

General Electric Aerospace



Model Evaluation

Flavio Ruvalcaba Leija - A01367631

Oscar Eduardo Nieto Espitia - A01705090

Eduardo Gonzalez Luna - A01658281

Cristian Rogelio Espinoza Diaz - A01702752

Anatanael Jesus Miranda Faustino - A01769232

16/10/2023

1. Evaluación de resultados.....	3
1.1 Hallazgos.....	3
1.2 Selección del modelo.....	4
1.3 Evaluación de los resultados de minería de datos.....	6
1.4 Modelos Aprobados.....	7
2. Revisión de procesos.....	7
2.1 Revisión de forma de trabajo.....	7
2.2 Revisión de objetivos de negocio con respecto a sus criterios de éxito.....	8
3. Siguiendo pasos.....	9
3.1 Lista de posibles acciones.....	9
3.2 Decisión.....	11

1. Evaluación de resultados

Con el objetivo de identificar las medidas registradas por los sensores que determinan la estabilidad de un avión durante la fase de 'Stable Cruise', hemos empleado los datos recopilados por estos sensores para entrenar un modelo predictivo. Nuestro propósito es lograr predicciones precisas que indiquen con certeza si un avión se encuentra en este estado deseado de estabilidad.

Se establecieron criterios de éxito para evaluar la eficacia de nuestro modelo. Para considerar nuestro esfuerzo de minería de datos exitoso, buscamos alcanzar una accuracy promedio superior al 70%. Esto no solo refleja la capacidad de nuestro modelo para predecir con precisión la Estabilidad de Crucero, sino que también garantizará su utilidad práctica en la identificación de este estado crucial de estabilidad en vuelo.

1.1 Hallazgos

1. **Hallazgo:** Variables con mayor importancia

Después de entrenar y evaluar nuestro modelo, hemos identificado las variables más influyentes en la predicción de la fase de "Stable Cruise". La importancia de las variables se evaluó utilizando el F-score, y los resultados revelaron que las siguientes variables tuvieron el mayor impacto en el rendimiento del modelo

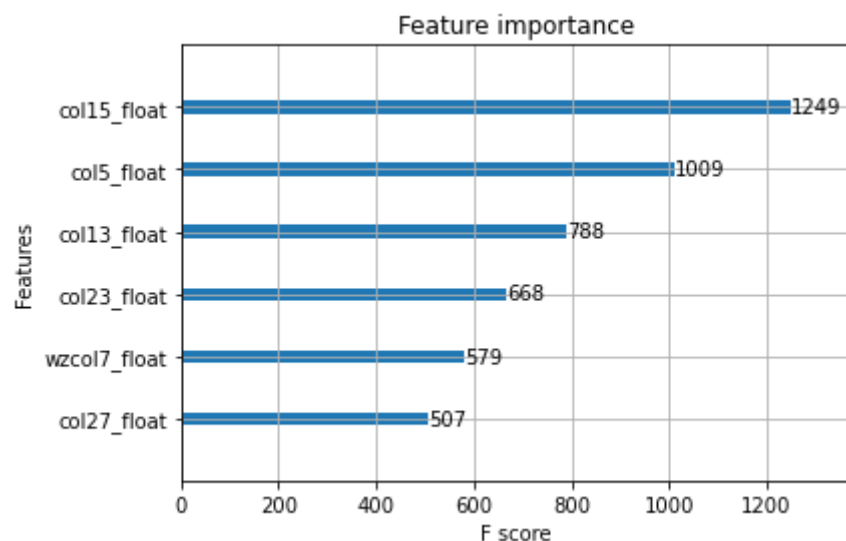


Fig. 1 Importancia de cada variable en F-score (XGBoost Undersampling L2)

Estos hallazgos destacan la importancia de estas variables específicas en la predicción de la fase de "Stable Cruise" y proporcionan perspectivas valiosas para comprender qué características del vuelo son más críticas para la estabilidad.

