SEMIMAR REPORT

ON

DATA SCIENCE

By

Mr. Ranjeet Walunj

CTO, B2X Care Solutions GMBH

Organised by,

Computer engineering Alumni committee,

## **RGIT, Mumbai.**

7th march 2018

ABOUT SEMINAR

Data Science is an emerging field that has grown out of the tremendous growth in statistics and computer science in the past few decades. Data can help one seek to understand patterns in the world in light of inherent uncertainty. In order to study and address many of the challenges in our society, one must have the tools to process data, perform computations, summarize, investigate, and communicate important findings from the information. Data science can play a critical role in many efforts to enhance the conditions of the human person and the world.

In this seminar we discussed how Data science is the core that drives new research in many areas from environmental to social sciences. There are many associated scientific challenges, ranging from data capture, creation, polishing, storage, search, sharing, modelling, analysis and visualization. The Seminar explores both theoretical and practical works on Data Science and Advanced Analytics with visionary speakers to outline different insights and views about today and future trends. The Seminar is also featured by a panel session to discuss challenges on selected topics and helped to solve real time problems By incorporating numerous disciplinary perspectives and relying heavily on domain knowledge and expertise, data science has emerged as an important new area that integrates statistics with computational knowledge, data collection, cleaning and processing, analysis methods, and visualization to produce actionable insights from big data.

Introduction to the speaker:

Mr. Ranjeet walunj has done his post graduation in MBA from Welingkar Institute of Management, information systems, in 2004 – 2007 and completed his bachelor's degree in computer engineering from Rajiv Gandhi Institute of Technology under affiliation of University of Mumbai in 1996-2000.
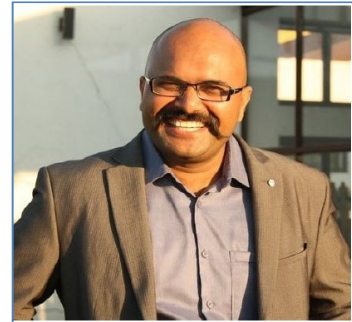
He is currently working as Chief Technology Officer at B2X Care Solutions GMBH, Investment Scout & Startup Advisor. He was Co-founder & CTO at The Service Solutions in duration of May 2011-September 2014. Also he'd worked as Senior Manager at Interactive Avenues and Tech - Mediaturf worldwide. He'd also worked as Senior Systems Analyst at EuroRSCG India. Also he was Programmer + System Analyst (Consultant) at Interpole Technologies Pvt. Ltd. in the period of June 1999 - April 2000 (11 months). InterPole, at the time, was a premier web design, web hosting and internet presence solutions provider. Where he'd worked on FilmTVIndia: Machine Learning based Movie Recommendation Engine and worked as Project head for the project team that delivered all Aditya Birla Group websites and online entities.

He has an Extensive international and domestic experience in technology and executive leadership roles and Recognised expert on the strategic application of technology and business strategy. He has Proven ability to envision and orchestrate large scale enterprise product development and solution projects that require complex business analysis, timely results and market performance. He has Equally adept at approaching, integrating, and solving problems from a business or technology perspective.

To take a walk down the corridors of nostalgia, the Annual Alumni Meet 2016, was inaugurated with by giving flower bouquet to Mr, Ranjeet walunj by computer engineering department's HOD, Dr. Satish Ket sir on the morning of 7th march 2018.

The event was compiled by computer engineering department's alumni committee under guidance of faculty of department. The occasion was graced by Alumni of computer department who is from pass out batch of 2000. Speaking on the occasion, Dr. Satish Ket and faculty of computer engineering department welcomed alumni with his compassionate words and mentioned the objective of holding seminar on data science for the advancement of our knowledge and giving brief idea about future scope and demands. While addressing alumni, committee member Shivani and Rudra underlined and praised the achievements of Alumni Mr. Ranjeet walunj. He felt that it was a matter of great pride for RGIT college that, its Alumni is holding positions of responsibility in corporate world and involved in various start-ups.

