SECTOR: A Neural Model for Coherent Topic Segmentation and Classification

Sebastian Arnold, Rudolf Schneider, Philippe Cudré-Mauroux*, Felix A. Gers, Alexander Löser

sarnold@beuth-hochschule.de
@sebastianarnold

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

Beuth University of Applied Sciences
Berlin, Germany

**eXascale Infolab*
University of Fribourg
Fribourg, Switzerland

Transactions of the Association for Computational Linguistics (TACL) Vol.7

ACL 2019, Florence, Italy

29.07.2019

Challenge: understand the topics and structure of a document

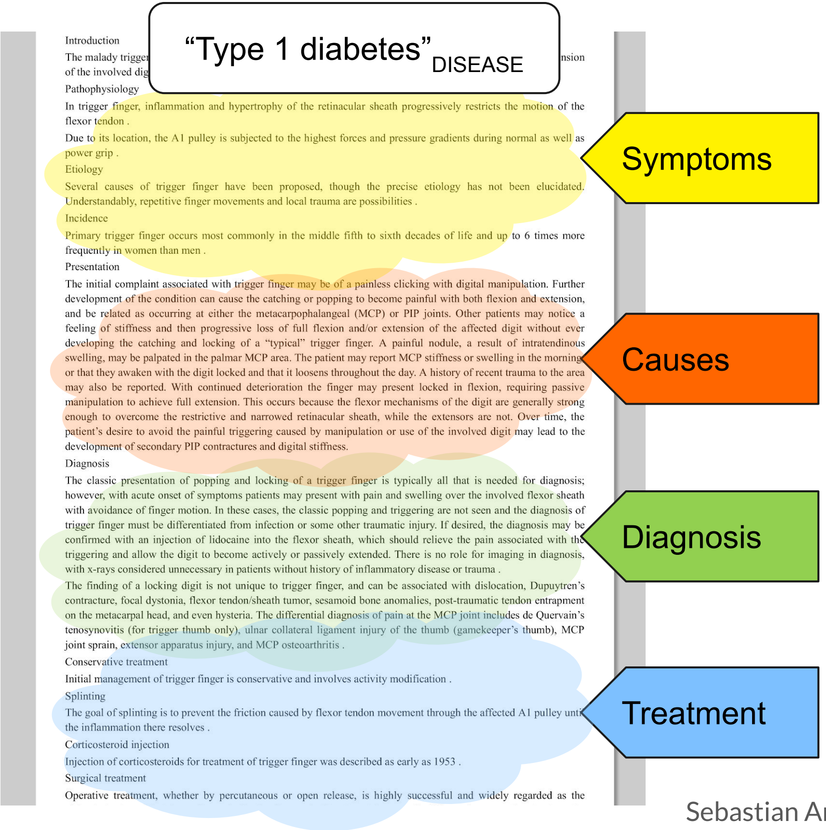
How can we represent a document with respect to the author's emphasis?

➔ **topical** information [Ma18] (e.g. semantic class labels)

➔ **structural** information [Ag09, Gla16] (e.g. coherent passages)

➔ in **latent vector space** [Le14, Bha16] (i.e. distributional embedding)

➔ required for **TDT**, **QA** & **IR** downstream tasks [All02, Di07, Coh18]


Introduction

The malady trigger... ...nsion of the involved dig...

Pathophysiology

In trigger finger, inflammation and hypertrophy of the retinacular sheath progressively restricts the motion of the flexor tendon .

Due to its location, the A1 pulley is subjected to the highest forces and pressure gradients during normal as well as power grip .

Etiology

Several causes of trigger finger have been proposed, though the precise etiology has not been elucidated. Understandably, repetitive finger movements and local trauma are possibilities .

Incidence

Primary trigger finger occurs most commonly in the middle fifth to sixth decades of life and up to 6 times more frequently in women than men .

Presentation

The initial complaint associated with trigger finger may be of a painless clicking with digital manipulation. Further development of the condition can cause the catching or popping to become painful with both flexion and extension, and be related as occurring at either the metacarpophalangeal (MCP) or PIP joints. Other patients may notice a feeling of stiffness and then progressive loss of full flexion and/or extension of the affected digit without ever developing the catching and locking of a "typical" trigger finger. A painful nodule, a result of intratendinous swelling, may be palpated in the palmar MCP area. The patient may report MCP stiffness or swelling in the morning or that they awaken with the digit locked and that it loosens throughout the day. A history of recent trauma to the area may also be reported. With continued deterioration the finger may present locked in flexion, requiring passive manipulation to achieve full extension. This occurs because the flexor mechanisms of the digit are generally strong enough to overcome the restrictive and narrowed retinacular sheath, while the extensors are not. Over time, the patient's desire to avoid the painful triggering caused by manipulation or use of the involved digit may lead to the development of secondary PIP contractures and digital stiffness.

Diagnosis

The classic presentation of popping and locking of a trigger finger is typically all that is needed for diagnosis; however, with acute onset of symptoms patients may present with pain and swelling over the involved flexor sheath with avoidance of finger motion. In these cases, the classic popping and triggering are not seen and the diagnosis of trigger finger must be differentiated from infection or some other traumatic injury. If desired, the diagnosis may be confirmed with an injection of lidocaine into the flexor sheath, which should relieve the pain associated with the triggering and allow the digit to become actively or passively extended. There is no role for imaging in diagnosis, with x-rays considered unnecessary in patients without history of inflammatory disease or trauma .

The finding of a locking digit is not unique to trigger finger, and can be associated with dislocation, Dupuytren's contracture, focal dystonia, flexor tendon/sheath tumor, sesamoid bone anomalies, post-traumatic tendon entrapment on the metacarpal head, and even hysteria. The differential diagnosis of pain at the MCP joint includes de Quervain's tenosynovitis (for trigger thumb only), ulnar collateral ligament injury of the thumb (gamekeeper's thumb), MCP joint sprain, extensor apparatus injury, and MCP osteoarthritis .

Conservative treatment

Initial management of trigger finger is conservative and involves activity modification .

Splinting

The goal of splinting is to prevent the friction caused by flexor tendon movement through the affected A1 pulley until the inflammation there resolves .

