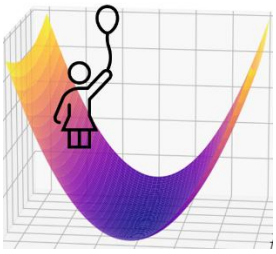
Aim

- Learn the differences between key optimization concepts and techniques.
- Apply your knowledge of different concepts to optimize a machine learning prediction model.
- Apply your knowledge of optimization techniques to solve global optimization problems.

Assessment

To pass this lab demonstrate your solutions at the lab.

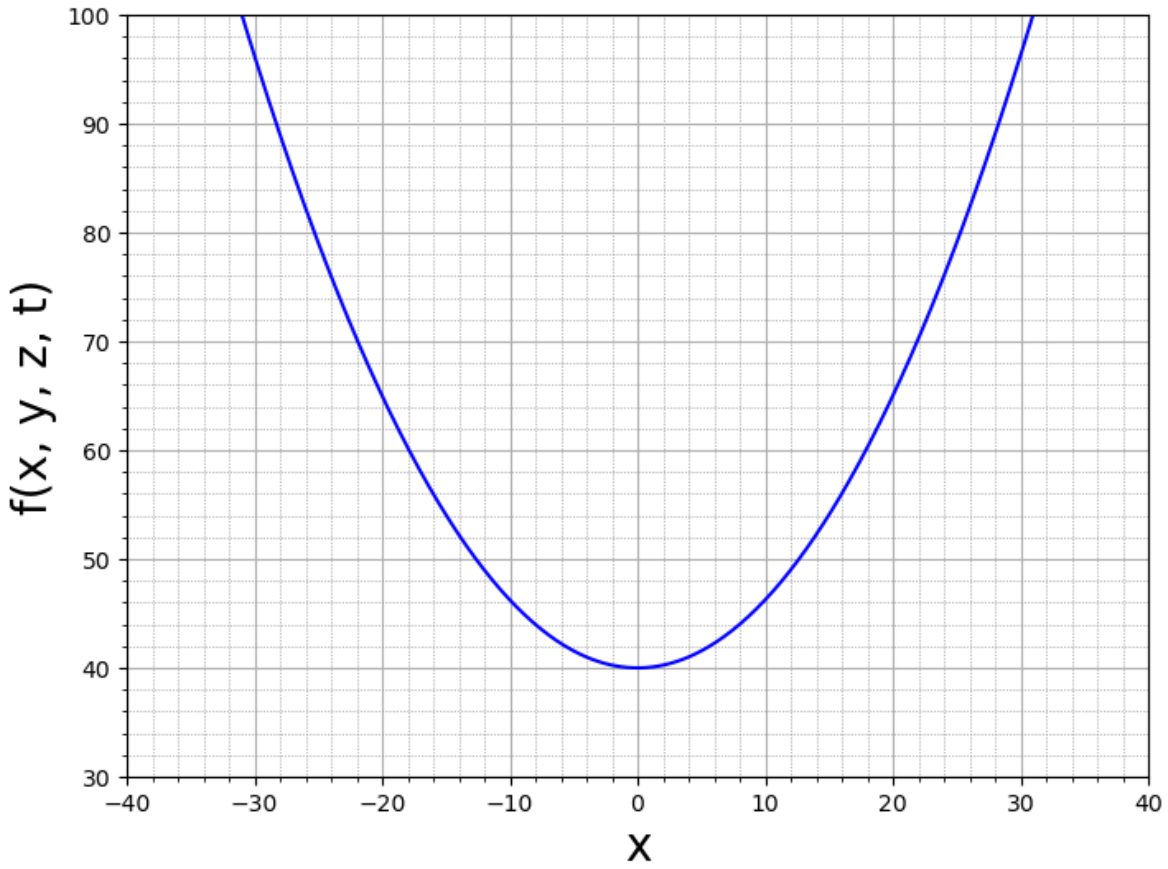
Requirements	Grade
Tasks 1, 2, 3	3
Tasks 1, 2, 3 and (4 or 5)	4
All tasks	5



Good Luck!

Task 1

The figure depicts the cost function $f(x, y, z, t) = \frac{x^2}{16} + 6y^2z + 12t^2 + 40$ across one feature of the training data.



To minimize the cost function across this feature, a novice ML engineer starts from the point $x = -16$ and performs three different sequences of steps as described in the following table.

1	3 iterations of steps along the gradient descent vector with learning rate = 20
2	3 iterations of steps along the gradient descent vector with learning rate = 10
3	4 iterations of steps along the gradient descent vector with learning rate = 4

Which strategy (or strategies) do you suggest the engineer to avoid and why?

Task 2

Specify which of the figures depict the following optimization-related phenomena?

