

¿Cómo predecir la cantidad de kilómetros que recorrerá un bus antes de presentar falla?

Hoy to predict the number of kilometers that a bus will travel before presenting a fault?



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Diego Fernando Aguilar Tigreros
Fundación universitaria los libertadores,
Facultad de Ingeniería y Ciencias Básicas
dfaguilart@libertadores.edu.co

Juan Carlos Rubriche Cárdenas
Fundación universitaria los libertadores,
Facultad de Ingeniería y Ciencias Básicas
jcrubrichec@libertadores.edu.co

RESUMEN

Unimetro S.A requiere gestionar el indicador de kilómetros a la falla dado que este impacta el Índice de estado de la flota y por ende el Indicador de calidad de desempeño, lo que afecta de forma directa los ingresos de la compañía poniendo en riesgo el contrato de concesión el cual es la razón de ser de la Compañía. Es por ello que realizar un modelo que permita pronosticar cuantos kilómetros recorre un autobús antes de presentar una falla cobra gran importancia.

La información que se obtenga producto del trabajo realizado será tomada en cuenta desde el área de operaciones para realizar una gestión adecuada de la flota y con ello disminuir las fallas que se presentan en la operación.

Palabras clave: Pronostico, kilómetros a la falla, concesionario operador, ente gestor, red neuronal, Validación Cruzada

ABSTRACT

Unimetro SA requires managing the indicator of kilometers to failure since it impacts the fleet status index and therefore the performance indicator, which directly affects the company's income, putting the design contract at risk. It is the reason for being of the Company. That is why it makes a model that allows predicting how many kilometers a bus travels before presenting a fault is of great importance.

The information obtained as a result of the work carried out will be taken into account from the operations area to carry out an adequate management of the fleet and thus be able to reduce the failures that occur in the operation.

Keywords: forecast, kilometers to failure, operator concessionaire, managing body, neural network

INTRODUCCIÓN

Unimetro S.A es una empresa operadora del sistema MIO en la ciudad de Cali, la cual cuenta con un contrato de concesión.

Dentro de las obligaciones contractuales se tiene que cumplir con el indicador de calidad de desempeño, el cual consta de cuatro sub índices los cuales tienen un peso del 25 % cada uno. Índice de puntualidad, cumplimiento, estado de flota y operación.

Para el caso del proyecto propuesto nos centraremos en el índice de Estado de la flota. Donde se debe gestionar el indicador “kilómetros a la falla”. El cual toma el kilometraje ejecutado en un periodo de tiempo y lo divide por el número de fallas presentadas en el mismo periodo.

Con el indicador de kilómetros a la falla se verifican los vehículos que se encuentran por debajo del estándar, el vehículo que no cumpla se le cancela el certificado de operación, un vehículo sin certificado de operación no podrá ser despachado por el concesionario. He iniciara un proceso sancionatorio, las sanciones podrá llegar hasta la cancelación del contrato. Además de afectar el índice de cumplimiento, puesto que al no contar con un vehículo se reduce la capacidad operativa del Concesionario operador de transporte. Tipantuña (2019) manifiesta “la movilización de pasajeros a lo largo de la ciudad es un tema importante puesto que garantiza las actividades cotidianas y productivas de los ciudadanos” (p.10).

Por lo anterior se precisa gestionar el indicador de forma efectiva, para ello requiere anticiparse a las novedades que afecten dicho indicador, puesto que el ente gestor mide la operación pasada, pero el concesionario requiere gestionar la operación futura.

Para ello se ha redefinido el indicador de kilómetros a la falla buscando ser proactivos en dicho trabajo con esto se hace necesario el desarrollo de un modelo de pronóstico que permita predecir con algún grado de incertidumbre cuantos kilómetros recorrerá un vehículo antes de presentar una falla. Rincón y Vargas (2019) afirman “La aplicación de redes neuronales en el diagnóstico de fallos tiene dos etapas. La etapa uno corresponde al proceso de entrenamiento, durante la cual los patrones de entrenamiento son suministrados a la red con el fin de que realice los cálculos o ajuste de algunos de sus parámetros. La etapa dos es el proceso de prueba, durante la cual se entrega un patrón de datos desconocido con el fin de verificar si la salida entregada por la red corresponde a la salida esperada”. (p.28)

Al pronosticar cuantos kilómetros recorrerá un bus antes de presentar una falla, se podrán tomar decisiones para realizar la asignación de la flota generando eficiencias operativas, al asignar los vehículos con mayor probabilidad de presentar una falla a un servicio cercano al patio. Esto se traduce en asignar cerca al patio los vehículos con menor confiabilidad. Núñez (2018) afirma “La confiabilidad es el indicador que los operarios de los diferentes mecanismos de los procesos, conocen y están a la expectativa que así se efectúe.”. (p.17).

