International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

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

Deep Learning-Infused Cascading Regression Approach to Predict the Academic Performance of Immigrant Students

Selvaprabu Jeganathan[1], Arun Raj Lakshminarayanan*[2]

Submitted: 04/02/2024 **Revised**: 12/03/2024 **Accepted**: 18/03/2024

Abstract: The academic performance of immigrant students is governed by a diverse range of resources and contexts, including the families of the students, the immigrant communities from which the students originate, and the social and educational attitudes that are held toward immigrants in the countries in which the students are currently residing. The Program for International Student Assessment, is an educational research initiative that is used to assess the knowledge and skills of students who are 15 years old. In this paper, the performance of immigrant students is predicted using the PISA dataset. There are a total of 35 attributes present in the dataset. Among these, the proposed method chooses three attributes(maths, science and reading) as target variables for performance prediction. This research constitutes a novel cascading regression framework designed to accurately forecast academic performance. Sequentially integrating CatBoost Regressor, Bidirectional Recurrent Neural Network (Bi-RNN), and Random Forest Meta Regressor optimizes predictive accuracy. Initiated by the CatBoost Regressor, its outputs serve as inputs for a Bi-RNN model, exploiting bidirectional sequential information. The ensuing predictions from Bi-RNN inform a Random Forest Meta-Regressor, refining the final outcome. Evaluation metrics, comprising MAPE, RMSE, and R2, substantiate the superior accuracy of the cascading model. The cascading ensemble significantly outperformed all individual models, achieving a MAPE reduction of 3.74%, an RMSE reduction of 20.70%, and an R-squared increase of 0.96.This research highlights the efficacy of cascading ensemble techniques for predicting student performance with enhanced accuracy. The method being proposed demonstrates the capacity to capture both fixed and changing characteristics, which may result in enhanced interventions and educational decision-making.

Keywords: *Academic performance, Bi-directional RNN, CatBoost, Deep Learning, Immigrant students, PISA dataset, Random Forest.*

1. Introduction

Immigration has spread globally, and there are now more immigrant pupils in educational systems all over the world. Because of their personal experiences with prejudice and low social standing in the host nation, immigrant parents desire social mobility for their children through education. Immigrant students may be more motivated in class if they subscribe to the notion of immigrant optimism, which refers to their lofty aspirations for success and their desire to further their education for the sake of their families [1, 2]. The majority of nations throughout the world have serious concerns about how to integrate immigrant children into their educational systems. Immigrant students frequently have inferior academic success in higher education and fewer job possibilities as a result [3,4]. So, the social and academic integration of immigrant students is a very important issue, especially for schools that take in a lot of foreign students. The current study, which aims to better understand immigrant student integration, examines the variables that affect how well immigrant students perform academically in the subjects of reading, science and math by assessing student attributes and interactions at school and home, school infrastructure, equity-focused regulations at the school, and the country's per capita gross domestic GDP [5,6,7].

In the analysis, there were two categories of immigrant students: those who were born abroad and those whose parents were born abroad. According to research, students who have recently moved to host nations experience greater difficulty transitioning to the dominant cultures and foreign educational systems as well as greater challenges while trying to acquire a new language. In comparison to native-born pupils, immigrant students generally do worse academically [8,9,10,11]. To further clarify, one of the main reasons for generally poor educational achievement is the number of immigrant pupils in a school. In particular, it has been discovered that the success of the relevant immigrant and indigenous student groupings is negatively impacted by the immigrant concentration at a school [12,13].

Based on Program for International Student Assessment findings, there is a considerable and ongoing achievement difference between immigrants and their native-born peers. On the other hand, the PISA study is in charge of gathering comprehensive information pertaining to the children's mathematics, reading and scientific literacy at the age of 15. Along with the general school and home setting, it also involves the student's motivation, learning styles, and

[1] *Department of Computer Science and Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Tamil nadu, India*
ORCID ID : 0000-0003-0004-3214
[2] *Department of Computer Science and Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Tamil nadu, India*
ORCID ID : 0000-0001-8181-5022
** Corresponding Author Email: arunraj@crescent.education*

outlooks or worldviews. PISA takes one step further by assessing the fundamental information and learning abilities that students have learned in their previous educational system and the degree to which they can apply them in the current world. [14,15]. PISA takes it a step further by assessing how well students can apply the fundamental information and learning abilities they have obtained in their previous educational system. The targeted kids should take a cognitive exam that evaluates their reading, arithmetic, and science abilities.

The paper's most significant contribution is as follows:

- In this paper, the PISA dataset is taken as the input data. The dataset is preprocessed using the min-max normalization method, which performs a linear transformation on the data and preserves the relationships among the data values.
- Next, the data segregation phase is executed. At this stage, the preprocessed data is split into two sets of data, called train data and test data. The small portion of training data is used for testing the large amount of test data.
- CatBoost regressor is identified as the best performing machine learning technique by passing the training and testing data among Catboost Regressor, XGBoost Regressor, Gradient Boosting Regressor, LGBM Regressor and Theil-Sen Regressor models.
- Train CatBoost on the dataset with student features (X_train) and academic outcomes (Y_train). Obtain initial predictions (Y_catboost).
- Concatenate CatBoost predictions (Y_catboost) with original features to create enriched training data (X_rnn_train). Train Bi-RNN on X_rnn_train with corresponding target variable (Y_train). Generate predictions (Y_rnn) capturing bidirectional sequential information.
- Concatenate CatBoost predictions (Y_catboost) and Bi-RNN predictions (Y_rnn) to form meta data (X_meta). Train Random Forest Meta-Regressor on X_meta with target variable (Y_train). Obtain final predictions (Y_final).
- Assess predictive accuracy using Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and R-squared (R2). In this phase, predictions are made based on three subjects, such as math, science, and reading skills.

