Improving the Prediction Accuracy of Student Performance in a Cross-Semester Scenario Utilizing a Domain Adaptation Approach

Matthew Bobea[1], Soyeong Park[1], Brendan Flanagan[2] and Owen H.T Lu*

[1] *Department of Management Information Systems, National Chengchi University, Taiwan*
[2] *Academic Center for Computing and Media Studies, Kyoto University, Japan*
[*] *International College of Innovation, National Chengchi University, Taiwan*

Abstract

One of the frequent tasks in learning analytics is predicting students' learning performance by using machine learning algorithms. A typical approach in student performance prediction is to train a machine learning model using a well-collected dataset. However, this approach requires dataset distribution consistency in past and future time periods. Otherwise, the risk prediction model will be unavailable once the distribution shifts. In practice, we can train a student performance prediction model using data collected from a past semester and use the model to predict students' risk in the following semester; yet the model's performance is restricted by how data distribution is consistent between each semester. This limitation makes the task of predicting learning behavior unfeasible. In this paper, we demonstrated that data distribution would shift between semesters even when this dataset showed that each sample has similar learning behaviors, activities, and subjects. Moreover, domain adaptation in transfer learning is introduced to solve the mentioned issues in cross-semester scenario. We applied methods of combining PCA and supervised domain adaptation (DAPCA) to preprocess our dataset for model classification. The datasets were analyzed to distinguish at-risk students and suggest behavior changes for preventing poor performance. Our research discovered 13 education features that are significant for predicting students' success. Additionally, our results show an accuracy improvement from 6.67% to 33.96% when preprocessing our data using DAPCA methodologies when compared to a traditional PCA approach.

Key words

Learning analytics, domain adaptation, risk student prediction

1. Introduction

Today, the education environment has changed significantly; data derived from a changing educational environment is emerging and has become a source that can help us to further understand the changes and conditions that are happening within our learning environment. Particularly advances in technology have improved various learning settings, such as web-based learning or remote learning. Covid-19's impact on education has further accelerated these changes in learning settings as interests in the online learning environment have increased. New data within the education environment has emerged, allowing for deeper understanding of Learning Analytics (LA) and student teacher interactions.

LA is the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance, and spot potential issues (Johnson et al., 2011). In this study we analyze the LA data found in the Learning Behavior and Learning Strategies data set (LBLS-160). This is a collection of learning behavior data regarding 160 students across three computer science semesters.

It is necessary to adapt newly developed methods to further enhance LA research because of the cross-semester scenario in risk student prediction, as shown in Figure 1. In practice, to distinguish between at-risk and non-risk students, the typical approach is to train a classification model using a dataset collected from a learning environment (JoungYoungRan, 2020). This dataset usually consists of students' learning behavior and their learning outcomes as the features and labels, respectively, for conducting supervised learning. In this supervised learning, algorithms such as Support-Vector Machines (SVM) search for a hyperplane on the data projected vector space. This hyperplane is highly reliant on the data distribution in the vector space, and the model would no longer be available if the distribution shifted. Compared to the real scenario in the classroom, we believe the data distribution will not be consistent, as it is not possible for a teacher to conduct exactly the same instructions in different semesters. We have named this issue the cross-semester issue to address it. Prior studies have proposed a domain adaptation approach, which is an option in transfer learning. Therefore, based on the concept of transfer learning, we can name the well-collected data from the previous semester as source data, its distribution as the source domain, and the model as the source model. On the contrary, the dataset without labels from the new semester is the target data, and the target domain.

