Indian Journal of Science and Technology, *Vol 9(35), DOI: 10.17485/ijst/2016/v9i35/98735, September 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

On Authorship Attribution of Telugu Text

S. Nagaprasad[1*], N. Krishnaveni[2], J. K. R. Sastry[3] and A. Vinayababu[4]

[1]Department of Computer Science, S.R.R.G.A.S.C. Karimnagar, Telangana, India; nagkanna80@gmail.com
[2]Government Degree College for Women, Karimnagar - 505001, Telangana, India
[3]Department of ECM, K L University, Vijayawada - 522502, Andhra Pradesh, India; drsastry@kluniversity.in
[4]Department of Computer Science and Engineering, JNTU, Hyderabad - 500085, Telangana, India; avb1222@gmail.com

Abstract

Background/Objectives: Authorship Attribution is one of the text classification methods. It is useful to find out the author with a given set of text based on author writing style. **Methods/Statistical Analysis:** Various methods that include Decision Tree, K-Nearest Neighbor, Naive Bayes and Support Vector Machine) have been used to find text patterns that exist within a text based database. The classification of the text patterns is deterministic whereas the authorship attribution to the text is un-deterministic. This paper presents a method that recognize a text pattern by using the authorship features using different phases of processing which include prior processing, extracting the features, feature selection, classifying the features and then finally leading to finding the author. **Findings:** The task of Authorship Attribution can be imposed to a range of exercises such as Scientific Analysis, Stealing Recognition and Authorship Recognition. Exploration in the part of Authorship Attribution is in view for more than 100 centuries, but the completed consequences were unacceptable. A range of provocations have been referred which include information collections, tokenizing of the content, applying Natural Language Tools, suitability of categorization methods and reorganization of a range of appearance which can discriminate one writer from the other writers. From the prevailing analysis, it can be concluded that the pronounced accruement are individual circumstance scene of situations, since it may not be useful to other *consequences* of Authorship Attribution associations. From the acquired inputs, it is recognized that the word "unigram" constituent acquired the finest record when assessed with all additional appearances for all classifiers. From among different classifiers, Support Vector Machine realized the best result when evaluated in conjunction with different classifiers such as Decision Tree, K-Nearest Neighbor and Naive Bayes classifiers. **Application/Improvements:** This authorship attribution method is used to find out authorship of vernacular language which in this case is TELUGU.

Keywords: Natural Language Processing, Text Classification

1. Introduction

The number of internet users are has been in raid increase. It has become easier for accessing information and producing the related statistics. The accessibility of the papers through internet is increasing. The papers with no authorship or provisioned authorship are also being figured on the Internet. The authorship attribution is being attacked due to availability and accessibility of the information on the Internet.

Authors follow different methods for writing the papers. Authors may write the papers by hand or uses active or passive voice while expressing an idea significantly. The idea may be celebrated or comprehended. Idea may be expressed in more number of sentences or in terms of little number of sentences.

Sometimes an author may express an idea in a more complicated way or some other author might express the same idea in a clearer way. Some of the authors find it easy to handle the papers by hand than using extensive tools available in the market. Hand written papers clearly reveal the identity, thoughts and views of the authors. However these documents are seldom placed on the internet making them in accessible. The writing of the papers by hand

**Author for correspondence*

is clearly dependent on academic experience, knowledge and the actual life of the author to which the author is subjected to.

One of the mathematical evaluation systems that handle the disputes that arise of using different writing approaches is called as Stylometry. Actual author who wrote a paper can be found by using stylometry. Stylometry can also be used to find the actual author of the paper when no author could be found from the documentation of the paper. The authorship of a document can also be found based on the exercises, problems cited etc. Authorship attribution is required for identifying the actual authors who wrote documents especially when authorship information is not included into the document.

Author ship attribution requires text mining especially by classifying the text. Stylometry is being used for authorship attribution. One main aim of the method is to select data related to the authors from the text. Authors don't understand the the methods like lexical analysis as it is complicated to understand and it is not easy to understand considering authors with different backgrounds. Prior knowledge of classification must be known in order to carry lexical analysis of the text.

Analysis of convergence can be used to recognize the author of a Telugu text who identity is not known. One can use appearance of the text in terms of Semantics, Lexemic and operational appearance and the associations that exist among those visualizations. Several machine learning techniques which include K-Nearest Neighbor, Naïve Bayes, support vector machines and decision tree can be used to determine the information related to authorship.