A synopsis of the paper's structure is as follows: The background information is provided in Section 2 along with an explanation of the current methods and the research related to them. Section 3 gives a description of the dataset and Section 4 portraits the proposed methodology. The preliminary experimental findings and analyses were presented in Section 5. The study's conclusion is noted in Section 6 along with suggestions for more research.

2. Literature Survey

F. Alivernini et al., [16] has suggested an adolescent immigrant's psychological well-being was predicted by school support for autonomy. which looked at the connection between immigrant teenagers' psychological well-being and teacher assistance at school meant to encourage students' autonomy. As control factors, the investigation examined gender, socioeconomic position, prior academic achievement, and immigrant generation. Females are more prone to experience mental health issues than males, gender appears to be the most significant contextual component.

Y. Kim, et al., [17] have recommended a meta-analysis of parental impacts on the academic attainment of immigrant pupils. The current meta-analysis of 14 studies found that three parental influence factors had marginal overall impacts on the achievement-related motivation of immigrant pupils. Parents' close proximity to their children had a significantly greater impact on their achievement-related motivation than did their distance from them or their educational background. The results show how important it is for immigrant children's motivation to do well in school that their parents are close by.

J.J. De Feyter, et al., [18] has suggested the early academic resiliency of youngsters from low-income immigrant households. Because a sizable portion of the immigrant students in our study were black and Latinx in origin, the findings may not apply to other immigrant student populations nationwide. That can be viewed as a weakness, but more recently, researchers have argued that in order to fully grasp what is happening in certain communities, it is necessary to investigate more homogenous community-based samples. Due to the heterogeneity of the immigrant community, there is a wide range of socioeconomic levels.

D. Valero et al., [19] introduced an interactive group for immigrant students that was an important component of their success. That study offers qualitative data demonstrating that in the examined schools. The kinds of interactions that take place in IGs help immigrant students succeed academically and increase coexistence. Because they encourage varied interactions between children and adults during the learning process and include all of the circumstances in which the kids learn and grow, IGs are successful in the schools under study.

A. Miyamoto et al., [20] have invented the accomplishments of immigrant students in relation to their goals for their education and motivation in the classroom uses Data from Germany's National Educational Panel Study on how different pedagogical philosophies affect students' socioeconomic status, educational aspirations, and levels of intrinsic motivation to learn.. Children from some migrant groups have a worse link between educational orientation and reading achievement than their majority-race colleagues, which is an example of the attitude-performance paradox.

In periods of economic and political unrest, immigrant students' motivation and success are examined by Urdan, T., et al., [21]. That paper examines the potential impact of anti-immigrant sentiment on immigrant and refugee students' motivation and academic performance. Then, utilizing Maehr's (1984) theory of personal investment (PI) as a framework, We integrate the various aspects of motivation applicable to the modern context of immigrant and refugee students. Educators will take action in order to combat these challenges and assist children in becoming future individuals who are highly engaged and successful in academics.

F. Borgonovi et al. [22] has recommended the significance of linguistic distance in both the academic success and the feeling of belonging experienced by immigrant students who are not native speakers of English. Utilizing information on 15-year-old children taking part in the Program for International Student Assessment (N =21,618), this study investigates the relationship between the language spoken by non-native-speaking immigrant students and the language of instruction and their outcomes. Reading, math, and scientific success are all correlated with linguistic distance, but a sense of belonging to the school community is not.

M. Karakus, et al., [23] has suggested an examination of the PISA 2018 data at several levels to comprehend the academic performance of students who are first- and second-generation immigrants. Using data from PISA 2018, the current study examines the causes of immigrant kids' academic achievement. The study examines the effects of student characteristics, peer and family support as perceived by the student, school provisions, and equity-focused school policies on immigrant students' academic progress. Utilizing the PISA data collection, the current analysis was limited to testing student, school, and country-level determinants of accomplishment.

According to Lilla, N., et al., [24] the study of academic self-concepts according to an acculturation profile An investigation into a potential factor for immigrant students' academic success. The paper investigated experimentally whether the acculturation profiles of immigrant scholars' general and domain-specific academic self-concept components differ from those of non-immigrant students, with the hypothesis that it is important to take into account the acculturation orientations of immigrant students. We uncover preliminary evidence, based on data from the German National Educational Panel Study (NEPS), that acculturation characteristics have an impact on the aspects of immigrant students' academic self-concept.

Higher education expectations, problems, and obstacles faced by indigenous and immigrant students have been discussed by J. Shankar et al. [25]. The present study used a qualitative technique in order to better understand the concerns and difficulties encountered by students from indigenous communities and immigrant backgrounds who are enrolled in postsecondary human services studies in Western Canada. That paper has argued that initiatives like financial aid and loan programmes won't be effective in achieving their objectives unless they are combined with institutional-level structural changes and the development of transformative learning environments where each student feels heard, valued, supported, and empowered. According to Pallathadka et al., [26], there may be a correlation between the talents and interests of students and their academic performance. Such analysis enables educators to allocate greater attention to students who require assistance the most. Success as an educator is often evaluated based on the academic achievements of his pupils. Every institution should evaluate the caliber of its faculty. Educators may undergo evaluation based on students' outcomes, remarks, and similar factors.

