ISITIA
2018

PROCEEDING

International Seminar on Intelligent Technology and Its Application 2018

August 30 – 31, 2018
Swiss-Belresort Watu Jimbar, Bali, Indonesia

IEEE Catalog Number : CFP18TIA-PRT
ISBN 978-1-5386-7652-3

PROCEEDING

2018 International Seminar on Intelligent Technology and Its Application

ISITIA 2018

August 30 – 31, 2018

Swiss-Belresort Watu Jimbar, Bali, Indonesia

IEEE Catalog Number : CFP18TIA-PRT

ISBN 978-1-5386-7652-3

Proceeding

2018 International Seminar on Intelligent Technology and Its Application

ISITIA 2018

IEEE Catalog Number : CFP18TIA-PRT

ISBN : 978-1-5386-7652-3

Additional copies of this proceeding may be ordered to:

Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember
Gd. B, C, & AJ, Kampus ITS Sukolilo
Surabaya, Indonesia 60111

Preface

Welcome to the 2018 International Seminar on Intelligent Systems and its Applications, or ISITIA. We are very grateful to all attendees today, including all of you who have submitted your recent research work to our conference and present your findings in this event.

In our records, the total submissions to ISITIA 2018 received 152 submissions with authors affiliated with institutions from 13 different countries. We accepted 92 papers for presentation. These papers belong to various topics such as power systems, telecommunications, electronics, control systems, biomedical engineering, and intelligent systems.

ISITIA has its roots from SITIA, or simply Seminar on Intelligent Systems and its Applications, a conference that had started 19 years ago, and have been held annually by our department ever since. We wish to provide a forum where researchers, academics, students, and industry to meet and discuss the latest development in the broad field of electrical engineering, telecommunications, and intelligent systems. In this age of the next industrial revolution, we feel that it is very important that we can extend our research to practical aspects. Hence, “Practical Prospect on New Technologies: From Theory to Industrial Challenges and Business Opportunities” becomes our conference theme this year.

This conference has received tremendous help and support, therefore we would like to thank all reviewers, mainly from three different countries, for their contributions in selecting high quality papers. We would also like to thank Toulouse INP and in particular the LAPLACE laboratory, as well as Udayana University, for their support to this conference. Our gratitude also goes to to Institut Teknologi Sepuluh Nopember, Surabaya, and members of our local organizing committee for the support and help for the conference.

Lastly, please have a great time at the conference, and we wish you a very pleasant stay in Bali, Indonesia.

Dr. Istas Pratomo

ISITIA 2018 General Chair

ORGANIZING COMMITTEE

GENERAL CHAIR
Dr. Istas Pratomo, ST. MT.

CO-CHAIR
Dr. Astria Nur Irfansyah

TREASURER
Nurlita Gamayanti, MT

REGISTRATION & SECRETARIAT
Muhammad Attamimi, Ph.D

PUBLICATION & PROCEEDING
Dr. Ir. Totok Mujiono

LOCAL ORGANIZING COMMITTEE (UDAYANA UNIVERSITY)
Ir. Linawati, M.Eng.Sc,Ph.D
Ni Made Ary Esta Dewi Wirastuti, ST, M.Sc, Ph.D
Wayan Gede Ariastina, S.T., M.Eng.Sc., Ph.D
Ir. I Wayan Sukerayasa, MT
Ir. I Nyoman Setiawan, MT
I Nyoman Satya Kumara, ST, M.Sc, Ph.D

CONTACT DETAILS
Electrical Engineering Department
B Building, Institut Teknologi Sepuluh Nopember
Surabaya, Jawa Timur, Indonesia, 60111
(+62 31) 594 7302
(+62 31) 593 1237
Email: committee@isitia.its.ac.id
website: http://isitia.its.ac.id/

TECHNICAL PROGRAM COMMITTEE & REVIEWERS

Prof. Mauridhi Hery Purnomo, *Institut Teknologi Sepuluh Nopember, Indonesia*
Prof. Gamantyo Hendrantoro, *Institut Teknologi Sepuluh Nopember, Indonesia*
Prof. Maurice Fadel, *Laplace University, France*
Prof. Francois Pigache, *University of Toulouse, France*
Prof. Andani Achmad, *Universitas Hasanuddin, Indonesia*
Prof. Stephane Caux, *INPT/ENSEEIHT, France*
Prof. Nathalie Raveu, *Laplace - Université de Toulouse - UPS INPT CNRS, France*
Prof. Widodo Budiharto, *Bina Nusantara University, Indonesia*
Prof. Mochamad Ashari, *Institut Teknologi Sepuluh Nopember, Indonesia*
Prof. Adi Soeprijanto, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Reza Fuad Rachmadi, *Kumamoto University, Japan, & Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Nurul Ain Mohd Said, *Universiti Teknikal Malaysia Melaka, Malaysia*
Dr. Hana Halim, *Universiti Malaysia Perlis, Malaysia*
Dr. Razaidi Hussin, *Universiti Malaysia Perlis, Malaysia*
Dr. Muhammad Ramli, *Universiti Malaysia Perlis, Malaysia*
Dr. Linawati, *Universitas Udayana, Indonesia*
Dr. Wayan Ariastina, *Universitas Udayana, Indonesia*
Dr. Lie Jasa, *Universitas Udayana, Indonesia*
Dr. Rony Teguh, *University of Palangkaraya, Indonesia*
Dr. Khairul Anam, *University of Jember, Indonesia*
Dr. Delta Prima, *Universitas Surabaya, Indonesia*
Dr. Muhammad Abdillah, *Universitas Pertamina, Indonesia*
Dr. Ramadoni Syahputra, *Universitas Muhammadiyah Yogyakarta, Indonesia*
Dr. Marwan Rosyadi, *Universitas Muhammadiyah Surabaya, Indonesia*
Dr. Mukhammad Andri Setiawan, *Universitas Islam Indonesia, Indonesia*
Dwi Ana Ratna Wati, *Universitas Islam Indonesia, Indonesia*
Dr. Muhammad Anshar, *Universitas Hasanuddin, Indonesia*
Dani Adhipta, *Universitas Gadjah Mada, Indonesia*
Dr. Hanung Adi Nugroho, *Universitas Gadjah Mada, Indonesia*
Dr. Igi Ardiyanto, *Universitas Gadjah Mada, Indonesia*
Dr. Adha Cahyadi, *Universitas Gadjah Mada, Indonesia*
Dr. Ridi Ferdiana, *Universitas Gadjah Mada, Indonesia*
Dr. Eka Firmansyah, *Universitas Gadjah Mada, Indonesia*
Dr. Paulus Insap Santosa, *Universitas Gadjah Mada, Indonesia*
Ahmad Nasikun, *Universitas Gadjah Mada, Indonesia*
Avrin Nur Widiastuti, *Universitas Gadjah Mada, Indonesia*
Anugerah Galang Persada, *Universitas Gadjah Mada, Indonesia*
Harry Prabowo, *Universitas Gadjah Mada, Indonesia*
Dr. Agfianto Eko Putra, *Universitas Gadjah Mada, Indonesia*
Dr. Lesnanto Multa Putranto, *Universitas Gadjah Mada, Indonesia*
Dr. Noor Akhmad Setiawan, *Universitas Gadjah Mada, Indonesia*
Eny Sukani Rahayu, *Universitas Gadjah Mada, Indonesia*
Dr. Selo Sulistyo, *Universitas Gadjah Mada, Indonesia*
Dr. Sunu Wibirama, *Universitas Gadjah Mada, Indonesia*
Dr. Iwan Setiawan, *Universitas Diponegoro, Indonesia*
Dr. Hermawan, *Universitas Diponegoro, Indonesia*
Dr. Linda Sartika, *Universitas Borneo Tarakan, Indonesia*
Dr. Nasaruddin, *Syiah Kuala University, Indonesia*
Dr. Robertus Nugroho, *Soegijapranata Catholic University, Indonesia*
Prof. Nicodimus Retdian, *Shibaura University, Japan*

