

ECE157A Final Project Phase2

ECE157A TAs

November 2021

Phase 2 Due: Friday, 10th of December
Submission: schedule live demo with TAs
Group: allow working as a group of 2-3 people

1 Introduction

You have set up the file uploading interface for your simple ML app in phase 1. You should have a good idea on how to spin up a web application with the Django framework now. In phase 2, you will continue to expand the back-end python code and the front-end HTML code to complete the analytic app. Remember our goal is to read in the CSV files and perform three kinds of ML analytics on them respectively. You can always refer back to the *Final Project Overview* PDF released in phase 1.

2 Run and Display Analytic

The idea here is to use the trained models (the classifier, the regressor, and the outlier model) to produce informative results such as predictions and then visualize the results in plots. You can be creative on what kinds of plots you want to show to the users (in this case demo to the TAs).

We also organize the three csv files used before in the folder *csv_files* for you to upload.

You can refer to the following example plots:

- **Classification (Homework 1):** Figure 1 plots the binary classification results on the 100 unknown data with two selected features, e.g., Blood Pressure and Glucose. The yellow and purple dots indicate the positive and negative predictions respectively.
- **Regression (Homework 2):** Figure 2 shows the predicted quality of 997 unknown wine data versus the ground truth labels. Note that for a perfect regression model, the plot should look like a line with slope one, meaning that the predicted labels are almost the same with the ground truth with little residuals.

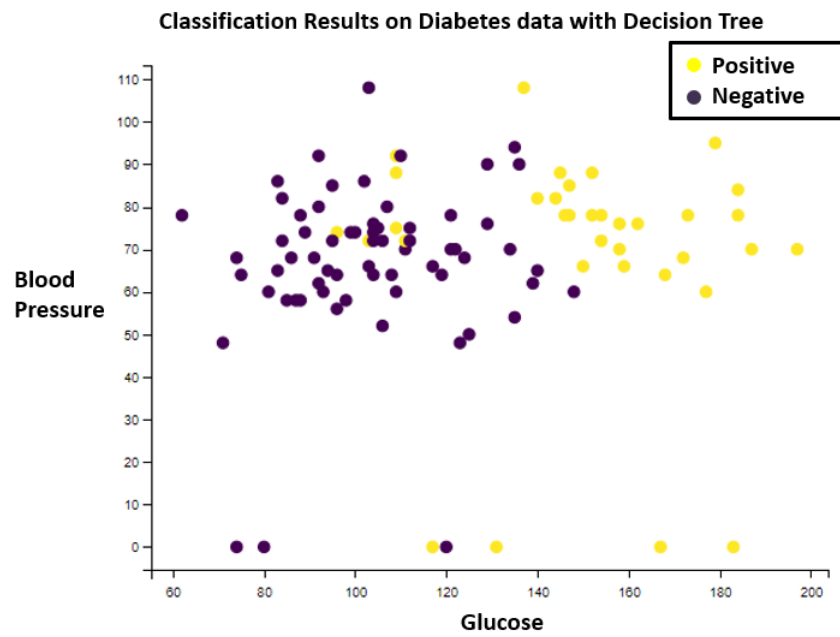


Figure 1: Example analytic plot for classification (hw1)

