MachineLearnAthon - Microlecture
Types of Classification Models

Classification

15.10.2024

Recap previous microlecture Classification-1

- Types of Machine Learning Tasks
- Stages of Supervised Machine Learning Pipeline
- Proper Machine Learning Modeling

Learning outcomes of today

After successfully completing this micro-lecture, you are able to….

- Identify the different types of machine learning models.
- Understand the distinction of Deep Learning from Machine Learning.

Agenda for today

- Defining a supervised ML task
- Linear Models
- Non-Linear Models
  - Deep Neural Networks

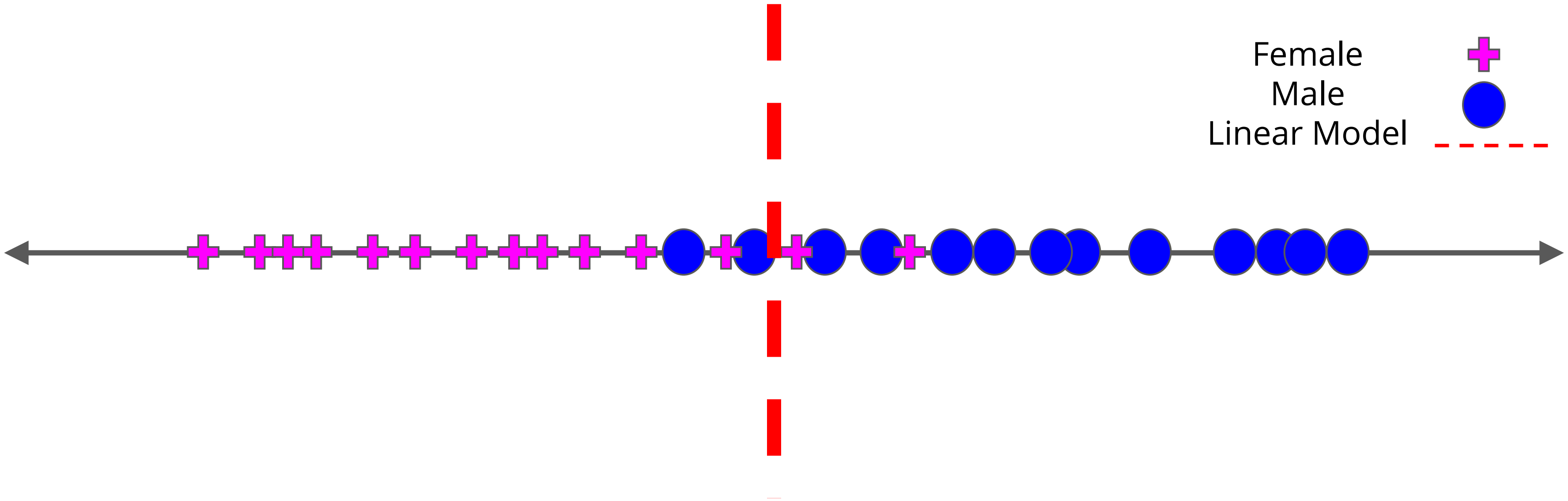
What is ML Model?

NB! Notations can vary across authors

- **X** - input space (set of all possible instances)
- Y - output space (all possible labels / target)
- f : **X** → Y **- any such function is a classifier/regressor**
- **x** ∈ **X** - instance
- y ∈ Y - actual / true label/target of an instance **x**
- f(**x**) = ŷ - predicted label/target of instance **x**