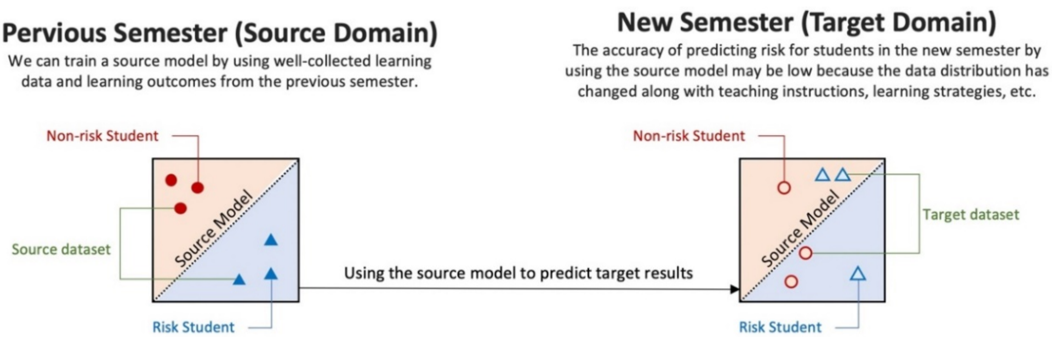


Figure 1: Cross-semester scenario in prediction risk student

This study has two aims. First, we wish identify the features that foresee students that may be 'at-risk' of poor performance while identifying features which correlate with successful students. Second, we wish to adapt newly developed methods of to further enhance LA research. This paper is separated in five sections. First, in this Introduction we introduce the problems commonly found in LA data today. Second, in the literature review section we discuss adapted methods from other research which we will use in the preprocessing and modeling of our data. The application process of these methods as other data processing methods is further discussed in the Methodologies section. Lastly, in Results and Discussion we show the outcomes of our applied methods. Finally, we summarize our finding and discuss future research opportunities in the Conclusion section.

2. Literature Review

2.1 Domain Adaptation (DA)

Recently, data mining technology has rapidly developed, expanding its uses diversely. However, training data (source) for developing a model to predict a model's unlabeled data (target) may be insufficient. Transfer learning was proposed in this case. The concept is to compensate for the lack of source data is to use trained knowledge from an alternative domain which may be richer in source data. Traditional machine learning techniques try to learn each task from scratch while transfer learning techniques try to transfer the knowledge from previous task to a target task when the latter has fewer high-quality source data (Pan and Yang, 2010). This allows saves time collecting source data. Compared to the traditional approach, the new model can be trained with a higher slope of the learning curve and higher asymptote of model accuracy (Torrey and Shavlik, 2010).

Domain adaptation (DA) (Ganin et al., 2016) is required when providing source knowledge from one domain to another. Specifically, domain adaptation is design for addressing issues with bias. This is referred to as named as domain shift (Stacke et al., 2021). Supervised machine learning methods only perform well when the given extensive labeled data are from the same distribution as the target distribution(Sun et al., n.d.). Domain shift is a serious problem for generalization of machine learning models and it is well-established that a domain shift between the source and target sets may cause a drastic drop in the model’s performance (Ugurlu et al., 2022). For these reasons we need to deliver source knowledge by means of domain adaptation.

The main objective of domain adaptation is to apply classifiers trained in the source domain to the target domain for pursuing better performance (Zhang et al., 2017). Furthermore, recently applied methods of combining PCA and supervised domain adaption (DAPCA) have been proposed by Gorban et al. as a solution to domain shift. In this paper, they introduced DAPCA along with methodologies such as semi-supervised PCA and TCA. Additionally, they advert DAPCA as a way to resolve the classical distance concentration problem (Gorban et al., 2021). Evgeny et al. use DAPCA to mitigate the data shift problem. Moreover, their study found that DAPCA can be used as a helpful tool for simultaneously integrating datasets from different origins and reducing their dimensionality (Mirkes et al., 2023). DAPCA is a topic that has been studied for a relatively short period compared to DA or PCA. As a result, DAPCA was the focus of a limited number of researchers. Nevertheless, DAPCA is an effective countermeasure to the domain shift problem found LBLS-160. This is because as previous research suggests, the DAPCA algorithm is generally valuable for decreasing domain shift problems in datasets such in LBLS-160.

