



Tecnológico de Monterrey

Campus: Monterrey, México.

7. Resumen ejecutivo

4.4 Establecer supuestos racionales para presentar cifras sobre los costos/beneficios.

4.5 Enumerar los riesgos o desafíos potenciales de la solución propuesta.

Estudiantes

Andrés Julián López Hurtado - A01793899

Nathalia Milena Prada Hernández - A01793999

Víctor Alejandro Regueira Romero - A01794404

Profesor titular

Dr. Horacio Martínez Alfaro

Profesora asistente

Mtra. Verónica Sandra Guzmán de Valle

Fecha: 10 de noviembre de 2024

Contenido

1	El resumen ejecutivo.....	1
1.1	Síntesis del problema.....	1
1.2	Hallazgos más importantes del análisis exploratorio de datos (Avance 1).....	2
1.3	Modelos generados y razones de la elección del modelo final (Avance 5)	3
1.4	Recomendaciones clave para implementar la solución (Avance 6)	4
2	Análisis costo – beneficio.....	7
2.1	Enumerar todos los costos incurridos	7
2.2	Costos esperados para la operación y mantenimiento del modelo	8
2.3	Beneficios potenciales	8
3	Riesgos y desafíos de la solución.....	9
3.1	¿Cuáles son los riesgos o desafíos potenciales de la solución propuesta?.....	9
3.2	Clasificar los riesgos identificados en las categorías provistas en los recursos de la semana (relacionados con los datos, ataques, prueba y confianza, cumplimiento).....	10
4	Bibliografía	11

1 El resumen ejecutivo

1.1 Síntesis del problema

Forza Transportation, especializada en el servicio de Full Truck Load (FTL), enfrenta el desafío de optimizar el consumo de combustible en su flota. Dado que este representa uno de los principales costos operativos, es clave identificar patrones y factores clave que impacten la eficiencia de su flota en rutas que abarcan toda Norteamérica, particularmente en el marco del T-MEC. Este proyecto tiene como objetivo **desarrollar un modelo predictivo que permita a Forza Transportation predecir el consumo de combustible**, ayudando a minimizar costos y reducir su huella de carbono .

Para lograr este objetivo, se ha aplicado la metodología CRISP a partir del análisis de datos recolectados a través de sensores GeoTab, que incluyen eventos de seguridad, tiempos de ralentí y consumo de combustible de dos meses de operación de la flota. Posteriormente, en el preprocesamiento, se han aplicado técnicas de limpieza exhaustiva y de ingeniería de características para asegurar la calidad de los datos y capturar tendencias relevantes a lo largo del tiempo. Finalmente, y luego de tratar con varios modelos de clasificación, el proyecto aporta tres entregables que aseguran valor agregado a la organización:

1. Modelo XGBoost optimizado para la predicción continua del consumo.
2. Un dashboard en PowerBI para visualización de datos de la flota
3. Caracterizaciones SIPOC de los procesos involucrados para gestión del conocimiento y posteriores trabajos.

1.2 Hallazgos más importantes del análisis exploratorio de datos (Avance 1)

El análisis exploratorio de datos (EDA) realizado para Forza Transportation se centró en examinar exhaustivamente las bases de datos recolectadas durante un período de dos meses a través del sistema de telemetría GeoTab. Estas bases de datos incluyen información clave sobre el rendimiento del combustible, tiempos de ralentí y eventos de seguridad, distribuidos en tres tablas principales. La tabla **bronze_safety** contiene datos de eventos relacionados con la seguridad, tales como excesos de velocidad y otros comportamientos de los conductores que podrían afectar indirectamente el consumo de combustible. La segunda tabla, **silver_idling**, almacena información detallada sobre el tiempo de ralentí de los vehículos, incluyendo las duraciones específicas y los costos asociados. Por último, **silver_consumption** es la tabla de referencia para la variable objetivo *consumption*, medida en millas por galón, e incluye además la distancia total recorrida y el total de combustible utilizado, aportando las métricas principales para evaluar la eficiencia de combustible en la flota de Forza .

