la aeronave y los datos climatológicos, respectivamente.

Una vez que se han ingresado los datos en cada una de las páginas, hemos utilizado una página final donde hemos combinado y analizado los resultados obtenidos por cada modelo de forma conjunta. En esta página final, interpretamos los resultados de cada algoritmo y los presentamos de manera global. Si varios modelos indican la posibilidad de un accidente, consideramos que existe una mayor probabilidad de que dicho accidente ocurra.

Esta metodología colaborativa nos ha permitido aprovechar la experiencia y conocimientos individuales de cada miembro del equipo en diferentes aspectos del problema de predicción de accidentes de aviones. Al combinar nuestros esfuerzos, hemos logrado desarrollar un sistema más completo y robusto, que abarca aspectos cruciales como la tripulación, la aeronave y los datos climatológicos.

## Estructura

La estructura utilizada en nuestro trabajo de Fin de Máster (TFM) se basó en una metodología colaborativa entre tres personas. A continuación, describiré brevemente la estructura adoptada:

-División de responsabilidades: Cada miembro del equipo asumió la responsabilidad de trabajar en una parte específica del proyecto. Uno se enfocó en los datos de la tripulación, otro en los datos de la aeronave y el tercero en los datos climatológicos. Esta división permitió un enfoque más especializado y eficiente en cada área.

-Construcción de modelos individuales: Cada miembro del equipo trabajó de manera individual en la construcción de un modelo de predicción utilizando las agrupaciones de columnas asignadas. Utilizamos algoritmos y técnicas de aprendizaje automático para entrenar y ajustar estos modelos, con el objetivo de predecir la probabilidad de accidentes de aviones.

- Desarrollo de páginas individuales en Streamlit: Utilizamos la biblioteca Streamlit para crear páginas individuales en las que cada miembro del equipo desarrolló un formulario para introducir los datos específicos de su modelo. Estas páginas permitieron a los usuarios ingresar información relacionada con la tripulación, la aeronave y los datos climatológicos, respectivamente.

-Página final de análisis global: Además de las páginas individuales, creamos una página final en Streamlit donde combinamos y analizamos los resultados obtenidos por cada modelo de forma conjunta. En esta página, interpretamos los resultados de cada algoritmo y presentamos una visión global de la probabilidad de ocurrencia de accidentes de aviones, considerando la información proporcionada por todos los modelos.

Esta estructura nos permitió aprovechar la experiencia y conocimientos individuales de cada miembro del equipo, mientras trabajábamos de manera colaborativa para desarrollar un sistema completo de predicción de accidentes de aviones.

# Planteamiento del problema

En el planteamiento del problema de nuestro Trabajo de Fin de Máster (TFM), nos enfrentamos al desafío de mejorar la seguridad en la industria de la aviación mediante la predicción de accidentes de aviones. Los accidentes de aviones tienen un impacto significativo en términos de pérdida de vidas humanas, daños materiales y consecuencias económicas. Por lo tanto, es crucial desarrollar herramientas que permitan predecir la ocurrencia de estos eventos y tomar medidas preventivas para evitarlos.

Aunque se han implementado múltiples medidas de seguridad y protocolos en la industria de la aviación, los accidentes todavía ocurren y es necesario abordar este problema de manera proactiva. Actualmente, la toma de decisiones en la gestión de la seguridad aérea se basa en datos históricos y en análisis manuales, lo que puede ser limitante y no proporcionar una visión holística de los riesgos.

El objetivo principal de nuestro TFM ha sido desarrollar un sistema de predicción de accidentes de aviones utilizando técnicas de aprendizaje automático y datos relevantes. La idea es aprovechar la disponibilidad de conjuntos de datos de accidentes de aviones previamente recopilados y aplicar algoritmos de aprendizaje automático para construir un modelo predictivo.

Este modelo nos permitirá evaluar la probabilidad de que ocurra un accidente de avión en función de diversos factores, como información de la tripulación, datos de la aeronave y condiciones climatológicas. Al predecir la probabilidad de accidentes, se podrían tomar medidas proactivas para mejorar la seguridad, como ajustar procedimientos, implementar cambios en el diseño de las aeronaves o fortalecer la formación y capacitación de la tripulación.

