EURASIAN JOURNAL OF MATHEMATICAL AND COMPUTER APPLICATIONS
ISSN 2306–6172
Volume 14, Issue 2 (2026) 95 – 114

DEEP LEARNING AND GRADIENT BOOSTING FOR PREDICTING EDUCATIONAL ATTAINMENT BASED ON SOCIOECONOMIC AND REGIONAL FEATURES

Najjar-Ghabel S. [1], **Yousefimehr B.**
Daneshpour A., **Kozhanov M.**, **Amirov A.**

Abstract Educational attainment plays a critical role in societal outcomes. It affects employment prospects, economic development, and social mobility. However, the prediction of educational achievement remains difficult due to complex socioeconomic and regional factors. Many previous studies focus mainly on academic variables. The combined effects of demographic, economic, and geographical factors are often ignored. In addition, many studies rely on single learning models with limited generalizability. To address these gaps, a novel meta–model blending approach is proposed in this study. The approach combines the strengths of a Deep Neural Network (DNN) and Gradient Boosting (GB). The aim is to improve the accuracy and robustness of educational attainment prediction. A preprocessing pipeline is applied before model development. Noise and missing values are handled. Feature refinement is performed. Min–Max normalization is also applied. After these steps, a hybrid ensemble model is constructed. The model is evaluated with a dataset of 1100 samples. The dataset includes 22 socioeconomic and regional features from towns in England. The proposed approach presents high performance with $R^2 = 0.9975$, Mean Absolute Error (MAE) = 0.1582, and Root Mean Squared Error (RMSE) = 0.2058, which exceeds that of single DNN, single GB, and simple blending methods.

Keywords: Deep Neural Networks, Computational Mathematics, Educational Attainment, Gradient Boosting, Socioeconomic, Regional Features, Machine Learning, Information Systems, Applied Mathematics.

AMS Mathematics Subject Classification: 68T05, 68T07, 68M11, 62H30.

DOI: 10.32523/2306-6172-2026-14-2-95–114

1 Introduction

Educational attainment is a key determinant of individual and collective prosperity in modern societies [1]. Higher levels of education are associated with better employment prospects, higher income, improved health outcomes, and stronger civic participation. At the societal level, a well educated population supports economic growth, innovation, and long term social stability [2]. However, large disparities in educational attainment remain across countries and regions. We observed that such disparities may tend to reflect socioeconomic and geographic inequalities. Such differences reduce opportunities for the young generation, and/or reinforce poverty cycles, and/or restrict social mobility. Thus, an in-dept understanding of educational attainment patterns is would be essential. For that we require robust and predictive models to be able to better study the groups and regions that face higher educational risks. Such knowledge supports the design of interventions that improve equity to expand opportunities, and support stable societies [3]. In addition, it is well studied that the socioeconomic

[1]Corresponding Author.

and regional factors can influence educational attainment through resource availability and learning conditions [4]. Nevertheless, the traditional prediction methods rely on manual analysis or statistical models [3, 5]. It has been well researched that methods as such cannot represent complex nonlinear relationships. Fortunately,artificial intelligence with the help of machine learning methods can provide reliable anticipations [6, 7, 8] where several variables can be analyzed simultaneously, via undestanding the hidden patterns [9, 10]. More accurate predictions are therefore achieved. These methods support educational planning and policy decisions [11, 12]. Recent studies apply statistical and machine learning approaches for educational attainment prediction [6, 13, 14, 15]. However, most models rely mainly on academic performance indicators. Socioeconomic and regional factors receive limited attention. Single model structures also reduce robustness, generalizability, and interpretability. We therefore require an integrated approaches that combine diverse data sources and complementary learning methods for reliable educational attainment prediction. This study contribute by facilitating this process through a novel approach. Concequntly, in this study, a hybrid modeling approach is proposed. The approach integrates Deep Learning (DL) and Gradient Boosting (GB) through integration of a meta model. To do that, we first apply an initial preprocessing step where the noise and missing values are handled. In addition, in this first step irrelevant features are removed, incomplete records are excluded, and data normalization is further conducted. In the next step, during model development, two predictive models are trained. The proposed models include a Deep Neural Network (DNN) and a GB model. The models' outputs are combined through a linear regression meta model to improve prediction accuracy. The proposed approach generates generalizable predictions through the inclusion of socioeconomic and regional variables in the modeling process. The proposed model further contributes by providing analytical support for policymakers in education systems. Data based insights can guide decisions related to educational equity and strategic planning. The objective is to implement a hybrid ensemble model for educational attainment prediction. For that a deep neural networks, gradient boosting, and meta model blending are integrated. This implementation provides a process where the socioeconomic, regional, and academic variables can be analyzed. Furthermore the prediction accuracy and model interpretability can be constantly improved. Policy relevance is also strengthened through broader data integration. A systematic preprocessing is applied for noise treatment and feature refinement. Data normalization improves model stability and prediction reliability. Extensive experiments validate the proposed approach across multiple evaluation settings. The proposed model outperforms conventional single model approaches. Better robustness is achieved under different data conditions. High scalability is also demonstrated across diverse educational datasets and regional contexts. We further evaluate model adaptability across different urban regions. Reliable predictions support educational planning and policy development. Data driven evidence improves resource allocation and targeted educational interventions.

2 Related Work

In the recent decades, predicting educational attainment has gained popularity. The relevance and applicability had been also increased worldwide with the increasing use of AI. This section reviews the most relevant studies on student performance prediction which highlights general challenges in the literature. In this context, Alyahyan and Düştegör [2] have applied ML to predict academic success in higher education, which provides a systematic meta analsyis and review. The authors review ML methods and detail best practices to define success, select student attributes, and to choose appropriate models. Ghorbani and Ghousi [14]

have also compared the effectiveness of different resampling methods in imbalanced datasets in predicting student performance using ML. The authors evaluate several resampling approaches, including Borderline SMOTE, Random Over Sampler, SMOTE, SMOTE-ENN, SVM-SMOTE, and SMOTE-Tomek, applied to two educational datasets and evaluate their performance with a range of classifiers such as Random Forest (RF), K-Nearest Neighbor (K-NN), Artificial Neural Network (ANN), XGBoost, Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), and Naïve Bayes (NB) They also use cross-validation and statistical significance testing (Friedman test). Its strength is the systematic comparison, which reveals that SVM-SMOTE combined with RF yields the best results for imbalanced data. However, its performance decreases as the number of classes increases because greater class diversity intensifies data imbalance and feature overlap. Indeed, it becomes difficult for classifiers to identify patterns in datasets with varying distributions, sizes, and feature importance. Zeineddine et al. [16] have improved the prediction of student success in higher education by employing Automated ML (AutoML) to identify the most effective predictive models using pre-admission data. It utilizes AutoML frameworks for model selection and data balancing for accurate prediction. The key advantage of the paper is that AutoML streamlines the modeling process and leads to more accurate, data-driven interventions for at-risk students, while a disadvantage is the potential "black box" nature of AutoML, which can make interpretation and trust in the predictions more challenging for educational stakeholders. Moreover, its reliance on pre-admission data and automated model selection, which restricts the integration of broader socioeconomic and regional factors, highlights the need for a more transparent and context-aware framework to achieve interpretable educational attainment prediction. Yağcı [17] has predicted undergraduate students' final exam grades using ML. He studies the midterm grades, department, and faculty information. His study compares the performance of several ML, e.g., RF, K-NN, SVM, LR, and NB, on a dataset of 1854 students, which achieved classification accuracy rates between 70–75%. A key strength of his approach was its capacity to detect at-risk students and support intervention efforts. However, his limitation is that relying on a small number of features which overlooks socioeconomic and regional factors may reduce the accuracy and applicability of the models in different educational environments. On the other hand, Pallathadka et al. [18] predicted student academic performance using ML algorithms on real data. The authors compare several algorithms, including NB, ID3, C4.5, and SVM, by evaluating the accuracy and error rates on a student performance dataset from the UCI repository. The main advantage of this paper is its systematic comparison of widely-used algorithms and practical demonstration on real data, while its key disadvantage is the reliance on past academic records, which may overlook broader social or contextual factors affecting performance. Kuadey et al. [19] have examined how different sources of technostress affect student learning burnout and perceived academic performance using ML. It collects survey data from university students, applies structural equation modeling, and uses ML algorithms, including RF and SVM, to analyze the relationships between technostress creators, burnout, and performance. This study combines behavioural variables and ML-based models to gain deeper insights into student well-being. However, it depends on psychological and self-reported variables, which limit its ability to capture socioeconomic and regional influences that affect educational attainments.