A) Convergence of the cost function to the global minimum.
B) Equal progress of the cost function convergence across two features.
C) Premature convergence of the cost function at a saddle point
D) Overshoot
E) Escape from a saddle point
F) Very slow convergence

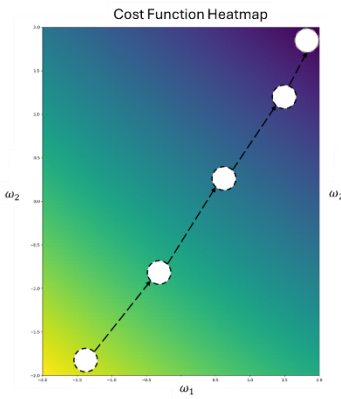


Figure 1

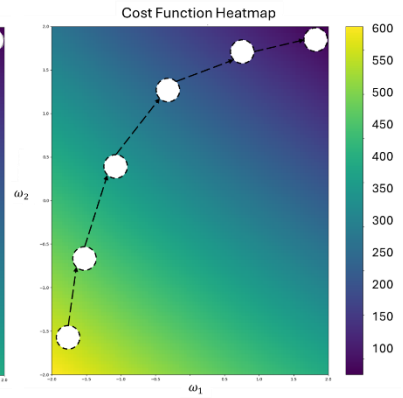


Figure 2

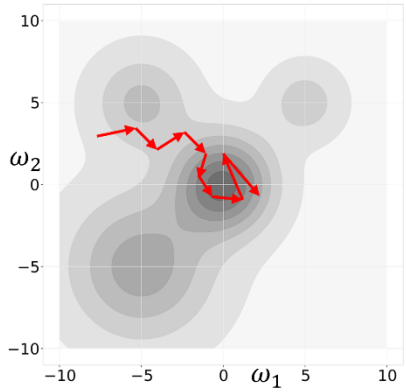


Figure 3

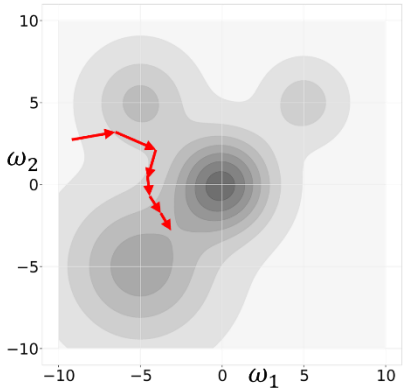


Figure 4

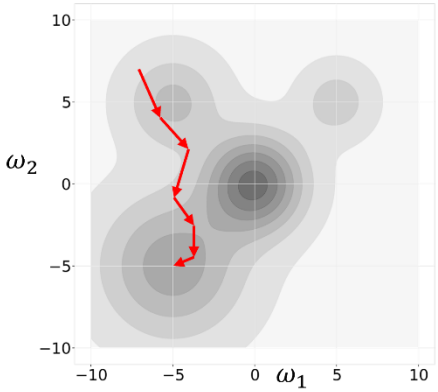


Figure 5

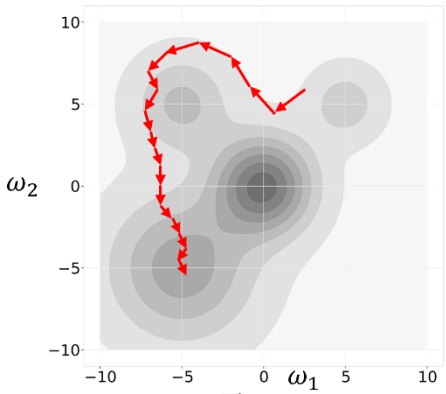


Figure 6

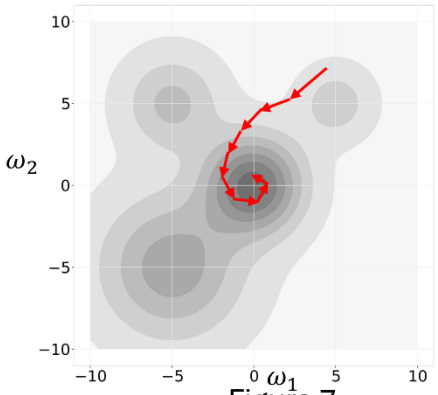


Figure 7

Task 3

Download the Python code and data for task 3.

1) Load the data, specify the number of features and convert the data into numpy arrays. (consider the last column as labels or values of the y)

2) Rename the functions to the correct optimization technique being applied.

3) Which of the techniques lead to the convergence of the cost function in less than 20 iterations?

4) For those that will not converge modify the hyperparameters so that they converge before 20 iterations.

Task 4

Download the data for Task 4.

1) Write a Python program that fits a linear regression model by minimizing the mean squared error (MSE) using the gradient descent optimization algorithm.
2) Present the final form of your linear regression model.
3) Report on the number of iterations required for the algorithm to converge (i.e., when the cost function converges).

Task 5

Download the data for Task 5.

Note that this data is saved in parquet format. To be able to load the data you need to install fastparquet using "conda install fastparquet" and then run the following command:

```
Import pandas as pd

pd.read_parquet('Task_5_data.parquet')
```

1) Write a Python program that fits a linear regression model by minimizing the mean squared error (MSE) using the mini-batch gradient descent optimization algorithm.
2) Present the final form of your linear regression model.
3) Report on the number of iterations required for the algorithm to converge (i.e., when the cost function converges).