Mentorship Programme

Mini Modules

Mini Overviews

1. Git
2. Web server and Rest API
3. Databases - NoSQL, SQL
4. Pipelines - Data Engineering
5. Visualization - Packages and Tools
6. Cloud Computing Basics
7. Machine Learning Overview
    a. Supervised Machine Learning
    b. Unsupervised Machine Learning
    c. Reinforcement Learning
    d. NLP
    e. Computer Vision
    f. Causal Inference &| Bayesian Models
    g. Survival Analysis

How does these software topics help me as a Data Scientist?

1. **Git** - Allows you to collaborate with others and work together in a technical project
2. **Web server and Rest API** - If you want to make your model accessible via the internet, you need this.
3. **Databases** - If you want to store data in a structured or unstructured format in disk and be able to retrieve and query the data.
4. **Pipelines/Data Engineering** - If you want to repeatedly run some scripts in some order, these tools help you orchestrate.
5. **Visualization** - Packages and tools that help me in visualizing data.
6. **Cloud Computing** - Code works in my local. What next ? How do i make it work for everyone?. Use cloud or painstakingly set up your own server and manage it everyday. Your choice.

How does these ML topics help me as a Data Scientist?

1. **Supervised Machine Learning** - Models built using training/labelled data.
2. **Unsupervised Machine Learning** - Models that doesn't have training/labelled data.
3. **Reinforcement Learning** - What is the optimal strategy i need to take in an environment so that i maximize my returns?
4. **NLP** - Strategies/ models built for textual data. (maybe labelled / unlabelled)
5. **Computer Vision** - Strategies/models built for images. (maybe labelled/unlabelled)
6. **Bayesian Models** - Models that make use of prior knowledge about the system we are trying to predict. Uses bayes law and we can easily measure uncertainty of our predictions.
7. **Survival Analysis** - What is the probability of an event occurring given that 'X' number of days has passed by ? P(Event == 'SomeValue' | Time == 'X')