Analysis and Design of Decision Support System Dashboard for Predicting Student Graduation Time

Satrio Wibowo¹
Information System Study Program
Telkom University
Bandung, Indonesia
satrio.bws@gmail.com

Rachmadita Andreswari²
Information System Study Program
Telkom University
Bandung, Indonesia
andreswari@gmail.com

Muhammad Azani Hasibuan³
Information System Study Program
Telkom University
Bandung, Indonesia
muhammad.azani@gmail.com

Abstract—Information Systems is one of the existing study program at Telkom University that has produced many graduates since it was established in 2008. However, not all graduates produced successfully completed the study period during the four years of normal study. The percentage of graduates on time has some decline between the target and the achievement of the study program. From academic year 2014/2015 to 2016/2017 decrease annually about 1% every year, which is it becomes problems for the credibility and existence of study program and also for academic planners who may have an impact on accreditation assessment process of the study program when it is audited. One of the efforts that can be done by the study program to increase the students on time graduation rate is by making decision support system dashboard that giving early warning to the lecturer or the head of the study program if there are students who are predicted not to graduate on time. By using the C4.5 algorithm to perform the data analysis by looking at the causes of student’s graduation time and pureshape methodology to perform dashboard development method. The result of this study is a prototype of decision support system dashboard, because there are lack of analysis in decision making and the dashboard only showing information and temporary prediction. The data model that used on this research is labeling data that has been processed using C4.5 algorithm and data that has been through data cleansing process using Pentaho Data Integration. This prototype is expected to be used as a reference base to support academic planners in order to make this application run with real time data.

Keywords—early warning, dashboard, decision support system, student’s graduation target

I. INTRODUCTION

The graduation rate of students is very important for an university especially for the study program, because it will affect the credibility and existence of study program. Information System is the one of study program at Telkom University that has 62% graduation rate on time. This certainly needs to be evaluated, as improvements in graduation time and attendance methodology to perform dashboard development method. The annual report of Information Systems Program shows that the on time graduates each year decreased by 1% from 2014/2015 until 2016/2017. The decrease is seen based on comparison of target and achievement of Information System program. But in 2017/2018 there was an increase of 0.2% with a percentage of 62% in this year. There are several factors that can affect student’s graduation time. Based on previous academic records, the researchers conducted an analysis to determine factors that may affect students in achieving on time graduation and design a business intelligence dashboard that used by head of study program and lecture of Information System as an early warning system of student graduation target. In this study, the dashboard is developed by using pureshape dashboard development method.

In order to identify the user requirement, define the key metric, and design the dashboard layout the researchers conducted an interview to the head of study program. The researcher using the previous research to collect data through academic data records of Information system student from 2009 to 2016. Based on the data of previous research, the criteria that affect student’s graduation on time are GPA score on 4th Semester, senior high school, the parent’s job, the parent's income, selection path, number of credits and SAT points. SAT (student activities transcript) points are obtained from organizational activities that students take as a condition for graduation. But, GPA (Grade-Point Average) is the most influential criterion in student on time graduation.
II. THEORY AND RELATED WORK

A. Business Intelligence (BI)

Business intelligence is a set of methodologies, processes, architectures, and technologies that convert raw data into meaningful and useful information used to foster strategic insight and strategic, tactical, and operational effective decision making [2]. The result is a wealth of information and leads to the development of new opportunities for the organization. If these opportunities are identified, and strategies implemented effectively, BI can give the organization a competitive advantage in the marketplace.

There are five dimensions of data quality in BI. First completeness that ensures availability of all required information. Second, consistency that ensures uniformity of information in a table. Third, validity that ensures the correctness of information. Fourth, accuracy that ensures data can be trusted and in accordance with objects in the real world. And the last is integrity that ensures that data are inter-related so as to reduce data duplication rates.

B. Decision Support System (DSS)

Decision support system is a computer-based system that helps users in assessing and selecting activities. DSS provides data storage and retrieval but to improve traditional access and retrieval function, DSS provides support to create reasoning-based models[3]. Decision support systems can present graphical information and may include artificial intelligence [4]. The decision support system was first used as a Marketing Decision Support System (MDSS), which is as a collection of data, systems, tools and technologies coordinated with software and hardware whereby organizations collect and interpret information from businesses and the environment and transform it into the basis of marketing actions [5].

According to [4] there are five models of decision support systems, that are communication model, data model, document model, knowledge model, and DSS driven model. The communication-based DSS model are targeted to internal teams including partners to produce a set of decisions to implement a strategy or solution. The most common technology used by this DSS model is web based or client server [4]. The data-based DSS model is emphasizes access and manipulation on the internal and external data in real time to fit the needs of the decision maker. This DSS model is used to run database queries or data warehouses to look for specific answers for specific purposes. Data-based DSS is usually created through the main frame system, client server link or via the web [4]. The document-based DSS model uses computer storage and processing technology to provide document retrieval and analysis. The used documents such as text documents, spreadsheets, and database records to generate decisions and manipulate information to improve strategy. The usual technology used to prepare such decision support systems is through web or client server systems [4]. The knowledge-based DSS model can recommend actions to managers to provide management advice in order to choose a product or service. Typical technologies commonly used to set up such this model can be client server systems, web, or software running on stand-alone PCs [4]. The DSS-driven model are emphasizes access, optimization manipulation and simulation on the financial side. This model are complex systems that help to analyze decisions or choose between different options. The DSS-driven model generally does not require a large database, because it uses only limited data and parameters to assist decision makers in analyzing a situation [4].

