Mid Term Presentation of BTP project

Real Time Battery Monitoring System Using Machine Learning

Presented by

Tathya Bhatt – 191030011041

Gurpreet Singh - 191030012005

Under the guidance of

Dr. Jagat Rath

Department of Mechanical and Aerospace Engineering

Institute of Infrastructure, Technology, Research and Management

Table of contents

Recap

Problem Statement

- Currently, battery monitoring system rely mostly on measurement of current and voltage and ultimately SOC through hardware systems.
- These methods neglect to consider temperature of surroundings as a parameter, which are pretty critical in understanding the true runtime of the battery
- Current algorithms (such as Coulomb Counting, Voltage Method) suffer from Hysterisis — a condition where there is a delay in the output of a system —causing a certain error in measurements.

Objective

- One of the best solution to overcome the challenges aforementioned is to use data driven approach in order to be efficient and accurate.
- To study different machine learning approach and to find the best approach.
- To bring into account the effect of battery temperature in order to get the optimum result.

Literature Review

Sr.No	Paper Title	Authors	Summary
1.	Battery state-of-health modelling by multiple linear regression *Journal of Cleaner Production*, *290*, 125700.	Søren B. Vilsen Daniel-Ioan Stroe	• There were many features which were reduced by principle component analysis • The lasso regression performs better which was evaluated by MAE evaluation metric
2.	Lithium-ion Battery Life Cycle Prediction with Deep Learning Regression Mode*l* *EEE Applied Power Electronics Conference and Exposition (APEC)*, 3346–3351	Huawei Yang Yuan Cao Huaiqi Xie Shuai Shao J ie Zhao Tianyi Gao J iajun Zhang Binghua Zhang	• The five-key-feature-basedbattery model is evaluated and validated all through the paper and is proved to have high prediction accuracy ofbattery life cycles. • The proposed regression model can be utilized for real-time life cycle prediction without the need of historical data
3.	*An Introduction to Variable and Feature Selection* - Journal of Machine Learning Research 3 (2003) 1157-1182	Isabeller Guyon Andre Elliseff	• Variable selection is important which can make the model cost effective.

Progress

Mid Sem 7th Semester

- Understand the working of a monitoring system and identify the problem statement
- Choose necessary sensors and hardware required for data acquisition

End Sem 7th Semester

- Understand machine learning algorithm and its working
- Understood regression and compared with other types of regression models

Mid Sem 8th Semester

- Increasing the amount of the dataset for better model performance.
- Using Machine Learning and Deep Learning to implement complex algorithms

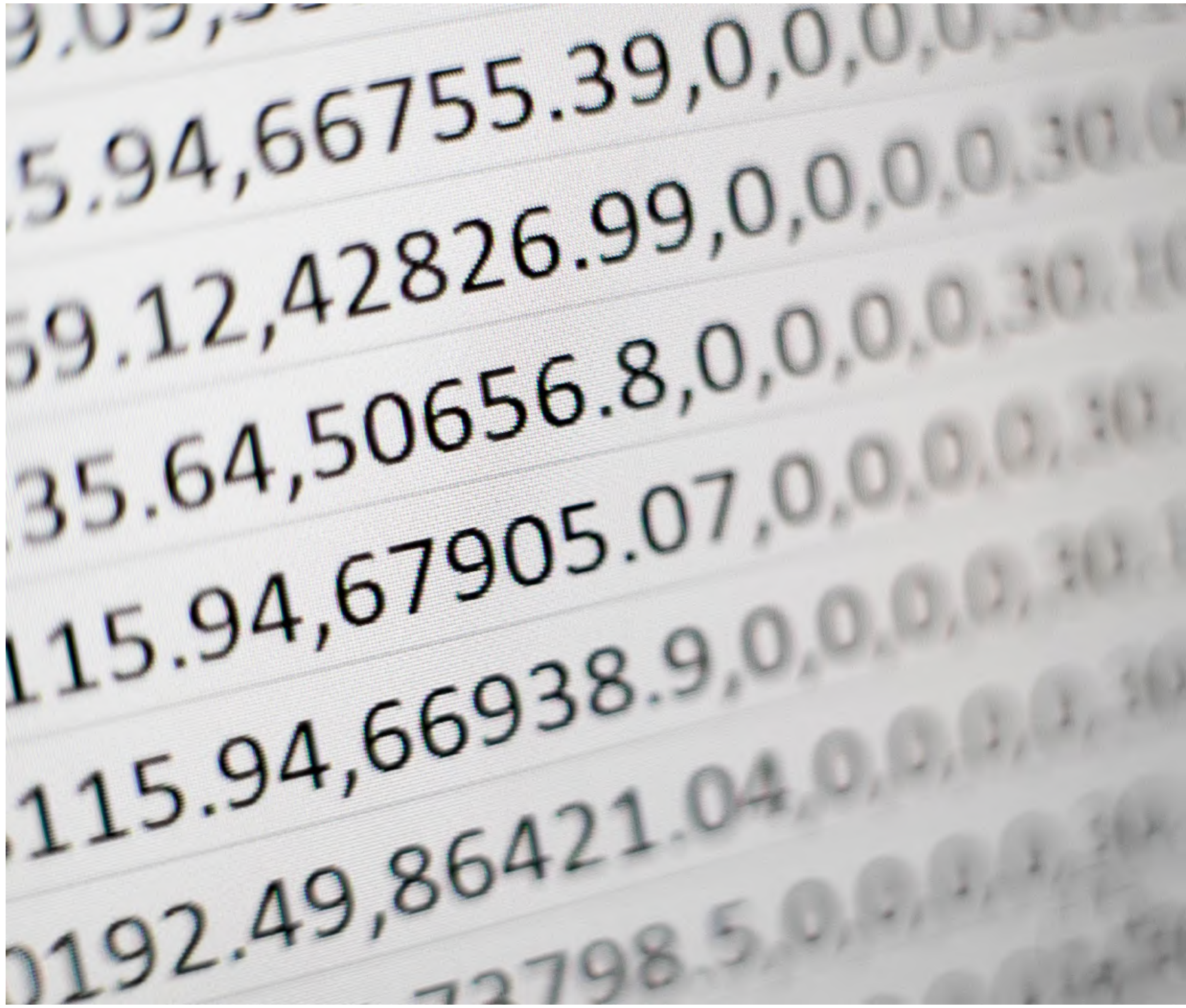
End Sem 8th Semester

- Integrate hardware and Machine Learning to get the desired result of a battery driven device

Data Selection and preprocessing

Dataset details

- Required continuous voltage, current, temperature and battery capacity measured every second accurately.
- LG 18650HG2 Li-ion Battery Data is used to train the model
- HPPC Tests (Hybrid Pulse Power Characterization) performed at McMaster University in Hamilton, Ontario, Canada by Dr. Phillip Kollmeyer.
- A brand new 3Ah LG HG2 cell was tested.
- A series of tests were performed at different temperatures (10℃, 25℃, 40℃) and the battery was charged after each test at 1C rate to 4.2V
- The discharge tests were performed with different C ratings in the data provided



Data Pre-processing

- Cleaned the data and removed unneccesary columns
- Calculated more features to the dataset like DC Resistance, Mean of variables and Standard Deviation of the variables
- Further the null values and zeros were removed in order to avoid error in the algorithm
- Combined data at various temperatures like 10, 25, 40 Degrees
- Combined the data at different charging rates to make the model more versatile.

