International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN: 2147-6799 www.ijisae.org Original Research Paper

Employability Prediction of Lateral Entry Engineering Students: A Deep Learning Based Inductive Reasoning and Interpretable Framework

Mousoomi Bora[1*], Rupam Baruah[2]

Submitted: 05/02/2024 **Revised :** 12/03/2024 **Accepted:** 19/03/2024

Abstract Numerous technology advancements have contributed significantly to the nation's economic stimulation. Proactive steps have been taken by the technical education industry to boost the employment rate of young engineers. Higher Education Institutions (HEI) have given equal opportunities to the Lateral Entry (LE) engineering students to compete and survive in the competitive job market as Regular Engineering (RE) students. A Deep Learning (DL) based inductive reasoning framework has been developed to forecast the employability rate of LE students after their graduation. A detailed time-series analysis is conducted to examine the trend of academic performance in terms of employability. An interpretable module is integrated with the model to interpret the major contributing features in the overall prediction process. The proposed framework outperforms the existing models in terms of Mean Square Error (MSE) and Mean Absolute Error (MAE).

Keywords *- Time series dataset, Gated Recurrent Neural Network, Attention mechanism, Interpretability framework, Lambda layer, Performance metrics*

I. Introduction

The application made by Lateral Entry (LE) engineering students has been granted by the All-India Council for Technical Education (AICTE) enabling them to be admitted to any engineering branch of an Indian technical universities. According to the AICTE chairman, having an additional degree increases their employability rate and facilitate career changes. In today's multimodal educational system, LE students are anticipated to get substantial support, according to the National Education Policy (NEP) 2020. The goal of the action taken by AICTE affiliated Indian technical colleges and universities is to provide LE students with the same employment opportunities as Regular Engineering (RE) students **[1].** Based on their preferences, country's technical institutions accept LE admission through counselling based on their final percentage marks in diploma program. In order to facilitate direct admission to Bachelor of Technology (BTech) programs, AICTE established an entrance examination that HEIs must administer as an eligibility test for BTech or BE degree. The aspiring students would be chosen on the basis of their merit. The authors - G. S. Nandakumar, L. M. Palanivelu have stated that due to lesser effort on theoretical understanding in polytechnic education, LE students struggle to keep up with their peers. For students from rural polytechnics, where they are technically enrolled in an English-medium program but are mostly taught in vernacular, the language utilized in the classroom also poses a hurdle. They find it challenging to interact socially with the usual stream pupils in their new educational institution. **[2].** In this study, we have examined the academic performance of LE students in terms of their employability at the end of undergraduate studies in Computer Science and Engineering (CSE). A Deep Learning (DL) based time-series forecasting model is developed with an objective of employability prediction of LE students. For this study, a dataset was prepared based on the academic information of 334 students from the Department of CSE of an anonymous technical university from India. As a method of Inductive Reasoning (IR), time series analysis has been carried out to forecast students' employability starting from the 3rd semester till the end of 8th semester in a sequential fashion. The rest of the paper is organized as follows. The relevant works in the field of student's performance prediction are discussed in the Section II. The data set is explained in the Section III. The Section IV explains methodology and proposed framework. The results and discussions are presented in the Section V. The Section VI offers the conclusion and some directions for further research.

II. Related Works

A great deal of research in the arena of Educational Data Mining (EDM) is conducted for employability prediction of students. The researchers have developed various prediction frameworks based on Machine learning (ML) and Deep learning (DL) to carry out their research in the area related to student's performance prediction. However, a research gap is observed where few researchers have explored methodology to predict the performance of LE students to best of our knowledge. R. Costa-Mendes, T.

[1]Department of Computer Science and Engineering,*
Assam Science and Technology University, Guwahati, India.
mousoomi@gmail.com
[2] Department of Computer Science and Engineering,
Assam Science and Technology University, Guwahati, India
rupam.baruah.jec@gmail.com

***Corresponding Author:** Mousoomi Bora*
**Department of Computer Science and Engineering, Assam Science and Technology University, Guwahati, India.*
mousoomi@gmail.com

