Mark Your Calendar!

2019 IEEE 8th Global Conference on Consumer Electronics

IEEE GCCE 2019

will be held
on October 15th to 18th, 2019
at Osaka, Japan.

More information will be announced at
http://www.ieee-gcce.org/2019/

IEEE GCCE 2018

Abstract Book

2018 IEEE 7th Global Conference on Consumer Electronics

October 9th - 12th, 2018

Nara Royal Hotel, Nara, JAPAN

Sponsored by

IEEE

Advancing Technology for Humanity

2018 IEEE 7th Global Conference on Consumer Electronics (GCCE 2018)

Nara, Japan
9 – 12 October 2018

IEEE

IEEE Catalog Number: CFP18GCB-POD
ISBN: 978-1-5386-6310-3

parallel corpus is prepared and introduces the Myanmar to English Neural Machine Translation system. Nowadays, neural machine translation models became a popular research field and it reaches good results in some languages. In this work, we did the experiment on the word-level model and character-level model based on neural method for Myanmar to English translation. The evaluation results show that neural machine translation models lead to improve the performance of Myanmar to English translation.

☐ Comparison of Naive Bayes and Support Vector Machine Classifiers on Document Classification (pp. 438 - 439)

Zun Moe, Thida San and Mie Mie Khin (Myanmar Institute of Information Technology, Myanmar); Hlaing Tin (Myanmar Institute of Information Technology & MIIT, Myanmar)

The main objective is to classify the field from IT research papers and compare the accuracy in two classifiers. Classification is the form of data analysis that can be used to extract models describing important data class or to predict future data trends. The most important features are selected and data are prepared for learning and classification. Text classification is the process of assigning a document to one or more target categories based on its contents. Training and classification are performed using Naive Bayes and Support Vector Machine (SVM) classifiers. Experimental results show that the methods are favorable in terms of their effectiveness and efficiency. This system classifies text on ten categories such as "Big Data", "Image Processing","Data Mining", "Artificial Intelligent", "Ontology", "Data Base Management System", "Management Information System" and "Software Engineering" and so on. This system calculates the accuracy of testing data using holdout method.

☐ A Study on Indoor Positioning System Based on Attitude Estimation by Sensors and Distance Estimation by SOKUIKI Sensor (pp. 440 - 442)

Shunnosuke Sasaki and Masahiro Fujii (Utsunomiya University, Japan)

In this paper, we propose a new positioning method based on an attitude estimation by the accelerometer, gyroscope and so on, a distance estimation by the SOKUIKI sensor and a reflection surface information. The proposed system estimates the attitude of the terminal by using measurement data from the sensors. At the same time, the SOKUIKI sensor provides an estimate of the distance between the terminal and reflection points. We estimate the location of the terminal so as to minimize squared error between the measurements and a hypothesis based on reflection surface information of surrounding structure.

☐ Design of Compact and Vibration Resistant Receiver for Visible Light Communication Using MEMS Mirror (pp. 443 - 445)

Haruka Iiyama (Tsukuba, Japan); Ryunosuke Kurimoto, Yoshihito Imai, Tadashi Ebihara, Koichi Mizutani and Naoto Wakatsuki (University of Tsukuba, Japan)

We propose a compact and vibration-resistant receiver for visible light communication (VLC) between vehicles, where the dynamic motion of the vehicle becomes a barrier for reliable communication. To address this problem, we designed a VLC receiver with a micro mirror, acceleration sensor, and image sensor. A fast-blinking light is scanned by micro mirror and captured by image sensor efficiently, and the micro mirror is controlled by acceleration sensor to cancel the effect of motion of the vehicle. The performance of the proposed receiver was evaluated in experiments. The obtained results suggest that the proposed receiver is suitable for dynamic environment.