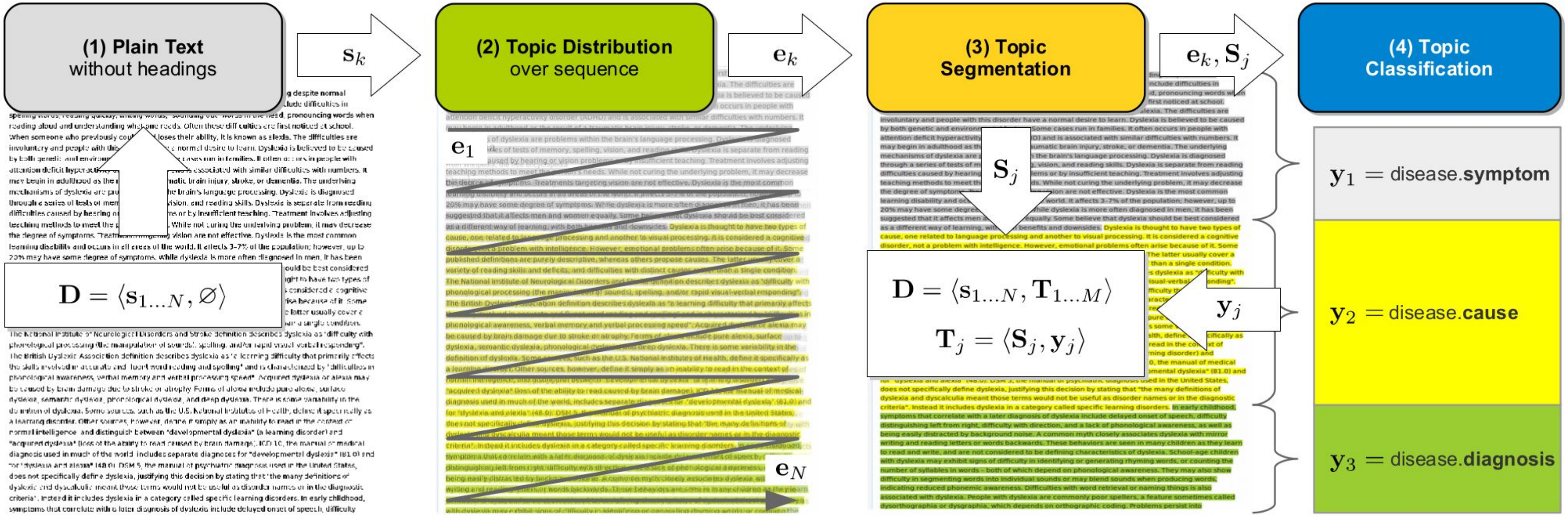
Corticosteroid injection

Injection of corticosteroids for treatment of trigger finger was described as early as 1953 .

Surgical treatment

Operative treatment, whether by percutaneous or open release, is highly successful and widely regarded as the

Task: split a document into coherent sections with topic labels



We aim to detect **topics** in a document that are expressed by the author as a **coherent sequence of sentences** (e.g., a passage or book chapter).

WikiSection: Wiki authors provide topics as section headings

Contents [hide]

1 Signs and symptoms
- 1.1 Diabetic emergencies
- 1.2 Complications

2 Causes
- 2.1 Type 1
- 2.2 Type 2
- 2.3 Gestational diabetes
- 2.4 Maturity onset diabetes of the young
- 2.5 Other types

3 Pathophysiology

4 Diagnosis

5 Prevention

6 Management
- 6.1 Lifestyle
- 6.2 Medications
- 6.3 Surgery
- 6.4 Support

7 Epidemiology

8 History
- 8.1 Etymology

9 Society and culture
- 9.1 Naming

10 Other animals

11 Research

12 References

13 Further reading

14 External links

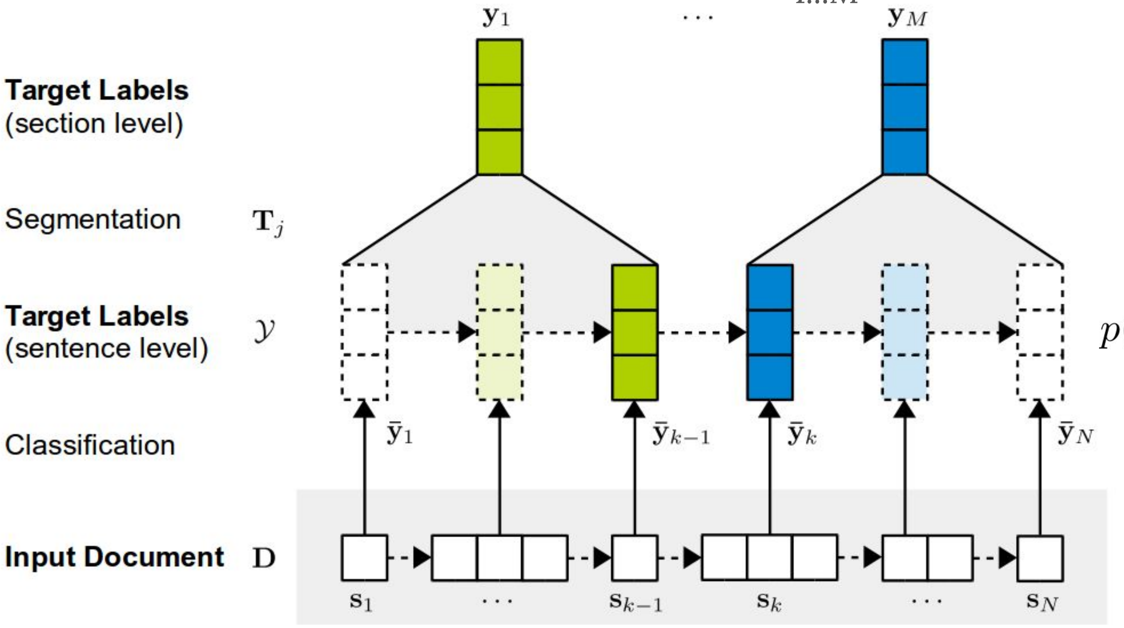
en_disease (27)	de_disease (25)
treatment	therapie
symptom	diagnose
diagnosis	symptom
cause	ursache
classification	kategorisierung
epidemiology	verlauf
history	epidemiologie
prognosis	geschichte
management	prognose
pathophysiology	praevalenz
mechanism	vorbeugung
prevention	fauna
research	terminologie
genetics	pathologie
tomography	definition
culture	klinik
etymology	komplikation
infection	genetik
fauna	infektion
risk	risiko
pathology	forschung
surgery	geographie
screening	mensch
medication	organe
geography	sonstiges
complication	
other	

en_disease	de_disease	en_city	de_city
3.6k English articles	2.3k German articles	19.5k English articles	12.5k German articles
8.5k headings	6.1k headings	23.0k headings	12.2k headings
27 topics (94.6%)	**25 topics** (89.5%)	**30 topics** (96.6%)	**27 topics** (96.1%)

https://github.com/sebastianarnold/WikiSection

SECTOR sequential prediction approach

- Transform a document of **N** sentences $\mathbf{s}_{1...N}$ into N topic distributions $\bar{\mathbf{y}}_{1...N}$
- Predict **M** sections $\mathbf{T}_{1...M}$ based on coherence of the network's weights
- Assign section-level topic labels $\mathbf{y}_{1...M}$



Number and length of sections is unknown!

