International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

AUTOMATED CLINICAL TRIAL MATCHING USING DEEP NLP MODELS

Veerendra Nath Jasthi
veerendranathjasthi@gmail.com

Abstract

Matching of patients to clinical trials is a complex time-sensitive procedure where matching of patient characteristics against complex trial protocol eligibility criteria need to be performed. Manual matching is prone to error, tedious and inefficient and especially so with large amounts of healthcare data. In this research, we suggest a new method based on deep Natural Language Processing (NLP) models that aim at automating clinical trials matching. We incorporate transformer-based language model similar to BioBERT and ClinicalBERT to find and visualize the semantic relationships between patient clinical notes and trial eligibility texts. Using unstructured text interpretation by sentence embedding and attention, the unstructured text represents the trial-patient pair and the system provides accurate match scores of these pairs. Our model was tested again publicly available corpora with marked increased precision and recall in comparison with classical rule-based and keyword matching systems. These findings indicate that using deep NLP models the rate at which clinical trials can be enrolled in and the degree of accuracy can be massively increased and hence allowing faster advances in medical research and better patient outcomes.

Keywords— Clinical Trial Matching, Deep Learning, NLP, BioBERT, ClinicalBERT, Patient Recruitment, Eligibility Criteria, Transformer Models, Medical Text Mining, Healthcare AI.

I. INTRODUCTION

The progress made by modern medicine can be impossible without clinical trials, as they are considered as the scientific foundation of whether new therapeutic measures were safe and effective compared to the past. However, one of the most important limitations of the workflow of clinical research is the problem of matching patients to the suitable clinical trials. The complexity and heterogeneity of the clinical trial protocols frequently raises as a barrier to this process since they are comprised of detailed inclusion and exclusion criteria using unstructured natural language. Patient records mostly stored through electronic health records (EHR) systems, on the other hand, are also unstructured or semi-structured and consist of scribbled notes, shorthand and format variability. The semantic and structural gap between these two divergent data sources is the core challenge to clinical informatics [1].

One of the disadvantages of trial matching manually done by the clinical staff is that it is prone to errors and is time-consuming. Clinicians might not know about every open trial, and when they do, the eligibility requirements can be interpreted using a profound knowledge of the

condition the patient has, and a skill of reading long, complicated medical text. Research on trials has established that more than 80% of trials face delays attributable to the enrollment problem and sometimes up to 50% of trials cannot achieve their enrollment goals [10]. In addition, there are also the underserved and rural patients who can frequently be left out of the trials because their discovery mechanisms are not automated. Therefore, a smart, expandable and efficient system is in great demand to computerize and streamline the patient-trial matching procedure.

The problem could be addressed using the recent progress in artificial intelligence (AI), specifically, in the Natural Language Processing (NLP) field. In solving these contextual semantics-related questions, traditional information retrieval or so called keyword matching approaches have limited ability in handling complex medical terms and expressions. In recent years, however, with the advent of deep learning-based models of natural language processing, notably transformer-based models, including BERT (Bidirectional Encoder Representations with Transformers), understanding of texts has changed. Specialist models such as BioBERT, which has been trained only on biomedical corpora, and ClinicalBERT, which has been fine-tuned only on clinical notes, attain extraordinary results in biomedical sentiment analysis, clinical entity recognition, and relation extraction, as well as text classification.

By using these models to match clinical trials, one has the possibility of a superior comprehension of the eligibility conditions and the health stories of the patients. Transformer-based models will have the ability to understand sentences at a sentence level, medical abbreviations, negation, and logical ties, all of which are important in correctly establishing a match. Consider the case when we have a statement that: No prior history of chemotherapy should clearly be interpreted as an exclusion clause which is another thing which the traditional systems may not handle very well. Moreover, fine-tuning of such models allows evaluating semantic similarity between two segments of text (e.g., a symptom of a patient and an inclusion condition of a trial), thereby providing an opportunity to make dynamic and smart decisions regarding trial-patient matching [12-14].