2.2 Modeling with DA in Learning Analytics

There has been research results on applying Big Data and AI technologies on the evaluation of individual learning for education (Lho, 2021). Those methods attract attention along with education data mining in the education era; a field which is referred to as LA. LA is a relatively new field developed to assist teaching and learning practices (Kew and Tasir, n.d.). It is the measurement, collection, analysis, and reporting of data about learners and their contexts for understanding and optimizing learning and the environments in which it occurs (“Penetrating the Fog: Analytics in Learning and Education | EDUCAUSE,” n.d.).

Recently, much attention has been paid to LA to understand different learning types, predict learners’ performances, and to further develop various teaching strategies under those learning settings (Ahn et al., 2015). Since 2010, the Horizon Report (Johnson et al., 2011) has consistently emphasized LA as one of the core technologies for digital data analysis in the educational environment. LA research in past studies have suggested many valuable motives for studying in LA. Especially analyzing student performance, the subject of our study. However, this paper differs from these previous studies in that they rarely consider the possibilities of a domain shift problems when they analyze data across different study groups. Nevertheless, the domain shift problem should always be considered as it is always present whenever an experiment is conducted which analyze different groups of data through cross validation.

Luan et al. conducted an examination of 40 pieces of research that employed machine learning to enhance education instruction (Luan and Tsai, 2021). In this research the papers were evaluated for computational strategies, assessment criteria, and confirmation techniques utilized in precision teaching with machine learning. Here it was discovered that SVM was the predominate machine learning algorithms selected. Additionally, cross validation was the prevailing form of validation and accuracy was the most common form of model evaluation. For these reasons, we will follow the same approach in selecting our model and parameter tuning.

Others have argued that artificial neural networks may be used to handle the domain shift problem. Shallow models have several benefits over artificial neural networks, such as shorter training periods, simpler debugging, and a smaller number of parameters to adjust. Furthermore, it facilitates more precise forecasts and improved generalization of the model. In addition, it may lessen the need for a large training dataset, as well as minimize the complexity of the model. Since the LBLS-160 data set is limited in sample size, SVMs are be advantageous in this circumstance.

While SVM may be ideal for LA analysis, is may be affected by domain shift problems. Tang et al. proposed utilizing incremental SVM classification with DA (Tang et al., 2022). Consequently, our study paid attention to the influence of the domain shift problem while conducting SVM analysis in an attempt to minimalize the domain shift problem. Cases of conducting LA research related to cross-semester student analysis utilizing DAPCA are sporadic in LA academics. Despite DAPCA's potential for use in LA academia, as its methodology has been overlooked in terms of its potential. In previous studies, machine learning through the combination of DAPCA was not actively studied. When limited to learning analytics, even fewer studies were conducted. Therefore, by using the DAPCA methodology, our study will shed light on the domain shift problem in the aspect of LA.

3. Methodology

3.1 Data Set Characteristic

We used the LBLS-160 dataset (Flanagan and Ogata, 2017; Lu et al., 2022; Ogata et al., 2017) for analyzing student behavior. The data in LBLS-160 was collected from students across three programming language semesters. Each semester contains 63, 56, and 41 students respectively, totaling 160 samples overall. The data set also contained samples from a fourth semester however this data was excluded as the dataset was incomplete. While each course was held during different times during the school year, all learning material (e.g., program, homework, materials, semester duration, and contents) remained the same across the three semesters. The students participating in this experiment is restricted to students who are not studying computer-related majors.

LBLS-160 is separated into two parts. In the first part each student's learning behavior was collected by means of two online learning environments, BookRoll (Ogata et al., 2015) and VisCode (Lu et al., 2016). These two learning environments are initially designed for teaching as they provide functionality which tracks the learning record of each individual student. Therefore, it allows for the students learning behavior data to be collected and logged. BookRoll also enables teachers to manage their teaching materials; learners can read, add memos, and highlight content while recording such contents as a log file. VisCode is a programing environment developed for the lecturer. They can upload sample code, and the students can code, execute and test their code data in this environment. It records the code data of each individual students.

