WINTECHCON 2018

September 28, 2018

**Akshata Kishore Moharir, Srishti Gautam, Seema Chopra, Naveen Kumar Megharaj**
**Boeing Company**

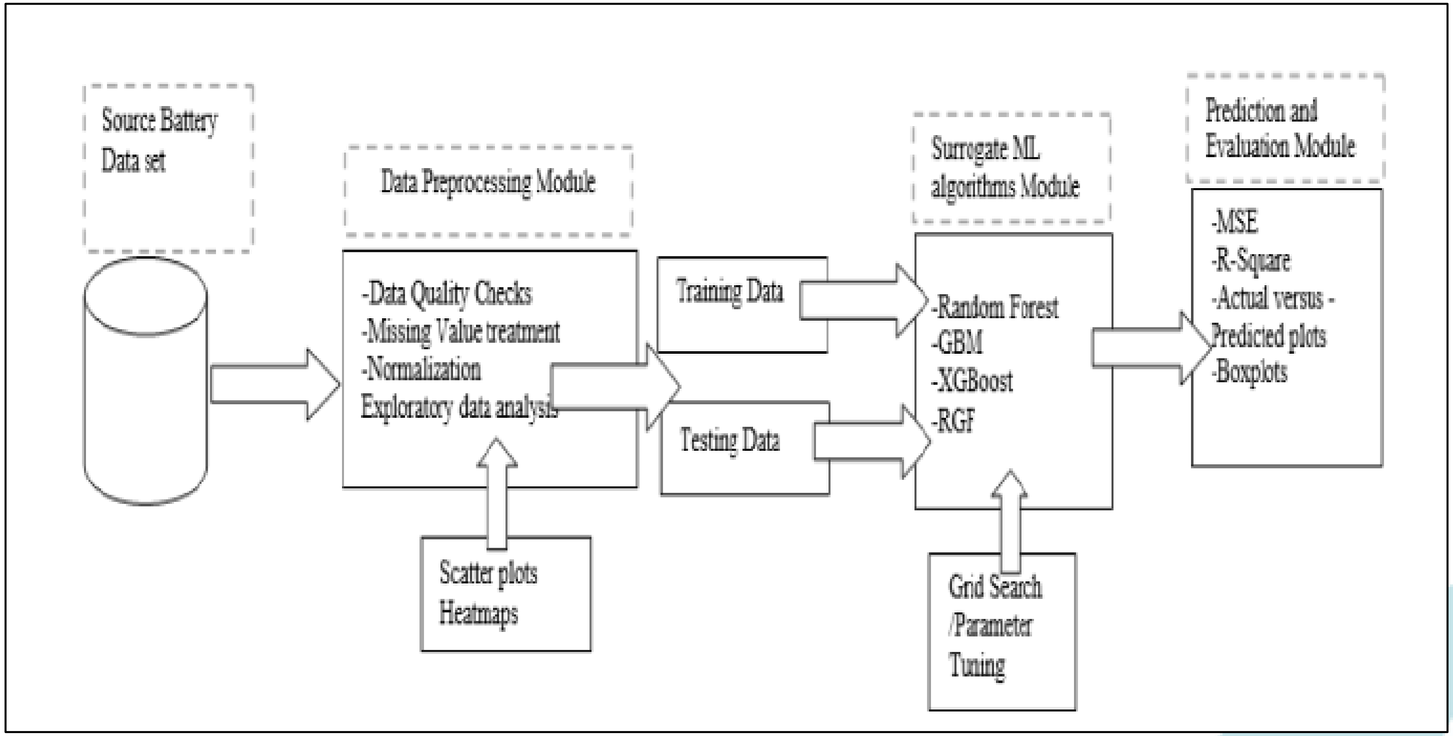


Presented at WinTechCon-2018, organized by IEEE CAS Bangalore Chapter, IEEE Bangalore Section, and IEEE WiE Council

# Problem Statement/Abstract


- *Abstract- A novel data driven approach is developed for predicting the battery degradation behavior for lithium-ion batteries using Regularized Greedy Forests (RGF). The aim of this novel approach is to improve and compare the prediction accuracy of the lithium ion (Li-ion) batteries using physics based models and surrogate models which are machine learning (ML) models. First, the approach takes in state of charge, state of health, and power of the battery as inputs and predicts the solid electrolyte interface of the Li-ion batteries which lead to the capacity fade which can be the additional cost of batteries. Second, the RGF is proposed as the new ML regressor for failure diagnostics to determine the battery degradation behavior. The prediction is based on the assumption of the availability of real-time observation and historical data. The feasibility of the regressors is validated using Li-ion battery diagnostic data. The experimental results show the following: (1) Fewer data dimensions for the input data are required compared to traditional state-of-art ML models (2) Physics based models describe the battery behavior accurately but they are computationally costly, therefore we have developed surrogate ML models for physic based models which are cost effective (3) The proposed ML regressor provides an effective way of diagnosing battery degradation behavior of the Li-ion battery compared to state-of art Surrogate ML algorithms with small error and has high prediction accuracy.*

