Machine Learning
For Natural Language Processing

Abdelhak Mahmoudi
abdelhak.mahmoudi@um5.ac.ma

2020

Content

1. Introduction

2. Machine Learning

   1. Supervised Learning, 2. Unsupervised Learning

3. Natural Language Processing

   1. Regular Expressions, 2. Tokenization, 3. Character Encoding, 4. Part-of-Speech Tagging, 5. Chunking, 6. Stemming and Lemmatization, 7. Parsing, 8. Named Entity Recognition, 9. Topic Segmentation

4. Introduction to Deep Learning for NLP

   1. Sequence models, 2. Embeddings, 3. BERT models

Supervised Learning

- Linear Regression

- Logistic Regression

- Support Vector Machines

- Trees (Decision and Regression)

- Random Forests

- Boosting

- **Artificial Neural Networks**

History

- Logical Neuron (1943 - Warren McCulloch and Walter Pitts)
  - Logical operations
  - No activation function

- Linear Threshold Unit
  - real numbers, weight,
  - step activation function

- Perceptron (1957 - Rosenblatt)
  - single layer of LTU
  - Trained with Hebb's rule
  - Weakness of perceptron (1969 - Marvin Minsky and Seymour Papert)

- Multi Layer Perceptron (MLP) (1986 – D. Rumelhart, G. Hinton, R. Williams)
  - Artificial Neural Network
  - Sigmoid activation function
  - Backpropagation

- Deep Neural Networks
  - More data, more power
  - More layers
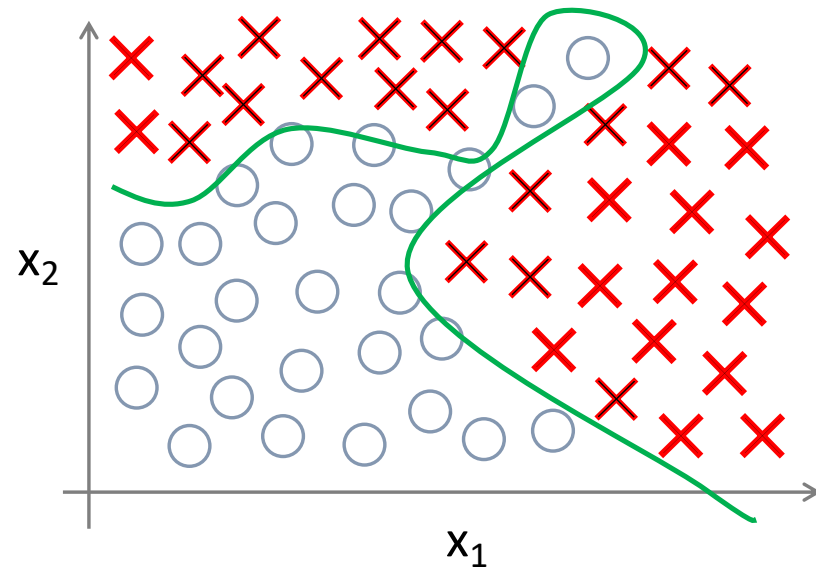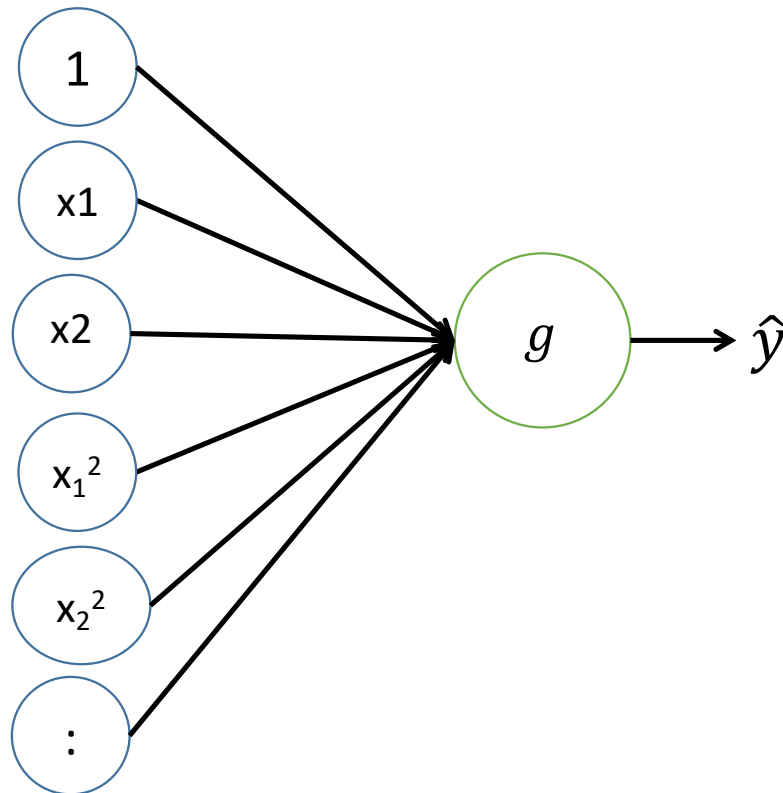 - Different activation functions
  - Different Architectures

