



# PubMeSH: Extreme multi-label classification of biomedical research

CS224N

Kevin Thomas<sup>1</sup>, Rohan Paul<sup>1</sup>, Mia Kanzawa<sup>1</sup>

1. Department of Biomedical Data Science, Stanford University, Stanford, CA.

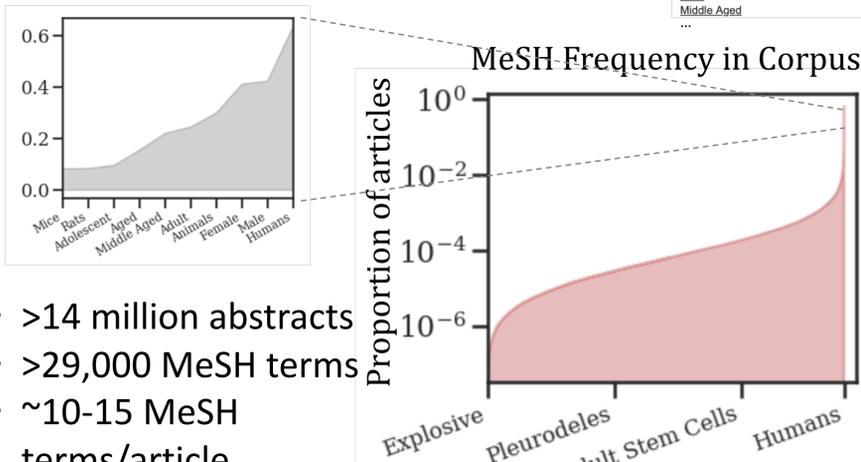
## Problem

- PubMed/MEDLINE is central hub for biomedical and life sciences journal articles and critical for dissemination of scientific knowledge
  - Contains **>29 million articles**
  - 3.3 billion searches in 2017
- Searches rely on MeSH labels which are manually assigned to each article
  - Estimated **>\$7 million** in labor costs/year
- Automatic labelling could significantly reduce costs
- Previous state-of-the-art: LSTM with attention
  - MiF of 0.6880 → requires human review
  - Limit # of possible labels with masking approach (misses up to 5% of labels)
  - Naïve representation of MeSH labels

