[ICRTCET-2018] ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

TEXT ANALYSIS AND SPEECH TO SIGN LANGUAGE TRANSLATION

Narendra Bhusakhare, Mausumi Goswami, Ashvinee, Jitesh Singh & BS Purkayastha

ABSTRACT

American Sign Language is the most widely known sign language. There are more sign languages used as well. The current scenario is that deaf individuals have already learned the language and use it for their daily communication. The only hurdle is that normal person have to learn the sign language In this paper an architecture is proposed based on machine learning. The system is designed in 3 modules - speech to text, text to sign and sign to animation. The first module is implemented using a speech recognition API. The second through a machine learning algorithm. The end module consists of a 3D animated avatar.

.

Keywords: *American Sign Language, KNN, SVM, Naïve Bayes, Speech to text , text to sign.*

I. INTRODUCTION

The deaf community faces a lot of difficulties in communication. It becomes more difficult to talk to them being a normal person. Thus deaf communities have come up with sign languages. American Sign Language is the most widely known sign language. There are more sign languages used as well.

This paper implements a translator which is going to empower deaf person to communicate with others those who don't know or have not learnt any sign language. The application listens to others' speech and convert it to sign language and also demonstrate the conversion through a humanoid avatar.

II. PROBLEM FORMULATION

The current scenario is that deaf individuals have already learned the language and use it for their daily communication. The only hurdle is that normal person have to learn the sign language to visually receive it as well as to convey the message. So, this idea eases the part of the normal person. They don't need to learn the sign language if they make use of this paper. The major area of research is the use of machine translation while translating from English texts to American Sign Language. The translation should precisely analyze the sentences (in English) to easily convert it to sign languages.

III. PROBLEM IDENTIFICATION

The paper is designed in 3 modules - speech to text, text to sign and sign to animation. The first module is implemented using a speech recognition API. The second through a machine learning algorithm and the end module consists of a 3D animated avatar. This paper uses an approach based on Machine Learning for text

Now comes the classification which is applied onto the pre-processed text data. There are several classifiers such as Decision Tree classifiers, Rule Based classifiers, Support vector machine, Naive Bayes classifier, k- nearest neighbors algorithm etc. All the classifiers have its own good and bad making them suitable for specific models.

IV. LIMITATIONS

This is an academic research based paper and has its own limitations. This system accepts only a single input language, i.e. English. Also, this system gives output in only one sign language, i.e. ASL. Few complex sentences have been tested to retrieve the meaning but this natural language is more complex in terms of context and sentence formation. Thus, this paper may not be able to perform for highly complex sentences or the lengthy ones.

ISSN 2348 – 8034
Impact Factor- 5.070

V. LITERATURE REVIEW

This part of the paper consists of a literature survey about the papers that has been done till now in this domain. Various pros and cons of each system has been studied. Some important works that has been done so far in this area has been discussed below.

TEAM: This is the first system to use machine translation for translation of English texts to American Sign Language using 3D animation. It takes into account all the required details to display animation. This system shows the minute details of Sign Language. This system can be implement to other languages also because this system is universal. To display American Sign Language sentences elaborate 3D human model is necessary.

Tessa: TESSA system based on direct translation approach. It is a speech to British Sign Language translation system which provides communication between a deaf person and normal people who doesn't know about the sign language. TESSA takes English as input text, lookup each word of the English string in the English-to-Sign dictionary, concatenates those signs together, and blends them into an animation.

An animation system to display Chinese Sign Language, after translation from Chinese texts, has been developed by Jing Wang, Yanfeng Sun and Lichun Wang, particularly for mobile users. The system works by using 3D animation on mobile devices. Mobile users are increasing every day. Their algorithm segments the Chinese text into words, known as Chinese-oriented semantic analysis algorithm. They have developed the avatar on their own and named it as Blue -Jane. It is a 3D virtual model to perform the animation on the mobile screen. It uses rendering to display the animation on the screen. The animation displayed in this system lacks continuity. The authors have used an approach Double Buffer based fast rendering to resolve this issue.

