Journal of Computer Science

Original Research Paper

Customized Named Entity Recognition Using Bert for the Social Learning Management System Platform CourseNetworking

[1]Kayal Padmanandam, [2]KVN Sunitha, [3]Behafarid Mohammad Jafari, [4]Ali Jafari, [5]Mengyuan Zhao and [1]Nikitha Pitla

[1]*Department of Information Technology, BVRIT HYDERABAD College of Engineering for Women, Hyderabad, India*
[2]*Department of Computer Science and Engineering, BVRIT HYDERABAD College of Engineering for Women, Hyderabad, India*
[3]*Department of Electrical and Computer Engineering, Indiana University-Purdue University Indianapolis, Indianapolis, USA*
[4]*Department of Computer and Information Technology, Indiana University-Purdue University Indianapolis, Indianapolis, USA*
[5]*Department of Instructional System Technology, School of Education, Indiana University, Bloomington, USA*

Article history
Received: 08-02-2023
Revised: 07-07-2023
Accepted: 04-08-2023

Corresponding Author:
Behafarid Mohammad Jafari
Department of Electrical and Computer Engineering, Indiana University-Purdue University Indianapolis, Indianapolis, USA
Email: bmohamma@purdue.edu

Abstract: Named Entity Recognition (NER) is an information extraction task and one of the most researched applications to extract knowledge from massive data. Conventional NER systems identify predefined entities like name, person, location, organization, time, etc. However, there is a limitation to identifying user-defined entities that are specific to an application. This challenge introduces the concept of customized NER. For instance, if a learning management system like CourseNetworking (CN) needs to identify the skill set of a user from their posts, the existing pre-trained NER models cannot be used. To overcome this information extraction limitation, we propose a customized named entity recognition system for the CN platform using the deep learning model, Bi-Directional Encoder Representation from Transformer (BERT) which is a transformer-based deep learning technique where all output elements are connected to all input elements with dynamic weight connections. The proposed customization model can be employed to train any entity of user choice with a decent amount of training dataset. The model shows 70-72% recall and F1-Score varied on the number of epochs trained. This model is used in various applications like fraudulent detection, recommendation systems, and business intelligence.

Keywords: Customized Named Entity Recognition, Named Entity Recognition (NER), Information Extraction, BERT, Simple Transformers, CourseNetworking

Introduction

With the rapid growth of the internet and online platforms, Social Networking Sites (SNSs) have developed significantly. Facebook, Twitter, and Instagram are a few examples of SNSs. Users can share content using these platforms. Some types of learning materials, such as teaching videos, can be delivered through public SNSs. However, these platforms have limited efficiency for academic purposes due to various reasons including the distractions they may cause for the students since they also serve other purposes Nkhoma *et al.* (2015); Rambe (2012). CourseNetworking (CN) is a Learning Management System (LMS) that focuses on academic social networking and sociocultural learning theory. CN is designed to be an academic networking platform to effectively enhance user communication and engagement. CN is an assortment of SNS and LMS, benefitting the fortes from both platforms by supporting course management as in LMSs and by connecting learners for several opportunities as in Jafari and Baylor (2012). The CN platform was seed-funded by Indiana University. In 2012 CN started as a social LMS and then developed and commercialized at IUPUI cyber lab until today. CN has millions of users from over 160 countries. Users can sign up for a free account by visiting thecn.com and then build a lifelong ePortfolio or share free personal courses Jafari *et al.* (2022). CN is a successful platform for engaging students and building helpful communities among them (Berlin and Weavera, 2022; Cooney *et al*., 2020; Gavrin and Lindell, 2017; Scherzinger, 2021). CN is continuously developed with new services which strengthen the CN environment. After multiple years of extensive research, Rumi was launched in October 2020 to CN end users for being a digital personal assistant powered by machine learning intelligence. Rumi enhances and customizes the services on CN. It proposes different services to the end user such as recommendations for connections, community, learning tips, teaching tips, user announcements, and many more Jafari *et al*. (2022).