Voltage	Current	Temperature	Capacity	Standard Deviation	DC Resistance	MeanofVariables	StateofCharge
4.17604	-0.15069	23.97615	2.99746	11.49291418	-0.039153228	7.00101	99.91533333
4.17014	-0.15069	23.97615	2.99239	11.49236934	-0.016789435	7.0008025	99.74633333
4.16761	-0.15069	23.76583	2.98986	11.38852164	-0.016443719	6.9482225	99.662
4.16509	-0.15325	23.66067	2.98732	11.3369821	-0.015399674	6.9212975	99.57733333
4.16273	-0.15325	23.76583	2.98478	11.38841991	-0.014599509	6.9476325	99.49266667
4.16053	-0.15069	23.76583	2.98225	11.38755401	-0.014533148	6.948355	99.40833333
4.15834	-0.15069	23.97615	2.97971	11.49077043	-0.013405004	7.0010225	99.32366667
4.15632	-0.15069	23.87099	2.97719	11.43864559	-0.013405004	6.9748575	99.23966667
4.1543	-0.15069	23.76583	2.97465	11.38652195	-0.012137031	6.9486975	99.155

Why it is needed

- To train an optimal model, we must utilise only the most important features.
- If there are too many features, the model may learn from noise and capture unimportant patterns.
- Data set should be formatted in such a way that more than one Machine Learning and Deep Learning algorithm are executed in one data set, and best out of them is chosen.

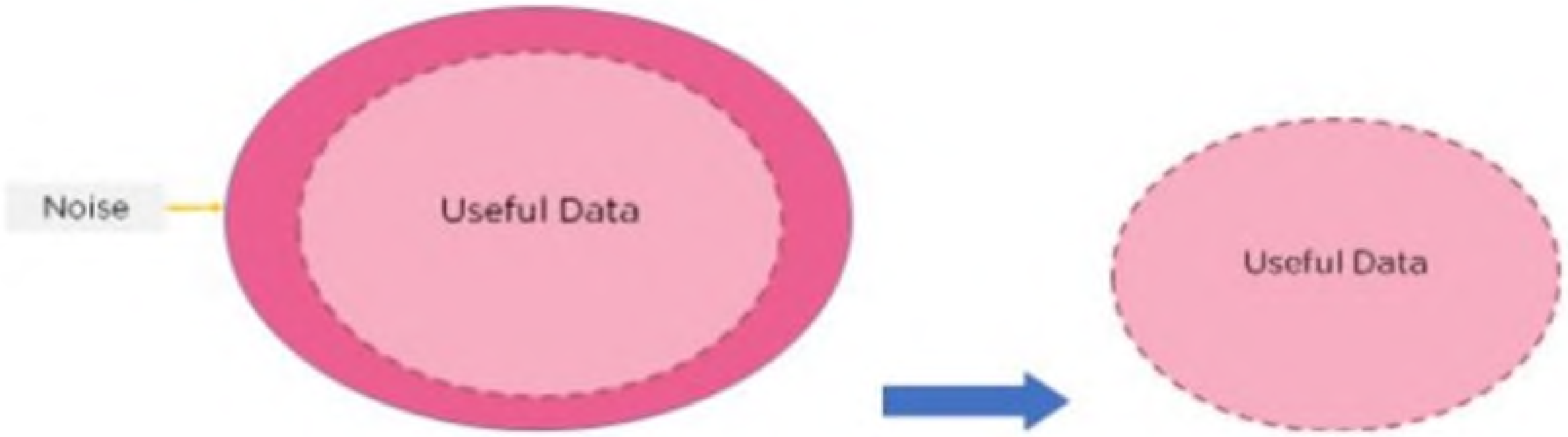


Fig. 2

Data Selection and preprocessing

Data Splitting

- Inadequate training and testing sets can have unpredictable effects on the model's output
- Ideally, the data should be divided into three sets: train, test, and validation/dev.

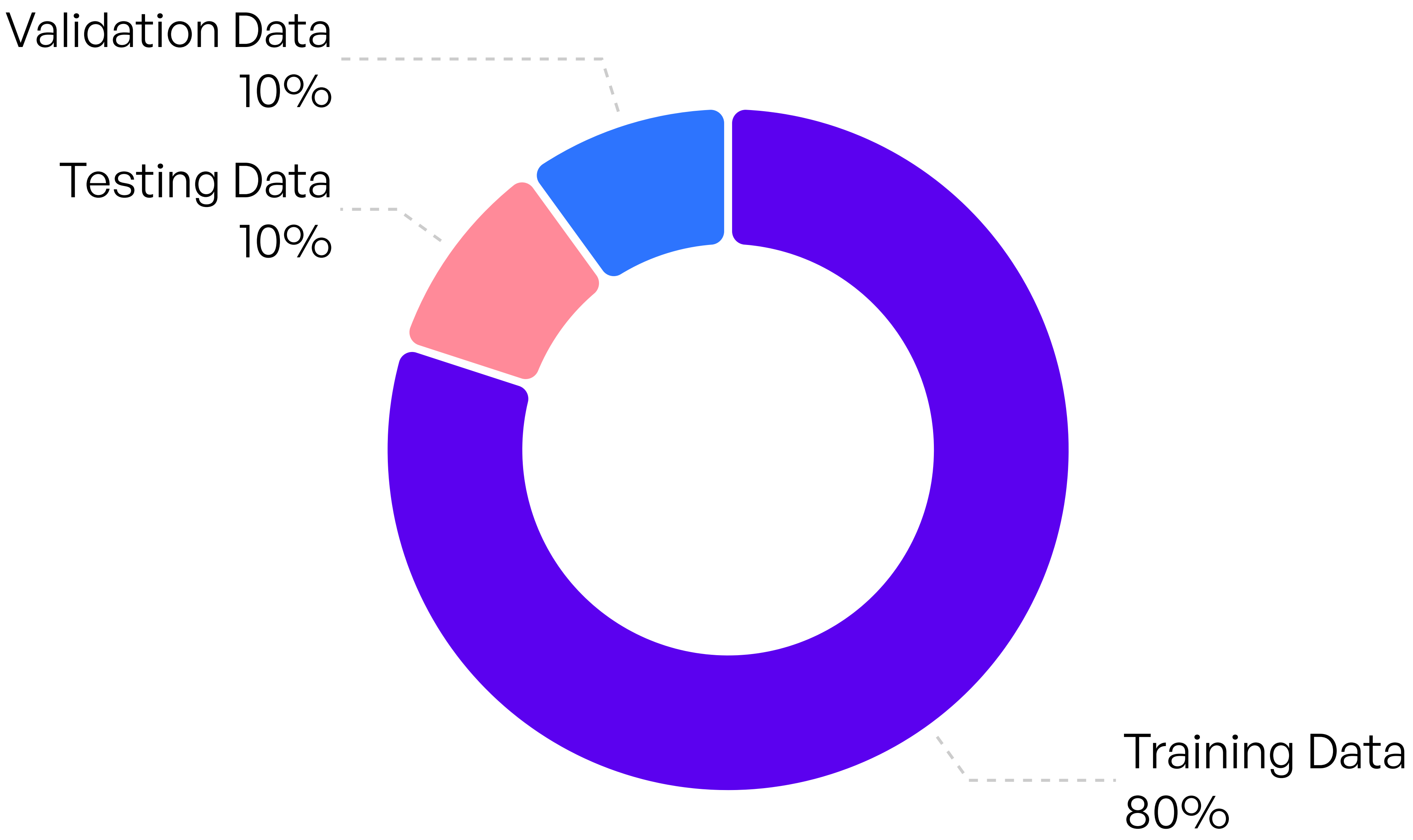


Fig. 3

Data Splitting

- Inadequate training and testing sets can have unpredictable effects on the model's output
- Ideally, the data should be divided into three sets: train, test, and validation/dev.
- **Training Set** - The data which will be fed into the model will be contained within the train set.

