# Influence of machine learning techniques on Authorship attribution for Telugu text features

**S.NAGAPRASAD**
Research Scholar
Dept. of CSE
Aacharya Nagarjuna University,
Guntur

**T.RAGHUNADHA REDDY**
Associate Professor
Head, Dept. of CSE
Swarnandhra Institute of Engineering & Technology
Narsapur

**Dr. P.VIJAYAPAL REDDY**
Professor
Dept. of CSE
Gokaraju Rangaraju Institute of Engineering & Technology
Hyderabad

**Dr. A.VINAYA BABU**
Professor
Dept. of CSE
J.N.T.U.College of Engineering
Hyderabad

**Dr. B. Vishnu Vardhan**
Professor & Head
Dept. of IT
J.N.T.U.College of Engineering
Nachupally
Karimnagar

*Abstract:* **Authorship Attribution (AA) deals with identify the author of an anonymous text from known author set. The Authorship Attribution problem is can be viewed as a classification problem. The different steps involved in Authorship Attribution are data preprocessing for vector representation of the text, feature extraction for quantitative representation of the text, feature selection is to reduce the dimensionality feature space, classification algorithms for pattern generation and finally author identification for the given unknown document. There are four categories of features such as lexical, character, syntactic, and semantic features. In this paper character level features and lexical features are considered for feature extraction. Dimensionality of the feature space is reduced using chi-square measure. Classifiers such as Naive Bayes, K-Nearest Neighbour, Support Vector Machine and Decision Tree are used to learn the training document set and to identify the author of a unknown text. The performance of these classifiers in combination with character and lexical features in the context of AA is empirically evaluated on Telugu Texts.**

*Keywords:* **Authorship attribution, Text preprocessing, Stemming, Feature extraction and Machine learning classifier.**

## 1. INTRODUCTION

From the last decade, research on authorship identification is extensively explored. In the beginning authorship attribution was manually conducted by observing the linguistic information embedded in a text corpus as there were no sophisticated natural languages processing tools available. These techniques were based on linguistic markers like term frequency and word similarity [1]. Most of the proposed methods are centered on the detection of authorship for literary texts. Well known study in the field of authorship attribution is the identification of an author of Federalist papers where there was a dispute about twelve of the authors [2]. Another study to identify the authorship of Shakespeare's plays in question [3].

Authorship attribution is a kind of text classification (TC) problem but it is different from categorization. AA is different from text classification because the writing style is also important in AA apart from the text content which is the only factor used in text classification. The features in TC are deterministic where as in AA not deterministic. Based on the size of the data set and number of authors, classifiers and feature sets may behaves differently in AA [4]. Hence these differences make AA task more challenging compared with TC. In text classification the texts are assigned to one or more predefined classes based on the categories where as in AA the texts are assigned to one or more predefined classes based on the author set [5]. Thus the texts in AA are categorized into different classes based on the given set of authors.

In this paper the main focus is on identifying the author of a given text using various steps. The various steps which include are data preprocessing, feature extraction,

  

feature selection, classification and author identification. Data preprocessing contains text tokenizing, stop word removal and text stemming. Feature extraction involves the process of extracting various features such as lexical, syntactic, structural and character level features. In this paper character level and word level features are considered for feature extraction. The dimensionality of the feature space is reduced using chi-square measure. Classification algorithms are used to make generalizations and discover rules from feature set. Classifiers such as Naive Bayes, Support vector machine, K nearest neighour and decision tree are used for pattern generation from the training author sets. Performances of different classifiers in combination with different features are evaluated on Telugu data set.

