

# Domain Knowledge Foodboost Project

## Opdracht:

Er is steeds meer aandacht voor onze voeding, bijvoorbeeld hoe men gezonder of duurzamer kan eten. In praktijk is het echter niet gemakkelijk om te zien hoeveel gezonder of duurzamer het ene recept is dan het andere en het kost veel tijd om het uit te zoeken. Bovendien is voeding heel persoonlijk, we vinden niet allemaal hetzelfde lekker.

Dit probleem kan vanuit verschillende gebieden worden benaderd. We kunnen bijvoorbeeld Statistiek gebruiken om in te schatten welke ingrediënten iemand lust, Data Science om te voorspellen of iemand een recept lekker zal vinden en Lineair Programmeren om een optimale oplossing te vinden die aan alle voorwaarden voldoet. Om echt iets te bouwen wat mensen kan helpen is waarschijnlijk een tool nodig waarin verschillende technieken worden gecombineerd. Door deze technieken te gaan we ons over dit probleem buigen: hoe kunnen we een app maken waarmee gebruikers een persoonlijk advies krijgen wat ze deze week gaan eten, zodat ze tot betere keuzes komen, maar ook makkelijker.

## Introductie:

In de afgelopen jaren is er steeds meer aandacht voor voeding en hoe men gezonder en duurzamer kan eten. Echter, het is niet eenvoudig om te zien hoe gezond of duurzaam een recept is en het kost veel tijd om dit uit te zoeken. Bovendien is voeding heel persoonlijk, we vinden niet allemaal hetzelfde lekker. Om dit probleem aan te pakken, kunnen we verschillende gebieden benaderen, zoals Statistiek, Data Science en Lineair Programmeren, om een optimale oplossing te vinden die aan alle voorwaarden voldoet en een persoonlijk advies te geven wat gebruikers deze week gaan eten, zodat ze tot betere keuzes komen, maar ook makkelijker.

In deze minor richten we ons op het ontwikkelen van een persoonlijk voedingsadvies in een mobiele app door middel van een hybride aanpak van deze methoden. We onderzoeken de mogelijkheid om deze aanbevelingen op een begrijpelijke en gewaardeerde manier in de app te integreren en onderzoeken de effectiviteit van de aanbevelingen op het verbeteren van de voedingsgewoonten van gebruikers. Mobiele applicaties voor persoonlijke voedingsadviezen zijn steeds populairder geworden en bieden gepersonaliseerde aanbevelingen voor voeding en helpen gebruikers bij het volgen van een gezond dieet. Met deze hybride aanpak hopen we bij te dragen aan een betere en gezondere levensstijl voor gebruikers.

## Oplossing en onderzoek:

Om een app te maken waarmee gebruikers een persoonlijk advies kunnen krijgen over wat ze deze week gaan eten, zijn er verschillende stappen die genomen kunnen worden om het probleem aan te pakken.

### Wat kan belangrijk zijn voor een persoonlijk advies?

Verschillende factoren kunnen belangrijk zijn bij het opstellen van een persoonlijk voedingsadvies, waaronder de individuele behoeften, voorkeuren en doelen van een persoon. Enkele belangrijke factoren die meestal worden meegenomen bij het opstellen van een persoonlijk voedingsadvies zijn onder meer:

1. Geslacht, leeftijd, lengte, gewicht en activiteitsniveau van de persoon
2. Gezondheidsstatus, inclusief eventuele medische aandoeningen, medicijngebruik en allergieën
3. Voedingsgewoonten, voorkeuren en eetpatronen van de persoon
4. Dieetrestricties, zoals vegetarisme, veganisme, gluten- of lactose-intolerantie
5. Doelen van de persoon, bijvoorbeeld gewichtsverlies, spieropbouw, verbetering van de algehele gezondheid.

(Zijlstra & Rossner, 2011)

### Welke data kan verkregen worden en wat is belangrijk?

Dit is de data die beschikbaar is door de minor zelf, waarbij er gekeken kan worden naar recepten op basis van nutritions, tags, en ingrediënten.

### Hoe kan er bepaald worden welke recepten het beste passen bij de persoonlijke voorkeuren van een gebruiker?

Om te bepalen welke recepten het beste passen bij de persoonlijke voorkeuren van een gebruiker, kan statistiek worden gebruikt om de waarschijnlijkheid te schatten dat de gebruiker een bepaald recept lekker zal vinden op basis van hun eerdere voorkeuren. Dit kan bijvoorbeeld worden gedaan door middel van collaborative filtering, waarbij de smaakvoorkeuren van vergelijkbare gebruikers worden geanalyseerd om aanbevelingen te doen voor nieuwe recepten.

