UNIVERSITY OF MADRAS
B.Sc. DEGREE COURSE IN COMPUTER SCIENCE
SYLLABUS WITH EFFECT FROM 2020-2021

BCE-CSC15

CORE-XV: INTRODUCTION TO DATA SCIENCE
(Common paper to B.Sc.Software Applications as Elective)

III YEAR / VI SEM

OBJECTIVES:
- To introduce the concepts, techniques and tools with respect to the various facets of data science practice, including data collection and integration, exploratory data analysis, predictive modeling, descriptive modeling and effective communication.

OUTCOMES:
- To describe what Data Science is, what Statistical Inference means, identify probability distributions, fit a model to data and use tools for basic analysis and communication

UNIT-I
Introduction to Data Science – Benefits and uses – Facets of data – Data science process – Big data ecosystem and data science

UNIT-II
The Data science process – Overview – research goals - retrieving data - transformation – Exploratory Data Analysis – Model building

UNIT-III
Algorithms - Machine learning algorithms – Modeling process – Types – Supervised – Unsupervised - Semi-supervised

UNIT-IV
Introduction to Hadoop – framework – Spark – replacing MapReduce– NoSQL – ACID – CAP – BASE – types

UNIT-V
Case Study – Prediction of Disease - Setting research goals - Data retrieval – preparation - exploration - Disease profiling - presentation and automation

TEXT BOOK
1. Davy Cielen, Arno D. B. Meysman, Mohamed Ali, "*Introducing Data Science*", manning publications 2016.

REFERENCE BOOKS
1. Roger Peng, "*The Art of Data Science*", lulu.com 2016.
2. MurtazaHaider, "*Getting Started with Data Science – Making Sense of Data with Analytics*", IBM press, E-book.
3. Davy Cielen, Arno D.B. Meysman, Mohamed Ali,"*Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools*", Dreamtech Press 2016.
4. Annalyn Ng, Kenneth Soo, "*Numsense! Data Science for the Layman: No Math Added*", 2017,1st Edition.
5. Cathy O'Neil, Rachel Schutt, "*Doing Data Science Straight Talk from the Frontline*", O'Reilly Media 2013.
6. Lillian Pierson, "*Data Science for Dummies*", 2017, 2nd Edition.

WEB REFERENCES
- NPTEL online course– Data Science for Engineers - https://nptel.ac.in/courses/106106179/