ISSN 2356-4407

www.STIKI.ac.id

Bersama memajukan ICT Indonesia

# PROCEEDING

# IC - ITECHS 2014

The 1st International Conference on Information Technology and Security

Malang, November 27, 2014

*Published by:*

**Lembaga Penelitian dan Pengabdian pada Masyarakat**

Sekolah Tinggi Informatika dan Komputer Indonesia

# PROCEEDING

## The 1st International Conference on Information Technology and Security (IC-ITechs) November 27, 2014


**Editors & Reviewers:**

Tri Y. Evelina, SE, MM Daniel
Rudiaman, S.T, M.Kom Jozua
F. Palandi, M.Kom

**Layout Editor:**

Eka Widya Sari



**LEMBAGA PENELITIAN & PENGABDIAN KEPADA MASYARAKAT**
**Sekolah Tinggi Informatika & Komputer Indonesia (STIKI) – Malang**
**Website: itechs.stiki.ac.id E-mail: itechs@stiki.ac.id**

# PROCEEDING

## The 1st International Conference on Information Technology and Security (IC-ITechs) November 27, 2014


## ISSN 2356 - 4407

viii + 276 hlm; 21 X 29,7 cm


### Reviewers & Editors:

Tri Y. Evelina, SE, MM
Daniel Rudiaman, S.T, M.Kom
Jozua F. Palandi, M.Kom

### Layout Editor:

Eka Widya Sari


Published by:
LEMBAGA PENELITIAN & PENGABDIAN KEPADA MASYARAKAT
Sekolah Tinggi Informatika & Komputer Indonesia (STIKI) – Malang
Jl. Raya Tidar 100 Malang 65146, Tel. +62-341 560823, Fax. +62-341 562525
Website: itechs.stiki.ac.id E-mail: itechs@stiki.ac.id

# GREETINGS

## Head of Committee IC-Itechs

For all delegation participants and invited guest, welcome to International Conference on Information Technology and Security (IC-Itechs) 2014 in Malang, Indonesia.

This conference is part of the framework of ICT development and security system that became one of the activities in STIKI and STTAR. this forum resulted in some references on the application of ICT. This activity is related to the movement of ICT development for Indonesia.

IC-Itechs aims to be a forum for communication between researchers, activists, system developers, industrial players and all communications ICT Indonesia and abroad.

The forum is expected to continue to be held continuously and periodically, so we hope this conference give real contribution and direct impact for ICT development.

Finally, we would like to say thanks for all participant and event organizer who involved in the held of the IC-Itechs 2014. We hope all participant and keynote speakers got benefit from this conference.

ISSN 2356 - 4407
The 1st International Conference
on Information Technology and Security

# LIST OF CONTENT

Implementation, Challenges, and Cost Model for Calculating Investment Solutions of Business Process Intelligence ..... 1 – 8
***Arta M. Sundjaja***

Bisecting Divisive Clustering Algorithm Based On Forest Graph ..... 9 – 14
***Achmad Maududie, Wahyu Catur Wibowo***

3D Interaction in Augmented Reality Environment With Reprojection Improvement on Active and Passive Stereo ..... 15 – 23
***Eko Budi Cahyono, Ilyas Nuryasin, Aminudin***

Traditional Exercises as a Practical Solution in Health Problems For Computer Users ..... 24 -29
***Laurentius Noer Andoyo, Jozua Palandi, Zusana Pudyastuti***

Baum-Welch Algorithm Implementation For Knowing Data Characteristics Related Attacks on Web Server Log ..... 25 -36
***Triawan Adi Cahyanto***

Lighting System with Hybrid Energy Supply for Energy Efficiency and Security Feature Of The Building ..... 37 – 44
***Renny Rakhmawati, Safira Nur Hanifah***

Interviewer BOT Design to Help Student Learning English for Job Interview ..... 45 – 50
***M. Junus, M. Sarosa, Martin Fatnuriyah, Mariana Ulfah Hoesny, Zamah Sari***

Design and Development of Sight-Reading Application for Kids ..... 51 -55
***Christina Theodora Loman, Trianggoro Wiradinata***

