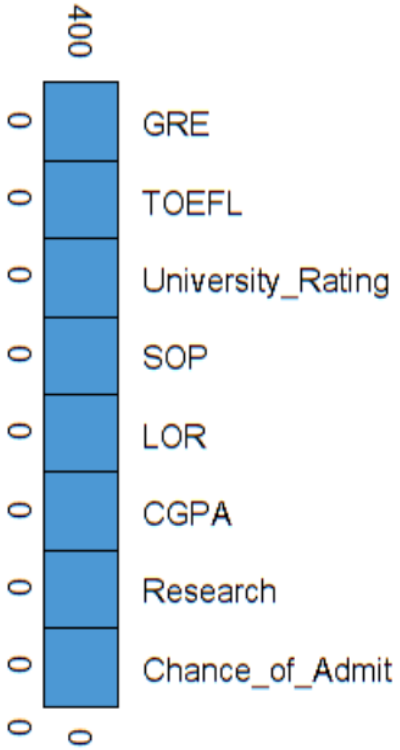
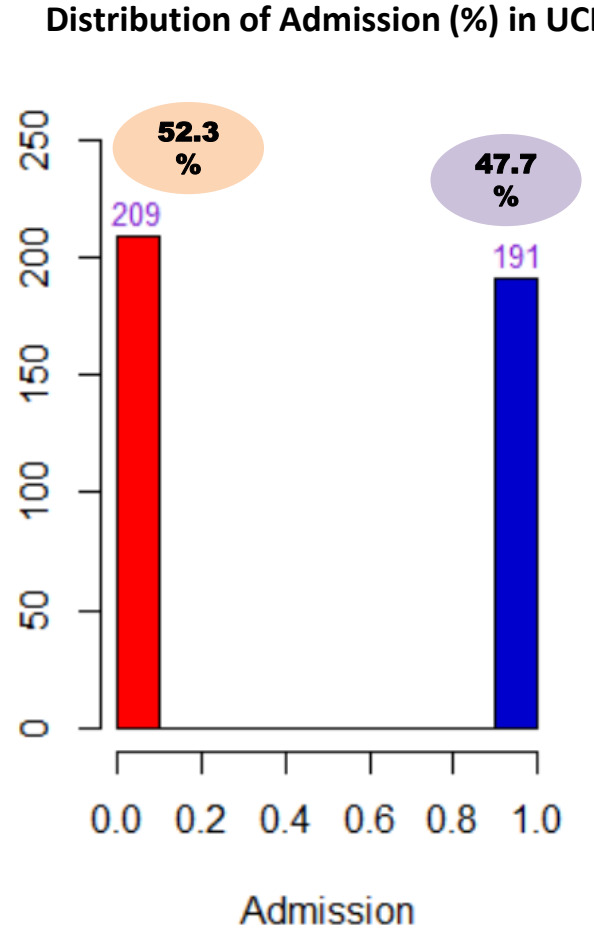
COLLEGE ADMISSION