3. Dataset

The goal of the paper is to use a deep learning infused cascading algorithm to forecast the academic success of immigrant students residing in the UAE, Australia , Spain ,Canada, and Qatar. There were 1119 attributes and 612,004 rows in the dataset. Only 35 of the 1119 variables were shown to be significantly related to students' performance in Math, Science, and Reading. The top five nations that accept immigrant students are determined by taking the top five counts of students based on the immigration status defined in the Index Immigration Status attribute, which comprises 29894 rows.

3.1. Dataset Description

The samples were from the 2018 PISA administration conducted by the Organization for Economic Co-operation and Development (OECD). The PISA is a study program used to assess the knowledge and Skills of students aged 15. It assesses a student's aptitude in reading, arithmetic, and science. PISA 2018 will involve 540,000 students from 72 participating nations and economies that are in the 15-year-old age group. The PISA 2018 dataset of the nations that made the shortlist was used to implement the Educational Data Mining (EDM) approach. For the purpose of investigating correlations between students' reading, learning, scientific, and mathematics skills, the PISA 2018 dataset for the UAE, Australia , Spain ,Canada, and Qatar have been assessed. On the PISA website, the created dataset is made available to the general public. PISA countries that have schools and students represent their total population as well as the characteristics of their educational methodology. Nearly 30,000 students' worth of data that includes information about the students and the legal systems of all the listed countries have been carefully examined. The PISA data that is being used is divided into two sections: the first component includes science, reading, and mathematics; the second section includes socioeconomic factors that affect student performance. The responses provided by the students in the questionnaire are encoded using a number of factors. Each student is given a certain weight in order to align the sample with the genuine population because the immigrant pupils must be 15 years old by law.

The complexity of secondary analysis using PISA data is highlighted by the fact that few of the student qualities match the real or original features. Instead, they are available as derived variables after being preprocessed using a variety of advanced methods. For the cognitive exam, for instance, the PISA datasets do not contain a single performance score. This implies that "plausible values" (PVs) are created for each student and their assessment area, which includes reading, math, and science. PISA reports and learning analyses do not offer a straight or single score. As an alternative, the student's abilities and outputs are given PVs. Each student is given a set and constrained amount of testing time. Additionally, the assessment covers a broad spectrum of knowledge and abilities, thus only a few pupils are able to complete each work that is given to them. PISA uses a broad strategy known as the Rasch Model to compare student performance even if they haven't exactly completed the same set of exams, in a reliable manner (OECD PISA 2009). The PISA dataset was developed to evaluate the educational literacies of immigrant students including their preferences and opinions about their school experiences. The PISA dataset results reveal that student learning proficiency levels vary widely, and that these differences may be defined by multitude of factors, including demographic attributes and socio-economic attributes involvement in extracurricular activities, and access to formal education. This study tests literacy using the PISA dataset, which was used to measure levels of reading competence in students. The PISA dataset was then used to analyze the results.

3.2. Dataset Preprocessing

Data preparation is extremely important to any model and is crucial to its success. Additionally, all machine learning approaches are thought to rely heavily on preprocessing data for their models. When executing linear changes on the original data, the Min-Max normalization approach creates a balance of value comparisons between the data before and after the procedure. The following formula is used for this procedure.

$$a_{new} = \frac{a - a_{min}}{a_{max} - a_{min}} \qquad (1)$$

The results of the PISA dataset indicate that student abilities of learning proficiency vary considerably and that this variance is defined through diverse range of factors, including demographic features, socioeconomic status, students' participation in educational activities, and the opportunity to study in school.

Table 1. Attribute details of the dataset

Attribute	Description
CNTRYID	Identifier for Country
CNTSCHID	Identifier for School
IMMIG	Immigration status indicator
ST004D01T	Gender of Students
ST011Q03TA	At your house: a secluded spot for study
ST011Q04TA	Having access to a home computer that can be used for academic purposes
ST013Q01TA	The number of books in your house.
ST097Q01TA	How frequently: students neglect the instructions provided by instructor during class.
ST102Q01TA	How frequently in class: The instructor emphasises specific learning objectives.
ST104Q04NA	How often does teacher provide me with suggestions on how to do better in class?
ST161Q03HA	I am able to read well.
ST161Q07HA	I need to study something multiple times before I can fully comprehend it.
ST059Q02TA	Number of Maths Classes Per Week That Are Usually Required To Be Attended
ST059Q03TA	Number of Science Classes Per Week That Are Usually Required To Be Attended
IC152Q02HA	Education-related digital media consumption within the past month: Mathematics
IC152Q03HA	Education-related digital media consumption within the past month: Science
IC008Q02TA	Utilization of digital devices for activities outside of the classroom, such as team-based online gaming
PV1MATH	Credible Score 1 in Mathematics
PV1READ	Credible Score 1 in Reading
PV1SCIE	Credible Score 1 in Science
MMINS	Studying timeframe (minutes per week) - Mathematics
LMINS	Studying timeframe (minutes per week) - Test Language
SMINS	Studying timeframe (minutes per week)- Science
REPEAT	Repeatedly attending the same class multiple times
MISCED	Education of Mother based on ISCED
FISCED	Education of Father based on ISCED
SC017Q08NA	Impact of School Infrastructure on teaching

