Predictive Modeling of the First Year Evaluation Based on Demographics Data: Case Study Students of Telkom University, Indonesia

Tora Fahrudin[1, 2], Joko Lianto Buliali[1], Chastine Fatichah[1]

torafahrudin@telkomuniversity.ac.id, joko@cs.its.ac.id, chastine@cs.its.ac.id

[1] Department of Informatics, Institut Teknologi Sepuluh November

[2] School of Applied Science, Telkom University

***Abstract*—Student academic failure prediction is still interesting topics in the Educational Data Mining. One of the challenges is how to predict student academic failure as early as possible. This research focuses on predictive modeling of unsuccessful students in the first year evaluation. We propose a new concept of predictive modeling of the first year evaluation which combines 3 input data: demographics, academic and social media. The modeling can be divided into two sub modeling (normal period and extra period). In this paper, we focus on demographic data modeling (first sub-modeling) which correlated with the probability of a student to pass the first year evaluation on normal period. A Weka tool is used to get a pattern of data by using white box classifier (decision tree and rule base). Meanwhile, to solve the problem of unbalanced in our training data, we use data balancing scenario using same portion oversampling, random oversampling and SMOTE. From the testing result, we choose the best three student failure pattern of the F-Measure minor class value which obtained from "One R" and "ADTree" algorithms using Balancing scenario, the reason is because F-Measure describes the smallest error rate both FP (False Positive) and also FN (False Negative). From the best three of student failure pattern, we found that gender, selection path, study program and age are the attributes that are most correlated with the probability to pass the first year evaluation on extra period.**

Keywords*—educational data mining, first year evaluation, weka, decision tree, rule base, data balancing, SMOTE

I. Introduction

Research topics in the field of Educational Data Mining (EDM) are still interesting to study. This is indicated by the number of publications about EDM is increasing exponentially in the last year [1]. EDM aims to convert the raw data from the world of education to obtain patterns or information useful for stakeholders to develop the academic environment.

Romero, et al categorized EDM into three domain areas of research environment [1]:

1. Offline Education (traditional / face to face)
2. E-learning and learning management system (Educational use of electronic media and learning management application)
3. Intelligent tutoring system and Adaptive Educational Hypermedia System (Intelligence-based learning system)

Of the three categories above, this research focuses on the domain area 1 (Offline Education) because most of the learning process at the Telkom University is done traditionally or by face to face.

One of the challenges is how to predict academic failure as early as possible. Predictive modeling on the first year evaluation stage of the student is considered as the best way to predict such failure. That is because success of students in the first year evaluation describes their successful study in a campus. It is also confirmed in the study [2] [3], which states that the greatest risk of student dropping out is in the first year of study, and decreasing after that. Research [3] [4] also mentioned that the best way to avoid such failure is maximizes the handling of the first year students.

We propose a new concept of predictive modeling first year evaluation which combines 3 input data: demographics, academic and social media. Section 2 provides the related topic about EDM and predicting modeling of dropout out first year. The main frame of the prediction modeling will be introduced in section 3. The focus of the discussion is modeling the part of that main frame which concern only on demographic data. That demographic data will be correlated with the probability of a student to pass the first year evaluation.

II. Related Topic

A. State of The Art Educational Data Mining

In a survey in the state of the art of EDM, Romero, et al [1], categorizes studies that have been done in into 11 tasks:

TABLE 1. 11 TASKS OF RESEARCH IN THE DOMAIN OF EDM

No	Data Mining Task
1	Analysis and Visualization of Data
2	Providing Feedback for Supporting Instructors
3	Recommendations for Students
4	Predicting Student's Performance
5	Student Modeling

978-1-5090-5671-2/16/$31.00 ©2016 IEEE

No	Data Mining Task
6	Detecting Undesirable Student Behaviors
7	Grouping Students
8	Social Network Analysis
9	Developing Concept Maps
10	Constructing Courseware
11	Planning and Scheduling

Of the 11 tasks above, our predictive modeling belong to the "Predicting Student's Performance" which aims to estimate the relationship of demographic variables that contributes to student academic performance.

