

# Framework para la Detección Anticipada de Fallas de Equipos Mineros Mediante el Uso de Machine Learning

Pablo Samillan Ramos, Bachiller<sup>1</sup> y Eveling Castro Gutierrez, Magister<sup>1</sup>  
Universidad Nacional de San Agustín de Arequipa, Peru  
psamillan@unsa.edu.pe, ecastro@unsa.edu.pe

**Abstract**—El costo de reparación por fallas en equipos pesados fluctúan entre el 15% y el 60% representando en la industria minera un mantenimiento de millones de dólares en reparación. Es por tal motivo que se propone un framework para predecir fallas tomando en consideración el registro de eventos de falla y el costo de mantenimiento debido a la facilidad de obtención y disponibilidad de dichos datos. Este paper propone la estructura de un framework que hace uso de los registros de eventos de fallas y los costos de mantenimiento para la predicción de fallas y evidenciar los beneficios del uso de Machine Learning en las actividades de mantenimiento.

**Index Terms**—Machine Learning, Long Short Term Memory, Mantenimiento Preventivo, Redes Neuronales Recurrentes

## I. INTRODUCCIÓN

La extracción de mineral esta compuesto de dos procesos: *Proceso de Minado*, es el proceso de extracción del material que contiene mineral y desmonte. Para tal efecto, es necesario el uso palas para el carguio y volquetes para el traslado del material desde la zona de extracción hacia la planta concentradora para continuar con el *Proceso de Molienda*, donde el material es pulverizado con el fin de separar el mineral del desmonte, para tal efecto, es requerido el uso de chancadoras y molinos [13]. La compra de equipos, requeridos en los procesos de minado, tiene un valor que oscila entre \$USD 4.5M a \$USD 21M.

A su vez, la compra de componentes principales de dichos equipos tiene un valor promedio de \$USD 300,000. Este es uno de los motivos de la importancia de predecir fallas que es la de disminuir costos de reparación.

El mantenimiento de equipos pesados es una actividad relevante, su función es mantener el estado óptimo de los sistemas, equipos y herramientas para realizar el proceso de producción y transformación de un bien (p. ej. minerales). En el proceso de extracción de minerales, el mantenimiento de equipos pesados representa un porcentaje considerable del costo de operación que fluctúa entre el 15% y 60% de los costos [1].

Las empresas del sector minero, con el fin de disminuir los costos de mantenimiento, han implantado distintos enfoques

de mantenimiento dentro de sus organizaciones.

El mantenimiento predictivo es uno de los enfoques de mantenimiento cuya finalidad es evaluar el estado de un equipo con el objetivo de predecir la falla antes que ocurra [6]. Las técnicas de mantenimiento predictivo permiten preveer fallas que puedan causar periodos largos de reparación y penalizaciones económicas elevadas por la falta de obtención del mineral.

Para el empleo de técnicas predictivas es necesario el uso de sensores y dispositivos de monitoreo que proveen gran cantidad de datos ya sean almacenados en memoria o en tiempo real.

La creciente disponibilidad de datos esta cambiando la manera en cómo las industrias toman las decisiones en importantes áreas como logística, gestión del mantenimiento y calidad [4]. Los métodos para generar modelos predictivos se pueden clasificar en dos categorías [3], como: a) modelos basados en la Física, que hacen el uso de métodos numéricos y ecuaciones diferenciales, b) modelos basados en datos, que intentan generar modelos a partir de los datos con la ayuda de técnicas de Machine Learning (ML).

Machine Learning (ML), hace el uso de métodos estadísticos para construir modelos predictivos a partir de datos, es adecuado para trabajar con la gran cantidad de dimensiones y características que poseen los datos. Así tenemos a: [14] que muestra el uso de técnicas de Deep Learning (DL) para identificar fallas en turbinas de manera temprana, [4] muestra el uso de algoritmos de clasificación para detección de fallas en dispositivos semiconductores, [6] muestra un modelo para predecir falla de motores usando redes neuronales recurrentes (RNN), [15] propone un modelo híbrido de LSTM para predecir las fallas en vehiculos de transporte urbano, [2] muestra a predicción de epitaxia en circuitos integrados. Los trabajos mencionados hacen el uso de datos (p. ej. registro de eventos de falla, voltaje, presión, temperatura, caudal, vibracion, sonido, etc.) [4] que son también encontrados en la industria de extracción de minerales.

