International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Performance Graduation Student Predicting Using One-Class Support Vector Machine Algorithm

[1]M. Ramaddan Julianti, [2]Yaya Heryadi, [3]Budi Yulianto, [4]Widodo Budiharto

Submitted:12/03/2024 **Revised**: 27/04/2024 **Accepted**: 04/05/2024

Abstract: This study explores the prediction of student graduation performance using the One-Class Support Vector Machine (OCSVM) algorithm. The objective is to accurately forecast the time and success rate of students graduating from academic programs. Predicting student performance has become increasingly vital for educational institutions aiming to improve retention rates and support academic planning. The research employs the OCSVM due to its effectiveness in handling imbalanced datasets, which are common in academic performance data. By focusing on a single class, the algorithm can detect anomalies and patterns that signify potential delays or failures in graduation. The dataset comprises various academic and demographic attributes of students from a private university in Indonesia. Data preprocessing techniques such as normalization and transformation were applied to enhance the model's accuracy. The results demonstrate that the OCSVM algorithm can effectively predict student graduation performance with a high degree of accuracy, offering educational institutions a robust tool for early intervention. This approach not only helps in identifying at-risk students but also facilitates the development of targeted support strategies, thereby enhancing overall academic outcomes.

Keywords*: Algorithm; one-class support vector machine; prediction of student*

I. Introduction

Predicting student graduation performance is a critical task for educational institutions to improve outcomes and operational efficiency. With the increasing volume of student data, advanced data mining techniques are being employed for analysis and prediction[1]. While Support Vector Machines (SVM) have shown significant promise in classification tasks due to their robustness and high accuracy, other methods like deep learning models are also gaining traction for predicting academic outcomes based on diverse student data sources[2]. Additionally, studies have highlighted the importance of utilizing machine learning models, including SVM, in educational data mining to enhance the learning process and inform decision-making [3]. By leveraging SVM and other advanced techniques, institutions can effectively analyze student data to predict graduation performance, ultimately leading to improved educational practices and student success.

Support Vector Machines (SVM) play a crucial role in enhancing student graduation prediction by providing accurate classification models. SVM techniques have been utilized in various studies to predict outcomes early in academic settings, such as identifying students at risk of performing poorly in courses [4]. Additionally, SVM models have been successfully applied in predicting employment status for college students, considering biases and academic performance, to aid in personalized interventions and early support [5]. Furthermore, SVM models have been used in epidemiological studies with small sample sizes, particularly in biomedical research, to analyze data efficiently, especially when handling binary outcomes and time-to-event data, showcasing superior performance compared to traditional statistical methods [6], [7]. These applications demonstrate the versatility and effectiveness of SVM in enhancing student graduation prediction by leveraging data-driven classification techniques across various educational and health-related domains.

The objective of this research is to develop and evaluate a predictive model using the One-Class Support Vector Machine (OC-SVM) algorithm specifically tailored for forecasting student graduation performance. Unlike traditional SVM, which is commonly used for binary classification, OC-SVM is designed for scenarios where only one class (e.g., students expected to graduate on time) is well-represented in the training data. This makes OC-SVM particularly suitable for identifying students at risk of delayed graduation, as it can effectively distinguish the normal pattern of on-time graduation from anomalies[8].

This research will leverage historical student data, encompassing academic records, demographic information, and other relevant factors, to build a robust predictive model. The ultimate goal is to provide educational institutions with a tool that can aid in early identification and intervention for

[1]*Doctor of Computer Science Bina Nusantara University West Jakarta, Indonesia*
mohammad.julianti@binus.ac.id
[2]*Computer Science Department Bina Nusantara University West Jakarta, Indonesia*
yayaheryadi@binus.ac.id
[3]*Computer Science Department Bina Nusantara University West Jakarta, Indonesia*
budi.yulianto@binus.ac.id
[4]*School of Computer Science Bina Nusantara University West Jakarta, Indonesia*
wbudiharto@binus.ac.id

students who might face challenges in graduating on time, thereby enhancing overall educational outcomes and student success rates [9]. To address the problem of predicting students' passing performance, we propose a comprehensive solution through data collection and processing, selection of the One-Class Support Vector Machine (OC-SVM) algorithm due to its suitability to scenarios where the training data contains only one class and development of an OC-SVM model on the training dataset, focusing on patterns of students graduating on time. The novelty of this research lies in the application of the One-Class Support Vector Machine (OC-SVM) algorithm to the problem of predicting student graduation performance. Although traditional SVM and other machine learning algorithms have been used in educational data mining and use of OC-SVM. This research not only develops the OC-SVM model but also rigorously evaluates its performance against other machine learning algorithms. By providing a comprehensive comparison, the study highlights the strengths and weaknesses of each approach, offering valuable insights for future research and practical applications.

