Proceedings on Engineering Sciences

NAMED ENTITY RECOGNITION FOR MEDICAL DATA EXTRACTION USING BIOBERT

Kayal Padmanandam[1]
Nikitha Pitla
Yeshasvi Mogula

A B S T R A C T

Technological advancements have caused widespread shifts in the medical industry. A vast quantity of information may be found in the medical literature publications released by researchers. Natural language processing innovations have made it simple to extract information on drugs, illnesses, symptoms, routes doses, species, and routes of administration from a documented source. This proposed research is used to identify the named entities from the medical literature. The BioBERT model is used to train the corpus that has been annotated. The proposed framework can outperform many state-of-the-art baselines and provide state-of-the-art results for BioNER. When compared to the existing model, the accuracy provided by the proposed system is satisfactory. The BioBERT is used for the extraction of medical entities like drugs, chemicals, genes, etc and to train the corpus that has been annotated. The BioBERT has in-built data of the medical entities, that will identify all the medical-related data or entities from the given statement. More entities can be found by this method than by the current standard model.When compared to the existing model, the accuracy provided by the trained version is satisfactory.The proposed system comes with a GUI for users to type the clinical words or components for analysis.

1. INTRODUCTION

Subsequent words in a sentence can be predicted and generated with the help of Natural Language Processing by analyzing and learning from the previous words in the sentence. Like many NLP processes, Named Entity Recognition (NER) extracts important entities like people, places, organizations, and medical terminology from the given text or document. Python libraries like Spacy, SciSpacy, and *BioBERT* are used for NER practices. Medical Information on drugs, illnesses, and symptoms can be extracted from the textual source using the SciSpacy or the *BioBERT* model. However, there is a limitation that users cannot create user-defined entities. To overcome this, the BlankSpacy model can be used to train the user-defined annotated entities (Alam & Schmeier, 2021). Many clinical Natural Language Processing (NLP) tools and systems have been published (Tarcar et al., 2019; Xiaodong Liu et al., 2019). The growing quantity of textual biomedical information allows for the pre-training of language (Tawfik & Spruit, 2020), which may then be used for a wide range of tasks in the biomedical domain. The practice of automatically recognizing and categorizing names of

---
[1] Corresponding author: Kayal Padmanandam
Email: kayalpaddu@gmail.com

entities like the persons, locations, organizations, nations, day date, time of the day or year, numbers, ordinal numbers, and cardinal numbers in the text which have given is known as Named-Entity Recognition (NER) (Devopedia, 2020). In the biomedical discipline, a large variety of entities such as diseases, drugs, genes, and proteins are available and can be recognized using the BioBERT Model. BIONER is used in several text mining applications in the healthcare industry, such as drug-drug interaction extraction and disease-treatment relationship discovery (BioNER). Models used in biomedical text mining have recently been improved and there is development using deep learning (Ramachandran & Arutchelvan, 2021) techniques in NLP. In recent years, several different methods have been applied to improve BioNER (Yadav & Bethard, 2019).