2. **Hallazgo:** Como lidiamos con multicolinealidad (L2)

En el proyecto desarrollamos un análisis profundo para verificar que las variables no estuvieran altamente relacionadas entre sí.

Se ha optado por implementar técnicas de regularización para abordar los problemas de multicolinealidad. Estos métodos de regularización imponen penalizaciones a las variables con mayor multicolinealidad, lo que repercute en la penalización de los pesos durante el proceso de entrenamiento del modelo.

3. **Hallazgo:** col17_float es muy similar a col15_float

Las variables "col5_float" y "col13_float" exhiben una significativa multicolinealidad, siendo comúnmente un valor por debajo de 5 un indicativo de ausencia de dicho fenómeno.

La decisión se tomó de eliminar la variable "col17_float" debido a su elevado índice de correlación con "col15_float", alcanzando una correlación de 0.96. Esta correlación representa la máxima obtenida después de filtrar las variables, sugiriendo esencialmente que ambas variables son prácticamente idénticas debido a su extremadamente alta correlación.

4. **Hallazgo:** Undersampling mejor que Oversampling

Identificamos una desproporción significativa de 400:1 en los datos de "stable_cruise", lo que podría afectar negativamente el rendimiento del modelo de clasificación binaria. Tras evaluar distintas opciones, determinamos que el undersampling, que consiste en reducir instancias de la clase mayoritaria para equilibrarla con la clase minoritaria, fue la estrategia más efectiva. Descartamos alternativas como SMOTE y ajuste de parámetros debido a la magnitud del desequilibrio y la duración del entrenamiento del modelo. Este hallazgo destaca la importancia de adaptar las técnicas de manejo de desequilibrios a la naturaleza específica de los datos para mejorar la capacidad predictiva del modelo.

5. **Hallazgo:** XGBoost buen modelo para este problema

XGBoost se destacó como la mejor alternativa entre los modelos entrenados debido a su arquitectura y características específicas que lo hacen idóneo para este proyecto. XGBoost es un algoritmo de ensamblado de árboles de decisión que ha demostrado ser eficaz en problemas de clasificación, como el nuestro.

La arquitectura de XGBoost permite la paralelización eficiente de la construcción de árboles, lo que acelera el tiempo de entrenamiento, y su capacidad para manejar multicolinealidad ayuda a evitar problemas asociados con la correlación entre variables.

1.2 Selección del modelo

En la fase de la creación de los modelos usamos diversas métricas de desempeño como:

- **Accuracy:** La accuracy es una métrica que muestra el porcentaje de predicciones correctas realizadas por el modelo. Esta métrica es especialmente útil para evaluar el rendimiento general de un modelo en un contexto amplio, donde el objetivo principal es conocer la proporción de predicciones correctas en comparación con el total de predicciones.

- **Precisión:** La precisión mide la capacidad del modelo para realizar predicciones positivas correctas en relación con el total de predicciones positivas hechas. Es especialmente relevante cuando minimizar los falsos positivos es crítico. Por ejemplo, en aplicaciones médicas o de seguridad, la precisión es esencial para garantizar que las predicciones positivas sean altamente confiables y no se generen falsas alarmas. La precisión es una métrica valiosa cuando el costo de los falsos positivos es alto.
- **F1-Score:** El valor F1-Score combina tanto la precisión como el recall en una sola métrica. Ofrece un equilibrio entre la capacidad del modelo para identificar positivos verdaderos y su capacidad para evitar falsos positivos. El F1-Score es especialmente efectivo cuando los costos de los errores en las clases son diferentes, y se necesita un equilibrio entre la precisión y la sensibilidad. Esta métrica proporciona una evaluación completa del rendimiento del modelo, considerando tanto la capacidad de hacer predicciones precisas como la capacidad de detectar positivos reales.