II. Related Work

The prediction of student graduation outcomes and academic performance has been a significant area of research in educational data mining. Various machine learning techniques have been applied to tackle this problem, with Support Vector Machine (SVM) being one of the most prominent methods due to its robustness and accuracy in classification tasks. This literature review provides an in-depth analysis of previous studies that have utilized SVM and other machine learning algorithms for predicting student performance and graduation outcomes.

A. One-class Classsofocation Model

One-class classification model can detect new data by forming a spherical boundary around the data set which can be created flexibly using the kernel function to outlier data or new data in the training dataset which is then described as negative data[10]. One of the one-class classification models used is a one-class support vector machine (OCSVM). The OCSVM model approach was initially proposed simultaneously with the use of a Gaussian kernel, where this kernel can project data in a vector space located on the surface of a unit radius hypersphere centered on the dataset collection. In this case, the OCSVM model can be illustrated geometrically when used in conjunction with a gaussian kernel, i.e. all functions $\exp(-x2)$ have values in the range (0,1) for all $x \in \mathbb{R}$, where the angle between two sample images in the projection space is between 0 and $\pi/2$, so that when the Gaussian kernel is implemented then the entire representation space (not just the target sample) in the projection space lies in a fraction of half the hypersphere. The projection space is as shown in Figure 1.

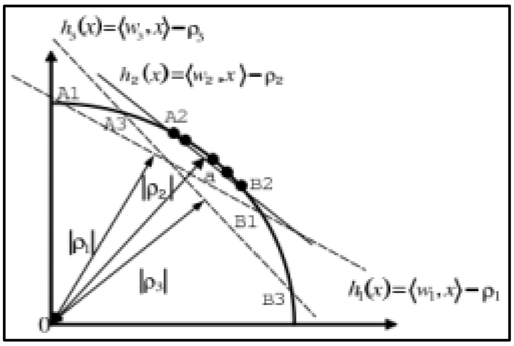

Fig. 1. Geometrical illustration of OCSVM Algorithm

Figure 1 shows the overall projection of the space representation covered by a partial circle of unit radius bounded by the angle $\pi/2$, where the target class samples are represented by solid circles. To separate these points from other points in space, three possible hyperplanes h1(x), h2(x), and h3(x) are displayed, in each hyperplane all data points lie on one side, however, there are several parts of the transformed space where there may be outlier samples.

Of the three hyperplanes, h2(x) is the best visually because it provides the most significant representation of the target data in terms of sample data points and covers the smallest portion of the remaining space And also the hyperplanes h1(x), h2(x) and h3(x) respectively intersect the surface of the unit hypersphere at points (A1, B1), (A2, B2) and (A3, B3)[11]. Support Vector Machines (SVM) have been commonly used as a solution to the binary classifier problem of the prediction process by classifying data on different class labels[12]. The goal of a one-class support vector machine (OCSVM) model is to discover unusual data sets, clean and identify new data sets in the training dataset, and distinguish common or original data classes from new input data sets[13]. One-class SVM is a single-class classification or novelty detection which can classify objects in one class only, by distinguishing the positive class from all other objects[14]. A strategy for carrying out outlier detection based on One-Class Support Vector Machines (OCSVM) results shows high performance on simulation and experimental data in detecting[15]. The one-class support vector machine (OC-SVM) trains the model to produce accurate values and determine the position of the points in vector space, this can be seen in the following equation[16]:

$$\mathrm{Min}_{w,\rho,\zeta}\frac{1}{2}\| w \|^2 + \frac{1}{vN}\sum_{i=1}^{N}\xi i - \rho$$

$$\text{Subject to } \langle w.\Phi(x_i)\rangle \geq \rho - \zeta_i, \forall_i \text{ and } \zeta_i \geq 0, \forall_i \quad (1)$$

The equation above, that N is the number of data samples in the training data set, w is the hyper-plane decision weight value, x_i is the number of training samples ke-i, $\Phi(.)$ is a function that transforms data $X \subseteq \mathbb{R}^d$ from the initial space of points to a new feature space. $F \subseteq \mathbb{R}^d$ which is done by the kernel $\Phi(x_i)$. $\Phi(x_j) = K(x_i, x_j)$, ξ_i is a slack variable that regulates various errors found during the training stage and $v \in [0, 1]$ adjust the outlier data (that is, training data that is outside the vector space) as well as the number of existing support vectors[17].

