The Department of Advance Computing

POST GRADUATE COURSE

Course Title – MSc (Big Data Analytics)

With the rapid proliferation of internet and smartphones, the world has turned digital where every aspects of our life driven by digital technology.

This, in turn, is generating humungous amounts of data. Every corporation is trying to make meaningful use of this data to service their customers better, bring out newer products, or increase their efficiencies. This has opened up a large market for IT services organizations thereby increasing the demand for qualified professionals who will play a pivotal role in satisfying the needs of this rapidly expanding digital economy.

With a desire to be the leading IT services player in this area, TCS identified the need for industry-academia collaboration to create an ecosystem that will ensure a steady supply of such highly qualified professionals. Accordingly, TCS in collaboration with St. Joseph's College (Autonomous) Bangalore introduced a two-year full-time Post Graduate course in Big Data Analytics starting from the academic year 2016-17.

Duration – 2 years(Compulsory Internship at Industry in fourth semester)

Eligibility –Graduates who have studied mathematics or statistics as cognate subjects and secured 40% in the cognate subject and 50% marks in aggregate of all optional subjects at the degree level or B.C.A with Mathematics up to +2 are eligible.
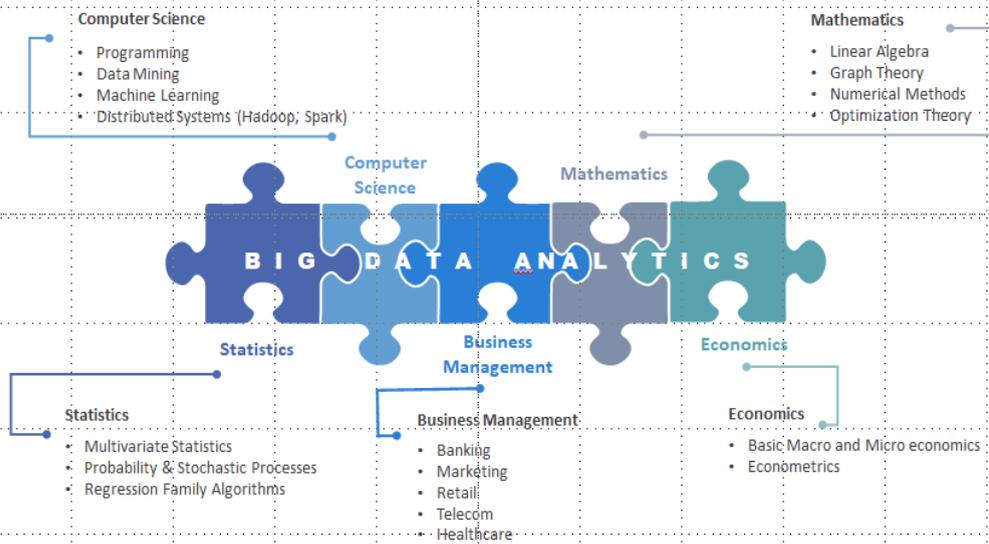
Outcome – Students will be industry ready with following knowledge:

1. Mathematical foundation, statistical and machine learning theories for analytics

2. Programming skills to implement the analytics and machine learning concepts

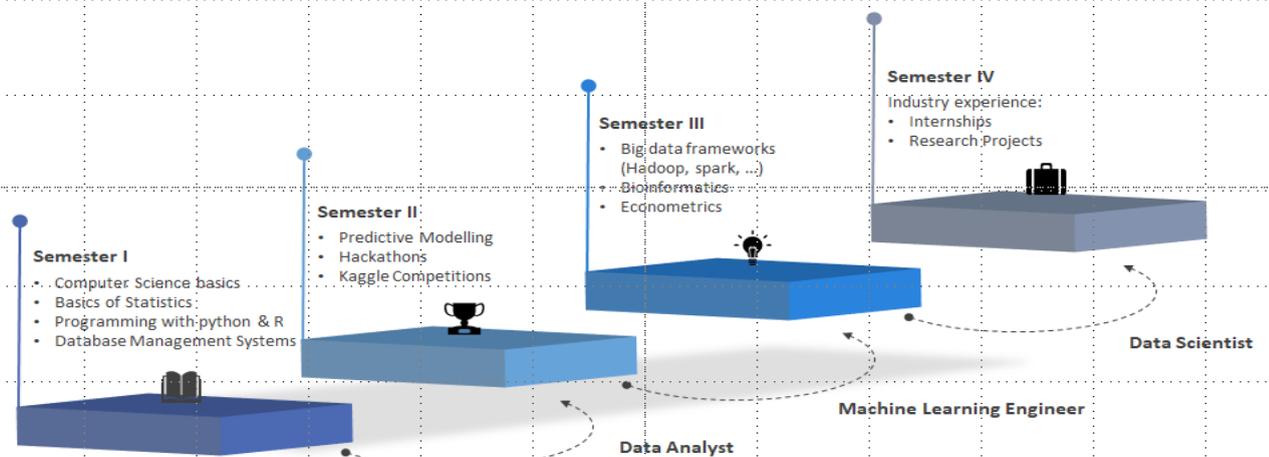3. Basic concepts of economics, accounts, ethical aspects of data analytics and so on

