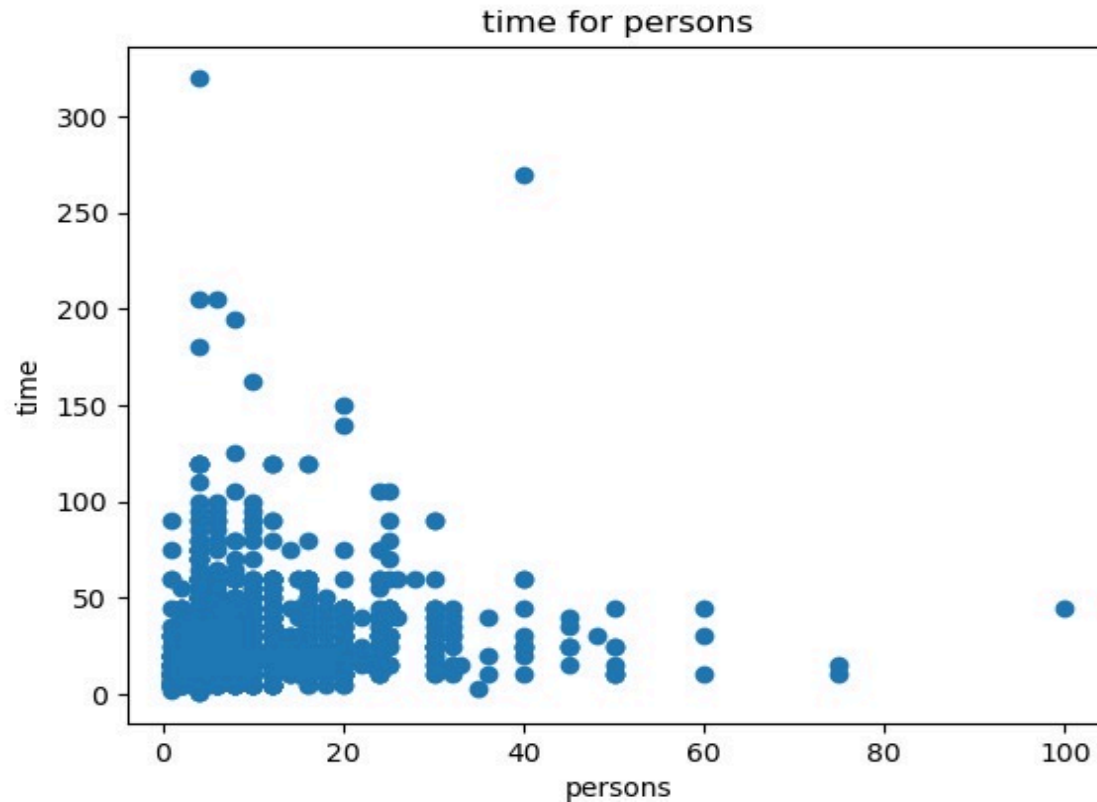




Interne Presentatie

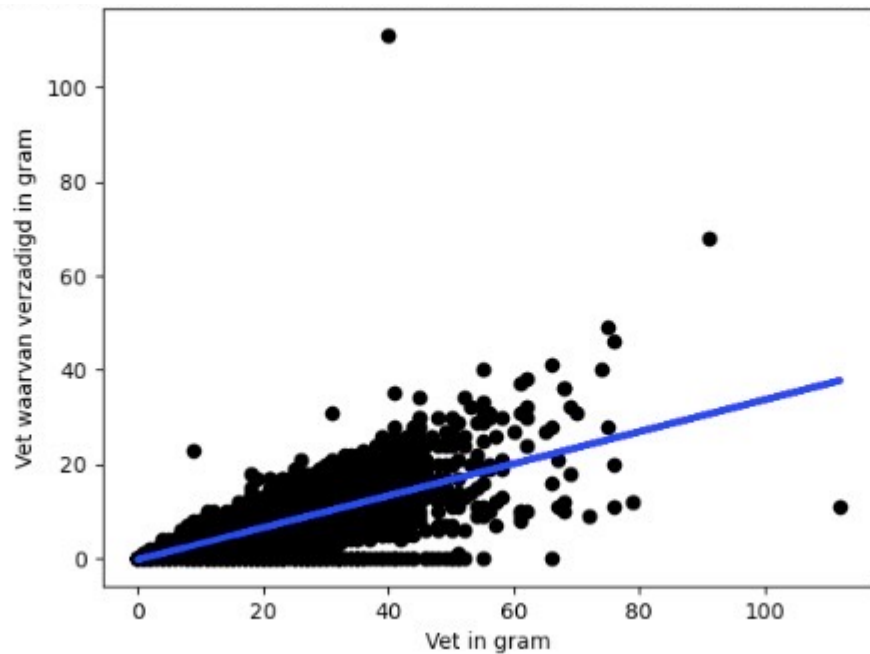
Project: Foodboost
groep 1

Data en Machine Learning

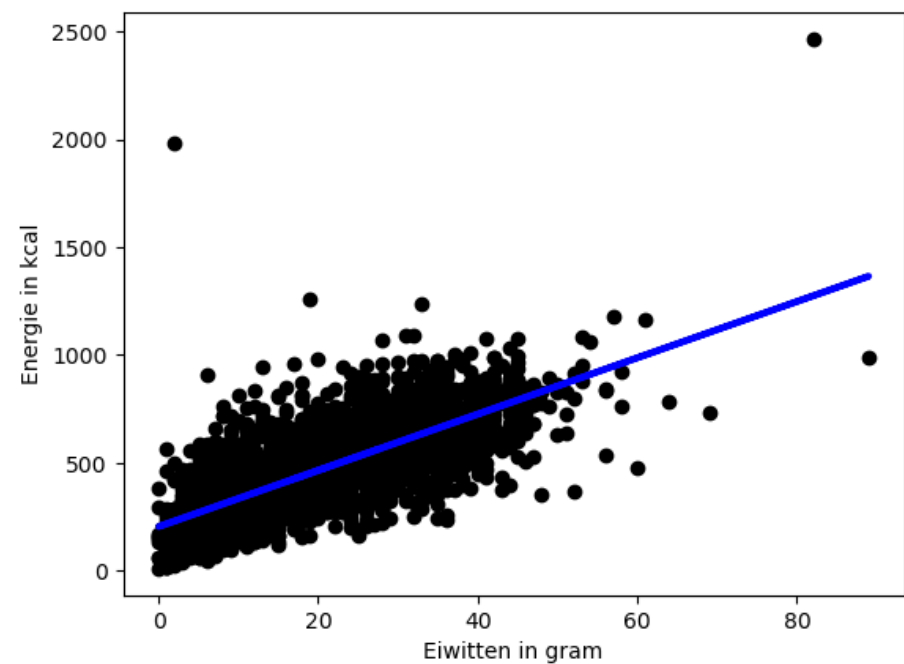


Data en Machine Learning

Lineaire Regressie: Voorspellen verzadigd vet op basis van hoeveelheid vet

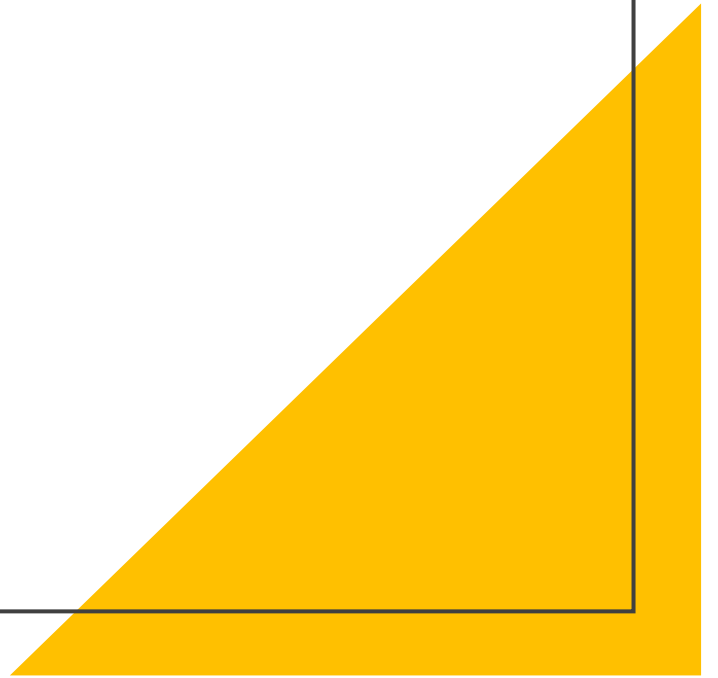


Lineaire Regressie: Voorspellen aantal kcal op basis van gram eiwitten

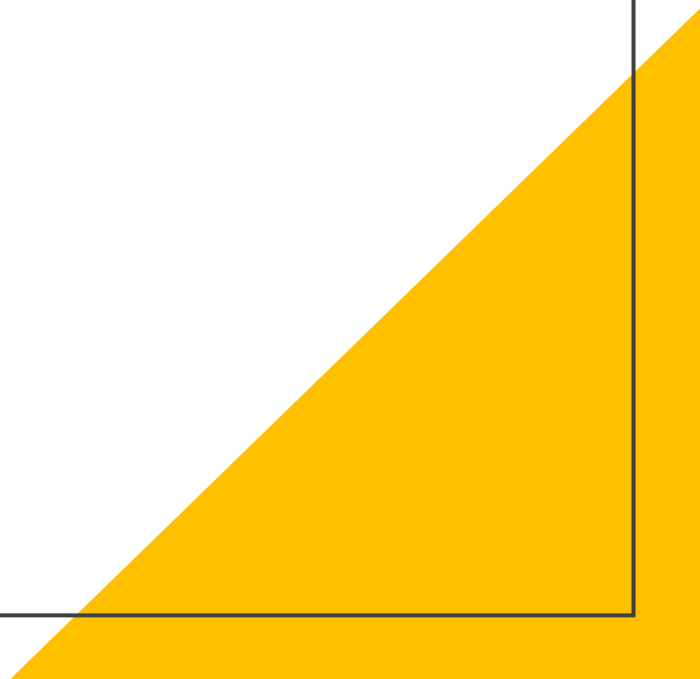


Aanpak

- Data krimpen
- Gesimuleerde gebruikers
 - Tags
 - Favorieten
- Tabellen met 1 of 0
- Trainen en testen




Gesimuleerde gebruikers



1. Data laden
2. Main menu
3. Model trainen
4. Gesimuleerde gebruikers

```
===== DATAFRAME INIT =====
