



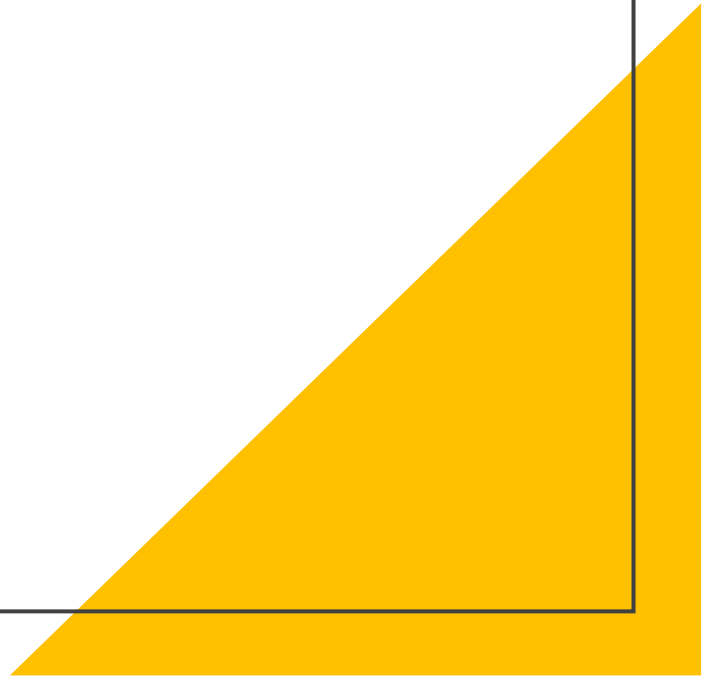
Foodboost

Externe presentatie

Groep 1

Inhoud

- Intro project
- Simulated Users
- Train set, Leave-One-Out
- Test set
- Model en Validatie
- Evaluatie resultaten



Project

- Data
- Onze doelstelling
- Favorieten voorspellen

```
In [12]: tagsdf.head()
```

```
Out[12]:
```

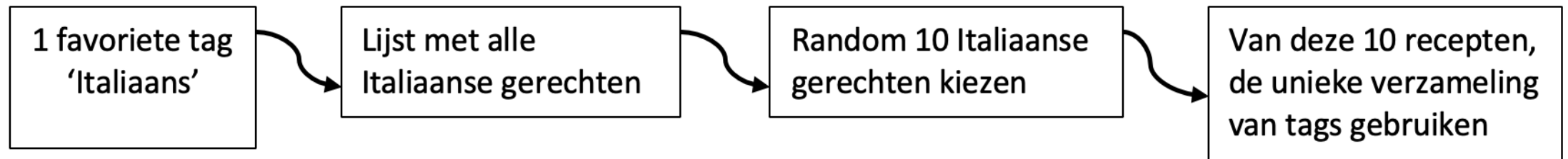
	recipe	tag
0	Kruidnoten met choco-discodip	hollands
1	Kruidnoten met choco-discodip	gebak
2	Kruidnoten met choco-discodip	gebak
3	Kruidnoten met choco-discodip	sinterklaas
4	Kruidnoten met choco-discodip	sinterklaasavond

```
In [11]: tags_bij_recept('Volkorenpasta met zoete-aardappelspaghetti')
```

```
Out[11]: array(['gezond', 'lactosevrij', 'hoofdgerecht', 'wat eten we vandaag',  
                'koken', 'eiwitrijk', 'suikerbewust'], dtype=object)
```

Simulated User

- Eén tag
- Random 10 favorieten



Trainset: Tag = 'italiaans'

		italiaans	sinterklaas	avondeten	...	2-italiaans	2-sinterklaas	2-avondeten	...	y	random	One-out
1	User1	1	-	-	...	1	-	-	...	1	NaN	Rec1
2	User1	1	-	-	...	0	-	-	...	0	rndm1	Rec1
3	User1	1	-	-	...	1	-	-	...	1	NaN	Rec2
4	User1	1	-	-	...	0	-	-	...	0	rndm2	Rec2
5	User1	1	-	-	...	1	-	-	...	1	NaN	Rec3
6	User1	1	-	-	...	0	-	-	...	0	rndm3	Rec3
...

Tags van 7 favorieten

Van L.O.O. of Random recept

Testset: Tag = 'italiaans'

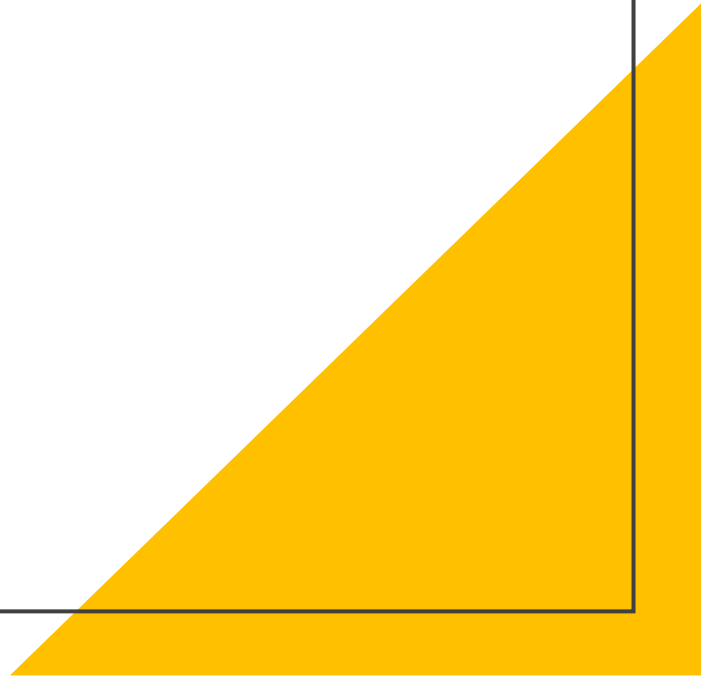
		italiaans	sinterklaas	avondeten	...	2-italiaans	2-sinterklaas	2-avondeten	...	y
1	User1	1	-	-	...	1	-	-	...	?
2	User1	1	-	-	...	-	-	-	...	?
3	User1	1	-	-	...	1	-	-	...	?
4	User1	1	-	-	...	-	-	-	...	?