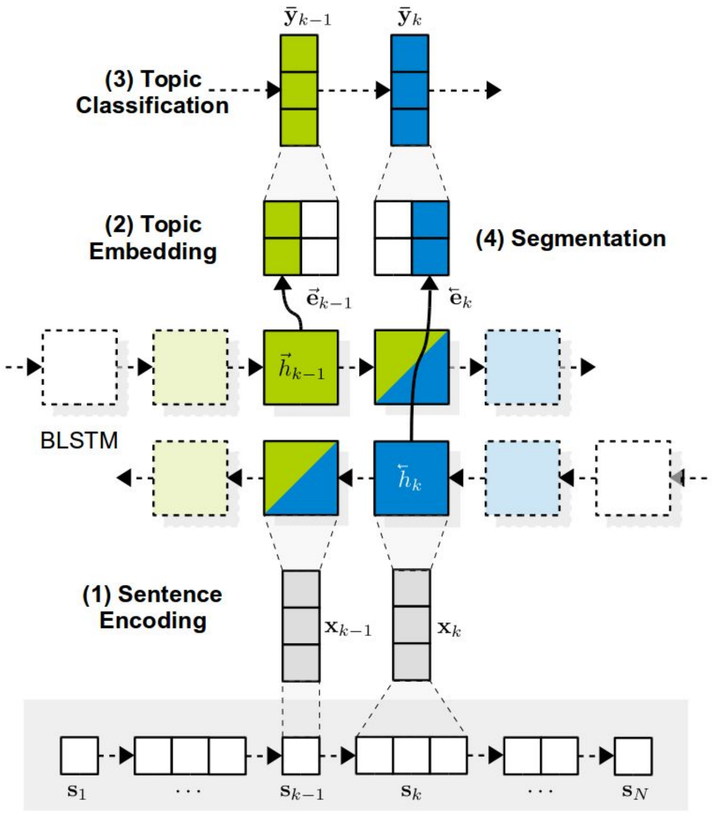
$$p(\bar{\mathbf{y}}_1, \dots, \bar{\mathbf{y}}_N \mid \mathbf{D}) = \prod_{k=1}^{N} p(\bar{\mathbf{y}}_k \mid \mathbf{s}_1, \dots, \mathbf{s}_N)$$

Network architecture (0/4) – Overview

Objective: maximize the log likelihood of model parameters Θ per document on sentence-level

$$\bar{\mathcal{L}}(\Theta) = \sum_{k=1}^{N} \log p(\bar{\mathbf{y}}_k \mid \mathbf{s}_1, \dots, \mathbf{s}_N; \Theta)$$

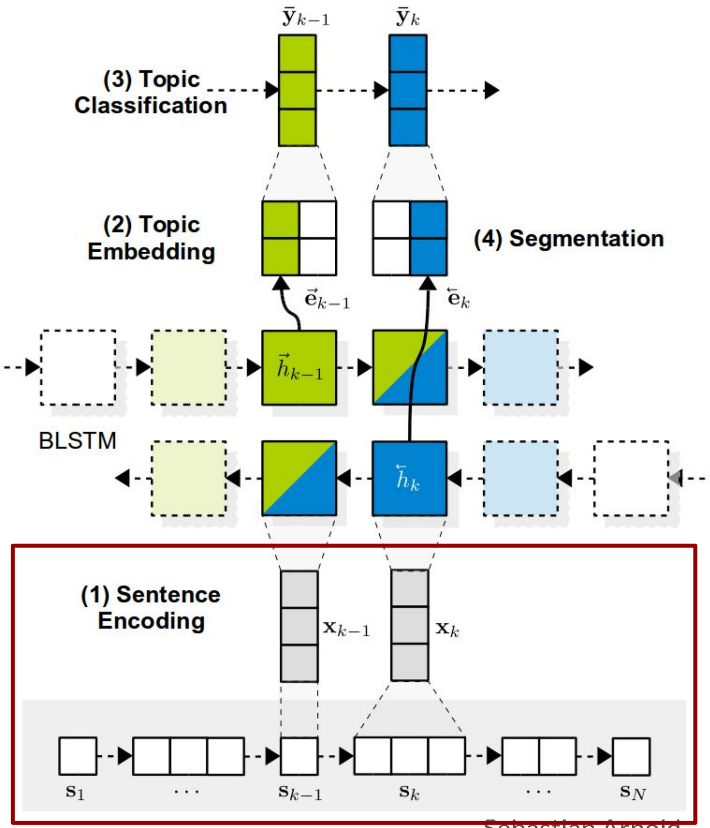
- Requires the entire document as input
- Long range dependencies
- Focus on sharp distinction at topic shifts



Network architecture (1/4) – Sentence encoding

Input: Vector representation of a full document

- Split text into sequence of sentences $\mathbf{s}_{1...N}$
- Encode sentence vectors $\mathbf{x}_{1...N}$ using
  - Bag-of-words (~56k english words)
  - Bloom filter (4096 bits) [Se17] or
  - Pre-trained sentence embeddings [Mik13, Aro17] (128 dim)
- Use sentences as time-steps



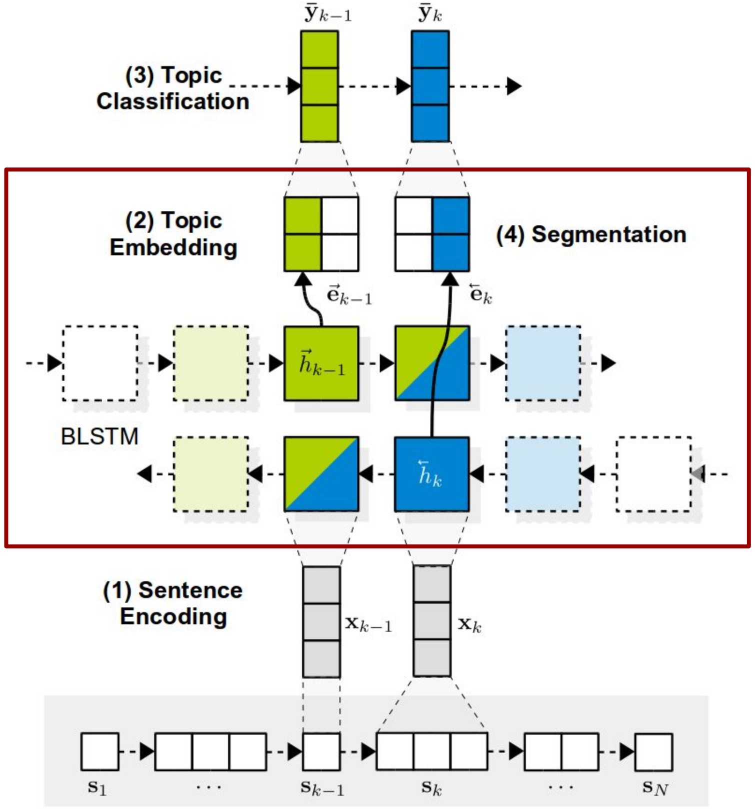
Network architecture (2/4) – Topic embedding