Dr. Gunawan, *Sekolah Tinggi Teknik Surabaya, Indonesia*
Dr. Mat Syaiin, *Politeknik Perkapalan Negeri Surabaya, Indonesia*
Dr. Agus Indra Gunawan, *Politeknik Elektronika Negeri Surabaya, PENS, Indonesia*
Dr. Prima Kristalina, *Politeknik Elektronika Negeri Surabaya, PENS, Indonesia*
Dr. Anang Tjahjono, *Politeknik Elektronika Negeri Surabaya, Indonesia*
Handy Wicaksono, *Petra Christian University, Indonesia*
Yusak Tanoto, *Petra Christian University, Indonesia*
Dr. Irwan Purnama, *Lembaga Ilmu Pengetahuan Indonesia, LIPI, Indonesia*
Prasetiyono Hari Mukti, *Institut Teknologi Sepuluh Nopermber, Indonesia*
Dr. Adhi Wibawa, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Achmad Affandi, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Trihastuti Agustinah, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Dhany Arifianto, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Achmad Arifin, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Ni Ketut Aryani, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Dimas Anton Asfani, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Muhammad Attamimi, *Institut Teknologi Sepuluh Nopember, Indonesia*
Yusuf Bilfaqih, *Institut Teknologi Sepuluh Nopember, Indonesia*
Fajar Budiman, *Institut Teknologi Sepuluh Nopember, Indonesia*
Rudy Dikairono, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Endroyono, *Institut Teknologi Sepuluh Nopember, Indonesia*
Muhammad Hilman Fatoni, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Puji Handayani, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Ridho Hantoro, *Institut Teknologi Sepuluh Nopember, Indonesia*
Zulkifli Hidayat, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Hendra Kusuma, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Vita Lystianingrum, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Ronny Mardiyanto, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Totok Mujiono, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Ardyono Priyadi, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Ketut Purnama, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Djoko Purwanto, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Dedet Candra Riawan, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Muhammad Rivai, *Institut Teknologi Sepuluh Nopember, Indonesia*
I Gusti Satriyadi, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Eko Setijadi, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Rony Seto Wibowo, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Surya Sumpeno, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Titiek Suryani, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Heri Suryoatmojo, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Supeno Mardi Susiki Nugroho, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Suwadi, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Dimas Uman Putra, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. I Made Yulistya Negara, *Institut Teknologi Sepuluh Nopember, Indonesia*
Ahmad Zaini, *Institut Teknologi Sepuluh Nopember, Indonesia*
Dr. Suharyanto, *PT Brainmatics Cipta Informatika, Indonesia*

TABLE OF CONTENTS

Information Systems

KEYNOTE LECTURE

Biomedical Engineering

Prof. Dr. Mohammad Nuh, DEA
Institut Teknologi Sepuluh Nopember
nuh@ee.its.ac.id

"Biomedical Engineering Development in Indonesia: Challenges and Opportunity"

Indonesia will experience a new era in health services with the enacted Universal Health Coverage (UHC) for all Indonesian people. The program was implemented gradually starting January 1, 2014 for 116 million people and by 2019 will implement for all Indonesians (260 million people). The implementation of UHC is carried out by The Implementing Agency of Social Security or Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS-Kesehatan) as a mandate of the Law Number 36 of 2009 concerning Health and Law Number 24 of 2011 concerning The Implementing Agency of Social Security.

The implementation of UHC promotes the significant increase in demand of healthcare services. This includes the number of beds, doctors, nurses, medical equipments, pharmaceuticals and support services. In addition, the number of middle classes (middle income) is growing rapidly. From 130 million people (55.2%) in 2010 to 215 million people (80%) by 2020. The increasing population of middle class will improve the awareness of the importance of health. This factor could be market driven in the biomedical engineering development.

The development of science and technology, especially material technology, the Internet of Things, technological convergence, the need to improve quality of life and increased purchasing power capacity are appropriate ecosystems in developing Biomedical Engineering.

