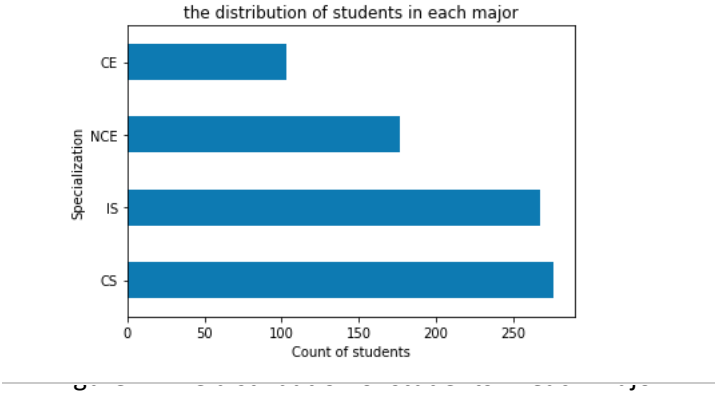
Predicting Student Academic in Advance

This experiment was conducted on a sample of students estimated at 824 students with 28 attributes. The number of attributes has been shortened to the number we need to reach the results. The university was also contacted and requested a larger number of data for the purpose of studying it. Students vary in their specializations between
(Computer Science, Computer Engineering, Information Systems, and Network and Communications Engineering) as shown in Figure 1



We tried to divide the students in these majors according to their academic level, then focused on the students at risk only because they are the most in need of help. At King Khalid University, a student at risk can be defined as a student who have passed the regular period of study, and graduated with a cumulative average of less than or equal to 2.75 out of 5 (with a good average) and who have exceeded the normal period of study of 10 semesters, or who have a cumulative average of less than 2.75. Therefore, this sample was sorted and then applied two algorithms, namely *Random Forest* and *Decision Tree*, to find out the extent of its potential in predicting students' academic performance. The preliminary results can be observed as follows:

	precision	recall	f1-score	support
0	0.79	0.70	0.74	74
1	0.78	0.85	0.81	91
accuracy			0.78	165
macro avg	0.78	0.77	0.78	165
weighted avg	0.78	0.78	0.78	165

Figure 2: Rndom Forest result

	precision	recall	f1-score	support
0	0.62	0.72	0.67	74
1	0.74	0.65	0.69	91
accuracy			0.68	165
macro avg	0.68	0.68	0.68	165
weighted avg	0.69	0.68	0.68	165

Figure 2: Decision tree result

Initially, by focusing on the F1 score that creates a balance between precision and recall, we find that random forests give the best result than the decision tree. Therefore, we believe that it is better than a decision tree in predicting at-risk students who represent Class 0.