

Master of Data Science and Artificial Intelligence

Group – 8

- Bo Peng
- Haotian Zhao
- Harshil Patel
- Shuyue Zhang
- Tuhin Ranjan
- Tushar Aggarwal

December 1, 2020
MIE-1624

Introduction

Objective:

- Re-design MIE 1624: Introduction to Data Science and Analytics Curriculum
- Design a new Master's of Data Science and Artificial Intelligence Program

Data:

- Web Scrapping for skills: LinkedIn and Indeed
- Web Scrapping for competitive Courses: Coursera
- University Course Review: North American Universities
- Kaggle Data Science Questionnaire : 2018 and 2019



[This Photo](#) by Unknown Author is licensed under [CC BY](#)



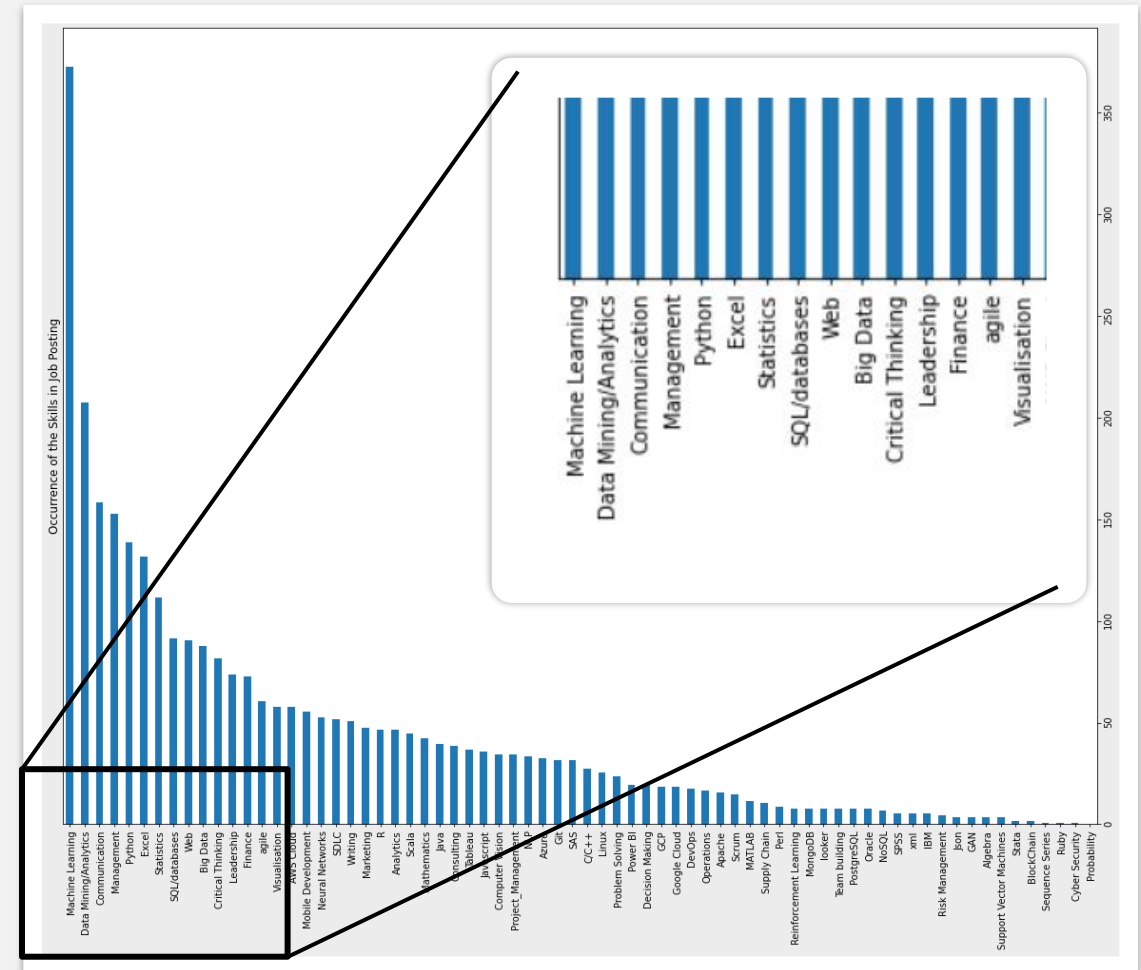
Exploratory Data Analysis

Indeed and LinkedIn

- Combined **Indeed** and **LinkedIn** web scrapped data was combined.
- A dictionary with skills as keywords was created with values as different references/names for that keyword.
- Count for keywords in the web scrapped data was then plotted to check the popular in demand skills in the field of Data Science and Machine Learning

Top in demand skills:

