JSiI | Jurnal Sistem Informasi
Vol. 12 No. 1 | Maret 2025, Hal. 25-32

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

Prediction Model For Students' On-Time Graduation Using Algorithm Support Vector Machine (SVM) Based Particle Swarm Optimization (PSO)

Syarif Hidayatulloh[1], Gandung Triyono [2], Kiki Ari Suwandi kosasih[3]

[1,2,3]*Program Studi Magister Ilmu Komputer, Fakultas Teknologi Informasi,Universitas Budi Luhur*
Jl. Ciledug Raya, RT.10/RW.2, Petukangan Utara, Kec. Pesanggrahan, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12260

[1]2211601154@student.budiluhur.ac.id
[2]gandung.triyono@budiluhur.ac.id
[3]2211600834@student.budiluhur.ac.id

Abstract

One of the indicators in the assessment of the tridharma of university output and achievement is the timeliness of student graduates, but some students experience delays in completing their studies, the number of active semesters that exceed the normal limit, the grades of courses that have not passed, GPA and social studies that are substandard, and insufficient total credits cause some students to experience delays in completing their studies, As well as the limited information and assistance of study programs for students, this is also a factor. In this study, we developed a model of predicting students' timely graduation using Support Vector Machine (SVM) and optimized with Particle Swarm Optimization (PSO) or the selection of Information Gain features to improve prediction accuracy. The selected attributes include Student Achievement Index 1 up to Student Achievement Index 4, Grade Point Average 1 up to Grade Point Average4, as well as Semester Credit Units 1 and Semester Credit Units 4, the model achieves an accuracy of 0.799, precision of 0.851, recall of 0.605 and AUC of 0.86. This approach shows a significant performance improvement compared to the SVM method without optimization. This can help the program in finding students who are at risk of not graduating on time and starting the warning early.

Keywords: On-time graduation, students, classification, SVM, PSO, Information Gain.

I. INTRODUCTION

One of the indicators in the assessment of the tridharma of higher education outputs and achievements is the timely graduation of students [1]. According to the Regulation of the Minister of Research, Technology and Higher Education Number 44 of 2015 concerning National Standards for Higher Education, the learning period for undergraduate students must not exceed seven years [2], Therefore, as part of improving academic quality, colleges strive to increase the number of students who graduate on time.

According to the 2021–2026 Strategic Plan, the Faculty of Islamic Religion and Teacher Education (FAIPG) sets a target for each study program to have 75% of students graduate on time. However, data shows that this achievement is still fluctuating. In the academic period 2018/2019 to 2022/2023. The percentage of on-time graduation varies, with upward and downward trends; In the 2022/2023 academic year, this percentage reached 67% However, this percentage has not met the indicator target with a percentage value of 73% for 2022/2023 [3].

Data-driven methods are needed to solve the problem of predicting students' graduation times. Data mining Data mining means extracting implicit information that is unknown and may be useful from data [4]. Data mining is widely used in various fields, such as marketing that uses data mining methods for customer relationship management [5], in the industrial field to help make decisions about clothing patterns [6], in the field of Education, which uses data mining to analyze students in predicting achievement and so on. In predictive analysis, data mining can be used to find patterns and components that affect student graduation rates. Previous studies have used a variety of classification techniques, including Naive Bayes, C4.5, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM), with varying degrees of accuracy.

In several previous studies, classification has been used to predict students' on-time graduation by using various classification methods, including in predicting student graduation using the C4.5 algorithm, error-based pruning, confidence with a number of 3 attributes including Regional Origin, GPA, TOEFL, and Study Length with an accuracy of 60.52% [7], Previous research that compared C4.5, Naïve Bayes, KNN, and SVM Algorithms in predicting students' grades and graduation time with a total of 5 attributes including JK, IPS3, IPS4, IPS5, IPS6, and produced the best accuracy, namely using the Naïve Bayes algorithm with an accuracy of

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

76.79% [8], Another study in predicting student graduation the algorithm used is K-Nearest Neighbor (MK-NN) attributes used amounting to 6 attributes including IPS1, IPS2, IPS3, IPS4, IPS5, IPK5, with an accuracy level of 84% [9]. *Support Vector Machine (SVM) has a competitive performance in the classification of academic data*. Another study also showed an increase in SVM accuracy with Particle Swarm Optimization (PSO) optimization, which succeeded in increasing the accuracy from 85.81% to 86.43% [10]. Based on the findings of this previous research, this study will concentrate on the application of the Particle Swarm Optimization (PSO)-based Support Vector Machine (SVM) method to improve the accuracy of students' on-time graduation predictions.

