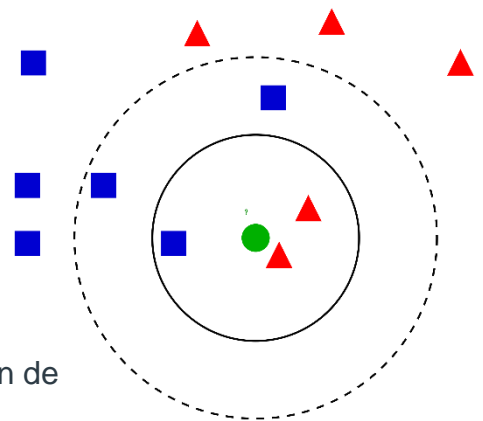


### Een probleembeschrijving, vanuit het perspectief van je (hypothetische) klant;

De klant zit op een terras en vraagt aan de ober een biertje. Na het drinken van het biertje weet hij niet meer wat die precies gedronken heeft. Hij weet alleen dat het biertje een bepaalde percentage alcohol heeft en wat voor soort biertje het was ( BIJV een IPA) Hij gaat naar huis en vult de waardes die hij weet in het algoritme. Het algoritme zoekt uit welk biertje het meest lijkt op de eigenschappen die hij heeft ingevuld, en hieruit blijkt dat het een oud vertrouwde groene soldaat was ( Heineken).

### Een beschrijving van het gekozen algoritme, die refereert aan minstens 1 geldige en toepasselijke bron;

Het KNN algoritme is een algoritme waarbij een ontbrekende of niet bekende waarde in de dataset, wordt geclassificeerd door te kijken naar welke waarde het dichtst in de buurt/het meest lijkt op de ontbrekende waarde. Hier in het onderstaande plaatje staat uitgelegd hoe het algoritme in zijn werking gaat. De groene cirkel ( X ) heeft een waarde waarvan onbekend is of die rood of blauw moet zijn. Door een bepaalde waarde K te nemen ( afstand van waarde tot omtrek circkel ) kan gekeken worden naar waardes die het meest lijken op die van X en zo kan X geclassificeerd worden. De waarde van K is zoals je ziet heel belangrijk voor de uitkomst. Deze moet dan ook goed gekozen worden om een betrouwbaar model op te stellen. Deze waarde kiest men door veel te testen met verschillende waarden voor X en de uitkomsten te vergelijken.



Sutton, C. (2012). Nearest-neighbor methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(3), 307–309. <https://doi.org/10.1002/wics.1195>

### De taken die je gaat uitvoeren, met aangegeven welke noodzakelijk en welke optioneel zijn;

#### Noodzakelijk:

- Repository opzetten
- Werkomgeving klaar zetten → welke taal, welke IDE etc.
- KNN goed bestuderen voor een duidelijk beeld om zo min mogelijk verwarring te voorkomen

#### - Database:

- Controleren en bestuderen → Welke info kan ik er al uit halen?
- Database verbinden met code → Hoe ga ik de data inladen?
- Data weergeven dmv code

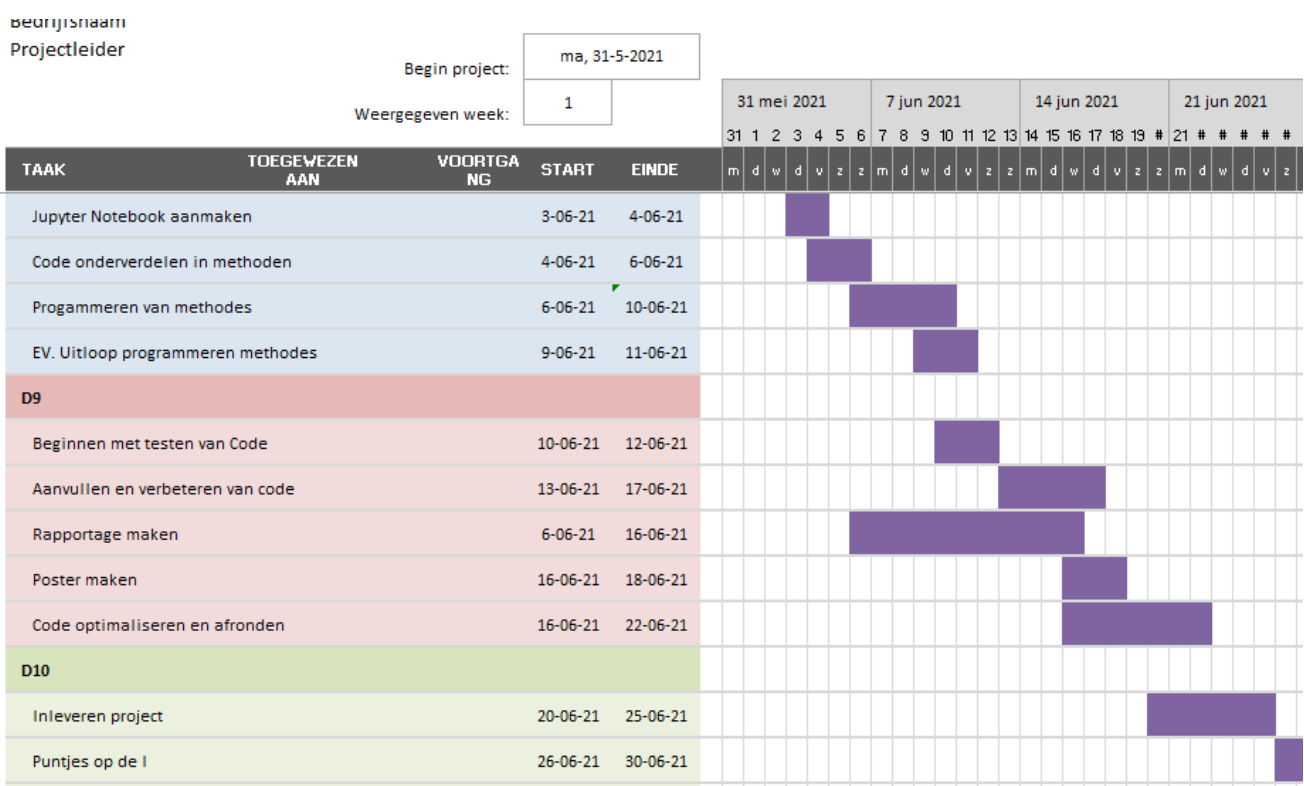
## - Algoritme

- Werking van algoritme → Hoe en wat is er nodig voor een werkend algoritme?
- Programmeren van de code
- Het testen van de code/algoritme

## -Visualiseren van het algoritme

- Jupyter notebook aanmaken voor grafisch beeld
- E.v. frontend bouwen bij tijd over

Een heldere planning van de projectweken, liefst schematisch weergegeven;



Risico's die je voorziet, en welke maatregelen je gaat treffen als die zich voordoen.

Een risico die zich kan vormen gedurende project is tijd tekort. Door te lang bezig te zijn met een onderdeel waar ik moeite mee heb, kan het project zomaar uitlopen. Ik weet van mijzelf dat ik ook niet snel om hulp vraag aan mensen, omdat ik het graag zelf wil uitzoeken. Een maatregel die ik hiervoor ga treffen is dat ik zodra ik ergens vastloop meteen aan iemand vraag om hulp.