1.	Title of the course	Foundations of Data Science
2.	Course number	CS534L
3.	Structure of credits	3-0-0-3
4.	Offered to	PG
5.	New course/modification to	Modification To CS5233/20
6.	To be offered by	Department of Computer Science and Engineering
7.	To take effect from	January 2022
8.	Prerequisite	Nil
9.	**Course Objective(s):** To understand the theory behind the computational aspects of data science and engineering. To develop tools and techniques from probability, statistics, and large-scale numerical computing for tackling problems in data sciences.	
10.	**Course Content:** Geometry of high dimensional space: volume of spheres and cubes in high dimensions; Topics from probability and statistics: probability mass in a unit sphere under spherically symmetric Gaussian, scales of multidimensional Gaussian in higher dimensions, concentration of measures and distances, computation of tail bounds; Deterministic dimension reduction: PCA review, nonlinear dimension reduction, SVD, K-rank approximation; Randomized algorithm for dimension reduction: random projections algorithms, properties, and applications; Randomized algorithm for linear algebra: random sampling, random projections and design of algorithms for low rank approximation of large matrices and large scale least square problems.	
11.	**Textbook(s):** 1. Michael W M, *Randomized Algorithms for Matrices and Data*, 1st Edition, Now (2011). 2. Vladimir S and David M, *Mathematical Foundations of Big Data Analytics*, 1st Edition, Springer (2021).	
12.	**Reference(s):** 1. Avrim B, John H and Ravindran K, *Foundations of Data Science*, 1st Edition, Cambridge (2020). 2. Ravindran K and Santosh V, *Spectral Algorithms*, 1st Edition, Now (2008).	