



TRABAJO DE GRADO
Opción Seminario-Diplomado.

**ALGORITMO COMPUTACIONAL PARA PROYECTAR EL CONSUMO DE
GALONES DE COMBUSTIBLE NECESARIOS PARA LOS MESES SIGUIENTES,
UTILIZANDO ESTRATEGIAS DE MACHINE LEARNING**

Corporación Universitaria Remington.
Facultad de Ingenierías
Programa Académico: Ingeniería Industrial

Estudiante:
Monica Yadira López Linares
Tutor: Juan Carlos Briñez de León
Opción de Trabajo de grado Seminario-Diplomado.
2024.

Dedicatoria

A mi madre, quien me motiva día a día para superarme. Su ejemplo y su amor incondicional iluminan mis días en este camino llamado vida.

Agradecimientos

Estoy muy agradecida con Dios, mis Padres, mis Hermanos y mi Esposo por su apoyo y su paciencia en este proceso. Resulta importante reconocer que contar con el soporte espiritual, económico y moral de ustedes me permitió cumplir esta meta.

Gracias al manejo de datos, a la inteligencia artificial y a herramientas computacionales con modelos de aprendizaje automático que facilitan el pronóstico de presupuesto en diferentes áreas organizacionales.

Tabla de Contenidos

Resumen	5
Marco conceptual y contextual	6
Pregunta problema	6
Acercamiento a los datos	6
Descripción de variables	7
Aproximaciones con gráficos - analítica	8
Objetivos	10
Desarrollo e implementación del aprendizaje	10
Preparación de los datos	11
Modelo de toma de decisiones	12
Validación del modelo	13
Conclusiones	13
Referencias	14

Resumen

Este trabajo aborda el análisis de datos registrado por una empresa de transporte de carga seca, basados en el consumo semestral de combustible por tractocamiones tipo cama alta. Cuando se considera proyectar algoritmos computacionales para pronosticar datos futuros basados en machine learning se tiene una importancia significativa en el análisis de datos para la toma de decisiones.

La empresa dedicada al transporte de carga por carretera, basa sus presupuestos semestrales para la compra de galones de combustible, de acuerdo a las rutas de los vehículos y a una estadística establecida de consumo, con un promedio de 10 galones por cada 100 km de recorrido.

Para trabajar en este tipo de algoritmos con modelo de regresión, debemos validar su ejecución en condiciones similares y aspectos que no influyan de forma negativa en el proceso, porque podrían suponer una proyección alejada de la realidad.

En los últimos años, se han desarrollado estudios para la gestión eficiente del consumo de combustible, mediante la creación de un modelo de optimización en la trayectoria de los vehículos, mientras se presta el servicio de transporte de carga seca. Daza, P. F. M., Gómez, J. A. B., Medina, S. V. G., Tafur, C. L., & Rodriguez, S. E. F. (2024).

Palabras clave

Análisis de datos, clasificación, consumo de combustible, datasets, machine learning, métodos, modelos de predicción, proyección, pronostico, python, regresión, variables.

Marco conceptual y contextual

En las empresas de transporte de carga del país puede evidenciarse cambios en el consumo de combustible de los vehículos, por factores directos como rutas, recorridos, eficiencia en el motor, buen mantenimiento de los equipos, estado de la carga, entre otros. Para conocer sus variaciones existe la posibilidad de utilizar algoritmos de machine learning como métodos de proyección de información basados en datos particulares de cada organización.

Las características de los modelos hacen imprescindible, por las condiciones normales bajo las cuales son creados, que sean adaptados a las condiciones del lugar donde se utilizarán para lo cual es necesario realizar investigaciones que permitan efectuar en forma adecuada tal adaptación y así los resultados obtenidos sean fiables. Posada Henao, J. J., & González-Calderón, C. A. (2013).