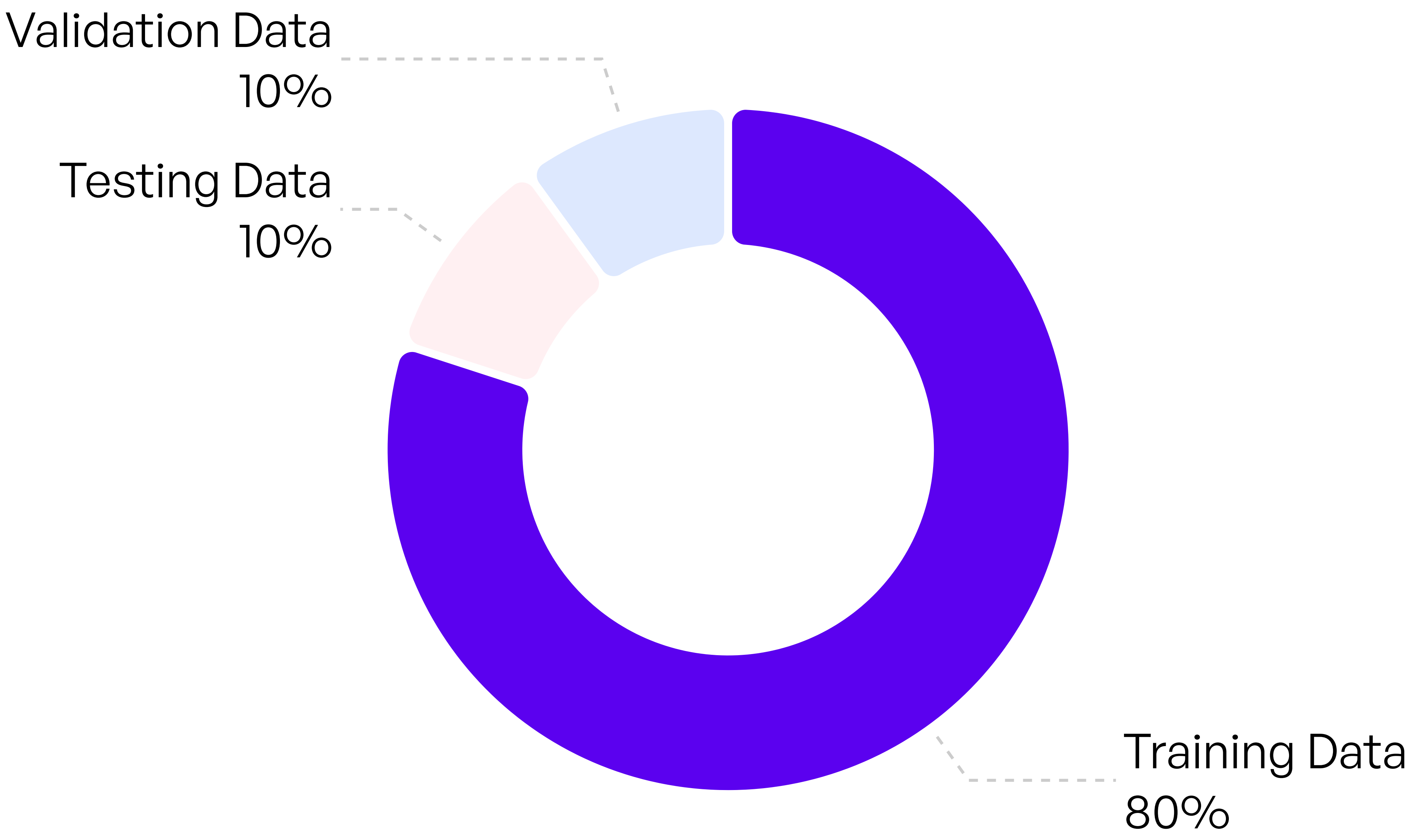


Fig. 3

Data Splitting

- Inadequate training and testing sets can have unpredictable effects on the model's output

- Ideally, the data should be divided into three sets: train, test, and validation/dev.

- **Training Set** - The data which will be fed into the model will be contained within the train set. The model would gain knowledge from this data.

- **Validation Set** - Validate the trained model with the development set. This is the most crucial setting because it will serve as the foundation for our model evaluation.

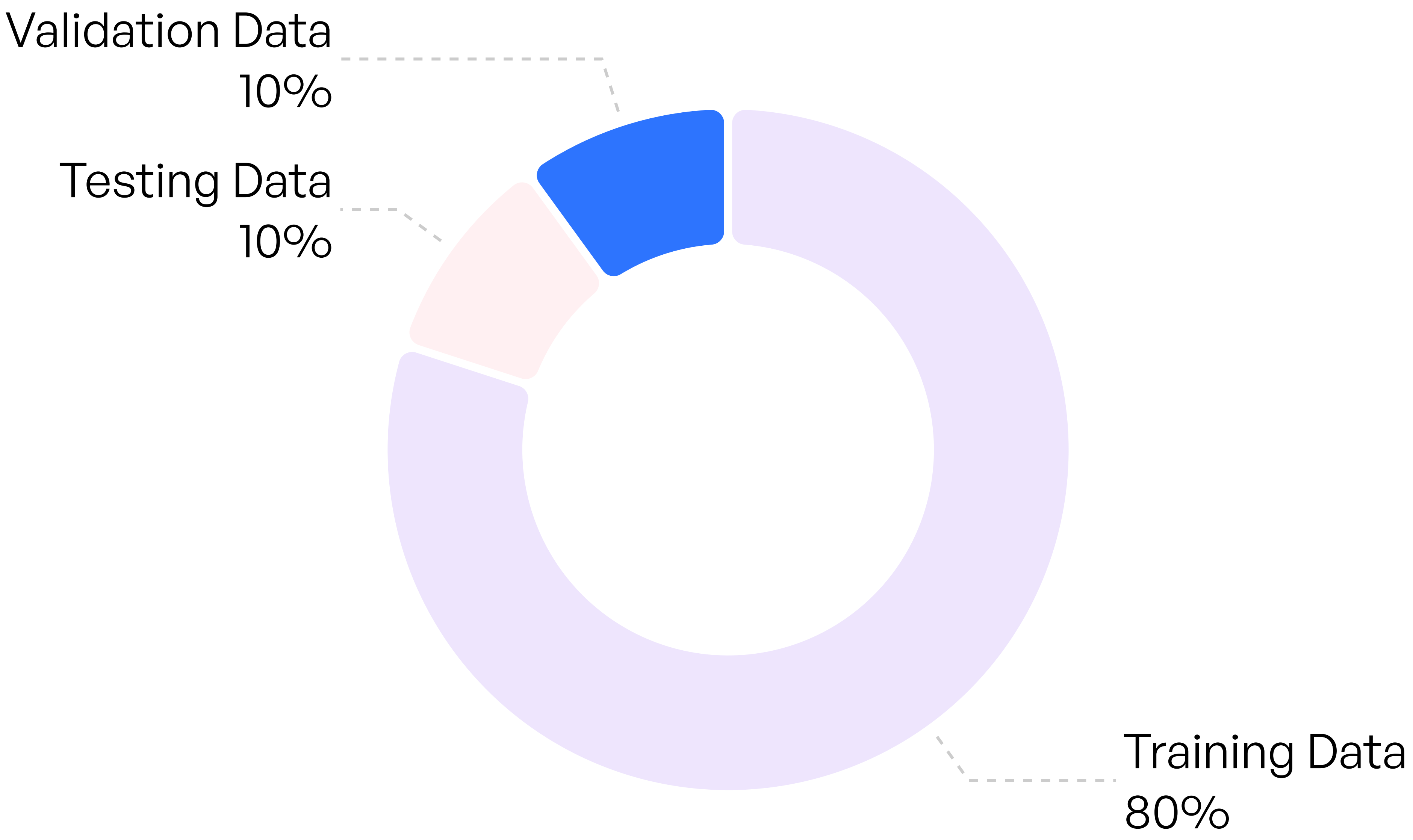


Fig. 3

Data Selection and preprocessing

Data Splitting

- Inadequate training and testing sets can have unpredictable effects on the model's output

- Ideally, the data should be divided into three sets: train, test, and validation/dev.