El EDA tuvo como objetivos principales seleccionar características relevantes para el análisis de consumo y realizar una limpieza exhaustiva de los datos para asegurar su calidad y consistencia. En este proceso, se identificaron valores faltantes y ceros en métricas importantes, como *total_distance_used_sum* y *total_fuel_used_sum*, que son fundamentales para el cálculo de consumo. Los valores cero en estas métricas señalaban posibles errores o datos incompletos que planteaban desafíos al análisis. Para resolver esto, se optó por sustituir los valores cero por *NaN* en una copia del conjunto de datos y, tras su identificación, se eliminaron las filas incompletas. Esto permitió mantener un conjunto de datos de alta calidad que pudiera utilizarse de manera confiable en la construcción del modelo predictivo.

Para comprender las relaciones entre las variables y apoyar la selección de características, se utilizó el coeficiente de correlación de Pearson y se generaron mapas de correlación que facilitaron la visualización de patrones de asociación. En la tabla **silver_consumption**, se observó una fuerte correlación entre *total_distance_used_sum* y *total_fuel_used_sum*, lo que respalda la relación directa entre la distancia recorrida y el consumo de combustible. Este hallazgo fue clave para orientar la selección de variables relevantes en el modelo de predicción. En **silver_idling**, las variables *idlingDuration* y *idlingPercent* mostraron una correlación positiva significativa con *IdlingCost*, confirmando que tiempos de inactividad prolongados incrementan los costos operativos. Este hallazgo sugiere que, al reducir el tiempo de ralentí, Forza podría mejorar significativamente la eficiencia de combustible. Por otro lado, la tabla **bronze_safety** reveló que ciertos eventos de seguridad, como infracciones de velocidad y frenadas bruscas, tienden a ocurrir conjuntamente. Aunque estas variables no mostraron una relación directa fuerte con el consumo de combustible, proporcionan información valiosa para entender los patrones de comportamiento de los conductores.

En suma, el EDA nos llevó a concluir que las variables de distancia total y consumo total de combustible eran los factores de mayor relevancia, ya que las demás variables no mostraron relaciones significativas con la variable de salida, *Consumption*. Este hallazgo fue determinante en la orientación del modelo de clasificación y en la decisión de explorar un

modelo de forecast. El EDA, apoyado en visualizaciones y análisis detallado, proporcionó una base sólida para el enfoque del proyecto y preparó el contexto para la evaluación de los modelos desarrollados.