Se busca aprovechar la información de los datos que se pueden extraer del monitoreo de procesos mineros: a) registro de eventos de fallas y b) costo de que implica la falla. Dichos datos tienen disponibilidad, facilidad y bajo costo de obtención, ya que no es necesario el uso de sensores ni de servicio de instalación de los mismos en los puntos importantes del proceso de extracción de mineral. Dichos datos proveen información valiosa que se puede aprovechar mediante las técnicas de ML para predecir fallas futuras así como el costo que implicaría si llegan a ocurrir.

Las siguientes secciones de este paper están organizadas de la siguiente manera: Sección II describe el proceso del mantenimiento y la importancia en la industria. La Sección III nos define el concepto de Machine Learning, los tipos de aprendizaje y una descripción de la base de datos. La Sección IV el framework propuesto y finalmente la Sección V presenta las conclusiones.

## II. MANTENIMIENTO

El eficiente manejo y control del mantenimiento de equipos pesados en un sector productivo como la minería, es esencial para disminuir los costos asociados a reparaciones y paradas de equipos. El mantenimiento puede ser dividido en cuatro categorías, con distintos niveles de eficiencia y complejidad:

### A. *Mantenimiento Correctivo (R2F)*

Es ejecutado después de ocurrida la falla. Este enfoque es el más simple, pero también el más costoso debido a la gran cantidad de fallas posteriores a la falla original [2].

### B. *Mantenimiento Preventivo (PVM)*

Es ejecutado periódicamente en base a un programa de reparación con el objetivo de anticiparse al proceso de falla. En este enfoque las fallas son evitadas, pero conlleva que algunas veces el mantenimiento sea innecesario ya que no toma en cuenta el estado del mismo [2].

### C. *Mantenimiento basado en la condición*

Es ejecutado después de haber tomado verificaciones de una o más condiciones que indiquen la degradación del sistema o equipo. Se basa en el monitoreo continuo del equipo o sistema y permite realizar el mantenimiento cuando realmente se necesita [2].

### D. *Mantenimiento Predictivo (PdM)*

Similar al mantenimiento basado en la condición, es realizado solo cuando es necesario. Hace uso de herramientas estadísticas para implementar cronogramas de mantenimiento basado en la condición del equipo o sistema [2]. El objetivo del mantenimiento predictivo (PdM) es hacer una estimación precisa de la vida útil restante (RUL - Remaining Useful Life) del componente o sistema [6].

Con el fin de evaluar la RUL, el mantenimiento predictivo hace el uso de sensores y dispositivos de monitoreo instalados

en los principales componentes que comprometen la continuidad de la operación. Para tal efecto, el mantenimiento predictivo analiza el historial de datos proporcionados por los sensores para hallar patrones de falla que pueden ser utilizados para optimizar las estrategias de mantenimiento y reducir las paradas de equipo [6].

Cuando los sistemas empiezan a tornarse más complejos, se incrementan los sensores y por consiguiente la data de cada uno de los equipos pesados, es cada vez más inmanejable encontrar patrones de falla y estimar la RUL. Es por esta razón que en los últimos años las técnicas de ML están siendo usadas cada vez más para predecir las condiciones de trabajo de un componente[6].

## III. MACHINE LEARNING

Son un conjunto de algoritmos que permiten representar el aprendizaje mediante modelos estadísticos que al ser alimentados con datos que posean características relevantes, descubren automáticamente patrones para la clasificación y detección [7].

ML es hacer que las computadoras modifiquen o adapten sus acciones de tal manera que las decisiones tomadas por el modelo sean precisas con respecto a los datos de entrenamiento[10].

En general los modelos de ML generalizan los aspectos más importantes de los datos. Los algoritmos de ML están divididos en dos categorías de aprendizaje:

- Aprendizaje No Supervisado, el aprendizaje es obtenido a partir de un conjunto de datos con gran cantidad de características no etiquetadas [9]. Generalmente, están diseñados para problemas de agrupamiento (Clustering) [10].
- Aprendizaje Supervisado, es el aprendizaje obtenido a partir de un conjunto de datos cuyas características están identificadas y debidamente etiquetadas, asociando las entradas con las salidas [9]. Generalmente están diseñados para resolver problemas de clasificación y regresión. La fase de entrenamiento depende de las características seleccionadas y de la información de supervisión [10].