ViSiCast: ViSiCAST project, developed by European Union, aims to meet the needs of people with hearing disabilities. This system translates English words to American Sign Language and Dutch or German Sign Languages as well. In addition to broadcast subtitles, the system is innovative in the field of recognition of a limited range of gestures that allow people with disabilities to communicate with hearing in a social context, such as post offices. CMU Link Parser is used which is an approach for analysis of English text. The data produced as an output of analysis is transformed using the grammar of decorative language.

Text to Sign Language Synthesis Tool: The main aim of this application is to provide a computer-based sign language output for the deaf community. The application takes English text as its input and generates 3D animated VRML sequences which can be rendered in any VRML-compliant browser. The technique initially converts all individual symbols which are present in sign box to sequences of MPEG-4 Face and Body animation parameters. These generated sequences animates the H-anim complaint VRML avatar with SNHC BAP and FAP players. Hence, this application can be used as a learning tool for general people and also for people who are keen to understand the sign language.

VI. BACKGROUND

Text classification is the process of classifying text documents into a predefined set of classes. It is a supervised learning approach. Text classification has gained its importance, resulting in the application of various data mining algorithms for text domain. All the algorithms stated has its own pros and cons and a good performance on classification demands the right choice of classifier for the right problem.

A. KNN Classification

The k-nearest neighbors algorithm is classification technique. It is also used for regression technique. k-NN is also known as instance-based learning or lazy learning. The k-NN algorithm is simplest machine learning algorithms. The k-nearest neighbors algorithm is an extension of the nearest neighbors algorithm. Euclidean distance metric is commonly used for continuous variable. The overlap metric which is also known as Hamming distance is used for discrete variables .The discrete variables are the text classification. k -NN is a linear search method. It compute distance of each data presented in the dataset with respect to all the data present in the dataset and it returns the k closest points.

The Euclidean distance mathematical formula:

 àà ô

R

B. SVM Classifier

The Support Vector Machine is a classification technique as well as regression technique. SVM is supervised learning model. SVM gives a cleaner and more powerful way of learning complex nonlinear functions.

SVM provide a fast and effective classification. In SVM, all the points of the datasets are separated by gaps and a new value is predicted based on, in which category it will fall. The SVM provide classification decisions for the new input.

C. Naive Bayes Classifiers

Naive Bayes Classifiers is one of the highly practical learning method. It is based on strong (naive) assumptions that values are conditionally independent of attribute values and the target value.

Naive Bayes Classifiers is a popular methods for classifying natural language text documents. It can be trained for supervised learning. Parameter estimation for naive Bayes models uses the method of maximum likelihood

Naive Bayes Classifier:

  à

where V N B denotes the target value output by the naive Bayes classifier.

Table I. speech to sign systems

S. No.	Paper	Gap Analysis
1	Maria Papadogiorgaki, Nikos Grammalidis, DimitrioTzovars as and Michael G. Strintzis,"Text-to-sign language synthesis tool", (2002)	This system faces difficulty in animating complex movements. Complex movements includes contact symbols such as the touch between the hands or the touch between face and hand.
	Julia S. Manueva and Mikhail G. Grif,"Russian Sign Langua Machin	The system does not considers the semanti component The