B. *Prediction Modeling of Drop Out First Year*

Table 2 discuss the results of the analysis of literature about the predictive modeling of students drop out in the first year

TABLE 2. LITERATURE SURVEY OF DROP OUT PREDICTION MODELING

Year	Ref Paper	Data
2014	[5]	Family background, previous academic achievements, entry examination score, score of the students at the end of the first semester
2015	[6]	Reason for applying, student opinions about their studies, student academic performance
2015	[4]	History of student, Student's involved in studies, Student's perception
2016	[2]	Student's academic performance in first semester, student social behavior, personal background and education background

III. MODEL OVERVIEW

A. *The Main Modeling of The Prediction*

Many universities apply academic evaluation to student at the end of their first year. At Telkom University, the evaluation is conducted as follows:

1) For student who passes all credit (for courses in semester one and two) in 2 semesters, the evaluation is done at the end of semester 2. In this research such student called as normal student which can pass the evaluation in "normal period". We divided this "normal period" evaluation into 2 sub modeling:

- ✓ First Sub Modeling was built at the beginning before the course
- ✓ Second Sub Modeling was built at the beginning of second semester

2) For student who can't passes all credit (for courses in semester one and two) in 2 semesters, the evaluation is done at the end of semester 4. In this research such student called as extra student which can pass the evaluation in "extra period". Student, who can't pass all credit in the end of fourth semester, will be classified as Drop Out student. We divided this "extra period" evaluation into 2 sub modeling:

- ✓ Third Sub Modeling was built at the beginning of the third semester
- ✓ Fourth Sub Modeling was built at the beginning of the fourth semester

The result of each sub modeling was the student failure pattern in the form of decision tree and rule based such as Table 7, Figure 9 and Figure 10 and also prediction result of student failure in the form of table.

Figure 1 clarifies the correlation about sub modeling, data input and objective for each sub modeling. We mark the picture in dotted line to show our focus on demographics data.

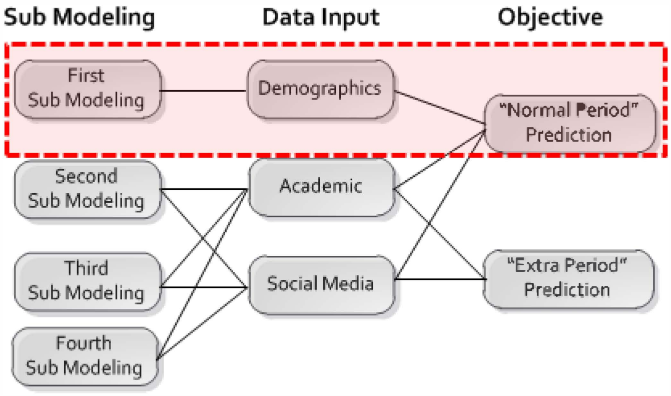


Figure 1. Correlation Diagram

Figure 2 illustrates that a modeling created in sequence of each semester.

