MEDSTONE

Development & evaluation of a NER model for recognition of important study characteristics

SURUS

Casper Peeters
Product development scientist
MSc Drug Innovation
casper.peeters@medstone.nl

MEDSTONE

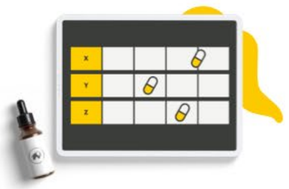
Company Structure

MEDSTONE HOLDING

SCIENCE

State of the Art Artifical Intelligence solutions for healthcare

PHARMACISTS PLATFORM



A fast growing pharmaceutical community

WHOLESALE (FRESCO FARMA)

Globally operating pharmaceutical wholesale division

SURUS assists systematic reviewers with screening literature

MEDSTONE

Systematic Literature Review (SLR)

- Crucial in clinical decision making
- Selection of the complete scientific body of evidence
- Highly laborious, manually performed by experts

Literature selection

- Literature is selected based on descriptive characteristics: “PICOS”
- PICOS determine relevance of scientific study to reviewers’ research question
- SURUS automatically extracts PICOS elements

Annotation strategy is designed for extraction of PICOS elements

Annotation for Named Entity Recognition (NER)

- 26 annotation subcategories
- Annotated by experts in the field
- Inter annotator agreement of 0.78

Annotation for segment prediction

- Automatically labelled based on author segmentation

In this **16 - week** , **randomized** , **open - label** **trial** , participants (mean : **45.8 years old** , **A 1 C 8.4 %** , **fasting plasma glucose [FPG] 9.9 mmol / L** , **BMI 26.9 kg / m (2)**) received **subcutaneous injections** of **IDeg** (A) (**600** **µ mol / L** ; **n = 59**) , **IDeg** (B) (**900** **µ mol / L** ; **n = 60**) , or **insulin glargine** (**IGlar** ; **n = 59**)

Annotations for the NER prediction model

Annotations for the segmentation model

Corpus of annotated PubMed abstracts

MEDSTONE

Dataset for NER
- 850 scientific abstracts (80:10:10 split)
- 7 major disease categories
- 80,908 manual annotations

Dataset for NER
- 10,000 scientific abstracts (80:20 split)

NER category	# subcategories	PICOS relevance	# annotations
Drug	6	Intervention (I), Comparison (C)	19,620
Methodology	7	Population (P), Outcome (O), Study design (S)	15,251
Parameter	3	Outcome (O)	8247
Result	6		27,878
Therapy	2	Intervention (I), Comparison (C)	6839
Identifier	1		710
Disease	1	Population (P)	2363

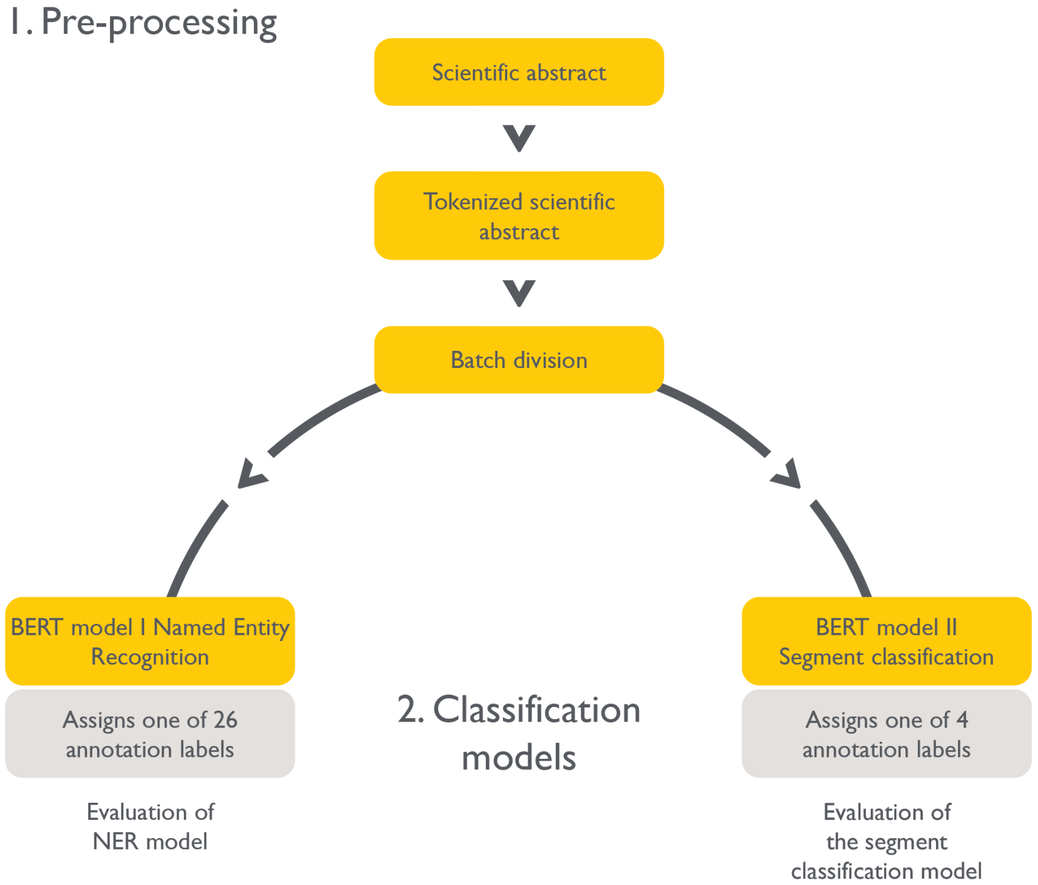
SURUS consists of 2 separate, but complementary, BERT models

