

Modelo de predicción para la vida útil restante de una batería a partir de redes CNN-LSTM en R

Catherine Rincón, Carmen Patiño, Olga Úsuga

Palabras clave: Batería; vida útil restante; predicción; LSTM; CNN

Resumen

Actualmente las redes neuronales se han convertido en un potencial campo de investigación en diferentes áreas de la ingeniería, ya que las redes neuronales tienen la habilidad de calibrarse y generalizar. Esto permite trabajar con problemas complejos que presentan no linealidad y ruido en sus datos, al reconocer patrones y predecir comportamientos. Esto ha permitido a los diferentes campos de la ingeniería implementar metodologías basadas en redes neuronales con resultados que permiten predecir adecuadamente la realidad. Las redes neuronales han sido implementadas en el diagnóstico de fallas, en mantenimiento, en sistemas predictivos, entre otros.

En esta investigación, se presenta una metodología de predicción que pretende modelar y predecir la degradación de una batería a lo largo del tiempo. La metodología propuesta, se implementa en R a partir de librerías como Tensorflow y Keras, que permiten generar diferentes modelos implementando redes neuronales. La metodología de predicción considera una estrategia de división de los datos en cuatro partes: entrenamiento, prueba, validación y degradación, que permiten al modelo obtener datos de patrones frecuentes y predecir con precisión la degradación de la batería a lo largo de los ciclos de carga y descarga. La arquitectura del modelo predictivo combina dos tipos de redes neuronales: redes convolucionales (CNN), que permiten reducir el ruido de los datos obtenidos a partir de sensores, y una red de corto y largo plazo (LSTM), que captura las dependencias a largo plazo para la predicción de la vida útil restante de la batería.

Para lograr una buena predicción en el modelo predictivo, es importante elegir cuidadosamente la arquitectura de la red LSTM y definir cada uno de los valores de los hiperparámetros asociados a la red, tales como: número de neuronas, tamaño del lote (*batch size*), épocas, tasa de aprendizaje, entre otros. Debido a la importancia de la elección de los hiperparámetros en el modelo predictivo, se implementa dentro de la metodología, un diseño de experimentos a partir de diferentes funciones de R tales como *aov*, que permitan definir los mejores valores para cada uno de los hiperparámetros mejorando la precisión de la predicción y evitando el sobreajuste que se puede presentar en los datos a partir de técnicas de regularización implementadas en la arquitectura de la red LSTM.

Para validar el modelo propuesto, se usan los datos públicos obtenidos de la NASA, conocidos como: *Battery Data Set*, los cuales han sido ampliamente usados para predecir la vida útil restante de una batería. La degradación de la batería se lleva a cabo a partir de diferentes reacciones fisicoquímicas y variación de parámetros en el tiempo que presentan un comportamiento complejo y no lineal que dificulta su modelado y precisión. Por lo cual, la predicción de la vida útil restante de una batería es un reto fundamental, ya que la batería es el principal suministro de energía y componente de almacenamiento de la electricidad que permite a un equipo funcionar correctamente. Una falla en la batería puede ocasionar incendios, fugas y explosiones y, en consecuencia, se puede generar una falla

MODELO DE PREDICCIÓN PARA LA VIDA ÚTIL RESTANTE DE UNA BATERÍA A PARTIR DE REDES CNN-LSTM EN R

en el correcto funcionamiento de todo el equipo. Por lo cual, la predicción de la vida útil restante de la batería es un factor clave para obtener información que permita garantizar el funcionamiento seguro y confiable de la batería.

Se considera un conjunto de datos de 4 baterías que fueron sometidas a una temperatura ambiente de 24 °C, las cuales fueron instrumentadas para capturar la información de las siguientes variables: ciclo de carga/descarga, fecha y hora de inicio del ciclo, voltaje, corriente, temperatura, corriente de carga, voltaje de carga, tiempo de ciclo y capacidad de la batería. En conjunto, se tiene una base de datos de 185.721 registros y 10 columnas. En la Figura 1 se observa la matriz de correlaciones de las variables y la gráfica de la capacidad de la batería a lo largo del tiempo, obtenidos del análisis exploratorio. Se observa una relación lineal fuerte inversa entre variables como: ciclo de carga/descarga y capacidad de la batería, temperatura y voltaje, y tiempo y voltaje. Así mismo, se ve reflejado una relación lineal fuerte directa entre las variables tiempo y temperatura.

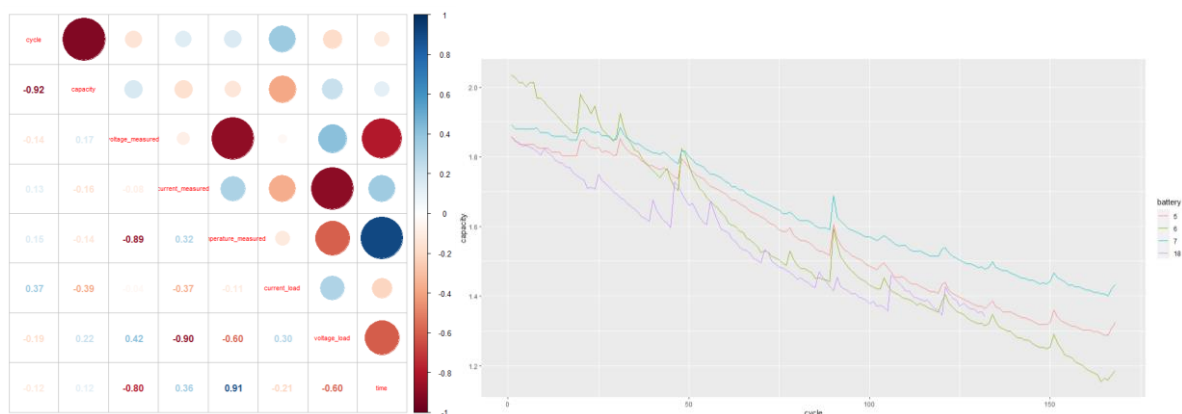


Figura 1. Resultados del análisis exploratorio con el conjunto de datos *Battery Data Set*

En la Tabla 1, se observan los resultados obtenidos de diferentes modelos de predicción: LSTM y CNN-LSTM con 3 conjuntos de datos: train, test y validación, además de modelos LSTM y CNN-LSTM con 4 conjuntos de datos: train, test, validación y degradación. Cada modelo fue evaluado a partir del RMSE. Estos modelos implementados para la predicción de la degradación de la capacidad a lo largo de los ciclos de carga/descarga obtienen valores promedio de RMSE para las 4 baterías inferiores al 5%.

Tabla 1. Resultados de los modelos de predicción implementados

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
RMSE	0.072	0.050	0.080	0.056

Esta investigación implementa aspectos diferenciadores en la división de los datos, considerando cuatro conjuntos de datos de degradación que permiten identificar patrones frecuentes y factores de degradación. Además, se implementa en el modelo una arquitectura CNN que permite el mejoramiento del ruido de los datos, ya que son datos obtenidos a partir de sensores. Por estas razones, a través de la combinación de diferentes metodologías se obtiene un modelo que permite realizar predicciones con mayor precisión.

Catherine Rincón Maya
 Universidad de Antioquia
 Catherine.rincon@udea.edu.co