

Small Project 1: Logistic Regression and SVM

CS4275: Machine Learning Foundations, Fall 2024

Weight: 15%

Submission Deadline: September 17, 2024, 11.59 PM

Logistic Regression vs. SVM with wine dataset:

1. Use the shared code ML_3.py and ML_4.py as base code. You can use the attached dataset or the following code:

```
from sklearn.datasets import load_wine
wine = load_wine()
wine.target[[10, 80, 140]]
print(list(wine.target_names))
```

```
X = wine.data[:, [0, 2]]
y = wine.target
```

```
print("X ", X)
print("y ", y)
```

2. Load (all features) and Split the Iris dataset to training and test sets with ratio 70% and 30%, respectively.
3. Fit your data using Logistic Regression in sklearn. Then plot the decision regions with $C=10, 100, 1000, 5000$ (generate one figure for each value of C). The x-axis and y-axis of each figure correspond to feature "Alcohol" and "acid", correspondingly.
4. Plot the accuracy_score (in sklearn) of your predictor with respect to $C=10-x$, $x=-4, -3, \dots, 3, 4$.
5. Repeat steps 1~4 above, using SVM instead of logistic regression, with kernel "RBF", for two cases of γ (gamma)=0.1 and 10.
6. Compare two types of predictors (give insight comments)

Deliverables:

1. Error free source code (5%).

2. Figures: Generate results for all combinations (5%).
3. Report: all results will be accumulated in a document file with technical insight and **your own comments** (5%).

Submission guidelines:

Option one: You can upload ipynb file code and report.

Option two: You can upload py file code and report.

Create a zip file that will contain all of the files compiling from either option one or option two.

Zip file name: COMP4275_SM1_YourNAME_YourID.zip

Upload the zip file to assignment Small Project 1 in Moodle.

Deadline: September 17, 2024, 11.59 PM

If you have any questions or discussion utilize the classroom meeting and my office hours.