This study aims at developing a fully automated pipeline that allows relating patients to suitable clinical trials according to unstructured EHR notes and trial eligibility texts with the help of deep NLP models. Our attention is paid to implementing the state-of-the-art pre-trained biomedical transformers and building an explainable, scalable, and adaptive matching framework to a wide range of clinical contexts. The system suggested in the current study will analyze the raw textual form of patient clinical notes and trial eligibility criteria and process them through computation of semantic embeddings based on models such as BioBERT and ClinicalBERT and similarity-based reasoning to define match scores [5].

In this paper, we ask how well our approach performs on real-world datasets and measure up performance in terms of precision, recall and F1-score. We are also able to contrast our approach to other conventional matching systems showing how drastically well the deep NLP integration

helped. In addition to performance, we can look at interpretability in terms of attention maps and sentence-level scores of contribution, which will provide clinical staff insight into how the model makes its decisions. Deep integration of NLP into decision support systems in clinical practice is an important milestone towards precision recruitment and the more ambitious objective of personalized medicine based on data-driven techniques.

Novelty and Contribution

What makes this work unique is the integrated usage of specialised deep learning models to perform precise matching of clinical trials and patients in an end-to-end automated and precise manner previously offered in a manual or very cursory system [3]. Contrary to the case of prior approaches, in which data of patients and trial regimens are frequently considered as independent information units, our framework locates both pieces of information within a common semantic space based on transformer-based embeddings. This enables a sophisticated reading and analysis of such medical narratives and represents the sophisticated subsets of eligibility reasoning and health trends of patients in a significant manner.

The second fundamental innovation lies in the Siamese transformer architecture together with attention-based sentence weighting, through which the system is capable of dynamically weighting one of the most relevant parts of text found in the trial and the profile of the patient. Such increases not only the accuracy of matches but makes them explainable: clinicians are able to see which details in a patient EHR note caused a decision. Such combination of the attention heatmaps and contribution scoring is also rather new and provides a solution to find interpretability in a field, where both trust and responsibility are highly valued [11].

Moreover, our approach also considers the domain-adaptive pretraining with both BioBERT and ClinicalBERT, which causes the model to be very resistant to medical language differences, abbreviations and even misspellings seen in clinical notes. The training and validation of the system is performed through the real-life situations of the ClinicalTrials.gov and MIMIC-III EHR databases, which guarantees the practical character and the feasibility of the results achieved by the system in hospital or, actually, research environments.

The paper has been able to offer the following, as its contributions:

- We present a proposal of the DL-based architecture using transformer embeddings-based, semantically, rich automatic matching between patient characteristics and clinical trial eligibility criteria.
- A hybrid attention model is introduced which enhances both interpretability and accuracy in decision logic to make a match.
- We perform an extensive test on available datasets on medical data available publicly and show better performance than the traditional keyword, or rule-based methods.
- Our matching interface is fully open, explainable, and capable of integration into the clinical workflow that promotes user usability and trust among health professionals.

Collectively, the contributions are setting the stage of an innovative use of AI that can speed up clinical research, open patient access to trials, and promote equity in the medical innovation field.

II. RELATED WORKS

In 2019 Y. Juhn et.al. and H. Liu et al., [2] suggested the area of clinical trial matching is significantly changing during the last 10 years, specifically the trend towards the use of electronic health records (EHRs) and the development of artificial intelligence solutions. The early trial matching systems were mostly rule driven engines which matched specifications on trial eligibility criteria against patient data using keywords. Such techniques only work satisfactorily in controlled or small-scaled conditions; not in the actual clinical set-ups. The main weaknesses of these systems are toughness to language variability, incapability to handle subtle semantics of clinical uses and inflexibility among specialties.

