The application of perceptron method in predicting student graduation based on several identified key factors

Wafi Sabiq Masykuri¹, Elis Khatizah¹, Fahren Bukhari¹

¹Department of Mathematics, Faculty of Mathematics and Science, IPB University, Indonesia

*Corresponding author: elis_khatizah@apps.ipb.ac.id

Abstract. In this paper, we predicted whether a student succeed to pass their higher education on time or not based on Grade Point Average (GPA) and several indicators related to final year project activities. It was also analysed which indicator that is dominant based on its resulted weight. The process was conducted by using perceptron method, while the data was taken from students of Department of Mathematics, IPB University, enrolled in 2013, 2014 and 2015. Furthermore, the results showed that learning proses accuracy stand at 80.5 % and all mentioned indicators give a positive contribution to students in pursing on-time graduation except supervised by the lecturer whose field of interest is in line with the student. Additionally, based on the resulted weight for each indicator, student’s GPA has a bigger impact for student to graduate on time.

Keywords: perceptron, on-time graduation, GPA.

1. Introduction

On-time graduation of students is one of the quality indicators of higher institution which also indicates students’ success in pursuing their degree [1]. In detail, students are called graduating on time if they can complete their study up to four years, while over-four-year study length leading to ones with not on-time graduation [2]. There are a lot of benefits can be obtained by completing the education before the deadline of study duration in most Indonesian university. For instance, students exempt from tuition fees and can start their carrier earlier. However, due to some reasons, the vast majority of students can not complete their studies in the university or college referring to a certain limit of time [3].

There is an internal factor and external factor which influence student graduation [4,5]. The inability of students in writing and the difficulty in communicating with supervisor were two main problems in completing the thesis, while supporting factors involve a supportive environment and management system [6]. In addition, for students who are active in organizations, there is a significant positive correlation between hope of graduating on time and self-regulation. The higher the hope someone to graduate on time, the higher someone’s self-regulation ability to pursue the goal [7].

Providing system information about the factors by which students graduate on-time is quite helpful both for students and universities. One of the research in this field usually focus on classifying student as who are graduate on time or not by using Fuzzy K-Nearest Neighbour Algorithm [8,9] or predicting the length of time study based on some indicators by comparing some methods [10,11]. Referring to Meinanda et al., 2009, the best model to predict time study was constructed by Artificial Neural Network where academic factors (Grade Point Average (GPA) and the numbers of taken subjects) were the main indicators [12].
In this paper, we identified a dominant time-study factor among GPA and several indicators related to the way students finish their final project. It was obtained by analysing close-relation factors which have a certain pattern in which we applied a learning process based on perceptron method, a simple method in artificial neural network. In spite of its simplicity, this method not only can identify the dominant indicators but also can be used to predict the graduation of one student, whether he is on-time or not.

2. Research Methodology

2.1 Collecting data
The data used in this paper are primary data which was collected from questionnaires involving the respondents who are students of Department of Mathematics, IPB University, enrolled in 2013, 2014, and 2015. We used six given indicators related to student capability in finishing final year project as a measurement of on-time graduation.

1. GPA equal to or more than 3.00 (out of 4.00).
2. Supervised by a lecturer whose field of interest is in line with the student.
3. Relatively easy to search for references.
4. Regularly meet with the supervisor to report the progress at least once in two weeks.
5. The student does not have other main activities such as fulltime work, internship, or has gotten married.
6. Student has his own timeline and applies it at once.

The resulted data was then divided into training data and testing data.

2.2 Literature study
In this stage, literature study relating to both graduation prediction and the method was conducted. Some papers were taken into account to support the legality of six mentioned indicators, while several books were become references for the used method, perceptron method.

2.3 Implementing the algorithm
Based on perceptron algorithm, the training data were processed to get the optimal weight and the resulted weight are then used to run the testing data. Following this, the result of testing process was compared with the real target to look for its accuracy. At the same time, the previous weights value was also used to identified the dominant indicator.

3. Results and Discussion
The total data obtained from the questionnaire are 104, where the number of inputs from students enrolled in 2013, 2014, and 2015 is 33, 35 and 36 respectively. We use the binary form for this data, where “1” means the indicator appears, in contrast to “0” which state that the indicator does not appear to the respondent. This binary number is also used as a target notation, where “1” denote that respondent graduated on-time and not on-time graduate student is denoted with “0”.

