

## CS F320 – Foundations of Data Science – Assignment #1

### Curse of Dimensionality (CoD) & Gradient Descent

Total Marks – 20 (weightage 10%)

Submission Deadline: 31<sup>st</sup> Mar. 2020

---

#### Aim:

The aim of this assignment is to understand problems related to:

1. high dimensional data and their solutions
2. gradient descent optimization technique and their solutions

#### Curse of Dimensionality (CoD)

Identify all the problems related to curse of dimensionality and illustrate them by carefully chose examples/situations.

Problems include (not an exhaustive list):

- exponential increase in computations
- concentration effect on  $L_p$ -norm
- data sparsity

What are the different ways in which CoD problems can be handled? Implement the solutions in C or R.

#### Gradient Descent

Implement GD and identify all problems related to it. What solutions have been proposed? Implement the solutions in C or R. Compare GD and its variants, qualitatively and quantitatively.

**Group Information** – Max. 2 students per group (with documented work division among group members – differential marking is applicable). It is strongly recommended that both members of the group should develop an understanding of the entire assignment and not just of the part he/she is responsible for.

**Deliverable 1 (due date: 19<sup>th</sup> March)** – Group Information, work division, identification of classification applications and datasets (submit a single handwritten page) [2 marks]

**Deliverable 2 (due date: 31<sup>st</sup> March)** – Final handwritten report (covering all theoretical aspects) and implementation details. [18 marks].

Theoretical rigor and understanding will fetch you credits. Good implementation of algorithms/techniques will also help you in scoring well.