Personalized and Explainable Course Recommendations for Students at Risk of Dropping out

Introduction

- **Course recommendation** to support struggling students
- Design developed with respect to insights obtained from a semi-structured group discussion conducted with 25 students → **User-centered design**
- Recommendations based on k-Nearest Neighbors (k-NN) → **Explainability**

Data

- 6 semester bachelor program with 1,484 students
- Program start: winter 2012 – summer 2019
- After preprocessing: 578 students with 9,500 records: 134 students who dropped out (D), 444 who graduated (G)
- **Performance data only**

Research Questions

RQ1: How large is the intersection between the set of recommended courses and the set of courses a student has passed? Are there differences between struggling (dropouts) and well-performing students (graduates)?

RQ2: Do the recommendations lower the risk of dropping out, and if so, how much?

Dropout Risk = Proportion of Students Predicted "Dropout"

Course Recommendations

- Recommendations for 578 students
- Recommendations based on **all neighbors** → **baseline**
- Recommendations based on **graduated neighbors** only

Recommended Courses

S		1				2						
C	M01	M02	M03	M04	M05	M05	M06	M07	M08	M09	M10	E01
0	G	3.3	2.7	3.0	2.0	5.0	2.7	5.0	7.0	7.0	7.0	1.7	1.7
1	G	3.3	3.3	2.3	1.7	5.0	3.0	2.3	7.0	2.7	3.3	2.7	1.0
2	G	3.7	3.3	1.7	2.3	5.0	2.7	3.0	6.0	2.7	7.0	2.0	1.7
3	G	2.7	2.0	1.7	2.0	5.0	3.0	3.3	6.0	3.7	7.0	6.0	1.7

Dropout Prediction

1st Step (P1)

- Model training with 80% of actual data (students who began their studies the farthest in the past)
- Dropout predicted for the remaining 20% of the students
- **Best models** out of a set of different algorithms + hyperparameter optimization + algorithmic-independent parameters
- Evaluated by **balanced accuracy**

2nd Step (P2)

- Dropout predicted using the same best models for the same students as above but with recommended courses **assuming that the courses will be passed**
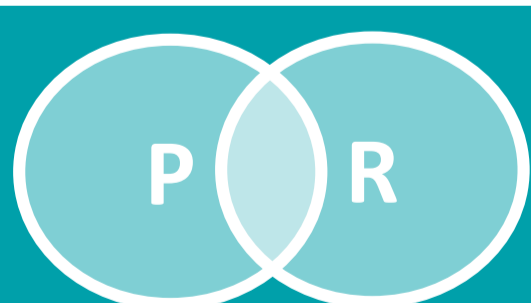
RQ1: Courses' Intersection

- **Mean F1 score is high for graduates**: this confirms our expectation that recommend courses (R) closely match the courses passed (P) by graduates
- **Mean F1 score tends to be low for dropouts**: the recommendations show another way of studying

P / R **F1 Score** **0.861 / 0.509**

For **graduated** / **dropped out** students after their 1st semester

RQ2: Changes in Dropout Risk

- **Dropout risk is lower when the predictions are based on graduated neighbors than using actual data**
- Graduated neighbors' recommendations also provide a lower dropout risk than the baseline

Dropout Risk Reduction **-0.112 / -0.591**

For dropped out students after their 1st / 2nd semester

Future Work

- Evaluate different approaches of **multilabel learning** and approaches that are **equally visualizable and explainable** to students
- Examine the number of recommended courses, especially cases with **no or few recommendations**
- Investigate suitability for **planning over several semesters**

BHT Berliner Hochschule für Technik

Kerstin Wagner, Agathe Merceron, Petra Sauer, Niels Pinkwart

Project „Students Advice"
kerstin.wagner@bht-berlin.de