National University of Computer and Emerging Sciences,
Islamabad, Pakistan

Machine Learning Basics

CS-4025: Deep Learning for Perception

Muhammad Atif Saeed

Department of AI & DS

Slides Credit: Dr. Akhtar Jamil

Department of Computer Science

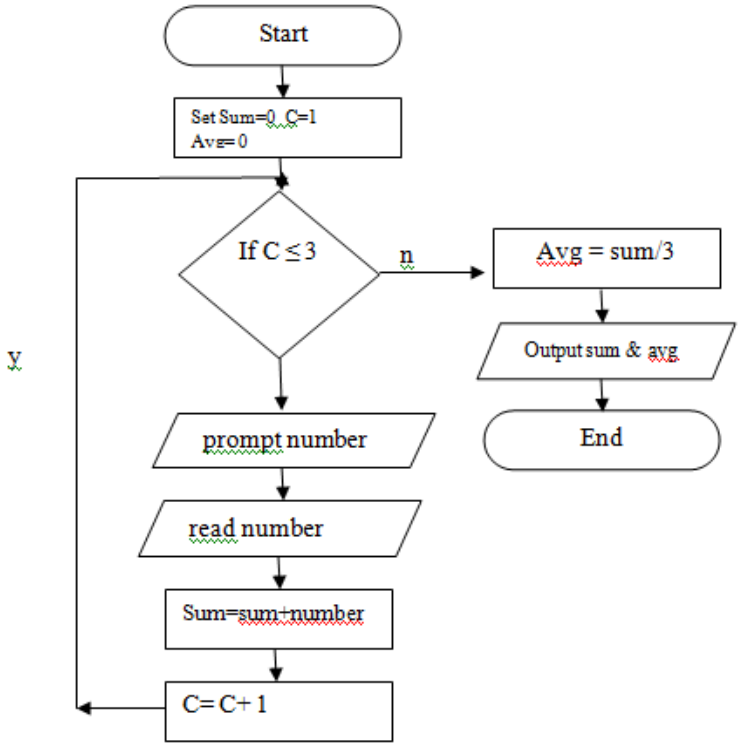
Goals

- Review of Previous Lecture
- Today's Lecture
  - Machine Learning Overview
  - Linear models
  - Loss functions, linear regression, gradient descent, overfitting, underfitting generalization, regularization, cross-validation

Recap of Previous Lecture

What is Learning?

- How can we solve a specific problem?
  - We write a program with a set of rules that are useful to solve the problem.
  - Example: Find average of three numbers



What is Learning?