We process the data by using perceptron method to get the weight of each mentioned indicators and also to predict the graduation time. Perceptron is the simplest method in neural network which is built by a single layer and a hard-limited threshold activation function. This method can only be applied to classify linearly separable data [13], meaning that we can draw a line to separate two classes of data in 2D illustration. Furthermore, there are two process in the method; supervised learning process and testing process [14]. In the former process, resulted outputs have to equal with the real target by which the weights and bias are determined to minimize the cost function, while the final value of weights and bias are then executed in the latter process. In short, a simple illustration of perceptron method [15] can be shown in Figure 1.
Due to the fact that perceptron can only be applied for separable data pattern [16-19], the learning process will not converge [20] if the method faces inseparable pattern which means not all of the expected outputs equal with its real targets. In this case, we need other aspects as development of perceptron or we need to delete several data to make the pattern separable. In 2D case, a linear separation function in perceptron with two inputs can be formulated by Equation (1) as follow

$$w_1 x_1 + w_2 x_2 + \text{bias} = \theta$$  

where $w_1$, $w_2$, and $\text{bias}$ are the final value of weights and bias resulted from the learning process, whereas $\theta$ state a threshold value [21]. In general, we have several steps to apply the perceptron method [22-25], generalized as follow:

1. Declaring $(n+1)$ input, and $m$ iterations to obtain the expected output which has the same value as the real target. In this step, we choose random numbers for the initial value of weights $w_{ij}$, where $i = 0, 1, 2, \ldots, n$, and $j = 1, 2, \ldots, m$ for input-ith and iteration-jth.

2. Calculating the input vector $x$ by using input function $u_j$ formulated as $u_j = \sum x_i w_{ij}$. As for the value of bias, which is $x_0 = 1$, we can simplify the formulation of input function, stated as $u_j = \sum x_i w_{ij} + b_{ij}$.

3. Using a hard-limited threshold function as an activation function to restrict the output $f(u_j) = y_j$. In this paper, we use binary hard-limited threshold function which can be seen on Equation (2) below.

$$f(u_j) = \begin{cases} 0, & u_j \leq \theta \\ 1, & u_j > \theta \end{cases}$$

4. Calculating error value as a distraction of the real target of input and the resulted output from the previous step or $\text{Err}_j = \text{target}_j - y_j$.

5. Renewing the weights and bias by applying the formula $w_{ij}(j+1) = w_{ij} + \alpha \text{Err}_j$ and $b_{ij}(j+1) = b_{ij} + \alpha x_0 \text{Err}_j$ respectively, where $\alpha$ state a learning rate with $0 < \alpha \leq 1$.

6. Repeating steps 2 to 5 so as we get $\text{Err}_j = 0$.

### 3.1 Learning and testing process

The learning process is firstly conducted by using the data from students enrolled in 2013 and 2014, while we use the rest data for the testing process. Unfortunately, the former process can not be done...
owing to the fact that the data pattern seems inseparable as a result of anomaly-data existence. In this term, anomaly data refers to respondents who have the same value of all indicators but their targets are different. To deal with it, we eliminate anomaly data in order to get a learning process convergency. From this step, in total there are 60 (out of 68) data used in the learning process where the first five respondents can be seen on Table 1 and the complete one is on Attachment 1.

| Respondent | Indicator-1 | Indicator-2 | Indicator-3 | Indicator-4 | Indicator-5 | Indicator-6 | Target |
|------------|-------------|-------------|-------------|-------------|-------------|-------------|--------|
| 1          | 0           | 0           | 1           | 0           | 0           | 0           | 0      |
| 2          | 0           | 1           | 0           | 0           | 0           | 0           | 0      |
| 3          | 0           | 1           | 0           | 0           | 0           | 0           | 0      |
| 4          | 0           | 1           | 0           | 0           | 0           | 0           | 0      |
| 5          | 0           | 1           | 1           | 0           | 1           | 0           | 0      |

