A study of ML/DL based Remaining Useful Life Prediction of

Lithium-ion Battery for preventive maintenance.

DSECLZG628T PROJECT WORK

by

AMBUJ KUMAR

2019HC04622

Project work carried out at
Qolsys software India Pvt Ltd, Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
Pilani (Rajasthan) INDIA February 2022

DSE CL ZG628T PROJECT WORK

A study of ML/DL based Remaining Useful Life Prediction of Lithium-ion Battery for preventive maintenance.

Submitted in partial fulfillment of the requirements of the

M. Tech. Software Engineering Degree programme

By

AMBUJ KUMAR

2019HC04622

Under the supervision of

Jagadisha Gangulli

Technical Manager

Qolsys software India Pvt Ltd., Hyderabad

Project work carried out at

Qolsys software India Pvt Ltd,

Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE

PILANI (RAJASTHAN)

January 2022

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE,
PILANI

CERTIFICATE

"This is to certify that the Project Work entitled **A study of ML/DL based Remaining Useful Life Prediction of Lithium-ion Battery for preventive maintenance** and submitted by AMBUJ KUMAR IDNo.2019HC04622 in partial fulfillment of the requirements of DSECL ZG628T Project Work embodies the work done by him under my supervision.

Jagadisha. G

Signature of the Supervisor
Name: Jagadisha Gangulli
Designation: Technical Manager,
Qolsys software India Pvt Ltd, Hyderabad

Place: Hyderabad

Date: 21-Feb-2022

ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to my supervisor Mr. **Jagadisha Gangulli**, for guiding me throughout the effort of this project and directing this project to reach its best form.

I would also like to thank my evaluator, Ms. **C V KRISHNAVENI** for her valuable guidance throughout this project. You provided me with the tools that I needed to choose the right direction and successfully complete my dissertation.

I would also like to thank my friends and family who helped me a lot in finalizing this project and supporting me in my endeavors in ways both big and small, but always significant.

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

FOURTH SEMESTER 2021-22

DSECLZG628TDISSERTATION

Dissertation Title : A study of ML/DL based Remaining Useful Life Prediction of Lithium-ion Battery for preventive maintenance.

Name of Supervisor : Jagadisha Gangulli

Name of Student : AMBUJ KUMAR

ID of Student : 2019HC0462

Abstract

In the present time, lithium-ion batteries have gained a lot of popularity because of their high charge density, longer time span and portability. It is one of the options for green energy. As many of the green energy applications like Home automation, Electrical Vehicle, Wind Energy and Solar Energy use Li-ion battery as their energy storage device. So, a better and intelligent Remaining Useful Life (RUL) prediction model will improve the reliability of these systems.

Remaining useful life prediction of lithium-ion batteries can reduce the risk of battery failure by predicting the end of life. In this project, I propose RUL prediction techniques based on long short-term memory (LSTM). To estimate RUL even in the presence of capacity regeneration phenomenon, I consider multiple measurable data from battery management system such as voltage, current and temperature charging profiles whose patterns vary as aging.

NASA battery degradation data set is used for this analysis and target data are extracted based on the geometric features. The model evaluation is based on root mean square error.

The main objective of this project can be outlined as follows.

- Conduct Survey and study of current work in this field.
- To understand the charging, discharging cycle of a battery in order to develop a data-driven model to predict remaining time before the failure of the battery.
- Feature Extraction from raw data.
- Implementation and evaluation of a model with LSTM.
- Identify the future course of action research scope.

Prognostic and health management (PHM) can ensure that a lithium-ion battery is working safely and reliably. The main approach of PHM evaluation of the battery is to determine the State of Health (SoH) and the Remaining Useful Life (RUL) of the battery. This dissertation project presents the preliminary development of the data-driven prognostic, using a Deep Learning, LSTM model approach to predict the SoH and the RUL of the lithium-ion battery. The effectiveness of the proposed approach is implemented with a battery dataset obtained from the NASA Ames Prognostics Center of Excellence (PCoE) database. The experimental results reveal that the performance of the LSTM algorithm could outweigh other machine learning algorithms such as Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Artificial Neural Networks (ANN), and Linear Regression (LR). Further, the presented results could serve as a benchmark of SoH and RUL prediction using machine learning approaches specifically for lithium-ion batteries application.

Table of Contents

1. Introduction

At present time, we are heavily dependent on machines, devices and equipment in our day-to-day life. Many of these systems have an energy storage device, battery as their integral part. Various batteries are used in these systems, but Lithium-ion batteries are emerged as one of the best performing option because of its high charge density, portability and long life span. These machines are meant to falter after the continuous use over a period. To ensure the smooth functioning of battery-based equipment, maintenance is performed. It is essential to monitor the health of batteries to ensure the proper functioning of these systems. There are different maintenance programs like reactive maintenance, preventive maintenance and predictive maintenance which are adopted for these systems.

However, battery degradation begins immediately after batteries are manufactured, and when 70% or 80% of initial capacity remains, batteries need be replaced for safe operation. Thus, it is important to predict when battery life will be over.

Reactive maintenance is performed when equipment stops working. It is the traditional way which may leads to downtime period and there is always an uncertainty about the current health of the equipment.

In the preventive maintenance, periodic maintenance is scheduled to ensure that machine does not brake while in operation. In this process the selection of maintenance period is of utmost importance otherwise it may result into over maintenance or under maintenance. These both scenarios are risky and possess substantial time and economic loss.

In predictive maintenance one collects the data regarding the health of the equipment in terms of various sensors and operational data, then accordingly maintenance of the device can be scheduled. This method is cost effective, and it reduces the chances of downtime. Predictive maintenance in batteries is used for their health monitoring, which includes prediction of battery's Remaining Useful Life.

This dissertation project investigates how to make use of Deep Learning based mechanism to solve the prediction of battery's RUL by programming.

2. The objective of the Project

The main objective this project is to design and train/learn Deep Learning model to explore and perform the possibility of using ML/DL based approach for Remaining Useful Life Prediction of Lithium-ion Battery on the battery charge/ discharge dataset obtained from NASA.

