Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 7
July 2023

Machine Translation Approach for Resolution of Part of Speech Ambiguity from English to the Sanskrit Language

Archana Sachindeo Maurya

Institute of Technology (IoT), Department of DCSIS, SRMU, Lucknow, UP, India

*(Corresponding Author)

Promila Bahadur*

Associate Professor, Institute of Engineering Technology (IET), Lucknow, UP, India

Divakar Yadav

Professor, Institute of Engineering Technology (IET), Lucknow, UP, India

Abstract

Machine Translation is one of the most important techniques of Natural Language Processing (NLP). It is an automated process of translation through a computer system. Now, due to advancement, there exist many efficient ways to express machine translation. Machine Learning (ML) is being extensively used in MT and has become an interesting area of research in the last three years. Resolving semantic ambiguity in natural language is a major challenge in MT. ML has given promising results in terms of system learning and predicting results. The text classification technique in ML is considered one of the most important methods to resolve Word Sense Disambiguation (WSD). The role of the dataset both as training and test data is important to predict the required results. We collected a total of 2,000 sentences, which we divided into training and testing data. The dataset plays an essential role in validating the output of the system. We have also done an analysis on supervised machine learning text classification algorithms, namely Naïve Bayes', Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Neural Network, Logistic Regression, and Random Forest. The accuracy of the given algorithms ranges between sixty-eight to eighty-four percent. Further, we have also developed a hybrid model. In this proposed model, we have combined the Naïve Bayes', Support Vector Machine, and Decision Tree algorithms to achieve better results.

We have analyzed our proposed "hybrid model" for prediction of POS ambiguity. The proposed model has a reported success rate of eighty-five percent. All the algorithms and hybrid model are tested on the ML tool WEKA. The "hybrid model" is also tested on the programming language Java. The accuracy of the algorithms and model is reported using the ten-fold cross-validation method.

The model has also reported high precision, recall, and F-score in comparison to all other supervised machine learning classification algorithms. The correctness of the algorithm and model is tested in terms of the total number of correct POS predicted. The algorithms and proposed model are analyzed with the help of a machine learning tool named as "AmbiF".

Keywords: Artificial Intelligence, Machine Translation, Machine Learning, Hybrid Approach for WSD, Sanskrit Translation, Bayes' Network, Naïve Bayes', SVM, Decision Tree, Random Forest, KNN.

1. Introduction

Nowadays, Machine Learning (ML) is an important aspect of the rising field of Artificial Intelligence (AI). ML is one of the most exciting and challenging technologies because it tends to make computers similar to human beings. We can say that ML makes computers learn. ML is actively being used today in every aspect of life. The term "ML" was defined by Arthur Samuel and he described ML as "a field of study that trains computers to enable an automatic and improving learning process from experience without any human intervention, i.e. the capability to

learn without being explicitly programmed" [1].

ML is useful in many real-life applications, like- web search engines, photo tagging applications, spam detectors, machine translation, etc.

Machine Learning Applications

Machine Translation (MT), Information Retrieval (IR), Information Extraction (IE), and Speech and Language Recognition (SLR) are some of the most common machine learning applications in Natural Language Processing (NLP). Among all these applications, MT is the most important, where we translate the text and speech from one language into another. Machine learning algorithms play a pivotal role in machine translation [2]. MT is the translation of text from a source language to a target language with the use of computers. MT is the automatic translation of the text from the source language (SL) to the text in the target language (TL) without any human intervention. The translation process is depicted in Fig. 1.

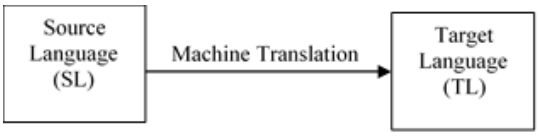

Fig. 1 Machine Translation Process

India is a multilingual country, and approximately three percent people of the population can understand the English language [3]. Therefore, we need the MT system to translate any information from the English language to the target language. On the basis of their translation methodology, The MT system can be broadly classified into four categories. These are direct, rule-based, knowledge-based, and corpus-based. MT plays an important role in reducing the language barrier among different regions due to the increased web data [4].