Syllabus

The focus areas of this course are mathematical foundation, statistical and machine learning theories for analytics, engineering aspects of analysing huge volumes of data and application of analytics in functional areas. While the three semesters (each of 4 months) are classroom-based, the final semester is earmarked for an internship at industry.

Course Structure

Computer Science

- Programming
- Data Mining
- Machine Learning
- Distributed Systems (Hadoop, Spark)

Mathematics

- Linear Algebra
- Graph Theory
- Numerical Methods
- Optimization Theory

Computer Science   Mathematics

BIG DATA ANALYTICS

Statistics   Business Management   Economics

Statistics

- Multivariate Statistics
- Probability & Stochastic Processes
- Regression Family Algorithms

Business Management

- Banking
- Marketing
- Retail
- Telecom
- Healthcare

Economics

- Basic Macro and Micro economics
- Econometrics

Theory to Practical

Semester IV

Industry experience:
- Internships
- Research Projects

Semester III

- Big data frameworks (Hadoop, spark, ...)
- Bioinformatics
- Econometrics

Semester II

- Predictive Modelling
- Hackathons
- Kaggle Competitions

Semester I

- Computer Science basics
- Basics of Statistics
- Programming with python & R
- Database Management Systems

Data Analyst

Machine Learning Engineer

Data Scientist

The first semester deals with the very basics of statistics and computer science. This acts as a bridge connecting theoretical knowledge to the practical aspects of the concepts. The second semester introduces one of the pillars of data science: machine learning. From advanced statistical methods and multivariate statistical aspects to data visualisation, it covers the essential background needed to delve deeper into the subject. The third semester polishes students by implementing modelling and more into the business and industrial sectors. This prepares students for their challenging internship in their fourth and final semester. At the end, students are instructed to prepare a presentation highlighting their contributions.

The college prepared itself well to deal with the new course. With constant support of the college management, this included a strong team of full-time faculty members led by Dr. Jayati Bhadra and is

complemented by a network of leading experts as guest faculties. The guest faculty includes Dr. Srinivas Bhogle, Honorary Scientist, CSIR-Fourth Paradigm Institute, who teaches probability and multivariate statistics. and Dr. Kaushik Basu from the National Law School, Bangalore who lectures on time series and panel data econometrics, In addition, Dr. Saurav Sen Gupta, faculty at Nanyang Technological University, Singapore, Prof. Vani from ISEC and Dr. Rituparna from ISI, Bangalore, also deliver occasional lectures in data science and analytics.

The College also invites eminent professors to talk about data science and related topics. This includes Dr. Bimal Roy, Chairman, National Statistical Commission, and former, Director, ISI and Prof. Steven Tanimoto, Washington University, who addressed the students on "Application of Machine Learning on Large Datasets" and "Transparency in Image Processing". Experts from TCS, IBM, Capgemini, Tech Mahindra also take time off to interact with and motivate the students.

The College prides itself in its state-of-the-art laboratory that was established in collaboration with TCS. The laboratory ensures computing infrastructure to all students and is capable of handling big data loads. It is also equipped with high end communication infrastructure including a high definition video conferencing facility that is used regularly by the students and the faculty to interact with TCS experts thereby increasing the collaboration as well as exposure for the students to the industry requirements. Overall, the laboratory infrastructure provides a conducive atmosphere for students to get themselves honed with the skills for data analysis. The lab even functions for extended hours to fulfil the needs of the students.

The students are given plenty of books to refer to in the college library, with many books specially ordered to stimulate intellectual growth. Students attend conferences and this year they are preparing to publish papers in reputed journals. Three projects on Toycathon are selected for the grand finale. One student's project received funding from the Karnataka state Govt. Value added courses are included to shape students in all aspects of their lives.

Students have been consistently performing well, with the four-month internship in industry being especially successful. They are inquisitive, hard working and eager to learn. They have formed an association called 'Sigma Squared' that conducts various programmes throughout the academic year to groom the students as data analysts. An Inter-collegiate fest named Datagram is held every year. Events like What if, Lecture contest, Hackathon, Panel discussions, Life skill training and workshops were much appreciated.

In the past, organizations like TCS, EY, Nielson, Amazon, Aditya Birla Group, Federal Bank etc. have recruited graduates from this course.