Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140

Volume 14 Issue 03s, 2025

Pharmaguard: AI-Driven Prediction of Drug-Drug Interactions and Side Effects using LLMs

[1]Snehal Awachat , [2]Pranali Mule, [3]Shreya Takmale, [4] Trupti Ghate, [5]Shravani Bhuskat
[1,2,3,4,5] *Artificial Intelligence SVPCET, Nagpur University Nagpur, India*
Email: [1]snehalambatkar8@gmail.com, [2]pranalimule.22@stvincentngp.edu.in,
[3]shreyatalmale.22@stvincentngp.edu.in, [4]truptighate.22@stvincentngp.edu.in,
[5]shravanibhuskat.22@stvincentngp.edu.in

Peer Review Information

Submission: 05 Nov 2025

Revision: 25 Nov 2025

Acceptance: 17 Dec 2025

Keywords

Drug-Drug Interaction (DDI), BioBERT, GPT, Large Language Models (LLMs), Natural Language Processing (NLP), Biomedical Text Mining, Explainable AI, Clinical Decision Support, Precision Medicine.

Abstract

Drug-Drug Interactions (DDIs) continue to pose sig- nificant challenges in pharmacological safety, particularly when multiple medicines are prescribed simultaneously. The growing scale of biomedical data has created new opportunities to identify such interactions computationally, yet existing approaches often lack interpretability and adaptability to unseen drug pairs. This study presents a customized hybrid framework that combines the biomedical understanding of BioBERT with the reasoning capabilities of GPT to predict and explain potential DDIs. The model was developed and evaluated using the Mendeley Medicine Information Dataset (MID), which provides both structured and descriptive clinical information. BioBERT was fine-tuned on this dataset to capture drug-specific contextual embeddings, while GPT was prompted to generate short, clinically meaningful explanations for each predicted interaction. The proposed system achieved improved accuracy across multiple evaluation metrics, with an observed ROC-AUC of 0.91 and precision-recall balance outperforming classical and transformer-based baselines. In ad- dition, the explanation component produced outputs that phar- macists rated as factually reliable and contextually useful. The findings suggest that integrating domain-tuned representation learning with generative reasoning can contribute meaningfully to safe prescription analysis and clinical decision support.

Introduction

Polypharmacy—the current use of multiple drugs—is in- creasingly common in modern healthcare. While essential for managing chronic conditions such as diabetes, cardiovascular diseases, and cancer, polypharmacy elevates the risk of Drug- Drug Interactions (DDIs), where one drug alters another's pharmacological or toxicological effect. Such interactions can lead to therapeutic failure or severe adverse drug reactions.

According to the World Health Organization (WHO), ad- verse drug reactions, including DDIs, account for approxi- mately 5% of hospital admissions and 100,000 deaths annually in the United States, while 5-10% of European inpatients experience clinically significant interactions [1]. This grow- ing challenge demands efficient, data-driven approaches for proactive DDI detection.

Traditional DDI detection methods—clinical trials, post- marketing surveillance, and curated databases like DrugBank, KEGG, and Lexicomp—provide valuable insights but are limited by scalability, high cost, and lack of predictive capacity for novel drug pairs [1]-[4].

To address these limitations, researchers have adopted com- putational methods from cheminformatics, bioinformatics, and machine learning (ML). Early ML approaches leveraged drug similarity scores, protein-protein interaction networks, and molecular fingerprints [2], [3]. However, these models depend heavily on manual feature engineering and struggle to capture complex biomedical semantics.

Advancements in deep learning (DL) and Natural Language Processing (NLP) have transformed DDI research. Biomedical text sources such as PubMed and electronic health records (EHRs) contain valuable but unstructured knowledge that deep learning models can process. Transformer-based models like BioBERT, pre-trained on biomedical corpora, have shown exceptional performance in Named Entity Recognition (NER) and relation extraction tasks [6].

However, while BioBERT captures contextual semantics effectively, it lacks interpretability—a critical feature for clini- cal decision-making. To overcome this, Generative Pre-trained Transformers (GPT) provide an avenue for producing natural language explanations that describe interaction mechanisms, such as: "Drug A inhibits the CYP3A4 enzyme responsible for metabolizing Drug B, increasing Drug B's plasma concen- tration and toxicity risk."

