

MMA classifier

If it looks like a duck, swims like a duck, and quacks like a duck, then it probably *is* a duck.

Duck test

From Wikipedia, the free encyclopedia

Mixed Martial Arts = Judo + boxing + karate + kickboxing + greco-roman wrestling + jiu-jitsu + sambo...



CONOR MCGREGOR

"NOTORIOUS"



BORN: 1988-07-14
AGE: 29

DUBLIN
IRELAND 

HEIGHT: **5'8"** (172.72 CM)
WEIGHT: **155 LBS** (70.31 KG)
ASSOCIATION: **SBG IRELAND**
CLASS: **LIGHTWEIGHT**



FIGHT HISTORY - PRO

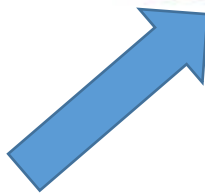
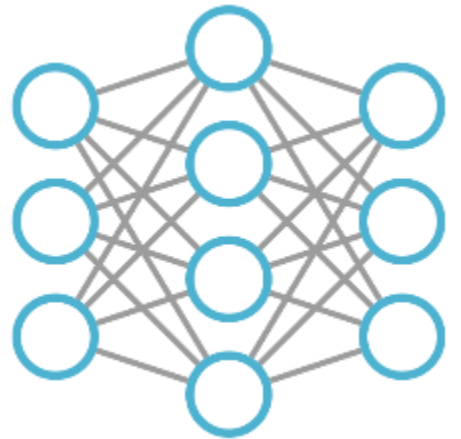
RESULT	FIGHTER	EVENT	METHOD/REFEREE	R	TIME
WIN	Eddie Alvarez	UFC 205 - Alvarez vs. McGregor Nov / 12 / 2016	TKO (Punches) John McCarthy	2	3:04
WIN	Nate Diaz	UFC 202 - Diaz vs. McGregor 2 Aug / 20 / 2016	Decision (Majority) John McCarthy	5	5:00
LOSS	Nate Diaz	UFC 196 - McGregor vs. Diaz Mar / 05 / 2016	Submission (Rear-Naked Choke) Herb Dean	2	4:12
WIN	Jose Aldo	UFC 194 - Aldo vs. McGregor Dec / 12 / 2015	KO (Punch) John McCarthy	1	0:13
WIN	Chad Mendes	UFC 189 - Mendes vs. McGregor Jul / 11 / 2015	TKO (Punches) Herb Dean	2	4:57

UFC Fight Night 59 -

Het idee:



UNIBET



Wat weet ik?

- Basale kennis Machine Learning: Masterclass van Arno Knobbe in 2012, Leiden Institute of Advanced Computer Science, UL
- Geen Python ervaring, wel 20 jaar programmeerervaring in o.a. Java, C, C#...

Doel

Rijk worden 😊

maar ook leren...

Python

API, JSON

Toepassing ML

Python libraries

Web scraper

BeautifulSoup

ML classifier

Scikit-learn

Automatic betting

Urllib (Betfair Exchange API via Jason-RPC)

```
- D:\users\igor\.spyder-py3\mma_webscraper.py
gubbins_ng.py x settings.py x api_ng.py x getting_started.py x sitecust
0     fighter_weight=""
1     fighter_age = ""
2     fighter_height = ""
3
4     #raise
5     try:
6
7         req1 = urllib.request.Request( url2, headers={'User-Agent': 'Mozilla/5.0'})
8         response1 = urllib.request.urlopen(req1)
9         page1 = response1.read()
0         soup1 = BeautifulSoup(page1)
1         strong=soup1.find_all("strong" )
2         for child in strong:
3             #print(child)
4             if child.text.find('lbs') != -1:
5                 #print(child.string)
6                 fighter_weight = child
7             if child.text.find('AGE') != -1:
8                 #print(str(child))
9                 fighter_age = child
0             if child.text.find('') != -1:
1                 #print(str(child))
2                 fighter_height = child
3             #print(fighter_weight)
4
5         con = db.connect('DRIVER={ODBC Driver 13 for SQL Server}')
6
7         cur_fighter = con.cursor()
8         qry_fighter = '''USE [mma_magic] INSERT INTO [dbo].[mma_
9             ([fighter_name]
0             , [fighter_weight]
1             , [fighter_age]
2             )
3             VALUES (? ,?,?) '''
4         try:
5             param_values_fighter = [ str(fighter_name),str(fighte
6             #print("inside database insert!")
7             cur_fighter.execute(qry_fighter, param_values_fighte
8             cur_fighter.commit()
9             #print('{0} row inserted fighter successfully.'.form
0         except:
1             e = sys.exc_info()[0]
2             print("error in insert fighter!" )
```