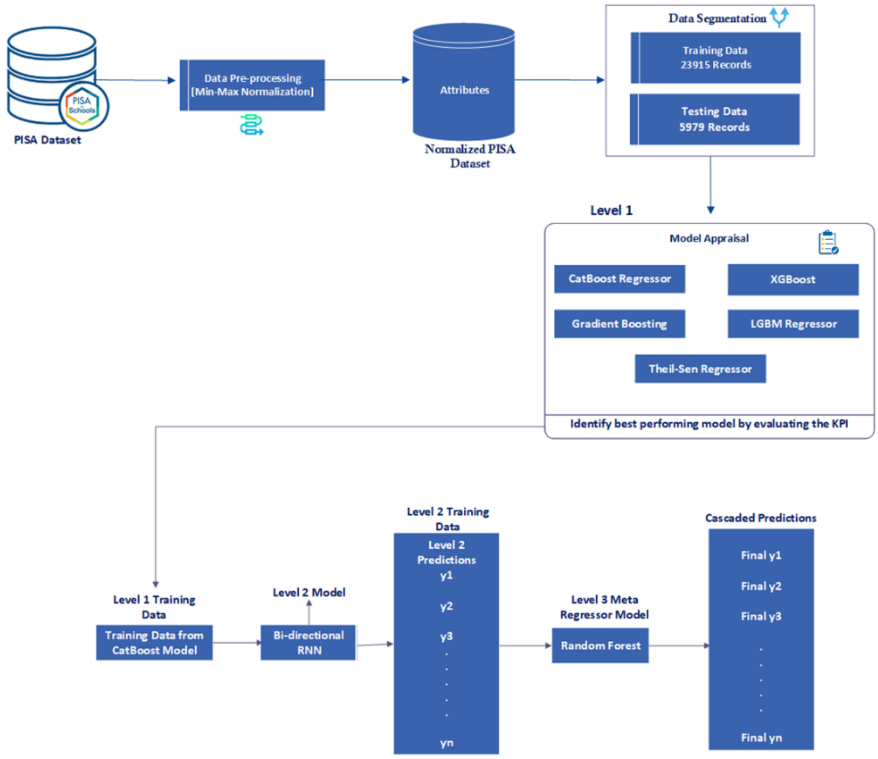
SC001Q01TA	Location of School(Rural, Urban)
SC053Q12IA	Does the School Offer a Reading Group to Encourage Students to Read?
SC150Q05IA	Policies of the school that prioritise equity include reducing class sizes in order to better meet the specific needs of these students.
SC164Q01HA	In the most recent academic year, what percentage of students in final class left school without a certificate?
SC152Q01HA	Is your school providing extra test language sessions outside of regular school hours?
CLSIZE	Size of the Class Room
SCHSIZE	Number of Students in a School
SCHLTYPE	Possession of a School

3.3. Data segregation

The preprocessed data is divided into training data and testing data during the data segregation step. The parameters are altered as specified; firstly, PV1Science and PV1Read are dropped, while the target variable PV1MATH is kept. The procedure is then repeated for the PV1Science and PV1Read variables, with PV1Read and PV1Math being withdrawn and PV1Science being added as the target variable.

Since the majority of the qualities used to describe the students in PISA data are not the same as the true or original attributes, secondary analysis reveals complexity. Instead, they are available as derived variables after being pre processed using a number of complex approaches. In the PISA datasets, for instance, the cognitive exam does not have a single performance score.

Fig. 1. Workflow of the proposed methodology



This indicates that reading, mathematics, and science evaluation PVs are made for each individual kid. In PISA reports and learning analyses, there is no explicit or single score given. A different option is to attribute PV to the student's abilities and performances. Each student will only have a certain amount of time for testing. Additionally, the assessment covers a broad spectrum of knowledge and abilities, so only a few pupils are able to complete each work that is given to them.

4. Proposed Methodology

Accurately forecasting student performance is critical for specific approaches and optimal utilization of resources in education.

This study investigates a cascading ensemble model that uses the strengths of CatBoost, Bidirectional RNNs, and Random Forests to address the multifaceted nature of student data and improve prediction accuracy. Figure 1 illustrates the workflow of the proposed model. This research paper presents a pioneering approach – a cascading regression model that intricately integrates CatBoost Regressor, Bidirectional Recurrent Neural Network (Bi-RNN), and Random Forest Meta Regressor.

5. Experiment & Results

Real-valued variables are utilized as the objective in regression models. The target variables consist of the scores acquired in PV1Math, PV1READ, and PV1SCIE. Level 1 of the model is processed using the following regression algorithms: CatBoost Regressor, XGBoost Regressor, Gradient Boosting Regressor, LGBM Regressor, and Theil-Sen Regressor. CatBoost exhibited superior performance in terms of MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Square Error), and R-Square when compared to all other regression techniques. The variables are diversified as previously mentioned. Initially, the target parameter is PV1Math, while PV1Science and PV1Read are assessed. Subsequently, PV1Science and PV1Math are evaluated while maintaining PV1Read as the objective variable; the procedure is repeated with PV1Science and PV1Math being measured. Table 2, 3 and 4 depicts the performance of the algorithms.