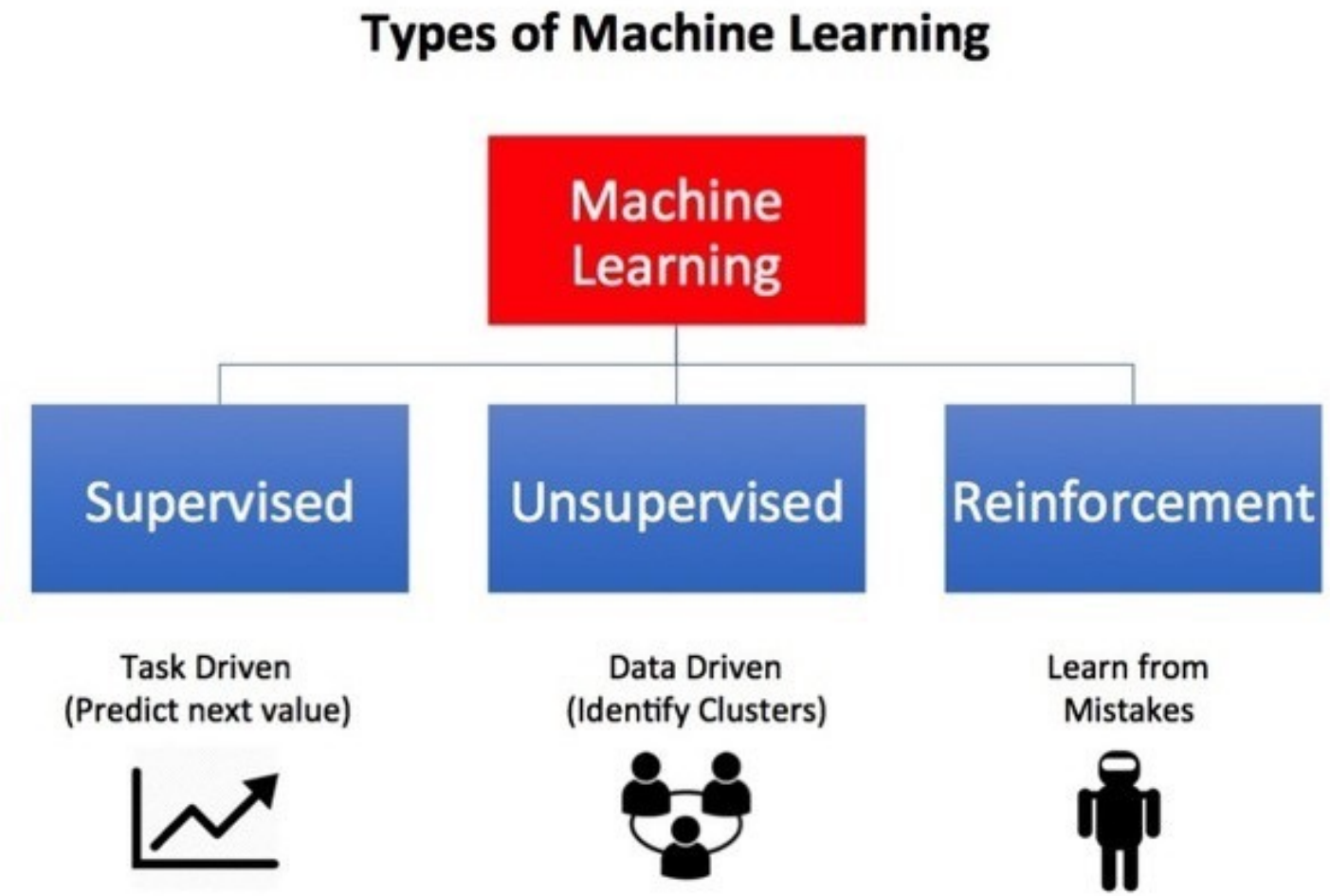
- In many situations it is very difficult to specify those rules to solve a problem.
- For example, given a picture determine whether there is a cat in the image

What is Machine Learning?