B. Machine Learning in Education

Machine learning techniques have significantly impacted the field of educational data mining by enabling the analysis of vast datasets to uncover patterns and make predictions [1]. In educational settings, various supervised learning algorithms such as Decision Trees, Random Forests, Logistic Regression, and Support Vector Machines (SVM) are applied to enhance decision-making processes, improve student support systems, and ultimately increase student success rates. Among these algorithms, SVM has shown high performance in solving classification problems in various fields, including bioinformatics[5]. SVM is known for its robustness in classification and regression problems, making it a commonly used algorithm in machine learning and data mining [18]. Additionally, machine learning offers a range of flexible statistical procedures to identify key indicators of a response variable from a large collection of potential predictor variables [3].

Recent studies have highlighted the increasing use of machine and deep learning models to predict academic outcomes based on student-related data, aiming to enhance the learning process comprehensively [4]. These models are trained on data from previous terms to predict outcomes in subsequent terms, showcasing the potential of machine learning in predicting student performance. The application of supervised learning algorithms like SVM, along with other machine learning techniques, has revolutionized educational data mining by enabling the extraction of valuable insights from large datasets to improve decision-making processes and student support systems, ultimately leading to increased student success rates.

C. Support Vector Machine (SVM) in Predicting Student Performance

Support Vector Machine (SVM) is a supervised machine learning algorithm known for its effectiveness in classification and regression tasks due to its ability to handle high-dimensional data and maintain high accuracy[19]. SVM has been widely used in various fields, including bioinformatics, medicine, and anomaly detection, showcasing its versatility and robustness in solving classification problems. In educational data mining, SVM has played a crucial role in predicting student performance by analyzing various student-related data sources. By leveraging SVM, educational institutions can make informed decisions to enhance student support systems and improve academic outcomes. SVM's capability to handle complex datasets and identify patterns makes it a valuable tool for predicting academic success and tailoring interventions to meet individual student needs effectively.

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

Additionally, SVM has been successfully employed in predicting common diseases like diabetes and pre-diabetes, demonstrating its efficacy in diverse domains. The algorithm's ability to discern patterns and make accurate predictions has made it a popular choice for researchers and practitioners seeking reliable predictive modeling solutions. SVM has found applications in anomaly detection scenarios, such as identifying overdose and underdose prescriptions in healthcare settings [20]. The one-class support vector machine (OCSVM) has been particularly effective in detecting anomalies and outliers in data, showcasing its adaptability to different applications beyond traditional classification tasks.

In the realm of academic performance prediction, SVM has been integrated with other machine learning models like Random Forest (RF) and K-Nearest Neighbors (K-NN) in ensemble models to evaluate and enhance academic performance [21]. This integration highlights SVM's versatility in being part of comprehensive predictive analytics frameworks. Support Vector Machine (SVM) is a powerful tool in educational data mining and beyond, offering accurate predictions, effective classification, and robust anomaly detection capabilities. Its application in predicting student performance underscores its significance in enhancing decision-making processes and improving educational outcomes.

D. Machine Learning Algorithms

Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), and Support Vector Machine (SVM) are commonly used machine learning algorithms for predicting student performance [22]. Each algorithm has unique strengths and weaknesses, and their performance is influenced by the dataset characteristics. Random Forest (RF) is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes as the prediction [23]. RF is suitable for handling large datasets with high dimensionality and is less prone to overfitting compared to individual decision trees, making it a robust choice for predicting student performance.

Decision Tree (DT) algorithms are intuitive and easy to interpret, aiding in understanding the factors influencing student performance [24]. While DTs can overfit, techniques like pruning can enhance their performance in predicting student outcomes. Logistic Regression (LR) is a linear model used for binary classification tasks, making it suitable for predicting student success or failure based on input features [25]. LR estimates the probability of a binary outcome by fitting the data to a logistic curve. LR can be effective when the relationship between predictors and outcomes is linear or when interpretability is crucial.

The performance of these algorithms in predicting student performance varies based on the dataset nature and educational context [26]. Random Forest excels in handling complex datasets, Decision Trees offer interpretability, Logistic Regression is suitable for linear relationships, and Support Vector Machine is effective in handling high-dimensional data and non-linear relationships. In educational data mining, selecting the most appropriate algorithm depends on analysis goals, data nature, and desired interpretability level[27]. Researchers and practitioners often experiment with different algorithms to optimize predictive performance and gain insights into student outcomes. Random Forest, Decision Tree, Logistic Regression, and Support Vector Machine are valuable tools for predicting student performance, each offering unique advantages based on the educational data mining task requirements.