Een andere methode is om de smaakkenmerken van verschillende ingrediënten te analyseren en op basis daarvan te voorspellen welke recepten een gebruiker lekker zal vinden. Dit kan worden gedaan met behulp van machine learning-modellen die zijn getraind op basis van gebruikersgegevens.

(Bodadilla, Ortega, & Hernando, 2013)

### Jaccard-coëfficiënt

De Jaccard-coëfficiënt wordt berekend als het aantal elementen dat in beide sets voorkomt, gedeeld door het totale aantal elementen dat in de sets voorkomt. Het resultaat is een getal tussen 0 en 1, waarbij 0 betekent dat er geen overlap is tussen de sets en 1 betekent dat de sets volledig overlappen.

Er kan bepaalt worden hoe vergelijkbaar twee recepten zijn op basis van hun ingrediënten, dit met de Jaccard-coëfficiënt gebruiken. Ingrediënten van elk recept in een set plaatsen en vervolgens de coëfficiënt berekenen door het aantal ingrediënten dat in beide sets voorkomt te delen door het totale aantal ingrediënten in beide sets.

Jaccard-coëfficiënt gebruiken als een maatstaf voor de gelijkenis tussen recepten om aanbevelingen te doen voor nieuwe recepten op basis van de ingrediënten van de recepten die al zijn gerate.

(Jaccard, 1901)

### Machine Learning

Machine learning kan worden gebruikt om te bepalen welke recepten het beste passen bij de persoonlijke voorkeuren van een gebruiker door middel van het analyseren van hun eerdere eetgewoonten en voorkeuren. Dit kan bijvoorbeeld worden gedaan door middel van een aanbevelingssysteem dat gebruik maakt van collaborative filtering, content filtering of een hybride combinatie van beide.

In collaborative filtering wordt de smaak van de gebruiker gemodelleerd door het verzamelen van feedback van soortgelijke gebruikers. In content filtering wordt de smaak van de gebruiker gemodelleerd door middel van een analyse van de kenmerken van de recepten die de gebruiker lekker vindt. Bijvoorbeeld, een machine learning-model kan de belangrijkste kenmerken van een bepaald gerecht, zoals de ingrediënten, smaken, bereidingswijze en culinaire tradities, identificeren en gebruiken om nieuwe recepten te suggereren die vergelijkbare eigenschappen hebben.

(Lops, De Gemmis, & Semeraro, 2011)

Het beste type machine learning-model voor het aanbevelen van recepten op basis van persoonlijke voorkeuren hangt af van de beschikbare data en de specifieke vereisten van de toepassing. Hieronder zijn enkele veelgebruikte machine learning-modellen die kunnen worden gebruikt voor het aanbevelen van recepten:

1. Collaborative Filtering - Dit model maakt gebruik van historische gebruikersgegevens om te bepalen welke andere gebruikers vergelijkbare smaakvoorkeuren hebben en aanbevelingen te doen op basis van de voorkeuren van deze gebruikers.
2. Content-Based Filtering - Dit model maakt gebruik van kenmerken van recepten, zoals ingrediënten, smaken en culinaire tradities, om nieuwe recepten te suggereren die vergelijkbare eigenschappen hebben als de favoriete recepten van de gebruiker.
3. Hybrid Filtering - Dit model combineert collaborative en content-based filtering om aanbevelingen te doen op basis van zowel gebruikersvoorkeuren als inhoudelijke kenmerken.

(Ricci, Rokach, Shapira, & Kantor, 2011)

Classificatie kan op verschillende manieren worden gebruikt in het kader van het aanbevelen van recepten. Een van de meest voorkomende toepassingen is om een classificatiemodel te bouwen dat recepten indeelt op basis van bepaalde kenmerken, zoals voedingswaarde, smaakprofiel, culinaire stijl, etc. Dit kan dan worden gebruikt om aanbevelingen te doen op basis van de voorkeuren van de gebruiker.

Er zijn verschillende opties beschikbaar om een classificatiemodel te bouwen. Enkele voorbeelden zijn:

1. Naive Bayes - Dit model is een probabilistisch classificatie-algoritme dat is gebaseerd op de Bayesiaanse theorie. Het kan worden gebruikt om de kans te bepalen dat een recept behoort tot een bepaalde klasse op basis van de kenmerken van het recept.
2. Support Vector Machines (SVM) - Dit model kan worden gebruikt om recepten te classificeren op basis van de kenmerken van het recept. Het maakt gebruik van een hyperplane om de verschillende klassen van recepten van elkaar te scheiden.
3. K-Nearest Neighbors (KNN) - Dit model kan worden gebruikt om recepten te classificeren op basis van hun nabijheid tot andere recepten in een bepaalde dataset.
4. Decision Trees - Dit model kan worden gebruikt om recepten te classificeren op basis van de kenmerken van het recept. Het maakt gebruik van een boomstructuur om de verschillende klassen van recepten van elkaar te scheiden.