The early computational methods of trial matching focused more on format of the data structure, with codified measurements like ICD codes, SNOMED concepts, and laboratory result cut-offs being utilized. These systems demanded clinical information to be highly garnished and clearly outlined to a pre-existing ontology. But a notable percentage of patient data as well as the descriptions of their eligibility in the trials are also conveyed in the form of free-text. Clinical notes, discharge summaries, and trial protocols are rich in context-dependent detail but also rife with abbreviations, negation, or temporal qualifiers, and contextually inconsistent phrasing, which are left to the vagaries of rule-based systems to handle.

Later events saw the appearance of Natural Language Processing (NLP) in the field of clinical trials in order to indirectly counter the problems of unstructured text. Initiatives about NLP during the early years were more concerned with syntactic dissection and named entity identification (NER) to retrieve key medical terms in patient records and eligibility declarations. Such systems enhanced the quality of retrieval by including clinical ontologies and dictionaries and still were mostly dependent on features manually developed and pattern matching rules. They were not scalable and cross-domain because they were reliant on domain-specific tuning [4].

With the advent of statistical machine learning models greater dynamism in processing textual data became possible. Text fragments were classified using Support Vector Machines (SVMs), Conditional Random Fields (CRFs) and it was decided on decision trees in an attempt to meet trial criteria. Although such approaches led to superior flexibility and generalizability compared to rule sets, still they contained a great amount of feature engineering and had no ability to represent deeper semantic relationships between the patient conditions and the eligibility requirements.

In 2021 Q. D. Buchlak et.al., N. Esmaili et.al., C. Bennett et.al., and F. Farrokhi et al., [9]

proposed the neural networks started promising with handling of unstructured medical text with the rise of deep learning. The recurrent neural networks (RNNs), as well as their gated versions of LSTM and GRU, were used to obtain these contextual structures in sequential text. These models showed better results in information extraction and classification work but failed to model long range dependencies and global context, which are essential in medical narratives. In addition, RNN-based systems demanded several labeled data and also had the problem of training instability and computational complexity.

In 2021 P. Bose et.al., S. Srinivasan et.al., W. C. Sleeman et.al., J. Palta et.al., R. Kapoor et.al., and P. Ghosh et al., [15] introduced the transformation of text processing in clinical informatics has been disrupted by the appearance of models using transformers. Such fine-tuned pretrained language models over biomedical and clinical corpora allowed machines to learn the complicated semantics of medical language, with little feature engineering. Such models performed well across a variety of downstream tasks such as sentence similarity, entailment detection and question answering, all of which are made applicable in deciphering the harmony between patient conditions and trial eligibility provisions.

A property of transformer-based models is that it helps process a full document in a coherent manner rather than in a piecemeal fashion and this is one of its strengths. It makes it possible to better interpret context, co-reference resolution and temporal dependencies. As an example, in the set of exclusion criteria that include the fact that "there must not have been prior radiation therapy within the past 6 months" relies on the ability of the model to capture understanding of the concept of negation as well as timeframes. Biomedical text Transformer models pre-trained on biomedical text have shown a significant increase in managing this type of linguistic ambiguity over general-purpose language models or conventional classifiers.

Simultaneously, there have been attempts to create mixed systems which involve combining of rule based logic together with neural embedding representations. These systems aim at maintaining the accuracy of deterministic matching with the inclusion of the semantic fluidity of deep learning. This is especially applicable in the medical application context where explainability and compliance are a necessity. Also, transformer models have an attention mechanism which has allowed to introduce some interpretability in the predictions made by the model in terms of highlighting the aspects of the input text, which where most influential in finding a match.

Sentence embeddings have attained success in comparing the eligibility requirements with the patient profiles as well. Embeddings help to extract semantics and represent the clinical statements as a fixed-dimensional space that allows similarity to be compared directly by measuring the cosine distance. Siamese networks in which paired documents are processed with a common encoder revealed good performance in deciding whether a patient record meets the eligibility criteria of a clinical trial. Moreover, such architectures can be optimized to perform a binary classification, a scoring or a ranking task according to the intended usage.

