Machine Learning Overview

7 Steps of Machine Learning

1: Gathering Data

2: Preparing The Data

3: Choosing a Model

4: Training

5: Evaluation

6: Hyperparameter Tuning

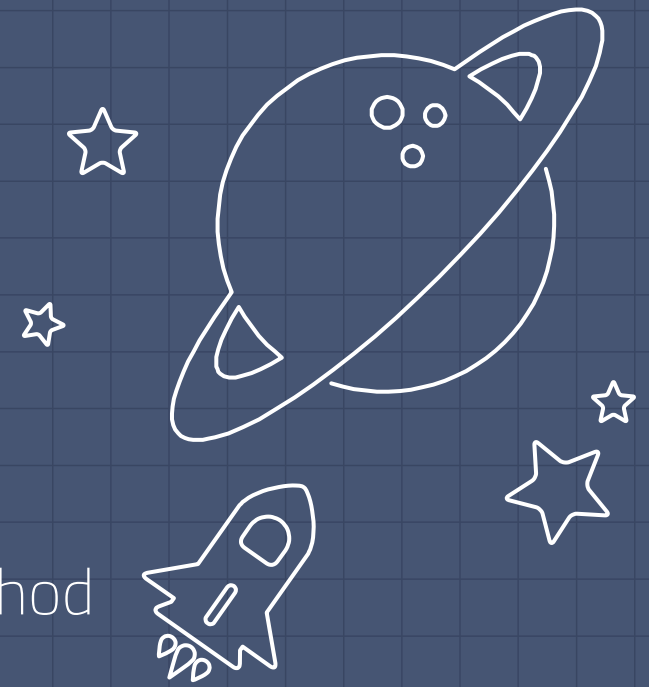
7: Prediction

Gathering Data

1

Steps To Collect Data

1. Determine What Information You Want to Collect
2. Set a Timeframe for Data Collection
3. Determine Your Data Collection Method
4. Collect the Data

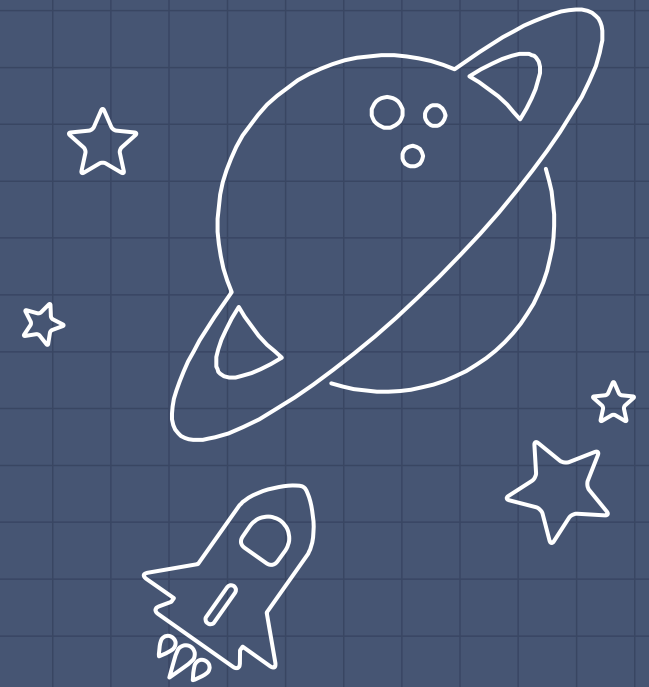


Preparing The Data

2

Uses of Data Collection

1. Improving Your Understanding of Your Audience
2. Identifying Areas for Improvement or Expansion
3. Predicting Future Patterns
4. Better Personalizing Your Content and Messaging



Data Cleaning

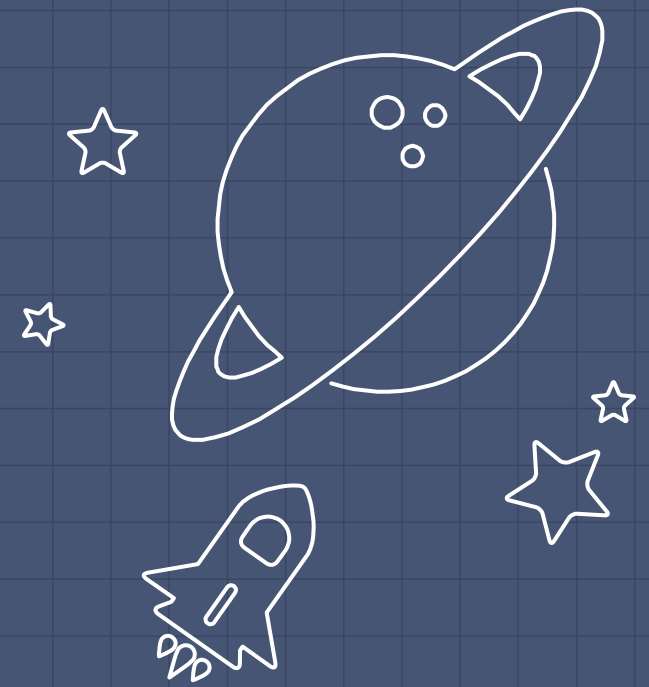
1: Remove duplicate or irrelevant observations