More of the information has to be used and maintained when Lexemic methods are to be employed for authorship attribution through machine learning. The authorship of Text written in European and Chinese language can be found better using the natural language processing techniques. No mathematical methods as such are recognized as on today that can be comfortable employed for determining the authorship of the text written in Indian languages especially written in Telugu.

In India Telugu is used in a southern state of India. The language as such has complex Agglutinative Morphology. Novel texts can be generated by adding affixes and adjuncts. Machine learning is required for automatically determining the author attribution. Mathematical methods as such are not quite suitable for assessing the attributes related to authors. Significant analysis is carried to be able to identify the authorship attribution as Telugu is quite Agglutinative Morphological language. So far no work has been done to in providing a platform that helps in determining the authorship attribution for text written Telugu language.

2. Related Works

Authorship Attribution can also be viewed in two classifications. Such as Traditional Authorship Attribution (TAA) and Modern Authorship Attribution (MAA). Traditional Authorship Attribution views each Internal and External verifies in direction to recognize the writer of an absolute document. Internal view process gathering the authorship verifications of a specific author from the given unidentified document itself.

Modern Authorship Attribution is also known as Non-Traditional Authorship Attribution. This method is used for Machine Learning techniques to recognize author's handwriting approach. Author's handwriting approach can be considered by getting at the fact of arrivals which are normally extracted from the papers. Handwriting method of an author can be featured in two ways. They are conscious impressions on paper and unconscious impressions on the paper. The conscious handwriting approach can be measured by authors where as the unconscious handwriting approach is by itself of author's awareness.

The initial analysis of Text Categorization techniques commenced from repeated paper arranging for Information Retrieval systems. Each paper content is described by assigning single or more than single key words or key phrases. The measured thesauruses subsist of arthematic hierarchical thesaurus. A controlled dictionary associated to automatic Meta information and indexing. Several Text Classifiers absolutely matured for text indication is discussed in collected work[2,3].

K-Nearest Neighbors algorithm and Nearest Neighbors algorithm are the simplest algorithms which can be used to store innovative and accessible contexts based on the frequency of association, In the 1970s the K-Nearest Neighbors algorithm is used for recognizing the patterns existing in the text through use of statistical estimation methods.

The text patterns are hierarchically organized and the pattern is assigned to an existing class based on the most of the votes assigned by neighboring text patterns. K-NN calculated for each of the activity separately if

K = 1. The pattern is assigned to the class related to its Nearest Neighbor (NN). This method is based on the premises that K articles which are in the neighborhood of the known text pattern. K-NN has to be applied in the beginning for categorisation[4].

Decision Tree (DT) can also be used for undertaking the classification. In this method, the inner connections are labelled using weights of the branches passing the connections. The Leaf nodes of the tree are the classes itself. The text is run through the tree until such that the leaf node is arrived at. Text data cannot be completely be fitted into decision trees constructed in the memory. Range operation are thus required ID3 algorithm is most suitable when range operations are required[5]. C4.5[6] and C5[7] are the other algorithms that can be used for range based decision tree classification.

Different type of document classifiers has been presented which are either reference line classifiers[8], or as associates different types of classifiers[9].

Text classification can also be undertaken by using vector machine algorithm[10]. Vector machines are created on sample data set. SVMs are constructed based on the convention of risk minimization which is based on the computational theory of erudition. VMs have been proved to be efficient for undertaking the training and classifying the text. The words available in the text are sued in the context of its direction and position in the text.

The Vector Space Model(VSM) is an easy mode to represent papers based on the texts in the paper. Paper is converted within a direction in the method of word period, d=(w1,…..,w|T|). Position |T| is the place of term range.

In VSM, word period contains about all the conditions in all papers in the entire collections which ever indications to the trial of measurement, which methods that the internet cost condition improve exponentially with the measurement of the difficulty. Wiener in[11] used Latent Semantic Indexing for local Latent Semantic Indexing whichever building each and every classification-detailed Latent Semantic Indexing interpretations, and for global Latent Semantic Indexing, building a soul Latent Semantic Indexing demonstration for the total classification group. The input directed that the local Latent Semantic Indexing method performs more than the global Latent Semantic Indexing and every method perform superior than simple word collection contingent on element collection scores, for example $\chi 2$, data achieve and so on. Such outputs can also be creating in the current study[12]. For further Text Categorization works that have used Latent Semantic Indexing techniques, see[13-16].

