[2024] Pattern Recognition Project (AI)

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 apply pre-processing, feature engineering, regression, and classification methods. Each project will be delivered in milestones

- The best three teams for this project will be honored.
- Registration starts: Wednesday 3/4/2024 11:59 PM.
- Registration ends: Saturday 6/4/2024.
- Delivering Milestone 1: 23/4/2024
- Delivering Milestone 2: Practical exam
- Minimum number of members is **3** and the maximum is **5**.
- Note: you can register as a team of **6 members**, but there will be an **additional** graded task which will be considered mandatory.
- You must deliver a detailed report for each milestone, containing all your work (feature analysis, algorithms used in each module and the accuracy achieved 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.

Milestone1: 50%

- Preprocessing and Regression

Milestone 1 Report Must Include:

- ❖ You must explain in details 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 (score/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.

Project: Doctor Fee Prediction

The "Doctor Fee prediction" projects aim to leverage advanced machine learning algorithms to accurately predict the Fee of the doctor based on some attributes.

Dataset Snapshot:

	Doctor Name	City	Specialization	Doctor Qualification	Experience(Years)	Total_Reviews	Patient Satisfaction Rate(%age)	Avg Time to Patients(mins)	Wait Time(mins)	Hospital Address
0	Prof. Dr. Syed Shamsuddin	QUETTA	Dermatologist	MBBS, MCPS, FCPS	30.0	827.0	98.0	14.0	13.0	National Hospital, Near Quetta Laboratory Prin...
1	Dr. Jaffar Ali	QUETTA	Dermatologist	MBBS, MCPS (Dermatology), FDV (Austria)	32.0	1250.0	97.0	16.0	13.0	Quetta Hospital, Quetta City, Quetta
2	Asst. Prof. Dr. Syed Bilal Shams	QUETTA	Dermatologist, Cosmetic Surgeon	MBBS, FCPS (Dermatology), Fellowship In Dermat...	15.0	1294.0	99.0	15.0	15.0	Skin Center Dermatologist Dr. Syed Bilal Ahmed...
3	Asst. Prof. Dr. Surjeet Kumar	QUETTA	Dermatologist	MBBS, FCPS(Dermatology)	12.0	114.0	99.0	14.0	7.0	Doctors Hospital, Quetta City, Quetta
4	Dr. Palwasha Jalil	QUETTA	Dermatologist	MBBS, FCPS (Dermatology) Agha Khan University	8.0	24.0	100.0	19.0	16.0	Aria Institute Of Medical And Health Sciences,...

~ Dataset Header Cont'd:

Hospital Address	Doctors Link	Fee(PKR)
National Hospital, Near Quetta Laboratory Prin...	https://www.marham.pk/doctors/quetta/dermatolo...	5000.0
Quetta Hospital, Quetta City, Quetta	https://www.marham.pk/doctors/quetta/dermatolo...	3000.0
Skin Center Dermatologist Dr. Syed Bilal Ahmed...	https://www.marham.pk/doctors/quetta/dermatolo...	5000.0
Doctors Hospital, Quetta City, Quetta	https://www.marham.pk/doctors/quetta/dermatolo...	2000.0
Aria Institute Of Medical And Health Sciences,...	https://www.marham.pk/doctors/quetta/dermatolo...	1500.0

Milestone 1 tasks:

1- Apply pre-processing on the provided dataset. You must preprocess all the features even if you will not include them in the model. You must use appropriate encoding techniques based on the situation (you should encode one column at least with one hot encoding and you can also try

other techniques that is appropriate). Note: using any non-logical encoding technique in order to train the model will cause you a penality.

2- Apply feature selection and experiment with regression techniques to reduce the error on prediction of the “Fee” (deliver at least two regression models with significant difference)
3- Apply hyper-parameter tuning in order to achieve better results for each experimented model.
4- Finish Milestone 1 Report.

Note: You must preprocess all features, but model and feature selection can be done after that (i.e. You can drop a feature only after preprocessing and with valid reason)