

MoF-DAC

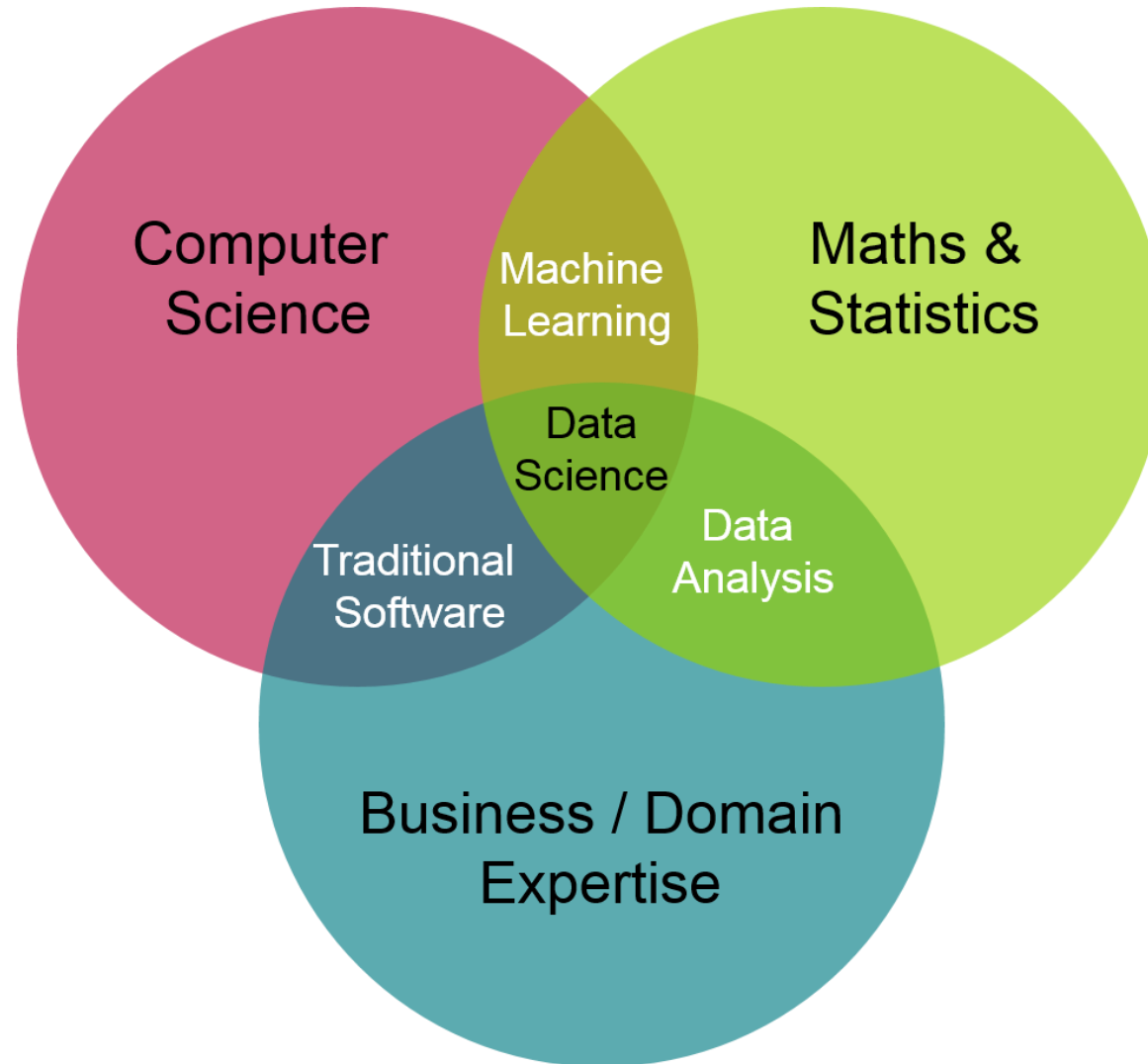
Ministry of Finance
Data Analytics Community

