Dataset that provides information on a local bank's marketing campaign. Our task is to predict whether someone will open a savings account or Not.

Build the best model you can .

Dataset Description
You will get two datasets. train.csv is for training your model, and test.csv contains the information to predict

Dataset description

Files

- **train.csv** - the training set
- **test.csv** - the test set

Columns

Client information

- `id` - client id (numeric)
- `age` - age of client (numeric)
- `job` - type of job (categorical: "admin.","artisan","entrepreneur", "housemaid", "management", "retired", "self-employed", "services", "student", "technician", "unemployed", "unknown")
- `civil` - marital status of client (categorical: "divorced", "married", "single","unknown"; note: "divorced" means divorced or widowed)
- `education` - education of client (categorical: "4K", "6K", "K9", "K12", "illiterate", "apprenticeship", "university", "unknown")
- `credit` - has credit in default? (categorical: "no","yes","unknown")
- `hloan` - has housing loan? (categorical: "no","yes","unknown")
- `ploan` - has personal loan? (categorical: "no","yes","unknown")

Campaign details

- `ctype` - contact communication type (categorical: "cellular","telephone")
- `month` - last contact month of year (categorical: "jan", "feb", "mar", …, "nov", "dec")
- `day` - last contact day of the week (categorical: "mon","tue","wed","thu","fri")
- `ccontact` - current number of contacts performed during this campaign and for this client (numeric, includes last contact)
- `lcdays` - number of days that passed by since client was last contacted by a previous campaign (numeric; 999 means client was not previously contacted)
- `pcontact` - number of contacts performed before this campaign and for this client (numeric)
- `presult` - outcome previous marketing campaigns (categorical: "failure","nonexistent","success")

Socioeconomic indicators

- `employment` - employment variation rate - quarterly indicator (numeric)
- `cprice` - consumer price index - monthly indicator (numeric)
- `cconf` - consumer confidence index - monthly indicator (numeric)
- `euri3` - euribor 3 month rate - daily indicator (numeric)
- `employees` - number of employees - quarterly indicator (numeric)

Outcome variable (target)

- `outcome` - has the client opened a saving account? (binary: 1 = "yes", 0 = "no")
-

Task to do:

EDA:

Data understanding (visualization, etc.)

Data preparation (variable treatment, feature creation)

Modeling: (Using Decision tree)

Evaluation methodology (how did you evaluate your model)

Plot confusion matrix diagram.

Model Evaluation

We will evaluate models using:

Accuracy

Area under ROC (AUC).

AUC value should be greater than 80%

Finally choose the model evaluation(Evaluation methodology to be implemented).