Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Cite as: AIP Conference Proceedings **2470**, 010001 (2022); https://doi.org/10.1063/12.0009021
Published Online: 25 April 2022

Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Welcome to the 3rd biannual International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021). This year, InCITE is organized from Surabaya, Indonesia and researchers from around the world discuss current challenges and technological solutions via an online platform on August 25-26, 2021.

InCITE 2021 presents "Leveraging Smart Engineering" as its central theme. In response to several world challenges, such as sustainable development, global convergence of information and communications technologies, along with smart systems as opportunities as well as challenges in developments for better industries, it is considered important to discover innovative approaches from science and engineering perspectives. Innovation suggests the introduction of novelty to create better solutions. Innovation in engineering and science requires contributions from multidisciplinary sectors, involving industries, practitioners, researchers, and academics.

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 010001-1–010001-1; https://doi.org/10.1063/12.0009021
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

Committees: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Cite as: AIP Conference Proceedings **2470**, 010002 (2022); https://doi.org/10.1063/12.0009708
Published Online: 25 April 2022

WELCOME FROM InCITE 2021 STEERING COMMITTEE

It is a great pleasure to welcome all of you to the 3rd Bi-Annual International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021) held by the Faculty of Engineering, University of Surabaya (UBAYA). The first and second InCITE have been successfully held in Bali, Indonesia in 2017 and 2019. Hence, now we are delighted to host the third InCITE through online media due to the Covid-19 pandemic situation.

There are 37 papers that have been selected to be presented in InCITE 2021. The papers were written by experts not only from Indonesia, but also from different parts of the world. The main theme of this conference is Leveraging Smart Engineering in response to the current and future Industrial Revolution 4.0 that should be handled by every country in the world. We hope through this conference, all participants will be able to know each other and interact to develop future collaboration.

We would like to express our sincere gratitude to the Keynote speakers, International Scientific Committee, Steering Committee, and Organizing Committee for their huge efforts to make this conference successful.

Thank you all for your support and attendance at InCITE 2021. Please enjoy the conference!

Asst. Prof. Doddy Sutrisna, Ph.D.
Steering Committee

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 010002-1–010002-2; https://doi.org/10.1063/12.0009708
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

WELCOME FROM InCITE 2021 ORGANIZING COMMITTEE

Welcome to InCITE 2021! The third bi-annual international conference on engineering domain conducted by the Faculty of Engineering, The University of Surabaya (UBAYA). Due to the COVID-19 pandemic, InCITE 2021 is held as an online conference. Online conference opens the opportunity for many researchers around the globe to share their findings and learn from other global researchers with less restrictions.

InCITE 2021 invites three keynote speakers, well reputable global researchers in their research domain from Australia and Taiwan. Following each keynote session are two presentation sessions run in parallel.

This year, we received 66 papers submitted by researchers from four distinct countries (i.e., first author's country of origin): Indonesia, Australia, Taiwan, and Kazakhstan.

We employed a double-blind review to ensure a high standard and a minimum level of bias in the reviewing processes. This resulted in 56% of the submissions were accepted and will be published to the AIP Conference Proceedings.

Authors of all accepted papers are to disseminate their findings during InCITE 2021 conference between 25 to 26 of August 2021. This presents a great opportunity for everyone, including the researchers, to discuss and further improve current achievements.

We thank all keynote speakers, presenters, and reviewers/scientific committees for the generous supports. We thank the University of Surabaya, the Faculty of Engineering UBAYA, and all InCITE 2021 committees that enable InCITE 2021.

We wish you a very pleasant and rich conference experience in InCITE 2021 and look forward to seeing you again on InCITE 2023! Thank you.

Yours sincerely,
Asst. Prof. Dr. Jimmy
InCITE 2021 Organizing Committee

AIP Conference Proceedings

Table of Contents

INTERNATIONAL CONFERENCE ON INFORMATICS, TECHNOLOGY, AND ENGINEERING 2021 (InCITE 2021): Leveraging Smart Engineering

PREV

Conference date: 25–26 August 2021
Location: Surabaya, Indonesia
ISBN: 978-0-7354-4180-4
Editors: Joniarto Parung, Nemuel Daniel Pah, Putu Doddy Sutrisna and Marcellinus Ferdinand Suciadi Suciadi
Volume number: 2470
Published: Apr 25, 2022

DISPLAY : 20 50 100 all

PRELIMINARY

BROWSE VOLUMES

Free . April 2022

Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

AIP Conference Proceedings **2470**, 010001 (2022); https://doi.org/10.1063/12.0009021

Free . April 2022

Committees: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

AIP Conference Proceedings **2470**, 010002 (2022); https://doi.org/10.1063/12.0009708

SMART DESIGN AND INNOVATION

Free . April 2022

How energy efficiency, smart factory, and mass personalization affect companies in industry 4.0

Yu-Chung Tsao, Felix Arril Simbara Barus and Chien-Wei Ho

AIP Conference Proceedings **2470**, 020001 (2022); https://doi.org/10.1063/5.0080526

SHOW ABSTRACT

BROWSE VOLUMES

Free . April 2022

Kansei engineering and product-service systems (KEPSS) integrative framework for customer-centered experience

Markus Hartono

AIP Conference Proceedings **2470**, 020002 (2022); https://doi.org/10.1063/5.0080177