1.3 Modelos generados y razones de la elección del modelo final (Avance 5)

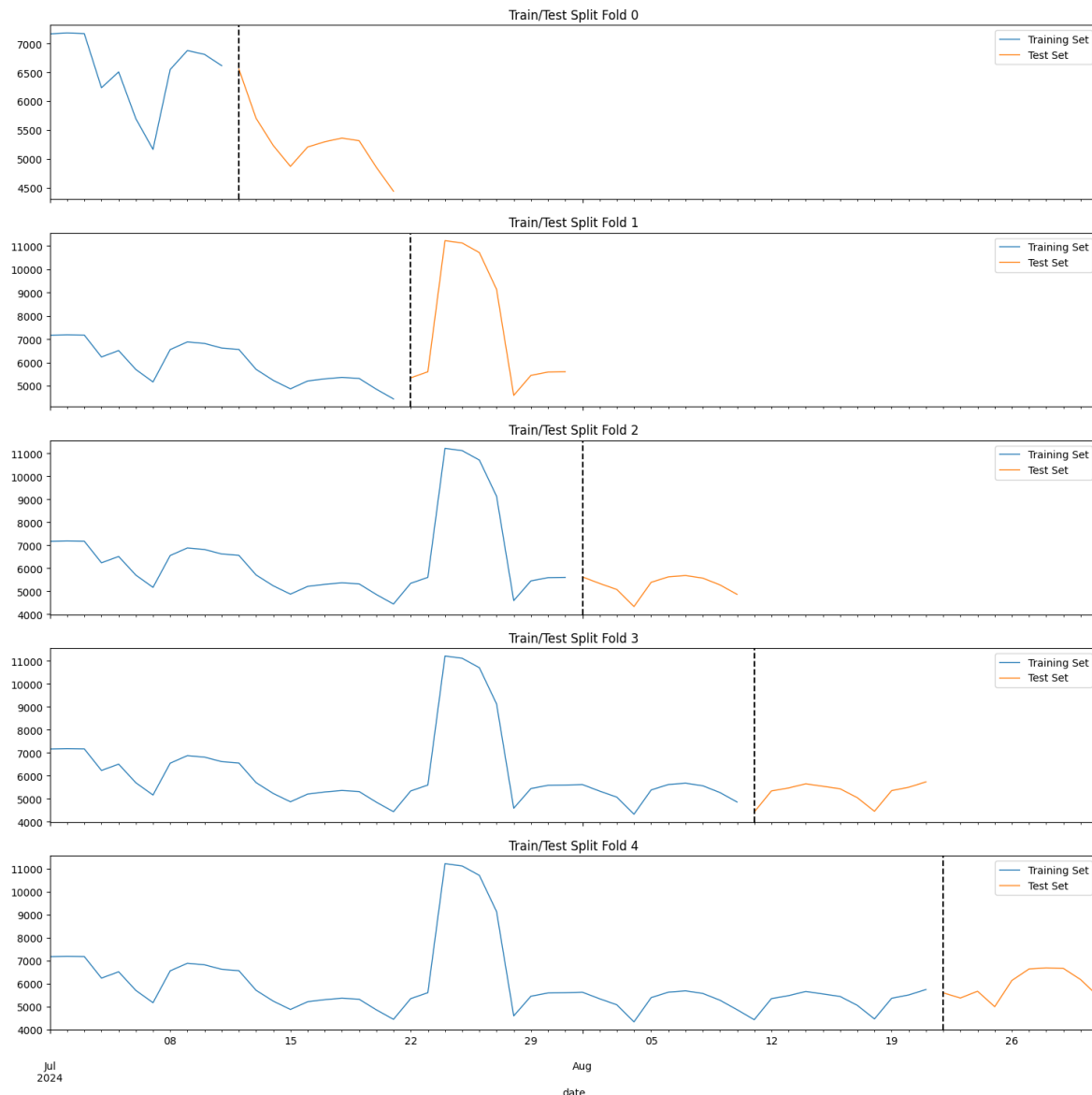
En la fase inicial del desarrollo de modelos, se construyeron y evaluaron dos enfoques de clasificación para determinar los factores clave en el consumo de combustible de la flota de Forza Transportation. Los modelos seleccionados fueron un árbol de decisión y un modelo XGBoost, cada uno optimizado para maximizar el rendimiento en clasificación. El árbol de decisión se eligió debido a su facilidad para interpretar las relaciones entre variables a través de una estructura jerárquica, mientras que XGBoost fue seleccionado por su capacidad para abordar problemas de clasificación complejos utilizando un enfoque de boosting, lo que mejora iterativamente el modelo al reducir los errores en cada etapa.

Cada modelo se entrenó con un conjunto de datos consolidado, que integraba información sobre consumo, tiempos de ralentí y eventos de seguridad. Para optimizar el rendimiento, se realizaron ajustes de hiperparámetros específicos. En el caso del árbol de decisión, se ajustaron la profundidad máxima del árbol y el criterio de partición, logrando una precisión promedio del 82.3% y un puntaje F1 de 81.0% en validación cruzada. En el modelo XGBoost, se optimizaron hiperparámetros como el *learning_rate*, el número de estimadores y la profundidad máxima de los árboles, alcanzando una precisión promedio de 87.2% y un puntaje F1 de 86.8%, superando en todas las métricas clave al árbol de decisión.

No obstante, ambos modelos mostraron limitaciones. La matriz de correlación, obtenida en el análisis exploratorio de datos, reveló que la mayoría de las variables, salvo *total_distance_used_sum* y *total_fuel_used_sum*, tenían una correlación débil con el consumo de combustible. Esto impactó el rendimiento, limitando la capacidad predictiva de ambos modelos. Aunque XGBoost demostró ser superior en precisión y consistencia, esta falta de variables adicionales significativas dificultó la clasificación precisa del consumo. Este hallazgo motivó la transición hacia un modelo de *forecasting*, que permitiera realizar predicciones continuas, capturando mejor las tendencias temporales y ofreciendo una mayor precisión en la toma de decisiones operativas.

