

CSC411: Assignment 3

Due on Sunday, December 3rd, 2017

Student Name: **Gokul K. Kaushik**

Student Number: **999878191**

Table of Contents

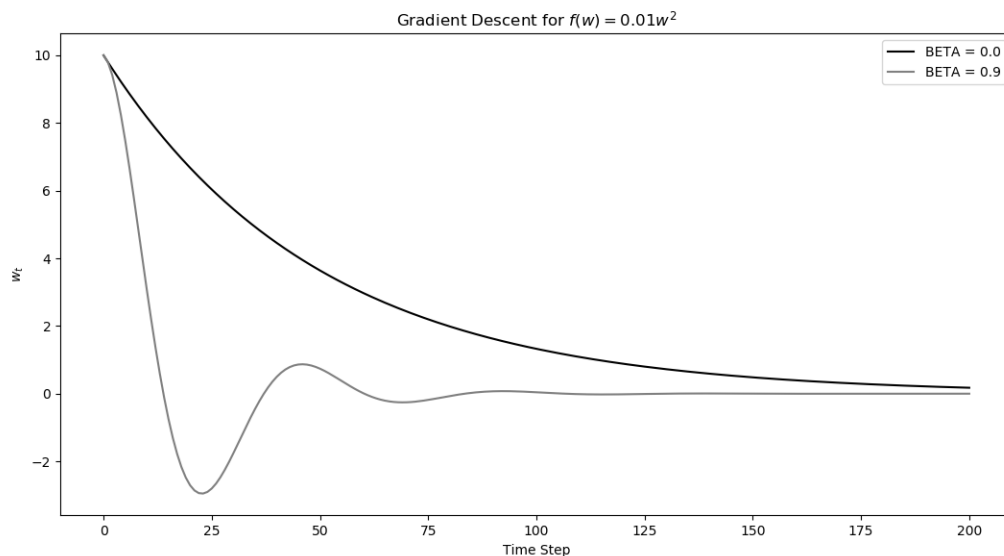
1 - 20 Newsgroup Predictions	1
2 - Training SVM with SGD	1
2.1 - SGD with Momentum	1
2.2 - Training SVM	2
2.3 - Apply 4-vs-9 Digits on MNIST	2
2.3.1 - Training Loss	2
2.3.2 - Test Loss	2
2.3.3 - Classification Accuracy on the Training Set	2
2.3.4 - Classification Accuracy on the Test Set	2
2.3.5 - Plot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

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 28×28 image

3 - Kernels

3.1 - Positive Semi definite and Quadratic Form

3.2 - Kernel Properties

3.2.1 - Prove Property

3.2.2 - Prove Property

3.2.3 - Prove Property

3.2.4 - Prove Property