

Marleny Barrueta¹, Karina Torres² y Jefferson Burga³

¹ Autor: Antamina, Av. El Derby 055 - Torre 1 (oficina 801), Lima, Perú (mbarrueta@antamina.com y 995904710)

² Coautor 1: Antamina, Av. El Derby 055 - Torre 1 (oficina 801), Lima, Perú, (ktorres@antamina.com y 926687570)

³ Coautor 2: Antamina, Av. El Derby 055 - Torre 1 (oficina 801), Lima, Perú, (jburga@antamina.com y 947650850)

RESUMEN

El presente trabajo aborda la problemática asociada a la asignación de camiones de acarreo en las estaciones de servicio (grifos) en la operación minera de Antamina, cuya flota actual asciende a 150 unidades, compuesta por equipos Komatsu 930 y 980, así como CAT 793 y 798. Durante el periodo 2023-2024, se identificó que el 60% de las asignaciones se realizaban de forma manual, lo cual generaba ineficiencias operativas evidenciadas en tiempos prolongados de espera y formación de colas, impactando negativamente en la disponibilidad y productividad de la flota.

Ante esta situación, se planteó como objetivo principal la optimización del sistema de asignación de camiones a grifos mediante el análisis de rutas, así como la reducción de tiempos de espera sin comprometer el abastecimiento continuo de combustible. La propuesta consideró variables operativas críticas y se apoyó en herramientas de análisis de datos y modelamiento operativo. Los resultados obtenidos reflejan mejoras significativas en la eficiencia del proceso de asignación para el abastecimiento de combustible en grifo, reduciendo significativamente el uso de cisternas en las vías, logrando un impacto positivo en la utilización de la flota y en los indicadores globales de operación. Asimismo, este trabajo tiene como propósito documentar la incorporación de tecnología en cada una de las etapas del proceso desarrollado, así como identificar y analizar los puntos críticos que deben ser considerados para una implementación eficaz a esta escala.

1. Introducción

En una operación minera de alta complejidad como la de Antamina, donde cada minuto cuenta y la eficiencia es sinónimo de seguridad y rentabilidad, la gestión inteligente del abastecimiento de combustible representa un desafío estratégico. Con una flota de 150 camiones de acarreo —entre

Komatsu 930/980 y CAT 793/798—, el abastecimiento eficiente se convierte en un factor crítico para evitar pérdidas operativas significativas.

Durante el periodo 2023-2024, un cuello de botella recurrente afectó directamente la productividad de la operación: la asignación manual de camiones a las estaciones de servicio (grifos), generando largas colas, tiempos muertos y una utilización subóptima de los activos más valiosos de la mina. Frente a este escenario, nació la necesidad de transformar un proceso tradicional en uno inteligente y automatizado.

Este proyecto propone una solución disruptiva basada en técnicas avanzadas de Machine Learning (ML) e Inteligencia Artificial (IA), con el objetivo de predecir con precisión el nivel de combustible de cada camión y asignar dinámicamente la ruta óptima hacia el grifo más adecuado. A través de modelos predictivos como LightGBM y algoritmos de optimización como Dijkstra, se logró no solo reducir los tiempos de espera, sino también liberar capacidad operativa equivalente a entre 250,000 y 350,000 toneladas adicionales por año.

Más que una mejora tecnológica, esta solución representa un cambio de paradigma en la operación minera: de decisiones reactivas a decisiones inteligentes, basadas en datos. Una iniciativa que eleva la eficiencia, mejora la seguridad del personal y sienta las bases para una operación cada vez más autónoma y sostenible.

2. Objetivos

- ❖ Optimizar la asignación de camiones en las estaciones de servicio mediante un sistema inteligente que reduzca los tiempos de espera, mejorando así la disponibilidad y eficiencia de la flota.

- ❖ Automatizar los procesos de asignación reemplazando métodos manuales por algoritmos basados en inteligencia artificial y machine learning que permitan decisiones rápidas, precisas y predictivas sobre el nivel de combustible y la programación del abastecimiento.
- ❖ Convertir los datos operativos en información estratégica mediante analítica predictiva avanzada, facilitando la toma de decisiones informadas y en tiempo real para maximizar la productividad y minimizar riesgos en la operación.
- ❖ Brindar al lector una comprensión clara del diseño, implementación y beneficios del aplicativo Refueling, resaltando su impacto en la optimización operativa y la gestión eficiente de recursos en entornos mineros.

3. Compilación de Datos y Desarrollo del Trabajo

Actualmente, la operación en Antamina cuenta con tres estaciones de abastecimiento de combustible, distribuidas estratégicamente para soportar la flota de camiones de acarreo. Sin embargo, una restricción importante radica en que no todos los camiones pueden abastecerse en cualquier estación, debido a factores como el tipo de toma de combustible y otras condiciones operativas específicas. Esta limitación implica que ciertos vehículos estén asignados exclusivamente a determinadas estaciones, lo que genera un reto adicional para la planificación eficiente del reabastecimiento.

3.1 Asignación de camiones a grifos

La asignación eficiente de los camiones a las estaciones de grifo es un proceso fundamental para garantizar que la flota se mantenga operativa y sin interrupciones durante el ciclo de acarreo. Para ello, el sistema Dispatch incorpora un módulo automatizado de gestión de combustible que facilita esta tarea. Sin embargo, cerca del 60% de estas asignaciones aún se realizan manualmente. Esta situación se debe a problemas como lecturas inexactas del nivel de combustible provocadas por fallas en los sensores, problemas de conectividad y otras limitaciones propias de la operación. Estas dificultades afectan la correcta priorización de los

camiones que necesitan reabastecerse, lo que a su vez genera mayores tiempos de espera en las estaciones y reduce la eficiencia general del proceso.

Tal como se observa en la Figura 1, a lo largo de los meses se evidenció un incremento sostenido en el tiempo de espera de los camiones en cola, lo que refuerza la necesidad de implementar una solución que revierta esta situación



Figura 1. Duración de tiempo de cola en grifo (verde año 2022, rojo año 2023)

3.2 Metodología

Frente a los desafíos identificados en la asignación y gestión del reabastecimiento de combustible, la metodología desarrollada para este proyecto combina técnicas avanzadas de optimización e inteligencia artificial, estructuradas en dos etapas complementarias.

En la primera etapa, se diseñó un **modelo de inferencia** capaz de predecir con mayor precisión el nivel de combustible en los tanques de los camiones mineros. Este modelo se basa en el análisis de variables clave que afectan el consumo, lo que permite anticipar la necesidad de reabastecimiento con mayor exactitud que los métodos tradicionales.

La segunda etapa involucra la construcción de un **modelo de optimización basado en grafos**, que utiliza las estimaciones del nivel de combustible junto con las condiciones operativas y logísticas propias de la mina. Este modelo asigna de forma dinámica y eficiente los camiones a las estaciones de grifo, priorizando la reducción de tiempos de espera y maximizando la continuidad del ciclo de acarreo.

La integración de ambos enfoques no solo busca disminuir los retrasos evidenciados en la Figura 1, sino también elevar la eficiencia operativa general y fortalecer la seguridad dentro de la operación

minera, transformando datos operativos en decisiones inteligentes y oportunas.

