NOTES ON DATA EXPLORATION AND ANALYSIS

The Open University UK Learning Analytics Dataset (OULAD) contains data about roughly 30 thousand students participating in 7 selected courses (called modules) over 2 years (4 semesters in total) and their interactions with Virtual Learning Environment (VLE).

The dataset and more detailed information can be found on:
https://analyse.kmi.open.ac.uk/open_dataset

Kuzilek J., Hlosta M., Zdrahal Z. Open University Learning Analytics dataset Sci. Data 4:170171 doi: 10.1038/sdata.2017.171 (2017).

The goal of this research is to accurately predict student performance. Specifically, whether they fail, pass or pass with distinction. Traditional machine learning models such as support vector machines, naïve Bayes classifiers, random forests and logistic regression are most commonly used for this purpose. I plan on assessing the performance of long short-term memory (LSTM) networks, a specific type of deep learning model, for this task in comparison to the traditional models mentioned previously.

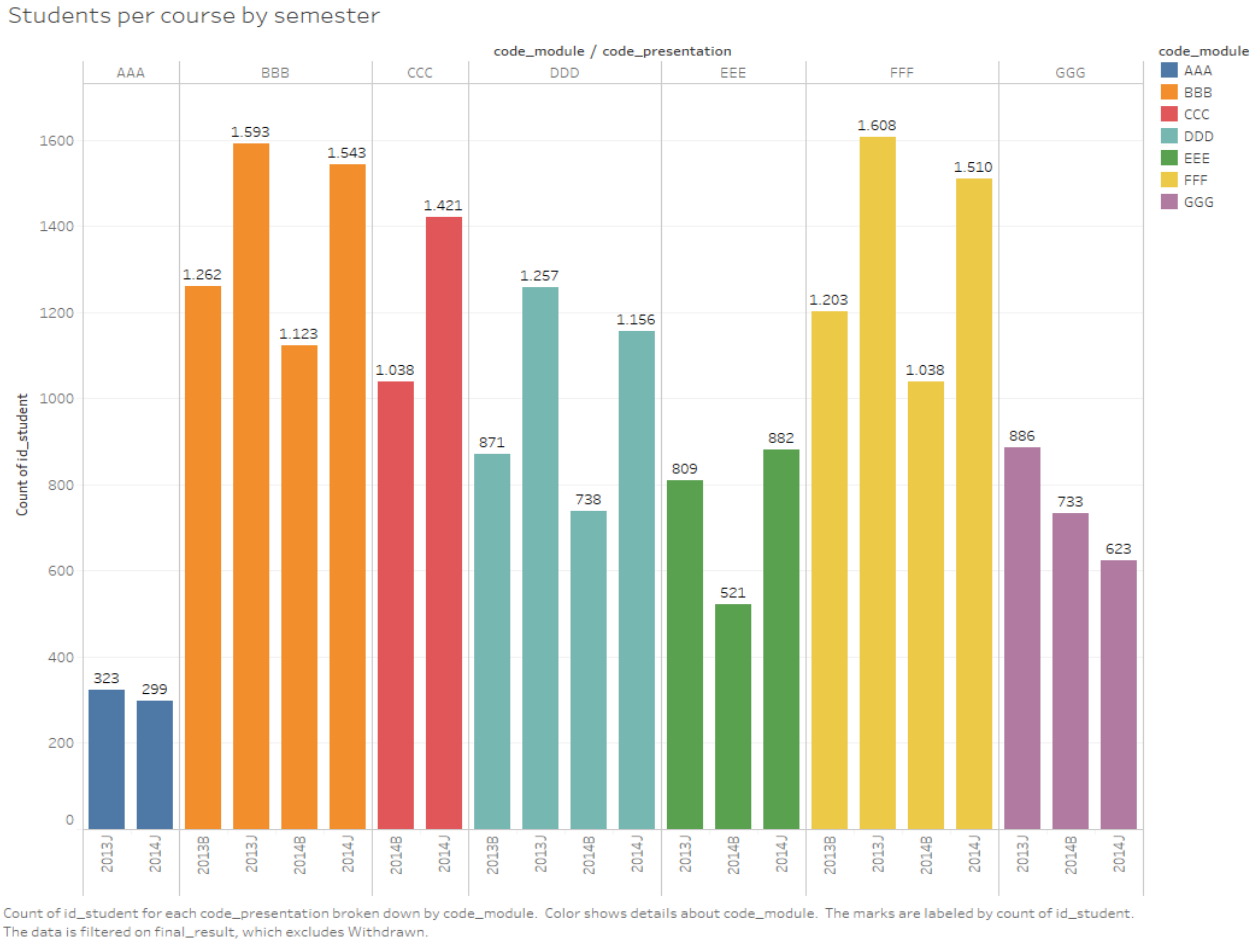
I hypothesize LSTM networks will prove to be more accurate for two reasons. First, in contrast to the traditional classifiers, LSTM networks are designed to take into account a time-dimension. Because of the presence of a time component in university courses that students participate in, LSTMs can be expected to better identify the existing patterns in the data, leading to a better performance. Second, deep learning models are known for their ability to recognize and capture complex non-linear relationships in data, which might also prove valuable when predicting student performance.

DATA EXPLORATION

First, it is important to explore the data that will be analyzed.

We only focus on students that did not drop out of the course, because they do not possess information on their final result.

The figure below shows the number of students per course per semester where “B” illustrates that the course starts in February while “J” represents courses starting in October.



Because the analysis warrants a considerable sample size, we only include the courses that are taught (or have data) spanning over all 4 semesters. Therefore, we will exclusively focus on courses BBB, DDD and FFF.

The tables below lists the assignments in the three courses over time in addition to their respective weights and dates. It should also be taken into account that each course has a different duration according to the year and semester. Each course starts at day 0 and ends after the number of days that is reported in the last row of the table. Note that the dates of some of the final exams are not reported in the dataset. However, the official documentation on the website states that the exam is at the end of the last course week if the information about the final exam date is missing. Therefore, I impute the final course-day as date for the final exam whenever the exam date is missing.

Course BBB Assignments:

2013B Asgmts	Date	Weight	2013J Asgmts	Date	Weight	2014B Asgmts	Date	Weight	2014J Asgmts	Date	Weight
TMA 1	19	5	**TMA 1**	19	5	**TMA 1**	12	5	**TMA 1**	19	0
TMA 2	47	18	**TMA 2**	47	18	**TMA 2**	40	18	**TMA 2**	54	10
CMA 1	54	1	**CMA 1**	54	1	**CMA 1**	47	1	**TMA 3**	110	20
TMA 3	89	18	**TMA 3**	96	18	**TMA 3**	82	18	**TMA 4**	152	35
CMA 2	89	1	**CMA 2**	96	1	**CMA 2**	82	1	**TMA 5**	201	35
TMA 4	124	18	**TMA 4**	131	18	**TMA 4**	117	18	**Exam**	NA	100
CMA 3	124	1	**CMA 3**	131	1	**CMA 3**	117	1	**Durat-ion**	262 days	-
TMA 5	159	18	**TMA 5**	166	18	**TMA 5**	152	18			
CMA 4	159	1	**CMA 4**	166	1	**CMA 4**	152	1			
TMA 6	187	18	**TMA 6**	208	18	**TMA 6**	194	18			
CMA 5	187	1	**CMA 5**	208	1	**CMA 5**	194	1			
Exam	NA	100	**Exam**	NA	100	**Exam**	NA	100			
Durat-ion	240 days	-	**Durat-ion**	268 days	-	**Durat-ion**	234 days	-			

The table above shows that the structure of course BBB in terms of assignments has remained the same over the first 3 semesters. However, the assignment structure of the course changed significantly in the last semester. It will be interesting to examine to what extent a model trained on the first 3 semesters can predict student performance in the fourth semester in which the course has been changed. This structural change will most likely cause a lower performance from the models than when trained and applied to a course of which the structure remained the same (such as FFF), as overfitting will become a bigger obstacle. The patterns that the models find in the first 3 semesters will probably not generalize as well to the fourth semester, because of the structural differences.

Course DDD Assignments:

2013B Asgmts	Date	Weight	2013J Asgmts	Date	Weight	2014B Asgmts	Date	Weight	2014J Asgmts	Date	Weight
CMA 1	23	2	TMA 1	25	10	TMA 1	25	10	TMA 1	20	5
TMA 1	25	7.5	TMA 2	53	12.5	TMA 2	53	12.5	TMA 2	41	10
CMA 2	51	3	TMA 3	88	17.5	TMA 3	74	17.5	TMA 3	62	10
TMA 2	53	10	TMA 4	123	20	TMA 4	116	20	TMA 4	111	25
CMA 3	79	3	TMA 5	165	20	TMA 5	158	20	TMA 5	146	25
TMA 3	81	12.5	TMA 6	207	20	TMA 6	200	20	TMA 6	195	25
CMA 4	114	4	Exam	261	100	Exam	241	100	Exam	NA	100
TMA 4	116	15	Duration	261 days	-	Duration	241 days	-	Duration	262 days	-
CMA 5	149	4									
TMA 5	151	15									
CMA 6	170	3									
TMA 6	200	15									
CMA 7	206	6									
Exam	240	100									
Duration	240 days	-									

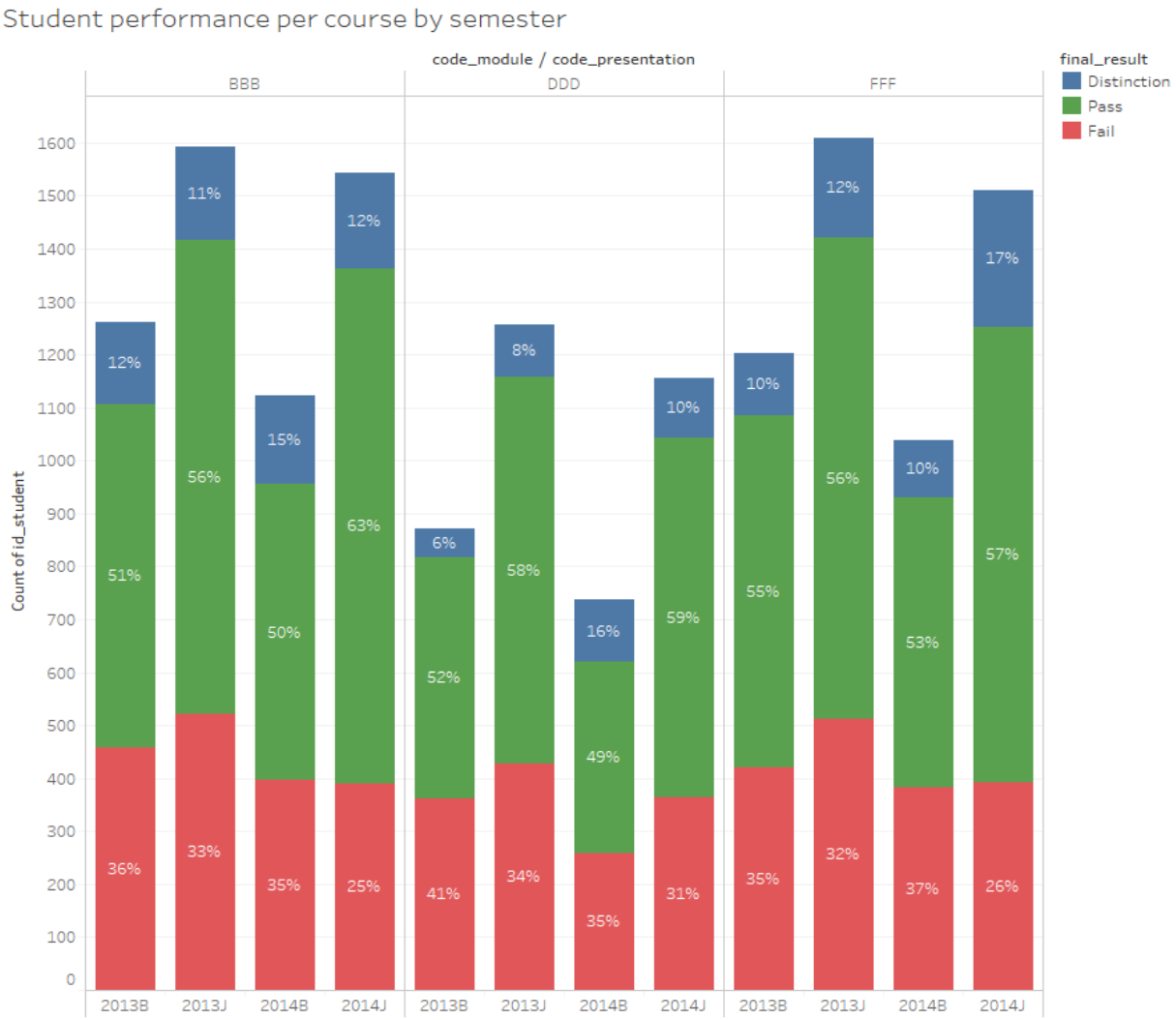
The table above shows that the structure of course DDD in terms of assignments changed significantly in the second semesters and slightly in the fourth semester (in terms of weightings). This will likely come at the costs of predictive performance when the first 3 periods are used to train the model in contrast to a course without structural changes (course FFF). However, the models will likely be more accurate than in the case of course BBB.