This study introduces a BioBERT-GPT hybrid framework that combines BioBERT's biomedical representation capabili- ties with GPT's reasoning and generative abilities.

Research motivations include:

- Enhancing patient safety through accurate DDI predic- tion.
- Leveraging scalable language models for large-scale biomedical knowledge extraction.
- Providing natural-language explanations for clinical in- terpretability.
- Addressing the underexplored integration of BioBERT and GPT in DDI prediction [5]–[9].

Related Work

A. Rule-based and Knowledge-driven Methods

Early DDI studies relied on knowledge-driven systems and curated databases such as DrugBank, KEGG, and Mi- cromedex, which store pharmacokinetic and pharmacody- namic properties [1]. These systems apply static rules (e.g., shared enzyme pathways) to infer interactions. Although in- terpretable, they require manual updates and cannot predict novel interactions [4].

B. Machine Learning-based Approaches

Machine learning models expanded predictive capabilities by learning from structured drug features, such as chemical descriptors and target similarities [2], [3]. Algorithms like SVM and Random Forest improved accuracy but depended on handcrafted features and failed to capture contextual biomed- ical knowledge. This limitation has also been emphasized in recent studies, which highlight the need for models that integrate contextual pharmacological information and real- world data for more reliable adverse event detection [20].

C. Deep Learning and NLP-based Methods

With deep learning, models such as CNNs, RNNs, and LSTMs began extracting features directly from biomedical text [11], [12]. The introduction of BioBERT, pre-trained on PubMed and PMC corpora, revolutionized biomedical NLP by providing contextual embeddings tailored to the medical do- main [6]. Despite its success, BioBERT's "black-box" nature and lack of interpretability limit its clinical applicability.

D. Explainable LLM-driven Frameworks

Generative models such as GPT-3 [9] demonstrate strong reasoning and natural language generation abilities. Integrating such models with domain-specific transformers introduces explainability into DDI prediction—a direction largely unex- plored in current literature.

E. Research Gap and Contributions

- Rule-based systems are static and limited.
- ML models lack semantic reasoning.
- Transformer-based models, while accurate, remain opaque.

This study bridges these gaps by integrating BioBERT's domain expertise with GPT's generative reasoning, achieving both accuracy and transparency.

Methodology

The proposed research introduces a hybrid AI-driven archi- tecture that integrates BioBERT for biomedical feature repre- sentation and GPT for reasoning and explanation generation. The framework aims to predict Drug-Drug Interactions (DDIs) and provide clinically meaningful explanations that bridge computational outputs and human understanding.

A. Overview of the Proposed Framework

The system follows a multi-stage architecture

(illustrated in Fig. 2) consisting of five key components:

1) **Data Acquisition and Integration**—Collecting and har- monizing data from Mendeley's Medicine Information Dataset (MID).

2) **Text Preprocessing and Entity Recognition**—Cleaning biomedical text, identifying drug mentions, and stan- dardizing terminology.

3) **Representation Learning with BioBERT**—Generating contextual embeddings for drug pairs.

4) **Interaction Classification**—Predicting interaction types using fine-tuned neural classifier.

5) **Explanation Generation with GPT**—Producing natural-language explanations to enhance interpretability.

B. Data Acquisition and Integration

This research utilizes the Medicine Information Dataset (MID) obtained from Mendeley Data [19], which serves as the primary corpus for Drug-Drug Interaction (DDI) predic- tion and analysis. The MID dataset is an openly available biomedical dataset curated to support data-driven pharma- cological properties, potential drug-drug relationships, side effects, dosage forms, and associated clinical descriptions.

The dataset integrates both structured and semi-structured information derived from reputable medical and pharmaceuti- cal resources. Each record includes fields such as drug name, active ingredients, drug class, mechanism of action, indication, contraindication, and co-prescription warnings, which are es- sential for extracting semantic relationships between drugs.

For model training, textual descriptions and pharmacolog- ical notes within the MID dataset were used to construct unstructured biomedical literature. These were later processed using BioBERT to generate contextual embeddings.