Encoder: Bidirectional Long Short-Term Memory (BLSTM) [Ho97, Ge00, Gra12] + dense embedding layer

- independent *fw* and *bw* parameters $\vec{\Theta}, \overleftarrow{\Theta}$ helps to sharpen left/right context
- embedding layer captures latent topics

$$\mathcal{L}(\Theta) = \sum_{k=1}^{N} \left(\log p(\bar{\mathbf{y}}_k \mid \mathbf{x}_{1\ldots k-1}; \vec{\Theta}, \Theta') + \log p(\bar{\mathbf{y}}_k \mid \mathbf{x}_{k+1\ldots N}; \overleftarrow{\Theta}, \Theta') \right)$$

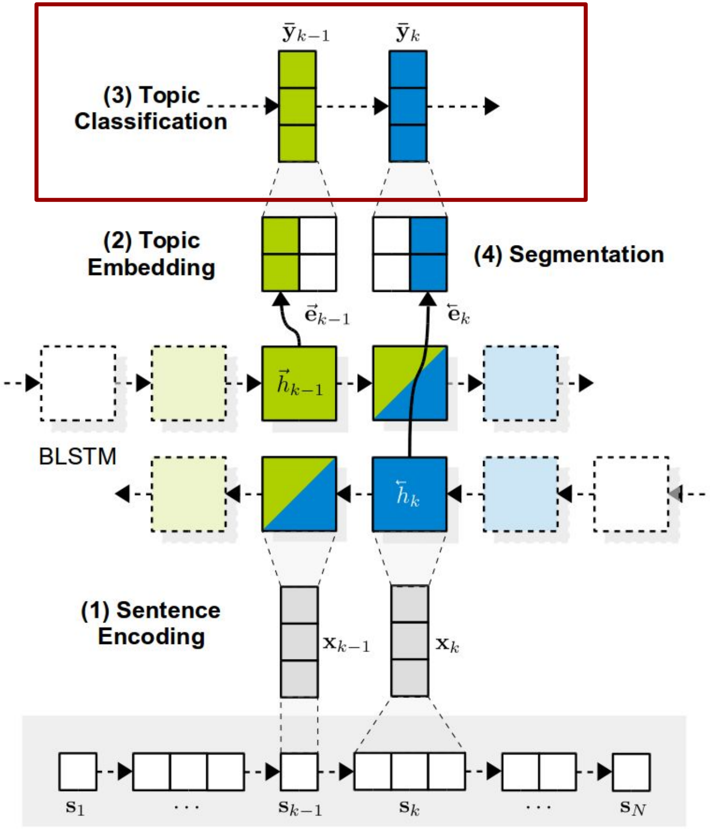
- 2x256 LSTM cells, 128 dim embedding layer, 16 docs per batch, 0.5 dropout, ADAM opt.



Network architecture (3/4) – Topic classification

Output layer: Classification

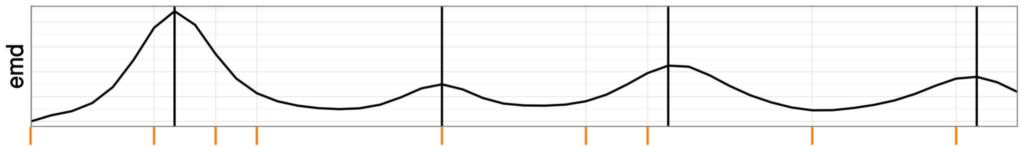
- Decodes target probabilities
- Human-readable topic labels for 2 Tasks:
  - **topic classes $\bar{\mathbf{y}}_{1...N}$** (25–30 topics)
    disease.symptom
    $$\hat{\bar{\mathbf{y}}}_k = \mathrm{softmax}(W_{ye}\vec{\mathbf{e}}_k + W_{ye}\overleftarrow{\mathbf{e}}_k + b_y)$$
  - **headline words $\bar{\mathbf{z}}_{1...N}$** (1.5–2.8k words)
    [signs, symptoms]
    $$\hat{\bar{\mathbf{z}}}_k = \mathrm{sigmoid}(W_{ze}\vec{\mathbf{e}}_k + W_{ze}\overleftarrow{\mathbf{e}}_k + b_z)$$



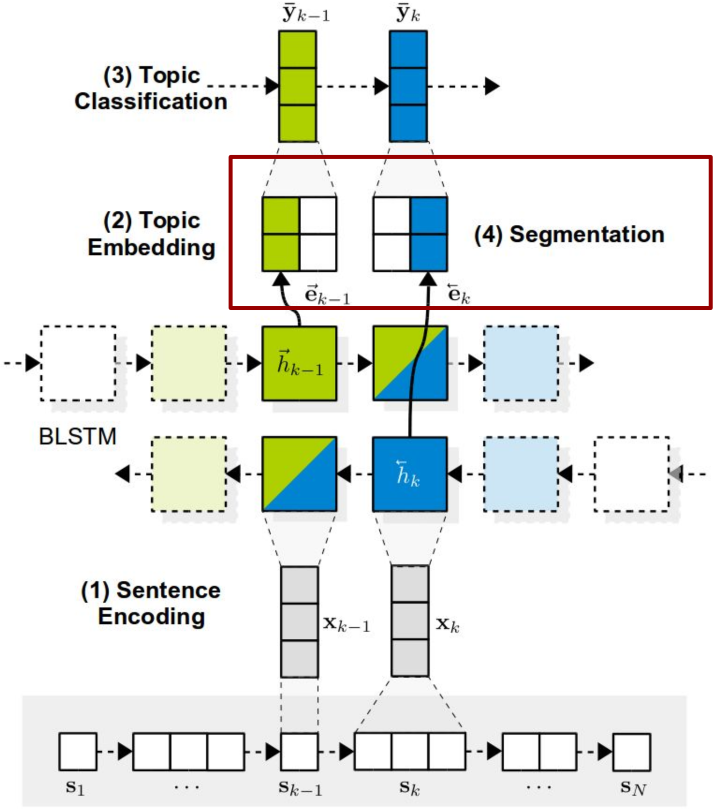
Network architecture (4/4) – Segmentation