No obstante, hubo una medida de desempeño que resalta más debido a que uno de los criterios que posee el socio con el proyecto es la eficiencia en la detección del parámetro de “Stable Cruise”. Esta medida es:

- **Recall (Sensibilidad):** El recall, también conocido como sensibilidad, evalúa la capacidad del modelo para identificar con precisión los casos positivos reales en relación con el total de casos positivos presentes. Esta métrica es fundamental en problemas donde la detección de casos positivos es de máxima importancia. Por ejemplo, en pruebas médicas para enfermedades graves, el recall se convierte en una métrica crítica, ya que se busca identificar a todos los pacientes que realmente padecen la enfermedad. El recall es especialmente relevante en situaciones de desequilibrio entre clases, como el caso de "Stable Cruise," donde se centra en la detección de los casos positivos.

Al mismo tiempo, otro parámetro que utilizamos para seleccionar el modelo final es:

Verdaderos Positivos (True Positives, TP): Este término se refiere a la cantidad de observaciones o instancias positivas que fueron correctamente identificadas por el modelo como positivas. En otras palabras, son los casos en los que el modelo predice correctamente que una instancia es positiva y la etiqueta real también es positiva.

A lo largo de los diversos modelos que creamos y evaluamos en la fase de “Modeling”, obtuvimos dos modelos con un desempeño excepcional.

Ambos utilizan el algoritmo de XGBoost y la técnica de regularización L2:

- XGBoost, que significa "eXtreme Gradient Boosting", es una biblioteca de software de código abierto que implementa el algoritmo de Gradient Boosting. Es una técnica

popular en el aprendizaje supervisado y se utiliza comúnmente para problemas de clasificación y regresión. XGBoost ha demostrado ser eficaz en una amplia gama de tareas y es conocido por su velocidad y rendimiento superior.

XGBoost se basa en el principio de Gradient Boosting, que es una técnica de ensamblaje de modelos. En lugar de construir un solo modelo fuerte, Gradient Boosting origina un conjunto de modelos más débiles (por lo general, árboles de decisión) y los combina para mejorar la predicción general.

- La técnica de regularización L2, también conocida como Ridge regularization o Tikhonov regularization, es una técnica utilizada en el aprendizaje automático para evitar el sobreajuste (overfitting) y mejorar la generalización de un modelo, especialmente en problemas de regresión. Esta técnica se aplica comúnmente en modelos lineales, como la regresión lineal, la regresión logística y otras formas de modelos lineales generalizados.

La diferencia que existe entre ambos modelos es la técnica que se usó para manejar el desequilibrio de las clases 0 y 1 en el conjunto de datos brindados por el socio formador.

1.2.1 Modelo 1. XGBoost con L2 y oversampling.

El primer modelo que estaremos evaluando utilizó la técnica de oversampling. El oversampling es una técnica utilizada en el ámbito del aprendizaje automático, la cual consiste en aumentar la frecuencia de la clase minoritaria, generando instancias sintéticas o replicando instancias existentes. Hay varias formas de implementar el oversampling, la forma en la que implementamos fue por medio de:

- **SMOTE (Synthetic Minority Over-sampling Technique):** es una técnica más avanzada que crea instancias sintéticas de la clase minoritaria mediante la interpolación entre instancias existentes. En lugar de duplicar instancias, Smote origina nuevas instancias generando puntos en el espacio de características entre instancias de la clase minoritaria vecinas.

```

Training set accuracy: 0.9061541540830916
Test set accuracy: 0.9060667743873259

Validation set accuracy: 0.9059730104887903
Confusion Matrix for the test set:
[[10467892  1087745]
 [   1218    36093]]

Classification Report for the test set:

```

	precision	recall	f1-score	support
0.0	1.00	0.91	0.95	11555637
1.0	0.03	0.97	0.06	37311
accuracy			0.91	11592948
macro avg	0.52	0.94	0.51	11592948
weighted avg	1.00	0.91	0.95	11592948

Fig 3. Resultados de la ejecución del modelo.

