

Introduction to Machine Learning (Gradient Descent, Logistic Regression)

Lab assignment #3 (worth 15%; deadline: Friday, November 08 at 11:59pm)

Learning objectives

After completing this lab assignment students should be able to:

- Understand and implement gradient descent
- Understand and apply logistic regression for binary classification
- Be familiar with common techniques in python

Preliminary Task – (25 Marks)

Similarly to Lab Assignment 1 and 2, each group must fill in the **certificate of work**, which is in the given notebook of this lab assignment (5 marks). Each group is responsible for maintaining a clean Git repository similar to the Bookkeeping tasks from Assignment 1, E.g., clear project structure, consistent version control, include only necessary files etc. (10 marks). Make sure the code is readable, E.g., consistent naming conventions, proper commenting, clear code structure, consistent Indentation and Spacing etc. (10 marks)

Task 1 – Gradient Descent (25 Marks)

Similar to Assignment 2, this task involves performing Single Point Positioning (SPP) using Gradient Descent instead of Least Squares Estimation (LSE). To build this algorithm:

- Please go through the Gradient Descent section in the notebook 'Assignment 3' and complete the related questions.
- For groups of 2 and 3, you can choose to omit this task, and you will be graded out of 75 marks in total. However, Gradient Descent **IS AN INTEGRAL PART OF QUIZ 2**, and it is **STRONGLY RECOMMENDED** to attempt the task, as it will greatly enhance your understanding of the concept and better prepare you for the quiz.
- Every team is required to use *Git* to collaborate on this assignment and ensure proper distribution of work.

Task 2 – Logistic Regression (50 marks)

Phishing websites are very common in the internet age, where the goal is to deceive users into providing sensitive information such as passwords, credit card numbers, or personal details by posing as trusted

entities. This task involves using a dataset containing samples of phishing and non-phishing websites and training a logistic regression model to correctly predict the phishing websites (binary classification).

- Please go through the Logistic Regression section in the notebook 'Assignment 3' and complete the related questions.
- Every team is required to use *Git* to collaborate on this assignment and ensure proper distribution of work.

Submission Materials

All the listed materials **must** be submitted in the *Lab Assignment 3 D2L* dropbox by the deadline. **Late** and **incomplete** submissions will result in a deduction of marks.

- A zip file of the `Lab_Assignment3` folder, which should include the Assignment 3 *.ipynb* file
- A link to the group GitHub repository
- **NO REPORT** is required for this assignment

As stated from "*Additional Lab Information*", please remember that:

- Each member is **required** to make at least one commit and push to the group repository. Failure to do so will result in a 1% **individual** course grade deduction.
- The D2L submission dropbox will close 24 hours after the deadline. All submissions in this 24-hour window are subject to a 20% lab grade deduction.
- There must not be any commits to the `Lab_Assignment3` folder in your GitHub repository after the deadline, or there will be a 20% lab grade deduction.