SHOW ABSTRACT

Free . April 2022

The role of supply chain performance factors to determine the firm performance

Yoni Kristiawan, Edi Purwanto and Rustono Farady Marta

AIP Conference Proceedings **2470**, 020003 (2022); https://doi.org/10.1063/5.0080235

SHOW ABSTRACT

Free . April 2022

Real-time monitoring design for make-to-order industry

Senapati C. Adisasmito, Pradipta Deffinika Pamungkas and Anas Ma'ruf

AIP Conference Proceedings **2470**, 020004 (2022); https://doi.org/10.1063/5.0080747

SHOW ABSTRACT

Free . April 2022

BROWSE VOLUMES

Mapping and Kaizen approaches: A case study in the musical instrument industry

Indra Setiawan and Hernadewita

AIP Conference Proceedings **2470**, 020005 (2022); https://doi.org/10.1063/5.0080159

SHOW ABSTRACT

Free . April 2022

Simple heuristics for scheduling apache airflow: A case study at PT. X

Hans Tanto, Indriati N. Bisono and Hanijanto Soewandi

AIP Conference Proceedings **2470**, 020006 (2022); https://doi.org/10.1063/5.0081042

SHOW ABSTRACT

Free . April 2022

Quality improvement with PDCA approach and design of experiment method in single socks industry in Indonesia

Hibarkah Kurnia, Choesnul Jaqin and Humiras Hardi Purba

AIP Conference Proceedings **2470**, 020007 (2022); https://doi.org/10.1063/5.0080179

SHOW ABSTRACT

Free . April 2022

BROWSE VOLUMES

company performance: Evidence from the Indonesian manufacturing industry

Esti Dwi Rinawiyanti, Xueli Huang and Sharif As-Saber

AIP Conference Proceedings **2470**, 020008 (2022); https://doi.org/10.1063/5.0080731

SHOW ABSTRACT

Free . April 2022

Community mobility during Covid-19 pandemic and tourism performance: Data mining approach

Gunawan

AIP Conference Proceedings **2470**, 020009 (2022); https://doi.org/10.1063/5.0080170

SHOW ABSTRACT

Free . April 2022

Inventory system improvement for poultry

Noverta Brilly Leksana Putra, Indri Hapsari and Dina Natalia Prayogo

AIP Conference Proceedings **2470**, 020010 (2022); https://doi.org/10.1063/5.0080160

SHOW ABSTRACT

Free . April 2022

Users' perception of digital prototypes in Indonesian fashion

BROWSE VOLUMES

Christabel Parung and Prayogo Waluyo

AIP Conference Proceedings **2470**, 020011 (2022); https://doi.org/10.1063/5.0080178

SHOW ABSTRACT

SMART MANUFACTURING AND PROCESSES

Free . April 2022

Dynamic study of batch milk cooling process at KUD SAE Pujon

Rudy Agustriyanto, Puguh Setyopratomo and Endang Srihari

AIP Conference Proceedings **2470**, 030001 (2022); https://doi.org/10.1063/5.0080176

SHOW ABSTRACT

Free . April 2022

Recent advances and application of Selective Laser Melting (SLM) technology in the aerospace industry

Anel Yerubayeva, Essam Shehab and Md. Hazrat Ali

AIP Conference Proceedings **2470**, 030002 (2022); https://doi.org/10.1063/5.0080173

SHOW ABSTRACT

POWER SYSTEM AND SMART ENERGY MANAGEMENT

BROWSE VOLUMES

Free . April 2022

Immobilization of xylanase on acid pretreatment bentonite as green biocatalyst

Lieke Riadi, Ruth Chrisnasari, Joshua Kristanto, Cahaya Caesar Bigravida and Meyta Sanoe

AIP Conference Proceedings **2470**, 040001 (2022); https://doi.org/10.1063/5.0080318

SHOW ABSTRACT

Free . April 2022

Rheological behavior and antioxidant activity of carrageenan extracted from Green seaweed (*Eucheuma cottonii*) using alkaline solution at low temperature

Puguh Setyopratomo and Lanny Sapei

AIP Conference Proceedings **2470**, 040002 (2022); https://doi.org/10.1063/5.0080286

SHOW ABSTRACT

Free . April 2022

Simulation of impact azimuth angle on specific energy output of a fixed mounting rooftop PV system in Jakarta, Indonesia

Elieser Tarigan

AIP Conference Proceedings **2470**, 040003 (2022); https://doi.org/10.1063/5.0080228

SHOW ABSTRACT

BROWSE VOLUMES

Free . April 2022

Rice husk ash for the stabilization of the outer interfacial layer of W/O/W double emulsion

Lanny Sapei, Rudy Agustriyanto, Endang Wahyu Fitriani, Zerravym Levy and Cindy Sumampouw

AIP Conference Proceedings **2470**, 040004 (2022); https://doi.org/10.1063/5.0080285

SHOW ABSTRACT

Free . April 2022

The effects of material to solvent ratio on the performances of natural dyes extraction

Putu Doddy Sutrisna, Meyta Sanoe, Rifando Gogo Adiyaksa, Kristina Wahyu Agustine, Hadiyatni Rita Priyantini, Prayogo Widyastoto Waluyo and I. Made Ronyastra

AIP Conference Proceedings **2470**, 040005 (2022); https://doi.org/10.1063/5.0080171

SHOW ABSTRACT

BROWSE VOLUMES

Free . April 2022

Carbon based sulfonated catalyst as an environment friendly material: A review

Putu Padmareka Deandra, Herry Santoso and Judy Retti B. Witono

AIP Conference Proceedings **2470**, 040006 (2022); https://doi.org/10.1063/5.0080728