Pembuatan Sistem E-Commerce Produk Meubel Berbasis Komponen ..... 66 – 74
***Sandy Kosasi***

Crowd sourcing Web Model of Product Review and Rating Based on Consumer Behaviour Model Using Mixed Service-Oriented System Design ..... 75 – 80
***Yuli Adam Prasetyo***

Predict Of Lost Time at Traffic Lights Intersection Road Using Image Processing ..... 81 – 88
***Yoyok Heru Prasetyo Isnomo***

Questions Classification Software Based on Bloom's Cognitive Levels Using Naive Bayes Classifier Method ..... 89 – 96
***M. Fachrurrozi, Lidya Irfiyani Silaban, Novi Yusliani***

A Robust Metahuiristic-Based Feature Selection Approach for Classification ..... 97 – 102
***Alna Musdholifah, Erick***

Building a Spatio-Temporal Ontology for Artifacts Knowledge Management ..... 103 - 110
***Nurul Fajrin Ariyani, Daniel Oranova Siahaan***

Decision Support on Supply Chain Management System using Apriori Data Mining Algorithm ..... 111-117
***Eka Widya Sari, Ahmad Rianto, Siska Diatinari Andarawarih***

Object Recognation Based on Genetic Algorithm With Color Segmentation ..... 118-128
***Evy Poerbaningtyas, Zusana E. Pudyastuti***

Developing Computer-Based Educational Game to Support Cooperative Learning Strategy ..... 129-133
***Eva Handriyantini***

The Use of Smartphone to Process Personal Medical Record by using Geographical Information System Technology ..... 134-142
***Subari, Go Frendi Gunawan***

Implementasi Metode Integer Programming untuk Penjadualan Tenaga Medis Pada Situasi Darurat Berbasis Aplikasi Mobile ..... 143-148
***Ahmad Saikhu, Laili Rochmah***

News Sentiment Analysis Using Naive Bayes and Adaboost ..... 149-158
***Erna Daniati***

Penerapan Sistem Informasi Akutansi pada Toko Panca Jaya Menggunakan *Integrated System* ..... 159-163
***Michael Andrianto T, Rinabi Tanamal, B.Bus, M.Com***

Implementation of Accurate Accounting Information Systems To Mid-Scale Wholesale Company ..... 164-168
***Aloysius A. P. Putra, Adi Suryaputra P.***

Conceptual Methodology for Requirement Engineering based on GORE and BPM ..... 169-174
***Ahmad Nurulfajar, Imam M Shofi***

Pengolahan Data Indeks Kepuasan Masyarakat (IKM) Pada Balai Besar Pengembangan Budidaya Air Tawar (BBPBAT) Sukabumi dengan Metode Weight Average Index (WAI) ..... 175-182
***Iwan Rizal Setiawan, Yanti Nurkhalifah***

Perangkat Lunak Keamanan Informasi pada Mobile Menggunakan Metode Stream dan Generator Cipher ..... 183-189
***Asep Budiman Kusdinar, Mohamad Ridwan***

*Analisys Design Intrusion Prevention System (IPS) Based Suricata* ... 190-193
***Dwi Kuswanto***

Sistem Monitoring dan Pengendalian Kinerja Dosen Pada Proses Perkuliahan Berbasis *Radio Frequency Identification (RFID)* Di Lingkungan Universitas Kanjuruhan Malang ..... 194-205
***Moh.Sulhan***

Multiple And Single Haar Classifier For Face Recognition ..... 206-213
***Go Frendi Gunawan, Subari***

Sistem Penunjang Keputusan Untuk Menentukan Rangking Taraf Hidup Masyarakat Dengan Metode Simple Additive Weighting ..... 214-224
***Anita, Daniel Rudiaman Sijabat***

Optical Character Recognition for Indonesian Electronic Id-Card Image ..... 225-232
***Sugeng Widodo***

Active Noise Cancellation for Underwater Environment using Raspberry PI ..... 233-239
***Nanang syahroni, Widya Andi P., Hariwahjuningrat S, R. Henggar B***