Biomedical Engineering development in Indonesia is based on point of view science and technology, market opportunity for medical equipment and healthcare system, resources availability, humanity value and the spirit of self-fulfilling in the field of healthcare services. Assistive technology, mobile healthcare, wearable technology, medical imaging, and artificial intelligence are the focus areas but not limited for developing Biomedical Engineering in Indonesia.

As the pioneering stage, the Ministry of Education and Culture in 2012 provided a mandatory assignment to Institut Teknologi Sepuluh Nopember (ITS), Institut Teknologi Bandung (ITB), Universitas Indonesia (UI), Universitas Gadjah Mada (UGM) and Universitas Airlangga (UNAIR) to establish Department and Research Center of Biomedical Engineering.

Muhammad Nuh (born 17 June 1959 in Surabaya) is the former Minister of Education and Culture of Indonesia in the Second United Indonesia Cabinet of Susilo Bambang Yudhoyono. Prof. Nuh was born on 17 June 1959 in Surabaya into a large farming family. By profession an electrical engineer, he was educated at Sepuluh Nopember Institute of Technology and Montpellier 2 University, France. More info: https://opensciencemeeting.org/prof-dr-ir-mohammad-nuh-dea/.

KEYNOTE LECTURE

Modeling and the control of the electric systems

Prof. Maurice Fadel
National Polytechnic Institute of Toulouse - LAPLACE
maurice.fadel@laplace.univ-tlse.fr

"Some Advances of Control in Power Electronics: Predictive and Allocation Method"

Static converter based power electronic systems and electrical power conversion systems using electrical machines require high performance while preserving simple designs and easy implementations. This situation becomes critical when the processed powers increase and the converters have a large number of switches for topologies to increase power through serialization and / or paralleling. In this context it is important to develop new control approaches such as predictive control or allocation control.

The predictive control applied to static converters offers simple and often effective solutions for conventional structures. When the number of switches increases the combinatorial explosion of the controls induces a limitation especially when the operating frequency is high. For this purpose the allocation approach based on the minimization of a criterion in real time is proposed as an alternative solution allowing further reconfiguration of the system during the appearance of faults.

As for static converters, electrical machines are now associated in cooperative systems, which produces multi-machine and multi-converter systems. The design of control laws becomes more complex especially if we take into account the improvement of energy efficiency. This presentation provides an overview of recent developments in this area, drawing on examples of multilevel static converters and synchronous magnet machines alone or in cooperation for high speed aeronautical or automotive applications. Implementation procedures will also be discussed and illustrated through examples deployed on DSPACE or implemented on FPGAs.

Maurice Fadel was born in Toulouse (France). He got the PhD degree at the Institut National Polytechnique de Toulouse in 1988, in the domain of the Control in Electric Engineering. He is currently a Professor in the Ecole Nationale Supérieure d'Ingénieurs en Electrotechnique, d'Electronique, d'Informatique, d'Hydraulique et de Télécommunications of Toulouse (ENSEEIHT). More info: http://www.laplace.univ-tlse.fr/FADEL-Maurice-1464?lang=en.

KEYNOTE LECTURE

Modeling modal method (plasma cavity, SIW circuits, transmission line) closer to the physical, wave concept iterative process, numerical analysis

Prof. Nathalie Raveu
National Polytechnic Institute of Toulouse - LAPLACE
raveu@laplace.univ-tlse.fr

"Propagation in Waveguides with Metamaterial Walls"

With the increasing needs of communications equipment for space applications, it is important to minimize the size and mass of satellite equipment. This leads to a reduction in the launch costs of the satellites in their orbit or allows the possibility of adding equipment to the rocket. The objective is to reduce the antenna or waveguide dimensions without deteriorating their performances (directivity, cross polarization, monomode band, etc.).

It is possible to control the propagation of electromagnetic waves in horn antennas and waveguides using anisotropic walls (corrugations, metamaterials). Thus, expected propagation constants and radiation properties of electromagnetic waves differ from in classical horn antennas and waveguides using anisotropic walls in such structures: for the waveguide, cutoff frequency reduction is possible; for the horn antenna, improvement of the directivity or reduction of side lobe level.

A new design methodology based on an Expanded Modal Theory Theory (TME) is proposed to characterize waveguides with anisotropic walls in collaboration with CNES agency and the MVG Company. It makes it possible to dimension very quickly metamaterial surfaces most adapted to the required applications. A prototype waveguide was designed, manufactured and measured using this methodology. The results obtained demonstrate the interest, efficiency and general character of the proposed method for the design of guided microwave devices with anisotropic walls.

Nathalie Raveu received the M.S. degree in electronics and signal processing in 2000 and the Ph.D. degree in 2003.She is a Professor with the National Polytechnic Institute of Toulouse (INPT) and a Research Fellow with the LAPLACE—CNRS (LAboratory of Plasma and Energy Conversion). More info: http://www.laplace.univ-tlse.fr/RAVEU-Nathalie-1119?lang=en.

KEYNOTE LECTURE

Relay Protection, Power Transmission, Power Systems,
Transformer designers

Ir. Eko Yudo Pramono, M.T.
General Manager Dispath Center
PT PLN (Persero)
ekoyudopramono@gmail.com

"Implementation of HVDC System for Improving Java Bali System"

Java Bali Grid is the largest electricity system in Indonesia, with its peak load reaching 25.880 MW on Tuesday, 08th May 2018 and total installed generating capacity of 36.211 MW. The transmission lines consist of 500 kV as the backbone with total length of 5.074 kmc. Power plants are connected to various transmission voltage, but the trend is that more bigger plants are connected to 500 kV main backbone. Currently the biggest plant unit is 815 MW, but starting on 2021 units of 1.000 MW ultra super critical coal-fired power plant will come on-line. Investment plan as stipulated in the RUPTL 2018-2027 dictates more than 12.000 MW of 1GW units will be commisioned within the next 10 years, and all will be connected to 500 kV transmission lines. Although these units will run at higher efficiency and lower generation cost, concern arises on the aspect of 500 kV system stability and capability to accommodate such amount of 1 GW units into the Java-Bali Grid. The short circuit level contribution from these plants will run significantly high above the existing breaker rating, at some point will reach more than 90 kA. Not only that, stability of the system after losing one or two 1 GW unit in contingencies also will be of concern. Therefore the study of grid impact of these 1 GW units is needed, including alternatives to improve grid stability and robustness. Possible solutions include new higher voltage of backbone such as 765 kV AC, or developing an HVDC system in Java Bali either it is DC link or HVDC back to back system. The DC system will be able to split system into several areas hence lowering the short circuit level, also isolate any fault propagating across grid therefore avoid the possibility of total black out. This will in turn increases the robustness of Java Bali grid. This lecture looks at the implementation of HVDC system to improve Java Bali grid, and proposes a distict merit of apply HVDC technology into the grid.