Para la predicción de la RUL, los métodos de aprendizaje supervisado como: support vector machine (SVM), árboles de decisión (DT), Random Forest (RF) [6] y Deep Learning [7] han sido propuestas en la revisión de la literatura.

### A. *Deep Learning*

Deep Learning (DL) es una red neuronal de múltiples capas intermedias, que realiza la extracción de características

de manera automática, permitiendo a los investigadores obtener características relevantes con un mínimo esfuerzo y sin necesidad de poseer un profundo dominio del tema a resolver. DL es popular en la clasificación de imágenes, videos, análisis de redes sociales y minería de datos [8].

De acuerdo con [8], los algoritmos de DL se clasifican en:

- 1) *Recursive Neural Networks (RvNN)*, se caracterizan por el uso de estructuras jerárquicas para la clasificación. Usado usualmente para el Natural Language Processing (NLP) [8].
- 2) *Convolution Neural Networks (CNN)*, se caracterizan por el uso de las múltiples capas para la convolución de los datos de entrada, compartiendo los pesos y conexiones locales [7]. Originalmente usadas para el reconocimiento de imágenes, es extendida para el NLP, reconocimiento del habla y visión computacional[8].
- 3) *Recurrent Neural Networks (RNN)*, se caracterizan por poder resolver problemas con información secuencial. Esta red suele ser elegida para el procesamiento y recocimiento del lenguaje [8].

### B. Recurrent Neural Networks (RNN)

Son redes neuronales cuya especialidad es el procesamiento de datos distribuidos secuencialmente  $x^1, \dots, x^n$ . [11].

Las RNN procesan un elemento de la secuencia de datos a la vez, manteniendo en sus celdas ocultas un 'vector estado' que implícitamente contiene información acerca del historial de todos los elementos de la secuencia que ya pasaron por la RNN.

Las RNN son modelos dinámicos y presentan el problema de desvanecimiento de la gradiente de aprendizaje durante el entrenamiento del modelo, ocasionando que la gradiente se incremente de manera abrupta o se desvanezca después de varios ciclos de entrenamiento [7]. La arquitectura general se aprecia en la Fig. 1.

$$\begin{aligned} h_t &= q(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \\ y_t &= r(W_{hy}h_t + b_y) \end{aligned}$$

Donde:

- $x_t$  = Es la entrada en el tiempo  $t$ .
- $h_t$  = Es el estado oculto (memoria) en el tiempo  $t$ .
- $y_t$  = Es la salida en el tiempo  $t$ .
- $W$  = Son los distintos pesos de las neuronas.

Para corregir problema de desvanecimiento de la gradiente, se propuso el uso de redes Long Short Term Memory (LSTM) en el estado del arte. La arquitectura general se aprecia en la Fig. 1.

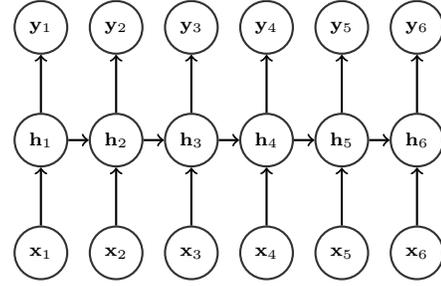


Fig. 1. Red neuronal recurrente

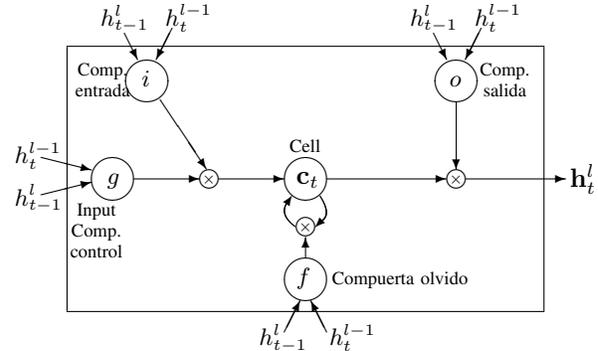


Fig. 2. Representación gráfica de una neurona LSTM usada en este paper adaptada de [11]