Villar and de Andrade [6] used different algorithms are such as Decision Tree (DT), Support Vector Machine (SVM) and Random Forest (RF). We also use advanced ensemble methods such as Gradient Boosting (GB), XGBoost, CatBoost and LightGBM. Synthetic oversampling is used to solve the class imbalance and parameters of models are optimised with Optuna. A major strength of this study is the comprehensive evaluation of multiple ML

techniques together with a systematic treatment of imbalanced data. However, the dataset is obtained from a single institution, which limits the generalizability of the findings. In addition, socioeconomic and regional variables are not included in the analysis. These factors have a known influence on educational attainment. This limitation indicates the need for models that integrate contextual variables to improve prediction accuracy, interpretability, and applicability across regions. In the other research by Kordbagheri et al. [13] they predict university students' completion of academic majors through ML methods. The study focuses on personality traits and other nontraditional predictors. Probabilistic evaluation metrics such as pseudo R^2 are applied. Model tuning is conducted through resampling and cross validation. Several algorithms are compared to identify the most influential factors associated with academic persistence. Their results indicate that particular factors, i.e., agreeableness, conscientiousness, and emotional stability are important predictors. A major strength of the study lies in the use of rigorous validation procedures and probabilistic evaluation to improve prediction reliability. However, strong emphasis on personality traits limits the broader applicability of the results. Important socioeconomic and environmental variables are not considered in the analysis. A summary of literature on educational attainment prediction is presented in Table 1, where the table includes a wide range of ML models that has been applied to identify students at academic risk and to support data based decision making in education. Although many approaches report promising performance, most rely on limited academic or institutional variables. Broader socioeconomic and regional factors that influence educational attainment are often excluded. Several challenges also remain unresolved, including data imbalance, limited feature diversity, and weak generalizability across different educational contexts. In addition, many existing models do not achieve both interpretability and robustness when applied to new datasets. These limitations present the research gap. A hybrid predictive approach is required that integrates heterogeneous data sources and advanced learning methods to improve accuracy, scalability, interpretability, and practical applicability.

3 Proposed Method

As for the limitation of the literature on educational attainment we have identified in our review, we will try to suggest a machine learning approach for data integration from different sources. Moreover, we should make sure about the accuracy, interpretability, and generalizability of predictions made by the proposed model. This model uses strict data pre-processing and ensembles of models construction. Table 2 presents the key symbols and their descriptions used throughout the experimental design and analysis of the proposed method. This gives consistent terminology for describing the data set structure, partitioning ratio, and parameters of model training. It should be stressed that the Learning Rate (LR) shown in Table 2 is a hyper-parameter controlling the amount by which the weights of the model are modified in each training step. This will help establish the pace and stability of the learning process.

The architecture of the proposed method is illustrated in Figure 1. It has two main phases: 1) data preparation and 2) model construction. In the data preparation phase, the dataset is refined in a systematic way to ensure consistency and readiness for ML analysis. First, noisy and missing values are identified and corrected to reduce bias and information loss. Irrelevant or redundant features are then removed to reduce dimensionality and improve computational efficiency. Incomplete rows are also excluded because they may distort model training. The Min Max normalization is applied to standardize all input features. A common value range is established across variables. Feature dominance is reduced through this process. Model

Table 1: Comparative summary of recent supervised studies on educational attainment prediction using ML

Authors/ Year	**Objective**	**Algorithms**	**Advantages**	**Disadvantages**
Alyahyan & Düştegör (2020) [2]	Provide guidelines for predicting student success in higher education	ML algorithms	Offer a step-by-step process for educators to use ML	Require some technical knowledge, guidelines may not fit all contexts
Ghorbani & Ghousi (2020) [14]	Compare different resampling methods for predicting student performance	RF, K-NN, ANN, XGBoost, SVM, DT, LR, NB	Comparison of classifiers	Performance drops as the number of classes increases, dataset dependency
Zeineddine et al. (2021) [16]	Improve the prediction of student success	AutoML	Streamline model selection, accurate interventions	AutoML, as a "black box," makes model interpretation difficult
Yağcı (2022) [17]	Predict students' final exam grades	RF, K-NN, SVM, LR, NB	Identify at-risk students, help with intervention planning in the real world	Limited features restrict model power and generalizability
Pallathadka et al. (2023) [18]	Predict student performance	NB, ID3, C4.5, SVM	Comparison of ML algorithms, tested on real data	Focus on academic records, misses broader contextual influences
Kuadey et al. (2024) [19]	Examine the impact of technostress creators on student performance	RF, SVM, Structural Equation Modeling	Integrate behavioural factors with ML	Rely on self-reported data, limited generalizability
Villar and de Andrade (2024) [6]	Compare supervised ML algorithms for predicting academic success	DT, SVM, RF, GB, XGBoost, CatBoost, LightGBM	Comprehensive evaluation of ML methods, effective management of imbalanced data	Results based on a dataset of a single institution, limited generalizability
Kordbagheri et al. (2025) [13]	Predict academic major completion using personality traits	ML algorithms with probabilistic assessment	Strong validation and probabilistic evaluation enhance prediction reliability	Focus on personality traits, limited generalizability, neglect key influencing factors

convergence and training stability are also improved. Clean, balanced, and standardized data are therefore obtained. This preparation supports reliable model performance. A hybrid machine learning model is then developed through deep neural networks, gradient boosting, and meta model blending. Complex nonlinear and structured relationships are effectively captured. Batch normalization and dropout are applied to reduce overfitting. Gradient boosting improves prediction accuracy through iterative error reduction. A meta model combines predictions from both learning models. We use this strategy to exploit the strengths of each algorithm. Deep neural networks capture complex feature relationships. Gradient boosting improves precision and feature interpretation. Meta model blending increases prediction stability and generalizability. Accurate and interpretable educational attainment prediction is therefore achieved across diverse socioeconomic and regional contexts.

Table 2: Summary of symbols and their descriptions

Symbol	Description
N	The total number of data records (rows)
k	Number of features
P_{train}	Percentage of the train set
P_{test}	Percentage of the test set
P_{val}	Percentage of the validation set
n_{train}	The number of records in the training set
n_{test}	The number of records in the test set
n_{val}	The number of records in the validation set
LR	Learning Rate

3.1 Data Preparation

The data preparation phase converts raw data into a suitable format for machine learning analysis. Data quality is improved through cleaning, transformation, and standardization. Reliable and consistent datasets are therefore obtained. We use this process to support accurate and reliable predictive model development.

3.1.1 Treatment of Noisy and Missing Values

Treatment of noisy and missing values is essential for reliable educational attainment prediction. Outliers, inconsistent records, and input errors are identified and corrected. Missing values are treated through imputation or row exclusion. Appropriate methods are selected according to feature characteristics. High quality and unbiased data are therefore obtained. The cleaned dataset reflects socioeconomic and regional conditions more accurately. Consistent data are then provided to deep neural networks and gradient boosting models. Stable feature representations are learned through this process. Overfitting to noisy patterns is also reduced. Model integrity, robustness, and interpretability are improved. Nonlinear relationships among educational, socioeconomic, and regional factors are more accurately identified. Reliable predictive evidence is therefore provided for educational planning and policy decisions.

3.1.2 Removal of Irrelevant Features

Elimination of the redundant features is an essential step in data preparation process. Many educational datasets contain identifiers, repeated attributes, and weak indicators that add little value to educational attainment prediction. If such features are retained, model complexity may increase. Computational cost may also rise. Multicollinearity and noise may be introduced, and prediction precision may be reduced. Concequently, the proposed method improves efficiency and interpretability through the exclusion of these features. The models can then focus on the most important socioeconomic, regional, and academic predictors. This refinement step helps the DNN extract more meaningful high-level representations from relevant data patterns. It also helps GB create efficient decision boundaries and reduce residual errors. Therefore, targeted feature selection improves model clarity, computational stability, and practical use.

