

BERT MODELS FOR BIOMEDICAL RELATION EXTRACTION

Murali Mohana Krishna Dandu¹, Vanitha Sivasankaran Balasubramaniam², A Renuka³, Om Goel⁴,

Prof.(Dr) Punit Goel⁵ & Dr. Alok Gupta⁶

¹Independent Researcher, Satyanarayana Puram, Vijayawada, Andhra Pradesh, India

²Independent Researcher, PT Rajan Salai, KK Nagar, Chennai, Tamil Nadu, India

*³Independent Researcher, Maharaja Agrasen Himalayan Garhwal University, Dhaid Gaon, Block Pokhra,
Uttarakhand, India*

⁴Independent Researcher, Abes Engineering College, Ghaziabad,

⁵Research Supervisor, Maharaja Agrasen Himalayan Garhwal University, Uttarakhand, India

⁶Ph.D, COT, G.B. Pant University of Agriculture & Technology, Pantanagar, India

ABSTRACT

Biomedical relation extraction is pivotal for advancing knowledge discovery and supporting decision-making in healthcare and research. Leveraging the capabilities of Bidirectional Encoder Representations from Transformers (BERT) has significantly enhanced the accuracy and efficiency of extracting complex relationships from vast biomedical literature. This study explores the application of BERT-based models in identifying and classifying relationships among biomedical entities such as genes, proteins, diseases, and drugs. By fine-tuning pre-trained BERT models on specialized biomedical corpora, the research addresses the nuanced linguistic patterns and domain-specific terminologies inherent in biomedical texts. Comparative analyses demonstrate that BERT models outperform traditional machine learning approaches and earlier deep learning frameworks in tasks like protein-protein interaction extraction, drug-disease association identification, and gene-disease relationship mapping. Additionally, the study investigates the integration of domain-adaptive pre-training and the incorporation of external knowledge bases to further enhance model performance. Challenges such as data scarcity, ambiguity in biomedical language, and the need for extensive computational resources are discussed, alongside strategies to mitigate these issues. The findings underscore the potential of BERT-based models to facilitate more accurate and scalable biomedical information extraction, thereby supporting the acceleration of biomedical research and the development of innovative healthcare solutions. Future directions include the exploration of more advanced transformer architectures, the expansion of annotated biomedical datasets, and the implementation of real-time relation extraction systems. This research contributes to the growing body of knowledge on natural language processing in the biomedical domain and highlights the transformative impact of BERT models on extracting meaningful relationships from complex biomedical data.

KEYWORDS: *BERT Models, Biomedical Relation Extraction, Natural Language Processing, Transformer Architectures, Gene-Disease Relationships, Protein-Protein Interactions, Drug-Disease Associations, Deep Learning in Biomedicine, Domain-Specific Fine-Tuning, Knowledge base Integration, Biomedical Text Mining, machine Learning in Healthcare, Relation Classification, Information Extraction*

Article History

Received: 04 Mar 2022 | Revised: 12 Mar 2022 | Accepted: 18 Mar 2022
