**PROBLEMA**: Análisis de tráfico Aéreo y predicción de retrasos

**OBJETIVOS SMART:**

* ***Specific (Específico):***

Desarrollar y validar un modelo de Machine Learning capaz de predecir retrasos en vuelos comerciales superiores a 30 minutos, generando la predicción con al menos 3 horas de antelación a la hora de salida programada e identificando los factores clave que los ocasionan.

* ***Measurable (Medible):***

Alcanzar una precisión de predicción superior al 85% para la identificación de retrasos y reducir el error cuadrático medio (RMSE) a menos de 15 minutos en la estimación del tiempo de demora, evaluado sobre un conjunto de datos de prueba.

* ***Achievable (Alcanzable):***

Utilizar conjuntos de datos históricos de tráfico aéreo, información meteorológica y datos operativos de vuelos, aplicando técnicas avanzadas de Machine Learning y herramientas de análisis de datos disponibles.

* ***Relevant (Relevante):***

Optimizar la gestión del tráfico aéreo y la eficiencia operativa de aerolíneas y aeropuertos, proporcionar información anticipada a los pasajeros y permitir una planificación proactiva que reduzca el impacto de las disrupciones.

* ***Time-bound (Acotado en el tiempo):***

Completar el desarrollo, validación y presentación del modelo, junto con el informe final de resultados, en un plazo máximo de 6 meses.

**Planificación de Riesgos del Proyecto: Análisis de Tráfico Aéreo y Predicción de Retrasos**

**1. Introducción**

Como Project Manager, este documento identifica y documenta los riesgos asociados al proyecto de desarrollo de un modelo de Machine Learning para predecir retrasos en vuelos comerciales superiores a 30 minutos, con al menos 3 horas de antelación. Se incluye una lista de riesgos evaluada según las escalas de probabilidad (Baja: 1, Media: 2, Alta: 3), impacto (Bajo: 1, Medio: 2, Alto: 3), nivel de riesgo (P × I), y clasificación (Bajo: 1–3, Medio: 4–6, Alto: 7–9), junto con un plan de mitigación para cada riesgo, cumpliendo con los criterios de aceptación de la Historia 6.1.

**2. Lista de Riesgos**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **ID** | **Riesgo** | **Probabilidad (P)** | **Impacto (I)** | **Nivel (P×I)** | **Clasificación** | **Plan de Mitigación** | | **R1** | **Datos incompletos o con baja calidad** | 3 | 3 | 9 | Alto | Establecer un proceso de validación y limpieza desde el Sprint 1; usar fuentes alternativas como BTS y OpenSky Network | | **R2** | **Falta de integración entre datos de vuelos y meteorológicos** | 2 | 3 | 6 | Medio | Definir un esquema de claves únicas para unión de datasets (aeropuerto-fecha-hora) antes del modelado | | **R3** | **Limitaciones de hardware o procesamiento** | 2 | 3 | 6 | Medio | Usar entornos en la nube (Google Colab, AWS) y optimizar el código para procesamiento por lotes | | **R4** | **Modelo no alcanza precisión mínima (85%)** | 2 | 3 | 6 | Medio | Probar múltiples algoritmos, ajustar hiperparámetros y aplicar ingeniería de características | | **R5** | **Cambios en disponibilidad de APIs o fuentes de datos** | 1 | 3 | 3 | Bajo | Descargar copias locales de los datos y documentar el proceso de obtención | | **R6** | **Retrasos en la entrega por sobrecarga del equipo** | 2 | 2 | 4 | Medio | Priorizar tareas Must Have en el backlog; realizar revisiones de carga de trabajo en Daily Scrum | | **R7** | **Problemas en despliegue de predicciones en tiempo real** | 2 | 3 | 6 | Medio | Realizar pruebas de API en un entorno de staging antes del despliegue en producción | | **R8** | **Falta de aceptación por parte de stakeholders** | 1 | 3 | 3 | Bajo | Involucrar a stakeholders en Sprint Reviews y ajustar según feedback | |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

**3. Metodología de Gestión de Riesgos**

* **Identificación:** Los riesgos se identificaron mediante un análisis del backlog, reuniones con el equipo (Data Scientists, Ingenieros de Datos, Analistas de Negocio, Product Owner) y la evaluación de los objetivos SMART.
* **Evaluación:** Se utilizaron las escalas proporcionadas: Probabilidad (Baja: 1, Media: 2, Alta: 3), Impacto (Bajo: 1, Medio: 2, Alto: 3), Nivel de Riesgo (P × I), y Clasificación (Bajo: 1–3, Medio: 4–6, Alto: 7–9).
* **Mitigación:** Se diseñaron estrategias específicas para reducir la probabilidad o el impacto de cada riesgo, alineadas con las actividades del backlog.
* **Monitoreo:** Los riesgos se revisarán en las reuniones de planificación y revisión de cada sprint (Sprints 2-6), con actualizaciones al plan según sea necesario.

**4. Conclusión**

Este plan de riesgos aborda los principales desafíos del proyecto, desde la calidad de los datos hasta la sobrecarga del equipo en el Sprint 6. Las estrategias de mitigación propuestas aseguran que el proyecto pueda avanzar dentro del plazo establecido (23 de julio al 30 de agosto de 2025) y cumplir con los objetivos SMART, especialmente la precisión del modelo (≥85%), el RMSE (<15 minutos), y la entrega de dashboards funcionales en Power BI. Se prestará especial atención al riesgo R1 (datos incompletos o con baja calidad) debido a su alta clasificación (9) y su impacto en todas las fases del proyecto.