En resumen, el planteamiento del problema de nuestro TFM se centra en la necesidad de mejorar la seguridad en la industria de la aviación mediante la predicción de accidentes de aviones. Al utilizar técnicas de aprendizaje automático y datos relevantes, buscamos desarrollar un sistema predictivo que proporcione información valiosa para la toma de decisiones y la implementación de medidas preventivas, contribuyendo así a reducir los riesgos y garantizar la seguridad en el ámbito aeronáutico.

# Objetivos del trabajo

El objetivo general de nuestro trabajo consiste en desarrollar un sistema de predicción de accidentes de aviones utilizando técnicas de aprendizaje automático y datos relevantes. Esta iniciativa tiene como finalidad mejorar la seguridad en la industria de la aviación, abordando de manera proactiva la prevención de accidentes y sus consecuencias.

Para lograr este objetivo general, hemos establecido varios objetivos secundarios que se derivan del principal y que justifican y permiten su desarrollo. En primer lugar, nos planteamos realizar un análisis exploratorio exhaustivo de los datos de accidentes de aviones disponibles. A través de este análisis, buscamos comprender la estructura de los datos, identificar patrones significativos, detectar valores atípicos y establecer posibles correlaciones entre las variables.

El segundo objetivo se centra en el preprocesamiento de los datos. Aquí, nos enfocamos en llevar a cabo tareas de limpieza, manejo de valores faltantes, normalización de variables y transformaciones necesarias para asegurar la calidad y coherencia de los datos utilizados en el modelo de predicción.

Como tercer objetivo, cada miembro del equipo se encargó de construir un modelo de predicción individual utilizando agrupaciones específicas de columnas de datos. Estas agrupaciones abarcaron la información de la tripulación, los datos de la aeronave y los datos climatológicos. Para ello, aplicamos algoritmos de aprendizaje automático con el fin de entrenar y ajustar los modelos, buscando predecir la probabilidad de accidentes de aviones de manera individual.

El siguiente objetivo fue implementar una interfaz de usuario utilizando la biblioteca Streamlit. Esta interfaz fue diseñada de manera intuitiva y amigable, permitiendo a los usuarios introducir los datos relevantes en cada uno de los modelos de predicción construidos por los miembros del equipo. Esta implementación facilitó la interacción con el sistema y mejoró su usabilidad.

Finalmente, el último objetivo consistió en integrar los resultados de los modelos individuales y realizar un análisis global. Este análisis nos permitió interpretar la probabilidad de ocurrencia de accidentes de aviones de manera conjunta, considerando la combinación de los resultados individuales. Esta aproximación nos proporcionó una visión más precisa y confiable de los posibles accidentes.

En resumen, nuestro trabajo se basó en el desarrollo de un sistema de predicción de accidentes de aviones, donde los objetivos secundarios de realizar un análisis exploratorio, preprocesar los datos, construir modelos individuales, implementar una interfaz de usuario y realizar un análisis global de resultados se complementaron para alcanzar el objetivo general de mejorar la seguridad en la industria de la aviación.

# Metodología

## 1.-Preprocesamiento del dataset

### 1.1.-Forecast

### 1.2.-Crew

### 1.3.-Airplane

## 2.-Generar el del modelo con pycaret

### 2.1.-Forecast

### 2.2.-Crew

### 2.3.-Airplane

## 3.-Generar estructura de Streamlit

### 2.1.-Forecast

### 2.2.-Crew

### 2.3.-Airplane

## 4.-Union de las partes y prueba en local

## 5.-Subida a la SteamCloud

# Evaluación de los resultados (¿o quiere decir que evaluemos que variables hacen que existan mayor riesgo de accidentes? , se podría usar la matriz de coorelación usada cuando se generó el modelo)

En primer lugar, se realizó un análisis exploratorio detallado de los datos de accidentes de aviones. A través de este análisis, se identificaron patrones y tendencias significativas en los datos, lo que permitió comprender mejor la naturaleza de los accidentes y los factores que podrían contribuir a su ocurrencia.

Posteriormente, se procedió al preprocesamiento de los datos, abordando cuestiones como la limpieza de datos, el manejo de valores faltantes y la normalización de variables. Este proceso fue fundamental para garantizar la calidad y consistencia de los datos utilizados en los modelos de predicción.