Each MT approach has its merits and demerits. Today, these approaches are most widely used, where different combinations are tested. The main aim of MT is to obtain an error-free translation, but errors can occur due to stylistic and structural differences among languages [5]. These stylistic and structural differences are of different types, like word order, word sense, pronoun resolution, idioms, and ambiguity. Out of these ambiguities, is the main issue in machine translation is how to resolve them.

Ambiguity Challenge in Machine Translation

Ambiguity is an open challenge in natural languages. Every evolved language has ambiguous words, i.e. the words that have more than one meaning in the source language; the problem is called ambiguity problem. Before translation, it is necessary to resolve this challenge. Numerous words exist in the source language that has different meanings and they can be translated in various senses. For example, in English-to-Sanskrit translation, the term bear appears in the English sentences "I can bear (verb) the responsibility of my class" and "He is going out on a bear (noun) hunt, which could translate to either "carry the weight of "(□ □ □ □ □) or "a bear is an animal" (□ □ □ □ □).

There are various kinds of ambiguity, like lexical, syntactic, pragmatic, semantic, and part of speech ambiguity. In this paper we will focus only on part of speech ambiguity. Part of speech ambiguity can be defined as "when a single word has more than one part-of-speech, like noun, verb, adjective, adverb, etc." For example, consider two sentences, S1 and S2:

S1: "She rides the skateboard very well (adverb)."

S2: "This well (noun) is very deep and dry."

In the first sentence, the word "well" means "in a good sense", which is an adverb, and in the second sentence, the word "well" means "hole for oil and water", which is a noun.

Ambiguity in the English Language

All Indian languages face various types of ambiguities. English is assumed to be the most important language in the world. Due to the rationale that English is the maternal language in many countries. There is a limitless vocabulary in the English language that has separate denotations in different frameworks, referred to as "ambiguous words", Words in the English language that have nouns, adjectives, verbs, or adverbs are a topic of concern because they can create part-of-speech ambiguity [6, 7]. When we translate from English to any other natural language, these ambiguous words need to be disambiguated properly for the correct translation in the target language. An example of ambiguous English words and their senses with their

respective parts of speech is shown in Fig. 2.

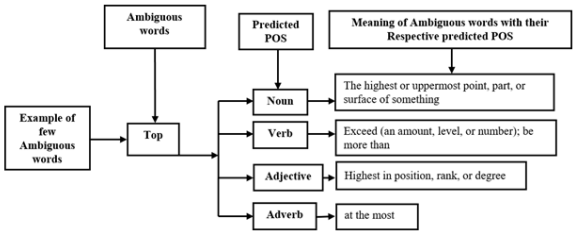

Fig. 2 Few English ambiguous words with their POS and meaning

2. Ambiguity Resolution And Word Sense Disambiguation (Wsd)

In machine translation, data can be translated from one dialect to another dialect with the help of machines. Almost all natural languages have reported different kinds of ambiguities. When we translate from one language to another, these ambiguous words need to be properly disambiguated properly for the appropriate translation into either the source or target language. The ambiguity problem can be resolved with the disambiguation process. WSD is the disambiguation method for selecting the correct meaning of an ambiguous word.

WSD is frequently defined as the process of determining the POS and specific meaning or right sense of an ambiguous word in a given sentence. There are various approaches that can be applied to WSD process. These approaches are knowledge-based, supervised learning, and unsupervised learning.

In this paper, we have analyzed common supervised learning approaches. All the discussed supervised learning algorithms and hybrid model are tested using the programming language Java, and the accuracy of each is reported below.

Developed Model AmbiF

Tool of hybrid model is developed using programming language JAVA named as “AmbiF”. “AmbiF” stands for “Ambiguity Free” word analysis. The programme is developed to detect and resolve POS ambiguity. This plugin is implemented in EtranS software to provide solutions for ambiguous words. Fig.3. shows the screen shot of the developed model with EtranS software:

Fig. 3. Screenshot of EtranS software powered by AmbiF

3. Organization Of The Paper

This paper is structured from sections 3 to 8 as follows:

Section 3: Here we have discussed the role of training and testing data in ML.

Section 4: In this section, we have discussed different approaches for the prediction and segmentation of test data.

Section 5: In this section, we have discussed our proposed hybrid model along with its functionalities.

Section 6: In this section, we have done the result analysis.