Table 2. Performance of algorithms – Maths as Target variable

Regressors	MAPE	RMSE	R^2
CatBoost	13.812	75.724	0.409
LGBM	14.052	76.957	0.389
XGBoost	14.169	77.461	0.381
Gradient Boosting	14.637	79.506	0.348
Theil-Sen	15.718	84.497	0.264

Table 3. Performance of algorithms – Science as Target variable

Regressors	MAPE	RMSE	R^2
CatBoost	13.761	75.921	0.4264
LGBM	14.065	77.520	0.402
XGBoost	14.027	77.600	0.400
Gradient Boosting	14.599	79.900	0.364
Theil-Sen	15.897	85.770	0.268

Table 2. Performance of algorithms – Reading as Target variable

Regressors	MAPE	RMSE	R^2
CatBoost	13.761	75.921	0.426
LGBM	14.065	77.520	0.402
XGBoost	14.027	77.600	0.400
Gradient Boosting	14.599	79.900	0.364
Theil-Sen	15.897	85.770	0.268

The performance of the CatBoost regressor was higher to that of the other chosen regression models as measured by Loss Function RMSE, MAPE, and R-Square. So in order to increase the performance, the dataset is being trained using the CatBoost and passed on to Bi-directional RNN algorithm for further processing. We have chosen bi-directional as a Level 2 learning model as it is much better at finding temporal relationships and sequential patterns in time series data than one-way RNNs. Unlike ANNs and CNNs, which process data in a unidirectional manner, Bi-RNNs process sequences both forwards and backwards. This bidirectional processing is particularly relevant in academic performance prediction, where historical grades, behavioral trends, and evolving study habits contribute to a student's current performance. By considering information from both past and future time steps, Bi-RNNs enhance the model's ability to capture the nuanced temporal dynamics of academic data.

While LSTM and GRU architectures are renowned for their ability to capture long-term dependencies, their unidirectional nature limits their exposure to future information. In contrast, Saravanan et al., portraits the bidirectional processing of Bi-RNNs allows them to leverage information from both directions, providing a more comprehensive understanding of the context surrounding each data point. In academic scenarios, where the influence of past and future events on current performance is significant, the bidirectional nature of Bi-RNNs aligns well with the underlying dynamics of the data[28].

Bidirectional processing in Bi-RNNs involves training two separate RNNs: one processing the input sequence in its natural order, and the other processing it in reverse. The outputs of these two RNNs are then concatenated or combined in some manner. This bidirectional architecture effectively doubles the capacity of the model to capture context, enabling it to discern patterns that may be missed by unidirectional models. Radhoush et al., used Random forest as a Meta Regressor model as it is good when predicting student performance, the utilization of Random Forest as a Meta-Regressor exhibits significant benefits in comparison to alternative approaches including XGBoost, SVM, and decision trees. Random Forest is a highly effective algorithm for managing complex, nonlinear associations that exist in the data [27]. Its ensemble structure ensures stability against overfitting. By aggregating multiple decision trees, the variance and biases associated with individual models are mitigated, thereby improving the overall predictive stability.

The procedure initiates with a CatBoost Regressor that has been trained on the PISA dataset which contains the academic details of immigrant students (Xtrain) and academic outcomes (Ytrain), producing initial predictions(Ycatboost). For the steps that follow, these are essential inputs. Enriching the context, the Bi-RNN is introduced, leveraging the bidirectional sequential information by concatenating CatBoost predictions with the original features, creating (Xrnn-train). Trained on this enriched dataset, Bi-RNN generates predictions (Yrnn) incorporating nuanced temporal dependencies. The cascading process continues with a Meta Regressor Model, implemented as a Random Forest Regressor, refining predictions further. Final predictions (Yfinal) undergo rigorous evaluation using metrics like MAPE, RMSE, and R2, establishing the accuracy of the cascading regression model. This collaborative approach not only outperforms individual techniques but also provides a holistic understanding of intricate patterns within educational data, opening new avenues for accurate and insightful academic forecasting.

Algorithm 1. Pseudocode Of Proposed - Deep Learning-Infused Cascading Regression Model

Step 1. Data Preprocessing:

- ***X ∈ ℝ^(n×m):*** Input feature matrix, where n is the number of samples and m is the number of features.
- ***y ∈ ℝ^n:*** Target vector of student performance scores.
- ***X_train, y_train***: Training set, a subset of X and y used for model training.
- ***X_test, y_test***: Testing set, a separate subset of X and y used for model evaluation.

Step 2. Level 1: CatBoost Regressor:

- ***f_catboost : ℝ^(n×m) → ℝ^n***: CatBoost prediction function.
- ***ŷ_catboost_train = f_catboost(X_train):*** CatBoost predictions on the training set.
- ***ŷ_catboost_test = f_catboost(X_test):*** CatBoost predictions on the testing set.

Step 3. Level 2: Bidirectional RNN:

- ***f_birnn : ℝ^n → ℝ^n:*** Bidirectional RNN prediction function, taking CatBoost predictions as input.
- ***ŷ_birnn_train = f_birnn(ŷ_catboost_train):*** RNN predictions on the training set.
- ***ŷ_birnn_test = f_birnn(ŷ_catboost_test):*** RNN predictions on the testing set.

