# DEEP LEARNING-BASED EDUCATION DECISION SUPPORT SYSTEM FOR STUDENT E-LEARNING PERFORMANCE PREDICTION

SUDHA PRATHYUSHA JAKKALADIKI,* MARTINA JANEČKOVÁ,† JAN KRUNČÍK,‡ FILIP MALÝ§ AND TEREZA OTČENÁŠKOVÁ¶

**Abstract.** Information Technology (IT) and its advancements change the education environment. Conventional classroom education has been transformed into a modernized form. Education field decision-makers are always searching for new technologies that provide fast solutions to support Education Decision Support Systems (EDSS). There is a significant need for an effective decision support system to utilize student data which helps the university in making the right decisions. The Electronic learning system (e-learning) provides a live forum for faculties and students to connect with learning portals and virtually execute educational activities. Even though these modern approaches support the education system, active student participation still needs to be improved. Moreover, accurately measuring student performance using collected attributes remains difficult for parents and teachers. Therefore, this paper seeks to understand and predict student performance using effective data processing and a deep learning-based decision model. The implementation of EDSS starts with data preprocessing, Extraction-Transformation-Load (ETL), a data mart area to store the extracted data with Online Analytical Processing (OLAP) processing, and decision-making using Deep Graph Convolutional Neural Network (DGCNN). The statistical evaluation is based on the student dataset from the Kaggle repository. The analyzed results depict that the proposed EDSS model on an independent data mart with efficient decision support and OLAP provides a better platform to make academic decisions and help educators to make necessary decisions notified to the students.

**Key words:** data mart, decision support system, deep learning, e-learning, ETL, OLAP

**1. Introduction.** A decision support system is challenging in present-day applications like education, recommendation systems, etc. They provide a framework for decision-making that enables users to analyze and evaluate information from multiple sources, models, and databases to help them make informed decisions. After the pandemic, most learning platforms have developed online services to provide practical learning opportunities to students. Online courses are offered through e-learning (electronic or online learning) to interested learners who can learn from their particular places. The shift from traditional education to e-learning has resulted in several prospects, objectives, and difficulties for educational institutions, instructors, and students. Additionally, the quality of content delivered to students continually improves due to various technological advancements in the animation industry [2]. When student populations increase beyond the e-learning platform, multiple challenges exist in measuring student performance, content quality, etc. Major challenges pointed out by researchers are communication between teachers and learners, online assessment of students, technical issues, student impacts on materials and studies, etc..

A learning management system requires practical tools in an e-learning platform to provide a strong communication link between teachers, students, and administrative staff. E-learning tools help to digitize the classroom environment with good communication. Some applications processed through the e-learning platform are digital class teaching, screen sharing, attendance, assessment, and reports [3]. Recent developments in learning systems like high-speed internet and improved networking structures help increase students' e-learning

---

*Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czech Republic (sudha.jakkaladiki@uhk.cz)

†Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czech Republic (martina.janeckova@uhk.cz)

‡Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czech Republic (jan.kruncik@uhk.cz)

§Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czech Republic (email: filip.maly@uhk.cz)

¶Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czech Republic (tereza.otcenaskova@uhk.cz)

usage. Student performance can be measured with multi-perspective features from the management system. There are several approaches to measuring student performance and reporting instructions daily in online systems. A decision support system helps mine the relevant data features and provides decisions in time. Deep learning, Artificial Intelligence (AI), and machine learning techniques help to achieve the greater performance of the system.

**Background:** In the field of education, advancements in information technology (IT) have transformed conventional classroom education into a modernized form of education. Educational decision-makers are constantly searching for new technologies. The problem being addressed in this research is the need for an effective decision support system to utilize student data and help universities make informed decisions regarding student performance. While electronic learning systems (e-learning) have provided a live forum for faculties and students to connect and virtually execute educational activities, there is still room for improvement in terms of active student participation and accurately measuring student performance using collected attributes.

The significance of this research lies in its proposed solution, which is a deep learning-based decision model that utilizes effective data processing and decision-making techniques. The implementation of this Education Decision Support System (EDSS) involves data pre-processing, Extraction-Transformation-Load (ETL), a data mart area to store the extracted data with Online Analytical Processing (OLAP), and decision-making using Deep Graph Convolutional Neural Network (DGCNN).