Implementasi Content Based Image Retrieval untuk Menganalisa Kemiripan Bakteri Yoghurt Menggunakan Metode Latent Semantic Indexing ..... 240-245
***Meivi Kartikasari, Chaulina Alfianti Oktavia***

Software Requirements Specification of Database Roads and Bridges in East Java Province Based on Geographic Information System ..... 246-255
***Yoyok Seby Dwanoko***

Functional Model of RFID-Based Students Attendance Management System in Higher Education Institution ..... 256-262
***Koko Wahyu Prasetyo, Setiabudi Sakaria***

*Assessment of Implementation Health Center Management Information System with Technology Acceptance Model (TAM) Method And Spearman Rank Test in Jember Regional Health* ..... 263-267
***Sustin Farlinda***

Relay Node Candidate Selection to Forwarding Emergency Message In Vehicular Ad Hoc Network ..... 268-273
***Johan Ericka***

*Defining Influencing Success Factors In Global Software Development (GSD) Projects* ..... 274-276
***Anna Yuliarti Khodijah, Dr. Andreas Drechsler***

# Questions Classification Software based on Bloom's Cognitive Levels using Naive Bayes Classifier Method

**M. Fachrurrozi, Lidya Irfiyani Silaban,Novi Yusliani**
Faculty of Computer Science, Sriwijaya University
obetsobets@gmail.com , lidyairfiyanisilaban@gmail.com , novi.yusliani@gmail.com

**Abstract**

*Questions Classification is one way to know how the student understanding some lessons. Those questions can be collected and classified based on cognitive Bloom level. Bloom Cognitive Level organized question in groups that represents contents of those questions. Words contained in every question have certain meaning and can be used as base to determine category of question. This study aims to classify amounts of questions based on cognitive Bloom level with Naive Bayes Classifier method. According to Bloom's taxonomy of cognitive level divided into six levels of the Knowledge (C1), Comprehension (C2), Application (C3), Analysis (C4), Synthesis (C5), and Evaluation (C6). In this study, questions classified into 3 classes based on cognitive Bloom level, Knowledge (C1), Comprehension (C2), Application (C3). The amount of collective data used for training process is 240 questions. Result of this study generates accuracy of 75 % from 60 question samples tested. Susceptibility often occured because of same vocabularies from each categories, thus cause mistakes in classification.*

***Keywords:*** *text classification, Bloom's taxonomy, Machine learning ,naive bayes classifier, natural language processing.*

## I. INTRODUCTION

An evaluation is needed for determining success rate achieved by the students after following an activity of learning. Student assessment is the result of the assessment and/or measurement of the activity of learning where the success rate is marked by the scale value such as letter, word, or symbol [1].

The questions used in the evaluation of year to year are increasingly. Those question can be classified based on Bloom's cognitive levels. By classification, the question will be organized into groups which describing the contain of questions.

The purpose of classification and taxonomy Bloom's cognitive levels is cited by Dimyanti and Mudjiono (2013 ) suggested there are six classes/levels but this paper only used three classes that are :

- Knowledge (C1), is the low level of Bloom's cognitive levels such as recall and notice. These function are study how to know the knowledge term, for example fact, terminology, and the principal.
- Comprehension (C2) is the continue level of Bloom's cognitive such as the ability to understand the meaning of the lesson without related to other lesson.
- Application (C3) is abilityby using the generalof other abstraction which related to correct or new situation.

## II. RELATED WORK

Clasification of questions techniques have been applied to many application areas, however, there is still research that needs to be done on the best method of clasification of questions. The previous research carried out by the method of Quan Zhao, Zhengtao Yu, Lei Su, Jianyi Guo, and Yu Mao [2], with the title of Question Feature Extraction and Semi-supervised Classification Based on Terms Relevance. This research applies, Term Relavance used to calculate the rate of occurrence of words and used Cosine Similarity to calculate word proximity. This study used the Chinese language. The second studies is Question Classification Using Naive Bayes Machine Learning Approach by Rishika Yadav and Megha Mishra [3]. This research applies Naive Bayes for classificition questions by Abbreviation, Description, Entity, Human, Location, and Numeric. The third is Automatic Classification of Questions into Bloom's Cognitive Levels by using Support Vector Machines by Anwar Ali Yahya and Addin Osman [4].