loaded recipes.csv 8706 rows...
loaded ingredients.csv 71809 rows...
loaded nutritions.csv 58765 rows...
loaded tags.csv 46040 rows...
DATAFRAME_INIT TOOK 0.13 SECONDS
===== DATAFRAME CLEANUP =====
===== CLEANING RECIPES =====
_RENAME_RECIPE_TITLE_COLUMN TOOK 0.00 SECONDS
===== CLEANING RECIPES =====
_DROP_COLUMNS TOOK 0.00 SECONDS
===== CLEANING RECIPES =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING INGREDIENTS =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING INGREDIENTS =====
Grouped 71809 rows into 8706 rows
_GROUPBY TOOK 0.27 SECONDS
===== CLEANING INGREDIENTS =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING INGREDIENTS =====
_REPLACE_INGREDIENT_UNIT_NAN TOOK 0.87 SECONDS
===== CLEANING NUTRITIONS =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING NUTRITIONS =====
Grouped 58765 rows into 8706 rows
_GROUPBY TOOK 0.18 SECONDS
===== CLEANING NUTRITIONS =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING TAGS =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING TAGS =====
Grouped 46040 rows into 8705 rows
_GROUPBY TOOK 0.09 SECONDS
===== CLEANING TAGS =====
_RESET_INDEX TOOK 0.00 SECONDS
===== CLEANING TAGS =====
_CLEAN_LIST_TAGS TOOK 0.01 SECONDS
CLEANER_EXEC TOOK 1.42 SECONDS
Press Enter to continue
```

- 1). Dataframes
- 2). Recept submenu
- 3). Model training

- 1). Import existing user data
- 2). Sim users 
- 3). Export sim user data
- 4). Train model
- 0). Vorige menu