SHOW ABSTRACT

Free . April 2022

A performance study of magnetite-lignin composites as photothermal materials in solar steam generation system

Marta Devega Yuharma, W. Widiyastuti, Ni Made Intan Putri Suari and Heru Setyawan

AIP Conference Proceedings **2470**, 040007 (2022); https://doi.org/10.1063/5.0080989

SHOW ABSTRACT

Free . April 2022

Organosolv lignin from coconut coir as potential biomaterials for sunscreen

Diana Novita Sari, Mahardika F. Rois, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040008 (2022); https://doi.org/10.1063/5.0080768

SHOW ABSTRACT

BROWSE VOLUMES

Free . April 2022

One-step electrochemical synthesis of silica-coated magnetite nanofluids

Delyana Ratnasari, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040009 (2022); https://doi.org/10.1063/5.0080982

SHOW ABSTRACT

Free . April 2022

Synthesis and characterization of chitosan-Allium sativum film

Emma Savitri, Maria Irine Tjahayani and Anastasia Apriyani Ngene Say

AIP Conference Proceedings **2470**, 040010 (2022); https://doi.org/10.1063/5.0080203

SHOW ABSTRACT

Free . April 2022

The effect of sodium lauryl sulfate on silica nanofluid stabilization using microbubble method

Ratri Sekaringgalih, I. Made Joni, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040011 (2022); https://doi.org/10.1063/5.0080985

SHOW ABSTRACT

Free . April 2022

BROWSE VOLUMES

Rezha Carina Rachmady, Ari Rahman and Wahyu Kunto Wibowo

AIP Conference Proceedings **2470**, 040012 (2022); https://doi.org/10.1063/5.0080740

SHOW ABSTRACT

THE ROLE OF IT IN INNOVATION ENHANCEMENT

Free . April 2022

Hate speech content detection system on Twitter using K-nearest neighbor method

Vincentius Riandaru Prasetyo and Anton Hendrik Samudra

AIP Conference Proceedings **2470**, 050001 (2022); https://doi.org/10.1063/5.0080185

SHOW ABSTRACT

Free . April 2022

School finder, intelligent recommendation system for elementary school selection

Susana Limanto, Endah Asmawati and Yusuf Wira Kencana Putra

AIP Conference Proceedings **2470**, 050002 (2022); https://doi.org/10.1063/5.0080461

SHOW ABSTRACT

BROWSE VOLUMES

Designing a recommender system based on the application of decision tree algorithm in data mining with KNIME (for recommending the topic of undergraduate's thesis)

Yenny Sari, Vincentius Riandaru Prasetyo and Kevin Liyansah

AIP Conference Proceedings **2470**, 050003 (2022); https://doi.org/10.1063/5.0081214

SHOW ABSTRACT

Free . April 2022

Weather image classification using convolutional neural network with transfer learning

Mohammad Farid Naufal and Selvia Ferdiana Kusuma

AIP Conference Proceedings **2470**, 050004 (2022); https://doi.org/10.1063/5.0080195

SHOW ABSTRACT

Free . April 2022

Student performance prediction in higher education: A comprehensive review

Ellysa Tjandra, Sri Suning Kusumawardani and Ridi Ferdiana

AIP Conference Proceedings **2470**, 050005 (2022); https://doi.org/10.1063/5.0080187

SHOW ABSTRACT

BROWSE VOLUMES

Sentiment analysis on feedback of higher education teaching conduct: An empirical evaluation of methods

Jimmy and Vincentius Riandaru Prasetyo

AIP Conference Proceedings **2470**, 050006 (2022); https://doi.org/10.1063/5.0080182

SHOW ABSTRACT

Free . April 2022

Implementing directed pairwise judgement approach in web-based AHP survey application to reduce inconsistency ratio

Daniel Hary Prasetyo

AIP Conference Proceedings **2470**, 050007 (2022); https://doi.org/10.1063/5.0080169

SHOW ABSTRACT

Free . April 2022

Leaf geometric properties measurement using computer vision system based on camera parameters

Joko Siswantoro, Ida Bagus Made Artadana and Muhammad Z. F. N. Siswantoro

AIP Conference Proceedings **2470**, 050008 (2022); https://doi.org/10.1063/5.0080190

SHOW ABSTRACT

Free . April 2022

BROWSE VOLUMES

blockchain technology

Alexander Yohan, Nai-Wei Lo and Kevin Valentino

AIP Conference Proceedings **2470**, 050009 (2022); https://doi.org/10.1063/5.0080266

SHOW ABSTRACT

Free . April 2022

Implementation of behavior tree for creating an in-game cut-scene

Delta Ardy Prima

AIP Conference Proceedings **2470**, 050010 (2022); https://doi.org/10.1063/5.0080233

SHOW ABSTRACT

Free . April 2022

The online attendance system models for educational institutions

Andre and Marcellinus Ferdinand Suciadi

AIP Conference Proceedings **2470**, 050011 (2022); https://doi.org/10.1063/5.0080180

SHOW ABSTRACT

Free . April 2022

Genetic algorithm with adaptive diversification and

BROWSE VOLUMES

Eric Wibisono, Iris Martin and Dina Natalia Prayogo

AIP Conference Proceedings **2470**, 050012 (2022); https://doi.org/10.1063/5.0080197

