Q1.

**Naïve Bayes Model:**

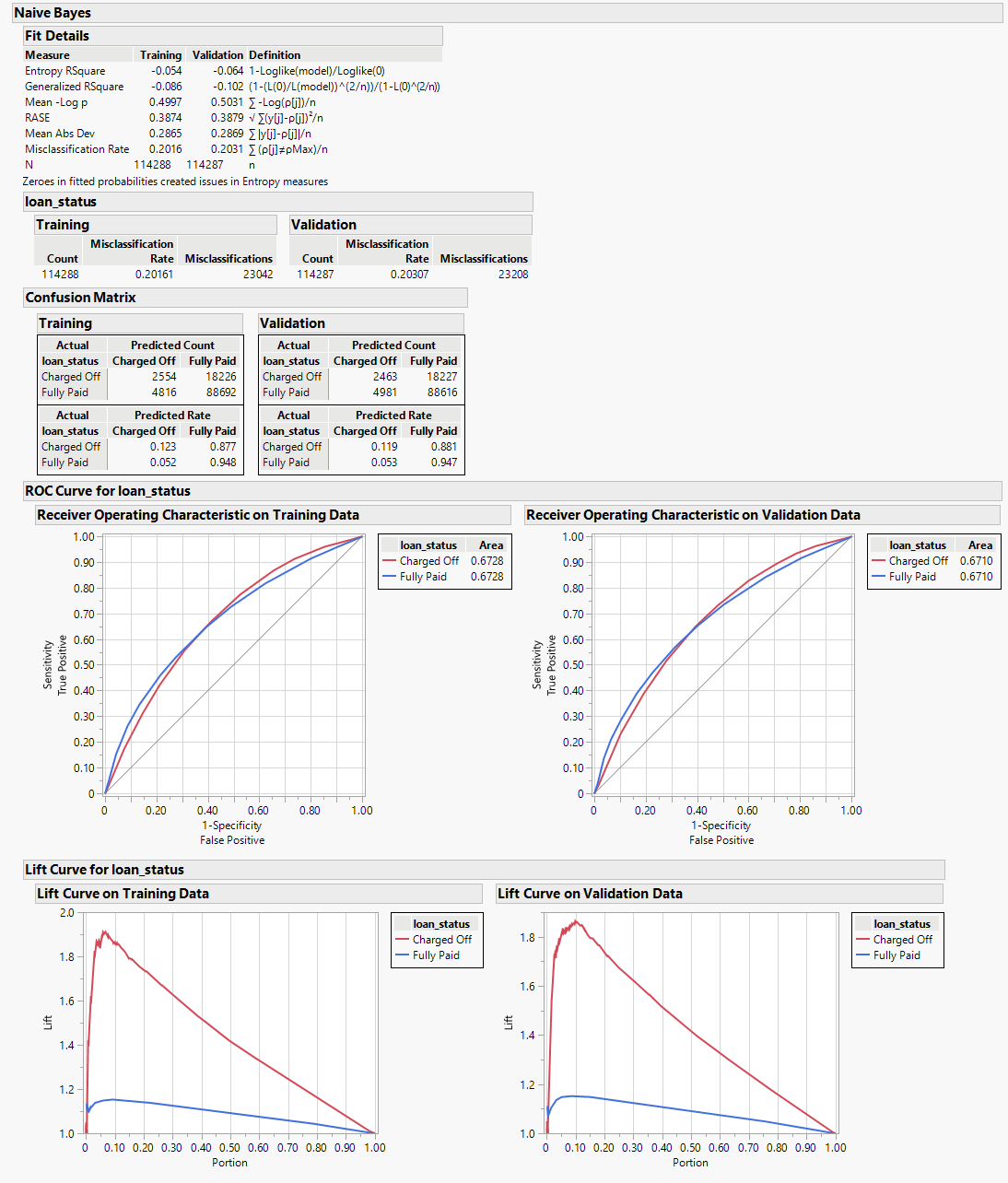
****

Figure 1. Naïve Bayes Model

AUC for naïve bayes final model is 67.10% for my model with logistic model 1st iteration from homework 5. With -6.4% Entropy RSquare, Generalized RSquare -10.2%, RMSE of 38.79% for the validation of naïve bayes model. As you can see, the misclassification rate is 20.31% for the validation column

