

Neural Facet Detection on Medical Resources

Thomas Steffek, WS 18/19

NCBI Resources How To Sign in to NCBI

PubMed.gov US National Library of Medicine National Institutes of Health

PubMed measles Search

Create RSS Create alert Advanced Help

Article types: Clinical Trial, Review, Customize ...

Text availability: Abstract, Free full text, Full text

Publication dates: 5 years, 10 years, Custom range...

Species: Humans, Other Animals

Clear all Show additional filters

Format: Summary Sort by: Most Recent Per page: 20

Send to Filters: Manage Filters

Sort by: Best match Most recent

Search results

Items: 1 to 20 of 27856 << First < Prev Page 1 of 1393 Next > Last >>

1. [Measles: low uptake blamed on "incredulity and hostility" towards doctors.](#)
Mahase E. BMJ. 2019 Apr 26;365:11932. doi: 10.1136/bmj.11932. No abstract available. PMID: 31028094 [Similar articles](#)

2. [Missed Opportunities for Rotavirus Vaccination.](#)
Sederdahl BK, Orenstein WA, Yi J, Anderson EJ, Bednarczyk RA. Pediatrics. 2019 Apr 25. pii: e20182498. doi: 10.1542/peds.2018-2498. [Epub ahead of print] PMID: 31023829 [Similar articles](#)

3. [Being unvaccinated and having a contact history increased the risk of measles infection during an outbreak: a finding from measles outbreak investigation in rural district of Ethiopia.](#)
Girmay A, Dadi AF. BMC Infect Dis. 2019 Apr 25;19(1):345. doi: 10.1186/s12879-019-3973-8. PMID: 31023269 [Similar articles](#)

4. [Histopathological and Immunohistochemical Characteristics of Measles Exanthema: A Study of a Series of 13 Adult Cases and Review of the Literature.](#)
Liersch J, Omaj R, Schaller J. Am J Dermatopathol. 2019 Apr 18. doi: 10.1097/DAD.0000000000001431. [Epub ahead of print] PMID: 31021834 [Similar articles](#)

5. [Type I Interferon receptor signaling drives selective permissiveness of astrocytes and microglia to measles virus during brain infection.](#)
Welsch JC, Charvet B, Dussurgey S, Allatif O, Aurine N, Horvat B, Gerlier D, Mathieu C. J Virol. 2019 Apr 24. pii: JVI.00618-19. doi: 10.1128/JVI.00618-19. [Epub ahead of print] PMID: 31019048 [Similar articles](#)

6. [Updated NACI recommendations for measles post-exposure prophylaxis.](#)
Tunis MC, Salvadori MI, Dubey V, Baclic O; National Advisory Committee on Immunization (NACI)*. Can Commun Dis Rep. 2018 Sep 6;44(9):226-230. doi: 10.14745/ccdr.v44i09a07. eCollection 2018 Sep 6. PMID: 31015814 Free PMC Article [Similar articles](#)

7. [Are community health workers cost-effective for childhood vaccination in India?](#)
Bettampadi D, Boulton ML, Power LE, Hutton DW. Vaccine. 2019 Apr 19. pii: S0264-410X(19)30500-6. doi: 10.1016/j.vaccine.2019.04.038. [Epub ahead of print] PMID: 31015814 [Similar articles](#)

Results by year

Download CSV

Related searches

PMC Images search for measles

See more (6025)...

Titles with your search terms

Vaccinating in the age of apathy: measles vaccination in Canada, 1963-1998. [CMAJ. 2018]

Measles, magic and misidentifications: a defence of the two-factor theory [Cogn Neuropsychiatry. 2019]

Burden of measles in Nigeria: a five-year review of casebased surveillance in [Pan Afr Med J. 2019]

See more...

Find related data

Database: Select

Smart-MD
Neural Paragraph Retrieval

Lyme × treatment × Search... 1

2

case (120) treatment (105) mechanism (90) risk (85) management (83) genotype (50) symptom (43) diagnosis (42) prognosis (38)