Section 7: In this section, we have discussed the social and material impact of the research.

Section 8: In this section, we have discussed the conclusion and future scope of the research.

4. Role Of Dataset In Machine Learning

Datasets play a significant role in ML. A dataset can be classified as the training data, and the test Data. **Training data** contain domain specific-information. In the current scenario, Training data contains ambiguous words. For example, as shown in Table 1, the given sentences have “watch” as an ambiguous word. The system is trained to extract features in the vicinity of the given word to predict the POS.

Table 1: Sample of Training Dataset with POS category

Training Sentence No.	Training Sentences	POS Category
1	This watch is ten minute fast.	Noun
2	Watch this funny movie today.	Verb
3	I had my watch stolen yesterday.	Noun

Test data is a set of sentences that are tested on a given model. The accuracy of the prediction is directly proportional to the training provided to the model. The 'better trained a model is, the better the results are', the sample of the test data is shown in Table 2.

Table 2. Test data set with Prediction of POS

Test Sentence No.	Test Sentences	POS Category
1	My watch is five minute slow.	Noun
2	Watch this movie with your family.	Verb

To check the efficiency of different models, we have taken a sample size of two-thousand sentences and testing has been done under ten-fold cross-validation. In the ten-fold cross-validation test method, the data set is divided into ten equal parts, and in each iteration, nine parts of the given data set are used for training purposes and tenth the part is used to predict the test data.

Format of the training data-set for the predictions of the test data-set

For the preparation of the training data-set, first of all, we have extracted the neighboring words around the target ambiguous word; for example, training sentences in Table 1 contain ambiguous word "watch". The features for the ambiguous word watch are extracted using window size 3. It means three words to the left and three words to the right of the target ambiguous word, as shown in Table 3. We have also stored position-specific information, like POS of the neighboring words. Neighboring words are considered features and used for further processing.

Table 3. Extracted features of the ambiguous word "watch" from the Training dataset taking window size 3

S.No.	Sentences	W-3	POS(W-3)	W-2	POS(W-2)	W-1	POS(W-1)	W	POS(W)	W+1	POS(W+1)	W+2	POS(W+2)	W+3	POS(W+3)	Category
1	This watch is ten minutes fast.	nil	nil	nil	nil	this	detminer	watch	noun	ten	noun	minutes	noun	fast	adjective	Noun
2	Watch this funny movie.	nil	nil	nil	nil	nil	nil	watch	verb	this	determiner	funny	adjective	movie	noun	Verb
3	I had my watch stolen yesterday	nil	nil	I	pronoun	my	pronoun	watch	noun	stolen	verb	yesterday	noun	nil	nil	Noun

5. Supervised Wsd Approaches For Machine Learning

There are many ML algorithms that can help computers make correct classifications based on the training data and extracted features. The training data are manually created and used to train the system. The training dataset consists of sentences with the target word, their features, and a specified class for that word. In the process of WSD, these algorithms are helpful in identifying the correct sense of an ambiguous word [8]. Various commonly used supervised machine learning approaches are discussed below:

Decision Tree Learning

Decision tree learning is a tree-based learning method that contains non-leafy nodes, leafy nodes, and branches. The different parts of the tree contain the following information:

1. Rules are defined in the form of the tree structure.
2. Non-leaf nodes denote tests.
3. The branches denote the test results.
4. The leaf nodes carry the different senses.

Random Forest Algorithm

Random Forest (RF) is also one of the most powerful and widely used methods for data exploration, data modelling, and predictive modelling. This classifier is an ensemble algorithm in which various decision trees are grouped together. An ensemble decision tree will have a low variance and a high accuracy value in comparison to a single decision tree.

A random forest can be built by using the decision trees for the same dataset, but the trees cannot be correlated. The result of this algorithm will be a tree constructed from the results of separate decision trees [9].

Each decision tree contributes votes for the segmentation of a new item based on its qualities or attributes, and the tree with the most votes is chosen for segmentation.

Neural Network Learning

The brain is one of the most complicated organs in the human body, and the main task of the brain is to learn new things. Neural networks are computing devices, which are parallel to each other and are capable of attempting brain functions. The main objective of this learning method is to develop a model of the brain so that a system can perform the computational tasks much faster than the current system. These computational tasks include segmentation, data clustering, pattern recognition, optimization, etc.