Curriculum vitae:

2016 – present	: General Manager Dispath Center, PT. PLN (Persero) P2B
2014 – 2016	: General Manager Transmission And Dispath Center, PT. PLN (Persero) P3B Sumatera
2012 – 2016	: Senior Manager Transmission Asset Management, PT. PLN (Persero) Head Office
2009 – 2012	: Planning Manager PT PLN (Persero) P3B Sumatera

Searching Cheapest Product On Three Different E-Commerce Using K-Means Algorithm

Vincentius Riandaru Prasetyo
Department of Informatics Engineering
University of Surabaya
Surabaya, Indonesia
vincent@staff.ubaya.ac.id

***Abstract*— The development of e-commerce is starting to change people's lifestyles, not least the people of Indonesia. The existence of e-commerce is very helpful user in buying and selling products. There are many e-commerce that can be found today. Some of the famous e-commerce in Indonesia are Bukalapak, Lazada, and Blibli. A large number of existing e-commerce makes users, especially buyers, have difficulty when looking for products at the cheapest price. This happens because each e-commerce offers different prices for the same product. This research aims to make the cheapest product search system in Bukalapak, Lazada, and Blibli using K-Means algorithm. The results of experiments showed that K-Means algorithm can be used to classify product data from Bukalapak, Lazada, and Blibli well. The results of the clustering process can also help for searching the cheapest products from the three e-commerce becomes faster. However, the number of clusters used will affect the effectiveness of the search process on the system.**

Keywords*— *e-commerce, clustering, k-means, cosine similarity, precision, recall

I. INTRODUCTION

E-Commerce is the distribution, sales, purchasing, marketing of products, and services that rely on electronic systems, such as the Internet, television, or other computer networks. E-commerce involves the transfer of funds and electronic data exchange, management systems and data collection automatically. E-commerce is one of the most popular businesses in Indonesia because it provides a promising advantage. The advantage of e-commerce is that we can sell products or services online without having to set up shop or big office. Another advantage is the ease of communicating between the seller and the buyer. Marketing of products is also much more profitable because we do not need to spend high cost to do the promotion [1].

Bukalapak, Lazada, and Blibli are famous e-commerce in Indonesia. Based on data from Google Trends for top 5 e-commerce in Indonesia which has the highest search in 2017, Bukalapak, Lazada, and Blibli included [2]. Bukalapak was founded by Achmad Zaky in early 2010 as a digital agency division named Suitmedia located in Jakarta. Bukalapak is one of the fastest growing e-commerce in Indonesia. Starting from a safe system in the process of buying and selling, has a wide and growing community in every city in Indonesia, rapid disbursement of funds for sellers, pioneer transaction security in the world of online trading, and many more advantages Bukalapak [3]. Lazada Indonesia was established in 2012 and is one of the branches of Lazada's online retail network in Southeast Asia. Lazada Southeast Asia is a subsidiary of the Rocket Internet network, a German company. As a newcomer to the Indonesian e-commerce business, Lazada is able to grab the attention of most internet users [4]. Blibli.com is an e-commerce that comes with the concept of an online mall in Indonesia since 2011. Blibli.com not only provide completeness of choice, but also the flexibility of payment method, delivery method, and service [5].

The number of e-commerce in Indonesia, causing competition among them, especially the selling price of products. There are often price differences across e-commerce sites even in the same stuff and the same sellers. The buyers will prefer the cheapest price on the same item. It will be difficult for the buyer because they have to check the price of an item from one e-commerce to another e-commerce. Therefore, this research aims to make a useful application to compare the price of an item sold through Bukalapak, Lazada, and Blibli. This application implements K-Means algorithm for grouping similar items.

K-means is non-hierarchical clustering method that categorize the data in the form of one or more clusters/groups. The data has same characteristics will be grouped in one cluster/group and the different characteristics will be grouped in other clusters/groups so the data has a small level of variation in one cluster/group [6].

The remaining of this paper is organized as follows. Section II discusses related works. Section III and IV explain our method and the result of experiments, respectively. Section V contains conclusions and future works.

II. RELATED WORKS

K-Means algorithm has been used in many previous kinds of research or cases. The research by Dhuhita [7] used K-Mean algorithm to determine the nutritional status of children in Indonesia. In addition, Dhuhita [7] also compares the results of grouping between by using k-means and Growth Chart First. K-Means algorithm is also used by Wardhani [8] to classify the patient's disease in Puskesmas Kajen Pekalongan. Patient's disease will be grouped into 2 main categories: "acute" and "not acute".

The research by Asroni and Adrian [9] examines the data already in the data warehouse of Muhammadiyah University of Magelang to find 5 students in Department of Informatics Engineering in order to select students who deserve to follow Cyberjawara competition. The initial phase of the system will perform grouping based on the criteria, such as the grade of Algorithm Programming, Basic Physics, Calculus, and GPA. The criteria are processed using the K-Means method.

The other related research on e-commerce, Muningsih and Kiswati [10] implement K-Means to determine the inventory on the online shop *Ragam Jogja*. The variables

used by Muningsih and Kiswati [10] are product code, transaction amount, sales volume, and average sales. The result of the research is an application that can classify products into categories of stock quantities like many, medium, and slightly based on sales transactions.