This study aims to develop a new Rumi feature that will intelligently tag skill-related keywords based on user content on their CN ePortfolio using the custom NER model building. The proposed tagging is significant for the CN application as it extracts the skill set of the users. This extraction will be helpful for our future research to build a social network community that can dynamically identify and form clusters of skilled personnel based on skill similarities.

Related Works

An enormous amount of data is generated every day on social media. Usefully structured information can be gained by processing this unstructured data. However, taking proper processing steps to fetch useful information can be very challenging Jafari *et al*. (2023). Named Entity Recognition (NER) is a primary yet challenging step in knowledge acquisition (Akhtyamova, 2020) and the core part of Natural Language Processing (NLP) Shelar *et al*. (2020).

In this section, we discuss NER and custom NER (user-defined NER). NER, as a sub-task of NLP, refers to the process of data extraction to recognize and classify named entities in unstructured text data into their appropriate categories including locations, organizations, times, etc., Shah and Bhadka (2017); Stepanyan (2020). The applicability of NER models is broad, ranging from recognizing dates and cities in chatbots to open-domain question answering Tarcar *et al*. (2019). These models are often used in the fields of Information Retrieval System (IRS) Brandsen *et al*. (2022), and Machine Translations (Xie *et al*., 2022). Question and answering Banerjee *et al.* (2021) text summarization Khademi and Fakhredanesh (2020), etc.

Research on NER has been in existence for 20 years Sun *et al*. (2018) and several models have been developed to recognize named entities in this previous research. Initial NER models were designed based on user-defined rules, language orthographic features, lexicons, and their relationships (Yadav and Bethard, 2019). Machine learning was then used to improve the early models and create new NER models. A few years later, NER with neural network and minimized feature engineering gained attention as they tend to exclude specialized resources and make them more domain-independent Yadav and Bethard (2019). In previous classical NER models, proper nouns alike person, organization, and location are the only entities highlighted in the text. However, it is possible to develop novel NER models that recognize different types of entities. Such entities need not be a noun at all Brandt (2021). Various customized NER models based on user-specified entities are possible Shelar *et al*. (2020).

Custom (user-defined) NER models are created and trained to serve different purposes. NER models can be customized to identify drug names, diseases, or symptoms in medical texts and prescriptions Ramachandran and Arutchelvan (2021), or they can be designed in a way that they recognize entities based on the job description to rank resumes Satheesh *et al*. (2020). In this row, our paper aims to develop a customized (user-defined) named entity recognition model to recognize various skills of CN users.

NER Approaches

There are different approaches to executing the NER tasks. Earlier NER systems used a rule-based approach where all the rules are handcrafted, whereas nowadays machine learning approaches are employed for NER tasks. To go further with the dominant approach and better results, neural network approaches are the most helpful ones today.

Rule-Based Approaches

These kinds of approaches require building the patterns annually and they follow regular expression format to match word sequences. Tokenizers are used to simplify this process with finite definitions for each word in the sequence.

Data-Driven Approaches

The limitation of the rule-based approach is that it cannot be applied to all different tasks as they are using finite definitions. However, data-driven approaches support automatic rule building using machine-learning approaches. Supervised approaches use a massive amount of data that are annotated, to study the details from the corpus data for auto rule generation. The approach used for sequence labeling can be the HMM (Bikel *et al*., 1999). Support Vector Machine (SVM) method (Takeuchi and Collier, 2002) maximum entropy model (McCallum *et al*., 2000), or Conditional Random Field (CRF) (McCallum and Li, 2003). Semi-supervised learning makes use of data with annotation and data without labeling. Initially, it uses the data that is annotated to boot the machine learning. Unsupervised learning uses only unlabeled data and instrument clustering approaches for categorization. Few research Alfonseca and Manandhar (2002) have assigned appropriate named entity types in WordNet to unseen entities.

DOI: 10.3844/jcssp.2024.88.95

Neural Network Approaches

Recently, neural networks have been relatively prevalent in NER. Supervised learning approaches require linguistic researchers to provide a large number of features to generate rules. But with the advent of the neural network, which is widely used for feature extraction, it is probable for anyone to leverage the NER approach with less or no linguistic acquaintance. LSTM, BI-LSTM, and CNN methods are used to achieve better performance.

