**Complete Data Science Roadmap**

**Beginners guide**

* **Programming**
  + **Python**
    - **Data structures**
    - **Pandas**
    - **Numpy**
    - **Matplotlib**
    - **Seaborn**

* **R**
* **Statistics**
  + **Mean, Median, Mode**
  + **Confidence Interval**
  + **Null hypothesis, Alternate hypothesis**

* **Probability**
  + **Distributions**
  + **pdf, pmf, cdf**
  + **QQ-plot**
  + **KL-divergence**

* **Machine learning**
  + **Exploratory Data Analysis**
    - **Data cleaning**
    - **Preprocessing**
    - **Handling missing and null values**
    - **Outliers**

* **Supervised**
  + **Regression**
    - **Linear Regression**
  + **Classification**
    - **Logistic Regression**
    - **K-nn**
    - **Decision Tree**
    - **SVM**
    - **Naive Bayes**
    - **Random Forest**
    - **Gradient Boosting**

* **Regularization**
* **Dimensionality Reduction**
* **Loss & metrics**
* **Unsupervised**
  + **Clustering**
    - **K-means**

* **Reinforcement**

* **Interview Preparation**
  + **Multiple choice questions**
  + **Scenario based questions**

* **Solving use cases**
  + **Binary Classification**
    - **Text based**
    - **Tabular**
  + **Multiclass classification**
    - **Tabular**

* **Regression**

**Advanced**

* **Linear Algebra**
  + **Vectors and matrices**
  + **Eigen vectors & eigen values**
  + **SVD**
  + **NMF**

* **Deep learning**
  + **Perceptron**
  + **Neurons and activations**
  + **CNN**
  + **RNN**
  + **Popular Architectures**
    - **Alexnet**
    - **VGGNet**
    - **UNet**

* **Capsule Net**
* **Understanding popular applications**
  + **ChatGPT**
  + **Dall e-2**

* **Transformers**
  + **Bert**
  + **GPT**

* **Interview Prep**
* **Use Cases**
  + **Text based**
  + **Image based**

* **Productionisation**
  + **Docker**
  + **Airflow**
  + **Flask and Rest API**