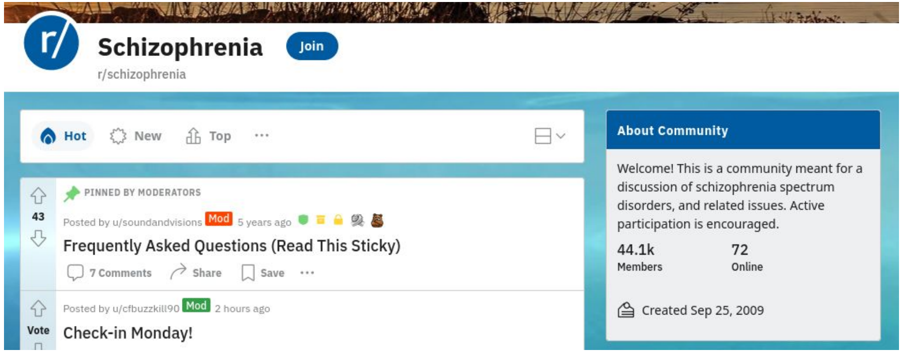
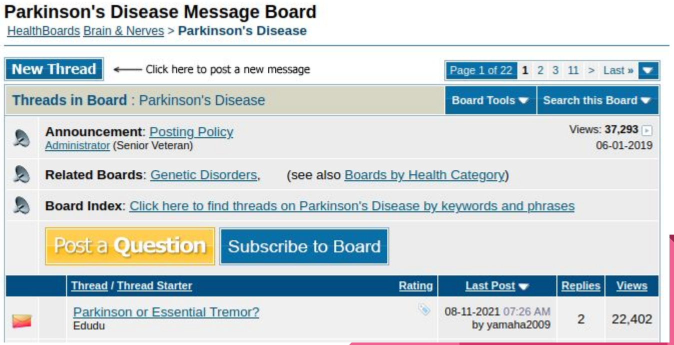
Knowledge-Aware Neural Networks for Medical Forum Question Classification

Soumyadeep Roy*, Sudip Chakraborty, Aishik Mandal, Gunjan Balde, Prakhar Sharma, Niloy Ganguly*, Shamik Sural (IIT Kharagpur, India)
Megha Khosla (Leibniz AI Lab, L3S Research Center, Germany)*
Anandhavelu Natarajan (Adobe Research, India)

Why online medical forums?

- Reliable source of health-related information, advice or support (Sinha et al. 2018)
- Rich interaction between medical professionals and patients
- Contains first-person accounts of patients, care-givers





Medical Forum Question Classification - Example

Title	Description	Health Info. Need Category
Bladder removal with neo-bladder surgery	My sister had invasive bladder cancer and had her bladder removed … is he correct that surgery can't be done now? Should my sister be going to another doctor who knows more about female bladder removal? … I am so sick about all this.	Disease
My 3 yr old is uncontrollable	My three year old girl is my only child. I am trying to figure out what I can do to help her out of control behavior … I am at a loss, … I love my child and I want to help her, this just doesn't seem like normal three year old behavior	Family Support

Sample data point of ICHI 2016 Shared Task dataset
Red: Medical words, Yellow: Relevant context with non-medical words

MFQC Task Formulation - Document Level

- Given a medical forum question (only post title or title + description), the task is to predict its **health information need** category
  - *Treatment* class: specific medical procedures like surgery, or taking a dose of medicines
  - *Family Support* class: issues related to caregiver (not the patient) like how to support one's spouse or an ill child
- Multi-class prediction task
  - Both single-label and multi-label