Course FFF Assignments

2013B Asgmts	Date	Weight	2013J Asgmts	Date	Weight	2014B Asgmts	Date	Weight	2014J Asgmts	Date	Weight
TMA 1	19	12.5	TMA 1	19	12.5	TMA 1	24	12.5	TMA 1	24	12.5
TMA 2	47	12.5	TMA 2	47	12.5	TMA 2	52	12.5	TMA 2	52	12.5
TMA 3	89	25	TMA 3	96	25	TMA 3	87	25	TMA 3	94	25
TMA 4	131	25	TMA 4	131	25	TMA 4	129	25	TMA 4	136	25
TMA 5	166	25	TMA 5	173	25	TMA 5	171	25	TMA 5	199	25
CMA 1	222	0	CMA 1	236	0	CMA 1	227	0	CMA 1	241	0
CMA 2	222	0	CMA 2	236	0	CMA 2	227	0	CMA 2	241	0
CMA 3	222	0	CMA 3	236	0	CMA 3	227	0	CMA 3	241	0
CMA 4	222	0	CMA 4	236	0	CMA 4	227	0	CMA 4	241	0
CMA 5	222	0	CMA 5	236	0	CMA 5	227	0	CMA 5	241	0
CMA 6	222	0	CMA 6	236	0	CMA 6	227	0	CMA 6	241	0
CMA 7	222	0	CMA 7	236	0	CMA 7	227	0	CMA 7	241	0
Exam	222	100	Exam	236	100	Exam	227	100	Exam	241	100
Duration	240 days	-	Duration	268 days	-	Duration	241 days	-	Duration	269 days	-

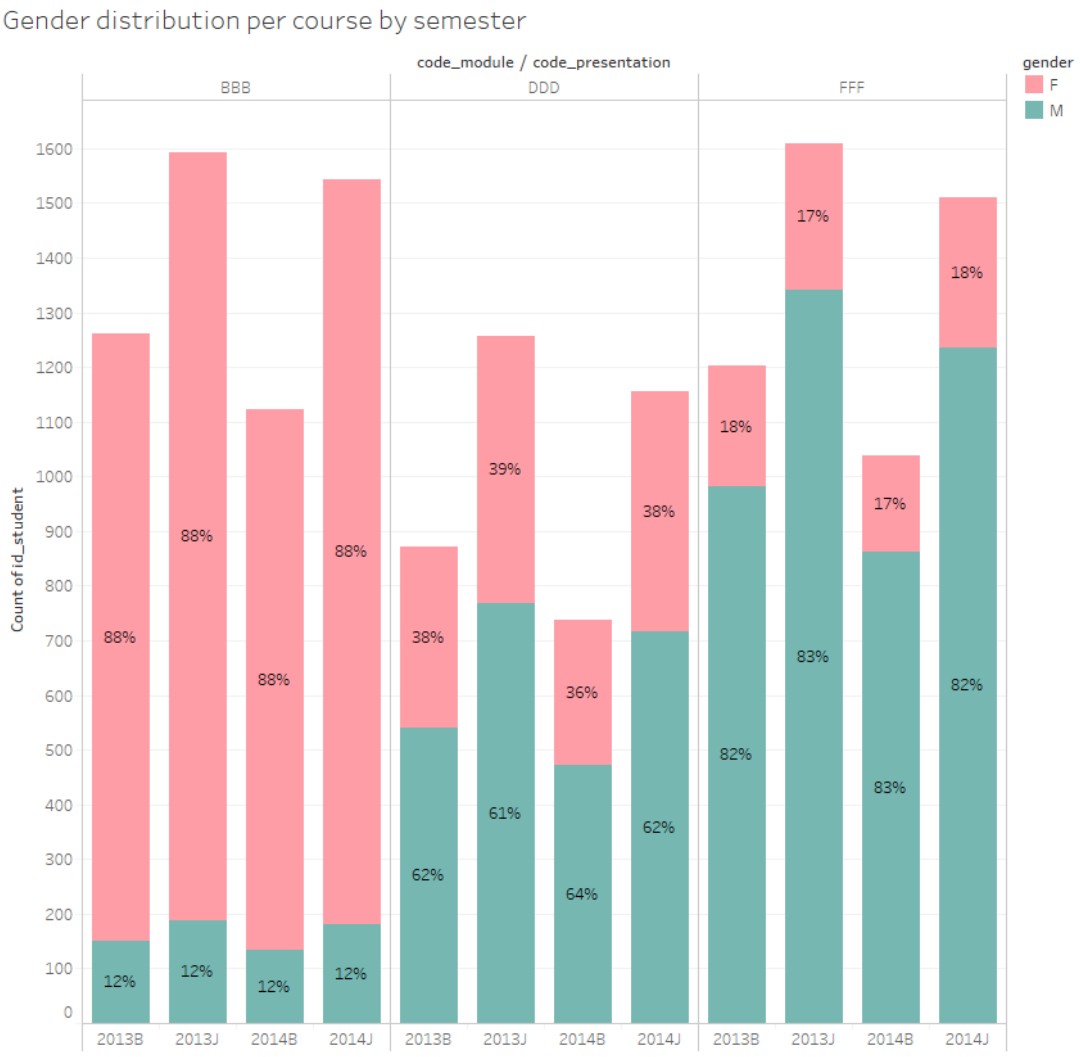
No structural changes in terms of assignments can be found in course FFF as shown in the table above. This makes course FFF the ideal candidate for model-building and an interesting control group to test the effect of structural changes in terms of assignments.

The next graph breaks down the student performance (e.g. the distribution of fail, pass and pass with distinction) per course by semester. From the graph can be inferred that for each course in every semester the majority of the students passes. Additionally, the proportion of students that fails the course is always bigger than the proportion that passes with distinction. It is interesting to note that for all three courses, the last semester has a higher proportion of students passing relative to student failing compared to the previous three semesters. This might indicate a change in school or grading policy. However, the exact reason is difficult to ascertain because there are only 4 semesters to glean insights from.



Count of id_student for each code_presentation broken down by code_module. Color shows details about final_result. The marks are labeled by % of Total Count of id_student. The view is filtered on final_result and code_module. The final_result filter excludes Withdrawn. The code_module filter keeps BBB, DDD and FFF.

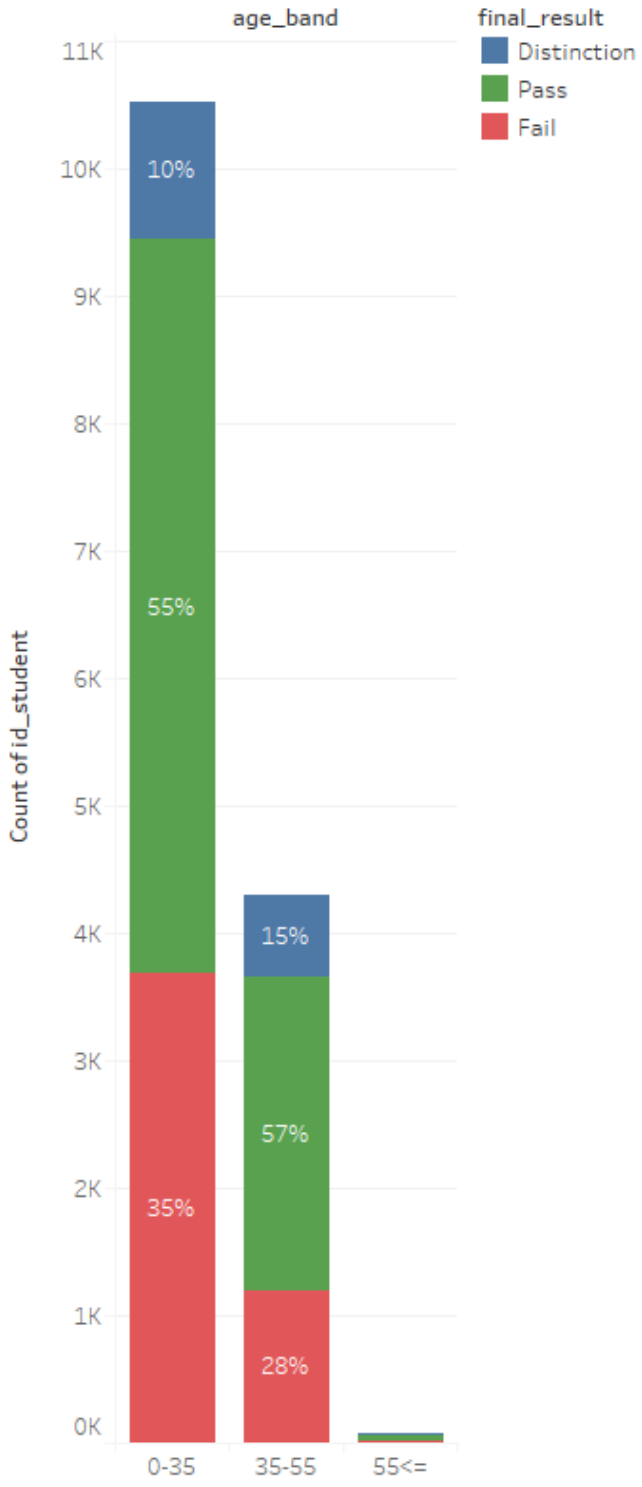
The graph below illustrates the gender distribution per course for each semester. Course BBB has a relatively large share of female students while course FFF has a predominantly large share of male students.



Count of id_student for each code_presentation broken down by code_module. Color shows details about gender. The marks are labeled by % of Total Count of id_student. The data is filtered on final_result, which excludes Withdrawn. The view is filtered on code_module, which keeps BBB, DDD and FFF.

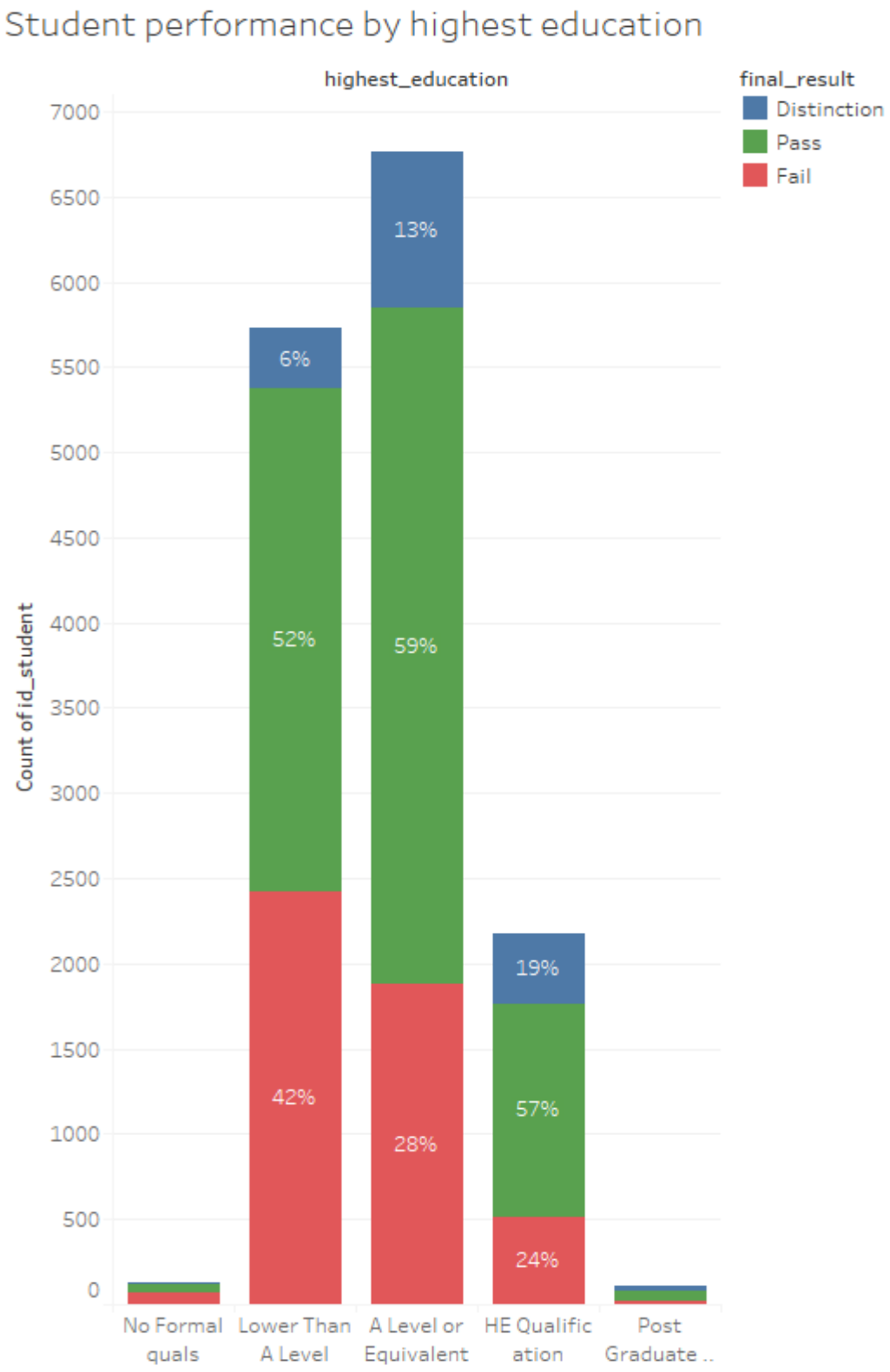
Most students in the dataset are below the age of 55 as can be seen in the graph below. The graph also shows that older students perform better on average than younger students.

Student Performance by age group



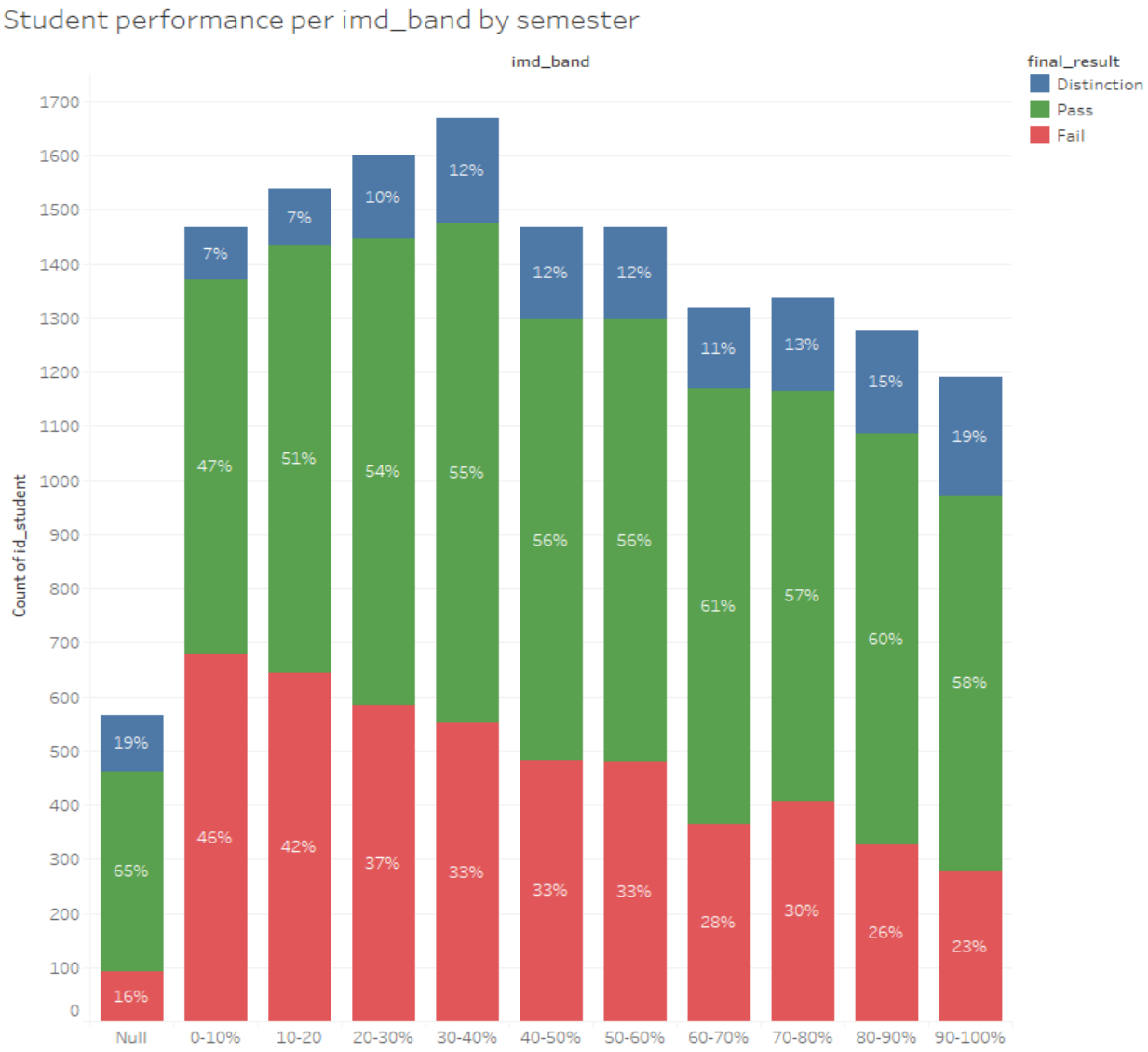
Count of id_student for each age_band. Color shows details about final_result. The marks are labeled by % of Total Count of id_student. The data is filtered on code_module, which keeps BBB, DDD and FFF. The view is filtered on final_result, which excludes Withdrawn.

