

# Machine Learning Project 1 - Fall 2020

Julien Hu, Matthieu Masouye, Sebastien Ollquist  
*Department of Computer Science, EPFL, Switzerland*

**Abstract—Implement basic Machine Learning methods on a given data set and analyze the predictions from it.**

## I. INTRODUCTION

The goal of this first mini project is to implement all basic Machine Learning methods on a given data set and analyze the results we obtained from running these algorithms. Essentially, the demanded algorithms were:

- 1) Linear regression using Gradient Descent and Stochastic Gradient descent
- 2) Least squares regression and ridge regression using normal equations
- 3) Logistic regression using Gradient Descent
- 4) Regularized logistic regression using Gradient Descent

## II. ALGORITHMS IMPLEMENTATION DETAILS

### A. Linear regression

This first algorithm is essential to Machine Learning. It consists of taking a data set that often contains two different data point types and split them using a line described by a linear function in order to divide the points the best way possible. We have performed two different implementations of it: one using Gradient Descent and the other one using Stochastic Gradient Descent.

Note that the Gradient Descent implementation does not work due to the fact that we are treating a big amount of data, so the Stochastic Gradient Descent will help us resolve that problem by only taking a batch of for example 50 randomly selected data samples.

If we only import the data as it is, we have a problem: the loss function does not converge. To resolve the problem, we essentially have to standardize the data, that is given a variable  $X$ , we compute the value  $Y = (X - \mu)/\sigma$  where  $\mu$  is the computed mean and  $\sigma$  the standard deviation.

### B. Least squares regression

### C. Ridge regression

### D. Logistic regression

## III. RESULTS OBTAINED

## IV. CONCLUSION