El consumo de combustible representa uno de los insumos más importantes dentro del costo de operación de un vehículo, Según la Federación Colombiana de Transportadores de Carga por Carretera –Colfecar-, para Colombia el combustible corresponde al 40% del costo total de operación para vehículos tipo tractocamión.

Pregunta problema:

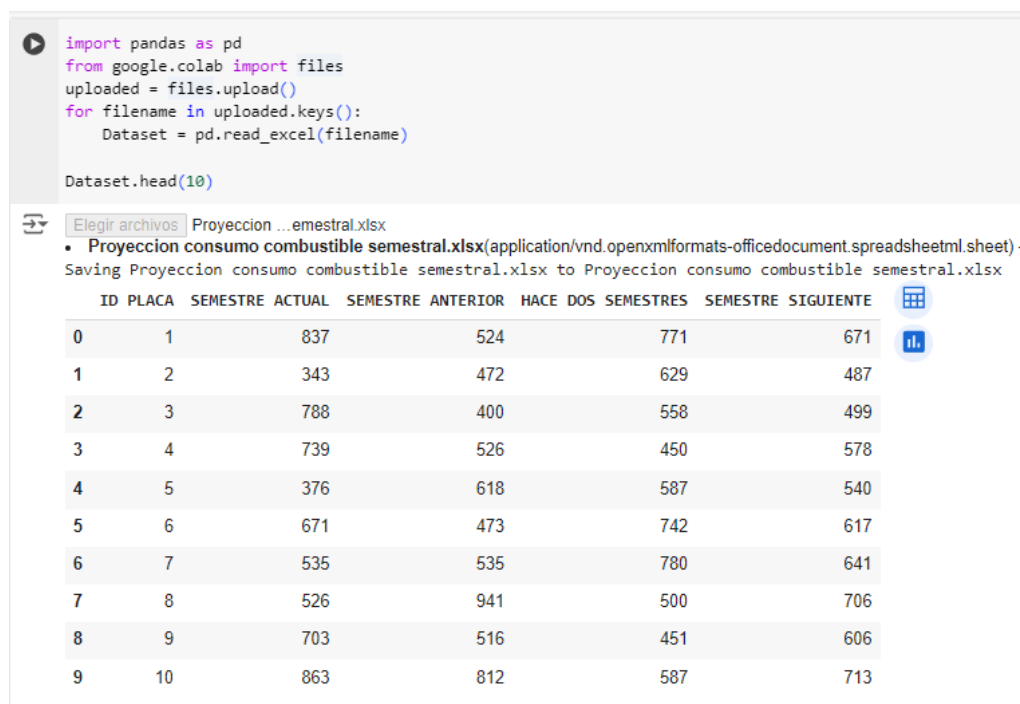
¿Cómo desarrollar una estrategia computacional para pronosticar el consumo de combustible en tractocamiones, a partir de algoritmos de Machine Learning?

Acercamiento a los datos: Estos datos vienen siendo recolectados desde hace dos años, por la empresa Transportes GM que se dedica al transporte de carga seca en tractocamiones

tipo cama alta en jurisdicción del municipio de Villanueva. Se trata de un histórico de datos que permite conocer el consumo de acpm de los vehículos mes a mes. Esto con el fin de lograr identificar el total de galones de combustible utilizados y establecer un pronóstico que permita conocer con anticipación dicho requerimiento, mediante la utilización de un modelo de regresión.

Descripción de variables.

La base de datos utilizada cuenta con 5 columnas (id placa 1; semestres, tomando como base de semestre actual del 1 de julio al 30 de diciembre 2023) y 40 filas con información contenida acerca de la cantidad de galones de combustible consumido en los recorridos realizados por cada vehículo. La siguiente imagen evidencia la lista de variables contenidas en el conjunto de datos.



```

import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Dataset = pd.read_excel(filename)

Dataset.head(10)

```

Elegir archivos: Proyeccion ...emestral.xlsx

- Proyeccion consumo combustible semestral.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - Saving Proyeccion consumo combustible semestral.xlsx to Proyeccion consumo combustible semestral.xlsx

	ID PLACA	SEMESTRE ACTUAL	SEMESTRE ANTERIOR	HACE DOS SEMESTRES	SEMESTRE SIGUIENTE
0	1	837	524	771	671
1	2	343	472	629	487
2	3	788	400	558	499
3	4	739	526	450	578
4	5	376	618	587	540
5	6	671	473	742	617
6	7	535	535	780	641
7	8	526	941	500	706
8	9	703	516	451	606
9	10	863	812	587	713

Figura 1. Lista de variables en el conjunto de datos.