List of Variables after Cleaning

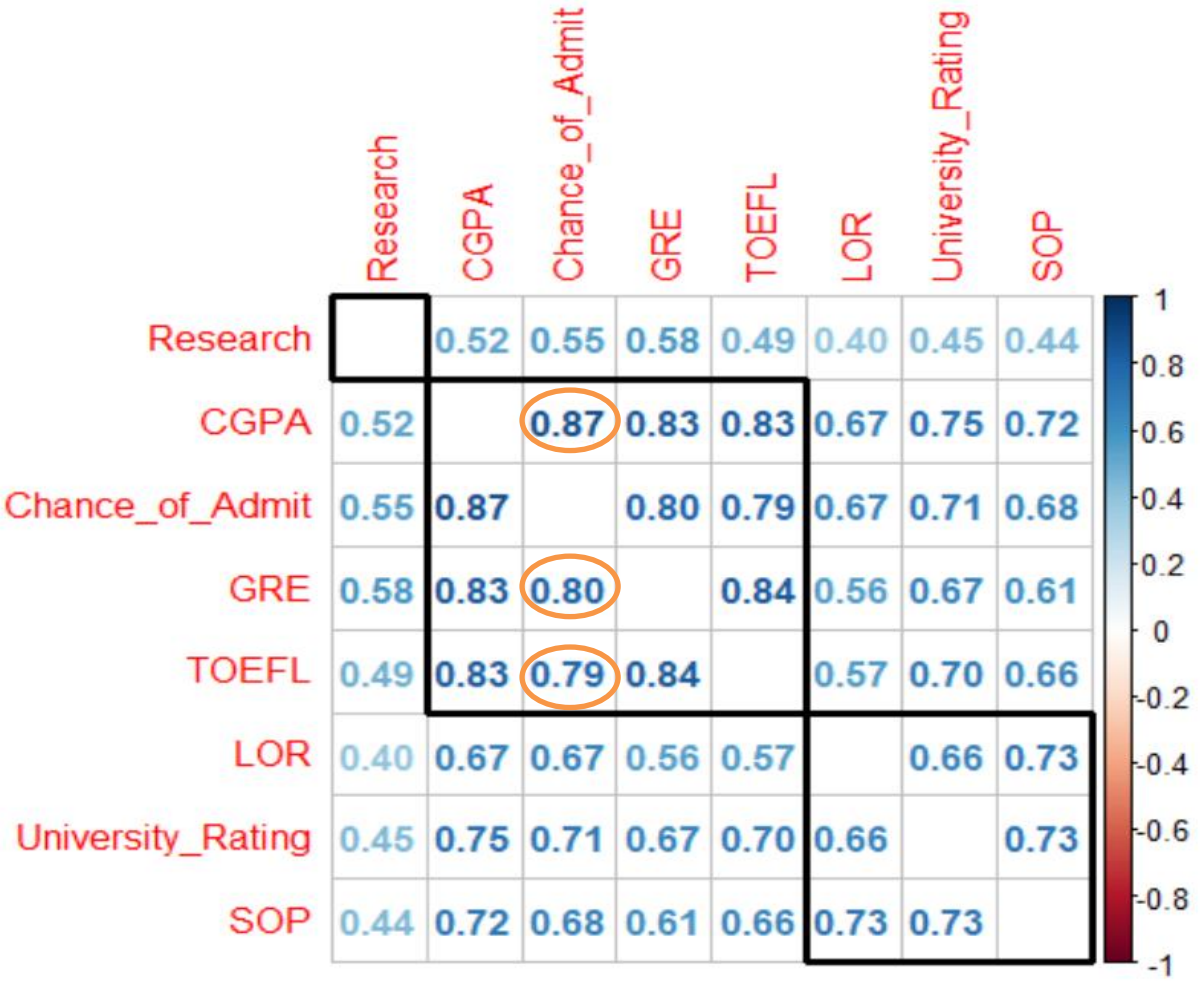


Preliminary Data



