**Assignment 2**

**Case 2: Energy Forecasting**

**FLOWCHART:-**

Compute the performance Metrix

Build model using Neural Network

Predict power usage

Compute Confusion Matrix and Error rate

Build classification models

Calculate average kWh

Read sampleFormat.csv

Use the above model to predict the performance

Convert to the sample format

Read forecastNewData.csv

Compute the performance Metrix

Build model using Regression Tree

Predict power usage

Predict power usage

Use functions from package zoo to fill NAs. Build Regression model and Check Performance

Replace zeros with NA

Build Regression model and check Performance

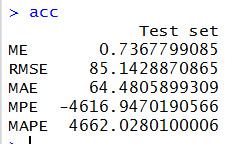
Remove all Zero entries

Import NewData.csv

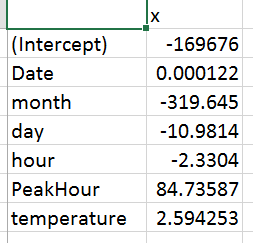
**Algorithm Implementation:-**

**Data wrangling and cleansing and Multiple linear regression :**

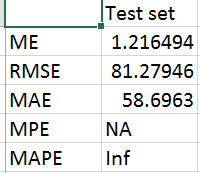
1. Remove all the zero-entries and use only the non-zero entries for KWH to build a regression model



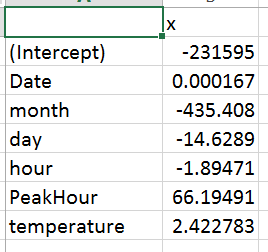
Model is as below:



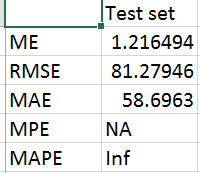
b. Raw data including zero:



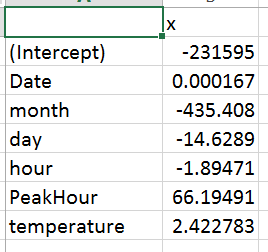
Model as below:



1. Use this model to predict power usage. Again build model on combined data. Check performance.
2. Raw data including zero:

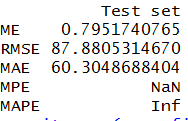


Model as below:



1. Using package zoo we replace zeros with NA and use functions na.approx, na.locf and na.fill to replace NAs.

Performance matrix after using functions na.approx or na.locf or na.fill to replace NAs