MACHINE LEARNING BASIC CLASSIFICATION

Aris Budi Santoso

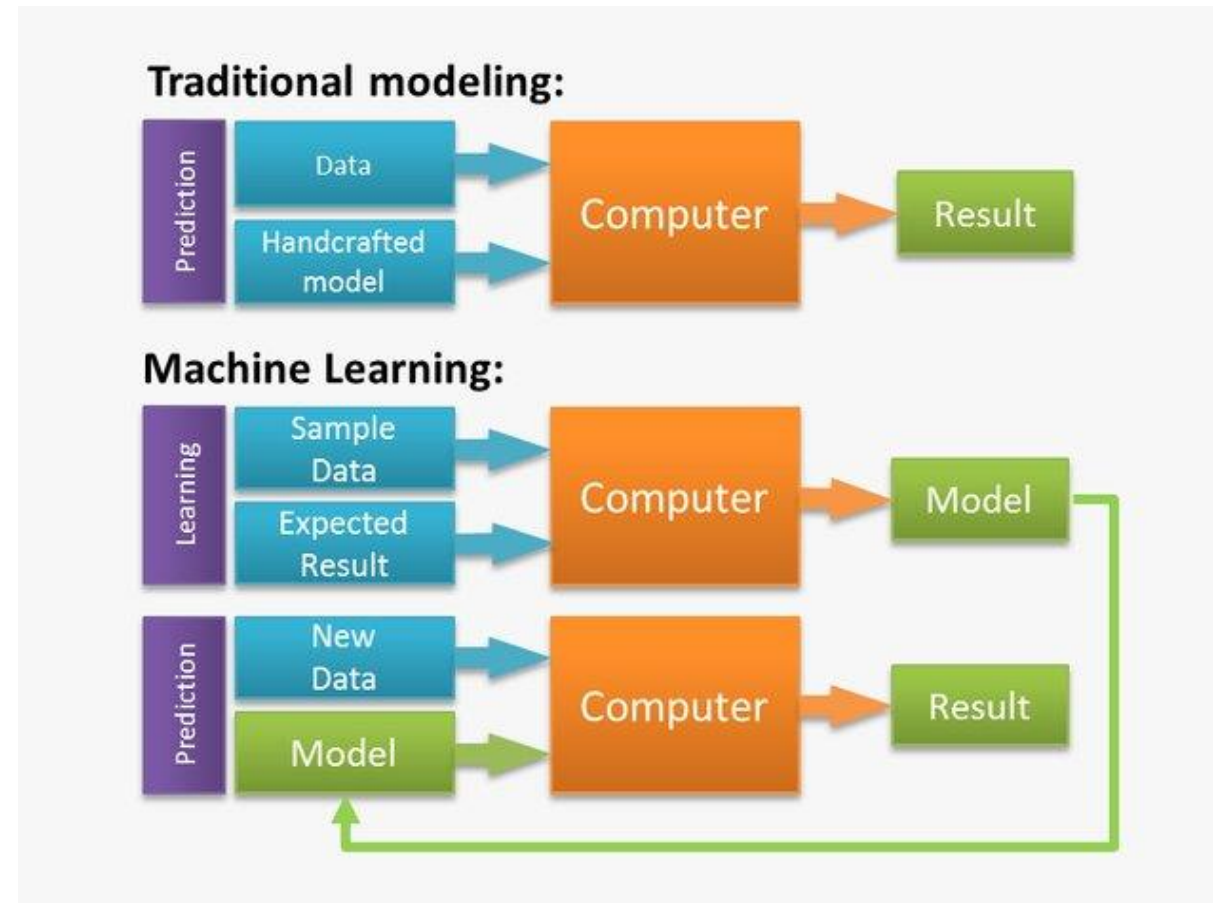
Reza Rizky Pratama

Data Science



Machine Learning

Machine learning is a field of computer science that aims to teach computers how to learn and act without being explicitly programmed [1]



Mehra, Sidharth & Hasanuzzaman, Mohammed. (2020). Detection of Offensive Language in Social Media Posts

[1] <https://deepai.org/machine-learning-glossary-and-terms/machine-learning>

JENIS MACHINE LEARNING



Supervised Learning

Model ini dilatih dengan menggunakan data yang sudah diberikan label terlebih dahulu. Algoritmanya akan mempelajari pola dari pasangan data dan label dan kemudian



Unsupervised Learning

Model ini dilatih hanya dengan data input tanpa diberikan arahan mengenai data output. Salah satu algoritma yang untuk machine learning ini adalah *dimensionality reduction*.