Webscraping

http://www.sherdog.com/robots.txt

```
def get_mma_stats(url, fighter_name):  
    print(url)  
    user_agent = 'Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.7) Gecko/2009021910 Firefox/3.0.7'  
    #url = 'http://www.sherdog.com/stats/fightfinder?SearchTxt=&weight=6&association='  
  
    try:  
  
        req1 = urllib.request.Request(url, headers={'User-Agent':user_agent,})  
        response1 = urllib.request.urlopen(req1)  
        page1 = response1.read()  
        soup1 = BeautifulSoup(page1)  
  
        all_links = soup1.find_all('a')
```

FIGHT FINDER

ronda rousey

FILTER SEARCH BY WEIGHT CLASS: Select Weight Class

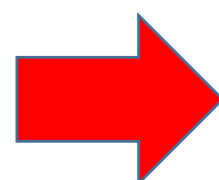
FILTER SEARCH BY ASSOCIATION: Filter by Association

TOP TRENDING FIGHTERS: Conor McGregor | Khabib Nurmagomedov | Fedor Emelianenko | Tony Ferguson | Francis Ngannou | Anderson Silva

OR SELECT AN ORGANIZATION: Search by organization

FIGHTER RESULTS LIST

FIGHTER	NICKNAME	HEIGHT	WEIGHT	ASSOCIATION
	Ronda Rousey "Rowdy"	5'6" (1.68 m)	135 lbs (61.23 kg)	Team Hayastan / Glendale Fighting Club



RONDA ROUSEY
"ROWDY"

BORN: 1987-02-01
AGE: 31
SANTA MONICA, CALIFORNIA
UNITED STATES

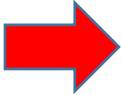
HEIGHT: 5'6" (167.64 CM)
WEIGHT: 135 LBS (61.23 KG)
ASSOCIATION: TEAM HAYASTAN / GLENDALE FIGHTING CLUB
CLASS: BANTAMWEIGHT

WINS: 12 (3 KO/TKO (25%), 8 SUBMISSIONS (75%), 0 DECISIONS (0%))

LOSSES: 2 (2 KO/TKO (100%), 0 SUBMISSIONS (0%), 0 DECISIONS (0%))

FIGHT HISTORY - PRO

RESULT	FIGHTER	EVENT	METHOD/REFEREE	R	TIME
LOSS	Amanda Nunes	UFC 207 - Nunes vs. Rousey Dec / 30 / 2016	TKO (Punches) Herb Dean	1	0:48
LOSS	Holly Holm	UFC 193 - Rousey vs. Holm	KO (Head Kick and Punches)	2	0:59