## III. PERFORMANCE

Natural Language Processing (NLP) is one of the disciplines of artificial intelligence and linguistics, which aims to make computer understand a variety of statements written in human language [5].

### a. Preprocessing

Preprocessing is a process of managing the data before the processing data [6]. Preprocessing consist of case folding and tokenizing. Case folding is a process of changing all the letters in a document / sentence to lowercase. Only the letters 'a' through 'z' received [7] while the characters than letters received are considered delimiter. Examples delimiter can be seen in Table I.

Table I. Daftar *Delimiter*

| **Daftar *Delimiter*** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 5 | [ | % | ` | . | ? | \| | ) | ≥ |
| **1** | 6 | ] | ^ | ~ | , | : | ! | - | ∞ |
| **2** | 7 | { | & | \\ | / | ; | @ | _ | π |
| **3** | 8 | } | * | £ | < | ‘ | # | + | ± |
| **4** | 9 | \ | ( | € | > | ‘’ | $ | = | ϕ |

Tokenizing is a process of identification the smallest units (tokens) of a sentence structure (Triawati, 2009). Breaking sentences into single words is performed by scanning a sentence using white space separators such as spaces, tabs, and newline. Filtering is taking important words from the tokens. Stopword is a word that is not descriptive.Schematic of the process of folding and tokenizing case can be seen in Table II.

Tabel II. Preprocessing Sentences Scheme

| **Preprocessing Sentences** | |
|---|---|
| **Question :** | Jelaskan pengertian sejarah sebagai ilmu ! |
| ***Case folding :*** | jelaskan pengertian sejarah sebagai ilmu |
| ***Tokenizing :*** | jelaskan \| pengertian \| sejarah \| sebagai \| ilmu |
| ***Filtering :*** | jelaskan pengertian sejarah ilmu |
| | |

**b. Analysis of Term Frequency Weighting Process**

Weighting TF ( Term Frequency) issued to count the number of occurrences of a particular word in a question. Word weighting values found in each word that appears in a obtained from the question.

$W_{ij} = tf_{ij}$ (1)

Specification:
$W_{ij}$ : the weight of the word / term ti of the document dj
$tf_{ij}$ : number of occurrences of the word / term ti of the document dj

Table III Term Frequency

| No | Term | Frequency | | |
|---|---|---|---|---|
| | | C1 | C2 | C3 |
| 1 | definisi | 1 | 0 | 0 |
| 2 | sejarah | 1 | 1 | 0 |
| 3 | adalah | 1 | 0 | 0 |
| 4 | jelaskan | 0 | 1 | 0 |
| 5 | pengertian | 0 | 1 | 0 |
| 6 | ilmu | 0 | 1 | 0 |
| 7 | tuliskan | 0 | 0 | 1 |
| 8 | kronologi | 0 | 0 | 1 |
| 9 | singkat | 0 | 0 | 1 |
| 10 | peristiwa | 0 | 0 | 1 |
| | Total | 3 | 4 | 4 |

**c. Analysis of Naive Bayes Classifier Method**

This method use calculation of probability, don't pay attention to the sequence of occurrences of words in a text document and consider a text document as the collection of words that arrange the text document [8]. Naive Bayes is one example of method supervised document classification that means it requires training in conducting classification data.

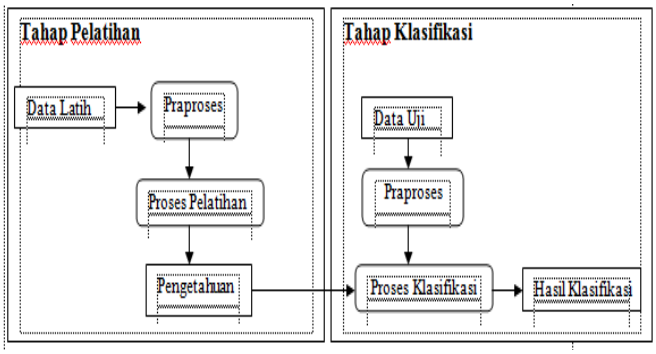


Figure I Software Development Process

**d. Analysis of Question Classification**

Classification of news is a process grouping news coressponding to categories that it has [9]. A document news can be grouped to specific category based on the words and sentences