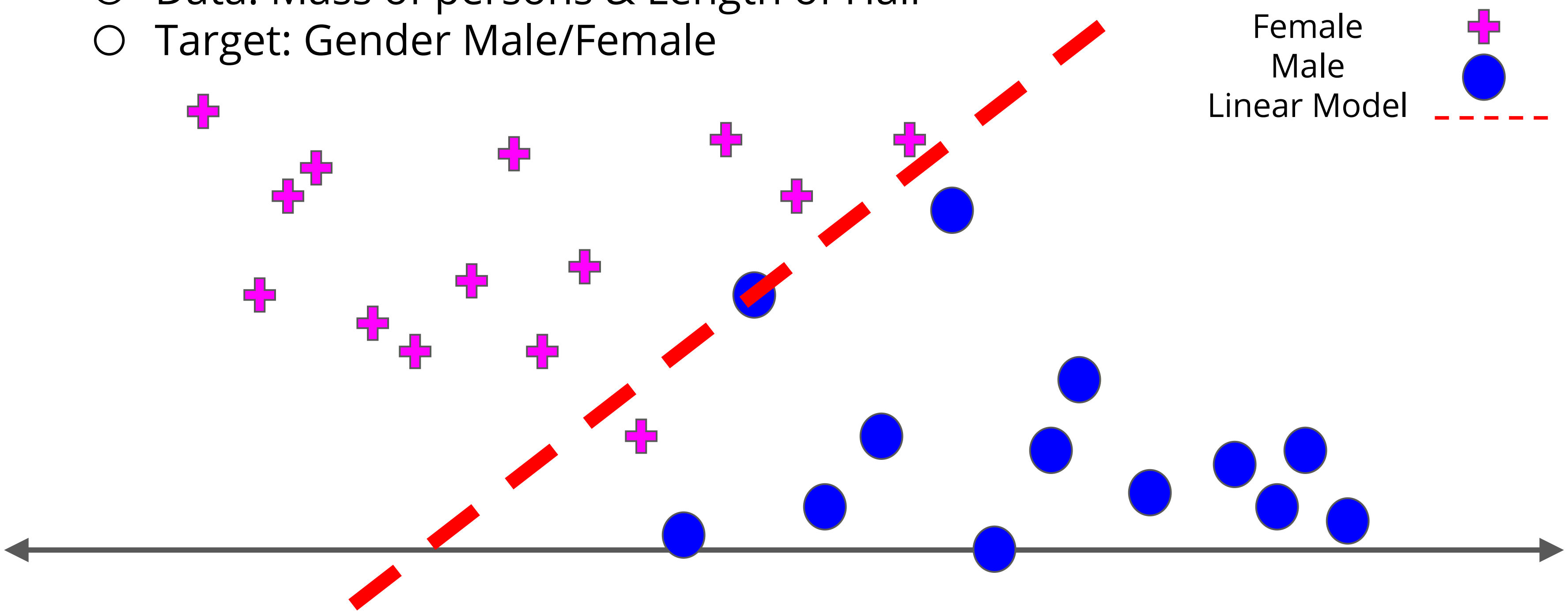
Linear Models in 1D

- Linear Models: f(**x**) = Linear Function of **x**
- Suppose:
  - Data: Mass of persons
  - Target: Gender Male/Female



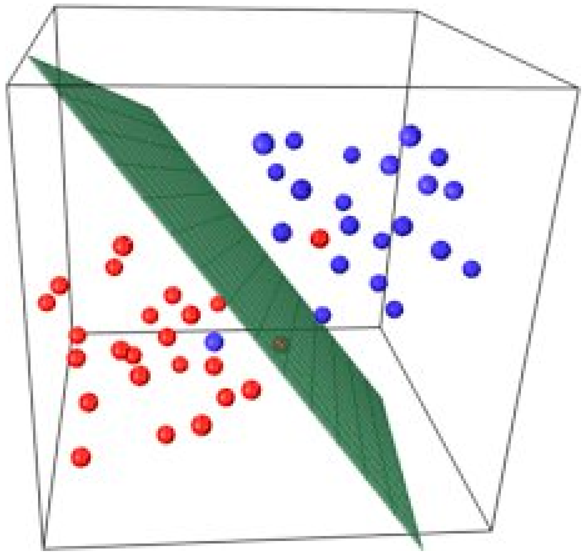
Linear Models in 2D

- Linear Models: f(x) = Linear Function of x
- Suppose:
  - Data: Mass of persons & Length of Hair
  - Target: Gender Male/Female



Linear Models in 3D

- Linear Models: f(x) = Linear Function of x
- Suppose:
  - Data: Mass of persons & Length of Hair & Height
  - Target: Gender Male/Female



- **Finds the best fitting hyper-plane in n-dimensional space**

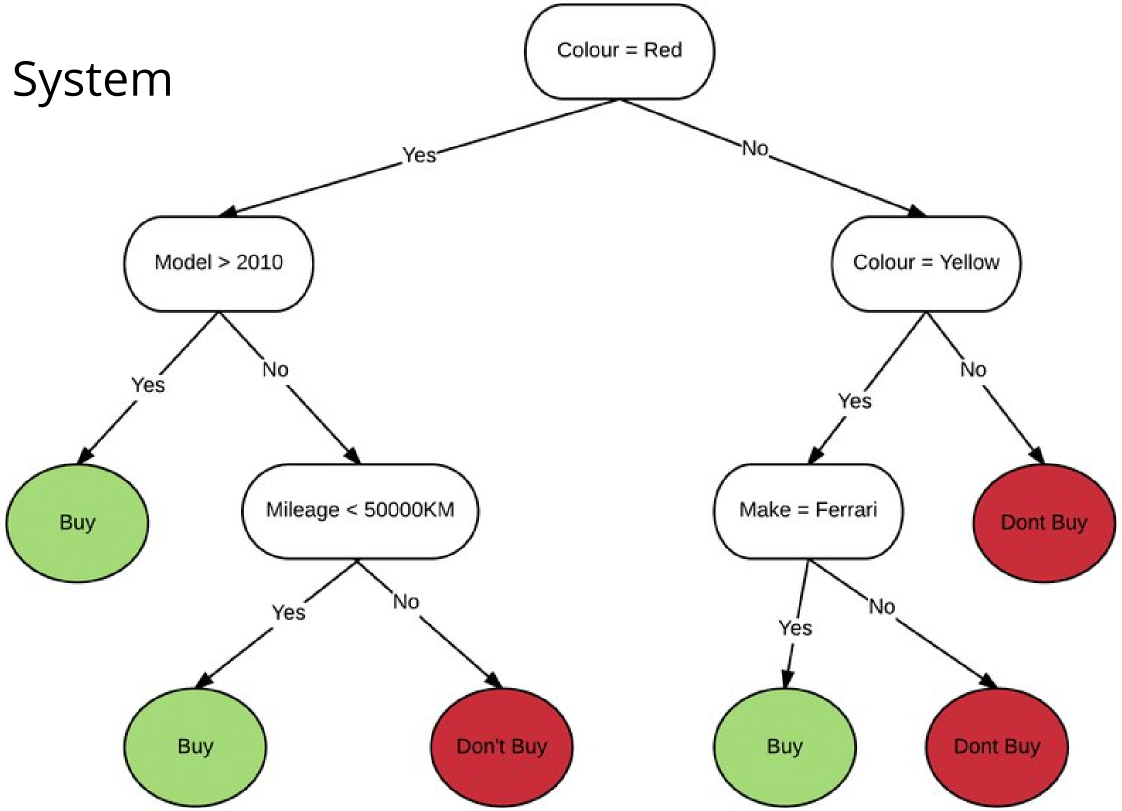
- **Examples:**
  - **SVM:** find best linear function maximizing margins with support vector instances
  - Logistic Regression
  - Perceptron

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.