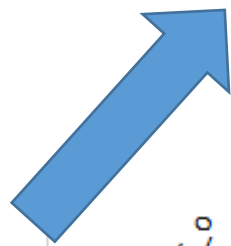
ML: Data, data, data

- Feature selection
- Imputation
- Normalization

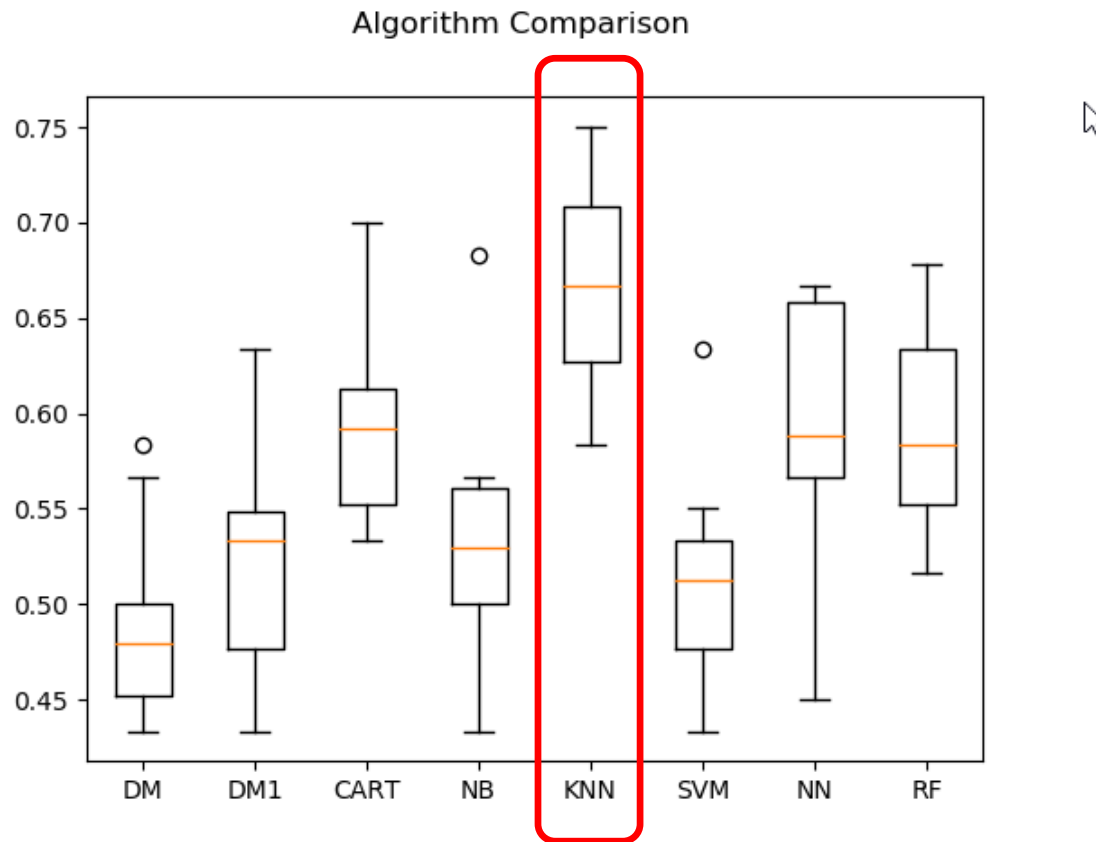
```

In [6]: print(rescaledX[0:5,:])
[[0.         0.         0.22222222 0.1         0.13333333 0.22222222
  [0.         0.3         0.22222222 0.55        0.36666667 0.16666667]
 [0.41666667 0.2         0.         0.4         0.36666667 0.
  [0.66666667 0.2         0.66666667 0.4         1.         0.11111111]
 [0.25        0.3         0.22222222 0.2         0.1         0.44444444
  [0.16666667 0.3         0.22222222 0.4         0.2         0.27777778]
 [0.16666667 0.3         0.22222222 0.4         0.2         0.27777778
  [0.25        0.3         0.22222222 0.2         0.1         0.44444444]
 [0.66666667 0.2         0.22222222 0.5         0.13333333 0.77777778
  [0.41666667 0.1         0.22222222 0.55        0.16666667 0.5         ]]
```

LOSSES_DECISIONS	LOSSES_KOTKO	LOSSES_SUBMISSIONS	WINS_DECISIONS	WINS_KOTKO	WINS_SUBMISSIONS	LOSSES_DECISIONS_o	LOSSES_KOTKO_o	LOSSES_SUBMISSIONS_o	WINS_DECISIONS_o	WINS_KOTKO_o	WINS_SUBMISSIONS_o	win_loss
0	0	2	2	4	4	0	3	2	11	11	3	0
5	2	0	8	11	0	8	2	6	8	30	2	0
3	3	2	4	3	8	2	3	2	8	6	5	1
2	3	2	8	6	5	3	3	2	4	3	8	1
8	2	2	10	4	14	5	1	2	11	5	9	0
0	0	1	6	1	5	4	2	2	7	5	12	1
4	2	2	7	5	12	0	0	1	6	1	5	0
2	2	0	2	5	3	0	3	1	2	4	3	1
2	1	3	7	25	12	0	3	0	4	6	4	1
2	1	0	7	8	4	1	0	0	1	4	6	1



Accuracy per classifier, 10 splits Cross validation



Dummy (uniform)

Dummy (most frequent)