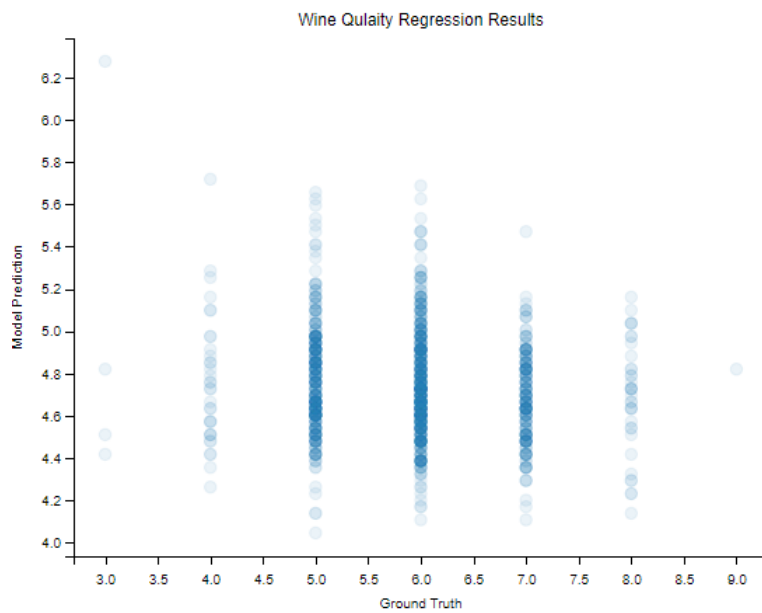


Figure 2: Example analytic plot for regression (hw2)

Outlier Detection on NBA Data with Elliptic Envelop

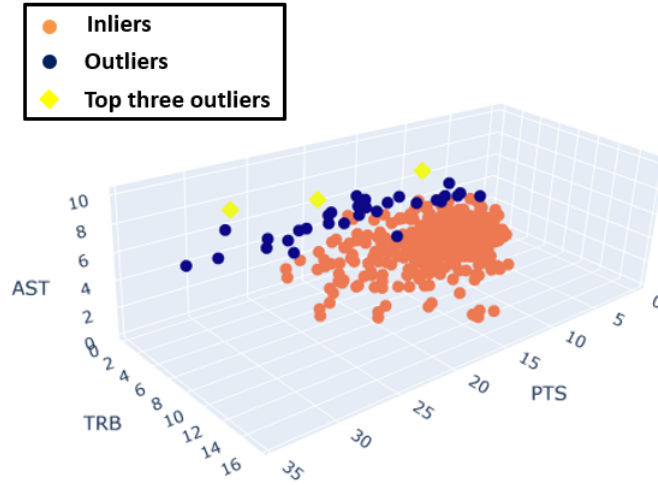


Figure 3: Example analytic plot for outlier detection (hw3)

- **Outlier Model (Homework 3):** Figure 3 shows the outlying versus inlying points as predicted by the outlier model, the same as in homework 3.

Note that undergrads need to demonstrate the interactivity of the first two plots during demo, in the way that you can zoom in and out the figure and drag the canvas to focus on specific region.

You can call functions/modules from the python scripts you wrote for the homework. Keep in mind that when the client sends the request to the server asking for running a specific analytic, it takes time for the server to process the task and reply to the client. Our assumption here is that the analytic (both training and inference) are running very fast within seconds so we can block our server and wait for the ML modules return. In reality, we will need to implement multi-threading to separate the analytic running into another thread, so it won't interfere with other routines of the server. For example, the user can upload new files, review analytic history etc. while the analytic is running at the back-end.

Refer to the phase 2 guide for implementation detail.

3 Rubric

3.1 Undergrads

- List uploaded CSVs (1 pts each, 3 max)
- List three analytics
 - Homework 1 diabetes predictor (1 pt)
 - Homework 2 wine quality regressor (1 pt)
 - Homework 3 NBA outlier model (1 pt)
- Display *interactive* result plots for Homework 1 and Homework 2 analytic results *on uploaded CSV* (5 pts each)
 - Missing labels/axes deducts one point per
 - Missing interactivity deducts two points
- Display 3d plot (may be static) for Homework 3 analytic (5 pts)
 - Missing labels/axes deducts one point per

Undergrad max points: 21

3.2 Grad additional requirements:

- Check file extensions before uploads (1 pt)
- Demonstrate deleting of uploaded files
 - From model (1 pt)
 - From disk (1 pt)
- Demonstrate algorithm management in Django /admin screens (2 pts)
- Demonstrate running analytic on *most recent* copy of a CSV. That is, if two CSVs are uploaded with the same name, only the latest one should be used for analytic running. (1 pt)
- All of your plots must have interactivity. This includes the 3d plot. (-2 pts if missing.)
- Show a list of recently run analytic results (2 pts)
- Show results of recently run analytics on click of list (3 pts)

Grad max points: 32

4 Demo

TA will send out the Google sheet for signing up the final demo. Stay tuned on Piazza!