Materials and Methods

A corpus is a collective document that contains annotations with entity types. Since 2005 datasets have been developed on diversified text sources ranging from Wikipedia articles to natural conversations, social media, and many different generated texts. Quite a lot of NER tools are accessible on the web which are available as trained models. NER models (English) developed by industry and academia are discussed in Li *et al.* (2020). The general process of corpus creation is shown in Fig. 3.

The proposed research uses the dataset created by the authors, named CN skillset. It is real-time data with extracted entities, based on CN users' showcases on their portfolios. The raw dataset has around 50,000 entities. The dataset attributes are skill and its usage (represented as a number of references) as shown in Table 1. The skill column denotes the admin-created skillset and the number of occurrences explains how many times each skill is used in the user-generated texts on CN. Each skill must have been used a minimum of two times to be extracted to the CN skillset library. CN skillset is shown in Fig. 1. As per the CN skillset, the NER Dataset (as shown in Fig. 2) is created with three attributes including sentence #, word, and tag. The attribute1 'sentence#' acts as an identifier and attribute2 'word' is tokenized and mapped to the appropriate 'Tag' (attribute 3) which has the target either as Skill (Skill) or Others (O).

Table 1: Skill reference table (CN_skillset)

Skill	Number_of_references
GardenDesign	2
Gardening	45
GCCS	23
GelElectrophoresis	11
GelExtraction	2
GenderAnalysis	5
GenderEquality	14
GenderStudies	23
GeneticCounseling	2

```
{'label': 'Skill', 'pattern': 'GardenDesign'},
{'label': 'Skill', 'pattern': 'Gardening'},
{'label': 'Skill', 'pattern': 'GCCS'},
{'label': 'Skill', 'pattern': GelElectrophoresis'},
{'label': 'Skill', 'pattern': 'GelExtraction'},
{'label': 'Skill', 'pattern': 'GenderAnalysis'},
{'label': 'Skill', 'pattern': 'GenderEquality'},
{'label': 'Skill', 'pattern': 'GenderStudies'},
{'label': 'Skill', 'pattern': 'GeneticCounseling'}
```

Fig. 1: CN_skillset data

Sentence #	Word	Tag
Sentence: 1	GeneticCounseling	Skill
	informs	O
	genetic	O
	conditions	O
	,	O
	GenderStudies	Skill
	is	O
	a	O
	study	O
	of	O
	academic	O
	field	O
	.	O
Sentence: 2	GardenDesign	Skill
	is	O
	an	O
	upcoming	O
	field	O
	in	O
	recent	O
	trends	O

Fig. 2: NER_dataset

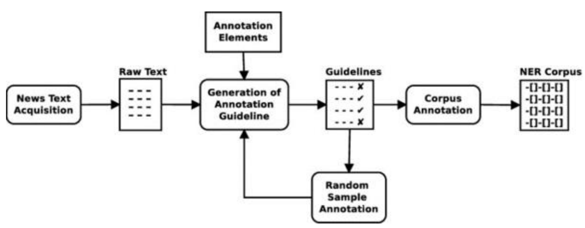

Fig. 3: The process of corpus creation

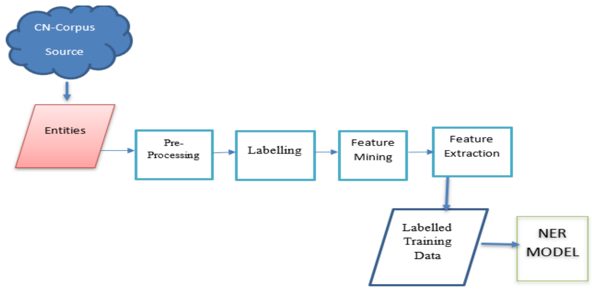

Fig. 4: NER architecture

NER Architecture

Named entity recognition is a significant application of NLP which is capable of working on unstructured data. There are several different pre-trained models publicly available to implement the NER based on the task requirements. The general architecture of the NER model is depicted in Fig. 4.