- Definition: “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. Mitchel

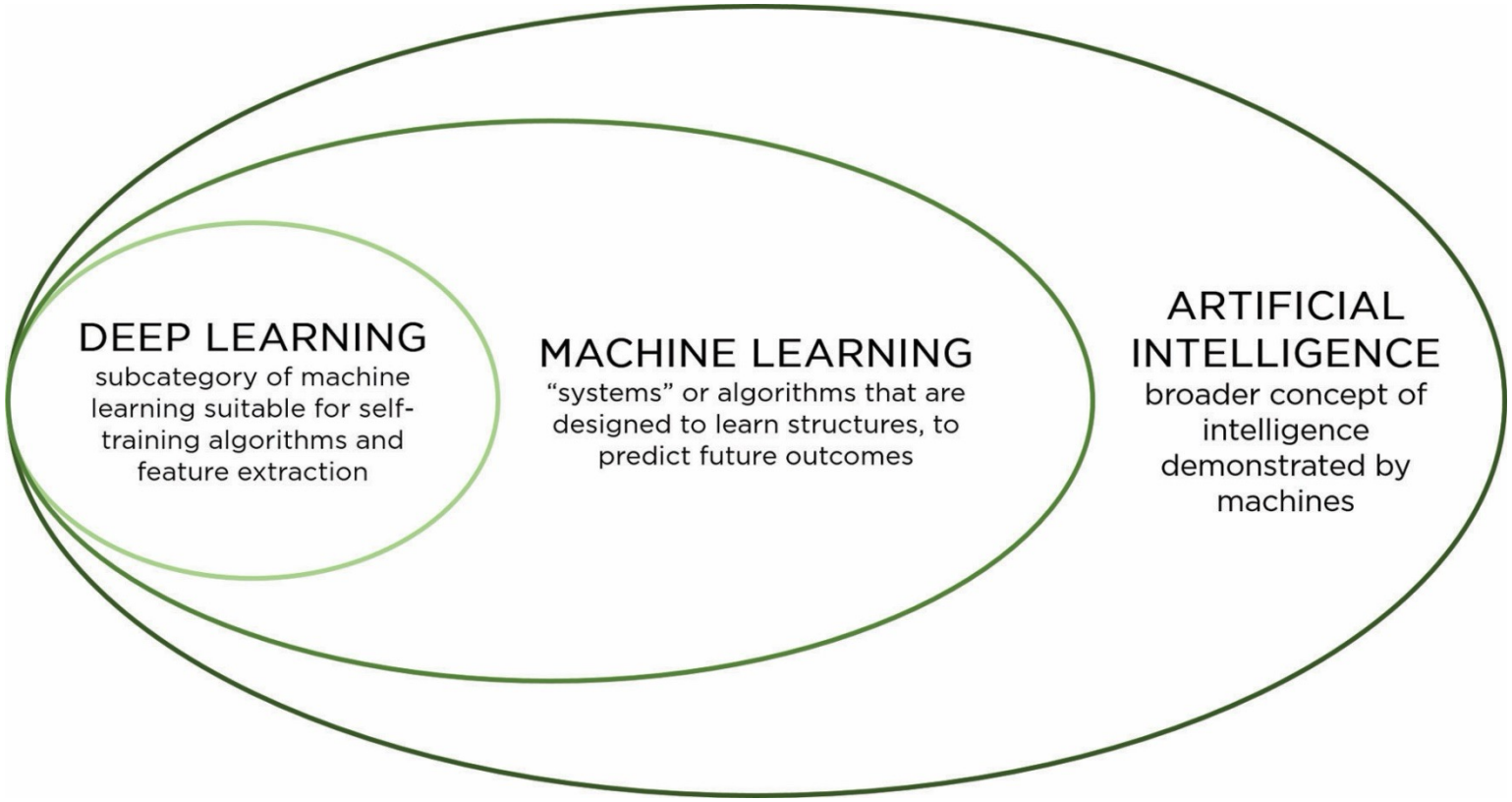
Types of Machine Learning…



Today's Lecture

Machine Learning Basics

- Machine learning is a subset of artificial intelligence (AI)
  - Development of algorithms that enable computer systems to learn through learning from data
  - Improve their performance on a specific task
  - No explicit programming
- Deep learning is a specific kind of machine learning.
- To understand deep learning well, one must have a solid understanding of the basic principles of machine learning.



Deep Learning

Specific subfield of machine learning.

Machine Learning

- For different learning algorithms, e.g. the linear regression algorithm, fitting the training data differs
- Models can have many hyperparameters
- Learning algorithm statistically estimate complicated functions to learn about data
- Uses some optimization algorithm to learn about data, most widely used algorithm called gradient descent.

Machine Learning

- Machine learning algorithms are complex
- A machine learning algorithm is an algorithm that is able to learn from data.
- ML algorithms combines various algorithm components, such as an optimization algorithm, a cost function, a model, and a dataset, to build a machine learning algorithm.