El modelo final seleccionado fue un XGBoost optimizado para forecasting. Se aplicó una transformación de raíz cúbica para reducir la varianza de las variables y se estableció un pipeline de preprocesamiento que incluyó el escalado de variables numéricas y la ingeniería de características temporales, como el día de la semana y el mes, para capturar patrones cíclicos. Este modelo alcanzó un MAE promedio de 790.22 en la validación cruzada, manteniendo un MAE de 1390.40 en el conjunto de prueba final, lo que confirma su robustez y capacidad de generalización en datos no vistos.

La siguiente gráfica muestra la capacidad predictiva del modelo final.



1.4 Recomendaciones clave para implementar la solución (Avance 6)

Recomendaciones para la Optimización del Consumo de Combustible

Para asegurar que el modelo de forecasting de consumo de combustible de Forza Transportation sea implementado con éxito y mantenga su efectividad a lo largo del tiempo, es fundamental seguir una serie de recomendaciones estratégicas en las áreas de automatización de datos, monitoreo, actualización continua y gestión en la nube. Cada una de estas recomendaciones permitirá optimizar tanto la precisión como la sostenibilidad operativa del modelo.

En primer lugar, se recomienda establecer una automatización completa del pipeline de datos, abarcando desde la recopilación hasta el procesamiento y la integración de datos en tiempo

real. La automatización del pipeline permitirá que el modelo reciba datos frescos de manera constante y sin intervención manual, lo cual es esencial para mantener la precisión de las predicciones y minimizar el riesgo de errores humanos. Este flujo automatizado debe estar diseñado para procesar información de múltiples fuentes y unificarla, adaptándose a variaciones de datos como eventos de seguridad y condiciones operativas de la flota. Un pipeline automatizado no solo reduce la carga de trabajo en el mantenimiento de datos, sino que asegura que el modelo esté siempre actualizado, maximizando su utilidad en la toma de decisiones diarias.

Además de la automatización, es crucial implementar un sistema de monitoreo en tiempo real que permita evaluar el rendimiento del modelo continuamente. Este sistema de monitoreo debe supervisar métricas clave, como el Mean Absolute Error (MAE) y la precisión en las tendencias de consumo predichas. La implementación de alertas automáticas para cualquier desviación significativa en el MAE o en otras métricas ayudará a identificar y corregir problemas antes de que afecten la operación. De esta manera, el equipo de Forza puede responder rápidamente a cambios inesperados en el rendimiento del modelo, garantizando que este continúe siendo una herramienta fiable para gestionar el consumo de combustible. Además, el monitoreo debe incluir visualizaciones y reportes que faciliten la interpretación de las métricas y su impacto en el negocio, permitiendo un análisis detallado de los patrones de consumo de combustible y su evolución.

Para asegurar que el modelo permanezca vigente y adaptable a condiciones cambiantes, es fundamental planificar un ciclo de actualización continua del modelo. Esto implica realizar reentrenamientos periódicos del modelo con datos recientes, lo cual le permitirá adaptarse a cambios en rutas, condiciones de tráfico, características climáticas y otros factores externos que influyen en el consumo de combustible. En una industria tan dinámica como la del transporte, donde las condiciones operativas pueden variar de forma significativa, es esencial que el modelo tenga la capacidad de ajustarse y mantener la precisión en sus predicciones. Este enfoque proactivo en la actualización del modelo garantiza que las predicciones sean relevantes y que la herramienta de forecasting continúe generando valor para Forza a medida que evolucionan sus necesidades y el entorno de operación.

Detalles de infraestructura para la implementación:

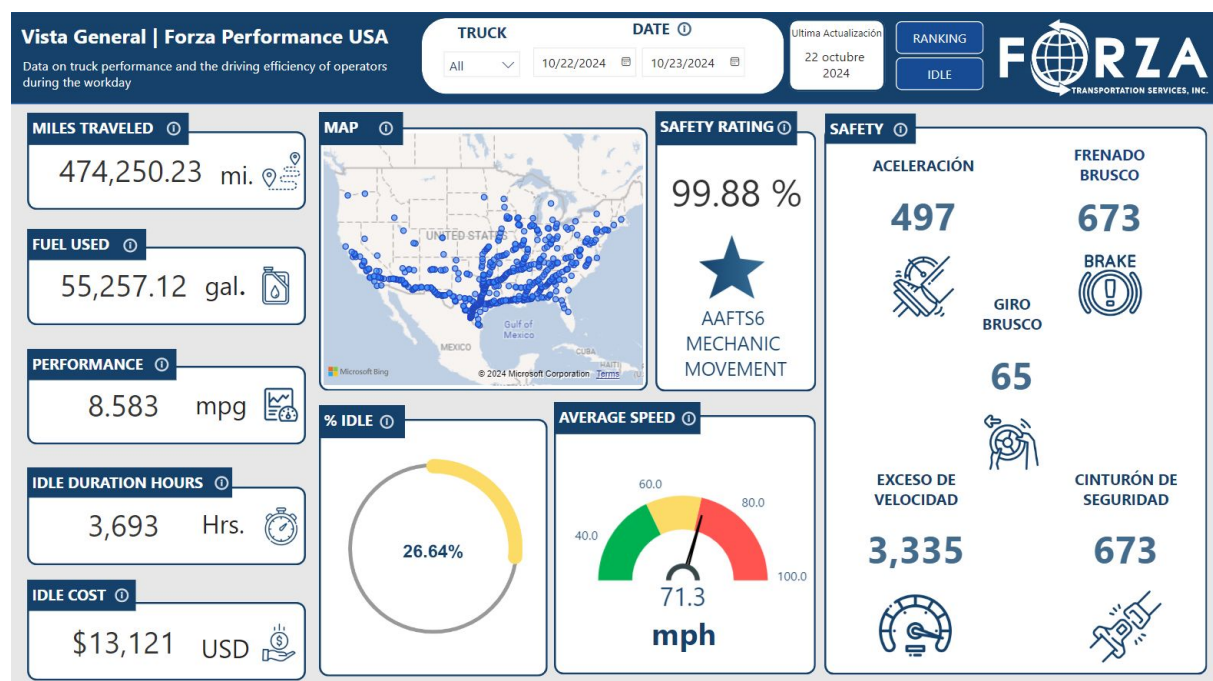
A nivel de infraestructura, se recomienda la utilización de Amazon Web Services (AWS) como plataforma en la nube para alojar y gestionar el modelo en producción. AWS ofrece una suite completa de servicios de machine learning, incluyendo Amazon SageMaker, que facilita no solo el entrenamiento y despliegue del modelo, sino también su monitoreo y mantenimiento. AWS también proporciona una infraestructura global robusta y escalable, lo que permitirá a Forza ajustar los recursos de acuerdo con la demanda, optimizando el uso de la infraestructura en función de las necesidades operativas. Esta flexibilidad es particularmente importante para Forza, dado el tamaño y la variabilidad de su flota, y la capacidad de escalabilidad de AWS asegura que el modelo pueda responder eficazmente a picos de uso sin comprometer su rendimiento. Adicionalmente, el uso de servicios de monitoreo en AWS, como Amazon CloudWatch, facilita la supervisión de la infraestructura, permitiendo detectar posibles problemas y actuar de inmediato.

Gestión de cambio con las personas de la empresa.

Para complementar esta infraestructura, es clave que el equipo de operaciones y técnicos de Forza reciba una capacitación integral en el uso y monitoreo del modelo. Esta capacitación debe cubrir desde el uso básico de la plataforma hasta la interpretación de las métricas y reportes de rendimiento. Es importante que el equipo no solo se familiarice con el funcionamiento técnico del modelo, sino que también comprenda cómo interpretar sus resultados y cómo estas predicciones pueden apoyar en la optimización de las rutas y en la toma de decisiones estratégicas. Además, se recomienda que esta capacitación incluya guías de mejores prácticas para responder a alertas y ajustar configuraciones del modelo en caso de cambios en las condiciones operativas. Así, el equipo podrá aprovechar al máximo las capacidades de la solución en la nube, optimizando la gestión de combustible y promoviendo una cultura de análisis de datos dentro de la organización.

Dashboard para análisis de información en tiempo real.