Classical ML Models
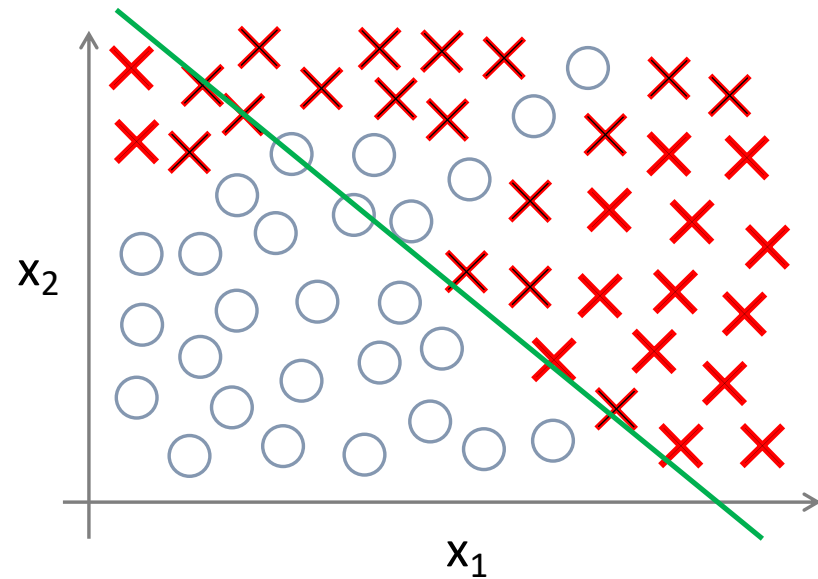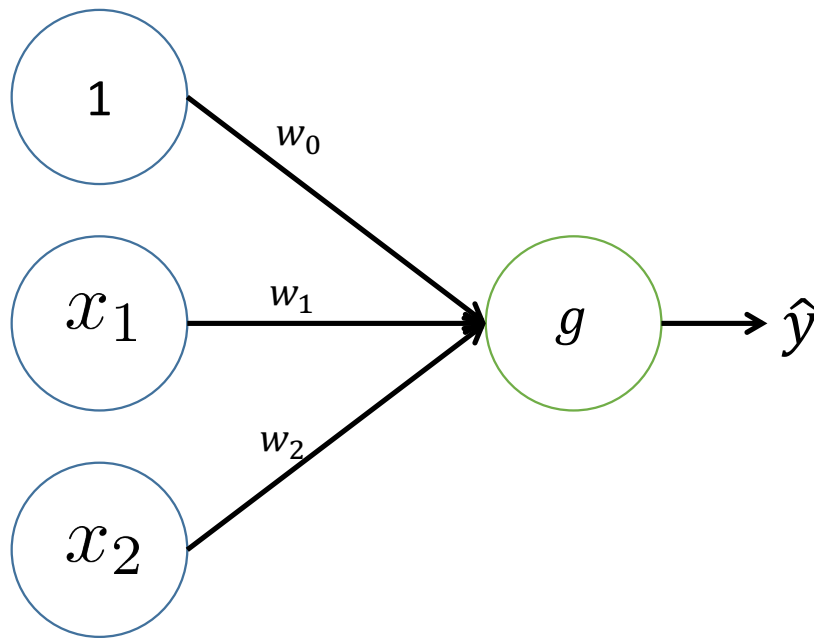
$$g(w_0 + w_1 x_1 + w_2 x_2)$$