A few of the methods are LSTM (Long Short-Term Memory) and Conditional Random Field (CRF). Although there are several implementations, limitations to biomedical text mining when using state-of-the-art natural language processing techniques and graph-based algorithms (Peters et al., 2018) exist. Word2Vec (Devlin et al., 2019), ELMo (Lan et al., 2019), BERT (Beltagy et al., 2019), and ALBERT (Mikolov et al., 2013) are a few word representation approaches that provide their own set of difficulties when evaluating their efficacy on biological text datasets.
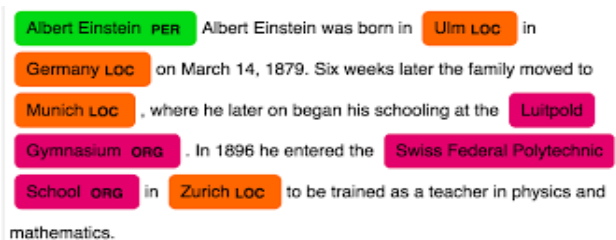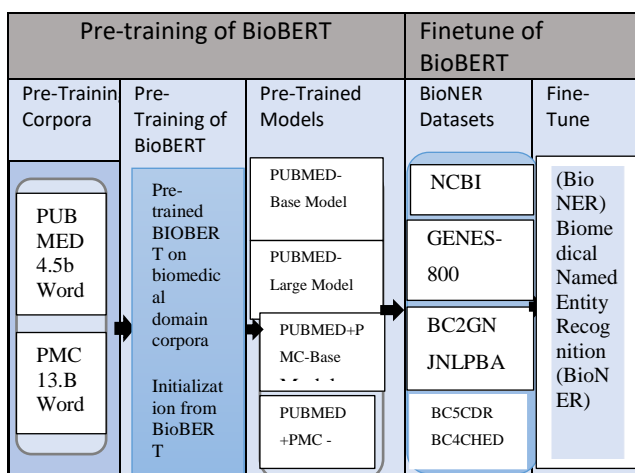

Figure 1. Sample output of NER

Figure 2. Overview of the pre-training and fine-tuning of BioBERT on NER

NLP Libraries:

Figure 1 shows the named entity recognized, where the person's name is identified as PER, the Location is identified as Loc, and the organization is identified as ORG. Figure 2 depicts the overview of the pre-training and fine-tuning of BioBERT on NER.

Spacy:

Python's Spacy is a powerful Natural Language Processing (NLP) module that is available and open-source. Spacy is optimized for production use and may be used to create programs that understand and analyze enormous amounts of text. It may be utilized as a pre-processing step for deep learning and the development of information extraction and natural language comprehension systems.

Scispacy:

An open-source and free library for sophisticated Natural Language Processing, Scispacy was developed in Python and CPython. Statistical neural network models for processing biomedical, scientific, or clinical literature are now available in the library, which is licensed under the MIT license.

BioBERT:

Bidirectional Encoder Representations from Transformers (BERT) is one of the deep learning models that takes its cues from Transformers, in which each output is linked to each input and the weights between them are created on the fly depending on their relationship. Figure 3 shows how the BioBERT Architecture works.
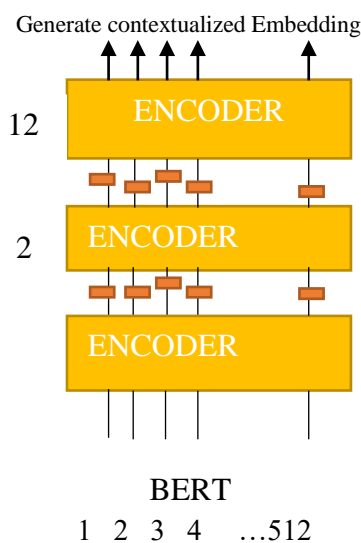

Figure 3. BioBERT Architecture

2. RELATED WORKS

An effective method for learning high-level distributed vector representation for the semantic word relations is captured by a large number of words. The regular word representation is used to speed up and learn the frequent words by subsampling (Mikolov et al., 2013). BioSentVec represents the initial launch of an open set of sentence embeddings, developed using a vast corpus exceeding 30 million scientific documents from the PubMed database, in conjunction with clinical data derived from the MIMIC II Clinical Database. In this study, BioSentVecs embeddings are evaluated using two separate phrase pair similarity tasks in a variety of biomedical text genres (Chen et al., 2019). Rather than viewing BioNER tasks through the lens of sequence labelling, the study by (Pyysalo et al., 2013) reframes them as problems of machine reading comprehension (MRC). This conceptual reorientation allows for the integration of additional prior knowledge through the utilization of carefully crafted queries, thereby eliminating the necessity for decoding procedures such as conditional random fields (CRF). The technique of bacterial-named entity recognition is also discussed, which incorporates domain-specific features into a sophisticated deep learning framework. This framework merges the capabilities of a convolutional neural network with a bidirectional long short-term memory network. (Zhu et al., 2018) can attain superior performance in bacterial NER (Devopedia, 2020). Yuqi Si et al. (2019) obtained superior performance in bacterial NER by utilizing the domain properties (Yuqi Si et al., 2019). In the study referenced by (Lee et al., 2016), the authors found that contextual embeddings, which were pre-trained on an extensive clinical dataset, set new benchmarks in all concept extraction tasks. The top-performing model outperformed all existing leading methods, which had F1 scores ranging between 80.74 and 81.65, with 80.74 being the highest score previously recorded.