Gunawan, et al. [11] in their research compares K-Means and the Apriori method to find out the items that are often purchased on an e-commerce and also sees a faster time in analyzing sales transactions. The results of this research conclude that the k-means algorithm is faster than the apriori algorithm for determining the products often purchased by consumers.

III. RESEARCH METHODS

A. Crawling Data

The purpose of data crawling is to search and collect information about a particular product that will be used on the application. Data crawling is done using *jaunt* library. *Jaunt* is java library for web-scraping & web-automation, including JSON querying. The library provides an ultra-light headless browser. By using *Jaunt* Java programs can easily perform browser-level, document-level, and DOM-level operations. Informations obtained from the crawling process such as title, description, price, URL, and pictures taken through HTML tags from Bukalapak, Lazada, and Blibli.

Bukalapak, Lazada, and Blibli have different template URLs. These differences cause the process of data crawling to be different because the location of the required information resides in different html tags.

There are many product categories in each e-commerce. However, only a few categories will be used: mobile phones, computers, CCTV, cameras, and televisions. The results of the crawling process are stored in a file with the format *.csv*.

B. Preprocessing

Preprocessing is performed for the result of the crawling process in order to get clean data so that the process of clustering and searching product are more accurate. The preprocessing consists of several steps and is shown in Fig. 1. Steps of preprocessing are as follows:

1. *Case Folding*
   In this step, all letters in a word are changed into lowercase.
2. *Remove Stopword*
   This step used to eliminate words that had no effect in the process of categorization, such as: *yang, dan, atau, ke, dari*, etc.
3. *Stemming*
   This step is a process to find the root of a word that will be implemented with Nazief Andriani algorithm (*Sastrawi 1.0.1*). The process of stemming consists of two main phase, ie: checking of the basic word and elimination of affixes, prefixes, suffixes [12].
4. *Tokenization*
   Tokenization is a process to split a sentence into words. The results of tokenization will be saved as clean data.

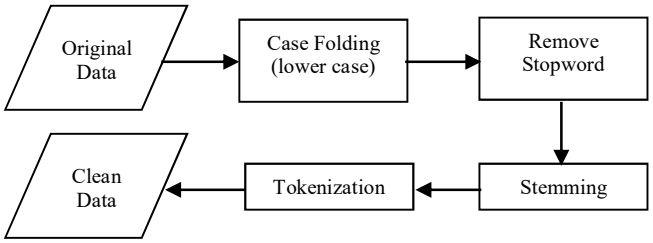

Fig. 1. Preprocessing Steps

C. Term Extraction and Term Weighting

Term extraction aims to eliminate the duplication of words contained in the preprocessing results so that will be obtained a set of unique words [12]. Term weighting will be calculated using term frequency and TF-IDF. Term frequency (TF) is the number of a term or a word that appears in a document. Term weighting with TF-IDF is calculated based on (1).

$$tf.idf_{t,d} = tf_{t,d} \; x \; \log(\frac{N}{df_t}) \qquad (1)$$

The value of $tf_{t,d}$ is the term frequency value of a term t in document d. The N value based on the number of data that are used, while df_t is the number of documents containing term t. As an example, we will show term weighting process. Examples of document that are used, as follows:

- Doc1: *samsung galaxy j2 smartphone gold*
- Doc2: *samsung galaxy note3 plus casing samsung*
- Doc3: *samsung galaxy j2 second bonus casing*

The first process is calculating the term frequency of document. For example, the appearance frequency of the term ”samsung” at *Doc1* and *Doc3* is 1 and at *Doc2* is 2. After the term frequency value is obtained, the next step is calculating TF-IDF value for each term. TF-IDF value of each term in *Doc1, Doc2*, and *Doc3* are shown in Table I.

TABLE I. TF-IDF OF DOCUMENTS

Term	Term Frequency			df_t	TF-IDF (*N=3*)		
	Doc1	*Doc2*	*Doc3*		*Doc1*	*Doc2*	*Doc3*
samsung	1	2	1	3	0	0	0
galaxy	1	1	1	3	0	0	0
j2	1	0	1	2	0,58	0	0,58
smartphone	1	0	0	1	1,58	0	0
gold	1	0	0	1	1,58	0	0
note3	0	1	0	1	0	1,58	0
plus	0	1	0	1	0	1,58	0
casing	0	1	1	2	0	0,58	0,58
second	0	0	1	1	0	0	1,58
bonus	0	0	1	1	0	0	1,58

D. Clustering Data

Clustering is a method used to divide objects into groups based on predefined similarities. Objects will be grouped into one or more clusters so that the objects in one cluster will have a high similarity with each other [13]. One of the most commonly used algorithms for clustering is K-Means.

K-Means is a method of data clustering with unsupervised learning. This method will group data with the

partition system. K-Means aims to minimize data variations on a cluster. The main principle of this method is to compile a center of the cluster (centroid) from the dataset. Iteratively, the cluster is continually updated until convergent. Illustration of data clustering by K-Means method is shown in Fig. 2.

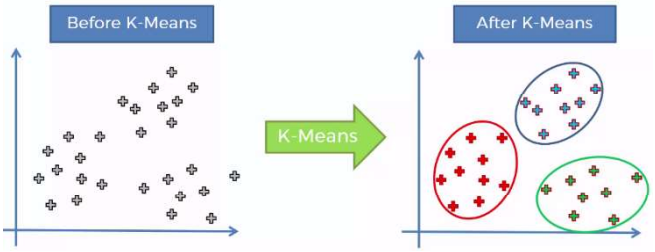

Fig. 2. Ilustration of K-Means

K-Means algorithm consists of several steps which start from determining the number of clusters to group the data based on the value of similarity to the centroid. The steps of the k-means algorithm are shown in Fig. 3.