- Machine Learning
- Communication
- Management
- Python Programming
- SQL

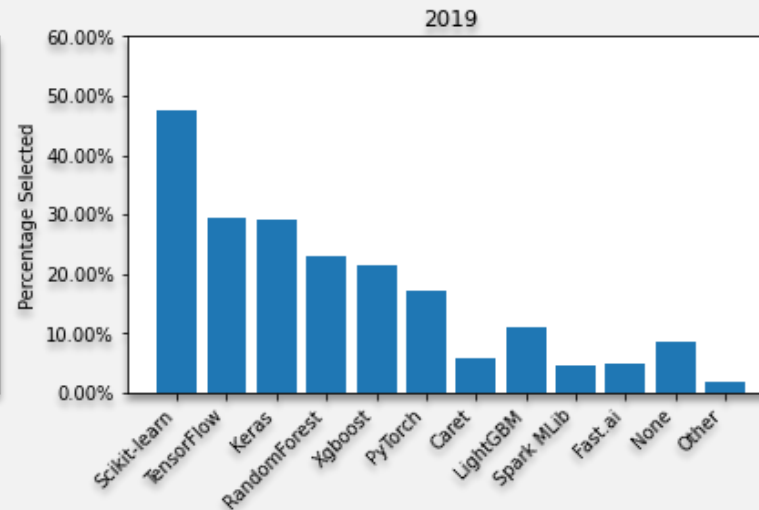
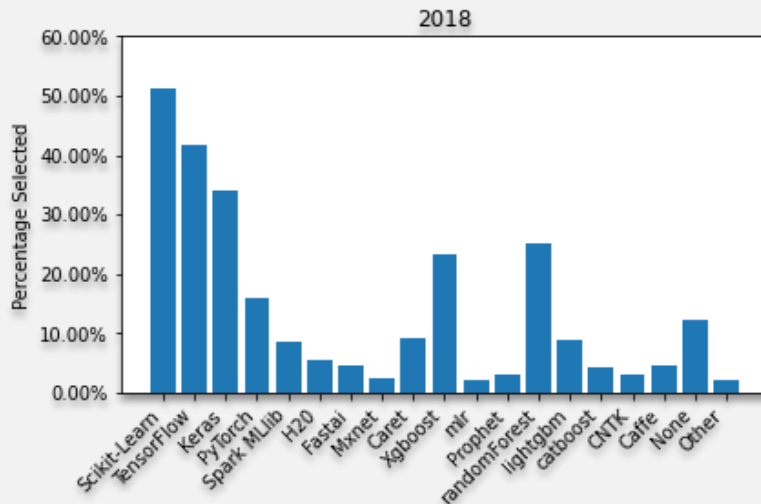


Exploratory Data Analysis

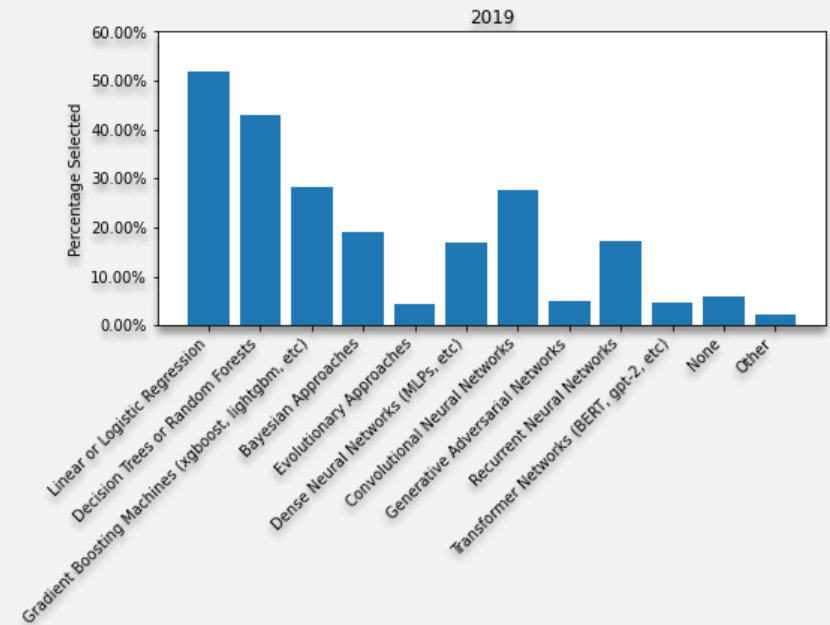
Kaggle 2018 and 2019 Data

- Kaggle survey was visualized for all the questions.
- Questions helped to narrow down our course design.
- Coursera was one of the popular website to learn data science and artificial intelligence skills

Which of the following machine learning frameworks do you use on a regular basis?



Which of the following ML algorithms do you use on a regular basis?



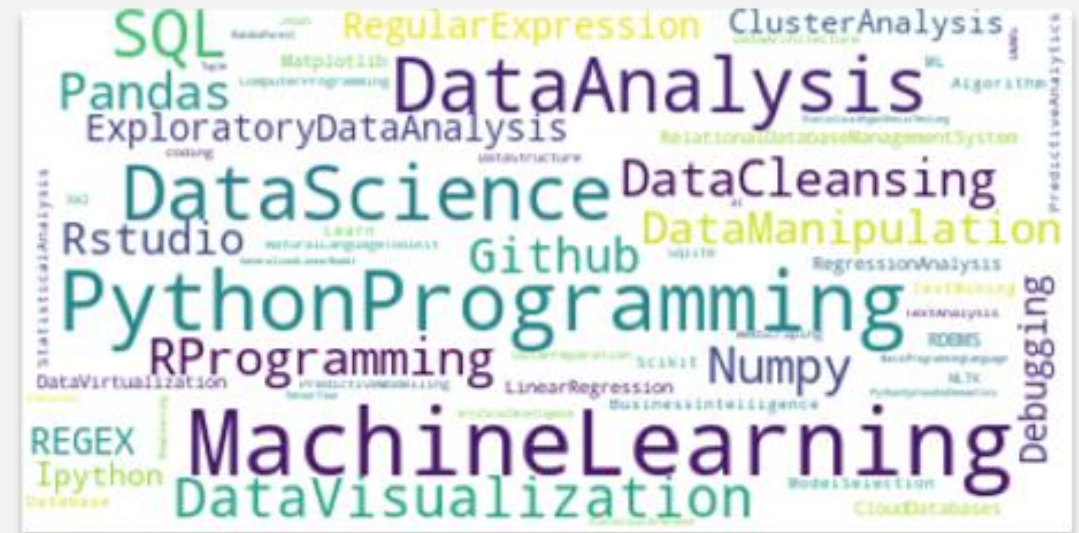
Exploratory Data Analysis

Coursera

- Coursera was selected as it was one of the popular websites to learn Data Science skills as per Kaggle surveys.
- Coursera web data was explored to see top courses and skills in data science that are taught on Coursera.