The next figure shows that students with a higher previous education level perform better on average.



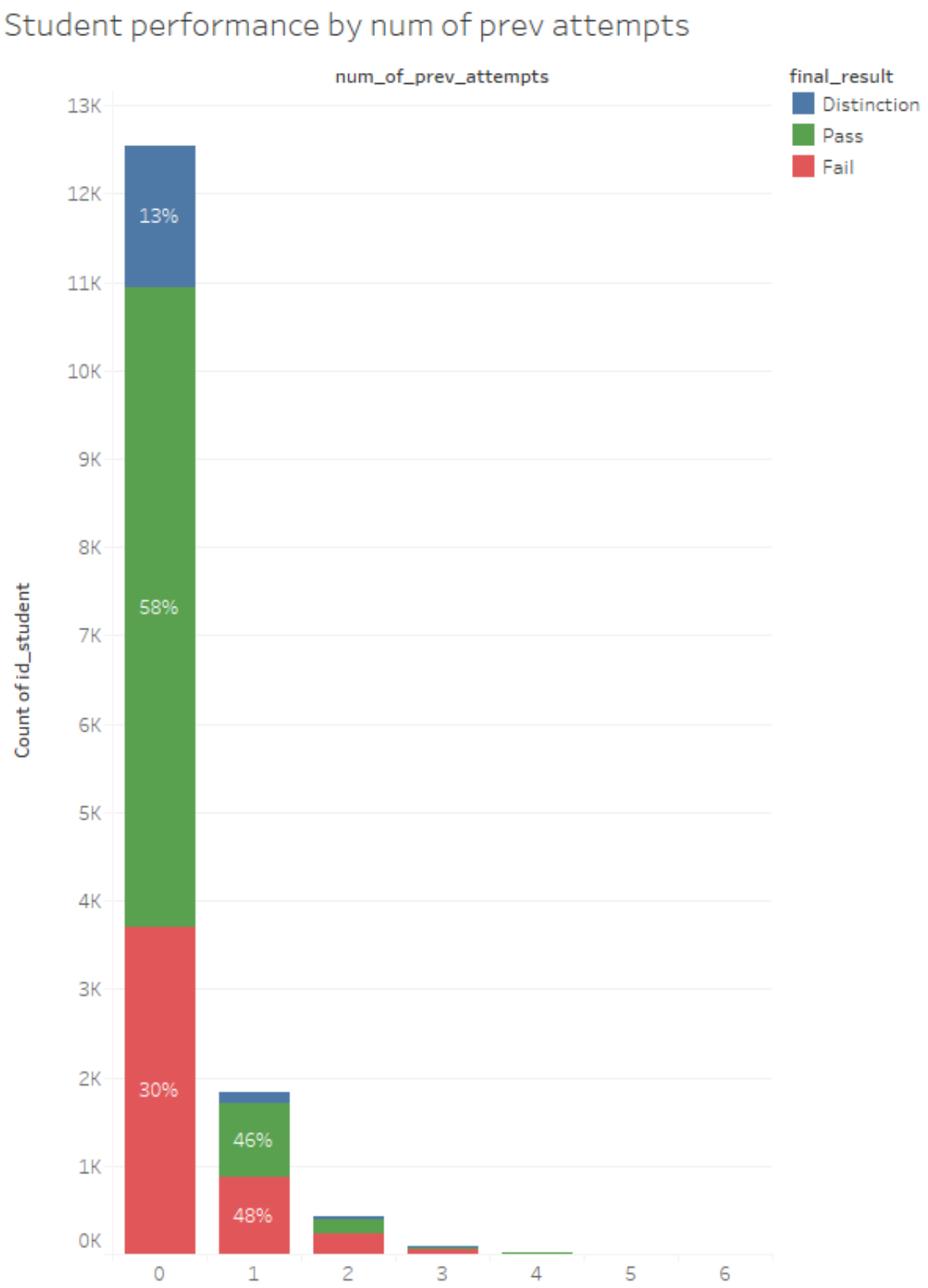
Count of id_student for each highest_education. Color shows details about final_result. The marks are labeled by % of Total Count of id_student. The data is filtered on code_module, which keeps BBB, DDD and FFF. The view is filtered on final_result, which excludes Withdrawn.

Another variable that could be of value when predicting student performance is the Index of Multiple Deprivation (IMD) which measures the deprivation of neighborhoods in the UK. The IMD-band of the place where the student lived during the course are presented in the dataset. The graph below indicates that on average the students that live in a more prosperous neighborhood (higher IMD-band) have a higher pass (with distinction) / fail ratio. Interestingly, the students that do not have any values reported for IMD-band have a surprisingly high pass (with distinction) / fail ratio. These results imply that IMD-band could carry significant predictive power for predicting student performance



Count of id_student for each imd_band. Color shows details about final_result. The marks are labeled by % of Total Count of id_student. The data is filtered on code_module, which keeps BBB, DDD and FFF. The view is filtered on final_result, which has multiple members selected.

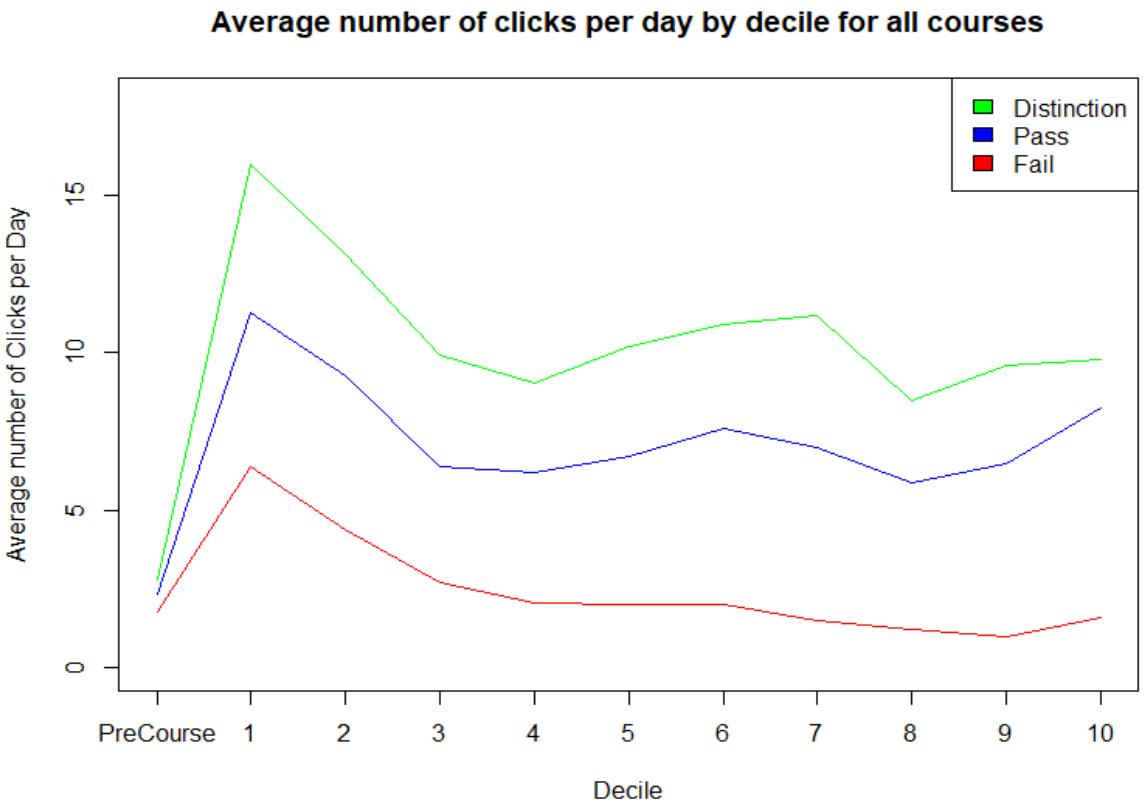
The next graph provides an indication of how many attempts student needed to finish the course. The higher the amount of attempts needed, the lower the pass/fail ratio.



Count of id_student for each num_of_prev_attempts. Color shows details about final_result. The marks are labeled by % of Total Count of id_student. The data is filtered on code_module, which keeps BBB, DDD and FFF. The view is filtered on final_result, which excludes Withdrawn.

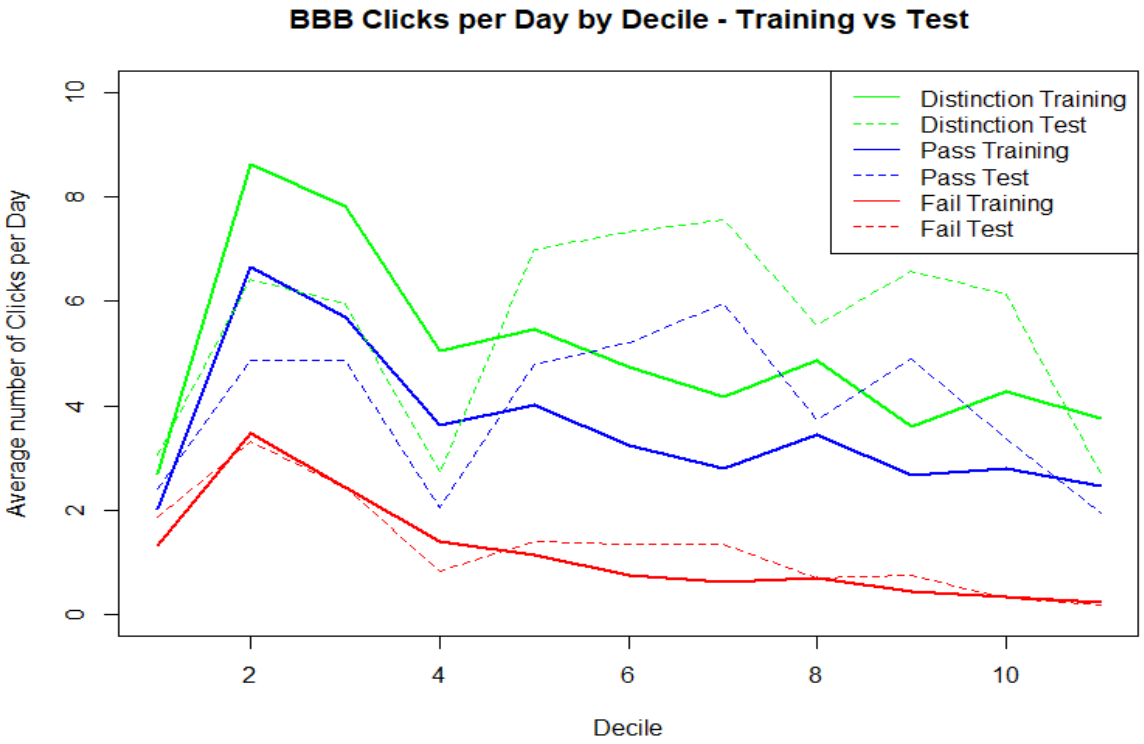
Besides from student background information, the dataset also contains information on how many times students clicked on certain elements in the virtual learning environment (VLE). The effort that a student puts into a course in terms of studying is arguably one of the most important predictors of student performance. However, effort is hard to measure. Universities that operate mostly online, such as the Open University are in a unique position to collect and analyze students' clicks in the online learning environment, which are expected to be a valuable proxy for student effort.

