Uitleg LanguageRecognizer

Om correcte werking te bevestigen is er eerst een simpel Java bestand gemaakt, die het bestand leest en verwerkt. Dit bestand is te vinden op: <https://github.com/opasta/LanguageRecognizer/blob/main/src/Main.java>> Dit bestand is volledig voorzien van commentaar. Om de output op waarde te kunnen controleren is voor elke zin bepaald of deze Nederlands of Engels is. Dit is in het bestand <https://github.com/opasta/LanguageRecognizer/blob/main/src/sentences.nl-enCHECKED.txt> te vinden. Dit is omgezet naar een array zodat de verwachte output kwa taal kon worden vergeleken met de daadwerkelijke output. De output in dit bestand is:

Afbeelding met tekst, Lettertype, schermopname, ontwerp

Automatisch gegenereerde beschrijving

Dit had 73 om 199 moeten zijn, en doormiddel van de array kan worden concludeert dat 173/192 (90.104%) correct zijn (9,9% afwijking).

Om de werking van het systeem te bewijzen is de volgende video gemaakt: [Video](https://youtu.be/BNQ-BYPlphY).

**Hadoop systeem:**

Het systeem wordt gedraait met het volgende commando: 

(er zijn 2 systemen, m3 is hoofd waar het meeste op gebeurt, m5 helpt)

Wanneer het bestand wordt gedraaid wordt de volgende output getoond:

Afbeelding met tekst, schermopname, Lettertype

Automatisch gegenereerde beschrijving

Dezelfde output als in het test java systeem!