Project Phases 4 and 5

You designed and implemented your project in the previous phases. This phase enables you to enhance the performance of your model, evaluate your outcomes, and write a formal report. To reach these goals, please follow the three steps below:

1. It is a good idea to reevaluate your model after the initial set up with the aim of finding a way to improve the performance of your model. This process is commonly known as **optimization** process and can be consider as an umbrella term. The following bullet points describes a few ways that probably can enhance the performance of your model:
    a. Feature engineering
        i. Different splitting strategy [1]
        ii. Creating/ using new features/samples
        iii. Feature reduction
            1. Feature selection
                a. You can use either feature importance or any feature selection, say sequential feature selection. You can find variety of these algorithms through [2]
                b. Dimension reduction, e.g., PCA [3]
                c. Removing undiscovered outliers
        iv. Changing the structure of your model
            1. For traditional classifiers you can add weight to minority category of samples, e.g., weighted SVM [4]
            2. For regression methods you can go with higher regularization penalty. For instance, you can use elastic net instead of ordinary least squared model [5]
            3. For deep learning models, you have variety of options such as changing the structure of layers, the initialization function, and activation function. E.g., [6]
    b. Tunning the hyper-parameters of your model. E.g., adjustments for SVM [7]
    c. Adjusting performance metrics: For instance, it would be a great idea to use F1 score instead of accuracy. Code: [8]
    d. Using alternative models. Please help yourself with all models provided in sklearn [9]

2. **Model Evaluation**
    a. Feature importance: You can measure the role of each feature in your final model using either built-in or external feature importance algorithms. E.g., [10]
    b. Decision boundary visualization. E.g, [11]
    c. Evaluating the performance in detail
        i. Classification Task: Precision, recall and F-measures [12]
        ii. Regression Task: Residual Plot, e.g., [13]
        iii. Deep learning
            1. The performance visualization in different epochs
            2. Visualization of patterns in perceptron

a. Please refer to Deep learning lecture regarding the code

iv. Investigating the patterns among samples that are predicted
    1. Strongly successful
    2. Weakly successful
    3. Weakly wrong
    4. Far away that we expected

3. **Final project report**
    a. **Please complete your previous report by adding the sections of Optimization, Model Evaluation, and Conclusion and Future Work.** Please report your final project report using the following suggested structure:

Introduction
- **General Description**
- **Research Question**
- **GitHub Repository Address**

Dataset Description
- **URL of your dataset**
- **Where, when, and how the data were collected**
- **The name, definition, and characteristics of features**

Related Work
- **Advantages and disadvantages of 3 related works in comparison to each other**

Project Plan

Data Exploration
- **Univariate Analysis**
    - **Descriptive Analysis**
    - **Distribution Analysis**
    - **Outlier Detection**
- **Bivariate Analysis**
    - **Person correlation**
    - **Pair plot**

Data Modeling
- **Preprocessing**
- **Data Splitting**
- **Fitting the model**
- **Measuring Performance**
    - **For a classification task**
        - **Accuracy/MAE**
        - **Confusion matrix**
        - **ROC curve**
    - **For a regression task**
        - **Mean absolute error regression loss (MAE)**
        - **Mean squared error regression loss (MSE)**
        - **Mean squared logarithmic error regression loss (MSLE)**

Optimization
- **Please use at least three strategies from all the strategies described above**

Model Evaluation

- **Please use at least two of the strategies described above**

Conclusion and Future Work

- **Please describe what did you learned through this project in one paragraph**
- **Please describe what the other features can you add to this project to provide a better application**

References

[1]. https://scikit-learn.org/stable/modules/classes.html#module-sklearn.model_selection
[2]. https://scikit-learn.org/stable/modules/feature_selection.html#feature-selection
[3]. https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html
[4]. https://scikit-learn.org/stable/auto_examples/svm/plot_weighted_samples.html
[5]. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ElasticNet.html
[6]. https://keras.io/examples/
[7]. https://scikit-learn.org/stable/modules/grid_search.html
[8]. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html
[9]. https://scikit-learn.org/stable/index.html
[10]. https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html
[11]. https://scikit-learn.org/stable/auto_examples/ensemble/plot_voting_decision_regions.html
[12]. https://scikit-learn.org/stable/modules/model_evaluation.html#precision-recall-f-measure-metrics
[13]. https://www.scikit-yb.org/en/latest/api/regressor/residuals.html

Notice:

>>>>>> Each group should have one documentation and one Jupyter Notebook. The head of your team should upload them on the private GitHub repository.

>>>>>>Each student must upload his/her team documentation and Jupyter notebook on Blackboard.

>>>>>>You can use any material or GitHub repository in this project, you must cite them. For each missing citation, you are subjected to 2 points penalty.

Cover Page Sample

Project Name

Artificial Intelligence
Your course section

Your Name

Sacred Heart University
School of Computer Science & Engineering
The Jack Welch College of Business & Technology

Submitted To:
Dr. Reza Sadeghi

Spring 2022

Title Page Sample

Final Project Report of Project Name

Your Name

Your SHU Email
.....@sacredheart.edu

Short Bio

Table of Content Sample

Table of Contents

Notice:

Please use the Writing Center facility to automatically get these points. Otherwise, **each typographical or grammatical error will cost -1 point**.

Questions and problem handling:

You can ask any questions regarding the project. You can ask your questions during class, or you can email your questions to your instructor sadeghir@sacredheart.edu.