Segmentation: based on topic coherence

- deviation d_k: stepwise "movement" of the embedding between two sentences

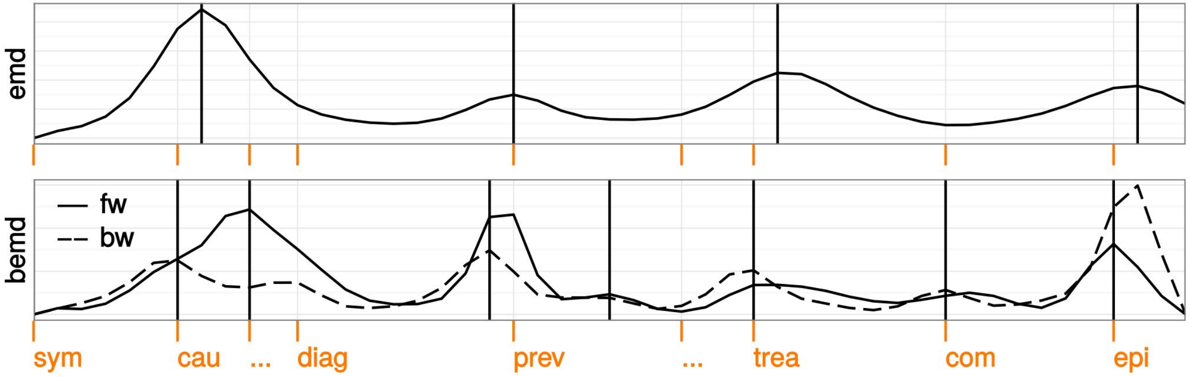


$$\mathbf{d}'_k = \sqrt{\cos(\vec{\mathbf{e}}'_{k-1}, \vec{\mathbf{e}}'_k) \cdot \cos(\overleftarrow{\mathbf{e}}'_k, \overleftarrow{\mathbf{e}}'_{k+1})}$$



Coherent segmentation using edge detection

We use the topic embedding deviation (emd) d_k to start new segments on peaks.



- Idea adapted from image processing: we apply *Laplacian-of-Gaussian edge detection* [Zi98] to find local maxima on the emd curve
- Steps: dimensionality reduction (PCA), Gaussian smoothing, local maxima
- Bidirectional deviation (bemd) on *fw* and *bw* layers allows for sharper separation

Experiments with 20 different models on 8 datasets

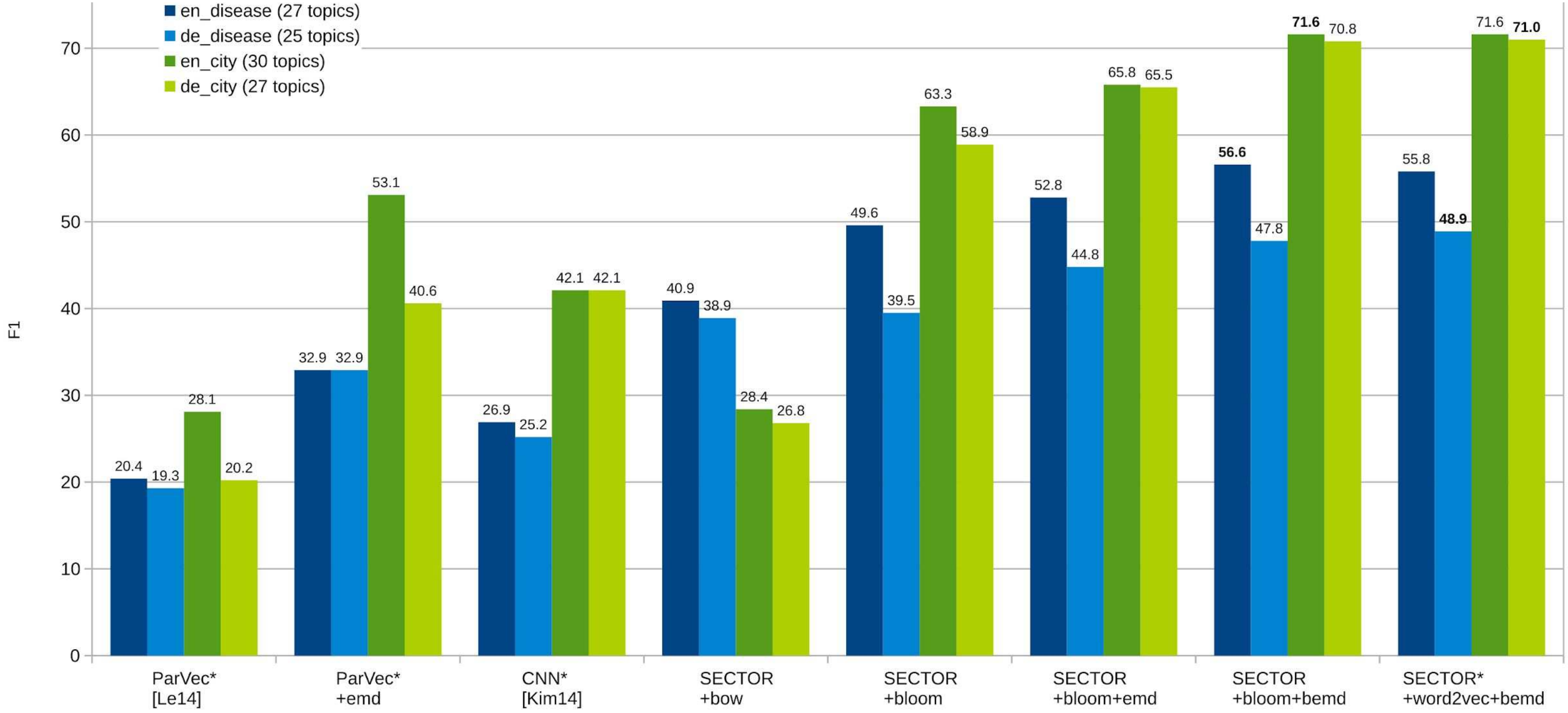
dataset	articles	article type	headings	topics	segments
WikiSection	38k train/test	German/English diseases and cities	X	X	X
Wiki-50 [Kosh18]	50 test	English generic	X		X
Cities/Elements [Chen09]	130 test	English cities and chemicals (lowercase)			X
Clinical Textbook [Eis08]	227 test	English clinical	X		X