The second part of the LBLS-160 data set measures each student's learning strategies via three questionnaires: SRL Motivation, SRL Strategy, and SILL. Data from the SRL questionnaires were not used the features did not show a high level of significance when compared to other features. SILL data was also not used since the data set is incomplete as it excluded data from semester C (see Table 1 (Lu et al., 2022)).

Table 1

Students and features in LBLS-160

		Learning behaviors		Learning strategies		
	Participants	**BookRoll**	**VisCode**	**SRL Motivation**	**SRL Strategy**	**SILL**
Semester A	63	√	√	√	√	√
Semester B	56	√	√	√	√	√
Semester C	41	√	√	√	√	

3.2 PCA

Principal component analysis (PCA) is an analytical technique that is applied to reduce the number of variables in a dataset. Many techniques have been developed for this purpose, but PCA is one of the oldest and most widely used (Jolliffe and Cadima, n.d.). The central idea of PCA is to find a low dimensional subspace which represents most of the variation within the sample data (Xia et al., 2007). PCA is also used to reduce the shift between domains. They can be used for solving domain adaptation problems without considering labels in the source domain. In particular, various generalizations of PCA's computing a joint linear representation of two and more datasets are widely exploited in machine learning (Mirkes et al., 2023). For this reason, PCA works in tangent with domain adaptation in analyzing the LBLS-160 data set.

The first step in transforming our data with PCA is to normalize the dataset. In the normalization process, we must ascertain the mean and standard deviation for each individual attribute. This is calculated by subtracting the mean from each value and dividing it by the standard deviation value for each value. This is expressed as $z = (x - \mu) / \sigma$. This step repeated for each variable in the matrix and is essential as it is used to calculate a symmetrical p × p covariance matrix (where p is the quantity of dimensions). This matrix is then used to calculate the eigenvectors and eigenvalues which will then be used to establish the main components. This is done by solving $(A-\lambda I)\nu = 0$ where A is our covariance matrix, ν a vector and λ a scalar (Nobles, 2020). Next, the eigenvalues are arranged and their associated eigenvectors in order. The next step is to select k eigenvalues are used to establish a matrix comprising the eigenvectors. The last step is to transform the original matrix by taking the feature matrix and multiplying by the eigenvector matrix. This results in a transformed LBLS-160 dataset.

3.3 Domain Adaptation

Domain adaptation techniques transfer knowledge from the source domain to the target domain, in the form of learned models and efficient feature representations, to learn effective classifiers on the target domain (Xie et al., 2018). This methodology trains a machine learning algorithm to make stable inferences in a source domain based on learned knowledge of a targeting domain. In many educational data in the real world, the data distribution is the same, and the violation rate is the same. Knowledge transfer or transfer learning between task domains would be advantageous in such circumstances. The source domain has labels can be used to build a classifier when the targeted domain does not have labels (Mirkes et al., 2023) For these reasons we will adapt domain adaptation to our model.

