

UTSolve at the NTCIR-18 MedNLP-CHAT: Leveraging BioBERT for Medical Text Classification

AUTHORS

Guanqi Cheng, Chang Qu, Ali Braytee
School of Computer Science, University of Technology Sydney

Emails: Guanqi.Cheng@student.uts.edu.au,
Chang.Qu-1@student.uts.edu.au,
ali.braytee@uts.edu.au

KEYWORDS

Medical NLP,
Large language models,
Question answering,
Risk classification

ABSTRACT

Our team, UTSolve, participated in the Medical Natural Language Processing for AI Chat (MedNLP-CHAT) task 1 at NTCIR-18. The task involved classifying various medical texts into medical, ethical, and legal risks. We utilized BioBERT to predict the risk level of medical texts, achieving the best performance among compared models.

METHODOLOGY

We used BioBERT v1.1 pre-trained on PubMed and PMC texts. Our dataset includes expert-annotated QA pairs. We fine-tuned the model with AdamW, learning rate of $3e-5$, batch size 16, and cross-entropy loss.

RESULTS

Risk Category	BioBERT F1	BioBERT Acc	MedBERT F1	MedBERT Acc	ClinicalBERT F1	ClinicalBERT Acc
Medical Risk	0.7812	0.8	0.6709	0.65	0.7	0.7
Ethical Risk	0.8629	0.85	0.719	0.65	0.8275	0.8
Legal Risk	0.7288	0.75	0.4875	0.6	0.5121	0.65

DISCUSSION&CONCLUSION

BioBERT showed strong performance on small datasets without external features. Future work includes integrating external knowledge and applying multi-task learning.

Pre-trained transformer-based models such as BioBERT are effective for medical risk classification. Future work may further enhance performance via ensemble methods.