In this study, there are main contributions, namely the *Support Vector Machine* (SVM) method based on *Particle Swarm Optimization* (PSO) to predict the graduation time of students, the use of relevant academic attributes based on previous studies, the evaluation of model performance in improving prediction accuracy compared to other approaches that have been used previously. This research is expected to help universities create academic strategies that will increase the number of students who graduate on time.

II. METHODS

In this research, the Cross Industry Standard Process for Data Mining (CRISP-DM) method is used as the basis for the steps of this research. This method follows a 6 (six) phase process, as shown in Figure 1.

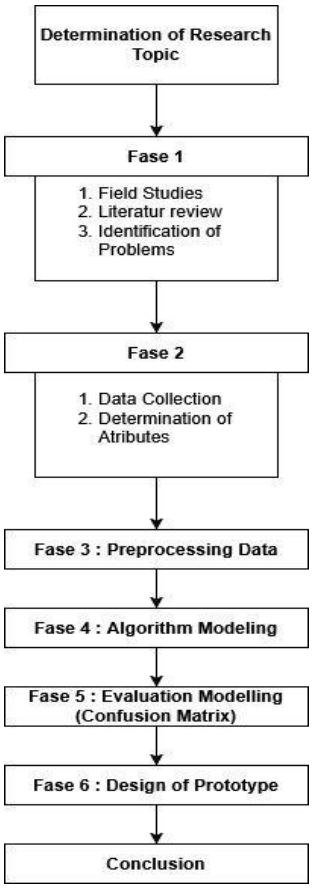

Figure1. steps of the research

The data lifecycle of a data mining project can be described with the current data mining process model. This contains the project phases, tasks and the relationships between them. It is not possible to identify all relationships at this level of description. The goals, background, and interests of the user, as well as most importantly the data, determine how the data mining tasks relate to each other [11].

Fase 1 (*Business Understanding)*
This initial phase centers on understanding the project goals and requirements from a business point of view. Next, this information is turned into data mining to define the problem and an initial plan to achieve the goal [11].

Fase 2 (Data Understanding)
The data understanding phase begins with initial data collection and continues with actions to familiarize yourself with the data, find quality issues, discover initial insights, or find interesting subsets to hypothesize about hidden information [11].

Fase 3 (Data Preparation)
All actions taken to create the final dataset, or the data to be fed into the modeling tool, from the initial raw data fall under the data preparation stage. The data preparation stage is usually performed repeatedly, and includes the selection of tables, records, and attributes, as well as the transformation and cleaning of data for the modeling tool [11].

Fase 4 (Modeling)
At this stage, various modeling methods are selected and applied, and their parameters are adjusted to ideal values. Many different techniques are used to solve the same data mining problem; some techniques require a specific form of data. Therefore, it is often necessary to return to the data preparation stage [11].

At this stage a comparison is made using Algorithms, The following are some classification techniques that basically use classification concepts.

- **Decision Tree**
  Decision Tree is an attempt to find a noise-resistant classification model is one of the popular and commonly used classification methods. Iterative Dichotomiser Version 3 (ID3) is one of the most popular decision tree methods; other popular variants are C4.5 and ASSISTANT. The ID3 method attempts to build a top-down decision tree classification model, by scoring each attribute using a statistical measure, usually information gain, to quantify the effectiveness of an attribute in classifying the set [12].

$$Entropy(S) = \sum_{i}^{c} \quad - P_i P_i$$

- **Support Vector Machine**
  To solve the quadratic programming problem, the support vector machine (SVM) learns a classification function with two target classes. This chapter discusses the theoretical basis of SVMs that bring quadratic programming problems to learn classifiers. After that, we introduce the SVM formulation for linear classifiers and linearly separable

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

problems, which is followed by the SVM formulation for linear classifiers and nonlinear and nonlinearly separable problems based on kernel functions. We also offer a method of applying SVM to classification functions involving more than two target classes. Data mining software that supports SVM is included here. Some SVM applications have references [13].

Support Vector Machine (SVM) is an effective classification technique for nonlinear problems introduced by Vapnik in 1992. SVM attempts to find a hyperplane by maximizing the class distance [12].