Este modelo posee los siguientes valores acorde a la matriz de confusión:

- **Verdaderos Positivos (True Positives, TP):** 36,093 predicciones correctas de la clase 1.
- **Falsos Negativos (False Negatives, FN):** 1,218 predicciones incorrectas de la clase 0.
- **Falsos Positivos (False Positives, FP):** 1,087,745 predicciones incorrectas de la clase 1.
- **Verdaderos Negativos (True Negatives, TN):** 10,467,892 predicciones correctas de la clase 0.

El modelo muestra un buen rendimiento en la clase 0 (negativa) según el informe de clasificación y la matriz de confusión. La precisión para la clase 0 es alta (1.00), y el recall es del 91%, lo que indica que el modelo es eficaz para identificar la mayoría de los casos negativos reales. La matriz de confusión muestra un número significativamente mayor de true negatives (TN) que false negatives (FN), lo que respalda la conclusión de un buen rendimiento para la clase 0.

Sin embargo, para la clase 1 (positiva), el rendimiento es menos satisfactorio. Aunque el recall es del 97%, lo que indica que el modelo identifica la mayoría de los casos positivos reales, la precisión es baja (0.03), lo que significa que la mayoría de las predicciones positivas son falsas. La matriz de confusión muestra que la cantidad de false positives (FP) es mayor que la de true positives (TP). Esto confirma que el modelo tiende a clasificar erróneamente muchos casos como positivos en esta clase.

A pesar de estas diferencias en el rendimiento entre las clases, la precisión general del modelo en el conjunto de prueba es alta (0.906). Esto se debe a la gran proporción de casos negativos en tus datos, lo que puede llevar a una alta precisión incluso si el modelo tiene dificultades para predecir la clase minoritaria.

En resumen, el modelo es efectivo para predecir la clase mayoritaria (clase 0), pero muestra limitaciones en la predicción de la clase minoritaria (clase 1), con un alto número de falsos positivos.

1.2.2 Modelo 2. XGBoost con L2 y undersampling.

El segundo modelo que estaremos evaluando utilizó la técnica de undersampling. El undersampling es una técnica utilizada en el ámbito de clasificación desequilibrada, donde una clase tiene muchas más instancias que la otra. En lugar de utilizar todas las instancias de la clase mayoritaria, submuestreo implica reducir la cantidad de instancias de la clase mayoritaria para igualarla a la cantidad de instancias de la clase minoritaria.

Nuestra implementación de esta técnica fue generar un dataset con la misma cantidad de ejemplos de la clase mayoritaria a la clase minoritaria.

```
Training set accuracy: 0.9442413208845606
Test set accuracy: 0.9420791212932926

Validation set accuracy: 0.9417703117428525
Confusion Matrix for the test set:
[[32972  3865]
 [  438 37016]]

Classification Report for the test set:
```

	precision	recall	f1-score	support
0.0	0.99	0.90	0.94	36837
1.0	0.91	0.99	0.95	37454
accuracy			0.94	74291
macro avg	0.95	0.94	0.94	74291
weighted avg	0.95	0.94	0.94	74291

Fig 3. Resultados de la ejecución del modelo.

Este modelo posee los siguientes valores acorde a la matriz de confusión:

- **Verdaderos Positivos (True Positives, TP):** 37,016 predicciones correctas de la clase 1.
- **Falsos Negativos (False Negatives, FN):** 438 predicciones incorrectas de la clase 0.
- **Falsos Positivos (False Positives, FP):** 3,865 predicciones incorrectas de la clase 1.
- **Verdaderos Negativos (True Negatives, TN):** 32,972 predicciones correctas de la clase 0.

En este conjunto de resultados, el modelo también muestra un rendimiento bastante bueno en general, pero con algunas diferencias en comparación con el caso anterior.