3.2.1 Modelo de Inferencia

i. Selección de fuentes de datos:

Para el desarrollo del modelo de inferencia, se utilizaron datos provenientes de dos sistemas fundamentales que soportan la operación en Antamina:

- Sistema Dispatch: Plataforma interna que gestiona de manera continua el flujo de información operativa. Este sistema integra datos en tiempo real y registros históricos para actualizar los estados de los camiones y facilitar la toma de decisiones en la asignación y planificación de las actividades de acarreo.
- Sistema Orpak: Sistema externo, administrado por el proveedor encargado del suministro de combustible en la mina. Este registra con alta precisión los eventos de abastecimiento de combustible para las unidades pesadas y ligeras, proporcionando datos fiables sobre el consumo y la logística de reabastecimiento.

La combinación de estos conjuntos de datos permitió construir un modelo robusto que aprovecha tanto la información operacional continua como los registros detallados de combustible, esenciales para predecir con mayor exactitud los niveles de combustible y optimizar las decisiones de reabastecimiento.

ii. Selección de variables:

Para el desarrollo del modelo de inferencia, se seleccionaron un conjunto de variables relevantes que influyen directamente en el nivel de combustible de los camiones, tanto del comportamiento operacional como de los registros de abastecimiento:

Variables operativas asociadas al camión:

- Distancia total recorrida desde el último reabastecimiento.
- Inclinación del trayecto a lo largo del recorrido.
- Velocidad promedio del camión durante el viaje.

- Carga transportada (en toneladas) en cada ciclo entre puntos de carga y descarga desde el último abastecimiento.
- Tipo de baliza utilizada para la comunicación y seguimiento.
- Nivel de combustible (%) reportado por el sensor instalado en el vehículo.

Variables de abastecimiento (registradas por el Sistema Orpak):

- Fecha y hora de inicio del abastecimiento.
- Fecha y hora de finalización del abastecimiento.
- Identificación del camión abastecido.
- Cantidad de combustible suministrada (medida en galones).

La integración de estas variables garantiza un modelo robusto, capaz de capturar las dinámicas reales del consumo y reabastecimiento, mejorando la precisión de las predicciones del nivel de combustible.

iii. Periodo de análisis

El conjunto de datos utilizado para el desarrollo y validación del modelo abarca un período comprendido entre el 1 de agosto y el 15 de octubre de 2023, incluyendo diversas condiciones operativas de la mina que permiten evaluar el desempeño del sistema bajo escenarios representativos y variables.

Durante este intervalo, se recopilaron variables críticas relacionadas con la flota, tales como la cantidad de camiones activos y el volumen total de combustible suministrado (ver tabla 1), segmentado por tipo de flota. Asimismo, se realizó un análisis detallado sobre la distribución del combustible remanente en los tanques justo antes del reabastecimiento en grifo (ver Figura 2), así como el tiempo promedio de abastecimiento por tipo de unidad.

Como parte del análisis operativo, se detectaron diferencias significativas en los tiempos promedio entre eventos de reabastecimiento según la clasificación de la flota. Los valores reflejados en la tabla 2 representan el intervalo promedio

transcurrido entre dos abastecimientos consecutivos para cada camión, lo cual es un indicador clave para optimizar la planificación del ciclo de reabastecimiento.

En cuanto a la periodicidad de los reabastecimientos, a partir del nivel de combustible medido por los sensores instalados en los camiones, se observa un patrón decreciente constante a lo largo del tiempo. Este comportamiento es interrumpido por incrementos abruptos en el nivel de combustible, que corresponden a eventos de reabastecimiento.

Flota	# Camiones	Capacidad de tanque
CAT 793	27	1200 y 1300
CAT 798	6	2000
KOM930	92	1400
KOM980	26	1400

Tabla 1. Distribución de camiones por flota y capacidad de tanque

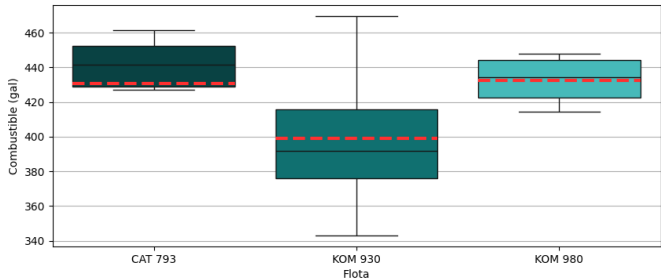


Figura 2. Distribución del combustible remanente previo al reabastecimiento

Flota	Intervalo de reabastecimiento (Hrs)
CAT 793	25.6
KOM930	19.72
KOM980	17.79

Tabla 2. Tiempo promedio de abastecimiento por flota

La Figura 3 muestra la evolución del nivel de combustible (%) para el camión HT320, con una capacidad de 1400 galones, durante el periodo comprendido entre el 30 de junio y el 3 de julio de 2025. Se aprecia un patrón cíclico descendente interrumpido por picos marcados con puntos rojos, que corresponden a los registros del sistema Orpak y señalan el momento exacto del reabastecimiento. Este análisis permite comprender el comportamiento de consumo y definir con precisión la duración de cada ciclo operativo, aportando datos valiosos para la calibración y mejora del modelo predictivo.

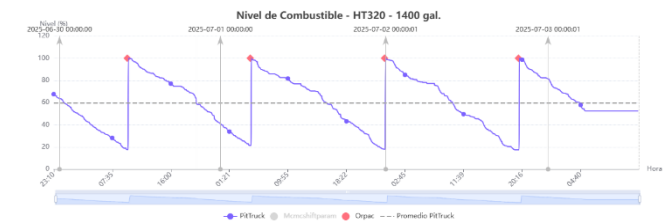


Figura 3. Evolución del nivel de combustible camión HT320

iv. Construcción del Target

Definición del objetivo

El propósito principal del modelo es estimar con precisión el nivel de combustible restante en el tanque de cada camión en tiempo real. Conociendo la capacidad total del tanque para cada unidad, es posible derivar el nivel remanente a partir del consumo acumulado utilizando una fórmula basada en la diferencia entre la capacidad total y el combustible consumido.

% Nivel combustible = $\left(\frac{Capacidad\ tanque - Combustible\ consumido}{Capacidad\ tanque} \right) \times 100\%$

Procedimiento para la construcción del target:

La generación de la variable objetivo (target) se realiza mediante una reconstrucción secuencial del consumo de combustible. Este proceso comienza en cada evento de abastecimiento registrado, retrocediendo en el tiempo hasta alcanzar el evento previo de recarga.

Los puntos de abastecimiento están claramente delimitados en los registros del sistema Orpak, que provee tanto la fecha y hora como la cantidad exacta de combustible suministrado. Esta información precisa permite establecer con exactitud el inicio de cada nuevo ciclo de consumo.

A partir de estos puntos de referencia, se calcula el consumo acumulado de combustible en cada instante intermedio, entre eventos consecutivos de reabastecimiento. Este cálculo secuencial genera la variable objetivo que será utilizada para el entrenamiento del modelo predictivo, asegurando que refleje el comportamiento real de consumo de cada camión.

