**Predict Students' Dropout and Academic Success**

The dataset includes information known at the time of student enrollment (academic path, demographics, and social-economic factors) and the students' academic performance at the end of the first and second semesters. The data is used to build classification models to predict students' dropout and academic success.

Feature: real, categorical, integer

Missing values: no

How to process the data?

Data selection: selecting the subset of all available data that you will be working with. You need to consider what data you need to address the question or problem you are working on. Make some assumptions about the data you require and be careful to record those assumptions so that you can test them later if needed.

Process data: After you have selected the data, you need to consider how you are going to use the data. The data you have selected may not be in a format that is suitable for you to work with. Cleaning data is the removal or fixing of missing data. There may be data instances that are incomplete and do not carry the data you believe you need to address the problem. There may be far more selected data available than you need to work with.

Data transformation: Transform preprocessed data ready for machine learning by engineering features using scaling, attribute decomposition and attribute aggregation.

machine learning techniques:

linear regression: is a supervised machine learning technique used for predicting and forecasting values that fall within a continuous range.

Training testing and validation