International Journal of Enhanced Research in Educational Development (IJERED)
ISSN: 2320-8708, Vol. 12 Issue 4, July.-August, 2024, Impact Factor: 8.376

Advancements and Challenges in Named Entity Recognition: A Comprehensive Survey

Mrs. Aditi Amit Malkar[1], Dr. Rahul Thour[2]

[1]Ph.D. (CSE) Scholar, Desh Bhagat University, Punjab
[2]Department of CSE, Desh Bhagat University, Punjab

ABSTRACT

Named Entity Recognition (NER) is the process of precisely identifying and extracting named entities or rigid designations, wherein preset semantic types such as Persons, Locations, and Organizations are highly valuable for information extraction from text. Efficient named entity recognition and extraction is crucial for resolving numerous issues in popular fields of study like bioinformatics, machine learning, information retrieval, video annotation, question answering and summarization systems, and semantic web search. NER is being used to identify names from text using a wide range of techniques, including Rule-base NER, Machine Learning-base NER and Hybrid NER. Thus, NER is the current hot topic and requires a lot of attention. As a result, here reviewing and analysing 25 research articles for this survey in order to analyse the techniques. In particular, we will be reviewing the gaps found in the current methods and analysing the different approaches used for the NER in this survey. The methods for thoroughly outlining the gaps will be described using the taxonomy. To highlight the demands, the analysis will be conducted by comparing the research accomplishments and the datasets used. Lastly, a demonstration of the future course that will inspire scholars to make meaningful progress in NER will be provided.

Keywords: Deep learning, Information Retrieval, Machine learning techniques, Named Entity Recognition and Transfer learning.

INTRODUCTION

The paradigm change from small data to large data has been observed by the natural language processing community through initiatives like Transformer [1-2] and its offspring. It is hardly unexpected that, given enough data, machine learning techniques can readily outperform human performance [3–4]. For many specialized domains, however, obtaining data might be a difficult undertaking. For instance, the privacy protection of not having enough Electronic Health Records has always been a problem for medical concept normalization, a fundamental subtask of named entity recognition (NER) in the medical field. Even though they appear insignificant to humans, small data sets with selection biases [5–6] can cause machine learning models to perform poorly when applied to inputs whose distribution differs from that of training data. Real understanding, dataset bias, and model robustness are other names that are used to discuss the same problems. Models trained on hypotheses-only (as opposed to hypotheses-premises) can produce better results in natural language inference than a majority-class baseline [7-8]. When it comes to reading comprehension, models trained on question-only or passage-only (as opposed to question-passage) still make accurate predictions [9], while models forecasted on a broken question (as opposed to the original question) still make accurate predictions [10–11].

This phenomenon's primary cause is misleading correlations of statistical learning. An example from computer vision can help clarify spurious correlations [12–13]: A classifier applied to a dataset of images showing cows and camels in their natural habitat will create erroneous connections between the image's geography (deserts, green pastures) and the output labels (cows, camels). Consequently, the classifier produces an incorrect prediction when it receives an image of cows taken on sandy beaches [16-17]. Given this context, we couldn't help but wonder if adding more human annotations to the data would be the only way to get rid of erroneous correlations. From a causal standpoint, confounding variables, as opposed to a direct or indirect causal pathway, are the source of misleading correlations. To some extent, spurious correlations can be eliminated from models by explicitly modifying the precursor variable in spurious correlations to produce counterfactual data [14–15].

The purpose of this survey is to identify named entities that are used in various methods, such as training datasets, evaluation metrics, benefits, and drawbacks. Additionally, this study looks into a number of restrictions, including misclassification issues, high computing costs, over fitting problems, and computational complexity. In this survey, about 25 research publications were examined with an emphasis on the difficulties encountered, the methods employed, the outcomes attained, and the parameters that were employed. An overview of existing methods and the benefits of deep learning, machine learning, and transfer learning methodologies were the goals of the study.

The following sections comprise this survey. The taxonomy of the NER literature review is explained in Section 2. An overview of the methods, datasets, metrics, successes, and drawbacks is given in Section 3. The research gaps are covered in Section 4, and the conclusion and next steps are covered in Section 5.