DecisionTree

Naive Bayes

K-Nearest Neighbor

Neural Network

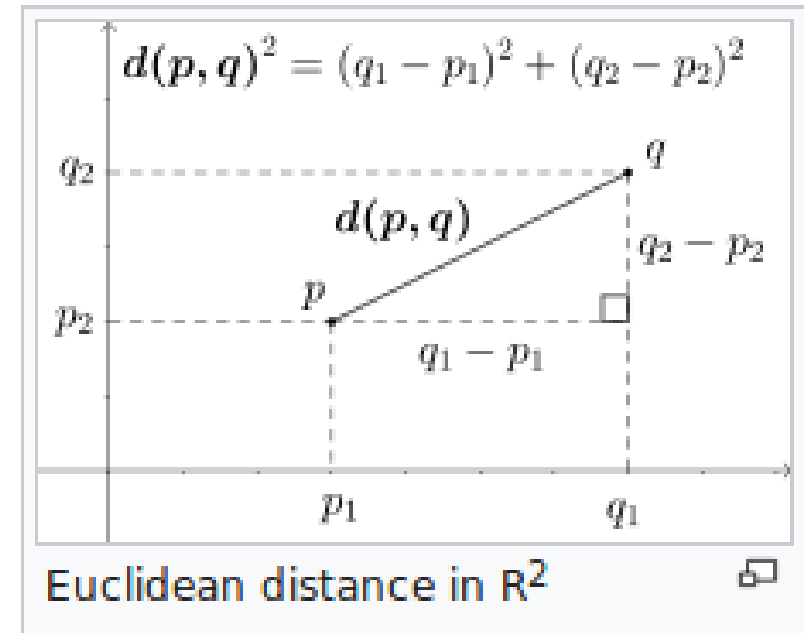
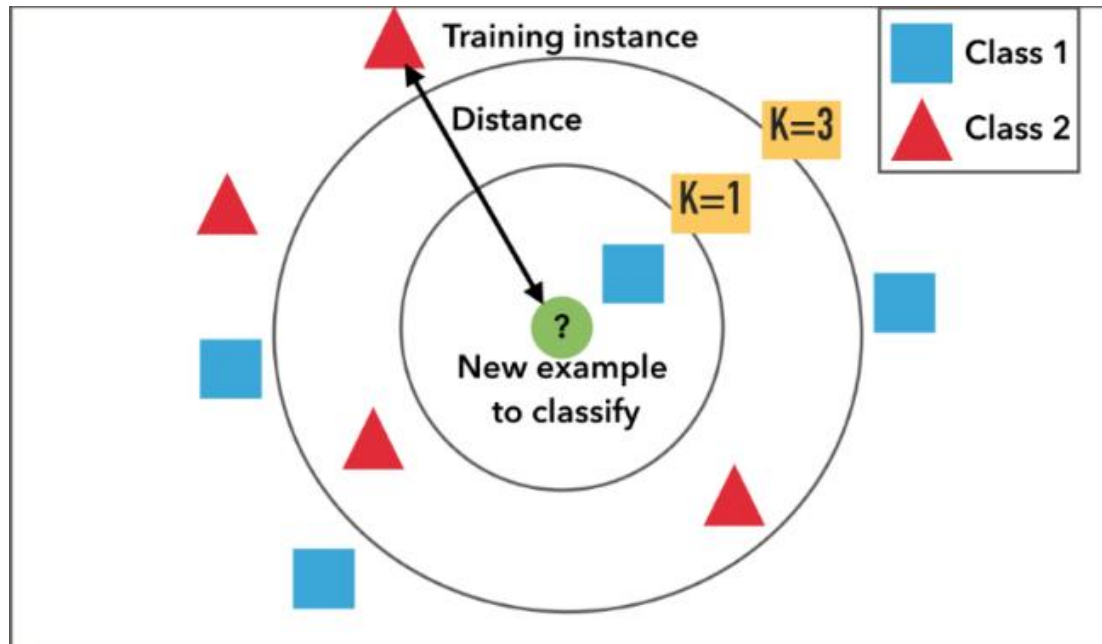
Support Vector Machine