Lo anterior se traduce en eficiencias operativas que garantizan la prestación del servicio y que apuntan directamente a la misión de la organización cuando esta se refiere a “Facilitar la movilidad de las personas mediante la administración y operación de sistemas de transporte, con seguridad, responsabilidad social” y en especial al apartado “apoyada en el talento humano”. Sandino y Moreno 2016 manifiestan “La confiabilidad se define como “la probabilidad de que un componente o equipo no falle estando en servicio durante un periodo determinado”, cuando es operado en condiciones razonablemente uniformes de voltaje, corriente, potencia, presión, temperatura, frecuencia, entre otros.”. (p.25)

Con la puesta a punto del modelo de pronóstico la organización será pionera en modelos de este tipo puesto que hasta el momento no hay casos conocidos de empresas de transporte que realicen la gestión de la operación haciendo uso de modelos de pronóstico.

REFERENTES TEÓRICOS

Unimetro S.A es una empresa operadora del sistema MIO en la ciudad de Cali, la cual cuenta con un contrato de concepción con el sistema Masivo integrado de Occidente.

Dentro sus procesos de gestión y obligaciones contractuales la empresa debe gestionar el ICD que hace referencia al indicador índice de calidad de desempeño, el cual está relacionado con los **indicadores de calidad** de los procesos mediante procesos estadísticos basados en cifras o ratios que se utilizan como criterio para juzgar y evaluar el desempeño de una organización, un sistema o un proceso. En ese sentido El ICD, está compuesto por cuatro subíndices los cuales tienen un peso del 25 % cada uno, los cuales se describen en la siguiente tabla.

INDICE DE CALIDAD DE DESEMPEÑO
IC: Índice De Cumplimiento: El cual mide el % de kilómetros ejecutados sobre los kilómetros programados.
IO: Índice de Operación: Mide la operación del concesionario y está ligado con la prestación eficiente del servicio siendo impactado por eventos que se presenten durante la operación.
IP: Índice de Puntualidad: Para calcular este indicador se establecen una serie de puntos de control “paradas” durante la operación. En estos puntos se calcula la diferencia entre la hora programada y la hora real de paso.
IE: Índice de Estado de la Flota. Mide condiciones mecánicas, eléctricas, estéticas de los vehículos.

Tabla 1. ICD. Elaboración propia.

Para el desarrollo del presente trabajo nos centraremos en el estudio del Índice de Estado de la Flota IE y más concretamente el indicador denominado kilómetro de falla.

Indicador kilómetro de falla

A partir de los indicadores de gestión el denominado kilómetro de falla está determinado por los kilómetros que recorre un autobús antes de presentar una falla. Podemos definir una falla como todo evento repentino que causa una anomalía en la operación de un autobús. Barrios y Martínez 2018 definen “Una falla es una interrupción temporal o definitiva del sistema, que genera una perturbación dentro del desarrollo de un procedimiento o proceso industrial”. (p.11).

El ente gestor dentro del IE evalúa todas las fallas asociadas al concesionario operador de transporte, las fallas que no se incluyen en el indicador están asociadas con el protocolo tecnológico siempre y cuando este no sea responsabilidad de concesionario operador de transporte. Las fallas asociadas al protocolo tecnológico son las de los elementos como IVU BOX y validador entre otros componentes que no son responsabilidad del Concesionario operador de transporte.

Las fallas que son calificadas son las que se son producto de la gestión de la empresa operadora y para el índice de kilómetros a la falla son las que se presentan dentro de la prestación del servicio y que afectan la calidad de este. Es por ello que se hace necesario generar la gestión desde el despacho de la flota y con ello poder anticiparse a la falla.