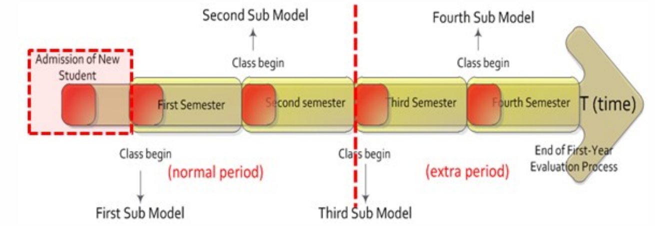


Figure 2. First year evaluation Main Model

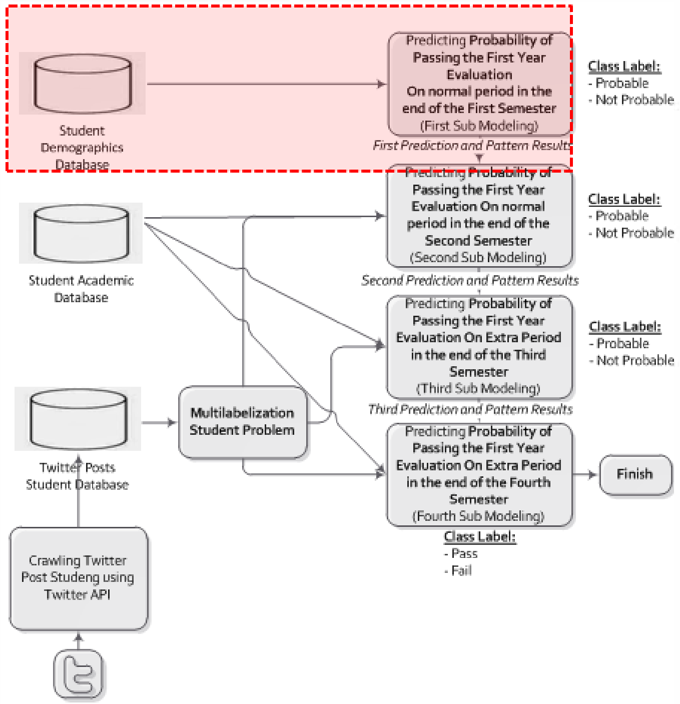

Student Demographics Database
Predicting Probability of Passing the First Year Evaluation On normal period in the end of the First Semester (First Sub Modeling)
First Prediction and Pattern Results
Class Label:
- Probable
- Not Probable
Student Academic Database
Predicting Probability of Passing the First Year Evaluation On normal period in the end of the Second Semester (Second Sub Modeling)
Second Prediction and Pattern Results
Class Label:
- Probable
- Not Probable
Predicting Probability of Passing the First Year Evaluation On Extra Period in the end of the Third Semester (Third Sub Modeling)
Third Prediction and Pattern Results
Class Label:
- Probable
- Not Probable
Twitter Posts Student Database
Multilabelization Student Problem
Predicting Probability of Passing the First Year Evaluation On Extra Period in the end of the Fourth Semester (Fourth Sub Modeling)
Class Label:
- Pass
- Fail
Finish
Crawling Twitter Post Studeng using Twitter API

Figure 3. Main Framework of First Year Evaluation Predictive Model

Based on the literature [7], student social media data will be processed into a multi-label student problems classification. Modeling social media posting such as twitter in Indonesian language domain have many challenges [8] such as informal language, ambiguity, out of vocabulary etc.

B. Demographics Modeling (First Sub Modeling)

The first sub modeling was built using demographic data which correlated with the probability of a student to pass the first year evaluation on normal period.

B1) Data Collection and Description

Data were obtained from 3 batches of student school year (2013, 2014 and 2015) which contains 19.048 data from seven faculties. There are 21 attribute used as inputs and 1 attribute is used as outputs for our modeling.

TABLE 3. DATA DESCRIPTION

Attribute name	Attribute type	Example
Student Gender	Nominal	Male, Female
Student Age	Numeric	16, 17, 18, 19, …
Student Religiom	Nominal	Islam, Kristen, …
Student Birthplace Province	Nominal	0_11, 0_12, …
Sudent Selection Path	Nominal	0_1, 0_2, 0_3, …
Student Province	Nominal	0_11, 0_12, …
Father Age	Numeric	36, 39, 40, …
Father Earning	Ordinal	<1JT, 1-3JT, …
Father Birthplace Province	Nominal	0_11, 0_12, …
Father Group Occupation	Nominal	GOL_0, GOL_1, …
Father Education	Ordinal	SD, SMP, SMA, …
Mother Age	Numeric	33, 35, 36, …
Mother Earning	Ordinal	<1JT, 1-3JT, …
Mother Birthplace Province	Nominal	0_11, 0_12, …
Mother Group Occupation	Nominal	GOL_0, GOL_1, …
Mother Education	Ordinal	-, SD, SMP, SMA, …
Number of Older Sibling	Numeric	0, 1, 2, 3, 4, 5, …
Number of Younger Sibling	Numeric	0, 1, 2, 3, 4, 5, …
Studyprogram receive	Numeric	1, 2, 3, 4, 5, …
Studyprogram Offer	Nominal	Y, T
Stduyprogram	Nominal	0_11, 0_12, 0_13, …
Output Label	Nominal	Probable, Not Probable

For modeling purposes, we choose 13.720 clean data from 19.048 data (not containing “Null” in all attributes).

B2) Derived class label

We derived the output label for our first sub model with reference to the “Telkom University Education Handbook, 2015”. Output label formed by following two rules regarding the requirements to pass the first year evaluation:

- ✓ Probable: End of first semester has GPA > = 2 and the rest of the credits in second semester <= 24 credits.
- ✓ Not Probable: End of first semester has GPA < 2 and the rest of the credits in second semester > 24 credits.