Correlation

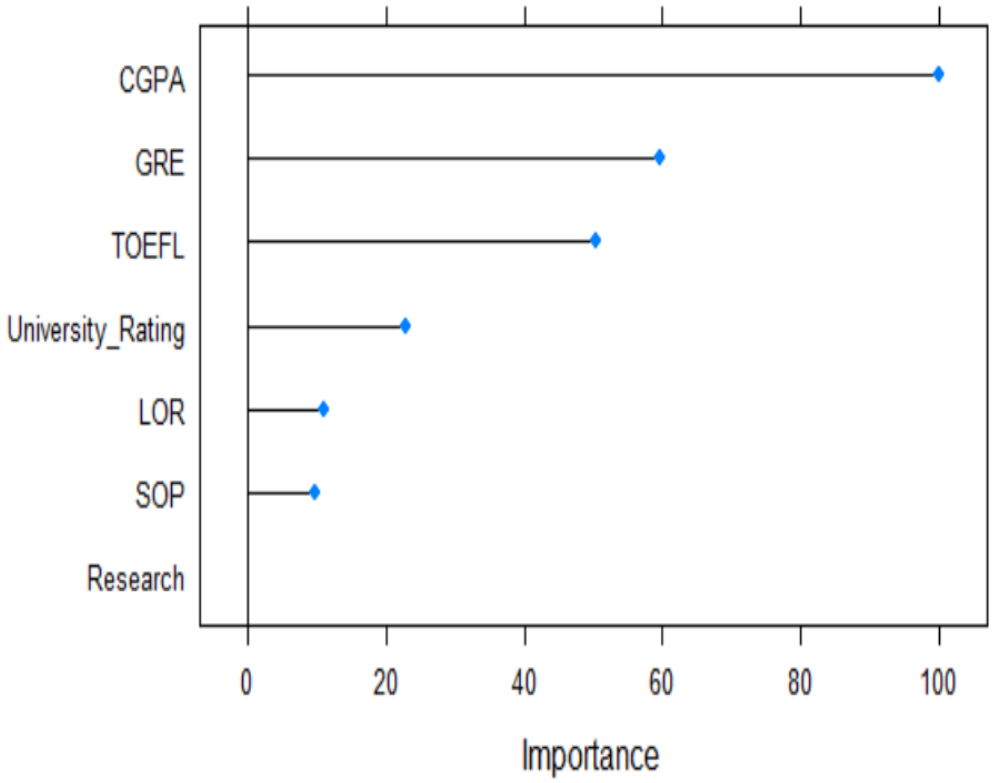
No. Observations = 400
No. Variables = 8