antibiotics 4 5

https://en.wikipedia.org/wiki/Lyme_disease

Antibiotics are the primary treatment. The specific approach to their use is dependent on the individual affected and the stage of the disease. For most people with early localized infection, oral administration of doxycycline is widely recommended as the first choice, as it is effective against not only "Borrelia" bacteria but also a variety of other illnesses carried by ticks. Doxycycline is contraindicated in children younger than eight years of age and women who are pregnant or breastfeeding; alternatives to doxycycline are amoxicillin, cefuroxime axetil, and azithromycin. Individuals with early disseminated or late infection may have symptomatic cardiac disease, refractory Lyme arthritis, or neurologic symptoms like meningitis or encephalitis. Intravenous administration of ceftriaxone is recommended as the first choice in these cases; cefotaxime and doxycycline are available as alternatives. These treatment regimens last from one to four weeks. If joint swelling persists or returns, a second round of antibiotics may be considered. Outside of that, a prolonged antibiotic regimen lasting more than 28 days is not recommended as no clinical evidence shows it to be effective. IgM and IgG antibody levels may be elevated for years after successful treatment with antibiotics. As antibody levels are not indicative of treatment success, testing for them is not recommended. 6 view

https://en.wikipedia.org/wiki/Relapsing_fever

Relapsing fever is easily treated with a one- to two-week-course of antibiotics, and most people improve within 24 hours. Complications and death due to relapsing fever are rare. Tetracycline-class antibiotics are most effective. These can, however, induce a Jarisch–Herxheimer reaction in over half those treated, producing anxiety, diaphoresis, fever, tachycardia and tachypnea with an initial pressor response followed rapidly by hypotension. Recent studies have shown tumor necrosis factor-alpha may be partly responsible for this reaction. view

cefuroxime axetil

Related Treatments of "Lyme" 3

Treatment	Count
other	25
antibiotics	20
cefuroxime axetil	8
doxycycline	8
cefotaxime	8
ceftriaxone	8

Figure 1: Screenshot of the Smart-MD user interface. The search bar (1) shows current query terms and offers a auto completion based on the neural entity and topic extractors. The fact distribution chart (3) and the topic tag bar (2) offer visual navigation

Hypotheses

In a novel usage, we apply Smart-MDs underlying machine learning model SECTOR on discharge summaries courtesy of *Charité Berlin's Medical Department, Division of Nephrology and Internal Intensive Care Medicine*.

We define two hypotheses:

- i. Specialized text embeddings perform better than general purpose text embeddings on medical domain
- ii. SECTOR as effective means of facet extraction on medical resources

Outline



Methodology

Bootstrapping Training Data
Facet Extraction with SECTOR



Evaluation

Quantitative Evaluation
Qualitative Evaluation



Conclusion

!!! Slide was removed for final presentation due to time restriction !!!

Challenges

Semantic Mismatch with WikiSection

- **Structural** mismatches due to differing medium and purpose
- **Vocabulary** mismatches due to differing intention of author

Missing Training Data

- **Privacy regulations** in Europe and Germany
- **Novel joined task** of facet segmentation and classification