Transforming the data set from the input space to a larger feature space can generally be described as[12].

$$\Phi: R^p \rightarrow R^q, dimana\ p < q$$

The above concept can be applied to the training data set with labels $x_i \in R^d$ and class labels with $y_i \in \{-1, +1\}$ to $i = 1,2, \dots, l$ where l is the amount of data. For example, it is assumed that the hyperplane is of dimension d, which is defined as

$$w.x + b = 0$$

A data x_i is classified as class -1 if

$$w.x_i + b \leq -1$$

And classified as class +1 if

$$w.x_i + b > 1$$

Maximizes the distance between the hyperplane and its closest point, i.e. $\frac{1}{\|w\|}$, is the best way to find the largest margin. This can be thought of as a Quadratic Programming problem (*QP*), which is to find the minimum point of.

$$min\ \tau(w) = \frac{1}{2}\|w\|^2$$

With Restrictions

$$y_i(x_i.w + b) - 1 \geq 0, \forall i$$

The Lagrange multiplier is one of the many methods that can be used to solve this problem

$$L(w, b, a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} a_i(y_i(x_i.w + b) - 1), i = 1,2, \dots,1$$

Where is the lagrange multiplier $a_i \geq 0$. The optimal value of the equation can be found by minimizing L against *w* and *b* and maximize *L* against a_i. Because the gradient optimal point *L* = o, The equation can be changed by maximizing *L* against a_i.

$$\sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i,i=1}^{l} a_i a_j y_i y_j x_i x_j$$

Subject to:

$$a_i \geq 0 (i = 1,2, \dots l) \sum_{i=l}^{l} a_i y_i = 0$$

A number of a_i with positive values are generated by this maximization. The support vector (sv) is the data associated with these a_i positiva.

- ***k-Nearest Neighbour Rule***

  This method works by finding the patterns or data objects that are closest to the input pattern; then, the class with the most patterns among the k patterns is selected.

  Here is the algorithm of kNNR

  - Training pattern <x,f(x)> the pattern is added to the training pattern
  - Input pattern x_q

    Suppose $x_1, x_2 \dots x_{k_}$k are the k patterns that have the closest distance to x_q.

    The decision class should be returned to the class with the most patterns out of the k patterns.

- **Bayes Classification**

  This method uses Bayes' theorem, which was discovered in the 18th century. Bayes' theorem describes probability, or conditional probability, as:

$$P(X) = \frac{P(H)P(H)}{P(X)},$$

  Where x is evidence, *H* is a hypothesis, *P(H|X)* is the probability that hypothesis *H* is true for evidence *X*, or in other words, the probability that evidence *X* is true for hypothesis *H* conditional on *X*, *P(X|H)* is the probability that evidence *X* is true for hypothesis *H* or the probability that evidence *X* is posterior to condition *H*, *P(H)* is the prior probability of hypothesis *H*, and *P(X)* is the prior probability of evidence *X*.

Furthermore, exploration is also carried out by adding the Particle swarm optimization algorithm, according to previous research, the Particle Swarm Optimization algorithm can increase the accuracy value [10].

Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population-based search algorithm based on simulating the social behavior of birds in flocks. Particles in PSO "flew" through the search space. Changes in a particle's position in the search space depend on the socio-psychological tendency of an individual to mimic the success of others. Therefore, the experience or knowledge of a particle's neighbors influences changes to a particle in the flock. A particle's search behavior is also influenced by other particles in the swarm (PSO is a cooperative symbiotic algorithm). The stochastic search of particles back to areas they have previously encountered in the search space is a result of modeling this social behavior [14].

$$x_i(t + 1) = x_i(t) + v_i(t + 1)$$

Fase 4 (Evaluation)

Before proceeding to the final implementation of the model, it is very important to thoroughly evaluate the model and review all the steps taken to build it to ensure that it can effectively achieve the business objectives. The main objective is to find out if there are any significant business issues that have not been adequately considered. At the end of this phase, decisions should be made about how the mined data will be used [11].