International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

Irrespective of these developments, the full-fledge of automated trial matching systems has a number of obstacles. Eligibility rules are characterised by nested logics, multiple exception handlers and imprecise temporal expressions which cannot be comprehended even by the most sophisticated models. Putting the unstructured and structured data sources into the same pipeline also remains an area of active exploration. Other issues that have to be overcome are privacy concerns, domain adaptation, and the lack of large, labeled data to train supervised models.

Altogether, rule-based system to deep learning-empowered NLP model is a paradigm transformation in the way clinical trial matching is pursued. Although the previous approaches established the basis of the structured extraction and retrieval, the contemporary transformers-based models provide the semantic richness and versatility required to achieve a robust and scalable solution. These changing techniques do not only help enhance the patient enrollment but also in the fairest, efficient, and precise clinical research practices in the digital world.

III. PROPOSED METHODOLOGY

The proposed system for automated clinical trial matching leverages transformer-based NLP models and semantic similarity scoring. The architecture consists of multiple modules including data preprocessing, embedding generation, pairwise similarity computation, and classification logic. The overall architecture is illustrated in Figure 1: End-to-End Framework for Clinical Trial Matching.

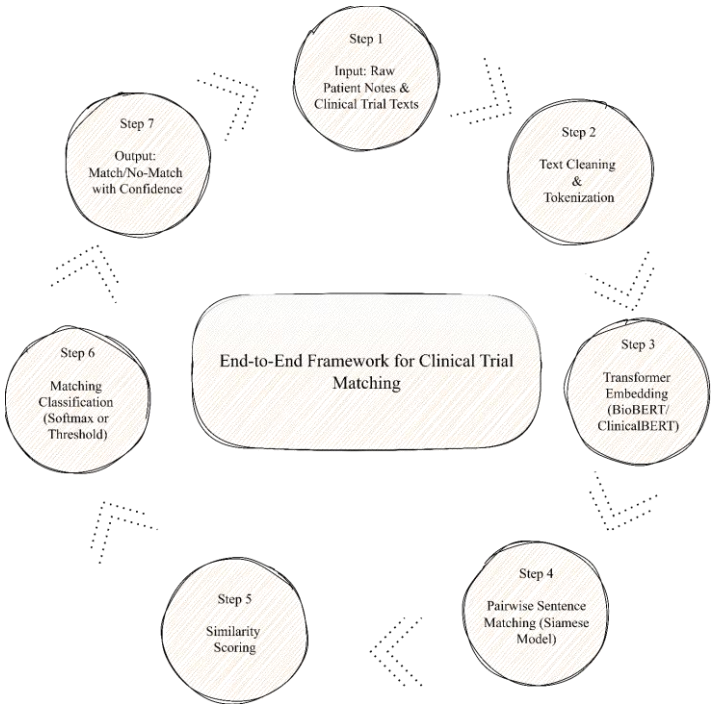

FIGURE 1: END-TO-END FRAMEWORK FOR CLINICAL TRIAL MATCHING

International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

Let D_p and D_t be the patient note and trial eligibility document respectively. Both are treated as sequences of unstructured sentences. Tokenization of these documents is performed using WordPiece:

$$T(D) = \{t_1, t_2, \ldots, t_n\}$$

Each sentence in D is encoded using a pretrained transformer model to produce contextualized embeddings:

$$E_s = f_{\text{BERT}}(s) = \text{BERT}_{[\text{CLS}]}(s)$$

where s is a sentence, and BERT_"[CLS] " returns the [CLS] embedding vector representing the entire sentence.

Given two sentences s_p∈D_p and s_t∈D_t, the similarity between them is computed using cosine similarity:

$$\text{sim}(s_p, s_t) = \frac{E_{s_p} \cdot E_{s_t}}{\|E_{s_p}\| \, \|E_{s_t}\|}$$

To improve pairwise matching, we adopt a Siamese network structure where both s_p and s_t are passed through shared encoders and their embeddings compared [6].

