

# **CSC411: Assignment 3**

Due on Sunday, December 3<sup>rd</sup>, 2017

Student Name: **Gokul K. Kaushik**

Student Number: **999878191**

## Table of Contents

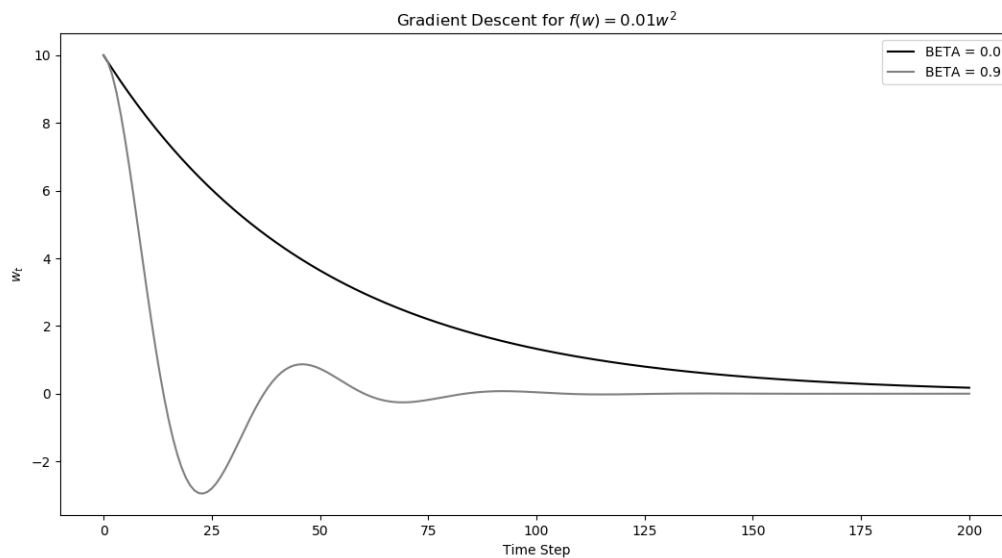
<b>1 - 20 Newsgroup Predictions</b>	<b>1</b>
<b>2 - Training SVM with SGD</b>	<b>1</b>
2.1 - SGD with Momentum . . . . .	1
2.2 - Training SVM . . . . .	2
2.3 - Apply 4-vs-9 Digits on MNIST . . . . .	2
2.3.1 - Training Loss . . . . .	2
2.3.2 - Test Loss . . . . .	2
2.3.3 - Classification Accuracy on the Training Set . . . . .	2
2.3.4 - Classification Accuracy on the Test Set . . . . .	2
2.3.5 - Plot $w$ as a $28 \times 28$ image . . . . .	2
<b>3 - Kernels</b>	<b>2</b>
3.1 - Positive Semi definite and Quadratic Form . . . . .	2
3.2 - Kernel Properties . . . . .	3
3.2.1 - Prove Property . . . . .	3
3.2.2 - Prove Property . . . . .	3
3.2.3 - Prove Property . . . . .	3
3.2.4 - Prove Property . . . . .	3

## 1 - 20 Newsgroup Predictions

## 2 - Training SVM with SGD

### 2.1 - SGD with Momentum

Plotting  $w_t$  for each time step  $t$  by applying the iterative stochastic gradient-descent on  $f(w) = 0.01w^2$ , we get the following graph (for  $\beta = 0.1$  and  $\beta = 0.9$  for upto 200 time-steps):



## 2.2 - Training SVM

## 2.3 - Apply 4-vs-9 Digits on MNIST

### 2.3.1 - Training Loss

### 2.3.2 - Test Loss

### 2.3.3 - Classification Accuracy on the Training Set

### 2.3.4 - Classification Accuracy on the Test Set

### 2.3.5 - Plot $w$ as a $28 \times 28$ image

## 3 - Kernels

### 3.1 - Positive Semi definite and Quadratic Form

Prove that a symmetric matrix  $K \in \mathbb{R}^{d \times d}$  is a positive semi definite iff for all vectors  $x$  we have  $\mathbf{x}^T K \mathbf{x} \geq 0$ .

Proof:

$$K \mathbf{x} = \lambda \mathbf{x}$$

where  $\lambda$  is the eigenvalue and  $v$  is the eigenvector.

Suppose  $\mathbf{x}$  is an eigenvector of  $K$  and replacing  $K \mathbf{v}$  with  $\lambda \mathbf{v}$  (from the definition of an eigenvector and eigenvalue above):

$$\mathbf{x}^T K \mathbf{x} = \mathbf{x}^T \mathbf{x} \lambda$$

$$\mathbf{x}^T K \mathbf{x} = |\mathbf{x}|^2 \lambda$$

Therefore, as  $|\mathbf{x}|^2 \geq 0$ , for the equation  $\mathbf{x}^T K \mathbf{x} \geq 0$ , the eigenvalue must be:  $\lambda \geq 0$ . Since it is a semi-definite matrix, where the eigenvalues  $\geq 0$ , this holds true.

## **3.2 - Kernel Properties**

**3.2.1 - Prove Property**

**3.2.2 - Prove Property**

**3.2.3 - Prove Property**

**3.2.4 - Prove Property**