University

- Courses in Data Science and Artificial Intelligence were explored for some of the competitive Universities in North America.
- Program design was explored for internship and capstone projects.



Top in demand skills as per Coursera:

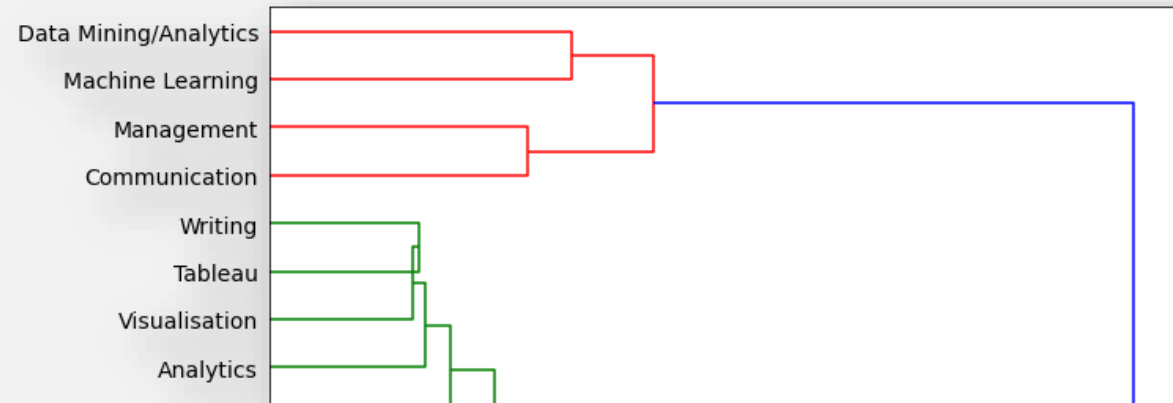
- Machine Learning
- Python Programming
- R Programming
- Data Analysis
- SQL



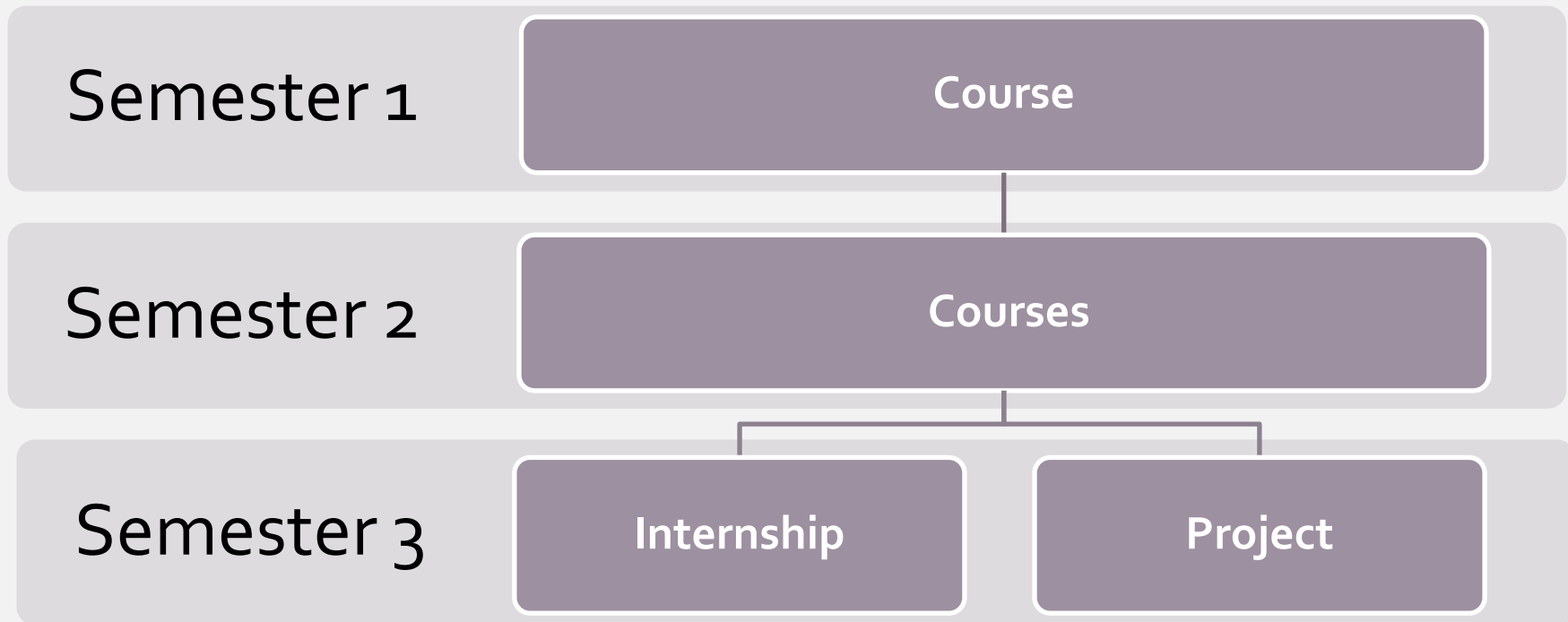
Exploratory Data Analysis

Hierarchical Clustering (HC)

- HC was performed at different stages of analysis
- Skills were grouped together using HC based on how many of them occur together in job descriptions.
- HC was also done on the Coursea data as well as on University data



Curriculum Design



The program is a full time Professional Master's Program with 3 4-month semesters. To receive the degree, student must take in the total of 7 courses and complete either a 4-month internship or an individual project in the last semester.