At this point, we take zero for the initial weight and bias, \((w_1, w_2, w_3, w_4, w_5, w_6) = (0, 0, 0, 0, 0, 0)\), and \(bias_0 = 0\). Furthermore, by using \(i = 1, 2, 3, 4, 5, 6\) and \(j = 1, 2, 3, \ldots, 60\), it is obtained that \((w_1, w_2, w_3, w_4, w_5, w_6) = (6, -3, 1, 4, 1, 5)\) and \(bias_0 = -8\) executing in 9 iterations and 2.1008 second of running time. The value of weights and bias is then used in the testing process to see its accuracy as illustrated in Table 2.

| Respondent | Indicator-1 | Indicator-2 | Indicator-3 | Indicator-4 | Indicator-5 | Indicator-6 | Output | Target |
|------------|-------------|-------------|-------------|-------------|-------------|-------------|--------|--------|
| 1          | 0           | 0           | 0           | 0           | 0           | 0           | 0      | 0      |
| 2          | 0           | 1           | 0           | 0           | 0           | 0           | 0      | 0      |
| 3          | 0           | 1           | 0           | 0           | 0           | 0           | 0      | 0      |
| 4          | 0           | 1           | 0           | 0           | 0           | 0           | 0      | 0      |
| 5          | 0           | 1           | 0           | 0           | 0           | 0           | 0      | 0      |

Table 1. The data of first five respondents, used in the learning process.

Table 2. The result of testing process.
Table 2 illustrates the result of testing process for the data of students who enrolled in 2015. The main point is spotted on the output column, while the target column consists of the real values. Comparing the output and its target, there are seven unequal data leading to obtaining 80.5% of rough accuracy. This accuracy is only fit for the data in this study which might differ if we apply to the other ones, for instance we can look at the comprehensive result by Meinanda et al., 2009 where the level of confidence was 95%. In spite of its exclusiveness, we can say that the application of perceptron method for our case relatively succeeds. Therefore, the result of learning process can be used to predict whether a student graduate on-time or not based on six previous indicators.

### 3.2 Learning process using entire data

In this part, we use all collected data in the learning process due to assumption that the more data uses, the better accuracy resulted. Owing to anomaly-data existence, the process involving 80 data (out of 104) which converge to weight vector, \((w_1, w_2, w_3, w_4, w_5, w_6) = (6, -2, 2, 3, 2, 5)\) and bias = -10. From this result, each of weight value is positive except \(w_2\) which is exactly the same as the previous result when we only use 60 data. Consequently, it can be said that all mentioned indicators give a positive contribution for students on pursuing on-time graduation except the second indicator, i.e. supervised by the lecturer whose field of interest is in line with the student. Moreover, the similar result also appears either in this part or in the previous learning process where the highest value of weights is \(w_1\). Hence, we can state that the GPA has a big impact on successfully getting the bachelor degree up to four years.

### 4. Conclusion

A single-layer perceptron method with hard-limited-threshold activation function has been applied to estimate on-time graduation of mathematics major students, IPB University. Based on twice learning process involving some and entire collected data, we have obtained a certain value of weights and bias that can be used to predict student graduation. Besides, the results are leading to the interpretation that all mentioned indicators give a positive contribution for students in pursuing on-time graduation except ‘supervised by the lecturer whose field of interest is in line with the student’. Additionally, a student’s GPA has a bigger impact on achieving a bachelor degree up to four years rather than other indicators. This result might give benefit information both for students and institutions in predicting the graduation, whether on-time or not, and identifying key factors to deal with completing a final year project in certain expected times.

### References

[1] BAN-PT. (2019). *Akreditasi Perguruan Tinggi Kreteria dan Prosedur 3.0*. 1–18.
[2] Statistika, D., Matematika, F., Ilmu, D. A. N., & Alam, P. (2015). *PROGRAM SARJANA IPB*. IPB University.
[3] Wardani, L. W., & Adil, A. (2019). Jurnal BITE Vol. 1 No. 1 (Juni) 2019, Hal 60-68 Jurnal BITE, 1(1), 60–68.

[4] Wahyudi, Frendy R. (2012). Faktor-faktor penyebab mahasiswa S1 Pendidikan Teknik Bangunan Universitas Negeri Malang tidak lulus tepat waktu. Skripsi.