Oliveira, M. Castelli, C. J. Frederico proposed a multilinear regression model associated with Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and Extreme Gradient Boosting (ExGB) to analyze the performance of undergraduate students. In comparison to the conventional multiple linear regression model, ML techniques manifestly possess a greater degree of predictive power **[3]**. The recent advancements in forecasting students' achievement in postsecondary education were studied by S. Alturki, I. Hulpuş and H. Stuckenschmidt **[4].** The authors have examined popular methods for forecasting academic achievement while highlighting the advantages of EDM. Political Fractional Competitive MultiVerse Optimization (PFCMVO) is used by A. J Baruah, S. Baruah to simplify the development of a Deep Neuro fuzzy network (DNFN) that is used to measure the performance of students in the Spark environment **[5].** M. Bora and R. Baruah developed a semi-supervised-sequence-prediction (SSSP) ML framework to predict students' academic achievement towards the completion of four-year degree course in outcome-based education system **[6]**. M. Elmasry developed a hybrid model that combines Decision Tree (DT) and Deep Neural Network (DNN) to predict students' final grades **[7].** N. Kumar, H. Krishna, S. Shubham, P.P. Rout developed a text classification framework to help classify students' queries **[8].** W. Tiancheng used Envelopment Analysis (DEA) to predict student success in online classrooms **[9].** B. Sekeroglu, K. Dimililer, K.Tuncal developed a model to predict student performance that relies on Support Vector Regression (SVR), Backpropagation (BP), and Long-Short Term Memory (LSTM) **[10].** A. Çetinkaya and O. Baykan have used ANN to evaluate the problem solving and programming skills of high school students **[11].** M. Tadayon and G.J. Pottie created a time series model using the Hidden Markov (HM) model to evaluate students' conceptual comprehension **[12].** T.T. Dien, T. -N. Nguyen, H. Nguyen, S. Luu devised an approach based on CNN and LSTM to forecast students' achievement in the forthcoming semesters based on their course accomplishment from the previous semesters **[13].** A hybrid 2D CNN model was developed by S. Poudyal, M. Mohammadi-Aragh, J. Ball to predict student's academic achievement **[14].** An attention-based GRU model based was proposed by S. Jung, J. Moon, S. Park, E. Hwang for automatic load forecasting **[15].** A CNN-LSTM-based model was proposed by Aljaloud et al. to estimate students' learning outcomes **[16].** K. Biswas, S. Kumar, S. Banerjee, A. Pandey have suggested a set of custom activation functions that surpass the activation functions that are already in use **[17].** L. Trottier, P. Gigu and B. Chaibdraa have mentioned that ReLU activation function suffers from the dying cells problem that deters the performance of the neural network model **[18].** Several explainable AI tools are explored by the researchers that aid in deciphering the rationale behind the prediction. As a result, the DL-based black box models are becoming more interpretable **[19].** An extensive correlation has been found between internships and students' employability, according to the model developed by O. Saidani, L. J. Menzli, A. Ksibi, N. Alturki, and A. S. Al-luhaidan **[20].** P. Guleria, M. Sood developed an AI enabled career counsellor that integrates the capabilities of Explainable AI (XAI) and ML to assess an educational dataset while monitoring the students' employability and skill development **[21].** An explainable fault analysis in mobile network has been carried out by M. Cilínio, M. Pereira, D. Duarte, L. Mata, and P. Vieira is based on boosting and TreeSHAP clustering **[22].**

III. Data Set Descrition

The dataset used in this study contains examination records of engineering students from a reputed university from India. The dataset is comprised of student's internal (sessional) and end semester examination marks starting from the time of enrolment in the degree program. Together with the examination marks, some preliminary information is also included as shown in Table 1.

TABLE 1. List Of Variables Of The Student's Dataset.

Columns	Column description
Sl. No.	Serial Number (of the students)
LPA	Yearly income of the student offered by the employers expressed in Lakh Per Annum
PLACED_ NOT_PLA CED	PLACED, NOT_PLACED (Whether the student got placement through on campus placement drive)
Regular_lateral	R, L (Admission through Regular or Lateral scheme)
HOSTELLERS	Y (Hostel accommodation), N (Student's own arrangement)
Type of Job offer	Highest (LPA >=5.0) High (LPA in the range of [3.0- 5.0] Low (LPA<3.0) 'NOT_PLACED_OR_OTHER_PLAN' (Either not placed or plan for higher studies or entrepreneurship)
CGPA	Cumulative Grade Point Average
BTECH_OVERALL_ PERENTAGE	Total marks received by the students in four years (in terms of percentage)
The percentage obtained in end term examination	3RD_SEM_PERCENTAGE till 8TH_SEM_PERCENTAGE
Gender	F (Female), M (Male)
PCM	Percentage obtained in Physics, Chemistry and Mathematics in Class XII board examination
ENGLISH	Percentage obtained in English in Class XII board examination
HS_Result	Division secured as 1^{st}, 2^{nd} or 3^{rd} in Class XII board examination
CATEGORY	SC, ST, STH, STP, OBC, General, NCC