Taxonomy of Named Entity Technique

Fig. 1 shows the taxonomy diagram for the named entity technique employing a variety of techniques, including machine learning, deep learning, and transfer learning.

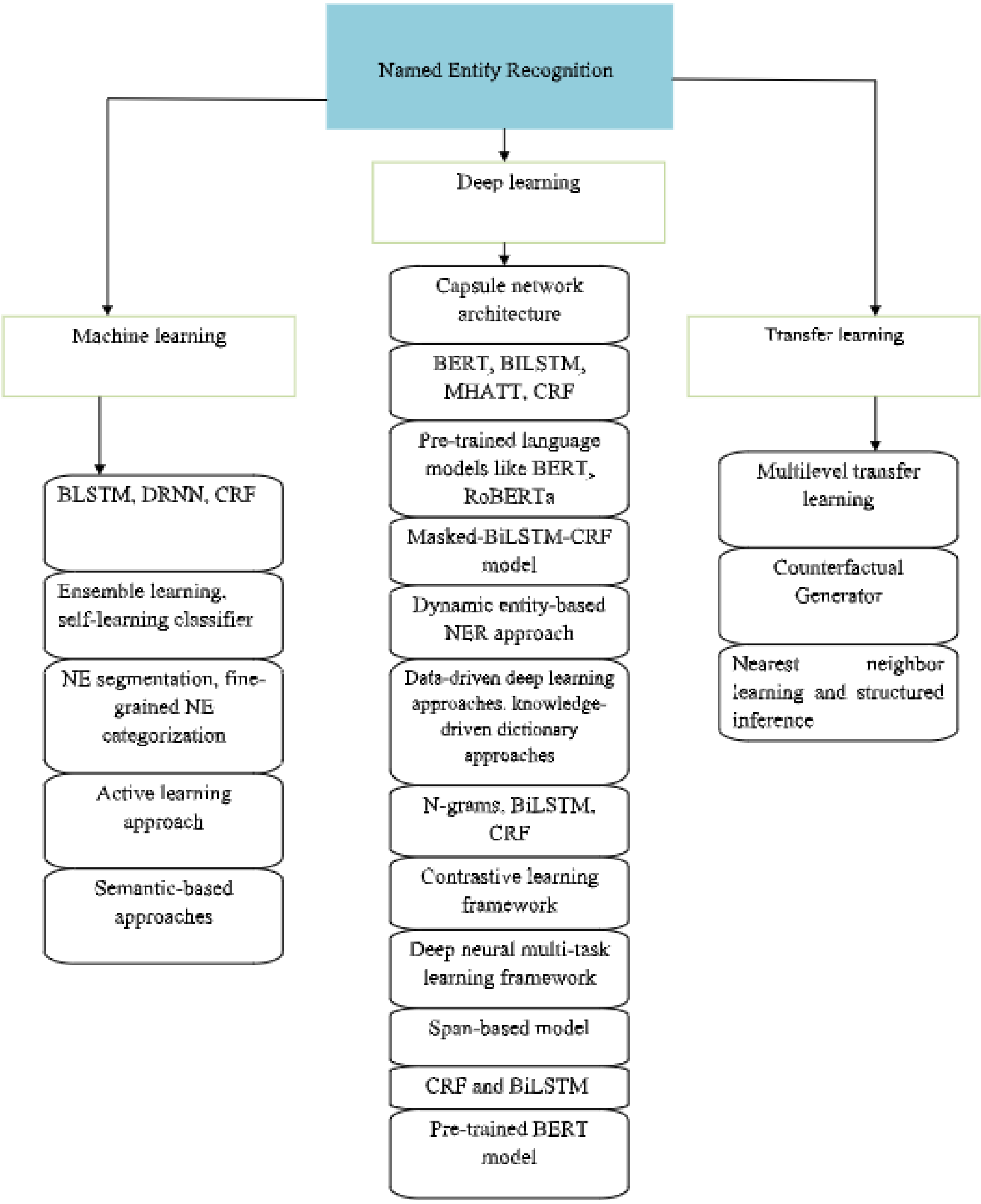

Fig. 1. Taxonomy diagram of named entity recognition techniques

Machine Learning Models:

In order to achieve biomedical named entity recognition, SudhakaranGajendran et al. [2] presented a novel neural network architecture that consists of three components: conditional random field (CRF), dynamic recurrent neural network (DRNN), and bidirectional long short-term memory (BLSTM). The only features used in this architecture are word level and character level embedding. With less intricately designed engineering elements, the suggested system performs better than the current ones. The majority of the chemical entities may be identified using the BiLSTM and the embedding approach. Nevertheless, over fitting problems persist with this model. The high level architecture of NER employing the ensemble learning method was introduced by R. Ramachandran et al. [4]. A self-learning classifier and an entity identifier based on a dictionary are features of the EL model. The suggested model gave excellent results and was quite accurate. Still, more work needs to be done to expand on the expansion of the abstracts and features. By combining the more general secondary goal of NE segmentation with the primary task of fine-grained NE categorization, Gustavo Aguilar et al. [16] introduced a novel multi-task technique. This model had very high categorization accuracy. However, extended entity recognition proved to be extremely challenging for this model. A straightforward yet powerful CNN-based network for NER, the gated relation network (GRN), was introduced by Hui Chen et al. [2]. Compared to other CNNs, GRN is better at collecting long-term context and effectively overcoming over-fitting problems. Nevertheless, the time required to train and test this model may increase. A unique Active learning (AL) approach for clinical NER was presented by Linh Le et al. [7]. It compares the train and test data distributions using low computing cost similarity metrics to determine which instances are best for training. However, this model had a very high computational cost. A technique for enhancing NER with unlabeled data and sparse labelled data was presented by Juae Kim et al. [12]. Without requiring extra manual labelling, the suggested NER systems perform better in both languages than the baseline systems. Nonetheless, the NER model's performance, which was trained using just machine-labelled data, is less impressive. A new target for Agriculture Entity Recognition (AGER) covering popular concepts in agriculture was presented by Quoc Hang et al. [11]. They also proposed a two-stage process for this task: in the first stage, build an annotated corpus of agricultural entities semi-automatically using semantic-based approaches for agricultural entity detection; in the second stage, use a machine learning approach to identify the agricultural entities from plain text and train on the annotated corpus, which have demonstrated high accuracy when evaluated in the AGER corpus. The model's computational complexity was extremely high, though.

Deep Learning Models:

Using deep learning techniques based on the capsule network architecture, Amit Kumar Jaiswal et al. [3] introduced a novel approach for corpus-based entity prediction. Different from convolutional neural networks and their "max-pooling" approach, this kind of network arranges neurons into so-called capsules to detect particular properties of an object without decreasing the original input. The problems with over fitting are successfully resolved by this model. That being said, the model's performance remained poor. A novel hybrid deep-learning model for NER in biochemistry was introduced by Jian Liu et al. [10]. The novel organic integration of BERT, BILSTM, multi-head attention (MHATT), and conditional random field (CRF) forms the foundation of the model. When compared to state-of-the-art techniques, the hybrid approach provides the best recognition performance; in particular, it significantly increases performance in detecting low-frequency entities, polysemes, and abbreviations. Nevertheless, this methodology continues to lead to inconsistent entity labels across the article. A new computational framework called BOND was introduced by Chen Liang et al. [24]. It uses pre-trained language models, such as BERT and RoBERTa, to enhance the prediction performance of NER models.

This approach significantly increases recall by preventing the model from over fitting the incompletely annotated labels. Nevertheless, producing high-quality labels with extensive coverage of the target corpus remained extremely difficult. A technique for weakly supervised learning was introduced by Nan Gao et al. [1] to identify complex named entities (that is, entities that are difficult to define boundaries around since they are typically made up of several tiny entity sequences, or CNEs) in the corpus. Masked-BiLSTM-CRF is a method that is proposed to improve the recognition accuracy by separating the entity boundary confirmation and context semantic relationship determination. The problems with over fitting are successfully resolved by this model. Nevertheless, this results in the addition of some interference features that lower the model's capacity for generalisation and impair its accuracy when extracting text sequence features.