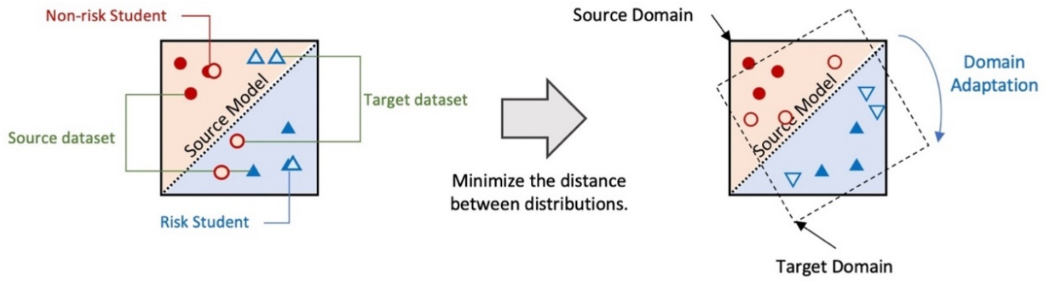


Figure 2: Domain adaptation learning

3.4 Domain Adaptation Principal Component Analysis (DAPCA)

One of domain adaptation's most significant computational problems is reducing the difference between source and target domain data distributions. Intuitively, discovering a good feature representation across domains is crucial (Pan et al., 2011). A good feature representation can reduce the difference in distributions between domains while preserving essential properties of the original one. PCA can help domain adaptation with this process. However, when we combine the effects of both PCA and DA, we can deal with domain shift problems more efficiently. Recently, a novel base linear method called DAPCA has been studied by scholars. In DAPCA we generalize the supervised PCA algorithm to the DA challenge. Firstly, the approach was outlined in the context of one- and few-shot learning problems. It relies on the definition of weights between some data points, both in the source and the target domains and between them those projections of data vectors onto the eigenvectors of a simple quadratic form would serve as a good feature concerning domain adaptation (Mirkes et al., 2023).

3.5 Feature Selection

The top 13 most significant features when determining a student's performance were selected for use in our model by computing the chi-squared statistic between each feature and our targets. We chose this methodology for variable selection because by measuring dependence between stochastic variables we can dispose of the features that are the most likely to be independent of our semester and therefore are less appropriate for classification. The selected features are as described in Table 2. All selected variables are numeric integers and do not have missing values. Each feature is labeled with the respective student's 'semester', a categorical variable consisting of A, B, or C.

Table 2

Independent Variables with Descriptions

Resource	Feature	Description
BookRoll	ADD MARKER	Added a marker to current page.
BookRoll	CHANGE MEMO	Modify the content of an existing memo on current page.
BookRoll	CLOSE	Closed the book.
BookRoll	DELETE MARKER	Deleted a marker on current page.
BookRoll	NEXT	Went to the next page
BookRoll	OPEN	Opened the book.
BookRoll	PREV	Went to the previous page.
VisCode	code_copy	Number of times a student copy code.
VisCode	code_execution	Number of times a student executes codes.
VisCode	code_paste	Number of times a student paste code.
VisCode	code_speed	Average input digits per minutes.
VisCode	notebook_open	Number of times a student open coding environment.
VisCode	code_length	Number of lines of code (LOC) coded in this semester.

4. Results and Discussion

4.1 Setting a Threshold for our Target Variable

When predicting which students are at-risk of underperforming, it is vital to find the optimal threshold with which to label students whom are outstanding or at-risk. While each student is a case-by-case scenario and may not be unequivocal in terms of being an at-risk student or not, a threshold needs to be drawn in order to support hyperparameter tuning within the SVM model. To accomplish this, we test rerunning the model at every threshold value between 70 and 90. The model was run on the full data set using a SVM model excluding DAPCA processing. We chose the 70-90 range as the score data follows a skew distribution with the majority of students scoring within the 80 point range. While modeling at each threshold, we also consider balancing the ratio of the students who are at-risk over the total number of students for each semester. The final results are visualized Figure 3 below.