The model that used in this research is knowledge model because the aim of this research is to provide management advice in order to choose a decision. And the DSS knowledge-based also can recommend any actions to managers.

C. Dashboard

The term of dashboard refers to the display of one-page information used to monitor what is going on in some aspects of the business. The dashboard shows key data that users must use efficiently to monitor what's happening in their area of responsibility. Generally, dashboards are used to monitor information on a daily basis, but some jobs require a dashboard to be monitored in real time as monitored activity is currently happening so delay in responding cannot be tolerated. The dashboard can be very useful because it has two capabilities, namely the visual power and the dashboard way to integrate everything the user should remember into one screen even though it is very complex [9]. There are several types of dashboard, namely:

1. Strategic dashboards provide information about the system in general and identify potential benefits for the system. This dashboard does not provide detailed information about what to do to make decisions, but this dashboard helps users identify future benefits for further analysis.

2. Analytical dashboards provide users with information about what might happen in the future using old data compared to various variables. Analytical dashboard gives more detail information than strategic dashboard and operational dashboard.

3. Operational dashboard is used to monitor the current operational conditions. This kind of dashboard needs to be updated in a short time, even real time.

There are several things to avoid in designing the dashboard such as too complicated information, too many warnings, indistinguishable warning, the visual that too big and bassist, inappropriate visual interpretation, incompatibility between information and visual representation, expression of indirect action, and the context are not enough.

In addition, to design an effective dashboard required some principles and positive practices, including using something that can blink and sound to attract attention, encourage users to active for thinking about data, not just passive reactions to alarm, not over-automate the action to the point where people become uninvolved, provide simple and easy way to respond and an overview for the entire team, support projection for proactive responses, and match the mental model.
D. Pureshape

Pureshape is one of the dashboard development methodologies developed by pureshape vendors. This method aims to facilitate projects related to the measurement and management of organizational performance [6]. In the pureshape development method, the steps taken are planning and design that understanding of user needs and identifying what features will be used on the dashboard. After the user need are identified, the next step is review the system and data to applying a bottom-up approach, such as identification of data sources, how to access data, and the size of a data. After the user needs and the data sources are identified, then the prototype design step is perform to apply a top-down and bottom-up approach together to provide an overview of the final look of the dashboard. After the prototype has a final look, the prototype are reviewed by potential users and perform development according to user needs on prototype refinement step. After the prototype is match with the user needs, the dashboard is implemented on the organization and socialize use through training in release step. The last step is continuous improvement to keep the quality of the dashboard, the dashboard must be developed in various areas of the organization repeatedly.

E. C4.5 Algorithm

Algorithm C4.5 is one of the algorithm used to perform the process of data classification by using decision tree technique. Algorithm C4.5 is the development of ID3 algorithm which is also an algorithm to build a decision tree. The C4.5 algorithm recursively visits each decision node, chooses optimal branching, until no more branches are possible [7].

The steps in building the decision tree using the C4.5 algorithm are as follows [8]:
1. Prepare training dataset. Training datasets are usually obtained from pre-existing data history and have been grouped into specific classes.
2. Specifies the root attribute of the decision tree.
3. Create a branch for each attribute.
4. Divide the case in the branch.
5. Repeat the process for each branch until all the cases on the branch have the same class.
6. The decision tree partition process will stop when:
   a. All records in node N get the same class.
   b. No attributes or variables in the record are partitioned again.
   c. No record in the empty branch.

F. Related Work

Researchers used the C4.5 algorithm to obtain criteria based on the case studies being conducted. Before the researchers used this method there is some evidence that has been done by other researchers.

And the other result of research that conducted by Adolfo Crespo Marquez & Carol Blanchar [5], in their journal entitled "A Decision Support System (DSS) For Evaluating Operations Investments in High-Technology Business" shows that decision support systems using dynamic simulation system models greatly enhance the analysis of go-to-market strategies, as they can integrate customer knowledge with simulations to analyze tradeoffs of expenditure in service features, support, integration, channel incentives, pricing, and advertising.

By the experiments that conducted by other researchers above, the dashboard decision support system is helpful to improve the organization values. So it can be applied to predicting the students graduation time at Information System study program.