III. Methodology

A. *Design Research*

Design Science Research Methodology (DSRM) has five (5) steps, including: (1). Problem Identification: This is the first step in which the scope of the problem is clearly defined. The problem in this research is how to predict graduation on student academic performance achievements in one course. (2). Design Objective of a Solution: In this phase, the researcher defines a problem formulation to provide a solution to the research problem using a one-class support vector machine approach. (3). Design and Development: This stage produces artifacts that build methods or models that will be implemented to solve research problems, in this case, the one-class support vector machine model to predict student graduation in one course. (4). Demonstration: In this stage, the one-class support vector machine (OCSVM) model is implemented to predict student graduation in one course, starting with training and testing on the dataset used, and calculating accuracy and precision values using the confusion matrix approach. (5). Evaluation: This stage is carried out to evaluate its effectiveness in predicting system behavior and measuring system performance using sigmoid kernels which produce accuracy values for system behavior in making predictions. (6). Communication: This stage, the results of the prediction process are published in the form of papers and journals as a contribution to knowledge about one-class support vector machines (OCSVM)[28].

B. *Data Collection and Pre-processing*

- Data Sources: Historical data will be collected from educational institutions, including academic records, demographic information, and other relevant factors. Data sources may include student information systems, academic transcripts, and demographic databases.
- Preprocessing Steps:
- Handling Missing Values: Missing values will be managed using techniques such as imputation or deletion, depending on the extent and nature of the missing data.
- Normalization: Numerical features will be normalized to a standard scale to ensure uniformity across the dataset.
- Encoding Categorical Variables: Categorical variables will be converted into a suitable format for machine learning algorithms using one-hot encoding or label encoding.
- Data Splitting: The dataset will be divided into training (80%) and testing (20%) sets to validate the model's performance [29].

These preprocessing steps are crucial for ensuring the quality and usability of the data for machine learning models.

C. *Algorithm Selection and Model Development*

The One-Class Support Vector Machine (OC-SVM) algorithm is chosen for its suitability in handling one-class data, particularly for identifying anomalies within a well-defined class (e.g., students who graduate on time). A radial basis function (RBF) kernel will be used due to its ability to capture non-linear relationships within the data. The OCSVM model in its implementation requires a kernel function to separate outlier data (anomalous data) from the majority data set (normal data) as seen in Figure 2.

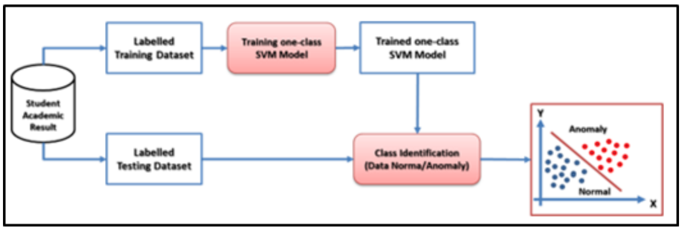

Fig. 2. One-Class Support Vector Machine (OCSVM) Method

The original data set (student pass data on courses) is divided into 2 categories, where data in the pass category is used as training data which is considered separately and data in the fail category will be considered as outlier (anomaly) data. The stages of implementing the OCSVM model begin by considering the data class in the dataset, where the data class is classified as normal data and abnormal/anomalous data (outliers). Next, the data group labelled training data will be trained using the OCSVM model approach. The next stage is evaluating the training data and testing data. After that, the performance of the OCSVM model is measured using a confusion matrix.

The OC-SVM (One-Class Support Vector Machine) model will be trained using a training dataset that specifically includes data on students who graduated on time. This focused approach helps in identifying the patterns and characteristics associated with on-time graduation. To, enhance the model's performance, cross-validation will be employed. This technique involves partitioning the dataset into multiple subsets, training the model on some subsets, and validating it on the remaining ones. This process helps in ensuring that the model generalizes well to new, unseen data. Key hyperparameters such as the kernel parameter and the regularization term will be optimized during cross-validation. Optimizing these parameters is crucial for improving the accuracy and effectiveness of the OC-SVM model[30].

IV. Results and Discussion

The One-Class Support Vector Machine (OC-SVM) model was evaluated on the testing dataset to assess its performance in predicting student graduation outcomes. The model achieved high accuracy, indicating its effectiveness in identifying students who are likely to graduate on time. At this stage, researchers will present a series of prediction results carried out using a one-class support vector machine approach which are represented in graphical form. The results obtained also use the sigmoid support vector machine kernel, where the kernel aims to speed up the process of training data with larger dimensions also in the sigmoid kernel the use of gamma can be adjusted to increase the accuracy value produced in the prediction process, the results of implementing the OCSVM model approach using a sigmoid kernel. The normal data class in this research dataset is 677 data, which is then trained using a one-class support vector machine (OCSVM) approach and a sigmoid kernel as in Figure 3.