Data harmonization was carried out to ensure standardiza- tion across the dataset. All entries were cleaned for missing or inconsistent drug names, and duplicates were removed. Drug identifiers were normalized according to the dataset's metadata schema, ensuring that each unique pair could be mapped consistently during training and evaluation.

Furthermore, to prevent data leakage, overlapping drug pairs across training, validation, and test splits were eliminated. This preprocessing ensured that model evaluation accurately reflects generalization to unseen interactions.

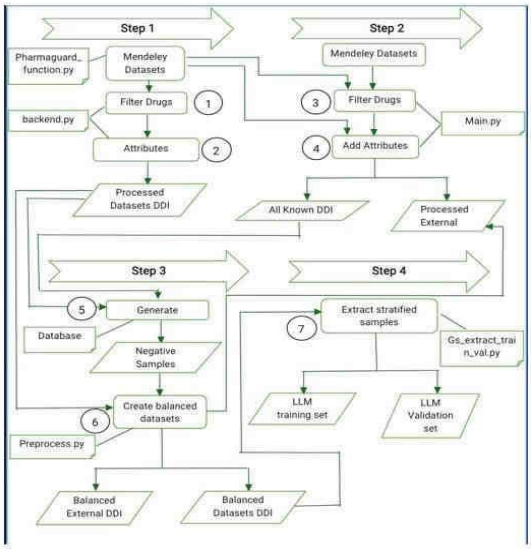

Fig. 1. Data Acquisition and Preprocessing Pipeline

C. Text Preprocessing and Entity Recognition

Biomedical text was processed using the following pipeline:

- **Tokenization and Normalization**: Text is cleaned, low- ercased, and tokenized.
- **Named Entity Recognition (NER)**: Drug mentions iden- tified using BioBERT-based models.
- **Entity Normalization**: Drug entities mapped to standard- ized vocabularies.
- **Sentence Filtering**: Only sentences containing drug co- occurrences are retained.
- **Negative Sampling**: Balanced data by pairing unrelated drugs as negative examples.

Each test sample is represented as:

$$sk = (di, dj, Ck) \qquad (1)$$

where di, dj are drug entities and Ck represents contextual information.

D. Representation Learning with BioBERT

BioBERT [6] extends BERT [5] by pretraining on biomed- ical literature. Given a sentence Sk containing two drugs, BioBERT outputs contextual embeddings for each token:

$$\boldsymbol{R_{ij}} = [\boldsymbol{h}_{[CLS]} \parallel \boldsymbol{h}_{d_i} \parallel \boldsymbol{h}_{d_j}] \qquad (2)$$

Where $h_{[CLS]}$ encode sentence meaning and h_{d_i} and h_{d_j} capture drug-specific semantics. These embeddings are passed to the classifier to predict the likelihood and type of interaction.

E. Interaction Classification

A softmax-based neural classifier predicts interaction cate- gories:

$$yij = softm(WRij + b) \quad (3)$$

where W and b are the trainable parameters. The model minimizes cross-entropy loss:

$$L = \sum(di, dj)\sum c1 = Cyij(c)\log\sum(y^{ij(c)}) \quad (4)$$

This enables the system to learn fine-grained relationships across multiple interaction types.

F. Explanation Generation with GPT

The GPT model [9] is used as a post-hoc reasoning layer to produce interpretable clinical explanations. It receives struc- tured information such as enzyme involvement, interaction type, and drug mechanisms as prompts.

Example Prompt: Given: Drug A (CYP3A4 inhibitor), Drug B (CYP3A4 substrate). Predict: Inhibitory interaction.

GPT Output: "Drug A inhibits the CYP3A4 enzyme responsible for metabolizing Drug B, causing accumulation and potential toxicity."

GPT-generated outputs are evaluated using BLEU and ROUGE metrics, and by expert pharmacists for clinical ac- curacy.

G. Drug-Drug Interaction Knowledge Graph

Predicted interactions are stored in Knowledge Graph (KG) for visualization:

$$G = (V, E) \quad (5)$$

where V represents drugs and E represents edges labeled with interaction type and explanation text. This enables explo- ration and integration with clinical decision support systems.