La precisión para la clase 0 es alta (0.99), lo que indica que la gran mayoría de las predicciones negativas son correctas. El recall es del 90%, lo que sugiere que el modelo identifica la mayoría de los casos negativos reales. La matriz de confusión muestra un número mayor de true negatives (TN) que de false negatives (FN), lo cual es positivo y respalda la conclusión de un buen rendimiento para la clase 0.

La precisión para la clase 1 es de 0.91, lo que indica que la mayoría de las predicciones positivas son correctas. Esta es una mejora significativa en comparación con el caso anterior. El recall es del 99%, lo que sugiere que el modelo identifica casi todos los casos positivos reales. También es una mejora sustancial en comparación con el caso anterior. La matriz de confusión muestra un número mayor de true positives (TP) que de false positives (FP), lo cual es positivo y respalda la mejora en el rendimiento para la clase 1.

La precisión general del modelo en el conjunto de prueba es alta (0.942), lo que indica un buen rendimiento en la clasificación general de ambos tipos de clases. Este modelo muestra una mejora en el rendimiento en comparación con el anterior, especialmente en la predicción de la clase 1. La precisión y el recall para ambas clases son notables, y la matriz de confusión refleja un equilibrio entre true positives y true negatives.

En resumen, este modelo demuestra una capacidad sólida para clasificar tanto la clase mayoritaria como la minoritaria, con una mejora significativa en la precisión y recall de la clase 1 en comparación con el modelo anterior.

Comparación inicial de los resultados de los modelos.

Métricas	XGBoost oversampling L2	XGBoost undersampling L2	Modelo Benchmark
Recall clase 1	97%	99%	68%
True Positives	36,093	37,016	25,340
Precisión clase 1	3%	91%	1%
Precisión clase 0	100%	99%	100%
Accuracy	≈ 90%	≈ 94%	≈ 72%

A simple vista, la elección del modelo final sería el modelo de XGBoost con la implementación de undersampling y L2. Debido a que fue el modelo que tuvo mejores resultados en todos los aspectos desde el accuracy en las fases de entrenamiento, pruebas y validación hasta la identificación de la clase 1 con métricas como recall y true positives.

No obstante, para evitar elecciones con posible sesgo dado que se pueden dar en la implementación de técnicas para el manejo de identificar la eficiencia del modelo de manera equitativa, se usó 100 archivos csv proporcionados por el socio formador. Estos datos se nos fueron proporcionados con este objetivo de que todos los equipos validarán el rendimiento de

los modelos debido a que son datos nunca vistos en los primeros 2718 csv que nos brindó para el entrenamiento y creación de los modelos.

Comparación de los resultados con 100 csv de los modelos.

Métricas	XG Boost undersampling L2	XG Boost oversampling L2
Recall	84%	77%
True Positives	16,543	15,138
Precisión clase 1	3%	3%
Precisión clase 0	100%	100%
Accuracy	96.21%	96.56%

En estos nuevos resultados se puede observar que el modelo con undersampling posee una accuracy ligeramente menor al modelo con oversampling, cualquiera de las dos accuracy cumple con nuestro objetivo de minería de datos. Sin embargo, al considerar que el objetivo de negocio consiste en la identificación de stable cruise, es decir, la detección de la clase 1.

El modelo XGBoost con la implementación de L2 y undersampling posee mayor rendimiento para la detección de la clase 1. Esta afirmación se puede corroborar por medio de las métricas que seleccionamos previamente donde el recall de la clase 1 es mayor por 13%, el número de true positives es mayor por 1,405.

En consecuencia, nuestra elección será el modelo con el uso de undersampling puesto que demostró un rendimiento ligeramente mayor al modelo de oversampling.

1.3 Evaluación de los resultados de minería de datos

Para evaluar los resultados de la minería de datos, es esencial abordar las preguntas específicas asociadas y determinar si se ha alcanzado el objetivo propuesto.