Tags van 8 favorieten

De andere 2 favorieten

Uitbreiding

- Meer favoriete recepten
- Meer Users onder elkaar



Train, Validate, Test

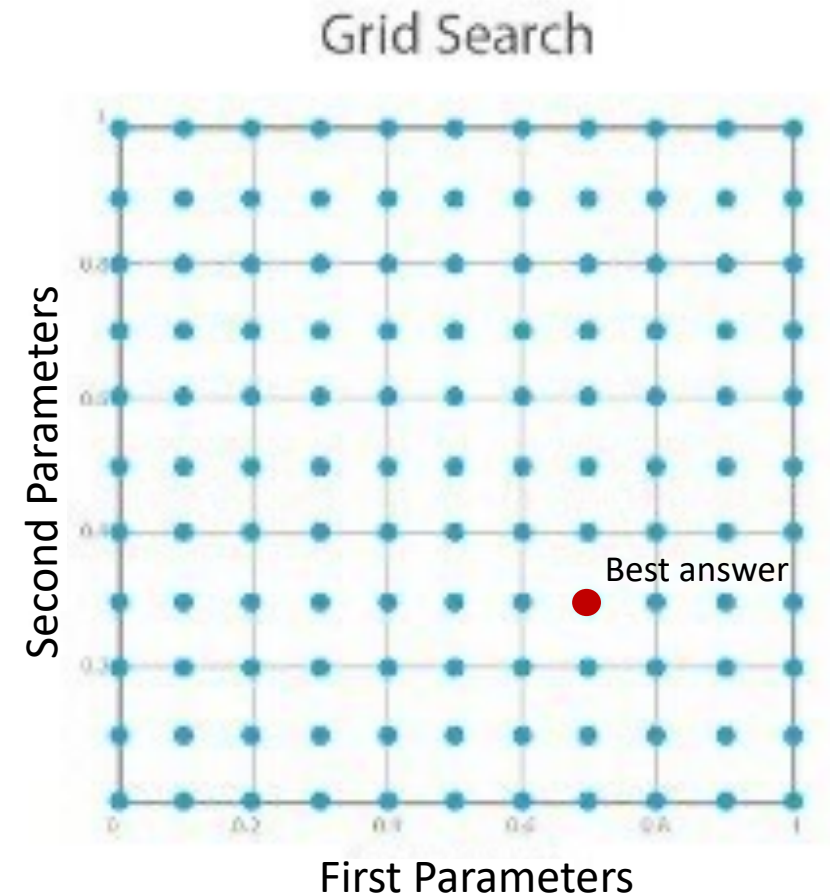
		italiaans	sinterklaas	avondeten	...	2-italiaans	2-sinterklaas	2-avondeten	...	y	random	One-out
1	User1	1	-	-	...	1	-	-	...	1	NaN	Rec1
2	User1	1	-	-	...	0	-	-	...	0	rndm1	Rec1
3	User1	1	-	-	...	1	-	-	...	1	NaN	Rec2
4	User1	1	-	-	...	0	-	-	...	0	rndm2	Rec2
5	User1	1	-	-	...	1	-	-	...	1	NaN	Rec3
6	User1	1	-	-	...	0	-	-	...	0	rndm3	Rec3
...

x

y

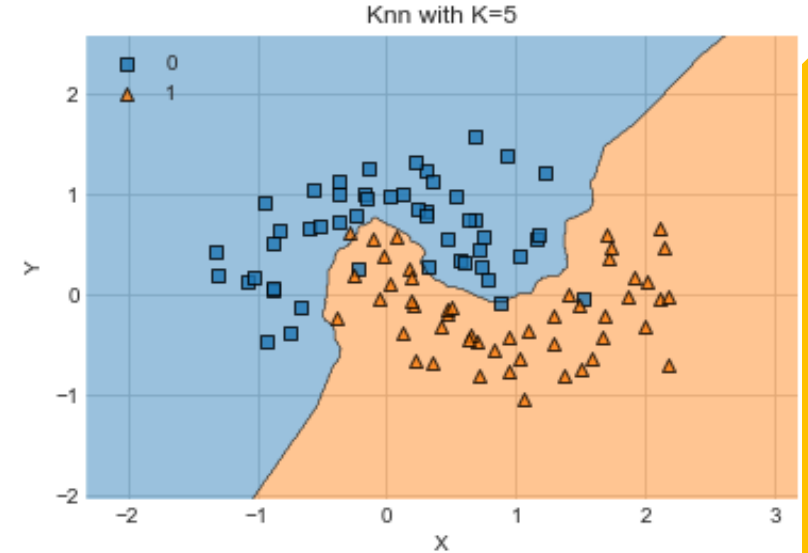
Welk model?

- **GridSearch:**
 - Logistic Regression
 - K-NearestNeighbors
 - Support Vector Classification (SCV)
 - Random Forest Classification



Testen & Validatie

- K-NearestNeighbors
- 7 Neighbors
- 79.1%



Discussie & Vragen