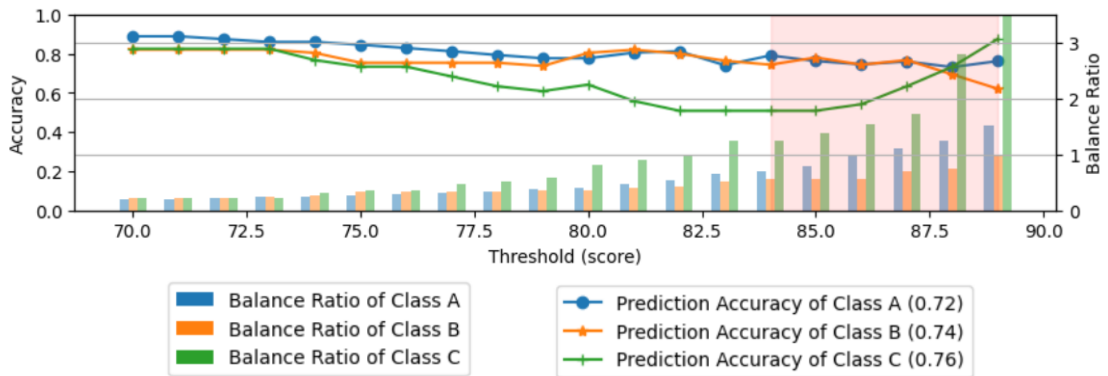


Figure 3: Semester A, B, and C distribution

Here, the figure shows that we have an ideal balance between the balance ration score and accuracy score in the vicinity of the 87 to 88 point range. A threshold of 87 was used because at a threshold of 88, the ratio of semester C to the other semesters were imbalanced. Within our data set, students who scored less than 88 we encoded '0' while students scoring 88 and above were encoded '1', where '0' represents a student whom is 'at-risk' and '1' represents an 'outstanding student'. This data is used as our target variable for our SVM model.

4.2 Data Mining and Analysis

The 13 variables selected in Table 2 are ran using a standard SVM algorithm with RBF kernel to predict at-risk students. Here, we apply cross validation to: 1) test the robustness of our model 2) evaluate how our model's accuracy will perform against a generalized data set and 3) asses the effects of using a DAPCA methodology to for better results. When conducting cross validation, we test the accuracy of our model using every combination of source and target data. In situations where the source and target are from different semesters, we can observe how our model is able to predict the performance of a student using a source data that may have been affected by different environmental factors. For example, the students from 'Semester A' will be used to train a model and make predictions regarding the performance of students from 'Semester B' (see Table 5: Source – Semester A, Target – Semester B). In situations where the source is from a semester which takes place after the target (for example: Source – Semester B, Target – Semester A), the analysis was still conducted so as to further assess the effects of DAPCA on the dataset as well as to continue to test if the model is generalized through cross validation.

To perform this experiment, the LBS-160 data set is divided into three separate data sets in order to isolate the data for semester A, B, and C respectively. Before applying any DAPCA methodologies to the data, we run the experiment using only PCA analysis as a control. Here, PCA is applied to the three data sets which are then each split between source and target data. Afterwards, we ran each source and target combination through a SVM model producing the results in Table 3 below.

Table 3
Non-Adopted DAPCA results

Target / Source	Semester A	Semester B	Semester C
Semester A	90%	75%	61%
Semester B	71%	84%	53%
Semester C	60%	68%	69%

Now that a baseline is established, we used the DAPCA approach to process the original semester A, B, and C datasets. Post processing, the three data sets are once again each split into source and target data. The last step is to rerun each DAPCA processed source and target data through the SVM algorithm. Final results are shown in Table 4; additionally, each percent change in the SVM's prediction accuracy shown in parentheses. This is excluded for the three scenarios where the model's source and target data are from the same semester. This is because DAPCA may not be necessary in same semester analysis. Nevertheless, they are still displayed in Table 4 for comparison purposes.

Table 4
Adopted DAPCA results

Target / Source	Semester A	Semester B	Semester C
Semester A	81%	80% (+6.67%)	66% (+8.20%)
Semester B	76% (+7.04%)	80%	71% (+33.96%)
Semester C	71% (+18.33%)	86% (+26.47%)	84%

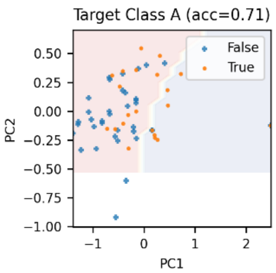
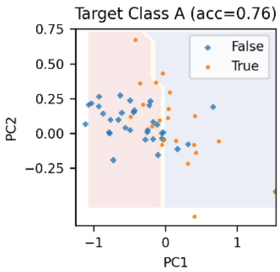
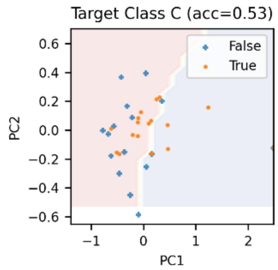
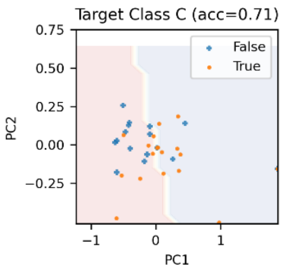
From Tables 3 and 4, it can be seen that in same semester scenario's, the prediction accuracy for the source domain decreases after adopting DAPCA. For example, semester A can obtain 90% accuracy when DAPCA is not adopted, but it drops to 81% post adaptation. The main reason is that after adopting