The dynamic entity-based NER technique was described by Feng Zhao et al. [6] with unconstrained tagging strategies. In order to get rid of the restrictions, we restructure popular tagging systems and suggest two unique unconstrained schemes: one where words and entities are labelled uniformly as chunks and one where tags are applied to words and entities independently. The suggested entity-based models are acceptable for identifying entities with lengthy durations, according to the recall rate vs entity length data. Yet, over fitting problems continue to impact this model. A new model that integrates knowledge-driven dictionary approaches which embed dictionaries into deep neural networks and data-driven deep learning approaches was described by Qi Wang et al. [17]. When compared to the most advanced deep learning techniques, our model's performance was quite competitive. To describe the structure of clinical named entities, it was challenging to enumerate every rule, and using such a handmade technique always results in a comparatively high system engineering cost. An NER system for biomedical entities was provided by Hyejin Cho and Hyunju Lee [19] by combining n-grams with BiLSTM and CRF; this system is known as a contextual LSTM network with CRF (CLSTM). The performance on biological NER problems is greatly enhanced by this strategy. The model's calculation time was extremely long, though. The contrastive learning framework CONTAINER, introduced by Sarkar SnigdhaSarathi Das et al. [22], optimises the inter-token distribution distance and models Gaussian embedding, resulting in fewer misclassifications and improved entity detection, showing stability and improved performance. The model, however, had an extremely long computation time.

A unique deep neural multi-task learning framework with explicit feedback techniques was presented by Sendong Zhao et al. [23] to jointly model recognition and normalization, achieving great accuracy compared to the current models. Nevertheless, the model came at a relatively high computational cost. A unique span-based model, which was highly competitive for overlapped and discontinuous NER, was presented by Fei Li1 and Zhichao Lin [25]. This model can recognize both overlapping and discontinuous entities concurrently. Nevertheless, this model had an extremely high computational complexity and training time.

A hybrid model that generates outputs for both intent classification (IE) and NER is provided by S. Rizou et al. [8]. It incorporates layers of conditional random fields (CRF) and BiLSTM technique. Regardless of the language and language model used, the suggested model exhibits excellent performance stability. But the model's performance was really poor. Two-stage entity identification was introduced by Yongliang Shen et al. [14]. To locate the entities, first generate span proposals by filtering and boundary-regression on the seed spans. Next, label the boundary-adjusted span proposals with the corresponding categories.

Despite this, the model was still plagued by high computational costs and a lack of boundary information. For biomedical named entity recognition, Wonjin Yoon et al. [15] presented a Collabo Net that combines many BiLSTM-Conditional Random Field (CRF) models. This should lead to even better biomedical named entity identification performance. Nevertheless, this model's precision tends to suffer from its inability to distinguish between different entity types. A pre-train BERT model that can take advantage of unlabeled domain-specific knowledge was presented by Xiangyang Li [18] using unlabeled Chinese healthcare records.

The model also included a novel approach for incorporating dictionary features. As a result, this strategy greatly increased the accuracy for long entities. On the other hand, the model required a very long training period. In a novel framework for Natural Language Understanding (NLU), proposed by Alberto Benayas et al. [14], the NER models' parameters and those of the intent classification (IC) are concurrently trained in a consolidated parameter space. This approach is straightforward yet highly effective. In terms of memory and processing costs, this approach was less expensive and could do several NLU tasks at once. Even yet, there are still a number of well-documented difficulties, including coping with novel or evolving domains, short phrases, long-range dependencies, out-of-vocabulary words, and unclear meanings.