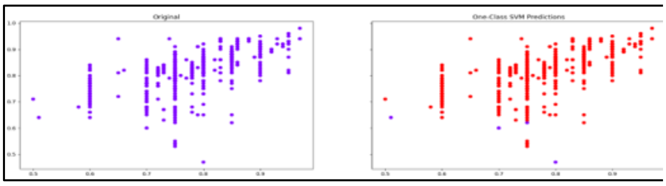

Fig. 3. Prediction Result Using One Class Support Vector Machine

The results show that there is a significant difference between before and after training, where the OCSVM model correctly (accuracy value of 99%) can predict most of the training data from the training data set as data in the "Pass" category. These results also show that the use of the OCSVM model in a single class classifier has a significant impact in determining the prediction evaluation of classifier performance in table 1.

TABLE I. Matrix Evaluation OCSVM

	Precision	*recall*	*F1-Score*	*Support*
Failed (0)	0.00	0.00	0.00	0
Passed (1)	1.00	0.99	1.00	677
Accuracy			0.99	677

That the performance results of the prediction model achieved with the one-class support vector machine approach produce an accuracy value of 99%, this shows that the OCSVM model has very good abilities in predicting the academic performance achievements of students in the "Pass" category in the group. training data, the same thing is also shown in the precision value, namely 100%, meaning that the OCSVM model is able to predict the positive class (pass category) well in the training data set and has very little prediction error in the negative class (fail category). In Figure 4, the prediction process is carried out on groups of data with failed categories or anomalous data (negative class) which are then considered as test data with a total of 49 data.

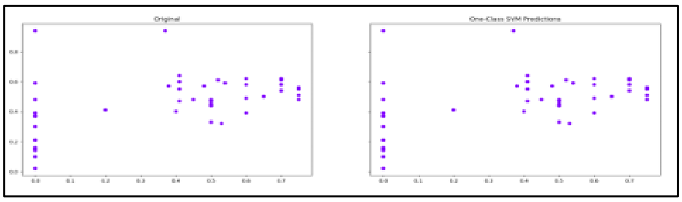

Fig. 4. Prediction Result Using OCSVM

The prediction results using the OCSVM model on the test data correctly predicted all data sets on the test data as data in the "Fail" category. These results show that the use of the OCSVM model in a class classifier has a significant impact in predicting student failure in achieving minimum competency standards in one course. In Table 2 it is shown that the performance results of the prediction model achieved with the one-class support vector machine approach produce an accuracy value of 100%, this shows that the OCSVM model has very good abilities in predicting the failed academic performance of students in the "Failure" category test data set, the same thing is also shown in the precision value, namely 100%, meaning that the OCSVM model can predict the negative class (failure category) well and there are no prediction errors for the positive class (pass category) in the testing data set.

TABLE II. Matrix Evaluation OCSVM

	Precision	*recall*	*F1-Score*	*Support*
Failed (0)	1.00	1.00	1.00	49
Accuracy			1.00	49

In this research, it is proposed to apply OCSVM to detect student academic performance achievements that do not meet minimum competency standards. The proposed approach has tested the performance of the prediction model using a confusion matrix and the results are very significant. With the OCSVM model approach in predicting student failure in one course, it is hoped that it can help universities to improve the quality of education, especially in online learning systems. As a proposal and interpretation, the researcher aims to expand the prediction performance assessment with the OCSVM model which is proposed as a prediction model to help detect student failure in one course in the learning process, especially online learning.

The OC-SVM model demonstrated high accuracy in predicting student graduation outcomes, underscoring its suitability for educational data mining tasks. The high precision and recall rates indicate that the model effectively identifies students at risk of delayed graduation, making it a reliable tool for early intervention. The model's performance metrics, including an accuracy of 99%, precision of 100%, recall of 99%, and an F1-score of 100%, validate its robustness and practical applicability in real-world educational settings.

The application of One-Class Support Vector Machine (OC-SVM) in predicting student graduation performance offers several key strengths that are crucial in educational data mining. Firstly, OC-SVM is effective in handling imbalanced data, a common issue in educational datasets where the majority of students graduate on time[31]. This capability is essential as it ensures that the model can accurately predict outcomes even when there is a significant class imbalance, thereby improving the reliability of the predictions. Secondly, OC-SVM's ability to identify at-risk students early on allows for timely interventions, potentially enhancing student retention and success rates. By detecting students who may be struggling or at risk of not graduating, educational institutions can provide targeted support and interventions to improve outcomes. Lastly, the scalability of the OC-SVM model is a significant advantage, as it can be applied to large datasets from multiple institutions, making it versatile and applicable across different educational contexts [32]. This scalability ensures that the model can handle varying data sizes and complexities, making it a valuable tool for educational institutions seeking to improve student outcomes.