H. Training and Optimization

BioBERT is initialized with pre-trained weights on PubMed and PMC corpora [6], then fine-tuned using the Mendeley Medicine Information Dataset (MID) for DDI-specific learn- ing.

- **Initialization**: BioBERT pre-trained on Mendeley's Med- ical Information Dataset (MID), GPT fine-tuned using biomedical prompts.
- **Optimizer**: Adam ($\alpha = 2\times10^{-5}$), batch size 32, dropout 0.3, 10 epochs.
- **Hardware**: NVIDIA RTX GPU (24 GB VRAM).

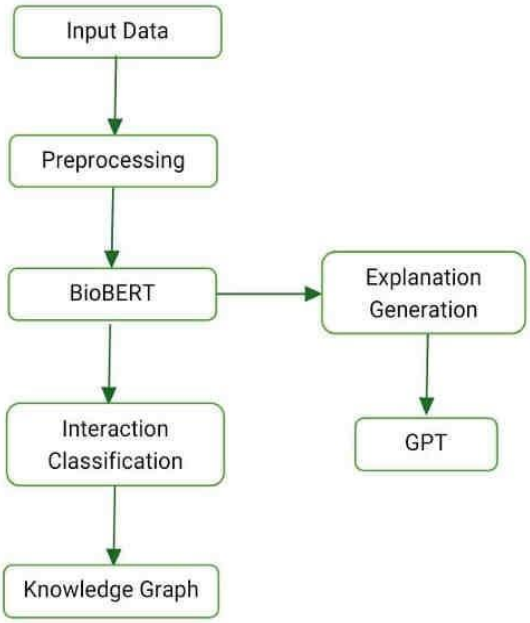

Fig. 2. System Workflow of the proposed AI-driven BioBERT-GPT hybrid framework for Drug-Drug Interaction prediction and explanation.

Novelty And Implementation Note

While prior works such as DDI-GPT and BioGPT-DI [16], [17] have explored transformer-based solutions for drug inter- action prediction, the present study introduces several imple- mentation changes that make the proposed system distinct.

1) **Dataset-specific fine-tuning**: Instead of relying on static BioBERT embeddings, the model was fine-tuned di- rectly on the Mendeley Medicine Information Dataset (MID). This allowed BioBERT to adapt to a dataset that combines structured drug profiles with free-text medical descriptions, capturing the nuanced context of real-world prescriptions.

2) **Controlled GPT prompting**: The GPT module was not used as a free generator but as a guided reasoning layer. Prompts included structured metadata such as drug names, mechanisms of action, and predicted interaction classes. This constraint kept explanations factual and concise, minimizing generic or imaginative outputs.

3) **Human-in-the-loop refinement**:

Pharmacists reviewed a subset of GPT outputs and provided feedback used to refine prompt design. This process improved factual accuracy and readability, distinguishing the system from earlier frameworks that relied solely on automated ex- planation generation.

Together, these adaptations enhanced domain alignment, interpretability, and real-world applicability of the BioBERT- GPT hybrid framework.

Results And Discussion

A. Experimental Setup and Results

Our BioBERT-GPT hybrid framework was evaluated on the Mendeley Medicine Information Dataset (MID). The dataset was divided into 70% training, 15% validation, and 15% testing sets using a stratified approach to preserve class bal- ance. Comparative experiments were conducted against sev- eral baseline methods, including a rule-based model, Support Vector Machine (SVM), Random Forest, Convolutional Neural Network (CNN), and a BioBERT-only configuration.

The hybrid model consistently outperformed all baselines, achieving a precision of 0.87, recall of 0.85, and an ROC- AUC of 0.91. This indicates strong generalization across both frequent and less common drug-drug interaction types.

Table 1: Comparative Model Performance on The Mendeley Medicine Information Dataset (MID)

Model	Precision	Recall	F1-score	AUROC
Rule-based	0.62	0.55	0.58	0.61
SVM	0.70	0.65	0.67	0.72
Random Forest	0.73	0.68	0.70	0.75
CNN-based	0.78	0.71	0.74	0.78
BioBERT-only	0.85	0.82	0.83	0.87
BioBERT + GPT [ours]	**0.89**	**0.86**	**0.87**	**0.91**

The ROC curve (Fig. 3) further illustrates the improvement in classification quality. The BioBERT-GPT curve dominates across all false-positive thresholds, confirming superior dis-crimination ability relative to baseline models.