# Solution Approach for battery degradation behavior using Surrogate ML models



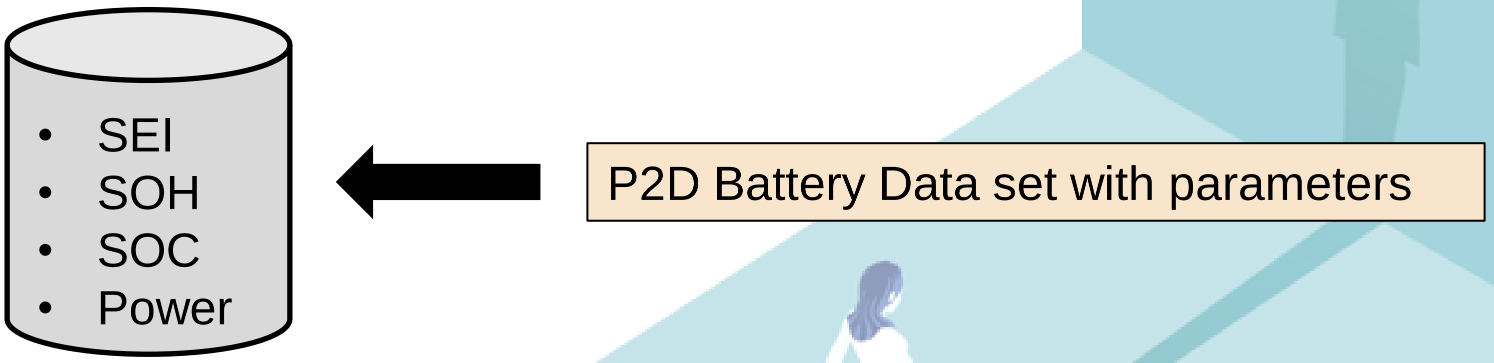
# Solution Approach:- Experimental Data Set

- Physics based model - P2D( Pseudo two dimensional) data set is considered for comparison of various Surrogate ML models.
- Battery capacity fade is directly impacted by the SEI layer growth which makes it a natural choice to track the variable that defines the end of life.
- SEI( solid electrolyte interface) layer growth is used for prediction and considered as a regression problem. Parameters used in Prediction of SEI layer which affects the capacity fade of the battery are:-

