ISSN No:-2456-2165

Approaches to Named Entity Recognition

Dr. A Debbarma
Department of CSE
Dasaratha Deb Memorial College
Khowai, Tripura, India

Abstract:- This study explores the evolution of Named Entity Recognition (NER) methods, encompassing rule-based, machine learning-based, and hybrid approaches. It emphasizes the significance of NER in North East India's linguistically diverse context, revealing digital disparities. Despite the region's modest size, NER initiatives have emerged for languages like Assamese, Manipuri, Mizo, and Kokborok. The journey spans rule-based systems to advanced machine learning techniques, highlighting the dynamic nature of the field. Successful hybrid approaches, combining rule-based and machine learning methods, are showcased. However, digital disparities persist, underscoring the need for continued research and technological advancements to bridge gaps in North East India's linguistic diversity.

Keywords:- *NLP, NER, Rule-based, ML, NER.*

I. INTRODUCTION

The named entity recognition empowers to identify specific predetermined elements within a text. An early instance of named entity recognition involves extracting company names from text using linguistic knowledge [1]. The exploration of named entity recognition gained traction after the Message Understanding Conference (MUC)-6 in 1995 [2]. The term "Named Entity (NE)" was first introduced during the Message Understanding Conferences (MUC-6)[3]. MUC-6 aimed to extract structured information regarding company activities and defense activities from unstructured text. The three subtasks of MUC-6 focused on extracting information for entities like ENAMEX (Entity name expression, e.g., Person, Location, Organization), NUMEX (Numerical expression, e.g., numbers, percentage), and TIMEX (Time expression, e.g., time, date, month, year).

Since the MUC-6 conference, research on named entity recognition has expanded, gaining broader acceptance for diverse named entities. Named entities now extend beyond ENAMEX, NUMEX, or TIMEX, encompassing various domains with their predefined entities. Numerous research groups have delved into the study of named entity recognition. The IREX (Information Retrieval and Extraction Exercise) project, as well as the shared tasks in CoNLL-2002 [4] and CoNLL-2003 [5], marked early instances where named entity recognition gained prominence. In India, research groups such as IJCNLP-2008 and NER - Named-Entity Recognition Indian Languages (FIRE-2013) focused on named entity recognition for Indian languages.

II. APPLICATION OF NER

Named entity recognition stands as a crucial element in various applications of natural language processing, particularly serving as a core component in question answering systems [6]. Its significance extends to multiple areas within NLP.

In information retrieval, the task involves extracting pertinent information from text, with approximately 17% of online web queries revolving around individuals' names, according to [7]. A clustering search results based on people's names and extracting biographical attributes, making named entities pivotal in relation extraction within information retrieval was proposed in [8].

Question answering systems, addressing queries related to "who," "where," "whom," etc., heavily rely on named entities, constituting around 80% of the questions in such systems [9]. Named entity recognition enhances the quality of answers provided by question answering systems [10]. In machine translation, an essential tool for language understanding and translation, the identification of named entities plays a crucial role [11]. Google Translate, a widely used machine translation platform, supports translation for 108 languages, including 12 Indian languages. Summarization of text benefits significantly from named entity recognition, as it enables the extraction of more meaningful segments [12]. This approach allows for effective sentence extraction in summarization [13]. When it comes to news classification, utilizing named entities has shown superior performance compared to systems that neglect this component [14]. Named entity recognition is also instrumental in extracting information from the biomedical domain.

III. RULE BASED APPROACH

In specialized applications, rule-based named entity recognition has been employed to extract food entities [15]. Additionally, a dietary recommendation system was developed based on a rule-based approach [16].

Rule-based systems are often regarded as the "simplest form of artificial intelligence" [17]. This approach relies on linguistic knowledge to formulate rules for named entity recognition (NER) systems. Morphological analysis of words provides insights into named entities, and rule-based NER systems adhere to specific linguistic patterns of a language.

In early entity extraction efforts, linguistic rules and heuristic features were combined to construct systems for extracting company names, surpassing human analysis [18]. Linguistic features form the foundation for establishing rule-based approaches in NER systems. These rich linguistic features are employed to design algorithms by studying linguistic patterns. Large-scale information extraction, such as the Large-scale Information Extraction (LaSIE) system developed as part of MUC-6, leverages linguistic knowledge [19]. Entity extraction benefits from graphological, syntactic, semantic, world knowledge, and discourse-level information within linguistic knowledge [20].