2: Fix structural errors

3: Filter unwanted outliers

4: Handle missing data

5: Validate and QA



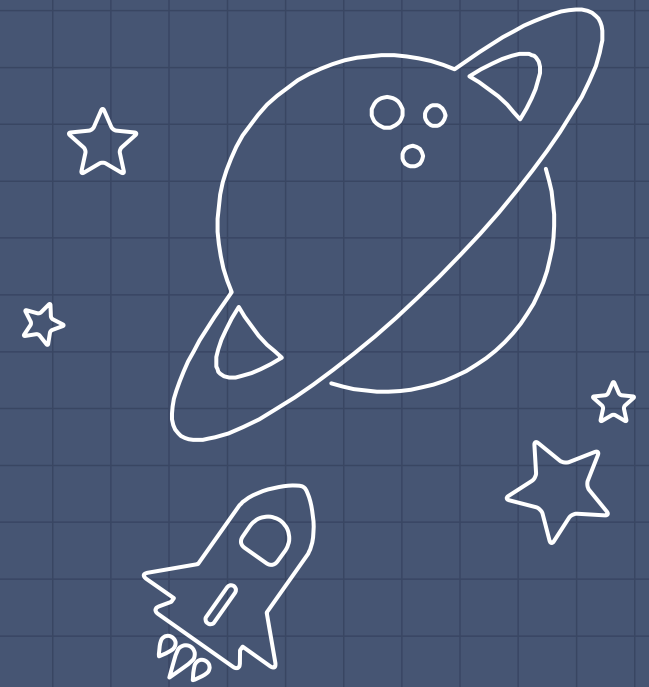
Data Visualization

Python Libraries

- Matplotlib
- Plotly
- Seaborn
- GGplot
- Bokeh
- Altair
- Geoplotlib

BI tools

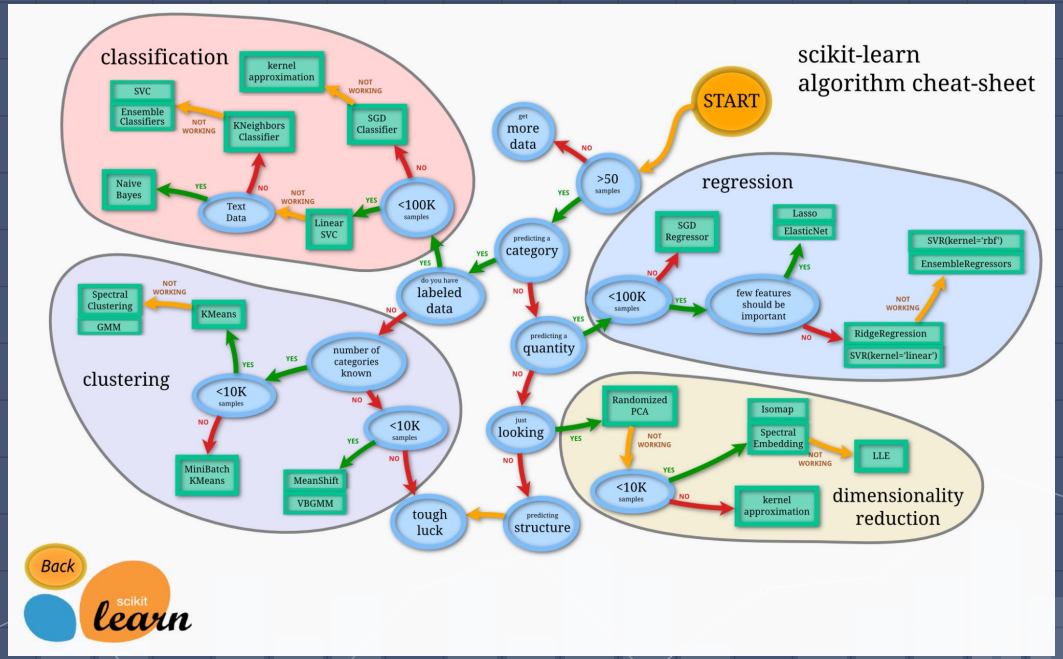
- Tableau
- Power BI

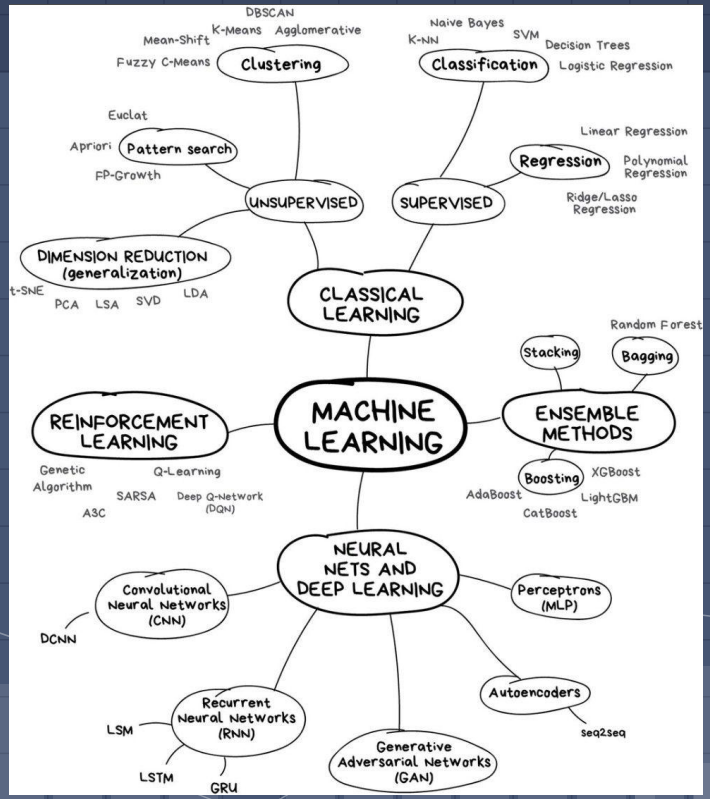


Choosing a Model

3

Algorithm Cheat Sheet

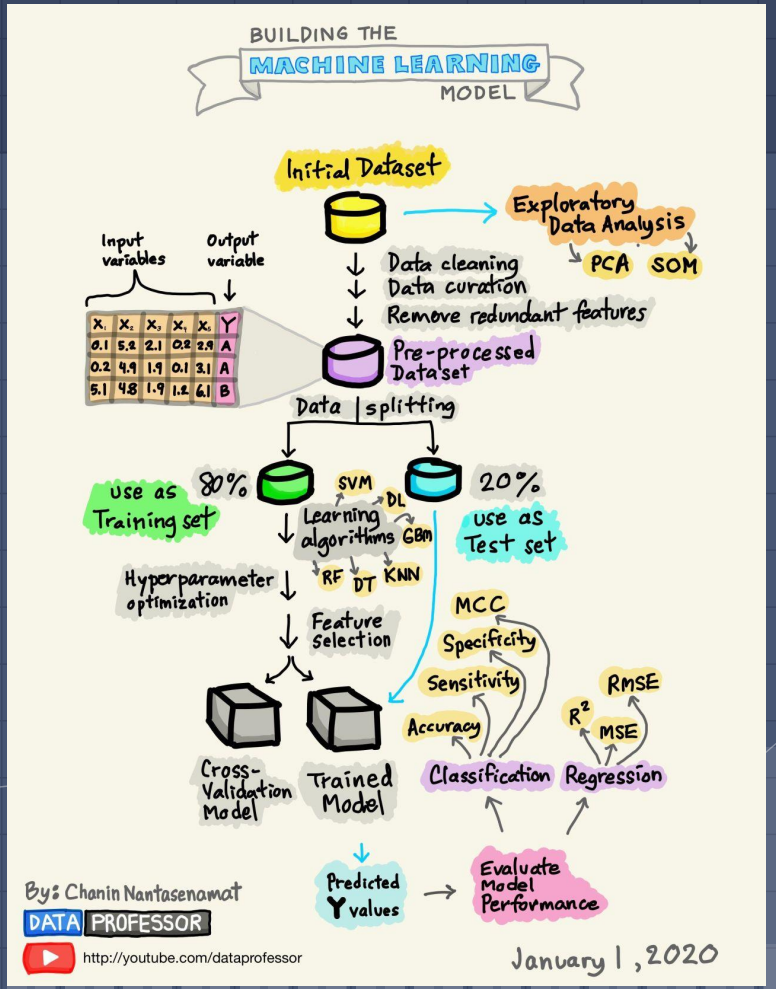




Training & Evaluation

4

Training Cheat Sheet



Hyperparameter Tuning & Prediction

5

GridSearchCV

1. Select the type of model we want to use like RandomForestClassifier, regressor or any other model
2. Check what are the parameters of the model
3. Select the methods for searching the hyperparameter
4. Select the cross-validation approach
5. Evaluate the model using the score

Accuracy Metric

1. Root Mean Square Error (RMSE)
2. or Mean Absolute Error (MAE)

Variance Bias Tradeoff

Low Variance	High Bias	Less Complex	Under Fitting
High Variance	Low Bias	More Complex	Over fitting