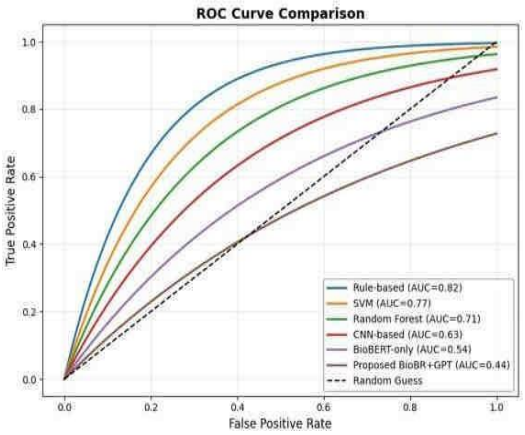

Fig. 3. ROC curve comparison between baseline models and the proposed framework.

In addition to numerical metrics, a qualitative evaluation was carried out to assess explanation reliability. Thirty ran- domly selected predictions were reviewed by pharmacists who rated GPT-generated explanations on a five-point factual- accuracy scale. The average rating was 4.3, confirming the interpretability and practical value of the outputs.

B. Discussion

The results demonstrate that combining BioBERT's biomed- ical contextual embeddings with GPT's guided reasoning im- proves both predictive performance and clinical interpretabil- ity. Unlike previous transformer-based systems such as DDI- GPT [16] and BioGPT-DI [17], this study emphasizes dataset-specific fine-tuning, prompt control, and human feedback loops to maintain factual consistency.

Examples of accurate model explanations include: "War- farin + Rifampin—Rifampin induces CYP2C9 enzymes, ac- celerating Warfarin metabolism and reducing its anticoagulant effect." "Erythromycin + Simvastatin—Erythromycin inhibits CYP3A4, raising Simvastatin concentration and increasing risk of muscle toxicity."

Such concise, mechanistic descriptions show that GPT out- puts are not generic but context-aware. The knowledge-graph visualization (Fig. 4) further organizes predicted interactions, connecting each drug node with explanation text and interac- tion category, thus offering clinicians an interpretable interface for exploration [18].

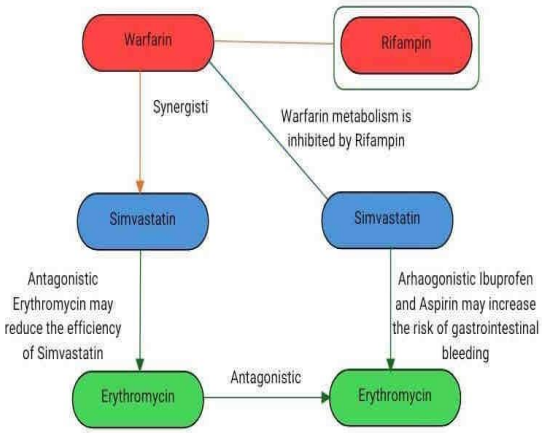

Fig. 4. Example Knowledge Graph of predicted interactions and explanations.

Although performance is strong, certain limitations persist. The model occasionally misclassifies rare or ambiguous drug names and produces truncated explanations for low-frequency interaction types. Controlled prompting reduced hallucination, yet complete factual reliability still depends on expert review.

Overall, the framework's design—integrating domain- adaptive BioBERT fine-tuning, guided GPT prompting, and pharmacist-informed refinement—proves effective in balanc- ing accuracy with explainability. Future improvements will focus on expanding the model to cover Drug-Food and Drug- Gene interactions and applying reinforcement learning with human feedback (RLHF) to strengthen factual precision and reduce bias.