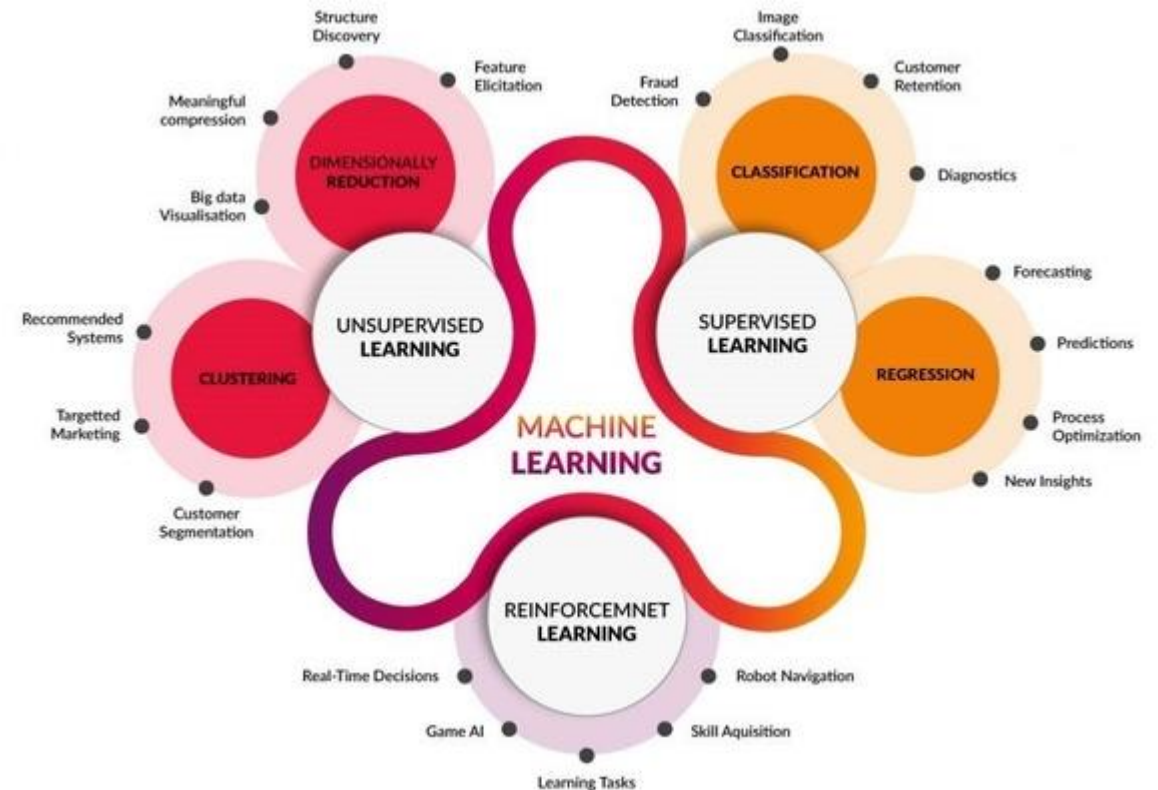
Semi-Supervised Learning

Merupakan gabungan antara Supervised dan Unsupervised Learning. Pada model ini, data latih sebagian diberikan label dan sebagian tidak. Contohnya adalah Google Photo.