2	ge e Interpreter System Based on the Analyses of Syntax and Semanti Constructi c on", (201 6)	c s. semantics of the sign language is needs to be addressed to develop a proper application.
3	Steph Cox Micha en , el Linc Jud oln and y Tryggva Melan son, ie Nakisa, Mark Wells, Marcus an San Abbo Tutt d ja tt, "TESS syste A, a m to aid communic dea ation with f people", (2004)	This system has a domain bounded applicat cove onl pos ion. It rs y t offic conversation offic e with an e clerk. Also, it requires more use of human effort to train the model.
4	Mate Ahmed Mujta en , ba Idre Abidee es, Zain ul n, Rafi Mumt an a az d Sana Khalique," Deaf Talk Using Animat Sig 3D ed n Language", (2016)	As the developed system is based on Kinect from Windows v2, so the syste compati onl wit m is ble y h those devices that support Kinect v2.Also, Microsoft has discontinued Kinect technology.
5	Lich Wang, Wan un Jing g, and Yanfeng Sun, "Chinese Langua Animati Sign ge on System On Mobile Devices", (201 0)	It can convert to Chinese sign language only. The used model in this project is not simplified and rendering nee mor algorithms ds e optimizati on.
	Geetha. Rahu M, l Muraleedh Kira aran, n A Pras and Ruvia J	This system has a domain

ISSN 2348 – 8034
Impact Factor- 5.070

6	ad nsh Raghav “Animatio an, n Syst Indi Sig em for an n Langua Communica ge tion usin LOT Notatio g S n”, (201 3)	bounded application. It has been developed for banking purposes only.
7	Parteek Kumar and Rupinder Kau “HamNoS r, ys Generation System for Sign Language”, (2014)	This paper doesn’t provide support fo glob system doesn r al s. It ’t transl t America Sig ate o n n Langua ge.
8	Ashok Kumar Sahoo, Gouri Sank Mishr an ar a d Syed Faraz Ali, “Domain Bounded Engli Indi sh To an Languag e Translation Model”, (2013)	For Indian Sign Language there is particul grammat no ar ical rule whic mak it difficult h es for the syntax semantanaly and ic sis as there are no rules to compare the English text with. The system that they have proposed is for railway reservation counters for enquiry.
9	Eva Safar and Ian Marshall “Th Architect e ure of an English-Text-to-Sign-Languag es Translation System”, (2001)	The text analysis phase of the project has been implemented with a lot of complexity.
10	Veronic Tap a ia, Raul Reinoso, Carrill Edison o, Dennis Cadas and Leonardo	The developed system is no longer available. Also, the use of robotic

	Criollo "Translation System Of Voice And Text To Language of Signs", (2017)	hand demonstration is expensive.
11	Malu S Nair, Nimitha and A P Idicul Sumam Mary a "Conver sion Malayalam of text to Indian sign language usin g synthetic animation", (201 6)	This paper only works for regional language. No support for ASL or any other globally accepted sign language.
12	Ian Marshall and E´ va Safar "A Prototype Text to British Sign Languag e (BSL) Translation System", (2003)	The main omission in the system currently is the absence of non-manual components of signing, though the SL generation has been designed to be extended in this direction in the future.

VII. METHODOLOGY

After the analysis of all the existing system thoroughly and going through the techniques of text analysis, an architecture has been proposed for this system as shown in Fig. 1. The standalone app can be casted into 3 phases; text to speech, speech to sign and avatar animation. The proposed architecture gives a more brief detail about it. Also, the sign language requires its own grammar, semantic structure and rules during the translation process which are all accessed through resources.

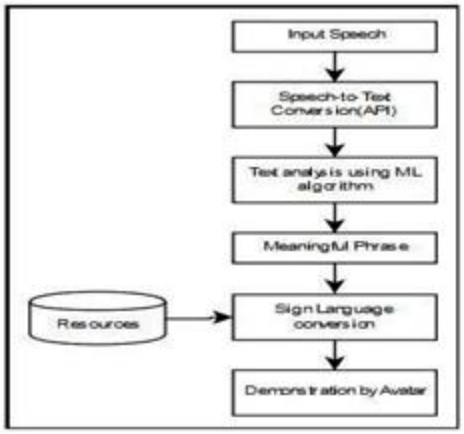

Fig. 1. Architecture of the proposed system

VIII. RESULTS

The Speech API when implemented in the Web platform morphs into an interface given in Figure 2. The 'Microphone' button enables the device microphone to start listening for input. The Speech API displays the recognized utterances in the text form in a text box visible to the user. The user can proceed wilfully after the text box shows the intended speech.

