

STAGE

Machine Learning Challenge 3 (Classification Problem) - Round 1

Save & Exit

00D : 02H : 09M : 50S

Questions

1. Classification Problem . ()

Attempts Remaining

✓ 0 / 15

Attempts made

Note:

- You can do multiple submissions.
- Your highest score will be considered

Classification Problem

The data given is of credit records of individuals with certain attributes. Please go through following to understand the variables involved:

a. **serial number** : unique identification key

b. ****account_info**** : Categorized details of existing accounts of the individuals. The balance of money in account provided is stated by this variable

A11 signifies 0 (excluding 0) or lesser amount credited to current checking account. (Amounts are in units of certain currency)
A12 signifies greater than 0 (including 0) and lesser than 200 (excluding 200) units of currency
A13 signifies amount greater than 200 (including 200) being recorded in the account
A14 signifies no account details provided

c. **duration_month** : Duration in months for which the credit is existing

d. **credit_history** : This categorical variable signifies the credit history of the individual who has taken the loan

A30 signifies that no previous loans have been taken or all loans taken have been paid back.
A31 signifies that all loans from the current bank have been paid off. Loan information of other banks are not available.
A32 signifies loan exists but till now regular installments have been paid back in full amount.
A33 signifies that significant delays have been seen in repayment of loan installments.
A34 signifies other loans exist at the same bank. Irregular behaviour in repayment.

e. **purpose**: This variable signifies why the loan was taken

Help



STAGE

Save & Exit

Machine Learning Challenge 3 (Classification)

Problem Statement

Round 1

A43 signifies that the loan is taken to buy furniture or equipment

A44 signifies that the loan is taken to buy radio or TV

A45 signifies that the loan is taken to buy domestic appliances

A46 signifies that the loan is taken for repairing purposes

A47 signifies that the loan is taken for education

A48 signifies that the loan is taken for vacation

A49 signifies that the loan is taken for re-skilling

A410 signifies that the loan is taken for business and establishment

A410 signifies other purposes

f. **credit_amount**: The numerical variable signifies the amount credited to the individual (in units of a certain currency)

g. **savings_account**: This variable signifies details of the amount present in savings account of the individual:

A61 signifies that less than 100 units (excluding 100) of currency is present

A62 signifies that greater than 100 units (including 100) and less than 500 (excluding 500) units of currency is present

A63 signifies that greater than 500 (including 500) and less than 1000 (excluding 1000) units of currency is present.

A64 signifies that greater than 1000 (including 1000) units of currency is present.

A65 signifies that no savings account details is present on record

h. **employment_st**: Categorical variable that signifies the employment status of everyone who has been allotted loans

A71 signifies that the individual is unemployed

A72 signifies that the individual has been employed for less than a year

A73 signifies that the individual has been employed for more than a year but less than four years

A74 signifies that the individual has been employed more than four years but less than seven years

A75 signifies that the individual has been employed for more than seven years

Help

i. **poi**: This numerical variable signifies what percentage of disposable income is spent on loan interest amount.

j. **personal_status**: This categorical variable signifies the personal status of the individual

A91 signifies that the individual is a separated or divorced male

A92 signifies female individuals who are separated or divorced

A93 signifies unmarried males

A94 signifies married or widowed males

A95 signifies single females

k. **gurantors**: Categorical variable which signifies if any other individual is involved with an individual loan case

A101 signifies that only a single individual is involved in the loan application

A102 signifies that one or more co-applicant is present in the loan application

A103 signifies that gurantors are present.

STAGE

Save & Exit

Machine Learning Challenge 3 (Classification Problem) - Round 1

A121 signifies that the individual holds real estate property
 A122 signifies that the individual holds a bond or mutual fund or life insurance
 A123 signifies that the individual holds cars or other properties
 A124 signifies that property information is not available

n. **age**: Numerical variable that signifies age in number of years

o. **installment_type**: This variable signifies other installment types taken

A141 signifies installment to bank
 A142 signifies installment to outlets or stores
 A143 signifies that no information is present

p. **housing_type**: This is a categorical variable that signifies which type of housing does a applicant have.

A151 signifies that the housing is on rent
 A152 signifies that the housing is owned by the applicant
 A153 signifies that no loan amount is present on the housing and there is no expense for the housing)

q. **credits_no**: Numerical variable for number of credits taken by the person

r. **job_type**: Signifies the employment status of the person

A171 signifies that the individual is unemployed or unskilled and is a non-resident

A172 signifies that the individual is unskilled but is a resident

A173 signifies that the individual is a skilled employee or official

A174 signifies that the individual is involved in management or is self-employed or a highly qualified employee or officer

s. **liables**: Signifies number of persons dependent on the applicant

t. **telephone**: Signifies if the individual has a telephone or not

A191 signifies that no telephonic records are present
 A192 signifies that a telephone is registered with the customer's name

u. **foreigner**: Signifies if the individual is a foreigner or not (considering the country of residence of the bank)

A201 signifies that the individual is a foreigner
 A202 signifies that the individual is a resident

Objective of the problem: As per predictions in the prediction problem. The objective of this problem is to predict the cluster number of serial number variable.

Cluster number 1 (Correct value is 1) : When the value of credit_amount is between 4000 and 20000

STAGE

Machine Learning Challenge 3 (Classification)


Save & Exit




~~Problem - Round 1~~
Evaluation Metric: Accuracy

Submission Limit: The submission limit for each team for each question is 15. Please note that individual submission limits would be shown to individuals. Though individual submissions are shown to individuals, if any other team mate make a submission, the team submission counter would increase.

Both questions of machine learning challenge 3 are to be answered.

Evaluation Algorithm

 Accuracy

	Training Set	(https://cdn.skillenza.com/files/c7e3d1c3-8fb3-4159-abc0-b2e44d0925ce/Yes_Bank_Train.csv)
	Test Set	(https://cdn.skillenza.com/files/104ca536-091a-46c1-ba9f-21e5f61fda8a/Yes_Bank_Test_int.csv)
	Sample Submission	(https://cdn.skillenza.com/files/26ee1105-4b49-4332-a91e-0cb3586a9ad8/sample_clusters.csv)

Upload file

Submit