**CS F320 - Foundations of Data Science**

**Assignment 1: Curse of Dimensionality and Gradient Descent**

**Group Information:-**

1. Pratik Kakade : 2017A7PS0086P
2. Nevin Thomas : 2017A7PS1175P

**Work division:-**

1. Report work on CoD : Pratik
2. Implementation of solutions of CoD problems : Pratik
3. Report work on Gradient Descent : Nevin
4. Implementation of Gradient Descent and its variants : Nevin
5. Comparison of Gradient Descent and its variants: Both

**Classification datasets and their applications :-**

1. Wine- Origins: Based on certain factors derived from chemical anaylsis of wine, we can determine the origin of the wine.
   1. Number of attributes: 13
   2. Number of instances: 178
2. Madelon dataset : An artificial dataset, which was part of the NIPS 2003 feature selection challenge. It is a two-class classification problem with continuous input variables.
   1. Number of attributes: 500
   2. Number of instances: 4400
3. Life Expectancy: Based on data from the World Health Organization, we can determine the life expectancy of people in different coutries based on certain f actors like whether it’s a developing country, the GDP, infant mortality rates etc.
   1. Number of attributes: 22
   2. Number of instances: 2938
4. Red-White Wine – Quality: A different wine dataset, the purpose of this is to look at certain factors and then determine the quality of it(a score from 1 to 10).
   1. Number of attributes: 11
   2. Number of instances: 6497
   3. Number of attributes: 500
   4. Number of instances: 4400

4. Red-White Wine – Quality: A different wine dataset, the purpose of this is to look at certain factors and then determine the quality of it(a score from 1 to 10).