**Pattern Recognition Project**

The objective of the projects is to prepare you to apply different machine learning algorithms to real-world tasks. This will help you to increase your knowledge about the workflow of the machine learning tasks. You will learn how to clean your data, applying pre-processing, feature engineering, regression, and classification methods. Each project will be delivered in milestones.

➢ The best three teams will get Bonus .

➢ Teams that Apply new techniques will get Bonus .

➢ Registration ends: 08/05/2021 11:59 PM .

➢ Delivering Milestone 1: To Be Announced.

➢ Delivering Milestone 2: Practical exam.

➢ Minimum number of members is 3 and the maximum is 6.

➢ You must deliver a detailed report for each milestone contains all your work (feature analysis, algorithms used in each module and the achieved accuracy for each one)

**Note :** Each report will be graded

In the first milestone, you will apply the followings :-

**Preprocessing:** Before building your models, you need to make sure that the dataset is clean and ready-to-use.

**Regression:** Apply different regression techniques (at least two) to find the model that fit your data with minimum error.

1

**Milestone 1:**

➢ Preprocessing, Regression.

**Milestone 1 Report Must Include:**

❖ You must explain in detail the **preprocessing techniques** you needed to apply on your dataset and how you implemented them.

❖ Perform **analysis** on the dataset as studied and explain how the features affect and relate to each other.

❖ You must explain what **regression techniques** you used (at least two). ❖ Mention the **differences** between each model and the acquired **results** (accuracy/error and so on).

❖ You must clearly mention **what features** you used or discarded to create your regression models.

❖ Explain what the **sizes** of your training, testing and validation sets are, if exist.

❖ Mention any further techniques that were used to **improve** the results (if exist).

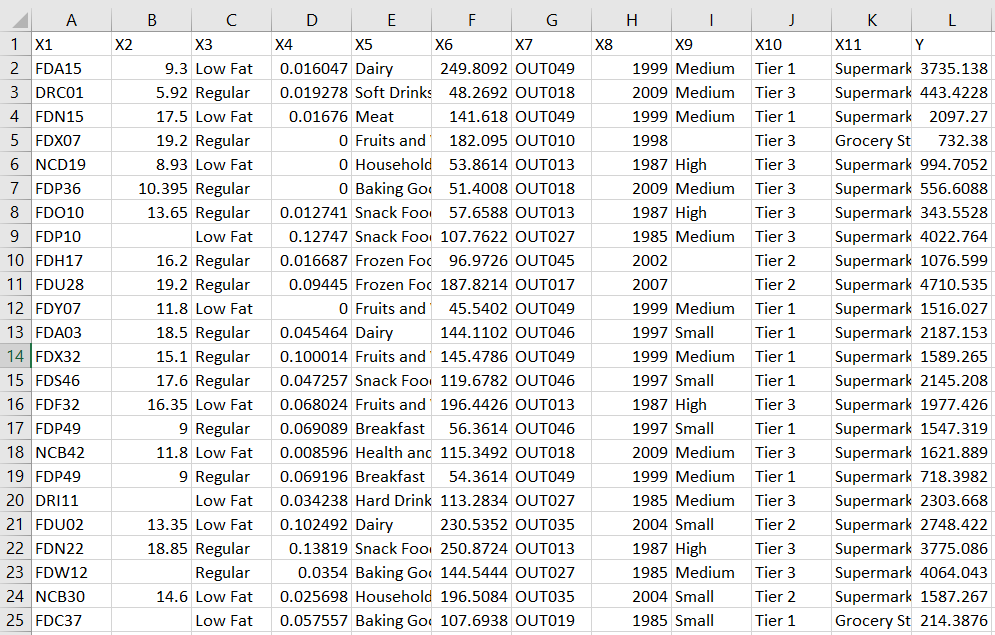
❖ You should include **screenshots** of the resultant(s) regression line plots. ❖ Finally, write a **conclusion** about this phase of the project and what intuition you had about your problem and how it was proved/disproved.

2

**Project : Predicting Store Item Sales**

Ever wonder why hypermarkets arrange the products in a similar way ? In fact this arrangement was done after a lot of research about the dependencies between buying some item X and another item Y . In this project you will investigate how the item and store features can affect sales of items .

**Dataset Snapshots:**

****

**Dataset Description:**

|  |  |
| --- | --- |
| **Feature** | **Description** |
| X1 | Item ID |
| X2 | Weight of Item |
| X3 | Amount of Fats in Item ( low fat or regular ) |
| X4 | The % of display area allocated for item in store |
| X5 | Item Category |
| X6 | Item Price |
| X7 | Store ID |
| X8 | Store Establishment Year |
| X9 | Store Size |
| X10 | Store Location Type (Type of City) |
| X11 | Store Type ( Supermarket , Grocery store , ..etc) |
| Y | Item Sales |

**Milestone 1 tasks:**

1. Apply pre-processing on the provided dataset.

2. Experiment with regression techniques to reduce the error on prediction of the sales of items in the store (Deliver at least two techniques).