Primary education to higher education faces significant challenges when the database is huge. Prediction and analysis of big data need more advanced deep learning algorithms in a learning management system. This research investigates a deep learning-based decision support system for analyzing students' performance using an e-learning system. Also, the authors aim to improve the e-learning system by giving the best performance to students. The significant contribution of this paper follows:

1. This paper developed an efficient EDSS for an e-learning environment that can improve the performance of the detection process of student performance.
2. This study consists of four phases: data preparation, ETL, data mart, and decision-making. The first process is data preparation, used to transfer the required dataset from the sources. Next, the extracted data is transformed into the required format using a series of transformation models.
3. The staged data is processed using the ETL transformation process and loaded into the data mart for further processing.
4. OLAP operations are executed in the data mart to make the retrieval process smoother, and then the decision-making process is performed using the proposed DGCNN model.
5. The efficiency of the proposed deep learning based EDSS system with ETL is experimented with and compared to conventional methods using the dataset from Kaggle.

This article is further organized as follows: Section 2 provides details of previous work on similar research, Section 3 presents the proposed work, including algorithms, figures, etc., Section 4 contains information on the results and outcomes, and Section 5 consists of the conclusion of the research.

**2. Related Work.** A methodology by Rujirayanyong and Shi [25] was developed to support universities in making admissions decisions using data mining techniques. The study's objective is to predict applicants' academic performance based on pre-admission criteria, such as high school grade average, Scholastic Achievement Admission Test score, and General Aptitude Test score [24]. The study used a dataset of 2,039 students enrolled in a Saudi public university's Computer Science and Information College from 2016 to 2019. The study results indicate that the applicants' early university performance can be predicted before admission based on specific pre-admission criteria. The Scholastic Achievement Admission Test score is the pre-admission criterion most accurately indicates future student performance. The study also found that the Artificial Neural Network technique has an accuracy rate above 79%, making it superior to other classification techniques [14]. Based on these findings, the authors suggest assigning more weight to the Scholastic Achievement Admission Test score in admissions systems.

The study [21, 26] addresses the challenges of constructing reliable Recommender Systems based on sparse, few in quantity, imbalanced, and anonymized data that may have been stored under poor conditions. This is an essential issue in Data Mining, where the quality and quantity of data can significantly affect the accuracy and reliability of predictive models. The study's success in creating a Recommender System using a real-world

dataset from a public Spanish university demonstrates the approach's applicability in natural environments.

The proposed research addresses the limitations of current online exam systems, which rely on traditional machine-learning methods that require handcrafted features and cannot learn hierarchical representations of objects from the data [12]. This limitation affects the efficiency and effectiveness of such systems in detecting cheating. The proposed approach utilizes deep learning models that can automatically extract useful features from visual images and speech using convolutional neural networks and statistical methods. The article's author proposes an e-learning system based on performance analysis using ensemble machine learning. Among other machine learning algorithms, random forest achieves the highest prediction rate in measuring student performance.

In Article [1], Al-Qahtani and Alanzi used student advice and monitoring as a dataset to analyze student performance levels in an e-learning platform. Others have used educational outcomes [6] and student achievement records [7] as sources of performance analysis. In recent years, there has been significant growth in the use of technology in various aspects of life, including education and healthcare. With the development of intelligent decision support systems and AI-based tools, decision-making has become more efficient and effective. Similarly, the use of technology in online examination systems has led to improved monitoring and cheating detection. This literature review aims to provide a summary of recent research in areas such as intelligent decision-making in healthcare, recommender systems for higher education, deep learning-based cheating detection approaches in online examination systems, AI-based learning style prediction in primary education, teaching machine learning in K-12 classrooms, and the use of AI-based online proctoring systems.

The article [18] proposes an interrelated decision-making model for an intelligent decision-support system in healthcare. The model combines fuzzy logic and neural network techniques to improve the accuracy of the decision-making process in healthcare. The work depicted in [5, 16] develops a recommender system to support higher education students in making enrollment decisions. The system uses a combination of collaborative filtering and content-based filtering techniques to provide personalized recommendations to students. Kadoura and Gumaei propose a deep learning-based cheating detection approach for online examination systems [12, 15]. The approach uses a convolutional neural network (CNN) to analyze students' behavior during an exam and detect any suspicious activity.

The study described in [23] aimed to develop an AI-based learning style prediction model for primary education students in online learning. The model utilizes a hybrid deep neural network (DNN) to predict students' learning styles and provide personalized recommendations to improve learning outcomes. In an article [27], Tedre et al. discussed the pedagogical and technological trajectories for teaching machine learning in K-12 classrooms. They emphasized the need for a comprehensive approach that includes curriculum design, teacher training, and the development of appropriate tools and resources. In their study [1], Al-Qahtani and Alanzi conducted a longitudinal cohort study to compare the predictive values of admission criteria for academic achievement among undergraduate students of health and non-health science professions. The study found that high school GPA was both groups' most significant predictor of academic achievement.