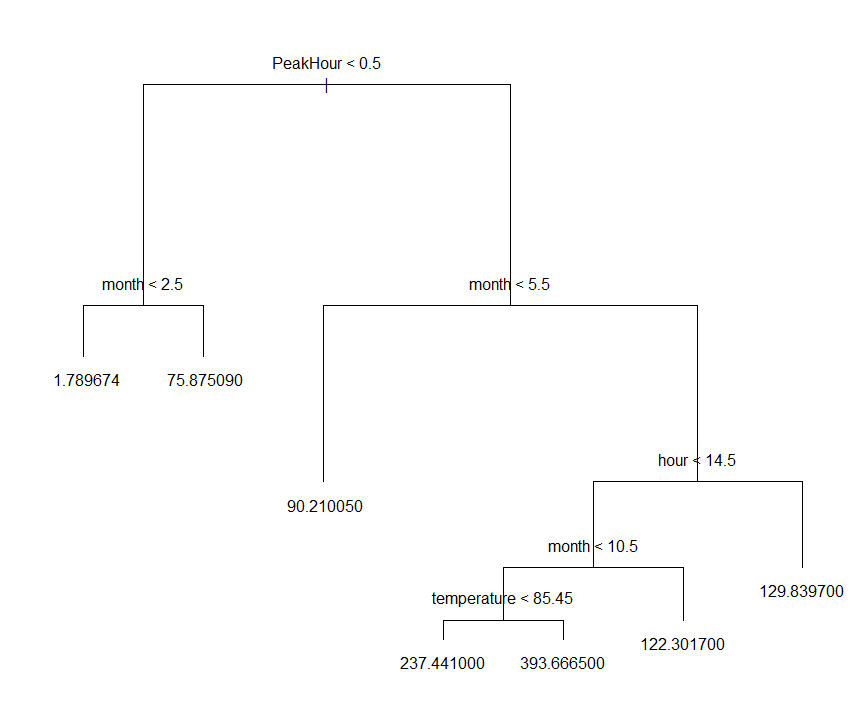


1. Selected the optimized model as combination of above 3 steps to get final data set for below models

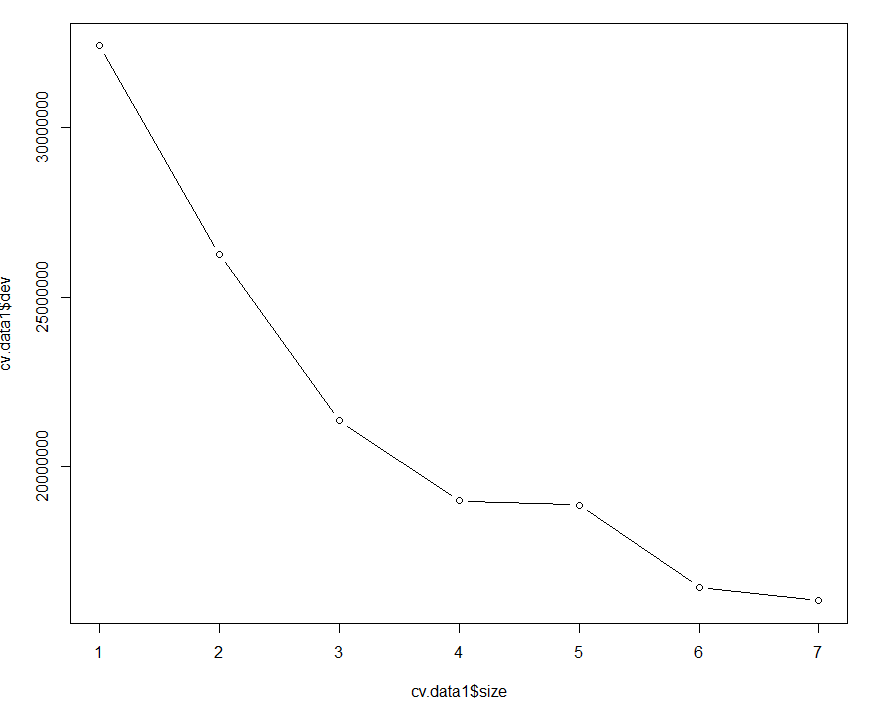
**Prediction:-**

1. **Regression Tree: Build a model using the tree() function and predict the power usage. Find performance value with the help of accuracy.**
2. Tree model without pruning

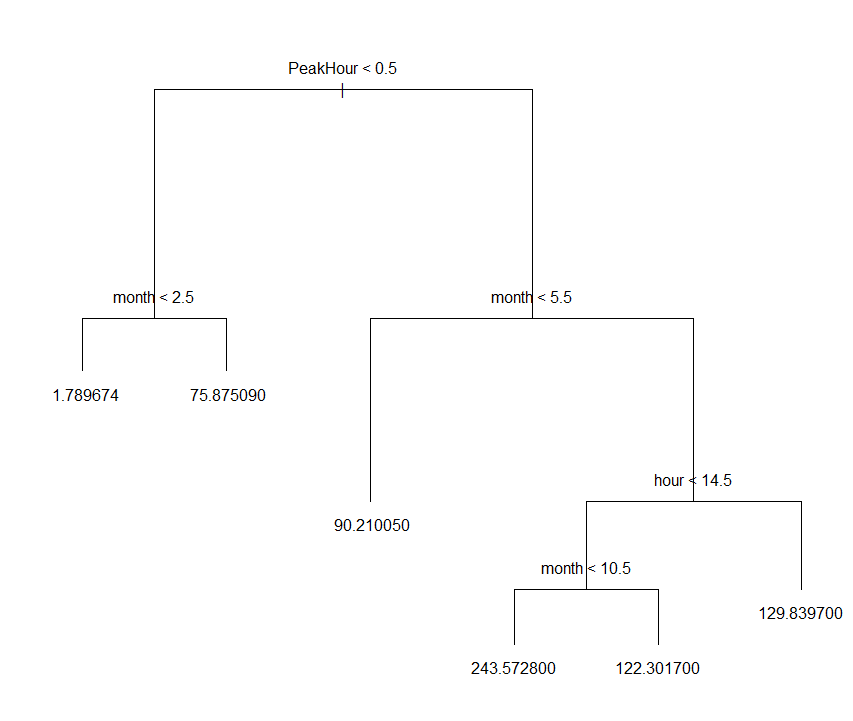
b.