B3) Imbalance problem on Data

After the output label was derived, we found skewed class distribution phenomena. That’s problem generally known as Imbalance data, which has characteristic of unequal distribution among its classes [9]. According to the literature [10] [11] [12] [13] [14], there are 2 kinds of solutions in dealing with imbalance problems: 1) data adaptation or 2) algorithm adaptation. To overcome our data imbalance problems, we choose oversampling as part of the data adaptation technique.

TABLE 4. DATA DISTRIBUTION FROM TWO CLASS OUTPUT LABEL

Class Label	Count of Data	Percentage of Data
Not Probable	1830	13.34%
Probable	11890	86.66%

B4) White box classifier

Based on how the result can be interpreted, classification algorithm is divided into white box and black box classifier [15]. The difference between the white-box and black-box classifier is: white-box classifier doesn’t allow student failure patterns to be viewed and understand by human, while the black-box classifier doesn’t give a clear explanation about the student failure pattern. In this paper, the white-box classifier was chosen because we need to be able to identify the student failure pattern. Table 5 gives a selected algorithm for each category of white-box classifier from Weka machine learning.

TABLE 5. WHITE BOX CLASSIFIER WHICH USED IN OUR MODEL

Category	Classifier
Decision Tree	-J48 -ADTree -Simple Cart
Rule based	-OneR -JRip -PART

B5) Block Diagram of 2 main scenarios

Figure 4 shows our first block diagram model which contains 2 main scenarios: 1) build a white-box classifier on original data and 2) build a white-box classifier on balance data. Each scenario tested with Decision Tree and Rule Based algorithms. Evaluation is done by using iterations on 5X 10 Fold Cross Validation. The model with the highest F-Measure accuracy is selected as the best model. The student pattern which extracted from our best model will be selected as student failure suspect model.

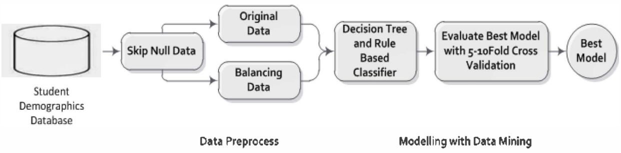

Student Demographics Database
Skip Null Data
Original Data
Balancing Data
Decision Tree and Rule Based Classifier
Evaluate Best Model with 5-10Fold Cross Validation
Best Model
Data Preprocess
Modelling with Data Mining

Figure 4. First Sub Model Block Diagram

Based on [16], c4.5 algorithms give the best results of accuracy for 50:50 learning distribution of major and minor. To get that best proportion of training data, we choose oversampling method. Oversampling is increasing the amount of the smaller class to achieve a certain balance scale. For balancing process we use three scenarios: same portion oversampling, random oversampling and SMOTE [17]. In our terms, Same Portion Oversampling is increasing training data of minor class by copy each minor class data with the same ratio until achieve 50:50 scale of distribution. Meanwhile, Random Oversampling (ROS) is increasing training data of minor class by using randomly sampling with replacement [18] of each minor class data until achieve 50:50 scale of distribution.

B6) Evaluation of model

Overall accuracy cannot be used as a measure of evaluation models in Imbalance problems. In imbalance domain, accuracy is calculated separately for each class (both major classes (+), and minor classes (-)). Moreover, Napierala [11] emphasize that accuracy in the minor class is more important than major class. So to get accuracy of each class, the performance of classifiers can be presented in a confusion matrix in two classes as in Table 6.

TABLE 6. CONFUSION MATRIX FOR TWO-CLASS PROBLEM

Actual	(+) Predicted	(-) Predicted
(+)	TP	FN
(-)	FP	TN

$$Recall\ (-) = \frac{TN}{FP+TN} \qquad (1)$$

$$Precision\ (-) = \frac{TN}{TN+FN} \qquad (2)$$

$$FMeasure\ (-) = \frac{(1+\beta^2).Recall\ .Precision}{\beta^2\ .Recall+Precision} \qquad (3)$$

Usually β = 1 [9], which means Recall and Precision have the same priority.

IV. EXPERIMENT RESULTS AND DISCUSSION

In this section, we report and discuss the performances of the each scenario from section 3. We observe average (Figure 5, 6) and maximum (Figure 7, 8) resulted from 5 X 10 fold cross validation for each scenario.