Gopane and Kotecha proposed an AI-based online exam monitoring system to improve cheating detection during exams [6]. The system utilizes a combination of image processing and machine learning techniques to detect any suspicious activity that may occur. In Article [7], Gumaei et al. developed a deep learning-based driver distraction identification framework that utilizes a CNN to detect various driving distractions. This framework can help prevent accidents by alerting drivers who may be distracted. Finally, Gumaei et al. in [8] proposed a decision-level fusion method for predicting the health status of Covid-19 patients. The technique combines multiple machine-learning models to improve the accuracy of health prediction for patients with Covid-19.

Jalali and Noorbehbahani evaluated various online proctoring tools and discussed their advantages and limitations [9]. They highlighted the need for a balance between privacy concerns and effective monitoring. Joshi et al. [10] proposed an automatic method for cheating detection in online exams by processing the students' webcam images. The method uses image processing and machine learning techniques to detect suspicious activity during an exam. Kadam et al. [11] reviewed the literature on explainability in multimodal deep neural networks and discussed various approaches for improving the interpretability of deep neural networks. Kaddoura et al. [13] proposed a method for detecting and localizing multiple image splicing using MobileNet

Fig. 3.1: Overview of proposed Education Decision Support System (EDSS) Source: Authors' elaboration

v1. The technique uses deep learning for accurate detection. Furthermore, AI-based decision-making systems for exams and university student performance are discussed with intelligent technologies and advanced tools [19, 20, 22, 28].

Some limitations of the above research are, Firstly, the use of AI-based systems for monitoring exams and identifying cheating has raised concerns about privacy invasion and ethical issues. It is essential to balance monitoring students' activities and respecting their privacy rights. Secondly, the proposed AI-based driver distraction identification framework has limitations in accurately detecting all types of driving distractions. For instance, it may not identify distractions that do not involve physical movements, such as daydreaming or inattention. Thirdly, the proposed decision-level fusion method for predicting Covid-19 patients' health status relies on multiple machine learning models, which may lead to increased complexity and computational costs. Additionally, the accuracy of predictions may depend on the quality and quantity of data used to train the models. Fourthly, while AI-based methods for detecting and localizing image splicing have shown promising results, these methods may face challenges in identifying more advanced forms of image manipulation, such as deepfakes.

Finally, while AI-based decision-making systems have shown a potential to improve university student performance, there is a need to develop explainable AI models that can provide interpretable results and enhance user trust. Moreover, there is a need for rigorous evaluation of these systems' effectiveness and ethical implications before implementing them on a large scale.

**3. Proposed E-learning Decision Support Methodology.** The overview of the proposed EDSS is depicted in Fig 3.1, which consists of four major components: data preparation, ETL, data mart area with OLAP processing, and decision-making using deep learning. The data preparation phase involves collecting and extracting basic data sources. The extracted input data is then transformed using various techniques to balance the data for further processing. The complete processed input data is loaded and staged in the data mart area for storage. The OLAP processing methods generate various access methods for the data stored in the data mart for analysis. Finally, the proposed EDSS model measures and predicts student performance.

1. **Dataset Preparation and Description**
   During the data preparation phase, various data formats, including structured, unstructured, and semi-structured, are collected and stored in a data warehouse for further processing and decision-making. The ETL process is applied to enhance the decision-making process of the stored data. The primary objective of this study is to evaluate the student's performance and participation during e-learning sessions. The proposed model is evaluated using the Kalboard 360 dataset from the Kaggle repository [2].

2. **ETL process**
   The execution of the ETL process will reduce the data warehouse development time, storage, and cost.

It involves various tasks to manipulate the data to obtain the final results. Below, only two first parts of the process are described, and the Load process follows in subsection 3.1.

- **Extraction**

  Extraction is the process of reading and understanding the data from various sources and extracting the necessary parts for further processing. The dataset has been extracted using experience API, consisting of 17 features and 480 instances of multiple characteristics. The attributes are classified into integer and categorical types. It comprises 175 female and 305 male students from various countries [2]. The numerous qualities called 'Visited resources', 'Discussions', 'Announcements', and 'Raise hands' were extracted. These attributes help measure the student's functional performance and participation during the class. The constant response to the questions and discussion forum enhances the prediction process. Based on the features, this data set is used for classification and prediction purposes and divided into three sections such as demographic, academic, and behavioral. The features such as raising a hand during the lecture, learning satisfaction, response to the survey, and using resources demonstrate learner engagement. Apart from this, the features such as gender, semester, nationality, discussion, and so on were selected. The classification of this data set is based on the label column of three categories: 'low', 'medium', and 'high' to indicate the student performance.

- **Transform**

  Once the data is extracted and necessary data has been selected, a series of processes are executed to convert the data meaningfully. It consists of five processes: data cleaning, balancing, normalization, and one hot encoding.