Modelos de pronóstico

Con el avance de la computación y el desarrollo de nuevas tecnologías se ha iniciado una carrera por el manejo de los datos generando gran valor en las organizaciones que buscan en la tecnología y en especial en el análisis de datos la solución a la problemática en cuanto a la toma decisiones. Vargas y Rincón (2019) sugieren “La aplicación de redes neuronales en el diagnóstico de fallos tiene dos etapas. La etapa uno corresponde al proceso de entrenamiento, durante la cual los patrones de entrenamiento son suministrados a la red con el fin de que realice los cálculos o ajuste de algunos de sus parámetros. La etapa dos es el proceso de prueba, durante la cual se entrega un patrón de datos desconocido con el fin de verificar si la salida entregada por la red corresponde a la salida esperada.”. (p.28)

Los modelos de pronóstico son usados en el transporte para la detección de fallas y con estos poder realizar los planes de mantenimiento. Díaz y Gave (2017) “Mantenimiento Predictivo (MPD): Este mantenimiento se basa en realizar una serie de mediciones o ensayos no destructivos con herramientas sofisticadas a las partes de la maquinaria sujeta a deterioro, de esta manera prever una falla importante.”.(p.7).

Para el presente trabajo se realizará un enfoque diferente al que se evidencia en la documentación se realizará un modelo pronóstico que permita determinar el número de kilómetros que recorre un bus antes de una falla, esto con el objetivo de ser usado desde la gestión de la operación para mejorar la toma de decisiones.

Donde se pueda asignar los vehículos con mayor kilometraje recorrido por kilómetro a servicios con mayor número de kilómetros programados. Con esto en busca aumentar los ingresos de la compañía y mejorar el indicador de estado de flota.

Entendiendo la aplicación de los modelos de pronóstico y el uso en diversas áreas del conocimiento consideramos que estos pueden ser de gran utilidad para lo que se requiere solucionar con el presente trabajo, haciendo uso de las herramientas tecnológicas con las que disponemos. Díaz 2019 afirma “El aprendizaje automático, o también conocido como Machine Learning, es la rama de la ciencia de la computación que permite la programación de máquinas para que sean capaces de aprender de un conjunto de datos.” (p.17)

METODOLOGÍA

El método utilizado para resolver el problema es descriptivo de corte interpretativo, que está relacionado

Fase 1. Consolidación y depuración de base de datos

Unimetro S.A lleva registros de la operación diaria en archivos de Excel. En estos se consignan las novedades que se presentan durante la operación. Novedades tales como salidas de operación por accidentes, vandalismos, conductores que no se presentan al turno y fallas de la flota. Adicional a esto se realiza una conciliación de kilómetros ejecutados con el ente gestor el cual lleva registros en Excel de la operación de los diferentes concesionarios. Con estos archivos se dirimen las diferencias que se puedan tener y son el insumo oficial para todo lo que tenga que ver con la gestión de la operación. “pagos, indicadores, sanciones a la empresa operadora o conductor”.

Estos archivos al estar almacenados de forma diaria son consolidados, se ordenan por fecha y hora de inicio de los servicios, con este se genera una variable que acumula el kilometraje recorrido por cada bus. Hasta que se presente una falla. Guardando los datos hasta ese punto y reiniciándose a cero la variable contador. Los datos depurados son almacenados en un archivo plano. El cual será usado para el presente análisis.

Fase 2. Selección de variables

Las variables seleccionadas están relacionadas con los índices que componen el indicador kilómetro de Falla. Las cuales se describen a continuación se conocerán las variables que serán usadas para el desarrollo del presente trabajo.

KmAcumDesvio: Total kilómetros realizados hasta presentar una falla

HorasAcum: Horas acumuladas hasta que se presenta una falla.

Paradas Plan: paradas planeadas.

Paradas Eje: paradas que se ejecutaron hasta la falla.

ART: Tipología de bus que presenta la falla, tipos Articulado

PAD: Tipología de bus que presenta la falla, tipos Padrón

COM: Tipología de bus que presenta la falla, tipos Complementaria

Fase 3. Selección del mejor Modelo.

Para seleccionar el mejor modelo, se recurre al ajuste de diferentes tipos de modelos de pronóstico. Evaluando el error cuadrático medio y dejando como el mejor modelo el realizado con redes neuronales el cual presento un menor error cuadrático.

Entre los modelos ajustados podemos destacar. Arboles de decisión con un error cuadrático medio de 714,903 para el árbol recortado y 530,8659 para el árbol sin recortar.

Bagguin arrojando un error cuadrático medio 383,6157 calculado con los datos de entrenamientos y 521,3865 con los datos de test.

RESULTADOS

Verificación de los datos iniciales.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
3.23	313.18	838.74	1836.94	2118.52	38057.71

Tabla 2. Análisis inicial de la variable KmAcumDesvio

La tabla muestra el resumen de la variable KmAcumDesvio, donde podemos ver que el kilometraje mínimo es de 3.23 el máximo de 38057,71, y la media es de 1836 kilómetros.