Preguntas clave a considerar:

- ¿Cuáles son las características más influyentes para mantener la estabilidad de crucero?
- ¿Existen diferencias estadísticamente significativas entre los datos registrados cuando se implementa redundancia?

Con el objetivo principal en mente:

- “Identificar las medidas capturadas por los sensores que determinan la estabilidad de una aeronave durante la fase conocida como ‘Stable Cruise’. Utilizaremos los datos recopilados por estos sensores para desarrollar un modelo que nos permita predecir con precisión si una aeronave se encuentra en este estado deseado de estabilidad.”

Al abordar la pregunta clave de si hemos logrado el objetivo, la respuesta es:

Sí, hemos logrado el resultado deseado

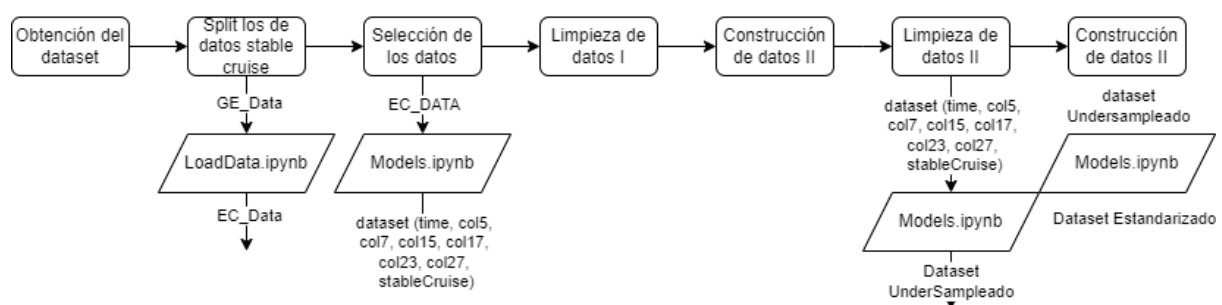
1.4 Modelos Aprobados

Todos los modelos aprobaron el criterio de minería de datos con un accuracy promedio mayor a un 70%, sin embargo, hubo un modelo que tuvo un accuracy que destacó sobre todos los demás con un valor de un 94%, este modelo fue XGBoost.

XGboost, como se puede ver en el documento de [Data Preparation](#), también nos apoyó en SMOTE, con la capacidad de sobre muestreo, un hallazgo importante y necesario para la realización de este caso.

2. Revisión de procesos

2.1 Revisión de forma de trabajo



En la fase inicial del proyecto, omitimos la limpieza de datos debido a la impresión inicial de que los datos estaban lo suficientemente limpios y no presentaban problemas significativos. Sin embargo, durante la revisión del proceso, identificamos una desproporción sustancial de 400:1 en los datos de 'stable_cruise', lo que podría afectar negativamente el rendimiento de nuestro modelo de clasificación binaria.

Considerando las implicaciones de este desequilibrio, reconocimos que podría haber un sesgo hacia la clase mayoritaria, lo que llevaría a que el modelo predijera predominantemente esa clase. Además, nos dimos cuenta de que las métricas de evaluación, como la precisión, podrían ser engañosas en este contexto.

Para abordar este problema y mejorar la robustez de nuestro modelo, implementamos cambios en nuestra estrategia. Optamos por la técnica de undersampling, imputando de manera aleatoria datos falsos para lograr un equilibrio más adecuado entre ambas clases.

Omitir la limpieza inicial de datos fue un error de planificación, ya que subestimamos su importancia. Este descuido afectó la fase de modelado al generar desafíos debido a la desproporción en los datos de 'stable_cruise'.

Tras revisar nuestro plan, hemos determinado que no hay actividades que deban repetirse. Esto indica que, a pesar de los desafíos encontrados durante la ejecución del proyecto, nuestras decisiones y acciones anteriores han sido sólidas.