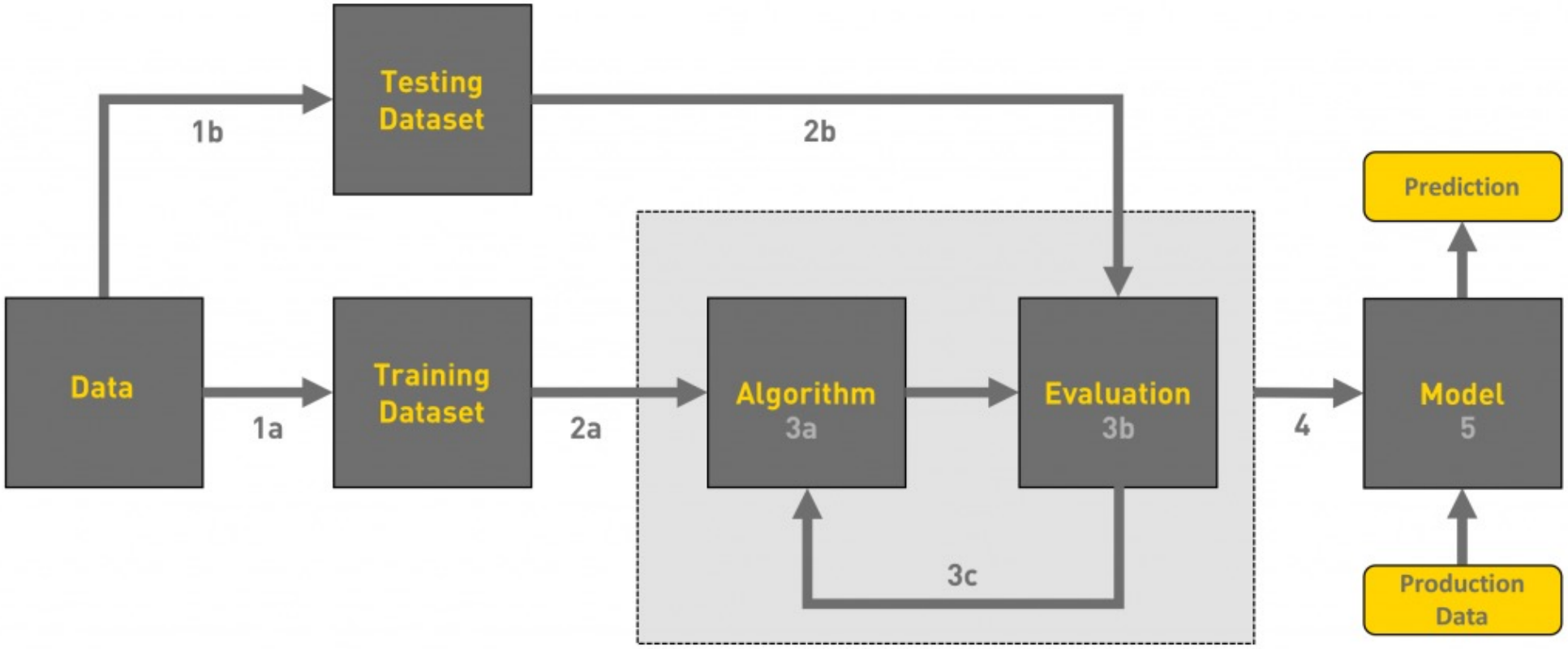
Hyperparameters vs Parameters

- Hyperparameters and parameters are both essential components of a machine learning model.
  - Have different purposes and distinct characteristics.
- **Parameters:**
  - Parameters are the internal variables of a machine learning model that are learned during the training process.
  - Model adjusts to fit the training data to understand the relationships in data.
  - For example, in a linear regression model, the parameters are the coefficients assigned to each feature, and in a neural network, the parameters include the weights and biases of the network's neurons.
  - Keep updating these parameters iteratively to minimize a chosen loss function

Hyperparameters vs Parameters

- **Hyperparameters:**
  - External parameters that are not learned by the model.
  - Set before the training process begins.
  - They have influence on learning algorithm
    - Model's generalization, convergence speed, and overall performance.
- These parameters are chosen by the model developer based on experimentation and domain knowledge.
- For Examples: For ANN learning rate, batch size, number of hidden layers, choice of optimization algorithm, etc.

Workflow of ML tasks



Workflow of ML Problem

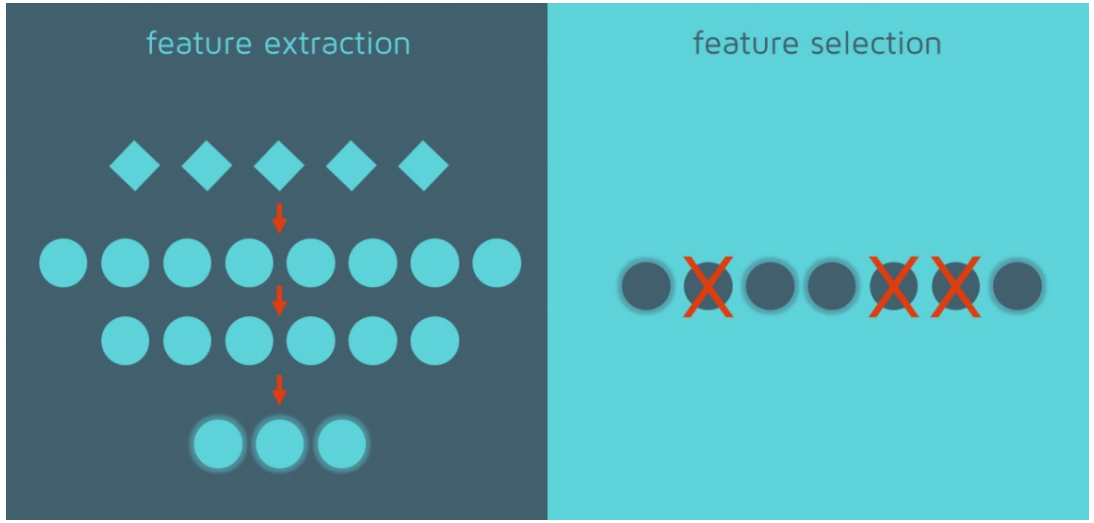
- The process of machine learning typically involves the following key steps:
  - **Data Gathering and Preparation**
  - **Feature Extraction/Selection**
  - **Model Selection and Development.**
  - **Train and Test model**
  - **Deploy your trained model.**
  - **Monitor and Manage models**