COLLEGE ADMISSION

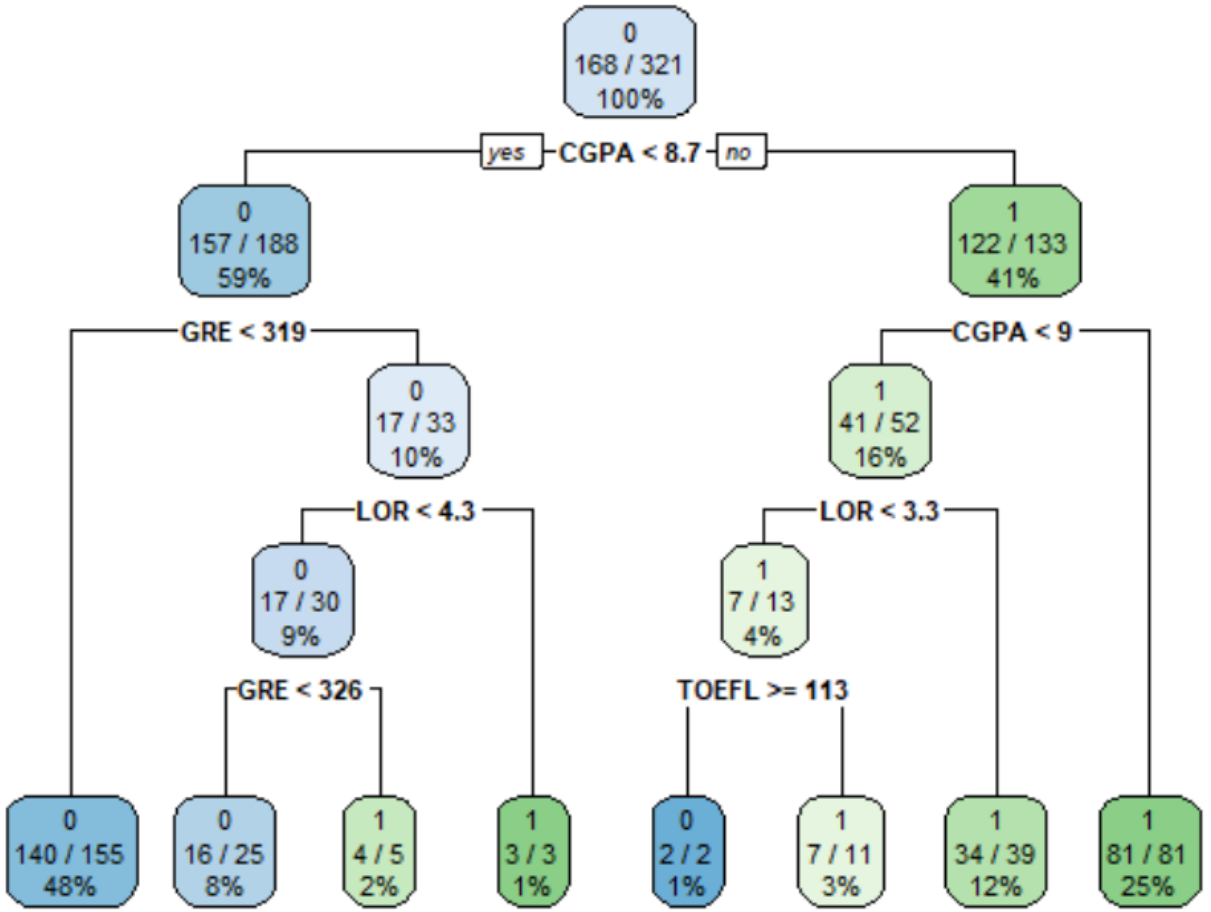
Classification/Regression : RandomForest(rf)