I. State of health

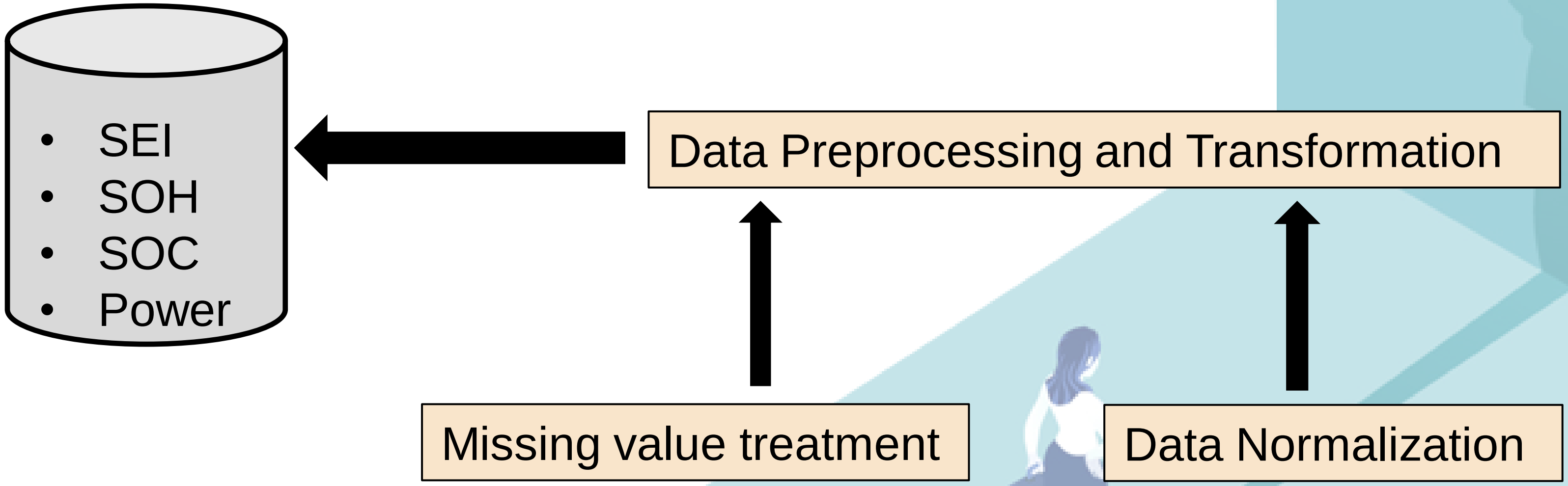
II. State of charge

III. Power



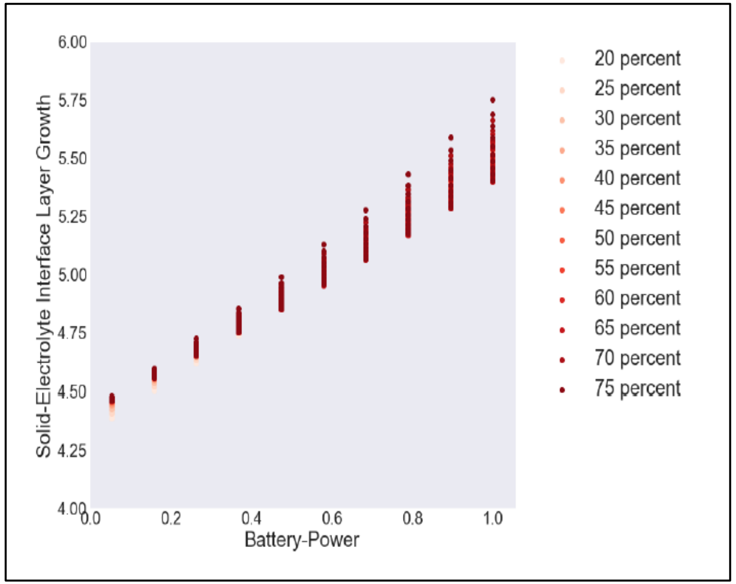
# Solution Approach:- Data Preprocessing and Transformation

- Data quality checks are performed to identify the missing values/ ‘NA’s and treated appropriately.
- Data Normalization is performed to scale the parameters to have same unit of measurement