Grades in various courses (both internal and end term examination marks)	**Programming Courses -** 'OOP Th','OOP Sess', 'OOP LAB PRACTICAL', 'OOP LAB Sess', 'Java Th', 'Java Sess', 'ITW LAB PRACTICAL', 'ITW Lab Sess', 'IWT Th', 'IWT LAB Sess' **Professional core courses** - 'MATH-III Th', 'MATH-III Sess', 'DS Th', 'DS Sess', 'DSA Th', 'DSA Sess', 'DSA LAB PRACTICAL', 'DSA LAB Sess', 'DM Th', 'DM Sess', 'COA Th', 'COA Sess', 'OS Th', 'OS Sess', 'GT Sess', 'GT Th', 'OS Lab Practical', 'OS Lab sess', 'DBMS Th', 'DBMS Sess', 'DAA Th', 'DAA Sess', 'FLAT Th', 'FLAT Sess', ''DBMS LAB PRACTICAL', 'DBMS LAB Sess', 'Compiler Design Th', 'Compiler Design Sess', 'Computer Networks Th', 'Computer Networks Sess', 'ACCOUNTANCY Th', 'ACCOUNTANCY Sess', 'Compiler Design Lab Practical', 'Compiler Design Lab Sess', 'Computer Networks Lab practical', 'Computer Networks Lab Sess', 'PM Th', 'PM Sess' **Elective courses** - 'PE1 Th', 'PE1 Sess', 'PE2 Th', 'PE2 Sess','PE3 Th', 'PE3 Sess', 'OE1 Th', 'OE1 Sess', 'PE4 Th', 'PE4 Sess','OE2 Th', 'OE2 Sess', 'OE3 Th', 'OE3 Sess', 'PE5 Th ', 'PE5 Sess', 'PE6 Th', 'PE6 Sess', 'OE4 Th','OE4 Sess', 'OE5 Th', 'OE5 Sess' **Internship and Projects** - 'INTERNSHIP I', 'INTERNSHIPI Sess', 'INTERNSHIP-II', 'INTERNSHIP-II Sess', 'MINOR PROJECT', 'MINOR PROJECT Sess', 'MAJOR PROJECT','MAJOR PROJECT Sess'

IV. Methodology And Proposed Framework

The proposed time series prediction framework is based on a hybridized approach leveraging two popular DL approaches – Convolutional Neural Network (CNN) and Gated Recurrent Network (GRU) and implemented on a time series dataset. The dataset is prepared with the data collected from three hundred and thirty-four BTech students pursuing BTech in CSE. It is composed of 93 features over a time period of 3 academic years of BTech program. The record of each student per semester acts as one of the sequences that creates total 1820 input sequences for the time series model as shown in the Fig. 1.

