CREDIT SCORE CLASSIFICATION GUIDELINE

JUN ZE HE

Introduction: Credit Score Classification Model Credit Score is an important factor that is how banks determine how much credit balance you can borrow from your credit account. To save manual efforts to identify whether a credit account has a good credit score, we want to develop a classification model, that can classify credit scores from peoples' credit accounts.

Credit Score Dataset

Week3: What is a classification model

Goal
(1) Get familiar with Google colab
(2) Know how to upload files and connect to your Google Drive on Google colab
(3) Take a look at my data science functions repository and know how to use functions
(4) recognize what the predicted variables and the response variable are in the dataset
(5) Get to know the basics of machine learning
Binary classification has two classes, but multi-class classification has 2 or more than two classes

∎

Resources
What is machine learning, and what is a classification model?

∎

Week4: Data Cleaning

Goal
The goal of data cleaning is to clean the dataset without NULL and strange values
(1) First step: check if there is any value in the correct form, for example, price is an object, any variable has an incorrect value
(2) Second step: identify how many null values are in the dataset, then try to correct them with median or mean in numerical variables and most frequent values in categorical variables

∎

Resources
Pandas Tutorials
numpy Tutorials
Handling missing values in categorical variables
Handling missing value in general

∎

Week5: statistical analysis / EDA

Goal

∎

Week6: Feature Selection / Feature Engineering

Goal

∎

Week7: Data Preprocessing / Models

Goal

∎

Week8: Conclusions/Making slides

Goal

∎