HOMEWORK 7
OPTIMIZATION AND SEARCH *

10-607 COMPUTATIONAL FOUNDATIONS FOR MACHINE LEARNING

START HERE: Instructions

- **Collaboration Policy**: Please read the collaboration policy in the syllabus.
- **Late Submission Policy:** See the late submission policy in the syllabus.
- **Submitting your work:** You will use Gradescope to submit answers to all questions and code.
  - **Written:** For written problems such as short answer, multiple choice, derivations, proofs, or plots, please use the provided template. Submissions can be handwritten onto the template, but should be labeled and clearly legible. If your writing is not legible, you will not be awarded marks. Alternatively, submissions can be written in LaTeX. Each derivation/proof should be completed in the boxes provided. To receive full credit, you are responsible for ensuring that your submission contains exactly the same number of pages and the same alignment as our PDF template.
  - **Latex Template:** `https://www.overleaf.com/read/tkffcnbrscpg`
  - **Programming:** You will submit your code for programming questions on the homework to Gradescope. After uploading your code, our grading scripts will autograde your assignment by running your program on a virtual machine (VM). Unless otherwise specified, you are only permitted to use the Python Standard Library modules and `numpy`.
- **Materials:** The data and reference output that you will need in order to complete this assignment is posted along with the writeup and template.

Question	Points
Search	5
Unconstrained Loss Optimization for ML (Programming)	15
Constrained Loss Optimization for ML (Theory)	6
Total:	26

*Compiled on Sunday 1st December, 2024 at 15:52

Instructions for Specific Problem Types

For "Select One" questions, please fill in the appropriate bubble completely:

Select One: Who taught this course?

- ● Matt Gormley
- ○ Marie Curie
- ○ Noam Chomsky

If you need to change your answer, you may cross out the previous answer and bubble in the new answer:

Select One: Who taught this course?

- ● Henry Chai
- ○ Marie Curie
- ⊗ Noam Chomsky

For "Select all that apply" questions, please fill in all appropriate squares completely:

Select all that apply: Which are scientists?

- ■ Stephen Hawking
- ■ Albert Einstein
- ■ Isaac Newton
- □ I don't know

Again, if you need to change your answer, you may cross out the previous answer(s) and bubble in the new answer(s):

Select all that apply: Which are scientists?

- ■ Stephen Hawking
- ■ Albert Einstein
- ■ Isaac Newton
- □ I don't know

For questions where you must fill in a blank, please make sure your final answer is fully included in the given space. You may cross out answers or parts of answers, but the final answer must still be within the given space.

Fill in the blank: What is the course number?

10-606 10-6067

1 Search (5 points)

For each statement, state which search algorithm it applies to. Multiple choices may be selected.

1. (1 point) Commonly used to decode a vector during machine translation

    - □ Beam Search
    - □ A* Search
    - □ Greedy Search
    - □ None of the above

2. (1 point) Requires use of heuristics in its algorithm

    - □ Beam Search
    - □ A* Search
    - □ Greedy Search
    - □ None of the above

3. (1 point) Involves using BFS to build a search tree

    - □ Beam Search
    - □ A* Search
    - □ Greedy Search
    - □ None of the above

4. (1 point) Guaranteed to find shortest path on a weighted graph regardless of the heuristic used

    - □ Beam Search
    - □ A* Search
    - □ Greedy Search
    - □ None of the above

5. (1 point) Guaranteed to find shortest path on a weighted graph given an admissible and consistent heuristic

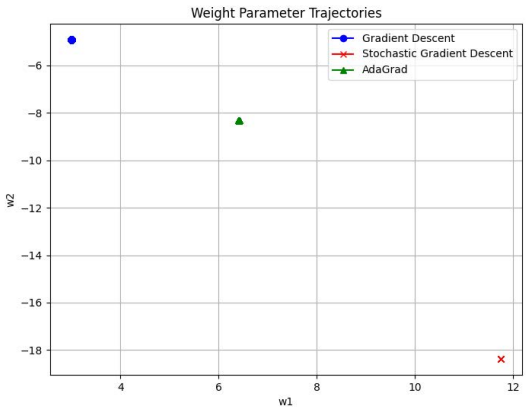
    - □ Beam Search
    - □ A* Search
    - □ Greedy Search
    - □ None of the above

2 Unconstrained Loss Optimization for ML (Programming) (15 points)

In this question, you will implement *Gradient Descent*, *Stochastic Gradient Descent*, and *AdaGrad*. Please code the update rules for each of these optimization techniques in the `loss_optimization.py` file provided to you. After you are done, you will submit your code to Gradescope, where we wil run your code on a full suite of tests and the autograder will assign your code points based on whether it passes the tests. After implementing the update rules, please answer the following questions:

