ΥΣ19 Artificial Intelligence II (Deep Learning for Natural Language Processing)
Fall Semester 2020
Homework I

Nikolaos Galanis - sdi1700019

OCtober 2020

1 Partial derivatives of the ridge regression loss function

Our goal is to compute the partial derivatives of the following function:

$$J(w) = MSE(w) + a\frac{1}{2}\sum_{i=1}^{n} w_i^2$$

where $MSE(w) = \frac{1}{m}\sum_{i=1}^{m}(h_w(x^{(i)} - y^{(i)})^2$, the minimum squared error loss function.

The partial derivatives of this loss function are given by the following equality:

$$\frac{\partial}{\partial w_j}MSE(w) = \frac{2}{m}\sum_{i=1}^{m}(w \cdot x^{(i)} - y^{(i)})x_j^{(i)}$$

The partial derivatives of the second term of the equation, $F(w) = a\frac{1}{2}\sum_{i=1}^{n} w_i^2$, are given by:

$$\frac{\partial}{\partial w_j}F = \frac{1}{2}a \cdot \frac{\partial}{\partial w_j}(w_1^2 + \cdots + w_m^2) = a * w_j$$

Thus the partial derivatives of the ridge regression loss function are:

$$\frac{\partial}{\partial w_j}J(w) = \frac{2}{m}\sum_{i=1}^{m}(w \cdot x^{(i)} - y^{(i)})x_j^{(i)} + a * w_j$$

Finally, the gradient of this loss function is:

$$\nabla_w J(w) = \frac{2}{m}(X^T(Xw - y)) + aw$$

2 Gradient descent implementations

The goal of this exercise is to implement gradient descent, stochastic gradient descent and mini-batch gradient descent methods for ridge regression. We re going to implement them by defining the necessary functions, and then demonstrate their performance using a specific dataset, and appropriate visualization techniques.

All of the code is written in the notebook GradientDescent, along with various plots and comments

3 Sentiment classifier using logistic regression

The goal of this exercise is to develop a sentiment classifier using logistic regression for the Twitter sentiment classification dataset that will be presented above. We are going to combine a plethora of features, and see how it performs on the given dataset, aiming to high accuracy. We are going to use tools from Scikit-Learn. Afterwards, we are going to evaluate the classifier using several metrics.

All of the code is written in the notebook sentiment_classification, along with various plots, comments, and metrics to evaluate the classifier