COMPARISON OF NAÏVE BAYES AND SUPPORT VECTOR MACHINE CLASSIFIERS ON DOCUMENT CLASSIFICATION

Zun Hlaing Moe[1]
Faculty of Information Science
[1]Myanmar Institute Of Information Technology, Myanmar
[1]zun_hlaing_moe@miit.edu.mm

Thida San[2]
Faculty of Information Science
[2]Myanmar Institute Of Information Technology, Myanmar
[2]thida_san@miit.edu.mm

Mie Mie Khin[3]
Faculty of Information Science
[3]Myanmar Institute Of Information Technology, Myanmar
[3]mie_mie_khin@miit.edu.mm

Hlaing May Tin[4]
Faculty of Computer System and Technology
[4]Myanmar Institute Of Information Technology, Myanmar
[4]hlaing_may_tin@miit.edu.mm

***Abstract*—The main objective is to classify the field from IT research papers and compare the accuracy in two classifiers. Classification is the form of data analysis that can be used to extract models describing important data class or to predict future data trends. The most important features are selected and data are prepared for learning and classification. Text classification is the process of assigning a document to one or more target categories based on its contents. Training and classification are performed using Naive Bayes and Support Vector Machine (SVM) classifiers. Experimental results show that the methods are favorable in terms of their effectiveness and efficiency. This system classifies text on ten categories such as "Big Data", "Image Processing","Data Mining", "Artificial Intelligent", "Ontology", "Data Base Management System", "Management Information System" and "Software Engineering" and so on. This system calculates the accuracy of testing data using holdout method.**

Keywords— Naïve Bayes, Classification, Text, SVM, Categories, Accuracy

I. INTRODUCTION

Document classification is the work of groping document into categories based on their content. There are many classification methods for documents. This system compared two classification algorithms that are Support Vector Machine and Naïve Bayes.

Classification can be defined as categorizing document into one of a fixed number of predefined classes with a single document belonging to only one class. Colas & Brazdil (2006) [1] sought about old classification algorithms in text categorization. They, also, found systematically the weaknesses and strength of SVM, naive Bayes and KNN algorithms in text categorization and examined how the number of attributes of the feature space effected on the performance. Bilski (2011) [2] described the most important techniques and methodologies used for the text classification. Effectiveness and advantages for contemporary algorithms are compared and their most applications presented. The used text classification algorithms are artificial neural networks, k Nearest Neighbor (kNN) approach, Naïve Bayes classifier, decision trees and rules induction algorithms. Gandhi& Prajapati (2012)[3] described and compared the three algorithms which are k-nearest neighbors classifier, naive Bayes and the Support Vector Machines. They defined the settings of the data which performed in experiments. Document is entered as input in this system, and then the system will do preprocessing steps. Main words are only retrieved. Finally, Main words are matched with all the data required for each fields stored in the database. And then count the important words using Term Frequency (TF) in feature selection. Finally calculate the accuracy by using holdout method.

II. NAÏVE BAYES CLASSIFIER VS. SUPPORT VECTOR MACHINE CLASSIFIER

A. Naive Bayes Classifier

The naïve Bayes classifier is a typical generative classifier, which can be regarded as a special case of Bayesian network classifiers. In general, Bayesian network classifier models first the joint distribution p(x,y) of the measured attributes x and the class labels y factorized in the form p(x|y)p(y), and then learns the parameters of the model through maximization of the likelihood given by p(x|y)p(y). Due to there is a fundamental assumption that the attributes are conditionally independent given a target class, the naïve Bayes classifier in fact learns the parameters of the model through maximization of the likelihood given by $p(y)\Pi j\ p(xj|y)$.

B. Support Vector Machine Classifier

The SVM classifier is a typical discriminative classifier. Different from generative classifier, it mainly focuses on how well they can separate the positives from the negatives, and does not try to understand the basic information of the individual classes. The SVM classifier maps first the instance x in a training set into a high dimensional space via a function Φ, then computes a decision function of the form $f(x) = \langle w, \Phi(x) \rangle + b$ by maximizing the distance between the set of points Φ(x) to the hyperplane or set of hyperplanes parameterized by (w, b) while being consistent on the training set. The SVM classifier builds a single model for all classes and hence it requires simultaneous consideration of all other classes.

III. SYSTEM OVERVIEW

In this section shows all the overview process of the system.