Random Forest

K Nearest Neighbors

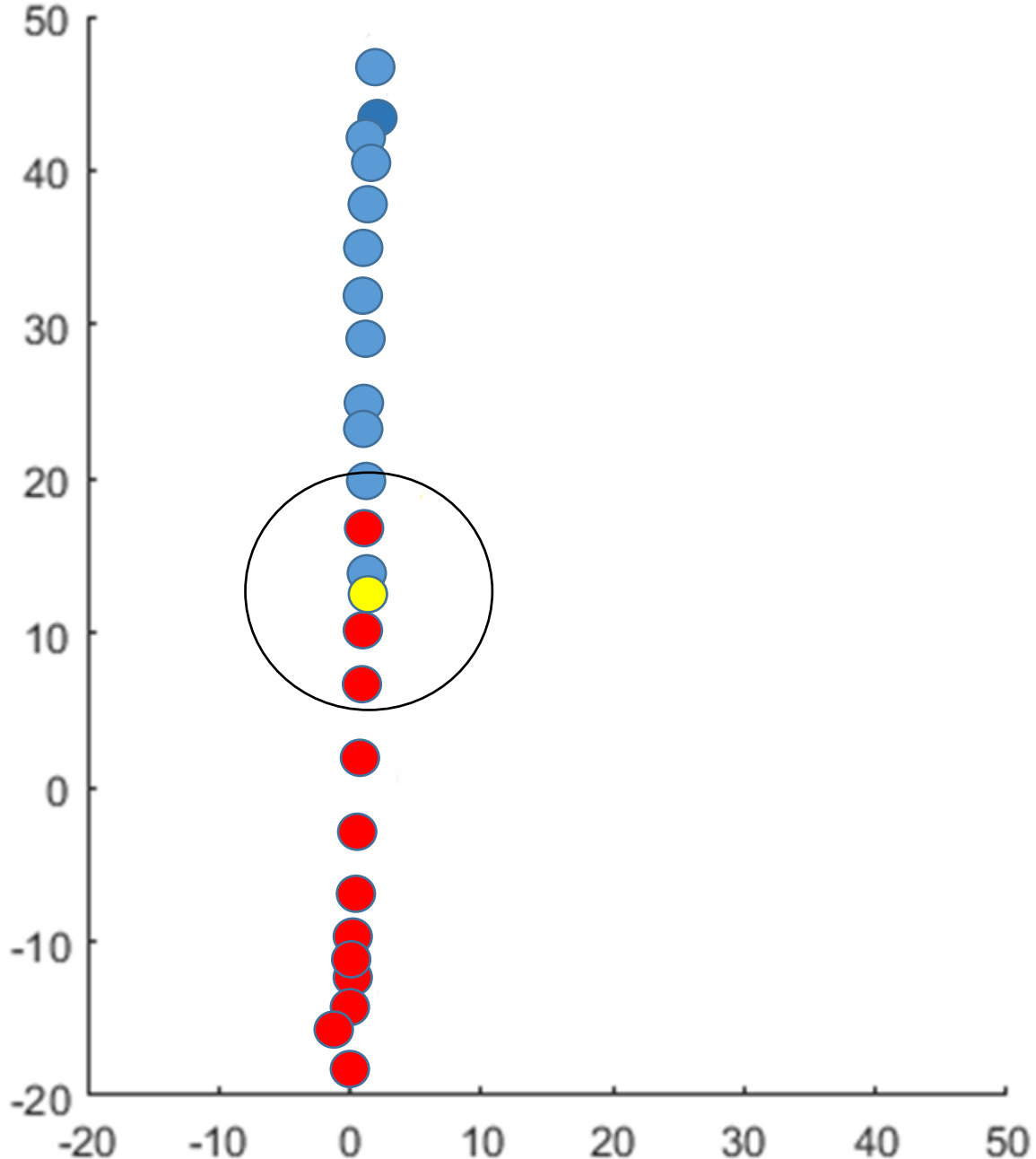
“Simple and lazy learner”

generaliseert (leert) niet, geen echte trainingsfase
heeft de hele training set nodig tijdens testing