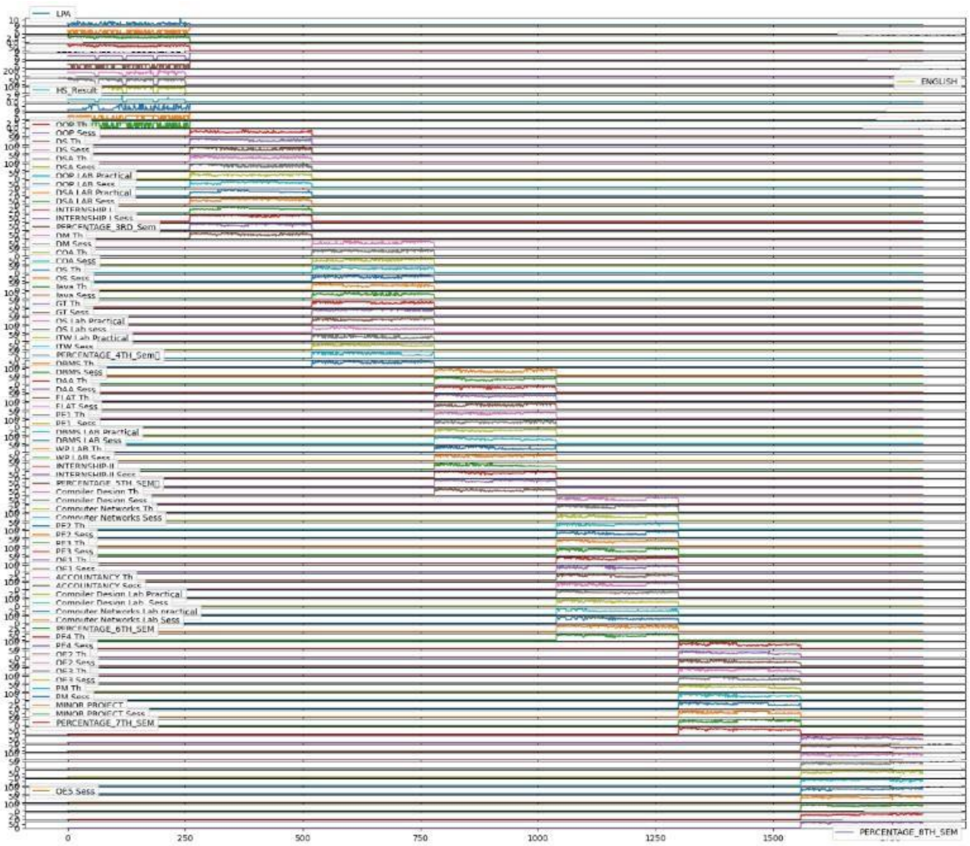

Fig. 1 Time series dataset of BTech student (3rd to 8th semester)

The Fig. 1 illustrates the formation of sequence for each student. There are 6 number of input sequences generated per student for 6 consecutive semesters – semester 3rd to semester 8th. The 1st semester is merged with the demographic information of the students. As a preprocessing step, an exploratory analysis is conducted to see the ratio of placed vs not placed LE students compared to RE students as shown in Fig. 2.

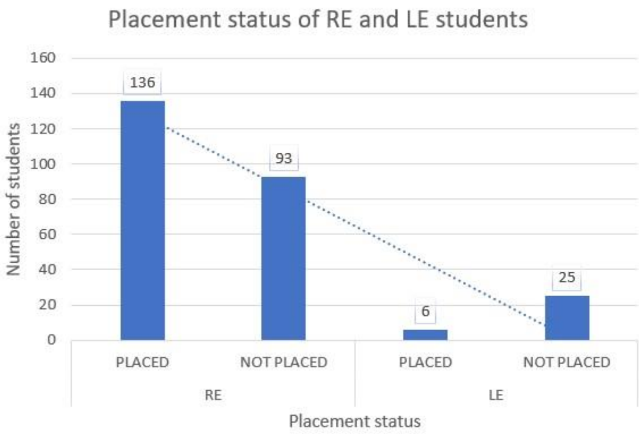

Fig. 2 Placement status of RE and LE student

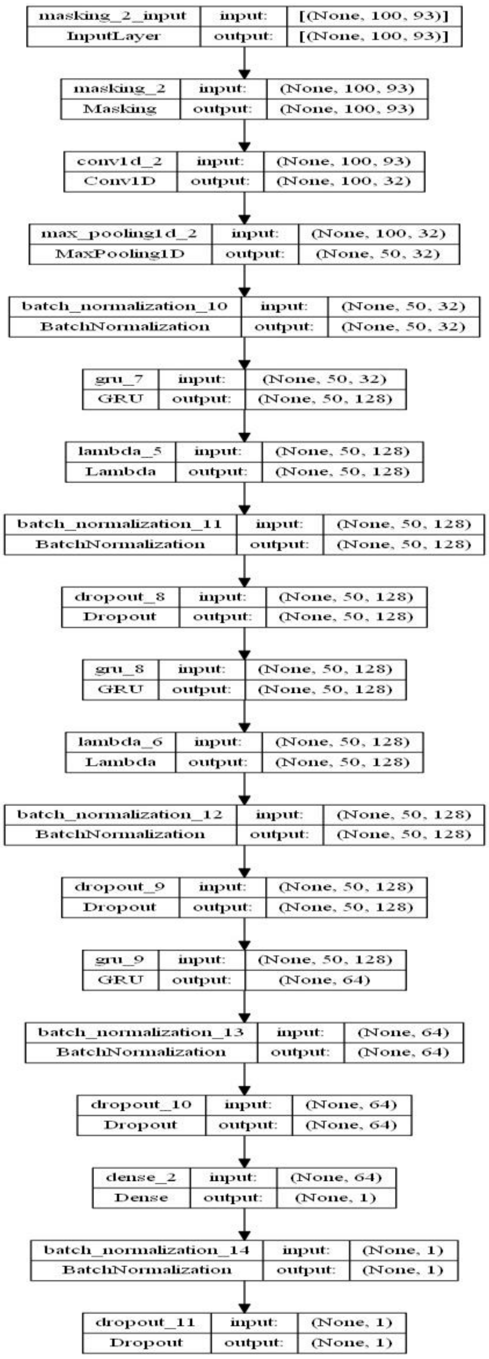

Fig. 3 Architecture of the DL based proposed framework

As shown in the Fig. 2, the number of placed students is 136 out of 229 students in the RE category. Whereas only 6 students are placed out of 31 students in LE side. The result shows that the employability rate for LE students is 19.35% compared to 9.38 % for LE students.