## 2. RELATED WORK

In [6] and [7], researchers used a variety of statistical methods to identify characteristics which are not variant for a given author but which varies from author to author. In the researcher working [8] on Federalist papers identified a set of functional words which are less frequent could serve as best features to identify authorial style. Yule in [7], [9] and [10] identified that complexity-based features such as sentence length, word length, type and token ratio, automated parsing, POS tagging and POS n-grams are useful features in authorship attribution. Peng in [11] modeled each author by a vector of most frequent n-grams in the text. Fung in [12], used SVM to determine the authors of Federalist papers. Stamatatos in [13] used Multiple Regression classifier in combination with syntactic style features. A character n-gram based method of author attribution has been proposed by Keselj [11]. N- gram models have been successfully applied in speech recognition [14], natural language processing [15] and spell correction [16]. It is also successfully applied in author attribution [17].

Short text authorship attribution is challenging compared with data set having long texts. Short texts requires reliable representation and Machine Learning (ML) algorithm that can handle with limited data. In [18], it is reported that the samples of texts should be long enough therefore the text representation features can sufficiently represent their style. In [19] showed that reducing the length of the training samples has a direct impact on performance. Some studies were shown promising results with short texts of 500 characters [20] or 500 words [21]. In [22] stated that the longer text results in identification.

In this paper the data set contains 300 different texts written by 12 authors. This paper focuses on employing various classifiers such as Naive Bayes classifier (NB), K-Nearest Neighbour (KNN), Support Vector Machine (SVM) and Decision Tree (DT) in the classification task. In contrast to NB, this study also implements SVM which is more suited to extremely big datasets [23]. In this paper, Author attribution of telugu text is focused on extracting various features and applying different classifiers. The rest of the paper is organized as section 3 discusses about the Authorship Attribution model which contains various steps such as data preprocessing, feature extraction, feature selection, classification and author identification. The characteristics of Telugu language are explained in the section 4. The experimental evaluations and discussions on the results are presented in section 5. Section 6 deals with the conclusions drawn from the results and also it contains the possible extensions to the proposed work.

## 3. PROPOSED MODEL

In this model various steps are as shown in Figure 1. The steps are data preprocessing, feature extraction, feature selection, classification and author identification. The data set is separated into training and testing set. In the first phase, features are extracted from the data and on the basis of text features training and test instances were created. In the second phase, an classification model is built from training data, so as to be tested on unknown test data. The training and test instances are numerical feature vectors which represent term frequency of every selected feature, followed by the author number. Labeled training data are used to train a machine learner, as it allows the evaluation of classification. The task of Authorship Attribution is conducted as multi class Authorship Attribution.

  