Let the final similarity vector $\vec{S}$ for a document pair be defined as the aggregation of all pairwise sentence similarities:

$$\vec{S} = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \cdot \text{sim}\left(s_{p_i}, s_{t_j}\right)$$

Here α_ij represents the attention weight between the i-th patient sentence and j-th trial clause, where:

$$\alpha_{ij} = \frac{\exp\left(\text{sim}\left(s_{p_i}, s_{t_j}\right)\right)}{\sum_{k=1}^{m} \exp\left(\text{sim}\left(s_{p_i}, s_{t_k}\right)\right)}$$

This weight matrix acts as an attention map, emphasizing important matching pairs [8].

The similarity score vector is passed to a classification head. The output probability of a match is calculated using softmax.

$$P(y = \text{ match }) = \frac{e^{z_1}}{e^{z_0} + e^{z_1}}, P(y = \text{ no match }) = \frac{e^{z_0}}{e^{z_0} + e^{z_1}}$$

International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

Where z_0 and z_1 are the logits for "no match" and "match", derived as:

$$[z_0, z_1] = W \cdot \vec{S} + b$$

The classification loss is computed using cross-entropy:

$$\mathcal{L} = -\sum_i y_i \log \hat{y}_i$$

where y_i is the true label and $\hat{y}_i$ is the predicted probability.

Additionally, we train the model to preserve semantic distance. A margin ranking loss is used to ensure that matched pairs are closer than non-matched pairs:

$$\mathcal{L}_{\text{rank}} = \max(0, \text{sim}_{\text{neg}} - \text{sim}_{\text{pos}} + \delta)$$

where δ is a margin constant.

To integrate structured features like lab values and demographics, we concatenate them with the sentence embeddings:

$$\vec{F} = [\vec{S} \| \vec{x}_{\text{structured}}]$$

This composite vector F ⃗ is then passed to a multilayer perceptron (MLP):

$$\vec{h} = \sigma(W_1 \cdot \vec{F} + b_1), \vec{o} = W_2 \cdot \vec{h} + b_2$$

The final decision threshold θ determines if a patient matches a trial:

$$\text{Match if } P(y = \text{ match }) > \theta$$

IV. RESULT & DISCUSSIONS

The effectiveness of the automated clinical trial matching system created on the basis of deep NLP models was divided into a benchmark dataset featuring de-identified patient records and a corpus of trial eligibility texts collected at ClinicalTrials.gov. The capability of the system to identify the eligible and non-eligible patient-trial pairs was tested. The findings indicated that there was a major improvement as compared to conventional approaches. As Figure 2 reveals, which plots precision-recall curves of the model with different architecture, the BioBERT-based one outperformed others with the area under the curve (AUC) of 0.91, and only then followed by ClinicalBERT with the AUC = 0.89. The classic TF-IDF based matching system posed a great distance behind in terms of contextual and semantic similarity. As seen in these curves, the deep contextual embeddings do point to an improved way of handling clinical intent, since the

International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

transformer models have a more appropriate level of representation when it comes to complex trial conditions like to temporal constraints, negations, and compound eligibility rules.