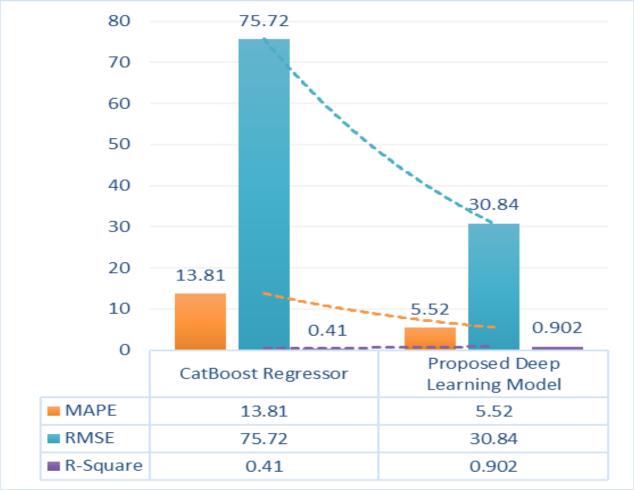
Step 4. Level 3: Random Forest Meta-Regressor:

- ***X_combined ∈ ℝ^(n×2):*** Concatenation of CatBoost and RNN predictions as input features.
- ***f_rf : ℝ^(n×2) → ℝ^n:*** Random Forest prediction function.
- ***ŷ_final = f_rf([ŷ_catboost_test, ŷ_birnn_test]):*** Final predictions on the testing set.

Step 5. Evaluation:

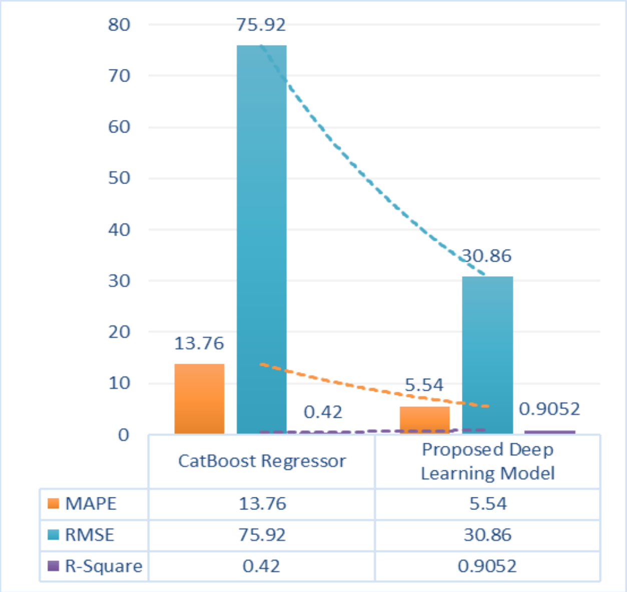
- ***MAPE(ŷ_final, y_test):*** Mean Absolute Percentage Error.
- ***RMSE(ŷ_final, y_test):*** Root Mean Squared Error.
- ***R^2(ŷ_final, y_test):*** Coefficient of Determination.
- Evaluate the metrics to quantify model accuracy.

Figure 2, Figure 3, and Figure 4 highlights the efficacy of the suggested model with Maths, Science, and Reading as the target variables, respectively.



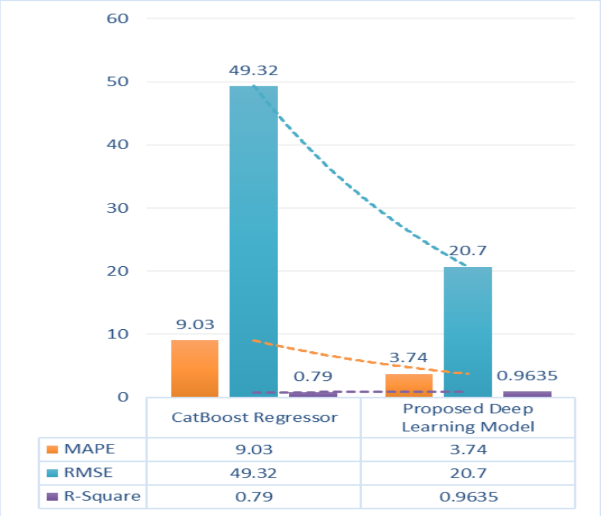
	CatBoost Regressor	Proposed Deep Learning Model
MAPE	13.81	5.52
RMSE	75.72	30.84
R-Square	0.41	0.902

Fig. 2. Performance of proposed model using Maths as target variable



	CatBoost Regressor	Proposed Deep Learning Model
MAPE	13.76	5.54
RMSE	75.92	30.86
R-Square	0.42	0.9052

Fig. 3. Performance of proposed model using Science as target variable



	CatBoost Regressor	Proposed Deep Learning Model
MAPE	9.03	3.74
RMSE	49.32	20.7
R-Square	0.79	0.9635

Fig. 4. Performance of proposed model using Reading as target variable

The question "out of all the predictions generated by our model, what percentage were accurate?" is intended to be the definition of accuracy. The target variable is always continuous in a regression model. Consequently, the model will get overfitted if we start evaluating its performance in terms of different accuracy parameters. The subsequent performance metrics are employed to assess the comparative merits of the proposed model.

5.1. Coefficient of Determination

The statistical measure R squared quantifies the degree of agreement between the observed data and the regression line, also known as the coefficient of determination [29]. R Squared, which is calculated as follows, represents the proportion of the variance in the dependent variable that can be explained by variations in the independent variable.

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{m}(Xi - Yi)^2}{\sum_{i=1}^{m}(\bar{Y} - Yi)^2} \qquad (2)$$

The proposed Deep Learning-Enhanced Cascading Regression model exhibited better results, as illustrated in Figures 2, 3, and 4. Furthermore, when PV1Math was used as the target variable, it generated a R Square value of 0.902. When PV1Science was the target variable, the R Square value was 0.9052 and When PV1Reading was the target variable, the R Square value was 0.9635.