Rule-based approaches involve handcrafting rules based on linguistic features of a language. While much research has been conducted for English and a few well-developed languages, fewer studies focus on other languages. For instance, a handcrafted lexicon resource was utilized to extract financial information for developing a Greek named entity recognition system, but errors were noted due to spelling and preprocessing issues [21]. Gazetteers and lemmatization enhance named entity recognition performance [22]. A rule-based Arabic NER system using a dictionary of names and regular expressions from Arabic grammar was developed [23].

English has dominated NLP research due to the availability of tools, but Indian languages, traditionally considered low-resource, are now developing NLP tools, reducing the research gap. Early work on named entity recognition for a South Asian language focused on Urdu, highlighting challenges in system development [24].

Inconsistencies in data hinder traditional approaches, emphasizing the importance of domain-specific data for improved named entity recognition. Gazetteers, which group named entities, are critical in NER systems and can be sourced from resources like Wikipedia and WordNet [25][26].

IV. STATISTICAL APPROACH

Statistical and machine learning-based approaches are pivotal in named entity recognition (NER) within the realm of natural language processing. Machine learning, a branch of artificial intelligence, involves learning from training input data to construct a model based on patterns and statistical information. This model is subsequently utilized to classify and recognize test data, providing desired output data. The Maximum Entropy Named Entity (MENE) system was introduced as part of the MUC-7 task, utilizing linguistic knowledge as features [27]. It was built upon the Maximum Entropy Toolkit. Hindi language employed orthographical and gazetteer features using the maximum entropy approach, achieving an F-value of 81.52% [28]. Hidden Markov Model (HMM) was employed to address NER, introducing the Identifinder system where each word is assigned a single label [29]. The Viterbi algorithm is used to generate the NER model. HMM, based on mutual information independence assumptions, incorporates gazetteers to enhance performance [30]. Conditional Random Field (CRF) is a model widely used in pattern recognition and machine learning. CRF has been applied to NER for Hindi, achieving an accuracy of 84% for the CoNLL shared task [31]. Gazetteers and features like noun tags have been incorporated to enhance performance in languages such as Telugu [32]. Support Vector Machines (SVM), introduced by Vladimir and Corinna [33], have been employed in NER for Arabic [34] and biomedical text [35]. SVM has also been utilized for Bengali language [36].

V. DEEP LEARNING APPROACH

Neural networks, or deep learning, have gained popularity in recent years. [37] proposed neural network architectures for NER. [38] used word-level representations in deep learning, while character-level representations were explored in [39]. Bidirectional RNN-LSTM was implemented for Hindi by [40], achieving an F1 score of 77.48%. Hybrid bidirectional LSTM and CNN architecture presented by [41] reached an F1 score of 91.62. Word embedding has been employed with Hindi bidirectional LSTM by [42]. The advantage of deep learning is its ability to operate independently of linguistic rules. Supervised learning, including SVM and CRF, necessitates a large collection of annotated data, which is a drawback [43]. In recent times, the deep learning approach, particularly neural networks, has gained prominence in solving the NER problem due to its ability to perform independently of linguistic rules. Various architectures, such as bidirectional RNN-LSTM and hybrid bidirectional LSTM and CNN, have shown promising results across languages. Word embedding is also recognized as an efficient representation in NLP, contributing to the success of deep learning approaches.

VI. HYBRID APPROACH

Hybrid approaches in named entity recognition (NER) aim to combine the strengths of both machine learning and rule-based methods. While machine learning-based NER requires substantial training data for optimal results, rule-based approaches demand the formulation of explicit rules for achieving excellence. In the context of a Bengali NER system [44] proposed a hybrid solution incorporating NER rules, a named entity dictionary, and statistical information on word occurrences. The inclusion of a dictionary significantly enhanced the system's ability to detect name entities (60%) and filter out non-named entities (90%). [45] employed a hybrid approach for Indian languages, combining the Maximum Entropy model (MaxEnt), named entity gazetteers,

and a rule-based approach. This combination sought to leverage the advantages of different methodologies. For Urdu [46] utilized a hybrid approach that integrated rule-based NER with dictionary lookup methods, yielding favorable accuracy. In another instance, a hybrid Bengali NER system was developed by combining machine learning and rule-based approaches, with the machine learning algorithm Hidden Markov Model (HMM) being employed [47]. Notably, many hybrid approaches tend to merge rule-based and machine learning-based methodologies to exploit their complementary strengths.