### C. Long Short Term Memory (LSTM)

Es un tipo de red neuronal recurrente creado por Hochreiter & Schmidhuber 1997 [11], hace uso de unidades especiales de memoria llamadas *celdas* que trabaja de manera conjunta con la neurona teniendo con compuertas definidas en [12], como: a) *compuerta de entrada*, b) *compuerta de salida*, c) *compuerta de control* y d) *compuerta de olvido*. El uso de las *celdas* en la red LSTM permite conservar o desechar valores que afectan la gradiente de aprendizaje a lo largo de la secuencia de datos.

La principal ventaja una red LSTM con respecto a una RNN es la capacidad que tiene la LSTM de conservar en memoria conceptos importantes entre largos periodos en una secuencia de datos [6]. En la Fig. 2 se muestra como esta organizada una red neuronal LSTM y sus cuatro compuertas internas, definidas en la siguiente ecuación:

$$\begin{aligned} i &= \sigma(W^i H + b^i) \\ f &= \sigma(W^f H + b^f) \\ o &= \sigma(W^o H + b^o) \\ c &= \tanh(W^c H + b^c) \\ m' &= f \odot m + i \odot c \\ h' &= \tanh(o \odot m') \end{aligned} \quad (1)$$

Donde  $\sigma$  es la función sigmoidea,  $W^i, W^f, W^o, W^c$  en  $\mathbb{R}^{d \times 2d}$  son las matrices de peso recurrente y  $b^i, b^f, b^o, b^c$  son

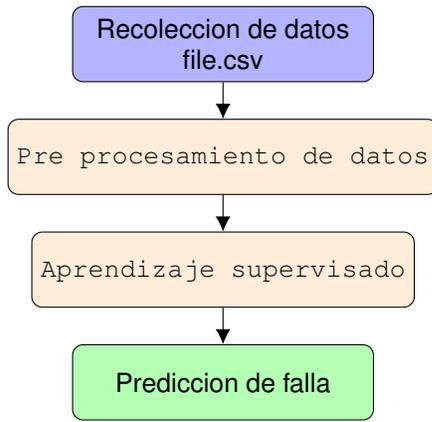


Fig. 3. Framework de detección de fallas

las correspondientes bias.  $H$  en  $\mathbb{R}^{2d}$  es la concatenación de la nueva entrada  $x_i$  y el previa vector oculto  $h_{i-1}$ :

$$H = \begin{bmatrix} Ix_i \\ h_{i-1} \end{bmatrix} \quad (2)$$

La diferencia entre una red LSTM y una RNN es el empleo de *celdas* estado, que pueden recordar información a largo plazo. La red LSTM tiene la habilidad de remover y agregar información a las *celdas* de estado, reguladas por estructuras llamadas compuertas.

Las compuertas en la ecuación (1) son  $i$ ,  $f$ ,  $o$ ,  $c$ , representando la *compuerta de entrada*, *compuerta de olvido*, *compuerta de control* y *compuerta de salida*. La compuerta de entrada puede decidir cuanta información ingresa a la actual celda. La compuerta de olvido decide cuanta información es olvidada de vectores previos. La puerta de control decide que nueva información es escrita en la celda de estado. La compuerta de salida decide que información se enviará a la siguiente neurona [3].

Ahora que conocemos las propiedades de la red Neuronal LSTM, podemos aplicar el modelo a los datos que serán usados en nuestra propuesta. En la siguiente sección describiremos a más detalle el framework propuesto.

#### IV. FRAMEWORK PROPUESTO PARA LA DETECCIÓN DE FALLAS

A continuación detallaremos el conjunto de datos de entrada, la recolección de datos, el pre procesamiento, el aprendizaje y la predicción de falla, como se aprecia en la Fig. 3.