The API also allows a user to change the given input instantly.

Fig. 2. Interface of Speech Recognition API

The API supplies its output to this part of the application. This works on the back-end of the application. The text analysis module implemented through python language generates a satisfactorily output. This module has been tested with few sentences that can give sign demonstrations. Some examples are mentioned here. Figure 3 and Figure 4 are the results and representation in graphical form respectively for the example sentence: "Charlie and Jenny are very good friends."

ISSN 2348 – 8034
Impact Factor- 5.070

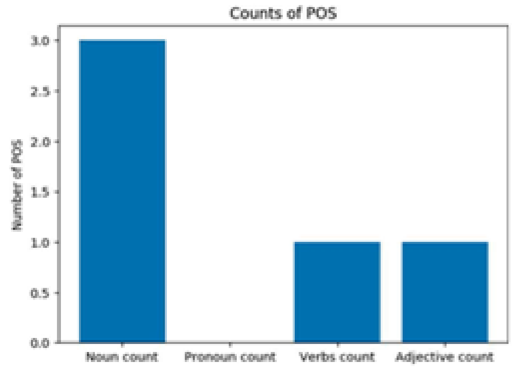

Fig. 3. Graph of the text tokenization

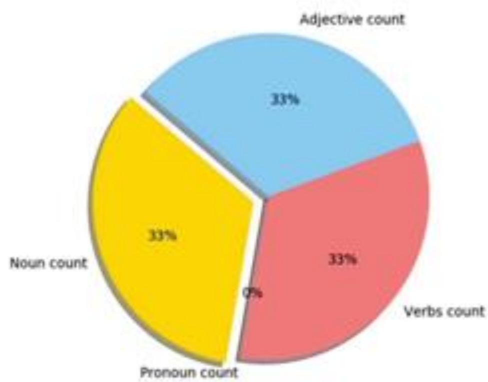

Fig. 4. Pie chart representation of the text tokenization

Figure 5 and Figure 6 give the output for another example sentence: "He is going to get a new house by this week."

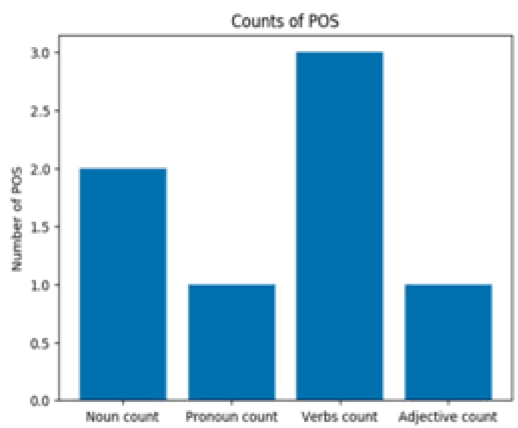

Fig. 5. Graph representation of the text tokenization

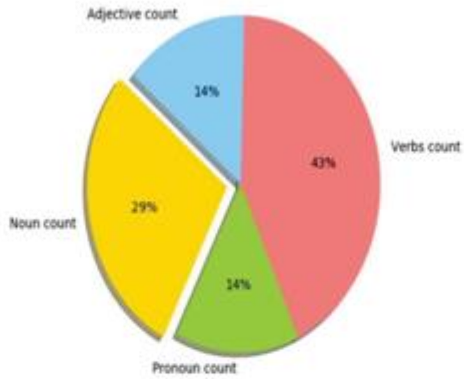

Fig. 6. Pie chart representation of the text tokenization

Figure 7 and Figure 8 give the output for another example sentence: "He is living his life."