Accuracy : 0.965 | Kappa : 0.9297



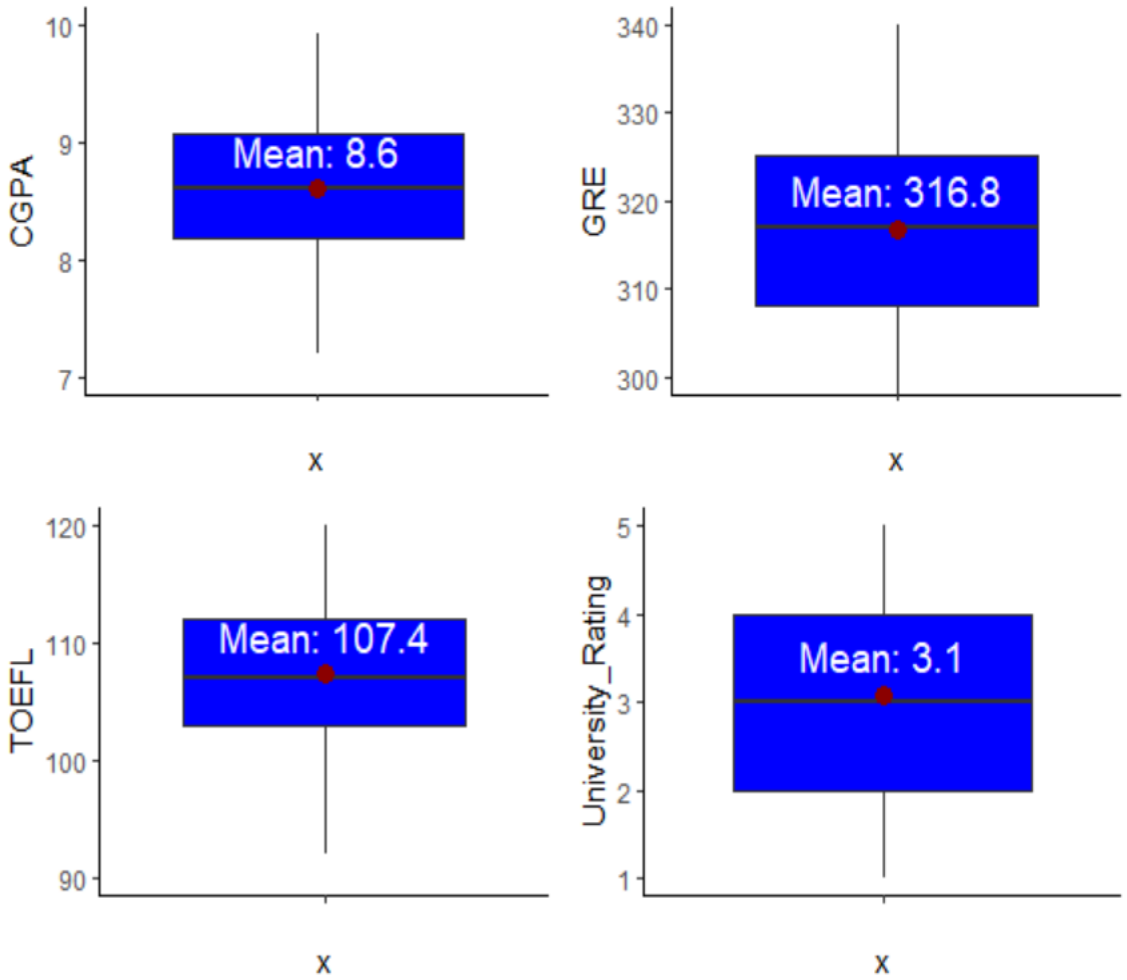
Ctree classification

Model Interpretation



COLLEGE ADMISSION

Box Plot



Random Forest Output

#Let's Answer the question (.90) through the model

GRE	TOEFL	University_Rating	SOP	LOR	CGPA	Research
330	115	5	4.5	3	9.34	1

Random Forest Regression

predict(model_fit_rf, df_pred1)

Answer: .90 (90% chance to get admission)

Random Forest classification

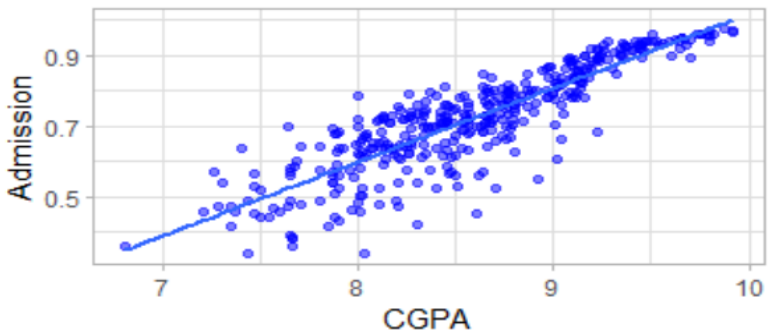
predict(model_fit_rf_class, df_pred1)

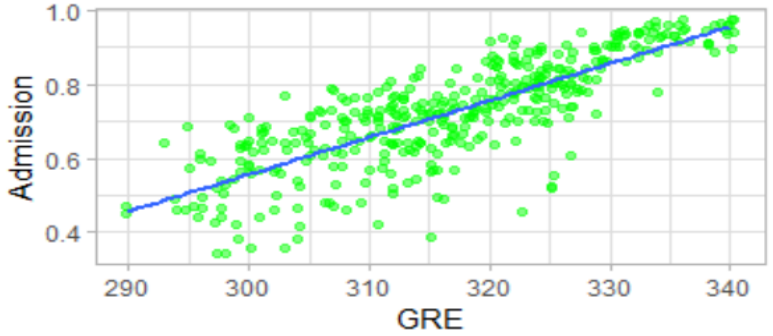
Answer: 1 (Student will get admission)