Transfer learning:

Multilevel transfer learning (mLTL) is a unique framework that was presented by Jason C. Hun and Jia-Wei Chang [9] and is based on the fine-tuning approach. By completing tasks for named entity recognition and facial emotion recognition, we were able to demonstrate the usefulness of this framework and draw important conclusions about the training sequence of related domain datasets. Based on the experimental findings, the deep neural network models with mLTL performed significantly better on the target tasks than the original models. This approach is effective in the majority of situations, but it has drawbacks, including over fitting.

The Counterfactual Generator was introduced by Zengxiangji et al. [5] with the goal of improving the original dataset by creating counterfactual examples by interventions on the current observational examples. The approach enhances the capacity of models to generalise under sparse observable examples, according to experiments conducted on three different datasets. Still, gathering data remained a difficult undertaking for many specialized fields.

Yi Yang and ArzooKatiyar [21] introduced a basic few-shot NER system that utilizes nearest neighbor learning and structured inference. By using a supervised NER model that was trained on the source domain as a feature extractor, our approach outperforms other models in terms of accuracy. Slot filling and part-of-speech tagging are two issues with few-shot sequence tagging that this approach still faces.

SUMMARIZED ANALYSIS

Table 1: Summarized Analysis of Named Entity Recognition Techniques

Sr. No.	Author	Metrics	Dataset	Research gaps
1	Nan Gao et al. [1]	Precision, recall and F1-score	-	Nevertheless, this results in the addition of some interference features that lower the model's capacity for generalization and impair its accuracy when extracting text sequence features.
2	SudhakaranGajendran et al. [2]	Precision, recall and F1-score	BioCreAtIvE II GM corpus, JNLPBA corpus and NCBI corpus.	Over fitting problems persist with this model
3	Amit Kumar Jaiswal et al. [3]	Precision, recall and F1-score	CoNLL-2003 and Rotten Tomatoes Movie Review dataset:	The model's performance remained poor
4	R. Ramachandran et al. [4]	Precision, recall and F1-score	-	Still, more work needs to be done to expand on the expansion of the abstracts and features
5	Zengxiangji et al. [5]	Precision, recall and F1-score	CNER, IDiag and CLUENER	Still, gathering data remained a difficult task
6	Feng Zhao et al. [6]	Precision, recall and F1-score	English, German, Dutch and Spanish dataset	Yet, over fitting problems continue to impact this model.
7	Linh Le et al. [7]	Precision, recall and F1-score	ShARe/CLEF 2013 and i2b2/VA 2010	However, this model had a very high computational cost.
8	S. Rizou et al. [8]	Precision, recall and F1-score	UniWay dataset and xSID datasets.	The model's performance was really poor.
9	Jason C. Hun and Jia-Wei Chang [9]	Precision, recall and F1-score	CK+, FER2013, LE and GFE2019.	It has drawbacks, including over fitting
10	Jian Liu et al. [10]	Precision, recall and F1-score	MULTIPLE and IDENTIFIER dataset	Nevertheless, this methodology continues to lead to inconsistent entity labels across the article
11	Quoc hung ngo et al. [11]	Precision, recall and F1-score	AGER TAGSET and AGER DATASET	The model's computational complexity was extremely high
12	Juae Kim et al. [12]	Precision, recall and F1-score	QA, wiki, news and coNLL	Nonetheless, the NER model's performance, which was trained using just machine-labelled data, is less impressive
13	Alberto benayas et al. [13]	Accuracy and F1-Score	ATIS and SNIPS NLU dataset	There are still a number of well-documented difficulties
14	Yongliang Shen et al. [14]	Precision, recall and F1-score	ACE 2004 and ACE 2005, KBP17 and GENIA	High computational costs and a lack of boundary information.
15	Wonjin Yoon et al. [15]	Precision, recall and F1-score	BC2GM, BC4CHEMD, BC5CDR, JNLPBA, NCBI.	The model's precision tends to suffer from its inability to distinguish between different entity types
16	Gustavo Aguilar et al. [16]	Precision, recall and F1-score	WNUT-2017	However, extended entity recognition proved to be extremely challenging for this model

