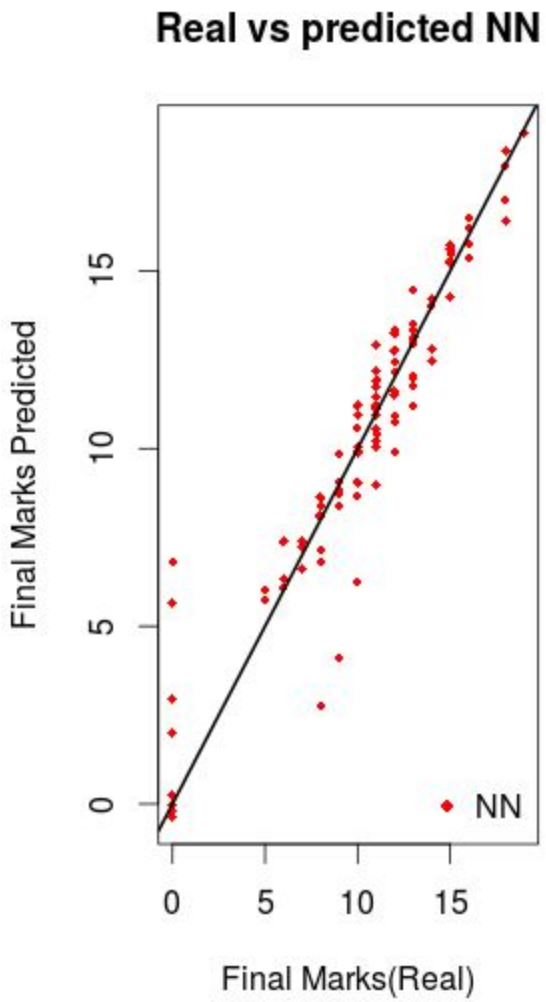
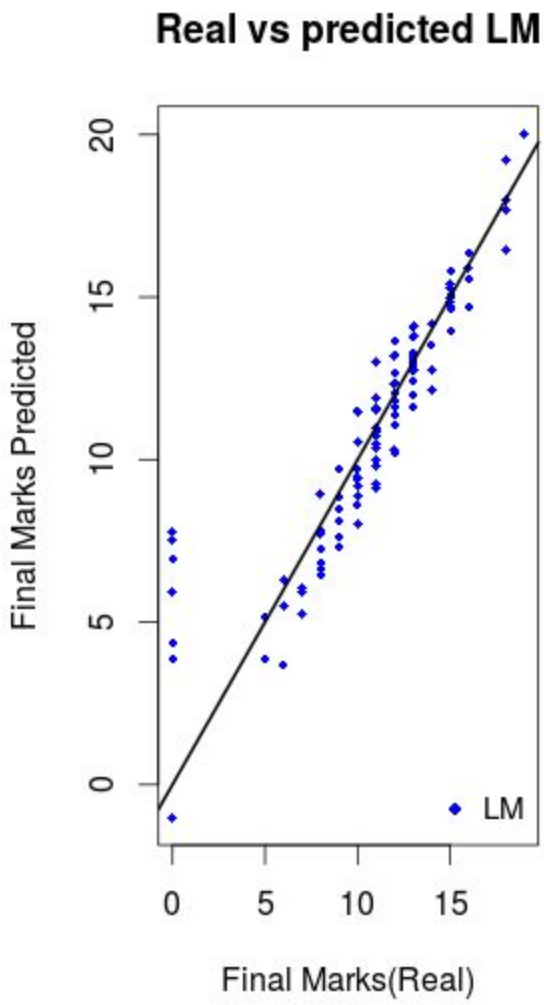
Experiments And Results

In the previous section we went into details with respect to the models that were built in the project to predict student performance. The UCI Student Exam Performance Dataset was split into training and test data. Training data consisted of 70% of the dataset and the remaining 30% of the dataset was used as test data. The neural network as well as the linear model were trained and tested. This was done for both the Math as well as the Portuguese courses present in the dataset.
The performance of the models were very good. This is shown in the visualizations below which is a scatter plot of Real vs Predicted values in the Neural Network as well as the Linear Model for the Mathematics course. Similar results were obtained for the Portuguese course as well.





We observed that both the linear model as well as the neural network performed in a similar fashion with the neural network having a slightly better RMSE value as compared to the linear model. In case of the Portuguese course, the linear model had a better RMSE value.However, on enhancing the neural network further using tensor flow as described in the previous section, the RMSE value further improved and beat the linear model in both the courses.

Model Validation

The next step was to validate our models. That is, we wanted to test how well it performs as compared to a predictive model that has already been applied to the dataset. This was done by comparing the RMSE values obtained by our models as

compared to those obtained in the paper “Using Data Mining To Predict Secondary School Student Performance” authored by P. Cortez and A. Silva.

The table below is a summary of the RMSE values of various models present in the paper for the Mathematics course.

Naive Predictor	Neural Network	Support Vector Machine	Decision Tree	Random Forest
2.01	2.05	2.09	1.94	**1.75***

* denotes best value

The table below is a summary of the RMSE values of the models built by us for the Mathematics course.

Linear Model	Neural Network (2 layers)	Enhanced Neural Network(3 layers)
1.93	1.48	**1.14***

* denotes best value

The table below is a summary of the RMSE values of various models present in the paper for the Portuguese course.

Naive Predictor	Neural Network	Support Vector Machine	Decision Tree	Random Forest
1.32*	1.36	1.35	1.46	**1.32***

* denotes best value

The table below is a summary of the RMSE values of the models built by us for the Portuguese course.

Linear Model	Neural Network (2 layers)	Enhanced Neural Network(3 layers)
1.21	1.51	**1.10***

* denotes best value

Thus, the results obtained show that through our enhanced neural network we have obtained a better RMSE value. This result provides us with a solid foundation to carry our research forward as with the addition of course dependency weights to our model which is our future enhancement we look to further improve the model and thus come one step closer to our aim of making an impact in the field of education through data analytics.