Workflow of ML Problem

- **Gathering and Preparation**
- Gather the data related to the problem you want to solve
  - File, databases, sensors or other sources
- May not be directly fed to ML model
  - e.g. textual data
  - Noisy data
  - Missing data
- Quality of data is important for learning of ML models
- Normalization (data scaling or feature scaling)
    - Preprocessing technique
    - Transforms the features (variables) of a dataset into a common scale.

Workflow of ML Problem

- **Feature Extraction/Selection:**
  - Identifying the most important and relevant features (attributes)
  - Used for training the machine learning model.
  - Improves the performance



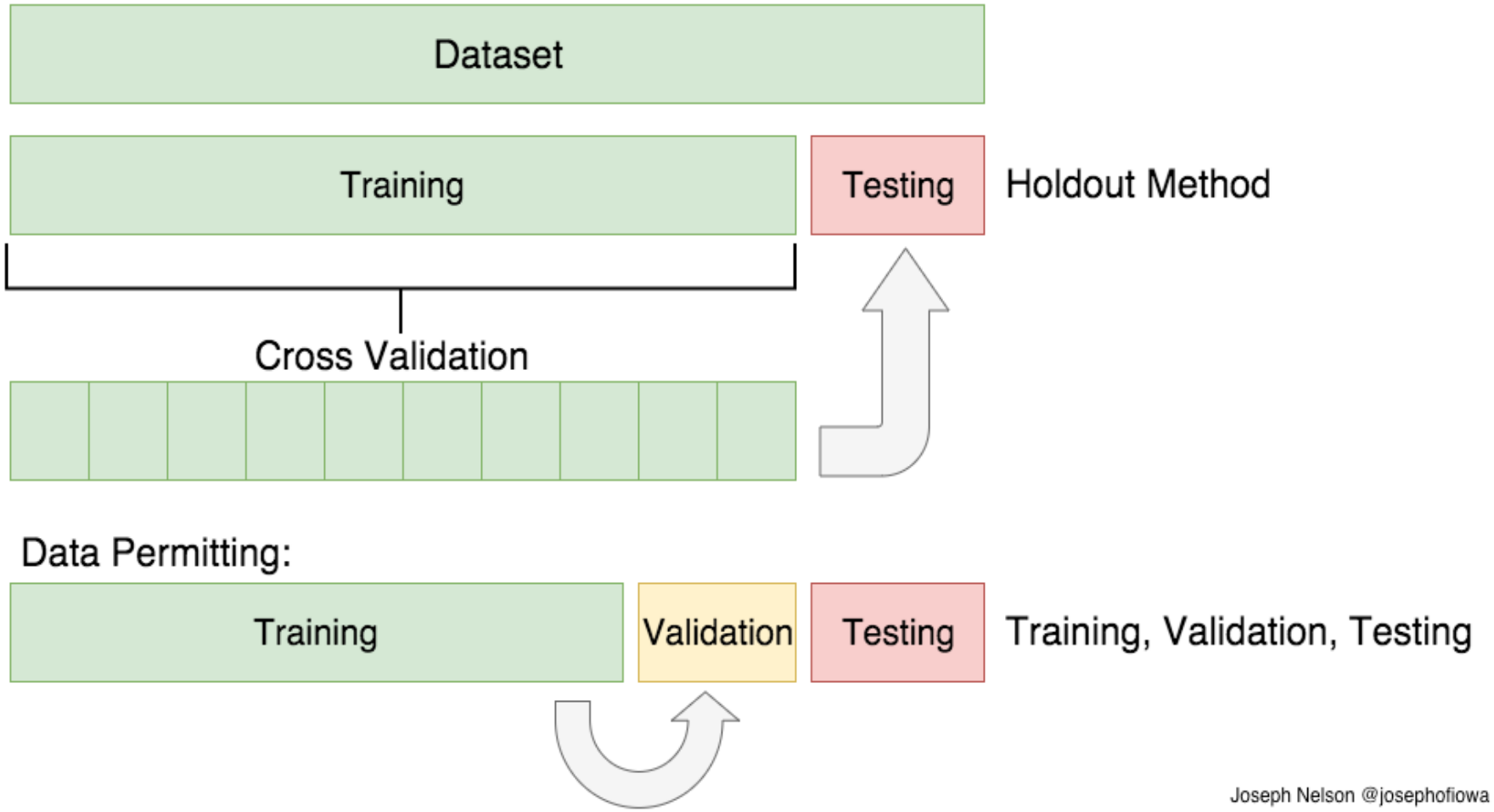
Model Selection and Development

- What kind of data you have?
- Select model accordingly.
  - Supervised or Unsupervised
  - Classification or Regression
  - Time series data or Image Data
  - Textual data
- Develop your model using established ML techniques
  - Scikit-Learn
  - Keras/Tensorflow

Train, Test and Evaluate model