En secciones posteriores se describirá el flujo completo para la creación de las variables predictoras y su vinculación directa con esta variable objetivo, completando así la estructura necesaria para el entrenamiento y validación del modelo.

v. Procedimiento para la construcción de variables predictoras y su relación con el target

La elaboración del conjunto de datos (dataset) sobre el cual se entrenará, validará y evaluará el modelo se llevó a cabo en varias etapas secuenciales, integrando técnicas avanzadas de ingeniería de características (feature engineering) para maximizar la capacidad predictiva del modelo con respecto al consumo de combustible.

Feature Engineering

A partir de los registros operativos provenientes del sistema Dispatch —que incluyen información como identificador del camión, fecha y hora del evento (timestamp), posición geográfica y carga transportada (en toneladas)— se aplicaron una serie de transformaciones para construir las variables explicativas del modelo. Este proceso se estructuró del siguiente modo:

- **Ordenamiento temporal:** Los registros fueron ordenados cronológicamente por timestamp para preservar la secuencia de eventos operativos y asegurar coherencia en la evolución del consumo.
- **Cálculo de distancias:** Se aplicaron métodos geoespaciales como Haversine (anexo 1) y proyecciones UTM (Universal Transverse Mercator) para calcular con mayor precisión la distancia recorrida entre puntos consecutivos de la trayectoria.
- **Generación de variables diferenciales:** Se calcularon variables clave entre cada par de puntos, tales como distancia recorrida, velocidad promedio, pendiente del trayecto y carga transportada. Estas variables permiten capturar el esfuerzo operativo y su relación con el consumo energético.
- **Acumulación desde el último reabastecimiento:** Se agregaron variables acumulativas por camión, que cuantifican el comportamiento desde el último evento de abastecimiento hasta el punto actual, alineando así las predicciones con los patrones de consumo observados.co
- **Preprocesamiento adicional:** Para garantizar la estabilidad del modelo, las variables fueron normalizadas y se aplicó la técnica de winsorización a fin de reducir la sensibilidad ante valores atípicos y mejorar la robustez del entrenamiento, especialmente en modelos basados en secuencias.

El resultado final de este procedimiento es un conjunto de datos estructurado, compuesto por variables predictoras acumuladas, transformadas y normalizadas, directamente vinculadas a la variable objetivo (nivel de combustible remanente). Este dataset constituye la base para el entrenamiento del modelo de inferencia y garantiza la coherencia entre los datos operativos y el objetivo de predicción.

vi. Suavizado de la data del sistema Dispatch

Como complemento a la construcción del modelo de inferencia, se incorporó una fuente adicional de datos sobre el nivel de combustible, proveniente directamente de una de las tablas internas del sistema Dispatch. No obstante, esta variable presenta un alto nivel de ruido debido a inconsistencias en la lectura de sensores, fluctuaciones operativas y posibles errores de transmisión, lo que afecta su utilidad directa para el entrenamiento del modelo.

Para abordar esta problemática, se implementó un proceso de suavizado orientado a estabilizar la señal y extraer una representación más fiel del comportamiento real del consumo de combustible. Este procedimiento incluyó los siguientes pasos:

Aplicación del filtro Savitzky–Golay: Se utilizó este filtro digital de regresión polinómica local, ampliamente reconocido en el procesamiento de señales, que permite suavizar la curva de datos sin distorsionar las características principales de la señal, como el ancho y la forma de los picos. Esta técnica preserva la estructura subyacente del consumo, facilitando su análisis sin introducir artefactos artificiales.

Monotonización decreciente: Como medida adicional, se aplicó una monotonización estricta con tendencia descendente sobre la señal suavizada, asegurando que la curva resultante refleje un consumo continuo y realista de combustible entre eventos de reabastecimiento, sin incrementos espurios.

Este proceso permitió mejorar significativamente la calidad de los datos provenientes del sistema Dispatch, facilitando su comparación con los registros de Orpak y permitiendo la validación cruzada del modelo de inferencia. A continuación, se presentan ejemplos visuales para dos camiones:

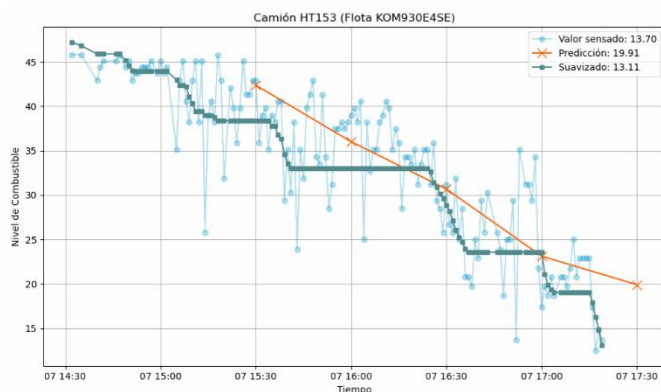


Figura 4. Resultado de la técnica de suavización aplicada al nivel de combustible del camión HT153

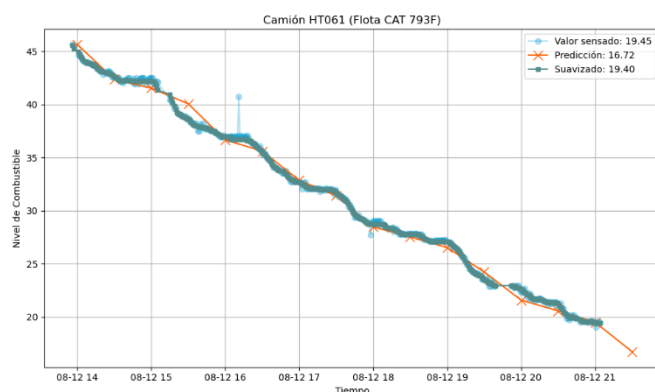


Figura 5. Resultado de la técnica de suavización aplicada al nivel de combustible del camión HT061

Aunque el proceso de suavización mejora significativamente la calidad del dato de nivel de combustible proveniente del sistema Dispatch, esta señal no es utilizada como variable objetivo (target) del modelo. Esto se debe a que su precisión está sujeta a la calidad de los puntos generados por el algoritmo interno del sistema, los cuales pueden diferir de los datos reales de reabastecimiento registrados en el sistema Orpak.

En consecuencia, esta curva suavizada se incorpora al conjunto de variables predictoras como una señal complementaria que aporta información útil sobre la tendencia de consumo, pero no se considera como referencia para la predicción del valor objetivo (target).

Asignación de la variable objetivo o target

Para construir el target, se parte del conjunto de variables acumuladas desde el último evento de abastecimiento, tales como la distancia recorrida, velocidad promedio, pendiente del terreno y carga transportada. Estas variables, consideradas en conjunto, se interpretan como un vector que representa el esfuerzo operativo del camión

durante el periodo comprendido entre dos eventos consecutivos de reabastecimiento.

A medida que el camión se aproxima a su siguiente abastecimiento, dicho vector crece en magnitud, y su módulo final corresponde al volumen total de combustible consumido en ese intervalo (por ejemplo, 980 galones). Utilizando este valor como referencia, se calcula el target para cada fila previa del historial, aplicando una proporción directa basada en la norma euclidiana del vector acumulado de esfuerzo. De esta manera, se asigna a cada instante una estimación del nivel de combustible restante, alineada con la dinámica operativa real del camión.

Este procedimiento se aplica sistemáticamente a todos los periodos históricos donde se dispone de información confiable de abastecimiento proveniente del sistema Orpak, asegurando consistencia y precisión en la generación del target.