## Data/Task

Goal: assign a set of MeSH terms to each abstract

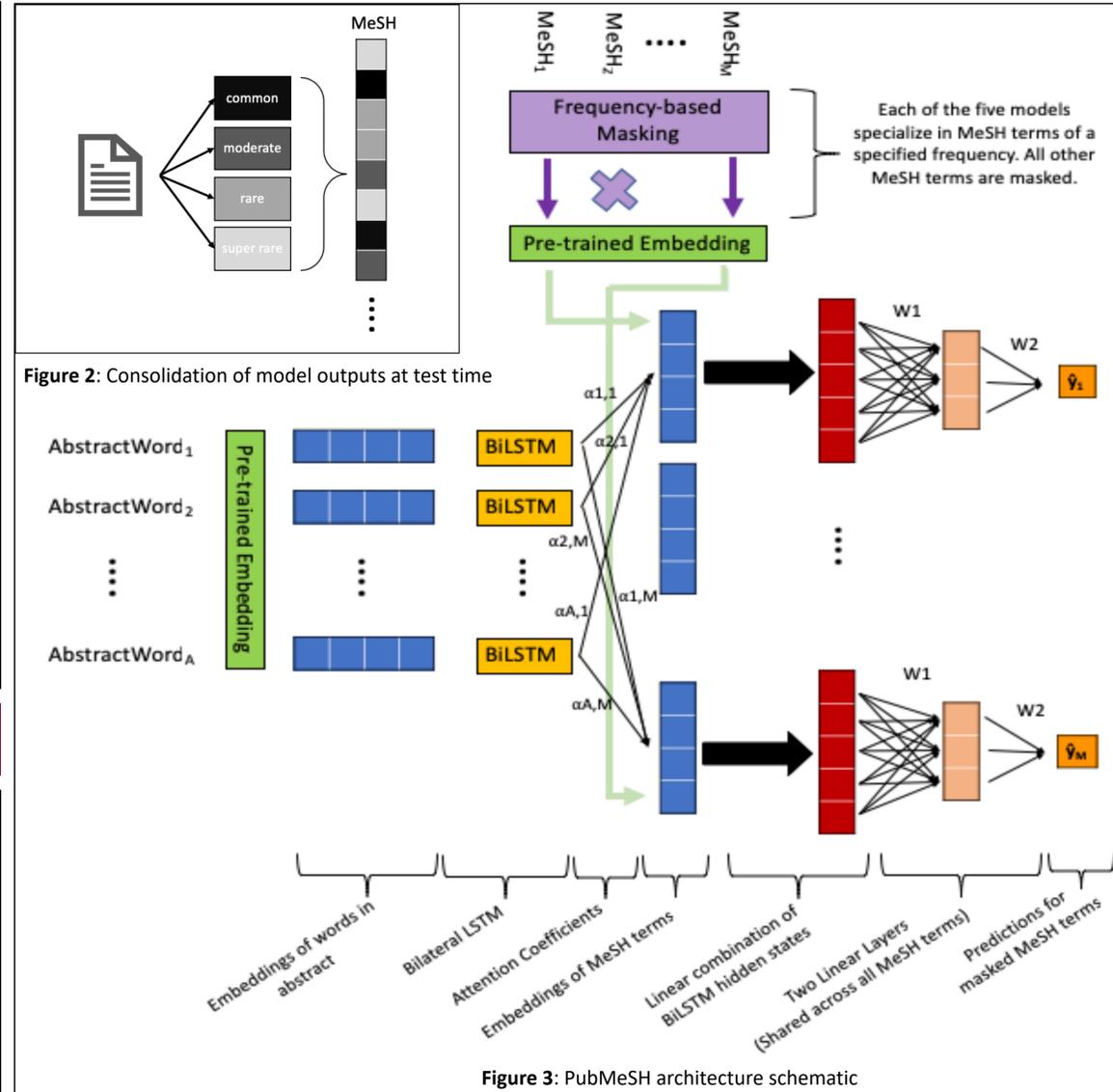
multi-label, multi-class classification



- >14 million abstracts
- >29,000 MeSH terms
- ~10-15 MeSH terms/article

Figure 1: Frequency distribution of MeSH labels among articles

## Approach



## Results

Model Name	MiF
Common	0.670
Moderate	0.384
Rare	0.229
Super Rare	0.171
Ensemble	0.484

## Analysis

Produces viable labels absent from ground truth

Model	Expert
[Curriculum, Education, Internship and Residency, Students]	[Clinical Competence, Education]
[Oxygen, Exercise]	[Oxygen Consumption, Physical Endurance, Heart Rate]
[Fracture Fixation, Osteoporosis]	[Lumbar Vertebrae, Pain Measurement]

Misses important labels

Model	Expert
[Brazil, Food, Fruit]	[Brazil, Fruit, Meat, Vegetables, Energy Intake, Eating, Feeding Behavior, Obesity, Body Mass Index]
[]	[Amiodarone]

## Conclusion

- Proposed novel methods for automated MeSH indexing do not yet surpass state-of-the-art
  - Pre-trained MeSH embeddings
  - Separate training for prediction of MeSH labels in different frequency categories
- Future directions:
  - Ablation studies for linear layers
  - Freezing embedding layers
  - Transfer learning
  - Hyperparameter tuning

## References

- Jin, Q., Dhingra, B., Cohen, W. & Lu, X. AttentionMeSH: Simple, Effective and Interpretable Automatic MeSH Indexer. in *Proceedings of the 6th BioASQ Workshop A challenge on large-scale biomedical semantic indexing and question answering* 47–56 (Association for Computational Linguistics, 2018).
- You, R., Dai, S., Zhang, Z., Mamitsuka, H. & Zhu, S. AttentionXML: Extreme Multi-Label Text Classification with Multi-Label Attention Based Recurrent Neural Networks. (2018).
- The Challenge | bioasq.org. Available at: <http://bioasq.org/>. (Accessed: 20th March 2019)