  (a) **Data cleaning:** Removes incorrect, missing, and duplicate data from the dataset. This step, also called data scrubbing, ensures the classification model's accuracy, effectiveness, reliability, and efficiency. The blank and white spaces of the labels are avoided, and the redundant features are removed.

  (b) **Data balancing:** Data imbalance is a significant problem in machine learning (ML). The data is rebalanced to improve the validation performance. The rebalance has two types: over-sampling (OS) and under-sampling (US). The US reduces the population of the most represented class, and OS increases the minimum defined people of the course. This paper used the best US model, the random US, and the best OS model, SMOTE (Synthetic Minority Over-Sampling Technique), for balancing. The SMOTE model combines the majority of the data with its closest neighbors.

  (c) **Data normalization:** The classifier's performance also depends on features with average or scaled values in the range [3]. This is not necessary to have all the features evenly distributed. This normalization process ensures the ML performance using the methods such as Yeo Johnson transformation, scaling, and one hot encoding.

      i. **Yeo Johnson:** It is from the Box-cox transformation and not restricted to positive values. It is mathematically stated as in Eqn 3.1

$$X_i^\rho = \begin{cases} \left( \frac{(X_i+1)^\delta - 1}{\delta} \right) & if \ \delta \neq 0, \delta \geq 0 \\ \log{(X_i + 1)} & if \ \delta = 0, \delta \geq 0 \\ \left( \frac{-[(X_i+1)^{2-\delta} - 1])}{(2-\delta)} \right) & if \ \delta \neq 0, \ \delta < 0 \\ -\log{(-X_i + 1)} & if \ \delta = 2, \delta < 0 \end{cases} \quad (3.1)$$

where $X$ is the input data and $\delta$ is the real number. If $\delta = 1$ means it denotes the identity transformation [17]. The Yeo-Johnson transformation is a mathematical formula that transforms a given set of input data, X, into a new set of transformed data, $X^\rho$, where $\rho$ is a parameter that can be varied. The Yeo-Johnson transformation allows for data that includes negative values, a limitation of the Box-Cox transformation. The parameter $\delta$ can take on any real value, including 0 and negative values, which makes it more flexible than the Box-Cox transformation. When $\delta$ is equal to 1, the transformation reduces to

the identity transformation, which means that the transformed data is the same as the original data.

Overall, the Yeo-Johnson transformation is a valuable tool in data analysis. It can transform skewed data into a more normal distribution, making it easier to analyze and model.

ii. **Scaling**: The imbalance feature scales will degrade the classifier performance. It is necessary to normalize it to an acceptable scale. The continuous data values are changed into the range 0 to 1 using the Min-Max method [4] denoted in Eqn 3.2

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \ \forall x \in X \tag{3.2}$$

where $x_{norm}$ is the normalized result, $x_{min}$ is the value of 0, and $x_{max}$ is the value of 1 and the remaining values are in the range of 0 to 1. This process makes the features to have the same range of values which reduces the training and testing time and overcome the fast convergence rate. This will increase the prediction system's reliability. The log transformation is applied before the normalization to handle the high variance features.

iii. **One Hot Encoding**: The discrete variables were handled using the label encoders, which encode the categorical data values into an integer. The behavioral attributes of the students are changed into integer values.

**3.1. Data Mart (DM) and OLAP.** The data mart consists of staging and data mart schema tables. The staging area contains the staging table for the transformation process and the final table called the fact table. The data mart schema used in this study is the star schema, where the records are connected to the five-dimension tables using data from the staging table. The staging table is the intermediate area where the storage occurs between the source and the data warehouse. It serves as the temporary storage area that is deleted after being uploaded into the repository. It is used in processes such as data source preparation, cleaning, conversion, extraction, loading, and updating the quality assurance [25], and is referred to by the ETL process. The staging area is prepared with the intention of performing OLAP queries. The data warehouse is implemented using three schemas: star, fact constellation, and snowflake. The star schema is famous for its simplicity, and it consists of a central table (fact table) that concatenates the keys to the dimensions and measurements [24]. Fig 3.2 illustrates the star schema for the Educational DM.

The DM schema is built using SSMS and it consists of five dimensions such as degree, participation, information, enrollment, behavior and degree. The fact tables have five keys which are used to concatenate the dimensional tables. The dimensions are used as the key factor for OLAP process and its responses. The OLAP operations including Roll_ Up, Drill_ Down, Dice, Pivot, and Slice were executed on these dimension tables. Fast responses for the OLAP queries are executed using the star schema. The overall process is stated in Fig 3.3.

