Intro to Machine Learning

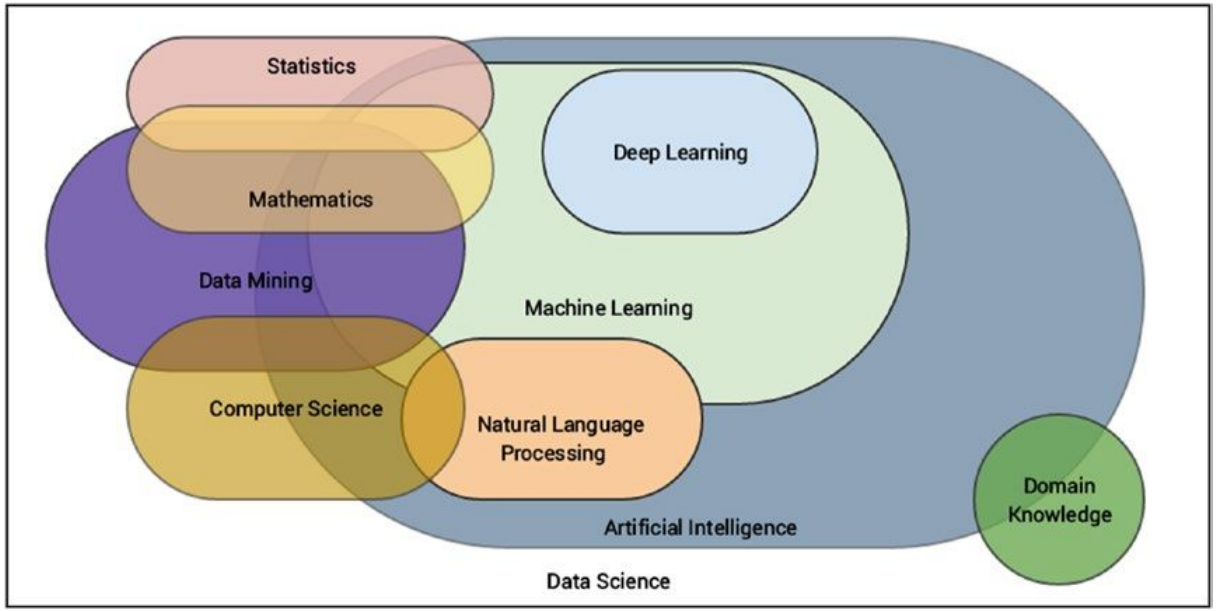
Andreas Chandra

Structure

- Intro to Machine Learning
- Practice Machine Learning Python | R
- Real Case 1 (Regression)
- Real Case 2 (Classification)

Note that we don’t discuss technical things

Concept

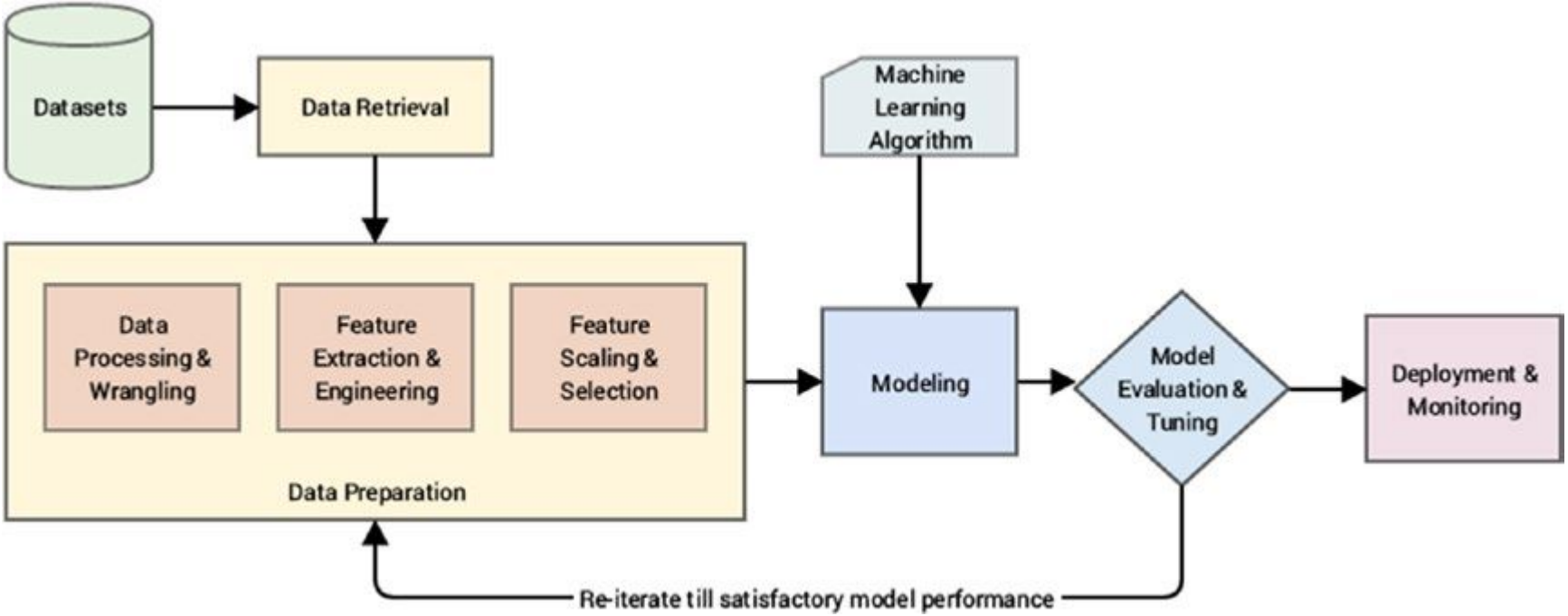


Practical machine learning with python, Sarkar et al

Concept

- Field of study that gives computers the ability to learn without being explicitly programmed - Arthur Samuel
- A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E. -Tom M. Mitchell.

Flow


Datasets
Data Retrieval
Machine Learning Algorithm
Data Processing & Wrangling
Feature Extraction & Engineering
Feature Scaling & Selection
Data Preparation
Modeling
Model Evaluation & Tuning
Deployment & Monitoring
Re-iterate till satisfactory model performance

Tabular, Text, Image, Speech, What’s Next?

Supervised Learning

Concept: take in data samples and associated outputs with each data sample during the training process

Objective: to learn a mapping or association between input data samples *x* and their corresponding output y’ based on multiple training data instances.

Algorithms: K-NN, SVM, Naive Bayes, Decision Tree (Random Fores & XGBoost)

Unsupervised Learning

Concept: the model don’t need the labels but tries to learn inherent latent structures, patterns and relationship from given data without any supervision.

Algorithms: K-Means, K-Median, DBSCAN, Association rule

Evaluation

- Cross Validation
- Dummy Classifier
- Accuracy, Precission, Recall, F1-Score

Exercise - Concept

Suppose your program watches which person classified you do or you do not mark as verified, and base on that learns how to better filter verified.

What is the experience, task, and performance in this condition?

Exercise - Evaluation

What situation should we consider for precission and recall?

		Actual	
		Positive	Negative
Predicted	Positive	**True Positive**	**False Positive**
	Negative	**False Negative**	**True Negative**

https://towardsdatascience.com/beyond-accuracy-precision-and-recall-3da06bea9f6c

Exercise - Use Cases

- E-Commerce
- Travel
- News & Media
- Entertainment
- Finance