In the study (Peng et al., 2019), it is posited that BioBERT, pre-trained on biomedical datasets, outstrips not only BERT but also other previous leading models. This can be credited to BioBERT's architecture which remains consistent across tasks. Though BERT matches the performance of other top-notch models, BioBERT markedly surpasses them in three standard biomedical text-mining tasks. Further research conducted by (Howard & Ruder, 2018) evaluated multiple baselines based on BERT and ELMo, concluding that the BERT model, pre-trained using PubMed abstracts and MIMIC-III clinical notes, delivered the optimal outcome. Meanwhile, (Yogatama et al., 2019) broach the subject of Universal Language Model Fine-tuning (ULMFiT), a transfer learning methodology efficacious for an array of tasks in natural language processing, and describe the pertinent strategies. The study identified (Qiu et al., 2020) executing a comprehensive empirical examination to ascertain performance against a crucial benchmark. The capacity to quickly adjust to new tasks by leveraging prior knowledge about a language's vocabulary, grammar, meaning, and practical rules is considered the essence of broad linguistic intelligence. Finally, Liu et al. (2016) present a succinct overview of language representation learning and its accompanying research trajectories. (Ling, 2023) research develops natural language processing (NLP) to analyze drug reviews of patient with their level of satisfaction.

3. DATASET FOR MEDICAL NAMED ENTITY RECOGNITION

The proposed system employs the Scispacy and Biobert models to identify medical phrases in text documents. It will categorize entities like Disease, drug/Chemical, Gene, etc., as shown in Figure 6, and highlight the entities with the help of displacy in the statements.

Dataset

The model has built on the following different datasets and the entities identified are the drug/chemical, the disease, and the Gene.

BC5CDR: This corpus, available through the Bio-Creative workshop (Li et al., 2016), is a challenge set by the Bio-Creative community for the extraction of chemical-disease relations. It aids in identifying chemical and disease entities from Medline abstracts.

BC4CHEMD: This dataset comprises chemical compounds and drugs extracted from PubMed abstracts. It's provided by the Bio-Creative challenge for the development and evaluation of tools for Chemical NER (Krallinger et al., 2015). The BioNER datasets include Annotations of the NCBI Disease, which contains 6,881 values.

NCBI Disease: This dataset provides information on disease-related data from the National Center for Biotechnology Information (NCBI). It is used to access biomedical and genomic data, as mentioned in the release by Leaman et al (Islamaj et al., 2014).

JNLPBA: This dataset, provided by (Kim et al., 2004), is employed to identify DNA, cell type, RNA, protein, and cell line.

BC2GM: Bio Creative II Gene Mention Recognition (BC2GM) is a semi-supervised learning task that uses alternating structure optimization. This dataset, provided by Ando (2007), holds the state-of-the-art in the BC2GM task.
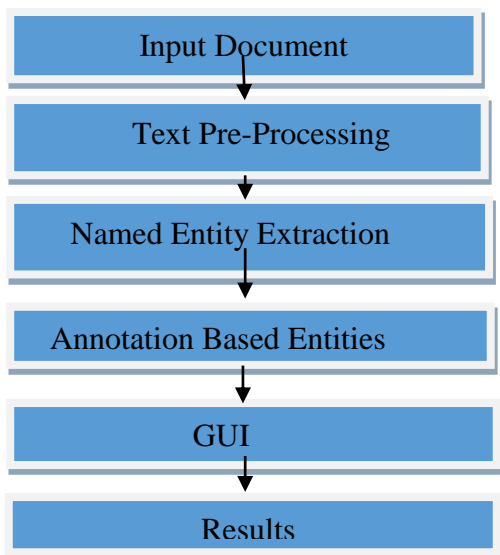
Figure 4. Proposed NER Flow Diagram

