# PROJECT PROPOSAL

## Student academic success and their future

NOVEMBER 6

**Fontys University of Applied Sciences**
**Authored by: Denitsa Goranova**
**Version: 1.3**

## Table of Contents

| Date | Version | Status |
|------|---------|--------|
| 16/09/2023 | 1.0 | Started document |
| 25/09/2023 | 1.1 | Made updates |
| 27/09/2023 | 1.2 | Added research questions, updated introduction, domain understanding based on feedback |
| 23/10/2023 | 1.3 | Updated based on feedback |

# Introduction

## WHAT

The pursuit of better student success rates inside educational institutions has never been more vital. The aim of this project is to improve student achievements and well-being by providing them with predictive models. The project represents forecasting final grades based on previous performance. I will build a machine learning model to predict if a student is likely to pass a course. I will collect relevant data, preprocess it, choose appropriate characteristics and develop a prediction model to accomplish this.

## WHY

Nowadays students are facing problems in their schools and sometimes they do not even know it. I believe that by predicting a student's academic performance and letting know those at risk of poor outcomes, I can make a difference and create a product which genuinely benefit students as well as educational institutions.

## WHO

The project will involve a stakeholder and end users of teachers, school administrators.

## WHEN

The first weeks (1-3), I will collect data and preprocess it, after that I will select the model and development, model evaluation, finalizing the predictive model.

## HOW

I will need access to student data from two Portuguese schools including grades, demographic information and other relevant data. There are 650 students data. I will use Python for data analysis and machine learning, along with libraries like Scikit-Learn, Pandas, Matplotlib. For my documentation and code development, I will use Jupyter Notebook.

# Stakeholders

The stakeholder in this project is a second-year student in a Bulgarian university.

# Phases

The project will take 12 weeks in order to be finished. The phases that we have are: Project proposal, Domain Understanding, Data Sourcing, Analytic Approach, Monitoring, Changes.

# Domain Understanding

## Methodology

The methods that I used are Interview and Community research. I decided to conduct an interview because it was recognized that real-world observations from students, such as the interview with Antonio Simeonov, give a detailed viewpoint. For community research- it allows me to find more complete findings, enables the investigation of communal experiences, behaviors.

## Research

To lay a solid foundation for my project I conducted a thorough investigation of student academic performance. My knowledge is founded on a combination of genuine exploratory research and ideas gained from interactions with educational stakeholders.

## Keywords

Student success prediction, academic achievement, performance, predictive analytics

## Main question:

How can predictive analytics models be leveraged to improve students success and reduce the likelihood of course retakes in educational institutions?

### Sub- questions:

1. How effective can predictive analytics models be in providing timely interventions and support to students identified at risk?
2. What specific data points are most indicative of a student's risk of retaking a course?
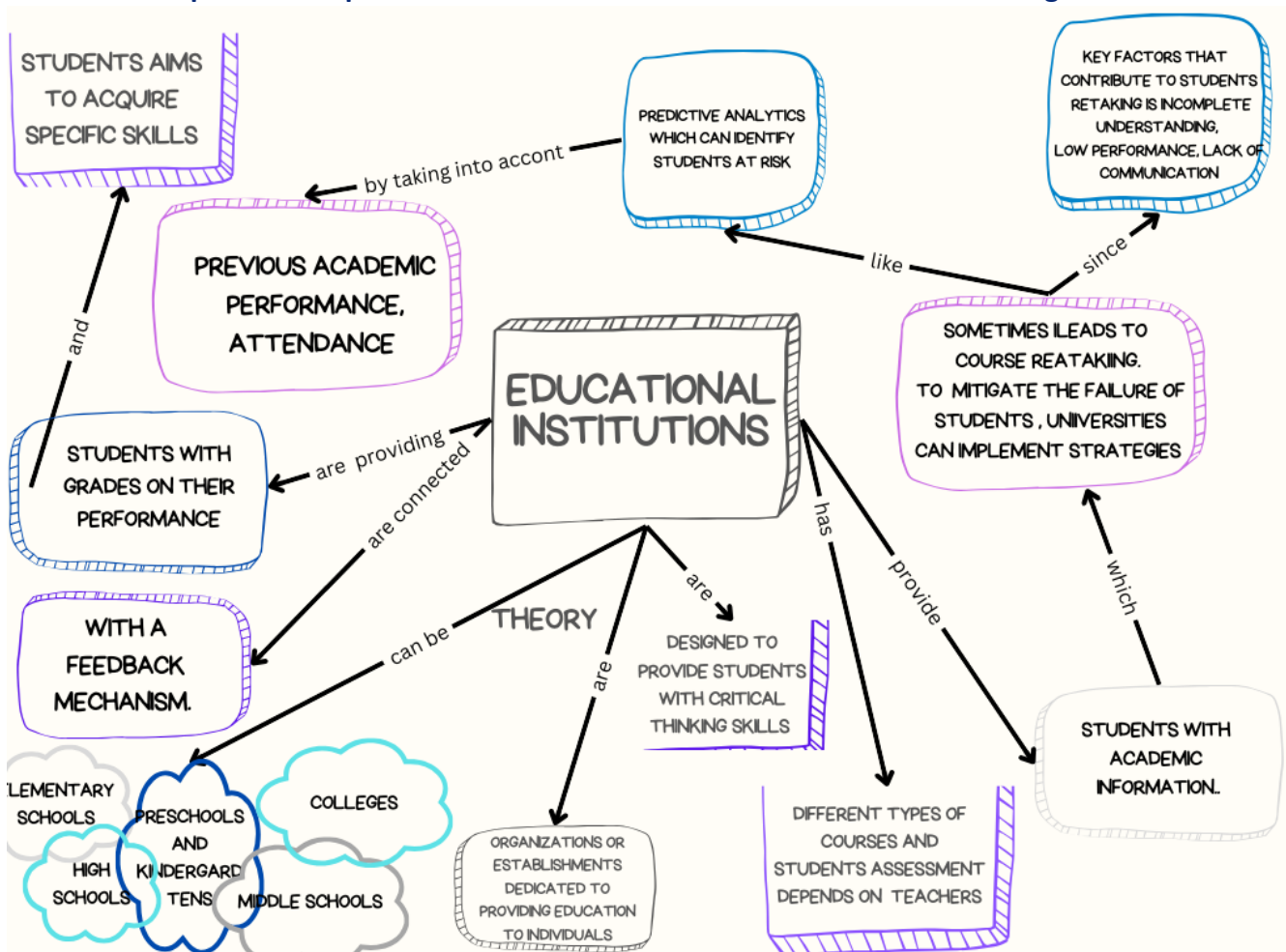