Technical/Business Pathways



Technical Pathway

- Introduction to Data Science and Analytics*
- Introduction to Machine Learning*
- Big Data*
- Introduction to Deep Learning
- Statistical Analysis
- Advance Data Analysis
- Full Stack Development

Business Pathway

- Introduction to Data Science and Analytics*
- Professional Communication and Leadership*
- Business Intelligence*
- Project Management
- Data Science Programming Basics
- Introduction to Computational Finance
- Operations Research

* Indicated that the course is mandatory for this option

MIE1624 Course Update

Revised Learning Outcomes:

- ☐ Emphasis on python programming
- ☐ Exposure to R and SQL usage in data science
- ☐ Focus on main advanced machine learning algorithms
- ☐ Less emphasis on deep learning
- ☐ Emphasis on technical presentations and communications
- ☐ Other Topics are briefly reviewed

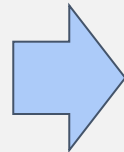
Proposed Grading Scheme Changes:

- ☐ Assignment #1: 14%
- ☐ Assignment #2: 14%
- ☐ Assignment #3: 14%



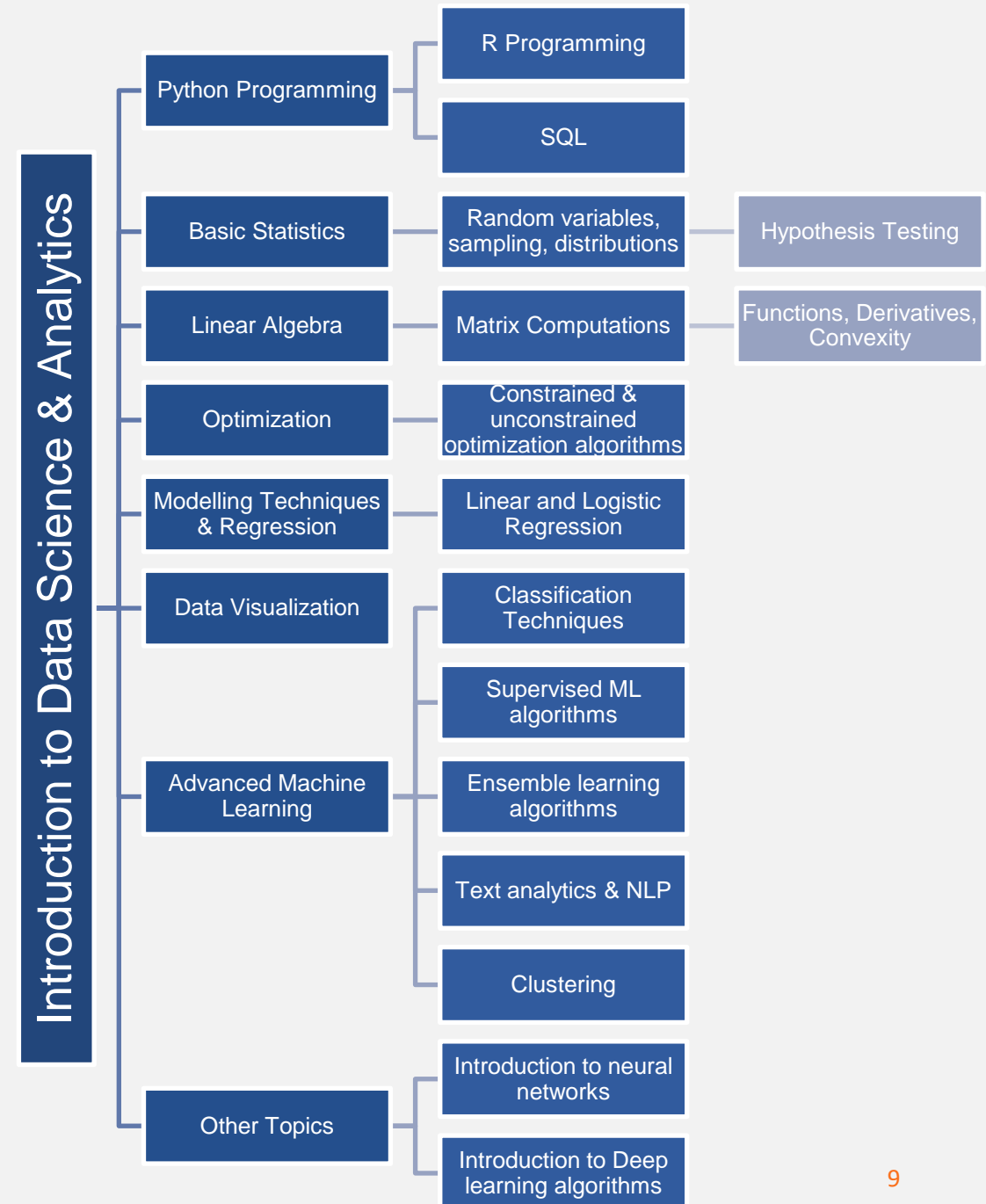
20% of each assignment allocated to 2-3 minute video presentation and report