Durante la etapa de inferencia en tiempo real, si bien el evento de reabastecimiento aún no ha ocurrido, las variables predictoras acumuladas se pueden calcular de forma análoga, permitiendo al modelo realizar una estimación del nivel actual de combustible en base a la progresión del consumo operativo.

Finalmente, se incorporan variables categóricas como el tipo de camión y el tipo de flota, que capturan características estructurales del comportamiento de consumo y fortalecen la capacidad explicativa del modelo.

vii. Técnicas adicionales en el preprocesamiento

Una vez definida la variable objetivo y construidas las variables predictoras —tanto acumulativas como categóricas—, se completó el conjunto de datos con un conjunto de transformaciones adicionales orientadas a mejorar la calidad de la información y la robustez del modelo durante el entrenamiento.

Estas técnicas de preprocesamiento permiten mitigar el efecto de registros anómalos y valores faltantes, así como asegurar la consistencia entre los datos operacionales y las estimaciones generadas por el modelo.

Para garantizar que el modelo trabaje con datos representativos y libres de distorsiones, se aplicaron las siguientes técnicas:

Winsorización: Esta técnica se utilizó para tratar valores atípicos detectados en algunas variables numéricas, que resultaban inconsistentes con la lógica operativa del desplazamiento del camión. Para cada variable, se calcularon los rangos intercuartílicos, y cualquier valor fuera de este rango fue reemplazado por el valor extremo más cercano dentro del umbral aceptable. Aunque los registros afectados representaban menos del 3% del total, su impacto era significativo, ya que podían introducir picos artificiales de consumo en cortos periodos, afectando la precisión del modelo.

Imputación de valores faltantes: En ciertos casos, la variable de carga reportaba valores nulos o igual a cero mientras el camión aún se desplazaba hacia un botadero, lo cual no era consistente con el ciclo de acarreo. En estas situaciones, se imputó la carga faltante utilizando la mediana histórica de carga asociada a cada camión, asegurando así una representación más realista de las condiciones de operación.

Estas acciones de limpieza y corrección de datos fueron fundamentales para preparar un dataset confiable y libre de sesgos extremos, favoreciendo un entrenamiento más estable y reduciendo el riesgo de sobreajuste a condiciones anómalas.

viii. Partición del conjunto de datos

Para garantizar una evaluación rigurosa del modelo de inferencia y evitar problemas de sobreajuste, el conjunto de datos fue dividido en tres subconjuntos independientes:

- Entrenamiento (Train): 70%
- Validación (Validation): 20%
- Prueba (Test): 10%

Esta partición no se realiza a nivel de registros individuales, sino sobre los periodos de consumo identificados para cada camión. Un periodo se define como la secuencia temporal de registros comprendida entre dos eventos consecutivos de reabastecimiento de combustible, iniciando justo después del fin de un evento de abastecimiento y

finalizando al inicio del siguiente. Esta estructura de partición por periodos completos permite preservar la naturaleza secuencial y acumulativa de las variables predictoras, evitando la fuga de información entre los conjuntos de entrenamiento y prueba.

Cada subconjunto tiene un propósito específico en el flujo de modelado:

El conjunto de entrenamiento se utiliza para ajustar los parámetros internos del modelo.

El conjunto de validación sirve para afinar los hiperparámetros y prevenir el sobreajuste, evaluando el rendimiento sobre datos no vistos durante el entrenamiento.

Finalmente, el conjunto de prueba proporciona una medida objetiva del desempeño del modelo sobre datos completamente nuevos, replicando condiciones reales de operación.

En la tabla 3, se presenta un ejemplo ilustrativo de cómo se estructura un periodo para un camión específico, indicando el inicio, fin, y la cantidad de registros contenidos en dicho intervalo:

Truck	# Periodos	Periodos para Train	Periodos para Validation	Periodos para Test
HT079	75	52	15	8
HT115	75	52	15	8
HT116	75	52	15	8
HT117	65	45	13	7
HT118	67	46.6	13.4	7
HT119	79	55.2	15.8	8
HT120	73	50.4	14.6	8
HT121	68	47.4	13.6	7

Tabla 3. Estructura temporal de datos por camión para modelado

ix. Resultados de los algoritmos aplicados

Para la construcción del modelo de inferencia del nivel de combustible remanente, se evaluaron diferentes algoritmos de regresión, priorizando tanto precisión como interpretabilidad. Tras un proceso comparativo, los modelos que presentaron el mejor desempeño durante la fase de validación fueron:

Regresión lineal multivariable, como modelo base, por su simplicidad y utilidad como línea base comparativa.

LightGBM (Light Gradient Boosting Machine), un modelo basado en árboles de decisión con enfoque en eficiencia y rendimiento, especialmente adecuado para grandes volúmenes de datos y estructuras secuenciales como las que se presentan en la operación minera.

El modelo LightGBM fue finalmente seleccionado como el modelo principal, debido a que logró un mejor equilibrio entre exactitud, capacidad de generalización y robustez ante datos heterogéneos. Su entrenamiento se realizó utilizando el conjunto de datos previamente procesado, incorporando variables acumuladas, transformadas y suavizadas, así como atributos categóricos relevantes.

Interpretabilidad del modelo: SHAP values
Para comprender el aporte relativo de cada variable en la predicción, se aplicó la técnica de interpretabilidad SHAP (SHapley Additive exPlanations). Esta metodología, basada en teoría de juegos, permite cuantificar el impacto de cada variable en la predicción individual del modelo.

En la tabla 4, los resultados mostraron que variables como distancia acumulada desde el último abastecimiento, carga transportada, pendiente promedio del trayecto, y el nivel de combustible suavizado (Dispatch) tuvieron una alta contribución al desempeño del modelo. También se identificó la influencia significativa de variables categóricas como el tipo de camión, lo cual reafirma la importancia de ajustar los modelos de consumo según las características específicas de cada unidad.

En el siguiente apartado se presentan los resultados cuantitativos de las visualizaciones SHAP y las métricas de desempeño que respaldan la selección final del modelo.

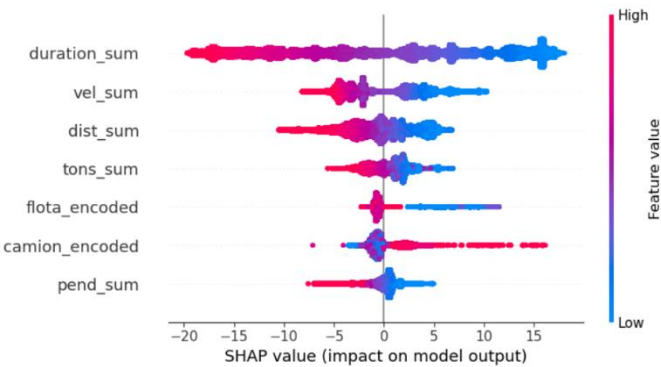


Figura 5. Valores SHAP (impacto en la salida del modelo)

Variables	Mean SHAP	Contribución (%)
duration_sum	9.57	46.51
vel_sum	3.55	17.28
dist_sum	2.59	12.58
tons_sum	1.53	7.45
flota_encoded	1.42	6.89
camion_encoded	1.12	5.43
pend_sum	0.79	3.85

Tabla 4. Importancia de variables según valores SHAP en el modelo LightGBM

x. Resultados del análisis del modelo de inferencia:

La variable duration_sum —tiempo acumulado en operación desde el último evento de reabastecimiento— resultó ser la más influyente en las predicciones del modelo, concentrando aproximadamente el 46.5% del peso total en la explicación de las salidas. Este resultado evidencia que el tiempo operativo transcurrido es un factor altamente representativo del consumo de combustible.

