

Informe de Análisis de Datos

Descripción del Conjunto de Datos

El conjunto de datos que se analizó comprende información sobre diferentes configuraciones de trenes y su impacto en el desgaste de las ruedas, medido a través de la métrica `max_Z_wear`. Las variables consideradas en el análisis incluyen tanto variables categóricas como continuas, que se describen a continuación:

- **Caso:** Tipo de modelo de tren (Modelo 1 o Modelo 2)
- **Carga:** Carga del pasajero (AW0 (TARE) o AW4 (CARGA MÁXIMA))
- **Configuración:** Tipo de tren (Conf1, Conf2 o Híbrido)
- **Perfil de Rueda:** Criterio de diseño del perfil de rueda (P1 o P2)
- **COF:** Coeficiente de fricción (Seco (DRY) o con lubricación de brida (WFL))
- **Pista:** Severidad de la geometría de la pista (1-Baja, 2-Media o 3-Alta)
- **Dirección:** Dirección de viaje (Unidireccional o Bidireccional)
- **Variables Continuas:** `F_KX1`, `F_KZ1`, `F_KZ2`, `F_S`
- **KPI Objetivo:** `max_Z_wear` (a minimizar)

Análisis Exploratorio

Estructura de los Datos

El conjunto de datos contiene 9068 entradas y 13 columnas. Se realizó un análisis descriptivo inicial para comprender la naturaleza de los datos, revelando que la variable objetivo `max_Z_wear` tiene un rango muy pequeño.

Correlación

Se calculó la matriz de correlación para evaluar la relación entre las variables continuas y la variable objetivo `max_Z_wear`. Sin embargo, se encontró que la correlación no se pudo calcular inicialmente debido a la presencia de variables categóricas. Se procedió a codificar estas variables utilizando la codificación one-hot.

Tras aplicar la codificación, se calculó nuevamente la matriz de correlación, obteniendo los siguientes resultados:

- `max_Z_wear` tiene una correlación positiva con `Wheel.Profile_P2` y `Track_3-High`, mientras que muestra una correlación negativa significativa con `COF_WFL`.

Modelo de Regresión

A pesar de que se intentó implementar un modelo de regresión lineal para predecir `max_Z_wear`, se encontró que la biblioteca `sklearn` no estaba disponible en el entorno. Por lo tanto, se optó por calcular manualmente los coeficientes utilizando el método de mínimos cuadrados ordinarios.

Los coeficientes obtenidos revelaron que `COF_WFL` tiene el efecto más fuerte en la reducción de `max_Z_wear`, seguido de `Wheel.Profile_P2`.

Análisis de Medias

Se calcularon las medias de `max_Z_wear` para diferentes categorías de las variables más significativas:

Por COF:

- **DRY:** 1.588213e-11
- **WFL:** 2.591649e-12

Por Perfil de Rueda:

- **P1:** 3.733496e-12
- **P2:** 1.409416e-11

Por Pista:

- **1-Baja:** 3.617308e-12
- **2-Media:** 9.757685e-12
- **3-Alta:** 1.296097e-11

Conclusiones y Recomendaciones

Los hallazgos del análisis sugieren varias estrategias para reducir `max_Z_wear`:

1. **Uso de Lubricación de Brida (WFL):** Dado que el uso de WFL reduce significativamente `max_Z_wear`, su implementación debe ser prioritaria en entornos operativos.
2. **Optimización del Perfil de Rueda:** Se recomienda adoptar el Perfil de Rueda P1, que resulta en menor desgaste en comparación con P2.
3. **Gestión de la Pista:** Es esencial mantener las pistas en la categoría de severidad más baja (1-Baja) para minimizar el desgaste.

Este análisis proporciona una base sólida para tomar decisiones informadas sobre cómo optimizar el rendimiento de los trenes y reducir el desgaste de las ruedas en el futuro.