**Model Comparison of 5 best models from HW5 till HW10 Best/Final Models:**

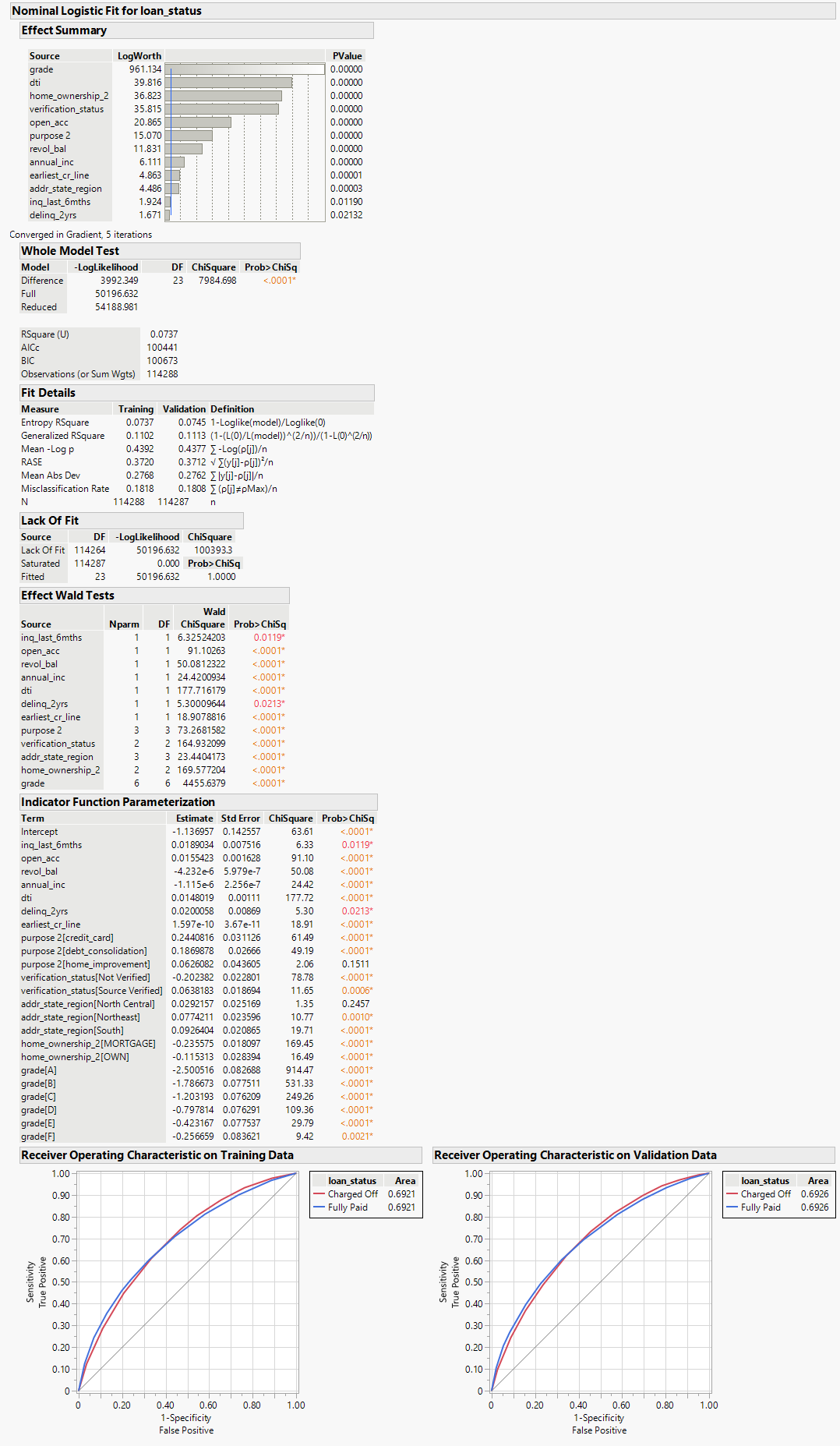
****

Figure 2. Logistic Regression Final Model from Homework 5

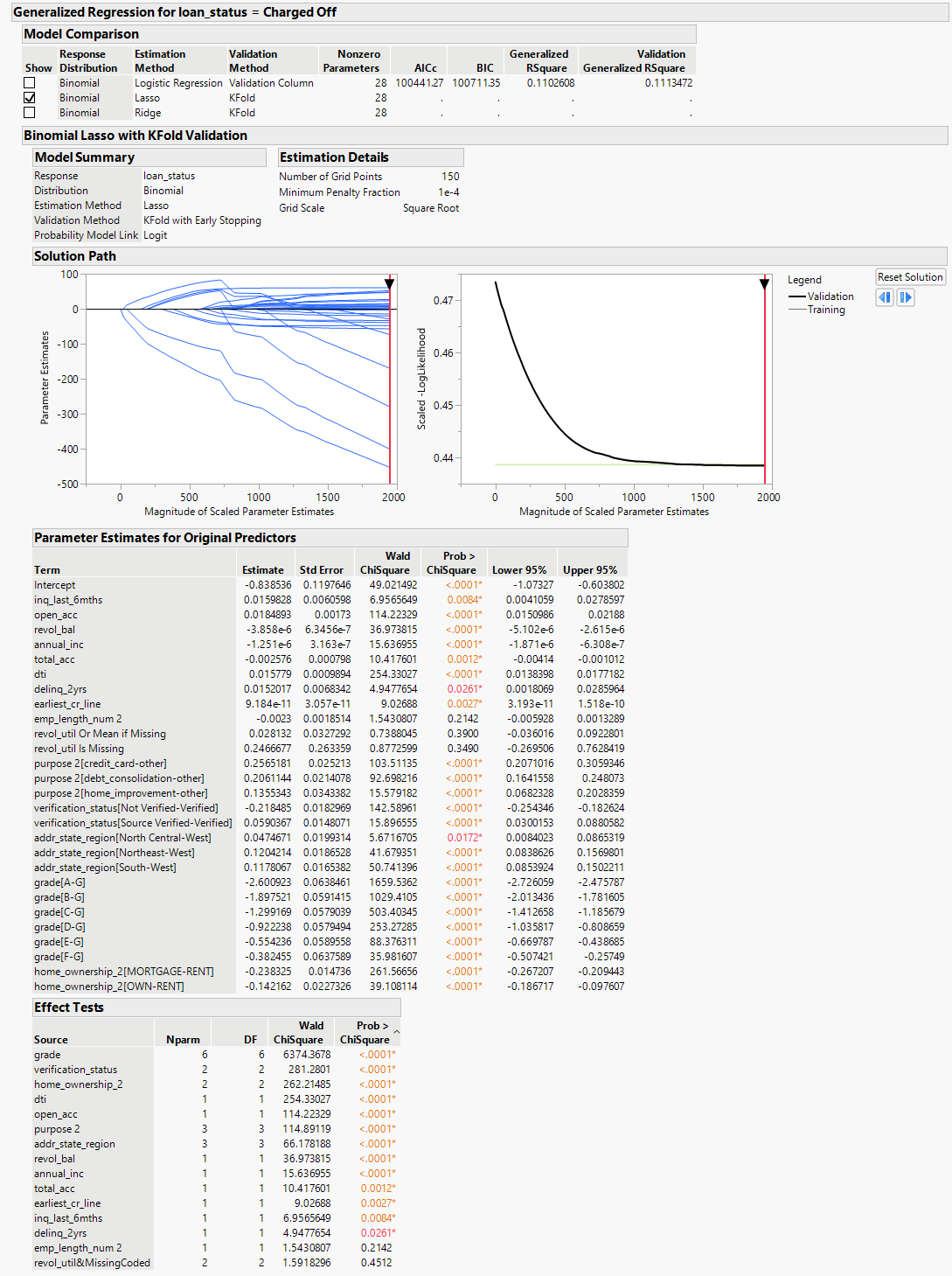


Figure 3. Lasso Regression Model

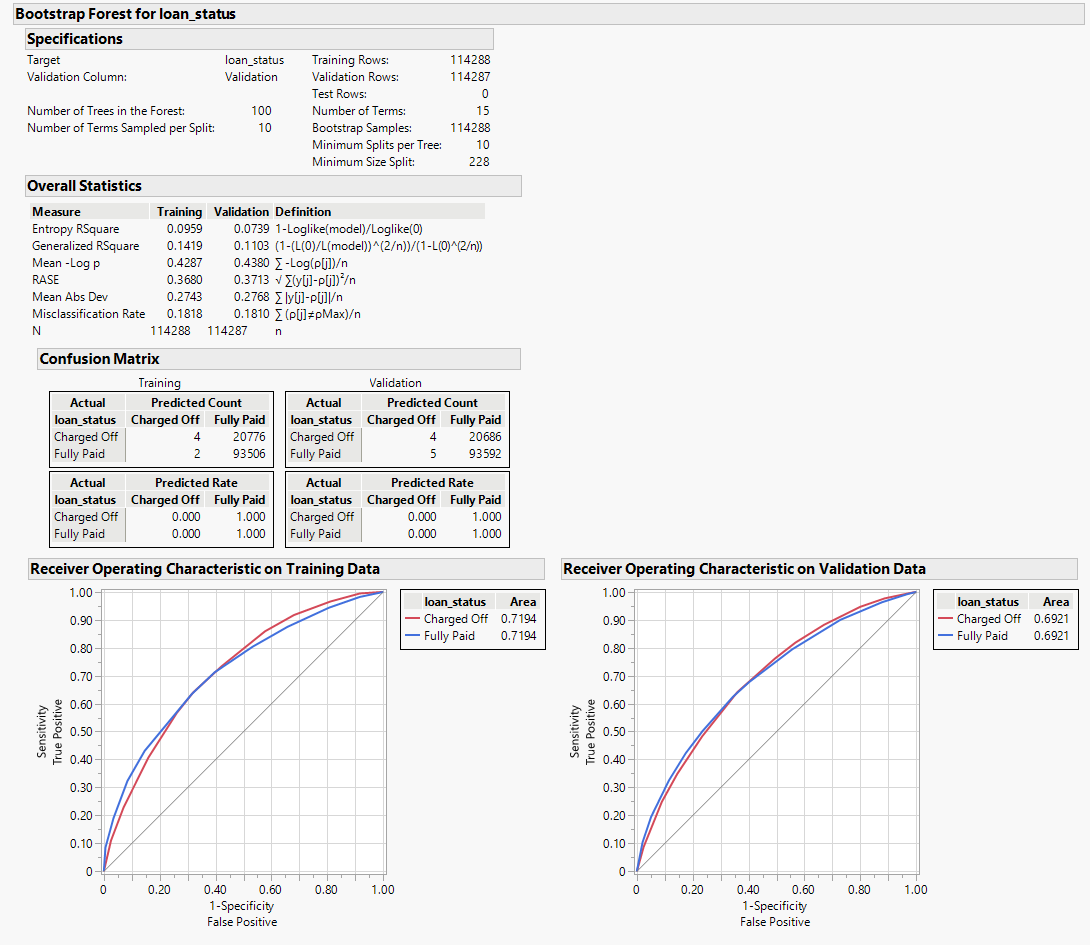


Figure 4. Bootstrap Forest Model

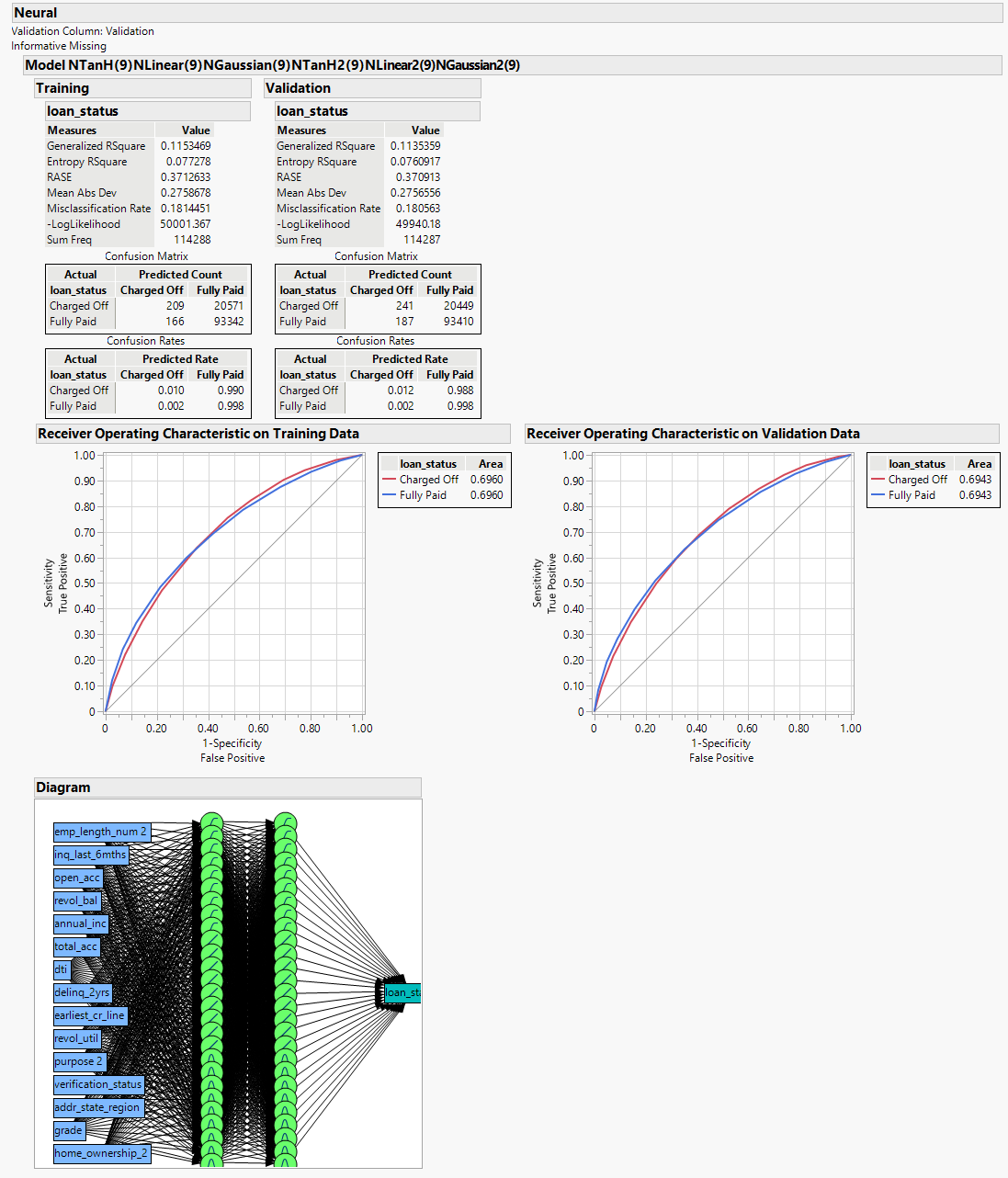


Figure 5. Best Neural Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Best Models from HW 5, 7, 8, 9, 10** | **AUC** | **Entropy R2** | **Generalized R2** | **RMSE** |