NER system can identify a specific entity from the raw data and categorize it. NER system reads the sentence and highlights particular entities in the sentences. Entities differ from one application to the other and might not be the same for different projects. General entities such as names, locations, organizations, dates, etc., can be identified using pre-trained models like Spacy and Stanford NER. On the other side, if the idea is domain-specific, the NER model must get customized training, which necessitates huge human participation. The annotation process handled is explained in Algorithm 1.

As an NER application for the CN platform, we must develop a customized NER that reads the text and highlights the entities called "Skill" from the sentences which is explained in Algorithm 2 and 3.

Algorithm 1: Corpus creation

Input: 5000 Raw Data Entities

Output: NER_Dataset corpus

1.**Filter entities using occurrences > 2 [**CN_skillset]
2.Map user sentence from CN for the entities filtered
3.**For each sentence [**user-generated CN statement]
 Create Sentence #, word, tag
  For each word in the sentence
   if the word is a (Skillset)
   Annotate '**Skill**"
   else
   Annonate '**O'**
  End
 End
End
4. **Return NER**_Dataset corpus

Algorithm 2: NER training model

Input: Pre-Processed CN User statement dataset

Output: NER Trained Model

1. **Initialize: Load CN dataset-**Ner_Dataset
2. **Begin**
3. LabelEncoder applied to the dataset
4. Separate X data, Y label
5. Spilt into Training(X) and Testing(Y) Data
6. Import NER Model, NERArgs
7. Set NERArgs arguments
8. Train Model using Bert(X)
9. Test model using eval_model(Y)
10. **Return** Model output saved
11. **End**

Algorithm 3: NER model

Input: Any real-time statement by the CN user

Output: Named Entity recognition for Skillset

1. **Algorithm 1**
2. **Algorithm 2**
3. Input user statement (Unseen statement)
4. NER model identifies the entity 'Skill'

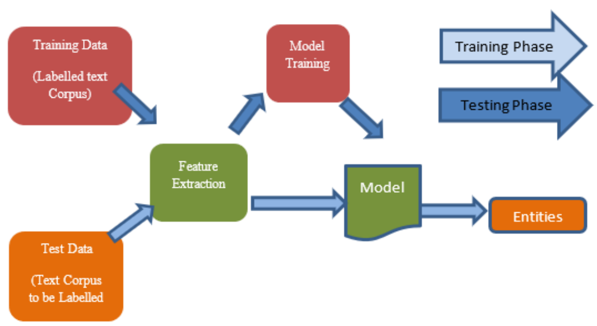

Fig. 5: NER training and testing workflow

During the training phase, the system understands and learns from the training data to fit the model as shown in Fig. 5. The idea behind this learning is to arbitrarily identify new entities from the set of samples learned. The training data is the human-annotated tags for the named entity "Skill". During the training phase, feature extraction plays a significant role in identifying the right entity based on the preceding 'n' words and the succeeding 'n' words in the sentence.

During the testing phase, the NER system should identify entities from an unseen text (sentence) that has not been trained during the model training. The testing phase reuses the feature extraction module from the training phase which should be a high-value pipeline for high expected throughput.

Table 2: Different tasks and their model names

Task	Model class name
Binary and multiclass text classification	ClassificationModel()
Language generation	LanguageGenerationModel()
Conversational AI (Chabot training)	ConvAIModel()
Multi-label text classification	MultiLabelClassificationModel()
DocumentRetrieval	RetrievalModel()
Language modeltraining /fine-tuning	LanguageModelingModel()
Named entityrecognition	NERModel()
Text Representation Generation	RepresentationModel()

The NER Simple Transformer Model

There are various Transformer models for NLP tasks. Each model is capable of doing the intended purpose with surpassed performance. For instance, the classification model performs binary and multiclass text classification, ConvAIModel does conversational chatbot training and the MultimodalClassificationModel performs multimode classification where users can work on text and image data. The simpleTransformersModel and their task are shown in Table 2.