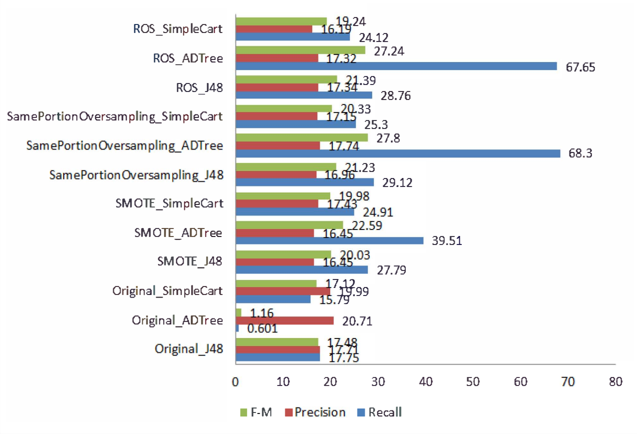


FIGURE 5. AVERAGE RESULT ON DECISION TREE

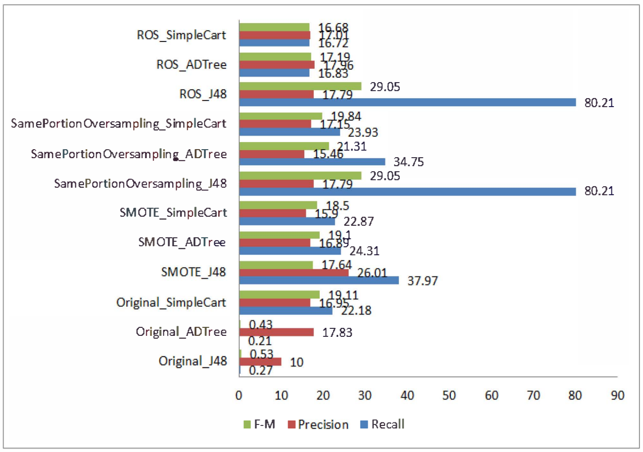


FIGURE 6. AVERAGE RESULT ON RULE BASED

In "Decision Tree" tested result, the highest average value of Recall is 68.3%, Precision 20.71% and F-Measure 27.8%. The best three Recall and F-Measure average value are achieved by using ADTree algorithms for all balancing scenario. On the other hand, "Rule Based" tested result gets the highest average 80.21% for Recall, 26.01% for Precision and 29.05% for F-Measure. OneR and JRip algorithms dominate the best achievement for average on "Rule Based" tested result.