DAPCA, the domains of both source and target are adapted simultaneously. As mentioned above in Figure 2, the goal of domain adaptation is to minimize the data distribution between the source and target domain and retrain the SVM model afterwards. Therefore, after adapting the source distribution, the source prediction accuracy from the retrained SVM model has dropped subsequently. However, in the cross-semester scenario, the source prediction accuracy can be ignored, as the source data set will come from a course that has already ended, so there is no need for risk prediction.

The effects of DAPCA on the data is visualized in Table 5. Non-Adopted Approach displays the data when only PCA processing was used and Adopted Approach displays the results after DAPCA processing. In these four examples, semester B source data is used to predict targets A and C. These semesters were chosen because in semester A we can see a scenario where using DAPCA had little effect. Alternatively, we can compare this to semester C where DAPCA's effect was substantial.

Table 5

Target semester A Non-adopted / Adopted DAPCA on source semester B, with SVM using RBF kernel

Source: Semester B / Target: Semester A		Source: Semester B / Target: Semester C	
Non-Adopted	**Adopted**	**Non-Adopted**	**Adopted**
			

5. Conclusion

We analyzed the LA data of 160 computer science students across three semesters. By doing so, we discovered 13 features which were relevant to forecasting at-risk students. Nevertheless, when making cross semester predictions we found that necessary statistical assumptions needed for using SVM models were violated. To counteract this issue, a recently developed DAPCA approach was used to preprocess our data. Non-adopted results in Table 3 averaged 64.66% while adopted results in Table 4 averaged 75.00%. Consequently, DAPCA was shown to increase cross semester prediction accuracy results by an overall 16% as a percent change when compared to traditional PCA methods. This statistic excludes cross semester predication because in these scenarios a DAPCA methodology may not be needed.

Recently, newer domain adaptation methods have been developed which may be more suitable for processing the LBLS-160 data set. For future research, we hypothesize that using an alternative DAPCA algorithm may more effectively separate our target data and therefore will improve our model's performance and produce better results.

References

Ahn, M. L., Choi, Y. Y., Ko, Y. M., & Bae, Y. H. (2015). An international literature review on learning analytics: focused on empirical studies. *Journal of Educational Information and Media*, 21(4), 601-643.

Flanagan, B., & Ogata, H. (2017, November). Integration of learning analytics research and production systems while protecting privacy. In *The 25th International Conference on Computers in Education, Christchurch*, New Zealand (pp. 333-338).

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1), 2096-2030.

Gorban, A.N., Grechuk, B., Mirkes, E.M., Stasenko, S.V., Tyukin, I.Y., 2021. High-Dimensional Separability for One- and Few-Shot Learning. *Entropy* 23, 1090.

Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, K., 2011. The 2011 Horizon Report, New Media Consortium. New Media Consortium.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical transactions of the royal society A: Mathematical, *Physical and Engineering Sciences*, 374(2065), 20150202.

JoungYoungRan, 2020. A prediction analysis on the dropout of cyber university based on learning analytics. *The Korean Journal of Educational Methodology Studies* 32, 205–232.

Kew, S. N., & Tasir, Z. (2022). Learning analytics in online learning environment: A systematic review on the focuses and the types of student-related analytics data. *Technology, Knowledge and Learning*, 1-23.