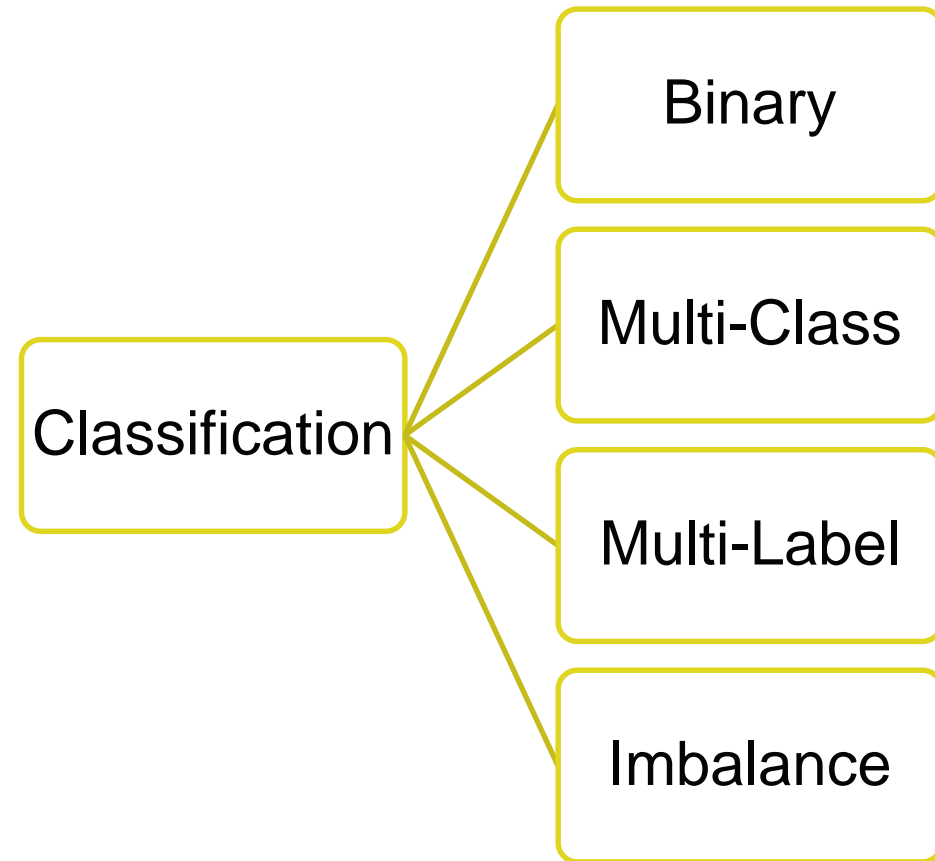
Reinforcement Learning

Dikenal sebagai model yang belajar menggunakan system *reward* dan *penalty*. Model ini dilatih untuk membuat keputusan terbaik secara berurutan untuk memaksimalkan kesuksesan. Contohnya adalah Alpha GO.



Classification

Classification is a task that requires the use of machine learning algorithms that learn how to assign a class label to examples from the problem domain [2]



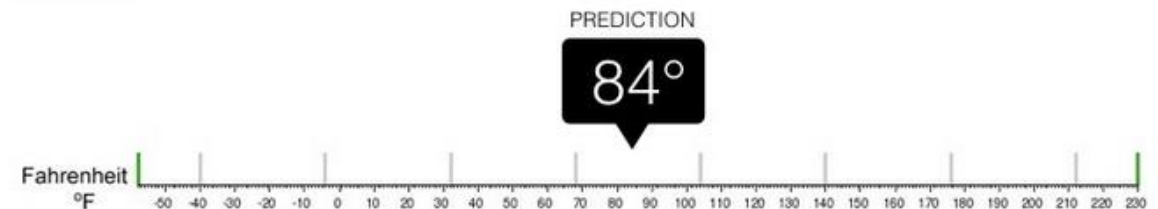
Classification vs Regression

Regression	Classification
Digunakan untuk memprediksi data kontinu (<i>continuous quantity</i>)	Digunakan untuk memprediksi label diskret pada suatu kelas (<i>discrete class label</i>)
Regresi dengan multiple input biasa disebut multivariate regression	Klasifikasi dengan 2 label kelas disebut binary dan lebih dari 2 kelas disebut dengan multi-class
Scoring yang umum digunakan : RMSE, R^2 , MAE, MAPE	Scoring yang umum digunakan : Accuracy, F1-score, ROC-AUC
Contoh : prediksi harga rumah, prediksi GDP, prediksi pertumbuhan penduduk.	Contoh : <i>fraud-detection</i> , <i>email spam filter</i> , <i>image classification</i> .



