**ENGR 421/DASC 521:** Introduction to Machine Learning
**Homework 6:** Are Under the ROC Curve
**Deadline:** January 20, 2024, 11:59 PM

In this homework, you will implement the receiver operating characteristics (ROC) curve in Python. Here are the steps you need to follow:

1. Read Section 20.7 from the textbook.

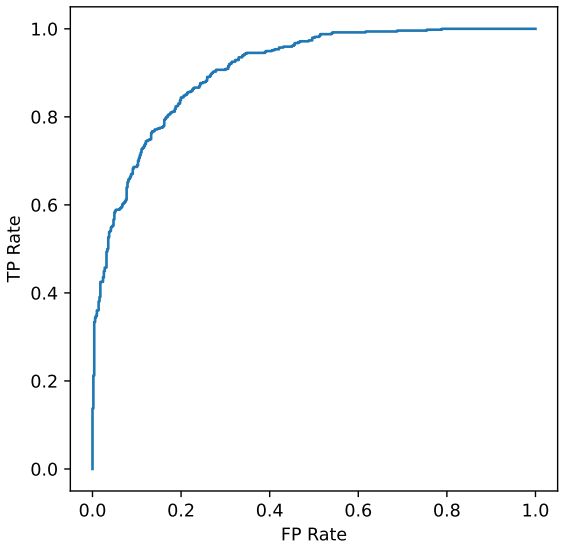
2. You are given the true labels of 1000 data points for a binary classification problem in the file named `hw06_true_labels.csv` and the predicted posterior probabilities of these 1000 data points for the positive class in the file named `hw06_predicted_probabilities.csv`.

3. Calculate possible classification thresholds using the predicted posterior probabilities. (20 points)

```
print(thresholds)
[0.00603665 0.01329955 0.02200585 ... 0.9655685  0.9670935  0.9836335 ]
```

4. Calculate FP and TP rates using the true labels, predicted posterior probabilities, and thresholds. (40 points)

```
print(fp_rates)
print(tp_rates)
[1.         0.99802372 0.99604743 ... 0.         0.         0.        ]
[1.         1.         1.         ... 0.00404858 0.00202429 0.        ]
```




5. Calculate the area under the ROC curve using the FP and TP rates. (40 points)

```
print("The area under the ROC curve is {}.".format(auroc))
The area under the ROC curve is 0.9064465283000738.
```

---

**What to submit:** You need to submit your source code in a single file (`.py` file). You are provided with a template file named as `0099999.py`, where `99999` should be replaced with your 5-digit student number. You are allowed to change the template file between the following lines.

```
# your implementation starts below

# your implementation ends above
```

**How to submit:** Submit the file you edited to Blackboard by following the exact style mentioned. Submissions that do not follow these guidelines will not be graded.

**Late submission policy:** Late submissions will not be graded.

**Cheating policy:** Very similar submissions will not be graded.

---