Lho, Y. U. (2021). A Survey on Deep Learning-based Analysis for Education Data. In *Proceedings of the Korean Institute of Information and Commutation Sciences Conference* (pp. 240-243). The Korea Institute of Information and Commutation Engineering.

Lu, O.H., Huang, A.Y., Flanagan, B., Ogata, H., Yang, S.J., 2022. A Quality Data Set for Data Challenge: Featuring 160 Students' Learning Behaviors and Learning Strategies in a Programming Course. *Asia-Pacific Society for Computers in Education* 30, 10.

Lu, O. H., Huang, A. Y., Huang, J. C., Huang, C. S., & Yang, S. J. (2016, September). Early-Stage Engagement: Applying Big Data Analytics on Collaborative Learning Environment for Measuring Learners' Engagement Rate. In *2016 International Conference on Educational Innovation through Technology (EITT)* (pp. 106-110). IEEE.

Luan, H., Tsai, C.-C., 2021. A Review of Using Machine Learning Approaches for Precision Education. *Educational Technology & Society* 24, 250–266.

Mirkes, E.M., Bac, J., Fouché, A., Stasenko, S.V., Zinovyev, A., Gorban, A.N., 2023. Domain Adaptation Principal Component Analysis: Base Linear Method for Learning with Out-of-Distribution Data. *Entropy* 25, 33.

Nobles, T., 2020. Understanding Principle Component Analysis(PCA) step by step. Analytics Vidhya. (accessed 1.7.23).

Ogata, H., Oi, M., Mohri, K., Okubo, F., Shimada, A., Yamada, M., ... & Hirokawa, S. (2017). Learning analytics for e-book-based educational big data in higher education. *Smart sensors at the IoT frontier*, 327-350.

Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K., & Yamada, M. (2015, November). E-Book-based learning analytics in university education. In *International conference on computer in education (ICCE 2015)* (pp. 401-406).

Pan, S. J., Tsang, I. W., Kwok, J. T., & Yang, Q. (2010). Domain adaptation via transfer component analysis. *IEEE transactions on neural networks,* 22(2), 199-210.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5), 30.

Stacke, K., Eilertsen, G., Unger, J., Lundström, C., 2021. Measuring Domain Shift for Deep Learning in Histopathology. *IEEE Journal of Biomedical and Health Informatics* 25, 325–336.

Sun, B., Feng, J., & Saenko, K. (2016, March). Return of frustratingly easy domain adaptation. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 30, No. 1).

Tang, J., Lin, K. Y., & Li, L. (2022). Using Domain Adaptation for Incremental SVM Classification of Drift Data. *Mathematics*, 10(19), 3579.

Torrey, L., Shavlik, J., 2010. Transfer learning, in: Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques. IGI global, pp. 242–264.

Ugurlu, D., Puyol-Antón, E., Ruijsink, B., Young, A., Machado, I., Hammernik, K., King, A.P., Schnabel, J.A., 2022. The Impact of Domain Shift on Left and Right Ventricle Segmentation in Short Axis Cardiac MR Images, in: Puyol Antón, E., Pop, M., Martín-Isla, C., Sermesant, M., Suinesiaputra, A., Camara, O., Lekadir, K., Young, A. (Eds.), Statistical Atlases and Computational Models of the Heart. Multi-Disease, Multi-View, and Multi-Center Right

Ventricular Segmentation in Cardiac MRI Challenge, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 57–65.

Xia, C. L., Shin, H. G., Park, M. C., & Ha, S. W. (2007). A Fast Method for Face Detection Based on PCA and SVM. *Journal of the Korea Institute of Information and Communication Engineering,* 11(6), 1129-1135.

Xie, X., Sun, S., Chen, H., & Qian, J. (2018). Domain adaptation with twin support vector machines. *Neural Processing Letters,* 48, 1213-1226.

Zhang, X., Wang, X., Du, Y., & Qin, X. (2017). Domain Adaptation Image Classification Based on Multi-sparse Representation. *KSII Transactions on Internet and Information Systems (TIIS)*, 11(5), 2590-2606.