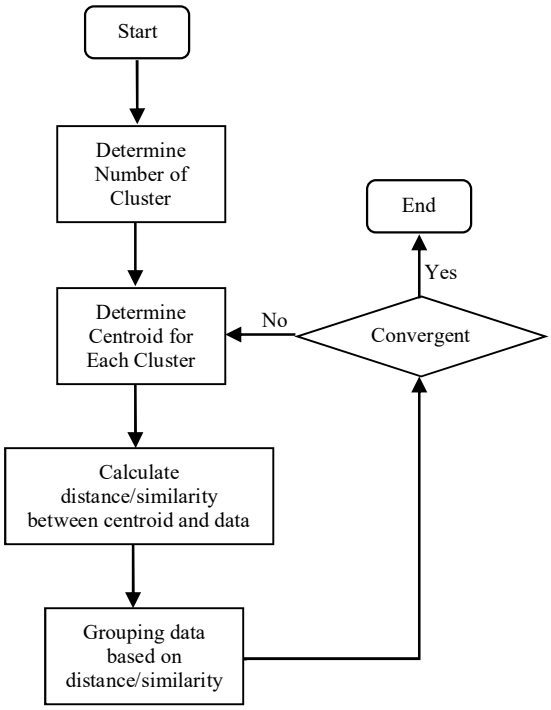

Fig. 3. K-Means Steps

At the first iteration, the centroid is determined randomly from the existing data. If clustering has not produced a convergent cluster, then to determine the centroid in the next iteration by calculating the mean of the cluster member. Wecan use any method to calculate distance/similarity between centroid and data, such as: Euclidean distance, Manhattan distance, cosine similarity, or Jaccard similarity. This research uses cosine similarity method for distance/similarity calculation.

Cosine similarity is a method to calculate the similarity between two objects expressed in two vectors by using keywords from a document as a measure. The formula used by cosine similarity is (2), where A_i is TF-IDF value of *Doc1* and B_i is TF-IDF value of *Doc2* that have similar term [14].

$$\cos \propto = \frac{A \cdot B}{|A||B|} = \frac{(\sum_{i=1}^{n} A_i \, x \, B_i)}{\sqrt{\sum_{i=1}^{n} (A_i)^2 x \sum_{i=1}^{n} (B_i)^2}} \quad (2)$$

When 2 documents are identical, the angle is 0° and the value of the similarity is 1, whereas when the documents and queries are not identical at all, the angle is 90° and the value of the similarity is 0. As an example, we will show K-Means clustering process. Examples of the document that are used, as in Table I. After TF-IDF value is obtained in Table I, the next step are determining the number of cluster and centroid for each cluster. For example, the number of clusters is 2 (*C1* and *C2*), *Doc1* as centroid *C1*, and *Doc3* as centroid *C2*. After that, we calculate the similarity value between centroid and document, which is calculated based on Equation (2). The results of the clustering shown in Table II.

TABLE II. CLUSTERING RESULT (1ST ITERATION)

Term	Term Frequency			df_t	TF-IDF (*N=3*)		
	Doc1	*Doc2*	*Doc3*		*Doc1*	*Doc2*	*Doc3*
samsung	1	2	1	3	0	0	0
galaxy	1	1	1	3	0	0	0
j2	1	0	1	2	0,58	0	0,58
smartphone	1	0	0	1	1,58	0	0
gold	1	0	0	1	1,58	0	0
note3	0	1	0	1	0	1,58	0
plus	0	1	0	1	0	1,58	0
casing	0	1	1	2	0	0,58	0,58
second	0	0	1	1	0	0	1,58
bonus	0	0	1	1	0	0	1,58
Similarity With *C1*					1	0	0,06
Similarity With *C2*					0,06	0,03	1
Clustering Result					C1	C2	C2

Based on Table II above can be seen that *Doc1* into the cluster 1 (C1) and *Doc2* and *Doc3* into the cluster 2 (*C2*). To ensure that the cluster is convergent, then it will be done again by clustering. The results of clustering at the second iteration shown in Table III.

TABLE III. CLUSTERING RESULT (2ND ITERATION)

Term	Cluster		TF-IDF (*N=3*)		
	C1	*C2*	*Doc1*	*Doc2*	*Doc3*
samsung	0	0	0	0	0
galaxy	0	0	0	0	0
j2	0,58	0,29	0,58	0	0,58
smartphone	1,58	0	1,58	0	0
gold	1,58	0	1,58	0	0
note3	0	0,79	0	1,58	0
plus	0	0,79	0	1,58	0
casing	0	0,58	0	0,58	0,58
second	0	0,79	0	0	1,58
bonus	0	0,79	0	0	1,58
Similarity With *C1*			1	0	0,06
Similarity With *C2*			0,04	0,72	0,74
Clustering Result			C1	C2	C2

The Based on Table II and III above that can be seen the cluster is already convergent. This is because Doc1, Doc2, and Doc3 remain on the same cluster in the first and second iterations.

IV. EXPERIMENTS AND RESULTS

A. System Testing

This section discussed the testing of the system that has been built. The purpose of this system is searching for the cheapest product from Bukalapak, Lazada, and Blibli. Before the user can search a product, crawling, term extraction, term weighting, and clustering will be run first. The crawling process will be run to get the desired data, as described in the previous section.

The number of product data that is used for testing are 2,280. The number of product data is divided into 10 titles that have been inputted previously by the user. Distribution of product data for each title is shown in Table IV.

TABLE IV. DISTRIBUTION OF PRODUCT DATA

Category	Title	Source	Number of Data
mobile phones	samsung galaxy	Blibli	72
		Bukalapak	99
		Lazada	72
	asus zenfone	Blibli	70
		Bukalapak	100
		Lazada	72
	iphone	Blibli	72
		Bukalapak	100
		Lazada	72
cameras	kamera dslr	Blibli	72
		Bukalapak	60
		Lazada	72
computers	asus notebook	Blibli	72
		Bukalapak	110
		Lazada	72
	acer notebook	Blibli	72
		Bukalapak	39
		Lazada	72
	apple macbook	Blibli	72
		Bukalapak	100
		Lazada	72
cctv	cctv	Blibli	72
		Bukalapak	100
		Lazada	72
televisions	samsung tv	Blibli	72
		Bukalapak	80
		Lazada	72
	sony tv	Blibli	66
		Bukalapak	60
		Lazada	72

The system will process the data in Table IV into preprocessing, term extraction, and term weighting before clustering. After the process is complete, clustering will be automatically run by the system. The user can input the number of clusters manually before clustering is executed. The results of clustering data with 5 clusters are shown in Fig. 4 and 10 clusters in Fig. 5.

