

Interactie van methoden en hypothesen

Serge ter Braake, Antske Fokkens en Niels Ockeloën

VU Amsterdam



Discussie over het delen van informatie en mogelijke biases in onze test case.

- Kennis: wat zou onze test case voor evaluatie kunnen verpesten?
- Hebben we een pilot case nodig, die we eventueel mogen verpesten? (zo ja, welke?)
- Waar we verder nog tijd voor hebben....

■ Kennis

- Wie mag wat wanneer weten?
- Wie moet wat wanneer weten (om te kunnen werken of om methode te controleren)

■ Vraagstelling

- Gedetailleerdheid (voorkennis van geschiedkundige vs. filter in methoden)

Evaluatie

Dit is afhankelijk van de evaluatie!

- Wat willen we precies evalueren?
 - betrouwbaarheid van informatie (waar het bioportaal de gold standaard is)
 - **de mogelijkheid informatie te kunnen combineren om een hypothese te beantwoorden (gaten in database opvullen/ vraag in deelvragen opdelen)**
 - **de mogelijkheid patronen te ontdekken of nieuwe informatie af te leiden**
 - gebruikersvriendelijkheid, vereiste voorkennis (geschiedkundig, over dit project) om de demonstrator te gebruiken

Taakverdeling

Serge	Niels	Antske
manual case vraagstelling, deelvragen	applicatie verwerkt informatie	informatie uit tekst tekst analyse

- **Hypothese:** In de VS komen mensen uit een bepaalde beroepsgroep voornamelijk uit de oostkust.
- **Data:** Veel plaatsnamen die in meerdere staten voorkomen (e.g. *Freemont*, *Springfield*)
- **Mogelijke bias:**
 - Coördinaten van Washington DC worden gebruikt om te kijken welke *Freemont* bedoeld wordt.
⇒ Duidelijke bias naar de hypothese (kan ook zonder voorkennis van hypothese zijn ingebouwd)
 - Geografische coördinaten worden gebruikt om te kijken welke *Freemont* bedoeld wordt.
⇒ Mogelijke bias voor deel van de hypothese en mogelijk versterkend effect op uitkomst (hypothese klopt, maar met lichte tendens bias maakt er een sterke tendens van)

- In principe: features mogen niet uit de gold standard gehaald worden (dus geen kennis tijdens het ontwikkelen van de applicatie/tekst analyse tools).
- In dit geval: ook als features niet uit de gold standard komen, zijn ze verboden als ze een bias introduceren
 - de *hypothese* is geen gold standard
 - het is geen training/test geval: we willen ook geen bias van andere cases in onze methoden
- Noodzakelijk:
 - Goede kennis over de methode, m.n. welke input wordt **in elk stadium** gebruikt
 - Documentatie van mogelijke invloeden op uitkomst
 - Voor publicaties: controle doorgang van alle elementen om mogelijke bias te voorkomen

- Gedetailleerde (deel)vragen naar standaard informatie in de database kunnen bias uit de applicatie voorkomen
- **Maar** in deze vragen kan een bias van de historicus zitten
 - Wie gebruiken de demonstrator?
 - Welke informatie komt van gebruikers, welke van ons?

- Kan overleg tussen Niels, Serge en Antske de evaluatie verpesten?
- Moeten we een *pilot case* definiëren, waar dit eventueel niet erg is?