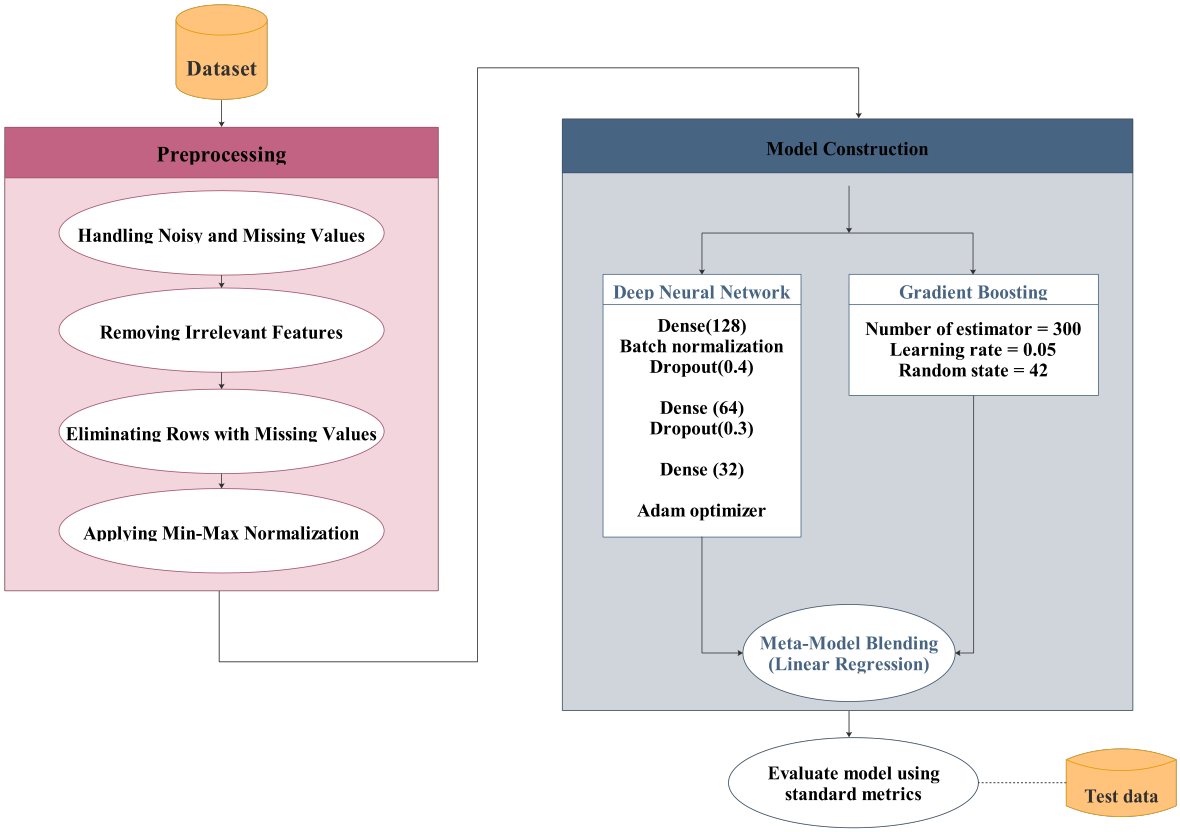

Figure 1: The overall scheme of the proposed educational attainment prediction method

3.1.3 Exclusion of Rows with Missing Values

Rows with missing values are excluded during data preparation. Dataset integrity and consistency are therefore improved. Incomplete records reduce prediction accuracy and introduce bias. Correlations among socioeconomic and regional variables are also distorted. Only complete observations are used for model training and validation. Reliable and unbiased learning is therefore achieved. This step is especially important for deep neural networks and gradient boosting models. Consistent data distributions improve model stability and learning quality. Deeper feature representations are learned by deep neural networks. More reliable decision trees are constructed by gradient boosting. Model robustness and validity are therefore improved. Complex relationships among educational, socioeconomic, and regional factors are more accurately identified.

3.1.4 Min–Max Normalization

The final step in the data preparation phase is Min–Max normalization. This method rescales all numerical features to a common range between 0 and 1 [20, 21]. This transformation ensures that each socioeconomic, regional, and academic variable contributes fairly to the ML process, regardless of its original scale. Without normalization, features with larger numerical ranges, such as population size and income indicators, may dominate the model process. They may also reduce the effect of smaller-scale features and limit balanced pattern detection. Min–

Max normalization improves model stability, convergence speed, and computational efficiency through standardized input magnitudes. It helps the DNN optimize weight updates and helps GB maintain consistent split criteria during tree construction. It also reduces training bias and ensures that all features contribute to prediction. Finally, Min–Max normalization helps ML algorithms identify complex interdependence among socioeconomic and regional determinants of educational attainment.

3.2 Model Construction

The phase of building the model combines multiple ML techniques to improve predictive generalizability. Our proposed method departs from the traditional single-model methods through the integration of DNN, GB, and meta-model blending. This combination captures complex patterns, reduces errors, and uses the strengths of different algorithms. Before model construction, the prepared dataset is divided into training, validation, and test subsets according to predefined proportions (P_{train}, P_{test}, P_{val}). This division supports fair model evaluation and prevents data leakage.

3.2.1 Deep Neural Network (DNN) Architecture

The first component in the model construction phase is the DNN. This method is developed to identify complex and nonlinear relationships among socioeconomic, regional, and academic variables. Our proposed architecture starts with a dense layer of 128 neurons. In addition, the batch normalization and a dropout rate of 0.4 are then applied. These components can help to standardize the activations and will reduce overfitting. The next layers include a dense layer with 64 neurons and a dropout rate of 0.3, followed by a final dense layer with 32 neurons. The network is trained with the Adam optimizer. Adam is known for its adaptive learning rate and efficient convergence, and it will enhance the training stability [7]. It is an architecture which facilitates the learning of high-level abstraction and complex feature interactions of the DNN which might be overlooked by simple machine learning algorithms. The architecture of the deep neural network is created to deal with advanced education datasets. The large number of nodes in the first hidden layer will be used for learning complex high-dimensional feature interaction. The dropout regularization will assist in decreasing overfitting and increase the generalization capability of the algorithm. In further hidden layers, the features will be downsized. As a result, the extracted feature representations will be compact and discriminative. Excessive use of dropout will keep the regularization and information retention in balance. The Adam optimizer is used due to its ability to provide stable convergence even with the presence of different feature scales. The choice of this architecture aims at increasing the precision of the prediction, robustness and learning stability. Thus, all of these aspects allow us to conclude that the DNN architecture provides advantages for predicting educational attainment. Therefore, the use of multiple layers and regularization techniques in the model will allow identifying and improving data patterns.Thus the structured DNN permits the final model to make good use of the available data. It provides reliable, data based, and interpretable predictions of educational attainment.

3.2.2 Gradient Boosting Architecture

The second component of the proposed model is the GB algorithm. GB is considered as an ensemble method that improves prediction performance by combining several weak learners [22, 23, 24]. In this study, GB is configured with 300 estimators, a learning rate of 0.05, and a fixed random state of 42. We must note that, these settings support stable convergence and

provide a balance between model complexity and convergence speed. The 300 estimators are selected to balance model complexity and computational efficiency. This number of boosting iterations allows the model to correct residual errors without overfitting or excessive training time. A learning rate of 0.05 is selected to support gradual and stable convergence. Each tree therefore makes a modest contribution to the final prediction, which improves generalization. The random state is fixed at 42 to ensure consistent results across different runs. This consistency is important for experimental validation. These hyperparameter settings allow the GB model to capture complex nonlinear relationships among socioeconomic and regional variables while it maintains high accuracy, stability, and robustness. The proposed method has several advantages in using GB. GB is good for modeling complex and non-linear relationships. It also performs well for heterogeneous data with different distribution. This algorithm is incorporated into the proposed method to improve the prediction accuracy, reduce the bias and variance, and increase the generalizability. GB's properties make it appropriate for prediction of educational achievement from a range of socioeconomic and regional factors.

3.2.3 Meta-Model Blending

A meta-model blending strategy is applied to further improve prediction accuracy and model robustness. This strategy combines the strengths of both DNN and GB models. In this approach, predictions produced by the DNN and GB models are used as input features for a secondary meta-model. Linear Regression is used as the meta-model. It learns how to combine the outputs into a final prediction [25]. This stacking strategy uses the complementary strengths of the base models. The DNN captures complex nonlinear patterns and deep feature interactions. GB manages tabular data effectively, corrects individual model errors, and reduces bias and variance. Therefore, the proposed method achieves higher accuracy, stability, and generalizability than any single model. This integration is useful for educational attainment prediction because it combines complex relationships within diverse socioeconomic and regional data. More reliable outcomes can therefore be obtained.