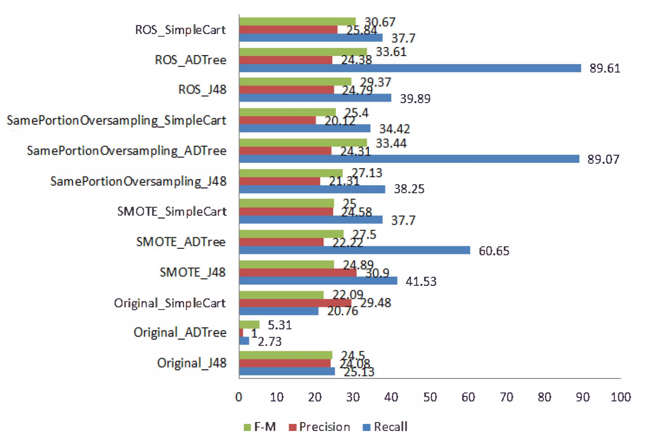


FIGURE 7. MAXIMUM RESULT ON DECISION TREE

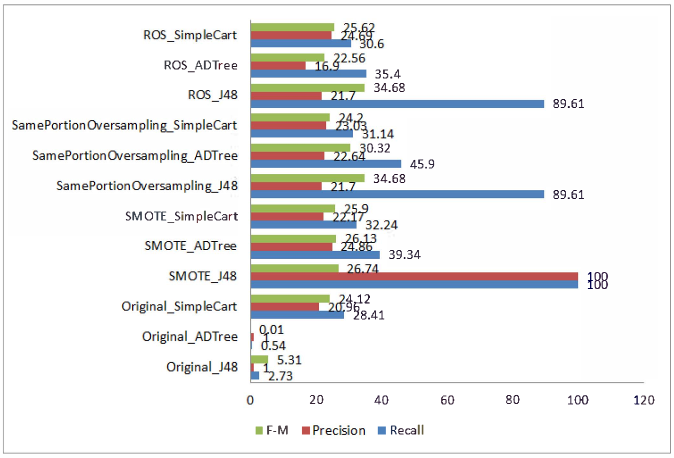


FIGURE 8. MAXIMUM RESULT ON RULE BASED

In "Decision Tree" tested result, the highest maximum value of Recall is 89.61%, Precision 30.09% and F-Measure 33.61%. The best three of Recall maximum best value are achieved by using ADTree algorithm for all balancing scenarios. F-Measure maximum best value has been achieved by ADTree algorithms using Same Portion Oversampling and Random Oversampling scenario.

In "Rule based" tested result, the highest maximum value of Recall is 100%, Precision 100% and F-Measure 34.68%. OneR and JRip algorithms dominate the best achievement for maximum Recall and F-Measure for all balancing scenario.

Instead of using original data scenario, balancing scenarios can improve average value of Recall, Precision and F-Measure accuracy (27.93% for Recall, 0.3% for Precision and 12.26% for the F-Measure). From three scenarios of balancing, highest average Recall improvement is achieved by Same Portion Oversampling (34.13%), second by Random Oversampling (29.58%) and the last by SMOTE (20.09%). The highest average Precision improvement is achieved by SMOTE (0.99%) and second Random Oversampling (0.07%). Meanwhile highest average F-Measure improvement is achieved by Same Portion Oversampling (13.95%), second by Random Oversampling (12.49%) and the last SMOTE (10.33%).

A. *Best Student Failure Pattern*

Of the three minor class parameters: Recall, Precision and F-Measure, we choose F-Measure as the main measure to be achieved. That is because the value of the F-Measure, describe the harmonization value of Recall and Precision [19]. In accordance with a formula of F-Measure, we believe that the value of F-Measure best "minor class" have described the smallest error rate for harmonization 2 parameter value, both FP (False Positive) and FN (False Negative).

Based on our experiment, best Recall value (100%) was obtained at fourth partition, where all data are predicted to "Not Probable" class. The opposite is for best Precision value (100%) obtained at third partition, where all data are predicted to "Probable" class. Therefore, Best Recall and Best Precision are not main measure to be achieved. That's why we choose the highest F-Measure.

The best F-Measure maximum value (34.68%) is obtained from third partition using OneR algorithms at Same Portion Oversampling scenario (Recall 86.33% and Precision 21.7%).

TABLE 7. STUDENT FAILURE PATTERN IN ONER ALGORITMS FOR BEST F-MEASURE

Gender	Predicted
Male	Not Probable
Female	Probable

Table 7 explains that male students have a lower probability to pass the evaluation of the first year on normal period than female students.

Second best F-Measure (33.61%) was achieved by ADTree algorithm using scenarios balancing Random Oversampling (Recall 54.1% and Precision 24.4%). The result of tree can be seen on Figure 9 below:

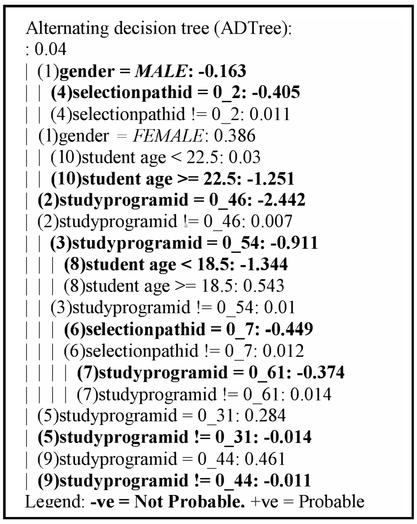
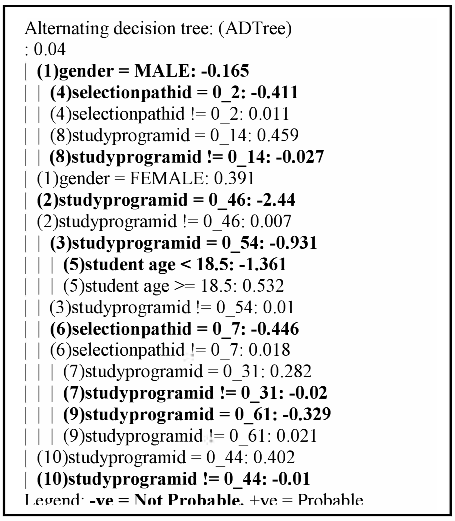

```
Alternating decision tree (ADTree):
: 0.04
| (1)gender = MALE: -0.163
| | (4)selectionpathid = 0_2: -0.405
| | (4)selectionpathid != 0_2: 0.011
| (1)gender = FEMALE: 0.386
| | (10)student age < 22.5: 0.03
| | (10)student age >= 22.5: -1.251
| (2)studyprogramid = 0_46: -2.442
| (2)studyprogramid != 0_46: 0.007
| | (3)studyprogramid = 0_54: -0.911
| | | (8)student age < 18.5: -1.344
| | | (8)student age >= 18.5: 0.543
| | (3)studyprogramid != 0_54: 0.01
| | | (6)selectionpathid = 0_7: -0.449
| | | (6)selectionpathid != 0_7: 0.012
| | | | (7)studyprogramid = 0_61: -0.374
| | | | (7)studyprogramid != 0_61: 0.014
| (5)studyprogramid = 0_31: 0.284
| (5)studyprogramid != 0_31: -0.014
| (9)studyprogramid = 0_44: 0.461
| (9)studyprogramid != 0_44: -0.011
Legend: -ve = Not Probable. +ve = Probable
```

FIGURE 9. ADTREE DECISION RESULTS IN SECOND BEST F-MEASURE

One example result of the student failure pattern (Figure 9) can be explained as follows: If a student has gender = Male and Selection Path = 0_2 than its predicted that he passes the evaluation of the first year in extra period.

Third best F-Measure (33.44%) was achieved by ADTree algorithm using scenarios balancing Same Portion Oversampling (Recall 53.55% and Precision 24.31%). The result of tree can be seen on Figure 10 below:



```
Alternating decision tree: (ADTree)
: 0.04
| (1)gender = MALE: -0.165
| | (4)selectionpathid = 0_2: -0.411
| | (4)selectionpathid != 0_2: 0.011
| | (8)studyprogramid = 0_14: 0.459
| | (8)studyprogramid != 0_14: -0.027
| (1)gender = FEMALE: 0.391
| (2)studyprogramid = 0_46: -2.44
| (2)studyprogramid != 0_46: 0.007
| | (3)studyprogramid = 0_54: -0.931
| | | (5)student age < 18.5: -1.361
| | | (5)student age >= 18.5: 0.532
| | (3)studyprogramid != 0_54: 0.01
| | (6)selectionpathid = 0_7: -0.446
| | (6)selectionpathid != 0_7: 0.018
| | | (7)studyprogramid = 0_31: 0.282
| | | (7)studyprogramid != 0_31: -0.02
| | | (9)studyprogramid = 0_61: -0.329
| | | (9)studyprogramid != 0_61: 0.021
| (10)studyprogramid = 0_44: 0.402
| (10)studyprogramid != 0_44: -0.01
Legend: -ve = Not Probable, +ve = Probable
```

FIGURE 10. ADTREE DECISION RESULTS IN THIRD BEST F-MEASURE

One example result of the student failure pattern (Figure 10) can be explained as follows: If a student has gender Male and (Selection Path = 0_2 or Study program != 0_14) than there are predict to pass the evaluation of the first year in extra period.

V. CONCLUSION AND FUTURE WORK

Predictive modeling using demographic data as the first sub modeling is an alternative solution to prevent Drop Out students as early as possible. When imbalance data problem is found (the number of Not Probable Data is less than Probable Data), oversampling with 3 scenarios (Same Portion Oversampling, Random Oversampling and SMOTE) are used. From the experimental results, it was found balancing data scenarios improve average accuracy in the amount of 27.93% for Recall, 0.3% for Precision and 12.26% for the F-Measure. Of the three balancing scenarios, Same Portion Oversampling is the best method to our data, because of the best achievement in increasing the average F-Measure (34.135%) and Recall value (13.95%). Among six classifiers, OneR and ADTree give the F-Measure and Recall best value using a data balancing scenario. From the best three of student failure pattern, we found that gender, selection path, study program and age are the attributes that are most correlated with the probability to pass the first year evaluation on extra period.