Artificial Neural Network Learning (ANNs) is one of the most important and significant features of neural networks because they have the ability to learn like the human brain [10]. ANN is made up of the processing units that are called neurons. A neuron has input units called dendrites and output units called synapses or axons.

Bayesian Learning

The Bayesian learning method is one of the most popular methods for document categorization. This is the probability based method and provides a quantitative approach for evaluating the performance of other algorithms. In Bayesian learning, the conditional probability of an event occurring is calculated, and that event is also correlated with some other events [11].

Naïve Bayes' Algorithm a supervised probability-based machine learning algorithm that works on the Bayes' theorem. This algorithm is mostly used for document segmentation problems. There are many variables present in the training dataset, and these variables are independent of each other. These independent variables are called "features." To compute the likelihood of certain features in a given class, this approach uses the Bayes' theorem [12, 13, and 14]. In this context, the uppermost value shows the most "significant" class. A Naïve Bayes' model is easy to implement and useful for high-dimensional training datasets, i.e. those with a large number of rows and columns. It is simple and easy to implement, and also outperforms all other classification methods.

The following formula can be used to compute the probability value for each class:

$$C = \underset{c}{\operatorname{argmax}} \quad P(c) \prod_{x=1}^{m} P\left(\frac{fx}{c}\right) \ldots\ldots\ldots (8)$$

$$C = \operatorname{argmax}_c P(c/f1, f2, \ldots\ldots\ldots fx) \ldots\ldots (6)$$

$$C = \underset{c}{\operatorname{argmax}} \frac{P\left(f1,f2,\ldots\ldots\frac{fx}{c}\right)P(c)}{P(f1,f2,\ldots\ldots fx)} \ldots\ldots\ldots\ldots (7)$$

Here,

C represents the POS of the ambiguous word.

The selected feature is represented by f1, f2, …….fx.

x the total number of retrieved features.

Instance-Based Learning

With the instance-based learning approach, the system can generalize new instances based on some shared properties after learning from the training data. Another name for instance-based learning is memory-based learning, or lazy-learning. The time

complexity of this learning algorithm depends on the size of the training data set. The most popular instance-based learning algorithm is K-Nearest Neighbor (KNN). The KNN method is a non-parametric grouping method that is basic but active in many situations [15]. This method saves all available cases and categories and creates new ones based on the votes of its k neighbours [16]. KNN is a popular statistical method for segmentation and is used for unlabeled observations after assigning them to the class. Features of observations are collected for the training dataset and the testing dataset. The Euclidean distance between two points can be calculated using the formula below:

$$= \sqrt{((X_2 - X_1)^2 + (Y_2 - Y_1)^2)} \ldots\ldots\ldots (4)$$

Here,

X1, and Y1, are the coordinates of the test word.

X2, and Y2 are the coordinates of matching feature.

Support Vector learning

Support Vector Machine is like a sharp knife, and it works on both small and complex datasets. It is a stronger and more powerful algorithm for building machine learning models. It is mostly used to solve categorization challenges. We can draw each data element as a [17] point in n-dimensional space with a value for each characteristic using this technique. After that, we can perform segmentation by finding the hyperplane. This hyperplane clearly differentiates the two classes [18].

Logistic Regression

This is a probability-based method supervised text segmentation. It is a predictive algorithm. This algorithm is used when the datasets are unconditional and the output will be in binary form. The segmentation problems based on the binary output are known as binary segment problems [19].

6. Proposed Hybrid Model Approach

We have presented a hybrid model based on ML algorithms. The hybrid model is used for obtaining improved and accurate results for the categorization of ambiguous words [20]. In this method, we have combined three supervised ML algorithms. These algorithms are Naïve Bayes', Support Vector Machine (SVM), and Decision Tree. On the basis of the training dataset, a hybrid model has been developed for removing the POS ambiguity and identifying the correct category of the ambiguous word on the basis of training data-set. In this model, we have combined the three supervised ML algorithms. These algorithms are Naïve Bayes', Support Vector Machine, and Decision Trees. The evaluation was done on the AmbiF machine learning tool [21]. The detailed hybrid model is depicted in Fig. 4.