As listed in the Table 1, the data set can be dived in two type of variables – independent and dependent variables. The student's demographic information and examination grades (both internal and end term) are prepared as independent variables and yearly income in terms of Lakh Per Annum (LPA) is used as target variable expressed. Although most of the features in the dataset are numerical, a few of those are categorical. The categorical features are viz.

Placed_Not_Placed,
Regular_Lateral, Hostellers, Type Of Job

OFFER, GENDER and CATEGORY. The categorical features are converted to numerical values though Label Encoding process. Data scaling is applied on each feature after generating the numerical dataset. This pre-processing step helps the algorithm to converge faster during the model training. Throughout this process feature selection step is conducted by the CNN. In the convolution layer of CNN, a custom activation function is applied that is based on Scale Exponential Linear Unit (SeLU). This custom activation processes the internal normalization of the model where each layer maintains the mean and variance from the previous layers. By modifying the mean and variance, SELU facilitates this normalizing process and accelerates internal normalization relative to external normalization **[23].** It optimizes the DL network to converge faster. After selecting the required features in the model, the GRU processes the input sequences, which is a recurrent neural network. In this process two Lambda layers are added to the GRU. The neurons in the custom Lambda layer are activated by another custom activation function based on hyperbolic sine function. The addition of new activation function and Lambda layer help to enhance the prediction performance of the network. Each layer of GRU is integrated with batch normalization and dropout. A custom attention layer is added as one of the hidden layers of GRU that helps the model to focus on inputs with higher weight. The attention mechanism helps the model to better fit the training data. The **Fig. 3** shows the overall architecture of the proposed model. As shown in the **Fig. 3,** the time series dataset is fed to the input layer followed by a masking layer. Masking is done to covert the variable length sequences to the fixed length and make the dataset ready for the convolution layer in CNN. The CNN is merged with GRU and Attention layer with the other helping companions – Lambda layer, batch normalization, activation function and drop out. The proposed model works as regression model where 'LPA' is chosen as the target variable which is continuous in nature. That is how the CNN-GRU is capable of interpreting the relationship between the student's employability status with the examination grade received from the previous semesters along with the demographic data. The model is compiled with 'Mean Squared Error (MSE)' as loss function and 'Mean Absolute Error (MAE)' as performance metric to evaluate the model's performance. One of the famous optimizers named as 'Adam' is used that helps the proposed model to automatically adjust the weights. The contribution of most impactful features in the prediction process are evaluated with the help of an interpretable framework – SHAP & LIME. The global interpretation of the input features is done through the mean of the SHAP (SHapley Additive exPlanations) values where SHAP values from multiple instances of the input dataset are combined to comprehend the model's overall behaviour. Thus, the global interpretation process identifies the most important features that influence the model's prediction. Another explainable AI tool knows as LIME (Local Interpretable Model-agnostic Explanations) focuses on illuminating the model's prediction for specific sequences rather than offering a comprehensive knowledge of the entire dataset. The results of this module help to derive the conclusion of Inductive Reasoning (IR) in terms of types of job offers received by the LE students based on their performance in specific courses. The query able framework is also prepared for information retrieval based on two queries i,e CEE marks and student's accommodation. The effectiveness of the proposed model is deliberated in the later section as Results and Discussion.

V. Results And Discussions

The outcomes of the suggested work are divided into the following categories for simpler analysis:

A. Performance evaluation of the proposed model As discussed in the previous section, a DL based predictive model is proposed in this study to meet the objective of student's performance prediction in their early academic days. As this research is specifically focused on the LE students in the domain of CSE, the evaluation starts at the end of 3rd semester. The internal and end term examination grades obtained by the students from BTech 3rd semester along with their demographic information are used to train the model under study. Due to a smaller number of input records, the MSE and MAE are expected to be more after 3rd semester. Subsequently, the next phase of evaluation is conducted based on their cumulative results taken from the 4th semester and 3rd semester. This way the evaluation process continues until 8th semester. The level of prediction performance improves as the student moves towards the next semester and the model is expected to provide the best accuracy at the end of 8th semester. As it is a nonlinear regression problem, the performance metrics used here are MSE and MAE as shown in the Table 2. The Table 2 shows the behaviour of the model at the end of each semester in an incremental approach. The MSE and MAE of the proposed framework decrease with the increase in training data collected after each successive semester. This is how the prediction performance gradually improves with the addition of a greater number of inputs. It indicates that the employability of a student can be predicted in their early academic days and the prediction accuracy improves as the student gets promoted to the next semester. This is the beauty of recurrent neural network like GRU that can nicely map the input – output relation.