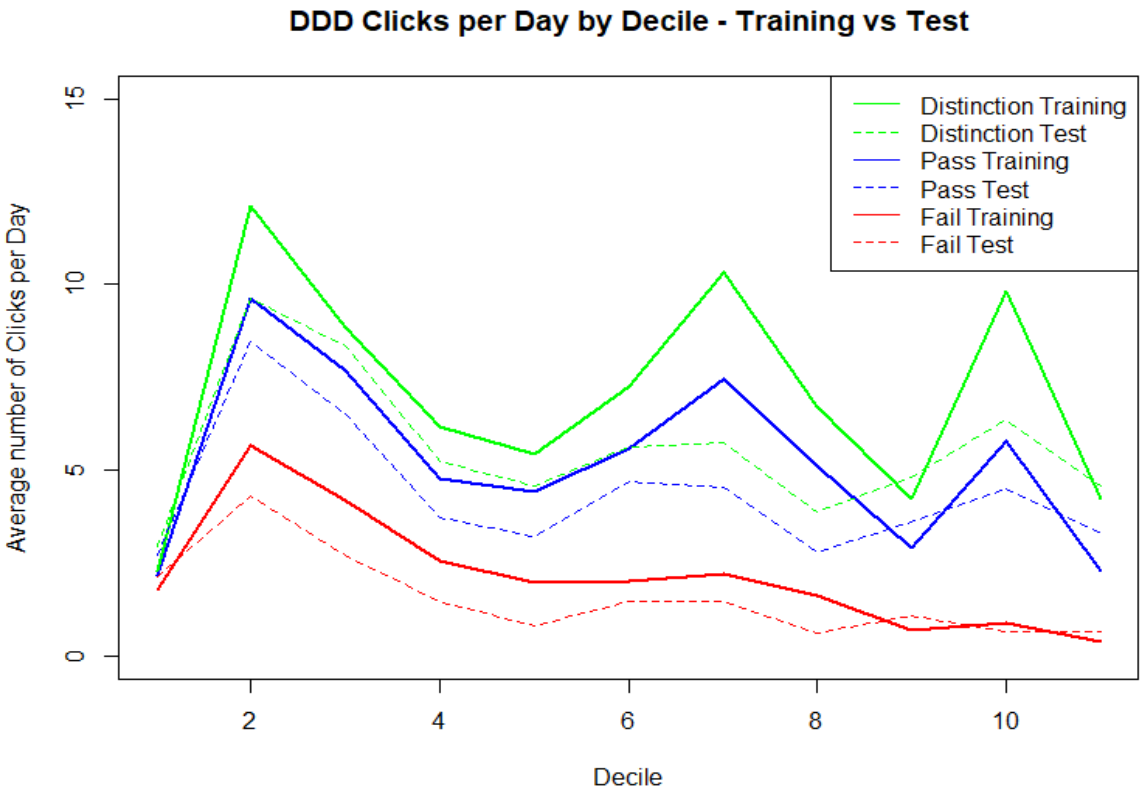
I split the duration of every course (the day before the final exam day) in each semester by 10 to create ten equal deciles each encompassing a range of days. Because the courses do not all have the same length, this relative division attempts to create relatively similar ranges of days. This resolves the problem of differing durations when the duration is split into months. This approach also makes it possible to extend the model to smaller or larger courses. I chose for 10-decile sequences because more could possibly introduce the "curse of dimensionality" where the amount of variables becomes too large relative to the sample size and the model starts to overfit. The clicks before the course starts are also recorded. I computed the average clicks per day for each decile and the period before the start of the course. This is shown in the graph below. This graph illustrates that the number of clicks clearly correlate with student performance. Students that passed with distinction clicked the most times per day in the VLE, while the students that failed clicked on average the least times in the VLE. This provides a clear indication of the predictive power that clicks have for predicting student performance.

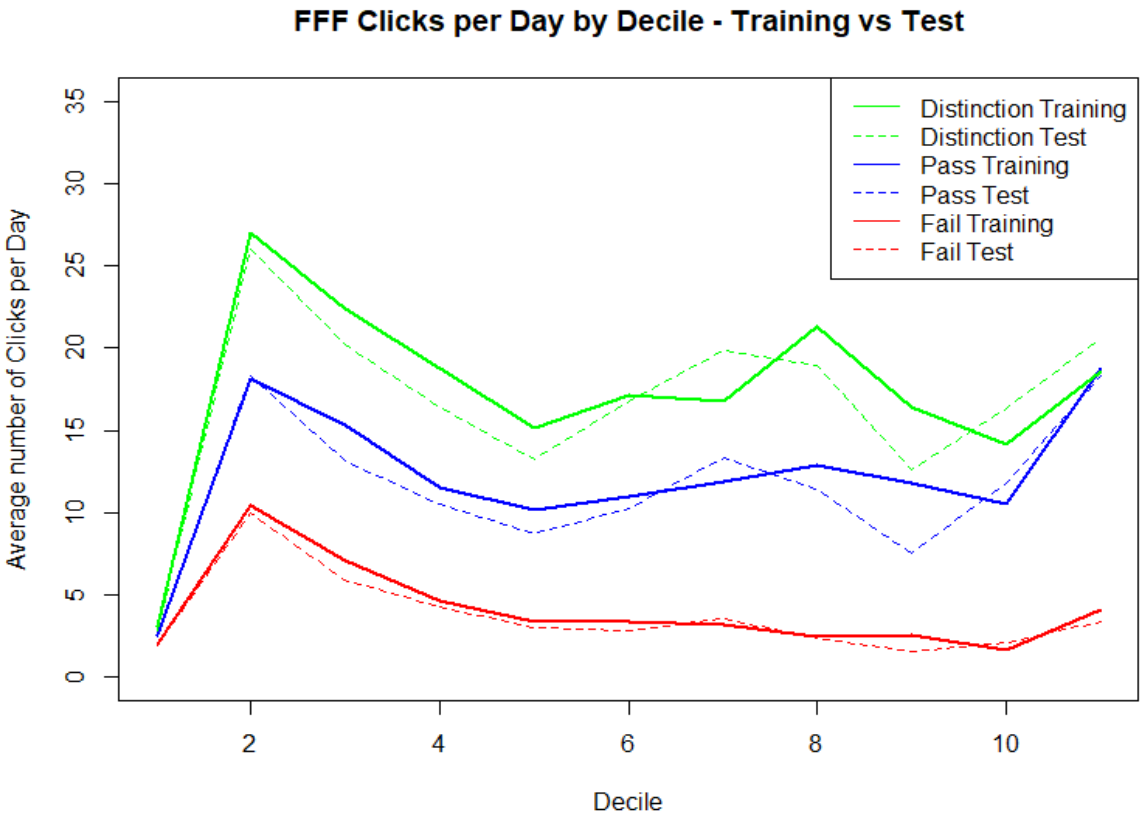


Because I will analyze three different courses, it might be insightful to look at the patterns in the number of clicks over time per course. Specifically, the difference between the training set (first 3 semesters) and the test set (last semester) will be important to examine, as two of the three courses endured structural changes in terms of assignments.

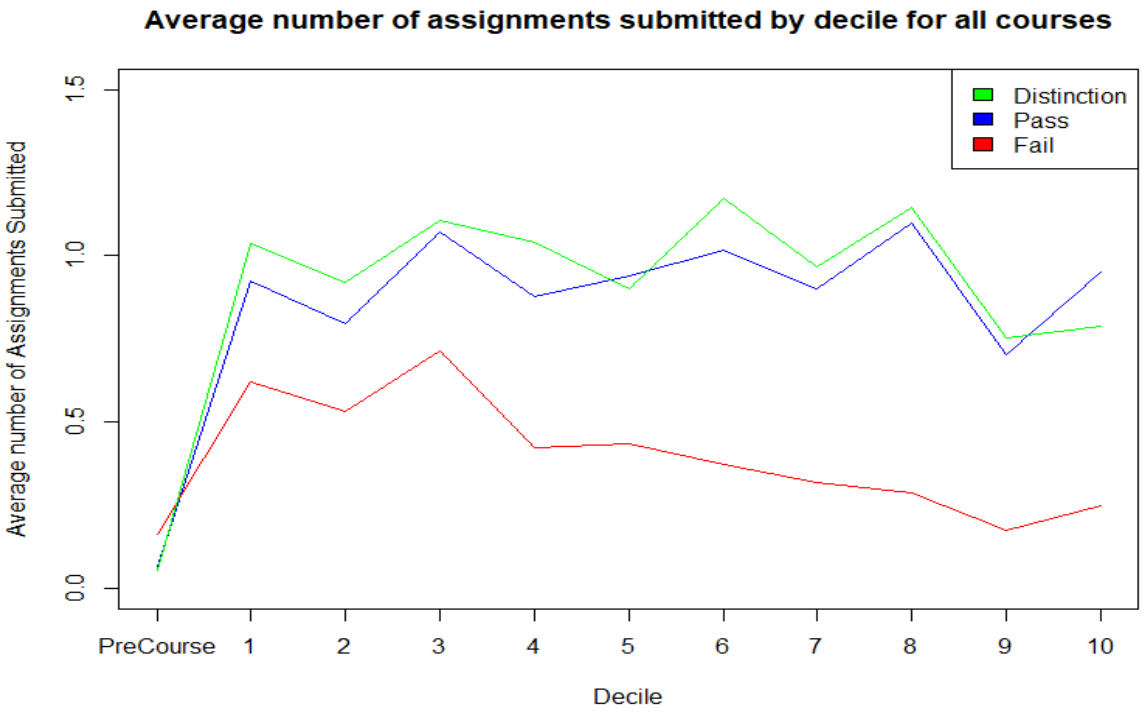
The next graph shows the average clicks per day in each decile for course BBB, split between test and training set. It can be seen that the pattern between the training and test sets are quite different. The same, but to a lesser extent, goes for the clicks made in the DDD course as shown in the subsequent graph. This is likely a consequence of the structural changes mentioned before, as the FFF course which did not endure any structural changes shows no significant changes in click patterns between the first three semesters and the last semester.

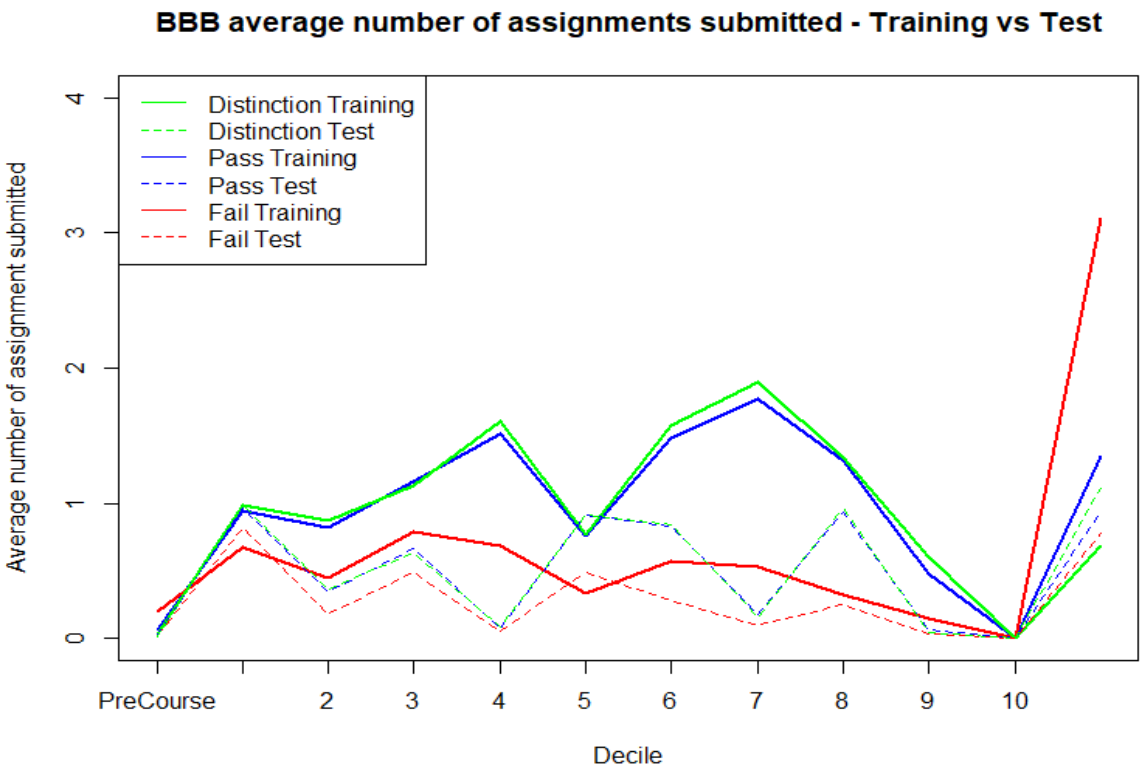


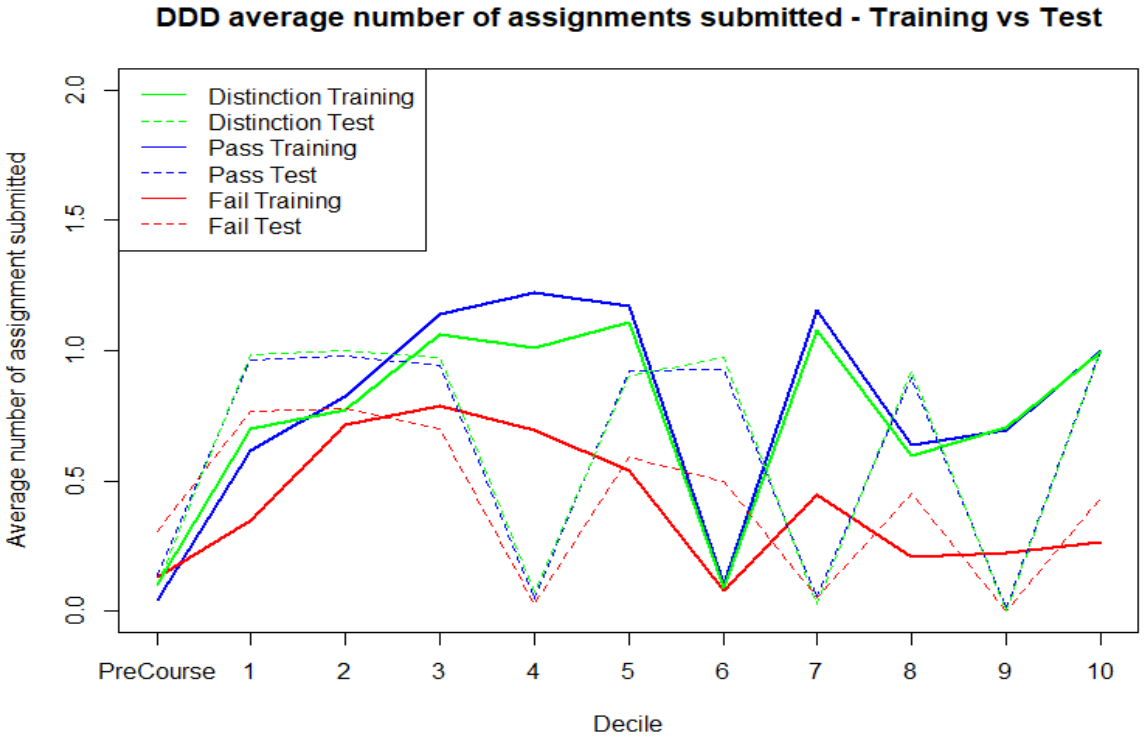

DDD Clicks per Day by Decile - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average number of Clicks per Day
0
5
10
15
2
4
6
8
10
Decile

