ENGR-UH 3332
Applied Machine Learning

Mini Project 2 - SVM for classification

Due Date: Refer to Brightspace

Introduction

Support vector machine is another simple algorithm that every machine learning expert should understand. Support vector machine is highly preferred by many as it produces significant accuracy with less computation power. Support Vector Machine abbreviated as SVM can be used for both regression and classification tasks. But it is widely used in classification objectives. The objective of the support vector machine algorithm is to find a hyperplane in N-dimensional space (N — the number of features) that distinctly classifies the data points. To separate the two classes of data points, many possible hyperplanes could be chosen. Our objective is to find a plane that has the maximum margin.

Dataset

Generate a synthetic dataset using scikit-learn's blob generator

```
X, y = make_classification(n_samples=500, n_features=3, n_informative=3,
                           n_redundant=0, n_clusters_per_class=1,
                           flip_y=0.1,  # adds a small amount of noise
                           class_sep=1.0,  # classes are separable but not too easily
                           random_state=40)
```

Requirements

1. Relabel the Y targets to +1/-1
2. Split the dataset into training and testing datasets
3. Implement soft margin SVM
4. Use mini batch gradient descent to minimize the loss function on the next page (shuffle the data first).
5. Return the optimal weights by minimizing the loss function
6. Perform some predictions on the test data
7. Calculate the accuracy score
8. Visualize the training data and decision boundary in 3D
9. Visualize the loss function over time during training

$$J(\mathrm{w}) = \frac{1}{2}\|\mathrm{w}\|^2 + C\left[\frac{1}{N}\sum_i^n \max\left(0, 1 - y_i * (\mathrm{w}\cdot x_i + b)\right)\right]$$

$$\nabla_w J(\mathrm{w}) = \frac{1}{N}\sum_i^n \begin{cases} \mathrm{w} & \text{if } \max(0,\ 1 - y_i * (w\ .\ x_i + b)) = 0 \\ \mathrm{w} - C y_i x_i & \text{otherwise} \end{cases}$$

gradient with respect to bias = 0 if $\max(0, 1 - y_i * (w . x_i + b)) = 0$
else: gradient with respect to bias = $-C * y_i$

Deliverables

1. A working project (Python Notebook or .py code)
2. Embed markup cells with detailed comments about your source code.
   i. Instructions on how to run your project
   ii. Answers to the programming questions.

Before submitting your project, please make sure to test your program on the given dataset.

Notes

You may discuss the general concepts in this project with other students, but you must implement the program on your own. ***No sharing of code or reports is allowed.*** *Violation of this policy can result in a grade penalty.*

Late submission is acceptable with the following penalty policy:
10 points deduction for every day after the deadline