Lesson 1 -07th August. 2021

OVERVIEW;

Understanding a problem / expectation of a client is important to any project.

The new technologies available in the market. Datascience entails different job profiles in the market.

Machine learning , deeplearning, natural language processing, robotic process automation (RPA) and artificial intelligence nothing but a technology that can be applied in any industry or domain, which is why career transition into datascience has been easy.

**How companies are utilizing these new technologies , in real time, to take them to the next level.**

By applying,

**Analytics** : the purpose of analytics is pulling out meaningful insights from data.

**Automation:** the purpose of automation is for revenue and time savings and to enable quick action and decision making

ML, AI, DL & NLP technologies are a subset of datascience. Lot of languages like python, r, sas etc. Python is simple yet has it’s own complexities. PYTHON & R are prime languages in datascience and AI

In 1950 Alok a businessman starts his business, he would sell different items . In 1965 , his sales dipped and he needed to effect some change . He approaches his mathematician friend Siddhu to help him with his financial issues.

Data is the raw material for any solution to be arrived at. Data is the new oil to provide solutions. Raising right questions on data is the essence of datascience.

For eg : Alok would have Volume of data with following parameters

**Date, Name , Product, Qty Rate Discount , Net price. Season**

This is the base theorem of probability and picked up to make a statement about future or early prediction or in other words, probability.

If a customer is buying a blanket when the season is summer/w inter/ at this discount/ from this location/ at this price/ . By raising multiple questions on data, and understanding the domain/ impact of the data. A probable solution framework can be arrived at by predictive analysis and applying mathematical knowledge, leads to analytics. At this stage there is no application of automation.

Analysis here is only pulling out of meaningful insights from data and applying mathematical equations to arrive at solutions

Naïve’s base theorem of probability could be one such statistical equation used to arrive at probable results.

P(a/b)= p(b/a) \* p(b)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

P(a)

But this cannot satisfy all requirements. Issues of volume,scalability,revenue savings,generation of business,speed of delivery necessitates automation

**Types of analytics** : Fraud , Supply chain, Risk, HR, Sales,

**If goal is automation one needs to get into python and for pure analysis, R would be more aligned.**

Lot of analysis can be done with python too.

**Three pillars**

Domain knowledge,Statistical Data Knowledge, Programming.

Datascientist combines all of the above and render great value to the client. Datascientist solves the unsolved problems,

Learning grid comprises of three elements : Learning rate, batches and iterations/ epoc

Iterations/ Epoch

Learning rate

Batches

|  |  |  |  |
| --- | --- | --- | --- |
| Designation | **Purpose** | **Skills** |  |
| Data Analysis | Insights from the data to take business decision and actions | Excel, VBA, Macros, SQL, Basics Python, BI Tools - Tableau/PowerBI, Data Viz skills, Story Telling Skills |  |
| Data Analytics | Insights from the data to take business decision and actions | Excel, VBA, Macros, SQL,  BI Tools - Tableau/PowerBI, Data Viz skills, Story Telling Skills |  |
| Business Analyst | Insights from the data to take business decision and actions | Excel, VBA, Macros, SQL,  BI Tools - Tableau/PowerBI, Data Viz skills, Story Telling Skills | **DATA ANALYSIS** |
| SQL Engineer | Insights from the data to take business decision and actions | Excel, VBA, Macros, **SQL,**  BI Tools - Tableau/PowerBI, Data Viz skills, Story Telling Skills |  |
| ETL Engineer (Extract, Load, Transform) | **Data Collection +  I**nsights from the data to take business decision and actions | Excel, VBA, Macros, **SQL, NoSQL, Postgre, MangoDB,** BI Tools - Tableau/PowerBI, Data Viz skills, Story Telling Skills |  |
|  |  |  |  |
| Data Scientist | Insights + Predictions + Follow Up Action | **Python, ML, DL, NLP,**  Excel, SQL, Story Telling Skills, Tableau, Hadoop, Django, Flask |  |
| ML Enginner | Insights + Predictions + Follow Up Action | Python, ML, DL, NLP | **DATA SCIENTIST** |
| AI Enginner | Insights + Predictions + Follow Up Action | Python, ML, DL, NLP |  |
| DL Engineer | Insights + Predictions + Follow Up Action | Python, ML, DL, NLP |  |
| NLP Engineer | Insights + Predictions + Follow Up Action | Python, ML, DL, NLP |  |
|  |  |  |  |
| Data Engineering | to build **pipeline** to collect and store the data fro the cloud | Spark, Hive, Pig, Hadoop, MapReduce |  |
| Big data Engineer | to build **pipeline** to collect and store the data fro the cloud | Spark, Hive, Pig, Hadoop, MapReduce | **DATA ENGINEER** |
| Hadoop Enginner | to build **pipeline** to collect and store the data fro the cloud | Spark, Hive, Pig, Hadoop, MapReduce |  |
| Data warehousing Enginner | to build **pipeline** to collect and store the data fro the cloud | Spark, Hive, Pig, Hadoop, MapReduce |  |
| Data Architect | to build **pipeline** to collect and store the data fro the cloud | Spark, Hive, Pig, Hadoop, MapReduce |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Full Stack Developer (Data Science) | Inisghts + Prediction + Follow Action + to build pipeline to collect and store the data fro the cloud | Python, ML, DL, NLP,  Excel, SQL, Story Telling Skills, Tableau, Hadoop, Django, Flask, Spark, Hive, Pig, Hadoop, MapReduce | **DATA SCIENTIST & DATA ENGINEER** |
|  |  |  |  |
| Backend | To build Web Application | Django, and other WebApplication Frameworks. | **PYTHON DEVELOPER** |
| Python developer | To build Web Application | Django, and other WebApplication Frameworks. |  |
|  |  |  |  |
|  |  |  |  |
| DevOps | Inisghts + Prediction + Follow Action + to build pipeline to collect and store the data fro the cloud | Python, ML, DL, NLP,  Excel, SQL, Story Telling Skills, Tableau, Hadoop, Django, Flask, Spark, Hive, Pig, Hadoop, MapReduce, **Cloud Platform Knowledge - AWS, Azure, GCP,** | **DEV OPS ENGINEER** |
| MLOPS Enginner | Inisghts + Prediction + Follow Action | **Python, ML, DL, NLP,**  Excel, SQL, Story Telling Skills, Tableau, Hadoop, Django, Flask, Spark, Hive, Pig, Hadoop, MapReduce, **Cloud Platform Knowledge - AWS, Azure, GCP,** |  |

**PYTHON DOWNLOAD**

Trending, easy, vastly used

PYTHON.ORG

3.9.5

Run executable installer,

Cmd python

Tutorial point.com / how to install python

**ANACONDA NAVIGATOR : INDIVIDUAL EDITION**, provide integrated development enterprise ( IDE)