In the evaluation stage, Confusion Matrix is used in this study to test the model and evaluate the classification performance. It is a useful tool to evaluate the ability of your classifier to find tuples of different classes

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

Table 1. Classification model evaluation measure

No	Size	Formulas
1	Accuracy	$\frac{TP+TN}{P+N}$
2	Error Rate	$\frac{FP+FN}{P+N}$
3	Recal	$\frac{TP}{P}$
4	Specificity	$\frac{TN}{N}$
5	Precision	$\frac{TP}{N}$
6	*F or* F_1 *or F-score*	$\frac{2x\ precision\ recall}{precision+recall}$
7	F_{β},*where* β is a non-negative rill.	$\frac{(1+\beta^2)\ x\ precision\ x\ recall}{\beta^2 precision+recall}$

Fase 5 (Deployment)

Even if the purpose of the model is to increase knowledge about the data, the creation of the model is not the end of the project. The client must be able to use this information properly. This usually involves using the model "live" in the organization's decision-making process such as reviewing marketing databases or changing web pages live. The deployment process can be simple by creating reports or complex by implementing an iterative data mining process across the enterprise, depending on the need. The customer usually does the deployment rather than the data analyst. Nevertheless, the customer should know what needs to be done upfront to utilize the created model [11].

III. RESULT AND DISCUSSION

In this research, a classification algorithm performance evaluation is carried out in making a prediction model for student on-time graduation. The data used is data on students graduating from the 2018/2019 academic year to the 2022/2023 academic year, lecture activity data, student active status data and class data so that from these data a dataset of 1270 data is generated with a total of 16 attributes including Gender, Class, Semester Achievement Index from semester 1 to semester 4 and GPA from semester 1 to semester 4, SKS semester 1 to semester 4, Total Active Status of students and also the pass label. At the comparison stage, the support vector machine (SVM) model based on Particle Swarm Optimization (PSO) and *information gain* selection features with selected features, namely Semester Achievement Index 1, Semester Achievement Index 2, Semester Achievement Index 3, Semester Achievement Index 4, Grade Point Average (GPA) 1, Grade Point Average (GPA) 2, Grade Point Average (GPA) 3, Grade Point Average (GPA) 4, Semester Credit Units 1, and Semester Credit Units 4 become the best model, and can predict graduation on time. At the evaluation stage, the support vector machine (SVM) algorithm model based on Particle Swarm Optimization (PSO) can produce an Accuracy value: 0.80, Precision: 0.85 Recall: 0.61 AUC: 0.86.

Business Understanding

At this stage, the problems related to the on-time graduation of students faced by the Faculty of Islamic Religion and Teacher Education, there is a problem that the study program does not have sufficient information related to the potential of students to graduate on time. The limitation of the faculty in assisting students in completing graduation on time results in the number of students who do not graduate on time is greater than those who graduate on time, this problem must be addressed because the number of students who do not graduate on time can also have an impact on the assessment of the success of student studies in university accreditation. In addition, it also has an impact on the target of the student on-time graduation indicator stated in the strategic plan of the Faculty of Islamic Religion and Teacher Education Although there are many reasons why a student fails to complete his education on time, the most obvious reason is still unknown. Processing and extracting data that is still hidden is necessary to generate new information and knowledge that can be used to address student issues during the current academic year.

Based on the existing problems, the problems can be formulated, namely the absence of a prediction model for students' on-time graduation to provide adequate information regarding potential, the lack of implementation of student study assistance in completing students' on-time graduation in completing students' on-time graduation. In research conducted by previous researchers, this classification algorithm can solve these problems and the Faculty can get information related to the potential of students who will graduate on time and provide assistance to students in completing their studies on time by using *machine learning* technology, especially data mining (data mining) by applying classification algorithms.

Data Understanding

The initial process in developing a student on-time graduation prediction model by collecting data from the academic department, the data includes 3 data, namely graduation data for the 2018/2019 academic year to 2022/2023, student activity data, student active status data and student class data.

The results of data collection obtained 1270 records of data from graduation data with 10 attributes including Number, Student Identification Number, Name, Study Program, M/F, Status, Exit Date, Diploma Number, Professional Certificate No, Remarks. then obtained also student activity data from the period 2012/2013 Odd to 2022/2023 Even with a total of 13 attributes including Number, Student Identification Number, Name, Study Program, Student Status, Semester Achievement Index, Total semester credits, Grade Point Average, Total credits, Tuition fees. Data on Active Status of lectures from class years 2013-2019 which has 6 attributes including Number, Student Identification Number, Name, Entry Period, Status, Lecture Period. Then also obtained Student class data with 8 attributes including Number, Student Identification Number, Student Name, Gender, Place of Birth, Date of Birth, Status, and Class.