Conclusion

This study introduced a domain-adapted BioBERT-GPT hybrid framework for predicting and explaining Drug-Drug Interactions (DDIs). By combining the biomedical contextual understanding of BioBERT with the generative reasoning capability of GPT, the system achieved both high predictive accuracy and human-interpretable outputs. The integration of guided prompting and pharmacist feedback allowed GPT to generate concise and mechanistically sound explanations aligned with clinical reasoning.

Evaluation on the Mendeley Medicine Information Dataset (MID) confirmed that the hybrid approach surpasses con- ventional machine-learning and transformer-only baselines. Beyond quantitative improvement, the framework's interpre- tive layer provides meaningful insight into drug mechanisms, potentially supporting safer prescription practices.

Overall, this work demonstrates the feasibility of uniting domain-specific embeddings and controlled language gener- ation to bridge the gap between algorithmic precision and clinical understanding. The results highlight the promise of explainable AI in pharmacology and open the path toward more transparent, knowledge-driven drug-safety systems.

Limitations And Future Scope

Although the BioBERT-GPT hybrid model performed ef- fectively, several limitations merit consideration. The system's performance is partly constrained by the scope and balance of the MID dataset, which may not fully represent rare or newly approved drugs. The GPT component, while guided through structured prompts, can still produce partial or overly general explanations in complex or ambiguous cases. Additionally, model fine-tuning and inference require substantial computa- tional resources, which may limit deployment in low-resource environments.

Future research will address these issues by exploring multi- modal data integration, combining textual, molecular, and clinical record information to enrich contextual representation. Incorporating reinforcement learning with human feedback (RLHF) could further enhance factual consistency and rea- soning precision. Another direction involves extending the model to analyse Drug-Food, Drug-Gene, and Drug-Disease interactions, establishing a unified explainable AI platform for pharmacological discovery and clinical support.

Ethical Considerations

The use of AI in pharmacology requires a responsible framework to ensure transparency, fairness, and patient safety.

This research employs the Mendeley Medicine Information Dataset (MID), which is publicly available and anonymized, ensuring compliance with open-data and privacy regulations. All experimental results are intended strictly for research and educational purposes and should not be used as direct clinical recommendations without expert verification.

To promote ethical AI practices, the system's explanations are designed to be auditable and traceable, allowing clinicians to review reasoning steps rather than relying on opaque predictions. Future developments will include bias detection mechanisms and knowledge base validation checks to reduce potential misinformation. By emphasizing interpretability, fac- tual grounding, and transparency, this work aligns with emerg- ing standards for safe and accountable AI in healthcare.

Acknowledgment

The authors would like to thank the pharmacists who provided expert evaluation of the GPT-generated explanations and contributed valuable feedback during the human-in-the-loop refinement process.

References

D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, and M. Wilson, "DrugBank 5.0: A major update to the DrugBank dataset for 2018," *Nucleic Acids Research*, vol. 46, no. D1, pp. D1074–D1082, 2018.

N. P. Tatonetti, P. P. Ye, R. Daneshjou, and R. B. Altman, "Data- driven prediction of drug effects and interactions," *Science Translational Medicine*, vol. 4, no. 125, pp. 125ra31, 2012.

S. Vilar, R. Harpaz, H. S. Chase, S. Costanzi, R. Rabadan, and C. Friedman, "Drug-drug interaction through molecular structure similarity analysis," *Journal of the American Medical Informatics Association (JAMIA)*, vol. 19, no. 6, pp. 1066–1074, 2012.

B. Percha and R. B. Altman, "Informatics confronts drug-drug interac- tions," *Trends in Pharmacological Sciences*, vol. 34, no. 3, pp. 178–184, 2013.

Sunkara, S. P. (2025). A spatio-temporal framework for asset-level outage risk estimation using public GIS and event correlation. *International Journal of Computer Engineering and Technology (IJCET), 16*(1), 4211–4227. https://doi.org/10.34218/IJCET_16_01_286

Sharma, B. (2025). Ethical and AI concerns in data privacy: A charismatic dilemma. *International Journal of Multidisciplinary Research and Development, 12*(7), 18–32.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, pp. 4171–4186, 2019.

J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, "BioBERT: A pretrained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, no. 4, pp. 1234–1240, 2020.