Otras variables operativas de tipo acumulativo, como vel_sum (velocidad promedio acumulada) y dist_sum (distancia total recorrida), también mostraron una contribución significativa, al capturar aspectos clave del esfuerzo mecánico y el desplazamiento efectivo de los camiones.

Factores adicionales como el tonelaje transportado y la pendiente acumulada del trayecto complementan la explicación del consumo, reflejando las condiciones físicas a las que estuvo expuesta la unidad durante el periodo analizado.

Finalmente, variables categóricas como el tipo de flota y el código del camión aportan valor adicional

al modelo, permitiéndole capturar diferencias estructurales entre tipos de vehículos y estrategias operativas específicas. Aunque su contribución individual es menor en comparación con las variables acumulativas, su inclusión mejora la capacidad del modelo para generalizar entre distintas unidades.

Este análisis respalda la lógica de construcción del conjunto de predictores y valida la relevancia operacional de las variables seleccionadas en el contexto minero.

xi. Evaluación de los modelos

Para evaluar los modelos se usaron las siguientes métricas:

Métricas	Qué indica	Ventaja	Desventaja
MAE	Error absoluto medio	Fácil interpretación	Sensible a outliers
MedianAE	Error absoluto típico (mediana)	Robusto a outliers	No refleja la magnitud promedio total
MAPE	Error porcentual absoluto medio	Interpretación relativa fácil	Inestable si el target es cercano a 0

Tabla 5. Descripción de métricas utilizadas para la evaluación de modelos de regresión

Los resultados obtenidos en el conjunto de datos de prueba para los dos modelos evaluados se muestran en la siguiente tabla:

Modelos	MAE	MedianAE	MAPE (%)
Linear regression	6	4.54	11
LightGBM	4.9	2.92	9.93

Tabla 6. Comparación de desempeño entre modelos en el conjunto de prueba

Los resultados evidencian que el modelo LightGBM supera al modelo de regresión lineal en todas las métricas evaluadas. Particularmente, presenta un menor MAE y MAPE, lo que indica una mayor precisión tanto en términos absolutos como relativos. Asimismo, la notable diferencia en MedianAE demuestra una mayor robustez frente a valores atípicos, lo que es especialmente relevante en entornos operativos con alta variabilidad como el minero.

En función de estos resultados, LightGBM fue seleccionado como el modelo óptimo para la estimación del nivel de combustible remanente en los camiones de acarreo, por su capacidad de adaptación a las condiciones operativas reales y su superioridad en desempeño frente a modelos base.

3.2.2 Modelo de optimización para la Asignación de Camiones a Grifos

Para la construcción del modelo de optimización, se utilizó información operativa en tiempo real proveniente del sistema Dispatch, la cual permitió capturar el estado dinámico de la flota de camiones y las condiciones de las estaciones de abastecimiento en cada instante de evaluación.

El objetivo principal del modelo es asignar de forma óptima los camiones que requieren abastecimiento a los grifos disponibles, minimizando el tiempo total de espera en cola y mejorando la eficiencia global del sistema. Para ello, se modela la asignación como un problema dinámico sujeto a múltiples restricciones operativas y logísticas propias de la mina.

Información utilizada para la optimización

El modelo considera un conjunto amplio de variables y parámetros en tiempo real, los cuales permiten representar el estado del sistema con precisión:

Camiones por asignar: Unidades que requieren abastecimiento, pero aún no tienen una estación asignada.

Camiones en ruta hacia grifo: Vehículos que ya han recibido asignación, pero aún se desplazan hacia el punto de abastecimiento.

Camiones en espera en grifo: Camiones que han llegado a la estación y esperan turno para ser atendidos.

Camiones actualmente abasteciendo: Unidades que se encuentran en proceso activo de carga de combustible.

Además, se integra información espacial y logística clave:

Posición actual de cada camión, utilizada para calcular rutas óptimas hacia los grifos mediante algoritmos de grafos.

Configuración vial de la mina, incluyendo:

- Restricciones operativas en ciertas rutas.
- Vías cerradas o inhabilitadas, que se modelan como filtros de exclusión en la red de caminos.
- Tipo de baliza o nodo actual del camión, que identifica su ubicación precisa en el mapa digital de la mina.
- Fase de operación asociada a la baliza, que permite restringir accesos según condiciones del frente o etapa minera.
- Tiempo de espera acumulado de cada camión en cola.
- Duración actual del proceso de abastecimiento en curso, para estimar el tiempo restante hasta la liberación de la estación.

Con estos elementos, el modelo construye un grafo dirigido y ponderado, donde los nodos representan posiciones posibles y los arcos modelan trayectos con pesos dinámicos según distancia, disponibilidad de ruta y tiempo estimado de desplazamiento. La solución del problema se formula como un proceso de optimización que considera prioridades operativas, restricciones físicas y objetivos logísticos.

Algoritmo de Ruta Óptima: Implementación de Dijkstra

Como parte del modelo de optimización, se implementó el algoritmo de Dijkstra para determinar las rutas más eficientes entre los camiones y los grifos disponibles dentro del entorno vial de la mina. Este algoritmo permite calcular la ruta más corta en un grafo ponderado, donde los pesos representan los tiempos estimados de viaje entre nodos, considerando las condiciones dinámicas de operación.

Funcionamiento del algoritmo:

El algoritmo de Dijkstra sigue los siguientes pasos para encontrar la ruta óptima desde un nodo origen hacia todos los demás nodos de la red:

1. Inicialización
 - ✓ Se asigna una distancia inicial a todos los nodos:
 - La distancia al nodo de inicio es 0.
 - La distancia al resto de los nodos es infinita
 - ✓ Se establece el nodo inicial como el nodo actual y se marca como visitado

- ✓ Se crea una cola de prioridad para gestionar los nodos pendientes de visitar

2. Proceso

- ✓ Para el nodo actual se evalúan todos sus vecinos
- ✓ Se calcula la distancia a cada vecino sumando:
 - La distancia acumulada hasta el nodo actual
 - El peso de la arista que lo conecta con el vecino
- ✓ Si la nueva distancia calculada es menor que la previamente almacenada, se actualiza
- ✓ Una vez procesada todos los vecinos, el nodo actual se marca como visitado
- ✓ Se selecciona como siguiente nodo a procesar aquel no visitado con la menor distancia registrada, utilizando la cola de prioridad

3. Repetición

- ✓ El proceso se repite hasta que todos los nodos hayan sido visitados o la cola de prioridad esté vacía.

4. Terminación

- ✓ Al finalizar, las distancias almacenadas representan la ruta más corta desde el nodo de inicio hacia todos los demás nodos del grafo.