Adoptability of Unicode conventionalized uncompensated for this empirical development in Asian text, retrieving ASCII (a popular coding scheme for Latin Script). Involvement in details of data text is found with a development measure of twenty six percent with indication to European languages. Indian languages, particularly south Indian languages are rich in morphology. Development directions occur in text setups. In Telugu, texts are for all time together with alternates. Rare texts as article components are covered in[17,18].

The aims for Text Categorization in Telugu described with surveys for competent paper demonstration[19] and Standard Machine Training methods[20,21]. Consonant n-grams are also preferred for Indian communication classification in[22] with bi-grams and tri grams on seven Indian communications specifically Bangle, Hindi, Kannada, Kashmiri, Malayalam, Telugu and Urdu. Experiments are effective in recognition all the communications exclude Hindi and Urdu, in view of every one bonds and for all expressions. Communication classification is adventured on Sanskrit word accepting n-gram technique in[23]. Residual shows the evident that although collective text methods are cheaper but n- gram method is extra given to vocabulary faults this time in text.

Consonant placed n-gram model is preferred for Bangla Text Categorization in[24]. N-grams of distance01, 02, 03 and 04 are covered as arranging appearance. N-Grams of distance02 and 03 are described to be functional for classification. Tri-grams are proposed as reasonable spare instead of elements for Telugu articles arranging. Reasonable bi-gram and reasonable tri-grams produced by comparing with collection of vocabularies are expended to organize Telugu Text papers. Consumptive bi-grams and tri-grams are removed with assessment to these vocabularies and since by compressing the measurement. Bi-grams alone are described to offer better competence. The confines of this pursue living the effective reduce of these vocabularies with trigrams and the process being incompatible for aspect n-gram graded papers with n measurement better than three. Reasonable bi-grams with scale assorted from five hundred to eight thousand in scale of five hundred and Tri-grams with a size of five hundred are consumed to basis papers in. Appearance is described to be expandingyetsynopsis size of three thousand and sub sequential endure constant. An assessment is built in the middle of to extend over bi- grams and

tri-grams for Telugu Text categorization in. Word-based pattern is used as a base line for assessment. To spread in view of bi-grams are recovered to allow successful appearance assessed to others.

Indian languages like many other groups of words in the domain have degree subject matter on the internet when evaluated with English. To type the dealings inferior, these group of words are normally stylistically more cultivated and are above decline group of words when domain to group of words like English, giving to the data approach a unruly. The unruly fronting the analysis society in the facing west countries workings on English and alternative group of words were proving to allocate with the data excess. In view of, Indian group of words questions, the unruly is of accessing applied documents are moderately some when assessed with documents in English group of words.

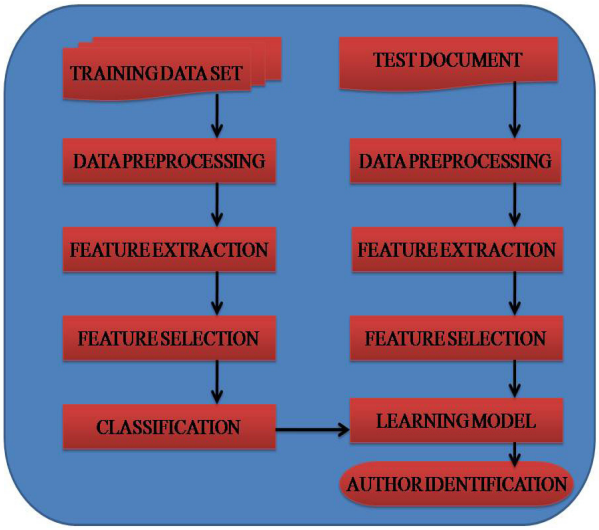

Figure 1. Authorship Attribution Text Model.

3. Problem Statements

Different types of classifiers such as KNN, NB, DT and SVM can be used for finding out the effect of machine learning approaches to determine Authorship attribution. This can be tested using a sample text.

Verbal data is required to understand and develop authorship attribution assignment to the Telugu text. Much number of subject, publications and many numbers of authors has to be considered for determining Authorship Attribution connected to Telugu text.