Al analizar la tabla de datos podemos ver una diferencia considerable entre el dato máximo y mínimo. Lo cual podría suponer una dificultad al intentar realizar una predicción.

Red Neuronal.

Con los datos que se recopilados se toman dos muestras una de entrenamiento con el 70% de los datos dejando el 30% restante para validación, se ajusta una red neuronal de una capa oculta con una capa de salida la cual es la predicción.

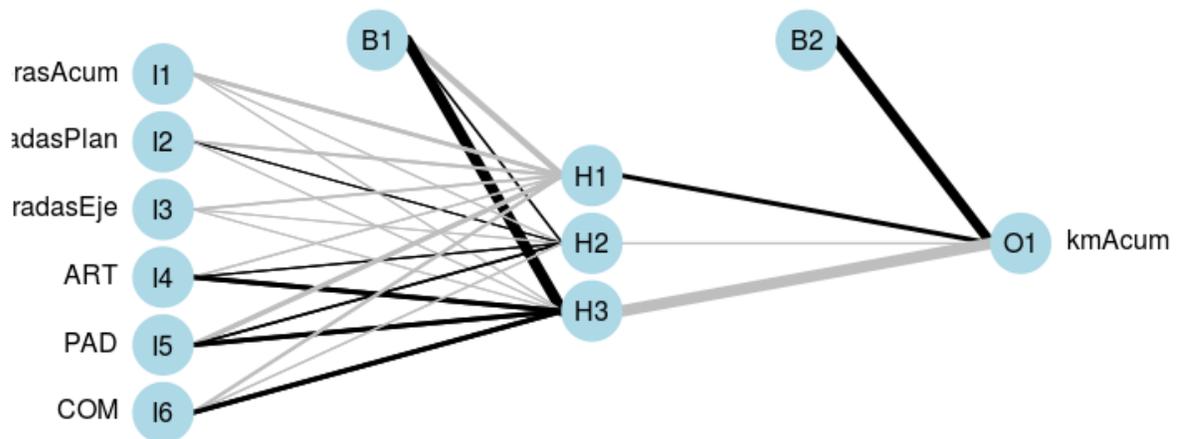


Figura 1. Red Neuronal. Elaboración propia.

La graficas muestra la red neuronal con seis nodos “parámetros de entrada” en la capa de entrada, en la capa oculta se pueden ver tres nodos y como nodos intermedios encontramos B1 y B2 y como nodo de salida O1 el cual representa el valor de la predicción.

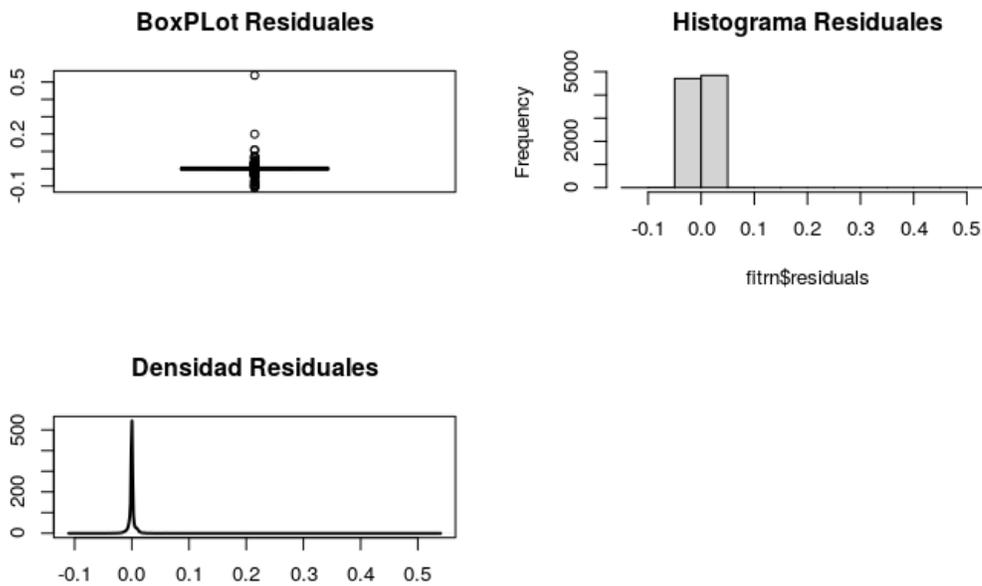


Figura 3: Normalidad de los residuos. Elaboración propia.