A partir de los datos generados por los sensores GeoTab y aplicando técnicas de limpieza y procesamiento de datos, se ha entregado, fruto de este proyecto, un dashboard para monitoreo en tiempo real utilizando Power BI. Este dashboard elimina la necesidad de procesamiento manual en Excel, proporcionando una visualización integral y en tiempo real de métricas críticas, tales como millas recorridas, consumo de combustible, velocidad promedio y eventos de seguridad. Gracias a esta herramienta, los responsables de Forza Transportation pueden tomar decisiones informadas de manera inmediata, optimizando así la eficiencia operativa y mejorando la seguridad en las rutas.



Análisis costo – beneficio

Establecer supuestos racionales para presentar algunas cifras sobre los costos/beneficios a las partes interesadas. Las siguientes preguntas pueden guiar el proceso:

1.5 Enumerar todos los costos incurridos

Para llevar a cabo el desarrollo del modelo de forecast y la implementación de la infraestructura necesaria, Forza incurrió en una serie de costos fundamentales que abarcan la adquisición de datos, infraestructura técnica y preparación de los recursos necesarios para el análisis.

- En primer lugar, la adquisición de datos de Geotab, incluyendo el acceso y la integración de los datos en tiempo real generados por los sensores de los camiones, tuvo un costo estimado de **\$20,000 USD**. Esta inversión inicial cubre la licencia de uso y los derechos de acceso a los datos durante los primeros seis meses del proyecto, lo que asegura que el equipo tenga acceso constante a la información necesaria para desarrollar y ajustar el modelo.
- En cuanto a la infraestructura técnica, Forza decidió emplear servidores en la nube para procesar el modelo de forecast y almacenar los datos históricos y en tiempo real generados por Geotab. Este servicio en la nube, contratado a través de un proveedor de servicios especializado, tiene un costo estimado de **\$15,000 USD** anuales, lo que incluye la capacidad de almacenamiento, el procesamiento de grandes volúmenes de datos, y las funciones de backup y recuperación de información.
- Además, se ha destinado un presupuesto de **\$5,000 USD** para la adquisición de software especializado para análisis de datos y visualización, incluyendo licencias para herramientas como Power BI, que se utilizarán en el dashboard de consulta rápida para el equipo de Forza.
- Para la preparación del modelo de forecast, se incurrieron en costos relacionados con la preprocesamiento y limpieza de los datos, la fase de modelado y la fase de evaluación. Estas actividades, distribuidas en las distintas etapas de la metodología CRISP-ML, implicaron un gasto total estimado de **\$10,000 USD** en términos de horas de trabajo del equipo de análisis y personal técnico. Esta inversión asegura que el modelo se entrene con datos precisos y que cumpla con los estándares necesarios para su puesta en producción y uso diario.

1.6 Costos esperados para la operación y mantenimiento del modelo

Para garantizar que el modelo de forecast se mantenga en funcionamiento y con alta precisión, Forza incurrirá en costos recurrentes relacionados con la operación y el mantenimiento de la infraestructura de datos y del modelo.

- En primer lugar, se estima que el costo anual de la infraestructura en la nube, que incluye el almacenamiento de datos y la capacidad de procesamiento, será de **\$15,000 USD**. Este costo cubrirá el almacenamiento de datos históricos y actuales de consumo de combustible, así como el procesamiento de las proyecciones del modelo en tiempo real, permitiendo a Forza contar con una infraestructura escalable y confiable.
- Además, el mantenimiento del modelo de forecast incluye la actualización y validación continua del modelo para asegurar que sus proyecciones se mantengan precisas y alineadas con las condiciones operativas de Forza. Este proceso de mantenimiento está estimado en **\$7,500 USD** anuales, lo que cubrirá los costos de reentrenamiento del modelo con datos nuevos, así como las pruebas de validación y la calibración del modelo para responder a cambios en los patrones de consumo de combustible.
- Forza también ha presupuestado un costo adicional de **\$5,000 USD** anuales para el monitoreo y supervisión del modelo en producción, integrando métricas y alertas para detectar desviaciones en el rendimiento del modelo. Estas actividades forman parte de la estrategia de MLOps, asegurando que el modelo se mantenga en condiciones óptimas y que los ajustes necesarios se realicen a tiempo.