1. *Figure 1-Cognitive scheme*

According to my research, academical retention is multidimensional issue with significant consequences for both students and educational institutions. Predicting Student Performance in Higher Educational Institutions Using Video Learning Analytics and Data Mining Techniques by Raza Hasan, Sellapan Pala, Salman Mahmood, Ali Abbas, Kamal Uddin Sarker and Mian Usman Sattar presented insights into the challenges that students faced. Furthermore, predictive analytics has evolved as a vital and important instrument in the field of higher education. Smith (2019) emphasizes its expanding importance, particularly in light of rising educational expenditures and the compelling need to shorten degree completion delays. These problems have a significant impact on students' academic journeys as well as their self-assurance.

A study of mental health of Canadian secondary school students [22] investigated whether symptoms of mental illness (depression and anxiety) and mental well-being are associated with self-reported grades and education behaviours. This was done using simple statistical measures such as direction, magnitude and significance. The results showed that lower depression and higher well-being were consistently associated with better academic performance and education behaviours. Some key findings from other research highlight that students' well-being varies substantially between schools, and that schools which have practices to promote well-being have better overall well-being levels

As a result, there is a compelling need for educational institutions to embrace predictive analytics as a method of identifying students who may need to repeat a course. We may build a more comprehensive and responsive strategy to improving students' academic achievement and reducing the need for course retakes by investigating the various and nuanced aspects that contribute to this issue.

**Interview**

In order to get clean picture of how students feel in their educational institution and more precisely if they know how they are performing during their educational time, I conducted an interview as well with student from Bulgarian university.

## Interview with Antonio Simeonov (student in Bulgaria)

**Age:** 23

**Gender:** Male

**Filed of study:** Second year student in a foreign language university

**On a scale from 1 to 10, how would you rate your current academic performance?**

**Answer:** I would say a solid 8. I have good grades and based on my previous performance, I am a student, who always seek to have new opportunities in order to develop myself.

**What do you believe are the most common challenges that students face in their educational setting?**

**Answer:** The most common challenges that students face in their educational setting are: adaption in their studies, management, deadlines and having a proper communication. What I know from my co-students, they are struggling with understanding the curriculum, specific subjects, stress, depression.

**Can you identify if you are at risk of retaking your course?**

**Answer:** In the beginning of my studies I have been struggling to identify if I am at risk or not, I was stressed about my deadlines, my communication with the teachers and mostly if I will succeed to reach the goal for the semester. For sure, it would be better to have common meetings with my teacher in order to identify if I am at risk, since in our university there are no individual meetings with the teachers. Based on my experience, a lot of students would be relieved if teachers give some clarification on how they are performing and if there is a chance of retaking. Since we study in a very big school, some of the teachers cannot give so many feedback sessions, because of the number of students.

**What factors do you believe that contribute students to fail?**

**Answer:** There are various reasons why a student can retake a course. For example: wasting time instead of studying, misalignment of priorities, misunderstanding of the tasks, miscommunication, focusing only on studying and forgetting about their social

life, which is an important part of the fully life. What I know from my personal experience is that miscommunication is the main contributor to people failing, Poor communication with teachers can hinder a student's ability to seek clarification, ask questions or address concerns. If students do not know where to turn for help, it can lead to academic struggles.