3. The benefit to the Organization

This is a research project. Here effort has been put to investigate how Deep Learning based method can be used to apply preventive maintenance for company's reliable operational warranty support to 5 million connected devices customers. It can benefit as -

- ✓ Reduce accruals by reducing financial risk of warranties
- ✓ Find optimal policy for warranties
- ✓ Increased productivity
- ✓ Increased uptime
- ✓ Increased lifetime of equipment's
- ✓ Increased safety standards
- ✓ Decreased cost of maintenance
- ✓ Decreased equipment failures/rate of failure
- ✓ Decreased stock in inventory

The experience of project can become strategically input to introduce ML/DL based methodology in development of preventive maintenance of connected device's battery. This project can be used to demonstrate management team the capabilities of ML/DL based approach/methodology in company engineering process.

This automated solution is going to reduce a huge cost of the company to serve the customers in a batch to a particular area of the city to replace the devices timely and planned way before its end of life. At present the cost is huge to serve the customers in ad-hoc and on-demand manner.

4. Methodology

In presented research, Knowledge Discovery in Database (KDD) methodology is followed for prediction of RUL of Li-ion battery. Its initial phase is focused on understanding the objective and requirement of the project from business point of view and in final stage full deployment in included but this project understood the objective from academic point of view while also knowing its business implication. This project derives the knowledge which shows it is useful way to predict RUL and it has applicability in real systems.

4.1. Data Selection

In order to predict RUL of any device or systems usually uses three kind of dataset

1. Dataset have reading from entire life span of the system.
2. Dataset have only failure reading.
3. Dataset have few readings and their threshold is already defined.

This dissertation uses the NASA PCoE Dataset. This dataset has four battery reading B005, B006, B007 and B0018. Each battery has charging, discharging and impedance profile of multiple cycles and given threshold condition is reduction in battery capacity by 30% of its rated capacity. Batteries in the dataset are 18650-size Li-ion cells that which are used in studies at Idaho National Laboratory. These tests were done under controlled ambient temperature. These files are in MATLAB format which are converted into json format for better readability and usability in python. The selection of this dataset is also motivated by the fact that this is one of the most used data sets for different deep learning network for health monitoring of the Li-ion battery.
Description of Charging, Discharging and Impedance cycle of Li-ion battery in selected data set:
Charging Cycle: Li-ion batteries are rechargeable so they are recharge by constant voltage or constant current source. In this data set Charging of the batteries are done under constant current of 1.5A until the voltage reached to 4.2V (single battery cell's maximum voltage) and then it is continued under this voltage until current dropped to 20mA.
Discharging Cycle: The process of using stored energy in Li-ion battery is called discharging cycle. In this data set discharging is done at constant current of 2A until battery voltages of B005 reached 2.7V, B006 reached 2.5V, B007 reached 2.2V and B0018 reached 2.5 V.
Impedance Cycle: Impedance measurements are taken by Electro-chemical Impedance Spectroscopy (EIS) and selected frequency is from 0.1 Hz to 5kHz.

Table 1: Different Batteries and their Number of Cycle

Batteries	**Number of Cycles**
B0005	168
B0006	168
B0007	168
B0018	132

4.2. Data Preparation

Original data set is in .mat format, in which data is stored in hierarchical format. For the easy usage and make it readable in python, it is converted into json file by enumerating data from MATLAB file to dictionary using loadmat from the SciPy library. For charging, discharging cycle separate json file is created which have cycle as key and inside each cycle various reading are captured regarding that particular cycle. Data is captured till the battery's capacity reached to the threshold condition.

4.3. Feature Selection

The dataset contains multiple charging and discharging cycle and each cycle have different data points. They cannot be used directly for the model creation; instead feature needs to be equal in each cycle. To address this, one option is to take some points randomly form each cycle, but it always has risk of losing the important data, so the approach needs to be backed up by the behavior of the Li-ion battery. Feature extracted must retain the battery's behavior so that a good prediction can be obtained from the data. Other than this, these extracted features requires to remain reliable for all operating condition and for other similar batteries, then only the batteries appropriate degradation can be mapped. The concept of geometric metric is described to estimate the capacity of battery based on their voltage, current and temperature profile. If the capacity estimation is correct one can say that the set of extracted features are accurately representing the charging and discharging of Li-ion battery over their lifetime. I have used same NASA battery dataset to define Geometric features of the Li-ion battery. These features were successful in mapping the capacity degradation. These features can be interpreted as time when voltage reached its maximum value, in our case it is 4.2V, time when current started to drop and time when max temperature reached under operating ambient conditions. These geometric features were able to depict the capacity degradation of Li-ion battery under various operating and aging condition. These geometric features can be seen, tracking the time-based relationship between internal parameters of battery with each charging discharging cycle. Similarly, for temperature as the battery get old it takes more and more time to reach same maximum temperature. During discharging similar trend is followed for voltage and current as the battery get old discharging process become quickly.

The entire features were extracted based on below equations.

For Charging Cycle, Batteries terminal voltage is according to equation (1):

$$(ti , vi), \text{ s.t. } vi \geq 4.2V \quad i = 1, 2, 3, 4 . . . n \qquad (1)$$

In the above equation (t) is a time when the battery voltage reaches the maximum value for the first time and (v) is the maximum voltage achieved by the battery during charging cycle, (i) is no of cycle up to (n), which represent sample size. Batteries terminal current is according to equation (2):

$$(ti , Ai), \text{ s.t. } Ai \leq 1.5Amp \quad i = 1, 2, 3, 4 . . . n \qquad (2)$$

In the above equation (t) represent the time when the current started to drop.(A) is the value of the current when it started to drop.(n) is total sample size.

$$(t_i, T_i) = t_i, T_i \text{ at } \max T_i \quad i = 1, 2, 3, 4 \ldots n \qquad (3)$$

In the above equation, (t) is the time when the temperature reaches the maximum value. (T) is the maximum temperature of the battery during charging. max T is the maximum temperature. (n) is sample size. Batteries current at load is according to the equation (4):

$$(t_i, A_i), \text{ s.t. } A_i \text{ just before it drops } i = 1, 2, 3, 4 \ldots n \qquad (4)$$