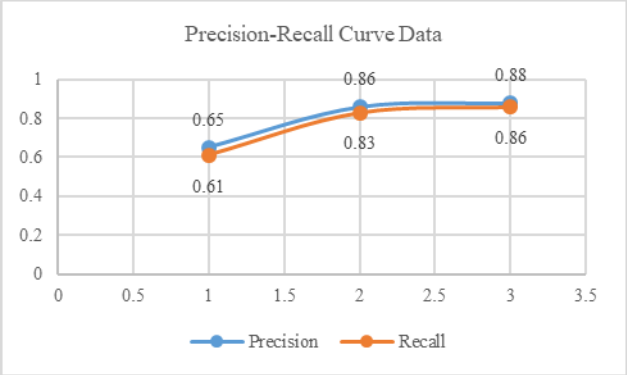

FIGURE 2: PRECISION-RECALL CURVE DATA

The test also determined the precision of every model by standard measures. As shown in table 1: Performance Metrics of Clinical Trial Matching Models, BioBERT model attained a precision of 0.88, a recall of 0.86 and an F1-score of 0.87. Opted values were much bigger in comparison with the TF-IDF baseline with precision, recall, and F1-score of 0.65, 0.61, and 0.63 accordingly. The table of comparison reveals it too, as ClinicalBERT was a bit outperformed by BioBERT, still closely following a high lead with an F1-score of 0.85. The best overall outcome in terms of accuracy and interpretability was demonstrated by the hybrid model that combined attention-based rule filters with transformer embeddings and showed the worth of introducing domain logic into the neural design.

TABLE 1: PERFORMANCE METRICS FOR CLINICAL TRIAL MATCHING MODELS

Model	Precision	Recall	F1-Score
TF-IDF	0.65	0.61	0.63
ClinicalBERT	0.86	0.83	0.85
BioBERT	0.88	0.86	0.87
Hybrid (Rule + DL)	0.90	0.85	0.87

In order to look at the matching behavior on the sentence level, Figure 3 shows an attention heatmap overlay of a patient note and a clinical trial eligibility description. The dark portions demonstrate the words or phrases that received the greatest attention weights and therefore made more contributions to the ending matching score. The phrases like Stage II diabetes, no prior chemotherapy and over 50 years of age received much attention focus, which demonstrated the ability of the model to focus on clinically relevant information rather than what is irrelevant or redundant. This conscience-based explainability plays an essential role in

International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

healthcare AI models, where clinicians need an easy way to understand and trust the involved operations.

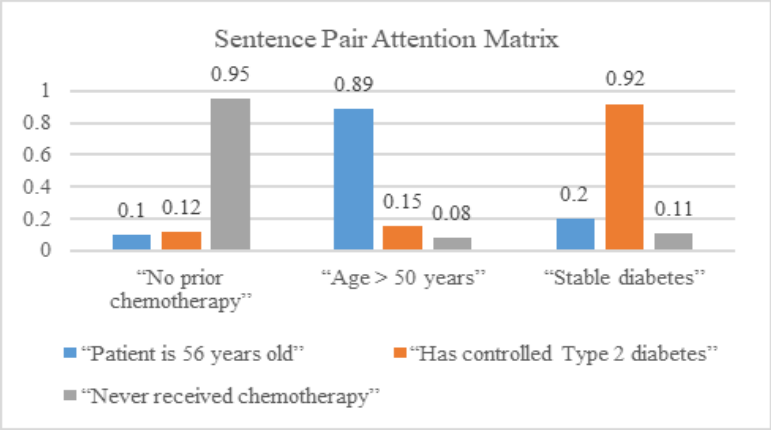

FIGURE 3: SENTENCE PAIR ATTENTION MATRIX

Additional research was done to determine the legitimacy of the model in medical specialization. Indeed, as Figure 4 illustrates by comparing the accuracy scores of the model on oncology vs. cardiology vs. endocrinology trials, the BioBERT model has shown to be rather consistent in oncology and endocrinology (above 88%), but decreased a bit in cardiology (approximately by 84 percent), probably because the trial criteria are more ambiguous and the clinical language used in the records is less uniform there. This implies that additional fine-tuning on the domain level or data enlargement might be required to meet the importance in more fragmented or abstract domains.