Sentence Classification Baselines: ParVec [Le14], CNN [Kim14]

Segmentation Models: C99 [Choi00], TopicTiling [Rie12], BayesSeg [Eis08], TextSeg [Kosh18]

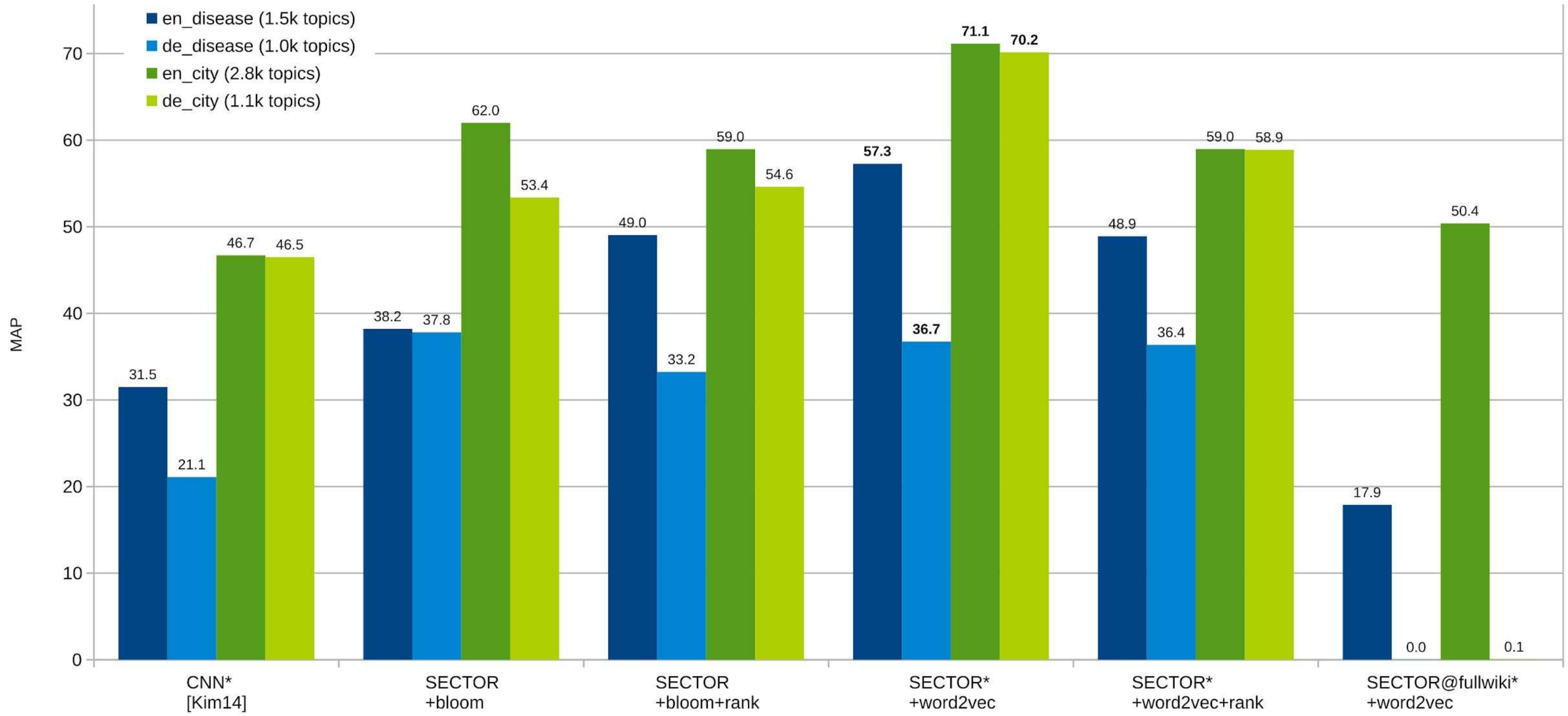
Experiment 1: segmentation and single-label classification

Segment on sentence-level and assign one of 25–30 supervised topic labels (F1)