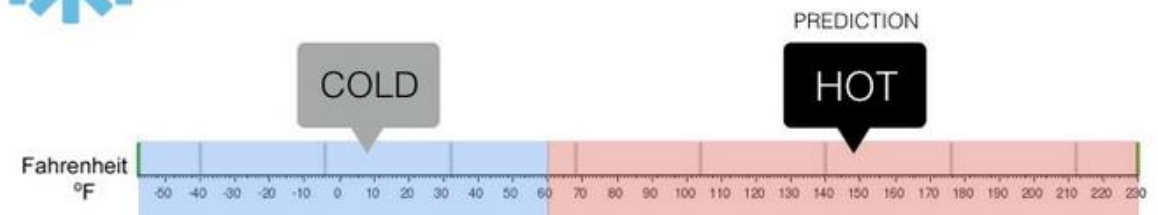
Regression

What is the temperature going to be tomorrow?

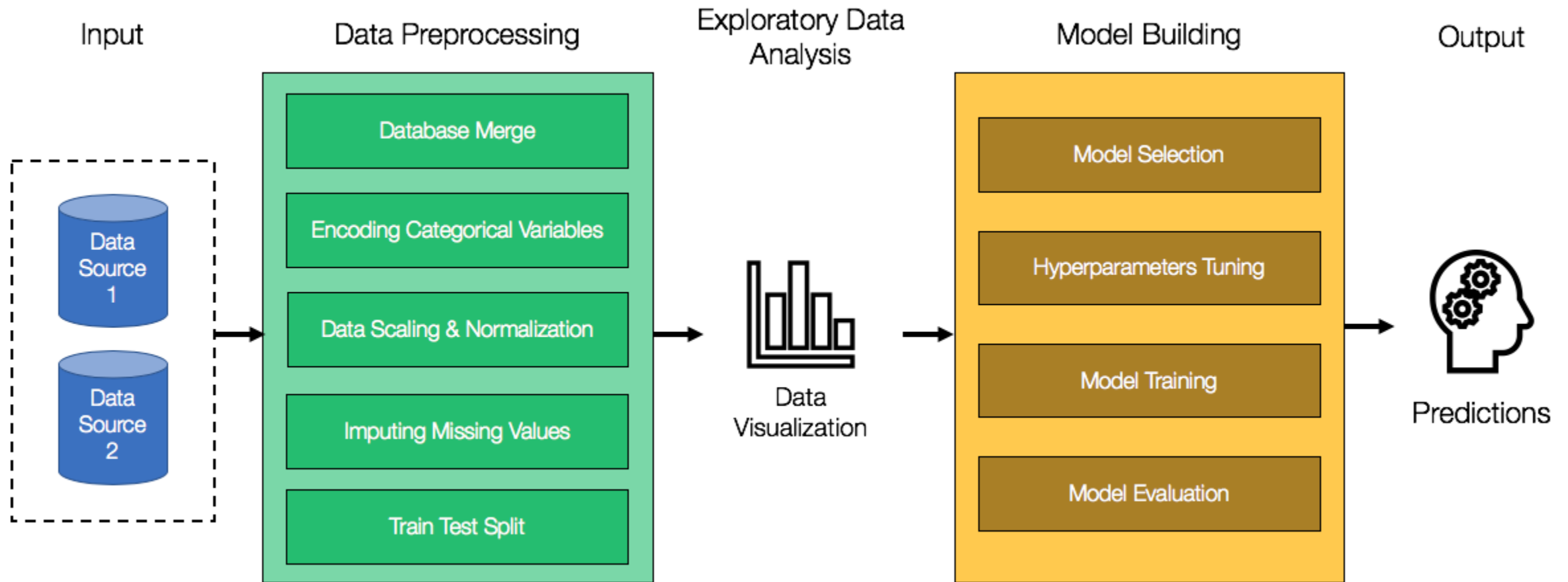


Classification

Will it be Cold or Hot tomorrow?



Machine Learning Workflow





THANK YOU.