which have in the news. It also applies in grouping questions. Word which have in a news question has a spesific meaning and can be used as basis to determine the category of news or questions. In this research, it will be built software to predict type the category question of History questions using algorithm for training and classification, which is a Naive Bayes Classifier algorithm [10]. On process classification consist of three process,ie preprocessing, training, and process of classification document.

For examples C1, C2, and C3 based on Bloom's cognitive shown in the following table.

Table IV Example Questions

| No | Question | Level |
|---|---|---|
| 1 | Secara umum definisi sejarah adalah.... | C1 |
| 2 | Yang bukan merupakan hak istimewa VOC yaitu.... | C1 |
| 3 | Proses pengumpulan sumber-sumber sejarah disebut.... | C1 |
| 4 | Mempelajari sejarah bangsa-bangsa lain berguna bagi bangsa kita.<br>SEBAB<br>Manusia sebagai pelaku sejarah terikat pada watak dasar yang sama. | C2 |
| 5 | Jelaskan pengertian sejarah sebagai ilmu | C2 |
| 6 | Jelaskan pengertian sejarah objektif dan sejarah subjektif | C2 |
| 7 | Sejarah memiliki tiga unsur pokok,yaitu manusia, ruang, dan waktu. Dapatkah Anda mengidentifikasi tiga unsur tersebut berdasarkan wacana di atas ? | C3 |
| 8 | Tuliskan kronologi singkat dari peristiwa di atas. | C3 |
| 9 | Uraikan secara ringkas tiga teori asal-usul manusia purba di Indonesia, dan tunjukkan teori mana yang paling populer dan diterima banyak kalangan. | C3 |
| 10 | Peristiwa sejarah merupakan peristiwa yang unik mengandung arti bahwa peristiwa sejarah adalah.... | ? |

| No | Term | Frequency | | | |
|---|---|---|---|---|---|
| | | C1 | C2 | C3 | test data |
| 1 | definisi | 1 | 0 | 0 | 0 |
| 2 | sejarah | 2 | 3 | 1 | 2 |
| 3 | adalah | 1 | 0 | 0 | 1 |
| 4 | bukan | 1 | 0 | 0 | 0 |
| 5 | merupakan | 1 | 0 | 0 | 1 |
| 5 | hak | 1 | 0 | 0 | 0 |
| 6 | istimewa | 1 | 0 | 0 | 0 |
| 7 | yaitu | 1 | 0 | 0 | 0 |
| | proses | 1 | 0 | 0 | 0 |
| 8 | pengumpulan | 1 | 0 | 0 | 0 |
| 9 | sumber | 2 | 0 | 0 | 0 |
| 10 | disebut | 1 | 0 | 0 | 0 |
| 11 | mempelajari | 0 | 1 | 0 | 0 |
| 12 | sebab | 0 | 1 | 0 | 0 |
| 13 | jelaskan | 0 | 2 | 0 | 0 |

| No | Term | Frequency | | | |
|---|---|---|---|---|---|
| | | C1 | C2 | C3 | test data |
| 14 | pengertian | 0 | 2 | 0 | 0 |
| 15 | ilmu | 0 | 1 | 0 | 0 |
| 16 | ojektif | 0 | 1 | 0 | 0 |
| 17 | subjektif | 0 | 1 | 0 | 0 |
| 18 | memiliki | 0 | 0 | 1 | 0 |
| 19 | unsur | 0 | 0 | 1 | 0 |
| 20 | pokok | 0 | 0 | 1 | 0 |
| 21 | dapatkah | 0 | 0 | 1 | 0 |
| 22 | mengidentifikasi | 0 | 0 | 1 | 0 |
| 23 | berdasarkan | 0 | 0 | 1 | 0 |
| 24 | tuliskan | 0 | 0 | 1 | 0 |
| 25 | kronologi | 0 | 0 | 1 | 0 |
| 26 | singkat | 0 | 0 | 1 | 0 |
| 27 | peristiwa | 0 | 0 | 1 | 3 |
| 28 | uraikan | 0 | 0 | 1 | 0 |
| 29 | ringkas | 0 | 0 | 1 | 0 |
| 30 | tunjukkan | 0 | 0 | 1 | 0 |
| 31 | paling | 0 | 0 | 1 | 0 |
| Total | | 14 | 12 | 15 | 7 |