In this row, the NERModel is used to work on a named entity recognition task. To work with the NERModel, the user must specify the model_type and a model_name. For the proposed customization we have used the Bert model_type and the model_name to specify the architecture of the NERModel which is already built and trained.

NER Metric Evaluation

NER systems evaluation is done by associating their output to the manually assigned actual output. This relative assessment can be measured by the exact match or related match evaluation. In this study, we evaluate our model based on the exact match method.

Exact-Match Evaluation

NER applies boundary detection and type identification to determine whether the instance is recognized correctly. With the help of the table of confusion matrix, we calculate the precision, F-score, and recall of the NER model.

Terms Used

1. False Positive (FP): The NER system determines that the entity is present but it is not
2. False Negative (FN): The NER system determines that the entity is absent but it is present
3. True Positive (TP): The NER system determines that the entity is present and it is present
4. Precision denotes the fraction of the results that are properly predicted by the NER model
5. Recall denotes the percentage of entire entities Properly predicted by the NER model:

$$Precision = \frac{No.\ of\ TP}{No.\ of\ (TP + FP)}$$

$$Recall = \frac{No.\ of\ TP}{No.\ of\ (TP + FN)}$$

6. F-score is the harmonic mean of precision and recall. It is computed as follows:

$$F - score = \frac{2 * Precision * Recall}{Precision * Recall}$$

Results and Discussion

The model is trained with the CN dataset. Each sentence is encoded with an identifier using the Label Encoder which is in the SK learn Library. It helps to normalize the non-numeric sentence label to a numeric value. In our case, for instance, it converts the sentence # from "sentence: 393" to number only "393" as shown in Fig. 6(a-b).

Sentence #	Word	Tag
Sentence :393	hardware	O
Sentence :393	and	O
Sentence :393	software,	Skill
Sentence :393	Zebra	Skill
Sentence :393	enables	O

(a)

Sentence #	Word	Tag
391	GeneticCounseling	Skill
391	informs	O
391	genetic	O
391	conditions	O
391	,	O

(b)

Fig. 6: (a) Before label encoder; (b) After label encoder

Table 3: Metric evaluation

	No. of epochs					
Metric	1	2	3	4	5	10
Eval_loss	0.309	0.306	0.312	0.361	0.411	0.743
Precision	0.677	0.676	0.642	0.725	0.647	0.635
Recall	0.675	0.562	0.654	0.618	0.723	0.716
F1_score	0.676	0.614	0.648	0.667	0.680	0.673

```
[[{'I': 'O'},
  {'love': 'O'},
  {'to': 'O'},
  {'read': 'O'},
  {'books': 'O'},
  {'related': 'O'},
  {'to': 'O'},
  {'any': 'O'},
  {'Engineering': 'SKILL'}]]
```
(a)

```
[[{'I': 'O'},
  {'love': 'O'},
  {'to': 'O'},
  {'read': 'O'},
  {'books': 'O'},
  {'related': 'O'},
  {'to': 'O'},
  {'ArtificalIntelligence': 'SKILL'}]]
```
(b)

```
[[{'GeneticCounseling': 'SKILL'},
  {'informs': 'O'},
  {'genetic': 'O'},
  {'conditions,': 'O'},
  {'GenderStudies': 'SKILL'},
  {'is': 'O'},
  {'a': 'O'},
  {'study': 'O'},
  {'of': 'O'},
  {'academic': 'O'},
  {'field': 'O'}]]
```
(c)

```
[[{'Design': 'SKILL'},
  {'is': 'O'},
  {'visual': 'O'},
  {'communication': 'SKILL'},
  {',so': 'O'},
  {'campaign': 'SKILL'},
  {'design': 'SKILL'},
  {'is': 'O'},
  {'the': 'O'},
  {'visual': 'O'},
  {'vocabulary': 'O'},
  {'of': 'O'},
  {'your': 'O'},
  {'ads': 'O'}]]
```
(d)

Fig. 7(a-d): The NER model identifies the entity

Next, the dataset is divided into training and testing with a proposition of 80 and 20 respectively. The NER model arguments are the number of training epochs, the learning rate, the training batch size, and the evil batch size, which are required to build the model. Once the argument values are defined, they can be passed as an argument in the NER model call. The model gets trained with the training data and evaluated on the test data. The metric evaluation based on the number of epochs is shown in Table 3.