As a discipline to use within Web science research, data science offers significant opportunities for uncovering trends in large Web-based datasets. A Web science observatory exemplifies this relationship by offering an online platform of tools for carrying out Web science research, allowing users to carry out data science techniques to produce insights into Web science issues such as community development, online behavior, and information propagation. The authors outline the similarities and differences of these two growing subject areas to demonstrate the important relationship developing between them.

Also we discussed various Platforms needed to learn data science. Along with it, Mr. Ranjeet Walunj had introduced us with hierarchy of Artificial Intelligence, various categories working under it and also made familiar with importance of statistics and linear algebra in computer science.

Throughout the seminar, we discussed about latest trending fields in market and corporate world and speaker guided us what should we learn and get valued knowledge before entering into corporate world and also encouraged about start-ups ideas with real time solutions,

RAJIV GANDHI INSTITUTE OF TECHNOLOGY, MUMBAI

ARTIFICIAL INTELLIGENCE:

Artificial intelligence is a branch of computer science that aims to create intelligent machines. It has become an essential part of the technology industry. Research associated with artificial intelligence is highly technical and specialized. The core problems of artificial intelligence include programming computers for certain traits such as:

- Knowledge
- Reasoning
- Problem solving
- Perception
- Learning
- Planning
- Data mining
- Ability to manipulate and move objects
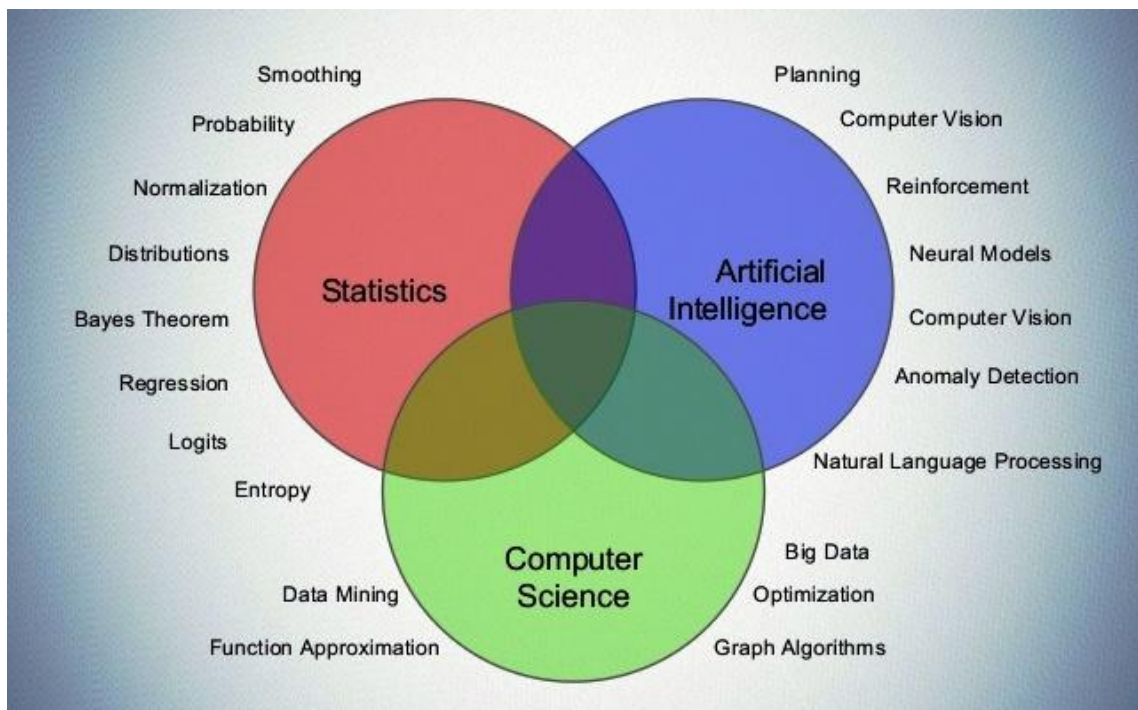
Knowledge engineering is a core part of AI research. Machines can often act and react like humans only if they have abundant information relating to the world. Artificial intelligence must have access to objects, categories, properties and relations between all of them to implement knowledge engineering.