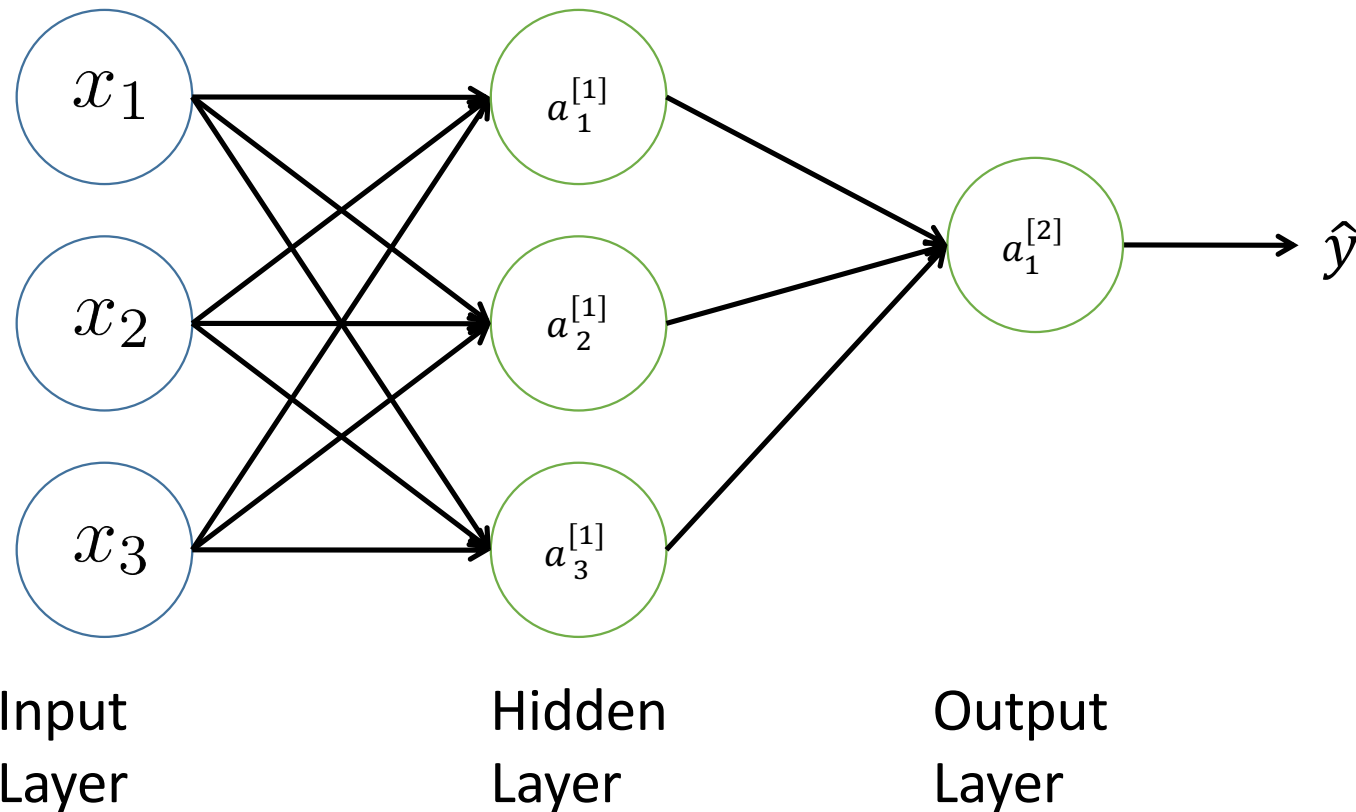A linear model will certainly underfit !