The NER model that is built for CN application works accurately, to correctly classify text into skills. Figure 7 shows how the customized NER in CN identifies the entity "Engineering", "Artificial Intelligence", and "Genetic Counselling" as 'SKILL'. In the below example, all the tokens are classified into ‘other Tag-O’ whereas the customizedentity is classified as ‘SKILL’.

Conclusion

This study focuses on customizing name entity recognition tasks for CourseNetworking platforms. NER model BERT was used to achieve the expected results. By using the CN datasets that are annotated to a large extent, it can be concluded that, the larger the training dataset better the result, which will be discussed in future works with benchmark comparisons.

The system is then tuned to optimize the hyper-parameters on training data and the best F1-score of 68% on test data was achieved. We also found that the model can possibly attain improved performance with a higher amount of training data and the epochs used for training the model.

Acknowledgment

This study is supported by CourseNetworking (thecn.com). We would like to thank the CN technical team for providing the skills dataset and the technical support whenever required.

Funding Information

This study is supported by CourseNetworking (thecn.com).

Author’s Contributions

Kayal Padmanandam: Participated in all experiments, coordinated the data analysis, and contributed to the written of the manuscript.

KVN Sunitha: Coordinated the research plan.

Behafarid Mohammad Jafari: Participated in defining the research, coordinated the data retrieval, and contributed to the written of the manuscript.

Ali Jafari: Participated in leading the research defined the goals and objectives and reviewed the content.

Mengyuan Zhao: Participated in leading the research, defined the tasks, and contributed to the written of the manuscript.

Nikitha Pitla: Participated in data extraction and data annotation.

Ethics

The authors own the research and the result isanalyzed in a truthful and complete manner.

References

Akhtyamova, L. (2020). Named entity recognition in Spanish biomedical literature: Short review and BERT model. In *2020 26th Conference of Open Innovations Association (FRUCT)* (pp. 1-7). IEEE. https://doi.org/10.23919/FRUCT48808.2020.9087359

Alfonseca, E., & Manandhar, S. (2002). Proposal for Evaluating Ontology Refinement Methods. In *LREC*. http://www.lrecconf.org/proceedings/lrec2002/pdf/38.pdf

Banerjee, P., Pal, K. K., Devarakonda, M., & Baral, C. (2021). Biomedical named entity recognition via knowledge guidance and question answering. *ACM Transactions on Computing for Healthcare*, *2*(4), 1-24. https://doi.org/10.1145/3465221

Berlin, K., & Weavera, K. V. (2022). Teaching Strategies Students Find Helpful in Online Learning Courses. *College Teaching*, *70*(3), 319-327. https://doi.org/10.1080/87567555.2021.1940814

Bikel, D. M., Schwartz, R., & Weischedel, R. M. (1999). An algorithm that learns what's in a name. *Machine Learning*, *34*, 211-231. https://doi.org/10.1023/A:1007558221122

Brandsen, A., Verberne, S., Lambers, K., & Wansleeben, M. (2022). Can BERT dig it? named entity recognition for information retrieval in the archaeology domain. *Journal on Computing and Cultural Heritage (JOCCH)*, *15*(3), 1-18. https://doi.org/10.1145/3497842

Brandt, A. (2021). *The Importance of Domain-Specific Expertise in Training Customized Named Entity Recognition Models* (Bachelor's thesis). https://studenttheses.uu.nl/handle/20.500.12932/41157

Cooney, E. M., Freije, E., & Zhao, M. A. (2020). Using ePortfolios to Facilitate Transfer Student Success. ASEE. https://doi.org/10.18260/1-2--35463

Gavrin, A., & Lindell, R. S. (2017). Connecting students' homework to their participation in a course-based social network. http://dx.doi.org/10.18260/1-2--28070

Jafari, A., & Baylor, A. L. (2012). From course management to coursenetworking: Conceptualizing a new learning environment based on social networking. *United States of America*.