TABLE 2. Performance Of The Model At The End Every Semester As Incremental Way

Model- CNN-GRU+ Attention+ Lambda layer +Custom activation				
Data Input	**MSE- Train**	**MSE - Test**	**MAE - Train**	**MAE - Test**
3rd semester	0.037 5	8.61E -04	0.121 3	0.027
(3rd+4th) semester	0.029 5	0.002 6	0.098	0.051 3
(3rd+4th+5th) semester	0.019 2	0.001	0.073 2	0.031 3
(3rd+4th+5th+6th) semester	0.014 9	1.89E -05	0.059 6	0.004 2
(3rd+4th+5th+6th+7th) semester	0.014 7	0.001 9	0.060 6	0.043 4
(3rd+4th+5th+6th +7 th+8th) semester	0.014	7.31E -04	0.059 2	0.029 4

A comparative analysis is conducted to evaluate the performance of the model with two other base line models – CNN-SimpleRNN and Multi-Layer Perceptron (MLP). The Fig.4 explains the result of comparative analysis.

CNN-GRU+Attention MSE (Train)	CNN-GRU+Attention MSE (Test)	CNN-GRU+Attention MAE (Train)	CNN-GRU+Attention MAE (Test)	CNN-SimpleRNN + Attention MSE (Train)	CNN-SimpleRNN + Attention MSE (Test)	CNN-SimpleRNN + Attention MAE (Train)	CNN-SimpleRNN + Attention MAE (Test)	MLP (Without Attention) MSE (Train)	MLP (Without Attention) MSE (Test)	MLP (Without Attention) MAE (Train)	MLP (Without Attention) MAE (Test)
0.014	7.31E-04	0.0592	0.0294	0.9274	0.4656	0.4438	0.6809	5.3447	6.3019	1.8596	1.8809

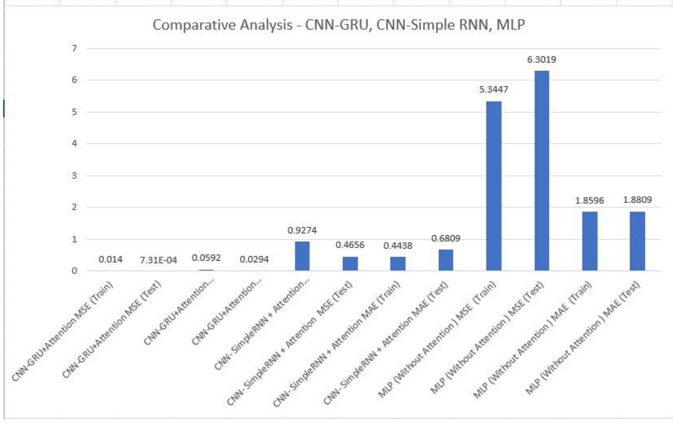

Fig. 4 Comparative Analysis - CNN-GRU, CNN-Simple RNN, MLP

From the Fig. 4, it is observed that among the three models, the losses are highest in MLP. The CNN-SimpleRNN model has shown moderate level of losses. Out of all these three models, the prediction performance is highest in the proposed model with CNN-GRU+ Attention.

B. Interpretation of the most influential feature(s) in the model prediction

After receiving both Diploma and BTech degree, the LE students aspire to be employable as soon as they receive the degree certificate. However, by means of employability the student's expectations may not be strictly to the jobs through placement drive only. They may wish to pursue either higher study or become an entrepreneur after graduation in the field of engineering. Through this study, a survey is conducted on employability of LE students to investigate whether LE students would be employable or not after completion of BTech program. If employable, then what kind of job offer they would be able to grab. That would be based on the rank of employers and yearly pay package expressed in LPA as shown in Table 1. An interpretable framework is integrated with the model (as shown in the Fig. 3) to assess different category of courses included in the course structure of BTech program. The solitary purpose of using this interpretability module is to determine the impact of different courses on students' employability. However, the mandatory courses (as mentioned in Table 1) are omitted from this research work as these are non-credit course. The contribution of programming, professional core courses, electives, internships and projects are discussed below with respect to employability of LE students.

