ECE ING4 MACHINE LEARNING

Jeremy Cohen

**Linear Regression**

Week 1 Review

Types of Machine Learning

**●** Supervised

**●** Output is a discrete number (0,1,2, ...), (SPAM/NOT SPAM), ...

**●** Supervised

**●** Output is a continuous number

**●** Unsupervised

**●** Unsupervised

**●** Outputs are clusters

**●** Outputs are clusters

**CLASSIFICATION REGRESSION CLUSTERING**

Regression

*Price*

label

feature

*Size of the*

**THE GOAL IS TO DETERMINE THE** *house*

**LINE OR CURVE THAT BEST FITS THE DATA**

Regression

**THE GOAL IS TO DETERMINE THE LINE OR CURVE THAT BEST FITS THE DATA**

230000

Freespace Regression Intro

in Practice

Regression in Practice

Outliers

**CLASSIFICATION REGRESSION**

*Price*

*Age*

*Size*

*Size of the house*

**OUTLIERS ARE TO BE REMOVED WHEN TRAINING**

Classification vs Regression

***TRAINING***

***PREDICTION***

Machine Learning Process

***DATA***

***NEW DATA PREDICTION TRAINED***

***MODEL***

***MACHINE LEARNING TRAINED***

***MODEL***

Training Linear Regression

*Price*

*Size of the house*

***DATA***

900

300

***MACHINE LEARNING TRAINED***

***MODEL***

50 75 100 125 150

*Price* I want to buy a house of 100 square meters. How much will it cost?

*Size of the house*

Testing Linear Regression

900

300 **915 k€**

***NEW DATA***

***PREDICTION***

50 75 100 125 150

***TRAINED MODEL***

3 Datasets

Validation Set Test Set

Training Set

~70% of the dataset ~20% of the dataset

~10% of the dataset

Used to **train the model**

Used to **test** the model and **generalize better to new data**

Used **only once** when both accuracies are good and ready for real-world

Used **only once** when both accuracies are good and ready for real-world

Train Test Process

Training Set

Test Set

Trained Model

Machine Learning Algorithm

TRAINING

TRAINING

Trained Model

TESTING

Prediction

Model

**House characteristics (size, neighborhood, ...): Feature X House Price : Label Y**

Simple Linear Regression

The Best Fit

Sum of Squared Errors (SSE)

The Best Fit

**The green line has the lowest SSE and the best fit!**

Mean Squared Error

**1/2m - simplifies the maths**

The Best Fit

The Best Fit

Gradient Descent

Notation - Simple Linear Regression

**X Size (m2) Y**

**Price (1000$)**

20 300

37 540

**m = 47** 88 986

... ...

m: Number of examples

x(i): Input of the i-th training example

**X**(2) **= [37]**

Parameter influence

Hypothesis : Θ0 = 0 hθ (x) =θ1x

hθ x for

Θ1 = 0.5

hθ x for Θ1 = 0

J(1) = 1/2m (02+02+02) = 0 J(0.5) = 1/2m ( (0.5-1)2+(1-2)2+(1.5-3)2 ) ~ 0.6

We want to minimize J(θ).

J(0) = 14/6 ~ 2.3

Parameter influence

J(Θ1)

hθ x for Θ1 = 1

Parameter influence

Θ0 ≠ 0

Gradient Descent

**● Start with random** Θ0 **and** Θ1

**● Change** Θ0 **and** Θ1 **to reduce J(**Θ0**,** Θ1**) until we find a minimum**

**Random Starting Point**

Gradient Descent

**For different initializations, we might have different results**

**Random Starting Point**

Local-Global minima

. . local minima

global minima

Gradient Descent

assignment

current value

learning

rate derivative

Gradient Descent

.

random starting point

optimal .

value

gradient / rate of change / slope / derivative. exemple : f’(x2) = 2x gradient for x = 2 is 4

gradient / rate of change / slope / derivative. exemple : f’(x2) = 2x gradient for x = 2 is 4

value

Gradient Descent

Learning Rate

GOOD LEARNING RATE BAD LEARNING RATE

● Not too low

● Allows the network to find the right parameters

● Too high

● Learns fast but quickly overshoots and ends up increasing the error

Gradient Descent Recap

Multivariate Linear Regression

Notation - Multivariate Linear Regression

**n=4**

**Size X**1 **Number of**

**Bedrooms X**2 **Floor Age (years) Price (1000$)**

**X**3 **X**4 **Y**

20 0 3 30 300

37 1 0 45 540

**m = 47** 88 3 4 60 986

... ...

m: Number of examples n: Number of features**X**(2) **= 37 1 0** x(i): Input (features) of the i-th training example **45**

xj(i): Value of feature j for the i-th training

**X**3(2) **= 0** example

Multivariate Linear Regression

Multivariate Linear Regression

For correct multiplication, we need ΘT X0 ΘT = [Θ0 Θ1 ... Θn] X1 ...Xn

=>1 X =

Multivariate Linear Regression

Gradient Descent

**Simple Linear Regression**

**Multivariate Linear Regression**

Polynomial Regression

Polynomial Regression

Polynomial Regression in Practice

Performance

Performance

**UNDERFITTING JUST RIGHT OVERFITTING**

ThankYou

jeremycohen.podia.com

https://www.linkedin.com/in/jeremycohen2626/