- **Training Set** - The data which will be fed into the model will be contained within the train set. The model would gain knowledge from this data.

- **Validation Set** - Validate the trained model with the development set. This is the most crucial setting because it will serve as the foundation for our model evaluation.

- **Testing Set** - The test set consists of the data used to evaluate the trained and validated model. It tells us how efficient our overall model

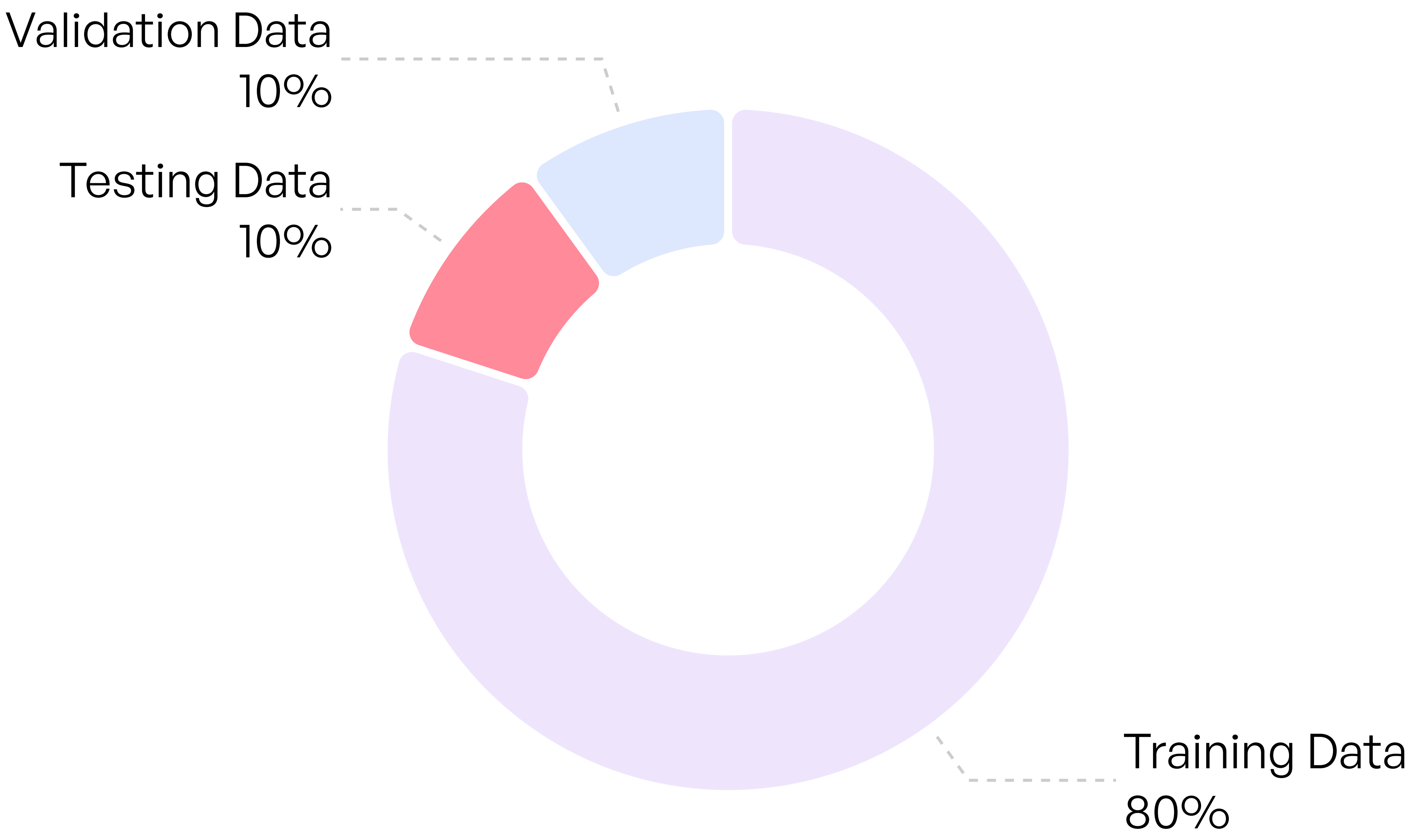


Fig. 3

Introduction to machine learning

Overview

Supervised Learning

- Training a model on a labelled dataset in which the correct outputs are provided for each example in the training data constitutes supervised learning
- The model uses this information to determine the relationship between input and output, and can then make predictions based on new, unobserved data.
- Regression and classification assignments are examples of supervised learning.

Un-Supervised Learning

- Unsupervised learning entails training a model with an unlabeled dataset in which the correct outputs are not provided.
- The model must self-learn to recognise patterns and structure in the data.
- Unsupervised learning includes tasks like clustering

Machine Learning Concept

Supervised Learning

Regression model

- Regression is a process of finding the correlations between dependent and independent variables. It helps in predicting the continuous variables such as prediction of Market Trends, prediction of House prices, etc.

- Used to predict the continuous values such as price, salary, age, etc.
- In Regression, we try to find the best fit line, which can predict the output more accurately.

Classification model

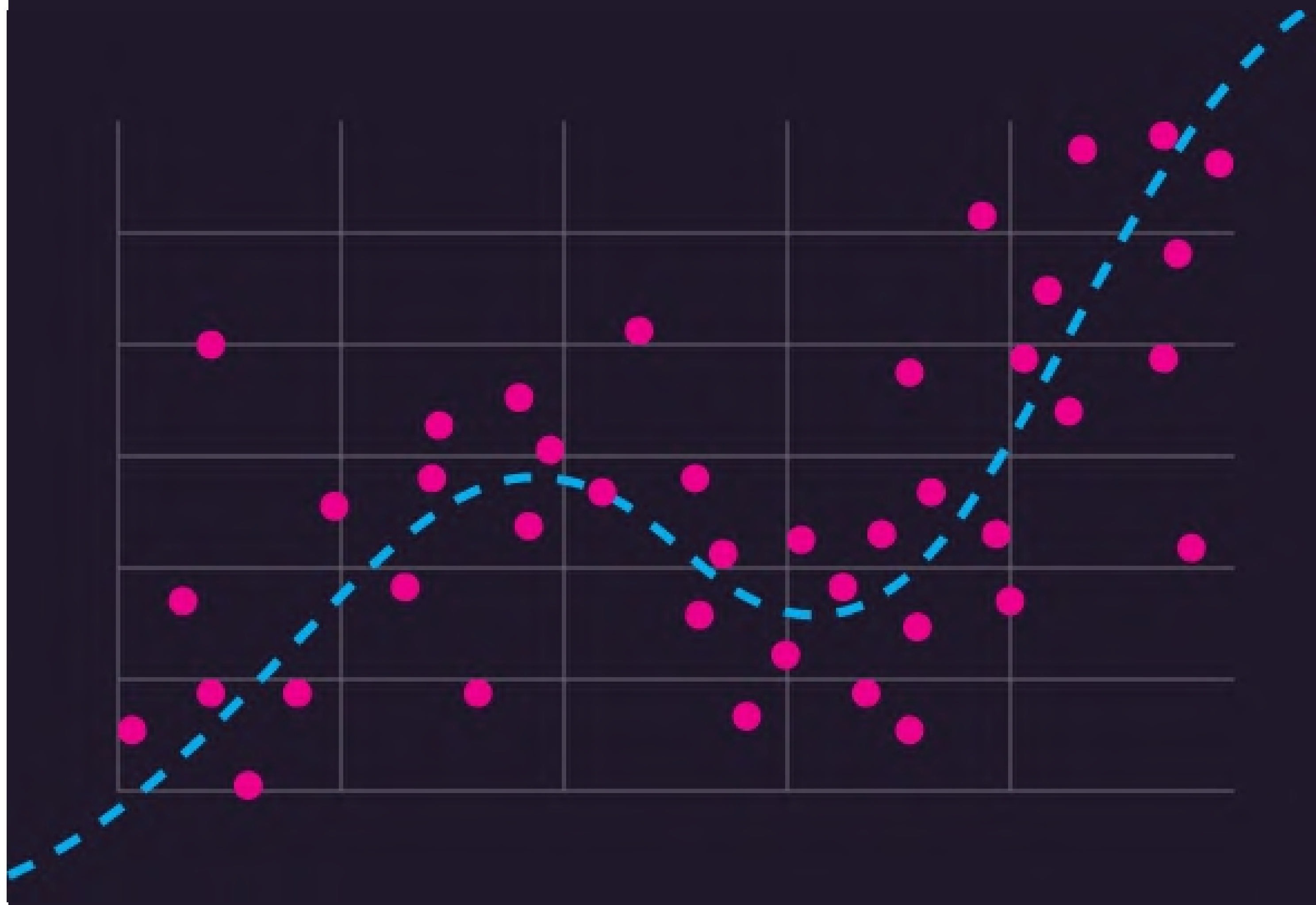
- Classification is a process of finding a function which helps in dividing the dataset into classes based on different parameters

