Course 1. What is data Science ?

Data Science is the field of exploring, manipulating and analyzing data and using data to answer questions or make recommendations.

Definition: Data Science can analyze structured and Unstructured data from many sources and depending on the nature of the problem, they can choose to analyze the data in different ways.

What Do Data Scientist Do?

- The typical work day for a Data Scientist varies depending on what type of project they are working on.
- Many algorithms are used to bring out insights from data.
- Accessing algorithms, tools, and data through the Cloud enables Data Scientists to stay up-to-date and collaborate easily.

Data Science Topics:

1. **Big Data And Data Mining :**

   Big Data: Big data refers to the dynamic, large and disparate volume of the data being created by people, tools and machine. It requires new innovative and scalable technology to collect, host and analytically process the vast amount of data gathered in order to derive real time business insights that relates to consumers, risks, profits, performance, productivity and enhanced shareholder value.

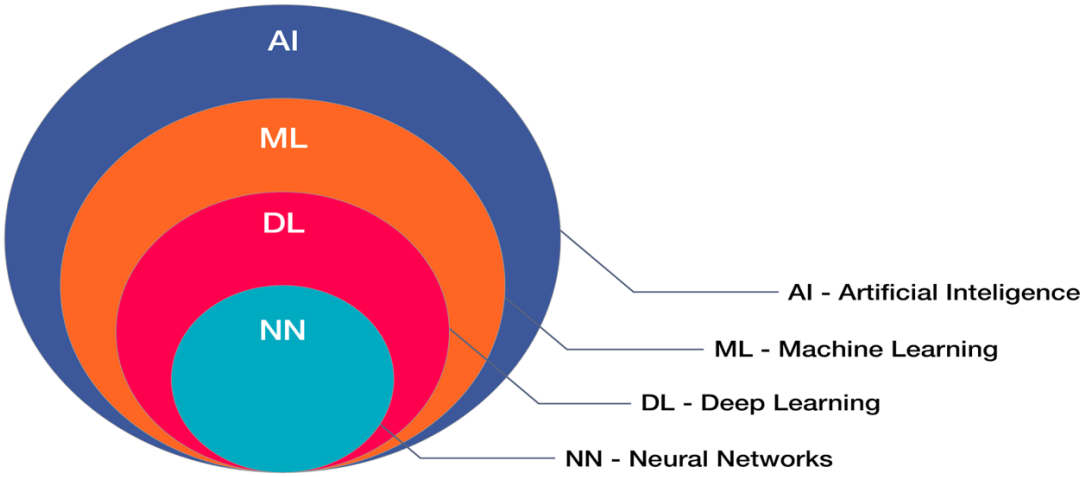
   The V's of the Big Data
   1. **Velocity**: Speed at which data is accumulated
   2. **Volume** : Scale of the data
   3. **Variety**: Diversity of the data
   4. **Veracity**: Quality and origin of data
   5. **Value**: Our ability and need to turn data into value.

   Tools such as *Apache Spark, Hadoop and its Ecosystem* provides ways to extracts, load, analyse and process the data across distribution compute resources, providing new insights and knowledge.

   Data Mining: The process of automatically searching and analysing data, discovering previously unrevealed patterns.

   - Pre-processing
   - Transformation.

2. Deep Learning And Machine Learning:



Once the pre-processing and transformation is done then insights and patterns and mined and extracted using the tools and techniques ranging from simple data visualization tools to machine learning and statistical models.

Machine Learning: ML is what enables machine to solve problems on their own and make accurate predictions using the provided data.

Application of Machine Learning:

1. Regression
2. Classification
3. Clustering
4. Predictive Analysis

Deep Learning: DL is the specialized subset of the ML that uses layered neural networks to simulate human decision making.

Deep Learning is " Neural Networks(NN) On Steroids"

DL just had multiple layers of NN and they use lots of computing power to solve them.

Applications of Deep Learning:

1. Recognition of speech
2. Recognition of People
3. Classification of Images.