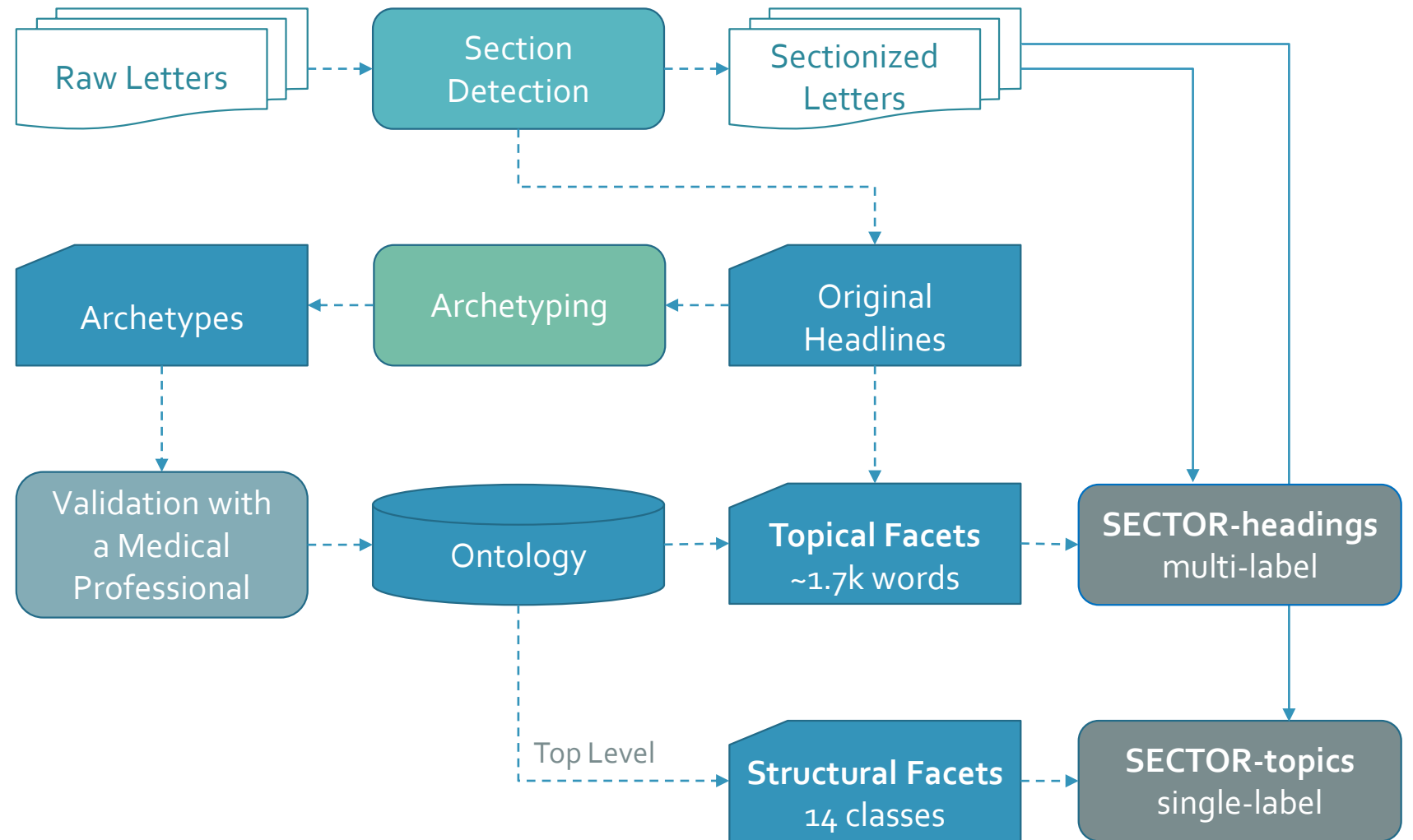
Ambique Medical Language

- Ambiguous medical **terms**
- Misleading **content** within sections
- Differentiation between structural and topical **facets**

Highly Specialized Domain Knowledge

- Medical work requires extensive studies and knowledge

Methodology Overview



Methodology

Bootstrapping

Section Detection

- Using regular expressions to segment sections and detect original headlines

Archotyping

- Aggregate original headlines to a manageable amount using a custom stemming algorithm

Validation with a Medical Professional

- Building an ontology on most common archetypes with the help of a medical professional

Ontology Example:

level 1	level 2	original title
Bildgebende Diagnostik	Röntgen	Röntgen-Thorax

Methodology

Facet Extraction

Structural Facets

- “...serve a structural purpose for an article — general question facets that could be asked about many similar topics” [Mac+18]

mutually exclusive



single-label problem

pre-defined generalized options



top level ontology

Example:

Röntgen-Thorax



Bildgebende Diagnostik

Methodology

Facet Extraction

Topical Facets

- “...describe details that are specific to the particular topic” [Mac+18]

ambiguous headings

reflect hierarchy

multi-label problem

all levels ontology

Example:

Röntgen-Thorax



Röntgen



Bildgebende Diagnostik

!!! Slide was removed for final presentation due to time restriction !!!

Evaluation of L2L-structural per Class

Class	#Examples	TP	FP	Acc	Prec	Rec	F1
Diagnose	2082	2032	84	97.60	96.03	97.60	96.81
Bildgebende Diagnostik	753	717	230	95.22	75.71	95.22	84.35
Status	981	575	61	58.61	90.41	58.61	71.12
Diagnostische Maßnahm	1732	1424	194	82.22	88.01	82.22	85.01
Labor	23131	23041	1439	99.61	94.12	99.61	96.79
Brief Kopf	3393	3393	0	100.00	100.00	100.00	100.00
Brief Anrede	491	476	3	96.95	99.37	96.95	98.14
Brief Schluss	1588	1588	4	100.00	99.75	100.00	99.87
Medikation	6431	6425	3	99.91	99.95	99.91	99.93
Verlauf und Therapie	888	699	17	78.72	97.63	78.72	87.16
other	799	328	23	41.05	93.45	41.05	57.04
Konsil	82	70	31	85.37	69.31	85.37	76.50
Beurteilung	458	62	8	13.54	88.57	13.54	23.48
Befund	276	137	21	49.64	86.71	49.64	63.13
[macro-avg]	43085	40967	2118	95.08	91.36	78.46	81.38

Evaluation
Qualitative

!!! Slide was removed for final presentation due to time restriction !!!

