**Faculty of Engineering**
**Computer Engineering Department**

Phase 1: Project Proposal

# Machine Learning Project

## Team 12

**Presented by**



| Name | Section | BN |
|---|---|---|
| Mohamed Yehia Alahmad | 2 | 14 |
| Moustafa Mohammed Elsayed | 2 | 23 |
| Mostafa Hani Mostafa | 2 | 24 |
| Mennatallah Ahmed Moustafa | 2 | 25 |

# Idea1: Loan Default Prediction

## Problem Definition:

Our finance company gives loans to people in cities, and we need to figure out if a new customer will pay us back or not. When someone asks for a loan, we look at their info—like how much they earn, their job, and past loans—to decide whether to say yes or no. There are two big worries: we lose money if we say yes to someone who won't pay us back (they "default"). If we say no to someone who would pay us back, we miss out on business. We have data about old customers and whether they paid or defaulted. The job is to use this data to build a machine learning model that guesses if a new person is "likely to default" (won't pay) or "likely to pay." This will help us decide who gets a loan, and who doesn't.

## Motivation:

We want to get better at giving loans so we don't lose money and can help more people. If too many customers don't pay us back, we lose cash from the loan and extra costs to chase them. But if we're too picky and say no to good customers, we miss chances to grow. Our data—like how much debt someone has or if they've missed payments before—can show us who's risky and who's safe. A smart machine learning tool can spot things we might miss, like someone with a big salary but bad habits. This project will help us lose less money by catching risky people, and make more money by saying yes to the right ones.

## Evaluation Metrics:

1. F1 score
2. Accuracy
3. Micro and Macro precision and recall

## Dataset:

| Link | #columns | # rows | Size |
|---|---|---|---|
| Dataset | 47 | 27k | 17 Mb train<br>7 Mb test |

# Idea2: Customer Segmentation

## Problem Definition

We need a machine learning model to predict if a customer will sign up for a term deposit after a phone campaign. A term deposit is when a customer puts money in the bank and agrees not to take it out for a set time, like 6 months or a year. The bank pays them extra interest for this, and it helps the bank because they can use that money to lend out and make profits. It's a yes-or-no question: 1 if they subscribe to a term deposit, 0 if they don't. We've got a training dataset with past results and a test dataset to predict. The data includes things like age, job, education, bank balance, loans, and how we contacted them. (We won't use the ID since it's just a label.) Our goal is to find patterns in this info to figure out who's likely to say yes to a term deposit. A good model will help us target the right customers, avoid wasting effort, and get more people to lock in their money with us.

## Motivation

Our bank is losing money because customers aren't depositing as much as they used to. Term deposits are a big deal for us—they let us hold onto money for a set time, so we can lend it out and make more profit. Plus, customers with term deposits are more likely to buy other stuff like insurance or funds, which boosts our revenue even more. The problem is, that we don't know who'll sign up for a term deposit after our marketing calls. We've got data from past phone campaigns—like customer age, job, loans, and more—and we can use it to predict who's likely to say yes. With a smart machine learning model, we can target the right people, save money on marketing, and get more deposits to grow our profits again.

## Evaluation Metrics:

1. F1 score
2. Accuracy
3. Micro and Macro precision and recall

## Dataset:

| Link | # columns | #rows | Size |
|---|---|---|---|
| **Dataset** | 18 columns | 31647 | 3 Mb Train<br>1.28 Mb Test |