SHOW ABSTRACT

Resources

AUTHOR

LIBRARIAN

ADVERTISER

General Information

ABOUT

CONTACT

HELP

PRIVACY POLICY

TERMS OF USE

FOLLOW AIP PUBLISHING:

BROWSE VOLUMES

Scitation

BROWSE VOLUMES

Designing a recommender system based on the application of decision tree algorithm in data mining with KNIME (for recommending the topic of undergraduate's thesis)

Cite as: AIP Conference Proceedings **2470**, 050003 (2022); https://doi.org/10.1063/5.0081214
Published Online: 25 April 2022

Yenny Sari, Vincentius Riandaru Prasetyo and Kevin Liyansah

Designing a Recommender System Based on the Application of Decision Tree Algorithm in Data Mining with KNIME (For Recommending the Topic of Undergraduate's Thesis)

Yenny Sari[1, a)], Vincentius Riandaru Prasetyo[2], and Kevin Liyansah[1]

[1]Department of Industrial Engineering, Faculty of Engineering, University of Surabaya, Surabaya 60293, Indonesia.
[2]Department of Informatics Engineering, University of Surabaya, Surabaya 60293, Indonesia.

[a)] Corresponding author: ysari@staff.ubaya.ac.id

Abstract. Data mining has many algorithms; one of the most frequently used is the decision tree algorithm. This article described the result of the research of a data mining approach that used the decision tree algorithm; it aimed to create and provide a recommendation system for university students to choose and decide the topic for their final projects. The final project (so-called as thesis) is one of the graduation requirements for undergraduate students; many university students are confused in choosing and deciding the right topic for their thesis, and this results in not completing it in the required time. The impact of overdue thesis is the extension of their period of studies. In this study, data mining was carried out based on 9 attribute data sets which were reduced from 21 derived factors. The data processing of the research has used the free software so-called KNIME. It was started with attribute correlation testing; it was continued by formulating KNIME workflow and resulted in the decision tree. Then, the tree diagrams were translated into the 65 IF-THEN rules. Based on these 65 IF-THEN rules, this study developed a recommender system using C# language; the recommendation system was tested on 45 students which gave its accuracy more than 70%.

Keywords: data mining, decision tree algorithm, recommender system, KNIME workflow

INTRODUCTION

A recommender system can be defined as a program that give recommendation of the most suitable option for users based of the defined decision factors; its ability to "guess" users' preferences or interests and generate personalized recommendation [1, 2]. Many recommender systems have been developed and implemented, the applications include recommending music/television programs, books, learning materials, etc.

The "final project" (or so-called as thesis) is one of the requirements for undergraduate student graduation. The object of the study is the department in Engineering Faculty where we work at. The thesis is one of the undergraduate subjects which need to be accomplished within one semester. Before working on their thesis, students must formulate the research proposal in advance. The selection of the right topic, the smooth data collection, the mastery of the research methodology and the ability to use the analytical tools are a bunch of factors that may cause the students accomplishing their thesis on time.

However, it is often found that many students feel anxious because it is difficult to determine the right topic according to their preferences and abilities. As a result, the thesis cannot be completed within the required time or even the students need to change the topic in the middle of the work which can cause the extension of their period of study. So, choosing the right topic plays an important role for students to complete the thesis on time. Thus, a recommender system becomes necessary so that it can give the recommend topic of the thesis for students so that they are expected to be able to minimize the obstacles because of incorrect topic choice. The recommendation

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 050003-1–050003-8; https://doi.org/10.1063/5.0081214
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

system will be developed using the result of data mining approach i.e. applying the decision tree algorithm. Decision tree [3] is a supervised algorithm that can be used in various types of problems related to the classification approach in data mining. Not only because this algorithm is the most frequently used, but the result of the decision tree can also contribute to the logic of human thinking in developing the recommender system.

The development of similar recommender systems has been found in several previous studies. Kumar and Ravi [4] applied data mining methods to predict customer credit card churn. Osmanbegovic and Suljic [5] used three supervised data mining algorithms to predict the graduation of a course and the performance of its learning methods. Kaunang [6] used the decision tree algorithm to analyze the poverty level based on the Human Development Index. The study conducted by Pagnotta and Hossain [7] was aimed to predict the level of alcohol consumption in students mining from more than 1,000 data, its prediction was based on 25 defined attributes; it also convinced that the factors that influenced students to consume alcohol are being male, affecting by social influence, and having a lot of free time. The decision tree algorithm used by Supriyanti et al. [8] was to determine the selection of the right student concentration; the attributes of the study covered the origin of the previous school, gender, achievement index, selected concentration, and length of study.

RESEARCH METHODOLOGIES

The research methodology used in this study was briefly described in the framework of Input-Process-Output-Objective, as shown in Fig. 1. Each stage of the methodology will be explained more in detail in the following section (see Section 4, the section described how the experiment and analysis of data mining were structured and executed as well as its results). The stages were divided into: (a) Input: attribute data set, (b) Process & Output: Data mining using KNIME workflow and Decision Tree, and (c) Objective: the developed recommender system.