- Used to classify the discrete values such as Male or Female, True or False, Spam or Not Spam, etc.
- In Classification, we try to find the decision boundary, which can divide the dataset into different classes.

Machine Learning Concept

Model Selection

- The goal is to train a model to make predictions or decisions based on labeled training data
- The voltage, current and temperature are continuous variable, such as a price or a probability.
- The output is the State of Charge included in the data
- The model uses this information to learn the relationship between the input and the output, and can then make predictions on new, unseen data.



Algorithm and analysis

Ordinal Least Square Regression

- Popular method for estimating the coefficients of linear regression equations that describe the relationship between one or more independent quantitative variables and a dependent variable (simple or multiple linear regression).

- Least squares represent the least squares error (SSE).

here

$\bar{Y}$ = mean of the dependent variable known as 'label'

$\overline{X_1}$ & $\overline{X_2}$ are the mean values of independent variables known as 'features'

β_0 is the intercept, β_1 & β_2 are the coefficients

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots . + \beta_k X_k + \varepsilon \tag{1}$$

For two variables, the slope can be calculated by:

$$\beta_0 = \bar{Y} - \beta_1 \overline{X_1} - \beta_2 \overline{X_2} \tag{2}$$

β_1 & β_2 are the coefficients of the variables

$$\beta_1 = \frac{(\sum x_2^2)(\sum x_1 y) - (\sum x_1 x_2)(\sum x_2 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \tag{3}$$

$$\beta_2 = \frac{(\sum x_1^2)(\sum x_2 y) - (\sum x_1 x_2)(\sum x_1 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \tag{4}$$

$$\sum x_1^2 = \sum X_1 X_1 - \frac{(\sum X_1)(\sum X_1)}{N} \tag{5}$$

$$\sum x_2^2 = \sum X_2 X_2 - \frac{(\sum X_2)(\sum X_2)}{N} \tag{6}$$

$$\sum x_1 y = \sum X_1 \mathrm{Y} - \frac{(\sum X_1)(\sum \mathrm{Y})}{N} \tag{7}$$

$$\sum x_2 y = \sum X_2 \mathrm{Y} - \frac{(\sum X_2)(\sum \mathrm{Y})}{N} \tag{8}$$

$$\sum x_1 x_2 = \sum X_1 X_2 - \frac{(\sum X_1)(\sum X_2)}{N} \tag{9}$$

Line of best fit

- The main purpose of the best fit line is that our predicted values should be closer to our actual or the observed values.

- We tend to minimize the difference between the values predicted by the model and the observed values, and which is actually termed as error.

- Variance is measure of how sensitive the model is to small changes in the training data.
- A model with high variance will have large predicted values for different training sets, while a model with low variance will have more consistent predicted values across different training sets.

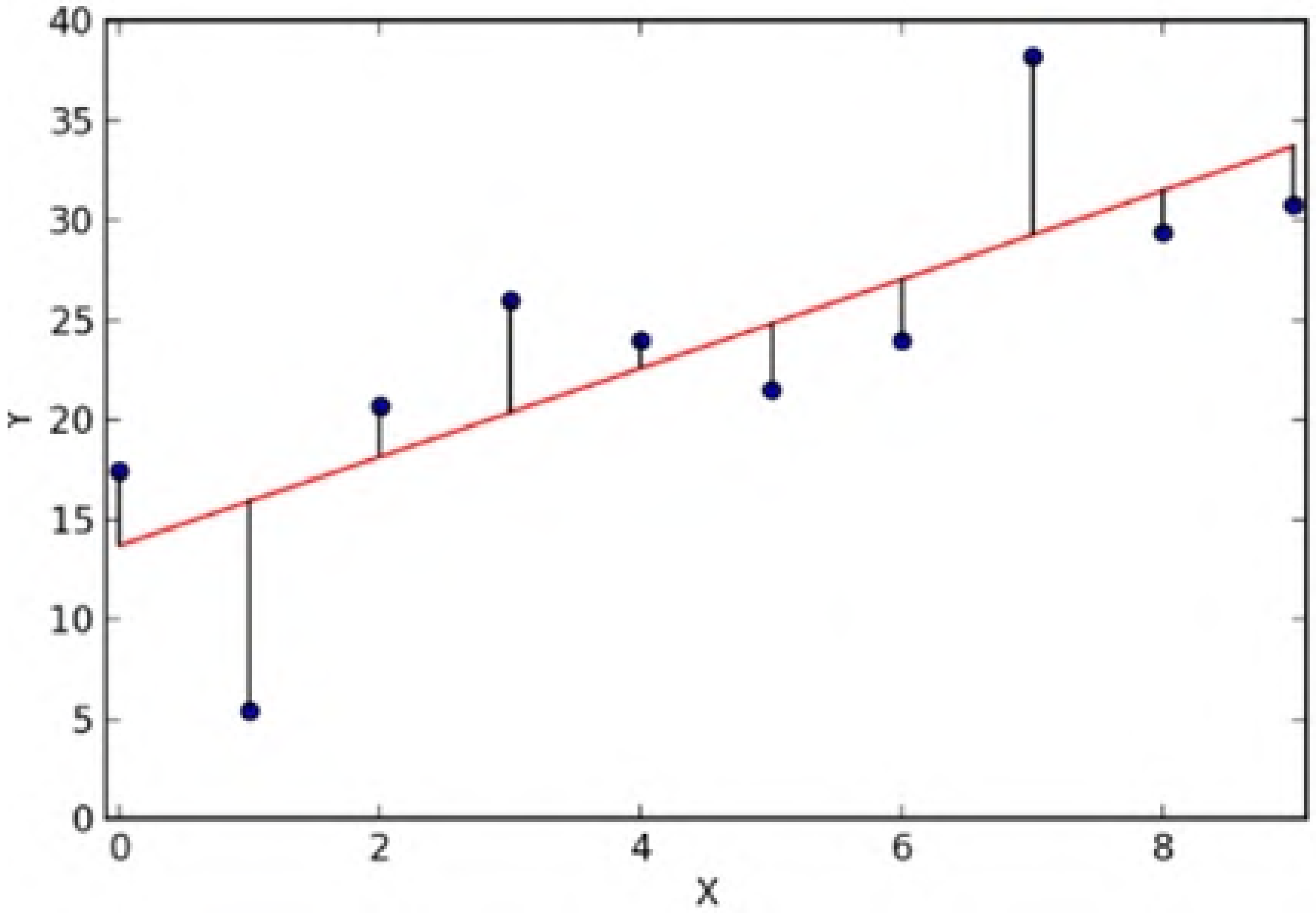


Fig. 4

https://tutorme.com/blog/post/how-to-find-line-of-best-fit/

Cost Function & Gradient Descent

- We need to check how well the model is performing.

