**Capstone project description**

For the capstone project, students are tasked with building and validating a predictive model to assess the likelihood of a loan becoming delinquent (90+ days overdue) within a 12-month period.

**Objective:**

* Use **train\_data.csv** to train a model that predicts the probability of a loan going into delinquency.
* Apply **val\_data.csv** solely for the validation process to evaluate the model's performance.

**Project Overview:**

* **Data preparation**: Students will analyze the features provided in the training data, identify any missing values, and perform necessary preprocessing.
* **Model selection**: They are expected to experiment with different models (e.g., logistic regression, decision trees, random forest) to find the most suitable one.
* **Evaluation metrics**: Performance metrics such as AUC-ROC, precision, recall, and F1-score should be used to evaluate the model's effectiveness.
* **Validation**: After model training, the validation dataset will be used to assess the generalization of the model and its predictive power.

The focus is on end-to-end development, from preprocessing data to model evaluation and interpreting the results.