Where, (t) is time just before current starts to drop. (A) is value of current in Amp when it started to drop. (n) is sample size.

$$(t_i, v_i) = (t_i, v_i) \text{ at } v_{imax} \quad i = 1, 2, 3, 4 \ldots n \qquad (5)$$

Where, (t) is time at which voltage at load reaches maximum value. (v) is value of maximum voltage at the load. (n) is number of samples. For Discharging cycles, Batteries terminal voltage is according to equation (6):

$$(t_i, v_i) = (t_i, v_i) \text{ at } v_{imin} \quad i = 1, 2, 3, 4 \ldots n \qquad (6)$$

Where,(t) is time when batteries voltage reaches its minimum value. (v) is minimum voltage of battery during discharging. (n) is sample size. Batteries terminal current is according to equation (7):

$$(t_i, A_i), \text{ s.t. } A_i > -2A \quad i = 1, 2, 3, 4 \ldots n \qquad (7)$$

Where,(t) is time when the terminal current gradually start increasing. (A) is value of current when it start increasing. (n) is sample size. Batteries Temperature is according to equation (8):

$$(t_i, T_i) = t_i, T_i \text{ at } \max T_i \quad i = 1, 2, 3, 4 \ldots n \qquad (8)$$

Where, (t) is time at which temperature reaches its maximum value. (T) maximum temperature value achieved by battery during discharging. (n) is sample size. Batteries current at load is according to equation (9):

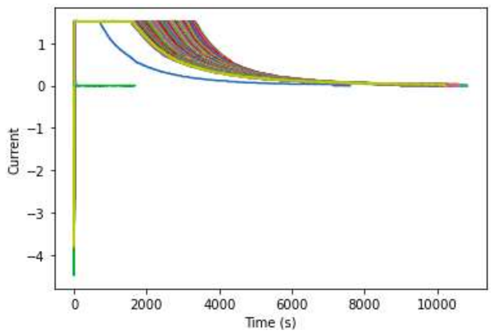
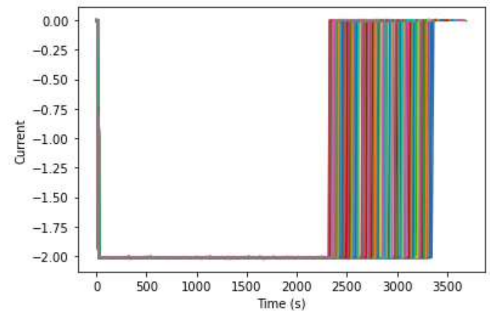
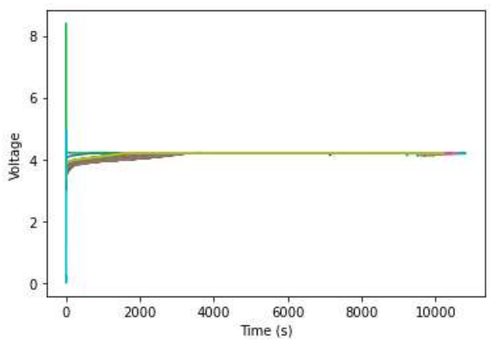
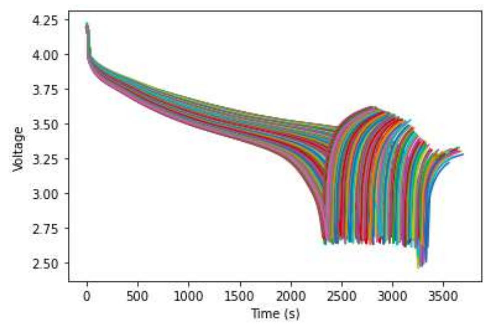
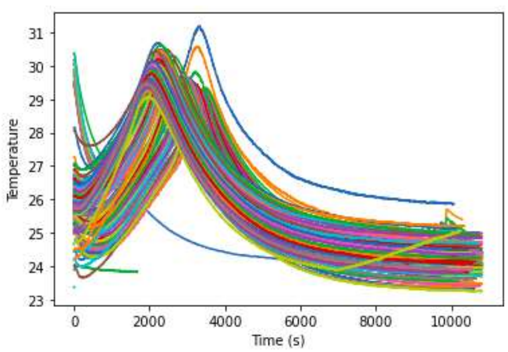
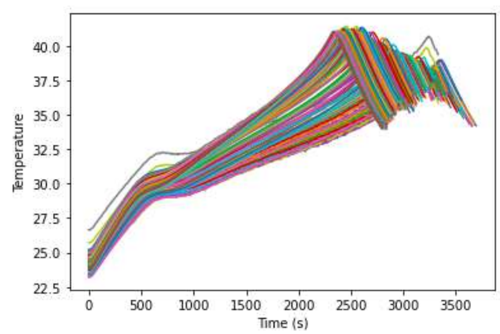
$$(t_i, A_i), \text{ s.t. } A_i > -2A \quad i = 1, 2, 3, 4 \ldots n \qquad (9)$$

Where,(t) is time when the current at load gradually start increasing. (A) is value of current measured at load when it start increasing. (n) is sample size. Batteries voltage at load during discharging is according to equation (10):

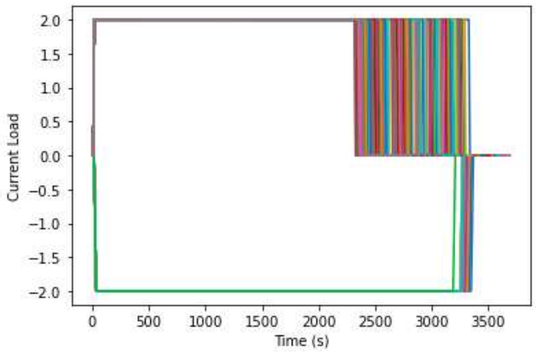
$$(t_i, v_i) = \{(t_i, v_i) \text{ at } \min(v_i) \text{ s.t. } v_i \, 6{=}\, 0\} \quad i = 1, 2, 3 \ldots n \qquad (10)$$

Where,(t) is time when voltage is minimum but not zero. (v) non zero minimum voltage value. (n) is sample size. This give 20-dimensional dataset, corresponding capacity of batteries also added into this so the final data is 21-dimensional.