During the blending process, the training dataset is first divided into K folds. Each base model, namely DNN and GB, is trained on $K-1$ folds. Predictions are then produced on the remaining fold. The out-of-fold predictions are concatenated to form a new dataset. This dataset is used as input for the meta-model. The Linear Regression meta-model learns an optimal weighting scheme. This scheme defines the contribution of each base model to the final prediction. Through this hierarchical process, the blending approach improves accuracy and reduces model uncertainty. This property is valuable for educational policy applications, where transparent and data based decision support is required. Algorithm 1 presents the overall workflow of the proposed method. The input is a raw data set with N records and k features. You get a trained model that predicts and metrics on how well it predicts. The process begins with the data preparation in Lines 1–5. In this phase, the data set is cleaned and standardized. Line 2 deals with noisy and missing data to enhance data quality. In Line 3 we eliminate irrelevant columns to reduce redundancy. The rows with missing values are left out in Line 4 to avoid biased learning. Line 5 performs Min–Max normalization to scale features to a common range and to avoid the dominance of variables with large values. The prepared dataset is then divided into training, test, and validation subsets according to the defined ratios (P_{train}, P_{test}, P_{val}). This division is performed in Line 6 to support fair performance evaluation. The model construction phase is described in Lines 7–17. In this phase, the hybrid predictive method is built through the integration of several ML components. The DNN is designed with successive layers, including `Dense(128)`, Batch Normalization, `Dropout(0.4)`, `Dense(64)`, `Dropout(0.3)`, and `Dense(32)`. The Adam optimizer is used to capture nonlinear

patterns, stabilize training, and reduce overfitting. These steps are shown in Lines 7–15. Next, a GB model is trained to reduce residual errors in Line 16. A Linear Regression meta-model then combines the predictions from DNN and GB in Line 17. The final model combines the complementary capabilities of DNN and GB to improve generalization. Finally, Lines 18–19 describe the evaluation phase. The trained model is evaluated on the test set through statistical indicators. This evaluation confirms the reliability of the proposed method across diverse educational contexts.

Algorithm 1 Pseudo-code of the proposed method for educational attainment prediction

Require: Raw dataset with k features and N records
Ensure: Trained model and evaluation metrics
1: **Preprocessing**
2: Handle noisy and missing values
3: Drop irrelevant columns
4: Eliminate rows with missing values
5: Apply Min–Max Normalization
6: Split the cleaned dataset into train (n_{train}), test (n_{test}), and validation (n_{val}) sets based on ratios (P_{train}, P_{test}, P_{val})
7: **Model Construction**
8: *Deep Neural Network Model:*
9: `Dense(128)`
10: `BatchNormalization`
11: `Dropout(0.4)`
12: `Dense(64)`
13: `Dropout(0.3)`
14: `Dense(32)`
15: Adam optimizer
16: *Gradient Boosting Model*
17: *Meta-Model Blending* (Linear Regression)
18: **Evaluate Model**
19: Compute performance metrics on the test set

4 Results

The models are developed in Python. Libraries for ML and preprocessing are Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow, and Keras. All experiments are run on the Apple M2 chip system with 8-core CPU, 10-core GPU, 8GB unified memory (RAM), and macOS Sonoma version 14.4.1. This system allows the learning components to be effectively trained and evaluated. Table 3 presents the full configuration of dataset parameters, DNN and GB hyperparameters, and evaluation metrics used in the proposed educational attainment prediction method. These settings are selected to balance predictive performance, computational efficiency, and model generalization. The dataset includes 1100 samples and 22 features. It is divided into training (91%), testing (5%), and validation (4%) sets. This split provides enough data for model learning while independent subsets are kept for unbiased evaluation. For the DNN model, the selected architecture employs Dense layers of sizes 128, 64 and 32 with ReLU activation. This architecture allows for the hierarchical learning of nonlinear relations between socioeconomic and regional variables. Dropout rates of 0.4 and 0.3 reduces

overfitting. The Adam optimiser with $LR = 0.001$ is used which results into adaptive gradient updates and fast convergence. The model is trained for 200 epochs with a batch size of 16, which lets it to refine its weights and achieve better accuracy. The loss function used is MSE and the evaluation metric used is MAE. These choices help to make accurate predictions in regression. For the GB model, 500 estimators and a learning rate of $LR = 0.05$ provide a suitable balance between accuracy and overfitting control. A maximum depth of 4 and a minimum samples split of 2 set a moderate level of model complexity. This setting allows the model to capture important data patterns without excess noise. A subsample value of 1 ensures that all data points contribute to each boosting iteration. The value `random_state = 42` ensures reproducibility. The model performance is evaluated with R^2, RMSE, and MAE, which measure prediction accuracy and error in the proposed method. These hyperparameter choices support the robustness of the model and its strong generalization across different educational and regional contexts.

Table 3: Summary of model hyperparameters, dataset partitioning, and evaluation metrics used in our model

Hyperparameter/Parameter		**Value/Method**
N		1100
k		22
P_{train}		91%
P_{test}		5%
P_{val}		4%
n_{train}		1003
n_{test}		55
n_{val}		42
DNN	Dense	[128, 64, 32]
	Dense Activation	ReLU
	Dropout	[0.4, 0.3]
	Optimizer	Adam (default LR=0.001)
	Loss Function	MSE
	Epochs	200
	Batch Size	16
	Metrics	MAE
GB	$n\_estimators$	500
	LR	0.05
	max_depth	4
	Subsample	1
	$min\_samples\_split$	2
	$random\_state$	42
Evaluation Metrics		R^2, RMSE, MAE

4.1 Dataset

The dataset used in this study is obtained from Kaggle. It contains detailed information on the educational attainment of young people across towns in England [26]. The original dataset includes 1104 records. Each record represents one town. The dataset also contains 31 features that cover demographic, socioeconomic, regional, and educational aspects. The variables include categorical identifiers, such as town codes, and quantitative measures, such as educational attainment stages, employment rates, and income indicators. The target variable is the "Education Score". It is a continuous value that represents the composite educational achievement of each town. Before model construction, data preparation is performed. Columns that are irrelevant to the prediction task are removed, such as unique identifiers and columns with

missing values. Rows with missing data are also removed to ensure consistency for ML implementation. A detailed feature analysis is presented in Table 4. The table demonstrates the procedure of feature selection and data cleansing. Out of the total 31 features, nine columns are dropped. The columns that have been dropped contain those variables with a lot of null values and those that are not very informative such as `TOWN11CD` and `TOWN11NM`. Variables of importance are included in the data such as population, region, coastal region, job density, income, education, and important landmarks of ages 18 and 19. They represent diversity from both socio-economic and regional aspects. The rows having missing values have been dropped too. Hence, the number of complete observations is reduced to 1100 along with a total of 22 features, excluding the target variable. Such preparations improve data integrity and diversity and also help the ML models to learn important relations that can be used in educational attainment prediction.

Table 4: A detailed analysis of the features

#	**Name**	**Feature/ Target**	**Data Type**	**Non-Null Count**	**Unique (Objects)**	**Used/ Drop**
1	TOWN11CD	Feature	Object	1100	1100	Drop
2	TOWN11NM	Feature	Object	1104	1104	Drop
3	POPULATION_2011	Feature	float64	1100	–	Used
4	SIZE_FLAG	Feature	Object	1104	8	Used
5	RGN11NM	Feature	Object	1102	9	Used
6	COASTAL	Feature	Object	1100	2	Used
7	COASTAL DETAILED	Feature	Object	1100	7	Used
8	TTWA11CD	Feature	Object	1100	154	Used
9	TTWA11NM	Feature	Object	1100	154	Used
10	TTWA CLASSIFICATION	Feature	Object	1100	4	Used
11	JOB DENSITY FLAG	Feature	Object	1100	3	Used
12	INCOME FLAG	Feature	Object	1100	4	Used
13	UNIVERSITY FLAG	Feature	Object	1100	2	Used
14	Level4Qual_residents35-64_2011	Feature	Object	1100	3	Used
15	KS4_2012-2013_counts	Feature	int64	1104	–	Used
16	Key Stage 2 attainment (2007–2008)	Feature	float64	1104	–	Used
17	Key Stage 4 attainment (2012–2013)	Feature	float64	1104	–	Used
18	Level 2 at age 18	Feature	float64	1104	–	Used
19	Level 3 at age 18	Feature	float64	1104	–	Used
20	Activity at age 19: full-time higher education	Feature	float64	1103	–	Used
21	Activity at age 19: sustained further education	Feature	float64	1046	–	Drop
22	Activity at age 19: apprenticeships	Feature	float64	909	–	Drop
23	Activity at age 19: employment (earnings > £0)	Feature	float64	1103	–	Used
24	Activity at age 19: employment (earnings > £10,000)	Feature	float64	1077	–	Used
25	Activity at age 19: out-of-work	Feature	float64	642	–	Drop
26	Highest qualification by age 22: less than level 1	Feature	float64	245	–	Drop
27	Highest qualification by age 22: level 1 to level 2	Feature	float64	1056	–	Drop
28	Highest qualification by age 22: level 3 to level 5	Feature	float64	1103	–	Used
29	Highest qualification by age 22: level 6 or above	Feature	float64	873	–	Drop
30	Highest qualification by age 22: average score	Feature	float64	1104	–	Used
31	**Education Score**	**Target**	**float64**	**1104**	–	**Used**

A Kernel Density Estimate (KDE) plot is a non-parametric method used to visualize the probability density function of a continuous variable. It provides a smooth representation of the data distribution [27]. The KDE plot of the Education Score is presented in Figure 2. The curve appears approximately symmetric and bell-shaped around zero, which indicates a near-normal distribution. Most towns are concentrated around the average education score, while the density gradually decreases toward both lower and higher values. This balanced

distribution indicates that extreme values are limited in the dataset. As a result, the prediction model is less likely to be influenced by outliers and can generalize more effectively across the observations.

