Machine Learning Foundations

Week-1 Revision

Arun Prakash A

Applications

Weather prediction

Automobiles and Robotics: Autonomous car

Chat bots , voice assistants : Alexa

Gaming : Alpha go

Recommendation : Amazon

ML or Not to ML:

Rules well defined,
known or not

Data

Data
x1 = [3,5,4]
x2 = [3,4,5]
x3 = [4,2,1]
x4 = [6,7,8]
x5 = [1,2,3]
x6 = [1,1,1]
x7 = [1,2,0]

Data	label
x1 = [3,5,4]	0
x2 = [3,4,5]	1
x3 = [4,2,1]	0
x4 = [6,7,8]	1
x5 = [1,2,3]	1
x6 = [1,1,1]	1
x7 = [1,2,0]	0

Terminology: Features (x_j^i), number of samples (n) , Labels (Ground truth) (y^i)

Features (x_j^i), Index starts from 1

number of samples (n=7),

Labels (Ground truth) (y^i), $y^2 = 1$

Train,Validation and Test Data

Total samples: 659

Train Set:

80% of total : 455

Validation Set:

20% of training :91

Test Set:

20% of total: 204

Types

Supervised (Data with labels)

Classification

Output: Discrete and Finite

Loss: 0-1 loss

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(f(x^i) \neq y^i)$$

Regression

Output: Continuous and infinite in general

Loss: MSE

$$\frac{1}{n}\sum_{i=1}^{n}(f(x^i) - y^i)^2$$

Unsupervised (Data without labels)

Dimensionality Reduction

Encoder, decoder (compressor or decompressor),

Loss (Reconstruction error)

$$\frac{1}{n}\sum_{i=1}^{n}||g(f(x^i)) - x^i||^2$$

$$x \in \mathbb{R}^d$$

$$f : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \quad \boxed{d' \ll d}$$

$$g : \mathbb{R}^{d'} \rightarrow \mathbb{R}^d$$

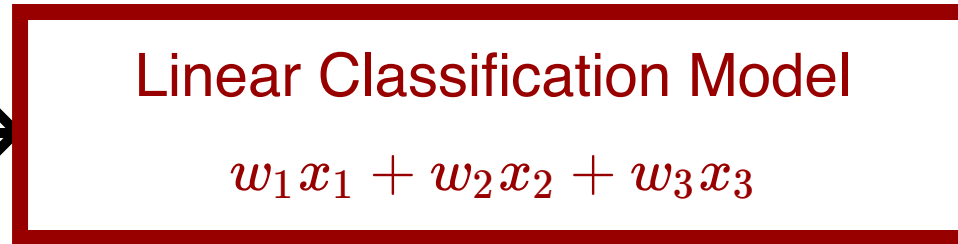### Density Estimation

Estimate PDF (Mean, variance),

Loss: Log-likelhood

$$\frac{1}{n}\sum_{i=1}^{n} -\log(P(x^i))$$

6

Training

w_0, w_1, w_2 are **parameters or weights** of the model. The **best values** for the parameters will be learned from the data

Data
x1 = [3,5,4]
x2 = [3,4,5]
x3 = [2,1,4]
x4 = [6,7,8]
x5 = [1,2,3]
x6 = [1,1,1]
x7 = [1,2,0]

Linear Classification Model

$$w_1 x_1 + w_2 x_2 + w_3 x_3$$

Label
0
1
0
1
1
1
0

Prediction:

Given a new sample, $x = [1, -1, 1]$, predict the output.

$$f(x) = 1x_1 + 0.5x_2 - 1x_3$$

Any Questions?