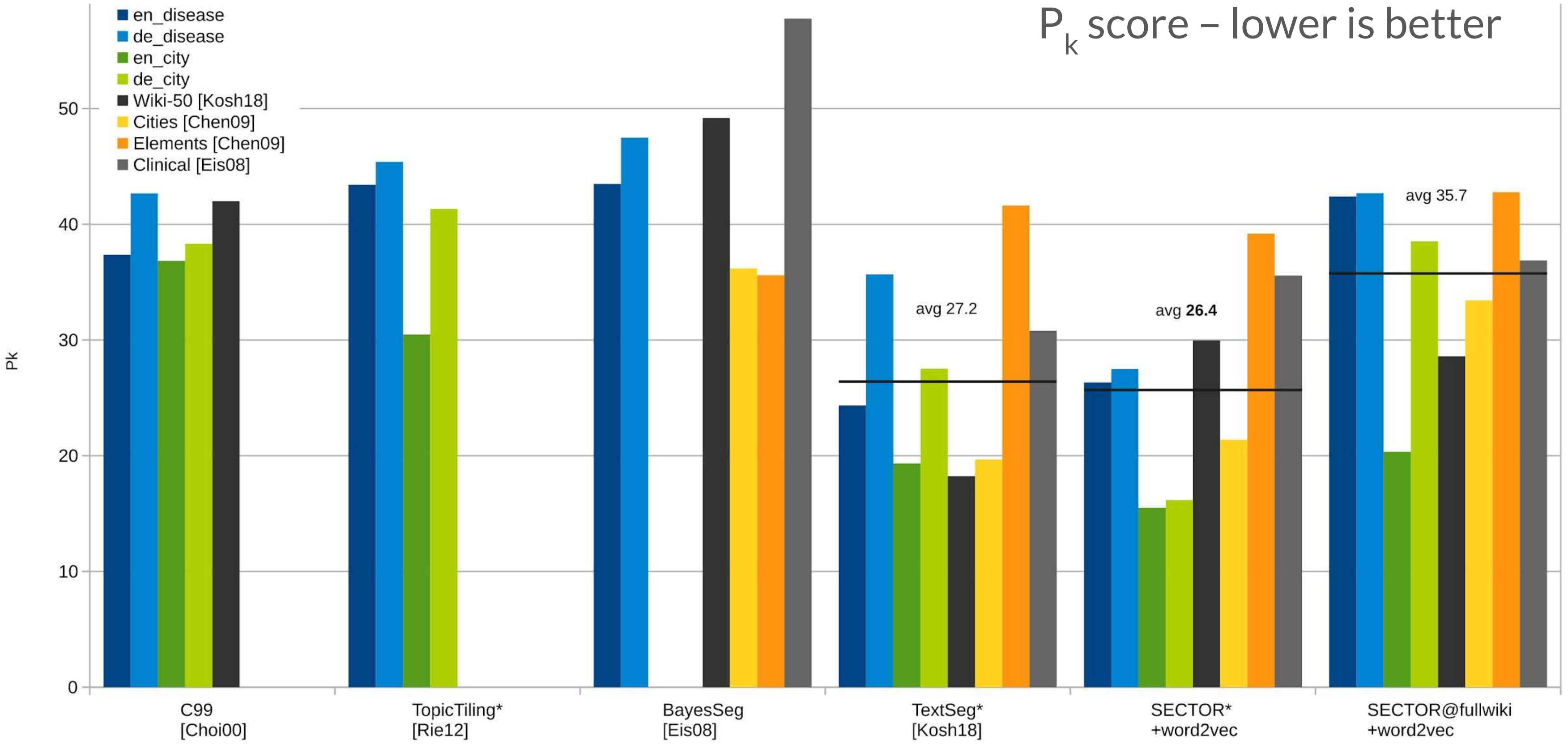


Experiment 2: segmentation and multi-label classification

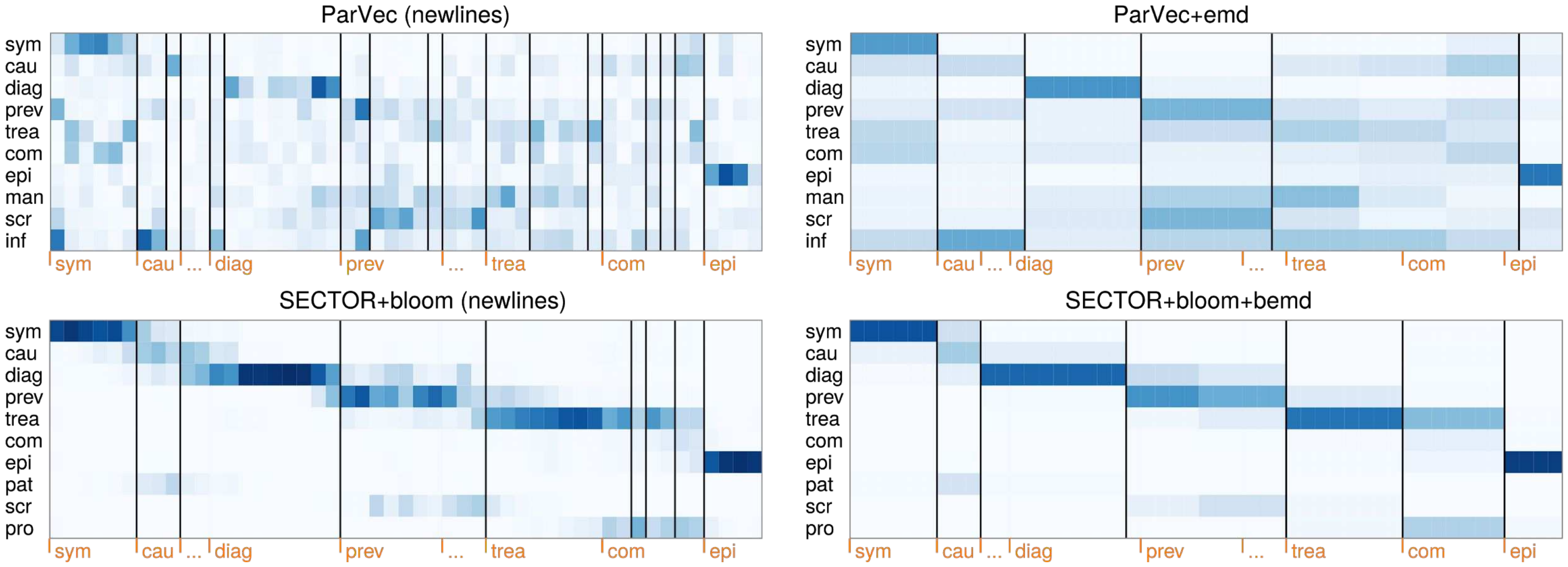
Segment on sentence-level and rank 1.0k–2.8k ‘noisy’ topic words per section (MAP)



Experiment 3: segmentation without topic prediction (cross-dataset)

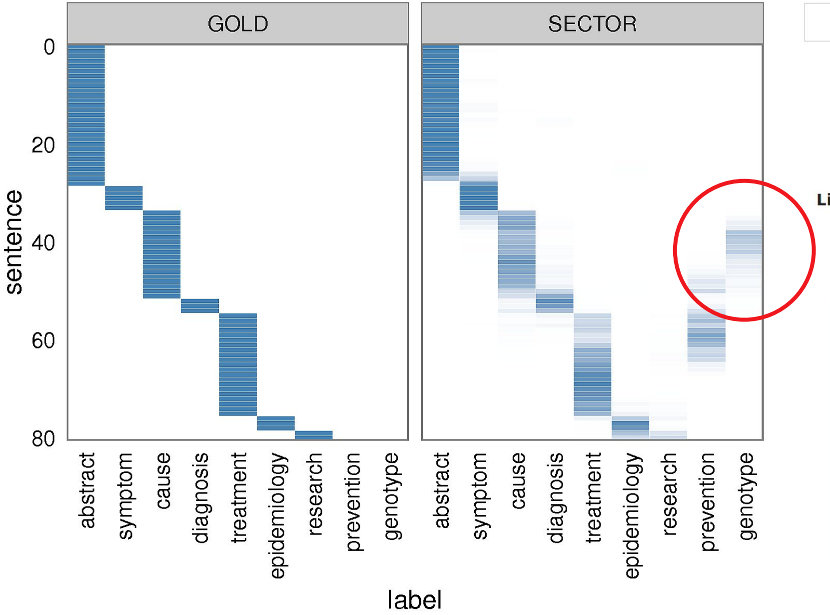


Insights: SECTOR captures topic distributions coherently



Topic predictions on sentence level – top: ParVec [Le14] – *bottom*: SECTOR
Segmentation – left: newlines in text (\n) – *right*: embedding deviation (emd)

SECTOR prediction on par with Wiki authors for “dermatitis”

