

Submission 8

Conference [News](#)

EasyChair


ICoCSIM 2021 Submission 8

Submission information updates are disabled.

For all questions related to processing your submission you should contact the conference organizers. [Click here to see information about this conference.](#)

All **reviews sent to you** can be found at the bottom of this page.

Submission 8

Title	Thesis Topic Classification Based on Abstract Using the Naïve Bayes Method
Paper:	 (Jun 08, 02:42 GMT) (previous versions)
Author keywords	naïve bayes TF-IDF weighting asbtrak classification text mining
EasyChair keyphrases	naive baye method (364), tf idf weighting (221), naive baye (215), text pre processing (142), baye method (115), design bumigora university (95), bumigora university mataram (95), text mining (80), thesis topic (80), weighting tf idf (63), tf idf method (63), pemberian pakan ternak (63), multinomial naive baye (63), term weighting tf (63), computer science (60), term weighting (50), teknologi augmented reality (47), sistem pakar diagnosis (47), penelitian ini adalah (47), pre processing (46), naive baye method classification (40), data collection (40), decision tree (40), actual sensitivity jaringan multimedia (40), multinomial naive baye method (40), tujuan pembuatan sistem pakar (40), bumigora university (40), true jaringan (40), predicted actual sensitivity jaringan (40)
Abstract	The thesis is a requirement for graduation from Bumigora university. The final year student's problem is determining the research topic because the undergraduate thesis collection of Computer Science is not grouped or classified based on student competencies. The purpose of this study was to compare the performance of the naïve Bayes method with TFIDF weighting and without TF-IDF weighting for the classification of thesis topics based on the abstract. The stages of this research are data collection, text pre-processing, term weighting with TF-IDF and without TF-IDF, Naïve Bayes method implementation, and result

	evaluation. Based on the results of the tests that have been done, the naïve Bayes method with TF-IDF has an accuracy of 81.74%, a precision of 86.1%, and a sensitivity of 80.15%. While the naïve Bayes method without TF-IDF weighting produces 88.69% accuracy, 89.76% precision, and 90.49% sensitivity. Thus, the naïve Bayes method without TF-IDF weighting has better performance than TF-IDF weighting for the classification of thesis topics based on the abstract.
Submitted	May 28, 02:02 GMT
Last update	May 28, 02:02 GMT

Authors						
first name	last name	email	country	affiliation	Web page	corresponding?
Hairani	Hairani	hairani@universitasbumigora.ac.id	Indonesia	Universitas Bumigora		✓
Ahmad Islahul	Wathan	wathanruna@gmail.com	Indonesia	Universitas Bumigora		
Kurniadin Abd	Latif	kurniadin@universitasbumigora.ac.id	Indonesia	Universitas Bumigora		
Khairan	Marzuki	khairan@universitasbumigora.ac.id	Indonesia	Universitas Bumigora		
Muhammad	Zulfikri	mzulfikri@universitasbumigora.ac.id	Indonesia	Universitas Bumigora		
Anthony	Anggrawan	anthony.anggrawan@universitasbumigora.ac.id	Indonesia	Universitas Bumigora		

Reviews

Review 1	
<i>Detailed Comments</i>	<p>It's quite a nice sequence and a good idea for a beginner to study. I would like you to follow the comments, to be more understandable for readers.</p> <p>Title: Which one you would like to classifier the thesis or Abstract? I suggest for you this title. The Abstract of thesis classifier by using Naïve Bayes method</p> <p>Abstract: thesis topics based on the abstract.... If you try to extract the topic will have a lot of features that would be a huge dataset used, but in this work seen only used the abstract as three specific topics, Why?... Please recheck again.</p> <p>With TF-IDF and without TF-IDF.... You mentioned twice.... Please recheck again and rewrite.</p>

Need to explain more about thesis topic what is it?

Keywords: check what you mean by asbtrak? Is it Abstract? Also, why use text mining as long as you didn't use it in your work properly? Please recheck....

Introduction: The end of it adds.... Paper organized by....

One of the solutions offered by this research??? In this research or Paper?? Please rewrite probably..... is to use the concept of text mining. Previous research used various methods for text mining-based thesis document analysis such as the k-means method..... In this Paragraph are you used text mining to cover the words as Vector from the topic??

Research methodology: Please describe your work steps how are going on, and refigure to show the pre-processing steps as you mentioned in this paper.

Usually, 10 Fold Cross Validation used with TF-IDF by using dataset an example 200 will be 150 training and 50 testings? So how you applied to get good accuracy with NB please prove...

Text Pre-Processing: Please make a subsection for each step to show your work off about pre-processing. 1. Tokenization 2. Stop word removal 3. Stemming including a figure for each one to show your work how to process it.... Prove it...

C. Term Weighting TF-IDF

Does the TF-IDF method combine two concepts?... What you mean by that, are you trying using concepts, but you extract from were to combine which tool are you used to combine from? Because you work extract topics as words... Please Prove that....

Data classified by the naïve Bayes method are grouped into training and testing data first. Multinomial (What you mean by multinomial) because this usually used for word pairs as using Multinomial logistic regression?

Please provide a figure that shows the input and output of your TF-IDF by using NB and how it affects your topics...

Table 1: What you mean by Confusion Matrix did you coding as a table? Please prove or add your epscode

The equations (5), (6), (7) I didn't see your equation are you apply your method, is that your own equation created?... Its blur does not show anything.... Please recheck.

I didn't see any section of Related Work of previous studies table?? What is the Research Gap in your work? Please must provide a table and section explain about your related work too....

Table 2: Why you extract only on the topic? How about others? And why you used Cross-Validation to get only one topic??.... Please explain?.

D. Naïve Bayes Method Classification

At this stage, the classification is carried out using the naïve Bayes multinomial?? In your paper used NB and now you

mentioned the naïve Bayes multinomial? What are you trying to do with multinomial? Please explain....

Conclusion: please add your future work...chi-square is used for two pairs of words as a topic, did you?

Please remove your Acknowledgment this is for funds if you have to add.....

The paper needs proofreading.

Good Luck!

Review 2

Detailed Comments

- it needs proofreading.

- References should be relevant, recent and readily retrievable