From the results of the literature study conducted by referring to several previous studies, the resulting attributes will be 1271 datasets with 16 attributes. After conducting interviews with experts in this case the vice dean of academic affairs, and references to previous research, 16 attributes were obtained that would be used, Gender, Class, Semester Achievement Index (SAI) from semester 1 to semester 4 and Grade Point Average from semester 1 to semester 4, Semester Credit Units 1 to semester 4, Total Active Status of students and also the graduation label.

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

Data Preparation

At this stage, the data cleaning process has been carried out to make the data used more qualified and ready for analysis. This data will be prepared according to the needs of the analysis using predetermined features. From the 1271 Dataset data, 31 missing or noisy data were found in the Class attribute totaling 2 missing value data, Semester Achievement Index 1 totaling 4 missing value data, Semester Achievement Index 2 totaling 4 missing value data, Grade Point Average 2 amounted to 3 missing value data, Semester Credit Units 1 amounted to 4 missing value data, Semester Credit Units 2 amounted to 5 missing value data, Semester Credit Units 3 amounted to 1 missing value data, and Semester Credit Units 4 amounted to 1 missing value data then cleaned the data that was missing or noisy on these attributes so that the cleaning data amounted to 1262 datasets.

After data preprocessing and data *preparation* , a dataset consisting of 16 attributes and 1261 data was produced. The division of the dataset is done using the stratified random sampling The first stage of the data is sorted based on the label Pass on Time and Pass Not on Time, then the percentage of each label is calculated. After that the data is divided into two consisting of 70% training data and 30% test data. The data is selected randomly while still paying attention to the label composition.

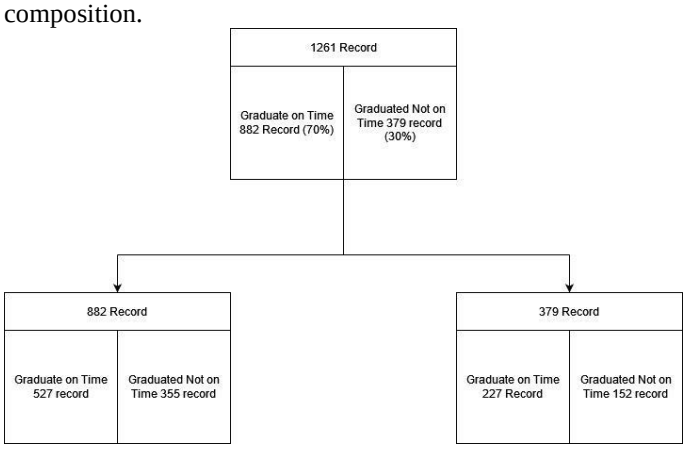

Figure 2. Composition of Data (Stratified Random Sampling)

From the results of data preprocessing and data preparation, the data composition is shown in Figure 2. The total number of data is 1261 *records.* The resulting data passed on time as many as 882 records and passed not on time as many as 379 records. Then the data is divided into 70% for training data, and 30% into test data. In the training data generated 882 records with the composition of the class passed on time 527 records and passed not on time 355 records, then for the training data generated 379 records with the composition of class data 227 to pass on time and 152 records for data passed not on time.

In this process, a comparison of the methods to be used is also carried out, the methods compared are 4 methods, namely Decision Tree, Support Vector Machine, K-Nearest Neighbor, Naive Bayes. The results of the accuracy comparison carried out using python on the jupyter notebook, can be seen in table 2.

Table 3. Comparison Results of Algorithm Models before using Particle swarm optimization and Feature Selection

No	Methods	*Accuracy*	*Recall*	*Precision*	*AUC*	*F1 Score*
1	Decision Tree	0.673	0.625	0.586	0.665	0.605
2	K-Nearest Neighbor	0.752	0.618	0.723	0.834	0.667
3	Support Vector Machine	0.710	0.309	0.904	0.818	0.461
4	Naive Bayes	0.747	0.493	0.798	0.784	0.610

The results of the comparison of algorithm models without using feature selection resulted in the best algorithm being K-Nearest Neighbor with an Accuracy value of 0 .752 , Recall 0.618, Precision 0.723, AUC 0.834 and f1 Score 0.667.

In exploring the comparison of algorithm models, feature selection is also carried out with the following types: PCA, Information gain, Chi Square, Forward Selection, and backward elimination [15].