Non-Linear Models

- Decision Trees
  - Car Recommendation System

- Examples:
  - ID3
  - C4.5
  - C5.0
  - CART



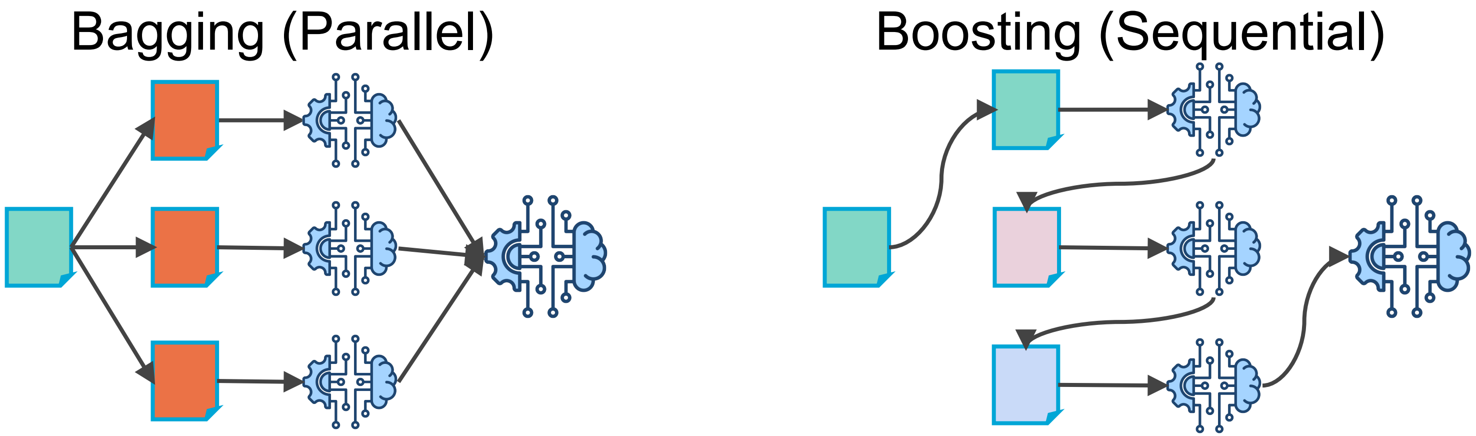
Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.

Non-Linear Models

- Ensemble Models

- **Majority Voting:** Weighted average of predictions of different models.
- **Bagging:** Average of multiple versions of the same model trained on random splits of the training data with replacement.
- **Boosting:** Sequential Learning from the same model where each model tries to fix the errors made by the previous models.

- Examples:
  - Random Forest
  - Xgboost
  - Gradient Boosting



Non-Linear Models

- Bayesian Machine Learning: Domain knowledge can be explicitly entered into learning in the form of the prior
- Based on Bayes Rule:

$$P(M \mid D) = \frac{P(D \mid M)P(M)}{P(D)}$$

  - **P(M):** How probable would we consider model M before seeing any data D (our inductive bias!)
  - **P(D|M):** model likelihood
  - **P(D):** normalizer, ensures P(M|D) adds up to 1
  - **P(M|D):** How probable is model M after seeing the data

- Generate Predictions:

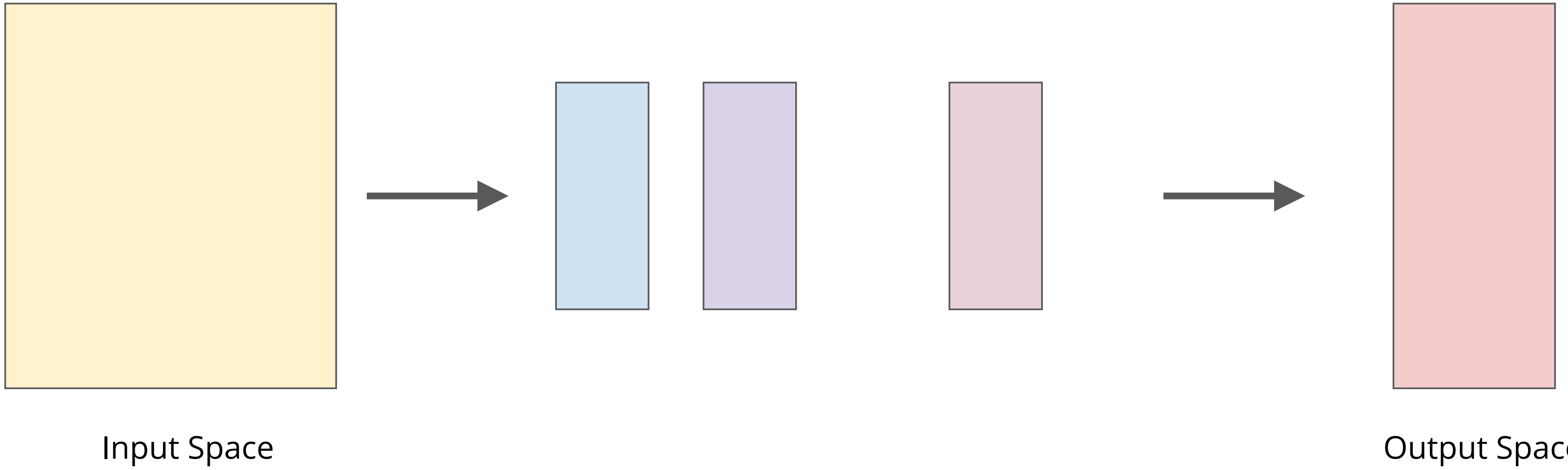
$$P(Y \mid D) = \sum_{M} P(Y \mid M)P(M \mid D)$$