- Generally, the data is split into 3 three subsets:
  - **Training**, **Validation** and **Testing data sets**.
- **Training data set**: Used to train the model
- **Validation data set:** Tune the parameters of the model.
- **Test data set:** Test the performance of classifier on unseen data

Training, Validation and Testing Data



Train, Test and Evaluate model

- Cross-Validation
- Set aside some portion of the data for validation and Train on rest of it.
- **LOOCV (Leave One Out Cross Validation)**
  - Perform training on the whole training data set but leaves only one sample for validation
- **K-Fold Cross Validation**
  - The data-set into split into k subsets(folds)
  - Perform training on the all the subsets but leave one(k-1)
  - Iterate for all folds

Monitor and Manage models

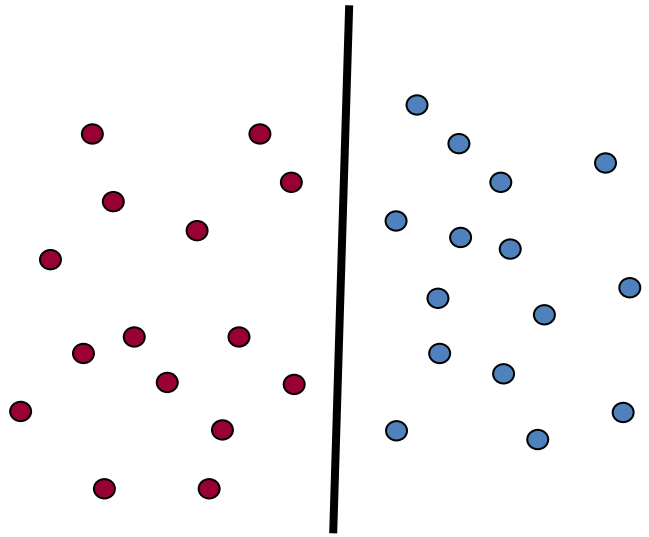
- Monitor the predictions on an ongoing basis.
- Manage your models and model versions
  - New models with versions numbers

Linear models for learning

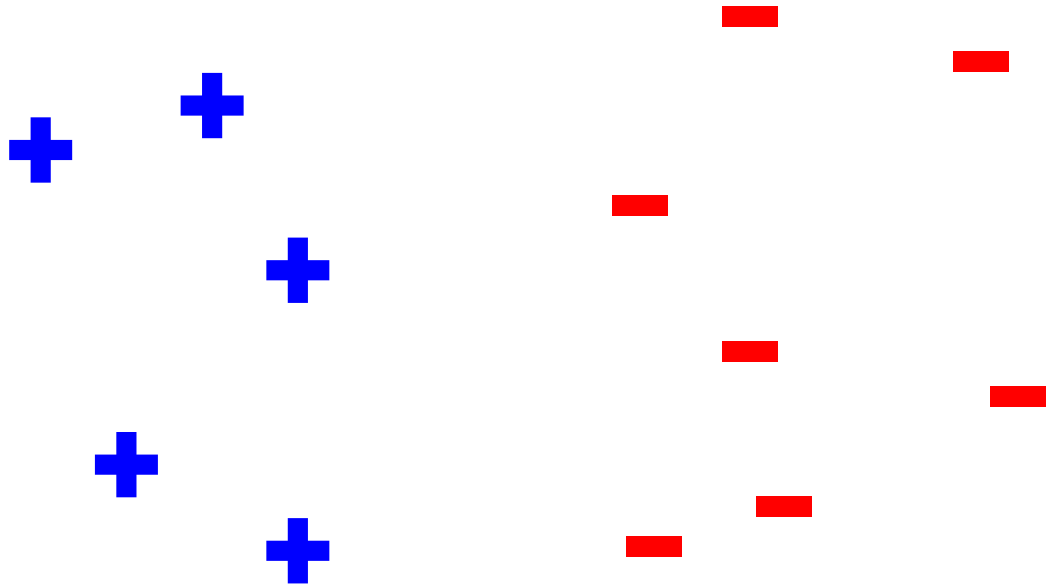
An assumption is *linear separability*:

- in 2 dimensions, can separate classes by a line
- in higher dimensions, need hyperplanes

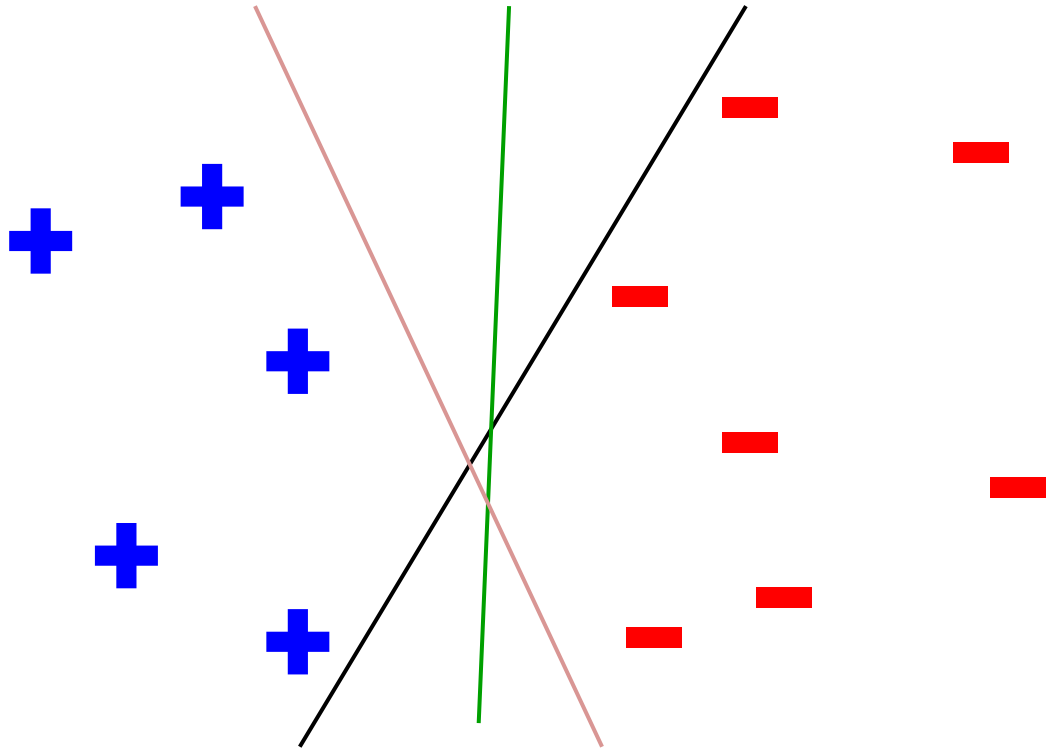
A *linear model* is a model that assumes the data is linearly separable



Which line will it find?



Which line will it find?



Only guaranteed to find ***some*** line that separates the data

Linear regression

- The goal of simple linear regression (univariate) model is to finds the relation between two variable.
  - A single feature (variable x) and a continuous valued response (target variable y).
  - X is called independent variable (predictor)
  - Y is called the dependent (target or response) variable.