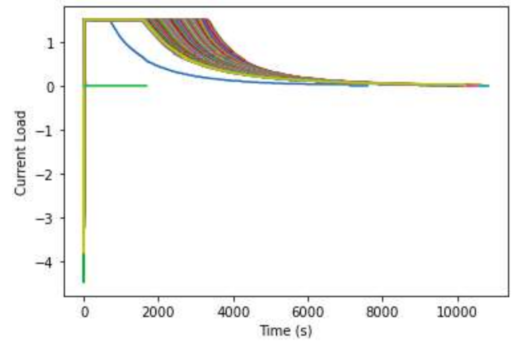
4.4. Visualization

<table>
<tr><td>Charging Cycle</td><td>Discharging Cycle</td></tr>
<tr><td>1. Current vs. Time
</td><td>1. Current vs. Time
</td></tr>
<tr><td>2. Voltage vs. Time
</td><td>2. Voltage vs. Time
</td></tr>
<tr><td>3. Temperature vs. Time
</td><td>3. Temperature vs. Time
</td></tr>
</table>

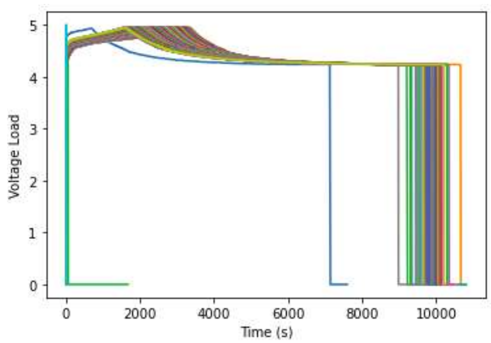
4. Current Load vs. Time



4. Current Load vs. Time



5. Voltage Load vs. Time



5. Voltage Load vs. Time

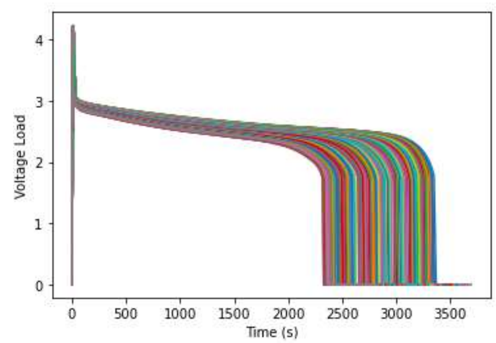


Figure 1: Changes in Different Battery Parameter over the different Cycles

4.5. Data Mining

This paper predicts the RUL of Li-ion battery in term of remaining cycle. Battery used is mostly having 168 cycles as their end of life cycle. So, it is a regression kind of problem where a number will be predicted for the RUL of battery.

4.6. RUL Calculation for Training of the Prediction Model

Before passing the data to prediction model, calculation of RUL is required, this dataset have reading till the battery reaches to their end of life criteria. For example B005 has 168 cycles let's call it (n), let's say battery for prediction is currently in ith cycle, RUL can be calculated as

$$RULi = n + 1 - i \text{ W here,}(0 < i < n) \qquad (11)$$

At this stage final data for prediction model is converted into supervised dataset with corresponding reading of RUL of the battery using equation (11).

4.7. Data Normalization

In process of model creation often feature have different ranges, which may introduce some error into the result. To avoid this paper uses data normalization technique, where all the data is normalized in the range of 0 to 1 using minimum maximum normalization technique.

$$X_{normalized} = x - \min(x)/\max(x) - \min(x) \qquad (12)$$

It is widely used normalization techniques it ensures that all the data are in same scale on other hand it is less effective with the outlier but in present case outlier are not the concern.

5. Design Specification

Process flow of this paper is presented below. The entire model developing process divided into three parts viz. Data definition, Model Training, Model Tuning, Model Testing and RUL prediction.