Aplicación del algoritmo a nuestro caso

En el sistema desarrollado para Antamina, el algoritmo de Dijkstra se aplica sobre un grafo espacial dinámico, representando la red vial de la mina. Las adaptaciones específicas incluyen:

Nodos:

Representan ubicaciones clave como la posición actual del camión, puntos de descarga, intersecciones y estaciones de abastecimiento (grifos).

Aristas:

Conectan nodos entre sí y están asociadas a trayectos viales disponibles.

Pesos:

Corresponden a los tiempos estimados de desplazamiento, que pueden variar en función de condiciones del terreno, restricciones viales, tráfico, o cierres temporales.

Restricciones operativas:

Se incorporan filtros de exclusión que eliminan rutas cerradas, nodos inaccesibles por configuración de flota, o caminos restringidos por seguridad u operación.

Resultado:

Se obtiene la ruta más eficiente hacia un grifo compatible con el camión, considerando su nivel estimado de combustible, su posición actual, y las condiciones operativas del sistema.

Esta implementación permite que cada decisión de asignación no solo considere el mejor destino en términos de disponibilidad de grifos, sino también la ruta más eficiente para alcanzarlo, lo que reduce tiempos muertos, mejora la continuidad del ciclo de acarreo y optimiza la logística de abastecimiento.

3.3 Flujo Integrado del Sistema Predictivo y de Optimización:

El presente apartado describe la integración de los modelos de inferencia y optimización para la asignación eficiente de camiones a estaciones de abastecimiento, buscando minimizar los tiempos de espera y optimizar la operación.

Paso 1: Identificación de camiones a asignar

Se comienza con la selección de los camiones que requieren ser asignados a un grifo. Esta identificación se realiza a partir de la información contenida en la tabla de inferencia, la cual almacena los niveles de combustible predichos para cada unidad de la flota.

Paso 2: Cálculo del tiempo total estimado de desplazamiento

Para cada camión identificado, se calcula el tiempo estimado desde su posición actual hasta el punto de descarga (t_1), en caso de que aún no haya alcanzado dicho punto. Posteriormente, se estima el tiempo de desplazamiento desde el punto de descarga hasta cada uno de los grifos disponibles (t_2). La suma de estos tiempos ($t_1 + t_2$) representa el tiempo total estimado para que el camión llegue a cada estación de abastecimiento.



Figura 6. Flujo del cálculo del tiempo total estimado ($t_1 + t_2$) hacia cada grifo

Paso 3: Estimación del tiempo de espera en estaciones de servicio

Se contabiliza la cantidad de camiones en ruta hacia un grifo (N_1), así como aquellos que se encuentran en espera o actualmente abasteciéndose (N_2). Esta información es utilizada para calcular el tiempo de espera esperado en cada estación.



Figura 7. Flujo para la estimación del tiempo de espera en grifos a partir del conteo de camiones en ruta (N_1), en espera y abasteciendo (N_2)

Finalmente, la asignación del grifo se realiza en función del menor tiempo de espera estimado. En situaciones donde un camión tenga acceso a un único grifo disponible, la asignación se realiza de forma inmediata.

La aplicación de este flujo integrado permite optimizar la asignación, como se ejemplifica en la siguiente tabla de resultados, donde, por ejemplo, se recomienda asignar el camión HT155 al Grifo 1, dada su menor espera estimada y la ausencia de restricciones operativas que impidan el ingreso.

Cod Camión	Flota	Time G1	Time G2	Time G3	Disponible G1	Disponible G2	Disponible G3
HT155	KOM930	881	1888	1041	1	0	1
HT128	KOM930	1372	189	1098	1	1	1
HT301	KOM980	909	1916	1069	1	0	1
HT111	KOM930	957	1964	1117	1	0	1
HT149	KOM930	772	1024	498	1	1	1

Tabla 7. Resultados de asignación de grifos por camión según tiempo estimado de espera.

Descripción de campos de la tabla de resultados:

Time G_i : Tiempo total estimado (en segundos) que el camión requerirá para desplazarse desde su posición actual (coordenadas x,y) hasta el **Grifo $_i$** . Este cálculo incluye el recorrido pendiente hasta el punto de descarga (en caso de que el camión aún no haya arribado) y el trayecto desde dicho punto hasta la estación de abastecimiento **G_i** .

Disponible G_i : Indicador binario que señala la disponibilidad del camión para abastecerse en el **Grifo $_i$** . Un valor de 1 en la columna "Disponible", indica ausencia de restricciones operativas que

impidan el ingreso; un valor de 0 denota que el acceso está restringido por condiciones logísticas, geográficas o de capacidad.

Donde $i = 1, 2, 3$

Paso 4: Asignación secuencial de camiones

La asignación de camiones a estaciones de abastecimiento se realiza de forma iterativa, priorizando aquellos con menor nivel de combustible predicho, siguiendo la lógica operacional descrita a continuación:

Asignación inicial: Si un camión dispone únicamente de un grifo accesible según las condiciones operativas, se le asigna dicho grifo de manera inmediata, sin necesidad de cálculos adicionales.

Actualización dinámica: Tras cada asignación, se actualiza la cantidad de camiones en ruta hacia los grifos correspondientes, considerando la incorporación del camión recién asignado. A continuación, se procede con la asignación del siguiente camión, repitiendo el proceso hasta completar la asignación de la flota.

Límite de atención por grifo: Para evitar saturación, se establece un límite máximo de tres camiones por estación: uno en proceso de abastecimiento, uno en espera y uno en ruta hacia el grifo.

Regla geográfica específica: Como criterio adicional, los camiones cuya ruta incluya la baliza denominada CHANCADORA y que cumplan las condiciones para ser asignados a un grifo, serán dirigidos de forma preferente al Grifo 2, facilitando la coordinación logística y reduciendo tiempos de desplazamiento.

3.4 Descripción del aplicativo resultante

Como culminación del flujo integrado entre el modelo de inferencia y el sistema de optimización basado en rutas y tiempos de espera, el aplicativo resultante denominado “Refueling”, representa la solución tecnológica desarrollada para optimizar la eficiencia en el desplazamiento y asignación de los camiones de acarreo a las estaciones de grifo dentro de la operación minera.

El diseño del aplicativo permite a los Despachadores (D2), en tiempo real, contar con:

- Un rango estimado del nivel de combustible en el tanque de cada camión, junto con la dirección hacia la que se dirige, mostrando un semáforo

de alerta según el porcentaje de combustible restante.

- Visibilidad del estado de los grifos.
- Opción de reportar incidencias en grifos o clima.
- Trazabilidad de las actividades de cada camión después de su último abastecimiento, así como un historial de movimientos según fecha.

Esta herramienta está diseñada para ser flexible y adaptable a los cambios en las condiciones de las vías y de la operación, facilitando una toma de decisiones eficiente y mejorando significativamente la asignación de camiones a las estaciones de grifo.