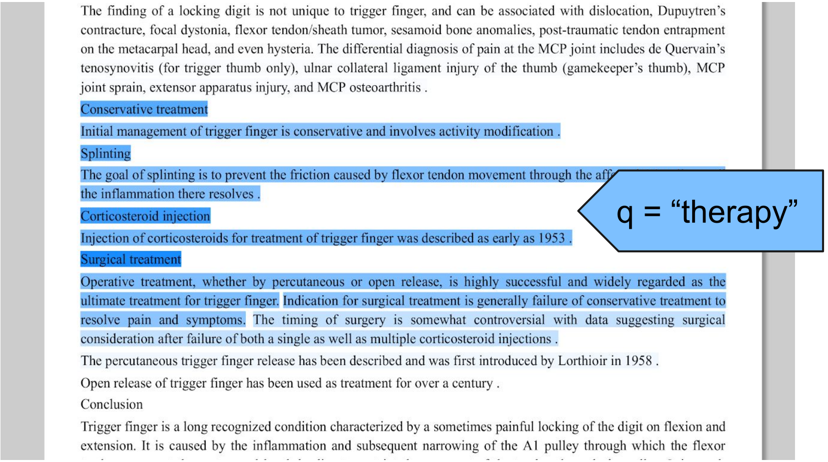


Source: https://en.wikipedia.org/w/index.php?title=Atopic_dermatitis&diff=786969806&oldid=772576326

Conclusion and future work

SECTOR is designed as a building block for **document-level knowledge representation**

- Reading sentences in document context is an important step to **capture both topical and structural information**
- Training the topic embedding with distant-supervised **complementary labels** improves performance over self-supervised word embeddings
- **In future work**, we aim to apply the topic embedding for unsupervised passage retrieval and QA tasks


The finding of a locking digit is not unique to trigger finger, and can be associated with dislocation, Dupuytren's contracture, focal dystonia, flexor tendon/sheath tumor, sesamoid bone anomalies, post-traumatic tendon entrapment on the metacarpal head, and even hysteria. The differential diagnosis of pain at the MCP joint includes de Quervain's tenosynovitis (for trigger thumb only), ulnar collateral ligament injury of the thumb (gamekeeper's thumb), MCP joint sprain, extensor apparatus injury, and MCP osteoarthritis .

Conservative treatment

Initial management of trigger finger is conservative and involves activity modification .

Splinting

The goal of splinting is to prevent the friction caused by flexor tendon movement through the aff
the inflammation there resolves .

Corticosteroid injection

Injection of corticosteroids for treatment of trigger finger was described as early as 1953 .

Surgical treatment

Operative treatment, whether by percutaneous or open release, is highly successful and widely regarded as the ultimate treatment for trigger finger. Indication for surgical treatment is generally failure of conservative treatment to resolve pain and symptoms. The timing of surgery is somewhat controversial with data suggesting surgical consideration after failure of both a single as well as multiple corticosteroid injections .

The percutaneous trigger finger release has been described and was first introduced by Lorthioir in 1958 .

Open release of trigger finger has been used as treatment for over a century .

Conclusion

Trigger finger is a long recognized condition characterized by a sometimes painful locking of the digit on flexion and extension. It is caused by the inflammation and subsequent narrowing of the A1 pulley through which the flexor

Thanks & Questions

SECTOR: A Neural Model for Coherent Topic Segmentation and Classification

Code and dataset available on GitHub:
https://github.com/sebastianarnold/SECTOR
https://github.com/sebastianarnold/WikiSection

Our work is funded by the German Federal Ministry of Economic Affairs and Energy (BMWi) under grant agreement 01MD16011E (Medical Allround-Care Service Solutions) and H2020 ICT-2016-1 grant agreement 732328 (FashionBrain).

Gefördert durch:

Bundesministerium für Wirtschaft und Energie

aufgrund eines Beschlusses des Deutschen Bundestages

Speaker: Sebastian Arnold

sarnold@beuth-hochschule.de
@sebastianarnold

Data Science and Text-based Information Systems (DATEXIS)
Beuth University of Applied Sciences
Berlin, Germany
www.datexis.de

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

References

[Ag09] Agarwal and Yu, 2009. **Automatically classifying sentences in full-text biomedical articles into introduction, methods, results and discussion.** Bioinformatics 25

[All02] Allan, 2002. **Introduction to topic detection and tracking.** Topic Detection and Tracking

[Aro17] Arora et al., 2017. **A simple but tough-to-beat baseline for sentence embeddings.** ICLR '17

[Bha16] Bhatia et al., 2016. **Automatic labelling of topics with neural embeddings.** COLING '16

[Chen09] Chen et al., 2009. **Global models of document structure using latent permutations.** HLT-NAACL '09

[Choi00] Choi, 2000. **Advances in domain independent linear text segmentation.** NAACL '00

[Coh18] Cohen et al., 2018. **WikiPassageQA: A benchmark collection for research on non-factoid answer passage retrieval.** SIGIR '18

[Di07] Dias et al., 2007. **Topic segmentation algorithms for text summarization and passage retrieval: An exhaustive evaluation.** AAAI '07

[Eis08] Eisenstein and Barzilay, 2008. **Bayesian unsupervised topic segmentation.** EMNLP '08

[Ge00] Gers et al., 2000. **Learning to forget: Continual prediction with LSTM.** Neural Computation 12

[Gla16] Glavaš et al., 2016. **Unsupervised text segmentation using semantic relatedness graphs.** SEM '16

[Gra12] Graves, 2012. **Supervised Sequence Labelling with Recurrent Neural Networks.**

[Ho97] Hochreiter and Schmidhuber, 1997. **Long short-term memory.** Neural Computation 9

[Kosh18] Koshorek at al., 2018. **Text segmentation as a supervised learning task.** NAACL-HLT '18

[Le14] Le and Mikolov, 2014. **Distributed representations of sentences and documents.** ICML '14

[Ma18] MacAvaney et al., 2018. **Characterizing question facets for complex answer retrieval.** SIGIR '18

[Mik13] Mikolov et al., 2013. **Efficient estimation of word representations in vector space.** CoRR, cs.CL/1301.3781v3.

[Rie12] Riedl and Biemann, 2012. **Topic-Tiling: A text segmentation algorithm based on LDA.** ACL '12 Student Research Workshop

[Se17] Serrà and Karatzoglou, 2017. **Getting deep recommenders fit: Bloom embeddings for sparse binary input/output networks.** RecSys '17

[Zi98] Ziou and Tabbone, 1998. **Edge detection techniques – An overview.** Pattern Recognition and Image Analysis 8