Classical ML Models

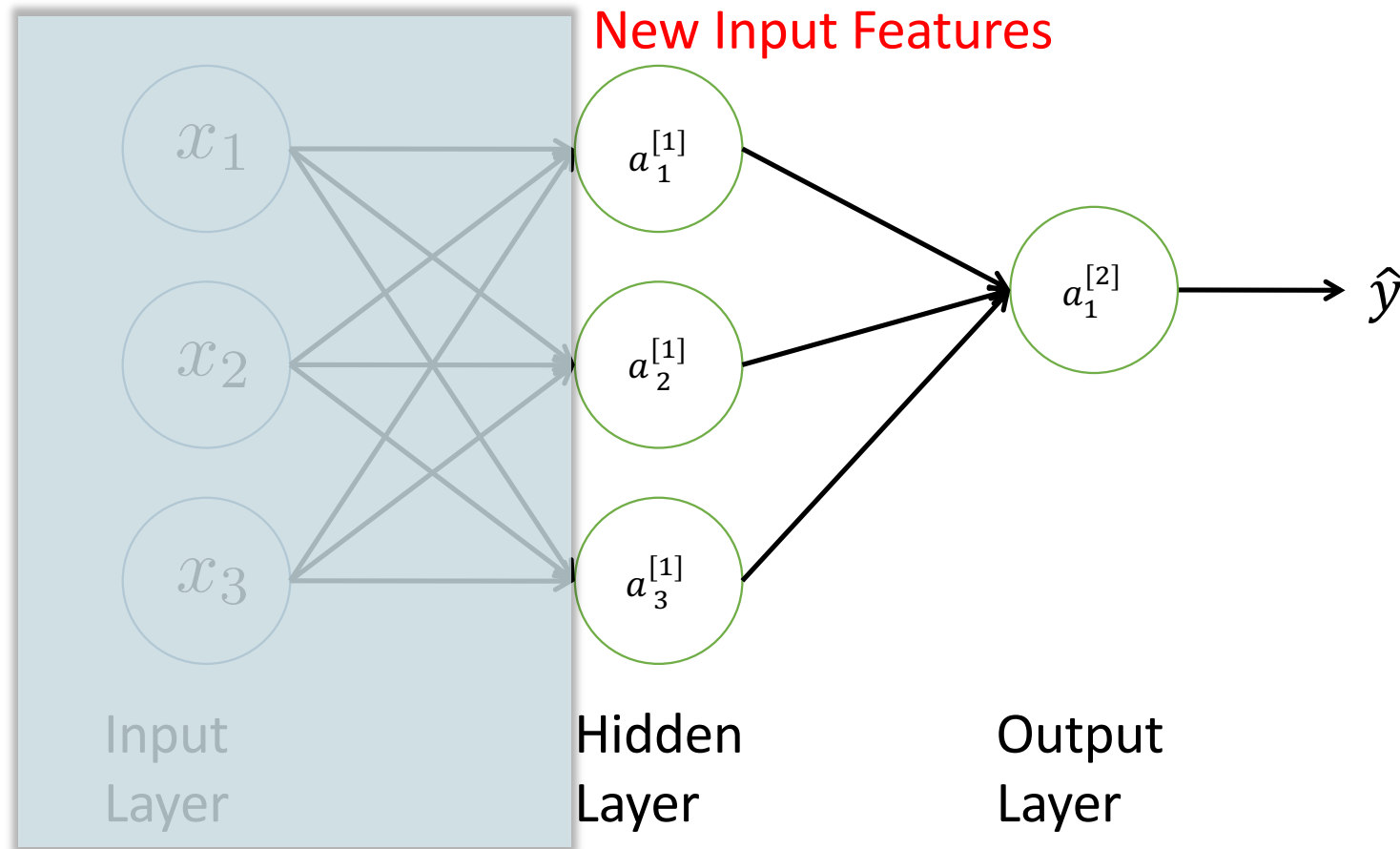$$g(w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_{12} x_1^2 + w_{13} x_2^2 \ldots + w_{25} x_1^d + w_{26} x_2^d)$$

… what if we have many many features?!