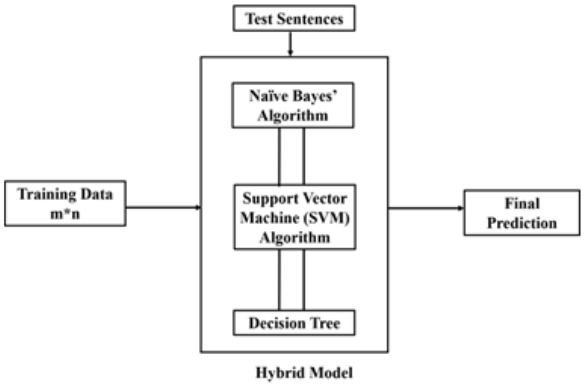

Fig. 4. The Hybrid Model

The complete work flow of the developed model is shown in Fig. 5, data flow in the model is depicted with the help of a flow chart.

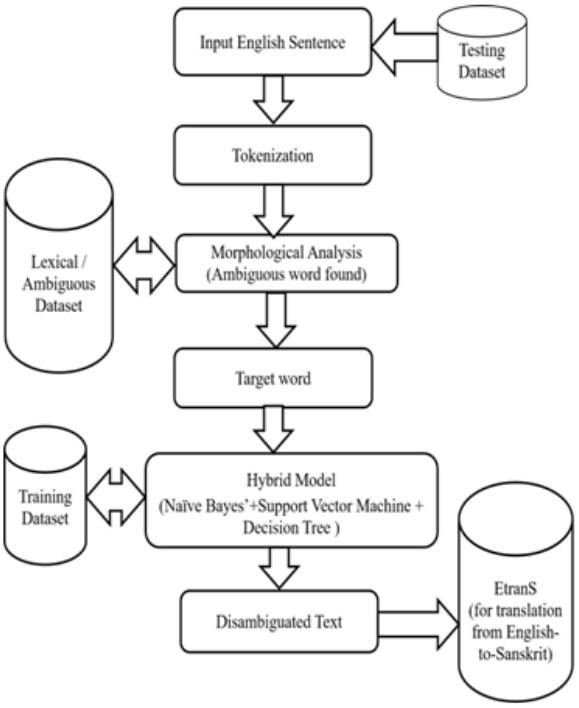

Fig. 5. The work flow of the Hybrid Model

The ambiguous words are grouped into five classes, namely nouns, verbs, adverbs, prepositions, and adjectives, on the basis of the part of speech of the ambiguous target word in the sentence. The selected attributes consist of nominal attribute types. This method of feature extraction is based on the collocation model of the Naïve Bayes' segmentation approach.

7. Testing Method

The experiment was done under the ten-fold cross-validation testing method. This testing method is the most common data resampling methodology for evaluating the model's prediction for the class. The dataset is randomly partitioned into ten parts for the cross-validation for ten folds. The data set is divided into ten equal parts. Out of these ten parts, nine are used for training purposes, while the tenth part is used for testing.. The process is repeated ten times, with the tenth part being tested each time [22].

For example, suppose we have a dataset of hundred sentences from S1 to S100, and the value of k is 10. The value of k determines the number of folds that are used to split the dataset. First of all, we shuffle the data and then split it into ten folds. We have a hundred sentences, then each fold will contain ten sentences, and a total of ten iterations will be performed.

The data is submitted for the Naïve Bayes', support vector machine, and decision tree algorithms. In all the ML algorithms, the training set was prepared in order to learn a model. So it can be capable of classifying the occurrences of data into well-known classes. The segmentation process of any algorithm involves the following steps:

1. Preparing the training dataset in ARFF or CSV format.

2. Identification of attributes and classes.

3. Using the training dataset, learning model is developed.

4. The final prediction for the class of new data through the hybrid model.

8. Result And Discussion

A dataset of two thousand sentences is evaluated in terms of the training and testing datasets. The dataset considered has a variety of POS like noun, verb, adjective, adverb, preposition classes, etc.

Different supervised ML algorithms, namely- Naïve Bayes', Support Vector Machines, Decision Tree, Neural Network, Random Forest, Logistic Regression, and K-Nearest Neighbor are tested on the given dataset using the plugin tool AmbiF. Hybrid Model is also tested using AmbiF. The result is generated on the basis of the prediction reported on the correct POS. The result analysis is given in table 4.