III. METHODOLOGY

The methodology that used on this research is pureshape methodology that have four steps, there are plan and design, system and data review, prototype refinement and release. At the plan and design stage, the KPI (key performance indicators) of the organization are identified to be improved, then identify the type of dashboard along with its user group and identify the meta information from KPI that has been made. Meta-information here in the form of factors that affect student graduation time. Then at the review system and data stage, the data source to be used are identified, then performs analysis to the data that has been obtained, then data cleansing are performs that can produce the expected information. After the KPI are identified and the data are “clean”, the design of the dashboard prototype can be made at prototype refinement stage. Then the prototype will be reviewed periodically and there will be renewal dashboard until the dashboard prototype is approved by the user. Once the dashboard prototype is approved by the user, the dashboard will be implemented to the user at release stage as shown in Fig. 2.

![Pureshape Methodology](image)

In this study, we used data records of students from 2009 to 2012 which is not the data of students who are “drop-out” or “resign” that obtained from the SISFO of Telkom...
University. The raw data will be divided into two parts as follows:

1. **Training set**: data are used in the system training process and consists of input data pair and target data. From the total data obtained, 70% - 80% part will be used as training data.
2. **Testing set**: data are used to test the ability of the system and also consists of input data pair and target data. Data testing used amounted to 20% - 30% of the data obtained.

The C4.5 algorithm is performed at “Identify the data sources” step on Fig. 2. The C4.5 algorithm train the set by calculating the gain ratio for all attributes that have been determined before, then select with the highest gain ratio to be the root node. And repeat the gain ratio calculation process and form a node that contains the attribute until all data has been included in the same class as shown in Fig 3.

**IV. RESULT AND ANALYSIS**

**A. Study Case Analysis**

In this case, the researcher uses academic record data (SISFO) Telkom University and questionnaires to expert judgment. Based on the academic record, the researcher divides into 7 criteria: GPA (Grade-Point Average) on 4th Semester, senior high school, the parent's job, the parent's income, selection path, number of credits and SAT (Student Activities Transcript) amount of points. Data processing is done based on academic record data from 2009 to 2012. There are 600 students who graduated on time, the percentage shown in the data of the annual report of Information Systems is 62%.

**B. Dashboard Design**

1. **Prediction page**

   The dashboard on Fig. 4 are consists of multiple bar chart that show the prediction of student’s graduation level and the target of on time student’s graduation. The yellow bar is describe the prediction of student’s graduation and the blue bar is describe the target of student’s graduation. The card view are used to inform the prediction of total student that may graduate on time, total student that may graduate not on time, total of the student that ever have a remedial on a subject, and the name of subject that most repeated. Beside the bar chart, there are the top 10 of subject that most repeated.

   ![Fig. 4. Prediction page](image)

When the prediction of students who graduated late card in the pressed, it will appear the name of student students who predicted to pass late with a grade 2 GPA of the student, as shown in Fig. 5.

![Fig. 5. Prediction of students who graduated late](image)

When the prediction of students who graduated on time card on tap, it will appear the name of student students are predicted to graduate on time along with a grade 2 GPA from the student, as in Fig. 6.

![Fig. 6. Prediction of students who graduated on time](image)
When the user presses one of the student's graduation prediction graphs, the user will go to the student's graduation prediction page that is grouped by class according to the student school year, as shown in Fig. 7.

Fig. 7. Prediction of student’s graduation that grouped by student school year

When the user presses the predicted graph of a student's graduation of a class, the student's graph will appear late and timely according to the pressed class, as shown in Fig. 8.

Fig. 8. Prediction of student’s graduation that grouped by class

2. Student’s Performance Page
   This page provides information on the performance of classes each semester grouped by force. This page is only accessible by caps. This graph contains data in the form of average grade GPA each semester. This graph provides users with significant increase / decrease in student performance in the third semester. This page is expected to be a reference for head of study program in analyzing the cause of student performance drop in third semester as shown in Fig. 9.

Fig. 9. Prediction of student’s graduation that grouped by class

3. Most Repeated Subject Page
   This page displays the 10 frequently repeated courses in each class. This page is only accessible by caps. To view details of each course, the user can press the desired course chart. With this page is expected to give head of study program action to the courses that are predicted to reduce the number of graduated on time students as shown in Fig. 10.

Fig. 10. Prediction of student’s graduation that grouped by class

4. Testing
   The testing method that used for this research are unit testing and usability testing that perform by user using Likert scale for the measurement method. The unit testing is performed by programmer to find bugs of the application. And the usability testing is performed only by head of information system study program, because this application is only used by that user. The usability testing is divided into two categories they are user interface and functionality. The result of unit testing is there are no bugs found in unit testing. And the result of the usability testing can be seen in Fig. 10.
Based on Fig. 10, user is satisfied with the user interface but less satisfied for the functionality to support the decision making.

V. CONCLUSION

Based on this research can be concluded this dashboard design is just a prototype, because there are lack of analysis in decision making and the dashboard only showing information and temporary prediction. The data model that used on this research is labeling data that has been processed using C4.5 algorithm and data that has been through data cleansing process using Pentaho Data Integration.

REFERENCES

[1] “About Us : Fact & Figure.”.