Aantal gebruikers:

10

Threshold voor favorieten (0.0 - 1.0):


.5

1. Data laden
2. Main menu
3. Model trainen
4. Gesimuleerde gebruikers
5. Menu exporteren











```
Aantal gebruikers:
10
Threshold voor favorieten (0.0 - 1.0):
.5
===== USER FAVORITE SIMULATION =====
Dataframe not merged yet, merging now..
===== DATAFRAME MERGE =====
DATAFRAME_MERGE_ALL TOOK 0.01 SECONDS
Aaron          43 United Arab Emi...  simmed with:    40/8706 favorites
Lenore         56 Sweden              simmed with:    28/8706 favorites
Samuel         63 Denmark              simmed with:   120/8706 favorites
Florine        46 Monaco                simmed with:   116/8706 favorites
Charles        25 Guam                  simmed with:    34/8706 favorites
Ronald         64 Mali                  simmed with:    24/8706 favorites
Harry         71 Saint Vincent a...  simmed with:    54/8706 favorites
Leanne         33 Costa Rica              simmed with:    17/8706 favorites
Donald         32 Jersey                  simmed with:    11/8706 favorites
Ernest         60 Burundi              simmed with:    11/8706 favorites
_CREATE_SIM_USERS_INTERNAL TOOK 24.88 SECONDS
done'd
Press Enter to continue
```

1. Data laden
2. Main menu
3. Model trainen
4. Gesimuleerde gebruikers
5. Menu exporteren
 1. Gesimuleerde gebruiker data
 2. Files

```
1). Import existing user data  
2). Sim users  
3). Export sim user data  
4). Train model  
0). Vorige menu
```



> _py > usergen_export > 20220918-153706

Name	Date modified	Type	Size
 Aaron.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,663 KB
 Charles.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,659 KB
 Donald.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,657 KB
 Ernest.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,653 KB
 Florine.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,686 KB
 Harry.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,655 KB
 Leanne.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,591 KB
 Lenore.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,646 KB
 Ronald.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,654 KB
 Samuel.csv	18 Sep 2022 15:37	Microsoft Excel C...	2,642 KB


```
scores = [
    calculate_similarity(row['tag'], user.favorite_tags),
    calculate_similarity(row['ingredient'], user.favorite_ingredients)
]

mean_score = np.mean(scores)
favorite_values.append(1 if mean_score >= user.favorite_threshold else 0)
```

(0.5)

