

Visual Analytics Proposal

Group 21 - Jayin Nejal, Nithish Gnanasekaran

Topic Student Performance and Learning Style

Dataset. For our visual analytics project, we have chosen our own challenge/dataset that explores the factors influencing student performance via analysing and predicting how student lifestyle can impact academic performance. The dataset [Student Performance and Learning Style](https://www.kaggle.com/jayaantanaath/student-habits-vs-academic-performance/data) is taken from kaggle and it includes information on 10,000 students (613kB) and looks at study habits, learning preferences, lifestyle choices, and participation levels related to academic outcomes like exam scores. This data set comes in the form of a table with 15 features which contains eight numerical, five categorical, and two ordinal attributes. Attributes like Study hours, sleep hours, and social media time are quantitative and have a naturally increasing ordering. Categorical fields like gender, learning style and use of tech are nominal with no inherent order.

Goal. The purpose of analysing this dataset is to improve overall student performance by evaluating habits which students are able to change. Struggling students might find it hard to see what they could improve or how they differ from well performing students. Many students are different, finding patterns in groups of students is therefore crucial. This particular dataset should be able to give a good overview due to its size and features.

User. The intended user of this system is an academic advisor who works directly with individual students, helping them improve performance and navigate their studies. They can use the insights from the tool to identify patterns in student behavior and performance in order to offer personalized guidance and improve academic outcomes by recommending study strategies or by suggesting behavioral changes. In addition, students who would like to improve their performance can use this tool to find their own strengths and weaknesses.

Tasks. The system supports both descriptive and predictive analytical tasks. Descriptive tasks include exploring student lifecycle clusters to understand how different (study) habits relate to academic performance, and identifying trends across various performance levels. Predictive tasks involve predicting a student's final grade based on their behaviors and characteristics, as well as recommending personalized study strategies tailored to their learning style and past outcomes. The system can support analytical questions like

- How do different learning styles and study behaviors correlate with academic performance?
- What personalized study strategies can be recommended to improve a student's outcome?

Tools. To develop an interactive visual analytics tool, we decided to go with Dash. As we are both familiar with machine learning in Python, Dash allows us to apply this into a web application which can be run locally. Alternatives we considered were using Typescript with D3.js for the frontend and then connecting this through a restAPI (like fastAPI) with a Python backend. However, using Dash seemed like a cleaner and more direct way, removing the need of a separate frontend and backend and combining all code into a single language. In addition, Dash is able to be flexible and update in real-time by for example interacting with the graphs or letting the user input their own data (see tasks above).

References

<https://www.kaggle.com/jayaantanaath/student-habits-vs-academic-performance/data>