VII. APPROACHES TO NER IN NE LANGUAGES

North East India comprises of eight states and constitutes approximately 7.9% of India's total geographical area. Despite its relatively small size, the region is home to more than 75% of the languages spoken in India [48]. However, the digital divide is evident, with many languages yet to make significant strides in the realm of digital technology. Only a few languages from the region have seen research initiatives in natural language processing (NLP), especially in named entity recognition (NER).

The first reported Assamese named entity recognition system was developed by [49]. Assamese, the official language of the state of Assam, spoken by over 15 million people, saw various approaches in NER. [50] proposed a suffix-stripping method for building a named entity recognition system in Assamese. Another method, a naive-based approach, achieved an 89% F-score [51]. A hybrid approach for Assamese, combining machine learning and rule-based methods, achieved an 85%-90% F-score [52]. Manipuri, the official language of Manipur, saw an NER approach by [53] using the SVM algorithm, achieving a 94.5% F-score on four named entities. A CRF approach for Manipuri achieved an 83.3% F-score [54]. A hybrid approach, combining CRF and rule-based methods, reached a 93.3% F-score [55]. For the Mizo language, the state language of Mizoram, a rule-based named entity recognition system was attempted by [56]. The approach involved creating a small corpus and designing an algorithm that considers the rich linguistic features of the Mizo language to achieve a robust NER result. Kokborok, the native language spoken in the state of Tripura and neighboring areas, faced challenges addressed in early research. The first Kokborok named entity recognition research utilized frequency analysis of the Kokborok named entity corpus [57]. Machine learning approaches using MIRA [58], CRF [59], and SVM [60] models were also explored. Additionally, a linguistic rule-based approach for Kokborok was developed [61].

VIII. CONCLUSIONS

The evolution of named entity recognition (NER) methods spans a diverse landscape, incorporating rule-based, machine learning-based, and hybrid approaches. The significance of NER extends beyond mainstream languages to encompass the rich linguistic diversity found in North East India, shedding light on the digital disparities and research gaps in this region. Despite its relatively small size, North East India harbors a vast linguistic wealth, with initiatives in NLP, especially NER, gradually emerging. The narrative unfolds with rule-based systems, where linguistic knowledge and heuristic features played a pivotal role, particularly in early works such as company name extraction. As research progressed, machine learning-based approaches, ranging from Maximum Entropy models to Hidden Markov Models, Conditional Random Fields, Support Vector Machines, and eventually deep learning, offered advanced techniques for NER. The hybrid approach, combining the strengths of rule-based and machine learning methods, emerged as a promising strategy, showcasing notable success in various linguistic contexts. The exploration then turns to the specific linguistic landscape of North East India, where languages like Assamese, Manipuri, Mizo, and Kokborok have witnessed pioneering efforts in NER. These endeavors reflect the challenges and opportunities unique to each language, addressing linguistic nuances and contributing to the broader goal of preserving and advancing linguistic diversity. Despite progress, digital disparities persist in North East India, emphasizing the need for continued research initiatives and technology development to bridge these gaps.

REFERENCES

[1]. L. F. Rau, “Extracting company names from text,” 1990, doi: 10.1109/caia.1991.120841.
[2]. B. M. Sundheim, “Overview of results of the MUC-6 evaluation,” 1995, doi: 10.3115/1119018.1119073.
[3]. R. Grishman and B. Sundheim, “Message understanding conference-6: A brief history. In New York,” *Assoc Comput Linguist*, vol. 1, pp. 466–471, 1996.
[4]. E. F. Tjong and K. Sang, “Introduction to the CoNLL-2002 Shared Task: Language-Independent Named Entity Recognition.” 2002, Accessed: Dec. 30, 2023. [Online]. Available: https://aclanthology.org/W02-2024.
[5]. E. F. Tjong, K. Sang, and F. De Meulder, “Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition.” pp. 142–147, 2003, Accessed: Dec. 30, 2023. [Online]. Available: https://aclanthology.org/W03-0419.
[6]. Liu, Y, and F Ren. 2011. "Japanese named entity recognition for question answering system." *IEEE International Conference on Cloud Computing and Intelligence Systems.*

[7]. J. Artiles, A. Borthwick, J. Gonzalo, S. Sekine, and E. Amigó, "WePS-3 Evaluation Campaign: Overview of the Web People Search Clustering and Attribute Extraction Tasks," Accessed: Dec. 30, 2023. [Online]. Available: http://nlp.uned.eshttp//search.intelius.com.