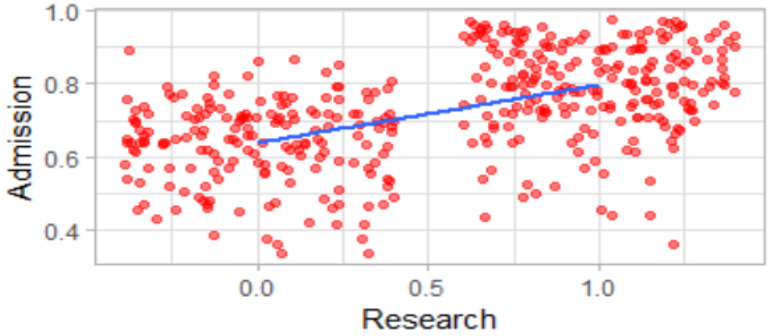
COLLEGE ADMISSION

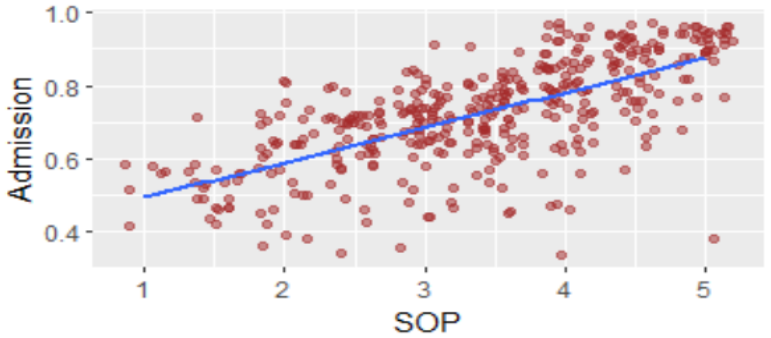
Additional Technical Details:

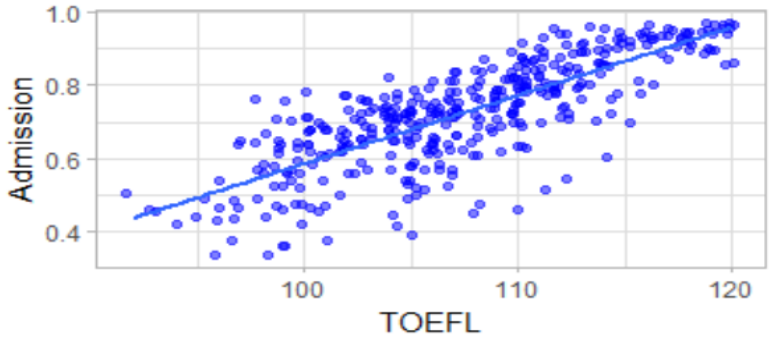
Admission Chance Vs Other Variables

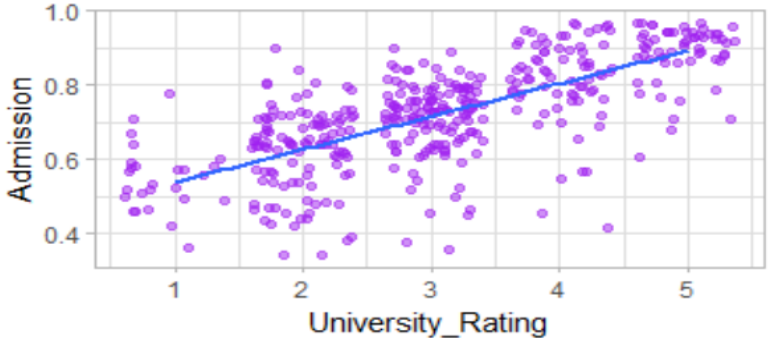


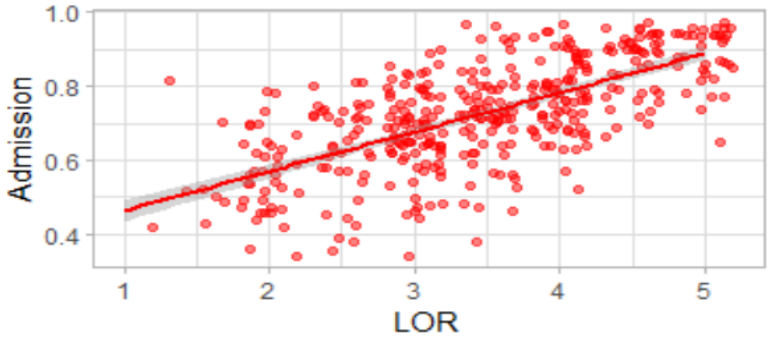










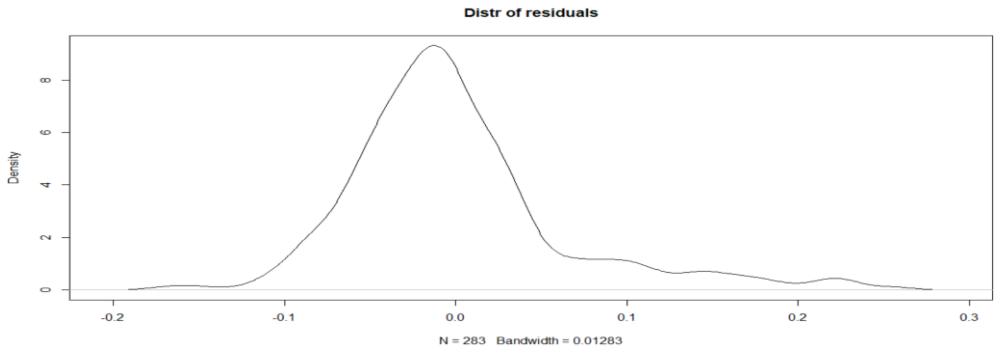


COLLEGE ADMISSION

Linear Regression

RMSE	Rsquared	MAE
0.06462765	0.8020547	0.04644549

Linear Regression of Residuals



Random Forest

mtry	RMSE	Rsquared	MAE
1	0.06357185	0.8122486	0.04733647
2	0.06127051	0.8178095	0.04459781
3	0.06144112	0.8150052	0.04454838
4	0.06219371	0.8096908	0.04522954
6	0.06360797	0.7995191	0.04622727

RMSE was used to select the optimal model using the smallest value.
The final value used for the model was mtry = 2

Random Forest Variable Importance

	Overall
CGPA	100.000
GRE	51.350
TOEFL	44.121
University_Rating	19.629
SOP	9.038
LOR	7.082
Research	0.000

COLLEGE ADMISSION

Summary

GRE	TOEFL	University_Rating	SOP
Min. :290.0	Min. : 92.0	Min. :1.000	Min. :1.0
1st Qu.:308.0	1st Qu.:103.0	1st Qu.:2.000	1st Qu.:2.5
Median :317.0	Median :107.0	Median :3.000	Median :3.5
Mean :316.8	Mean :107.4	Mean :3.087	Mean :3.4
3rd Qu.:325.0	3rd Qu.:112.0	3rd Qu.:4.000	3rd Qu.:4.0
Max. :340.0	Max. :120.0	Max. :5.000	Max. :5.0

LOR	CGPA	Research	Chance_of_Admit
Min. :1.000	Min. :6.800	Min. :0.0000	Min. :0.3400
1st Qu.:3.000	1st Qu.:8.170	1st Qu.:0.0000	1st Qu.:0.6400
Median :3.500	Median :8.610	Median :1.0000	Median :0.7300
Mean :3.453	Mean :8.599	Mean :0.5475	Mean :0.7244
3rd Qu.:4.000	3rd Qu.:9.062	3rd Qu.:1.0000	3rd Qu.:0.8300
Max. :5.000	Max. :9.920	Max. :1.0000	Max. :0.9700

Confusion Matrix and Statistics

Confusion Matrix and Statistics

	Reference	
Prediction	**0**	**1**
0	**207**	**12**
1	**2**	**179**

Accuracy : 0.965

95% CI : (0.942, 0.9807)

No Information Rate : 0.5225

P-Value [Acc > NIR] : < 2e-16

Kappa : 0.9297