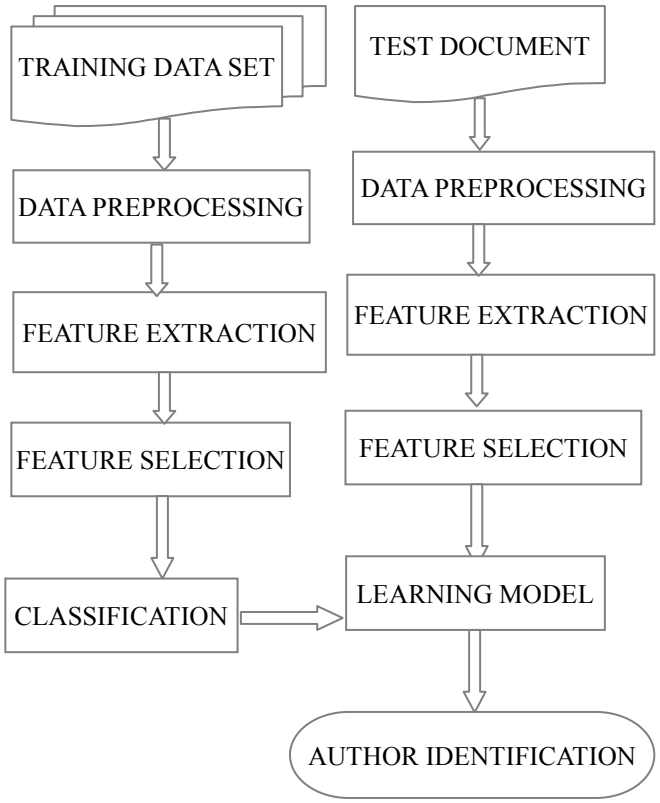


**Figure 1: TEXT AUTHORSHIP ATTRIBUTION MODEL**

### 3.1 Data pre-processing

Data pre-processing is a very important step in authorship attribution. Text documents in their original form are not suitable for meaning patterns generation. They must be converted into a suitable input format. It is then converted into a vector space since most of the learning algorithms use the attribute, value representation. This step is important for the next stages. Data preprocessing involves tokenization, stopword removal and stemming.

#### 3.1.1. Tokenization

Tokenization is the process of chopping a document into small units called tokens which usually results in a set of atomic words having a useful semantic meaning. This phase outputs the article as a set of words by removing the unnecessary symbols like semicolons, colons, exclamation marks, hyphens, bullets, parenthesis, numbers etc.

#### 3.1.2. Stopword removal

As in [27, 28] a stop list is a list of commonly repeated features which appear in every text document. The common features such as pronouns, conjunctions and prepositions need to be removed because they do not have effect on the classification process. For the same reason, if the feature is a special character or a number then that feature should be removed. Stop word list is identified using Parts of speech (POS) tagging done Telugu morphological Analyser (TMA).

#### 3.1.3. Stemming

Stemming is the process of removing affixes (prefixes and suffixes) from features as in [29]. This process is used to reduce the number of features in the feature space and improve the performance of the classifier when the different forms of features are stemmed into a single feature. By using the tool Telugu morphological analyser (TMA) stem forms of the inflected words are identified.

### 3.2 Feature Extraction

**N-gram model:**

There are four main kinds of features that contains authorial impressions for authorship. They are lexical, character, syntactic, and semantic features. In this paper, empirical evaluations are carried using lexical and character features as they are more reliable than semantic features. The different character level features considered in this paper are character unigram, bigram, trigram and tetragrams. Character unigram takes individual characters as tokens where as character bigram considers two consecutive characters, trigram consider three consecutive and tetragram considers four consecutive characters as features. As in [30, 31] character ngrams demonstrated that they are able to handle limited data effectively. Lexical ngrams features are the most widely used kind of features. Whereas researches done in [25] and [19] stated that lexical features are good for small datasets. As defined in [32] word unigram, word bigram, word trigram and word tetragram features are used as lexical features in this experiment. Word unigram takes single word as a feature, word bigram takes two consecutive words, word trigram takes three consecutive words and word tetragram considers four consecutive words as a feature.

### 3.3. Feature Selection

The aim of feature selection methods is to reduce the dimensionality of dataset by removing irrelevant features for the classification task. As in [33, 35], some types of features, such as character and lexical features can considerably increase the dimensionality of the features' set. In such case, feature selection methods can be used to reduce such dimensionality of the representation. Features which are not positively influencing the TC process is

  

removed without affecting the classifier performance, known as Dimensionality reduction (DR).

Feature selection deals with several measures such as document frequency, DIA association factor, chi-square, information gain, mutual information, odds ratio, relevancy score, GSS coefficient. These methods are applied to reduce the size of the full feature set. DR by feature extraction is to create a small set of artificial features from original feature set, which can be done using Term clustering and latent semantic indexing.

In Indian languages, the numbers of features are be even higher compared with English text because of richness in morphology. In this paper, we use chi-square metric [24] for feature selection, a most effective feature selection metric in the literature. Chi-square measures the correlation between feature and class. Let A be the times both feature t and author set c exists, B be the times feature t exists, but author set c doesn't exist, C be the times feature t doesn't exist, but author set c exists, D be the times both feature t and author set c doesn't exist, N be the total number of the training samples. Then CHI square statistics can be depicted as:

$$X^2(t,c) = \frac{N*(AD-BC)^2}{(A+C)*(B+D)*(A+B)*(C+D)}$$

### 3.4 Classification

The goal of Machine Learning (ML) is to construct programs that automatically learn from the training dataset. ML algorithms are able to discover rules from training examples [26]. Therer are two types of machine learning algorithms named as eager learning and lazy learning algorithms. The k-Nearest Neighbour algorithm is an example of a Lazy Learning algorithm. All other learning algorithms which are considered in this paper are eager learning algorithms [26].