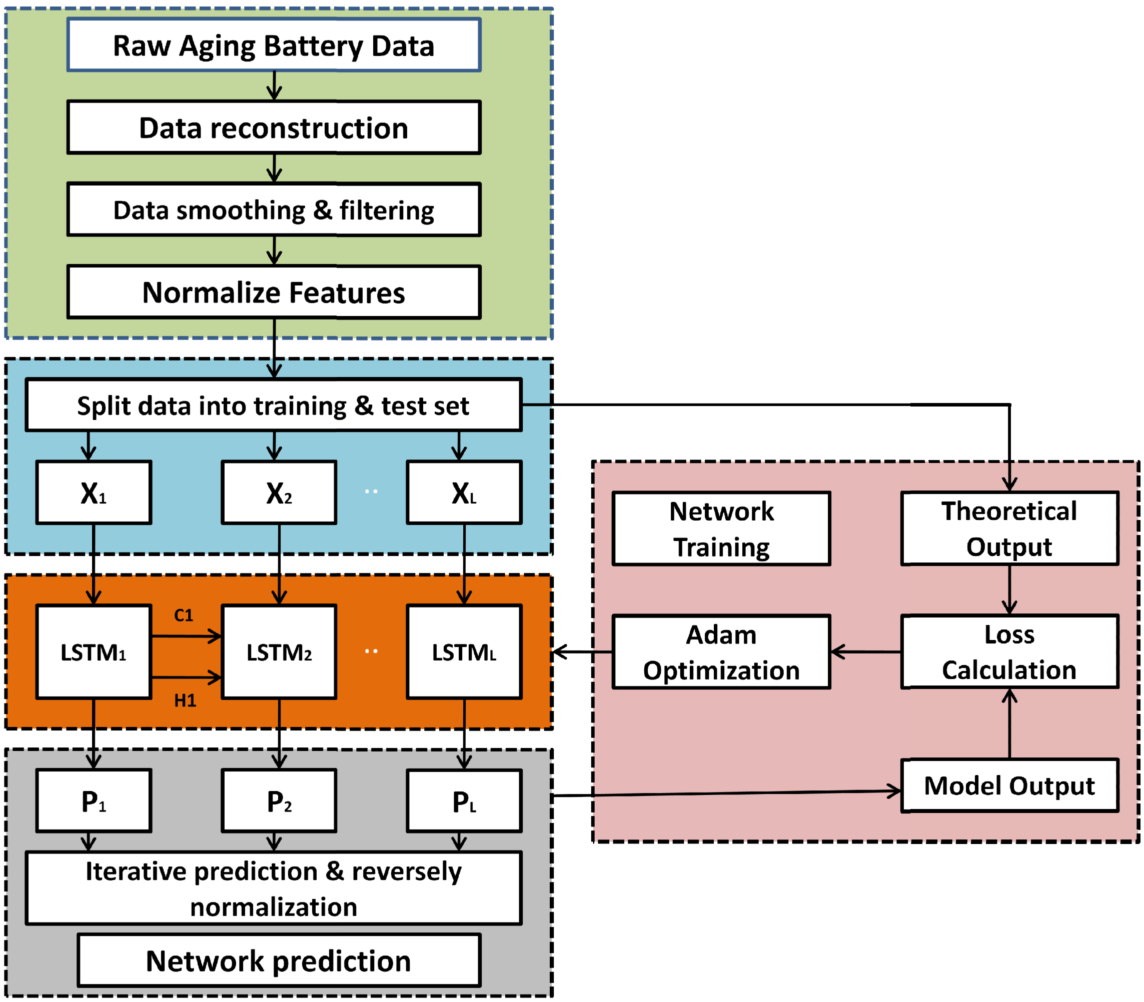


Figure2: The process of a prognostic framework using LSTM deep learning model

Deep learning is an improvement of a Multi-Layer Perceptron with a better power of learning representation, which holds the potential to overcome the deficiencies in traditional machine learning methods. The notable

advantage of deep learning is that it is able to capture the representation of information from raw data through multiple complex non-linear transformations and approximations. The main algorithms of deep learning include the Deep Neural Network, the Convolutional Neural Network (CNN), the Recurrent Neural Network (RNN), and the expansion of CNN and RNN, such as the Long short-term memory network (LSTM). However, RNN has a limitation in capturing the length of the data. This leads to the development of the LSTM network, which can capture longer sequences of information.
LSTMs deal with both Long Term Memory (LTM) and Short Term Memory (STM) and for making the

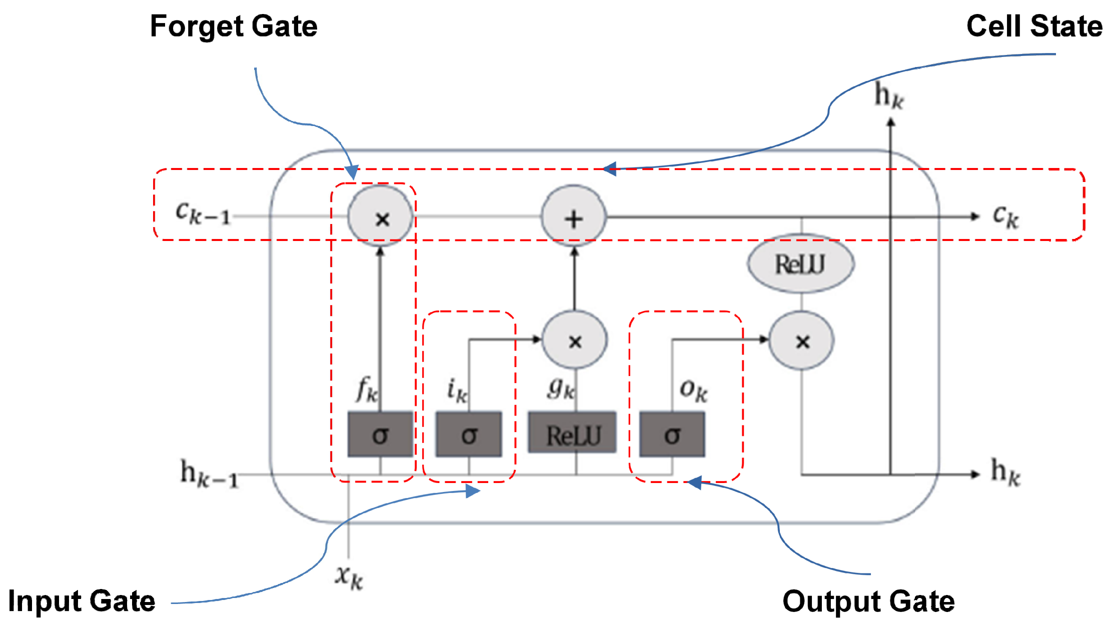


Figure 3: Cell structure of LSTM

Calculation:

$$
\begin{aligned}
f_k &= \sigma(W_x^f x_k + W_h^f h_{k-1} + b^f) \\
i_k &= \sigma(W_x^i x_k + W_h^i h_{k-1} + b^i) \\
g_k &= ReLU(W_x^g x_k + W_h^g h_{k-1} + b^g) \\
c_k &= f_k \odot c_{k-1} + i_k \odot g_k \\
o_k &= \sigma(W_x^o x_k + W_h^o h_{k-1} + b^o) \\
h_k &= o_k \odot ReLU(c_k)
\end{aligned}
$$

6. Implementation

The coding is done in python using jupyter notebook.

6.1. Data Processing

First thing was to convert MATLAB file into dataframe/json file so that it can be opened in Python.

As per the README of the dataset, the data is stored in several ".mat" files, each file corresponds to a specific battery and the data structure of each file is as follows:

Data Description:
A set of four Li-ion batteries (# 5, 6, 7 and 18) were run through 3 different operational profiles (charge, discharge and impedance) at room temperature. Charging was carried out in a constant current (CC) mode at 1.5A until the battery voltage reached 4.2V and then continued in a constant voltage (CV) mode until the charge current dropped to 20mA. Discharge was carried out at a constant current (CC) level of 2A until the battery voltage fell to 2.7V, 2.5V, 2.2V and 2.5V for batteries 5 6 7 and 18 respectively. Impedance measurement was carried out through an electrochemical impedance spectroscopy (EIS) frequency sweep from 0.1Hz to 5kHz. Repeated charge and discharge cycles result in accelerated aging of the batteries while impedance measurements provide insight into the internal battery parameters that change as aging progresses. The experiments were stopped when the batteries reached end-of-life (EOL) criteria, which was a 30% fade in rated capacity (from 2Ahr to 1.4Ahr). This dataset can be used for the prediction of both remaining charge (for a given discharge cycle) and remaining useful life (RUL).

