Probabilistic Programming for Voucher Information Extraction
Preliminary Practical Experiences
Al-Sibahi, Ahmad Salim; Hamelryck, Thomas Wim; Henglein, Fritz

Publication date:
2018

Document version
Early version, also known as pre-print

Citation for published version (APA):
Probabilistic Programming for Voucher Information Extraction

Finding Features w/Keywords

Features are usually located around identifying keywords. Keywords can be positive or negative depending on the feature to be found.

Finding Features with Keywords

Probabilistic model below tries to infer a latent score \( r \) from the vector of observed angles \( \theta^+ \) and distances \( d^+ \) from positive keywords to potential target features.

For 1000 sample vouchers we achieved 21 topics, and our next goal is to rely on these topics to construct more precise feature extraction models.

Practical Experiences

Sampling

- Ease of use ✓
- Precision ✓
- Scalability ✗

Variational Inference

- Scalability ✓
- Set-up ✗
- Precision ✓

GPU Support

- Discrete Latents ✗
- Ease of use ✓
- Precision ✗

The first author would like to thank Dan Rose and the Scanner Team at Skanned.com (Bjørn Kaae, Toke Reines, Danni Dromi) for discussions and feedback on the work!