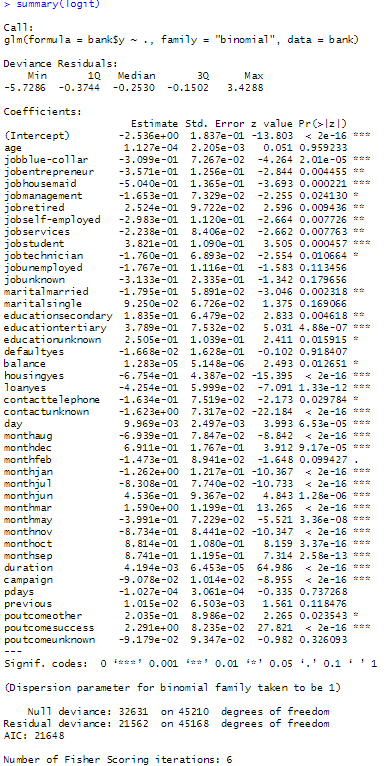
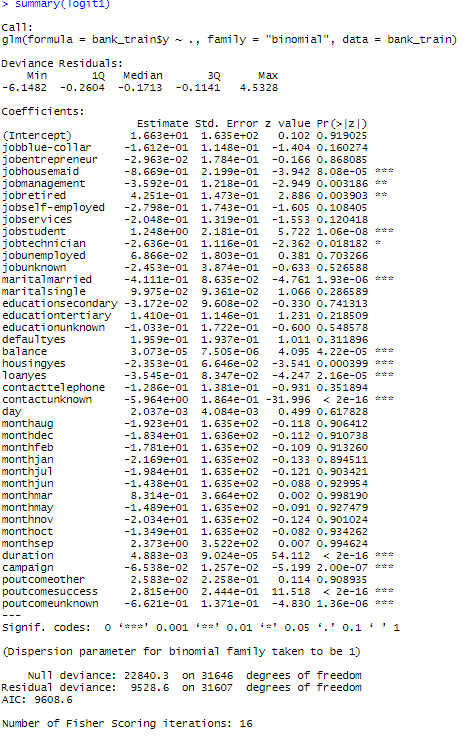
**Logistic Regression**

**Bank Term Deposit Subscription Prediction**

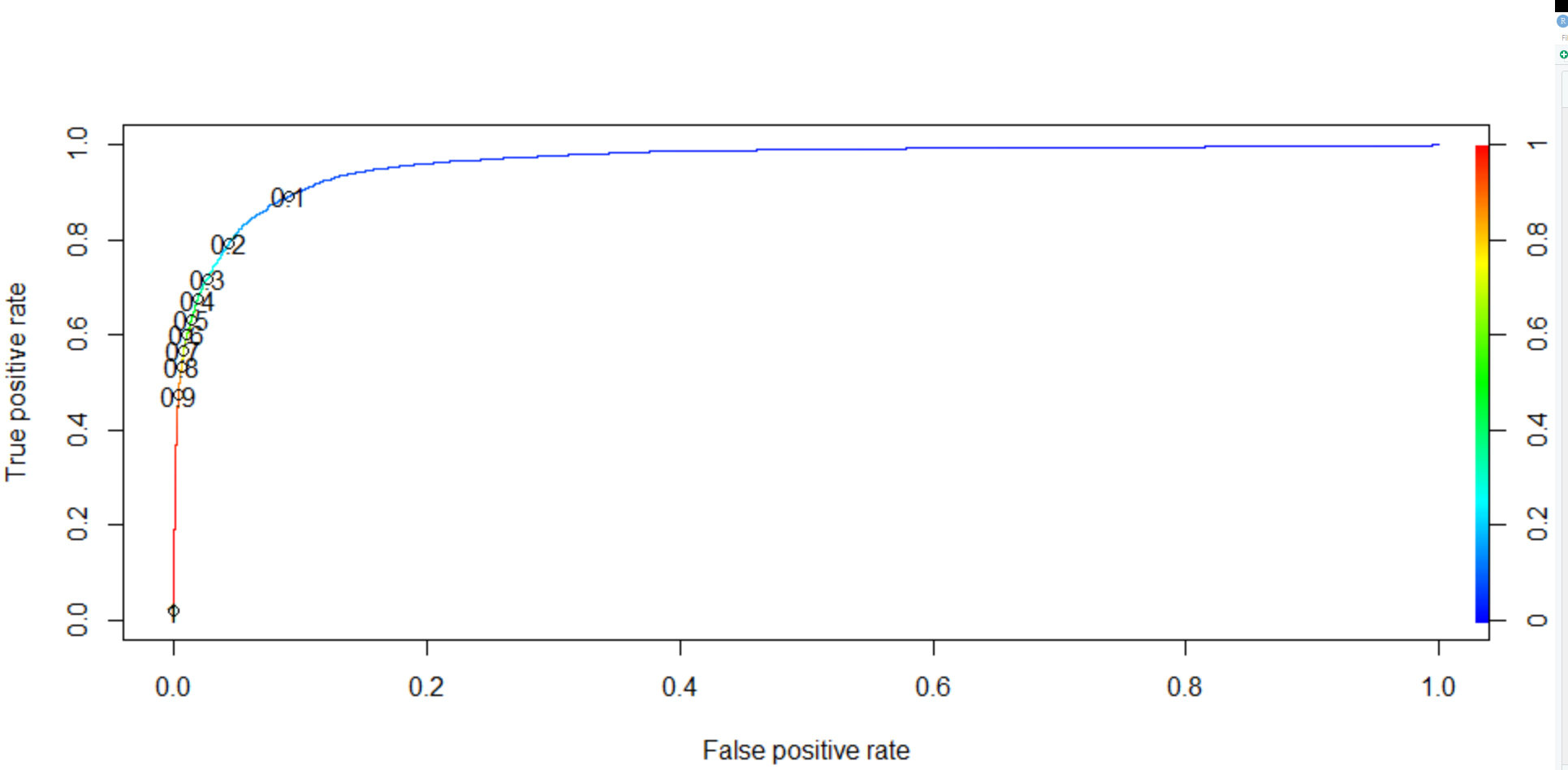
* After performing EDA and making the first logit model, I got the following result:



* From above summary, we can infer that the probability values of “age”, “pdays” and “previous” are not significant. So, I deleted those columns.
* After deleting those 3 columns I split the data and formed another model, for which following results were obtained:



* We can see that the AIC value has also improved from 21648 to 9608.6.
* For above model, with a cutoff of 0.5, I achieved a training accuracy of 94.48 % and testing accuracy of 32.16 %.
* Conducted a ROCR analysis to understand the optimum cutoff value by plotting the ROC curve:



* AUC for above model is 0.9618256 which is outstanding. From above plot we can infer that the cutoff value should be in between 0.1 and 0.2.
* For a cutoff value of 0.12 I got a training accuracy of 91.86% and testing accuracy of 29.50%