- Examples:
  - Gaussian Processes
  - Bayesian Ridge

Non-Linear Models

- Artificial Neural Network Methods that learn multiple levels of representation to model the relation between input data and the output = **Automatic Feature Selection+Extraction+Modeling**.

- ANN is a collection of simple trainable mathematical units that collaborate to compute a complicated function

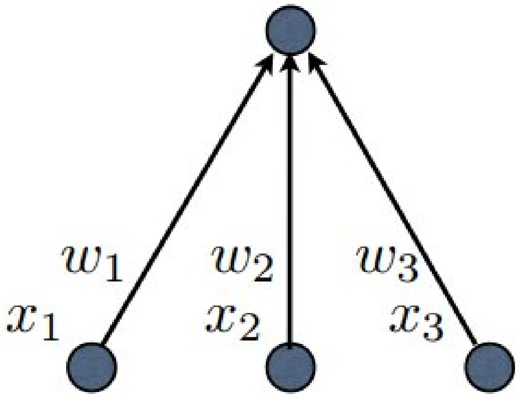


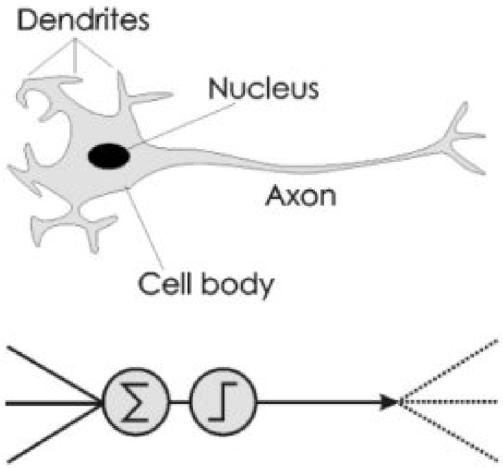
Non-Linear Models

- Linear part = sum [weights (Ws) * inputs (Xs)]
- Non-Linear part = Activation Function = F ( )