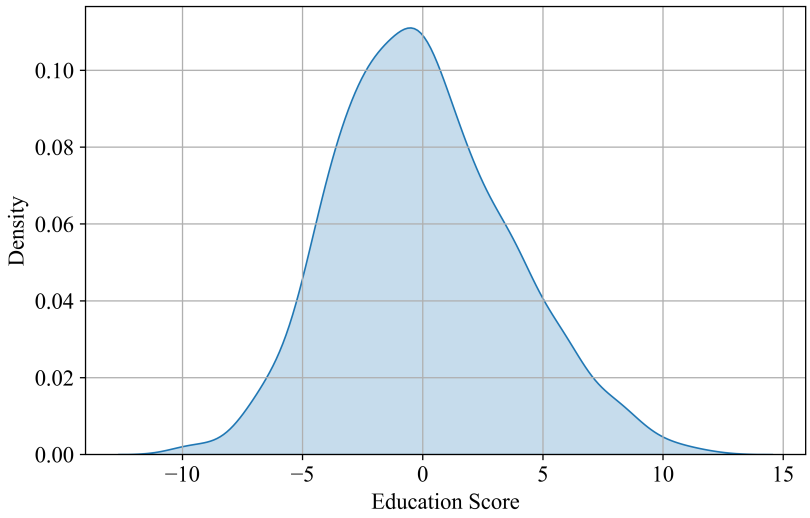

Figure 2: KDE plot of Education Score

A correlation matrix presents the pairwise correlation coefficients among variables, allowing the relationships between different features to be examined [28], [29,30]. The correlation matrix of the dataset is shown in Figure 3. A correlation matrix presents the pairwise correlation coefficients among variables, allowing the relationships between different features to be examined [28], [29,30]. The correlation matrix of the dataset is shown in Figure 3. In contrast, geographic indicators such as "COASTAL" and "SIZE_FLAG" show weak correlations with educational scores. This observation suggests that these factors have limited direct influence on education outcomes. The correlation analysis supports the identification of influential predictors and emphasizes the importance of examining multiple socioeconomic and educational variables in prediction tasks. In contrast, geographic indicators such as "COASTAL" and "SIZE_FLAG" show weak correlations with educational scores. This observation suggests that these factors have limited direct influence on education outcomes. The correlation analysis supports the identification of influential predictors and emphasizes the importance of examining multiple socioeconomic and educational variables in prediction tasks.

Figure 4 presents the correlation values between each feature and the Education Score, which is the target variable. Several features show strong positive associations with educational outcomes. Variables such as "Level 3 at age 18," "Highest level qualification achieved by age 22: average score," and "Key Stage 4 attainment (2012 to 2013)" have correlation values above 0.9. This result indicates that these academic milestones strongly contribute to the final Education Score. "Activity at age 19: full-time higher education" and earlier educational indicators such as "Key Stage 2 attainment" also display strong positive relationships, demonstrating that continued academic progress across different stages contributes to higher educational attainment. Moderate positive correlations are observed for socioeconomic indicators including "INCOME FLAG" and "COASTAL." In contrast, variables such as population, job density, and several employment activities at age 19 show weak or negative relationships with the target variable. The feature "Activity at age 19: employment with earnings above £10,000" shows the strongest negative correlation (-0.29). This relationship suggests that

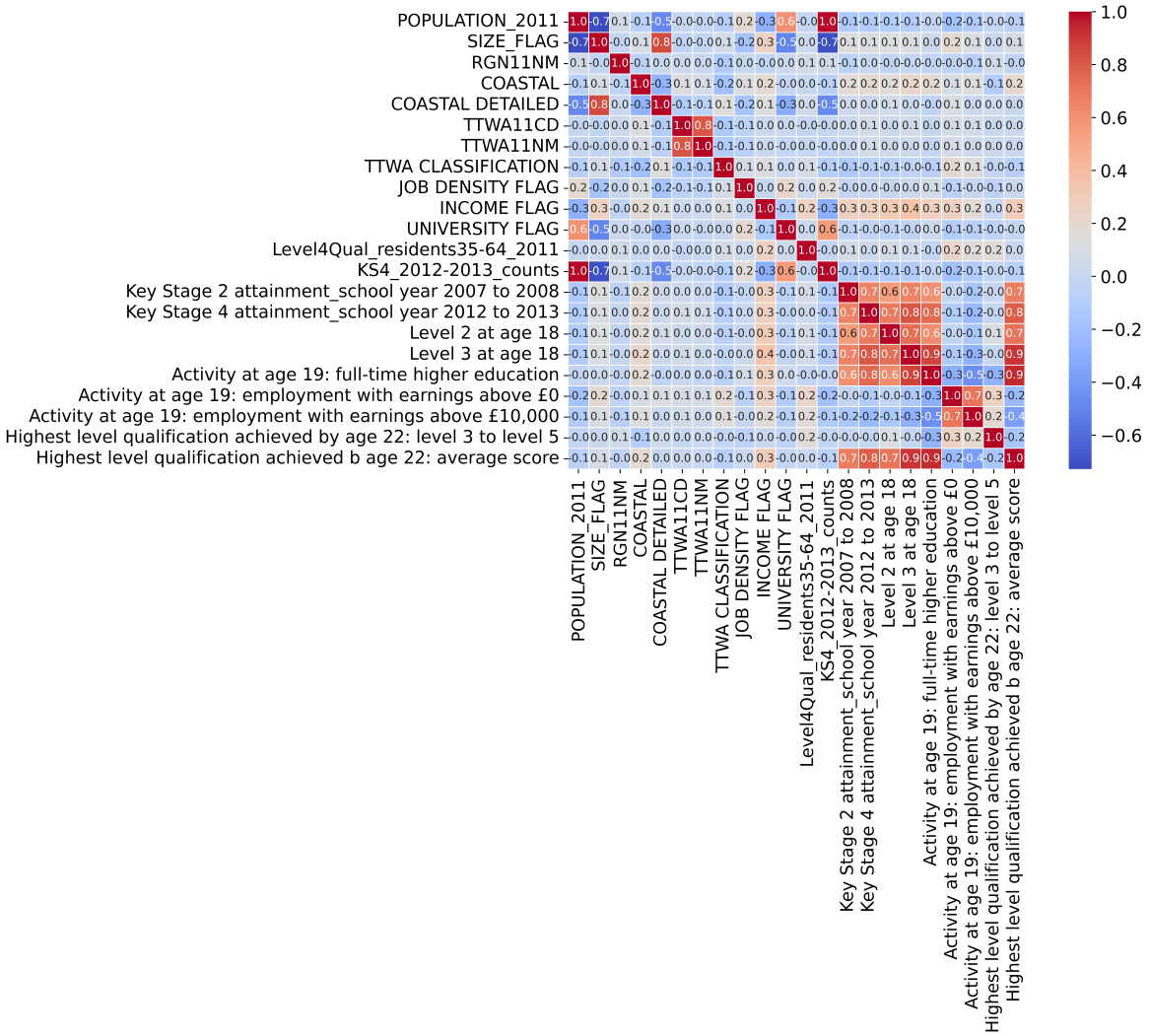

Figure 3: Correlation matrix

entering employment with higher earnings at an early stage may reduce the likelihood of continued academic progression. These findings assist in understanding the factors associated with educational outcomes and support informed feature selection for predictive modeling.

4.2 Experimental Results and Model Evaluation

Before model evaluation, the dataset is divided into training, validation, and test subsets to support reliable model development and evaluation. As presented in Table 5, the dataset contains 1100 samples. Among these, 1003 samples (91%) are assigned to the training set, allowing the model to learn patterns from a large portion of the data. The validation set contains 42 samples (4%), which are used for hyperparameter adjustment and monitoring of overfitting during training. The remaining 55 samples form the test set (5%), which is used for the final performance assessment of the trained model.