En total, los costos de operación y mantenimiento del modelo de forecast para Forza suman aproximadamente **\$27,500 USD** anuales, lo cual representa una inversión clave para maximizar la eficacia y la confiabilidad del modelo a lo largo del tiempo.

1.7 Beneficios potenciales

El modelo de forecast y las estrategias de optimización implementadas ofrecen a Forza una serie de beneficios potenciales significativos, tanto en términos financieros como operativos.

- Uno de los beneficios clave es la reducción de costos de combustible mediante una mejor planificación de compras. Se estima que, al prever períodos de alta y baja demanda y realizar compras estratégicas en función de estos patrones, Forza podría ahorrar alrededor del 5% en sus costos de combustible anual, lo que equivale a aproximadamente **\$50,000 USD al año**, considerando el gasto promedio actual en combustible.
- Otro beneficio potencial radica en la optimización de rutas y la reducción del tiempo de inactividad, que permitiría una operación más eficiente de la flota. Al maximizar el uso de rutas predecibles y minimizar el consumo en tramos de alta demanda, se estima que Forza podría mejorar su eficiencia en al menos un 3%. Esta mejora operativa se traduce en una reducción de aproximadamente **\$30,000 USD** anuales en costos asociados al desgaste y mantenimiento de los vehículos, además de contribuir a una mayor durabilidad de la flota.
- El modelo de forecast también aporta valor a la toma de decisiones, proporcionando datos más precisos y actualizados que respaldan decisiones informadas sobre el uso

y asignación de recursos. Este nivel de precisión permite a los gerentes de operaciones realizar ajustes en tiempo real y responder de manera proactiva a variaciones en el consumo de combustible. La capacidad de tomar decisiones informadas en un entorno competitivo contribuye a mejorar la posición de Forza en el mercado y a incrementar la satisfacción del cliente, al asegurar que los servicios se presten de manera confiable y con menores costos operativos.

- Finalmente, el uso de datos predictivos abre oportunidades de negocio adicionales para Forza, como la posibilidad de ofrecer servicios de consultoría en optimización de consumo a otras empresas del sector. Este enfoque diversificado podría generar ingresos adicionales estimados en **\$20,000 USD** anuales, posicionando a Forza como un referente en innovación en el sector de transporte.

En total, los beneficios potenciales del proyecto ascienden a aproximadamente **\$100,000 USD anuales**, integrando ahorros en combustible, eficiencia operativa y nuevas oportunidades de negocio. Estos beneficios consolidan el impacto del modelo de forecast en la competitividad y rentabilidad de Forza, asegurando un retorno sólido de la inversión en optimización.

2 Riesgos y desafíos de la solución

2.1 ¿Cuáles son los riesgos o desafíos potenciales de la solución propuesta?

La implementación de la solución basada en el modelo de forecast enfrenta varios desafíos técnicos y organizativos que Forza debe abordar para maximizar el impacto del proyecto. Uno de los desafíos más destacados es la integración del modelo en las operaciones diarias de Forza. Esto requiere una infraestructura de datos que permita el acceso y la actualización constante de la información de consumo y rutas. El modelo necesita un flujo continuo de datos de Geotab para que sus proyecciones sean precisas, y cualquier interrupción en este flujo podría afectar la capacidad de Forza para utilizar el forecast de manera efectiva.

Otro desafío importante es la capacitación del equipo operativo en el uso y la interpretación de las proyecciones generadas por el forecast. Aunque el modelo de forecast está diseñado para ofrecer una visión clara del consumo futuro, su correcta interpretación requiere que el equipo comprenda los conceptos básicos de predicción y los factores que afectan la precisión de las proyecciones. La falta de capacitación adecuada podría llevar a interpretaciones incorrectas de los datos, lo que afectaría la toma de decisiones operativas. Por ello, se recomienda realizar talleres de capacitación periódicos y proporcionar guías que faciliten la interpretación de las proyecciones y su aplicación práctica en la planificación de rutas y compras de combustible.