Como pudo ser visto las variables son de tipo numérico, en su totalidad números enteros. Estas variables informan la cantidad de galones de combustibles consumidos en cada semestre por cada uno de los tractocamiones tipo cama alta en la empresa.

Aproximaciones con gráficos - analítica.

El conjunto de datos cuenta con 40 entradas y 4 columnas divididas en tres variables de entradas y 1 variable final que corresponde a la proyección de consumo para el siguiente semestre. A manera de resumen la siguiente imagen muestra la lista de variables contenidas en el mismo.

```

▶ #Información de la estructura de datos
Conjunto_Datos.info()

↳ <class 'pandas.core.frame.DataFrame'>
RangeIndex: 40 entries, 0 to 39
Data columns (total 4 columns):
#   Column                Non-Null Count  Dtype
---  -
0   SEMESTRE ACTUAL       40 non-null    int64
1   SEMESTRE ANTERIOR     40 non-null    int64
2   HACE DOS SEMESTRES    40 non-null    int64
3   SEMESTRE SIGUIENTE    40 non-null    int64
dtypes: int64(4)
memory usage: 1.4 KB

```

Figura 2. Clasificación de las variables

En cuanto al consumo de acpm la variable del “semestre anterior” fue la que presento mayores cifras en comparación con las demás.

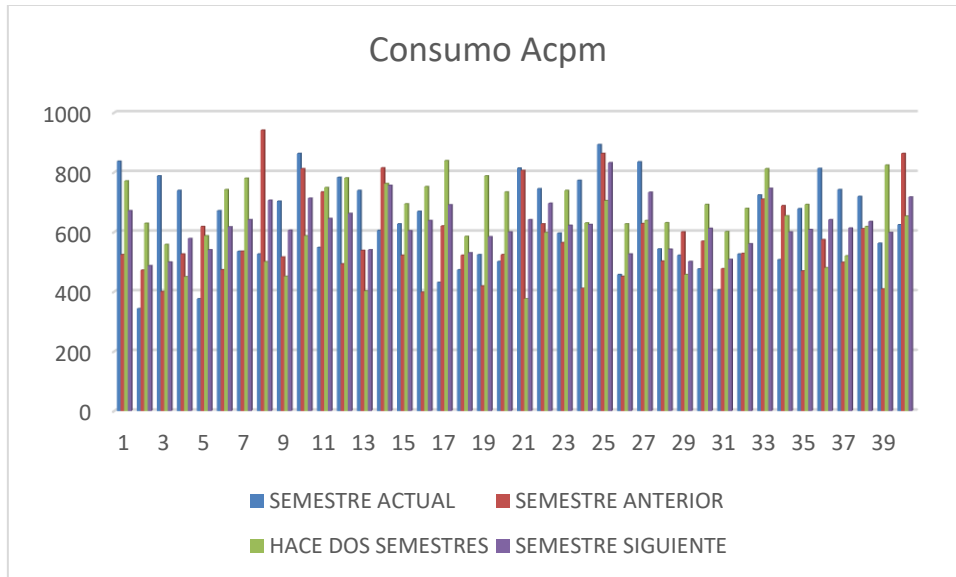


Figura 3. Gráfico consumo de combustible

En cuanto a el vehículo No 1 se evidencia que el mayor consumo de combustible se presentó en el semestre actual.