Table 4. Comparison Results of Algorithm Models using feature selection before using particle swarm optimization

No	Methods	*Feature Selection Method*	*Accuracy*	*Recall*	*Precision*	*AUC*	*F1 Score*
1.	Decision Tree	PCA	0.736	0.658	0.676	0.723	0.667
2.	KNN	PCA	0.739	0.625	0.693	0.831	0.657
3.	SVM	PCA	0.797	0.605	0.844	0.851	0.705
4.	Naive Bayes	PCA	0.741	0.467	0.807	0.785	0.592
5.	Decision Tree	Information Gain	0.723	0.618	0.667	0.706	0.642
6.	KNN	Information Gain	0.755	0.651	0.712	0.821	0.680
7.	SVM	Information Gain	0.702	0.296	0.882	0.813	0.443
8.	Naive Bayes	Information Gain	0.739	0.474	0.791	0.779	0.593
9.	Decision Tree	Chi-Square	0.662	0.605	0.575	0.653	0.590
10.	KNN	Chi-Square	0.728	0.618	0.676	0.780	0.646
11.	SVM	Chi-Square	0.720	0.368	0.848	0.806	0.514
12.	Naive Bayes	Chi-Square	0.736	0.480	0.777	0.764	0.593
13.	Decision Tree	Forward Selection	0.697	0.664	0.612	0.691	0.637
14.	KNN	Forward Selection	0.770	0.632	0.756	0.837	0.688
15.	SVM	Forward Selection	0.702	0.316	0.842	0.804	0.459
16.	Naive Bayes	Forward Selection	0.739	0.493	0.773	0.778	0.602
17.	Decision Tree	Backward Elimination	0.689	0.605	0.613	0.674	0.609
18.	KNN	Backward Elimination	0.763	0.625	0.742	0.827	0.679
19.	SVM	Backward Elimination	0.712	0.303	0.939	0.812	0.458
20.	Naive Bayes	Backward Elimination	0.744	0.474	0.809	0.786	0.598

Then after exploring the model with the addition of selection features, namely PCA, Information gain, chi square, Forward selection and Backward elimination, the best model was produced with an accuracy value of 0.797, Recall 0.605, precision 0.844, AUC 0.851, and f1-score 0.705, namely support vector machine (SVM) with Feature selection PCA.

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

Furthermore, exploration is also carried out by adding the *Particle swarm optimization* algorithm, according to previous research, the Particle Swarm Optimization algorithm can increase the accuracy value [10]. comparison results can be seen in table 4.

Tabel 5. Comparison Results of Support Vector Machine (SVM) Algorithm Model Feature Selection and Particle swarm optimization

No	Methods	*Feature Selection Method*	*Accuracy*	*Recall*	*Precision*	*AUC*	*F1 Score*
1.	Decision Tree	PCA	0.752	0.553	0.764	0.791	0.641
2.	KNN	PCA	0.784	0.559	0.850	0.857	0.675
3.	SVM	PCA	0.799	0.625	0.833	0.848	0.714
4.	Naive Bayes	PCA	0.741	0.467	0.807	0.785	0.592
5.	Decision Tree	Information Gain	0.765	0.658	0.730	0.759	0.692
6.	KNN	Information Gain	0.760	0.546	0.790	0.860	0.646
7.	SVM	Information Gain	0.826	0.664	0.871	0.863	0.754
8.	Naive Bayes	Information Gain	0.739	0.474	0.791	0.779	0.593
9.	Decision Tree	Chi-Square	0.726	0.513	0.722	0.746	0.600
10.	KNN	Chi-Square	0.755	0.566	0.761	0.806	0.649
11.	SVM	Chi-Square	0.739	0.493	0.773	0.812	0.602
12.	Naive Bayes	Chi-Square	0.736	0.480	0.777	0.764	0.593
13.	Decision Tree	Forward Selection	0.765	0.612	0.756	0.776	0.676
14.	KNN	Forward Selection	0.784	0.605	0.807	0.849	0.692
15.	SVM	Forward Selection	0.781	0.559	0.842	0.840	0.672
16.	Naive Bayes	Forward Selection	0.739	0.493	0.773	0.778	0.602
17.	Decision Tree	Backward Elimination	0.728	0.612	0.679	0.788	0.644
18.	KNN	Backward Elimination	0.794	0.605	0.836	0.851	0.702
19.	SVM	Backward Elimination	0.797	0.638	0.815	0.856	0.716
20.	Naive Bayes	Backward Elimination	0.744	0.474	0.809	0.786	0.598

From the results of the comparison of algorithm models by adding the *Particle Swarm Optimization* (PSO) algorithm, the best algorithm based on particle swarm optimization with an *Accuracy* value of 0.815, *Recall* 0.645, *Precision* 0.860, AUC 0.865 and *F1 Score* 0.737 is the *Support Vector Machine* algorithm selection feature using *Information gain.* At the algorithm comparison stage, it can be concluded that the algorithm model to be used is the *Particle Swarm Optimization* (PSO)-based *Support Vector Machine* (SVM) algorithm with the *information gain* selection feature .