5.2. Root Mean Squared Error (RMSE)

Providing an average measurement of the error, the root-mean-squared error (RMSE) is a polynomial ranking technique [30]. It is more precisely the root square of the sum of all squared disparities between the predicted value and the actual value. Calculating RMSE as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((Xi - Yi)^2)} \qquad (3)$$

The enhanced outcomes of the proposed Deep Learning-Enhanced Cascading Regression model are depicted in Figures 2, 3, and 4. Further, the selection of PV1Math as the target variable resulted in an RMSE value of 30.84. The root mean square error (RMSE) for the target variable PV1Science was 30.86, whereas for the target variable PV1Reading, it was 20.70.

5.3. Mean Absolute Percentage Error (MAPE)

Assigning the average absolute percentage error between the predicted and observed values, the MAPE statistic evaluates the precision of a regression model. Mean absolute percentage error can be utilized to assess the efficacy of a regression model. Calculated in the subsequent manner:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Yi - Xi}{Yi}\right| \qquad (4)$$

Where X_i – absolute value,

Y_i – Predicted value,

n – number of observations

Based on the performance metrics extracted from Figures 2, 3, and 4 the proposed Deep Learning-Enhanced Cascading Regression model is good with low MAPE of 5.52 when PV1Math as target variable, 5.54 when PV1Science as target variable and 3.74 when PV1Reading as target variable.

Our methodology presents a novel framework for forecasting immigrant student performance by employing a cascading regression model. By integrating CatBoost, a conventional gradient boosting algorithm, Bidirectional Recurrent Neural Network (Bi-RNN), a deep learning model, and Random Forest Meta-Regressor, an ensemble learning technique, this model achieves a strategic integration. By utilizing complementary strengths in a sequential fashion, these models can be cascaded, resulting in a more resilient and sophisticated predictive framework. In contrast to conventional standalone models, our hybrid approach effectively incorporates both linear and nonlinear relationships present in academic datasets by building on the interpretability of CatBoost and the capability of Bi-RNN to handle complexity. By collecting dependencies from both past and future data points, the integration of Bi-RNN introduces a sophisticated temporal dimension that enhances the model's comprehension of sequential patterns. The transparent insights into feature importance offered by the Random Forest Meta-Regressor further augment the interpretability of our model. This is a crucial aspect for stakeholders operating within educational contexts.

6. Conclusion & Future Work

Relocation can have a detrimental influence on academic achievement and the creation of human capital and the inadequate care that children receive after one or both parents move away is a contributing factor in these unfavorable outcomes. The number of family members who can assist students with their schoolwork is also likely to decline following migration. In this paper, our research introduces a sophisticated cascading regression model that leverages the collective strengths of CatBoost, Bidirectional Recurrent Neural Network (Bi-RNN), and Random Forest Meta-Regressor to enhance the precision of immigrant student performance prediction using PISA dataset. The predictive capabilities of deep learning and traditional regression techniques are enhanced through their collaborative synergy, surpassing those of standalone models. The Random Forest Meta-Regressor not only enhances prediction accuracy but also offers significant insights into the determinants of academic achievements through its interpretability. By combining well-established regression techniques with deep learning methodologies, we gain a more comprehensive understanding of academic data and gain access to a robust tool for educational analytics. The performance of the proposed model was outstanding compared to other models as it produced MAPE reduction of 3.74, an RMSE reduction of 20.70, and an R-squared increase of 0.96.

We anticipate that our model could be further enhanced by Investigating additional deep learning architectures or ensemble methods may result in a further improvement of predictive precision. Conducting an evaluation of the model's applicability to various educational settings and populations is critical in order to ascertain its resilience. The potential for enhanced predictive capability exists when additional characteristics associated with socio-economic factors, learning patterns, or engagement levels are integrated. Ongoing optimization of the cascading architecture continues to require the application of techniques such as feature engineering and hyperparameter tuning. In addition, by broadening the scope to encompass longitudinal data and real-time updates, a dynamic perspective on trends in academic performance could be achieved. Incorporating input and validation from educational practitioners and stakeholders in the real world will enhance the practicality and applicability of the proposed model.

References

[1] A.G Langenkamp, *Latino/a immigrant parents' educational aspirations for their children*. Race Ethnicity and Education, 22(2),2019, pp.231-249.

[2] A. Hadjar and J. Scharf, The value of education among immigrants and non-immigrants and how this translates into educational aspirations: a comparison of four European countries. Journal of Ethnic and Migration Studies, 45(5),2019, pp.711-734.

[3] J. Orupabo, I. Drange and B. Abrahamsen, Multiple frames of success: How second-generation immigrants experience educational support and belonging in higher education. Higher education, 79(5), 2020, pp.921-937.

[4] A.Kumi-Yeboah, Educational resilience and academic achievement of immigrant students from Ghana in an urban school environment. Urban Education, 55(5), 2020, pp.753-78.

[5] SH. Ham, H. Song and K.E. Yang, Towards a balanced multiculturalism? Immigrant integration policies and immigrant children's educational performance. Social Policy & Administration, 54(5), 2020, pp.630-645.

[6] M.Melkonian, S. Areepattamannil, L.Menano and P.Fildago, Examining acculturation orientations and perceived cultural distance among immigrant adolescents in Portugal: Links to performance in reading, mathematics, and science. Social Psychology of Education, 22(4),2019, pp.969-989.