La integración de prácticas de MLOps representa otro desafío. Aunque MLOps ofrece una metodología clara para gestionar el ciclo de vida del modelo, su implementación requiere una infraestructura y procesos adicionales que Forza puede no tener en su sistema actual. Establecer pipelines automáticos de reentrenamiento, monitoreo y alertas exige una inversión en recursos técnicos y talento especializado que garantice que el modelo se mantenga

actualizado y alineado con las condiciones operativas cambiantes. Para superar este desafío, Forza podría optar por soluciones modulares de MLOps o servicios externos de monitoreo y mantenimiento que complementen las capacidades de su equipo interno.

Finalmente, un desafío importante en la implementación es garantizar la aceptación del modelo y de las prácticas de forecast en la cultura organizacional de Forza. La adopción de soluciones basadas en datos y predicciones a menudo enfrenta resistencia en las organizaciones, especialmente en industrias donde las decisiones se han basado tradicionalmente en la experiencia y el juicio humano. Para asegurar la aceptación de la solución, se recomienda un enfoque de cambio cultural que muestre los beneficios tangibles del forecast y fomente la colaboración entre el equipo de operaciones y el equipo de datos.

2.2 Clasificar los riesgos identificados en las categorías provistas en los recursos de la semana (relacionados con los datos, ataques, prueba y confianza, cumplimiento)

Riesgos Relacionados con los Datos

Uno de los principales riesgos asociados con el modelo de forecast es la calidad y disponibilidad de los datos provenientes de Geotab. La precisión del modelo depende directamente de la fiabilidad y consistencia de estos datos, y cualquier variación en la calidad o interrupción en el flujo de información podría comprometer la capacidad del forecast para proyectar el consumo de combustible con exactitud. Además, la falta de datos adicionales que puedan capturar otros factores de consumo representa un riesgo, ya que limita la capacidad del modelo para adaptarse a cambios en las condiciones operativas. Para mitigar este riesgo, se recomienda implementar un sistema de monitoreo de calidad de datos que detecte y corrija posibles anomalías en tiempo real.

Riesgos de Seguridad (Ataques)

El uso de datos de telemetría y la dependencia en la infraestructura de la nube para el procesamiento de información exponen al sistema a posibles ciberataques. Estos ataques podrían comprometer tanto los datos como la infraestructura de procesamiento, afectando la disponibilidad y confidencialidad de la información. Un ataque exitoso podría interrumpir la operación del forecast y poner en riesgo datos sensibles de Forza, lo que podría derivar en pérdidas operativas y reputacionales. Para gestionar este riesgo, es fundamental implementar controles de seguridad robustos, como firewalls y sistemas de autenticación multifactor, además de realizar auditorías regulares de seguridad en la infraestructura.

Prueba y Confianza

Un desafío importante en la implementación del forecast es mantener la precisión del modelo en el tiempo. Los cambios en las condiciones de operación, como alteraciones en las rutas o variaciones en los precios de combustible, podrían disminuir la exactitud de las proyecciones

si el modelo no se ajusta adecuadamente. Esta falta de confianza en las predicciones afectaría su uso en la toma de decisiones operativas. Para mitigar este riesgo, se recomienda un programa de reentrenamiento regular del modelo y el monitoreo de métricas de rendimiento que permitan ajustar el forecast de manera oportuna.

Cumplimiento Normativo

El manejo de datos de telemetría y de consumo de combustible puede estar sujeto a regulaciones específicas en ciertas regiones, especialmente si Forza planea expandir el uso del modelo a nivel internacional. Las normativas de privacidad de datos y los requisitos de almacenamiento de información varían entre países, y el incumplimiento de estos reglamentos podría resultar en sanciones y restricciones operativas. Para prevenir este riesgo, Forza debería establecer políticas de cumplimiento que alineen la recopilación y uso de datos con las normativas locales y realizar auditorías periódicas de cumplimiento.

3 Bibliografía

- Visengeriyeva, L., Kammer, A., Bär, I., Kniesz, A., y Plöd, M. (2023). CRISP-ML(Q). The ML Lifecycle Process. MLOps. INNOQ. [https://ml-ops.org/content/crisp-mlLinks to an external site.](https://ml-ops.org/content/crisp-mlLinks%20to%20an%20external%20site)
- Mssaperla. (2024, 1 marzo). *What is the medallion lakehouse architecture?* - Azure Databricks. Microsoft Learn. <https://learn.microsoft.com/en-us/azure/databricks/lakehouse/medallion>