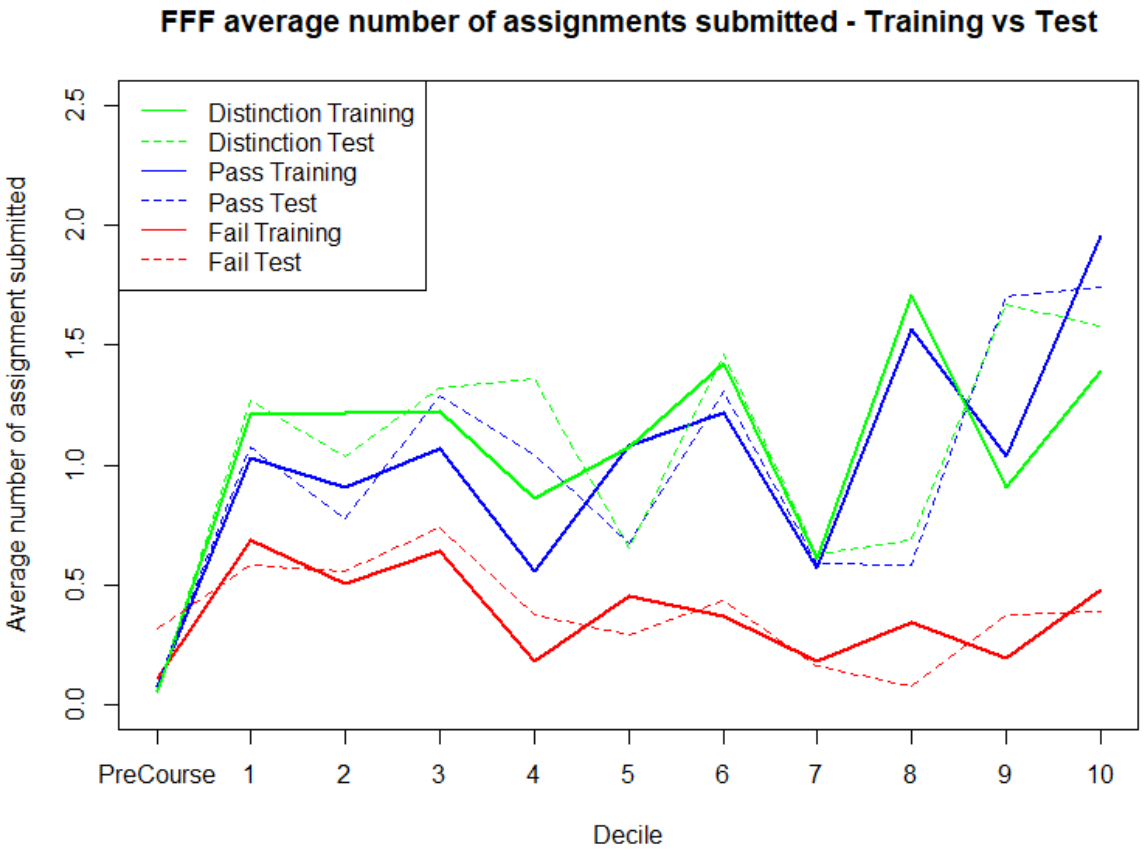

FFF Clicks per Day by Decile - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average number of Clicks per Day
0
5
10
15
20
25
30
35
2
4
6
8
10
Decile

The dataset also contains information on the assignments that students made as part of the course. I also computed the average number of assignments made in each decile for the 3 performance categories. From the next graph can be inferred that students that passed the course (with distinction) made on average more assignments before the final exam than student that failed the course. The 3 subsequent graphs show that the difference between assignment submission patterns are relatively more different between training and test set for course BBB and DDD than for FFF following the same logic as for clicks.

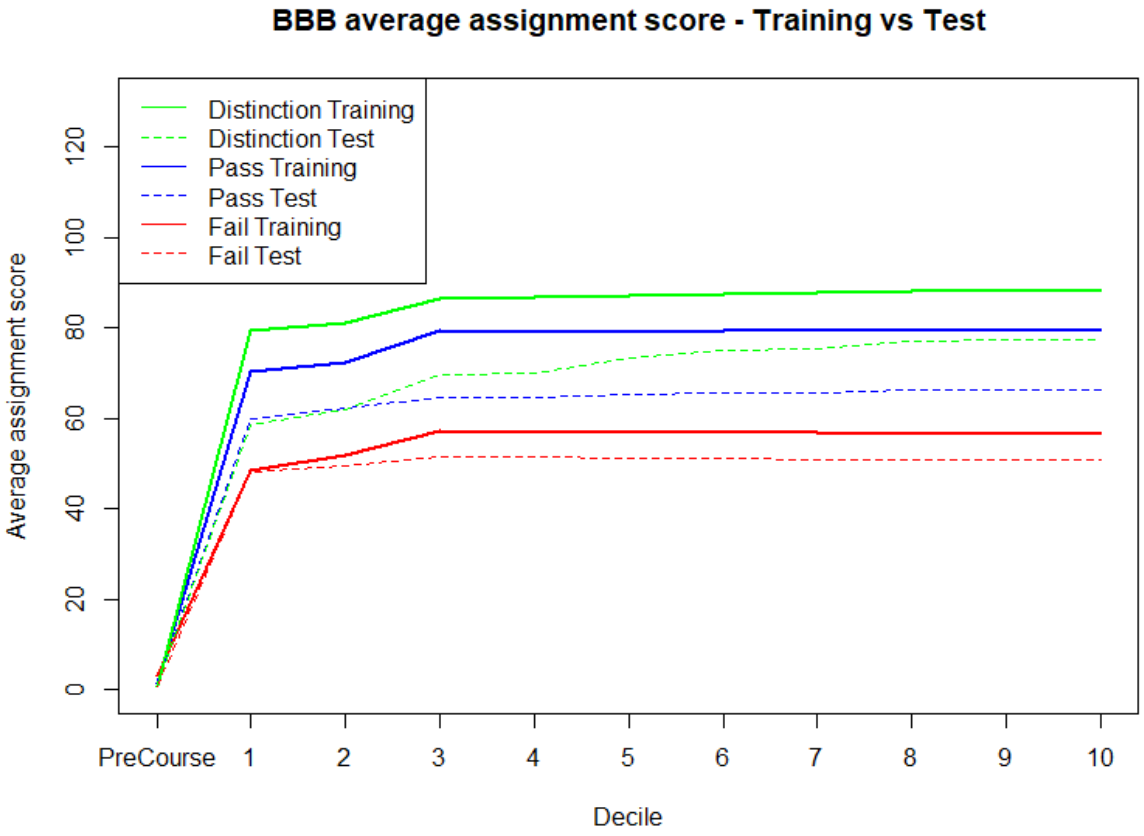


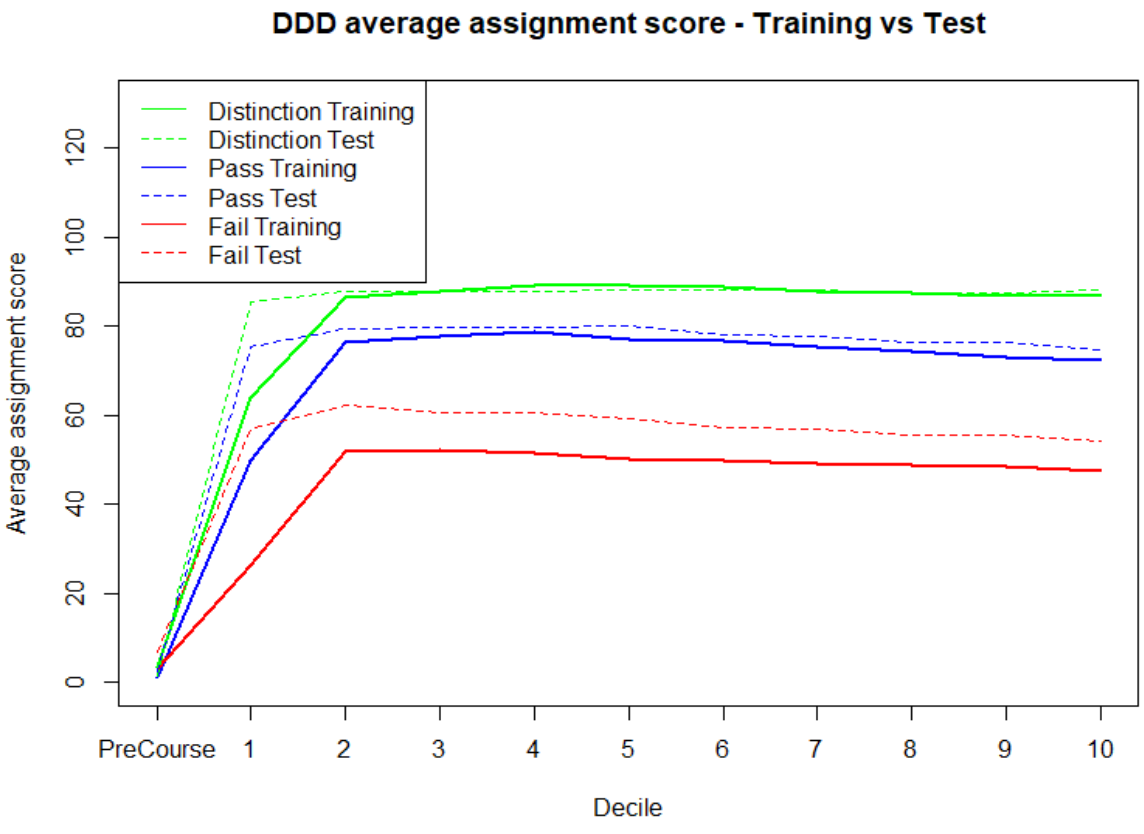

BBB average number of assignments submitted - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average number of assignment submitted
0
1
2
3
4
PreCourse
2
3
4
5
6
7
8
9
10
Decile

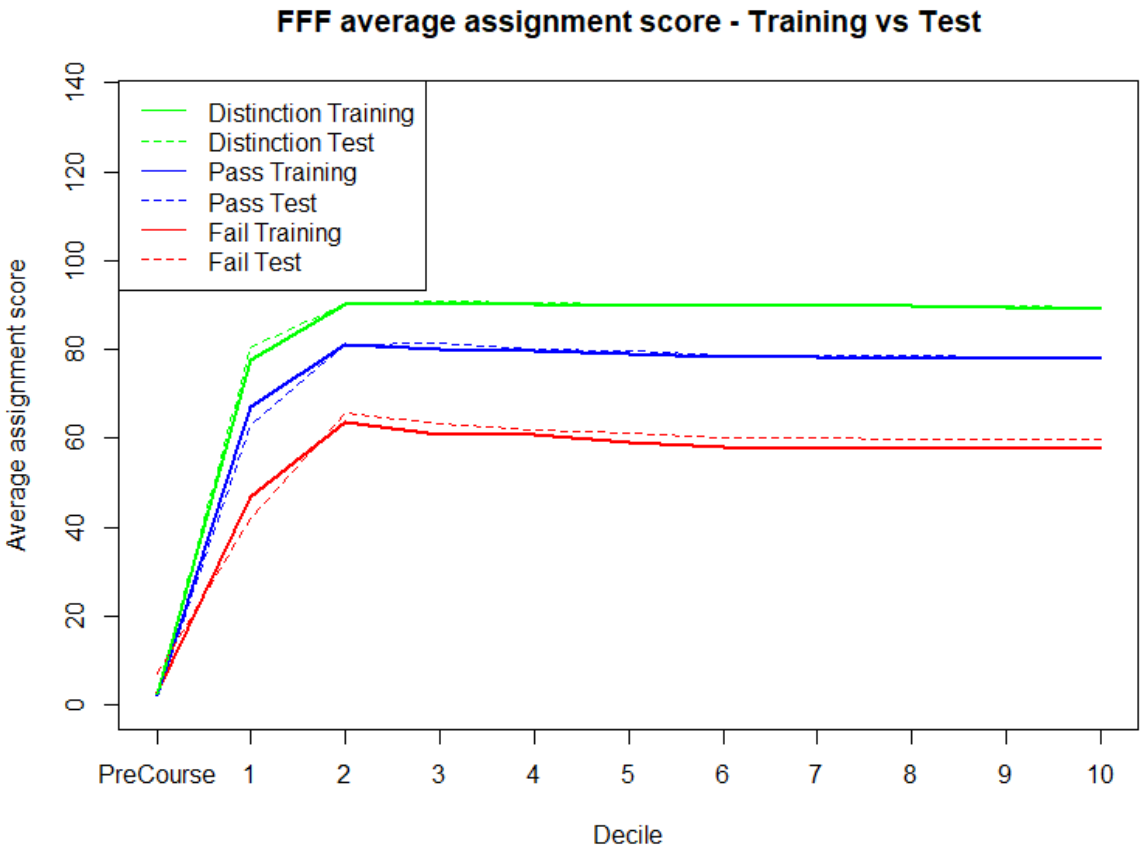

DDD average number of assignments submitted - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average number of assignment submitted
0.0
0.5
1.0
1.5
2.0
PreCourse
1
2
3
4
5
6
7
8
9
10
Decile



Below the progression of the average exam grade over time is displayed. The patterns stay relatively consistent across training and test sets.


BBB average assignment score - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average assignment score
0
20
40
60
80
100
120
PreCourse
1
2
3
4
5
6
7
8
9
10
Decile


DDD average assignment score - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average assignment score
0
20
40
60
80
100
120
PreCourse
1
2
3
4
5
6
7
8
9
10
Decile


FFF average assignment score - Training vs Test
Distinction Training
Distinction Test
Pass Training
Pass Test
Fail Training
Fail Test
Average assignment score
0
20
40
60
80
100
120
140
PreCourse
1
2
3
4
5
6
7
8
9
10
Decile

DATA ANALYSIS

Before I built the LSTM network, I developed a selection of traditional machine learning models in order to predict student performance. First, I will try to predict all three performance classes (fail, pass and pass with distinction). Except for the LSTM, I will try to predict this with a logistic regression, a support vector machine, a naive Bayes classifier and a random forest in order to see which model performs the best. Later on, I will decrease the target variables to just pass and fail by converting pass with distinction into a regular pass. Mainly, because predicting who will fail is often the most valuable goal and the difference between pass and pass with distinction is less pressing. Moreover, the accuracy is bound to increase when having to predict only two classes instead of three.