Con las gráficas se puede observar que se cuentan con datos atípicos esto se puede evidenciar especialmente en el grafico BoxPlot y la gráfica de densidad donde se observan unas colas pronunciadas. El histograma muestra una gran concentración de datos sobre media cero pero al detallar la gráfica se evidencia una cola pronunciada hacia el cinco lo que se suma al concepto de las gráficas Boxplot y la curva de densidad de los residuos.

Error de los residuos

Se realiza el cálculo del error cuadrático medio de los residuos de la red neuronal obteniendo el siguiente resultado: 320.3728.

Posterior a ello se calcula el error cuadrático medio de la predicción respecto al conjunto de test. Obteniendo como resultado 245.912. Donde se evidencia una reducción considerable.

Predicción

A continuación se presenta la tabla de predicciones

Cuadro predicciones vs datos de test					
Predicciones más ajustadas			Predicciones menos ajustadas		
Predicción	Datos de Test	Dif	Predicción	Datos de Test	Dif
289,711	289,786	0,075	4135,017	3,666	4131,351
11,757	11,619	0,138	3471,754	19,705	3452,049
201,431	201,643	0,212	25688,247	27829,995	2141,748
402,324	402,713	0,389	8141,020	9036,070	895,050
81,651	82,087	0,436	4890,663	5610,463	719,800
173,629	174,136	0,507	6527,118	7206,872	679,754
137,145	136,634	0,511	4328,353	4847,082	518,729
310,801	310,281	0,520	5157,215	5605,441	448,226
212,315	212,979	0,664	4165,453	4611,452	445,999
117,860	118,720	0,860	11692,329	11301,906	390,423
73,252	74,179	0,927	3822,323	4209,590	387,267
132,517	131,463	1,054	3761,075	4140,842	379,767
1004,791	1005,845	1,054	5517,327	5883,868	366,541
155,623	154,509	1,114	3666,850	4033,197	366,347
314,407	313,244	1,163	2726,409	3077,148	350,739
439,726	438,552	1,174	4170,950	3823,250	347,700
58,773	59,958	1,185	7904,956	8252,574	347,618
584,065	585,256	1,191	2956,103	3302,638	346,535
371,442	372,681	1,239	3375,781	3718,927	343,146
55,014	56,297	1,283	2437,458	2767,497	330,039
517,828	519,155	1,327	10950,819	11270,163	319,344
389,002	387,657	1,345	4386,522	4703,957	317,435

Tabla 4. Predicciones versus datos de validación.

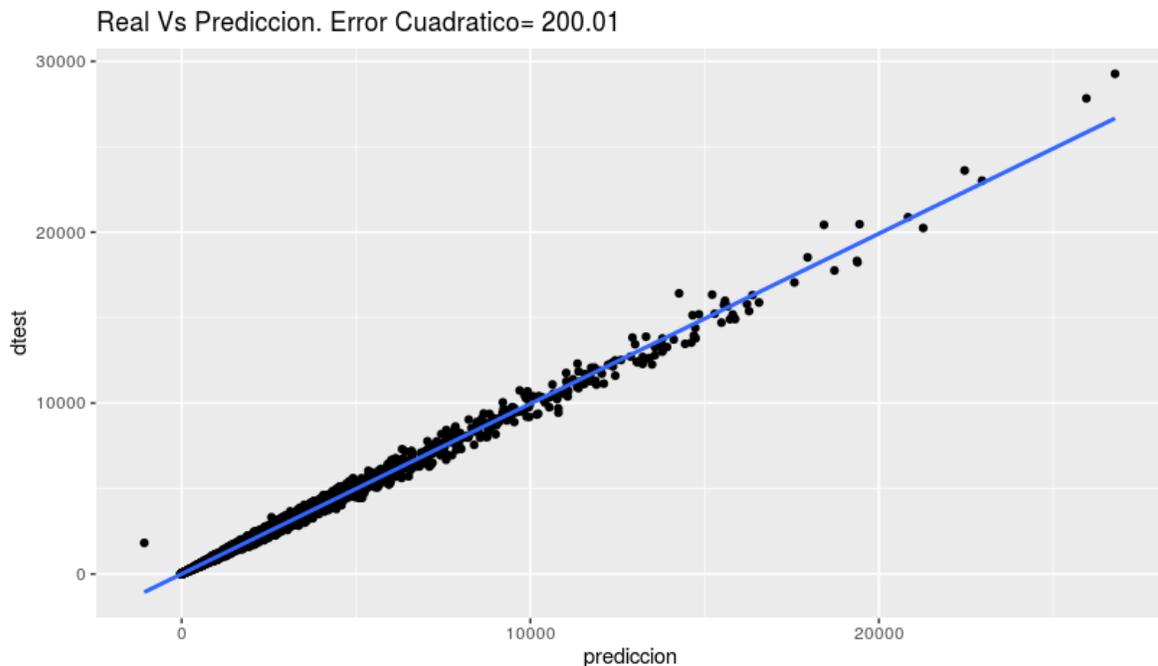


Figura 3. Gráfico de datos reales versus datos de la predicción.