Aantal gebruikers:
10
Threshold voor favorieten (0.0 - 1.0):
.5

Random geselecteerd

Op basis van de
gelijkenis van de tags
en ingrediënten

	A	B	C	D	E	F	G	H
			ingredient	tag	ingr_score	tag_score	mean_score	favorite
1		recipe						
2	0	Bananenbroodtulband met noten	['banaan', 'zelfrijzend bakmeel', 'gemalen l	['traktatie', 'gebak', 'oven', 'vooraf te make		1	1	1
3	1	Citymoms honnysalade	['traditionele olijfolie', 'gegrilde aubergine	['salade', 'bakken', 'lunch', 'glutenvrij']		1	1	1
4	2	Chileense empanada's de pino	['middelgrote ui', 'zonnebloemolie', 'runde	['chileens', 'borrelhapje', 'zuid-amerikaans		1	1	1
5	3	Snelle haastaart	['ongezouten roomboter', 'poedersuiker', 'l	['pasen', 'gebak', 'oven']		1	1	1
6	4	Chineeskoolreepjes met bosui	['Chinese kool', 'bosui', 'rijstolie', 'knoflook	['aziatisch', 'roerbakken/wokken', 'bijgere		1	1	1
7	5	Kruidendip met cottagecheese	['cottagecheese', 'kruidenboterkruiden', 'c	['borrelhapje', 'picknick', 'budget', 'borrel',		1	1	1
8	6	Pasta met blauwe kaas & pompoenpitten	['pompoenpitten', 'cherrytomaten', 'knoflc	['hoofdgerecht', 'pasta', 'koken', 'wat eten		1	1	1
9	7	Woknoedels met ei en pinda's	['mais scharrelei', 'eiernoedels', 'arachidec	['hoofdgerecht', 'noedels', 'aziatisch', 'roer		1	1	1
10	8	Mexicaanse wraps met kip, kidneybonen en paprika	['rode ui', 'paprikamix', 'verse koriander',	['hoofdgerecht', 'wrap', 'bakken', 'mexicaa		1	1	1
11	9	Spaghetti met homemade Italiaanse gehaktballetjes	['mix voor gehakt Italiaans', 'geraspte Parr	['hoofdgerecht', 'pasta', 'Italiaans', 'bakken		1	1	1
12	10	Kruidnoten met choco-discodip	['melkchocolade', 'kruidnoten', 'discodip']	['gebak', 'sinterklaasavond', 'sinterklaas', 'l	0	0	0	0
13	11	Kruidnoten in marsepein	['blanke marsepein', 'ongezouten roombot	['gebak', 'sinterklaasavond', 'sinterklaas', 'l	0.288675135	0	0.144337567	0
14	12	Kruidnoten met chocodips	['pure chocolade', 'ongezouten roomboter	['gebak', 'sinterklaasavond', 'sinterklaas', 'l	0.25819889	0	0.129099445	0
15	13	Pepernotentaart met marsepeinstrik	['ongezouten roomboter', 'witte basterdsu	['gebak', 'sinterklaasavond', 'oven', 'hollan	0.348155312	0.25819889	0.303177101	0
16	14	Perencake	['middelgroot ei', 'Griekse yoghurt 10%', 'fi	['oven', 'gebak', 'vooraf te maken']	0.19245009	0.333333333	0.262891712	0
17	15	Massaman curry	['koelverse sereh citroengras', 'kokosmelk']	['hoofdgerecht', 'koken', 'thais', 'curry', 'azi	0.107211253	0.23570226	0.171456757	0
18	16	Tartiflette	['vastkokende aardappel', 'middelgrote ui']	['hoofdgerecht', 'oven', 'frans', 'glutenvrij']	0.204124145	0.577350269	0.390737207	0
19	17	Carpaccio rolletjes	['grote tortillawrap', 'Greenfields carpacci	['traktatie', 'borrelhapje', 'pasen', 'wrap', 'c	0	0	0	0
20	18	Ajam pangang	['limoen', 'scharrel kippenpoten naturel',	['hoofdgerecht', 'aziatisch', 'oven', 'indone	0.522232968	0.577350269	0.549791619	1
21	19	Pepernotentaart	['middelgroot ei', 'mandarijn', 'vanille-extr	['sinterklaasavond', 'sinterklaas', 'gebak', 'l	0	0.288675135	0.144337567	0
22	20	Appelkruimeltaart	['middelgroot ei', 'citroen', 'tarwebloem', 'l	['oven', 'verjaardag', 'gebak', 'vooraf te ma	0.3086067	0.288675135	0.298640917	0
23	21	Bloemkoolschotel	['knoflook', 'bloemkool', 'ongezouten roon	['hoofdgerecht', 'wat eten we vandaag', 'h	0.182574186	0.577350269	0.379962228	0
24	
25	80	Stoofvlees met zoete-aardappelfriet en rodekoolsalade	['ovenfriet zoete aardappel', 'milde olijfoli	['hoofdgerecht', 'oven', 'wat eten we vand	0	0.516397779	0.25819889	0
26	81	Spaghetti met tartaar, tomaat en wortel	['bospeen', 'tomaten', 'verse platte peters	['hoofdgerecht', 'pasta', 'bakken', 'wat eter	0	0.471404521	0.23570226	0
27	82	Biefstuk met romige paddenstoelenjus (advertorial)	['iets kruimige aardappel', 'biefstuk', 'Vers	['hoofdgerecht', 'bakken']	0.40824829	0.816496581	0.612372436	1
28	83	Gestoorde pompoen met puntpaprika en kidneybonen (ai	['puntpaprika', 'arachideolie', 'koelverse m	['hoofdgerecht', 'vegetarisch', 'slank', 'gezc	0.19245009	0.23570226	0.214076175	0