# Solution Approach:- Exploratory Data Analysis

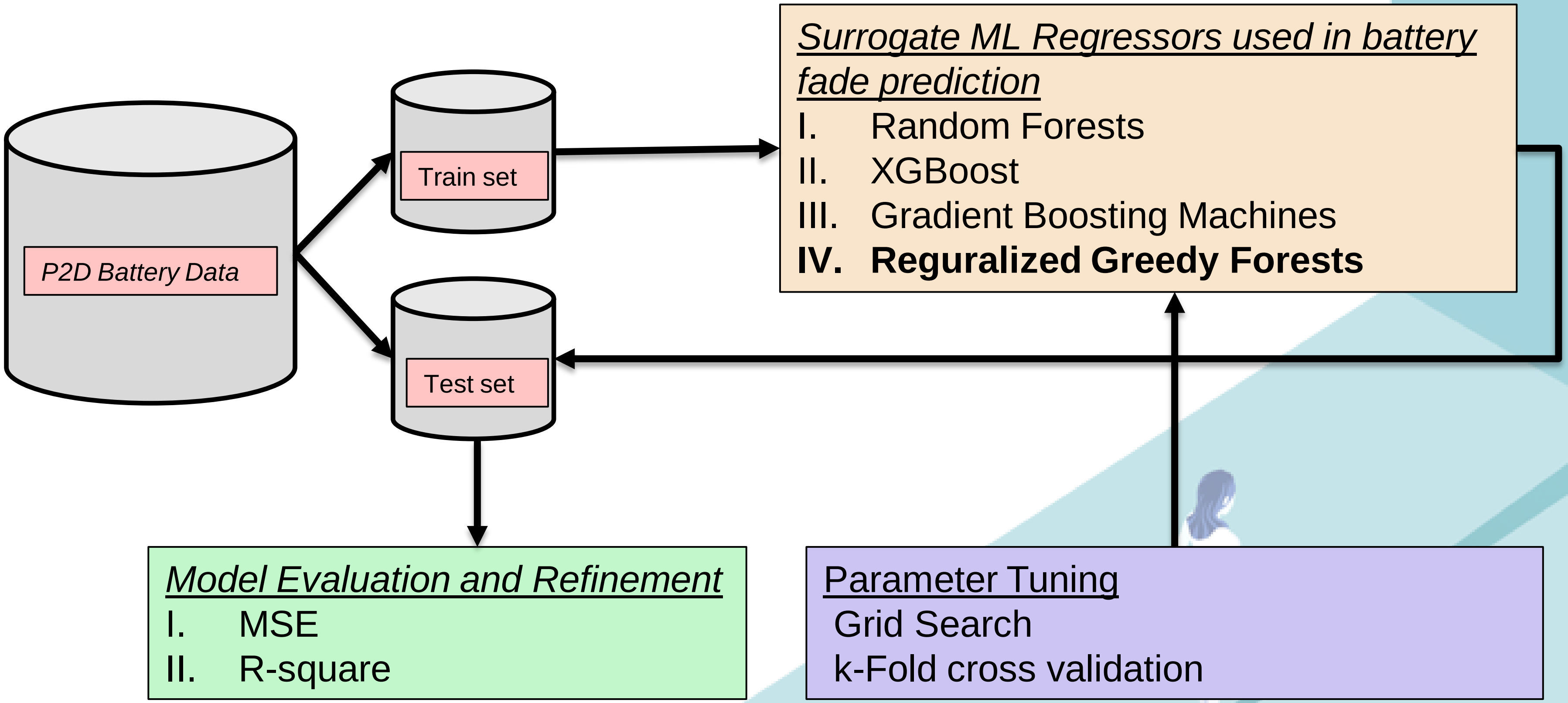
- The scatterplots show the effect of the SEI layer growth on the state of health of the battery.




*Power value are plotted against the SEI layer growth showing degradation of state of health of battery where the colors indicate the correlation with state of charge.*

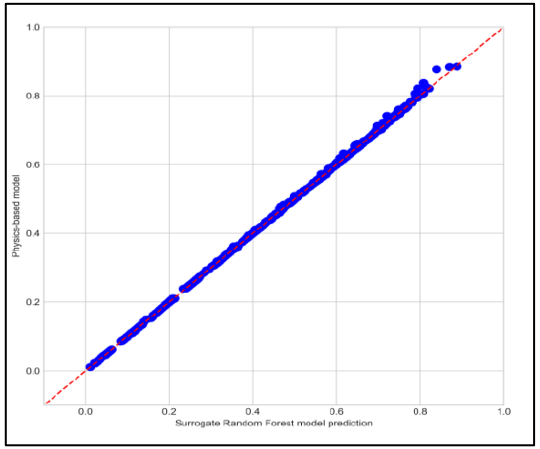
# Solution Approach:-Model Development and Training

- The data set is split in to train and test and various  machine learning algorithms are applied .

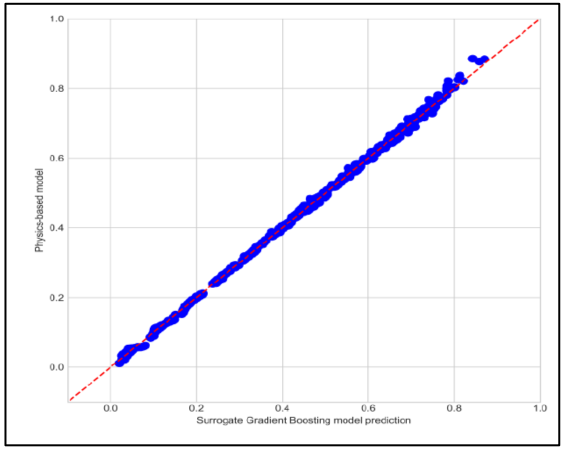


# Experimental Results

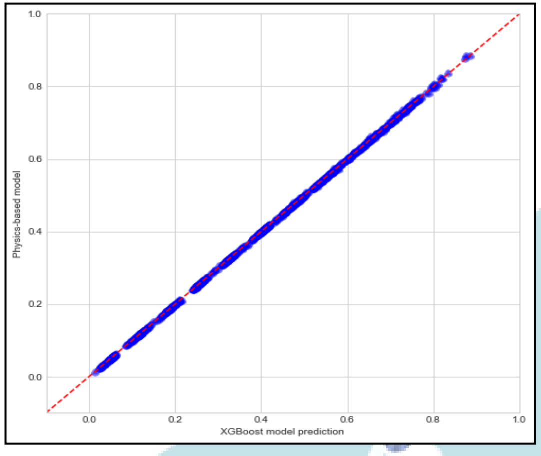
- Results of comparing the surrogate models for predicting the battery fade were noted down for different linear and nonlinear regressors with respect to their MSE and R-square values on the test data set.
- Regularized greedy forest was a clear winner compared to rest of the state-of-art regressors.