The variables used for the traditional machine learning models are:

Student Background	**Clicks**	**Assignments**
Gender	Avg clicks pre-course	Assignments pre-course
Highest education	Avg clicks decile 1	Assignments decile 1
IMD band	Avg clicks decile 2	Assignments decile 2
Age band	Avg clicks decile 3	Assignments decile 3
Num of previous attempts	Avg clicks decile 4	Assignments decile 4
Studied credits	Avg clicks decile 5	Assignments decile 5
Disability	Avg clicks decile 6	Assignments decile 6
Year	Avg clicks decile 7	Assignments decile 7
Semester	Avg clicks decile 8	Assignments decile 8
	Avg clicks decile 9	Assignments decile 9
	Avg clicks decile 10	Assignments decile 10
		Assignments score pre-course
		Assignments score decile 1
		Assignments score decile 2
		Assignments score decile 3
		Assignments score decile 4
		Assignments score decile 5
		Assignments score decile 6
		Assignments score decile 7
		Assignments score decile 8
		Assignments score decile 9
		Assignments score decile 10

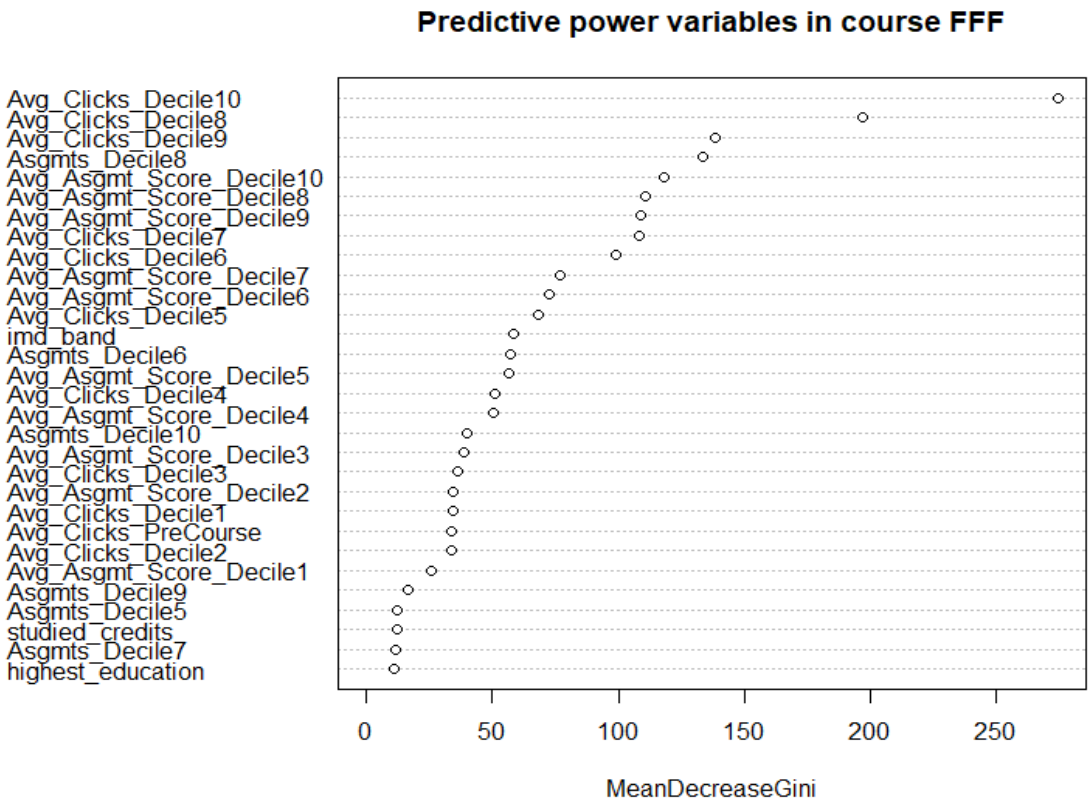
A new model is developed, trained and tested for every course and for every added decile.

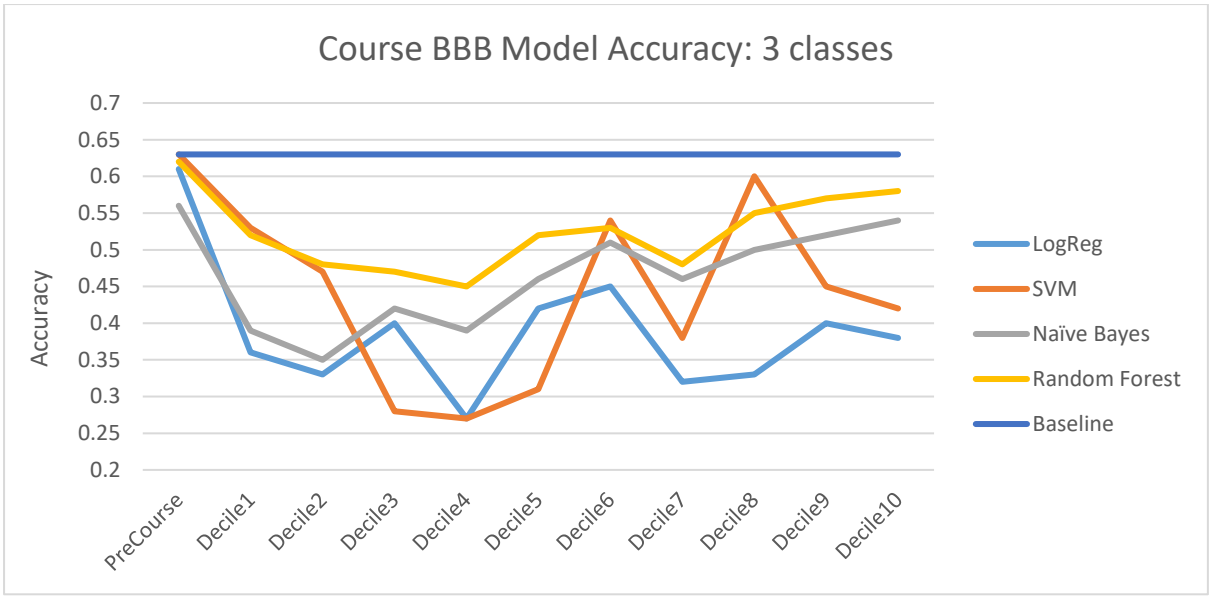
Predicting 3 classes (fail/pass/pass with distinction)

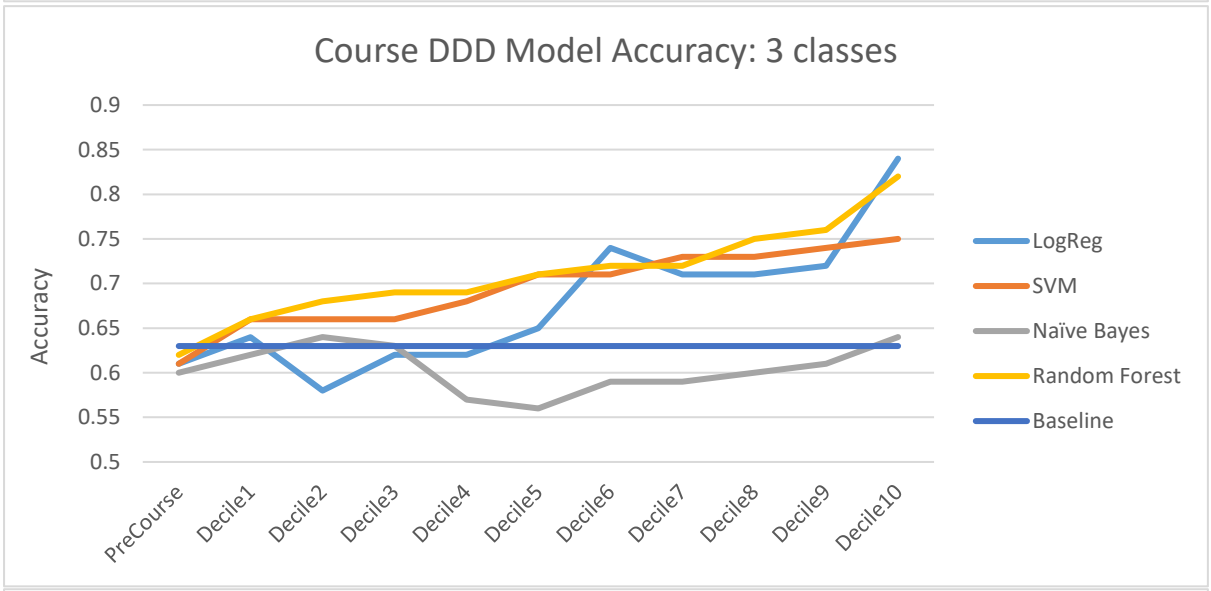
The graphs in the two pages below show the accuracy of each model for each course at different timestamps. Moreover, the "fail" sensitivity has been computed, showing how many of the people that actually failed the course were correctly predicted to fail by the model.

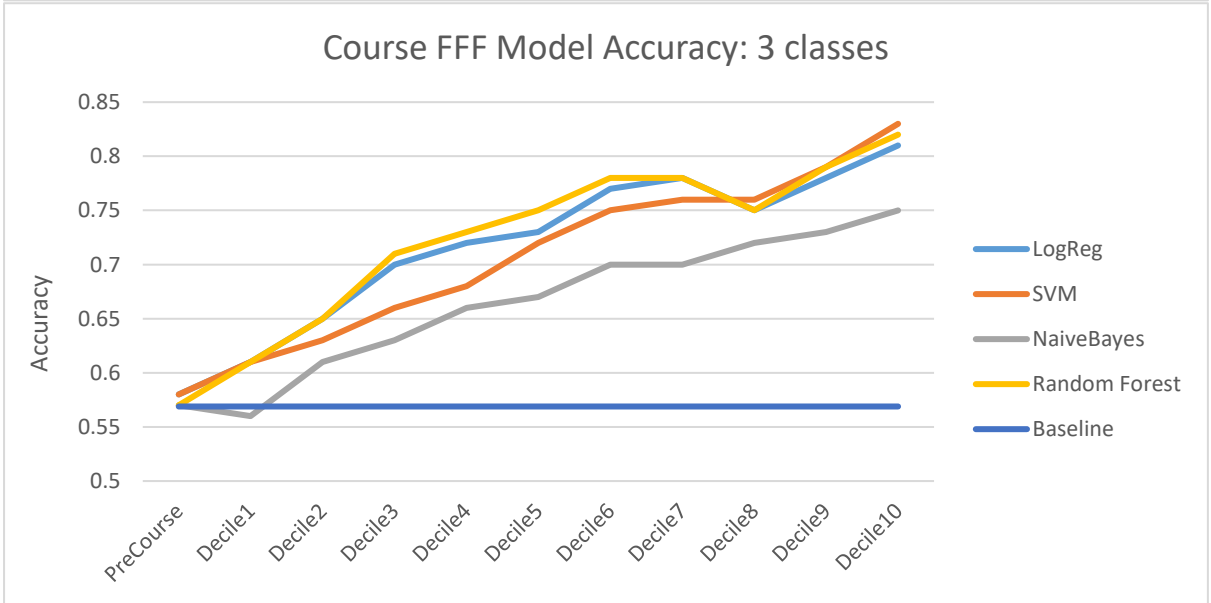
The graphs show that the models did not predict the performance of students in course BBB with much accuracy. They are all below the baseline in which all students are predicted to pass. This is because the models predict that the majority of the students will fail. This likely has to do with the structural changes in terms of assignments mentioned earlier. In one of the previous graphs can be seen that the average number of assignments submitted in the BBB course in the last semester (test set) was overall lower than in the first three semesters (training set). In fact, the number of assignments submitted by students that passed (with distinction) in the last semester was about as high as the number of assignments submitted by the students that failed the course in the first three semesters. This is expected to cause a disproportionate amount of false predictions of students that failed. This problem does not occur (to such an extent) in courses DDD and FFF where the accuracy is higher and fail sensitivity lower.

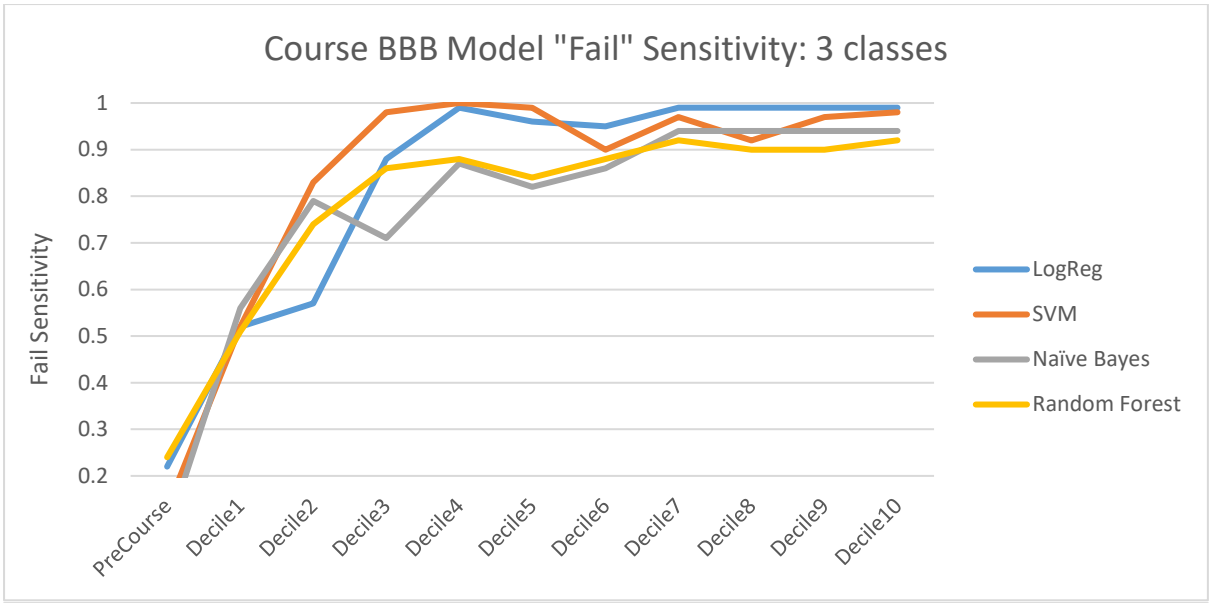
The random forest model in course FFF shows good potential as right before the exam it predict performance with 82% accuracy while predicting the students that failed with 86% precision right before the exam (in decile 10). It is interesting to examine which variables possess the most predictive power. For that purpose, I plotted the mean decrease in Gini-coefficients outputted by the random forest model in course FFF. This plot shows that the number of clicks in the last three deciles carried most of the predictive power of the model. A large portion of the most predictive variables were variables measured during the course. Therefore, the models get better as time goes on.