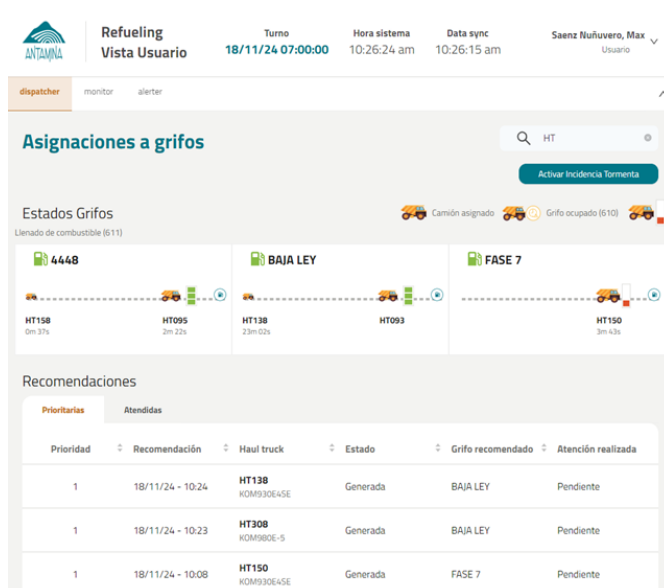


Figura 8. Pantalla de recomendación de Refueling



Figura 9. Listado de recomendaciones de asignación de camiones



Figura 10. Opción de reporte de incidencias



Figura 11. Tiempo de camiones en grifo

4. Presentación y discusión de resultados

Para validar la eficacia y robustez de la solución desarrollada, se implementaron dos pilotos operativos en la mina, orientados a evaluar aspectos críticos de la integración y desempeño del sistema Refueling en condiciones reales.

Objetivos del piloto:

Verificar la correcta implementación y operación de las lógicas de asignación, garantizando que las restricciones y particularidades de la operación minera sean respetadas.

Evaluar la integridad y fluidez del flujo de información en tiempo real, identificando posibles desfases o inconsistencias en la sincronización de datos entre sistemas.

Medir la aceptación, usabilidad y grado de adopción del aplicativo por parte de usuarios clave, principalmente supervisores de despacho (D3), y su integración en las rutinas operativas diarias.

Primer piloto

El primer piloto se desarrolló en condiciones operativas complejas, incluyendo la temporada de lluvias, que introdujo desafíos adicionales como interrupciones del sistema y desincronización temporal de datos. A pesar de estas adversidades,

se logró validar los objetivos planteados y detectar oportunidades de mejora cruciales.

El sistema operó simultáneamente con el módulo automático de asignación de combustible del sistema Dispatch, lo que generó escenarios donde no siempre se siguieron las recomendaciones del aplicativo Refueling.

La adherencia a las recomendaciones se cuantificó en un 55%, evidenciando una buena disposición inicial, aunque con margen para optimización.

Se recopilaron observaciones que fundamentaron ajustes significativos en las lógicas de optimización y manejo de datos para la siguiente iteración del sistema.

Los resultados preliminares demostraron una reducción tangible en los tiempos de espera en grifos, validando el potencial de la solución incluso en condiciones operativas adversas.

Segundo piloto

En esta fase, se implementaron mejoras clave para maximizar el impacto y la confiabilidad del sistema:

Se corrigieron los desfases en la sincronización de datos, asegurando la integridad y actualidad de la información procesada.

Se desactivó el sistema de asignación automática de Dispatch, permitiendo que el 100% de las decisiones provinieran exclusivamente del aplicativo Refueling, consolidando el control y la trazabilidad de las asignaciones.

Tras el segundo piloto, se procedió a un despliegue continuo y operativo durante cuatro meses, funcionando 24/7 para evaluar el rendimiento y la estabilidad del sistema bajo condiciones normales de operación sostenida.

Resultados globales del piloto extendido (4 meses)

El análisis cuantitativo de los datos recogidos durante el despliegue prolongado reveló:

Una reducción significativa y consistente en los tiempos de espera en estaciones de abastecimiento, optimizando el ciclo de acarreo y contribuyendo a la mejora de la productividad general.

Mayor eficiencia en la asignación dinámica de camiones a grifos, reduciendo cuellos de botella y

maximizando la utilización de los recursos disponibles.

Alta aceptación del sistema por parte de los supervisores, quienes valoraron positivamente la interfaz amigable, la información en tiempo real y la trazabilidad ofrecida.

Estos resultados confirman que la solución Refueling no solo cumple con las expectativas técnicas iniciales, sino que también aporta un valor operativo tangible, consolidándose como una herramienta innovadora y adaptable que fortalece la gestión operativa en entornos mineros complejos como es el caso de Antamina.

En la Figura 12 se presenta el análisis temporal de los tiempos de espera en estaciones de abastecimiento durante el período de despliegue continuo del sistema. Se observa una reducción promedio de 1.08 minutos por evento de abastecimiento, evidenciando una mejora significativa en la eficiencia operativa atribuible a la implementación de la solución Refueling.

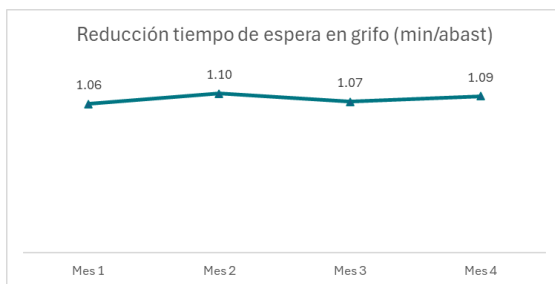


Figura 12. Reducción de tiempo de espera por abastecimiento

5. Conclusiones

El presente proyecto logró optimizar de manera efectiva la asignación de camiones y la gestión de tiempos de espera en las estaciones de servicio de Antamina, consolidando una solución desarrollada internamente que responde de forma precisa a las necesidades operativas del entorno.

A través del uso de técnicas avanzadas de Inteligencia Artificial (IA) y Machine Learning (ML), se diseñaron modelos predictivos y algoritmos inteligentes capaces de anticipar niveles de combustible y asignar rutas óptimas en tiempo real. Esta estrategia permitió reducir significativamente los tiempos improductivos y mejorar la eficiencia general de la flota.

La decisión de desarrollar la solución in-house, en lugar de optar por plataformas comerciales, resultó

clave para garantizar una integración flexible, escalable y adaptable a las condiciones dinámicas y específicas de la operación minera. Esta personalización no solo fortaleció el control sobre el sistema, sino que habilitó una evolución continua, basada en datos operativos reales y en la retroalimentación directa de los usuarios finales.

Los resultados obtenidos durante los pilotos y el despliegue continuo validan el enfoque metodológico adoptado, evidenciando mejoras sustanciales en la reducción de tiempos muertos y en la gestión logística de la flota. Además, el sistema demostró robustez para operar en condiciones adversas, como temporadas de lluvias y variabilidad en la infraestructura vial.

En perspectiva, la continuidad en el desarrollo y mantenimiento interno de esta solución tecnológica se presenta como una estrategia fundamental para preservar la adaptabilidad, el control operativo y la actualización constante de los modelos predictivos y de optimización. Esto permitirá sostener elevados estándares de eficiencia, seguridad y productividad, factores críticos para la operación minera en entornos altamente dinámicos y competitivos.

6. Anexos

Anexo 1: Método de Haversine

Los camiones se desplazan en un área amplia (mina, carreteras internas).

La tierra no es plana por lo que la distancia euclidiana (Pitagórica) entre latitudes/ longitudes no es exacta.