Diagrama de torta vehículo 1

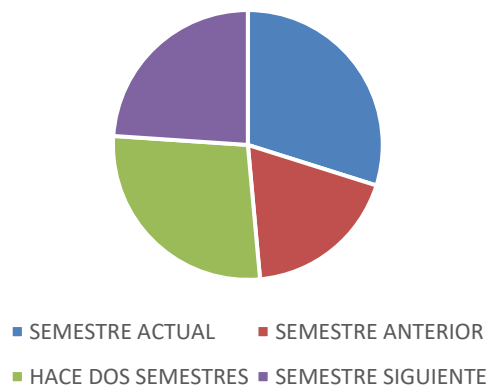


Figura 4. Gráfico tractocamiión 1

Objetivos:**Objetivo general.**

Implementar un algoritmo computacional de regresión para el análisis y toma de decisiones a partir de la proyección de combustible necesario en la empresa, utilizando estrategias de machine learning.

Objetivos específicos.

- Caracterizar y procesar los datos de interés, con miras a lograr pronosticar los galones de combustible necesarios para el siguiente mes.
- Generar una estrategia de decisión mediante Machine learning
- Implementar un algoritmo de Machine learning para la toma de decisiones a partir de los datos de interés.
- Validar el funcionamiento de toma de decisiones a partir de datos nuevos.

Desarrollo e implementación del aprendizaje

Para el desarrollo de este trabajo, inicialmente se obtuvo una base de datos recopilados en excel. Se procedió a organizar los datos de manera manual para que su estructura concordara con la requerida en python y así, utilizar un algoritmo computacional mediante la implementación de un modelo de machine learning basado en el método de regresión. Se evaluó el desempeño y se puso a prueba para pronosticar el consumo de combustible de los tractocamiones durante el siguiente semestre. A continuación, se desarrollan cada una de las etapas de la metodología.

Preparación de los datos

- Limpieza: Se opto por limpiar los datos de la librería de python y cargar la nueva base de datos para correr el código.

• Proyeccion consumo combustible semestral.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - Saving Proyeccion consumo combustible semestral.xlsx to Proyeccion consumo combustible semestral.xlsx

ID PLACA	SEMESTRE ACTUAL	SEMESTRE ANTERIOR	HACE DOS SEMESTRES	SEMESTRE SIGUIENTE
0	1	837	524	771
1	2	343	472	629
2	3	788	400	558
3	4	739	526	450
4	5	376	618	587
5	6	671	473	742
6	7	535	535	780
7	8	526	941	500
8	9	703	516	451
9	10	863	812	587

Figura 5. Imagen base de datos en Python.

- Se eliminaron filas con variables innecesarias al momento de ejecutar modelo de machine learning.

	SEMESTRE ACTUAL	SEMESTRE ANTERIOR	HACE DOS SEMESTRES	SEMESTRE SIGUIENTE
0	837	524	771	671
1	343	472	629	487
2	788	400	558	499
3	739	526	450	578
4	376	618	587	540
5	671	473	742	617
6	535	535	780	641
7	526	941	500	706
8	703	516	451	606
9	863	812	587	713

Figura 6. Imagen de las variables sin Id placa.

Modelo de toma de decisiones

Mediante el modelo de machine learning el método utilizado para el análisis de datos y tomas de decisiones fue el de regresión, ya que nos permite predecir y estimar la cantidad de galones de combustible necesarios para el consumo de cada vehículo durante el siguiente semestre.