No missing values are present in any subset after preprocessing. This condition ensures that both model training and evaluation are performed on complete data records, which im-

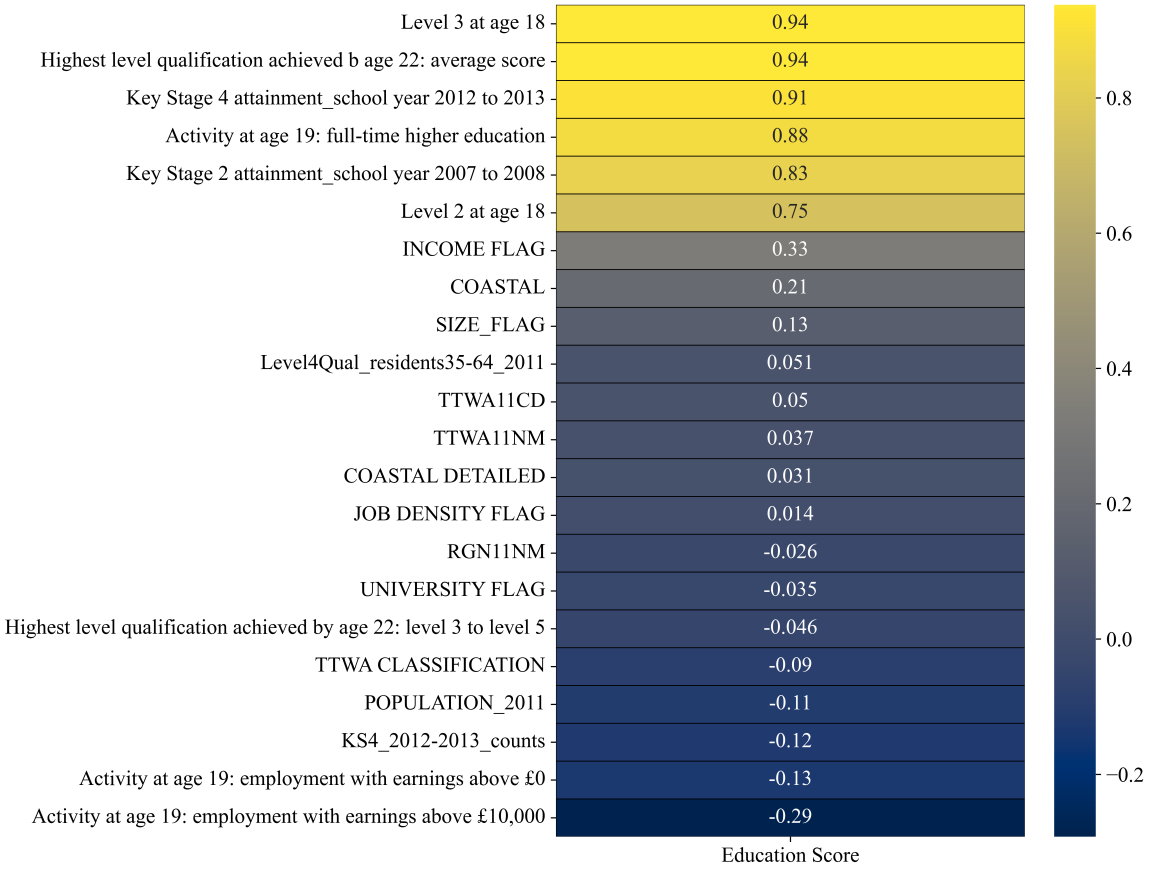

Figure 4: Correlation coefficients between each feature and the Education Score (target feature)

proves the reliability of the experimental analysis. The 91/5/4 split provides a large training portion for model learning while maintaining separate validation and test sets. These independent subsets support unbiased performance evaluation and help assess the generalization ability of the proposed prediction method on unseen data.

Table 5: Dataset partitioning into training, validation, and test sets

Subset	**Total Number of Data**	**Missing (%)**
Training	1003	0%
Validation	42	0%
Test	55	0%
Total	**1100**	**0%**

MSE measures the average squared differences between predicted and actual values, while MAE calculates the average absolute differences between predicted and actual values [29]. Figure 5 illustrates MSE (Figure 5(a)) and MAE (Figure 5(b)) curves for our DNN across 200 epochs, evaluating the learning dynamics and potential overfitting of our model. Both MSE and MAE curves exhibit a rapid initial decline, followed by low values for both training and validation sets, without significant divergence between them. This close alignment indicates that the model generalizes well to unseen data and does not suffer from overfitting, as the validation loss does not increase relative to the training loss. Low error metrics across epochs also

highlight the robustness of DNN in capturing the underlying patterns of the dataset. Thus, the results demonstrate the strength of our DNN model in achieving accurate educational attainment predictions while maintaining generalizability and stability.

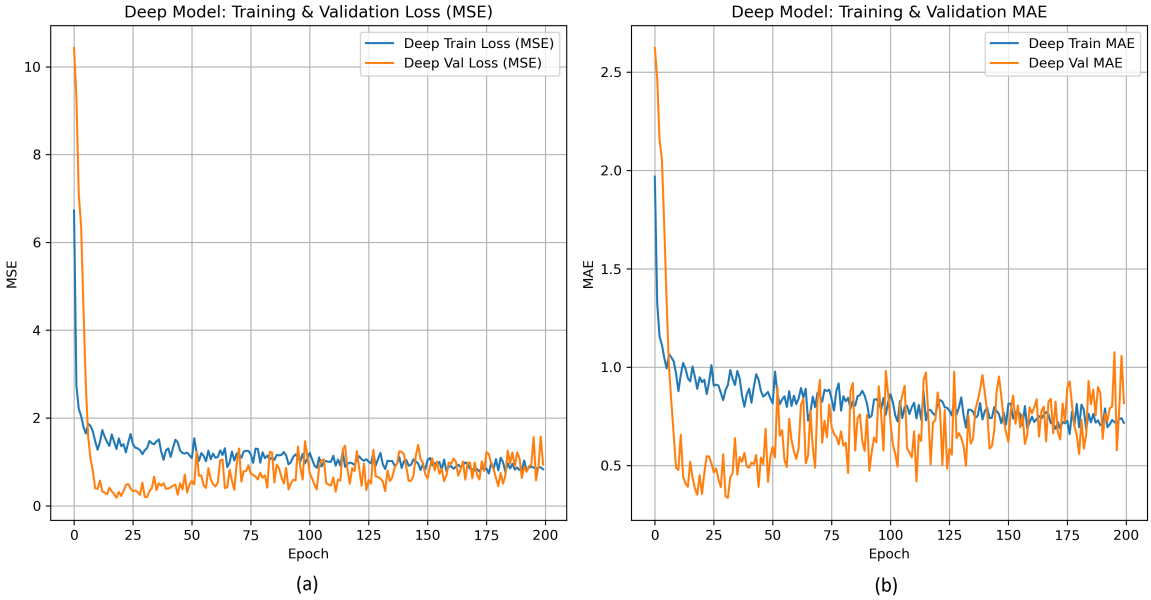

Figure 5: MSE and MAE curves for DNN: (a) MSE, (b) MAE

Figure 6 displays the training and validation loss curves for our GB model in terms of MSE (Figure 6(a)) and MAE (Figure 6(b)) across boosting stages. It demonstrates a rapid decrease in error as the number of boosting stages increases, with the validation and training losses tracking each other and plateauing at low values without any signs of divergence or overfitting. It indicates that our GB model is effective at capturing the underlying patterns in the data while maintaining strong generalization to unseen ones, reflecting both the robustness and predictive strength of our method. Indeed, the ability of our model to maintain such low and parallel error curves throughout the training process highlights its stability and suitability for predicting educational attainment outcomes.

Figure 7 includes four scatter plots comparing the true versus predicted values for different models (GB [6], Deep Learning [14], Simple Blending, and Meta-Model Blending (our proposed method)), with the red dashed line representing the ideal line where predicted values match the true values. From the plots, it is evident that all models exhibit strong predictive performance, as the points cluster around the ideal red line. However, our model and the simple blending model demonstrate the tightest alignment, indicating superior accuracy and reduced error compared to the individual models. The GB and Deep Learning models also perform well, but they show more deviation from the ideal line for extreme values. Overall, the results confirm that the ensemble (blending) approaches, particularly our meta-model model, yield the most reliable predictions by capturing the underlying patterns and minimizing discrepancies between the true and predicted educational attainment scores.

R^2, i.e., coefficient of determination measures the proportion of variance in the target variable that is explained by the model, while RMSE, i.e., Root Mean Squared Error quantifies the average magnitude of prediction errors in the same units as the target variable [29, 30]. As demonstrated in Table 6, our model outperforms the other approaches across all key metrics.