Jafari, B. M., Luo, X., & Jafari, A. (2023). Unsupervised Keyword Extraction for Hashtag Recommendation in social media. In *The International FLAIRS Conference Proceedings* (Vol. *36*). https://doi.org/10.32473/flairs.36.133280

Jafari, B. M., Zhao, M., & Jafari, A. (2022). Rumi: An intelligent agent enhancing learning management systems using machine learning techniques. *Journal of Software Engineering and Applications*, *15*(9), 325-343. https://doi.org/10.4236/jsea.2022.159019

Khademi, M. E., & Fakhredanesh, M. (2020). Persian automatic text summarization based on named entity recognition. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 1-12. https://doi.org/10.1007/s40998-020-00352-2

Li, J., Sun, A., Han, J., & Li, C. (2020). A survey on deep learning for named entity recognition. *IEEE Transactions on Knowledge and Data Engineering*, *34*(1), 50-70. https://doi.org/10.48550/arXiv.1812.09449

McCallum, A., & Li, W. (2003). Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons. https://doi.org/10.3115/1119176.1119206

McCallum, A., Freitag, D., & Pereira, F. C. (2000). Maximum entropy Markov models for information extraction and segmentation. In *ICML* (Vol. *17*, No. 2000, pp. 591-598). https://dl.acm.org/doi/10.5555/645529.658277

Nkhoma, M., Cong, H. P., Au, B., Lam, T., Richardson, J., Smith, R., & El-Den, J. (2015). Facebook as a tool for learning purposes: Analysis of the determinants leading to improved students' learning. *Active Learning in Higher Education*, *16*(2), 87-101. https://doi.org/10.1177/1469787415574180

Ramachandran, R., & Arutchelvan, K. (2021). Named entity recognition on bio-medical literature documents using hybrid-based approach. *Journal of Ambient Intelligence and Humanized Computing*, 1-10. https://doi.org/10.1007/s12652-021-03078-z

Rambe, P., (2012). Constructive disruptions for effective collaborative learning: Navigating the affordances of social media for meaningful engagement. *Electronic Journal of E-Learning*, *10*(1), pp132-146. https://academic publishing.org/index.php/ejel/article/view/1621

Satheesh, K., Jahnavi, A., Iswarya, L., Ayesha, K., Bhanusekhar, G., & Hanisha, K. (2020). Resume Ranking based on Job Description using SpaCy NER model. *International Research Journal of Engineering and Technology*, *7*(05), 74-77. ISSN: 2395-0056.

Scherzinger, L. N. (2021). CourseNetworking: How a Social Media Platform Can Connect Us in a Crisis. https://doi.org/10.14434/jotlt.v9i2.31176

Shah, D. N., & Bhadka, H. (2017). A survey on various approach used in named entity recognition for Indian languages. *International Journal of Computer Applications*, *167*(1), 11-18.

Shelar, H., Kaur, G., Heda, N., & Agrawal, P. (2020). Named entity recognition approaches and their comparison for custom NER model. *Science and Technology Libraries*, *39*(3), 324-337. https://doi.org/10.1080/0194262X.2020.1759479

Stepanyan, L. (2020). Automated custom named entity recognition and disambiguation. *International Journal of Signal Processing*, *5*. ISSN: 2367-8984.

Sun, P., Yang, X., Zhao, X., & Wang, Z. (2018, November). An overview of named entity recognition. In *2018 International Conference on Asian Language Processing (IALP)* (pp. 273-278). IEEE. http://dx.doi.org/10.1109/IALP.2018.8629225

Takeuchi, K., & Collier, N. (2002). Use of support vector machines in extended named entity recognition. In *Coling-02: The 6th Conference on Natural Language Learning 2002 (CoNLL-2002)*. https://doi.org/10.3115/1118853.1118882

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

Xie, S., Xia, Y., Wu, L., Huang, Y., Fan, Y., & Qin, T. (2022). End-to-end entity-aware neural machine translation. *Machine Learning*, 1-23. https://doi.org/10.1007/s10994-021-06073-9

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