##### A. Descripción del Conjunto de Datos (Dataset)

Se detalla el dataset que usaremos para entrenar nuestro modelo. El objetivo es predecir las futuras fallas y el costo de realizar el mantenimiento si ocurriese la falla. El costo de mantenimiento contiene mano de obra para realizar el trabajo, compra de componentes y herramientas para realizar

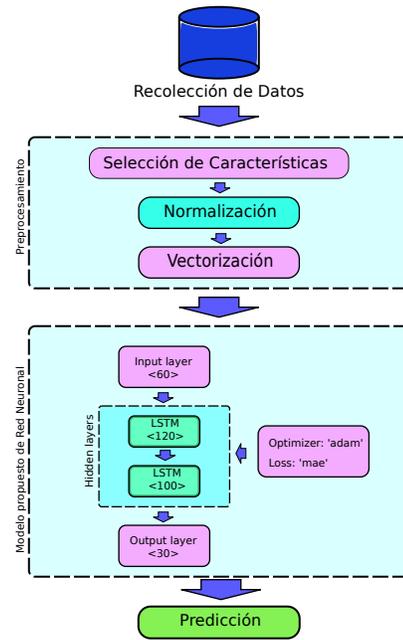


Fig. 4. Arquitectura propuesta: 1 Cap

la reparación o cambio según sea el caso.

Se enfrentará el problema con el enfoque de regresión supervisada. Esto quiere decir que los datos generados por el modelo son estimados de tiempo y costo de la próxima falla.

Como el enfoque usado es el aprendizaje supervisado necesitamos un conjunto de datos que contengan el historial de fallas, costos y características que sean identificadas y debidamente etiquetadas.

Los datos provienen del historial de fallas de una empresa minera de extracción de cobre. Entre los tipos de equipos registrados, en el proceso de minado tenemos: palas, perforadoras, volquetes, tractores.

En el proceso de molienda se utilizan maquinarias pesadas como: chancadoras y molinos. Los datos consisten en log de fallas diarias de los equipos pesados y los costos cargados a dichas fallas.

A continuación, se detalla las actividades de preprocesamiento y entrenamiento de la red neuronal de trabajo mostrado en la Fig. 4.

##### B. Recolección de datos

Los datos a recolectar que provienen de dos fuentes: a) *registro de eventos de falla*, que almacena equipo, fecha, hora, tipo de falla y breve descripción de la falla y b) *órdenes de mantenimiento*, que contiene detalles la avería como, componente y costos de compra de materiales de reparación. La

importancia de usar dichas fuentes radica en la *disponibilidad* y *facilidad* para obtenerlos. Los datos son organizados en una tabla y guardados en un archivo CSV (Comma separated value).

### C. Pre procesamiento de datos

Antes de introducir los datos al modelo de predicción de fallas (LSMT), es necesario realizar un trabajo previo de revisión de datos con la finalidad de incrementar el grado de precisión del modelo LSTM a entrenar retirando información incompleta o mal asignada, tenemos:

- 1) *Selección de características*, para el entrenamiento del modelo LSTM las características requeridas son: a) fecha de falla, b) descripción de equipo y c) costo de mantenimiento.
- 2) *Limpieza de datos*, es necesario eliminar datos con información incompleta o innecesaria tales como: duplicidad de eventos, asignación errónea de código de falla, ausencia de descripción de equipo.
- 3) *Normalización de datos*, con el método MinMax para que los datos estén en una escala entre 0 y 1 para una sencilla manipulación.
- 4) *vectorización*, preparar los datos de acuerdo en vectores que contengan el: 1) nombre del equipo, 2) subsistema del equipo y 3) duración de falla.

Con las características seleccionadas y limpieza de las mismas, nos disponemos a introducir la data al modelo LSTM para el entrenamiento.

### D. Aprendizaje supervisado

Para lograr que el modelo LSTM aprenda a predecir fallas mediante el uso del registro de eventos de falla, se propone el entrenamiento con un dataset conteniendo datos de eventos de fallas, el dataset puede variar por importancia y costo de reparación de equipos pesados.

Se propone inicialmente tomar datos de las palas debido a que el costo de los componentes de este tipo de equipo pesado es el más costoso. Como se mencionó inicialmente, los precios promedio de uno de sus componentes se sitúa entre los USD\$300,000.

El dataset está compuesto de 140,000 eventos registrados en un período de 4 años comprendidos entre el año 2016 y el año 2019. De los cuales el 70% será asignado al entrenamiento y el 30% servirá de supervisión para el ajuste de resultados.