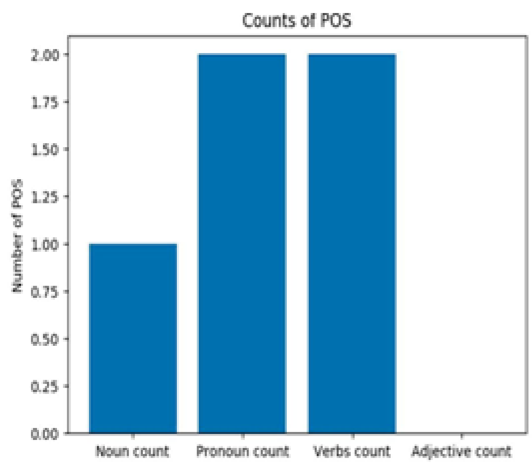

Fig. 7. Graph representation of the text tokenization

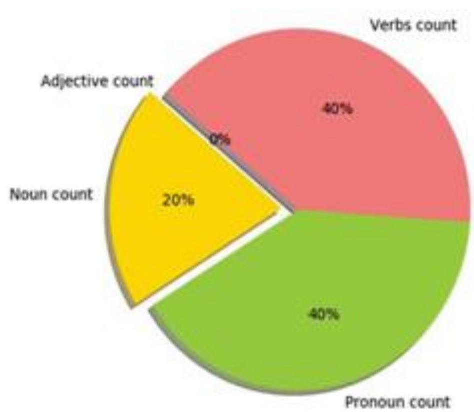

Fig. 8. Pie chart representation of the text tokenization

Figure 9 and Figure 10 give the output for another example sentence: "She knows David, David knows her."

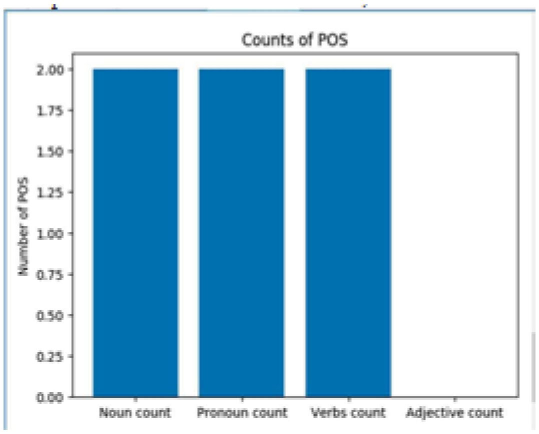

Fig. 9. Pie chart representation of the text tokenization

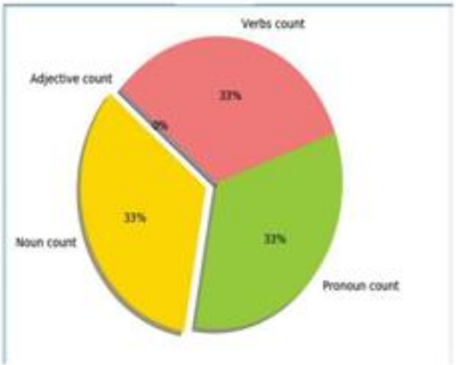

Fig. 10. Pie chart representation of the text tokenization

The animated avatar model has a lot of signs to test. Testing of the avatar demonstration depends on the sign database created for the prototype. The working avatar is given in the Figure 11 shows rendering for the sample word 'Scotland'

.

Since there are various text classification algorithms available, a comparison was needed to choose the most apt one. A comparative study was done on different machine learning algorithms to get better accuracy results for English language identification. Three different classifier used for English language identification are Naive Bayes, SVM and KNN.