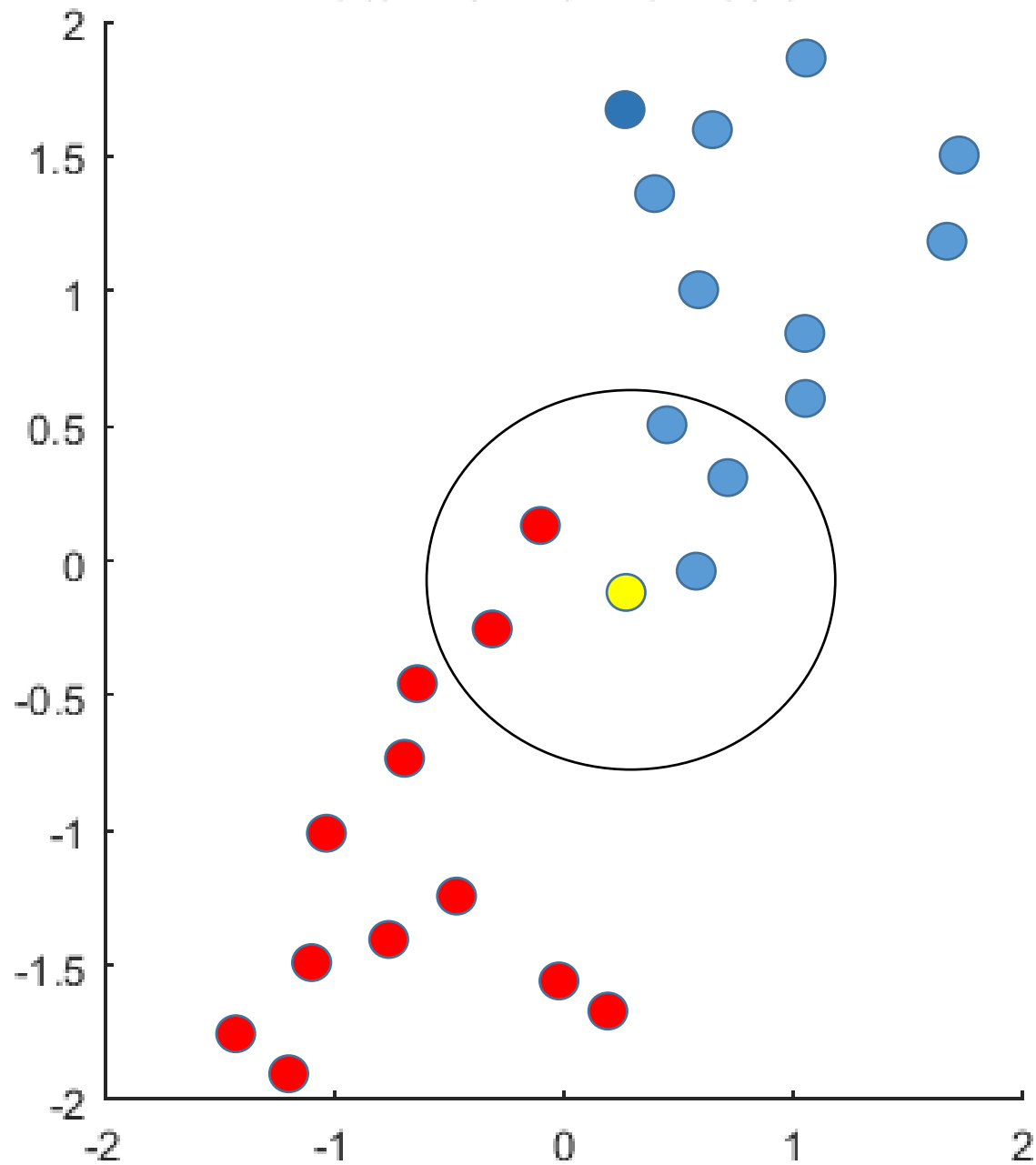


$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$
$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Data without normalization