#### 3.4.1 Naive Bayes classifier

Probabilistic classifiers are based on the Bayes' theorem as in [34], which computes the probability that a document, represented by a vector dj of (binary or weighted) terms, belongs to a category ci. P(dj) is the probability that a randomly picked document has vector dj as its representation and P(ci)is the probability that a randomly picked document belongs to ci:

$$P(c_i / d_j) = \frac{P(c_i)\,P(d_j / c_i)}{P(d_j)}$$

Naive Bayes classifiers are based on this theorem and use the joint probabilities of words and categories to estimate the probabilities of categories given a document. The words mentioned here are interpreted as the features that are selected to identify author of a document. Naive Bayes classifier builds a probabilistic model of each authorship class based on training data of that class. Then it calculates and multiplies the probabilities of all features to give the probability of test text. The highest probability among all authors is most likely an author of that anonymous or test text.

#### 3.4.2 KNN

K-nearest neighbor (KNN) which is also known as Text-to-Text Comparison (TTC), is a statistical approach, which is has been successfully applied to TC problem [13] and showed promising results. Given a test document to be attributed, the algorithm searches for the K nearest neighbors among the pre-classified training documents based on similarity measure and ranks those k- neighbors based on their similarity scores, the categories of the k-nearest neighbors are used to predict the authorship of the test document by using the ranked scores of each as the weight of the candidate authorships, if more than one neighbor belong to the same authorship category then the sum of their scores is used as the weight of that category, the authorship category with the highest score is assigned to the test document.

#### 3.4.4 Support Vector Machine

Support Vector Machines try to find the surface that best separates the positive from the negative training examples. On the basis of a small set of training examples, called the support vectors, the best decision surface is determined. First the support vector machine is trained using the training value sets prepared according to our authors and their texts. The SVM used here is multi SVM. It is because pools of authors are used. So more than one class is strictly required. After training, the trained machine is used to identify the author for a new input text data.

#### 3.4.5 Decision tree classifier

Decision Tree (DT) Classifiers are easily interpretable learning algorithms because they consist of

  

nested if-then rules. Internal nodes are labelled by terms, branches departing from them are labelled by tests on the weight that the term has in the test document, and leafs are labelled by categories. Every node represents a test of some attribute of the instance. leaf nodes which represent the categories. The goal of a DT learning algorithm is to find the tests that best separate the classification problem into the different categories. The one that gives the best separation of examples should be the root node or test. If not all the training examples have the same label new tests should be designed until all the training examples are separated in a number of categories Applied to an Authorship Attribution task, discriminatory features are represented by the DT nodes or tests. The values on the branches can be thresholds.

### 3.5 Author identification

In this step, for a given test document the name of the author will be returned. For this purpose four steps are performed as shown in figure 1. These steps are same steps that are performed on the training data set. Data preprocessing is performed which involves tokenizing, stopword removal and stemming of the input test document, feature extraction is performed after that reduce the dimensionality of the feature set then input the classifier with reduced feature set of the test document.

## 4. Telugu Language Characteristics

There are more than 150 different languages spoken in India today. Many of the languages have not yet been studied in any great detail in terms of Authorship attribution. According to the author knowledge there is no study on Telugu text in terms of Authorship Attribution. Indian languages are characterized by a rich system of inflectional morphology and a productive system of derivation [30]. This means that the number of surface words will be very large and so will be the raw feature space, leading to data sparsity. Dravidian morphology is in particular more complex. Dravidian languages such as Telugu and Kannada are morphologically among the most complex languages in the world, comparable only to languages like Finnish and Turkish. The main reason for richness in morphology of Telugu (and other Dravidian languages) is, a significant part of grammar that is handled by syntax in English (and other similar languages) is handled within morphology. Phrases including several words in English would be mapped on to a single word iwe may n Telugu. Hence there is a necessity to study the influence of features and different classification approaches on Indian context.