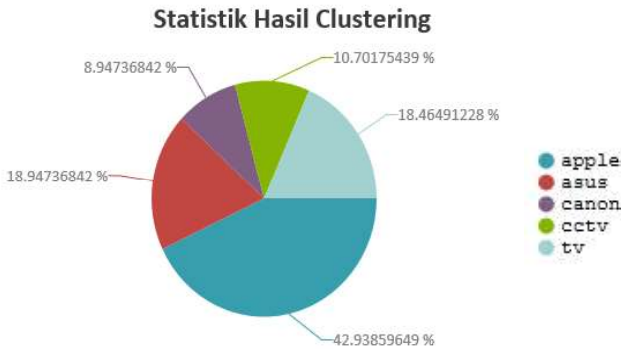

Fig. 4. Clustering With 5 Clusters

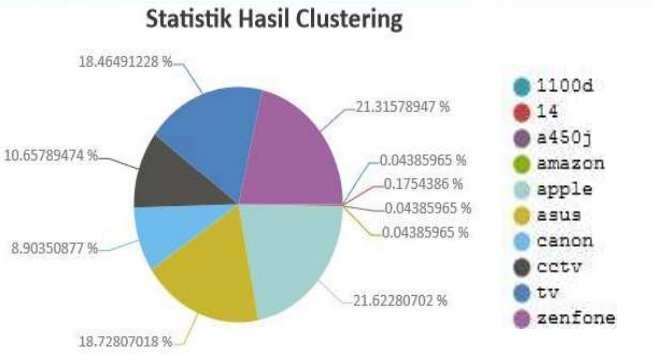

Fig. 5. Clustering With 10 Clusters

Based on Fig. 4 and Fig. 5 above can be seen that the name of each cluster is assigned automatically by the system. The name of the cluster is based on the most frequent terms in the cluster. After the system completes all the processes before, then the user can do a product search. The user will input the name, e-commerce, and category of the product, then the system will display the appropriate product data.

The system will search for product data based on the clustering results that have been done before. For example, the user inputs "apple iPhone" as the product name, "all e-commerce" as e-commerce, "all categories" as the category, and there are 5 product clusters as in Fig. 4. Based on Fig. 4, product data included in the "apple" cluster amounted to 42.94% of the total data, that means amounted to 979 data. First, the system will search on the "apple" cluster because there is "apple" word in the name of the product inputted by the user. It aims to speed up the search because the system does not have to search from all data. If the system does not find a matching word between user input and cluster name, then the system will search from all data. The display of search results performed by the user is shown in Fig. 6.

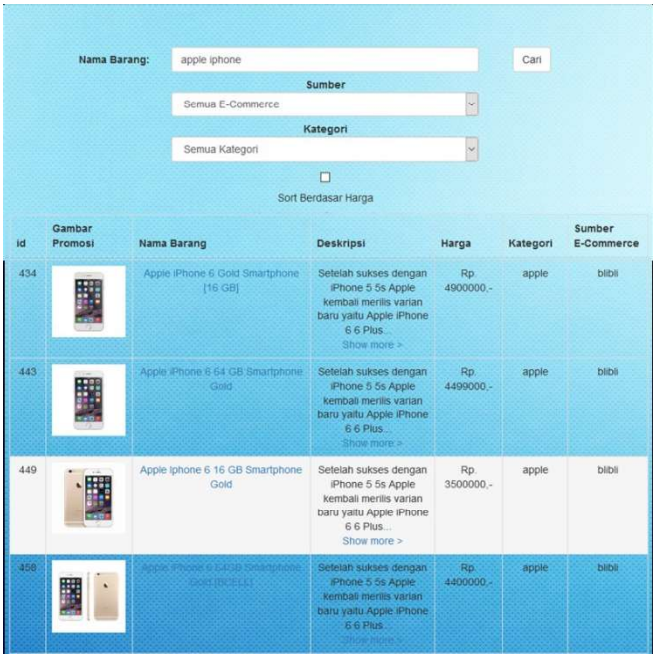

Fig. 6. The Display of Search Results

B. *System Evaluation*

To determine whether the system is working properly or not, it will be calculated the value of precision and recall. Precision is the level of accuracy between the information requested by the user and the answers provided by the system. While the recall is the level of success of the system in rediscovering an information. The Confusion matrix is used to facilitate the calculation of precision and recall values, as shown in Table V.

TABLE V. CONFUSION MATRIX

	Relevant	Not Relevant	Total
Retrieved	A	B	A+B
Not Retrieved	C	D	C+D
Total	A+C	B+D	A+B+C+D

The precision value is calculated from the number of relevant documents found by the system divided by the total document found by the system, whereas the recall is calculated from the number of relevant documents found by the system divided by the total of the relevant documents. Therefore, based on Table V above, it can be concluded that the precision value is *A/(A+B)* and the recall value is *A/(A+C)*.

The system evaluation uses 10 different keywords to search the product data, while for the number of clusters used are 5 and 10. In addition, searching will be performed on all e-commerce and categories. Table VI and VII show the results of calculating precision and recall values for 5 clusters and 10 clusters.