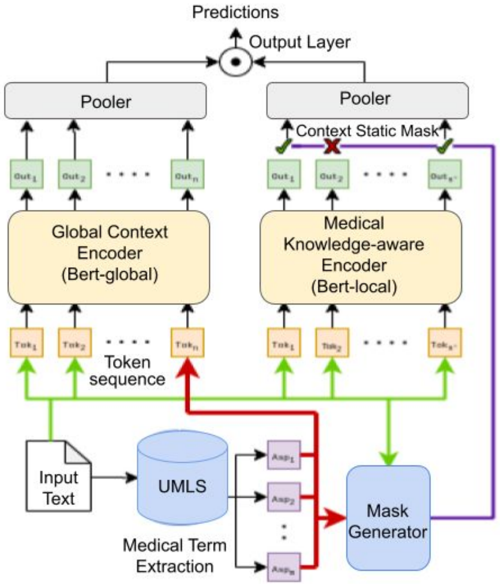
Research Background

- Existing models rely on hand-crafted features, complex ensemble models, embedding-based methods based on open-domain text (**limited medical domain knowledge**)
- Noisy text due to presence of contractions and misspellings -> extra effort required to **standardize patient vocabulary with that of medical professionals**
- **SoA-DN model** (Jalan et al. 2018) extracts medical concept-bearing words from text by MetaMap, and computes 'strength of association' between each word and target class

Jalan et al., Medical forum question classification using deep learning, European Conference on Information Retrieval, 2018, pp. 45 - 58

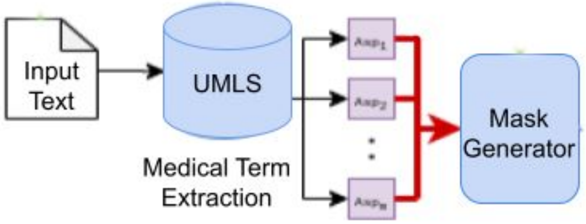
Key Contributions

- **MedBERT**: Novel application of dual-encoder model for MFQC task
- Additional medical domain-specific side information
  - No hand-crafted features, works well in low-resource setting
  - Explicit importance to medical concept-bearing words
- Introduce a labeled multi-label MFQC dataset



MedBERT: Extraction of medical aspects

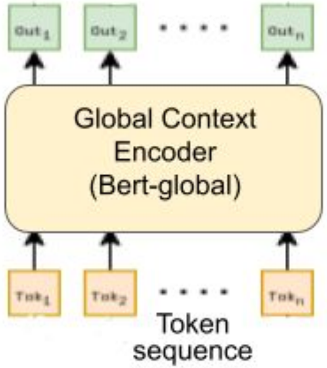
- Whole input text, tokenized into sequence of words (length: n)
- Extract medical concept-bearing tokens (aspects) using QuickUMLS tool
- Utilize a MedDRA Patient Friendly Term List lexicon to improve coverage of medical terms based on patient vocabulary
    - Aching in limb, crawling sensation of skin



MedBERT: Global Context Representation

- Use pre-trained BERT-base uncased model (Devlin et al. 2019) as an encoder
- Input: Append aspect sequence to token sequence as input

[CLS] + ‘I had surgery for retinal detachment in December’ + [SEP] + ‘surgery’ + ‘retinal detachment’ + [SEP]



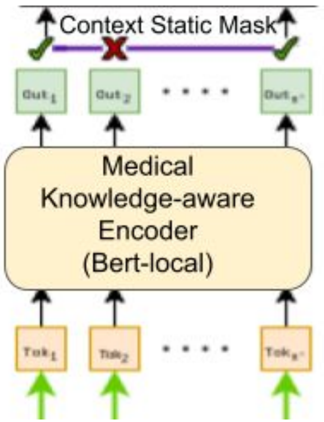
MedBERT: Knowledge-aware Representation

- Same BERT Encoder where the encoder output has a context static mask (CSM)
    - CSM is a binary vector, which zeroes all output tokens that are not medical words

[CLS] + ‘I had surgery for retinal detachment in December.’ + [SEP]