$$y_i = F\left(\sum_i w_i x_i\right)$$

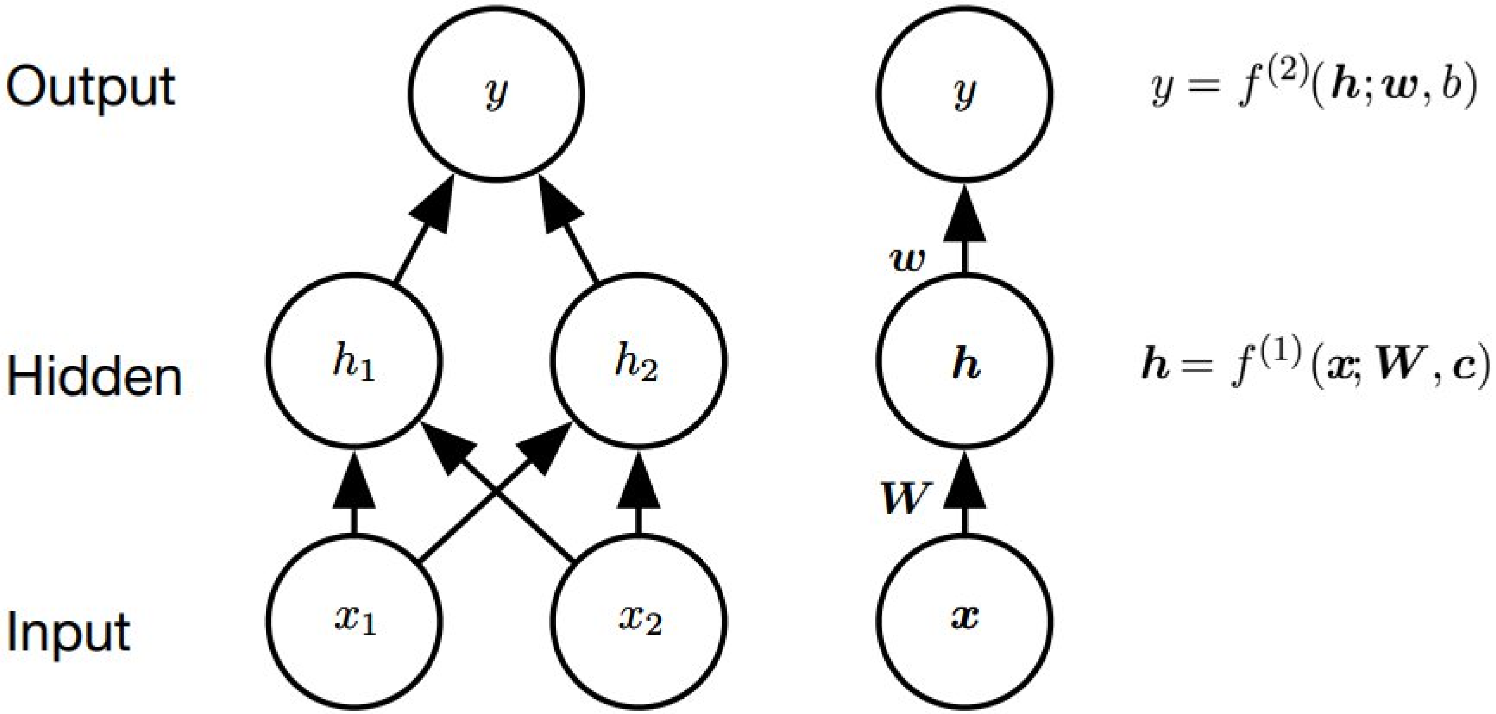
$$F(x) = \frac{1}{1 + \exp(-x)}$$





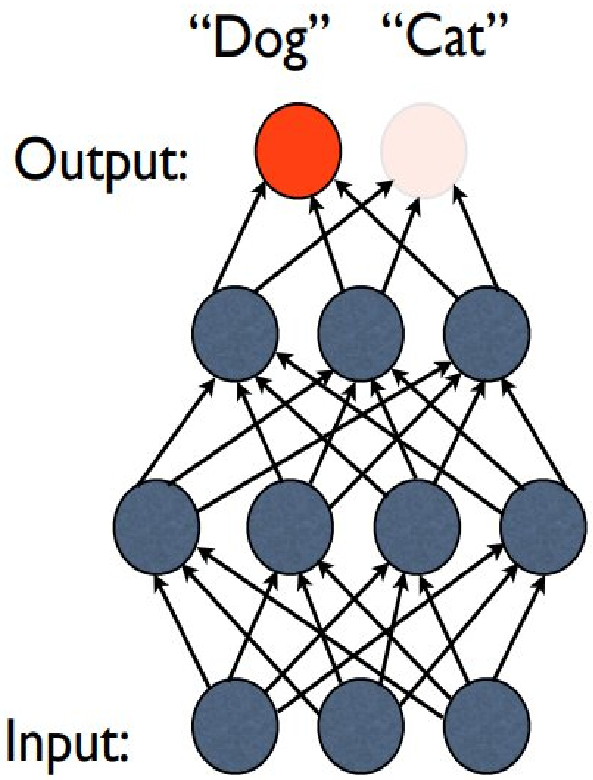
Norman, J. M. McCulloch & pitts publish the first mathematical model of a neural network.

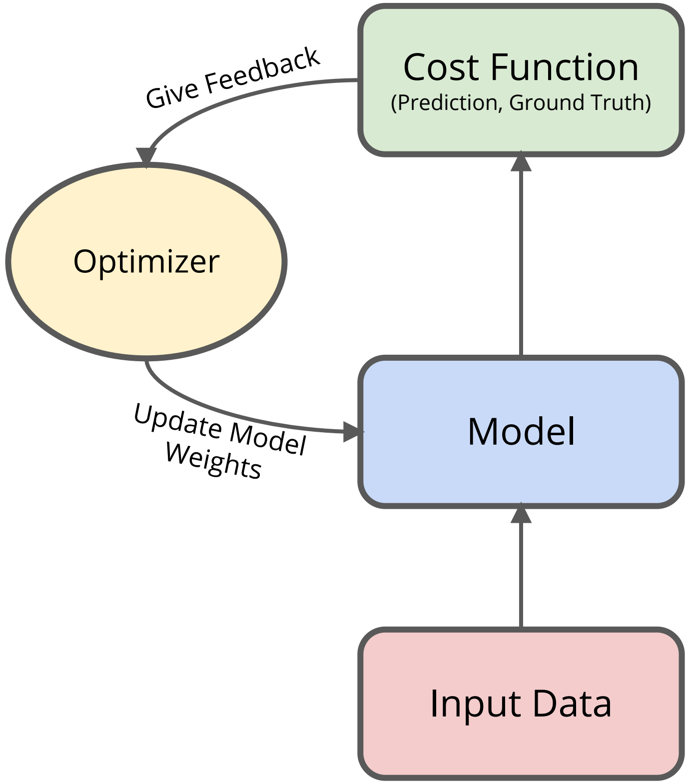
Non-Linear Models



$$f(\boldsymbol{x}; \boldsymbol{W}, \boldsymbol{c}, \boldsymbol{w}, b) = f^{(2)}(f^{(1)}(\boldsymbol{x}))$$