En retrospectiva, podemos concluir que la omisión de la fase inicial de limpieza de datos fue un error de planificación que tuvo implicaciones significativas en el desarrollo del proyecto. La implementación de la técnica de undersampling fue una decisión acertada para abordar el desequilibrio de clases y mejorar la robustez de nuestro modelo. Sin embargo, este ajuste pudo haberse evitado si hubiéramos realizado una limpieza más exhaustiva en la fase inicial del proyecto.

En cuanto a si todas las fases agregaron valor y se llevaron a cabo de manera óptima, podemos afirmar que, a pesar del contratiempo en la limpieza de datos, las fases subsiguientes demostraron ser efectivas.

2.2 Revisión de objetivos de negocio con respecto a sus criterios de éxito

Para aprobar o eliminar modelos es necesario conocer sus actuaciones para resolver el objetivo de negocio:

- Creación de un modelo destinado a clasificar la estabilidad de vuelo, con el fin de identificar los parámetros ideales que permitan a las aeronaves alcanzar su estado de crucero de forma eficiente.

Este enfoque es fundamental para General Electric, ya que contribuirá significativamente a su principal objetivo:

- Optimizar el consumo de combustible y reducir el desgaste del motor en sus aeronaves.

El criterio de éxito es obtener hallazgos significativos sobre la causalidad de los datos. El socio formador determinará si los hallazgos fueron significativos

- Análisis exploratorio de datos:
 - Al analizar los datos, comprendimos mejor la distribución, las relaciones y las posibles variables influyentes.
 - En el siguiente documento se encuentra la obtención de información del anterior punto: [Data Understanding](#)
 - Puntos a evaluar del anterior documento:
 - Estadísticas descriptivas
 - Correlación con “Stable Cruise”

- Gráficas de distribución

- Selección de variables:

- Identificamos las variables relevantes para el análisis. Utilizamos técnicas como el análisis de correlación y chi cuadrado para comprender las relaciones entre diferentes variables.
- En el siguiente documento se encuentra la obtención de información del anterior punto: [☰ Data Understanding](#)
- Puntos a evaluar del anterior documento:
 - Gráficas de distribución
 - Necesidad de datos adicionales
- En el siguiente documento se encuentra información sobre la decisión de selección de variables: [☰ Data Preparation](#)
- Puntos a evaluar del anterior documento:
 - Selección de datos
 - Análisis de multicolinealidad

- Modelado:

- En los diferentes modelos que creamos encontramos hallazgos que influyen en el uso de hiperparámetros, beneficios de undersampling, mejores modelos, etc.
- En el siguiente documento se encuentra la obtención de información del anterior punto: [☰ Modeling](#)
- Puntos a evaluar del anterior documento:
 - Construcción del modelo
 - Evaluación de los modelos
- En el siguiente documento se encuentra la obtención de información de los beneficios de undersampling: [☰ Data Preparation](#)
- Puntos a evaluar del anterior documento:
 - Calidad de los datos

- Minería de datos

- A través del objetivo y preguntas de la minería de datos respondimos los hallazgos que encontramos en las mismas.
- En el siguiente documento se encuentra la obtención de información del anterior punto: [☰ Data Understanding](#)
- Puntos a evaluar del anterior documento:
 - Interpretación de resultados