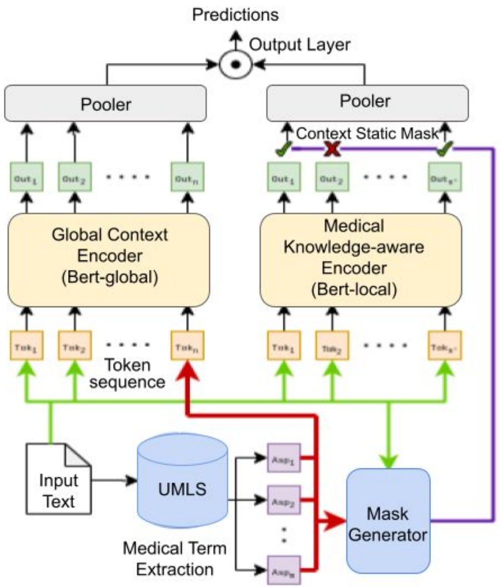
Medical tokens (aspects): surgery, retinal detachment

Context Static Mask: [0, 0, 1, 0, 1, 1, 0, 0]



MedBERT: Combining Global and Local Representation

- Concatenate global and local representation
- Use fully-connected layer to map representation to target classes
- Cross-entropy loss for multi-class prediction
- Binary cross-entropy loss for multi-label prediction



Experimental Setup

- Baselines
  - Neural models w/o medical knowledge: FastText, CNN, HAN, Bert-base
  - SoA-DN (Jalan et al. 2018): strength of association of medical entity with target class
  - SoA-DN + TFIDF-DN + Hier-BiLSTM (**current SOTA**): 3-component ensemble model
  - LCF-BERT, MedBERT (Global only), MedBERT (Local only): Modified by same medical side-information
- Datasets
  - ICHI (multi-class): 8000 train, 3000 test. Same data split used as previous works
    - Disease, Treatment, Family Support, Socializing, Demographic, Goal-oriented, Pregnancy
  - CADEC (multi-label): 942 train, 300 test, 4 classes

Jalan et al., Medical forum question classification using deep learning, European Conference on Information Retrieval, 2018, pp. 45 - 58

Multi-label Annotated Data - CADEC

- Annotate CADEC dataset, a benchmark of adverse drug event dataset based on online medical forum posts
  - High class imbalance issue exists
- Four health information search classes
  - Uncertainty of post-diagnosis (UPD)
  - Medical Assistance (MAS)
  - Diet and Maintenance (DM)
  - Information Source (IS)

Class	Pos.	Neg.	Word Count
UPD	210	1037	117.1
MAS	257	990	85.2
DM	144	1103	117.9
IS	88	1159	138.2

Karimi et al. 2015, Cadec: A corpus of adverse drug event annotations, Journal of Biomedical Informatics, 55, pp. 73 - 81

Performance comparison

Models	**ICHI (All)**	**UPD**	**MAS**	**DM**	**IS**	**All**
TFIDF + SVM	0.64	0.59	0.7	0.74	0.58	0.65
H-BiLSTM+TFIDF -DN + SoA-DN	**0.7**					
FastText	0.61	0.46	0.45	0.47	0.48	0.47
HAN	0.61	0.6	0.6	0.61	0.48	0.57
BERT-base	0.65	0.32	0.72	0.52	0.54	0.53
LCF-BERT	0.69	0.46	0.45	0.47	0.48	0.47
MedBERT	**0.7**	0.55	0.74	0.8	0.75	**0.71**