The results of the evaluation showed that the selection of SVM with the selection of the Gain Information feature had the best performance. With an accuracy of 81.5%, the model shows a good ability to classify the data as a whole. Nonetheless, the Recall value is lower than Precision at a ratio of 64.5% and 86%. shows that the model is better at identifying positive classes with high accuracy but still has some errors in capturing all the truly positive data. This can be overcome by adjusting hyperparameters or considering data balancing techniques in case of class imbalances. In addition, the AUC value (0.865) indicates that the model has good predictive ability in distinguishing between positive and negative classes. Meanwhile, a higher F1-Score (0.737) compared to Recall shows that there is a balance between precision and the model's ability to capture positive data well. Particle Swarm Optimization (PSO) was used in this study to find the optimal parameters in the SVM algorithm, which has an important role in finding the optimal class separation margin. Using PSO, the C (regularization parameter) and γ (kernel coefficient) parameters in the SVM can be adjusted automatically, so that the model is better able to handle non-linear data. PSOs find a better combination of parameters than manual or default searches, thus improving the generalization capabilities of the model. Feature selection using Information Gain plays an important role in improving model performance. Information Gain helps in selecting the attributes that have the highest contribution to classification, thus: Reducing Overfitting, Speeding up computational processes, Improving Accuracy

Modeling

At this stage the modeling stage of the algorithm used is the Support Vector Machine (SVM) based on *Particle Swarm Optimization* (PSO). From the optimization, the value of C (Penalty Parameter) is 7.8419 and also Gamma (Kernel Coefficient) 0.059. Where the C Parameter controls the penalty for misclassification. A higher value of C means the model tries harder to classify all the training examples correctly. and the Gamma parameter determines the extent to which the influence of one training example is far reaching. A low value means 'far' and a high value means 'near'.

Evaluation

The results of the evaluation carried out in this study by looking for accuracy values, in calculating the accuracy value, stratified random sampling is also used with 70% data division for training data and 30% for testing data and also Confusion Matrix. The results of the calculation process above can produce accuracy, precision, recall, and AUC values can be seen in the table below.

Tabel 6. Confusion Matrix

		On Time	Not on time
Actually	On Time	211	16
	Not on time	60	92
		On Time	Not on time

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

After evaluation or testing, the accuracy, precision, recall, and AUC values are as follows:

Table 7. Result Confusion Matrix

NO	Metrics	Value
1	Accuracy	0.80
2	Precision	0.85
3	Recall	0.61
4	AUC	0.86

Deployment

At the *deployment* stage, this is done by creating a prototype based on the results of the selected model, namely the *Support Vector Machine* (SVM) model. The following is a picture of the results of the prototype that was deployed.