Table 1. Dataset Table

Name	Entity Types	Train Sentence	Train entity	Test Sentenc	Test Entity
BC2GM	Gene/ Protein	14093	18257	5038	6325
BC4CHEMD	Drug/ Chem	10321	58964	26364	25346
BC5CDR-chem	Drug/ Chem	7041	10550	4797	5385
BC5CDR-disease	Disease	7031	8427	4797	4424
NCBI-Disease	Disease	5217	5921	.0822	940
JNLPBA	Gene/ Protein	12326	40753	3856	6241

Figure 5. Dataset accessed from https://www.kaggle.com/datasets/tboyle10/medicaltranscriptions

The above Figure 5 shows the dataset which is used with Scispacy. It has the statement description, transcripts, and keywords. According to the keywords, the entities will be identified and they will be annotated.

Figure 6. Biomedical entities identified

4. METHODOLOGY

4.1 NLP Pipeline

1. Sentence Segmentation: It is the foremost step in the Natural Language Processing (NLP) pipeline. It involves breaking down a large body of text into its individual sentence constituents for easier processing. Take, for instance, the sentence, "Hyderabad is the capital and most populous city of Telangana and also for AP before the Telangana formation. It is also referred to as the pearl city. It was once under the rule of the Nizams, who constructed the renowned Charminar." In this case, sentence segmentation would separate this text into distinct sentence units for further analysis.

2. Word Tokenization: Word tokenization divides the large statements to the small phases or the tokens breaks the sentence into separate phrases or tokens. This helps apprehend the context of the textual content. While tokenizing the above sentence " Hyderabad is the capital and most populous city of Telangana and the city ", the sentence will be separated into single words, i.e., "Hyderabad", "is", "the", "capital", "and", "most", "populous", "city", "of", "Telangana", "and", "the", "pearl", "City".

3. Stemming: It simplifies words to their root or base form, a process known as normalization thereby enabling more precise prediction of the parts of speech for each token.

4. Lemmatization: It refines words by removing inflectional endings and returning the base or dictionary form of a word, also known as the lemma. Unlike stemming, which may result in non-standard words, lemmatization ensures the output is a valid term or phrase found in the language.

5. Stop word analysis: Stop word analysis involves evaluating the relevance of each word in a sentence. Some words, like "is," "a," "the," and "and," are more common in English and often carry less meaningful information. These are designated as "stop words" in natural language processing (NLP). As they occur frequently but don't contribute much to the overall meaning, they can be filtered out to highlight more critical, information-rich words.

6. Dependency parsing: It is primarily utilized to decipher the interrelationships among words in a sentence. It works by creating a tree to visualize these relationships, with each word represented as a node. The central verb of the sentence serves as the root of this tree, signifying its foundational role in the overall sentence structure.

7. Part-of-speech tagging: It involves assigning grammatical categories, such as verbs, adverbs, nouns, and adjectives, to each word in a sentence. This process aids in comprehending the linguistic and semantic significance of words, thereby ensuring grammatical correctness and enhancing sentence understanding.