Files:
B0005.mat Data for Battery #5
B0006.mat Data for Battery #6
B0007.mat Data for Battery #7
B0018.mat Data for Battery #18

Data Structure:
cycle: top level structure array containing the charge, discharge and impedance operations
 type: operation type, can be charge, discharge or impedance
 ambient_temperature:ambient temperature (degree C)
 time: the date and time of the start of the cycle, in MATLAB date vector format
 data: data structure containing the measurements
  for charge the fields are:
   Voltage_measured: Battery terminal voltage (Volts)
   Current_measured: Battery output current (Amps)
   Temperature_measured: Battery temperature (degree C)
   Current_charge: Current measured at charger (Amps)
   Voltage_charge: Voltage measured at charger (Volts)
   Time: Time vector for the cycle (secs)
  for discharge the fields are:
   Voltage_measured: Battery terminal voltage (Volts)
   Current_measured: Battery output current (Amps)
   Temperature_measured: Battery temperature (degree C)
   Current_charge: Current measured at load (Amps)
   Voltage_charge: Voltage measured at load (Volts)
   Time: Time vector for the cycle (secs)
   Capacity: Battery capacity (Ahr) for discharge till 2.7V
  for impedance the fields are:

Sense_current:	Current in sense branch (Amps)
Battery_current:	Current in battery branch (Amps)
Current_ratio:	Ratio of the above currents
Battery_impedance:	Battery impedance (Ohms) computed from raw data
Rectified_impedance:	Calibrated and smoothed battery impedance (Ohms)
Re:	Estimated electrolyte resistance (Ohms)
Rct:	Estimated charge transfer resistance (Ohms)

For the proposed Deep Learning model, it is only necessary to collect the data related to the discharge of the battery. To pre-process the data in .mat format, a function is created in Python and storing it in memory in two pandas DataFrame for later access. After loading the dataset, a description of the data is made using panda functions to verify if the data loading was correct.
Another python function is created to process the three different types of data – charge, discharge and impedance for different visualization of data.

6.2. Model Training and Testing

The dataset is prepared in such a way that it can be used by Tensorflow in the training phase, for this, two structures are created corresponding to the input and output expected to be obtained. For the input data, the relevant characteristics of the dataset are filtered, which are:

- Battery capacity
- Voltage
- Current
- Temperature
- Charging voltage
- Charging current
- Instant of time

For the output data, the SoH of the battery is calculated and in both input and output cases, the values are normalized to a range of values between [0-1]. Preparation of the model, 3 dense layers is used, and the parameters are used. 3 dense layers and one dropout, and one of the ADAM type is used as optimizer

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 8)	64
dense_1 (Dense)	(None, 8)	72
dense_2 (Dense)	(None, 8)	72
dropout (Dropout)	(None, 8)	0
dense_3 (Dense)	(None, 1)	9

Total params: 217
Trainable params: 217
Non-trainable params: 0

Table 2: Model Summary of SoH.

A table is created containing the actual SoH and the predicted SoH by the network and the root of the mean square error is calculated.

	cycle	actual SoH	predicted SoH
0	1	1.000000	0.961570
1	2	0.994990	0.958874
2	3	0.989185	0.955776
3	4	0.989165	0.955771
4	5	0.982898	0.952375
5	6	0.989467	0.955924
6	7	0.989075	0.955696
7	8	0.967304	0.944013
8	9	0.966997	0.943855
9	10	0.961625	0.940945

Root Mean Square Error: 0.09117926405407686

Table 3: Actual SoH vs. Predicted SoH

In the same way that was done for the estimation of SoH, the training and testing dataset is prepared, in this particular case the battery capacity data is used using the first data of the first 50 cycles to predict the capacity in the following cycles in such a way as to be able to know when the threshold of the battery is reached and estimate the remaining cycles to reach the End of Life of the battery.

The model is trained, using LSTM-type networks.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 10, 200)	161600
dropout_1 (Dropout)	(None, 10, 200)	0
lstm_1 (LSTM)	(None, 10, 200)	320800
dropout_2 (Dropout)	(None, 10, 200)	0
lstm_2 (LSTM)	(None, 10, 200)	320800
dropout_3 (Dropout)	(None, 10, 200)	0
lstm_3 (LSTM)	(None, 200)	320800
dropout_4 (Dropout)	(None, 200)	0
dense_4 (Dense)	(None, 1)	201

Total params: 1,124,201
Trainable params: 1,124,201
Non-trainable params: 0

Table 4: Model Summary of RUL

From this trained model final prediction is done on the test data. R-square is calculated for the finally predicted RUL value.

6.3. Experimental Result and Observation##

6.3.1. Aging process graph

The following graph shows the aging process of the battery as the charge cycles progress. The horizontal line represents the threshold related to what can be considered the end of the battery's life cycle.