Elegir archivos Proyeccion ... emestral.xlsx

- Proyeccion consumo combustible semestral.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 12126 bytes, last modified: 16/7/2024 - 100% done

Saving Proyeccion consumo combustible semestral.xlsx to Proyeccion consumo combustible semestral (1).xlsx

SEMESTRE ACTUAL	SEMESTRE ANTERIOR	HACE DOS SEMESTRES	SEMESTRE SIGUIENTE
0	837	524	771
1	343	472	629
2	788	400	558
3	739	526	450
4	376	618	587
5	671	473	742
6	535	535	780
7	526	941	500
8	703	516	451
9	863	812	587

Entrenando los modelos (En este caso entrenamos dos modelos)

```
[8] #Modelo para predicciones de series de tiempo (KNN: Bueno si sabemos que los datos no se salen del rango)
from sklearn.neighbors import KNeighborsRegressor as KNNR
Modelo_1=KNNR()
Modelo_1.fit(Entradas,Salida)
```

KNeighborsRegressor
KNeighborsRegressor()

```
#Modelo para series de tiempo complicadas (Sólo se conoce la dinámica)
from sklearn.neural_network import MLPRegressor
Modelo_2 = MLPRegressor()
Modelo_2.fit(Entradas, Salida)
```

MLPRegressor
MLPRegressor()

Figura 7. Imagen de dos modelos de regresión, utilizados para la predicción.

Validación del modelo

Al comprobar el mecanismo computacional entrenado para anticipar datos, se logra evidenciar que el algoritmo de predicción más exacto fue el de KNN Regresor ya que, fue el que más se acercó al resultado real de 641 galones de combustible.

Utilizando el modelo entrenado

```

#Se ingresan los datos de entrada para generar una predicción
Nueva_entrada = np.zeros((1,3))
Nueva_entrada[0,0]=float(input('Ingrese el valor de SEMESTRE ACTUAL: '))
Nueva_entrada[0,1]=float(input('Ingrese el valor de SEMESTRE ANTERIOR: '))
Nueva_entrada[0,2]=float(input('Ingrese el valor de HACE DOS SEMESTRES: '))

Proyeccion_1 = Modelo_1.predict(Nueva_entrada)
Proyeccion_2 = Modelo_2.predict(Nueva_entrada)

print('')
print('')
print('Según los datos ingresados, la proyección de combustible para el siguiente semestre usando KNN será: ',Proyeccion_1[0])
print('')
print('Según los datos ingresados, la proyección de combustible para el siguiente semestre usando ANN será: ',Proyeccion_2[0])

```

Ingrese el valor de SEMESTRE ACTUAL: 450
 Ingrese el valor de SEMESTRE ANTERIOR: 639
 Ingrese el valor de HACE DOS SEMESTRES: 858

Según los datos ingresados, la proyección de combustible para el siguiente semestre usando KNN será: 637.8
 Según los datos ingresados, la proyección de combustible para el siguiente semestre usando ANN será: 677.4258721202303

Figura 8. Imagen de los resultados de acuerdo a los modelos de predicción utilizados.

Conclusiones y trabajos futuros

En este trabajo se materializó la construcción de un algoritmo para la toma de decisiones basadas en información real que permitió proyectar el consumo de combustible para los siguientes semestres, utilizando modelos de machine learning mediante el método de regresión.

En cuanto a los modelos utilizados, el algoritmo de mayor precisión fue el de KNN Regresor ya que nos permitió la correcta estimación del consumo de combustible.

Para futuros trabajos, es indispensable ejecutar el modelo directamente en la industria, para que otras empresas se vean beneficiadas con la utilización del mismo.

Referencias

- Posada Henao, J. J., & González-Calderón, C. A. (2013). Consumo de combustible en vehículos para transporte por carretera-Modelos predictivos. *Revista Ingenierías Universidad de Medellín*, 12(23), 35-46.
- Posada Henao, J. (2012). Efecto de la cantidad de carga en el consumo de combustible en camiones. Universidad Nacional de Colombia Sede Medellín Facultad de Minas programa de Sistemas.
- Daza, P. F. M., Gómez, J. A. B., Medina, S. V. G., Tafur, C. L., & Rodríguez, S. E. F. (2024). Aplicación de un modelo predictivo de consumo de combustible a partir de machine learning y random forest para una aerolínea comercial colombiana. *Ciencia y Poder Aéreo*, 19(2), 122-134.