The CRF, employed by both the NLP Speech Tagger and NLTK, represents a statistical model suitable for handling sequential data, like words. It is capable of capturing a comprehensive understanding of a sentence's context. In this model, the input is $X = \{\vec{x_1}, \vec{x_2}, \vec{x_3}, \cdots / \vec{x_T}\}$.

$$P(y - x) = \frac{1}{z(\vec{x})} \prod_{t=1}^{T} \quad exp\{\sum_{k=1}^{k} \quad \omega_k f_k(y_t, y_t-1, \vec{x_t})\} \qquad (1)$$
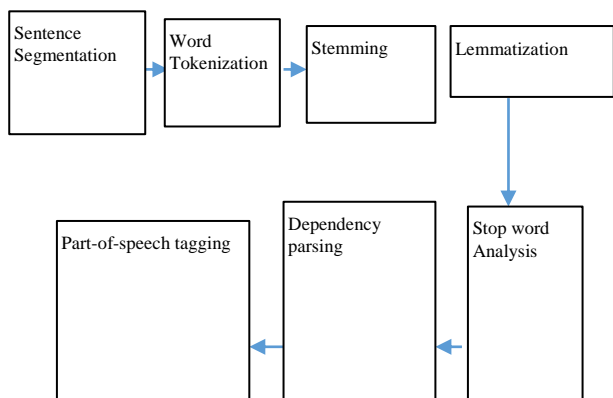
Figure 7. NLP Pipeline

4.2 Pre-trained models and hyper parameter tuning

Deep learning technique, the hyperparameters of the BERT model must be initialized. On the other hand, finding the sweet spot at the beginning might be a challenge. Modifications may be made after comparing the model's results with data. The improved model may be referenced from the NER database of the transformers. Like training any other deep learning network, the approach is easy. A loop that takes into consideration the intended training length, or epochs, and the availability of graphics processing units (GPUs), will be created.

4.3 NER Model Accuracy Prediction

A NER model's effectiveness may be estimated in a variety of ways. This is commonly done by the use of measurements like the Precision, Recall, F1 score, or relaxed match. The below Figure 8 show the relation between the drug and the dosage of the drug.

Figure 8. Relation between the drug and the dosage

5. EXPERIMENTAL RESULTS

BioBERT is a space-explicit language portrayal model optimized for large-scale biomedical corpora, which we call a "transformer for biomedical message mining." BioBERT, which uses almost identical engineering for different tasks, outperforms the BERT and other state-of-the-art models when trained on a large number of biomedical corpora, the biomedical text mining tasks. When compared to prior-generation models, BERT's performance is on par, but BioBERT outperforms them in three key areas, delegate biomedical message named entity recognition, connection extraction, and question-response rate. The study has shown that BERT can better understand difficult biomedical texts if it is first trained on relevant biomedical corpora. The system is evaluated using the performance measures such as precision, recall, and F1-score as given in the equation 1,2,3. The two different algorithms Scispacy and the BioBERT are compared. The model has been loaded with 3 different datasets. A comparison study between the built-in model Scispacy and the customized model using BioBERT is made and the results are shown in Table 2 and 3.

Table 2. BioBERT Training and Testing Result

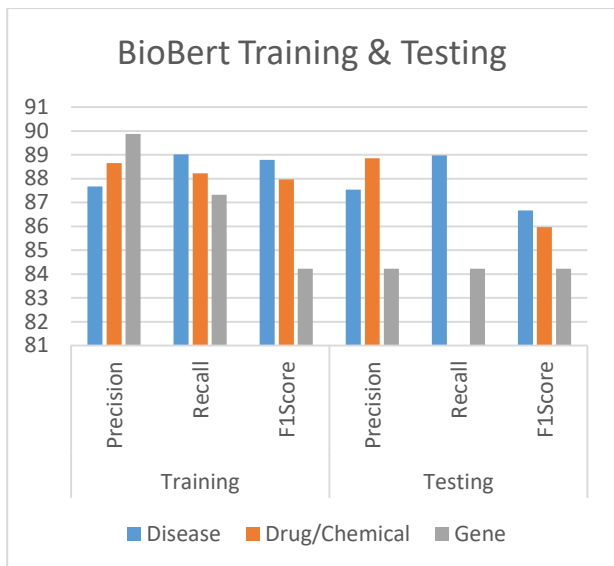| Bio Bert						
	Training			Testing		
Datasets	Precision	Recall	F1Score	Precision	Recall	F1Score
Disease	87.67	89.02	88.79	87.54	88.98	86.67
Drug/ Chemical	88.65	88.23	87.97	88.86	89.66	85.97
Gene	89.87	87.32	88.68	83.45	84.98	84.23

