**Predicting Credit Card Approvals (Logistic Regression + Cross Validation Implementation)**

**Project description and objective:**

**Approving Credit Card Process:** Approval of credit cards is typically done by financial institutions, which decide whether to grant an applicant access to a credit line by issuing a card. When an application is submitted, various factors determine whether to approve or deny it. In the project, these factors will be transformed into features, including the applicant’s income level, employment status, debt-to-income ratio, and credit scores, which can objectively estimate the risk level, and other financial information. Upon approval, the applicant receives a credit card, providing them with a line of credit and a predefined credit limit, enabling them to borrow funds as needed and repay later. The ML model for credit card approval can serve as an automated credit card approval system, assisting institutions in managing risks and maintaining an efficient workload.

**Logistic regression will be used to build a predictive model, while cross-validation will be used to evaluate the model's performance.**

**Datasets:**

1. A dataset containing general information about the applicant (birth date, gender, assets, education, etc.)

* 12 categorical variables
* 5 continuous variables
* ID

1. A dataset containing the applicant’s loan payment records.

* 1 categorical variable
* 1 continuous variable
* ID

(two datasets have a feature called ID, which will be used to join the tables.)

**Datasets Link:**

[**https://www.kaggle.com/datasets/rikdifos/credit-card-approval-prediction?resource=download&select=credit\_record.csv**](https://www.kaggle.com/datasets/rikdifos/credit-card-approval-prediction?resource=download&select=credit_record.csv)

**Datasets dictionary link:** [**https://www.kaggle.com/datasets/rikdifos/credit-card-approval-prediction/discussion/119320**](https://www.kaggle.com/datasets/rikdifos/credit-card-approval-prediction/discussion/119320)

**Logistic regression usage:** We either approve or deny the credit card, so we have 2 outcomes and let x = 1 if approved and x = 0 otherwise. So we have Bernoulli distribution and a linear classification problem. For this reason, we will use logistic regression.