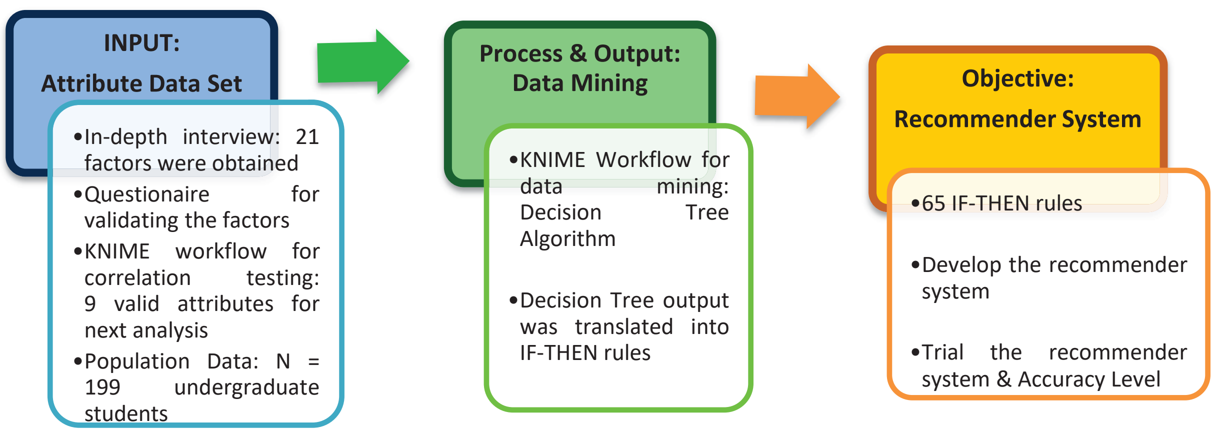


FIGURE 1. Designing a Recommender System: a framework of Input, Process, Output, and Objective

EXPERIMENTS & RESULTS

Attribute Data Set

As input to the data mining process, the population data of 199 undergraduate students in the selected department of Engineering Faculty was used in data mining processing and analysis. Initially, there were 21 factors that had been defined for the study, after the validation process, they became 11 attributes. These number of attributes had been reduced into 9 attributes due to the correlation testing, and the attribute data set were described in Table 1.

TABLE 1. Attribute Data Set

No	Attribute	Type of Data	Domains
1	Areas of Specialization	Binary	***2 domains:*** Sustainable Enterprise Systems and Sustainable Supply Chain Systems
2	Origin of Topic	String	***3 domains:*** student, lecturer, organization (or the research object)
3	Origin of Research Idea	Binary	***2 domains:*** student, lecturer

No	Attribute	Type of Data	Domains
4	Location of the organization (as the research object)	String	***4 domains:*** in the city (Surabaya), out of town (East Java), out of province, none
5	Semester when doing thesis	Numeric	7, 8, 9, 10
6	GPA	Numeric	≤ 4
7	Level of difficulty (of the topic)	String	***3 domains:*** Very challenging, challenging, less challenging
8	Expected 1st Advisor	String	***12 domains:*** List of possible advisors/supervisors
9	Expected 2nd Advisor	String	***17 domains:*** List of possible advisors/supervisors

The process of defining the attribute data set was explained as follows. Initially, a survey was conducted to identify what factors that influence students in determining the topic of the thesis. The survey was conducted through in-depth interviews among 5 lecturers, 7 alumni, and 16 students; and a total of 21 factors were obtained. To validate these factors, a survey using the questionnaire was distributed to 32 students; the questionnaire used the Likert scale of 1-4 (1: not relevant, 4: very relevant). Only the factors which have an average value of more than 2.5 will be considered as the influence factors. There were 11 attributes (out of 21 initial factors) had been tested as valid ones.

These attributes were again tested using correlation testing; it was done by using KNIME [9] workflow, as shown in Fig. 2. There are three nodes used in KNIME workflow diagram: (i) Node *Read Data*, (ii) Node *Number To String*, and (iii) Node *Linear Correlation*. The correlation test showed that attributes that had a value of linear correlation less than 0.35 will not be used as a data set [7, 9]. Finally, there were only 9 attributes which have correlation value greater than 0.35, so that these attributes were used as the significant attribute data set.

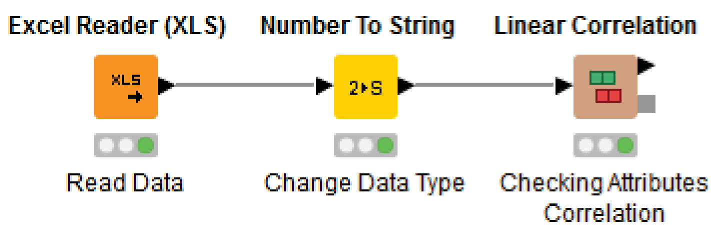


FIGURE 2. KNIME Workflow for Correlation Testing

Data Mining: KNIME Workflow and Decision Tree

Data mining was done by using free software so-called KNIME. KNIME [9] stands for *Konstanz Information Miner*; the software was created by the German University of Konstanz which can be accessed for free and downloaded via the website (https://www.knime.com/). The use of this software is relatively easy because users only need to drag and drop from the nodes that are already available in the library, so there is no need to use coding. In addition, Naik and Samant (2016) mentioned the accuracy of KNIME is the most superior regarding the usage of the decision tree algorithm when compared to other software. The accuracy of the decision tree algorithm is 95.37% while K-Nearest Neighbor is 86.58% and Naïve Bayes is 72.56% [10].