Evaluation of L2L-structural per Class

Class	#Examples	TP	FP	Acc	Prec	Rec	F1
Diagnose	2082	2032	84	97.60	96.03	97.60	96.81
Bildgebende Diagnostik	753	717	230	95.22	75.71	95.22	84.35
Status	981	575	61	58.61	90.41	58.61	71.12
Diagnostische Maßnahm	1732	1424	194	82.22	88.01	82.22	85.01
Labor	23131	23041	1439	99.61	94.12	99.61	96.79
Brief Kopf	3393	3393	0	100.00	100.00	100.00	100.00
Brief Anrede	491	476	3	96.95	99.37	96.95	98.14
Brief Schluss	1588	1588	4	100.00	99.75	100.00	99.87
Medikation	6431	6425	3	99.91	99.95	99.91	99.93
Verlauf und Therapie	888	699	17	78.72	97.63	78.72	87.16
other	799	328	23	41.05	93.45	41.05	57.04
Konsil	82	70	31	85.37	69.31	85.37	76.50
Beurteilung	458	62	8	13.54	88.57	13.54	23.48
Befund	276	137	21	49.64	86.71	49.64	63.13
[macro-avg]	43085	40967	2118	95.08	91.36	78.46	81.38

To address recall errors: Sampling false negatives.

Evaluation
Qualitative

!!! Slide was removed for final presentation due to time restriction !!!

Evaluation of L2L-structural per Class

Class	#Examples	TP	FP	Acc	Prec	Rec	F1
Diagnose	2082	2032	84	97.60	96.03	97.60	96.81
Bildgebende Diagnostik	753	717	230	95.22	75.71	95.22	84.35
Status	981	575	61	58.61	90.41	58.61	71.12
Diagnostische Maßnahm	1732	1424	194	82.22	88.01	82.22	85.01
Labor	23131	23041	1439	99.61	94.12	99.61	96.79
Brief Kopf	3393	3393	0	100.00	100.00	100.00	100.00
Brief Anrede	491	476	3	96.95	99.37	96.95	98.14
Brief Schluss	1588	1588	4	100.00	99.75	100.00	99.87
Medikation	6431	6425	3	99.91	99.95	99.91	99.93
Verlauf und Therapie	888	699	17	78.72	97.63	78.72	87.16
other	799	328	23	41.05	93.45	41.05	57.04
Konsil	82	70	31	85.37	69.31	85.37	76.50
Beurteilung	458	62	8	13.54	88.57	13.54	23.48
Befund	276	137	21	49.64	86.71	49.64	63.13
[macro-avg]	43085	40967	2118	95.08	91.36	78.46	81.38

To address precision errors: Sampling false positives.

Evaluation
Qualitative

Evaluation

Qualitative

Hierarchical Error

- Sections that are identified as atomic units, but actually constitute a subcategory of the preceding section
- Origins in **wrong assumptions** about the letters' content

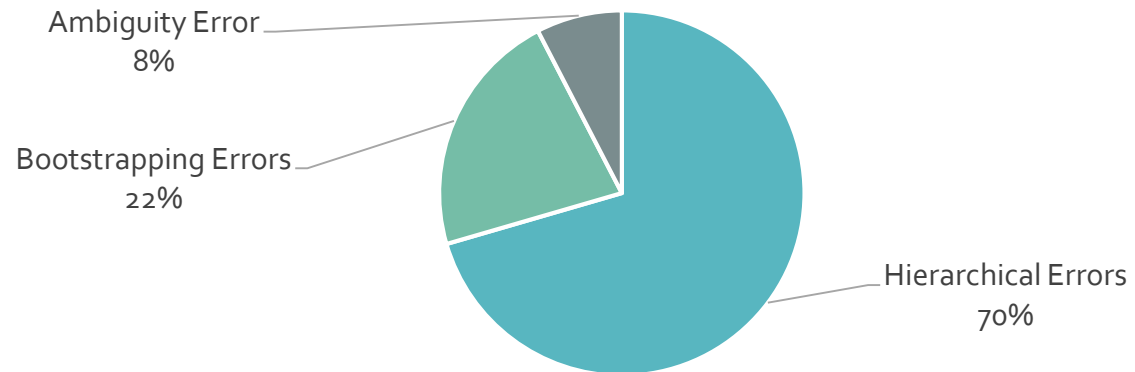
Bootstrapping Error

- Sections that are wrongfully labeled due to errors during bootstrapping process
- Origins in **bootstrapping algorithm**