[8]. S. Riedel, L. Yao, and A. Mccallum, "Modeling Relations and Their Mentions without Labeled Text."

[9]. Srihari, R, and W. Li. 2008. ",Information Extraction Supported Question Answering,." *In 8thText Retrieval Conference (TREC-8).*

[10]. Molla, Diego, Menno van Zaanen, and Daniel Smith. 2006. "Named Entity Recognition for Question Answering." *Proceedings of the Australasian Language Technology Workshop 2006.* Sydney, Australia.

[11]. Bhalla, Deepti, Joshi Nisheeth, and Iti Mathur. 2013. "Improving the quality of MT output using novel name entity translation scheme." In *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1548--1553. IEEE

[12]. Kabadjov, M, J Steinberger, and R Steinberger. 2013. "Multilingual Statistical News Summarization." In *Multi-source, Multilingual Information Extraction and Summarization. Theory and Applications of Natural Language Processing*. Springer, Berlin, Heidelberg.

[13]. Nobata, Chikashi, Satoshi Sekine, Hitoshi Isahara, and Ralph Grishman. 2002. "Summarization System Integrated with Named Entity Tagging and IE pattern Discovery." *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02).*

[14]. Cooley, Robert. 1999. "Classification of news stories using support vector machines." *In Proc. 16th International Joint Conference on Artificial Intelligence Text Mining Workshop,.* vol. 26, pp. 28-30.

[15]. Popovski, Gorjan, Stefan Kochev, Korousic-Seljak Barbara, and Tome Eftimov. 2019. "FoodIE: A Rule-based Named-entity Recognition Method for Food Information Extraction." *In ICPRAM.* 915-922.

[16]. Eftimov, Tome, Barbara Seljak Koroušić, Korošec, and Peter. 2017. "A rule-based named-entity recognition method for knowledge extraction of evidence-based dietary recommendations." *PloS ONE* 12(6).

[17]. Grosan, C, and A Abraham. 2011. "Rule-Based Expert Systems." *In: Intelligent Systems. Intelligent Systems Reference Library, Springer, Berlin, Heidelberg* 17.

[18]. Rau, L. F. 1991. "Extracting company names from text." *Proceedings. The Seventh IEEE Conference on Artificial Intelligence Application,* . Miami Beach, FL, USA: IEEE. 29-32.

[19]. Gaizauskas, Robert, Takahiro Wakao, Kevin Humphreys, Hamish Cunningham, and Yorick Wilks. 1995. "UNIVERSITY OF SHEFFIELD: DESCRIPTION OF THE LaSIE SYSTEMAS USED FOR MUC-6." *In Sixth Message Understanding Conference (MUC-6).* Columbia, Maryland,.

[20]. Wakao, Takahiro, Robert Gaizauskas, and Yorick Wilks. 1996. "Evaluation of an Algorithm for the Recognition and Classification of Proper Names." *COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics.*

[21]. Dimitra, Farmakiotou, Karkaletsis Vangelis, Koutsias John, and Sigletos George. 2000. "Rule-based named entity recognition for Greek financial texts." *In Proceedings of the Workshop on Computational lexicography and Multimedia Dictionaries (COMLEX 2000).*

[22]. Piskorski, Jakub. 2004. "Rule-based named-entity recognition for Polish." *In Proceedings of the Workshop on Named-Entity Recognition for NLP Applications held in conjunction with the 1st International Joint Conference on NLP.* Proceedings of the Workshop on Named-Entity Recognition for NLP Applications held in conjunction with the 1st International Joint Conference on NLP.

[23]. Shaalan, Khaled, and Raza Hafsa. 2008. "Arabic Named Entity Recognition from Diverse Text Types." *In International Conference on Natural Language Processing.* Springer, Berlin, Heidelberg. 440-451.

[24]. Becker, Dara, B Bennett, E Davis, D Panton, and K Riaz. 2002. "Named Entity Recognition in Urdu: A Progress Report." In *Proceedings of the International Conference on Internet Computing*, 757--761. Nevada, USA: CSREA Press.