There are three steps of applying the decision tree algorithm: (a) determine the attributes to become the root of the decision tree based on the highest gain value, (b) create a branch for each attribute based on the calculation of the entropy and gain values, and (c) repeat the branch creation process until there are no remaining attributes or there are no new branches. The formula of entropy and gain value [3] are defined as follows:

$$Entropy(S_i) = \sum_{i=1}^{n} -pi * log_2\, pi \text{ i)} \quad (1)$$

$$Gain\ (S, A) = S - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * S_i \quad (2)$$

The application of the decision tree algorithm using the KNIME workflow was shown in Fig. 3. There are three nodes used in the workflow diagram: (i) the *Read Data* node to read tabulated data that has been compiled previously, (ii) the Column Filter node to filter out attributes that are not used in the data set, while (iii) the Decision Tree Learner node to serve the results of tree diagrams formed from data. Because the department has implemented the new curriculum, there are only 199 graduated students, and data processing used 100% of the population data (N = 199). This historical data was used as the basis for creating the output or so-called as data training.

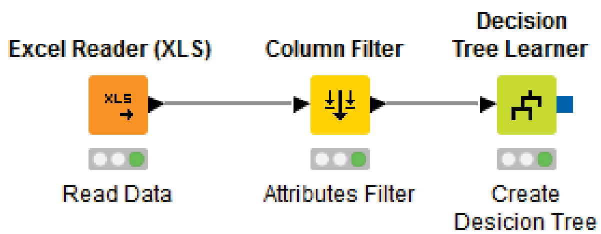


FIGURE 3. KNIME Workflow for Diagram Tree Algorithm

The Recommender System

The output of the diagram tree algorithm using KNIME workflow (as shown in Fig. 4) had been translated into the 65 IF-THEN rules. Because the tree diagram produces multi-level branches, only a few parts are shown in this article. Figure 5 showed how part of the diagram tree was translated into several IF-THEN rules. These IF-THEN rules had provided human logic and were used to develop the recommender system.

Several IF-THEN Rules were translated from the output of Decision Tree in Fig. 4 are:
Blue Circle : **if** *GPA<=2,6235* **then** *Marketing*
Green Circle : **if** *GPA >2,6235* **and** *1st Advisor="Eric Wibisono, S.T., M.Eng., Ph.D."* **then** *Lean*
Red Circle : **if** *GPA >2,6235* **and** *1st Advisor="Prof. Ir.Joniarto Parung, M.M.B.A.T., Ph.D."* **and** *2nd Advisor="Dina Natalia Prayogo, S.T., M.Sc."* **then** *SCM*
(SCM stands for Supply Chain Management, it is one of the thesis' topics)

The recommender system was developed by using C# language. The design of the recommender system will be made simple by prioritizing the easy-to-use function. The system uses a dropdown list, so the users can choose from the list of options (the option was defined based on the domains of its attribute data set). The display was designed by using a bright color background, black color font, and the type of font (Times New Roman) was applied. Figure 5 showed the interface and the mechanism of the recommender system that was developed based on the existing 65 IF-THEN rules.

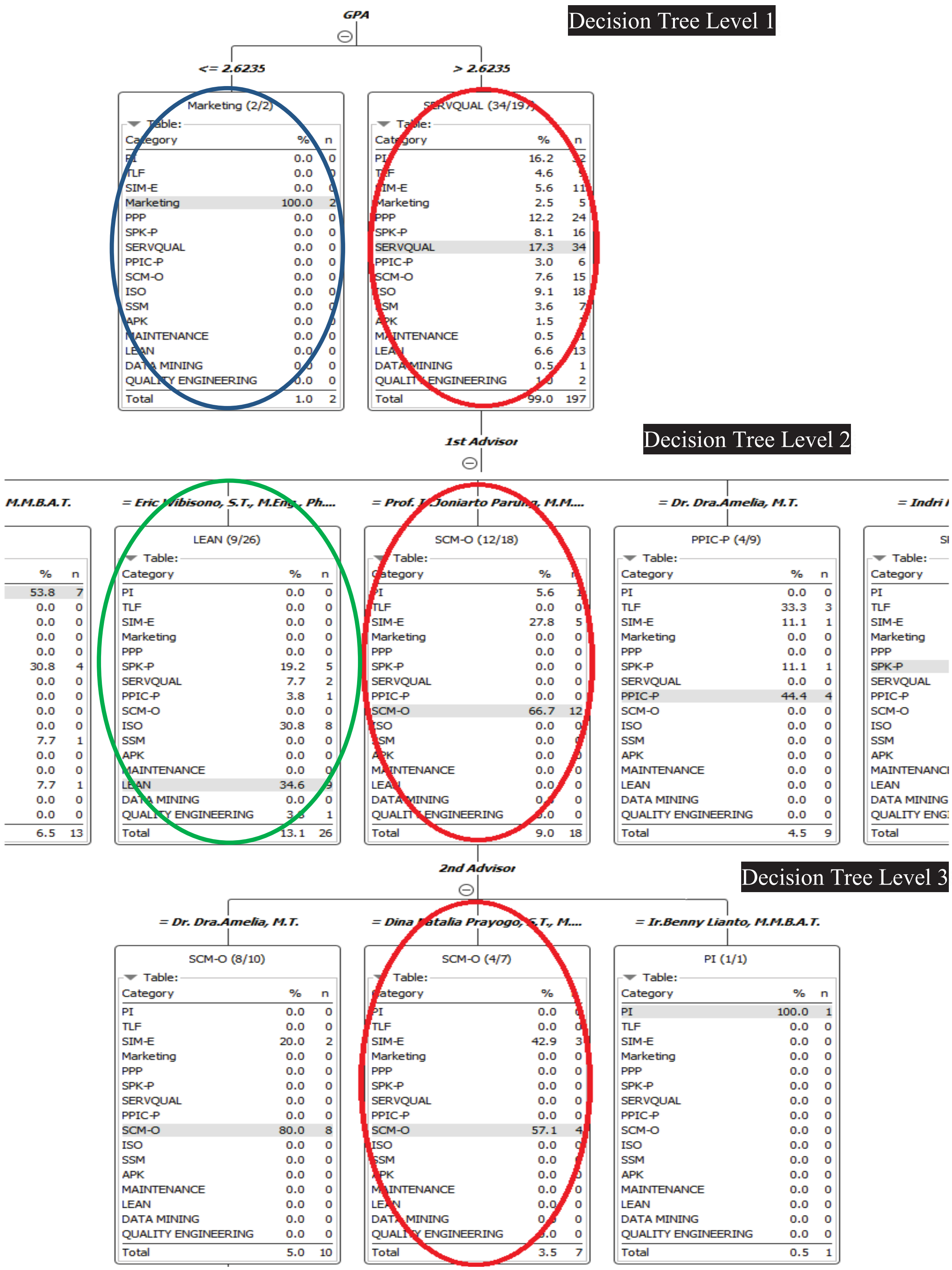
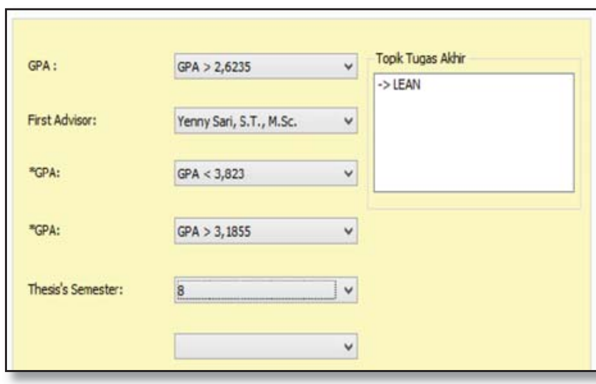


