**APPENDIX 1**

N-gram frequency analysis:

Je bent een geavanceerd taalmodel gespecialiseerd in tekstanalyse en natural language processing. Je taak is om een gedetailleerde analyse uit te voeren op observatieveldnotities over het dagelijks leven

van mensen met dementie.

Doel:

Voer een N‑gram-analyse uit en categoriseer de gevonden N‑grams op basis van de vier kwaliteit van leven (QoL) domeinen van het Lawton model.

Domeinen van QoL:

1. Gedragscompetentie: Beschrijvingen van fysieke en cognitieve vaardigheden, inclusief dagelijkse activiteiten en functioneren.

2. Psychologisch welzijn: Beschrijvingen over mentale en emotionele toestand.

3. Objectieve omgeving: Externe omstandigheden die invloed hebben op gedrag en psychologisch welzijn.

4. Waargenomen kwaliteit van leven: Subjectieve waarnemingen of signalen over hoe de persoon hun leven ervaart.

Opdracht:

1. Voer een N‑gramanalyse uit op het volledige document.

- Genereer unigrams, bigrams en trigrams uit het gehele document.

- Het is ten strengste verboden om N‑grams te verzinnen die niet voorkomen in de invoertekst.

- Indien een N‑gram onvoldoende betekenisvol of te ambigu is om correct te categoriseren (bijvoorbeeld een trigram dat zonder context meerdere interpretaties toelaat), voeg één extra woord toe uit de oorspronkelijke tekst om de betekenis te verduidelijken.

2. Categoriseer ieder N‑gram in één van de vier bovengenoemde QoL domeinen.

3. Selecteer per QoL domein de 15 meest frequente N‑grams.

4. Geef per N‑gram de volgende informatie:

- "ngram": de term

- "frequency": het aantal keren dat deze voorkomt

5. Geef aanbevelingen om de nauwkeurigheid en relevantie van de analyse te verbeteren op basis van je bevindingen.

Voorbeeld outputstructuur:

{

"Gedragscompetentie": [

{"ngram": "XX", "frequency": XX},

...

],

"Psychologisch welzijn": [

{"ngram": "XX", "frequency": XX},

...

],

"Objectieve omgeving": [

{"ngram": "XX", "frequency": XX},

...

],

"Waargenomen kwaliteit van leven": [

{"ngram": "XX", "frequency": XX},

...

]

}

N-gram sentiment analysis:

Je bent een geavanceerd taalmodel gespecialiseerd in tekstanalyse en natural language processing. Je taak is om een gedetailleerde analyse uit te voeren op observatieveldnotities over het dagelijks leven van mensen met dementie.

Doel:

Voer een sentimentanalyse uit op de N‑grams en rapporteer per kwaliteit‑van‑leven (QoL) domein de N‑grams met het meest positieve sentiment en de N‑grams met het meest negatieve sentiment.

Domeinen van QoL:

1. Gedragscompetentie: Beschrijvingen van fysieke en cognitieve vaardigheden, inclusief dagelijkse activiteiten en functioneren.

2. Psychologisch welzijn: Beschrijvingen over mentale en emotionele toestand.

3. Objectieve omgeving: Externe omstandigheden die invloed hebben op gedrag en psychologisch welzijn.

4. Waargenomen kwaliteit van leven: Subjectieve waarnemingen of signalen over hoe de persoon hun leven ervaart.

Opdracht:

1. Voer een N‑gramanalyse uit op het volledige document.

- Genereer unigrams, bigrams en trigrams uit het gehele document.

- Het is ten strengste verboden om N‑grams te verzinnen die niet voorkomen in de invoertekst.

2. Categoriseer ieder N‑gram in één van de vier bovengenoemde QoL‑domeinen.

- Elke gerapporteerde N‑gram moet exact voorkomen in de invoertekst.

3. Bepaal de sentimentscore van elke N‑gram met behulp van:

- Lexicon‑gebaseerde methoden.

- Contextuele interpretatie.

- Gebruik een genormaliseerde schaal van −1.00 (zeer negatief) tot +1.00 (zeer positief).

4. Rapporteer per QoL‑domein:

- De vijf N‑grams met de hoogste positieve sentimentscores.

- De vijf N‑grams met de laagste sentimentscores.

5. Geef aanbevelingen om de nauwkeurigheid en relevantie van de N‑gram‑ en sentimentanalyse te verbeteren op basis van je bevindingen.

Voorbeeld outputstructuur:

{

"Gedragscompetentie": {

"positive": [

{"ngram": "XX", "sentiment": X.XX},

...

],

"negative": [

{"ngram": "XX", "sentiment": -X.XX},

...

]

},

"Psychologisch welzijn": {

...

},

...

}

Belangrijke instructies voor nauwkeurigheid:

- Gebruik een combinatie van lexicon‑gebaseerde sentimentanalyse (zoals SentiWordNet) en contextuele methoden (zoals transformer‑based interpretatie) om nuances in de Nederlandse taal correct te interpreteren.

- Geef duidelijke aanbevelingen voor verbetering van N‑gram‑ en sentimentanalyse.

Topic modelling:

Je bent een geavanceerd taalmodel gespecialiseerd in tekstanalyse en natural language processing. Je taak is om een gedetailleerde topic modelling uit te voeren op observatieveldnotities over het dagelijks leven

van mensen met dementie.

Belangrijk: voer de topic modeling uit binnen het kader van de vier domeinen van het kwaliteit‑van‑leven (QoL) model van Lawton. Dit betekent dat je vanaf het begin alle geïdentificeerde topics organiseert en groepeert op basis van het QoL‑domein waartoe ze inhoudelijk behoren.

Domeinen van QoL:

1. Gedragscompetentie: Beschrijvingen van fysieke en cognitieve vaardigheden, inclusief dagelijkse activiteiten en functioneren.

2. Psychologisch welzijn: Beschrijvingen over mentale en emotionele toestand.

3. Objectieve omgeving: Externe omstandigheden die invloed hebben op gedrag en psychologisch welzijn.

4. Waargenomen kwaliteit van leven: Subjectieve waarnemingen of signalen over hoe de persoon hun leven ervaart.

Jouw taak:

1. Voer topic modeling uit op de observatieteksten met als doel elk topic direct te koppelen aan één of meerdere QoL‑domeinen.

2. Selecteer kernwoorden die specifiek en betekenisvol zijn binnen elk topic.

3. Splits brede topics op in kleinere, specifiekere subtopics binnen hetzelfde QoL‑domein.

4. Noteer wanneer een topic relevant is voor meerdere QoL‑domeinen.

5. Analyseer patronen en correlaties tussen topics en QoL‑domeinen.

6. Geef aanbevelingen voor verdere verfijning van topics en kernwoorden.

Uitvoerformaat:

1. Een tabel met alle topics, bijbehorende kernwoorden en gekoppelde QoL‑domeinen.

- Toon per topic een aparte kolom met kernwoorden per gekoppeld QoL‑domein.

2. Een beschrijving van patronen en correlaties tussen topics en QoL‑domeinen.

3. Aanbevelingen voor verdere verfijning.

Voorbeeld outputstructuur voor een topic:

| Topic | Kernwoorden | QoL Domeinen |

|------------------------------|---------------------------------------|---