[25]. Toral, Antonio. 2005. "DRAMNERI: a free knowledge based tool to Named Entity Recognition." *In Proceedings of the 1st Free Software Technologies Conference.* pp. 27-32.

[26]. Toral, Antonio, and Rafael. Munoz. 2006. "A proposal to automatically build and maintain gazetteers for Named Entity Recognition by using Wikipedia." *Proceedings of the Workshop on NEW TEXT Wikis and blogs and other dynamic text sources.*

[27]. Borthwick, Andrew, John Sterling, Eugene Agichtein, and Ralph Grishman. 1998. "NYU: Description of the MENE named entity system as used in MUC-7." *In Seventh Message Understanding Conference (MUC-7).* Fairfax, Virginia.

[28]. Saha, Sujan Kumar, Sudeshna Sarkar, and Pabitra Mitra. 2008. "A hybrid feature set based maximum entropy Hindi named entity recognition." In *Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-I.*

[29]. Bikel, D.M, R Schwartz, and Weischedel R.M. 1999. "An algorithm that learns whats in a name." *Machine Learning* 34: 211–231 .

[30]. Zhou, GuoDong, and Jian Su. July 2002. "Named Entity Recognition using an HMM-based Chunk Tagger." *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL).* Philadelphia. pp. 473-480.

[31]. McCallum, A, and W Li. 2003. " Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons." *In Proceedings of HLT-NAACL 2003 - Volume 4, CONLL '03, ACM*

[32]. Srikanth, P, and Kavi Narayana Murthy. 2008. "Named Entity Recognition for Telugu." *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages.* pp. 41–50.

[33]. Vladimir, Vapnik, and Cortes Corinna. 1995. "Support-vector networks." *Machine Learning* 20 (3): 273-297.

[34]. Benajiba, Yassine, Mona Diab, and Paolo Rosso. 2008. "Arabic named entity recognition: An svm-based approach." *Proceedings of 2008 Arab international conference on information technology (ACIT).* Amman, Jordan: Association of Arab Universities. pp. 16-18.

[35]. Ju, Zhenfei, Jian Wang, and Fei Zhu. Wuhan, 2011,. "Named Entity Recognition from Biomedical Text Using SVM." *5th International Conference on Bioinformatics and Biomedical Engineering.* pp. 1-4.

[36]. Ekbal, Asif, and Sivaji Bandyopadhyay. 2008. "Bengali named entity recognition using support vector machine." *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages.* pp 51–58.

[37]. Collobert, Ronan, and Jason Weston. 2008. "A unified architecture for natural language processing: Deep neuralnetworks with multitask learning." *In Proceedings of the 25th international conference on Machine learning, pages 160–167. ACM.*

[38]. Strubell, E, P Verga, D Belanger, and A McCallum. 2017. "Fast andaccurate entity recognition with iterated dilated convolutions." *in ACL, pp. 2670–2680.*

[39]. Kuru, O., O. A. Can, and D. Yuret. 2016. "Charner: Character-levelnamed entity recognition." *in COLING, pp. 911–921.*

[40]. Vinayak, Athavale, Bharadwaj Shreenivas, Pamecha Monik, and Prabhu, Manish Shrivastava Ameya. 2016. "Towards Deep Learning in Hindi NER: An approach to tackle the Labelled Data Sparsity." In *Proceedings of the 13th International Conference on Natural Language Processing*, 154-160.

[41]. Chiu, J. P., & Nichols, E. (2016). Named entity recognition with bidirectional LSTM-CNNs. *Transactions of the Association for Computational Linguistics*, *4*, 357-370.

[42]. Shah, Bansi, and Sunil Kumar Kopparapu. "A Deep Learning approach for Hindi Named Entity Recognition." *arXiv preprint arXiv:1911.01421* (2019).

[43]. Bandyopadhyay, Sivaji. 2008. "Invited Talk: Multilingual Named Entity Recognition." In *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages*. Asian Federation of Natural Language Processing.

[44]. Chaudhuri, Bidyut Baran, and Suvankar Bhattacharya. 2008. "An Experiment on Automatic Detection of Named Entities in Bangla." *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages.* 75–82.

[45]. Saha, Sujan Kumar, and Sandipan Dandapat, Sudeshna Sarkar and Pabitra Mitra Sanjay Chatterji. 2008. "A Hybrid Approach for Named Entity Recognition in Indian Lan." *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages.* 17-24.