Ambiguity Error

- Sections whose contents seem to belong to a specific class, but belong to another
- Origins in **neural network**

Error Distribution



Evaluation

Qualitative

Conclusions

- Ontology failed to recognize structural hierarchy
- Bootstrapping algorithms are a mere approximation

Evaluation

Quantitative

best performing model	P@1	P@3	R@1	R@3	F1	Pk	MAP
L2L dataset: 14 structural facets as single-label task							
SEC>T+bow	95.21	32.68	95.21	98.04	95.08	2.40	96.74
SEC>T+fT@CC	94.08	32.51	94.08	97.53	94.35	3.10	96.26
SEC>T+W ₂ V@WD+DL	94.72	32.60	94.72	97.79	94.83	2.56	96.55
SEC>T+fT@WD+DL	94.58	32.59	94.58	97.77	94.65	2.82	96.50
L2L dataset: 1,670 topical facets as multi-label-task							
SEC>H+bow	85.49	45.20	61.90	84.58	77.90	10.15	88.74
SEC>T+fT@CC	93.42	50.52	64.66	89.71	81.48	9.16	93.10
SEC>H+W ₂ V@WD+DL	95.16	52.20	65.22	91.19	82.25	8.91	94.45
SEC>H+fT@WD+DL	94.89	51.63	65.12	90.53	82.20	6.36	93.89
L2.1L dataset: 12 structural facets as single-label task							
SEC>T+bow	98.72	33.25	98.72	99.74	98.97	0.96	99.41
SEC>T+W ₂ V@WD+DL	98.68	33.25	98.68	99.75	95.60	3.21	97.59
SEC>T+fT@WD+DL	97.79	33.15	97.79	99.44	98.39	1.69	99.02
L2.1L dataset: 1,687 topical facets as multi-label task							
SEC>H+bow	99.13	52.90	69.33	93.92	87.07	5.80	97.36
SEC>H+W ₂ V@WD+DL	97.68	52.23	68.68	93.32	86.43	7.64	97.15
SEC>H+fT@WD+DL	97.50	51.51	68.67	92.58	86.45	7.15	96.70

Evaluation

Quantitative

best performing model	P@1	P@3	R@1	R@3	F1	Pk	MAP
L2L dataset: 14 structural facets as single-label task							
SEC>T+bow	95.21	32.68	95.21	98.04	95.08	2.40	96.74
SEC>T+fT@CC	94.08	32.51	94.08	97.53	94.35	3.10	96.26
SEC>T+W ₂ V@WD+DL	94.72	32.60	94.72	97.79	94.83	2.56	96.55
SEC>T+fT@WD+DL	94.58	32.59	94.58	97.77	94.65	2.82	96.50
L2L dataset: 1,670 topical facets as multi-label-task							
SEC>H+bow	85.49	45.20	61.90	84.58	77.90	10.15	88.74
SEC>T+fT@CC	93.42	50.52	64.66	89.71	81.48	9.16	93.10
SEC>H+W ₂ V@WD+DL	95.16	52.20	65.22	91.19	82.25	8.91	94.45
SEC>H+fT@WD+DL	94.89	51.63	65.12	90.53	82.20	6.36	93.89
L2.1L dataset: 12 structural facets as single-label task							
SEC>T+bow	98.72	33.25	98.72	99.74	98.97	0.96	99.41
SEC>T+W ₂ V@WD+DL	98.68	33.25	98.68	99.75	95.60	3.21	97.59
SEC>T+fT@WD+DL	97.79	33.15	97.79	99.44	98.39	1.69	99.02
L2.1L dataset: 1,687 topical facets as multi-label task							
SEC>H+bow	99.13	52.90	69.33	93.92	87.07	5.80	97.36
SEC>H+W ₂ V@WD+DL	97.68	52.23	68.68	93.32	86.43	7.64	97.15
SEC>H+fT@WD+DL	97.50	51.51	68.67	92.58	86.45	7.15	96.70