Figure 9. BioBERT training and testing

$$F1\ Score\ =\ \frac{2*precision*recall}{Precision+Recall} \tag{2}$$

$$Precision\ =\ \frac{TP}{TP+FP} \tag{3}$$

$$Recall\ =\ \frac{TP}{TP+FN} \tag{4}$$

Table 2 shows the model training & testing results of the data using BioBERT model to extract the entities that are related to the medical or clinical texts. The dataset is collected from PubMed and the PMC. Figure 9 depicts the comparison of training and testing accuracy.

Table 3 shows the comparison between the Scispacy and the BioBERT. For every dataset, the precision, recall, and F1-score of both models are displayed. The customized BioBERT model built and trained to extract the named entity of medical terms works equally good to the existing built in model Scispacy. Figure 10 depicts the comparison of precision, recall and F1Score results of the same.

Table 3. Comparison table

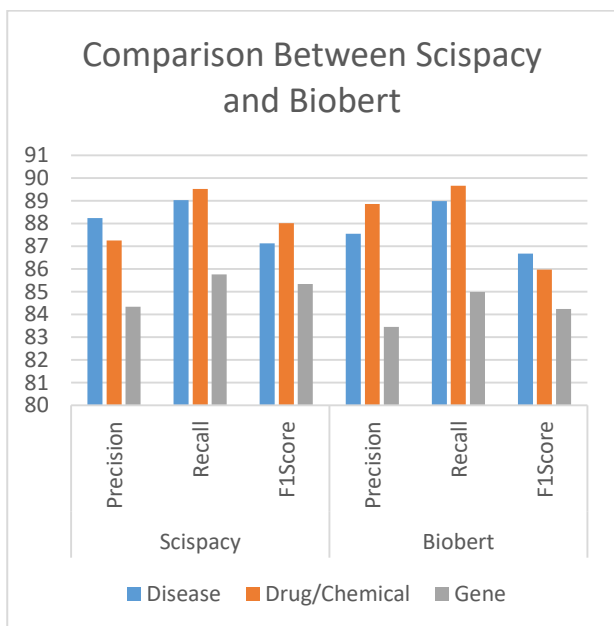| Comparison Between Scispacy and BioBERT						
	Scispacy			BioBERT		
Datasets	Precision	Recall	F1Score	Precision	Recall	F1Score
Disease	88.23	89.02	87.12	87.54	88.98	86.67
Drug/Chemical	87.25	89.52	88.01	88.86	89.66	85.97
Gene	84.34	85.76	85.34	83.45	84.98	84.23

Figure 10. Comparison between Scispacy and BioBERT

GUI:

In GUI the TensorFlow, PyTorch, and carafe, model encourage a site page where clients can get the nuances of the clinical words or components. Exactly when the client types or paste the substance in the text area and clicks analyse, the model identifies the trained named entity biomedical terms – Disease, Chemical, drug, and gene as shown in Figure.11.
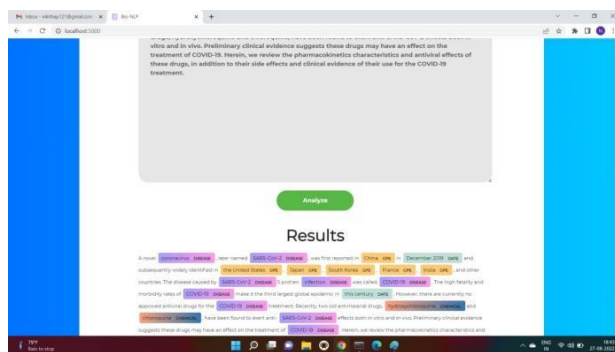