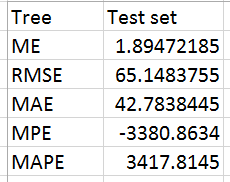


1. Decide best size for prune tree:



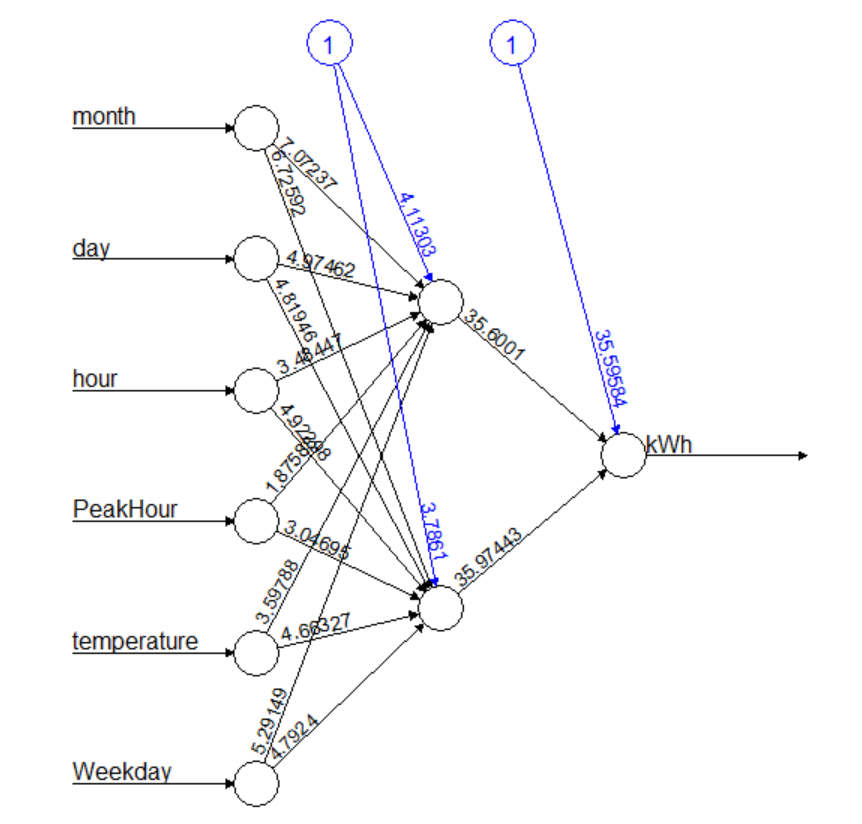
1. Prune tree:



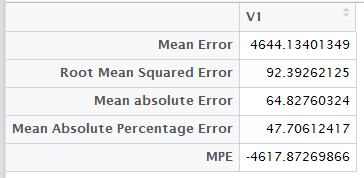


We have decided the tree based on performance matrix and we found that we get better result when we take the tree with prune of best 6

1. **Neural Networks: Build a model using the neuralnet() function and predict the power usage. Find performance value with the help of accuracy.**
2. Build neural network on our data, model is as below:

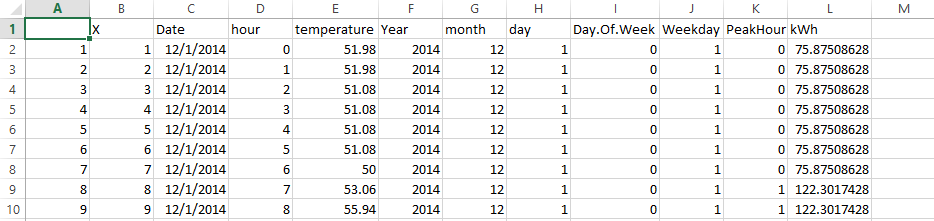


1. Performance matrix for above model:

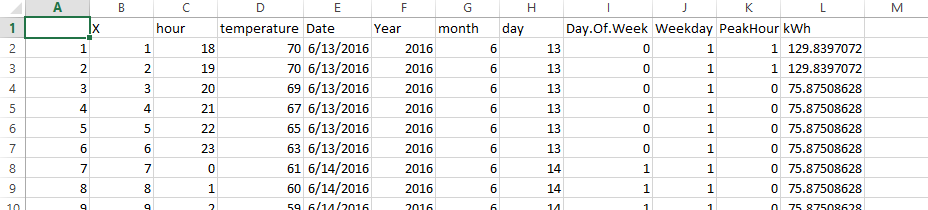


**Forecast:-**

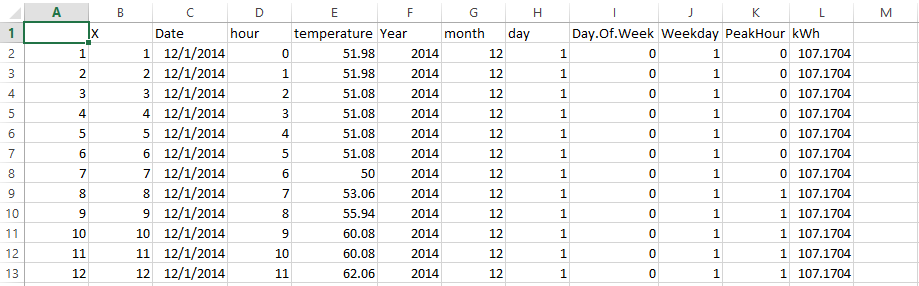
1. Converted forecastNewData.csv and forecastNewData2.csv into forecastInput.csv format
2. Applied Tree model and Neural Network models generated in step 2 to forecast power
3. a. After applying neural network model, below is power forecast for each hour for each hour for **forecastNewData.csv**:



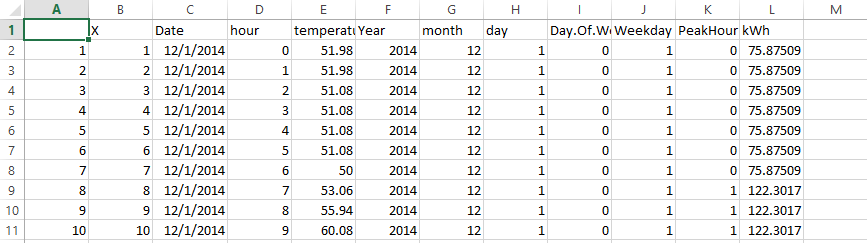
1. After applying tree model, below is power forecast for each hour for **forecastNewData.csv**:



1. a. After applying neural network model, below is power forecast for each hour for each hour for **forecastNewData2.csv**:



1. After applying tree model, below is power forecast for each hour for **forecastNewData2.csv**:

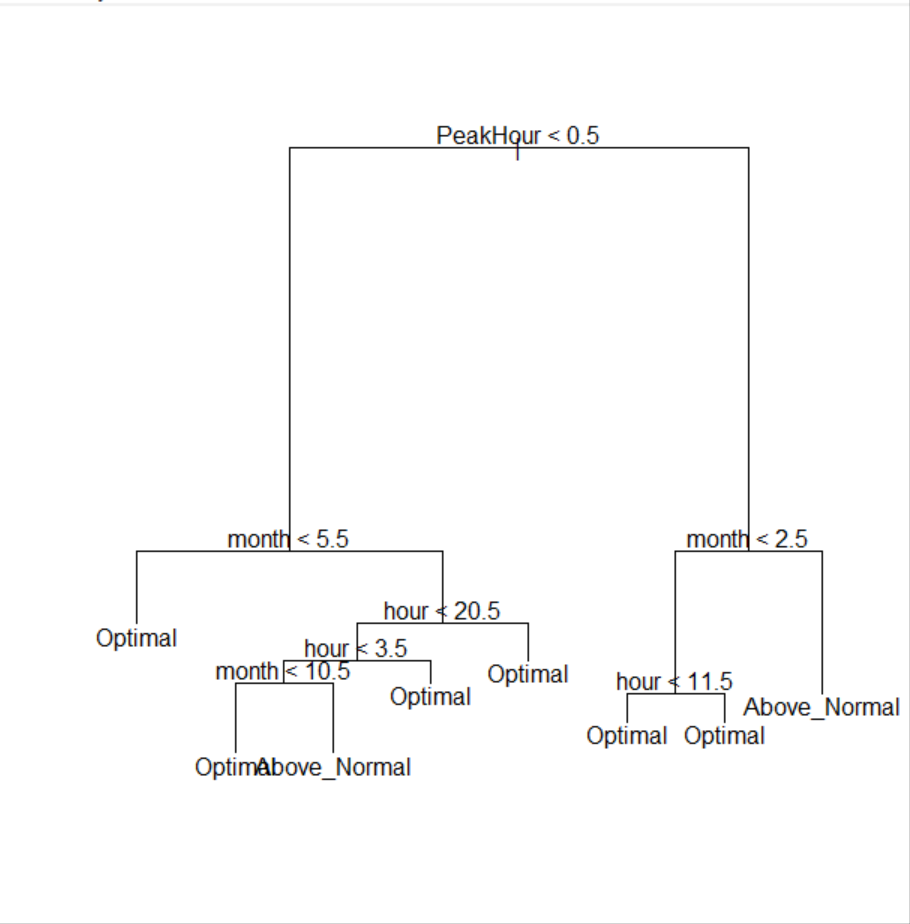


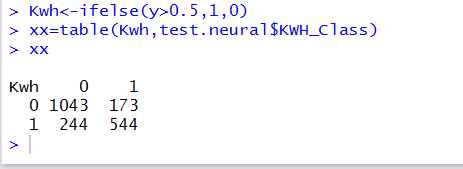
**Part 2: Classification:-**

1. In the Hourly\_filled\_data.csv file, compute the average KWH and add a new column, KWH\_Class.
2. If KWH > average KWH, KWH\_Class = “Above\_Normal” otherwise, KWH\_Class = “Optimal”