Table II. comparison of accuracy for 1-gram

Algorithm	Accuracy	TP	FP	TN	FN
NB	80.5%	800	7	0	2393
SVM	76.56%	763	6	37	2394
KNN	77.18%	795	10	5	2390

Table III. comparison of accuracy for 2-gram

Algorithm	Accuracy	TP	FP	TN	FN
NB	84.40%	800	2	0	2398
SVM	81.28%	800	1	0	2399
KNN	77.62%	797	1	3	2399

Table IV. comparison of accuracy for 3-gram

Algorithm	Accuracy	TP	FP	TN	FN
NB	87.28%	800	2	0	2398
SVM	79.15%	799	0	1	2400
KNN	49.78%	322	0	478	2400

Table V. comparison of accuracy for 4-gram

Algorithm	Accuracy	TP	FP	TN	FN
NB	88.5%	800	1	0	2399
SVM	68.96%	632	0	168	2400
KNN	27.96%	26	0	774	2400

Table VI. comparison of accuracy for 5-gram

Algorithm	Accuracy	TP	FP	TN	FN
NB	88.25%	800	3	0	2397
SVM	48.03%	296	0	504	2400
KNN	25.31%	1	0	799	2400

Table VII. comparison of accuracy for 1to 5-gram

Algorithm	Accuracy	TP	FP	TN	FN
NB	87.18%	800	2	0	2398
SVM	78.03%	799	1	1	2399
KNN	78.65%	799	1	1	2399

IX. DISCUSSIONS

Accuracy is calculated by applying these three algorithms on the dataset one by one. A model is trained on the sentences present in the dataset. The sentences are divided into n-gram where n is from 1 to 5. Accuracy of all three algorithms are compared. Naïve Bayes is having highest accuracy among all the three classifiers. SVM is having second priority among the three classifier and KNN is having third priority among the three classifier.

Precision and recall are also calculated for all three classifier ML algorithms. Precision is also known as positive predictive value and it is defined as [Precision = true positives/ (true positives + false positives)]. Recall is also known as sensitivity and it is defined as [Recall = true positives/ (true positives + false negative)]. Precision and recall are calculated with four parameter i.e. TP (True Positive), FP (False Positive), TN (True Negative) and FN (False Negative). All these four attributes are calculated for each n-gram models.

Table II to Table VII gives accuracy and prediction outcome of 1-gram to 5-gram models. The prediction outcome represents following results -

- True Positive (TP): Actual value is English Language and prediction is English Language.
- False Positive (FP): Actual value is not English Language but prediction is English Language.
- True Negative (TN): Actual value is English Language but prediction fails.
- False Negative (FN): Actual values is not English Language and prediction is not English Language. For Naive Bayes classifier, the True Positive value is 800 in all the n-gram models. The accuracy of Naive Bayes classifier is based on 3,600 sentences only. It means that Naive Bayes classifier are able to correctly classify the sentences based on languages. The Naive Bayes classifier is chosen for language identification with 4-gram as it has accuracy of 88.5%. It predicted correctly for all the 800 English sentences as an English language. For only single sentence it has predicted as English language which was actually not English language. All the sentences except a single sentence are predicted as not English language which was also not English language. Finally, the Naive Bayes classifier is chosen for language identification.

X. CONCLUSION

This paper describes the idea and method to implement a portable system to translate English language to American Sign Language in order to handle the communication between a normal person and deaf person. The difficulty is to learn sign language for a common person. This paper achieved some limitations of the previous papers. This paper too has its own limitations in terms of language complexity and language support. The main focus of the paper to cover as many complex sentences as possible. The proposed architecture defines the steps to develop the system. The system is developed on web platform. Speech to text part is done using API, which perform exact processing of user's speech to gives text as output. The system is implemented using various machine learning algorithms for language detection and language classification. Natural Language Processing has been used for processing the English text to convert it into the American Sign Language specific text format by tokenizing and stemming the sentences to get in the correct format for ASL input. The 3d Avatar takes the input which is in ASL format and demonstrates the signs of each respective word present in the text. The sign rendering is the final output of this paper which can be understood by deaf people.

