

Introducción

La empresa XY se especializa en la comercialización de vehículos. En el proceso de desarrollo del producto, participan diversos equipos, incluidos ingenieros de diseño, pruebas y manufactura. Cada etapa del desarrollo presenta desafíos específicos que deben abordarse rápidamente en cada ciclo de diseño para lanzar el producto al mercado antes que los competidores.

El equipo de pruebas se encarga de realizar mediciones en vehículos nuevos y existentes para generar datos que alimenten al equipo de diseño. Estos datos son fundamentales para tomar decisiones y realizar cálculos que garanticen la integridad estructural del vehículo. La calidad y rapidez en la entrega de estos datos son cruciales, ya que determinarán la arquitectura del vehículo desde las primeras etapas del proyecto.

Ignorar o malinterpretar los datos de pruebas debido a la falta de tiempo para obtener insights podría llevar a dos escenarios adversos. Por un lado, sobrediseñar el vehículo resultaría en costos innecesarios y un aumento de peso, lo que reduciría la eficiencia de combustible. Por otro lado, diseñar por debajo de las especificaciones podría provocar fallas estructurales en el vehículo, comprometiendo la seguridad del usuario y dañando la reputación de la marca. Es esencial evitar estos escenarios asegurando la precisión y prontitud en el análisis de los datos de pruebas.

El equipo de pruebas se dedica principalmente a colocar sensores en los vehículos, replicar pruebas que imiten el uso del usuario final y realizar cálculos para determinar si el vehículo probado cumple con las especificaciones de estándares o leyes aplicables. En este proyecto, se utilizan datos reales capturados en campo para una prueba de especial interés. A pesar de que los detalles del contexto de la prueba y el vehículo se omiten por razones de privacidad, el desconocimiento de estos detalles no es relevante para llevar a cabo un análisis estadístico y descriptivo de la información.

Los datos recopilados para este estudio consisten en una prueba realizada en un solo vehículo, que se divide en 41 subconjuntos de operaciones. Para cada operación, se recopila información de 70 sensores. Todos los datos capturados por los sensores se almacenan como series de tiempo con duraciones diferentes. En otras palabras, algunas operaciones son más largas que otras, y no hay dos operaciones iguales. La duración de estas operaciones varía desde 20 segundos hasta un máximo de 20 minutos. Cada uno de los 70 sensores mide con una frecuencia de 200Hz, lo que significa que se obtienen 200 datos por cada sensor por cada segundo transcurrido. Todos los sensores tienen la misma cantidad de mediciones, lo que convierte cada medición en una observación. Es importante destacar que los sensores están ubicados en diversas partes del vehículo. Debido a esto, a pesar de que se trata de la misma prueba, los sensores pueden mostrar diferentes magnitudes, valles, picos, máximos y mínimos, e incluso cambios de polaridad (positivos o negativos).

Objetivo / Objetivos

Principales

- Transcribir datos del programa especializado de pruebas Y a un formato CSV limpio y sin información no necesaria para iniciar el postprocesamiento
- Automatizar el post procesamiento para obtener: máximos, mínimos, estadísticas descriptivas, correlación entre canales, detección de valores atípicos y de períodos inactivos. Presentar los resultados de forma numérica y visual.
- Identificar el ciclo de la serie de tiempo con el máximo daño estructural por fatiga y determinar el tiempo en el nivel de posición de la maquinaria

Secundarios

- Desarrollar un panel en Power BI que muestre los resultados de los objetivos principales, permitiendo a los usuarios obtener conclusiones sin necesidad de programar

Metodología

Hasta el momento los pasos para poder obtener los resultados actuales consisten en:

1. Limpieza de datos
 - a. Extraer datos del programa Y especializado de pruebas.
 - b. Transformar a txt, transformar a csv.
 2. Estandarización de datos.
 - a. Debido a la naturaleza de unidad de medida, el marco de datos tiene que ser estandarizado ya que las celdas de carga, acelerómetros y sensores de estrés miden en Kilonewtons, mm/s² y Pascales respectivamente.
 3. Postproceso terminados
 - a. Maximos y minimos
 - i. De un sensor a través de todas las pruebas
 - ii. De todos los sensores una prueba.
 - iii. De ciertos sensores en una prueba.
- Detección de offset
- b. Valores iniciales y finales de prueba para un sensor a través de todas las pruebas.
4. Inspección visual de series de tiempo.
 5. Datos simultáneos
 - a. Dado un segundo interés obtener los valores de todos los sensores en una prueba.
 6. Estadística descriptiva
 - a. De todos los sensores hay una prueba.
 - b. De un sensor a través de todas las pruebas.
 - c. Diagramas de violín de varios sensores de interés en una prueba.
 7. Correlación
 - a. Correlación de un canal respecto al resto en una prueba.
 - b. Mapa de calor de correlación de todos los sensores.
 8. Daño por fatiga estructural
 - a. Determinar daño generado
 - b. Calcular daño acumulado
 - c. Calcular daño acumulado en porcentaje
 - d. Representación de barras de los ciclaje más dañino.

9. Outliers

- a. Detección de outliers con zscore.