In the context of predicting student graduation performance, the strengths of OC-SVM align well with the requirements of educational data mining. The ability to handle imbalanced data is particularly relevant in educational settings where student outcomes may not be evenly distributed. By effectively managing class imbalances, OC-SVM can provide more accurate predictions and insights into student performance. Early detection of at-risk students is crucial for implementing timely interventions that can positively impact student success rates. By leveraging OC-SVM's early detection capabilities, educational institutions can proactively support students who may be facing challenges, ultimately improving retention rates and graduation outcomes. Additionally, the scalability of OC-SVM ensures that it can be applied across diverse educational contexts, making it a versatile tool for predicting student graduation performance in various settings. The application of OC-SVM in predicting student graduation performance offers significant advantages in handling imbalanced data, enabling early detection of at-risk students, and providing scalability across different educational contexts. These strengths make OC-SVM a valuable tool in educational data mining for improving student outcomes and enhancing decision-making processes.

Despite its strengths, the OC-SVM model has certain limitations that need to be addressed:

- Feature Selection: The effectiveness of the model heavily depends on the quality and relevance of the input features. Future research should focus on identifying and integrating additional features that can enhance predictive accuracy[33].
- Computational Complexity: OC-SVM can be computationally intensive, particularly with large datasets. Optimizing the model to reduce computational requirements without compromising accuracy is an area for future exploration [34].

The successful application of the OC-SVM model in predicting student graduation outcomes has several practical implications:

- Improved Decision-Making: Educational institutions can use the model to make data-driven decisions regarding resource allocation and student support services.
- Targeted Interventions: By identifying students at risk of delayed graduation, institutions can implement targeted interventions such as academic counseling and tutoring, thereby enhancing student success rates.
- Policy Formulation: Insights from the model can inform policy formulation aimed at improving educational outcomes and addressing systemic issues affecting student performance.

Future research should focus on several key areas to enhance the predictive capabilities of the OC-SVM model:

- Integration of Additional Features: Incorporating additional data points, such as socio-economic factors, extracurricular activities, and psychological assessments, could improve the model's accuracy.
- Cross-Institutional Studies: Expanding the research to include data from multiple institutions can help validate the model's generalizability and effectiveness across different educational settings.
- Hybrid Models: Exploring hybrid models that combine OC-SVM with other machine learning algorithms could leverage the strengths of multiple approaches to improve predictive performance[35].

V. Conclusion

The research demonstrates the effectiveness of the One-Class Support Vector Machine (OC-SVM) algorithm in predicting student graduation outcomes. The OC-SVM model achieved high accuracy and reliability in identifying students at risk of delayed graduation, making it a valuable tool for educational institutions. The model's ability to handle imbalanced datasets and its robustness in early detection of at-risk students highlight its practical applicability in educational settings.

By leveraging historical student data, the OC-SVM model provides actionable insights that can inform targeted interventions, such as academic counseling and tutoring, thereby improving overall student success rates. The comparative analysis with other machine learning algorithms further underscores the superiority of OC-SVM in this context.

Future research should focus on integrating additional features to enhance the model's accuracy and exploring hybrid models that combine the strengths of multiple algorithms. Expanding the research to include data from multiple institutions can validate the model's generalizability and effectiveness across diverse educational environments. By continuously refining and expanding predictive models, educational institutions can better support their students and achieve higher success rates.

References

[1] N. Sghir, A. Adadi, and M. Lahmer, *Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022)*, vol. 28, no. 7. Springer US, 2023.

[2] X. Chen, Y. Peng, Y. Gao, and S. Cai, “A competition model for prediction of admission scores of colleges and universities in Chinese college entrance examination,” *PLoS One*, vol. 17, no. 10 October, pp. 1–17, 2022, doi: 10.1371/journal.pone.0274221.

[3] J. C. Immekus, T. sun Jeong, and J. E. Yoo, “Machine learning procedures for predictor variable selection for schoolwork-related anxiety: evidence from PISA 2015 mathematics, reading, and science assessments,” *Large-Scale Assessments Educ.*, vol. 10, no. 1, pp. 1–21, 2022, doi: 10.1186/s40536-022-00150-8.

[4] S. N. Liao, D. Zingaro, K. Thai, C. Alvarado, W. G. Griswold, and L. Porter, “A robust machine learning technique to predict low-performing students,” *ACM Trans. Comput. Educ.*, vol. 19, no. 3, pp. 1–19, 2019, doi: 10.1145/3277569.

[5] W. Yu, T. Liu, R. Valdez, M. Gwinn, and M. J. Khoury, “Application of support vector machine modeling for prediction of common diseases: The case of diabetes and pre-diabetes,” *BMC Med. Inform. Decis. Mak.*, vol. 10, no. 1, pp. 1–7, 2010, doi: 10.1186/1472-6947-10-16.