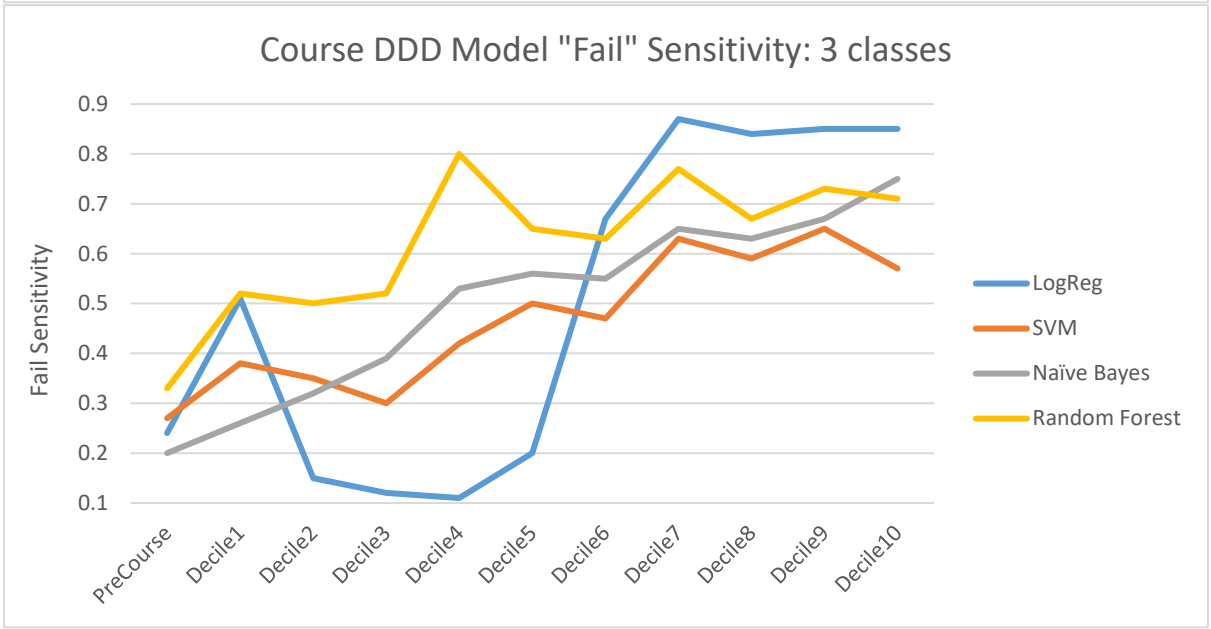


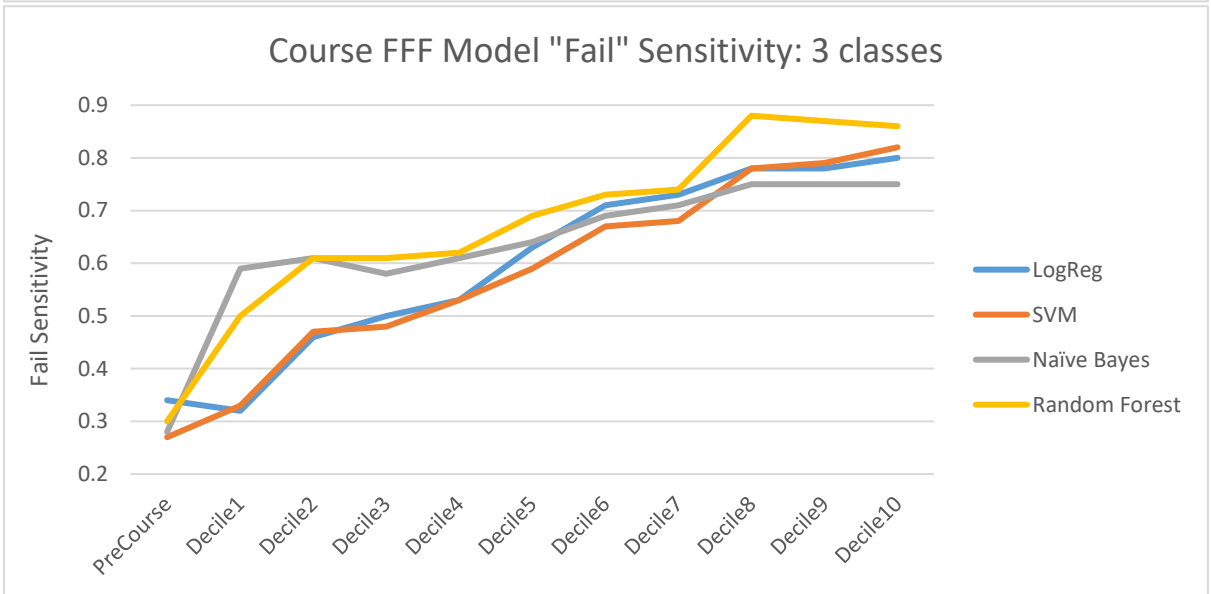

Course BBB Model Accuracy: 3 classes
Accuracy
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest
Baseline


Course DDD Model Accuracy: 3 classes
Accuracy
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest
Baseline


Course FFF Model Accuracy: 3 classes
Accuracy
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
NaiveBayes
Random Forest
Baseline

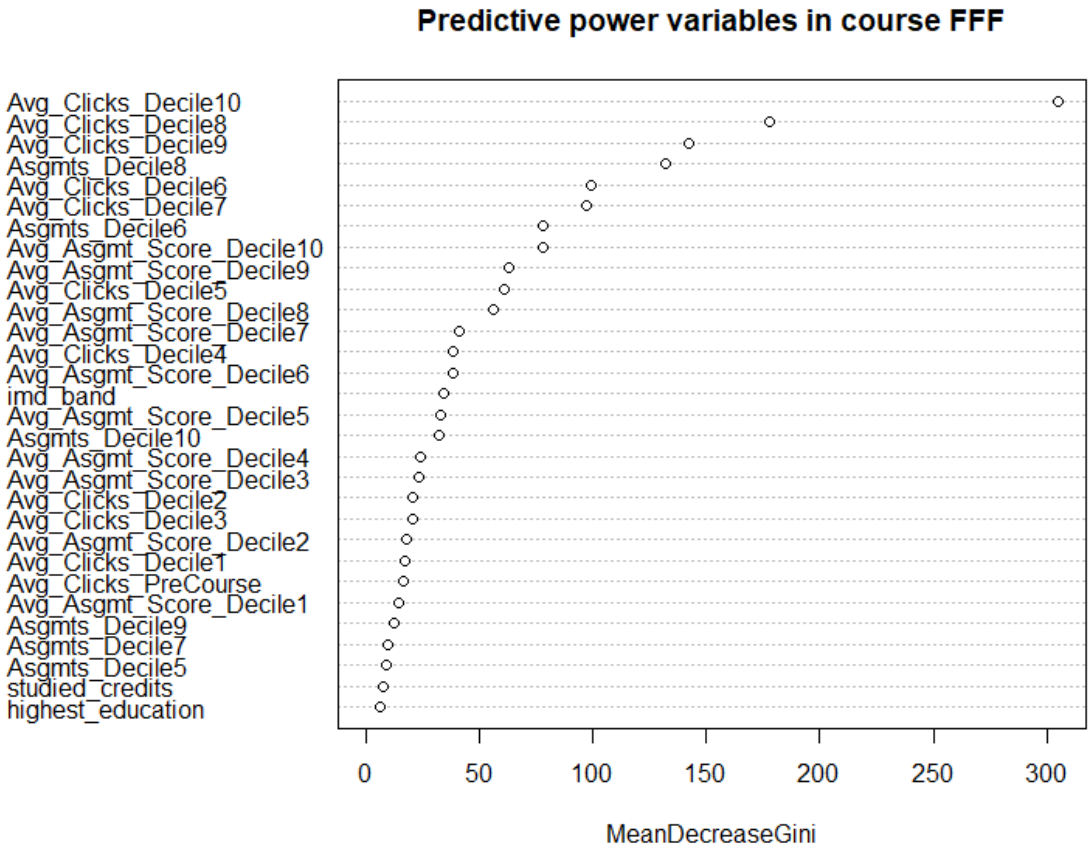

Course BBB Model "Fail" Sensitivity: 3 classes
Fail Sensitivity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest


Course DDD Model "Fail" Sensitivity: 3 classes
Fail Sensitivity
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest


Course FFF Model "Fail" Sensitivity: 3 classes
Fail Sensitivity
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest

Predicting 2 classes (fail/pass)

Next, we convert the “pass with distinction” target classes to the regular “pass” class, so we end up with just two classes to predict. As shown in the graphs on the following pages, this improves overall accuracy. However, we still see the same patterns as in the figures above. The random forest in course FFF now can predict student performance with 94% accuracy in decile 10, with a sensitivity measure of 86%. Meaning that 86% of all students that actually failed were correctly predicted to fail by the model. The mean decrease in Gini-coefficients is relatively similar to the random forest that predicted 3 classes and looks as follows for the model:

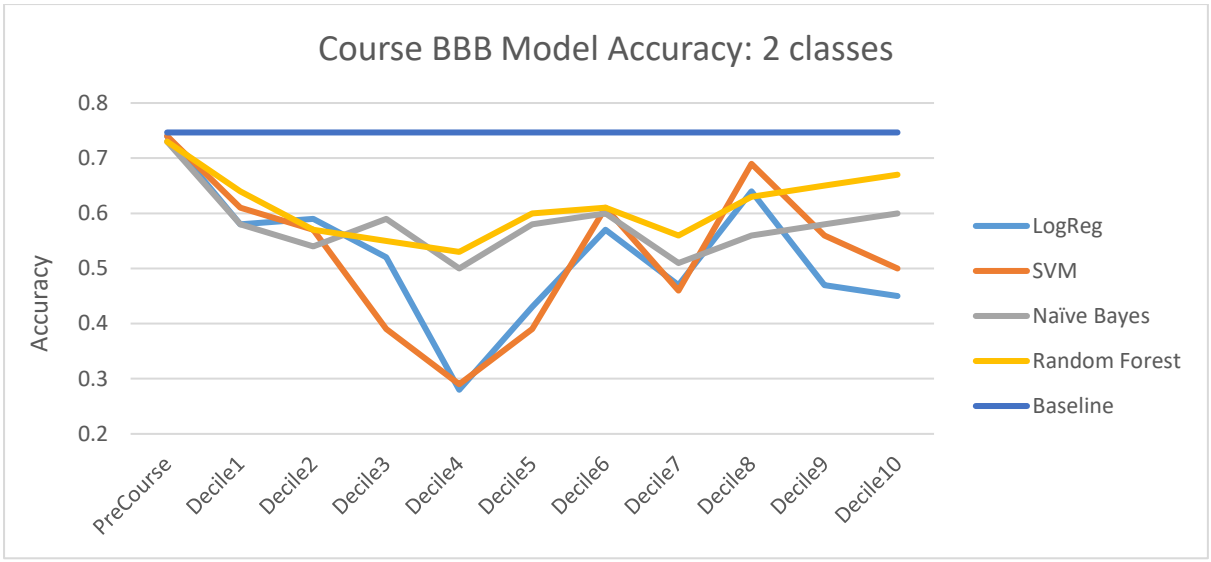


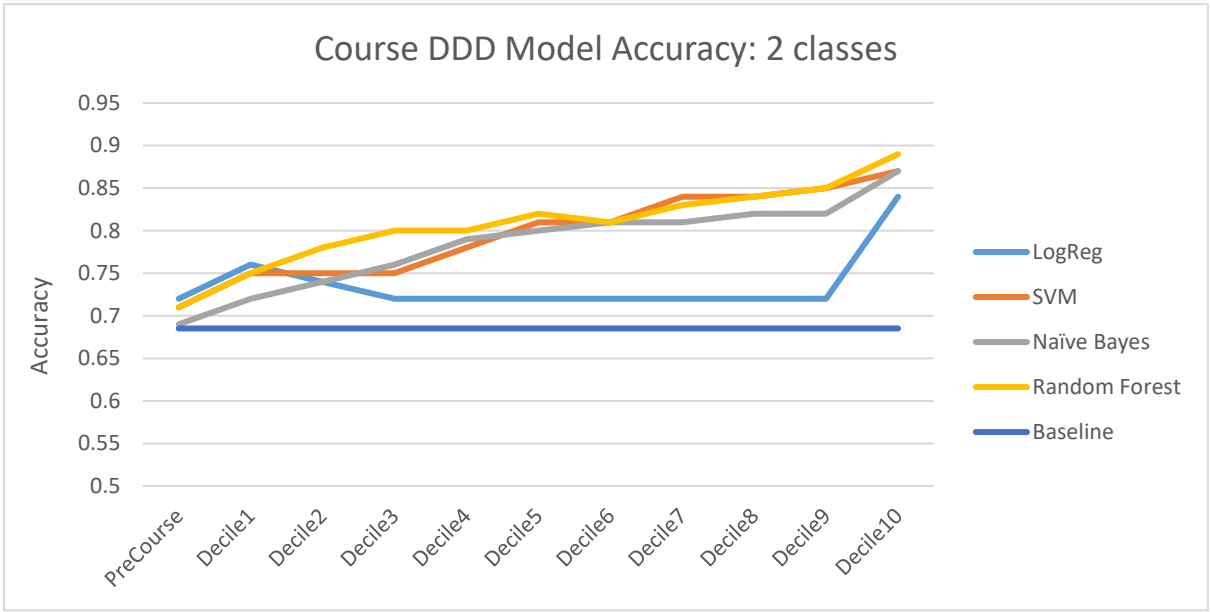
Besides the mean decrease in Gini-coefficients outputted by the random forest, it is also possible to use the logistic regression output to see whether the significant variables corroborate these findings.