**3.2. Decision-Making using DGCNN.** The issue with traditional CNNs is that if the input is not prominent, the CNN pooling layers lose the feature data, which reduces the network's learning ability. Therefore, conventional CNNs do not provide accurate predictions in some instances. Deep graph CNNs are deep learning models that provide an accurate classification for data mining, image processing, and cloud environment problems. They are represented as graphs and process complex data that is not processed by conventional CNNs. The nodes of the graph are connected hierarchically in a spatial area called convolution learning. The graph convolution is constructed with the spectral domain and inverse Fourier transform. The spatial domain is represented by an adjacency matrix $A \in R \wedge (N \times N)$, and the spatial relationship between the central node and the neighbor node is determined using spectral similarity. If the node $X_i$ is adjacent to $X_j$ then its edges have the weight as $X_{ij} \in A$. The node adjacency is represented in Eqn 3.3.

$$X_{ij} = \begin{cases} 0 & \text{if } X_i \text{ is not adjacent to } X_j \\ \exp\left(-\frac{||X_i - X_j||}{\sigma_W^2}\right) & \text{otherwise} \end{cases} \tag{3.3}$$

where $\sigma_W^2$ is the weight range. The DGCNN is declared as G (X, A) and the classification result is Y of the graph nodes. This process has three sub-phases such as: The characteristic of each node is extracted. From
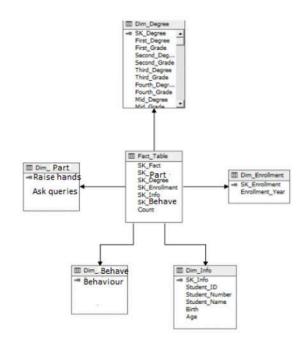
Fig. 3.2: Educational DM (Star Schema) Source: Authors' elaboration

the local structure of its neighbor, these details are collected. To increase the ability of the network, nonlinear transformation is carried from the previous data node. The graph convolution process aggregates the neighbor node data into a new form which is represented in Eqn 3.4

$$Y = g\left(\text{f}, A\right) \tag{3.4}$$

$$f^{(i+1)} = \sigma\left(L_g f^{(i)} w^{(i)}\right), L_g = \tilde{D}^{-\frac{1}{2}} \tilde{A}^{-\frac{1}{2}} \tag{3.5}$$

Where, $f^{(i)}$ is the extracted feature of the layer i from ETL process, $\sigma$ is the activation function, $w$ *is the* learnable parameter, $L_g$ is the Laplacian matrix, $\tilde{D}$ is thedegree matrix and $\tilde{A}$ denotes A+I and I denotes Identity matrix. During the training process, the network loss is rescued by adjusting the learnable parameter w using Eqn 3.6

$$O = \sum_{i=1}^{n} loss(g\left(X_i^g, A\right), g_i) \tag{3.6}$$

where $g_i$ is the sample ground truth and n is the number of features. Once the training is over, the feature from the graph and the matrix A has been used to classify the nodes. The DGCNN structure is shown in Fig 3.4.

The steps to execute the DGCNN are stated as follows:

Input: Dataset D with features f

Output: Classification results of student performance

Step 1: Collect features from the ETL process

Step 2: Construct the graph nodes called V and feature f are used to initialize the graph nodes as $f^G = [f_1^G, f_2^G, \dots f_k^G]^T \in R^{k \times D}$ .

Step 3: Construct the graph edges called E using the first order adjacency relationship by executing Eqn 3.3.

Step 4: Compute the edges w and form Eqn 3.4

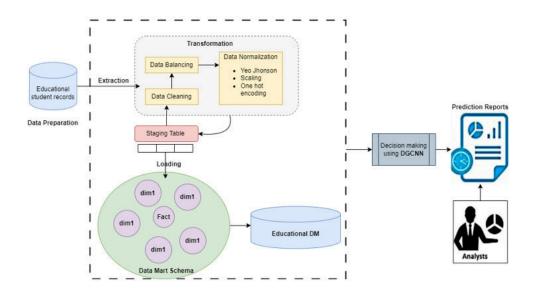Step 5: Return the prediction results.

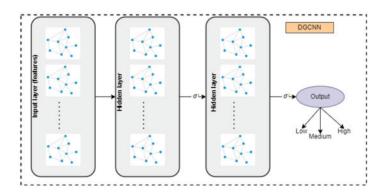Fig. 3.3: Detailed process of proposing EDSS Source: Authors' elaboration



Fig. 3.4: Structure of DGCNN (Source: Authors' elaboration

**4. Results and Discussions.** The proposed EDSS model has been experimented using the Kallboard dataset. This model is applied using 10 fold cross validation to obtain the optimal accuracy of the learning model. It consists of two parts such as training and testing. In the cross validation, the data set is partitioned into k folds where each iterations, one subset from testing or training has been used for evaluation. The proposed model is implemented using Python 3.4. from Scikit library using the evaluation metrics such as accuracy, recall, precision and F1 measure.