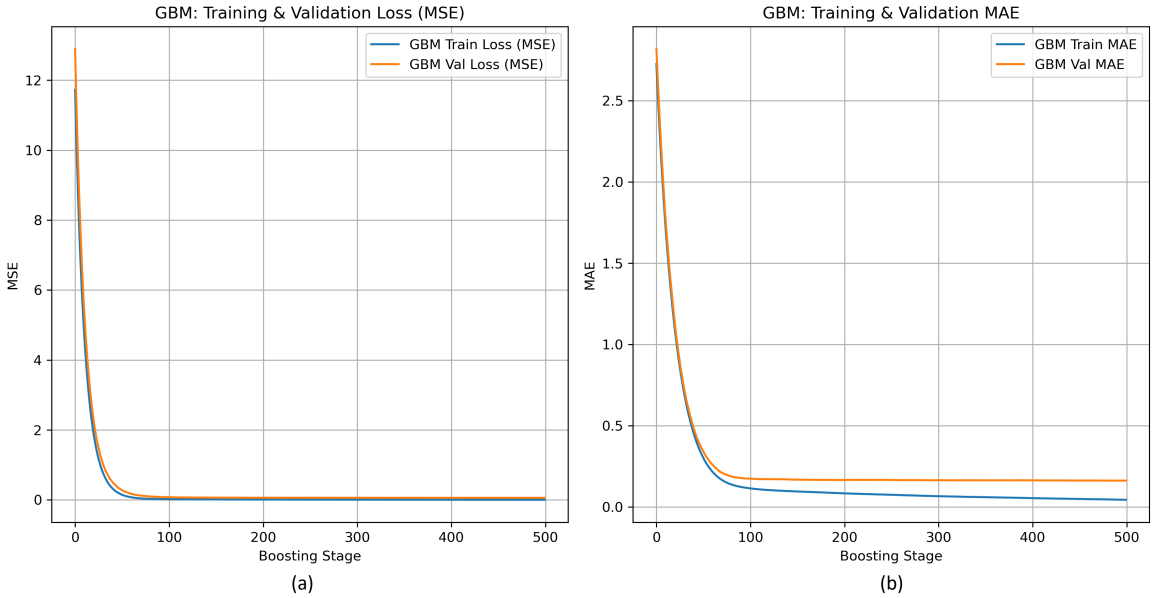

Figure 6: MSE and MAE curves for GB: (a) MSE, (b) MAE

In more detail, our ensemble model achieves superior accuracy and error reduction with an R^2 of 0.9975, an MAE of 0.1582, and an RMSE of 0.2058, indicating its strong ability to explain all the variance in educational attainment scores and to deliver accurate predictions. In contrast, the simple blending and the deep learning models show lower R^2 values and higher error rates, highlighting their limitations in capturing complex relationships within the dataset. On the other hand, when the individual models are compared, the Gradient Boosting model delivred better performance than the deep learning model. It achieves an R^2 value of 0.9957 and an RMSE of 0.2442. However, its performance remains slightly lower than that of the proposed method. The deep learning model and the simple blending approach produce higher error values and lower R^2 scores. We can note that a single deep learning architecture or a basic blending strategy does not fully capture the relationships present in the data. In contrast, the proposed method achieves superior predictive accuracy. This outcome demonstrates that combining multiple model architectures can improve prediction performance by capturing both nonlinear patterns and structured relationships within the dataset.

Table 6: Comparison of the performance of different predictive approaches

Approach	R^2	**MAE**	**RMSE**
Proposed Model (Meta-Model Blending)	**0.997498**	**0.158231**	**0.205836**
Simple Blending	0.992172	0.256678	0.329775
Deep Learning [14]	0.957035	0.612534	0.772589
Gradient Boosting [6]	0.995706	0.161122	0.244232

Experimental findings confirm high predictive power of the proposed model. High accuracy and low estimation error are provided. Hybrid ensemble uses deep neural networks, gradient boosting, and meta modeling approach. Deep neural networks learn complicated nonlinear dependencies. Gradient boosting learns structured tabular dependencies. Prediction stability and accuracy are enhanced by means of model integration. Sociodemographic

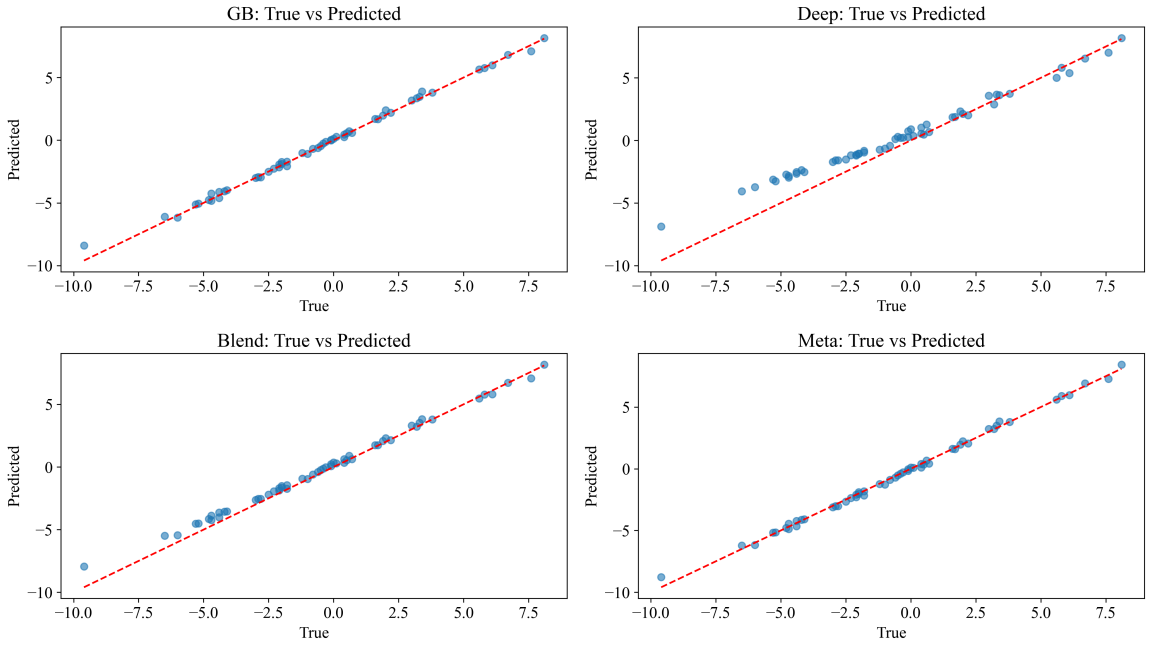

Figure 7: True versus predicted values for different models

and geographical features make it possible to widen the scope of the analysis from purely academic factors. Wider impact of external factors on education is therefore considered. We prove that the proposed model can be applied for educational planning and resources allocation in different regions. However, good prediction is conditional upon consistent input data. Regional differences in collecting data might influence generalizability of the model. Future research should consider temporal datasets and interpretable machine learning models.

5 Conclusion

This study proposes a hybrid ensemble model for educational attainment prediction. Deep neural networks, gradient boosting, and meta model blending are integrated. Socioeconomic and regional variables are analyzed with academic indicators. Prediction accuracy and model robustness are improved through complementary learning strategies. Experimental results confirm reliable and stable predictive performance. Broader contextual factors are also represented in educational outcome estimation. We show that the proposed model supports educational planning and resource allocation through data driven analysis. However, prediction quality depends on reliable and consistent input data. Future research should include longitudinal and behavioral datasets. Interpretable machine learning methods should also be incorporated. Greater transparency, adaptability, and responsible educational decision support are therefore achieved.

Acknowledgments

The authors declare a minor use of AI for language editing. The authors declare a record of prior joint coauthorship with an editorial board member of the journal.

References

[1] P. Vestfal and L. Seduikyte, "Systematic Review of Factors Influencing Students' Performance in Educational Buildings: Focus on LCA, IoT, and BIM," *Buildings*, vol. 14, no. 7, p. 2007, 2024, doi: 10.3390/buildings14072007.

[2] E. Alyahyan and D. Düştegör, "Predicting academic success in higher education: literature review and best practices," *Int. J. Educ. Technol. Higher Educ.*, vol. 17, no. 1, p. 3, 2020, doi: 10.1186/s41239-020-0177-7.

[3] S. Batool, J. Rashid, M. W. Nisar, J. Kim, H.-Y. Kwon, and A. Hussain, "Educational data mining to predict students' academic performance: A survey study," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 905–971, 2023, doi: 10.1007/s10639-022-11152-y.

[4] E. Alhazmi and A. Sheneamer, "Early Predicting of Students Performance in Higher Education," *IEEE Access*, vol. 11, pp. 27579–27589, 2023, doi: 10.1109/ACCESS.2023.3250702.