MEDSTONE

Two separate BERT models
- Sliding-window approach
- Evaluation of BERT and BioBERT

NER model
- Prediction of BILUO and NE labels
- BILU labels used for quality measurement

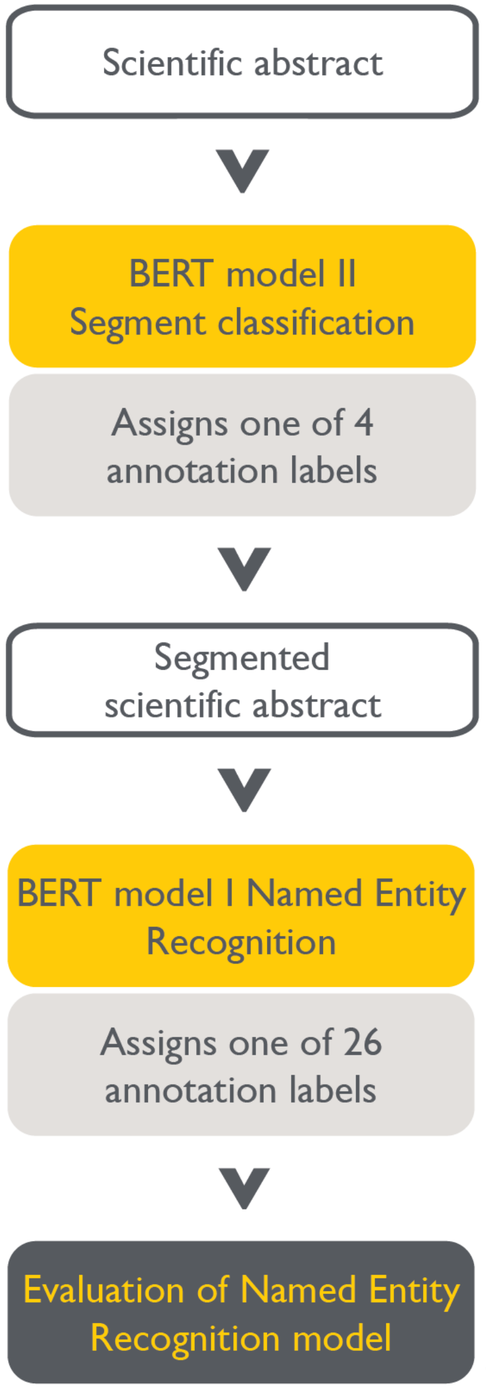
Segmentation model
- Predicts location of tokens in abstract
- Introduction, Methods, Results or Conclusion
- Rule-based separation of Aims and Introduction



Segment: feature incorporation approach

Incorporation of segment ID in NER prediction

- Addition of segment label as a feature (Introduction, Aims, Methods, Results, Conclusion)



SURUS performs well and consistently at NER

NER model
- Overall F1 of >0.94 (BioBERT)
- BioBERT improves performance over BERT

Segmentation model
- Accurate prediction of interventional study headers

Combination of models
- Adding a segment feature to NER training does not improve model quality

SURUS is a practical tool for SLR

Conclusion and SURUS application

- Reliable recognition of most PICOS elements
- Potential to speed up the SLR screening process

SURUS test case

- Selection of relevant abstracts
- Based on completed Cochrane review
- NE label-based filters

Test case results

- 89.4% of irrelevant articles excluded
- No exclusion of relevant articles