3. Siguiendo pasos

3.1 Lista de posibles acciones

Finalizar el proyecto y avanzar a la fase de entrega:	Iniciar más iteraciones:	Crear nuevos proyectos de minería de datos:
<p>Puntos a favor:</p> <ul style="list-style-type: none">- El proyecto ha alcanzado sus objetivos iniciales y ha logrado los resultados esperados, podría ser un indicativo de que el proyecto está listo para avanzar a la fase de entrega.- El modelo ha sido sometido a una validación rigurosa y ha demostrado un buen rendimiento en conjuntos de datos de prueba y validación, proporcionando confianza en su capacidad para generalizar a nuevos datos.	<p>Puntos a favor:</p> <ul style="list-style-type: none">- El modelo actual, como cualquier otro, aún tiene un porcentaje de mejoras que puede tener, una nueva iteración podría ser necesaria para abordar esas deficiencias y lograr un mejor rendimiento.	<p>Puntos a favor:</p> <ul style="list-style-type: none">- La minería de datos es un campo en constante evolución. Poner en marcha nuevos proyectos puede ser una oportunidad para aprender y aplicar las últimas técnicas y mejores prácticas en el análisis de datos.- Si la organización está interesada en explorar áreas adicionales de conocimiento o en ampliar su comprensión de los datos, arrancar nuevos proyectos puede ser una forma efectiva de obtener una visión más completa y holística.
<p>Puntos en contra:</p> <ul style="list-style-type: none">- Los usuarios finales aún no han visto el modelo en marcha en primera mano.	<p>Puntos en contra:</p> <ul style="list-style-type: none">- La falta de recursos, ya sea en términos de tiempo, presupuesto o capacidad computacional, es un factor limitante para iniciar una nueva iteración.- Los resultados obtenidos en la iteración actual cumplen con los objetivos y requisitos del proyecto de manera satisfactoria.	<p>Puntos en contra:</p> <ul style="list-style-type: none">- Los objetivos originales del proyecto de minería de datos se han alcanzado de manera satisfactoria y no hay necesidad inmediata de explorar nuevos patrones o conocimientos.- La asignación de recursos, como tiempo, personal y presupuesto, puede ser un factor limitante. Si los recursos disponibles no son suficientes para abordar nuevos proyectos sin comprometer la calidad, puede ser más prudente consolidar los resultados actuales.

Potencial de entrega de resultados: El análisis de los resultados y hallazgos del proyecto presenta un potencial significativo para contribuir al entendimiento y mejora de la predicción de la fase de "Stable Cruise". Se destaca la identificación de variables clave, proporcionando diversas perspectivas sobre las características de vuelo críticas para la estabilidad.

La gestión de la multicolinealidad mediante técnicas de regularización muestra una estrategia efectiva para mantener la calidad del modelo. El hallazgo sobre la preferencia del undersampling frente al oversampling frente a la desproporción de datos destaca la importancia de adaptar las estrategias de manejo de desequilibrios a las características específicas del conjunto de datos.

Posibles mejoras al proceso: Para mejorar el proceso, pensamos que se debe dar mayor importancia a la fase inicial de limpieza de datos, realizando análisis exploratorios detallados y evaluando posibles desafíos, como desequilibrios en las clases. Podríamos haber implementado técnicas avanzadas de visualización y análisis estadístico para detectar patrones y anomalías desde el principio. En el caso de que se optimizará la fase de limpieza de datos, nos hubiéramos evitado contratiempos mejorando el rendimiento general del modelo.

Posibles continuaciones al proyecto:

- Ingeniería de Características Avanzada: Se podría ejecutar una ingeniería de características más avanzada, creando nuevas variables derivadas de las existentes o extrayendo patrones más complejos de los datos.

3.2 Decisión

Cómo se puede observar anteriormente, una decisión en la lista de posibles acciones es finalizar el proyecto y avanzar a la entrega sin embargo necesitamos la validación del socio formador esto se logrará el próximo viernes 17 de noviembre una vez obtenida esa validación ya que es el único punto en contra.

4. Referencias

- Raeburn A. (2023). Accuracy vs. Precision: What's the Difference? Recuperado de <https://asana.com/es/resources/accuracy-vs-precision>
- Kundu R. (2022). F1 Score in Machine Learning: Intro & Calculation. Recuperado de <https://asana.com/es/resources/accuracy-vs-precision>
- Iguazio (s.f.). What is Recall in Machine Learning. Recuperado de <https://www.iguazio.com/glossary/recall/>