- ☐ In-class group presentation: 14%



- ☐ Course project: 24%
- ☐ Final Exam Project: 20%

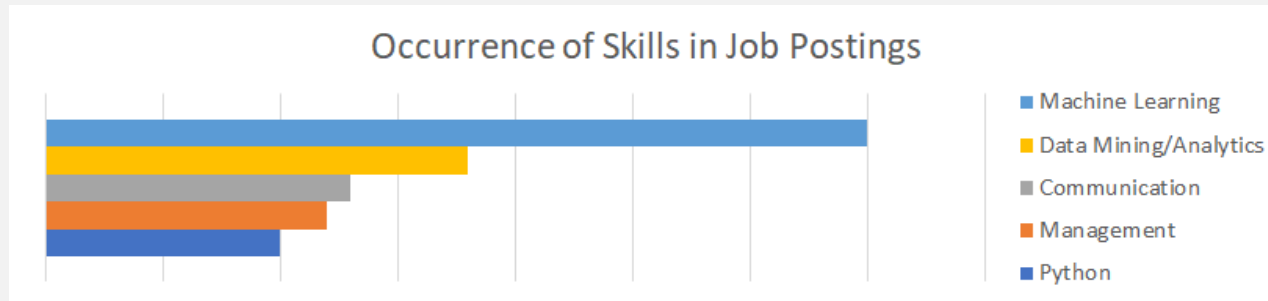
20% of each project allocated to 2-3 minute video presentation and technical report



Introduction to Machine Learning

Why is this a core course?

- ❑ Machine learning is a top 3 skill for data science related jobs

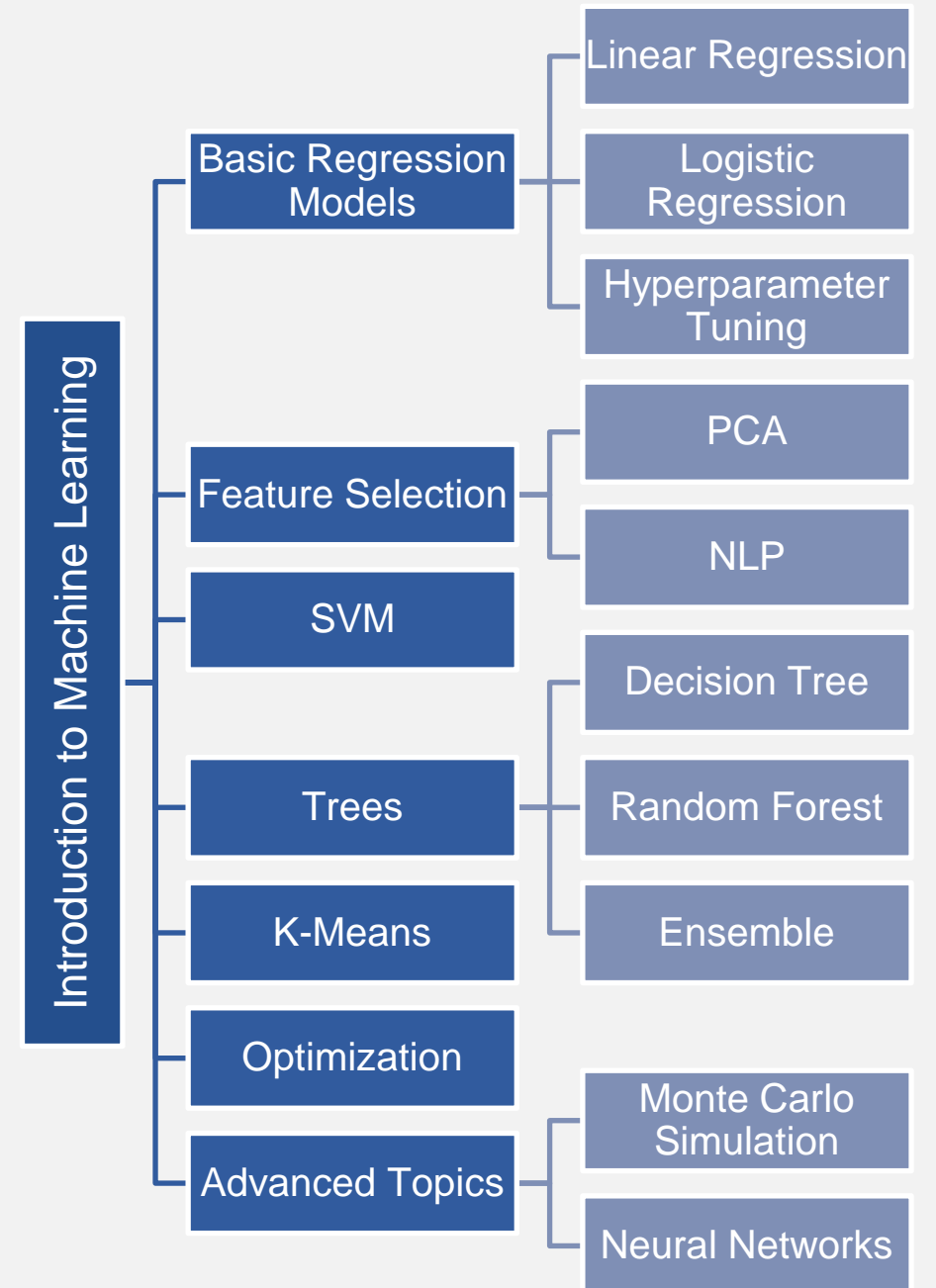


Key Aspects

- ❑ Python Programming language
- ❑ Emphasis on derivation of ML algorithms using mathematics

Pre-requisites:

- ❑ Higher undergraduate level statistics
- ❑ Probability theory
- ❑ Linear algebra



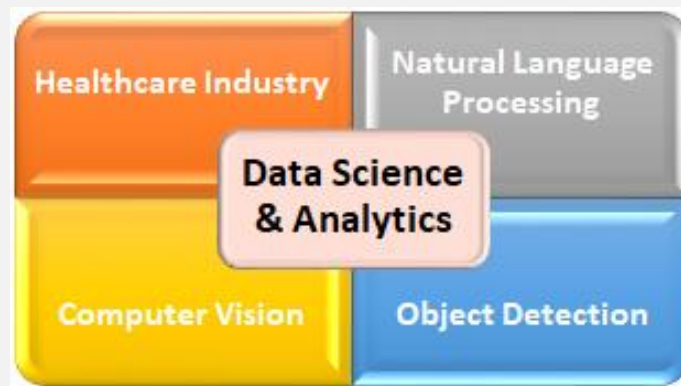
Introduction to Deep Learning

Why is this a core course?