**4.1. Evaluation Metrics.** The data mining models performance is measured using the factors such as True positive, false positive, true negative and false negative. Based on these factors, the model accuracy, recall, precision and F1score are measured using the following equations.

$$Accuracy = \frac{(Tp + Tn)}{(Tp + Tn + Fp + Fn)} \tag{4.1}$$

$$Recall = \frac{Tp}{(Tp + Fn)} \tag{4.2}$$

Table 4.1: Proposed DL based EDSS model performance

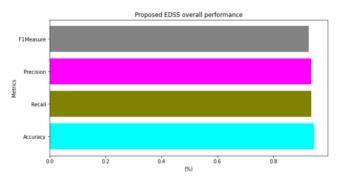| Class | Accuracy | Recall | Precision | F1measure |
|---|---|---|---|---|
| Low | 0.9348 | 0.9234 | 0.9317 | 0.9145 |
| Medium | 0.9472 | 0.9345 | 0.9215 | 0.9176 |
| High | 0.9572 | 0.9478 | 0.9522 | 0.9473 |



Fig. 4.1: Proposed model overall performance Source: Authors' elaboration

$$Precision = \frac{Tp}{(Tp + Fp)} \tag{4.3}$$

$$F1measure = \frac{2 * recall * precision}{(recall + precision)} \tag{4.4}$$

**4.2. Result Discussions.** The performance of the proposed model to classify the student performance under three categories such as low, medium and high is shown in Table 4.1. The overall performance of the proposed model is illustrated in Fig 4.1. As an average of all the classes, the proposed model secured the improved Accuracy of 94.6%, Recall of 93.52%, Precision of 93.51% and F1measure of 92.6%.

The efficiency of the proposed student performance prediction system is compared with the existing decision support system such as C4.5 [14], Artificial neural network [21] and Ensemble ML models [26]. The comparative results are shown in Table 4.2.

From this table, the average performance of proposed and existing approaches are computed and illustrated in Fig 4.2. Compared to the existing approaches, the proposed model secured the improved accuracy of 95%, recall of 94%, precision of 94% and f1measure of 93%. Whereas, the existing DSS model called C4.5 secured the accuracy, recall, precision and f1measure of 88%, 87%, 88% and 87% respectively. The Artificial Neural Networks model secured the accuracy, recall, precision and f1measure of 83%, 82%, 82% and 82% respectively and the Ensemble based DSS secured the accuracy, recall, precision and f1measure of 83%, 82%, 81% and 82% sequentially. The comparison in terms of ROC (Receiver Operator Characteristics) is shown in Fig 4.3 where it proves the efficiency of the proposed model on classifying the student performance as low, medium and high with improved ROC of 0.9 than other approaches. Hence, the proposed model is superior to other approaches on the prediction of student performance.

The average detection time of the student performance is compared, and it is illustrated in Fig 4.4. Due to the implementation of efficient ETL and OLAP operations, the data is stored in the data mart efficiently. The retrieval process is smoother because of this which will reduce the detection rate of the proposed model. Compared to the existing approaches, the proposed model secured reduced detection rate which improves the system integrity, and the required results are displayed in fraction of seconds. The proposed model obtained the

Table 4.2: Comparative analysis of proposed and existing DSS

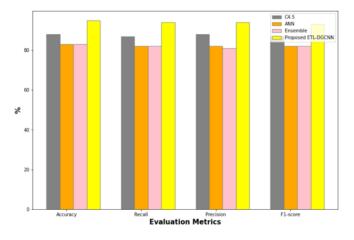| Models | Class labels | Accuracy | Recall | Precision | F1 Measure |
|--------|-------------|----------|--------|-----------|------------|
| C4.5 | Low | 0.8837 | 0.8729 | 0.8812 | 0.8777 |
| | Medium | 0.8729 | 0.8673 | 0.8714 | 0.8713 |
| | High | 0.8782 | 0.8711 | 0.8725 | 0.8673 |
| ANN | Low | 0.8627 | 0.8536 | 0.8614 | 0.8625 |
| | Medium | 0.7672 | 0.7727 | 0.7525 | 0.7618 |
| | High | 0.8534 | 0.8253 | 0.8531 | 0.8242 |
| Ensemble | Low | 0.8684 | 0.8534 | 0.8656 | 0.8684 |
| | Medium | 0.7836 | 0.7727 | 0.8095 | 0.7906 |
| | High | 0.8351 | 0.8250 | 0.7674 | 0.7951 |
| Proposed ETL+DGCNN | Low | 0.9348 | 0.9234 | 0.9317 | 0.9145 |
| | Medium | 0.9472 | 0.9345 | 0.9215 | 0.9176 |
| | High | 0.9572 | 0.9478 | 0.9522 | 0.9473 |