[5] S. Najjar-Ghabel, S. Yousefi, and E. H. Ismael, "Enhancing Melanoma Detection through Multi-Modal Learning: Integration of Dermoscopic Images and Clinical Data," in *Proc. 2025 IEEE 7th Symp. Comput. & Inform. (ISCI)*, 2025, pp. 70–75, doi: 10.1109/ISCI65687.2025.11167604.

[6] A. Villar and C. R. V. de Andrade, "Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study," *Discover Artif. Intell.*, vol. 4, no. 1, p. 2, 2024, doi: 10.1007/s44163-023-00079-z.

[7] S. Yousefi, S. Najjar-Ghabel, R. Danehchin, S. S. Band, C.-C. Hsu, and A. Mosavi, "Automatic melanoma detection using discrete cosine transform features and metadata on dermoscopic images," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 36, no. 2, p. 101944, 2024, doi: 10.1016/j.jksuci.2024.101944.

[8] N. Binesh and M. Ghatee, "Distance-aware optimization model for influential nodes identification in social networks with independent cascade diffusion," *Inf. Sci.*, vol. 581, pp. 88–105, 2021, doi: 10.1016/j.ins.2021.09.017.

[9] S. Yousefi, S. Najjar-Ghabel, and Z. S. Owaid, "A Supervised Learning-based Framework for Failure Detection in Toy Cars Using Acoustic Signal Analysis," in *Proc. 2025 IEEE 7th Symp. Comput. & Inform. (ISCI)*, 2025, pp. 76–81, doi: 10.1109/ISCI65687.2025.11167791.

[10] B. Yousefimehr, M. Ghatee, and R. Razavi-Far, "Multi-stage self-distillation for class-imbalanced anomaly detection," *IEEE Trans. Emerg. Topics Comput. Intell.*, 2026, publisher: IEEE.

[11] S. Saronian, B. Yousefimehr, M. Ghatee, and M. M. Bejani, "An Ensemble of Deep Clustering Models With Autoencoders to Mine Travel Patterns From Smart Card Data," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 12, pp. 20960–20969, 2024, doi: 10.1109/TITS.2024.3475295.

[12] B. Yousefimehr and M. Ghatee, "A systematic survey and empirical comparison of hybrid methods for imbalanced fraud detection: Combining resampling and machine learning," *AUT J. Math. Comput.*, vol. 7, no. 1, pp. 85–116, 2026, doi: 10.22060/ajmc.2025.24642.1446.

[13] A. Kordbagheri, M. Kordbagheri, N. Tayim, A. Fakhrou, and M. Davoudi, "Using advanced machine learning algorithms to predict academic major completion: A cross-sectional study," *Comput. Biol. Med.*, vol. 184, p. 109372, 2025, doi: 10.1016/j.compbiomed.2024.109372.

[14] R. Ghorbani and R. Ghousi, "Comparing Different Resampling Methods in Predicting Students' Performance Using Machine Learning Techniques," *IEEE Access*, vol. 8, pp. 67899–67911, 2020, doi: 10.1109/ACCESS.2020.2986809.

[15] A. R. Kenari, A. Montazerolghaem, Z. Zojaji, M. Ghatee, B. Yousefimehr, A. Rahmani *et al.*, "Isfahan Artificial Intelligence Event 2023: Reflux Detection Competition," *J. Med. Signals Sens.*, vol. 15, no. 2, art. no. 6, 2025, doi: 10.4103/jmss.jmss_46_24.

[16] H. Zeineddine, U. Braendle, and A. Farah, "Enhancing prediction of student success: Automated machine learning approach," *Comput. Electr. Eng.*, vol. 89, p. 106903, 2021, doi: 10.1016/j.compeleceng.2020.106903.

[17] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, p. 11, 2022, doi: 10.1186/s40561-022-00192-z.

[18] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data using various machine learning algorithms," *Mater. Today: Proc.*, vol. 80, pp. 3782–3785, 2023, doi: 10.1016/j.matpr.2021.07.382.

[19] N. A. Kuadey, C. Ankora, F. Tahiru, L. Bensah, C. C. M. Agbesi, and S. O. Bolatimi, "Using machine learning algorithms to examine the impact of technostress creators on student learning burnout and perceived academic performance," *Int. J. Inf. Technol.*, vol. 16, no. 4, pp. 2467–2482, 2024, doi: 10.1007/s41870-023-01655-3.

[20] K. Ravikumar, K. Saravanakumar, and A. Viswanathan, "Preemptive Min Max Optimal Cost Based Scheduling for Improving the Load Balancing in Virtualized Cloud Environment," *SN Comput. Sci.*, vol. 5, no. 6, p. 754, 2024, doi: 10.1007/s42979-024-03085-9.

[21] S. Najjar-Ghabel, S. Yousefi, and P. Habibi, "Comparative Analysis and Practical Implementation of Machine Learning Algorithms for Phishing Website Detection," in *Proc. 2024 9th Int. Conf. Comput. Sci. Eng. (UBMK)*, 2024, pp. 1–6, doi: 10.1109/UBMK63289.2024.10773479.

[22] S. B. Nadkarni, G. S. Vijay, and R. C. Kamath, "Comparative Study of Random Forest and Gradient Boosting Algorithms to Predict Airfoil Self-Noise," in *Proc. RAiSE-2023*, MDPI, 2023, p. 24, doi: 10.3390/engproc2023059024.

[23] S. Najjar-Ghabel, S. Yousefi, and H. Karimipour, "Blockchain Applications in the Industrial Internet of Things," in *AI-Enabled Threat Detection and Security Analysis for Industrial IoT*, Springer International Publishing, Cham, 2021, pp. 41–76, doi: 10.1007/978-3-030-76613-9_4.

[24] B. Yousefimehr, M. Ghatee, and A. Heydari, "Improving ADHD Detection with Cost-Sensitive LightGBM," in *Proc. 2024 14th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, 2024, pp. 109–113, doi: 10.1109/ICCKE65377.2024.10874782.

[25] X. Zhang, M. Usman, A. ur R. Irshad, M. Rashid, and A. Khattak, "Investigating Spatial Effects through Machine Learning and Leveraging Explainable AI for Child Malnutrition in Pakistan," *ISPRS Int. J. Geo-Inf.*, vol. 13, no. 9, p. 330, 2024, doi: 10.3390/ijgi13090330.

[26] S. Sachidanandan, "Educational Attainment," Kaggle. [Online]. Available: `https://www.kaggle.com/datasets/samithsachidanandan/educational-attainment-of-young-in-english-towns`

[27] C. A. Staunton, A. Kårström, H. Kock, M. S. Laaksonen, and G. Björklund, "Kernel Density Estimation: a novel tool for visualising training intensity distribution in biathlon," *Front. Sports Active Living*, vol. 7, 2025, doi: 10.3389/fspor.2025.1546909.

[28] S. Yousefi, S. Najjar-Ghabel, and H. Shafaei, "A Technical Analysis and Practical Implementation of Machine Learning Algorithms for Predicting Survival in Breast Cancer Patients," in *Proc. 2024 9th Int. Conf. Comput. Sci. Eng. (UBMK)*, 2024, pp. 1168–1173, doi: 10.1109/UBMK63289.2024.10773521.

[29] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, p. e623, 2021, doi: 10.7717/peerj-cs.623.

[30] M. Kozhanov, C. Mako, V. Poser, G. Eigner, and A. Mosavi, "Knowledge Augmented Generation for Curriculum Planning," *Eurasian J. Math. Comput. Appl.*, vol. 14, no. 1, pp. 17–34, 2026, doi: 10.32523/2306-6172-2026-14-1-17-34.

Samad Najjar-Ghabel*,
Department of Computer Engineering,
University of Mohaghegh Ardabili,
Ardabil, Iran;
e-mail: `S.Najjar@uma.ac.ir`.

Behnam Yousefimehr,
Department of Mathematics and Computer Science,
Amirkabir University of Technology,
Tehran, Iran;
e-mail: `behnam.y2010@aut.ac.ir`.

Ali Daneshpour,
Department of Mathematics and Computer Science,
Amirkabir University of Technology,
Tehran, Iran;
e-mail: `alidaneshpour@aut.ac.ir`.

Murat Kozhanov,
Doctoral School of Applied Informatics and Applied Mathematics,
Budapest, Hungary;
e-mail: `murat.kozhanov@uni-obuda.hu`.

Azamat Amirov,
Abylkas Saginov Karaganda Technical University,
Karaganda, Kazakhstan.

Received 04.11.2025 , Accepted 17.04.2026 Available online 30.06.2026.