| Final Logistic Linear Regression Model from HW5 | 0.6926 | 0.0745 | 0.1113 | 0.3712 |
| Naïve Bayes Model | 0.6710 | -0.064 | -0.102 | 0.3879 |
| Lasso Model | 0.6931 | 0.0750 | 0.1120 | 0.3711 |
| Bootstrap Forest | 0.6921 | 0.0739 | 0.1103 | 0.3713 |
| 2 Layers 9 Nodes Each for TanH, Linear, Gaussian Activation Function | 0.6943 | 0.0761 | 0.1135 | 0.3709 |

Looking at the above table, the Neural model with the 2 layers and 9 nodes each for TanH, Linear, and Gaussian Activation Function works best as per the AUC (0.6943) followed by Lasso model with the AUC (0.6931). As for the AUC, Naïve Bayes is worse at AUC 0.6710. by far, the Naïve Bayes Model is the worse and Neural Model is the best in all above outcomes. RMSE is also the lowest at 0.3709 for Neural Model and for Naïve it is worse at 0.3879 as the low RMSE is better. we can deduce that the Lasso Model and the Bootstrap Forest have the highest AUC scores, suggesting they are better at classifying instances correctly compared to the other models. The Final Logistic Linear Regression Model from HW5 has the highest Generalized R², indicating it might best explain the variance in the dataset after accounting for the model's complexity. It’s important to note that the choice of the best model could also depend on the specific context and requirements of the predictive task, such as the trade-off between model interpretability and accuracy.