## 5. Results and Discussion

### 5.1. Dataset Description

For authorship identification system, the dataset is collected from Telugu news papers. The topics are ranges from editorials, business and sports. The dataset contains 300 news articles written by 12 authors. The average number of words is 547 per document. In our experiments, we have separated our dataset into two groups: testing and training data. The training set contains 20 different texts for each one of 12 different authors. On the other hand, for the test set there are 5 different texts for each one of 12 different authors. The training data set is used to create data patterns of each author and was treated as data with a known author. We extracted the same profile from the testing data and this data was treated as data with unknown author. Taking the profiles of each testing document one by one, we compared them with each of the training profiles belonging to each author to identify author of unknown test document.

### 5.2. Evaluation measure

In order to compare the results of all possible features with classifiers, we computed the precision, recall and F1 measure in [29]. Precision is the proportion of examples labelled positive by the system that were truly positive, and recall is the proportion of truly positive examples that were labelled positive by the system. Where F1 is computed based on the following equation:

$$F_1 = \frac{2 * Recall * Precision}{Recall + Precision}$$ Where,

$$Precision = \frac{X}{X+Y}$$

$$Recall = \frac{X}{X+Z}$$

Where X is documents assigned and current, Y is documents assigned but not current and Z is documents not assigned but correct. The precision, recall and F1 values obtained for different classifiers in combination with different character level features and word level features are presented in the below tables.

  

| S.No | Feature | $F_1$ value | | | |
|---|---|---|---|---|---|
| | | NB | KNN | SVM | DT |
| 1 | Character Uni-gram | 0.57 | 0.61 | 0.65 | 0.54 |
| 2 | Character Bi-gram | 0.62 | 0.67 | 0.74 | 0.57 |
| 3 | Character Tri-gram | 0.74 | 0.76 | **0.81** | 0.68 |
| 4 | Character Tetra-gram | 0.68 | 0.71 | 0.78 | 0.65 |
| 5 | Word uni-gram | 0.76 | 0.81 | **0.83** | 0.70 |
| 6 | Word Bi-gram | 0.68 | 0.72 | 0.75 | 0.62 |
| 7 | Word Tri-gram | 0.58 | 0.64 | 0.67 | 0.54 |
| 8 | Word Tetra-gram | 0.54 | 0.61 | 0.63 | 0.49 |

**Table 1: The $F_1$ measure obtained for different features by applying the NB, KNN, SVM and DT classifiers**

Precision, recall and F1 values are calculated for various character level features and lexical features on different classifiers. From the values obtained it can be concluded that character trigram feature is performing well out of all character level features. From the view of classifiers, SVM performance is good compared with all other classifiers. Word unigram is identified as a best feature for authorship attribution when compared with all other character level features and word level features. Compared with word level features, on an average character level features can be considered as good indicators for authorship identification.

**6. Conclusions & Future Scope**

In this work, an Authorship Attribution task has been experimented on an Telugu dataset In this work different texts of various authors are selected. These texts are preprocessed. Several features have been tested for Telugu dataset. In our experiments tried to find the best authorship attribution conditions using character level features and lexical features. Emperical evaluations are carried on the text set using different machine learning classifiers such as naive bayes, k-nearest neighbour, support vector machine and decision tree classifiers in combination with different character and word level features. From the results it can be concluded that the character level features are better than the word based features. The word unigram feature gave the best F1 score obtained as 0.83 for authorship classification. The SVM classifier shows good performance in this experiment of Authorship Attribution compared with all other classifiers. This type of attempt to find the author ship attribution is of a new initiation applied on telugu text.