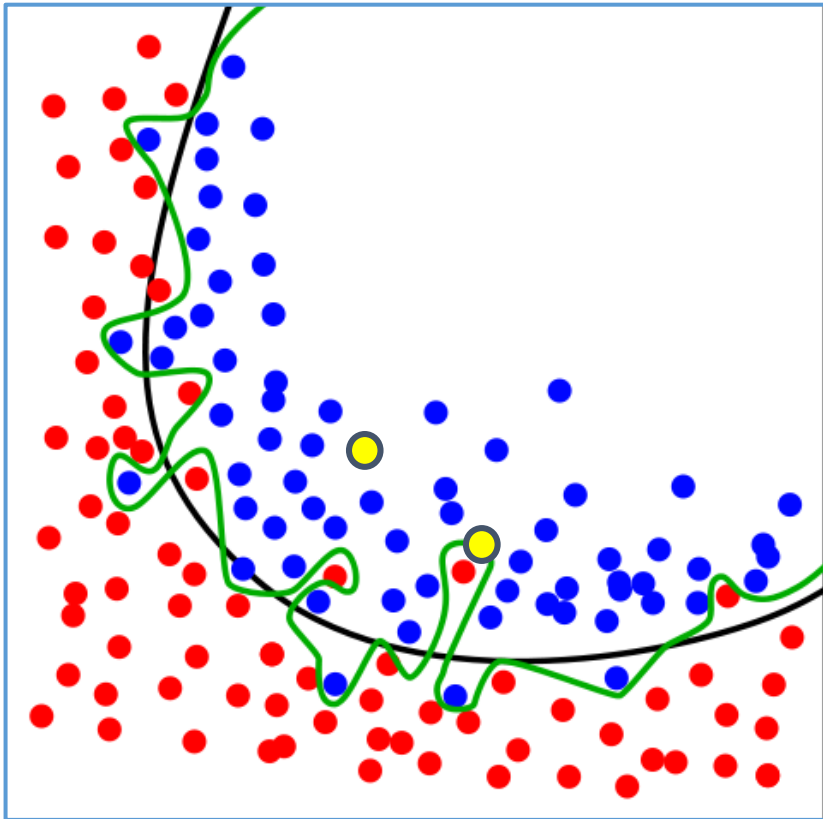


Data with normalization

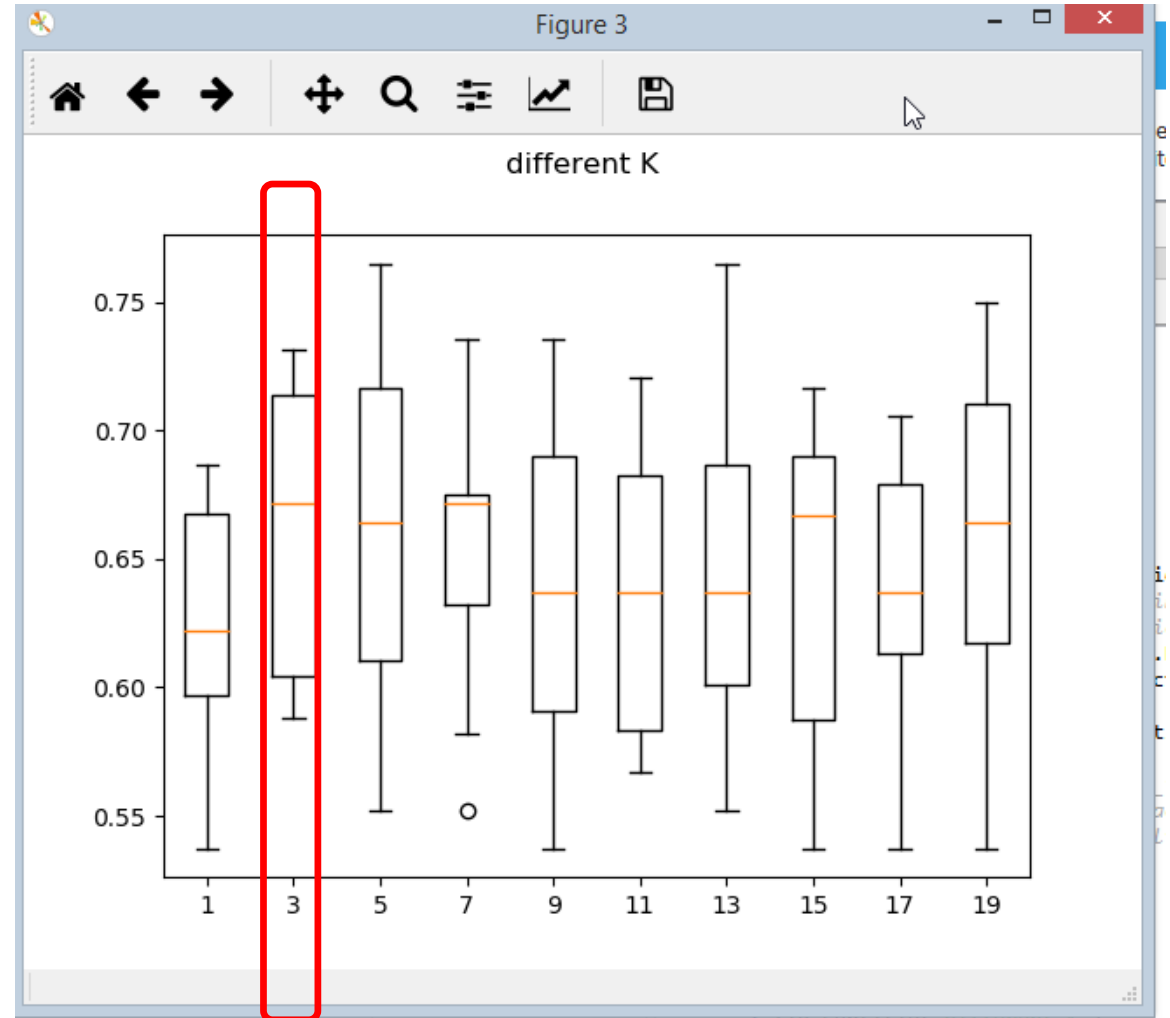


K - te klein: overfitting
K - te groot ?

weights = 'distance'



Cross validation voor KNN
K: 1..19



Confusion matrix

Accuracy: **0.60**

		Predicted			
Actual		0	1		
0		TN: 59	FP: 33		
1		FN: 41	TP: 54		
	`precision	recall	f1-score	support	
0	0.59	0.64	0.61	92	
1	0.62	0.57	0.59	95	
avg /					
total	0.61	0.60	0.60	187	

Metric	Formula
True positive rate, recall	$\frac{TP}{TP+FN}$
False positive rate	$\frac{FP}{FP+TN}$
Precision	$\frac{TP}{TP+FP}$
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
F-measure	$\frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$

Accuracy

It is the number of correct predictions made divided by the total number of predictions made

Precision

the number of positive predictions divided by the total number of positive class values predicted. It is also called the Positive Predictive Value (PPV).