[5] Zajacova, A., Lynch, S.M. & Espenshade, T.J. Self-Efficacy, Stress, and Academic Success in College. Res High Educ 46, 677–706 (2005). https://doi.org/10.1007/s11162-004-4139-z

[6] Musa, A. E., Studi, P., Akuntansi, P., Keguruan, F., Ilmu, D. A. N., & Surakarta, U. M. (2019). SKRIPSI PADA PROGRAM STUDI PENDIDIKAN AKUNTANSI. Universitas Muhammadiyah Surakarta.

[7] Febrina, L. (2018). Hubungan Antara Harapan Lulus Tepat Waktu dan Regulasi Diri pada Mahasiswa yang Aktif Berorganisasi. Skripsi.

[8] Anugerah, A. S. P., Indriati, & Dewi, C. (2017). Implementasi Algoritme Fuzzy K-Nearest Neighbor untuk Penentuan Lulus Tepat Waktu (Studi Kasus: Fakultas Ilmu Komputer Universitas Brawijaya). Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2(4), 1726–1732.

[9] Nikmatun, I. A., & Waspada, I. (2019). Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor. Jurnal SIMETRIS, 10(2), 421–432.

[10] Zainuddin, M. (2018). Metode Klasifikasi Berbasis Particle Swarm Optimization (Pso) Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa. Jurnal Ilmiah Teknologi Informasi Asia, 13(1), 1. https://doi.org/10.32815/ji.201811247

[11] Handayani, L., & Maulida, E. L. (2015). Perkiraan Waktu Studi Mahasiswa Menggunakan Metode Klasifikasi Dengan Algoritma Naive Bayes. Seminar Nasional Teknologi Informasi Komunikasi Dan Industri, November, 138–147.

[12] Meinanda, M. H., Annisa, M., Muhandri, N., & Suryadi, dan K. (2009). Prediksi masa studi sarjana dengan artificial neural network. Interntetworking Indonesia Journal, 1(2), 31–35.

[13] Desiani A dan Arhami M. (2006). Konsep Kecerdasan Buatan. Yogyakarta (INA): Andi Graha Ilmu.

[14] Heaton, J. (2012). Introduction to the Math of Neural Networks. https://doi.org/10.1016/j.neunet.2004.10.001

[15] Kasabov, N. K. (1997). Foundations of neural networks, fuzzy systems, and knowledge engineering. In Choice Reviews Online (Vol. 35, Issue 01). https://doi.org/10.5860/choice.35-0330

[16] Daniel G. 2013. Principles of Artificial Neural Network. Advanced Series in Circuits and System – Vol. 7. World Scientific Publishing Co. Pte. Ltd. USA

[17] Art, T., & Prologue, T. (2000). MACHINE LEARNING The Art and Science of Algorithms that Make Sense of Data As.

[18] Chakraverty, S & Susmita M. 2017. Artificial Neural Networks for Engineers and Scientists Solving Ordinary Differential Equations. CRC Press.

[19] Neapolitan, R. E., & Neapolitan, R. E. (2018). Neural Networks and Deep Learning. In Artificial Intelligence. https://doi.org/10.1201/b22400-15

[20] James M.K., Derong L & David B.F. (2016). Fundamentals of Computational Intelligence Neural Networks, Fuzzy Systems, and Evolutionary Computation. John Wiley & Sons. New Jersey

[21] Kusumadewi S. 2003. Artificial Intelligence (Teknik dan Aplikasinya). Yogyakarta (INA): Graha Ilmu.

[22] Mirjalili, S. (2019). Evolutionary Algorithms and Neural Networks. Springer.

[23] Turner, B. M., Forstmann, B. U., & Steyvers, M. (2019). Joint Models of Neural and Behavioral Data. https://doi.org/10.1007/978-3-030-03688-1

[24] Czum, J. M. (2020). Dive Into Deep Learning. Journal of the American College of Radiology, 17(5), 637–638. https://doi.org/10.1016/j.jacr.2020.02.005

[25] Putra, J. W. G. (2019). Pengenalan Konsep Pembelajaran Mesin dan Deep Learning. 1–235.