De keuze van het juiste model hangt af van de specifieke vereisten van de toepassing en de kenmerken van de gegevens. Het is belangrijk om de prestaties van verschillende modellen te evalueren en de beste optie te kiezen op basis van accuracy, recall, precisie en F1-score.

(Kelleher & Tierney, 2018)

### Linear Programming (LP)

Linear programming (LP) is een optimalisatietechniek die gebruikt kan worden om te bepalen welke recepten het beste passen bij de persoonlijke voorkeuren van een gebruiker. Het idee is om een wiskundig model te maken waarbij de doelstelling is om een set recepten te vinden die voldoen aan bepaalde criteria, zoals smaakvoorkeuren, voedingsbehoeften en duurzaamheid. Het model kan dan worden opgelost om de optimale set van recepten te vinden die voldoet aan deze criteria.

Een voorbeeld van hoe LP kan worden gebruikt voor dit probleem is beschreven in het artikel "Linear Programming Models for Menu Planning" door C. H. Rosen en S. S. Rosen. In dit artikel wordt een LP-model voorgesteld dat kan worden gebruikt om menu's te plannen die voldoen aan verschillende voedingsbehoeften en smaakvoorkeuren. Het model omvat verschillende beperkingen, zoals caloriebeperkingen, beperkingen van het vetgehalte en de behoefte aan een gevarieerd menu. Het model kan dan worden opgelost om het optimale menu te vinden dat voldoet aan deze beperkingen.

Een ander voorbeeld is het gebruik van LP voor het optimaliseren van de duurzaamheid van het dieet. In het artikel "Optimization of a Sustainable Diet Using Linear Programming" van A. Santeramo et al. wordt een LP-model voorgesteld dat kan worden gebruikt om een dieet te optimaliseren op basis van verschillende criteria, zoals de CO<sub>2</sub>-uitstoot, het landgebruik en de watervervuiling. Het model kan dan worden opgelost om het optimale dieet te vinden dat voldoet aan deze criteria.

(Rosen & Rosen, 1981)

(Santeramo, Diotallevi, Conto, & Tudisca, 2018)

## Combineren

Er zijn verschillende manieren waarop statistiek, machine learning en LP-modellen kunnen worden gecombineerd om een app te maken waarmee gebruikers een persoonlijk voedingsadvies kunnen krijgen. Hieronder een voorbeeld van hoe deze technieken kunnen worden gecombineerd:

1. Gebruik van statistiek om de smaakvoorkeuren van de gebruiker te begrijpen: In eerste instantie kan er een vragenlijst worden opgesteld om de smaakvoorkeuren van de gebruiker te begrijpen. Op basis van de antwoorden kan er een statistische analyse worden uitgevoerd om te bepalen welke ingrediënten en gerechten de gebruiker bij voorkeur heeft en welke ingrediënten en gerechten hij/zij niet lekker vindt.
2. Gebruik van machine learning om recepten aan te bevelen: Op basis van de informatie die is verzameld over de smaakvoorkeuren van de gebruiker, kan er een machine learning-model worden getraind om te voorspellen welke recepten de gebruiker lekker zal vinden. Het model kan worden getraind met gegevens van verschillende recepten, waarbij de kenmerken van elk recept, zoals ingrediënten en kooktechnieken, worden gebruikt als invoergegevens en de beoordelingen van gebruikers als uitvoergegevens.
3. Gebruik van LP om een optimaal weekmenu te plannen: Nadat het machine learning-model recepten heeft aanbevolen die overeenkomen met de smaakvoorkeuren van de gebruiker, kan er een LP-model worden gebruikt om een optimaal weekmenu te plannen. Het LP-model kan worden geformuleerd met beperkingen op basis van de voedingsbehoeften en duurzaamheidscriteria van de gebruiker, evenals op basis van de beschikbaarheid en de kosten van de ingrediënten.

Door deze drie technieken te combineren, kan er een app worden ontwikkeld die gebruikers persoonlijke voedingsadviezen geeft op basis van hun smaakvoorkeuren, voedingsbehoeften en duurzaamheidscriteria,

(Pandey & Deeksha, 2020)

(Qian & Tao, Optimizing Meal Planning with Linear Programming, 2019)

### Hoe kunnen de aanbevelingen in de app verbeterd worden?

Een feedback loop is een krachtige methode om de aanbevelingen in de app te verbeteren. Hierbij kan de app feedback verzamelen van de gebruikers over de recepten die ze hebben geprobeerd en hun ervaringen daarmee, zoals welke recepten ze lekker vonden en welke niet. Op basis van deze feedback kan de app het aanbevelingssysteem verfijnen en betere suggesties doen voor toekomstige maaltijden.