17	Qi Wang et al. [17]	Precision, recall and F-measure	CCKS-2017 Task 2 benchmark dataset	High system engineering cost
18	Xiangyang Li. [18]	Precision, recall and F1-score	CCKS-CNER 2017 and CCKS-CNER 2018 dataset	The model required a very long training period.
19	Hyejin Cho and Hyunju Lee [19]	Precision, recall and F-measure	Corpora, NCBI, GM and CDR	The model's calculation time was extremely long
20	Hui Chen et al. [20]	Precision, recall and F1-score	CoNLL-2003 English NER and Onto Notes 5.0	Nevertheless, the time required to train and test this model may increase
21	Yi Yang and ArzooKatiyar [21]	F1-score	Onto Notes , CoNLL'03, I2B2'14 and WNUT'17	Slot filling and part-of-speech tagging are two issues
22	Sarkar SnigdhaSarathi Das et al. [22]	F1-score	Onto Notes, CoNLL'03, WNUT '17, GUM and a new large scale Few-Shot NER dataset (Few-NERD)	The model, however, had an extremely long computation time
23	Sendong Zhao et al. [23]	F1-score	NCBI and BC5CDR	Nevertheless, the model came at a relatively high computational cost
24	Chen Liang et al. [24]	F1-score, precision and recall	Biomedical Domain NER Datasets	Nevertheless, producing high-quality labels with extensive coverage of the target corpus remained extremely difficult.
25	Fei Li1 and Zhichao Lin [25]	Precision, recall and F1-score	CLEF, GENIA and ACE05	Extremely high computational complexity and training time

RESEARCH GAPS

Machine Learning Model:

- Explore how the machine learning model's performance can be enhanced towards identifying and extracting entities that have a structural complexity like nested or overlapping entities and are generally observed in certain fields for instance biomedical or legal ones [2].
- Investigate ways to reduce effects of data scarcity on machine learning models especially when labeled data is scarce like in low resource languages or when focused in specific fields [4].
- Review the methods that can be used to enrich feature space of the algorithms for NER using a machine learning approach and incorporate domain knowledge or language characteristics to improve the already used methods for the tasks that involve multiple types of entities [6].
- Study to understand how the concepts of domain adaptation can enable better learning when the data sets of the training as well as the testing endure a holocaust of differences [6].
- Establish approaches for dealing with noisy or possibly ambiguous data, such as techniques for identifying entities when there are spelling mistakes, abbreviations or slight variations in the naming conventions regarding the mentioned entities [7].

Deep Learning Model:

- Consider ways to improve the interpretability and explain ability of deep learning models for NER, to make the mechanisms behind the model more transparent, which is necessary in case when the decision-making process is supervised by humans or regulated by legal norms [3].
- Research on ways to optimise the training and inference steps within the deep models for NER, such as methods aimed at reducing the computational and memory intensity of DL models, which can be required in resource-limited contexts during NER [3].
- Studied methods for acquiring and describing the long-range relations inside the text sequences that are critical for the identification of the entities in long documents and in cases where the entities may spread across numerous sentences or even paragraphs [10].
- Explore approaches to implement semi-supervised or unsupervised learning in deep learning models for NER, which allows the model to utilize unlabeled data or knowledge from other sources to enhance their performance, especially when there is a scarcity of annotated data [10].

- Research strategies for increasing resilience of deep learning models against adversarial perturbations or input disturbances, which is crucial for the model's usage in real-world applications where there is potential for adversarial manipulations or alterations of inputs [19].

Transfer Learning Model:

- Examine the approaches to apply unsupervised transfer learning in case NER, where the domain's models are applied in the target domain with no access to labeled dataset, thus solving the problem of restricted access to specific domains' labeled data [9].
- Explore the techniques of multi-task learning in transfer learning for NER where the models are trained to solve more than one related task at a time like POS tagging, NER, and Entity linking to improve the overall performance of the model across the tasks [9].
- It is important to find an adequate solution to the problem of domain shift and address the problem of constructing transfer learning approaches that can successfully transfer the knowledge onto other domains while reducing the negative effects of domain differences on target domains' model performance [5].
- Discover the methods of zero-shot and few-shot learning in the field of transfer learning for NER in which models are trained for entity recognition with the help of small to no labeled data by importing knowledge from the data sources with rich labeled data [21].
- It should be necessary to create dynamic transfer learning methods that adjust the transfer process according to the target domain or task characteristics allowing models to adjust to the differences in the data and task at hand [21].
- Explore privacy-preserving transfer learning approaches for NER in which knowledge is transferred from the source model while keeping the content of the target domain's data confidential; thus, this work shall address data privacy and security issues [21].
- Research for effective transfer learning with interpretable NER methods that allow users to deduce the type of knowledge being transferred and the transferred model's decisions with the aim of understanding how best adopted pre-existing models for the target tasks or domains [9].
- Investigate what has been accomplished in transfer learning for NER; what it means for models to learn how to learn in new tasks or domains by using transfer-information in the minimum amount of annotated data [9].

CONCLUSION

NER is vital in different applications in many fields, for instance, Question Answering, Information Retrieval, and Bioinformatics. As the relevance of correct and precise extraction of named entities increases in text, there is a substantial need to revisit the existing approaches and consider potential artifacts and limitations for opening new directions for research. Here in this survey, we have analysed 25 research articles, some of them discussed techniques for NER like deep learning Method, Machine Learning-based Method, and transfer learning Method. Through this literature review of these methods and their datasets, we have recognized some of the gaps for future research. The taxonomy description given in the survey helps to get a detailed notion of the techniques used in NER and the specific aspects that require enhancement. We have identified efforts and datasets employed in the prior research to stress the necessity of overcoming these gaps for the improvement of NER systems. In the future, the authors should therefore concentrate on researching and finding solutions to the already noted gaps in understanding NER methods. It is recommended that future work in this area focuses on deriving better and more reliable methods in NER because of the uncertainty that arises when addressing specific domains or tasks. If these challenges are to be met and new methods sought, the researchers themselves will be able to promote new NER development and demonstrate the subject's relevance in other fields.

REFERENCES

[1]. Gao, Nan, Zhenyang Zhu, ZhengqiuWeng, Guolang Chen, and Min Zhang. "A supervised named entity recognition method based on pattern matching and semantic verification." Journal of Internet Technology 21, no. 7 (2020): 1917-1928.

[2]. Gajendran, Sudhakaran, D. Manjula, and VijayanSugumaran. "Character level and word level embedding with bidirectional LSTM–Dynamic recurrent neural network for biomedical named entity recognition from literature." Journal of Biomedical Informatics 112 (2020): 103609.

[3]. Jaiswal, Amit Kumar, Prayag Tiwari, Sahil Garg, and M. Shamim Hossain. "Entity-aware capsule network for multi-class classification of big data: A deep learning approach." Future Generation Computer Systems 117 (2021): 1-11.

[4]. Selvan, Senthamizh. "Named Entity Recognition using Ensemble Learning." International Research Journal on Advanced Science Hub 2, no. 7 (2020): 44-48.

[5]. Zeng, Xiangji, Yunliang Li, YuchenZhai, and Yin Zhang. "Counterfactual generator: A weakly-supervised method for named entity recognition." In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 7270-7280. 2020.

[6]. Zhao, Feng, XiangyuGui, Yafan Huang, Hai Jin, and Laurence T. Yang. "Dynamic entity-based named entity

recognition under unconstrained tagging schemes." IEEE Transactions on Big Data 8, no. 4 (2020): 1059-1072.

[7]. Le, Linh, GianlucaDemartini, Guido Zuccon, Genghong Zhao, and Xia Zhang. "Active learning with feature matching for clinical named entity recognition." Natural Language Processing Journal 4 (2023): 100015.

[8]. Rizou, Sofia, AngelosTheofilatos, Antonia Paflioti, Eleni Pissari, IraklisVarlamis, George Sarigiannidis, and K. ChChatzisavvas. "Efficient intent classification and entity recognition for university administrative services employing deep learning models." Intelligent Systems with Applications 19 (2023): 200247.

