

AUTOMATISCHE SENTIMENTANALYSE IN TEKSTEN

Input = hetzelfde voor Lexicongebaseerd en Machine Learning

Hier kom ik nog eens terug !

De bediening is verschikkelijk, zo traag!

Een mooi, gezellig restaurant in hartje Gent ❤️

Ze sturen me van het kastje naar de muur ...

Preprocessing Inputtekst (instantie) wordt voorbereid voor verwerking = hetzelfde voor Lexicongebaseerd en Machine Learning

Lowercasen	Alle tekst omzetten naar kleine letters ("Gezellig resto in Gent" → "gezellig resto in gent")
Tokeniseren	Zinnen in woorden splitsen ("gezellig resto in gent" → "gezellig", "resto", "in", "gent")
Lemmatiseren	Woorden vervangen door lemma/woordenboekvorm ("is" --> "zijn", "hartje" → "hart")
Part-of-speech tagging	Automatisch woordsoorten toekennen ("zijn" = werkwoord, "hart" = zelfstandig naamwoord)

Lexicongebaseerd (woordenlijsten)

Woordenlijsten + matching

Een **mooi, gezellig** restaurant in hartje Gent ❤️

De bediening is **verschikkelijk**, zo traag!

POS/NEG/NEU woorden worden gematcht met woordenlijsten en een globale score wordt berekend voor de instantie

Classificatie

Een **mooi, gezellig** restaurant in hartje Gent ❤️ **+2**

De bediening is **verschikkelijk**, zo traag! **-1**

Output

Sentiment voorspelling	Testdata	Sentimentvoorspelling
POS	Een mooi, gezellig restaurant in hartje Gent ❤️	POS
NEUTRAL	Hier kom ik nog eens terug !	POS
NEG	De bediening is verschikkelijk, zo traag!	NEG
NEUTRAL	Ze sturen me van het kastje naar de muur ...	NEG

Machine Learning (lerende aanpak)

Feature engineering

Taaltechnologen voeden algoritmes met trainingsdata

Data wordt omgezet in feature vectoren

```

2:1, 37:1, 94:1, 124:1, 175:4, 208:0.7 [positief]
5:1, 29:1, 74:1, 142:1, 175:0, 208:0 [positief]
12:1, 23:1, 55:1, 89:1, 123:1, 176:2, 208:-1 [negatief]
21:1, 48:1, 53:1, 112:1, 175:0, 176:-, 208:-0.3 [negatief]
    
```

2:1, 3:7, 74:1, 124:5, 205:0,7 [positief|negatief|neutraal]?

Testdata wordt omgezet naar featurevectoren en het getraind model genereert een voorspelling

Een mooi, gezellig restaurant in hartje Gent ❤️