Initiating common sense, reasoning and problem-solving power in machines is a difficult and tedious approach.

Machine learning is another core part of AI. Learning without any kind of supervision requires an ability to identify patterns in streams of inputs, whereas learning with adequate supervision involves classification and numerical regressions. Classification determines the category an object belongs to and regression deals with obtaining a set of numerical input or output examples, thereby discovering functions enabling the generation of suitable outputs from respective inputs. Mathematical analysis of machine learning algorithms and their performance is a well-defined branch of theoretical computer science often referred to as computational learning theory.

Machine perception deals with the capability to use sensory inputs to deduce the different aspects of the world, while computer vision is the power to analyze visual inputs with a few sub-problems such as facial, object and gesture recognition. Robotics is also a major field related to AI. Robots require intelligence to handle tasks such as object manipulation and navigation, along with sub-problems of localization, motion planning and mapping.

DATA SCIENCE:

Data Science combines courses in Computer Science, Mathematics, and Statistics to equip students with the skills needed to analyze Big Data using advanced computing and statistical methods. Students begin with introductory classes in these areas to build their knowledge of the mathematical foundations of data science and their computational and statistical thinking skills.

Data science, also known as data-driven science, is an interdisciplinary field of scientific methods, processes, algorithms and systems to extract knowledge or insights from data in various forms, either structured or unstructured, similar to data mining.

Data Science is an interdisciplinary field about processes and systems to extract knowledge or insights from data in various forms. This means that data science helps AIs figure out solutions to problems by linking similar data for

future use. Fundamentally, data science allows for AIs to find appropriate and meaningful information from those huge pools faster and more efficiently.

An example of this is Facebook's facial recognition system which, over time, gathers a lot of data about existing users and applies the same techniques for facial recognition with new users. Another example is Google's self-driving cars which gathers data from its surroundings in real time and processes those information to make intelligent decisions on the road.

Machine Learning is likely the connection between data science and artificial intelligence since machine learning is the process of learning from data over time. However, it's not the only thing connecting those two together. But, machine learning is the branch of AI that works best with data science

THINGS NEEDED TO LEARN DATA SCIENCE:

Ranjeet sir gave us a keynotes the things required to get hold over data science and machine learning and gave us info how should we start. He shared following guidelines for the data science:

- Learn basic mathematics and statistics required for data science
- Develop a basic understanding of machine learning algorithms and solving real life problems from them
- Skills required to land you first data science internship / job
- Learn python/R language

In brief context, Ranjeet sir has discussed importance of probability, linear algebra, numerical methods, statistics in real time problem solutions.

Along with it, he encouraged us to join various online python courses running over internet free of cost which would help us to build understanding in data science and AI.

Also, following topics are must to learn before entering into data science or artificial intelligence:

Basic Machine Learning Algorithms.

- Linear Regression
- Logistic Regression
- Decision Trees
- KNN (K- Nearest Neighbours)
- K-Means
- Naïve Bayes
- Dimensionality Reduction

Advanced algorithms

- Random Forests
- Dimensionality Reduction Techniques

MISCELLANEOUS:

It is very important for a Data Scientist to have a GitHub profile to host all the codes of the project he/she has undertaken. Potential employers not only see what you have done, how you have coded and how frequently / how long you have been practicing data science.

Also, codes on GitHub open up avenues for open source projects which can highly boost your learning. If you don't know how to use Git, you can learn from Git and GitHub on Udacity. This is one of the best and easy to learn course to manage the repositories through terminal.

Ranjeet sir has also encouraged us to involve in competitive programming to enhance logical reasoning as well as help us to solve and gives vision to solve real time problems.

PICTURES:

❈   **THANK YOU**   ❈