Figure 11. Screenshot of GUI Application

6. CONCLUSION

In NER, it accomplice the named components to their legitimate sorts and the entitles are eliminated for the data. For a particular space model for clinical data, it will eliminate clinical terms like disease, drug, treatment, and so on. In this work the entitles are isolated from the sentence for clinical data. Here we proposed a comprehensive plan for the BioNER. suggested with a novel portrayal layer, which is processed by a BiLSTM-CRF layer, based on cognitive processes. Using a combination of BioBERT and substantial level word description, the proposed system created a layer that may be employed in a variety of contexts. This representation layer is trained on the biomedical corpora and is unique in its ability to deal with phonetic issues such as polysemy, meaning, accentuation, OOV, and disruption within the

biomedical composition. The proposed framework can outperform many state-of-the-art baselines and provide state-of-the-art results for BioNER. In the future, we want to extend our model to other geographies and investigate new methods for bolstering more data qualities, languages, and complexity.

References:

Alam, T., & Schmeier, S. (2021). Deep Learning in Biomedical Text Mining: Contributions and Challenges. In *Lecture Notes in Bioengineering* (pp. 169-184). (Lecture Notes in Bioengineering). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-67303-1_14

Ando, R. K. (2007). Biocreative II gene mention tagging system at IBM Watson. Retrieved from https://riejohnson.com/rie/Ando_biocreative_cr.pdf

Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A pretrained language model for scientific text. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. https://doi.org/10.48550/arXiv.1903.10676

Chen, Q., Peng, Y., & Lu, Z. (2019). BioSentVec: Creating sentence embeddings for biomedical texts. In *2019 IEEE International Conference on Healthcare Informatics (ICHI)* (pp. 1–5). https://doi.org/10.1109/ICHI.2019.8904728

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). https://doi.org/10.18653/v1/N19-1423

Devopedia. (2020). Named entity recognition. Version 5, February 4. Retrieved October 9, 2022, from https://devopedia.org/named-entity-recognition

Howard, J., & Ruder, S. (2018). Universal language model finetuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 328–339). Association for Computational Linguistics. https://doi.org/10.18653/v1/P18-1031

Islamaj, R. D., Leaman, R., & Lu, Z. (2014). NCBI disease corpus: A resource for disease name recognition and concept normalization. *Journal of Biomedical Informatics*, 47, 19–24. https://doi.org/10.1016/j.jbi.2013.12.006

Kim, J.-D., Ohta, T., Tsuruoka, Y., Tateisi, Y., & Collie, N. (2004). Introduction to the bio-entity recognition task at JNLPBA. In *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications, JNLPBA 04* (pp. 70–75). Association for Computational Linguistics. https://doi.org/10.3115/1567594.1567610

Krallinger, M., Rabal, O., Leitner, F., Vazquez, M., Salgado, D., Lu, Z., Leaman, R., Lu, Y., Ji, D., Lowe, D., Sayle, R., Batista-Navarro, R., Rak, R., Huber, T., Rocktschel, T., Matos, S., Campos, D., Tang, B., Xu, H., & Valencia, A. (2015). The ChemDNER corpus of chemicals and drugs and its annotation principles. *Journal of Cheminformatics*, 7(S1), S2. https://doi.org/10.1186/1758-2946-7-S1-S2

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2019). ALBERT: A lite BERT for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942*. https://doi.org/10.48550/arXiv.1909.11942

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2019). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *arXiv preprint arXiv:1901.08746*. https://doi.org/10.1093/bioinformatics/btz682

Li, J., Sun, Y., Johnson, R. J., Sciaky, D., Wei, C. H., Leaman, R., Davis, A. P., Mattingly, C. J., Wiegers, T. C., & Lu, Z. (2016). Biocreative v CDR task corpus: A resource for chemical disease relation extraction. *Database: The Journal of Biological Databases and Curation*. https://doi.org/10.1093/database/baw068