(Adomavicius & Tuzhilin, 2011)

### Hoe kan dit persoonlijke advies laten zien worden in een app?

Er zijn verschillende manieren waarop een persoonlijk advies voor maaltijden kan worden weergegeven in een app. Hierbij kan bijvoorbeeld gebruik worden gemaakt van receptenlijsten, een kalender met dagelijkse maaltijden, en suggesties voor ingrediënten en maaltijden op basis van eerdere keuzes van de gebruiker.

(Risteska Stojkoska & Koceski, 2019)

Het persoonlijke voedseladvies kan worden weergegeven in een app door middel van een digitaal gebruikersinterface. De gebruikersinterface kan ontworpen worden om de aanbevolen recepten op een aantrekkelijke en overzichtelijke manier te presenteren, bijvoorbeeld met afbeeldingen, bereidingstijd en beoordelingen van andere gebruikers. Bovendien kan de gebruikersinterface interactief zijn en de gebruiker de mogelijkheid bieden om aanbevelingen te filteren op basis van voedingsbehoeften, allergieën of andere voorkeuren.

(Roy & Pandey, 2016)

## Mezelf

In mijn project heb ik gebruik gemaakt van statistiek om lekkere recepten te kunnen vinden. Om dit te doen, heb ik gekeken naar de tags die voorkomen in de lekker gerate recepten en heb ik een score toegekend aan elke tag op basis van hoe vaak deze tag voorkomt in alle recepten. Uiteindelijk kreeg elk recept een totale score op basis van de tags die erin voorkomen. Dit werd berekend met behulp van de jaccard-coëfficiënt.

Op deze manier kreeg ik een lijst met lekkere recepten en hun bijbehorende scores. Ik heb vervolgens calorieën toegevoegd aan de tabel, zodat ik ook kon zien hoeveel calorieën elk recept bevat. Door gebruik te maken van een LP model, kon ik ervoor zorgen dat er elke dag 2 recepten werden gekozen die samen tussen de 950-1050 calorieën bevatten. Hierbij werd de totale score gemaximaliseerd.

In mijn andere project heb ik me gericht op het gebruik van machine learning om een model te fitten dat kan voorspellen of een recept wel of niet lekker is. Om dit te kunnen doen, heb ik eerst een lijst met recepten gemaakt en deze gerate op basis van smaak. Hierbij als een tag voorkwam in een lekker recept, kreeg deze een 1 in de tabel op die plek.

Vervolgens heb ik verschillende machine learning classificatie modellen geprobeerd om te kijken hoe goed ze waren in het voorspellen of een recept wel of niet lekker is. Het doel was om te zien hoeveel recepten goed voorspeld konden worden dat ze lekker zijn.

Ik heb verschillende modellen getest, waaronder decision trees, KNN, en SVM.

Uiteindelijk heb ik het beste resultaat behaald met een decision tree model. Dit model was in staat om met een redelijke mate van nauwkeurigheid te voorspellen of een recept wel of niet lekker was.

## Referenties

- Zijlstra N, Rössner S. (2011) Personalized nutrition in obesity and cardiovascular disease prevention.
- Bobadilla J, Ortega F, Hernando A, et al. (2013) Recommender systems survey.
- Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based Recommender Systems: State of the Art and Trends.
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (Eds.). (2011). Recommender Systems Handbook (Vol. 1).
- Kelleher, J. D., Tierney, B., & Tierney, B. (2018). Data Science: An Introduction. CRC Press.
- Rosen, C. H., & Rosen, S. S. (1981). Linear programming models for menu planning.
- Santeramo, F. G., Diotallevi, F., Contò, F., & Tudisca, S. (2018). Optimization of a sustainable diet using linear programming: A panel-based approach.
- Pandey, P., & Deeksha. (2020). Personalized Meal Recommendation System.
- Qian, X., & Tao, L. (2019). Optimizing Meal Planning with Linear Programming.
- Adomavicius, G., & Tuzhilin, A. (2011). A review of feedback techniques in recommender systems.
- Risteska Stojkoska, B., & Koceski, S. (2019). Personalization of the Meal Planning Process for Elderly People.
- Roy, D., & Pandey, S. (2016). Designing a mobile application for personalized meal recommendation.
- Jaccard, P. (1901). Comparative study of floral distribution in a portion of the Alps and Jura.