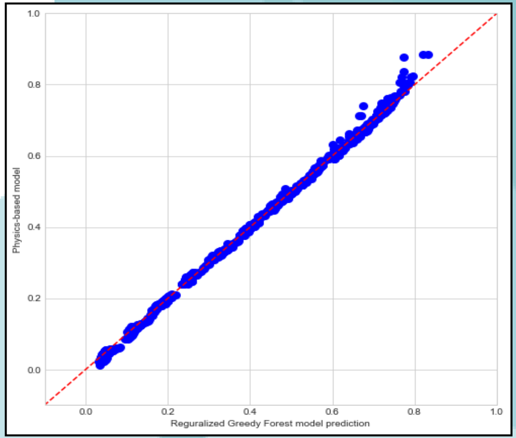

Actual versus predicted plot of SEI using the surrogate Random Forest model




Actual versus predicted plot of SEI using the surrogate Gradient Boosting model




Actual versus predicted plot of SEI using the surrogate XGBoost model.

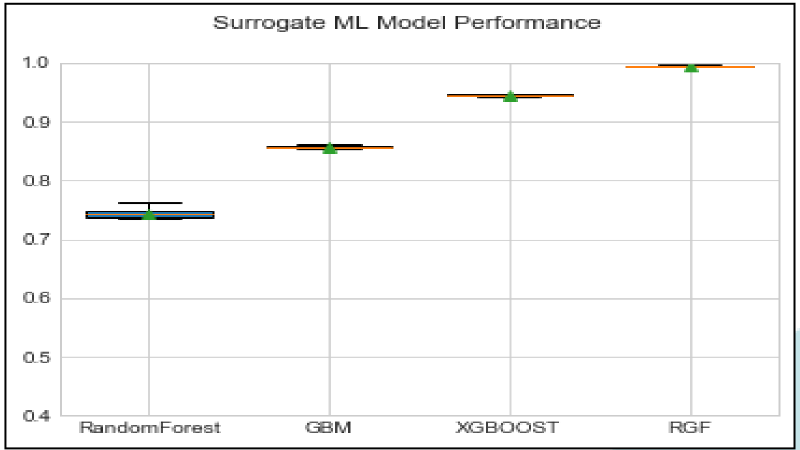



Actual versus predicted plot of SEI using the surrogate RGF

# Experimental Results

*MSE and R-Square of Regressors*

| *Surrogate ML Methods* | *MSE* | *R-square* |
|---|---|---|
| Random Forest | 0.009 | 77.4 |
| GBT | 0.003 | 85.7 |
| XGBoost | 0.002 | 94.5 |
| Regularized Greedy Forests | 0.001 | 98.9 |




*Boxplots showing the comparison of the Surrogate ML models plotted against their R-square values.*

# Conclusion

- The most important conclusion of this work is application of ensemble based technique Regularized Greedy forest which is still in research and compare its accuracy and performance with state-of-art algorithms for P2D Li-ion battery data set.
- Second important conclusion is RGF outperformed in terms of predicting the battery degradation behavior with least minimum error as compared to rest of the regression algorithms. Current regression accuracy is great compared to other machine learning algorithms discussed in the paper, but we can further increase it as we get access to more data and as we replace generic algorithms with more specialized ones like regularized greedy forests which are still under research.

# References


- http://web.mit.edu/braatzgroup/Modeling_and_simulation_of_lithium_ion_batteries_from_a_systems_engineering_perspective.pdf
- https://depts.washington.edu/maple/pubs/66_DataScience_Neal_2018.pdf
- https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1164&context=eng_etds
- https://depts.washington.edu/maple/pubs/53_JES_SEI_layer_growth_and_capacityfade.pdf
- https://www.sciencedirect.com/science/article/pii/S0378775318302350
- https://arxiv.org/ftp/arxiv/papers/1210/1210.3672.pdf