Haversine corrige esto, calculando la distancia real siguiendo la curvatura terrestre, con funciones trigonométricas:

- φ_1, λ_1 : *latitud y longitud del punto 1 en radianes*
- φ_2, λ_2 : *latitud y longitud del punto 2 en radianes*
- $\Delta \varphi = \varphi_2 - \varphi_1$
- $\Delta \lambda = \lambda_2 - \lambda_1$
- R : *radio de la tierra, 6371 km aprox*

$$a = \sin^2\left(\frac{\Delta \varphi}{2}\right) + \cos(\varphi_1) \cdot \cos(\varphi_2) \sin^2\left(\frac{\Delta \lambda}{2}\right)$$

$$c = 2 \cdot \arctan2(\sqrt{a}, \sqrt{1-a})$$

$$d = R \cdot c$$

A partir de esta función y el recorrido secuencial del camión se obtiene la distancia, velocidad media y pendiente del terreno

7. Referencias bibliográficas

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269-271.

Tito Surco, T. W. (2019). Estudio del consumo de combustible e implementación de un módulo de información en el sistema Dispatch para los camiones de la Cia Minera Antapaccay-Espinar-Cusco.

Souza, F. R., Câmara, T. R., Torres, V. F. N., Nader, B., & Galery, R. (2019). Mine fleet cost evaluation-Dijkstra's optimized path. *REM-International Engineering Journal*, 72(2), 321-328.

Wang, Q., Zhang, R., Lv, S., & Wang, Y. (2020). Open-pit mine truck fuel consumption pattern and application based on multi-dimensional features and XGBoost. *Sustainable Energy Technologies and Assessments*, 43.

Veness, C. (2012). Movable type scripts: calculate distance, bearing and more between latitude/longitude points. [online], Movable Type Scripts, nd. Available at: <http://www.movabletype.co.uk/scripts/latlong.html> [Accessed: 10 Dec. 2016]

Tukey, J. W. (1962). The future of data analysis. In *Breakthroughs in Statistics: Methodology and Distribution* (pp. 408-452). New York, NY: Springer New York.

Ibañez Asensio, S., Gisbert Blanquer, J. M., & Moreno Ramón, H. (2011). El sistema de coordenadas UTM.

8. Ilustraciones / Imágenes

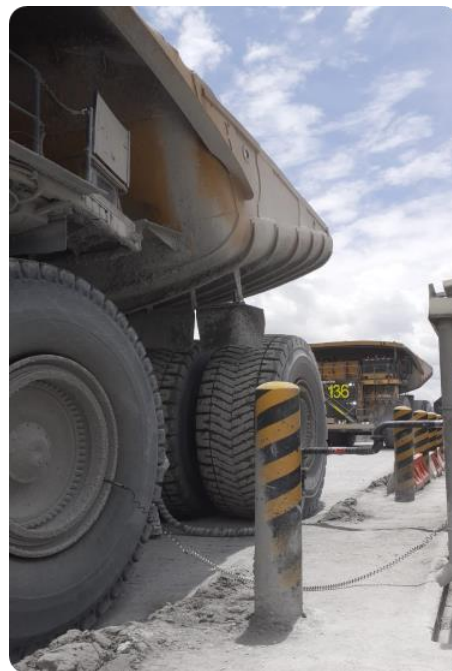


Figura 13. Camiones en cola en la estación de grifo



Figura 14. Abastecimiento del camión en la estación de grifo



Figura 15. Despachador (D2) usando el aplicativo – Piloto 2

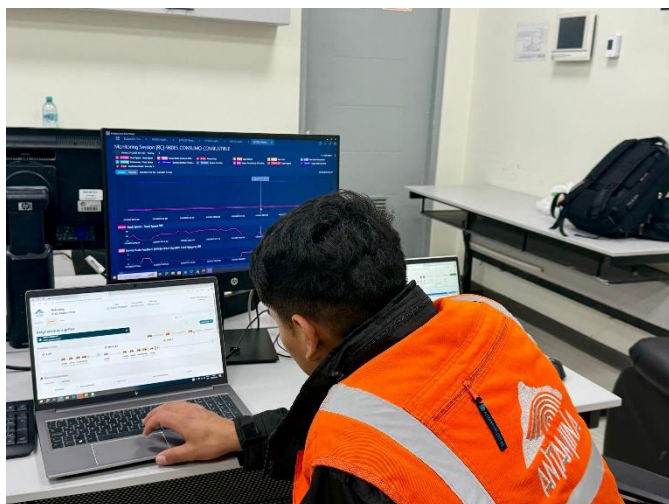


Figura 16. Despachador (D2) usando el aplicativo – Piloto Extendido

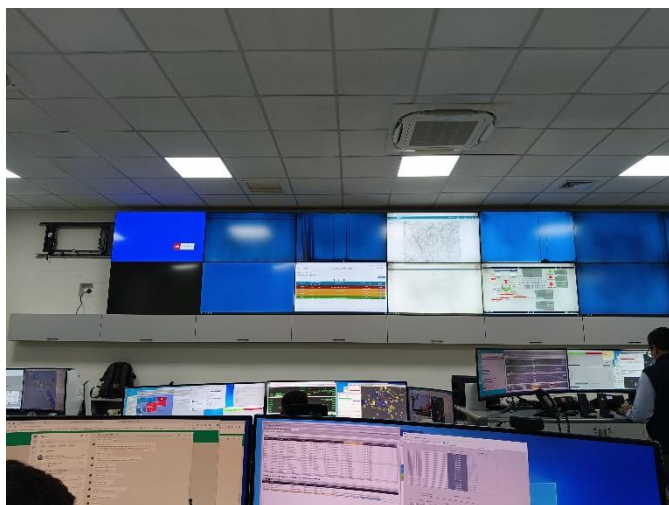


Figura 17. Aplicativo Refueling instalado en Sala Dispatch

9. Videos

Instructivo Refueling

https://drive.google.com/file/d/1IDigD9FoRqmYBj_YXGsluKg-Nq1f7The/view?usp=drive_link

Marleny Barrueta Mendoza

Ingeniero Estadístico con más de 14 años de experiencia en Ciencia de Datos, especializado en el desarrollo de modelos predictivos y prescriptivos para la mejora de procesos operativos. Gran expertise, centrado en la toma de decisiones estratégicas, con un fuerte enfoque en la eficiencia, la innovación y el trabajo colaborativo para generar resultados tangibles.

Con un MBA en Dirección Ejecutiva y cursando un MBA en Ingeniería de Minas, lo que complementa su habilidad para optimizar procesos y gestionar eficientemente en el sector.

Karina Hesi Torres

Ingeniero Estadístico con amplia experiencia en el desarrollo e implementación de proyectos de analítica avanzada y ciencia de datos, creando soluciones innovadoras para optimizar procesos clave. Con capacidad para trabajar en equipo, y con un enfoque orientado al logro de resultados y eficiencia en la ejecución de los proyectos.

Magister en Estadística Aplicada y cursando un MBA en Ciencia de Datos.

Jefferson Burga

Ingeniero Civil con más de 10 años de experiencia en el sector minero, con sólida formación técnica.

En los últimos años, ha desempeñado funciones clave en el área de despacho, coordinando operaciones logísticas de transporte de mineral, optimización de rutas y programación de cargas, asegurando el cumplimiento de los objetivos operativos y estándares de seguridad. Posee habilidades comprobadas en liderazgo de equipos, manejo de software de control logístico y comunicación efectiva con áreas operativas y administrativas.