[7] M. Triventi, E.Vlach and E.Pini, *Understanding why immigrant children underperform: evidence from Italian compulsory education*. Journal of Ethnic and Migration Studies, 48(10), 2022, pp.2324-2346.

[8] S. Rodríguez, A. Valle, L.M. Gironelli, E. Guerrero, B. Regueiro and I. Estévez, *Performance and well-being of native and immigrant students. Comparative analysis based on PISA 2018*. Journal of Adolescence, 85,2020, pp.96-105.

[9] X. Ding, X. Chen, R.Fu,, D. Li, and J. Liu, *Relations of shyness and unsociability with adjustment in migrant and non-migrant children in urban China*. Journal of Abnormal Child Psychology, 48(2), 2020, pp.289-300.

[10] R. Guerra, R.B. Rodrigues, C. Aguiar, M. Carmona, J. Alexandre and R.C Lopes, *School achievement and well-being of immigrant children: The role of acculturation orientations and perceived discrimination*. Journal of school psychology, 75, 2019, pp.104-118.

[11] A.A Ismail, *Immigrant children, educational performance and public policy: A capability approach*. Journal of international migration and integration, 20(3), 2019, pp.717-734.

[12] M.Pivovarova,and J.M. Powers, Generational status, immigrant concentration and academic achievement: comparing first and second-generation immigrants with third-plus generation students. Large-scale Assessments in Education, 7(1), 2019, pp.1-18.

[13] G.Gabrielli, S.Longobardi and S.Strozza, *The academic resilience of native and immigrant-origin students in selected European countries*. Journal of Ethnic and Migration Studies, 48(10), 2022, pp.2347-2368.

[14] J.M. Cordero, C. Polo and R. Simancas, Assessing the efficiency of secondary schools: evidence from OECD countries participating in PISA 2015. Socio-Economic Planning Sciences, 2022, p.100927.

[15] T. Delahunty, N. Seery and R. Lynch, *Exploring problem conceptualization and performance in STEM problem solving contexts*. Instructional Science, 48(4), 2020, pp.395-425.

[16] F. Alivernini, E. Cavicchiolo, S.Manganelli, A.Chirico and F.Lucidi, *Support for autonomy at school predicts immigrant adolescents' psychological well-being*. Journal of Immigrant and Minority Health, 21(4), 2019, pp.761-766.

[17] Y. Kim, S.Y. Mok and T.Seidel, Parental influences on immigrant students' achievement-related motivation and achievement: A meta-analysis. Educational Research Review, 30, 2020, p.100327.

[18] J.J. De Feyter, M.D. Parada, S.C Hartman, T.W Curby and A. Winsler, *The early academic resilience of children from low-income, immigrant families*. Early Childhood Research Quarterly, 51, 2020, pp.446-461.

[19] D. Valero, G. Redondo-Sama and C. Elboj, *Interactive groups for immigrant students: a factor for success in the path of immigrant students*. International Journal of Inclusive Education, 22(7), 2018, pp.787-802.

[20] A.Miyamoto, J.Seuring and C.Kristen, *Immigrant students' achievements in light of their educational aspirations and academic motivation*. Journal of Ethnic and Migration Studies, 46(7), 2020, pp.1348-1370.

[21] T.Urdan, N.Sharma and M.Dunn, Motivation and achievement of immigrant students in times of economic and political instability. In Motivation in education at a time of global change. Emerald Publishing Limited,Vol. 20, 2019, pp. 169-184.

[22] F.Borgonovi and A.Ferrara, Academic achievement and sense of belonging among non-native-speaking immigrant students: The role of linguistic distance. Learning and Individual Differences, 81,2020, p.101911.

[23] M. Karakus, M. Courtney and H.Aydin, Understanding the academic achievement of the first-and second-generation immigrant students: a multi-level analysis of PISA 2018 data. Educational Assessment, Evaluation and Accountability,2023, pp.1-46.

[24] N.Lilla, S.Thürer, W.Nieuwenboom and M.Schüpbach, Exploring Academic Self-Concepts Depending on Acculturation Profile. Investigation of a Possible Factor for Immigrant Students' School Success. Education Sciences, 11(8), 2021, p.432.

[25] J.Shankar, E.Ip and N.E. Khalema, Addressing academic aspirations, challenges, and barriers of indigenous and immigrant students in a postsecondary education setting. Journal of Ethnic & Cultural Diversity in Social Work, 29(5), 2020, pp.396-420.

[26] Pallathadka H, Wenda A, Ramirez-Asís E, Asís-López M, Flores-Albornoz J and Phasinam K. *Classification and prediction of student performance data using various machine learning algorithms*. Materials today: proceedings. 2023 Jan 1;80:3782-5.

[27] Radhoush S, Vannoy T, Whitaker BM, and Nehrir H. *Random Forest Meta Learner for Generating Pseudo-*

Measurements in Active Distribution Power Networks. In2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT) 2023 Jan 16 (pp. 1-5). IEEE.

[28] Saravanan, D and Kumar, KS, Improving air pollution detection accuracy and quality monitoring based on bidirectional RNN and the Internet of Things. Materials Today: Proceedings. 2023 Jan 1;81:791-6.

[29] Botchkarev A. Evaluating performance of regression machine learning models using multiple error metrics in azure machine learning studio. Available at SSRN 3177507. 2018 May 12.

[30] Chicco D, Warrens MJ, Jurman G, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science. 2021 Jul 5;7:e623.