1) Programming courses and employability

Programming is a vital part of the curriculum where many students from the technical background either fail or unable to conceptualize the basic terminologies. The programmers are in high demand in IT industry. The students often disqualify in the technical round of job interviews only in the Programming part. The **Fig. 5** shows the most impactful programming subjects for the students to qualify the interviews and achieve a high paid job through placement drive.

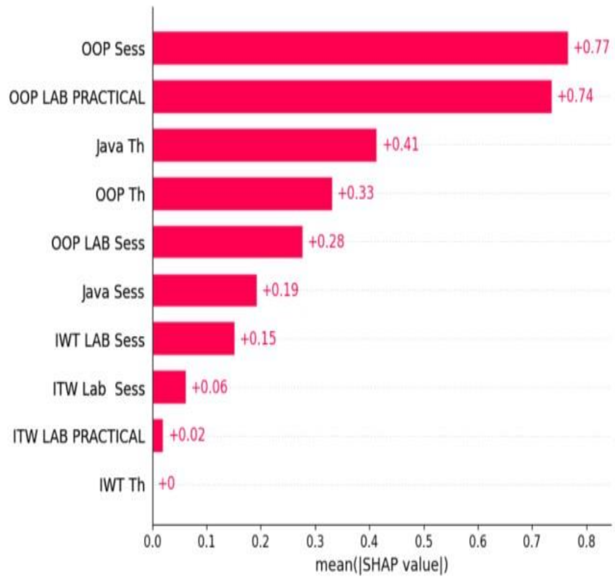

Fig. 5 Mean SHAP plot showing the contribution of important

Programming courses for employability

It is obvious from the Fig. 5 that the internal grades in OOP subjects have a significant influence on employability. Additionally, candidates should possess solid conceptual understanding in theoretical and practical areas of many programming subjects like Java Th, OOP LAB PRACTICAL, and so on.

Although it is not feasible to explain each student's performance on their internal and final exams, only two cases are selected for this study. The visual interpretation through the LIME plot as shown in the Fig. 6 explains the rationale behind "NOT_PLACED_OR_OTHER_PLAN" as placement status with a 96% likelihood.

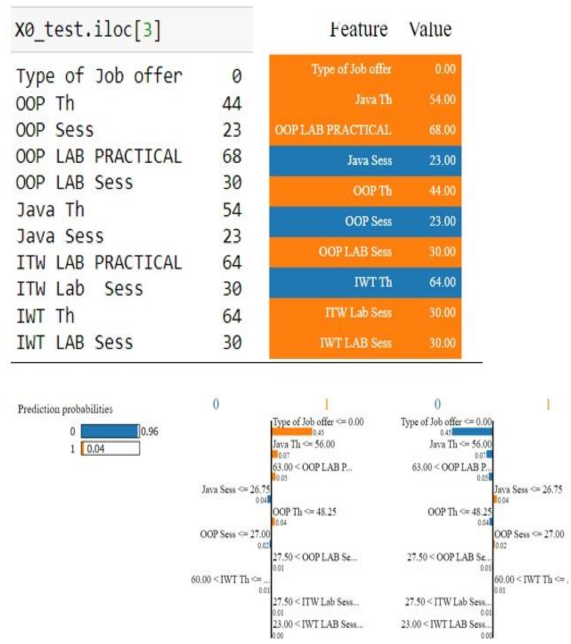

Fig. 6 LIME plot for the instance with Placement status:

NOT_PLACED_OR_OTHER_PLAN (Course category: Programming)

The Fig. 6 displays one of the instances from the test dataset under LE category. The features with blue background indicate probable predictors of the model with placement status i,e "NOT_PLACED_OR_OTHER_PLAN." It can be observed that the student is unable to get acceptable internal marks in - OOP (Object Oriented Programming using C++) and Java. Whereas both these two courses are important and are highly demanded for the job profile of software developers. The student's performance in IWT (Internet and Web Technology) Theory is also less than 70%. The remaining variables (with orange background) are not supposed to be the primary predictors for the placement status of "NOT_PLACED_OR_OTHER_PLAN."