TABLE VI. PRECISION AND RECALL FOR 5 CLUSTERS

Keyword	Precision	Recall
apple iphone	0,49	0,97
samsung smartphone	0,74	0,44
canon dslr	0,84	0,35
led tv	0,86	0,37
notebook ram 4gb	0,59	0,12
lcd tv	0,27	0,15
notebook 14 inch	0,65	0,81
smartphone asus	0,61	0,17
iphone free casing	0,51	0,48
samsung anti gores	0,49	0,33
Average	**0,605**	**0,418**

TABLE VII. PRECISION AND RECALL FOR 10 CLUSTERS

Keyword	Precison	Recall
apple iphone	0,60	0,51
samsung smartphone	0,65	0,78
canon dslr	0,52	0,86
led tv	0,70	0,78
notebook ram 4gb	0,91	0,57
lcd tv	0,52	0,57
notebook 14 inch	0,88	0,85
smartphone asus	0,98	0,84
iphone free casing	0,91	0,63
samsung anti gores	0,81	0,99
Average	**0,747**	**0,738**

A searching system has good effectiveness if it has the same or 1: 1 ratio of precision and recall value. Based on Table VI and VII above, itcan be seen that the ratio of precision and recall value for 10 clusters is close to 1:1. This shows that the number of clusters will affect the effectiveness level of the system. The system will search the data based on the similarity between the cluster name and the keyword before it searches from all data. If the number of clusters more and more, then the amount of data contained in the cluster will be more spesific. This causes the amount of relevant data and retrieved data to increase as well.

V. CONCLUSIONS AND FUTURE WORKS

Based on the test results on section IV, it can be concluded that the K-Means algorithm can be used to classify the product data from Bukalapak, Lazada, and Blibli sufficiently. The results of the clustering process can also help to search for the cheapest products from the three e-commerce faster. However, the number of clusters used will affect the effectiveness of the search process on the system. It can be seen that the ratio of precision and recall is better on 10 clusters. The system can not be implemented in the real world scenario which usually has a lot of products for now because K-means needs a predefined number of cluster.

For the future works, the system can be developed by using other methods, such as K-Nearest Neighbors. It aims to find out which method is suitable for the system in order to work more effectively. In addition, it can also be add more training data so there will be more variations of data on a cluster.

REFERENCES

[1] A.G. Khan, "Electronic Commerce: A Study on Benefits and Challenges in an Emerging Economy", *Global Journal of Management and Business Research: B Economics and Commerce*, 2016, Vol. 16, pp. 19-22.

[2] R. Haryanto, “E-Commerce Indonesia Competition in 2017, Who's Famous?”, “Persaingan e-Commerce Indonesia di 2017, Siapa Terpopuler?”, Transl. Available: https://inet.detik.com/cyberlife/d-3793019/persaingan-e-commerce-indonesia-di-2017-siapa-terpopuler. [Accessed: 1- May- 2018].

[3] C.F. Safira, A. Kusyanti, and H. Aryadita, “Analysis of Quality of Service Bukalapak For User Satisfaction Using E-S-QUAL”, “Analisis Kualitas Layanan Website Bukalapak Terhadap Kepuasan Pengguna Menggunakan E-S-QUAL”, Transl. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2017, Vol. 1, No. 12, pp. 1813-1821.

[4] G.S. Budhi, “Analysis Of E-Commerce System In Lazada Indonesia”, “Analisis Sistem E-Commerce Pada Perusahan Jual-Beli Online Lazada Indonesia”, Transl. *Jurnal Electronics, Informatics, and Vocational Education*, 2016, Vol. 1, No. 2, pp. 78-82.

[5] W. Sihombing, 2013, “Analysis Of Factors Affecting The Use Of Online Payment Service On E-Commerce”, “Analisis Faktor Yang Mempengaruhi Penggunaan Layanan Pembayaran Online Pada Ecommerce”, Transl. University of Indonesia in press.

[6] J.O. Ong, “Implementation Of K-Means Clustering Algorithm For Determining The Marketing President University Strategy”, “Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University”, Transl. *Jurnal Ilmiah Teknik Industri*, 2013, Vol. 12, No. 1, pp. 10-20.

[7] W.M.P. Dhuhita, “Clustering Using K-Means Method For Determining Nutrition Status Of Children”, “Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita”, Transl. *Jurnal Informatika*, 2015, Vol. 15, No. 2, pp. 160-174.

[8] A.K. Wardhani, ”Implementation Of K-Means Algorithm For Classification Of Patient Diseases In Puskesmas Kajen Pekalongan”, “Implementasi Algoritma K-Means Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Kajen Pekalongan”, Transl. Jurnal Transformatika, 2016, Vol. 14, No. 1, pp.30-37

[9] Asroni and R. Adrian, “Implementation Method for K-Means Clustering Based Student Value with Weka Interface a Case Study of Department of Information UMM Magelang”, “Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang”, Transl. *Jurnal Ilmiah Semesta Teknika*, 2015, Vol. 18, No. 1, pp.76-82.

[10] E. Muningsih and S. Kiswati, “Implementation of K-Means Method For Online Shop Product Clustering In Determination of Product Stocks”, “Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang”, Transl. *Jurnal Bianglala Informatika*, 2015, Vol. 3, No. 1, pp.10-17.

[11] Gunawan, F. Halim, T.S. Debataraja, and J.E. Peranginangin, “Development Of Application Determination Of Ecommerce Benefits Using K-Means And Apriori Algorithm”, “Pengembangan Aplikasi Penentuan Tingkat Keuntungan Pada Ecommerce Dengan Menggunakan Algoritma K-Means Dan Apriori”, Transl. *Proceedings of Seminar Nasional Teknologi Informasi dan Komunikasi 2016*, Yogyakarta, 2016.

[12] V. R. Prasetyo and E. Winarko, "Rating Of Indonesian sinetron based on public opinion in Twitter using Cosine similarity," *2016 2nd International Conference on Science and Technology-Computer (ICST)*, Yogyakarta, 2016, pp. 200-205.

[13] B.M. Metisen and H.L. Sari, “Clustering Analysis Using K-Means Method In The Groups Of Products Sales On Fadhila Support”, “Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila”, Transl. *Jurnal Media Infotama*, 2015, Vol. 11, No. 2, pp.110-118.

[14] R.T. Wahyuni, D. Prastiyanto, and E. Supraptono, “Implementation of Cosine Similarity Algorithm and TF-IDF Weighting on Thesis Document Classification System”, “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi”, Transl. *Jurnal Teknik Elektro*, 2017, Vol. 9, No. 1, pp.18-23.