$$y = w_0 + w_1 x$$

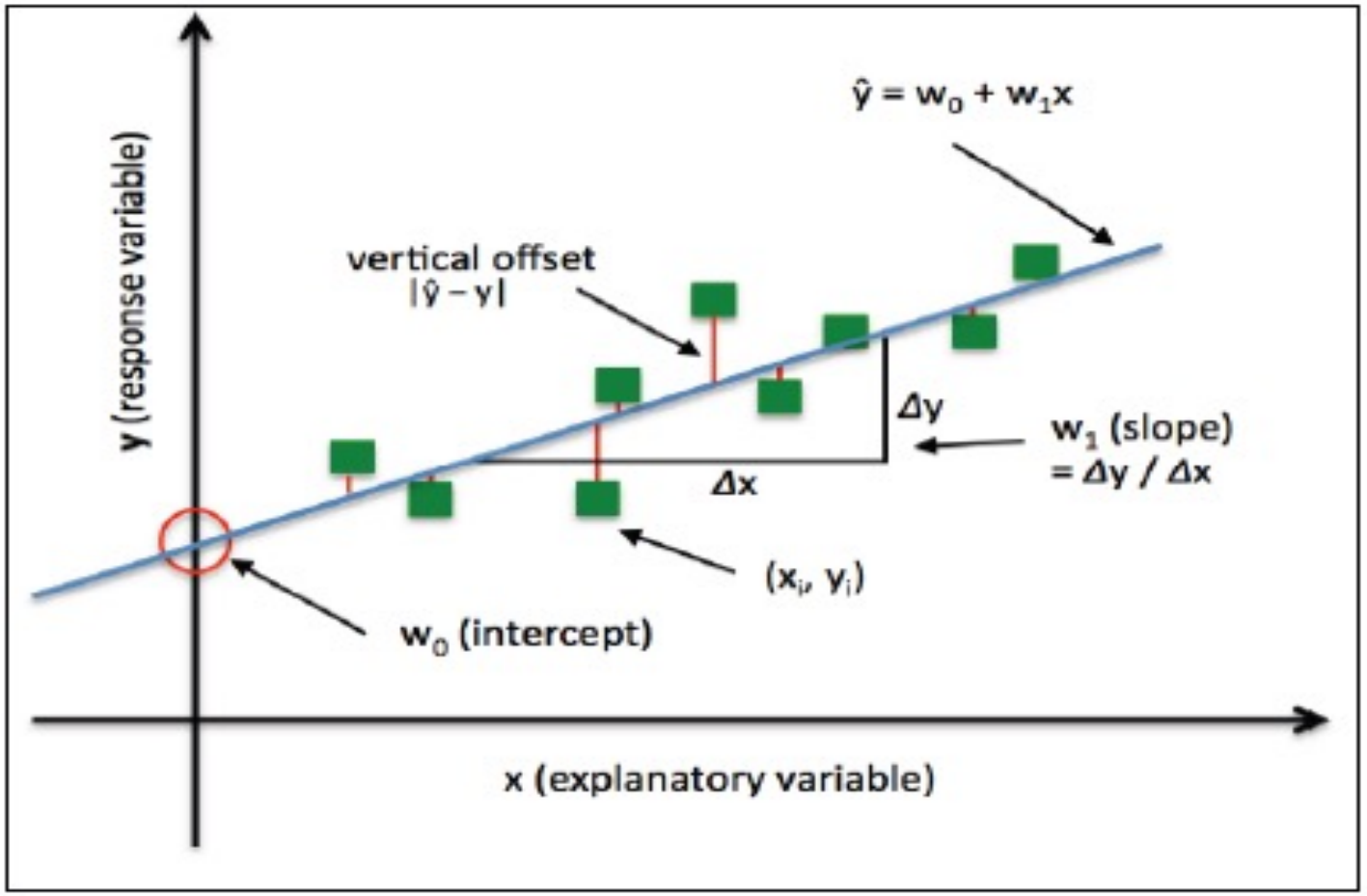
Linear regression

- The weight w_0 represents the y axis intercepts and w_1 is the coefficient of the feature (x variable).
  - w_0 and w_1 are unknown
- The goal of linear regression is to learn the weights of the linear equation
  - Describe the relationship between the **x** and **y**
- Then this relation can be used to predict the responses of new data

Linear regression

- Linear regression can be understood as finding the best-fitting straight line through the sample points.
- This best-fitting line is also called the **regression line.**
- The distance between the regression line to the sample points are the so-called **offsets** or **residuals**
  - The errors of our prediction.

Linear regression



Linear models in general

- For linear model:

$$y = w_0 + w_1 x$$

- These are the parameters we want to learn
- Need to define a criteria to optimize these parameters of the model
  - cost function (objective)
    - Minimize the cost function

Thank You ☺