In your experience, how important are factors like family support, extracurricular activities in a student's academic success?

**Answer:** As I previously mentioned in the previous question, social life and family life is a very important part of our lives. They are playing a crucial part in our lives, since it influences our mood, energy and motivation. If for example, you have problems at your home, I am sure that you would not be motivated enough to study about your Math test.

In your opinion, how can predictive analytics play a role in addressing those challenges?

**Answer:** Predictive analytics can help with expansion of the before mentioned challenges. For example, it will be very helpful, if I can see some predictions on my studies. I also believe that in the Bulgarian educational system would be very helpful to have it since the teachers are going to see on which students should focus more. Predictive analytics can support the development of adaptive learning systems. These systems can adjust the difficulty of coursework based on individual student progress, ensuring that students receive appropriate challenges and support to keep them engaged and motivated.

How comfortable are you with reaching out to your teachers when you have questions or feedback desires?

**Answer**: Sometimes, I am comfortable with seeking help from the teachers or my advisors, but it depends on the person in front of me. Sometimes the teachers are not open for meetings and prefer to give feedback only during feedback sessions.

**Is there something that you would like to add about your academic challenges, any suggestions for improvement?**

**Answer:** There is one thing that I miss in our field, especially in Bulgarian educational institutions- modern solutions. I would very much appreciate if there was an app which helps us improve our performance or at least some more info about our future in the institution.

# Conclusion

Based on the interview with the student, I can say that students are frequently challenged by lack of communication, deadlines, time management. There are some factors that are crucial in the situation with their performance like family status, previous performance, absence. The feedback also emphasizes how crucial it is for instructors and students to communicate effectively. A lot of students say they want feedback and communication that is clearer so they can see where they can grow and possible hazards. The absence of one-on-one contacts with teachers at bigger educational institutions may make it more difficult to comprehend their performance and possible hazards. In the end, predictive analytics may be extremely important for raising student achievement and lowering the chance of course retakes in academic settings. Educational establishments may establish a more encouraging and productive learning environment by utilizing data-driven insights to offer tailored interventions, better coordinate academic assistance, and improve communication. This method can assist students in staying on course, achieving their learning objectives, and performing well academically.

# Data Sourcing

I discovered an appropriate dataset with 650 rows for the data sourcing portion of this project. ([https://archive.ics.uci.edu/dataset/320/student+performance](https://archive.ics.uci.edu/dataset/320/student+performance)) This dataset focuses on secondary school students accomplishments and includes two Portuguese schools. It consists of a diverse set of data elements such as student grades, social aspects and school-related features. The information was gathered using a mix of school reports. It's essential to highlight that the target attribute, labeled as "G3," exhibits a strong correlation with the attributes "G2" and "G1." G3 represents the final year grade, while G1 and G2 correspond to grades from the first and second periods.  There are other values in my dataset. School which indicates the student's school (GP is Gabriel Pereira, MS -Moushino da Silveira), sex-student's gender, age, address, family size, parent's cohabitation status- describes the parent status(T – living together, A-apart), mother education and father education(on a scale from 0 to 4), mother's and father's job(categorizes the father occupation as "teacher", "health care related", "civil services", "at_home", "other". There are other features related to travel time, study time, past class failures, family support, activities, health status. These features will help me as a foundation for my predictive modeling and analysis to identify factors influencing student success and predicting the final grades.

Based on the interview and the research that I have made, I can conclude that academic retention challenges are confined by various factors. The knowledge that I gained from the domain understanding is evident that the implementation of predictive analytics offers a convenient solution for students. The research and the interview have already begun to provide me with answers and a predictive model will empower students and institutions alike to proactively address academic retention challenges, leading to improved student success and confidence.

# Analytic Approach

The problem is primarily a classification task. I will explore various machine learning algorithms suitable for classification tasks, such as logistic regression, decision trees, k-nearest neighbors and others. I will evaluate the models using appropriate metrics

for classification tasks including accuracy. I will conduct correlation analysis to identify which features have a significant influence on the target variable.  This will involve calculating correlation coefficients and identifying strong positive or negative correlations. I will analyze the dataset to identify relevant indicators that may influence a student's risk at failing. If the initial model results do not meet my desired performance, I will revisit the section and refine the proper approach.

# References

Smith, J. A. (2019). Predictive analytics in higher education, Educational Data Journal, 15(3), 225-240

Gardner, J., Brooks, C. Student success prediction in MOOCs. *User Model User-Adap Inter* 28, 127–203 (2018)

Hassan Zeineddine, Udo Braendle, Assaad Farah,

Enhancing prediction of student success: Automated machine learning approach,

Computers & Electrical Engineering,Volume 89,2021,106903