As a part of future work we may experiments with other machine learning algorithms with more number of authors and with small dataset of texts. It is also possible to experiment with other types of features such as syntactic and semantic and also with the combination of different types of features.

**References**

[1] B. Allison and L. Guthrie, " Authorship attribution of e-mail: comparing classifiers over a new corpus of evaluation," in Proceedings of the Sixth International Language Resources and Evaluation (LREC'08), May 28-30, 2008, Marrakech, Morocco.

[2] Klarreich, E. 2003. Bookish math. Science News 164(25)

[3] T. Merriam, "Heterogeneous authorship in early Shakespeare and the problem of Henry V," Literary and Linguistic Computing, vol. 13, no. 1, 1998, pp. 15-28

[4] Bozkurt, D., Baglıoglu, O., & Uyar, E. (2007), “Authorship Attribution: Performance of Various Features and Classification Methods” Computer and Information Sciences.

[5] Zhao, Y. (2007), “Effective authorship attribution in Large Document Collections”, PhD Thesis, School of Computer Science and Information Technology, RMIT University , Melbourne, Victoria, Australia.

[6] Holmes, D.: The Evolution of Stylometry in Humanities Scholarship Literary and Linguistic Computing, (1998) 13, 3, 111-117

[7] Koppel, M., Schler, J.: Exploiting Stylistic Idiosyncraises for Authorship Attribution, IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis, Acapulco, Mexico (2003)

[8] Mosteller, F., Wallace, D.L.: Inference and Disputed Authorship: The Federalist Reading, MA:Addison-Wesley (1964)

[9] Yule, G.U.: On sentence length as a statistical

  

characteristic of style in prose with application to two cases of disputed authorship, Biometrica (1938) 30, 363-390

[10] Stamatatos, E., Fakotakis, N., Kokkinakis, G.: Computer-Based Authorship Attribution without Lexical Measures, Computers and the Humanities (2001) 193-214

[11] Peng, F., Schuurmans, D., Keselj, V., Wang, S.: Language Independent Authorship Attribution using Character Level Language Models, 10th Conference of the European Chapter of the Association for Computational Linguistics, Budapest (2003) 267-274

[12] Fung, G., Mangasarian, O.:The Disputed Federalist Papers: SVM Feature Selection via Concave Minimization, Proceedings of the 2003 Conference of Diversity in Computing, Atlanta, Georgia, USA (2003) 42-46

[13] Stamatatos, E., Fakotakis, N., Kokkinakis, G.: Automatic Authorship Attribution, Nineth Conf. European Chap. Assoc. Computational Linguistics, Bergen, Norway (1999)

[14] F. Jelinek, Statistical Methods for Speech Recognition, MIT Press, Boston, Massachusetts, USA, ISBN: 0-262-10066-5, 1998

[15] Foundations of Statistical Natural Language Processing Christopher D. Manning and Hinrich Sch ¨ utze (Stanford University and Xerox Palo Alto Research Center) Cambridge, MA: The MIT Press, 1999, xxxvii + 680 pp; hardbound, ISBN 0-262-13360-1

[16] Mayes, E., F. Damerau, et al. (1991). "Context Based Spelling Correction." Information Processing and Management 27(5): 517-522.

[17] Abbasi, A. and Chen, H. 2008. Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace.

[18] Stamatatos, E. (2009), "A survey of Modern authorship attribution methods", Journal of the American Society for Information Science and Technology, 538-556.

[19] Luyckx, K. (2010), "Scalability Issues in Authorship Attribution", PhD Thesis, Faculty of Arts and Philosophy, Dutch UPA University.

[20] Sanderson,C. & Guenter, S. (2006), "Short text authorship Attribution via Sequence Kernels, Markov Chains and Author Unmasking", An Investigation. Proceeding of 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP), 482-491.