Fig. 4.2: Average performance of proposed and existing EDSS Source: Authors' elaboration
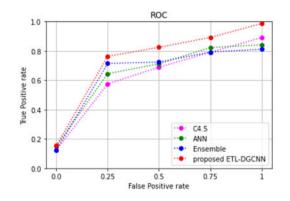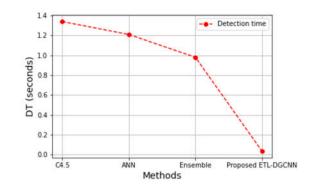


Fig. 4.3: ROC comparison Source: Authors' elaboration

Fig. 4.4: Detection Time comparison Source: Authors' elaboration

reduced detection time of 0.034seconds than other approaches such as C4.5 (1.34s), ANN (1.21s) and Ensemble MLmodel (0.98s) respectively.

The main aim of this study is to develop a new model that can be useful to predict the student performance of e-learning from online series data. The deep learning and efficient ETL process can effectively predict the performance of the students in a minimum detection time which will be helpful for their growth. Compared to the existing DSS models, the proposed model obtained the improved performance from 8 to 9.2% of accuracy. Overall, the proposed model results show the accuracy, reliability and efficiency for predicting the performance of the students in the e-learning environment.

**Limitation:** DGCNNs are sensitive to the quality and structure of the input data, which means they may not perform well on datasets with missing or noisy data, or on datasets with complex and irregular structures. DGCNNs are best suited for problems that can be represented as graphs, which limits their applicability to certain domains and types of data.

**5. Conclusion.** This paper proposes the development of an efficient Education Decision Support System for e-learning environments to improve the detection process of student performance. The study comprises four phases: data preparation, ETL, data mart, and decision making. The first phase involves collecting the required dataset from sources and transforming the extracted data into the required format using transformation models. The staged data is then processed using the ETL-transformation process and loaded into the data mart for further processing. In the data mart, OLAP operations are executed to make the retrieval process smoother, and then the decision-making process is executed using the proposed Deep Graph Convolutional Neural Network model. The efficiency of the proposed EDSS system with ETL is experimentally compared with conventional systems using a dataset from Kaggle. The proposed model achieved improved accuracy of 95% with a reduced detection time of 0.034 seconds compared to existing approaches. The timely decisions of the student performance detection can improve their performance in future courses. Hence, the proposed model is suited for the educational industry to make timely decisions with accurate prediction of student performance. In the future, the proposed model will be experimented with an increased number of records in the real-time environment using cloud and Internet of Things.

### REFERENCES

[1] M. F. AL-QAHTANI AND T. M. ALANZI, *Comparisons of the predictive values of admission criteria for academic achievement among undergraduate students of health and non-health science professions: a longitudinal cohort study*, Psychology Research and Behavior Management, (2018), pp. 1–6.

[2] A. A. ALSULAMI, A. S. A.-M. AL-GHAMDI, AND M. RAGAB, *Enhancement of e-learning student's performance based on ensemble techniques*, Electronics, 12 (2023), p. 1508.

[3] J. BROWNLEE, *How to use power transforms for machine learning*, Machine Learning Mastery [Internet], (2020).

[4] P. DEVAN AND N. KHARE, *An efficient xgboost–dnn-based classification model for network intrusion detection system*, Neural Computing and Applications, 32 (2020), pp. 12499–12514.

[5]   A. J. Fernández-García, R. Rodríguez-Echeverría, J. C. Preciado, J. M. C. Manzano, and F. Sánchez-Figueroa, *Creating a recommender system to support higher education students in the subject enrollment decision*, IEEE Access, 8 (2020), pp. 189069–189088.

[6]   S. Gopane and R. Kotecha, *Enhancing monitoring in online exams using artificial intelligence*, in Proceedings of International Conference on Data Science and Applications: ICDSA 2021, Volume 2, Springer, 2022, pp. 183–193.

[7]   A. Gumaei, M. Al-Rakhami, M. M. Hassan, A. Alamri, M. Alhussein, M. A. Razzaque, and G. Fortino, *A deep learning-based driver distraction identification framework over edge cloud*, Neural Computing and Applications, (2020), pp. 1–16.