[9]. Hung, Jason C., and Jia-Wei Chang. "Multi-level transfer learning for improving the performance of deep neural networks: Theory and practice from the tasks of facial emotion recognition and named entity recognition." Applied Soft Computing 109 (2021): 107491.

[10]. Liu, Jian, Lei Gao, SujieGuo, Rui Ding, Xin Huang, Long Ye, QinghuaMeng, AsefNazari, and DhananjayThiruvady. "A hybrid deep-learning approach for complex biochemical named entity recognition." Knowledge-Based Systems 221 (2021): 106958.

[11]. Ngo, Quoc Hung, TaharKechadi, and Nhien-An Le-Khac. "Domain specific entity recognition with semantic-based deep learning approach." IEEE Access 9 (2021): 152892-152902.

[12]. Kim, Juae, YoungjoongKo, and Jung Yun Seo. "Construction of machine-labeled data for improving named entity recognition by transfer learning." IEEE Access 8 (2020): 59684-59693.

[13]. Benayas, Alberto, ReyhanehHashempour, Damian Rumble, ShoaibJameel, and Renato Cordeiro De Amorim. "Unified transformer multi-task learning for intent classification with entity recognition." IEEE Access 9 (2021): 147306-147314.

[14]. Shen, Yongliang, Xinyin Ma, Zeqi Tan, Shuai Zhang, Wen Wang, and Weiming Lu. "Locate and label: A two-stage identifier for nested named entity recognition." arXiv preprint arXiv:2105.06804 (2021).

[15]. Yoon, Wonjin, Chan Ho So, Jinhyuk Lee, and Jaewoo Kang. "Collabonet: collaboration of deep neural networks for biomedical named entity recognition." BMC bioinformatics 20 (2019): 55-65.

[16]. Aguilar, Gustavo, SurajMaharjan, Adrian Pastor López-Monroy, and Thamar Solorio. "A multi-task approach for named entity recognition in social media data." arXiv preprint arXiv:1906.04135 (2019).

[17]. Wang, Qi, Yangming Zhou, Tong Ruan, DaqiGao, Yuhang Xia, and Ping He. "Incorporating dictionaries into deep neural networks for the Chinese clinical named entity recognition." Journal of biomedical informatics 92 (2019): 103133.

[18]. Li, Xiangyang, Huan Zhang, and Xiao-Hua Zhou. "Chinese clinical named entity recognition with variant neural structures based on BERT methods." Journal of biomedical informatics 107 (2020): 103422.

[19]. Cho, Hyejin, and Hyunju Lee. "Biomedical named entity recognition using deep neural networks with contextual information." BMC bioinformatics 20 (2019): 1-11.

[20]. Chen, Hui, Zijia Lin, Guiguang Ding, Jianguang Lou, Yusen Zhang, and BorjeKarlsson. "GRN: Gated relation network to enhance convolutional neural network for named entity recognition." In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, no. 01, pp. 6236-6243. 2019.

[21]. Yang, Yi, and ArzooKatiyar. "Simple and effective few-shot named entity recognition with structured nearest neighbor learning." arXiv preprint arXiv:2010.02405 (2020).

[22]. Das, Sarkar SnigdhaSarathi, ArzooKatiyar, Rebecca J. Passonneau, and Rui Zhang. "CONTaiNER: Few-shot named entity recognition via contrastive learning." arXiv preprint arXiv:2109.07589 (2021).

[23]. Zhao, Sendong, Ting Liu, Sicheng Zhao, and Fei Wang. "A neural multi-task learning framework to jointly model medical named entity recognition and normalization." In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, no. 01, pp. 817-824. 2019.

[24]. Liang, Chen, Yue Yu, Haoming Jiang, SiawpengEr, Ruijia Wang, Tuo Zhao, and Chao Zhang. "Bond: Bert-assisted open-domain named entity recognition with distant supervision." In Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, pp. 1054-1064. 2020.

[25]. Li, Fei, ZhiChao Lin, Meishan Zhang, and Donghong Ji. "A span-based model for joint overlapped and discontinuous named entity recognition." arXiv preprint arXiv:2106.14373 (2021).