Se puede observar que las predicciones se ajustan a la línea de regresión. Lo que podría suponer que con la red neuronal se pueden realizar predicciones de cuantos kilómetros recorrerá un bus antes de presentar una falla con un buen margen de ajuste.

Validación Cruzada.

Con el objetivo de validar que las predicciones obtenidas con la red neuronal se ajusten adecuadamente a lo que se requiere, se procede a realizar validación cruzada con los datos de entrenamiento.

El proceso de validación cruzada consiste en tomar los datos de entrenamiento del modelo dividirlos K veces, ajustando el mismo número de modelos con cada una de las divisiones creadas siendo validado cada modelo con los datos restantes. Ej. Si la base de datos contiene 1000 elementos y se desea un k igual a 5 el conjunto inicial de 1000 se dividirá en cinco subconjuntos de 200 elementos creando un modelo con el primer subconjunto y validando el modelo con el resto de subconjuntos. Esto se repite el número de K deseado.

A continuación se describen los pasos que se llevaron a cabo para realizar la validación cruzada.

- Se crean dos funciones encargadas de realizar la validación.
- La primera función lee la base de datos y el número de divisiones que el usuario desee crear, crea los subconjuntos de datos. Y llama la función encargada de la creación del modelo y el cálculo del error cuadrático medio
- La segunda función recibe la base de datos, el número de divisiones y el conjunto de datos. Crea el modelo de redes neuronales, con el primer subconjunto, calcula el error cuadrático medio y devuelve el promedio de los errores, esto es un proceso iterativo que se ejecuta el número de k definidos por el usuario. .
- La primera función muestra los datos por consola al usuario. Mostrando el error cuadrático medio de cada una de las iteraciones o subconjunto.

Obteniendo los siguientes resultados.

	Mod1	Mod2	Mod3	Mod4	Mod5
RMSE	3977,373	3984,549	3782,452	3814,544	3912,424

Tabla 5. Error cuadrático medio de los diferentes modelos de la validación cruzada.

El promedio de los errores cuadráticos 3894,268, calculando la desviación típica obtenemos. 92,52766

CONCLUSIONES

- El modelo que presenta un mejor ajuste es el realizado con redes neuronales.
- Las gráficas muestran que los datos no provienen de una distribución normal.
- Hay una reducción considerable entre el error cuadrático del modelo versus el error cuadrático medio del modelo respecto a los datos de test.
- La validación cruzada es un método para evaluar los resultados de un modelo de pronóstico. Con el objetivo de estimar la precisión del modelo.

REFERENCIAS BIBLIOGRÁFICAS

- Barros Leal, J. A., & Martínez Espinosa, J. C. (2018). Modelo para detección y simulación de fallas bajo la gestión de mantenimiento y proyectos. *Informador Técnico*, 82(1), 11-25. <https://doi.org/10.23850/22565035.1366>
- Díaz Chunga, C, G & Barja, Raúl A (2017). Implementación de un plan de mantenimiento preventivo para reducir las fallas de los buses Golden Dragón de la UNALM, en la ciudad de Lima 2017.
- Díaz Tipantuña, D. (2019). Aprendizaje automático para el mantenimiento predictivo en los buses del transporte público. Trabajo de titulación previo a la obtención del Título de Ingeniero Informático. Carrera de Ingeniería Informática. Quito: UCE.
- Núñez Palma, J (2018). Gestión de mantenimiento para mejorar la disponibilidad de la flota de transporte de la empresa “ángel divino”- Chiclayo
- Sandino Noguera, C, A & Moreno Duarte, K, F (2016) Gestión de activos para determinar el TPEF (tiempo promedio entre fallas) de una flota de buses, caso: empresa Masivo Capital SAS.
- Sarle, Warren S. (1994). Neural Networks and Statistical Models, Proceedings of the Nineteenth Annual SAS Users Group International Conference, pp. 1-13.
- Vargas Chavarro, D, A & Rincón Aristizábal L F (2019) Sistema de detección y diagnóstico de fallas de un proceso térmico mediante inteligencia artificial.