- Hence, the function which corrects the model is required in order to get accuracy

- Algorithm that is used to optimize the cost function or the error of the model. It is used to find the minimum value of error possible in your model.

- Gradient Descent can be thought of as the direction one have to take to reach the least possible error

$$\text{Cost Function} = \frac{1}{n}\sum_0^n\left(Y_{pred} - Y_{actual}\right)^2 \tag{10}$$

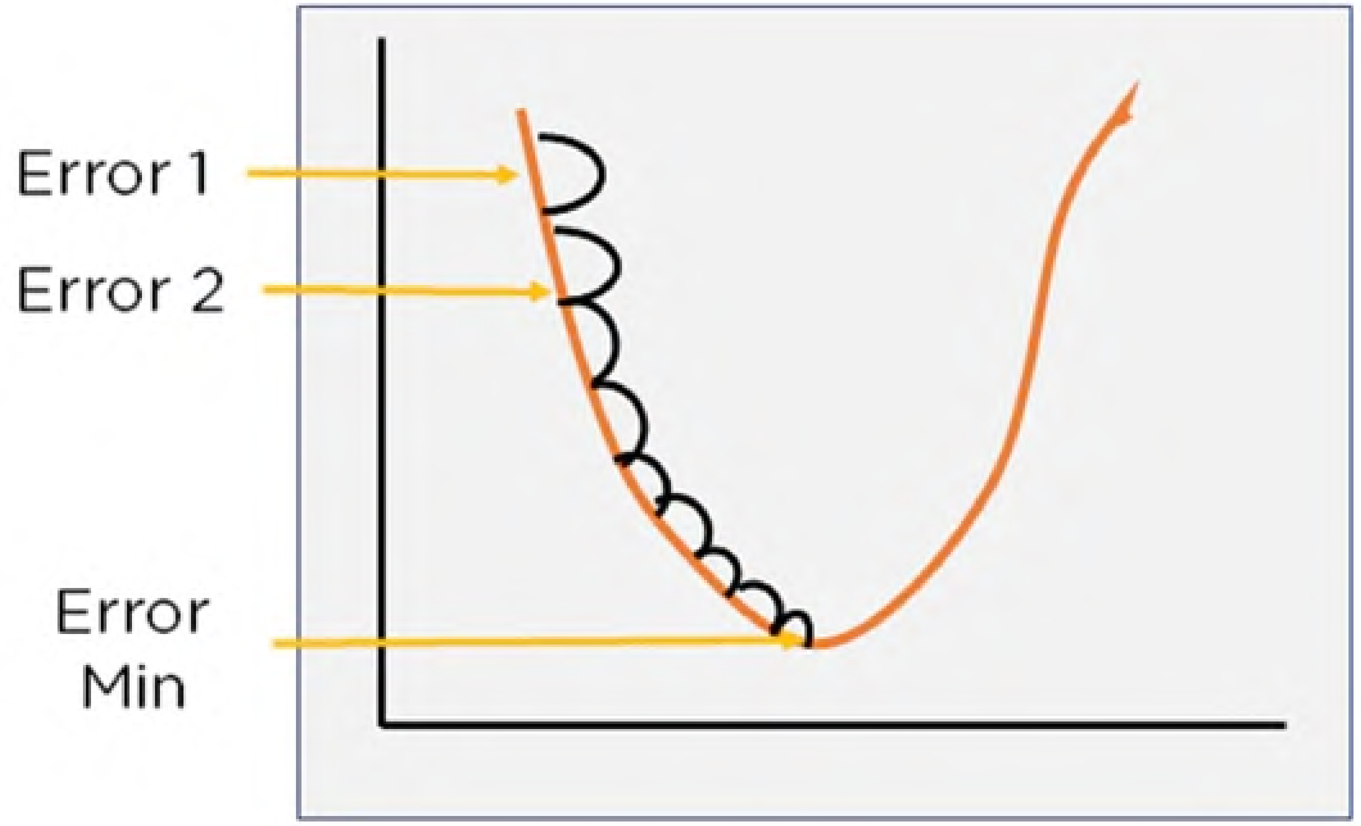


Fig. 5 - Gradient Descent Illustration

Algorithm and analysis

Ridge Regression

- A generalized model should have a low variance in order to be a good model.

- By introducing penalty term it helps to find the solution with smaller coefficient values
- Reduces overfitting and improves generalization

- Used when there is lot of noise in data and more correlated features.

$$\text{Cost Function} = \frac{1}{n}\sum_0^n\left(Y_{pred} - Y_{actual}\right)^2 + \lambda\left(\sum_0^n m_i^2\right)$$

Lasso Regression

- Lasso is similar to ridge regression instead it has a modulus which tries to make the slope zero
- Therefore, lasso selects the only some feature while reduces the coefficients of others to zero.
- It is generally used when we have more number of features, because it automatically does feature selection.
- Used in high dimensional data
- Reduces the irrelevant predictors

$$\text{Cost Function} = \frac{1}{n}\sum_0^n \left(Y_{pred} - Y_{actual}\right)^2 + \lambda\left(\left|\sum_0^n m_i\right|\right)$$

Results

Feature Correlation

- Heatmaps are used to identify the correlation between two or more variables.
- Used to understand the relationship between different features in a dataset and how they may affect the performance of a model.
- If two or more features in a dataset are highly correlated, it can cause problems with the model's ability to learn and make accurate predictions.
- If a model is trained on a dataset with correlated features, it may be less able to generalize to other datasets with different correlations between the features