Non-Linear Models

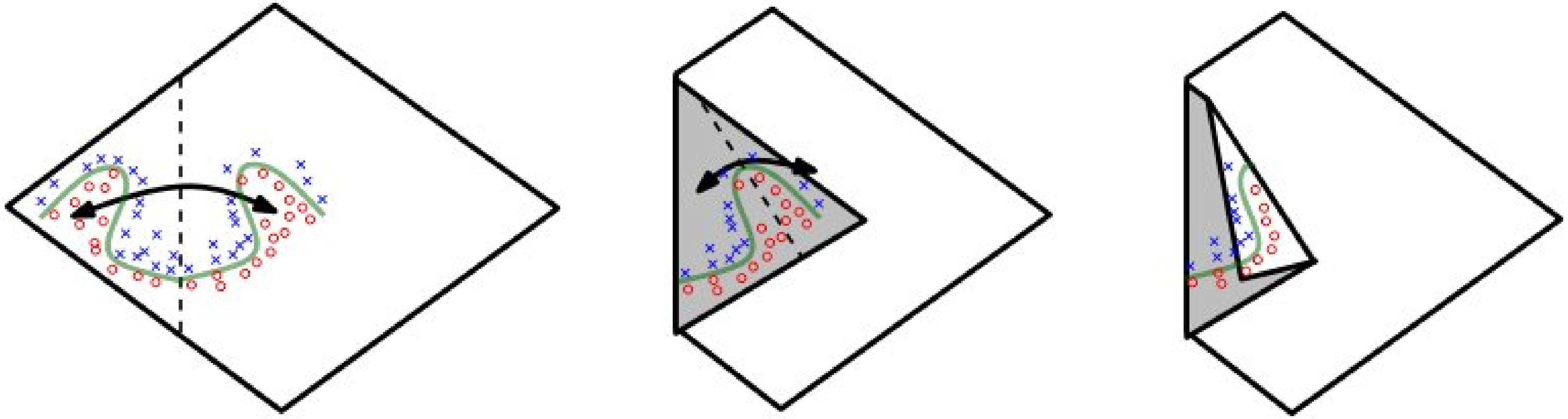




Non-Linear Models

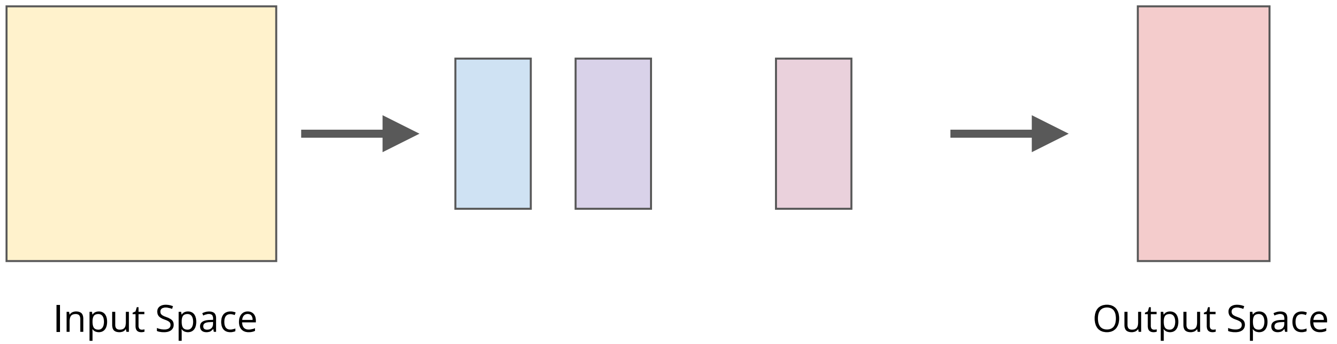
Universal approximation Theorem [WikiPedia]:
A feed-forward network with a ***single hidden layer*** containing a finite number of neurons can approximate continuous functions, under mild assumptions of the activation functions.

So, Why DEEP?

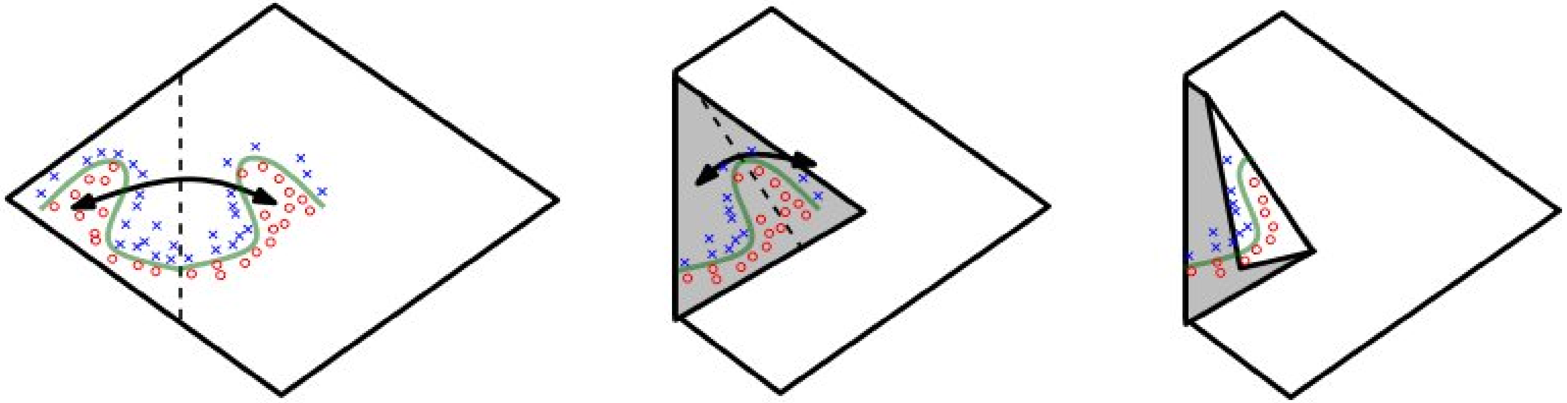


Heaton, J. (2018). Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. *Genetic programming and evolvable machines*, *19*(1), 305-307.