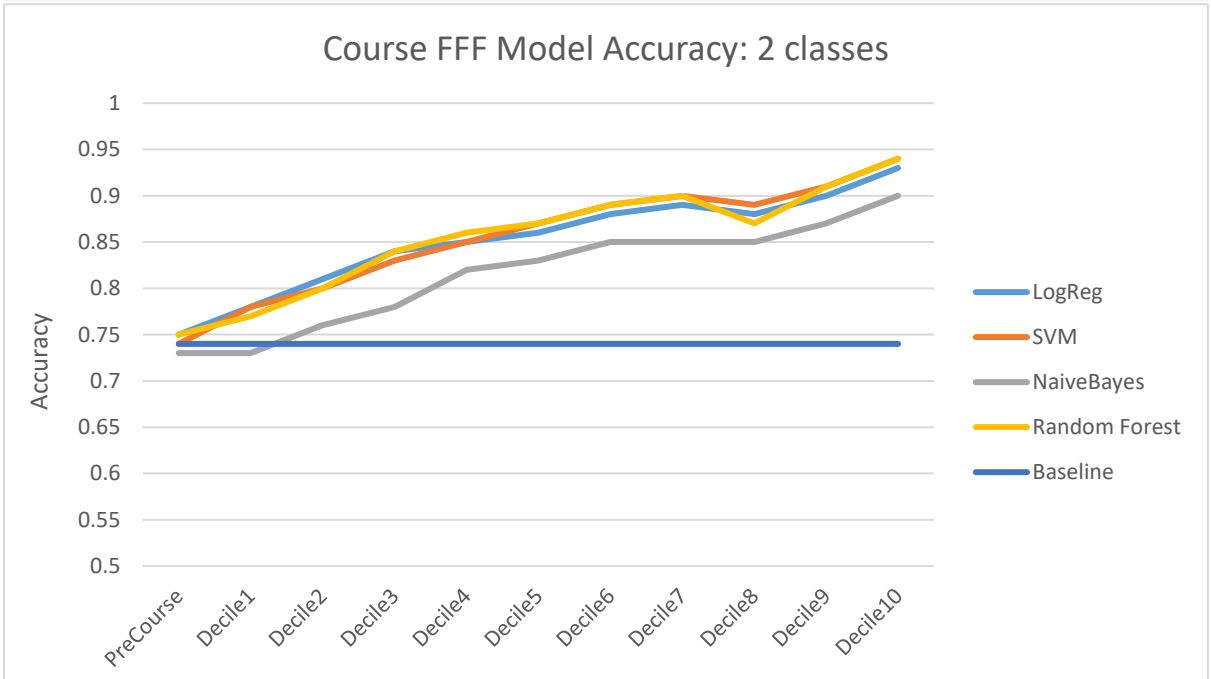
The regression results show similar significant variables. However, it seems does not seem to recognize the effect of the average clicks except for those in decile 9 and 10 in contrast to the random forest model.

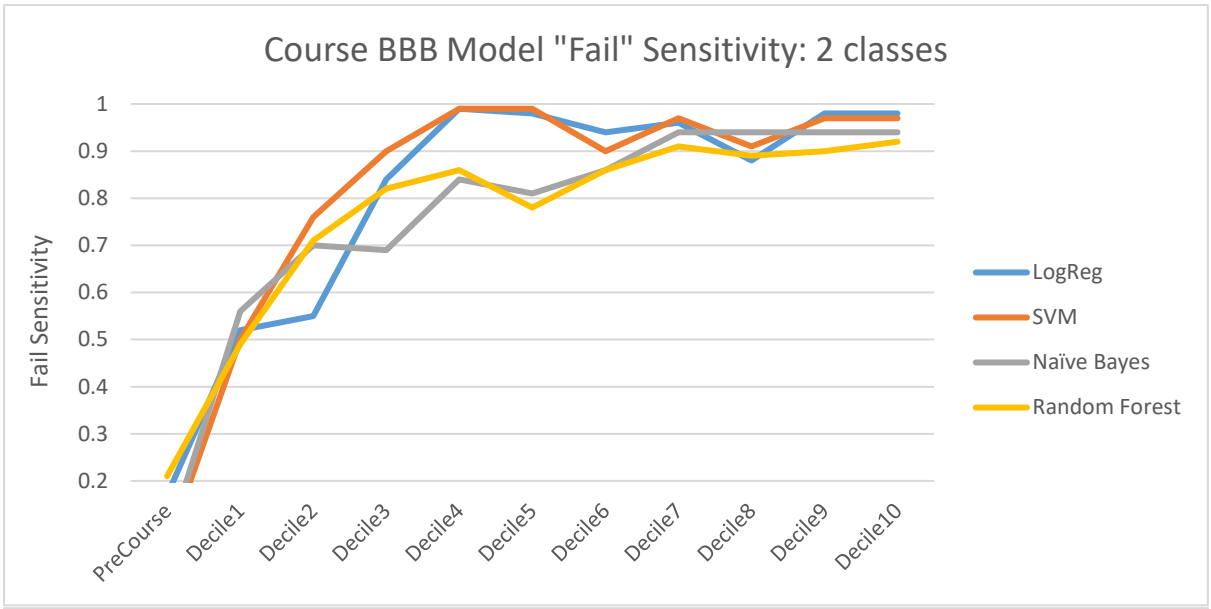
Coefficients:

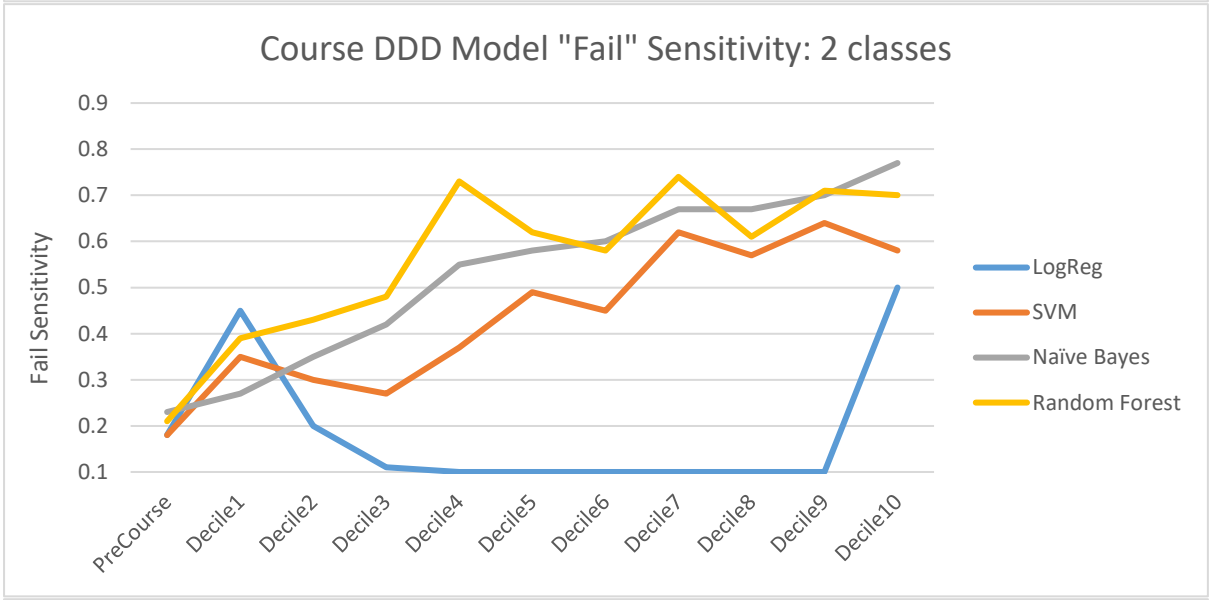
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-1.060e+03	4.738e+02	-2.237	0.025258	*
genderM	-4.325e-02	1.936e-01	-0.223	0.823245	
highest_educationHE Qualification	-2.580e-01	2.180e-01	-1.184	0.236499	
highest_educationLower Than A Level	-6.392e-01	1.524e-01	-4.196	2.72e-05	***
highest_educationNo Formal quals	-4.616e-01	6.812e-01	-0.678	0.498026	
highest_educationPost Graduate Qualification	-1.088e+00	7.658e-01	-1.421	0.155358	
imd_band0-10%	-1.288e+00	3.888e-01	-3.311	0.000929	***
imd_band10-20	-1.012e+00	3.811e-01	-2.655	0.007942	**
imd_band20-30%	-1.260e+00	3.924e-01	-3.212	0.001316	**
imd_band30-40%	-8.740e-01	3.807e-01	-2.296	0.021697	*
imd_band40-50%	-9.135e-01	4.072e-01	-2.244	0.024864	*
imd_band50-60%	-1.201e+00	3.837e-01	-3.130	0.001751	**
imd_band60-70%	-7.380e-01	4.055e-01	-1.820	0.068779	.
imd_band70-80%	-9.825e-01	3.987e-01	-2.465	0.013715	*
imd_band80-90%	-1.010e+00	4.173e-01	-2.419	0.015554	*
imd_band90-100%	-9.857e-01	4.040e-01	-2.440	0.014690	*
age_band35-55	-2.301e-01	1.733e-01	-1.328	0.184194	
age_band55<=	1.350e+00	1.969e+00	0.686	0.493003	
num_of_prev_attempts	-1.300e-02	1.485e-01	-0.088	0.930225	
studied_credits	-1.666e-03	1.791e-03	-0.930	0.352194	
disabilityY	-1.486e-01	2.501e-01	-0.594	0.552428	
Avg_Clicks_PreCourse	-3.926e-02	5.463e-02	-0.719	0.472406	
Avg_Clicks_Decile1	-9.678e-03	6.657e-03	-1.454	0.145956	
Avg_Clicks_Decile2	-1.719e-02	1.011e-02	-1.701	0.088907	.
Avg_Clicks_Decile3	-1.251e-02	1.161e-02	-1.078	0.281184	
Avg_Clicks_Decile4	-1.391e-02	1.253e-02	-1.110	0.266985	
Avg_Clicks_Decile5	-1.174e-02	1.223e-02	-0.959	0.337406	
Avg_Clicks_Decile6	-2.000e-02	1.413e-02	-1.415	0.156965	
Avg_Clicks_Decile7	2.415e-02	1.263e-02	1.912	0.055836	.
Avg_Clicks_Decile8	-9.320e-03	1.149e-02	-0.811	0.417172	
Avg_Clicks_Decile9	3.911e-02	1.322e-02	2.958	0.003099	**
Avg_Clicks_Decile10	1.048e-01	8.723e-03	12.020	< 2e-16	***
Asgmts_PreCourse	3.344e-01	1.119e-01	2.989	0.002801	**
Asgmts_Decile1	3.490e-01	1.567e-01	2.227	0.025928	*
Asgmts_Decile2	6.376e-01	1.156e-01	5.517	3.45e-08	***
Asgmts_Decile3	7.136e-01	1.258e-01	5.674	1.39e-08	***
Asgmts_Decile4	7.416e-01	1.411e-01	5.257	1.46e-07	***
Asgmts_Decile5	6.482e-01	1.129e-01	5.740	9.49e-09	***
Asgmts_Decile6	8.303e-01	1.265e-01	6.563	5.28e-11	***
Asgmts_Decile7	5.655e-01	1.058e-01	5.347	8.95e-08	***
Asgmts_Decile8	7.856e-01	9.035e-02	8.695	< 2e-16	***
Asgmts_Decile9	4.617e-01	7.573e-02	6.097	1.08e-09	***
Asgmts_Decile10	1.449e-01	4.278e-02	3.389	0.000703	***
Avg_Asgmt_Score_PreCourse	1.306e-04	6.468e-03	0.020	0.983892	
Avg_Asgmt_Score_Decile1	4.342e-03	3.583e-03	1.212	0.225590	
Avg_Asgmt_Score_Decile2	1.113e-02	9.263e-03	1.201	0.229573	
Avg_Asgmt_Score_Decile3	-8.293e-03	2.430e-02	-0.341	0.732863	
Avg_Asgmt_Score_Decile4	-5.385e-02	2.705e-02	-1.991	0.046494	*
Avg_Asgmt_Score_Decile5	-8.188e-03	3.171e-02	-0.258	0.796211	
Avg_Asgmt_Score_Decile6	7.501e-02	4.467e-02	1.679	0.093128	.
Avg_Asgmt_Score_Decile7	-6.497e-03	4.741e-02	-0.137	0.891010	
Avg_Asgmt_Score_Decile8	5.847e-02	4.414e-02	1.325	0.185245	
Avg_Asgmt_Score_Decile9	-7.898e-02	4.516e-02	-1.749	0.080307	.
Avg_Asgmt_Score_Decile10	1.306e-01	2.714e-02	4.811	1.50e-06	***
year	5.202e-01	2.353e-01	2.211	0.027028	*
semesterJ	9.905e-01	1.872e-01	5.292	1.21e-07	***

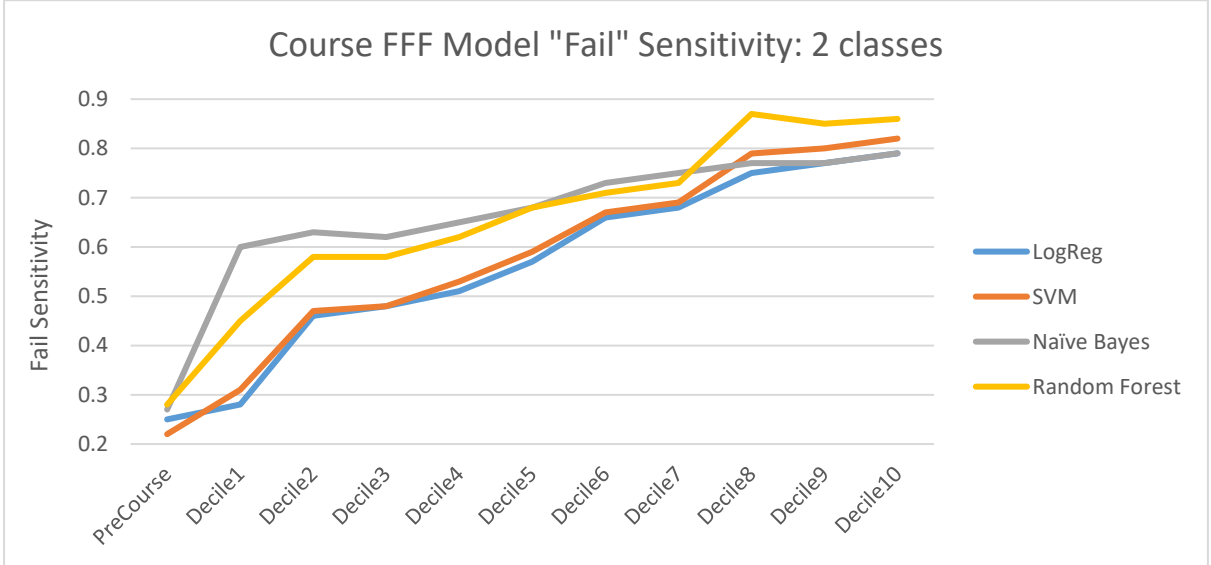

Course BBB Model Accuracy: 2 classes
Accuracy
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest
Baseline


Course DDD Model Accuracy: 2 classes
Accuracy
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest
Baseline


Course FFF Model Accuracy: 2 classes
Accuracy
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
NaiveBayes
Random Forest
Baseline


Course BBB Model "Fail" Sensitivity: 2 classes
Fail Sensitivity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest


Course DDD Model "Fail" Sensitivity: 2 classes
Fail Sensitivity
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest


Course FFF Model "Fail" Sensitivity: 2 classes
Fail Sensitivity
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
PreCourse
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10
LogReg
SVM
Naïve Bayes
Random Forest

Next step:

- Develop LSTM model