Standard Model Roadmap

DOMS | Page No.
Date / /

1. Define the Problem

- Problem statement
- objectives
- Scope
- Data Requirements
- Project Plan

2. Data Collection

- Data sources
- Data acquistion
- Data Quality
- Privacy
- Augmentation (if required)
  - increase dataset artificially.

3. Data Preprocessing

- Handling missing data
- Outliers and anomoly detection and Handling
- Cleaning
  - correcting typos, incosistency entries.
- Data splitting into train, validate & test.
- Augmentation (if Required.)

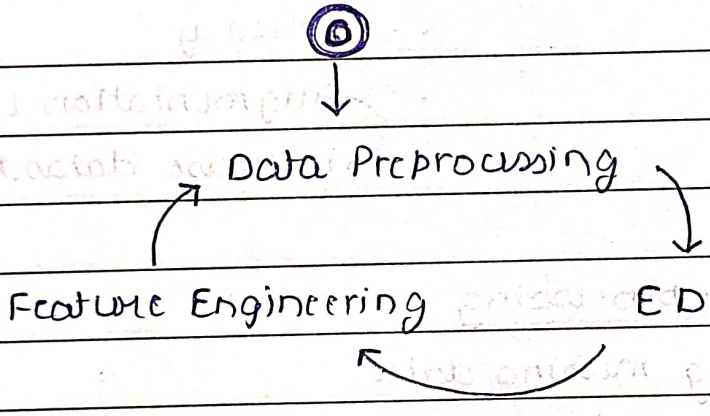
4. Exploratory Data Analysis [EDA]

- Summarization & visualization
- correlation analysis
- outlier & anomoly Detection
- Feature interaction.
  - analyse features via different plots.
- Hypothesis Testing
- Report finding & Making

DOMS | Page No.
Date / /

(5.) Feature Engineering

- New feature creation
- Transformation
  - Log transformation, square, square root, scale etc
- Encoding categorical variables
- Text Data Processing like TF-IDF, Word2Vec. etc
- Scaling.
- Normalisation & standardisation
- Dimensionality Reduction.
- ✓ Feature selection.

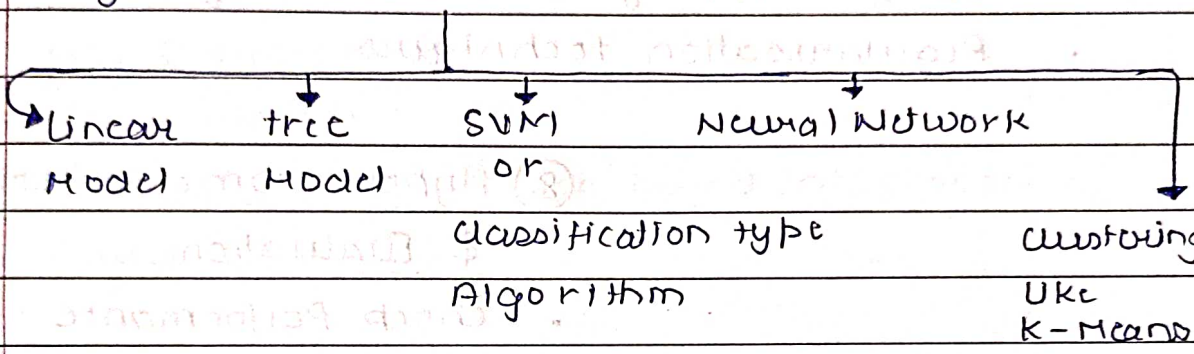


★ All these comes in a loop and any no. of cycles can be performed.

DOMS | Page No.
Date / /

(6.) Model Selection.

- Problem type identification.
  - classification, regression, clustering, recommend etc.
- Model selection criteria
  - identify speed, scalability performance requirement
- Algorithm choice



- Choose between ML & DL
- Ensemble Methods (if applicable)
- Create a custom Model (if applicable)
- Hyperparameter tuning & Regularisation.
- Cross Validation
- Model Evaluation Metrics
- Bench Marking.

★ It is very vast, above mentioned are very generalised and therefore experience via projects.

DOMS | Page No.
Date / /

(7.) DL Architecture

- Choose Loss function
- optimization algorithm
- Hyperparamter tuning
- Epochs and Batches
- Early stopping & checkpointing
- Regularisation techniques

(8.) Hyperparamter tuning & Evaluation.

- check Performance metrics.
- Model comparision
- Model Robustness
- Visualisation & Evaluate Metrics Compare

DOMS | Page No.
Date / /

(9.) Deployment

- Environment setup
  - cloud, infrastructure or hardware to host model.
- Model serialisation.
  - save trained model in pickle or onnx format

- API Development
  if applicable
  Develop a API or web based integration of model to web.

- Model containerisation. using docker, Kubernetes
- Integration with Real time data sources
- Model versioning via implementing version control system. for updates and roll backs.

- Monitoring and logging
- Model inference for new and unseen data
- Feed back loop (if applicable)
  - mechanism to continuously improve model by collecting prediction.

- Performance optimization.
- Load Balancing (if applicable)
- A/B Testing (if applicable)

- Rollout, Back up, Recovery strategy
- compliance and Governance.

any number of times depending upon Results and Objectives.