Table 4: The outcome of a ten-fold cross-validation test

S. No.	Algorithm	No. of Correct Classifications	Accuracy (Percentage)
1	Logistic Regression	1370	68.47
2	Naïve Bayes' Classifier	1525	76.24
3	IBK	1469	73.45
4	SMO	1665	83.23
5	Decision Tree (J48)	1692	84.33
6	Random Forest	1407	70.34
7	Neural Network	1520	76.00
8	Hybrid Approach	1696	**84.78**

Table 5 has further depicted the precision, recall, and F-score of the above mentioned algorithms and model [23, 24]. We can conclude that hybrid models have better precision, recall, and F-score in comparison to other ML algorithms. The correctly segmented dataset in the hybrid model is eighty-five percent, while in other analyzed methods it ranges from sixty-eight to eighty-four percent.

Table 5: Average accuracy for all the approaches

S.	Algorithm	Precision	Recall	F-

No.				**score**
1	Logistic Regression	0.63	0.672	0.652
2	Naïve Bayes' Classifier	0.763	0.762	0.761
3	IBK	0.737	0.734	0.735
4	SMO	0.832	0.832	0.832
5	Decision Tree (J48)	0.846	0.844	0.843
6	Random Forest	0.742	0.703	0.692
7	Neural Network	0.772	0.760	0.762
8	Hybrid Approach	0.849	0.848	**0.848**

The accuracy of the "F-measure" for the hybrid model is reported 0.848. Thus, Hybrid Model could be a better approach to resolve ambiguity for existing Machine Translation models for example EtranS. A comparative chart is shown in Fig. 6 with respect to a number of correct analysis.

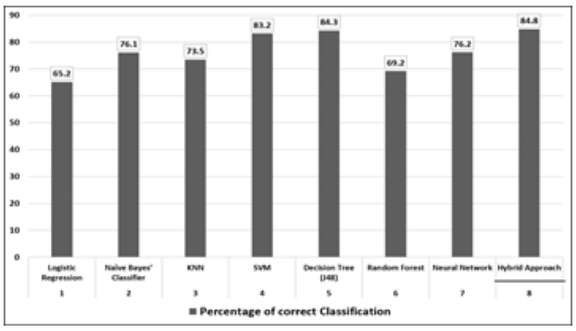

Fig. 6 Chart showing accuracy percentage of six analyzed supervised learning algorithm and hybrid approach

9. Conclusion And Future Scope

Dealing with semantic ambiguity in NLP has been an open challenge in the field of machine translation. In this paper, we have outlined and analyzed a hybrid model involving MT and subsequently analyzed it using ML techniques. Ambiguity refers to a word having multiple meanings, senses, POS, etc. The central idea of the paper is to provide a solution to the POS ambiguity issue using existing machine learning algorithms. Different ML algorithms have been tested in terms of accuracy and efficiency on the pre-processed data set. The sentences are classified as nouns, verbs, adjectives, adverbs, and prepositions on the basis of the respective parts of speech of the ambiguous words. All the algorithms are tested on the machine learning tool WEKA along with the proposed hybrid model. The proposed hybrid model is also tested on the programming language Java. The performance rate of the analyzed algorithm ranges from sixty-eight to eighty-four percent on the given data set.

The hybrid model has been proposed and developed on Java for better accuracy and has a reported success rate of 85 percent. The proposed hybrid model is implemented as a plugin tool on the machine translation software EtranS (currently providing English-Sanskrit translation). The proposed system helps make better sense of ambiguous words. The performance of the algorithms and model is evaluated on the basis of precision, recall, and F-measure. The efficiency of the system can also be improved by increasing the window size of the neighbouring words in the future.

References

[1] Samuel, Arthur L. (1959). "Some Studies in Machine Learning Using the Game of Checkers". IBM Journal of Research and Development. 44: 206 226. CiteSeerX 10.1.1.368.2254.doi:10.1147/rd.441.0206

[2] Ray, S. (2019, February). A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 35-39). IEEE.