[6] T. Guo *et al.*, “Graduate employment prediction with bias,” in *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, 2020, pp. 670–677, doi: 10.1609/aaai.v34i01.5408.

[7] H. Sanz, F. Reverter, and C. Valim, “Enhancing SVM for survival data using local invariances and weighting,” *BMC Bioinformatics*, vol. 21, no. 1, pp. 1–20, 2020, doi: 10.1186/s12859-020-3481-2.

[8] M. Freitas *et al.*, “Identification of Abnormal Behavior in Activities of Daily Life Using Novelty Detection,” in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 2023, vol. 492 LNICST, pp. 559–570, doi: 10.1007/978-3-031-34776-4_29.

[9] Y. Li, “Analysis of Strategies for Enhancing Informatization Teaching Ability of College Teachers Combined with Mathematical Statistics Multiple Regression Models,” *Appl. Math. Nonlinear Sci.*, vol. 9, no. 1, 2024, doi: 10.2478/amns.2023.2.01390.

[10] M. GhasemiGol, R. Monsefi, and H. S. Yazdi, “Ellipse support vector data description,” *Commun. Comput. Inf. Sci.*, vol. 43 CCIS, pp. 257–268, 2009, doi: 10.1007/978-3-642-03969-0_24.

[11] A. Bounsiar and M. G. Madden, “One-class support vector machines revisited,” in *ICISA 2014 - 2014 5th International Conference on Information Science and Applications*, 2014, pp. 31–34, doi: 10.1109/ICISA.2014.6847442.

[12] R. Ghiasi, M. A. Khan, D. Sorrentino, C. Diaine, and A. Malekjafarian, “An unsupervised anomaly detection framework for onboard monitoring of railway track geometrical defects using one-class support vector machine,” *Eng. Appl. Artif. Intell.*, vol. 133, no. PB, p. 108167, 2024, doi: 10.1016/j.engappai.2024.108167.

[13] A. Harish, P. Asok, and M. V. Jayan, “A comparative evaluation of Stacked Auto-Encoder neural network and Multi-Layer Extreme Learning Machine for detection and classification of faults in transmission lines using WAMS data,” *Energy AI*, vol. 14, no. December 2022, p. 100301, 2023, doi:

10.1016/j.egyai.2023.100301.
[14] H. Jia *et al.*, “Enabling temporal–spectral decoding in multi-class single-side upper limb classification,” *Eng. Appl. Artif. Intell.*, vol. 133, no. PE, p. 108473, 2024, doi: 10.1016/j.engappai.2024.108473.
[15] S. S. Todkar, V. Baltazart, A. Ihamouten, X. Dérobert, and D. Guilbert, “One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using Ground Penetrating Radar data,” *J. Appl. Geophys.*, vol. 192, no. March, 2021, doi: 10.1016/j.jappgeo.2021.104392.
[16] A. Pollastro, G. Testa, A. Bilotta, and R. Prevete, “Semi-Supervised Detection of Structural Damage Using Variational Autoencoder and a One-Class Support Vector Machine,” *IEEE Access*, vol. 11, no. June, pp. 67098–67112, 2023, doi: 10.1109/ACCESS.2023.3291674.
[17] A. A. Abdulhussein, M. faisun Nasrudin, S. M. Darwish, and Z. A. A. Alyasseri, “A Genetic Algorithm Based One Class Support Vector Machine Model for Arabic Skilled Forgery Signature Verification,” *J. Imaging*, vol. 9, no. 79, pp. 1–26, 2022, doi: 10.2139/ssrn.4303232.
[18] M. Alkhodari *et al.*, “Screening Cardiovascular Autonomic Neuropathy in Diabetic Patients with Microvascular Complications Using Machine Learning: A 24-Hour Heart Rate Variability Study,” *IEEE Access*, vol. 9, pp. 119171–119187, 2021, doi: 10.1109/ACCESS.2021.3107687.
[19] S. Revathy and S. S. Priya, “Enhancing the Efficiency of Attack Detection System Using Feature selection and Feature Discretization Methods,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. February, pp. 156–160, 2023, doi: 10.17762/ijritcc.v11i4s.6322.
[20] K. Nagata *et al.*, “Detection of overdose and underdose prescriptions—An unsupervised machine learning approach,” *PLoS One*, vol. 16, no. 11 November, 2021, doi: 10.1371/journal.pone.0260315.
[21] S. Sayadi, S. B. Rejeb, and Z. Choukair, “Anomaly detection model over blockchain electronic transactions,” in *2019 15th International Wireless Communications and Mobile Computing Conference, IWCMC 2019*, 2019, pp. 895–900, doi: 10.1109/IWCMC.2019.8766765.
[22] B. Chen, F. Wei, and C. Gu, “Bitcoin Theft Detection Based on Supervised Machine Learning Algorithms,” *Secur. Commun. Networks*, vol. 2021, no. August 2016, 2021, doi: 10.1155/2021/6643763.
[23] Z. Zhang, G. Wang, C. Liu, L. Cheng, and D. Sha, “Bagging-based positive-unlabeled learning algorithm with Bayesian hyperparameter optimization for three-dimensional mineral potential mapping,” *Comput. Geosci.*, vol. 154, 2021, doi: 10.1016/j.cageo.2021.104817.
[24] A. Z. Woldaregay, I. K. Launonen, D. Albers, J. Igual, E. Arsand, and G. Hartvigsen, “A novel approach for continuous health status monitoring and automatic detection of infection incidences in people with type 1 diabetes using machine learning algorithms (Part 2): A personalized digital infectious disease detection mechanism,” *J. Med. Internet Res.*, vol. 22, no. 8, 2020, doi: 10.2196/18912.
[25] Y. Li, “Analysis of Strategies for Enhancing Informatization Teaching Ability of College,” *Appl. Math. Nonlinear Sci.*, vol. 9, no. 1, pp. 1–19, 2024, doi: 10.2478/amns.2023.2.01390.
[26] I.-M. Chiu, C.-Y. Cheng, P.-K. Chang, C.-J. Li, F.-J. Cheng, and C.-H. R. Lin, “Utilization of Personalized Machine-Learning to Screen for Dysglycemia from Ambulatory ECG, toward Noninvasive Blood Glucose Monitoring,” *Biosensors*, vol. 13, no. 1, 2023, doi: 10.3390/bios13010023.
[27] Z. Güney, “Four-Component Instructional Design (4C/ID) Model Approach for Teaching Programming Skills,” *Int. J. Progress. Educ.*, vol. 15, no. 4, pp. 142–156, 2019, doi: 10.29329/ijpe.2019.203.11.
[28] M. Firdaus, E. Suryani, R. Nadlifatin, and A. Tjahyanto, “Enhancing Organizational Culture and Productivity through Innovative Performance Appraisal: A Design Science Research Approach,” *Procedia Comput. Sci.*, vol. 234, pp. 1128–1136, 2024, doi: 10.1016/j.procs.2024.03.108.
[29] D. A. Crake, N. C. Hambly, and R. G. Mann, “HEADSS: HiErArchical Data Splitting and Stitching software for non-distributed clustering algorithms,” *Astron. Comput.*, vol. 43, p. 100709, 2023, doi: 10.1016/j.ascom.2023.100709.
[30] P. L. dos Santos, T. P. Azevedo-Perdicoúlis, and P. A. Salgado, “Non-parametric Gaussian process kernel DMD and LS-SVM predictors revisited - A unifying approach,” *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 10533–10539, 2023, doi: 10.1016/j.ifacol.2023.10.1075.
[31] S. A. Priyambada, T. Usagawa, and M. ER, “Two-layer ensemble prediction of students’ performance using learning behavior and domain knowledge,” *Comput. Educ. Artif. Intell.*, vol. 5, no. June, p. 100149, 2023, doi: 10.1016/j.caeai.2023.100149.
[32] R. Isus, K. Kolesnikova, I. Khlevna, T. Oleksandr, and K. Liubov, “Development of a model of personal data protection in the context of digitalization of the educational sphere using information technology tools,” *Procedia Comput. Sci.*, vol. 231, no. 2023, pp. 347–352, 2024, doi: 10.1016/j.procs.2023.12.215.
[33] Y. Bouchlaghem, Y. Akhiat, K. Touchanti, and S. Amjad, “A novel feature selection method with transition similarity measure using reinforcement learning,” *Decis. Anal. J.*, vol. 11, no. May, p. 100477, 2024, doi: 10.1016/j.dajour.2024.100477.

[34] L. Ferreira and P. Cortez, “AutoOC: A Python module for automated multi-objective One-Class Classification[Formula presented],” *Softw. Impacts*, vol. 18, no. August, p. 100590, 2023, doi: 10.1016/j.simpa.2023.100590.

[35] A. Barbado, Ó. Corcho, and R. Benjamins, “Rule extraction in unsupervised anomaly detection for model explainability: Application to OneClass SVM[Formula presented],” *Expert Syst. Appl.*, vol. 189, no. October 2021, p. 116100, 2022, doi: 10.1016/j.eswa.2021.116100.