Student academic and social media modeling will be explored in our future work. Modeling student academic data will be focus on modeling student failure by academic activities student data such as student presence, exam score, etc. Modeling social media data will be focus on how to formalize the social media language and extract multi label student problem in Indonesian language domain. These two models can be combine into one comprehensive model to give better performance in predicting student failure possibility.

REFERENCES

[1] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," IEEE Trans. Syst. Man, Cybern. C Appl. Rev., vol. 40, no. X, pp. 601–618, 2010.

[2] P. Meedech, N. Iam-on, and T. Boongoen, "Prediction of Student Dropout Using Personal Profile and DataMining Approach," in Intelligent and Evolutionary Systems, vol. 5, 2016, pp. 143–155.

[3] R. Chen, "Institutional Characteristics and College Student Dropout Risks: A Multilevel Event History Analysis," Res. High. Educ., vol. 53, no. 5, pp. 487–505, 2012.

[4] M. A. Yehuala, "Application of Data mining Techniques for student success and failure prediction," Int. J. Sci. Technol. Res., vol. 4, no. 4, pp. 3–6, 2015.

[5] M. Goga, S. Kuyoro, and N. Goga, "A Recommender for Improving the Student Academic Performance," Procedia - Soc. Behav. Sci., vol. 180, no. November 2014, pp. 1481–1488, 2014.

[6] K. Kori, M. Pedaste, E. Tõnisson, H. Altin, and R. Rantsus, "First-year dropout in ICT studies," no. March, pp. 444–452, 2015.

[7] X. Chen, M. Vorvoreanu, and K. P. C. Madhavan, "Mining Social Media Data for Understanding Students' Learning Experiences," IEEE Trans. Learn. Technol., vol. 7, no. 3, pp. 246–259, 2014.

[8] A. R. Naradhipa and A. Purwarianti, "Sentiment classification for Indonesian message in social media," Proc. 2011 Int. Conf. Electr. Eng. Informatics, pp. 1–4, 2011.

[9] H. He and E. A. Garcia, "Learning from imbalanced data," IEEE Trans. Knowl. Data Eng., vol. 21, no. 9, pp. 1263–1284, 2009.

[10] D. Tomar and S. Agarwal, "A Survey on Pre-processing and Post-processing Techniques in Data Mining," vol. 7, no. 4, pp. 99–128, 2014.

[11] K. Napierała, "Improving Rule Classifiers For Imbalanced Data," Poznan University of Technology, 2012.

[12] Y. Dong and X. Wang, "A new over-sampling approach: Random-SMOTE for learning from imbalanced data sets," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 7091 LNAI, Springer Berlin Heidelberg, 2011, pp. 343–352.

[13] R. Longadge, S. Dongre, and M. Latesh, "Class Imbalance Problem in Data Mining : Review," Int. J. Comput. Sci. Netw., vol. 2, no. 1, 2013.

[14] L. Abdi and S. Hashemi, "To Combat Multi-Class Imbalanced Problems by Means of Over-Sampling Techniques," IEEE Trans. Knowl. Data Eng., vol. 28, no. 1, pp. 238–251, 2016.

[15] C. Márquez-Vera, A. Cano, C. Romero, and S. Ventura, "Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data," Appl. Intell., vol. 38, no. 3, pp. 315–330, 2013.

[16] S. Visa and A. Ralescu, "Issues in mining imbalanced data sets-a review paper," in Proceedings of the sixteen midwest artificial intelligence and cognitive science conference, 2005, no. August 2016, pp. 67–73.

[17] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE : Synthetic Minority Over-sampling Technique," vol. 16, pp. 321–357, 2002.

[18] R. Dubey, J. Zhou, Y. Wang, P. M. Thompson, and J. Ye, "Analysis of sampling techniques for imbalanced data: An n=648 ADNI study," Neuroimage, vol. 87, pp. 220–241, 2014.

[19] J. Makhoul, F. Kubala, R. Schwartz, and R. Weischedel, "Performance measures for information extraction," in Broadcast News Workshop '99 Proceedings, 1999, pp. 249–252.