[21] Koppel, M., Schler, J., & Bonchek-Dokow, E. (2007), "Measuring differentiability: Unmasking pseudonymous Authors", Journal of Machine Learning Research, 8, 1261-1276.

[22] Siham, O. & Halim, S. (2012), "Authorship Attribution of Ancient Texts Written by Ten Arabic Travelers Using a SMO-SVM Classifier", The 2nd International Conference on Communications and Information Technology (ICCIT): Digital Information Management, Hammamey, 44-47.

[23] Elayidom, M. S., Jose, C., Puthussery, A., & Sasi, N. K. (2013), "Text Classification for Authorship Attribution Analysis", Advanced Computing: An International Journal (ACIJ), Vol.4, No.5, 1-9.

[24] YANG Y, PEDERSEN J Q. A comparative study on feature selection in text categorization. Proceedings of the 14th International Conference on Machine Learning (ICML), 1997: 2-3.

[25] Hong, R., Tan, R., & Tsai, F, S. (2010), "Authorship Identification for Online Text", International Conference on Cyberworlds, 155-162.

[26] Mitchell T. (1997) Machine Learning. New York [a.o.]: The McGraw-Hill Companies, Inc.

[27]B.VishnuVardhan,P.Vijaypal Reddy,A.Govardhan"Analysis of BMW model for title word selection on Indic scripts", International Journal of Computer Application (IJCA) Vol 18 Number 8 March 2011 pp 21-25

[28] B.Vishnu Vardhan,P.Vijaypal Reddy, A.Govardhan"Corpus based Extractive summarization for Indic script", International Conference on Asian Language Processing (IALP) IEEE Computer Society (IALP 2011) pp 154-157

[29] P. Vijay pal Reddy, Vishnu Murthy.G, Dr. B. Vishnu Vardhan, K. Sarangam "A comparative study on term weighting methods for automated telugu text categorization with effective classifiers" International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.3, No.6, November 2013

[30] B.Vishnu Vardhan, L. Pratap Reddy, A. Vinay Babu, " A Model for Overlapping Trigram technique for Telugu Script" Journal of Theoritical and Applied Information Technology, Vol. 3, No. 3, Sep. 2007, pp 9-14.

[31] B. Vishnu Vardhan ,B. Padmaja Rani,, A. Kanaka Durga, L. Pratap Reddy, A. Vinay Babu, "Analysis of N-

  

gram model on Telugu Document Classification" Proceedings of 2008 IEEE Congress on Evolutionary Computation (CEC 2008), Hong Kong, 1-6 June, 2008, pp 3198-3202.

[32] B Vishnu Vardhan, Pratap Reddy L, A Vinay Babu – " Document Categorization Using Trigram Technique – A model for Telugu Script ", Proceedings of International Conference on Systemics, Cybernetics and Informatics(ICSCI 2007), Hyderabad, Jan., 2007, Vol. 1, pp 883-887

[33] B. Vishnu Vardhan ,B. Padmaja Rani, A. Kanaka Durga, A.Govardhan, L. Pratap Reddy, A. Vinay Babu, Impact of dimensionality reduction on the categorization of phonetic based language documents- A case study on Telugu" Geetham journal of Information and Communication Volume1 , Issue 1, July-December 2008, pp 93-98

[34] B. Vishnu Vardhan, L. Pratap Reddy, B. Padmaja Rani, A. Kanaka Durga, A. Govardhan, A. Vinaya Babu, "Telugu Document Classification using Bayes Probabilistic Model" Technology Spectrum, Journal of Jawaharlal Nehru Technological University, Vol.II, No. 1, March 2008, pp 26-30.

[35] B. Vishnu vardhan,P.Vijayapal reddy, B Sasidhar, B Harinatha reddy,L. Pratap reddy , A. Govardhan, " Approaches of Dimensionality Reduction for Telugu Document Classification" International Conference on Asian Language Processing (IALP) IEEE Computer Society 2009, pp 259-264.

  