It is necessary to find Syntactic, Lexical and Structural features and the interrelationships between then on the Authorship Attribution to Telugu Text

4. Proposed Model

The phases to be used for processing the text to recognize a pattern are shown in Figure 1. The phases used for extracting the pattern includes pre-processing, Feature extraction, feature selection, classification and then at the end the author selection

The entire sample text is divided into two sub-samples. One sample is related to Training and the other is related to testing. The training data has to be used for identifying major features. The major features are tested against the test data and the test situations were authored. In the next level a classification model is retrieved

Every author identified in the process is assigned with a number and this number is attached with frequency of Training and Testing instances and extraction. A machine learner has to be trained using the labeled classified data. The machine learner then used to determine authorship attribution.

Pro-processing helps in finding Authorship attribution. The sample text as it is not quite understandable or structured for generating the various kinds of precedents existing in the Text. Therefore there is a need to do pre-processing to determine the precedents. The sample text must be transformed into syntactic format. Further the formatted text must be converted to vector space as majority of the machine tools are quite based on vector based representation of attribute values. In the pre-processing stage, various kinds of processing are undertaken that include stemming, Tokenization and removal of stop words.

After pre-processing is done, the next thing to do is Tokenization. Tokenization is process of breaking entire file in to identifiable words called tokens. The tokens are words that have meaning. At this step the superfluous characters such as quotes, colons, punctuation, exclamation, hyphens, bullets, parenthesis etc.

In the next step a stop list is made[25,26] which are common features such as pronouns, adjectives, conjunctions, prepositions and verbs which generally do not affect the development of classification. If the token is symbol or a number, the same are also eliminated. The list of stop words can be used to identify tags of speech parts using analyzer called TMA (Telugu Morphological Analyzer.

The verbs contained in the Stop list are standardised by using a process called stemming. Alternative forms

of words are identified and the same are reduced into a single form by using the stemming process. The Stemming methods and machine learning methods are consolidated leads to a new technology all together. The Group of wards that are spelt in Indian languages[27] can be processed using constrained-based parsers, clunkers, statistical parsers, and semantic analysers and POS taggers tools. TMA is used to get all stems which are inflected.

The features of Indian languages are evaluated through English processing tools due to the huge processing capability especially for conducting the analysis.

Several metrics have been used which help in selecting the features that has the best fit. Chi-square selection metric is frequently used for measuring the association between Features and the classes arrived through the classification method.

Let **B** represent the feature t and writer set **D** that subsist, **C** represent subsistence of the feature t and the non-subsistence of writer **D** and **E** act as the non-subsistence of both the feature **t** and the writer **D. Let N** act as the total number of training tuples**. The Chi** Square statistics can be depicted as given under

$$Y^2(t\,c)=\frac{N*(B\ -D\ )^2}{(B+D)*(C+E)*(B+C)*(D+E)}$$

F1 and accuracy measures have been used for evaluating lexical features considering a range data set and writer set sizes.

The exactness of the total number of papers having the text that are accurately classified by the writer is computed by using the equations below:

$$\text{Accuracy}=\frac{No.\,of articles that are correctly deputed}{Whole\ number of test\ articles}$$

F1 is calculated as in equation II

$$F_1=\frac{2*exactness*recall}{exactness+recall}$$

Where

$$Exactness=\frac{No.\,of\ articles\ correctly\ writer\ deputed}{No.\,of articles\ writer\ deputed}$$

and

$$recall=\frac{No.\,of\ articles\ correctly\ deputed}{Whole number of test\ articles}$$

5. Authorship Attribution based on number of writers in the training set

120 text articles have been used as a training set and the accuracy of the same has been measured. 5 Articles have been used as test set. This means that 6 writers for a set of 20 articles have been considered in the training set.

The recall, F1 measure and exactness have been calculated considering different feature vectors through application of support vector machine have been calculated and the same are shown in Table 1. Further computations have been made considering 8, 10, and 12 writers have been considered in the training set with 15, 12 and 10 text articles for each of the writer respectively.

From the table it is concluded that the performance in terms of F1 and accuracy decreases as the number of writers in the training set increases. The key feature "unigram" has outperformed compared to all the other features. The next word that has performed well "trigram" compared to all the remaining features. The performance of character level features has been proved to be better in comparison to word level features.