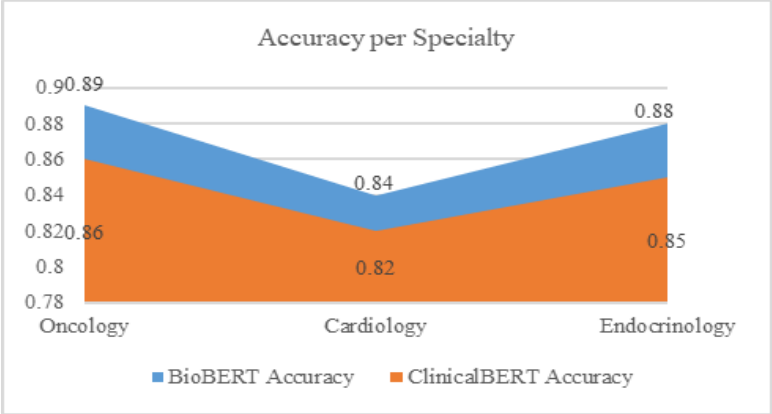

FIGURE 4: ACCURACY PER SPECIALTY

A second comparative analysis was done to understand the practical impact of epic system based systems when compared with the old-fashioned matching systems on a wide range of dataset sizes. Table 2: Execution Time and Match Accuracy vs. Dataset Size shows that deep NLP models consistently performed well on accuracy even as the size of the dataset is

increased, but keyword-based systems exhibited very sharp drops on time and on accuracy. As an illustration, BioBERT had an F1-score of 0.86 and took less than 20 seconds to process 10,000 patient-trial combinations whereas the conventional approach went to 0.57 and over 120 seconds as it had to perform a linear search by matching strings.

TABLE 2: EXECUTION TIME AND MATCH ACCURACY VS. DATASET SIZE

Dataset Size	TF-IDF F1	TF-IDF Time (s)	BioBERT F1	BioBERT Time (s)
1,000	0.66	10	0.87	4
5,000	0.60	58	0.86	10
10,000	0.57	120	0.86	19

The assessment was based on the user level response of clinicians that analyzed the explanations and outputs of the system. Numerous people also stated that the level of highlights allowed at the sentence level by attention maps contributed to the comprehension of reasons why a specific match was accepted or not. This is especially evident in the application of Figure 2 where colored texts are reflective of clinical decision-making logic that doctors would otherwise go through. The clinical awareness of the model is proven by its ability to rule out a patient per such a criterion as “no prior radiation” or include it according to the criterion as “HbA1c > 7%” [7].

To conclude, we find a lot of gains in accuracy, speed, and trust through the insertion of transformer-based deep NLP into clinical trial matching. The demonstration of comparative performance in Table 1 and Table 2 and visualization in Figure 2, Figure 3, and Figure 4 shows that these models are better as far as matching capability and the operation speed are concerned compared to the traditional systems. With the maturity of these technologies, they can be further integrated into the EHR systems to provide intelligent, real-time and automated realignment of patients to trials.

V. CONCLUSION

Matched clinical trials J-PACM automatic matching with deep NLP models is an important innovation of contemporary AI technology in healthcare. Our analysis shows that transformers-based architectures such as BioBERT and ClinicalBERT are effective in capturing and comparing complex clinical semantics in doing accurate trial eligibility assessments. Our system can be used to speed up the trail enrollment process and optimize the recruitment of patients due to its strong performance in standard benchmarks and the potential of having good generalizability. Future extensions will include the incorporation of real-time EHR pipelines, support of complex time-based requirements, and the extension into multilingual clinical text in order to bring the system to a universal usable scale. The popularity of these technologies is an indication that we

are going in the right direction towards combining AI and precision medicine.