- ❑ Deep learning algorithms are among the most used ML algorithms in the industry according to 2018 & 2019 Kaggle surveyors

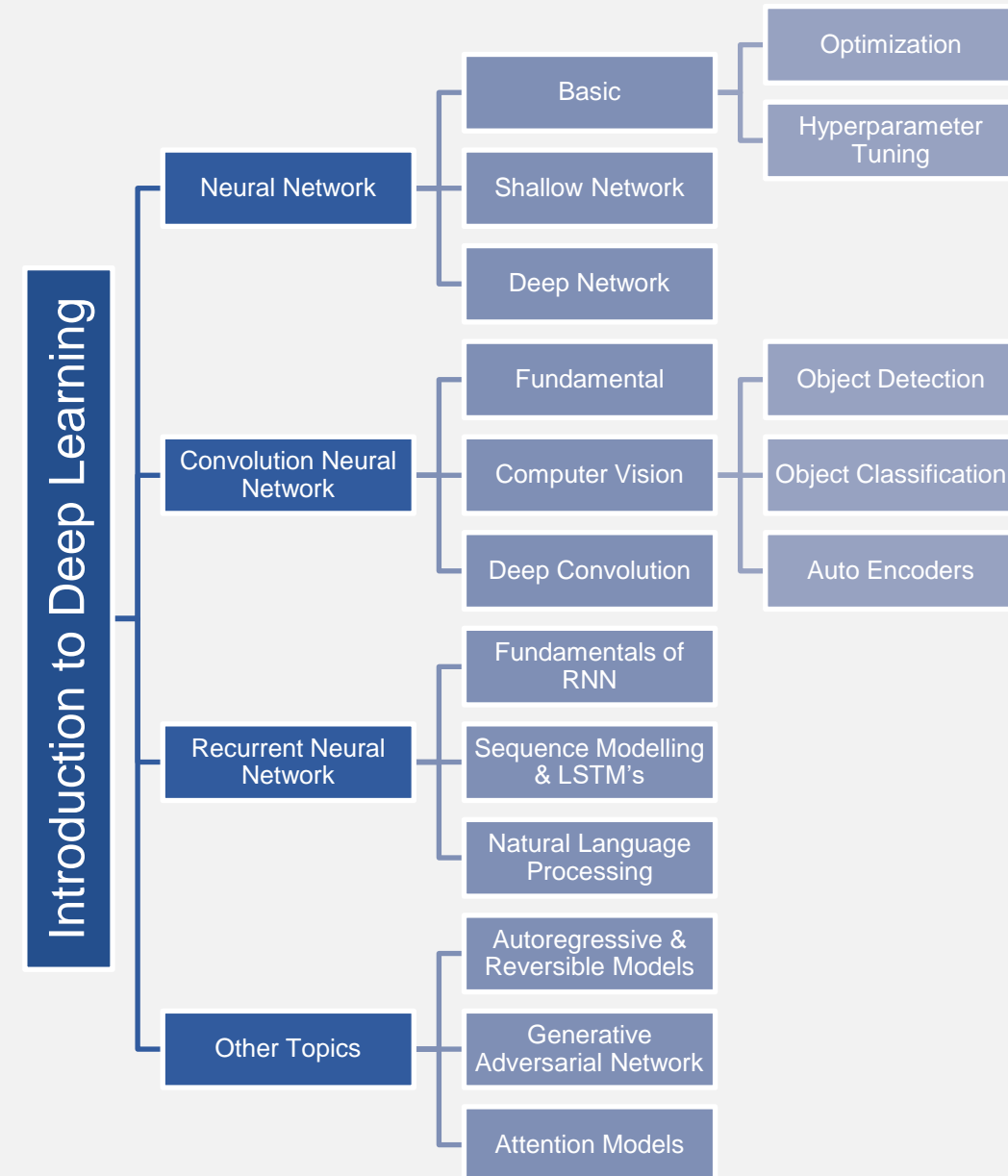
Key Aspects

- ❑ Comprehensive knowledge of advanced ML algorithms
- ❑ Exposure to industry-valued algorithms such as YoloV4, SIFT keypoint detectors, etc.



Pre-requisites:

- ❑ Introduction to Machine Learning
- ❑ Exposure to TensorFlow beneficial but not required
- ❑ Higher level calculus, linear algebra, and statistics courses



Big Data

Why is this a core course?