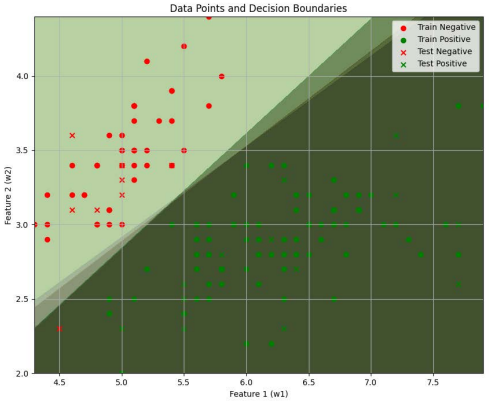
1. (3 points) Create a plot for the 2D parameter space $\mathbf{w}_1, \mathbf{w}_2$ corresponding to the non-bias weight parameters. These are the second and third elements in the weight vector since the first element $\mathbf{w}_0$ corresponds to the bias term. In this plot, show the trajectory of the weight parameters captured at the end of each epoch as they are updated during the training process. The plot should include three trajectories, one for each of the three parameter update rules. Include a legend indicating which trajectory corresponds to which optimizer.

Plot



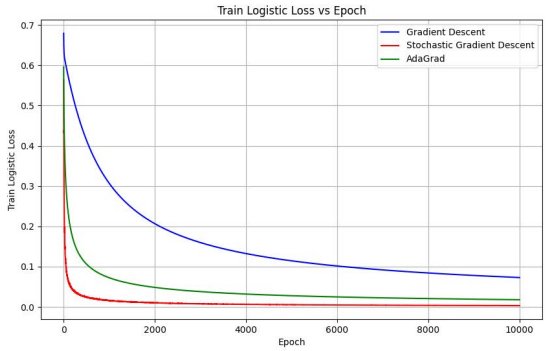
2. (3 points) Plot the two-dimensional data with the positive datapoints (label=1) in green and negative datapoints (label=0) in red. Display the train and test datapoints using different market shapes. Also include in the plot the classifier decision boundary learned by the three different optimizers. Include a legend indicating which decision boundary corresponds to which optimizer.

Plot



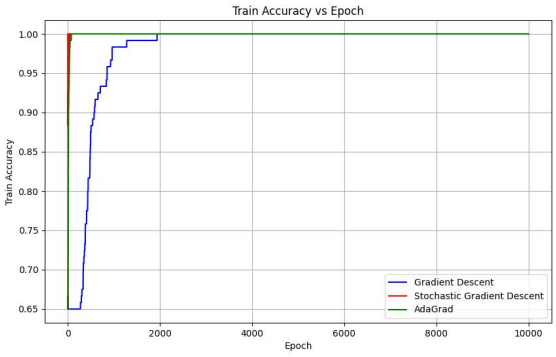
3. (3 points) Plot the curve for train logistic loss versus train epoch for each of the three optimizers in a single plot. Include a legend indicating which curve corresponds to which optimizer.

Plot



4. (3 points) Plot the curve for train accuracy versus train epoch for each of the three optimizers in a single plot. Include a legend indicating which curve corresponds to which optimizer.

Plot



5. (1 point) What is the accuracy obtained by the gradient descent optimizer on the test data?

Solution

96.67%

6. (1 point) What is the accuracy obtained by the stochastic gradient descent optimizer on the test data?

Solution

96.67%

7. (1 point) What is the accuracy obtained by the Adagrad optimizer on the test data?

Solution

96.67%

3 Constrained Loss Optimization for ML (Theory) (6 points)

State True/False for the statements provided below. Provide a reason for your choice is the box accompanying each statement.

1. (2 points) Simplex algorithm is an example of an interior point method for convex optimization. **Select one:**

   ◯ True

   ◯ False

   Solution

2. (2 points) In linear programming i.e. optimization problems with a linear objective and linear equality or inequality constraints, the optimal point lies at the boundary of the feasible set if the feasible set is bounded. Assume that all inequalities in the considered linear programming problem are non-strict i.e. of the form $\mathbf{Ax} \leq \mathbf{b}$. Feasible set is the set of all points that satisfy the constraints of the linear programming problem. **Select one:**

   ◯ True

   ◯ False

   Solution

3. (2 points) In linear models such as linear regression or logistic regression, stochastic gradient descent involves iteratively performing a gradient descent update on the scalar weight parameter associated with a randomly chosen feature. **Select one:**

   ◯ True

   ◯ False

   Solution

4 Collaboration Questions

After you have completed all other components of this assignment, report your answers to these questions regarding the collaboration policy. Details of the policy can be found in the syllabus.

1. Did you receive any help whatsoever from anyone in solving this assignment? If so, include full details.
2. Did you give any help whatsoever to anyone in solving this assignment? If so, include full details.
3. Did you find or come across code that implements any part of this assignment? If so, include full details.

Your Answer

1. I did not recieve any help whatsoever.
2. I did not give any help whatsoever.
3. I did not find or come across any code.