[3] Sinha, R. M. K., & Jain, A. (2003). AnglaHindi: an English to Hindi machine-aided translation system. In Proceedings of Machine Translation Summit IX: System Presentations.

[4] Zachary, W. W. (1979). A survey of approaches and issues in machine-aided translation systems. Computers and the Humanities, 17-28.

[5] Zelta, E.N. (2014). Ambiguity. In Stanford Encyclopedia of Philosophy. Retrieved from http://plato.stanford.edu.R. Navigli, "Word sense disambiguation: A survey," ACM Computing Surveys (CSUR), p. 10, 2009.

[6] Palanati, D. P., & Kolikipogu, R. (2013). Decision list algorithm for word sense disambiguation for TELUGU natural language processing. Int. J. Electron. Commun. Comput. Eng, 4(6), 176-180.

[7] Richard Laishram Singh, Krishnendu Ghosh, Kishorjit Nongmeikapam and Sivaji Bandyopadhyay,

”A DECISION TREE BASED WORD SENSE DISAMBIGUATION SYSTEM IN MANIPURI LANGUAGE”, Advanced Computing: An International Journal (ACIJ), Vol.5, No.4, July 2014, pp 17-22.

[8] Maurya, A. S., & Bahadur, P. (2022). A Detailed Analysis of Word Sense Disambiguation Algorithms and Approaches for Indian Languages. In Proceedings of Second Doctoral Symposium on Computational Intelligence (pp. 693-710). Springer, Singapore.

[9] Venkatesan, N., & Priya, G. (2015). A study of random forest algorithm with implementation using weka. International journal of innovative research in computer science and engineering, 1(6), 156-162.

[10] Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of pharmaceutical and biomedical analysis, 22(5), 717-727.

[11] Mitchell, T. M. (1999). Machine learning and data mining. Communications of the ACM, 42(11), 30-36.

[12] Szepesvári, C. (2010). Algorithms for reinforcement learning. Synthesis lectures on artificial intelligence and machine learning, 4(1), 1-103.

[13] Escudero, G., Màrquez, L., & Rigau, G. (2000). Naive Bayes and exemplar-based approaches to word sense disambiguation revisited. arXiv preprint cs/0007011.

[14] Nyein Thwet Thwet Aung, Khin Mar Soe, Ni Lar Thein, “A Word Sense Disambiguation System Using Naïve Bayesian Algorithm for Myanmar Language”, International Journal of Scientific & Engineering Research Volume 2, Issue 9, September-2011, pp. 1-7.

[15] Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003, November). KNN model-based approach in classification. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 986-996). Springer, Berlin, Heidelberg.

[16] Wang, L. (2019, December). Research and implementation of machine learning classifier based on knn. In IOP Conference Series: Materials Science and Engineering (Vol. 677, No. 5, p. 052038). IOP Publishing.

[17] Chauhan, N. S. (2020). Decision tree algorithm, explained—Kdnuggets. Available from: https://www. Kdnuggets. com/2020/01/decision-tree-algorithm-explained. html. [Accessed 9 May 2021].

[18] Jeyanthi, S. N. (2007). Efficient classification algorithms using SVMs for large datasets. Supercomputer Education and Research Center.

[19] Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of pharmaceutical and biomedical analysis, 22(5), 717-727.

[20] Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I. H., & Trigg, L. (2009). Weka-a machine learning workbench for data mining. In Data mining and knowledge discovery handbook (pp. 1269-1277). Springer, Boston, MA.

[21] Dhar, S., Roy, K., Dey, T., Datta, P., & Biswas, A. (2018, December). A hybrid machine learning approach for prediction of heart diseases. In 2018 4th International Conference on Computing Communication and Automation (ICCCA) (pp. 1-6). IEEE.

[22] Liu, Y., Wang, X., Wang, L., &Lv, Z. (2019). A Bayesian collocation method for static analysis of structures with unknown-but-bounded uncertainties. Computer Methods in Applied Mechanics and Engineering, 346, 727-745.

[23] Browne, M. W. (2000). Cross-validation methods. Journal of mathematical psychology, 44(1), 108-132.

[24] Tripathi, M., & Taneja, A. (2019). K-fold cross-validation machine learning approach on data imbalance for wireless sensor network. Int. J. Sci. Res. Eng. Trends, 5(5), 1590-1595.