FIGURE 4. Output of Data Mining: Multi-level branches of Diagram Tree

As shown in Fig. 5, on the left side (Code: A), the user needs to select the options in the dropdown lists to indicate their preferences or related personalized data (these are the influenced factors that have been selected as input data to determine the final project's topic). On the right side (Code: B), the recommender system will suggest the relevant topic based on the result of mining historical data using the application of decision tree algorithm.

WHAT is the TOPIC for your FINAL PROJECT?

(a)



(b)

FIGURE 5. Easy-to-use recommender system

(a) To validate the result, crosscheck it with the IF-THEN rule: **if** *GPA >2,6235* **and** *1stAdvisor="Prof. Ir.Joniarto Parung, M.M.B.A.T., Ph.D."* **and** *2ndAdvisor="Dina Natalia Prayogo, S.T., M.Sc."* **then** *SCM*

(b) To validate the result, crosscheck it with the IF-THEN rule: **if** *GPA>2,6235* **and** *1stAdvisor= "Yenny Sari, S.T., M.Sc."* **and** *GPA<=3,823* **and** *GPA>3,1855* **and** *Semester=8* **then** *LEAN*

According to the decision tree level 1 (in Fig. 4), the attribute *"GPA"* appears as the first basic attribute for the recommender system based on the calculation of the highest entropy or gain value. If there is any further levels of decision tree (i.e. level 2, 3, etc.), then the recommendation of the final project's topic requires other attribute inputs before giving the result. In contrast, if there is no further level of the decision tree, then the recommendation can be precisely given. As an illustration, as shown in Fig.4, the first attribute required as input is the student's GPA; if *"GPA <= 2.6235"*, then a direct topic recommendation can be given (i.e. topic of *"Marketing"*). However, when the input attribute is *"GPA> 2.6235"*, then a level-two input attribute was needed, namely *"First Advisor"*. For the level-two input attribute is *"Prof. Ir. Joniarto Parung, M.M.B.A.T., Ph.D."*, as shown in Fig. 4 and Fig. 5(a), the recommendation system requires next input attribute namely *"Second Advisor"*.

For some decision trees that have multiple levels, the recommendation system could only run if all input attributes were filled with data, for example in Fig. 5(b), "GPA" was the first basic attribute (with the option of "GPA > 2.6235") and "First Advisor" was the level-two input attribute (with the option of "Yenny Sari, S.T., M.Sc."). However, at level-three input attribute, it required more specified information about "GPA", undergraduate students with GPA between the value of 3.1855 and 3.823 are recommended to the topic of "LEAN" while GPA between 2.6235 and 3.1855 or GPA above 3.823 were suggesting to different topics.

When evaluating the test results in Table 2, the accuracy level from the 1st group of data estimation was 86.67% while the 2nd group was 40%. From both groups of data, a total of 45 data, the accuracy of the recommender level was 71%. The accuracy level of the recommender system can be increased if more historical data is used, and the number of topics can be reduced into either smaller or more concise options. To be noted, during the evaluation of the second group, students who were starting to work on their thesis are faced with the suggestion of several new research topics from the lecturers so that the new topics have not been accommodated in the recommendation system which was developed using the classification approach upon the historical data.

