SCIENCE

Prerequisite Course Recommendation Based on Course Description and Students' Grades

Student researcher: Haozhe Zhou, Junior

A prerequisite course is a class that a student needs to complete before taking an advanced course. Many studies have demonstrated that a student's experience in low-level prerequisite courses is crucial to the student's success in later advanced courses. Yet, current prerequisite courses are mainly decided based on faculty members' professional knowledge and experience in teaching, which has the risk of overlooking certain knowledge components or skills related to the following advanced courses. This gives rise to a concern that students may be underprepared when entering some advanced courses.
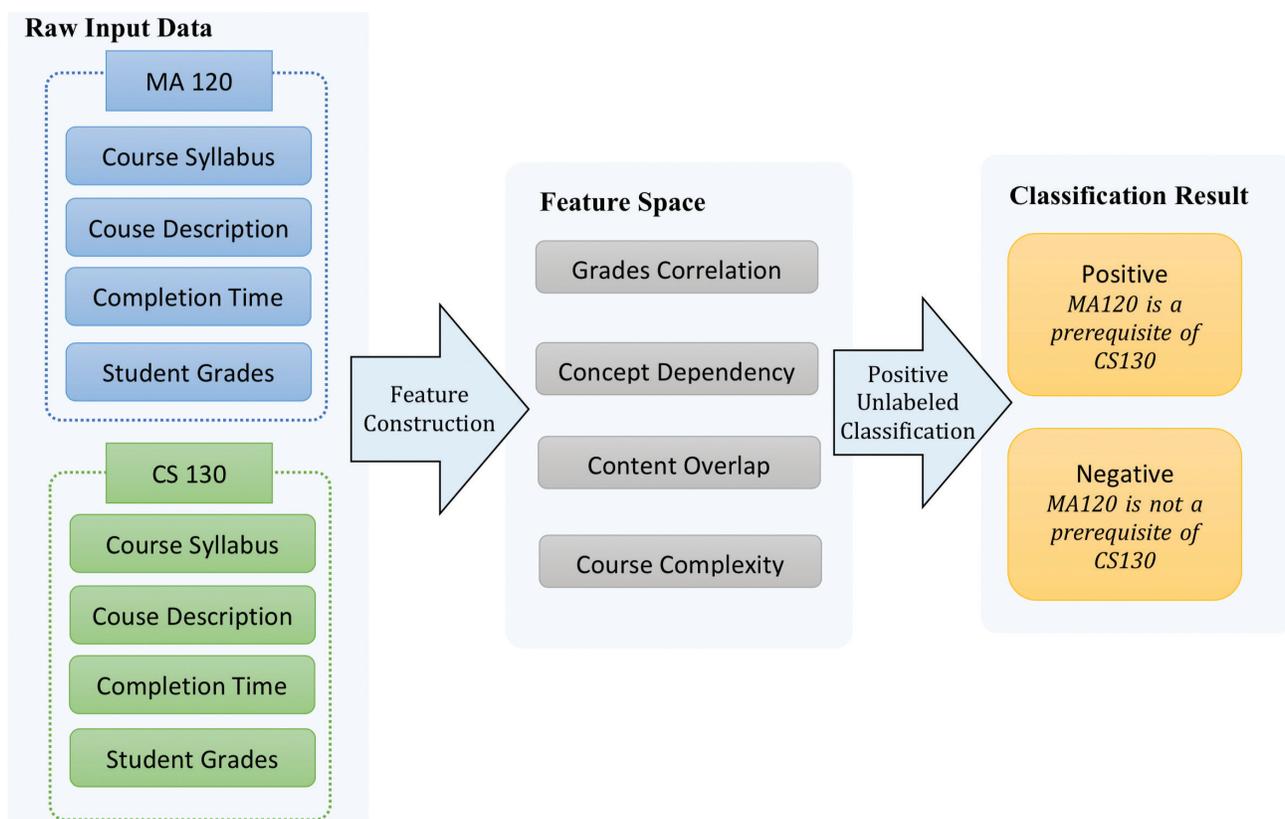
This study aims to construct a model to recommend prerequisite courses from a data mining perspective. The process of suggesting new prerequisites can be viewed as a positive-unlabeled learning problem, where a classification model is trained from only positive and unlabeled data. In this case, the current prerequisites are regarded as the positive labeled data and the goal is to identify other potential positives from the rest of the unlabeled pairs of courses. The model combines grades correlation, concept dependency, content overlap, and course difficulty between two courses to predict their prerequisite relationship.

The dataset comes from historical course grades of the College of Science from 2011 to 2018. Additionally, the course syllabus and textual description of each course are obtained from Purdue's Catalog. With the rich semantic data provided by Wikipedia, dependency relationships between each pair of two courses can be computed using the concepts extracted from the course syllabus and text descriptions.

The hypothesis is that this model will be able to suggest some new prerequisite courses that could have a positive effect on students' performance in the advanced courses. This project will potentially assist students and institutions to improve the current curriculum pathway and contribute to academic success.

Research advisor David Nelson writes: "Purdue College of Science has the most enrollments and service courses that support students. It also has the most varied pathways to degree within science majors, where students rarely follow a specific track. Haozhe explores an algorithm to help students choose the path most likely to lead to academic success."

Zhou, H. (2020). Prerequisite course recommendation based on course description and students' grades. *Journal of Purdue Undergraduate Research, 10,* 110–111.

Overview of the data flow. The input raw data from two example courses CS130 and MA120 are processed and mapped to a feature vector that covers grades correlation, concept dependency, content overlap, and course complexity relationship between the two courses. The classifier then predicts either positive or negative from these constructed features.

Zhou, H. (2020). Prerequisite course recommendation based on course description and students' grades. *Journal of Purdue Undergraduate Research, 10,* 110–111.