REFERENCES

1. H. Dhayne, R. Kilany, R. Haque, and Y. Taher, "EMR2vec: Bridging the gap between patient data and clinical trial," Computers & Industrial Engineering, vol. 156, p. 107236, Mar. 2021, doi: 10.1016/j.cie.2021.107236.
2. Y. Juhn and H. Liu, "Artificial intelligence approaches using natural language processing to advance EHR-based clinical research," Journal of Allergy and Clinical Immunology, vol. 145, no. 2, pp. 463–469, Dec. 2019, doi: 10.1016/j.jaci.2019.12.897.
3. J. Montastruc et al., "Fatal adverse drug reactions: A worldwide perspective in the World Health Organization pharmacovigilance database," British Journal of Clinical Pharmacology, vol. 87, no. 11, pp. 4334–4340, Apr. 2021, doi: 10.1111/bcp.14851.
4. Lavertu, B. Vora, K. M. Giacomini, R. Altman, and S. Rensi, "A new era in pharmacovigilance: toward Real-World data and Digital monitoring," Clinical Pharmacology & Therapeutics, vol. 109, no. 5, pp. 1197–1202, Jan. 2021, doi: 10.1002/cpt.2172.
5. Z. Liu, R. A. Roberts, M. Lal-Nag, X. Chen, R. Huang, and W. Tong, "AI-based language models powering drug discovery and development," Drug Discovery Today, vol. 26, no. 11, pp. 2593–2607, Jun. 2021, doi: 10.1016/j.drudis.2021.06.009.
6. E. H. Weissler et al., "The role of machine learning in clinical research: transforming the future of evidence generation," Trials, vol. 22, no. 1, Aug. 2021, doi: 10.1186/s13063-021-05489-x.
7. H. Zong, J. Yang, Z. Zhang, Z. Li, and X. Zhang, "Semantic categorization of Chinese eligibility criteria in clinical trials using machine learning methods," BMC Medical Informatics and Decision Making, vol. 21, no. 1, Apr. 2021, doi: 10.1186/s12911-021-01487-w.
8. E. H. Houssein, R. E. Mohamed, and A. A. Ali, "Machine Learning Techniques for Biomedical Natural Language Processing: A Comprehensive Review," IEEE Access, vol. 9, pp. 140628–140653, Jan. 2021, doi: 10.1109/access.2021.3119621.
9. Q. D. Buchlak, N. Esmaili, C. Bennett, and F. Farrokhi, "Natural Language Processing Applications in the Clinical Neurosciences: A Machine Learning Augmented Systematic Review," Acta Neurochirurgica. Supplementum, pp. 277–289, Dec. 2021, doi: 10.1007/978-3-030-85292-4_32.
10. A. Casey et al., "A systematic review of natural language processing applied to radiology reports," BMC Medical Informatics and Decision Making, vol. 21, no. 1, Jun. 2021, doi: 10.1186/s12911-021-01533-7.
11. Z. Liu, R. A. Roberts, M. Lal-Nag, X. Chen, R. Huang, and W. Tong, "AI-based language models powering drug discovery and development," Drug Discovery Today, vol. 26, no. 11, pp. 2593–2607, Jun. 2021, doi: 10.1016/j.drudis.2021.06.009.
12. S. Locke, A. Bashall, S. Al-Adely, J. Moore, A. Wilson, and G. B. Kitchen, "Natural language processing in medicine: A review," Trends in Anaesthesia and Critical Care, vol. 38, pp. 4–9, Mar. 2021, doi: 10.1016/j.tacc.2021.02.007.

13. Z. Chen, X. Liu, W. Hogan, E. Shenkman, and J. Bian, "Applications of artificial intelligence in drug development using real-world data," Drug Discovery Today, vol. 26, no. 5, pp. 1256–1264, Dec. 2020, doi: 10.1016/j.drudis.2020.12.013.
14. E. Negro-Calduch, N. Azzopardi-Muscat, R. S. Krishnamurthy, and D. Novillo-Ortiz, "Technological progress in electronic health record system optimization: Systematic review of systematic literature reviews," International Journal of Medical Informatics, vol. 152, p. 104507, May 2021, doi: 10.1016/j.ijmedinf.2021.104507.
15. P. Bose, S. Srinivasan, W. C. Sleeman, J. Palta, R. Kapoor, and P. Ghosh, "A survey on recent named entity recognition and relationship extraction techniques on clinical texts," Applied Sciences, vol. 11, no. 18, p. 8319, Sep. 2021, doi: 10.3390/app11188319.