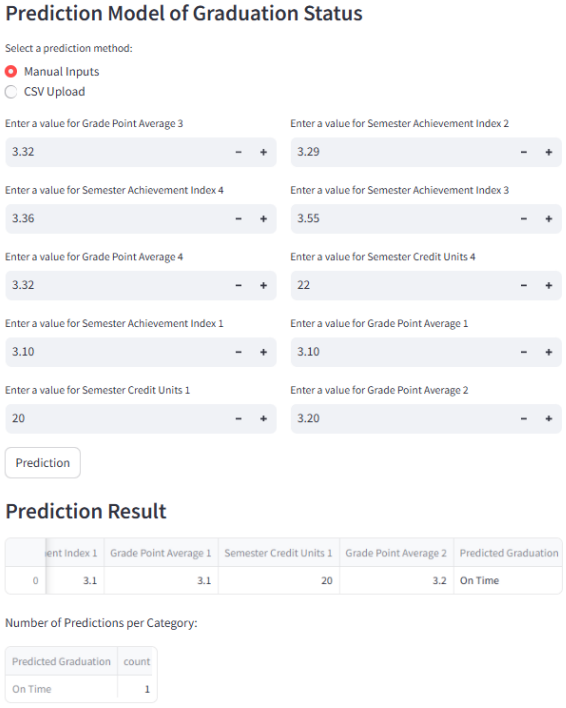

Figure 3. Prediction Application Model

IV. CONCLUSION

In this study, the information gain feature showed that it succeeded in filtering attributes that affect the prediction of students' on-time graduation. From the initial number of attributes, the selected attributes were Student Achievement Index 1 up to Student Achievement Index 4, Grade Point Average 1 up to Grade Point Average4. Particle Swarm Optimization (PSO) optimizing the Support Vector Machine Model (SVM) is proven to significantly improve performance with an optimal parameter of c of 7.8419 and a gamma parameter of 0.059. After the implementation of PSO, the model achieved an accuracy of 0.799, precision of 0.851, recall of 0.605, and an AUC of 0.86. This research shows that the performance of academic prediction models can be improved with optimization techniques such as PSO, which can help educational institutions find students who are at risk of not graduating on time. Institutions can make more accurate predictions to perform more beneficial actions, such as better academic guidance. As for this study, the dataset used is limited to academic attributes without considering non-academic factors, student motivation, social support, or economic conditions, the model developed is only tested on one dataset, so the generalization of results to other institutions needs to be studied further.

V. SUGGESTIONS

The results of this research can be used to improve future research, and prediction models can be applied in the academic world. In addition, for further research, use alternative methods to improve accuracy other than Particle Swarm Optimization (PSO). These systems can be integrated into university academic information systems to ensure these prediction models are useful in the real world. Advanced research should test this model on datasets from different universities with a variety of academic systems to ensure that it is reliable and can be applied to different institutions.

REFERENSI

[1] L. A. Mandiri, "Lampiran Peraturan BAN-PT No 10 Tahun 2021 tentang Instrumen Akreditasi Program Studi pada Program Sarjana Lingkup Kependidikan." 2021.

[2] T. Kementerian Riset, "Standar Nasional Pendidikan Tinggi (SN Dikti)," *Produk Hukum*, vol. 49. pp. 21–23, 2015.

[3] F. Agama, I. Dan, P. Guru, and U. Djuanda, "Rencana strategis 2021-2026." 2021.

[4] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. 2016.

[5] L. Zahrotun, "Implementation of data mining technique for customer relationship management (CRM) on online shop tokodiapers.com with fuzzy c-means clustering," in *Proceedings of the 2017 2nd International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2018, pp. 299–303. doi: 10.1109/ICITISEE.2017.8285515.

[6] D. Sartika and D. I. Sensuse, "Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian," *Jatisi*, vol. 1, no. 2, pp. 151–161, 2017.

[7] R. P. S. Putri and I. Waspada, "Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 4, no. 1, pp. 1–7, 2018, doi: 10.23917/khif.v4i1.5975.

[8] S. Widaningsih, "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4.5, Naïve Bayes, Knn Dan Svm," *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.

[9] I. D. Larasati, A. A. Supianto, and M. T. Furqon, "Prediksi Kelulusan Mahasiswa Berdasarkan Kinerja Akademik Menggunakan Metode Modified K-Nearest Neighbor (MK-NN)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 5, pp. 4558–4563, 2019.

Received: Jan 21st, 2025 | Published: March 21st, 2025

DOI 10.30656/jsii.v11i1.6936
p-ISSN: 2406-7768
e-ISSN: 2581-2181

[10] Suhardjono, W. Ganda, and H. Abdul, "Prediksi Kelulusan Menggunakan Svm Berbasis Pso," *Bianglala Inform.*, vol. 7, no. 2, pp. 97–101, 2019.

[11] P. C. N. et al., "Crisp-DM," *SPSS Inc.*, vol. 78. pp. 1–78, 2000.

[12] Suyanto, *Data Mining Untuk Klasifikasi dan Klasterisasi Data*, Cetakan Pe. Informatika Bandung, 2019.

[13] N. Ye, *Data Mining: Theories, Algorithms, and Examples*. CRC Press, 2013.

[14] A. P. Engelbrecht, "Chapter 16 - Particle Swarm Optimization," in *Computational Intelligence: An Introduction*, 2007, pp. 289–358.

[15] R. Chauhan and H. Kaur, *Predictive Analytics and Data Mining: A Framework for Optimizing Decisions with R Tool*. 2013. doi: 10.4018/978-1-4666-4940-8.ch004.

Received: Jan 21st, 2025 | Published: March 21st, 2025