A. Parsing

From the parse, we can find the relation of each words in the sentence to all the others, and typically also its function in the sentence (eg. Subject, object, etc).

B. Stop Words Removal

Removing stop words from the document are very common in information retrieval. Eliminating the stop words from the documents, this will lead the reduction in the dimensionality of feature space

C. Feature Selection

Feature selection is the process of selecting a subset of relevant features. It is important in classifying text. It improves the classification accuracy by eliminating the noise features.

D. Keywords Match

Match Words from the documents and keywords which are already stored in the database. Finally the system will display the name of the field category.

E. Accuracy

The systems calculate the accuracy of the user selected document.

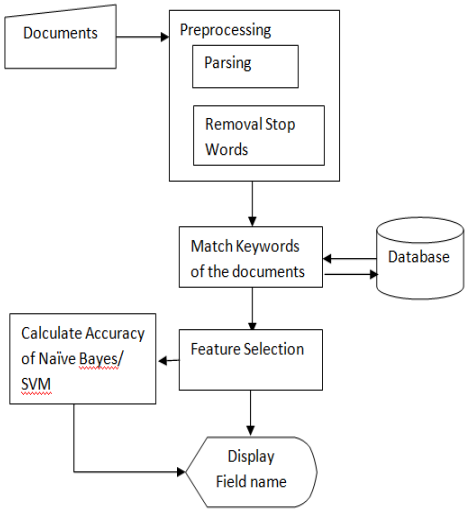

Fig. 1. Process of the system.

IV. EXPERIMENTAL EVALUATION

This section provides the characteristics and descriptions of the data sets used for performing our experiments. In this system, the user can know any field of document and calculate the probability of the words of document. Only words including more and more in a document is to display what kind of field. But if count of words is same, it needs to make accuracy so that the program displays what kind of field.

A. Data

In this system, we compared the performance of Naïve Bayes and Support Vector Machines in the text classification by using 10 main categories dataset. Table I shows the descriptions of ten categories. We collect approximately 1000 documents as the training datasets, which are constituted by extracting 100 documents from each of the ten categories.

TABLE I. CATEGORIE DESCRIPTION

Category Name	***Description***
BD	Big Data Analysis
IM	Image Processing
DM	Data Mining
AI	Artificial Intelligent
DBMS	Database Manement System
SE	Software Engineering
Ot	Ontology
MIS	Management Information System
ICS	Information Communication Systems
DL	Deep Learning

B. Accuracy

Accuracy is a measure of how close the results of the automatic classification match the true categories of the documents. Accuracy is estimated by applying the classifier to the testing dataset classified by domain experts. In this system, we compute the accuracy by using holdout method after preprocessing steps.

$$accuracy = sensitivity\frac{pos}{(pos+neg)} + specificity\frac{neg}{(pos+neg)}$$

In this paper we are comparing results of Support Vector Machine (SVM) and Naïve Bayes techniques. According to our research the SVM is more accurate the Naïve Bayesian Classifier.

V. CONCLUSION AND FUTURE WORK

In this paper, it was described extraction of fields from related IT research papers. It presented the comparison of Naïve Bayes and Support Vector Machine classifiers. It applied these algorithms to classify automatically IT research papers. These classifiers give correct and accurate result. The results demonstrate the validity of our approach. Support Vector Machine is more accurate than Naïve Bayes classifier

For future work, we intend to classify so many other categories such as business, environment, human and society and so on.

REFERENCES

[1] Korde, V., & Mahender, C. N. (2012). Text classification and classifiers: A survey.International Journal of Artificial Intelligence & Applications,3(2), 85.

[2] Pawar, P. Y., & Gawande, S. H. (2012). A comparative study on different types of approaches to text categorization. International Journal of Machine Learning and Computing,2(4), 423.

[3] Bhumika, P. S. S. S., & Nayyar, P. A. (2013). A review paper on algorithms used for text classification. International Journal of Application or Innovation in Engineering & Management, 3(2), 90-99.

[4] J.B.Mohammed, S.A.Mahmud and I.A.Adnan "Arabic Text Classification using K-NN and Naïve Bayes".