[8]   A. Gumaei, W. N. Ismail, M. R. Hassan, M. M. Hassan, E. Mohamed, A. Alelaiwi, and G. Fortino, *A decision-level fusion method for covid-19 patient health prediction*, Big Data Research, 27 (2022), p. 100287.

[9]   K. Jalali and F. Noorbehbahani, *An automatic method for cheating detection in online exams by processing the student's webcam images*, in 3rd conference on electrical and computer engineering technology (E-Tech , Tehran, Iran, 2017, pp. 1–6.

[10]  G. Joshi, R. Walambe, and K. Kotecha, *A review on explainability in multimodal deep neural nets*, IEEE Access, 9 (2021), pp. 59800–59821.

[11]  K. Kadam, S. Ahirrao, K. Kotecha, and S. Sahu, *Detection and localization of multiple image splicing using mobilenet v1*, IEEE Access, 9 (2021), pp. 162499–162519.

[12]  S. Kaddoura and A. Gumaei, *Towards effective and efficient online exam systems using deep learning-based cheating detection approach*, Intelligent Systems with Applications, 16 (2022), p. 200153.

[13]  S. Kaddoura, D. E. Popescu, and J. D. Hemanth, *A systematic review on machine learning models for online learning and examination systems*, PeerJ Computer Science, 8 (2022), p. e986.

[14]  S. Latif, F. XianWen, and L.-L. Wang, *Intelligent decision support system approach for predicting the performance of students based on three-level machine learning technique*, Journal of Intelligent Systems, 30 (2021), pp. 739–749.

[15]  S. Maheswaran, S. Sathesh, P. Priyadharshini, and B. Vivek, *Identification of artificially ripened fruits using smart phones*, in 2017 international conference on intelligent computing and control (I2C2), IEEE, 2017, pp. 1–6.

[16]  S. Maheswaran, S. Sathesh, G. Saran, and B. Vivek, *Automated coconut tree climber*, in 2017 International Conference on Intelligent Computing and Control (I2C2), IEEE, 2017, pp. 1–6.

[17]  A. Mahfouz, A. Abuhussein, D. Venugopal, and S. Shiva, *Ensemble classifiers for network intrusion detection using a novel network attack dataset*, Future Internet, 12 (2020), p. 180.

[18]  N. B. Mahiddin, Z. A. Othman, A. A. Bakar, and N. A. A. Rahim, *An interrelated decision-making model for an intelligent decision support system in healthcare*, IEEE Access, 10 (2022), pp. 31660–31676.

[19]  N. Malhotra, R. Suri, P. Verma, and R. Kumar, *Smart artificial intelligence based online proctoring system*, in 2022 IEEE Delhi section conference (DELCON), IEEE, 2022, pp. 1–5.

[20]  M. M. Masud, K. Hayawi, S. S. Mathew, T. Michael, and M. El Barachi, *Smart online exam proctoring assist for cheating detection*, in International conference on advanced data mining and applications, Springer, 2022, pp. 118–132.

[21]  H. A. Mengash, *Using data mining techniques to predict student performance to support decision making in university admission systems*, Ieee Access, 8 (2020), pp. 55462–55470.

[22]  A. S. Milone, A. M. Cortese, R. L. Balestrieri, and A. L. Pittenger, *The impact of proctored online exams on the educational experience*, Currents in Pharmacy Teaching and Learning, 9 (2017), pp. 108–114.

[23]  B. Pardamean, T. Suparyanto, T. W. Cenggoro, D. Sudigyo, and A. Anugrahana, *Ai-based learning style prediction in online learning for primary education*, IEEE Access, 10 (2022), pp. 35725–35735.

[24]  P. Ponniah, *Data warehousing fundamentals for IT professionals*, John Wiley & Sons, 2011.

[25]  T. Rujirayanyong and J. J. Shi, *A project-oriented data warehouse for construction*, Automation in Construction, 15 (2006), pp. 800–807.

[26]  F. Saleem, Z. Ullah, B. Fakieh, and F. Kateb, *Intelligent decision support system for predicting student's e-learning performance using ensemble machine learning*, Mathematics, 9 (2021), p. 2078.

[27]  M. Tedre, T. Toivonen, J. Kahila, H. Vartiainen, T. Valtonen, I. Jormanainen, and A. Pears, *Teaching machine learning in k–12 classroom: Pedagogical and technological trajectories for artificial intelligence education*, IEEE access, 9 (2021), pp. 110558–110572.

[28]  B. Vivek, S. Maheswaran, P. Keerthana, S. Sathesh, S. Bringeraj, R. A. Sri, and S. A. Sulthana, *Low-cost raspberry pi oscilloscope*, in 2018 international conference on intelligent computing and communication for smart World (I2C2SW), IEEE, 2018, pp. 386–390.