From preprocessing results, the next is calculate the probability of each category. First step ie :

1. Tingkat C1

$$p(c_{C1}) = \frac{fd(c_{C1})}{|D|} = \frac{3}{9} = 0.33$$

2. Tingkat C2

$$p(c_{C2}) = \frac{fd(c_{C2})}{|D|} = \frac{3}{9} = 0.33$$

3. Tingkat C3

$$p(c_{C3}) = \frac{fd(c_{C3})}{|D|} = \frac{3}{9} = 0.33$$

Next look for the value of $p(w_{kj} \mid c_i)$ of each term in each category that have been calculated . The calculation of $p(w_{kj} \mid c_i)$ as follows:

1. Term “sejarah”

$$p(w_{\text{”sejarah”}} \mid c_{C1}) = \frac{f(w_{kj} \mid c_i) + 1}{f(c_i) + |W|} = \frac{2+1}{41+14} = 0.054$$

2. Term “sejarah”

$$p(w_{\text{”sejarah”}} \mid c_{C2}) = \frac{f(w_{kj} \mid c_i) + 1}{f(c_i) + |W|} = \frac{3+1}{41+12} = 0.075$$

3. Term “sejarah”

$$p(w_{\text{”sejarah”}} \mid c_{C3}) = \frac{f(w_{kj} \mid c_i) + 1}{f(c_i) + |W|} = \frac{1+1}{41+15} = 0.035$$

calculation results $p( w_{kj} | c_i )$ :

| No | Term | Frequency | | | test data | $p( w_{kj} \mid c_i )$ | | |
|---|---|---|---|---|---|---|---|---|
| | | C1 | C2 | C3 | | C1 | C2 | C3 |
| **1** | sejarah | 2 | 3 | 1 | 2 | 0.054 | 0.075 | 0.035 |
| **2** | adalah | 1 | 0 | 0 | 1 | 0.036 | 0.018 | 0.017 |
| **3** | merupakan | 1 | 0 | 0 | 1 | 0.036 | 0.018 | 0.017 |
| **4** | peristiwa | 0 | 0 | 1 | 3 | 0.018 | 0.018 | 0.035 |

From the results of the above calculations can be determined by the level of test questions sought opportunities of each level.

P( C1 | SoalUji ) = 0.33 x 0.054 x 0.036 x 0.036 x 0.018
= 4.157 x 10-7

P( C2 | SoalUji ) = 0.33 x 0.075 x 0.018 x 0.018 x 0.018
= 1.443 x 10-7

P( C3 | SoalUji ) = 0.33 x 0.035 x 0.017 x 0.017 x 0.035
= 1.168 x 10-7

After three levels calculated chances, it is known that the level of C1 has a better chance. So test problems including the
**C1 level.**

## IV. EXPREMENTAL

The main problem of this research is how to build software to classification using Naive Bayes Classifier method, so as to identify the relationship between the matter in predicting grade / level. In this research, classification of questions from preprocessing which case folding, tokenizing and filtering. The results of the research can be seen in Tabel II, III, and IV.

This research used 240 questions training set and 60 questions testing set[11][12][13]. Based on the experiment results of the software by entering the 60 samples of question, obtained 15 sample that can not be classification accurately.