Se construyeron modelos de predicción individuales para cada una de las agrupaciones de columnas de datos (tripulación, aeronave y datos climatológicos). Se aplicaron diversos algoritmos de aprendizaje automático para entrenar y ajustar estos modelos, con el objetivo de predecir la probabilidad de accidentes de aviones.

A través de la implementación de una interfaz de usuario utilizando Streamlit, se permitió a los usuarios introducir los datos relevantes en cada uno de los modelos de predicción. Esto facilitó la interacción con el sistema y brindó una experiencia amigable al usuario.

Al integrar los resultados de los modelos individuales y realizar un análisis global, se obtuvieron conclusiones significativas sobre la probabilidad de ocurrencia de accidentes de aviones. Se pudo establecer una visión más precisa y confiable al considerar la combinación de los resultados individuales. Esto permitió identificar patrones comunes y factores de riesgo que influyen en la probabilidad de accidentes.

En resumen, los resultados obtenidos de nuestro TFM demostraron la eficacia de los modelos de predicción construidos utilizando técnicas de aprendizaje automático y datos relevantes. La integración de estos modelos y el análisis global de los resultados proporcionaron una comprensión más profunda de la probabilidad de ocurrencia de accidentes de aviones, lo que contribuye a la mejora de la seguridad en la industria de la aviación. Los aspectos destacados de esta evaluación incluyen la identificación de patrones, la validación de los modelos y la identificación de factores clave para la prevención de accidentes. Estos hallazgos son fundamentales para tomar medidas preventivas y mejorar la seguridad en el sector aeronáutico.

# Conclusiones

A lo largo de nuestro estudio, hemos logrado desarrollar un sistema de predicción de accidentes de aviones utilizando técnicas de aprendizaje automático y datos relevantes. Este sistema representa una herramienta efectiva para evaluar la probabilidad de ocurrencia de accidentes y contribuir a la mejora de la seguridad en la industria de la aviación.

Nuestro análisis exploratorio detallado de los datos de accidentes de aviones nos ha permitido identificar patrones y tendencias significativas. Esto ha brindado una comprensión más profunda de los factores que pueden influir en la ocurrencia de accidentes y, por lo tanto, ha proporcionado una base sólida para tomar medidas preventivas más efectivas.

El preprocesamiento exhaustivo de los datos ha sido esencial para asegurar la calidad y consistencia de los mismos. Mediante tareas de limpieza, manejo de valores faltantes y normalización de variables, hemos mejorado la precisión de los modelos de predicción y hemos aumentado la confiabilidad de los resultados obtenidos.

Nuestros modelos de predicción, construidos para cada agrupación de columnas de datos (tripulación, aeronave y datos climatológicos), han demostrado ser eficaces para predecir la probabilidad de accidentes de aviones de manera individual. Esta aproximación nos ha permitido obtener una visión detallada de los factores específicos que pueden influir en la ocurrencia de accidentes, lo que resulta fundamental para la prevención y toma de decisiones informadas.

La implementación de una interfaz de usuario intuitiva utilizando Streamlit ha mejorado la interacción con el sistema y ha facilitado el uso por parte de los usuarios. Esto ha permitido introducir los datos relevantes en cada modelo de predicción de manera sencilla y ha mejorado la experiencia global del usuario al utilizar nuestro sistema.

Al integrar los resultados individuales de los modelos de predicción y realizar un análisis global, hemos obtenido una visión más precisa y confiable de la probabilidad de ocurrencia de accidentes de aviones. Esta evaluación conjunta, considerando la combinación de los resultados individuales, nos ha permitido identificar la importancia de factores múltiples y obtener una evaluación más sólida de los riesgos asociados.

En resumen, a través de nuestro TFM hemos logrado desarrollar un sistema de predicción de accidentes de aviones que contribuye a mejorar la seguridad en la industria de la aviación. Nuestros resultados respaldan la eficacia de los modelos de predicción, la relevancia de los datos utilizados y la importancia de aplicar técnicas de aprendizaje automático para tomar decisiones informadas en el ámbito de la seguridad aérea.

# Referencias

La bibliografía irá al final del trabajo, ordenada alfabéticamente.

# Anexos TFM

Se trata de una sección que se utiliza para adjuntar la información en detalle.

En este caso, el fichero que contenga el código de su trabajo. Suele añadirse cualquier documento que por tamaño no se haya podido exponer en el trabajo y que aporte algún conocimiento o recurso adicional al resto del proyecto.