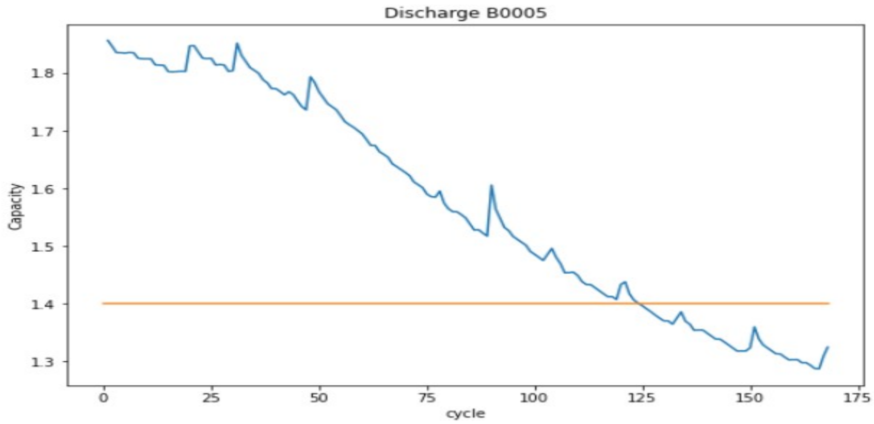


Figure 4: Battery aging process graph

6.3.2. SoH Graph

Similarly to what has been done previously, a graph of the SoH is made for each cycle, the horizontal line represents the threshold of 70% in which the battery already fulfills its life cycle and it is advisable to make the change.

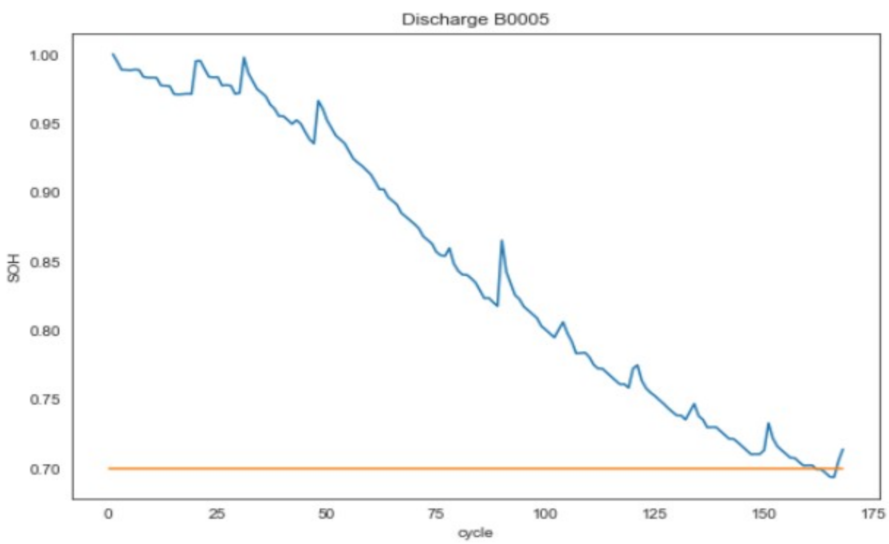


Figure 5: Battery SoH vs. Cycle graph

6.3.3. Actual SoH vs. Predicted SoH

For the estimation of SoH, it can be seen that the data pattern is learned by the model correctly, as predicted by the theory, since the shape of the curves is almost identical. The SoH shown has the same behavior as expected in theory, which is corroborated with the root mean square error value of the graph in illustration 8, whose value of 9% is very similar to that found previously. This reaffirms the precision when making the prediction.

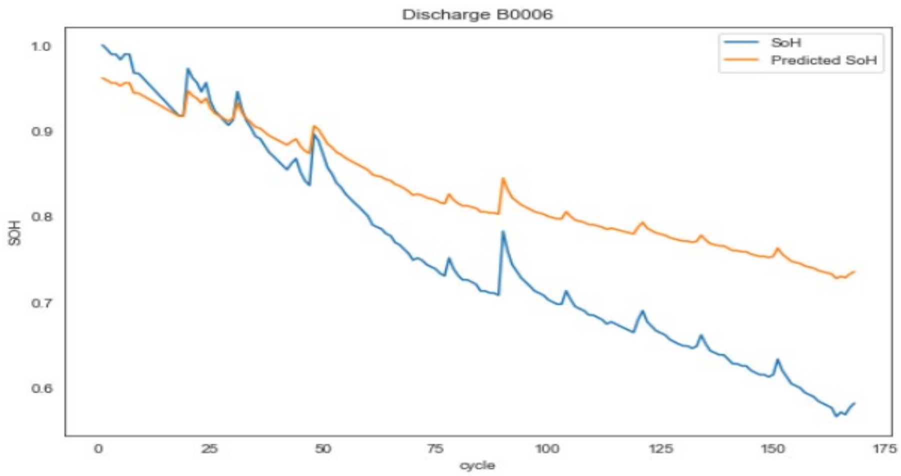


Figure 6: Battery actual SoH vs. predicted SoH graph

6.3.4. RUL Prediction

Finally, it can be seen in the graph that the capacity value and how it behaves over time is very close to the real value and supporting these data, the error in the estimation of the RUL was 13 which makes us understand that The model went behind by 13 cycles to estimate that the battery reached its end of life.

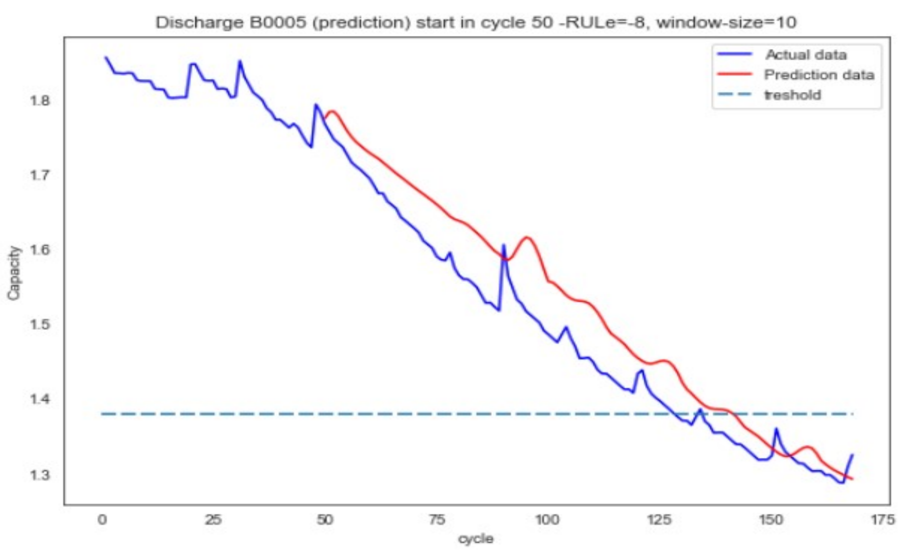


Figure 7: Battery RUL Prediction graph

.