Non-Linear Models

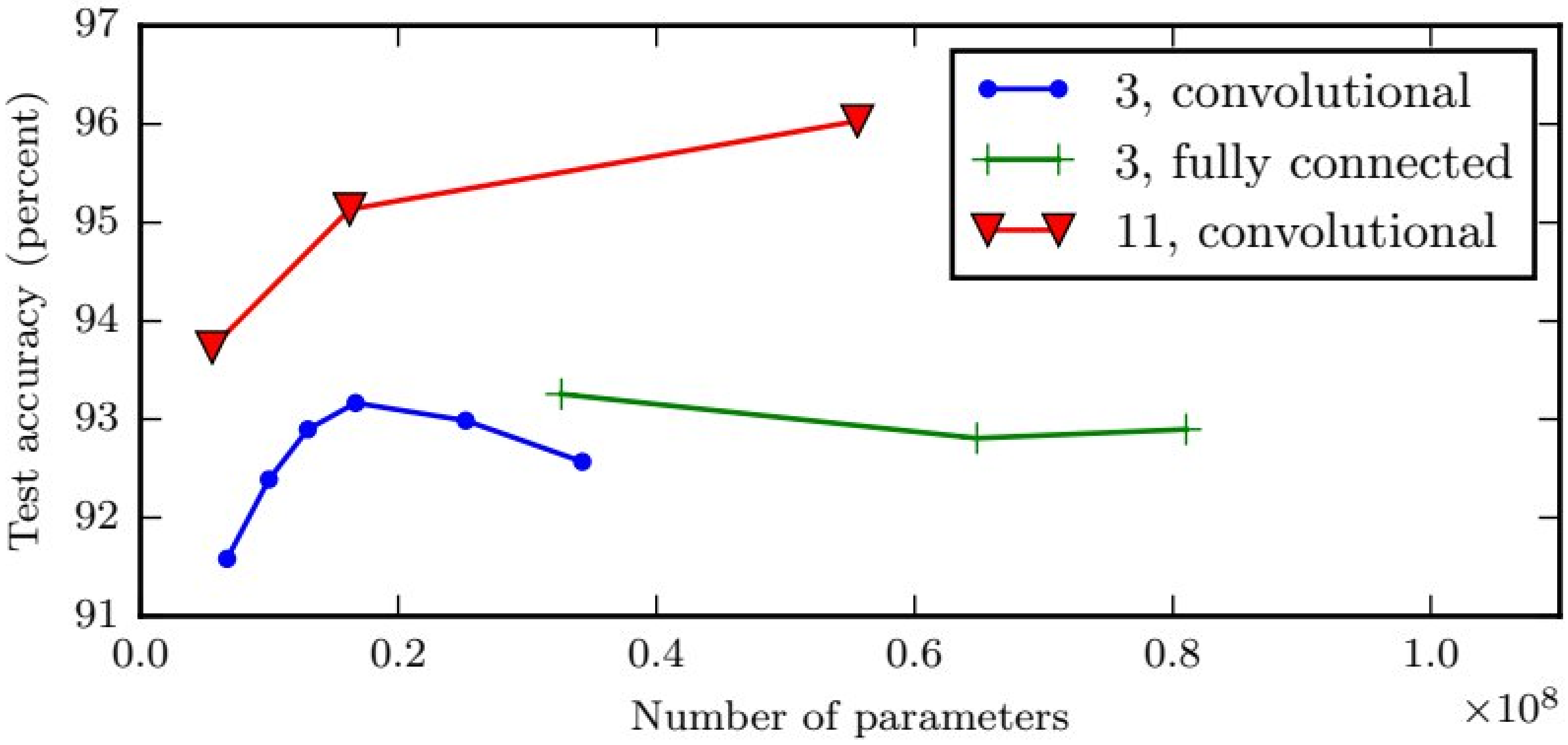


So, Why DEEP?



Heaton, J. (2018). Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. *Genetic programming and evolvable machines*, *19*(1), 305-307.

Non-Linear Models



Heaton, J. (2018). Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. *Genetic programming and evolvable machines*, *19*(1), 305-307.

Deep Learning: When to use ML? | When to use DL?