Precision can be thought of as a measure of a classifiers exactness. A low precision can also indicate a large number of False Positives.

Recall

Recall is the number of True Positives divided by the number of True Positives and the number of False Negatives. Put another way it is the number of positive predictions divided by the number of positive class values in the test data. It is also called Sensitivity or the True Positive Rate.

Recall can be thought of as a measure of a classifiers completeness. A low recall indicates many False Negatives.

F1 Score

The F1 Score is the $2 \cdot ((\text{precision} \cdot \text{recall}) / (\text{precision} + \text{recall}))$. It is also called the F Score or the F Measure. Put another way, the F1 score conveys the balance between the precision and the recall.

Real data...



fighter_name	LOSSES_DECISIONS	LOSSES_KOTKO	LOSSES_SUBMISSION	WINS_DECISIONS	WINS_KOTKO	WINS_SUBMISSIONS
Jason Knight	3	1	0	4	3	13
fighter_name_o	LOSSES_DECISIONS_o	LOSSES_KOTKO_o	LOSSES_SUBMISSION_o	WINS_DECISIONS_o	WINS_KOTKO_o	WINS_SUBMISSIONS_o
Makwan Amirkhani	2	0	1	3	1	9

- knn = KNeighborsClassifier(n_neighbors=3, weights = 'distance')
- knn.fit(rescaledX, Y)
- predictions = knn.predict(X_real)
- print(predictions)
- [0] ←

MMA Classifier: Amirkhani wins !

Home

Live

My Bets

8 Jun 2018



POPULAR

ATP

International Friendly Matches

MLB

NBA

NHL

WTA

French Open

French Open Women

Formula 1

World Cup 2018

World Cup Boosts

ALL SPORTS

NUMBER OF BETS

American Football 261

Australian Rules 793

Baseball 558

Basketball 1223

Beach Volley 16

Boxing 31

Chess 2

27 May 2018 07:36:17

Won

Stake: €9.00

Odds: 2.25

Knight, Jason - Amirkhani, Makwan

Payout: €20.25

Bout Odds: Amirkhani, Makwan



26 May 2018 14:27:40

Won

Stake: €10.00

Odds: 1.42

Allen, Arnold

Payout: €14.20

Bout Odds: Allen, Arnold

26 May 2018 14:10:54

Won

Stake: €1.00

Odds: 1.16

Magny, Neil - White, Craig

Payout: €1.16

Bout Odds: Magny, Neil

26 May 2018 13:56:09

Won

Stake: €10.00

Odds: 1.95

Thompson, Stephen - Till, Darren

Payout: €19.50

Bout Odds: Till, Darren

Jason

```
{  
  "params": {  
    "filter": {  
      "eventIds": [1]  
    }  
  },  
  "jsonrpc": "2.0",  
  "method": "SportsAPING/v1.0/listCompetitions",  
  "id": 1  
}
```

- Application key
- Session token
- Certificates

The screenshot shows the Betfair website interface. At the top, there's a navigation bar with 'betfair' logo, 'My Bets', 'My Bonuses', and a search bar. Below that, there's a menu with categories like 'Exchange', 'Sportsbook', 'Casino', etc. The main content area displays a fight result for 'Santiago v Ige' on 'Sun 10 Jun, 0:15'. The fight result section includes a 'Going In-Play' checkbox, 'Cash Out' button, and 'Rules' link. A summary box shows 'Liability: €2.00' and 'Cash Out €1.78 Profit: -€0.22'. Below this, there's a table with 2 selections and odds for Mike Santiago and Dan Ige.

2 selections		103.0%		Back all	Lay all	96.8%	
Mike Santiago €2.04	1.01 €200	1.99 €214	2.12 €44	2.28 €22	2.36 €40	2.48 €211	
Dan Ige -€2.00	1.74 €54	1.78 €9	1.79 €19	1.89 €2	1.9 €47	2.02 €211	

Conclusie

Python rules!

Long live Internet 😊

OSTAGRAM.RU