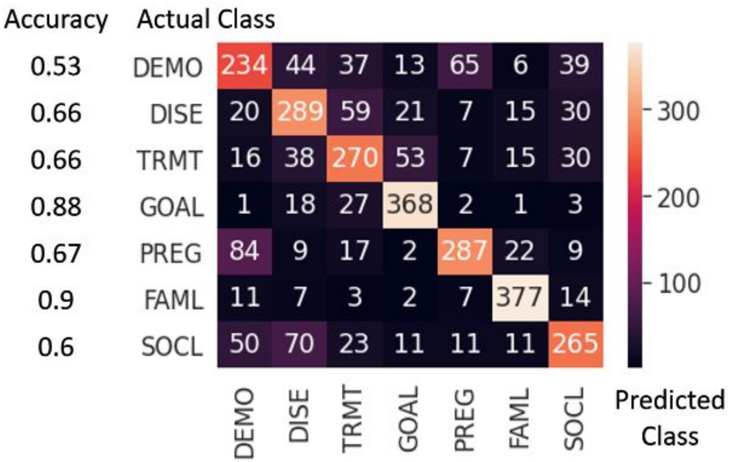
ICHI: Accuracy
CADEC: macro F1

Discussion of Results

- MedBERT achieves best performance on both ICHI and CADEC
- Performs comparably with ensemble model (Hier-BiLSTM + TFIDF-DN + SoA-DN)
- MedBERT outperforms MedBERT (Global only) and MedBERT (Local only) by 2.9%
- MedBERT improves over BERT-Base (like MedBERT w/o medical keywords) by 7.7%
    - Shows the importance of adding medical side-information

Error Analysis

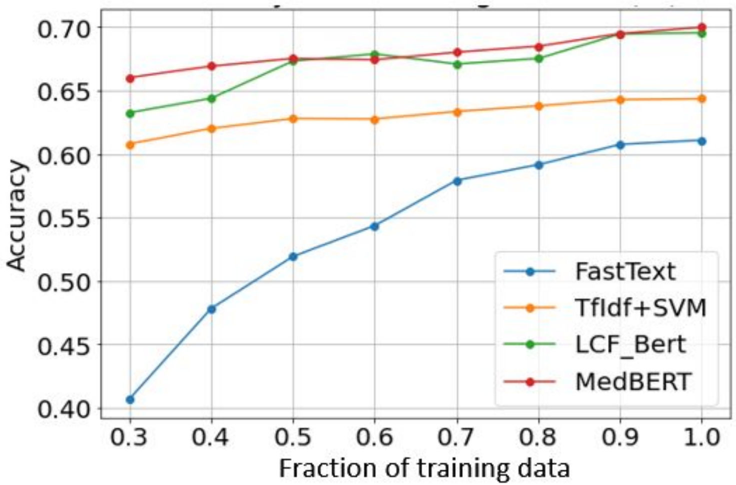
- Pregnancy to Demographics (vice-versa) is frequently misclassified
    - Both target a specific age range and gender



- Socializing to Disease
    - Socializing do not target health-related issues and rather focuses on hobbies/recreational activities
    - MedBERT focuses on medical concept-bearing words

Effect of Training Data Size

- LCF-Bert and MedBERT outperform competing baselines by a good margin
- MedBERT outperforms LCF-Bert in low training data regime
  - Outperforms LCF-Bert by 4.76% on 30% training data
- Medical domain-specific side-information helps to overcome limited training data issue



Conclusion and Key Takeaways

- Propose MedBert, a novel application of dual encoder model for MFQC task
- Contribute a multi-labeled MFQC dataset
- MedBERT generalizes well to low-resource setting

- Codebase available at https://github.com/roysoumya/knowledge-aware-med-classification

Acknowledgements

- Complex Network Research Group, IIT Kharagpur, India (@cnerg) and Leibniz AI Lab, L3S Research Center, Germany
- SIGIR Student Travel Grant for covering the conference registration costs

- Research Work supported in part by:
    - Institute PhD Fellowship, IIT Kharagpur, India
    - Federal Ministry of Education and Research (BMBF), Germany
    - Adobe Research, India
    - IMPRINT-1 Project RCO, India

References

- Healthcare data analytics challenge, in “2016 IEEE International Conference on Healthcare Informatics (ICHI)”
- Derksen et al. (2017), What say ye gout experts? A content analysis of questions posted on the social news website Reddit, BMC Musculoskelet Disord 18, 1
- Zeng et al. (2019), LCF: A Local Context Focus Mechanism for Aspect-Based Sentiment Classification, Applied Sciences 9, 16
- Mrabet et al. (2016), Combining open-domain and biomedical knowledge for topic recognition in consumer health questions, AMIA Annual Symposium Proceedings, pp. 1040 - 1049
- Sinha et al. (2018), The use of online health forums by patients with chronic cough: qualitative study, J. Med. Internet Research, 20, 1

Thank you for listening

Questions?

Please feel free to contact me at:
soumyadeep.roy9@iitkgp.ac.in, sroy@l3s.de
@roysoumya1