TABLE 2. Trial of the Recommender System and Prediction of Accuracy Level

No	Data ID/ Student ID	The real topic of final project (Reality)	Conformity between Recommendations versus Reality	No	Data ID/ Student ID	The real topic of final project (Reality)	Conformity between Recommendations versus Reality
1st Group: Graduates' Historical Data (Class 2014 & 2015)				**1st Group: Graduates' Historical Data (Class 2014 & 2015)**			
1	3	SIM-E	**Yes**	26	170	SERVQUAL	**Yes**
2	5	Marketing	**Yes**	27	181	SCM-O	**Yes**
3	16	SPK-P	**Yes**	28	189	SPK-P	**Yes**
4	19	PPP	**Yes**	29	196	ISO	**Yes**
5	24	PI	**Yes**	30	199	ISO	**Yes**
6	28	PPP	**Yes**	**2nd Group: Current Students' Ongoing Final Project (Class 2016)**			
7	32	PPIC-P	**No**	1	YA	PPIC	**No**
8	35	ISO	**Yes**	2	AP	PPIC	**No**
9	39	SPK-P	**No**	3	VE	PPP	**No**
10	40	Marketing	**Yes**	4	MO	PI	**Yes**
11	44	SSM	**Yes**	5	LO	TLF	**No**
12	57	SERVQUAL	**Yes**	6	FE	Lean	**Yes**
13	62	ISO	**Yes**	7	ER	SPK	**No**
14	74	PI	**Yes**	8	OL	Lean	**Yes**
15	81	LEAN	**Yes**	9	AL	Servqual	**Yes**
16	83	SCM-O	**Yes**	10	YO	SPK	**No**
17	92	SCM-O	**Yes**	11	ANT	PI	**Yes**
18	101	SIM-E	**Yes**	12	BO	DTS	**No**
19	108	SIM-E	**Yes**	13	JO	Servqual	**Yes**
20	116	SPK-P	**No**	14	ANG	Marketing/SSM	**No**
21	126	TLF	**No**	15	WI	DTS	**No**
22	139	ISO	**Yes**	*Summary*		*Yes*	*No*
23	146	LEAN	**Yes**	1st Group		26/30 (87%)	4/30 (33%)
24	153	TLF	**Yes**	2nd Group		6/15 (40%)	9/15 (60%)
25	162	SERVQUAL	**Yes**	**Total**		**32/45 (71%)**	**13/45 (29%)**

CONCLUSIONS

In creating the recommender system to help students decide their final project's topic, it is necessary to identify the factors that influence them to choose a certain topic. A combination of in-depth interviews, questionnaire surveys, and correlation tests between factors was conducted to select the attributes that had a significant role in determining the research topic in their research proposals. Nine influenced attributes had been set as the decision criteria in the data mining, namely GPA, 1st and 2nd Advisor, Areas of Specialization, Origin of Topic, Origin of Research Idea, Location of the organization, Semester, and Level of difficulty.

The data mining process was conducted using the KNIME software and run on 199 graduate data, this data is population data from Class 2014 and 2015 that owned by the department due to the implementation of the new curriculum. The thing to be noted is that the more data used for training data, the better the result can be. Because more data for input can increase the level of accuracy of the developed recommendation system. The use of a decision tree algorithm (based on a classification approach in data mining) had produced 65 IF-THEN rules and these were the basis for developing the recommendation system. The recommender system was created in an easy-to-use form and had passed the trial phase with an accuracy of over 70%.

For the domains of the objective attribute, it is recommended not to have too many options/domains, the current condition was there were more than 40 choices (domains) of attribute "Final Project's Topic". This number of destination domains can cause a low level of accuracy to be predicted. It is also recommended that topics should be clustered properly first so that the destination domain can be reduced. In addition, it is necessary to give a choice of "new topics" to trigger students to research new topics.

REFERENCES

1. Bobadilla J., Ortega F., Hernando A., Gutiérrez A., "Recommender systems survey," Knowledge-based systems **46**, 109–132 (2013).
2. Park, D. H., Kim, H. K., Choi, I. Y. and Kim, J. K., "A literature review and classification of recommender systems research," Expert systems with applications **39(11)**, 10059–10072 (2012).
3. Bramer, M., *Principles of data mining* (Springer, London, England, 2007).
4. Anil Kumar D and Ravi V., "Predicting credit card customer churn in banks using data mining," International Journal of Data Analysis Techniques and Strategies **1(1)**, 4–28 (2008).
5. Osmanbegovic E and Suljic M., "Data mining approach for predicting student performance," Economic Review: Journal of Economics and Business **10(1)**, 3–12 (2012).
6. Kaunang, F.J., "Application of the J48 Decision Tree Algorithm for Analysis the Poverty Levels in Indonesia," CogITo Smart Journal **4(2)**, 348–357 (2019).
7. Pagnotta, F. and Amran, H.M., "Using data mining to predict secondary school student alcohol consumption," Project Report, Department of Computer Science, University of Camerino, 2016.
8. Supriyanti, W., Kusrini, K., Amborowati, A., "Performance Comparison between C4.5 and Naive Bayes for the Accuracy of Selecting Student Concentration," Jurnal Informa **1(3)** 61–67 (2016).
9. Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K. and Wiswedel, B., "KNIME: The Konstanz Information Miner Studies in Classification," Data Analysis and Knowledge (2007).
10. Naik, A. and Samant, L., "Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and KNIME," Procedia Computer Science **85**, 662–668 (2016).