REFERENCES

1. *Maria Papadogiorgaki, Nikos Grammalidis, Dimitrios Tzovaras and Michael G. Strintzis,"Text-to-sign language synthesis tool," 2002.*
2. *Mikhail G. Grif and Julia S. Manueva,"Russian Sign Language Machine Interpreter System Based on the Analyses of Syntax and Semantic Construction*
3. *Stephen Cox, Michael Lincoln and Judy Tryggvason, Melanie Nakisa, Mark Wells, Marcus Tutt and Sanja Abbott, "TESSA, a system to aid communication with deaf people," 2004.*
4. *Mateen Ahmed, Mujtaba Idrees, Zain ul Abideen, Rafia Mumtaz and Sana Khalique," Deaf Talk Using 3D Animated Sign Language", 2016.*
5. *Jing Wang, Yanfeng Sun, Lichun Wang, "Chinese Sign Language Animation System on Mobile Devices," 2010.*
6. *Ruviansh J Raghavan, Kiran A Prasad, Rahul Muraleedharan,Geetha.M, "Animation System for Indian Sign Language Communication using LOTS Notation," 2013.*
7. *Rupinder Kaur and Parteek Kumar, "HamNoSys Generation System for Sign Language," 2014.*
8. *Syed Faraz Ali, Gouri Sankar Mishra and Ashok Kumar Sahoo "Domain Bounded English to Indian Sign Language Translation Model," 2013.*

9. *Eva Safar and Ian Marshall "The Architecture of an English-Text-to-Sign-Languages Translation System," 2001.*
10. *Veronica Tapia, Raul Reinoso, Edison Carrillo, Dennis Cadas and Leonardo Criollo "Translation System of Voice and Text to Language of Signs," 2017.*
11. *Malu S Nair, Nimitha and A P Sumam Mary Idicula "Conversion of Malayalam text to Indian sign language using synthetic animation," 2016.*
12. *Ian Marshall and E´ va Safar "A Prototype Text to British Sign Language (BSL) Translation System," 2003.*
13. *Achraf Othman and Mohamed Jemni, "Statistical Sign Language Machine Translation: from English written text to American Sign Language Gloss," 2011.*
14. *M.Jemni and O.Elghoul, "An Avatar Based Approach for automatic Interpretation of Text," 2007.*
15. *Wen Hua, Zhongyuan Wang, Haixun Wang, "Understand Short Texts by Harvesting and Analyzing Semantic Knowledge," 2016.*
16. *Mohammed Abdul Wajeed, T.Adilakshmi, "Semi-Supervised text classification using enhanced KNN algorithm," 2011.*
17. *Bo Tang, Halibo He, Paul M. Baggenstoss, Steven Kay, "A Bayesian Classification Approach using Class-Specific Features for Text Categorization," 2016.*
18. *Jutapuck Pugsee, Martha W. Evens & Wanchai Rivepiboon, "Analysis of thai sentences with a serial verb using a semantic lexicon," 2006.*
19. *Indrajit Mukherjee, Mohammad AL-Fayoumi, P.K. Mahanti, Ritesh Jha,"Content analysis based on text mining using genetic algorithm," 2010.*
20. *M. Krendzelak and F. Jakab,"Text Categorization with Machine Learning and Hierarchical Structures," 2015.*
21. *Vishwanath Bijalwan, Vinay Kumar, Pinki Kumari and Jordan Pascual, "KNN based Machine Learning Approach for Text and Document Mining," 2014.*
22. *Jan Bungeroth, Hermann Ney, "Statistical Sign Language Translation," 2004.*
23. *Pradnya Kumbhar, Manisha Mali and Dr. Mohammad Atique, "A Genetic-Fuzzy Approach for Automatic Text Categorization," 2017.*
24. *Michael Grif and Yuliya Manueva," Semantic Analyses of Text to Translate to Russian Sign Language," 2016.*
25. *Ashvinee, Narendra Bhusakhare, Jitesh Singh, Mausumi Goswami, "An architecture of machine translation from speech to sign language based on machine learning", 2018 International Conference on Economics, Business Management and Business Analytic.*