ANN Representation: Learning its own features

Input Features
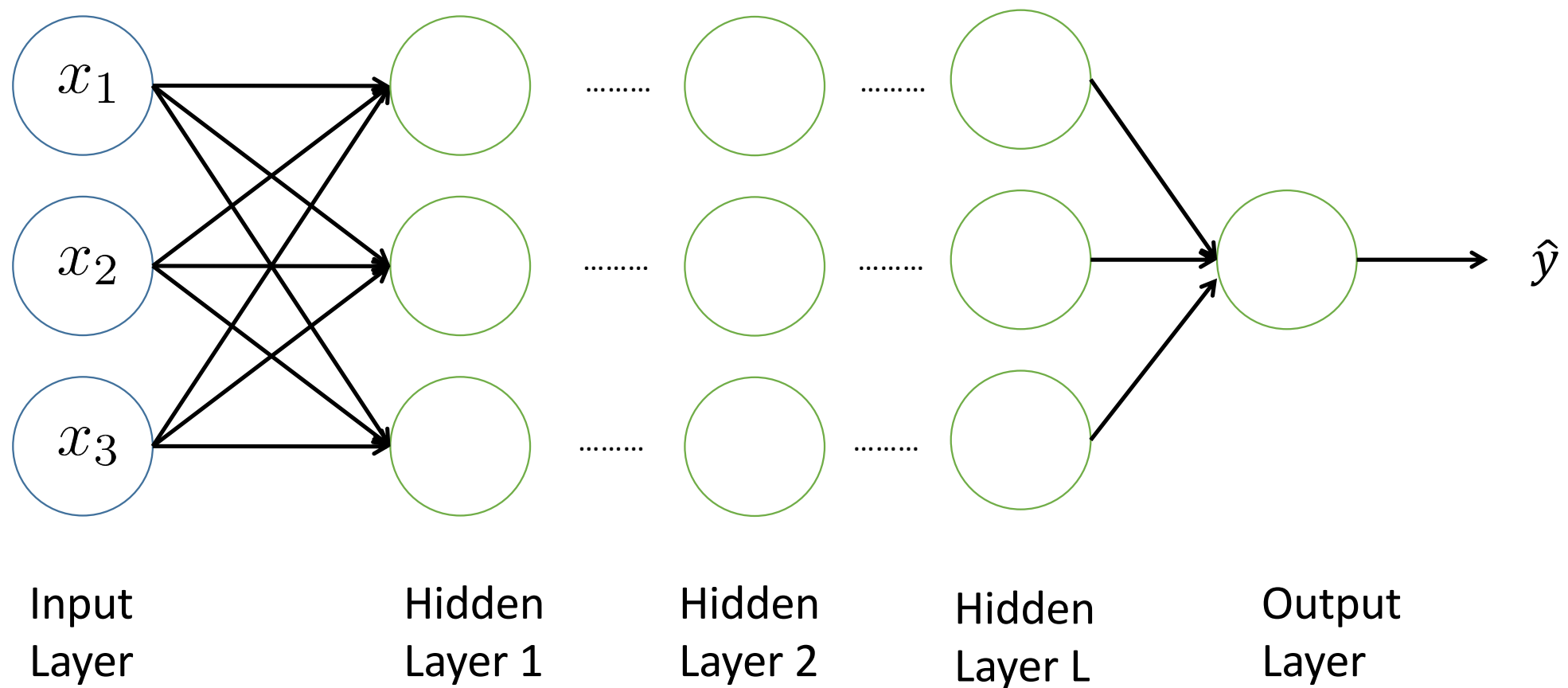

Input Layer   Hidden Layer   Output Layer

ANN Representation: Learning its own features

New Input Features

x_1

x_2

x_3

$a^{[1]}_1$

$a^{[1]}_2$

$a^{[1]}_3$

$a^{[2]}_1$

$\hat{y}$

Input Layer

Hidden Layer

Output Layer

ANN Representation: Deep NN

Input Layer — Hidden Layer 1 — Hidden Layer 2 — Hidden Layer L — Output Layer

x_1, x_2, x_3 → ... → $\hat{y}$

NLP Applications

- Sentiment classification
- Machine translation
- Name entity recognition
- Spam Detection
- Customer Queries Tagging
- Topics Classification
- Speech recognition
- Etc.