Machine Learning

- Relatively Small Volume of Data.
- Domain experience or knowledge how to do feature engineering.
- Better Interpretability / Explainability
- Faster execution

Deep Learning

- Large Volume of Data. Otherwise, highly susceptible to overfitting.
- Difficult to do feature engineering
- Complex Relation between inputs and outputs
  - Computer Vision Tasks
  - NLP Tasks

Types of Deep Learning Architectures

- Feed Forward Neural Networks
  - Suitable for tabular data

- Convolutional Neural Networks
  - Suitable for computer vision tasks
    - Image classification
    - Image Segmentation

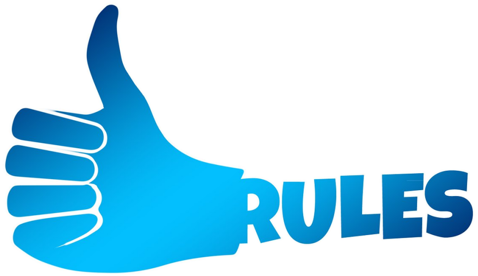
- Recurrent Neural Networks
  - Suitable for sequential data
    - Time-Series Data
    - Machine Translation

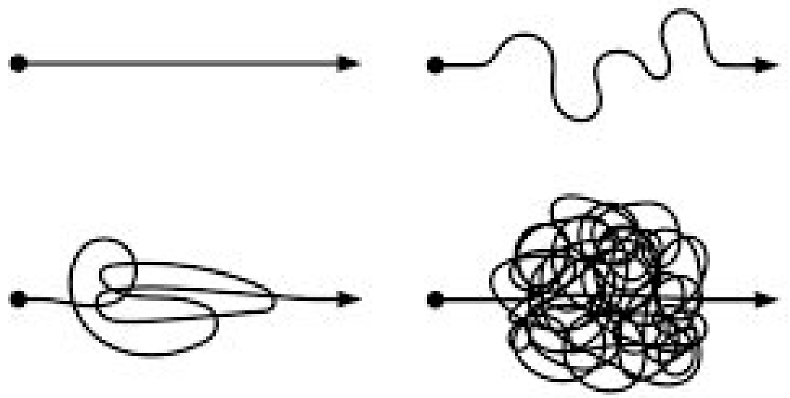
- Attention-based Neural Networks (Transformers)
  - Boosting performance of NLP tasks

- Graph Neural Networks
  - Suitable for graph data structures (E.g. Traffic Flow Networks)

Linear Vs Non-Linear Models

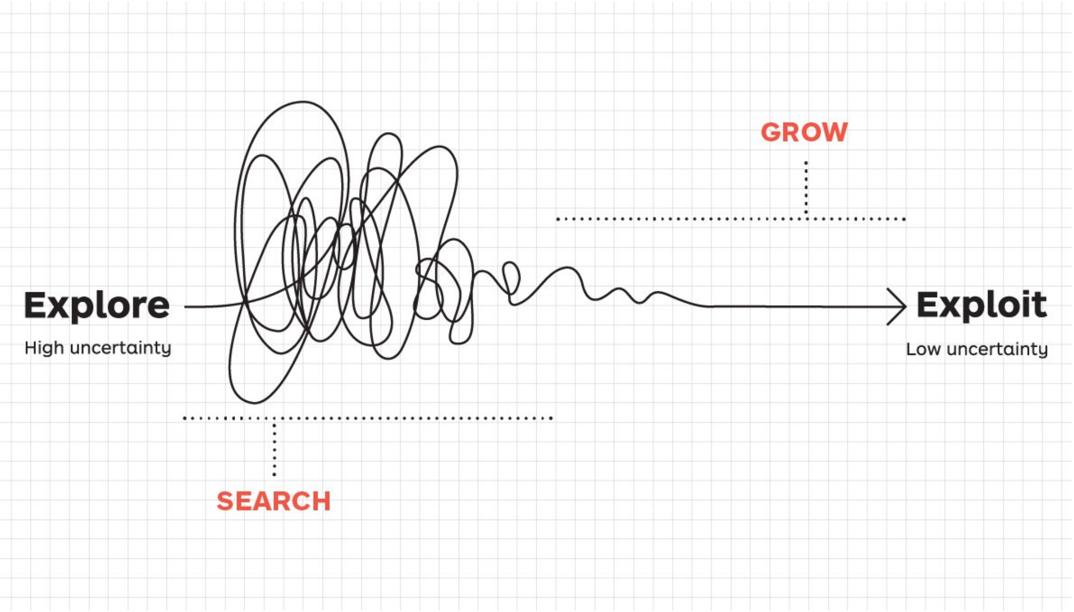
- Simpler models are more able to generalize well on new data.
- If linear models have good performance, Why to go for complex models?!
- Rule of Thumb:
  - If SVM = 98%,
  - If Deep Neural Network = 98.2%
  - **Use the SVM**





How to choose your model?

- Start with a single model out of each family. **(Explore)**
- Get recommendations from AutoML tools. **(Explore)**
- Start focusing and tuning on the best performing families. **(Exploit)**



Recap this lecture

After successfully completing this lecture, you are able to….

- Identify the different types of machine learning models.
- Understand the distinction of Deep Learning from Machine Learning.

Outlook: What will the tutorial be about?

- we'll explore how different classification models—from linear ones like Logistic Regression to non-linear ones like Decision Trees—learn to separate data points.
- You'll see how linear models draw straight boundary lines (hyper-planes),
- while non-linear models create more complex, flexible boundaries.
- By the end, you'll understand the strengths of each approach and know when to use linear vs. non-linear classifiers to solve your classification problems!

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.