[46]. Saha, Sujan Kumar, and Sandipan Dandapat, Sudeshna Sarkar and Pabitra Mitra Sanjay Chatterji. 2008. "A Hybrid Approach for Named Entity Recognition in Indian Lan." *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages.* 17-24.

[47]. Drovo, Mah Dian, Moithri Chowdhury, Islam Uday Saiful, and Amit Kumar Das. 2019 . "Named Entity Recognition in Bengali Text Using Merged Hidden Markov Model and Rule Base Approach." *7th International Conference on Smart Computing & Communications (ICSCC).*

[48]. Bareh, Curiously. 2013. "Languages of North East India: A Comparative and Contrastive Analysis on Some of the Phonological Features." *International Journal of Linguistics* 5 (5): pp 45-58.

[49]. Sharma, Padmaja, Utpal Sharma, and Jugal Kalita. 2011. *Machine Intelligence: Recent Advances.* Delhi: Narosa Publishing, ISBN 978-81-8487-140-1.

[50]. Sharma, P., Sharma, U., & Kalita, J. (2012). "Suffix stripping based NER in Assamese for location names." *2nd National Conference on Computational Intelligence and Signal Processing (CISP).* pp: 91-94.

[51]. Talukdar, Gitimoni, Pranjal Protim Borah, and Arup Baruah. 2014. "Assamese Named Entity Recognition System Using Naive Bayes Classifier." In *Advances in Computing and Data Sciences*, edited by Mayank Singh, P. K. Gupta, Jan Flusser and Tuncer Oren, pp 35-43. Springer Singapore.

[52]. Sharma, P., Sharma, U., & Kalita, J. (2016). "Named entity recognition in Assamese: A hybrid approach." International Conference on Advances in Computing, Communications and Informatics (ICACCI).

[53]. Singh, Thoudam Doren, Kishorjit Nongmeikapam, Asif Ekbal, and Sivaji Bandyopadhyay. 2009. "Named Entity Recognition for Manipuri Using Support Vector Machine." *Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation, Volume 2.* Hong Kong.

[54]. Nongmeikapam, Kishorjit & Shangkhunem, Tontang & Chanu, Ngariyanbam & Singh, Laishram & Salam, Bishworjit & Bandyopadhyay, Sivaji. 2011. "CRF based Name Entity Recognition (NER) in Manipuri: A highly agglutinative Indian Language." *2nd National Conference on Emerging Trends and Applications in Computer Science.* pp. 1-6.

[55]. Jimmy, L, and Darvinder Kaur. 2013. "Named Entity Recognition in Manipuri: A Hybrid Approach." In *Language Processing and Knowledge in the Web*, edited by Iryna Gurevych, Chris Biemann and Torsten Zesch, 104--110. Springer Berlin Heidelberg.

[56]. Bentham, J., P Pakray, S. Lalbiaknia G. Majumder, and A. Gelbukh. 2016. "Identification of Rules for Recognition of Named Entity Classes in Mizo Language." In *Fifteenth Mexican International Conference on Artificial Intelligence (MICAI)*.

[57]. Debbarma, A, Bipul Shyam Purkayastha, and Paritosh Bhattacharya. "Frequency Based Named Entity Recognition System for Under Resource Language." 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT).

[58]. Patra, B. G., N Debbarma, Aby Abahai T., D. Das, and S Bandyopadhyay. 2015. "Named Entity Recognizer for less resourced language Kokborok." *International Conference on Asian Language Processing (IALP).* International Conference on Asian Language Processing (IALP).

[59]. Debbarma, A, Bipul Shyam Purkayastha, and Paritosh Bhattacharya. 2016. "CRF Approaches to Kokborok Named Entity Recognition, a Low Resource." *International Journal of Control Theory and Applications* 9 (42).

[60]. Debbarma, A, Bipul Shyam Purkayastha, and Paritosh Bhattacharya. 2017. "SVM approach to name entity recognition for low resource language." *Journal of Advanced Research in Dynamical and Control Systems* (2): 738-747.

[61]. Debbarma, A, Bipul Shyam Purkayastha, and Paritosh Bhattacharya. 2019. "Named Entity Recognition for a Low Resource Language." *International Journal of Recent Technology and Engineering, ISSN: 2277-3878.*