Ling, Y. (2023). Bio+Clinical BERT, BERT base, and CNN performance comparison for predicting drug review satisfaction. *arXiv preprint arXiv:2308.03782*. https://doi.org/10.48550/arXiv.2308.03782

Liu, P., Qiu, X., & Huang, X. (2016). Recurrent neural network for text classification with multi-task learning. *arXiv preprint arXiv:1605.05101*. https://doi.org/10.48550/arXiv.1605.05101

Liu, X., He, P., Chen, W., & Gao, J. (2019). Multitask deep neural networks for natural language understanding. *arXiv preprint arXiv:1901.11504*. https://doi.org/10.48550/arXiv.1901.11504

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems* (pp. 3111–3119). https://doi.org/10.48550/arXiv.1310.4546

Peng, Y., Yan, S., & Lu, Z. (2019). Transfer learning in biomedical natural language processing: An evaluation of BERT and ELMo on ten benchmarking datasets. *arXiv preprint arXiv:1906.05474*. https://doi.org/10.48550/arXiv.1906.05474

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. *CoRR*, abs/1802.05365. https://doi.org/10.48550/arXiv.1802.05365

Pyysalo, S., Ginter, F., Moen, H., Salakoski, T., & Ananiadou, S. (2013). Distributional semantics resources for biomedical text processing. In *Proceedings of the 2013 Workshop on Biomedical Text Mining and its Applications*. Retrieved from http://lbm2013.biopathway.org/lbm2013proceedings.pdf

Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. *arXiv preprint arXiv:2003.08271*. https://doi.org/10.48550/arXiv.2003.08271

Ramachandran, R., & Arutchelvan, K. (2021). Named entity recognition on biomedical literature documents using a hybrid-based approach. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-021-03078-z

Si, Y., Wang, J., Xu, H., & Roberts, K. (2019). Enhancing clinical concept extraction with contextual embeddings. *Journal of the American Medical Informatics Association*, 26(11), 1297–1304. https://doi.org/10.1093/jamia/ocz096

Tarcar, A. K., Tiwari, A., Dhaimodker, V. N., Rebelo, P., Desai, R., & Rao, D. (2019). Healthcare NER models using language model pretraining. *arXiv preprint arXiv:1910.11241*. https://doi.org/10.48550/arXiv.1910.11241

Tawfik, N. S., & Spruit, M. R. (2020). Evaluating sentence representations for biomedical text: Methods and experimental results. *Journal of Biomedical Informatics*, 103, 103396. https://doi.org/10.1016/j.jbi.2020.103396

Yadav, V., & Bethard, S. (2019). A survey on recent advances in named entity recognition from deep learning models. *arXiv preprint arXiv:1910.11470*. https://doi.org/10.48550/arXiv.1910.11470

Yogatama, D., de Masson d'Autume, C., Connor, J., Kocisky, T., Chrzanowski, M., Kong, L., Lazaridou, A., Ling, W., Yu, L., Dyer, C., et al. (2019). Learning and evaluating general linguistic intelligence. *arXiv preprint arXiv:1901.11373*. https://doi.org/10.48550/arXiv.1901.11373

Zhu, H., Paschalidis, I. C., & Tahmasebi, A. M. (2018). Clinical concept extraction with contextual word embedding. In *NIPS Machine Learning for Health Workshop*. https://doi.org/10.48550/arXiv.1810.10566

Kayal Padmanandam
BVRIT HYDERABAD College of
Engineering for Women,
Hyderabad,
India
kayalpaddu@gmail.com
ORCID 0000-0002-3872-8422

Nikitha Pitla
BVRIT HYDERABAD College of
Engineering for Women,
Hyderabad,
India
nikithap121@gmail.com
ORCID 0009-0000-7997-2547

Yeshasvi Mogula
BVRIT HYDERABAD College of
Engineering for Women,
Hyderabad,
India
yeshasvireddy26@gmail.com
ORCID 0009-0007-2233-2320