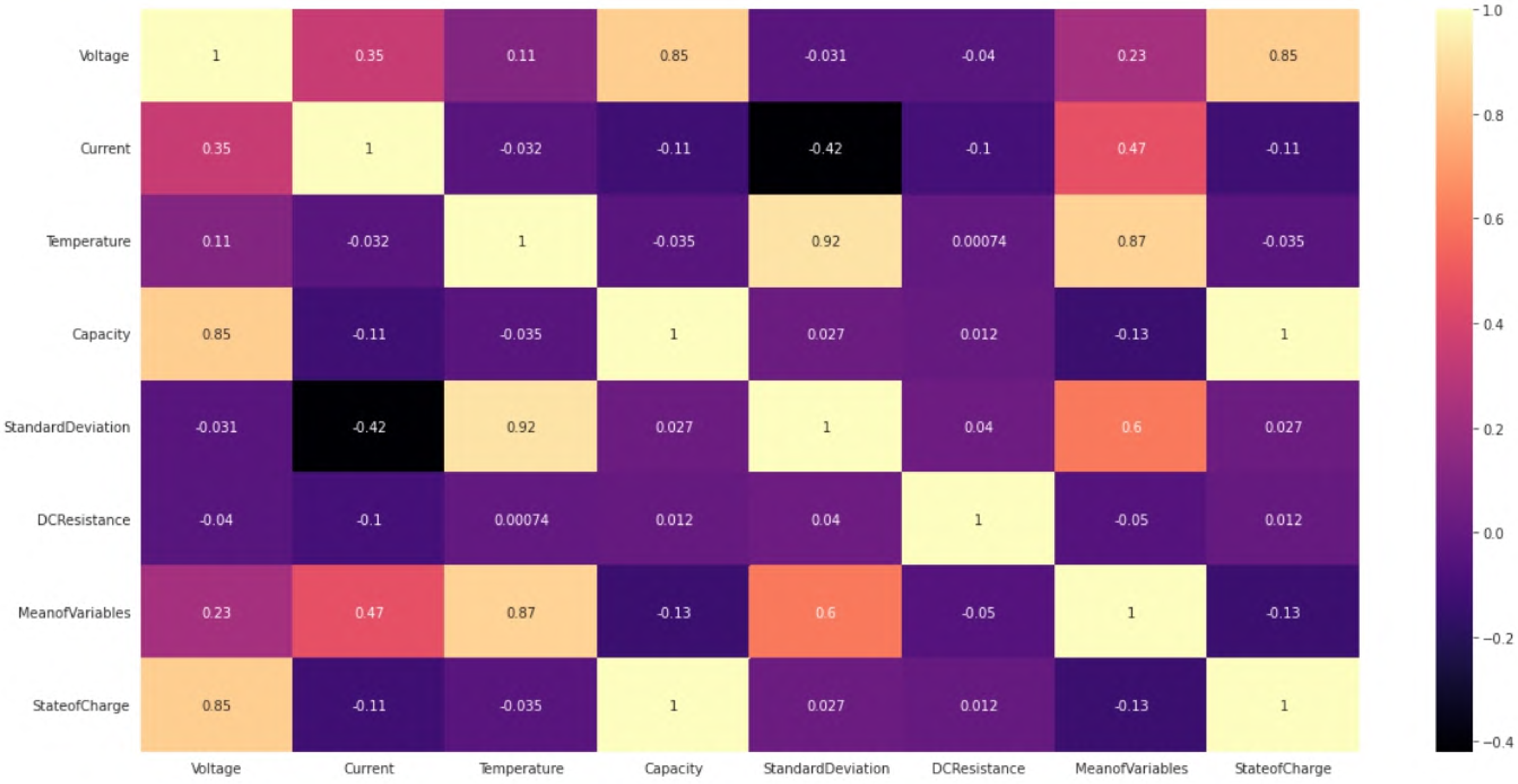


Fig. 6 - Correlation heatmap of our dataset

Results

Model Evaluation

A random unseen array of features was selected and then it's predicted output i.e State of Charge is shown by the algorithm

```
X_val.iloc[13].values.reshape(1,-1)

array([[ 3.69093000e+00, -1.80012200e+01,  3.85931800e+01,
         2.62954000e+00,  2.36311288e+01, -3.72000000e-05,
         6.16333750e+00]])

+ Code

[47] ols.predict(X_val.iloc[13].values.reshape(1,-1) )

/usr/local/lib/python3.8/dist-packages/sklearn/base.py:450:
  warnings.warn(
array([[87.65133333]])
```

Model Evaluation

A random unseen array of features was selected
and then it's predicted output i.e State of Charge is shown by the algorithm

```
[ ] X_val.iloc[13].values.reshape(1,-1)

array([[ 3.69093000e+00, -1.80012200e+01,  3.85931800e+01,
         2.62954000e+00,  2.36311288e+01, -3.72000000e-05,
         6.16333750e+00]])

[47] ols.predict(X_val.iloc[13].values.reshape(1,-1) )

/usr/local/lib/python3.8/dist-packages/sklearn/base.py:450:
  warnings.warn(
array([[87.65133333]])
```

The image below shows the original output and we can see that both the values are similar and thus, the model is working accurately

```
array([[87.65133333]])

[48] y_val.iloc[13]

StateofCharge    87.651333
Name: 7716, dtype: float64
```

Results

R-Squared Evaluation Metric

R-Squared Evaluation metric

- R-squared is a statistical measure that indicates how well a regression model fits the data. The ideal r-square value is 1
- The more closely r-square approaches 1, the better the model's fit
- R-square is the ratio between the residual sum of squares (SSres) and the total sum of squares

$$R - Squared = 1 - \frac{\Sigma(y_{act} - y_{pred})^2}{\Sigma(y_{act} - y_{avg})} \qquad (12)$$

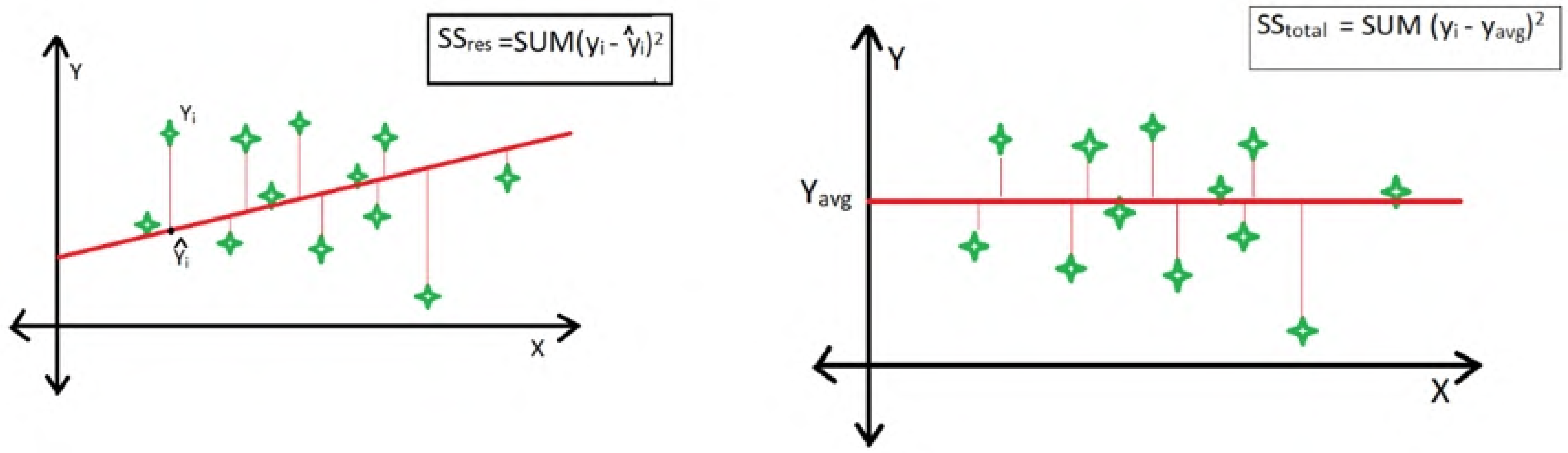


Fig. 5

https://statisticsbyjim.com/regression/interpret-r-squared-regression/

Results

Model Evaluation

Here, the R-Squared for all the three models is shown and thus the models accuracy is 99%.