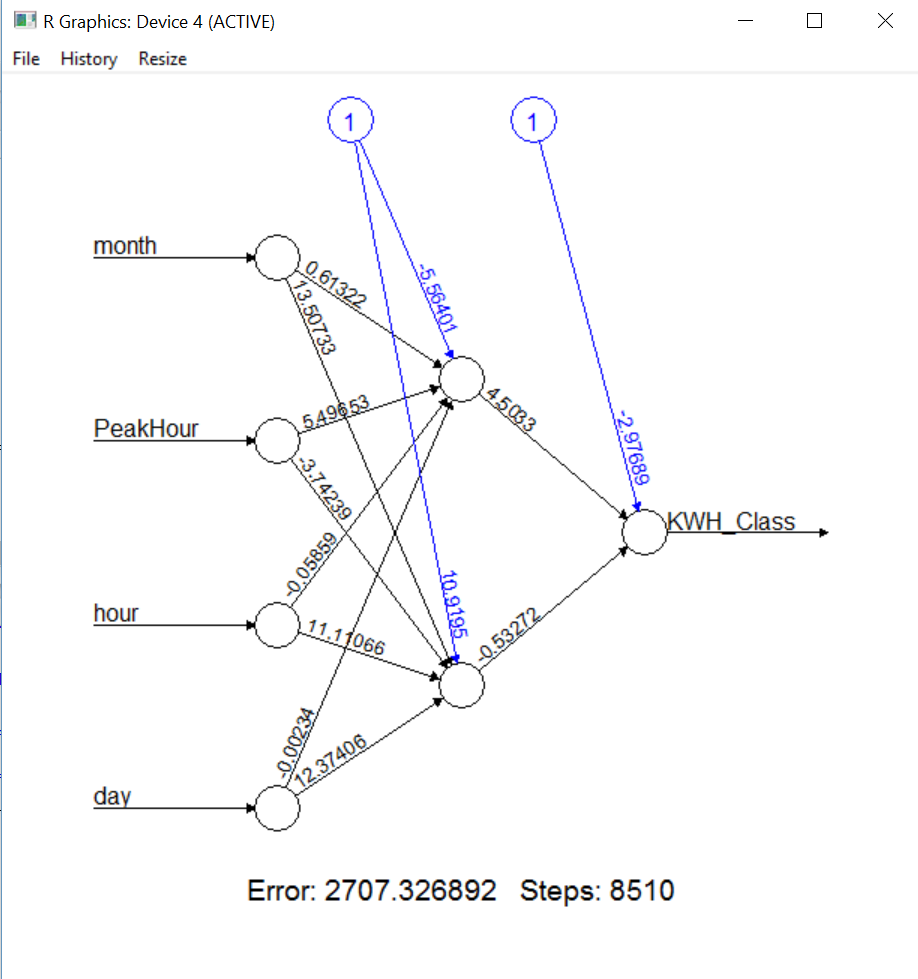
**i) Logistic Regression**

**ii) Classification Tree**

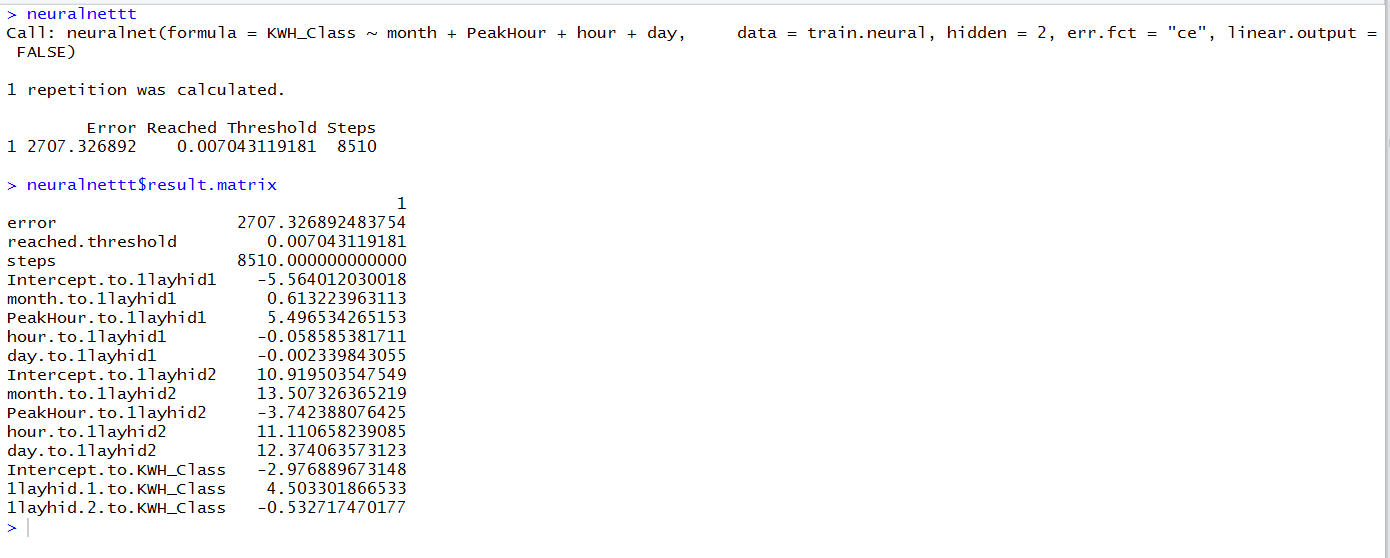




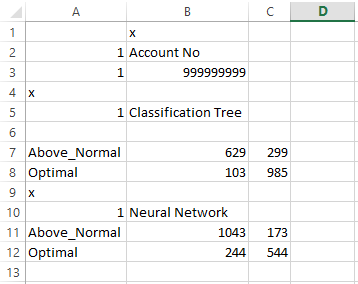
**iii) Neural Network**



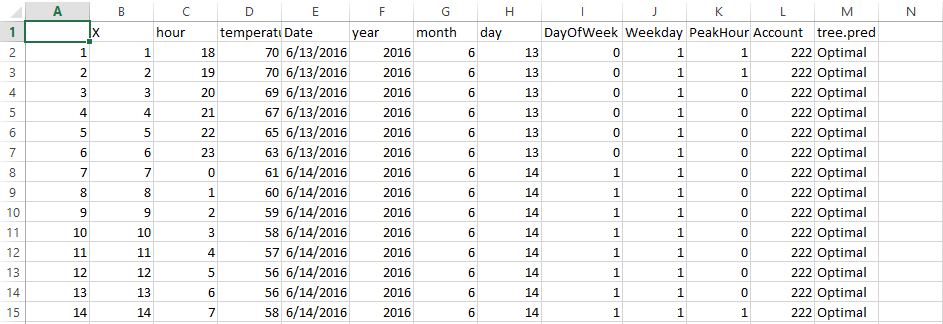
Neural net output:



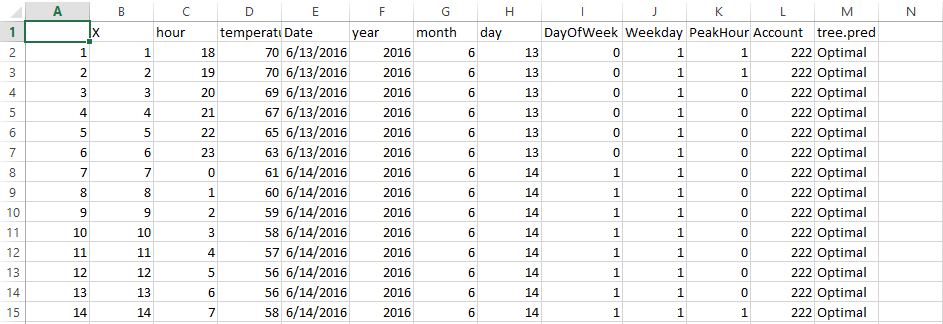
Confusion Metrix for classification using regression tree and neural network:



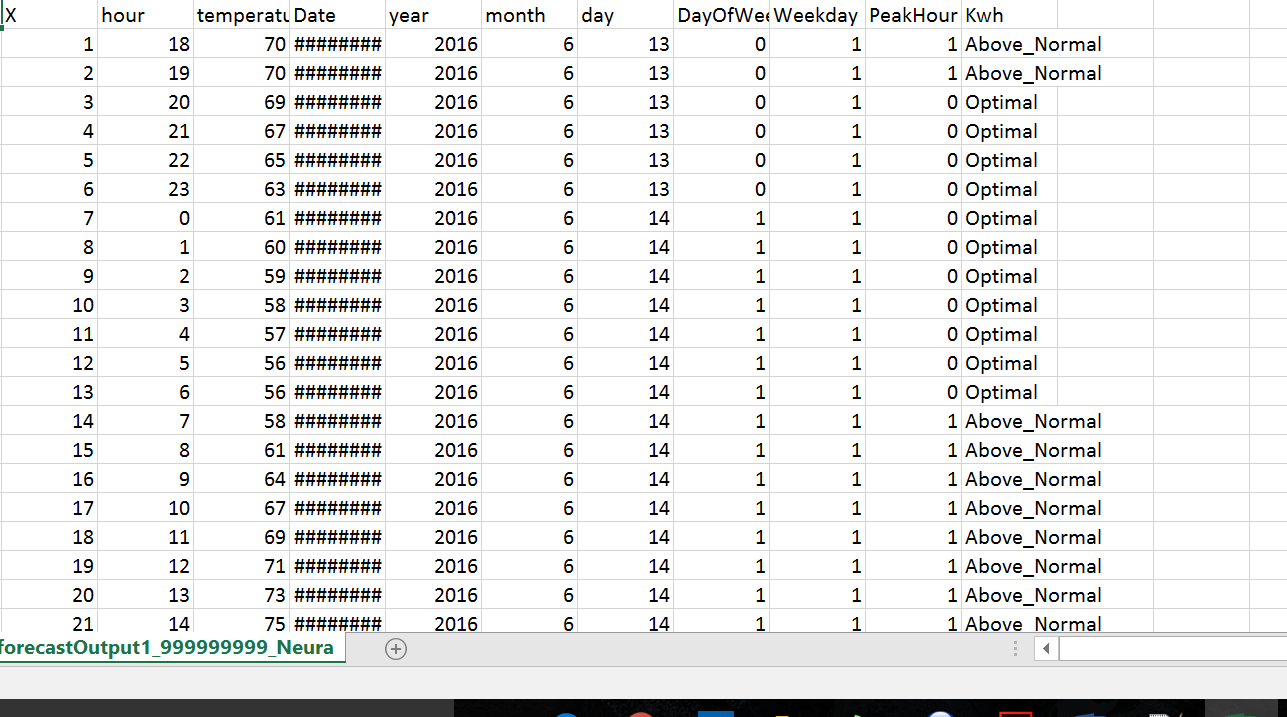
1. **Classification Forecast:**
2. **Forecast for Classification tree for forecastnewData.csv**



1. **Forecast for classification tree for forecastnewData2.csv**



1. **Forecast for classification neural net for forecastnewData.csv**



1. **Forecast for classification neural net for forecastnewData2.csv**