Table 1. F1 and Exactness detail for different number of writers applying SVM

	F1 Assessment				Exactness			
Feature	**Total number of Writers**							
	06	**08**	**10**	**12**	**06**	**08**	**10**	**12**
Character Unigram	0.68	0.64	0.61	0.58	0.74	0.71	0.69	0.65
Character Bigram	0.75	0.71	0.68	0.63	0.78	0.75	0.70	0.64
Character Trigram	**0.82**	0.78	0.75	0.69	**0.85**	0.81	0.74	0.71
Character Tetra gram	0.79	0.75	0.76	0.67	0.82	0.79	0.77	0.74
Word Unigram	**0.84**	0.81	0.77	0.71	**0.87**	0.83	0.79	0.76
Word Bigram	0.76	0.73	0.67	0.64	0.79	0.75	0.72	0.67
Word Trigram	0.68	0.66	0.64	0.61	0.73	0.71	0.68	0.64
Word Tetra gram	0.64	0.62	0.60	0.56	0.69	0.70	0.63	0.60

6. Authorship Attribution based on the amount of data per writer included into the training set

Data from different un-related documents composed of 12 authors have been selected. 300 text documents have been downloaded from the internet out of which 240 documents have been considered as training set

An equal interval of 5 documents each of 20 pages for each of the author has been considered for carrying at empirical assessment. F1 and accuracy have been computed for each of the feature by using support vector machines. The details of the results obtained are placed in Table II.

It has been observed that with increase in number of reports in the training set, substantial increase in growth, F1 and accuracy could be achieved.

The word "Unigrram" has been found to have been performed well. In the next level character "Trigram" and all other features have performed well. Some of the features in between the "Unigram" and "Trigram" showed great performance when assessed considering the word level features.

Table 2. F1 and Exactness ideals for number of Articles features applying SVM

	F1 Assessment				Exactness			
Feature	**Total number of testimony per writers**							
	05	**10**	**15**	**20**	**05**	**10**	**15**	**20**
Character Unigram	0.51	0.58	0.62	0.65	0.58	0.65	0.68	0.71
Character Bigram	0.55	0.63	0.65	0.71	0.61	0.64	0.70	0.75
Character Trigram	0.59	0.69	0.73	**0.75**	0.65	0.71	0.76	**0.81**
Character Tetra gram	0.60	0.67	0.70	0.73	0.63	0.74	0.73	0.78
Word Unigram	0.66	0.71	0.76	**0.85**	0.68	0.76	0.83	**0.89**
Word Bigram	0.58	0.64	0.69	0.76	0.62	0.67	0.69	0.77
Word Trigram	0.52	0.61	0.63	0.69	0.59	0.64	0.66	0.69
Word Tetra gram	0.49	0.56	0.59	0.66	0.58	0.60	0.64	0.67

It also has been observed that the character level features increases starting from hints to syntactic, lexical, and structural. The sparseness has also been found to have reduced when character level "Trigrams" have been considered.

7. Conclusion

Several implementations such as stealing recognition, authorship recognition and scientific analysis can be carried through the task of Authorship Attribution. Author ship exploration has been in use for more than 100 years but the results presented are more than unacceptable

A variety of processing's are to be carried which include information collection, Tokenization, Natural Language processing, categorization, recognition of various kinds of appearances that differentiates one kind of writer from other etc. From the analysis it can be recognized that the authorship attribution is circumstantial and may not be applicable for every kind of circumstance.

It has been observed the word "Unigram" has outperformed compared to any other kind of classifier. SVM has been quite effective in obtaining accurate results by considering variety of classifiers which include Decision tree, K-nearest neighbor, and Nave byes classifier.

8. References

1. Robertson SE, Harding P. Probabilistic automatic indexing by learning from human indexers**.** Journal of Documentation. 1984; 40(4):264–70.
2. Tzeras K, Hartmann S. Automatic indexing based on Bayesian inference networks. In SIGIR '93. Proceedings of the 16th annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY, USA, ACM Press. 1993. p. 22–35.
3. Debole F, Sebastiani F. Supervised term weighting for automated text categorization**.** In SAC '03: Proceedings of the 2003 ACM Symposium on Applied Computing, New York, NY, USA, ACM Press. 2003. P. 784–8.
4. Yang Y, Chute CG. An example–based mapping method for text categorization and retrieval. ACM Trans Inf Syst. 1994; 12(3):252–77.
5. Gvert N, Lalmas M, Fuhr N. **A** Probabilistic description-oriented approach for categorizing web documents**.** In CIKM. 1999; 475–82.
6. Cohen WW, Singer Y. Context-sensitive learning methods for text categorization**.** SIGIR '96: Proceedings of the 19th Annual International ACM SIGIR Conference on Research