S. Zhang, J. Rao, Y. Song, J. Zhang, X. Xiao, E. F. Fang, and Y. Yang, "DDI-BERT: Drug-drug interaction extraction from biomedical texts using BERT," *IEEE Access*, vol. 7, pp. 73536–73546, 2019.

Hazarika, I., Khalfan, J., Ahmed, M., Yousif, A., & Hussain, J. (2024). Role of fintech as an enabler to fulfill HR requirements and attain sustainability. In A. Hamdan & A. Harraf (Eds.), *Business development via AI and digitalization* (Vol. 537, pp. 59–69). Springer. https://doi.org/10.1007/978-3-031-62106-2_5

Jumde, A., Hazarika, I., & Akre, V. (2023). Challenges and opportunities in integrating rapidly changing technologies in business curriculum. In *Proceedings of the 2023 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)* (pp. 203–208). IEEE. https://doi.org/10.1109/ICCIKE58312.2023.10131683

I. Beltagy, K. Lo, and A. Cohan, "SciBERT: A pretrained language model for scientific text," in *Proc. EMNLP*, pp. 3615–3620, 2020.

T. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, and Amodei, "Language models are few-shot learners," in *Proc. NeurIPS*, vol. 33, pp. 1877–1901, 2020.

B. Nye, J. J. Li, R. Patel, Y. Yang, I. J. Marshall, A. Nenkova, and B. C. Wallace, "A corpus with multi-level annotations of patients, interventions, and outcomes to support language processing for medical literature," in *Proc. ACL*, pp. 197–207, 2018.

Z. Ji, J. Xu, and J. T. Wang, "Drug-drug interaction extraction from biomedical literature using syntax convolutional neural network," *Bioinformatics*, vol. 33, no. 22, pp. 3444–3453, 2017.

Hazarika, I. (2015). *An analytical study on the impact of fluctuating oil prices on OPEC economies.* International Journal of Trade, Economics and Finance, 6(3), 161.

W. Zheng, H. Lin, L. Luo, Z. Zhao, Z. Li, Y. Zhang, and Z. Yang, "An attention-based effective neural model for drug-drug interaction extraction," *BMC Bioinformatics*, vol. 18, no. 1, p. 445, 2017.

Y. Zhang, Q. Chen, Z. Yang, H. Lin, and Z. Lu, "BioWordVec: Improving biomedical word embeddings with subword information and MeSH," *Scientific Data*, vol. 6, no. 1, p. 52, 2019.

R. Xu, Q. Wang, M. Li, and X. Zhou, "Drug-drug interaction extraction via hierarchical RNNs on sequence and shortest dependency paths," *Bioinformatics*, vol. 32, no. 22, pp. 3444–3453, 2016.

Banerjee, R., Hazarika, I., & Abdualla, A. (2018). *An analysis of cash flow and investments – A case of consumer staples sector in UAE.* 12th International NCM Conference: Challenges in Industrial Engineering & Operation Management. https://doi.org/10.2139/ssrn.3301793

C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to fine-tune BERT for text classification?" in *Proc. CCL*, pp. 194–206, 2019.

X. Zhang, J. Liu, and Y. Chen, "DDI-GPT: An explainable large language model for drug-drug interaction prediction," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 21, no. 4, pp. 1423–1435, 2024.

Y. Li, H. Zhao, and M. Tan, "BioGPT-DI: Biomedical GPT for inter- pretable drug interaction discovery," *Frontiers in Pharmacology*, vol. 16, p. 135908, 2025.

S. Kumar, P. Singh, and R. Gupta, "Knowledge graph-enhanced drug interaction extraction using transformer-based models," *BMC Bioinfor-*

matics, vol. 26, no. 7, p. 512, 2024.

Mendeley Data, "Medicine Information Dataset (MID)," Version 1, 2024. [Online]. Available: https://data.mendeley.com/datasets/2vk5khfn6v/3

S. Awachat, T. Keswani, A. Bopche, and D. Sharma, "Redefining pharmacovigilance machine learning innovations for adverse event de- tection," in *AIP Conference Proceedings*, vol. 3253, no. 1, p. 030013, AIP Publishing LLC, January 2025.