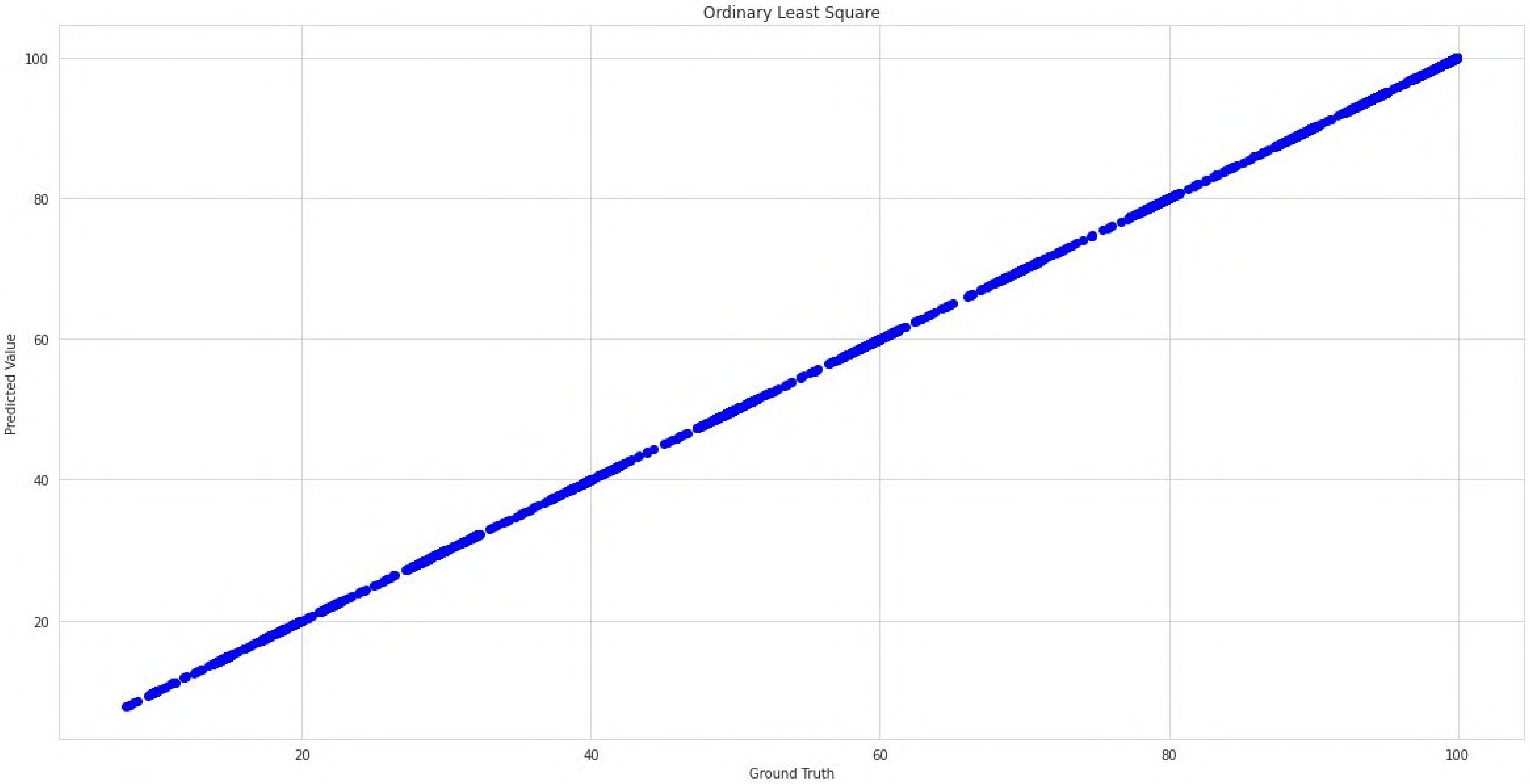
```
[19] # 2. R-squared

     print(cl('R-SQUARED:', attrs = ['bold']))
     print('----------------------------------------------------------------------------------')
     print(cl('R-Squared of OLS model is {}'.format(r2(y_val, ols_yhat_val)), attrs = ['bold']))
     print('----------------------------------------------------------------------------------')
     print(cl('R-Squared of Ridge model is {}'.format(r2(y_val, ridge_yhat_val)), attrs = ['bold']))
     print('----------------------------------------------------------------------------------')
     print(cl('R-Squared of Lasso model is {}'.format(r2(y_val, lasso_yhat_val)), attrs = ['bold']))
     print('----------------------------------------------------------------------------------')
```

```
R-SQUARED:
--------------------------------------------------------------------------------
R-Squared of OLS model i  1.0
--------------------------------------------------------------------------------
R-Squared of Ridge model i  0.999999917874719
--------------------------------------------------------------------------------
R-Squared of Lasso model i  0.9999998132931261
--------------------------------------------------------------------------------
```

Results

Predicted vs Actual Value



Try Pitch

BMS Module Circuit Diagram

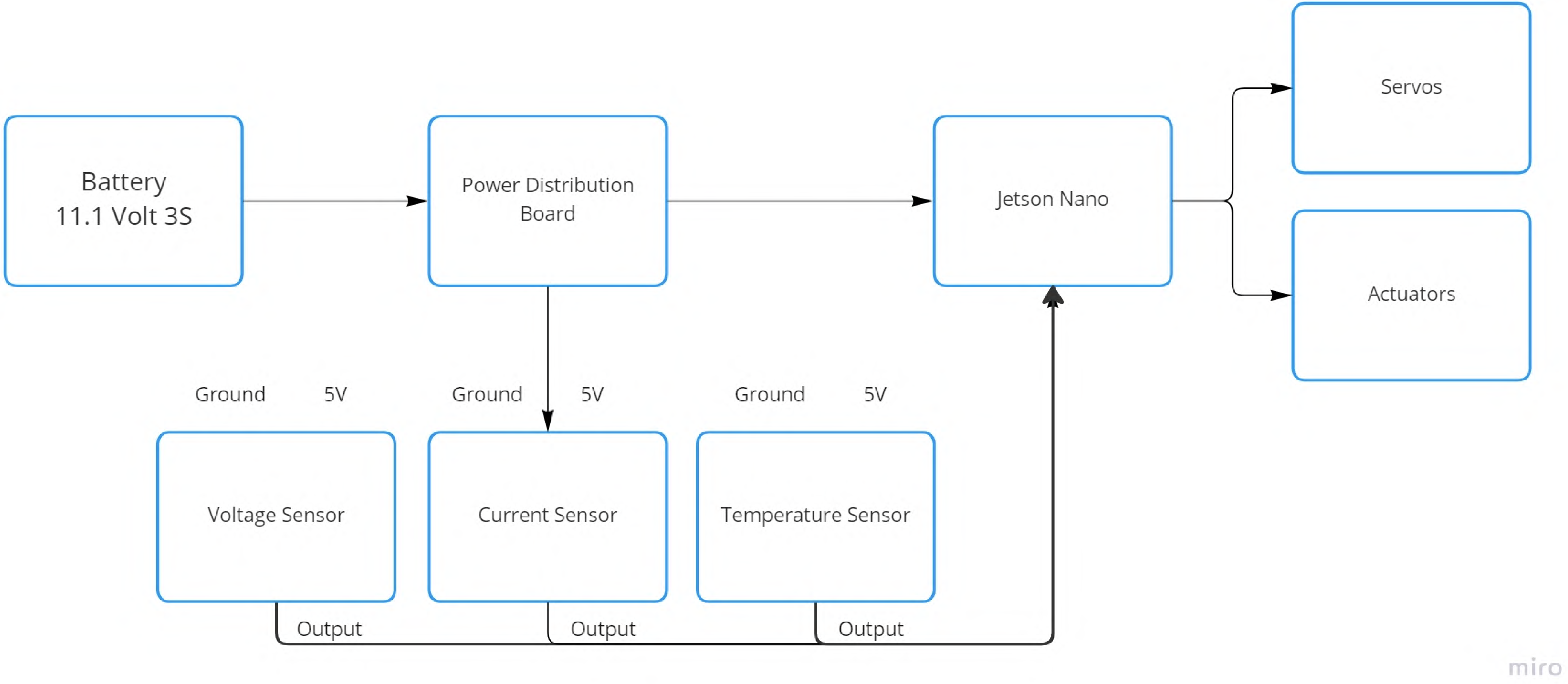


Fig. 1

Future Work

- Although, the model predicted accurately, regression cannot be used for large dataset which requires high computational power.

- Thus, a higher complex model such as neural networks and Support Vector Machine should be implemented to make the use generalized.

- In future, we will try to use Artificial Neural Networks and Support Vector Machine and compare the accuracy with more datapoints.

- After testing the model, our next goal would be to integrate the hardware system with this model and real time data can be collected and predicted with the help of this system.

References

1. Vilsen, S. B., & Stroe, D.-I. (2021). Battery state-of-health modelling by multiple linear regression. *Journal of Cleaner Production, 290*, 125700. https://doi.org/10.1016/j.jclepro.2020.125700
2. Yang, H., Cao, Y., Xie, H., Shao, S., Zhao, J., Gao, T., Zhang, J., & Zhang, B. (2020). Lithium-ion Battery Life Cycle Prediction with Deep Learning Regression Model. *2020 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 3346–3351. https://doi.org/10.1109/APEC39645.2020.9124049
3. https://www.ansys.com/en-in/blog/building-better-batteries
4. https://medium.com/analytics-vidhya/linear-regression-explained-in-simple-terms-yagnik-8f9eccb680ec

Thank you