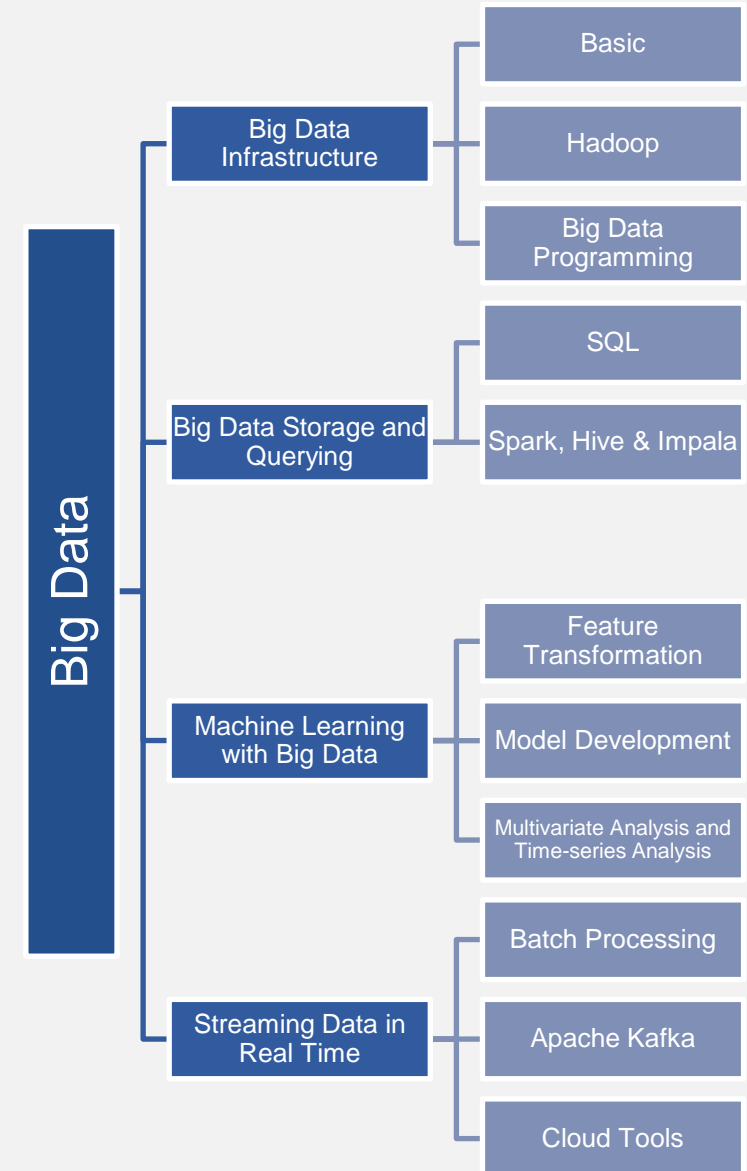
- ❑ Deep learning algorithms are among the most used ML algorithms in the industry according to 2018 & 2019 Kaggle surveyors

Key Aspects

- ❑ Comprehensive knowledge of advanced ML algorithms
- ❑ Exposure to industry-valued algorithms such as YoloV4, SIFT keypoint detectors, etc.

Pre-requisites:

- ❑ Experience with SQL and functional language should be helpful but not required



Estimation & Assessing Theories

三六



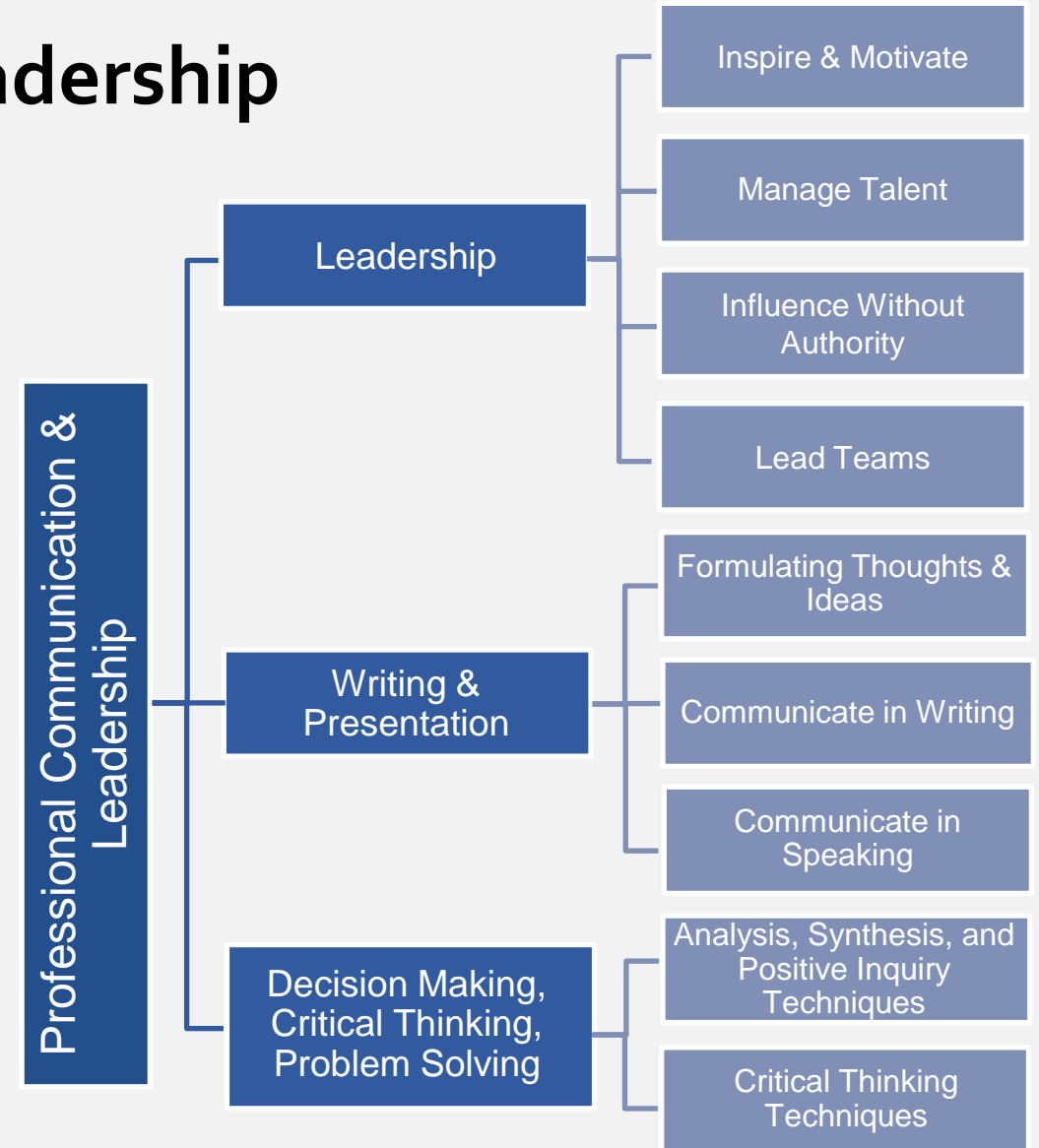
Professional Communication & Leadership

Why is this a core course?

