# Project Description

Choose one of these project

## Project 1

## Description

- Use the methods and techniques studied in lecture and lab in order to apply a regression model on your selected dataset you like.
- The total mark of the project will be evaluated according to
    - The project delivery
    - Individual oral discussion
- Project delivery date& discussion will be on the practical exam date

## The project implementation will include the following

- Search for a real dataset with multiple attributes ( a min of 10 attributes & and min 4000 rows)
- Describe the data attributes
- Understand the topic of dataset and scientifically define your project objectives
- Apply data wrangling methods (if needed)
- Apply Data visualization charts to observe more knowledge about your dataset
- Apply a regression algorithm using gradient descent
- Tune model parameters to get high accuracy

## Project deliverables

- Dataset(.csv file)
- Code files( python file)
- Presentation including short description of your dataset , data preprocessing, visualization and your selected model with screenshots for each.

## Bonus

- Implement gradient descent algorithm from scratch and apply this model on your selected dataset.

## Project 2

## Description

- Use the methods and techniques studied in lecture and lab in order to apply a genetic algorithm model.
- The total mark of the project will be evaluated according to
    - The project delivery
    - Individual oral discussion
- Project delivery date& discussion will be on the practical exam date

## The project implementation will include the following

- Apply a genetic algorithm on the following problem
- In Timetabling, we have to allocate time for the activities we have planned and coordinate resource in an orderly way so that we can obtain our intended results without having to violate any constraints.
    - For example, a school timetable would coordinate student, teachers, classrooms, subjects and time slots.
- Consider you are trying to come up with a weekly timetable for classes in a college for different groups of students. We have to arrange classes and come up with a timetable so that there are no clashes between classes.
- Here, our task is to search for the optimum timetable schedule

.

## Project deliverables

- Code files( python file)
- Presentation including short description of your code implementation with screenshots for each. And illustrate the genetic algorithm steps.

## Bonus

- Implement a GUI in any programming language.

**Project 3**

**Description**

- Use the methods and techniques studied in lecture and lab in order to apply a Particle Swarm Optimization (PSO) algorithm model on this dataset. Audit dataset
- The total mark of the project will be evaluated according to
    - The project delivery
    - Individual oral discussion
- Project delivery date& discussion will be on the practical exam date

**The project implementation will include the following**

- Apply a PSO algorithm on the following problem
- The goal of the dataset is to help the auditors by building a classification model that can predict the fraudulent firm on the basis the present and historical risk factors. The information about the sectors and the counts of firms are listed respectively as Irrigation (114), Public Health (77), Buildings and Roads (82), Forest (70), Corporate (47), Animal Husbandry (95), Communication (1), Electrical (4), Land (5), Science and Technology (3), Tourism (1), Fisheries (41), Industries (37), Agriculture (200).
- This a case study of an external government audit company which is also the external auditor of government firms of India. During audit-planning, auditors examine the business of different government offices but the target to visit the offices with very-high likelihood and significance of misstatements. This is calculated by assessing the risk relevant to the financial reporting goals (Houston, Peters, and Pratt 1999). The three main objective of the study are as follow:

    - To understand the audit risk analysis work-flow of the company by in-depth interview with the audit employees, and to propose a decision-making framework for risk assessment of firms during audit planning.

    - To examine the present and historical risk factors for determining the Risk Audit Score for 777 target firms, to implement the Particle Swarm Optimization (PSO) algorithm to rank examined risk factors,

and evaluating the Risk Audit Class (Fraud and No-Fraud) of nominated firms..

**Project deliverables**

- Code files( python file)
- Presentation including short description of your code implementation with screenshots for each. And illustrate the PSO steps.

**Bonus**

- Implement a GUI in any programming language.

**Project 4**

**Description**

- Use the methods and techniques studied in lecture and lab in order to apply a regression model and data wrangling techniques to get the highest accuracy on the following dataset.
  Airline Tickets
- The total mark of the project will be evaluated according to
  - The project delivery
  - Individual oral discussion
- Project delivery date& discussion will be on the practical exam date

**The project implementation will include the following**

- Apply the most suitable data wrangling methods.
- Apply Data visualization charts to observe more knowledge about your dataset
- Apply a regression algorithm using gradient descent algorithm or any regression algorithm.
- Tune model parameters to get high accuracy

**Project deliverables**

- Dataset(.csv file)
- Code files( python file)
- Presentation including short description of your dataset, data preprocessing, visualization and your selected model with screenshots for each.

**Bonus**

- For the team who will get highest accuracy.