Experiment result on this research using 60 sample of questions can be seen on appendix A .
The amount of the accuracy of the three categories

| Category of Question | Number of Accuracy | Number of Data Training | Number of Test Data | Ppercentages |
|---|---|---|---|---|
| C1 | 15 | 80 | 20 | 75 % |
| C2 | 14 | 80 | 20 | 70% |
| C3 | 16 | 80 | 20 | 80% |

Factors that cause of error is the effect of the appearance of a word in each category affect the determination of the question category. The training data are much more varied vocabulary will greatly affect the results about the category. Most of the matter in the training data in a matter of category C2 and C3 has almost the same vocabulary, causing errors during classification. The word "Jelaskan" belongs to the C2 and C3 can cause problems of test data

that contain the word "Jelaskan" can be entered into the C2 or C3 depending on the number of other vocabulary contained in the question. The word "Siapakah" does not exist in the training data so that when it occurs in a matter of test data, the word does not have a probability value. Therefore, the balance of the vocabulary in the training data is necessary, so that the data tested fit into the right category.

Based on the experimental results of 60 samples of question obtained 15 sample of question that can not be classification appropriately. Therefore, the percentage of success of software obtained for 75% of the software is built.

## CONCLUSION

The conclusion that can be take from this study are :

1. Naive Bayes classifier can be used to build software that can classify matter based on Bloom's cognitive domain, in this study only C1, C2, and C3.
2. The results of the classification depends on the amount of training data and the vocabulary used to define the matter falls within the categories C1, C2, or C3.
3. The software has an accuracy of 75% to 60 test data consisting of three categories C1, C2, and C3.
4. In this study, an error of 25% due classification sufficient number of the same words in categories C2 and C3.

**REFERENCES**


[1] Dimyati and Mudjiono, “Belajar dan Pembelajaran”,. Rineka Cipta, Jakarta, Indonesia, 2013.

[2] Daryanto. 2012. Evaluasi Pendidikan. Rineka Cipta, Jakarta, Indonesia.

[3] Zhao, Q., Yu, Z., Su, L., Guo, J., Mao, Y. ,”Question Feature Extraction and Semi-supervised Classification Based on Terms Relevance”, Proceedings of the International Symposium on Intelligent Information Systems and Applications (IISA’09), Qingdao, P. R. China, Oktober 28-30, 2009.

[4] Yadav, R. and Mishra M.,”Question Classification Using Naive Bayes Machine Learning Approach”, International Journal of Engineering and Innovative Technology (IJEIT), 2013.

[5] Yahya, A.A. and Osman, A.,”Automatic Classification of Questions into Bloom’s Cognitive Levels using Support Vector Machines”,In the proceeding of ACIT, 11-14 December 2011, Naif University of Security Sciences, Riyadh, Kingdom of Saudi Arabia.

[6] G. G. Chowdhury, "Natural Language Processing," *Annual Review of Information Science and Technology (ARIST)*, vol. 37, 2003.

[7] R. A. Sukamto, "Penguraian Bahasa Indonesia dengan Menggunakan Penguraian Collins," Magister, Program Magister Informatika, Institut Teknologi Bandung, Bandung, 2009.

[8] C. Triawati, "Metode Pembobotan Statistical Concept Based untuk Klastering dan Kategorisasi," Informatika, ITTELKOM, Bandung, 2009.

[9] Wibisono Y.,” Klasifikasi Berita Berbahasa Indonesia menggunakan Naive Bayes Classifier, Departemen Pendidikan Matematika, FMIPA UPI, 2005.

[10] Distiawan, B.,” Pemanfaatan Dokumen Unlabeled pada Klasifikasi Topik Berbasis Naive Bayes dengan Algoritma Expectation Maximization”,Universitas Indonesia,Depok,2009.

[11] Hapsari, R. dan Adil, M.,2014. Sejarah Indonesia untuk SMK/MAK Kelas XI. Erlangga, Jakarta, Indonesia.
[12] Hapsari, R. dan Adil, M.,2013. Sejarah Indonesia untuk SMA/MA Kelas X. Erlangga, Jakarta, Indonesia.
[13] Hapsari, R. dan Adil, M.,2014. Sejarah Kelompil Peminatan Ilmu-Ilmu Sosial untuk SMA/MA Kelas XI. Erlangga, Jakarta, Indonesia.