Table 5: Evaluation of Prediction Model

Evaluation Parameter	Value
RMSE	0.052
Actual fail at cycle number	128
Prediction fail at cycle number	141
Error of RUL (Cycles)	13

As can be seen, the mean RMSE is 0.05 (5%), which is very close to the values observed in the literature using this type of network.

Finally, it can be seen in the graph that the capacity value and how it behaves over time is very close to the real value and supporting these data, the error in the estimation of the RUL was 13 which makes us understand that The model went behind by 13 cycles to estimate that the battery reached its end of life.

7. Conclusion and Future Work

The goal of this dissertation is to develop data-driven model by using a deep learning algorithm for the prognostic data. The deep learning algorithm provides a promising outcome for predicting and modeling the battery prognostic and health management applications. Based on the accuracy archived, we also believe that our organization's engineering department may plan to deploy this data-driven model in the near future, in production IoT Data platform application. The reliable data-driven model has many advantages over the manual approach to support customer's edge devices battery replacement before failure. The first major advantage is that it overcomes the complexity of the manual approach. This attribute of less complexity in a data-driven model helps to reduce the involvement of the domain experts in particular fields. In the future, the predictive model might be able to be generated and constructed without any opinion or knowledge from experts at all. The second advantage is that data-driven models can be employed in real-time situations, due to the shorter computational time needed. The last point is that the data-driven model is more cost-effective to construct and to employ in real applications. As an example, a data-driven model can be generated and monitored by using only regular personal computing devices, without the need for exclusive and excessive resources. This future trend of data-driven models is in line with the recent achievement of deep learning algorithms and artificial intelligence. These methodologies are believed to be the main approaches in the further development of data-driven models. However, the accuracy of prediction and the higher performance of using deep learning algorithms also comes with the drawback of higher computational time. With rapid advancements in technology, the computational time could be substantially reduced. The future direction of this work will focus on developing a hybrid-deep learning model that could be universally applicable to multiple types of prognostic data.

8. Resource Needed for the Project

Jupyter Notebook at Anaconda setup is used for running the Deep Learning Model. The model is built using Tensorflow, Keras.

Following are the library/packages used in this project.

- Tensorflow
- Keras
- Scipy.io
- Numpy
- Pandas

- Sklearn
- Matplotlib
- Seaborn
- Json
- Os

Following are the Hardware used for this project.

- Laptop

9. Project Plan & Deliverables

Table 6: Project Plan and Deliverables

#	Task	Expected date of completion	Names of Deliverables
1.	Understand Battery Modeling Techniques	08/12/2021	Presenting Understanding
2.	Finalize optimal Model and Bringing up Experiment	08/12/2021	Finalizing Model
3.	Survey and Understand of paper in the relevant area	10/12/2021	Presentation
4.	Develop of First DL/ML model for stated goal	05/01/2022	Initial Model
5.	Prepare Mid Sem Evaluation Report for Review	11/01/2022	Mid Sem Report Draft
6.	Submit Mid Sem Report after Review with Supervisor	14/01/2022	Mid Sem Report
7.	Optimize the DL/ML model to achieve better performance	07/02/2022	Optimized Model
8.	Prepare Final Report for Review	20/02/2022	Final Report Draft
9.	Submit Report for Final Evaluation after review with Supervisor	25/02/2022	Final Report

10. Bibliography & Reference& Acknowledgement

10.1. Reference

10.1.1. Main Research Papers and Publication

[1.1] Introduction A Review on Battery Modeling Techniques
https://www.researchgate.net/publication/354466830_A_Review_on_Battery_Modelling_Techniques

[1.2] Understanding Architecture of LSTM
https://www.analyticsvidhya.com/blog/2021/01/understanding-architecture-of-lstm

[1.3] How to Develop LSTM Models for Time Series Forecasting
https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting

[1.4] Essentials of Deep Learning : Introduction to Long Short Term Memory
https://www.analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm/

[1.5] NASA PCoE Datasets [Battery Aging data set] https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/

10.1.2. Books

[3.1] Dive into Deep Learning: https://d2l.ai/
[3.2] Neural Networks and Deep Learning: http://neuralnetworksanddeeplearning.com/
[3.3] Deep Learning: https://www.deeplearningbook.org/

10.1.3. Additional Readings

[4.1] Deep Learning [NPTEL Course]: https://onlinecourses.nptel.ac.in/noc20_cs62/preview
[4.2] Data Science: Machine Learning [EDX Course]
https://learning.edx.org/course/course-v1:HarvardX+PH125.8x+2T2021/home

10.1.4. Source code of the project

[5.1] https://github.com/kumar-ambuj/dissertation.git

11. Appendix##

List of Figures

List of Tables

Abbreviations

LSTM - Long short-term memory
CNN - Convolutional Neural Network
DNN - Deep Neural Network
RNN - Reccurrent Neural Network
LTM - Long tem memory
STM - Short term memory
AI – Artificial Intelligence
DL – Deep Learning

12. Checklist of items for the final report

Table7: Checklist of items for the final report

#	Item	Y/N
1.	Is the Cover page in the proper format?	Y
2.	Is the Title page in the proper format?	Y
3	Is the Certificate from the Supervisor in the proper format? Has it been signed?	Y
4	Is Abstract included in the Report? Is it properly written?	Y
5	Does the Table of Contents' page include chapter page numbers?	Y
6	Is Introduction included in the report? Is it properly written?	Y
7	Are the Pages numbered properly?	Y
8	Are the Figures numbered properly?	Y
9	Are the Tables numbered properly?	Y
10	Are the Captions for the Figures and Tables proper?	Y
11	Are the Appendices numbered?	Y
12	Does the Report have Conclusions/ Recommendations of the work?	Y
13	Are References/ Bibliography given in the Report?	Y
14	Have the References been cited in the Report?	Y
15	Is the citation of References/ Bibliography in the proper format?	Y