Evaluation

Quantitative

best performing model	P@1	P@3	R@1	R@3	F1	Pk	MAP
L2L dataset: 14 structural facets as single-label task							
SEC>T+bow	95.21	32.68	95.21	98.04	95.08	2.40	96.74
SEC>T+fT@CC	94.08	32.51	94.08	97.53	94.35	3.10	96.26
SEC>T+W ₂ V@WD+DL	94.72	32.60	94.72	97.79	94.83	2.56	96.55
SEC>T+fT@WD+DL	94.58	32.59	94.58	97.77	94.65	2.82	96.50
L2L dataset: 1,670 topical facets as multi-label-task							
SEC>H+bow	85.49	45.20	61.90	84.58	77.90	10.15	88.74
SEC>T+fT@CC	93.42	50.52	64.66	89.71	81.48	9.16	93.10
SEC>H+W ₂ V@WD+DL	95.16	52.20	65.22	91.19	82.25	8.91	94.45
SEC>H+fT@WD+DL	94.89	51.63	65.12	90.53	82.20	6.36	93.89
L2.1L dataset: 12 structural facets as single-label task							
SEC>T+bow	98.72	33.25	98.72	99.74	98.97	0.96	99.41
SEC>T+W ₂ V@WD+DL	98.68	33.25	98.68	99.75	95.60	3.21	97.59
SEC>T+fT@WD+DL	97.79	33.15	97.79	99.44	98.39	1.69	99.02
L2.1L dataset: 1,687 topical facets as multi-label task							
SEC>H+bow	99.13	52.90	69.33	93.92	87.07	5.80	97.36
SEC>H+W ₂ V@WD+DL	97.68	52.23	68.68	93.32	86.43	7.64	97.15
SEC>H+fT@WD+DL	97.50	51.51	68.67	92.58	86.45	7.15	96.70

Evaluation

Quantitative

best performing model	P@1	P@3	R@1	R@3	F1	Pk	MAP
L2L dataset: 14 structural facets as single-label task							
SEC>T+bow	95.21	32.68	95.21	98.04	95.08	2.40	96.74
SEC>T+fT@CC	94.08	32.51	94.08	97.53	94.35	3.10	96.26
SEC>T+W ₂ V@WD+DL	94.72	32.60	94.72	97.79	94.83	2.56	96.55
SEC>T+fT@WD+DL	94.58	32.59	94.58	97.77	94.65	2.82	96.50
L2L dataset: 1,670 topical facets as multi-label-task							
SEC>H+bow	85.49	45.20	61.90	84.58	77.90	10.15	88.74
SEC>T+fT@CC	93.42	50.52	64.66	89.71	81.48	9.16	93.10
SEC>H+W ₂ V@WD+DL	95.16	52.20	65.22	91.19	82.25	8.91	94.45
SEC>H+fT@WD+DL	94.89	51.63	65.12	90.53	82.20	6.36	93.89
L2.1L dataset: 12 structural facets as single-label task							
SEC>T+bow	98.72	33.25	98.72	99.74	98.97	0.96	99.41
SEC>T+W ₂ V@WD+DL	98.68	33.25	98.68	99.75	95.60	3.21	97.59
SEC>T+fT@WD+DL	97.79	33.15	97.79	99.44	98.39	1.69	99.02
L2.1L dataset: 1,687 topical facets as multi-label task							
SEC>H+bow	99.13	52.90	69.33	93.92	87.07	5.80	97.36
SEC>H+W ₂ V@WD+DL	97.68	52.23	68.68	93.32	86.43	7.64	97.15
SEC>H+fT@WD+DL	97.50	51.51	68.67	92.58	86.45	7.15	96.70

Evaluation

Quantitative

Conclusions

- Bag-of-words with bloom filters outperforms word embeddings
- General purpose model performed worst
- Specialized word2vec and fastText perform on par

Conclusion

- Specialized Text Embeddings Perform Better than General Purposed Text Embeddings on Medical Domain
- SECTOR as Effective Means of Structural Facet Extraction
- SECTOR as Effective Means of Topical Facet Extraction

- Bag-of-words encoding with bloom filter performs better than word embeddings



Sources

- [Mac+18] Sean MacAvaney, Andrew Yates, Arman Cohan, et al. „Characterizing Question Facets for Complex Answer Retrieval“. In: *SIGIR*. arXiv: 1805.00791. May 2018
- [Sch+18] Rudolf Schneider, Sebastian Arnold, Tom Oberhauser, et al. „Smart-MD: Neural Paragraph Retrieval of Medical Topics“. In: *Companion Proceedings of the The Web Conference 2018*. WWW '18. event-place: Lyon, France. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 203–206
- [pub] *Pubmed*. URL: <https://www.ncbi.nlm.nih.gov/pubmed/?term=measles> (visited on Apr. 29, 2019)