- ❑ Critical thinking, communication and leadership all become essential skills of the data science jobs.
- ❑ Since this is a technical- and business-oriented degree, soft skills should also be included in the core courses.

Key Aspects

- ❑ Leadership
- ❑ Writing & Presentation
- ❑ Decision Making, Critical Thinking & Problem Solving



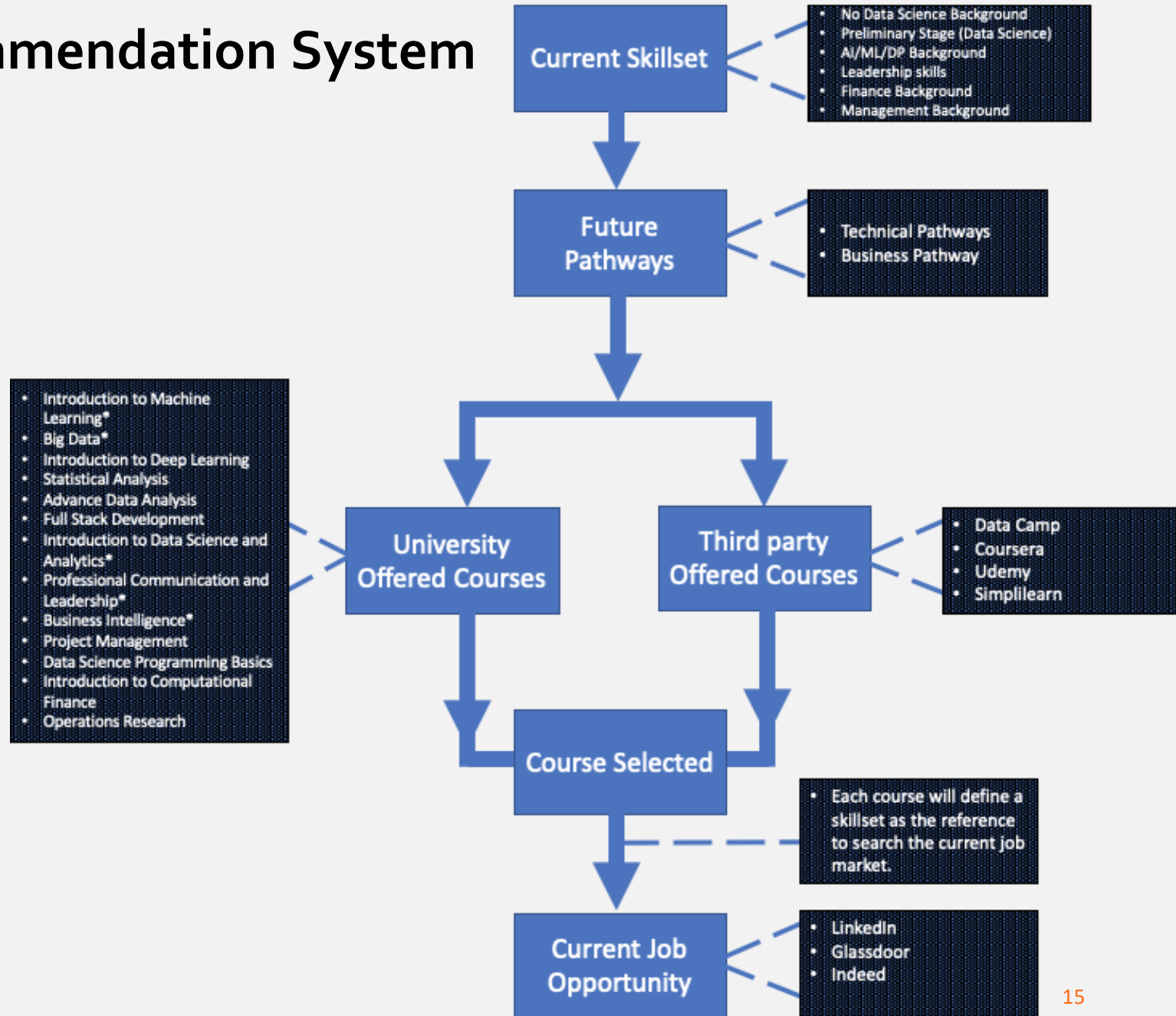
Data Science: EdTech Recommendation System

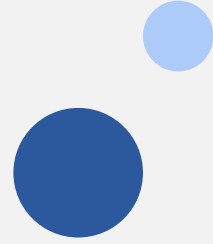
Conclusion:

- ❑ From the analysis of the different dataset we came to conclusion that the courses like Introduction to ML, Big Data, Introduction to DL, Statistical Analysis and Professional & Communication leadership provides the required skill set to get a job in current market.
- ❑ We already designed some course structures based on our research but the problem is how to provide easy access to the upcoming students to understand the need of the market and the courses they should opt for.
- ❑ We are suggesting a recommendation system on a website platform which will provide an easy understanding of the fast changing data science jobs in market. Refer the flow chart:

Advantages:

- ❑ Courses offered will be based on current and required skill set.
- ❑ Easy connect between the job opportunities and the offered course option.
- ❑ Time saving and easy to use.





Thanks for your time!

Any Questions???