and Development in Information Retrieval, New York, NY, USA, ACM Press. 1996. p. 307–15.
7. Li YH, Jain AK. Classification of text documents. The Computer Journal. 1998; 41(8):537–46.
8. Joachims T. Text categorization with support vector machines: learning with many relevant features. In: Dellece CN, Rouveirol C editors. Proceedings of ECML-98, 10th Euroepean Conference on Machine Learning, Number 1398 Chemnitz, DE, Springer Verlag, Heidelberg. 1998. P. 137–42.
9. Schapire RE, Singer Y. Boostexter: **a** boosting-based system for text categorization. Machine Learning**.** 2000; 39(2/3):135–68.
10. Dumais S, Platt J, Heckerman D, Sahami M. Inductive learning algorithms and representations for text categorization**.** Proceedings of the Seventh International Conference on Information and Knowledge Management, ACM Press. 1998. p. 148–55.
11. Wiener ED, Pedersen JO, Weigend AS. A neural network approach to topic spotting**.** Proceedings of SDAIR-95, 4th Annual Symposium on Document Analysis and Information Retrieval, Las Vegas, US. 1995. p. 317–32.
12. Liu T, Chen Z, Zhang B, Ma W-Y, Wu G. Improving text classification using local latent semantic indexing**.** In ICDM. 2004; 162–9.
13. Wan X, Yang J. Multi-Document Summarization using Cluster-based Link Analysis**.** Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'08, Singapore, ACM. 2008. p. 299–306. ISBN 978–1–60558–164–4.
14. Schutze H. Automatic word sense discrimination. Comput Linguist. 1998; 24(1):97–123.
15. Weigend AS, Wiener ED, Pedersen JO. Exploiting hierarchy in text categorization**.** Information Retrieval. 1999; 1(3):193– 216.
16. Carbonell J, Goldstein J. The Use of MMR, Diversity-Based Reranking For Reordering Documents and Producing Summaries**.** Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development.
17. Vishnu Vardhan B, Padmaja Rani B, Kanaka Durga A, Govardhan A, Pratap Reddy L, Vinay Babu A. Impact of dimensionality reduction on the categorization of phonetic based language documents- A case study on Telugu. Geetham Journal of Information and Communication. 2008 Jul-Dec; 1(1):93–8.
18. Murthy KN. Automatic Text categorization Proceedings of Semantic Web Workshop-DRTC-ISI-Banglore. 2003; 1–16.
19. Vishnu Vardhan B, Pratap Reddy L, Padmaja Rani B, Kanaka Durga A, Govardhan A, Vinaya Babu A. Telugu Document Classification using Bayes Probabilistic Model**.** Technology Spectrum, Journal of Jawaharlal Nehru Technological University. 2008 Mar; 2(1):26–30.
20. Raghuveer K, Murthy KN. Text Categorization in Indian languages using Machine learning Approaches Proceedings of IICAI-07. International Conference on Artificial Intelligence. 2007. p. 1864–83
21. Rahal, Perrizo W. An optimized approach for KNN text categorization using P-trees**.** Proceedings of ACM Symposium on Applied Computing. 2004; 613–7.
22. Majumder P, Mitra M, Chaudhuri BB. N – gram: a language independent approach to IR and NLP. 2000. Available from: http://www.unl.fi.upm.es/cosorcio/archivos/publicaciones/goa/ paper15.pdf
23. RCILTS publication Automatic Language Identification of Documents using Devanagri Script**.** 2003. Available from: www.tdil.mit.gov.in/ PaperAbstractJuly03.pdf
24. Mansur M, Uzzaman N, Khan M. Analysis of n – gram based text categorization for Bangla in a news paper corpus 2004, Available from: http://www.naushadzaman.com/textcat_ICCIT06.pdf.
25. Vishnu Vardhan B, Vijaypal Reddy P, Govardhan A. Analysis of BMW model for title word selection on Indic scripts. International Journal of Computer Application (IJCA). 2011 Mar; 18(8):21–5.
26. Vishnu Vardhan B, Vijaypal Reddy P, Govardhan A. Corpus based Extractive summarization for Indic script. International Conference on Asian Language Processing (IALP) IEEE Computer Society (IALP 2011). 154–7.
27. Vijay pal Reddy P, Vishnu Murthy G, Vishnu Vardhan B, Sarangam K. A comparative study on term weighting methods for automated Telugu text categorization with effective classifiers**.** International Journal of Data Mining and Knowledge Management Process (IJDKP). 2013 Nov; 3(6).