El modelo de ML a usar es LSTM, debido a su facilidad para trabajar con datos en serie. Nuestro modelo esta compuesta por una red neuronal con una capa de 60 entradas con 3 características: 1) identificación de equipo, 2) sistema

con el problema y 3) duración de la falla. Dos capas ocultas con 120 y 100 perceptrones recurrentes y una capa de salida con 30 perceptrones.<sup>1</sup>.

### E. Predicción de falla

Los resultados que se esperan obtener del modelo LSTM permitirán predecir la posible falla buscando que el margen de error sea menor a 7 días, ya que dicho plazo permite generar órdenes de mantenimiento programado en el plan semanal.

## V. RESULTADOS Y CONCLUSIONES

El framework propuesto, y actualmente en desarrollo de acuerdo a la arquitectura mostrada en la figura 4, busca la implantación de técnicas de ML en los procesos de mantenimiento. Se concluye que el uso de ML será beneficioso en el proceso de mantenimiento de equipos pesados ya que no requiere de gran cantidad de especialistas para poder generar modelos predictivos y con el crecimiento de datos disponibles, los modelos basados en datos serán los apropiados para dichas tareas.

## REFERENCES

- [1] Mobley, R Keith, "An introduction to predictive maintenance", 2002, Elsevier.
- [2] G. A. Susto and A. Beghi and C. De Luca, "A Predictive Maintenance System for Epitaxy Processes Based on Filtering and Prediction Techniques", IEEE Transactions on Semiconductor Manufacturing", vol. 25, no. 4, pp. 638-649, 2012, doi: 10.1109/TSM.2012.2209131.
- [3] Q. Zhang and H. Wang and J. Dong and G. Zhong and X. Sun, "Prediction of Sea Surface Temperature Using Long Short-Term Memory", IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 10, pp. 1745-1749, 2017, doi: 10.1109/LGRS.2017.2733548.
- [4] G. A. Susto and A. Schirru and S. Pampuri and S. McLoone and A. Beghi, "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach", IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812-820, 2015, doi: 10.1109/TII.2014.2349359.
- [5] Sipos, Ruben and Fradkin, Dmitriy and Moerchen, Fabian and Wang, Zhuang, "Log-Based Predictive Maintenance", Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, pp. 1867-1876, 2014, doi: 10.1145/2623330.2623340.
- [6] D. Bruneo and F. De Vita, "On the Use of LSTM Networks for Predictive Maintenance in Smart Industries", 2019 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 241-248, 2019, doi: 10.1109/SMARTCOMP.2019.00059.
- [7] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, 2015, doi: 10.1038/nature14539.
- [8] Pouyanfar, Samira and Sadiq, Saad and Yan, Yilin and Tian, Haiman and Tao, Yudong and Reyes, Maria Presa and Shyu, Mei-Ling and Chen, Shu-Ching and Iyengar, S. S., "A Survey on Deep Learning: Algorithms, Techniques, and Applications", ACM Comput. Surv., vol. 51, no. 5, 2018, doi: 10.1145/3234150.
- [9] Goodfellow, Ian and Bengio, Yoshua and Courville, Aaron, Deep learning, 2016, MIT press.
- [10] Li, Jundong and Cheng, Kewei and Wang, Suhang and Morstatter, Fred and Trevino, Robert P. and Tang, Jiliang and Liu, Huan, "Feature Selection: A Data Perspective", ACM Comput. Surv., vol. 50, no. 6, 2017, doi: 10.1145/3136625.
- [11] Hochreiter, Sepp and Schmidhuber, Jürgen, "Long short-term memory", Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.
- [12] Graves, A., "Supervised Sequence Labelling with Recurrent Neural Network", vol. 385, no. 9, 2013.
- [13] Harmant H. and Muntmanky J., "Introductory Mining Engineering", 2002, Wiley press.

- [14] W. Lu and Y. Li and Y. Cheng and D. Meng and B. Liang and P. Zhou, "Early Fault Detection Approach With Deep Architectures", IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 7, pp. 1679-1689, 2018.
- [15] G. Zhu and H. Zhao and H. Liu and H. Sun, "A Novel LSTM-GAN Algorithm for Time Series Anomaly Detection", 2019 Prognostics and System Health Management Conference (PHM-Qingdao), pp. 1-6, 2019.