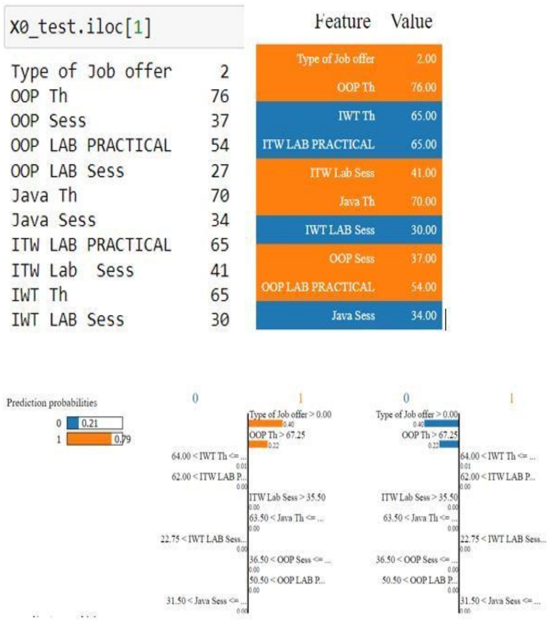

Fig. 7 LIME plot for the instance with Placement status:

PLACED (Course category: Programming, Type of Job Offer: High LPA)

The Fig. 7 shows the interpretability results for another test instance having placement status as ‘PLACED’. The student represented by this instance has received a high paid job offer indicated by the ‘Type of job offer’ as 2 with 79% probability as displayed in orange background**.** Looking at the marks obtained in various programming courses, it is observed that student has good command in most of the programming courses. However, the marks in the OOP’s practical are below 60%.

2) Professional core courses and employability

Apart from the hardcore programming courses, the recruiters evaluate the students based on their core subject’s competency level. The technical skills in the professional core CSE courses are very much essential so that they can be fit any of the IT domain in their future job. The Fig. 8 shows the list of impactful core courses in receiving a high paid job offer.

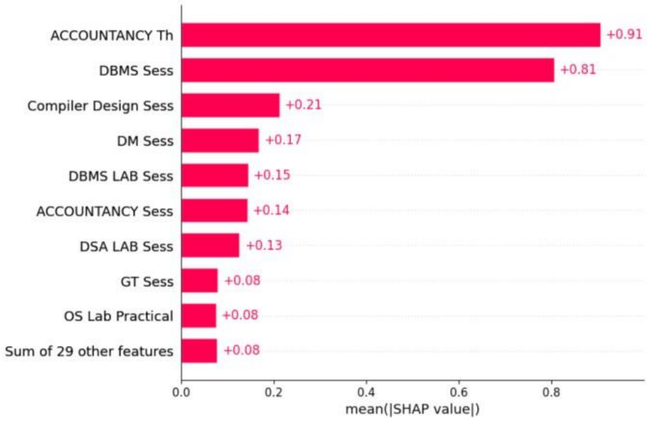

Fig. 8 Mean SHAP plot showing the contribution of important Professional core courses for employability

The Fig. 8 illustrates that although not directly used in the domain of IT, Accountancy theory has the significant importance on employability. The recruiters might give preferences in Accountancy subject for Business analyst or Program manager position. The internal marks in subjects like Data Base Management Systems (DBMS), Compiler Design, Discrete Mathematics (DM) etc have significant impact on employability.

The Fig. 9 indicates that the student has received good internal marks in DBMS subject. Due to the relatively lower internal marks in Computer Network Lab and Project Management (PM), the student was unable to obtain employment through placement drives conducted by the institution. However, the student might have different career plans, such as starting own start-up or pursuing higher study.

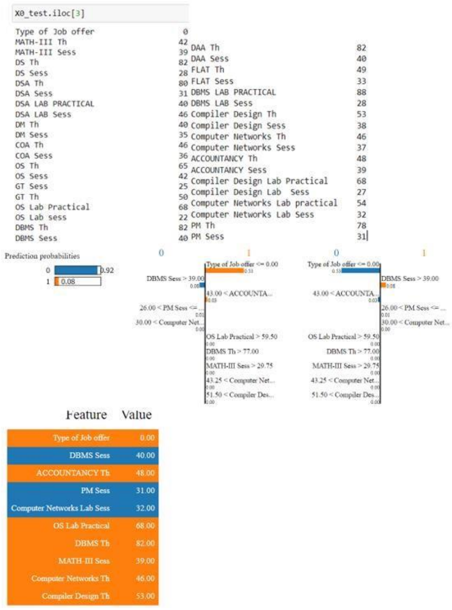

Fig. 9 LIME plot for the instance with Placement status: NOT_PLACED_OR_OTHER_PLAN (Course category: **Professional core courses**)

The instance under our investigation explained in the Fig. 7 is also examined in this section in order to determine the impact of other core professional courses on high-paying placements.

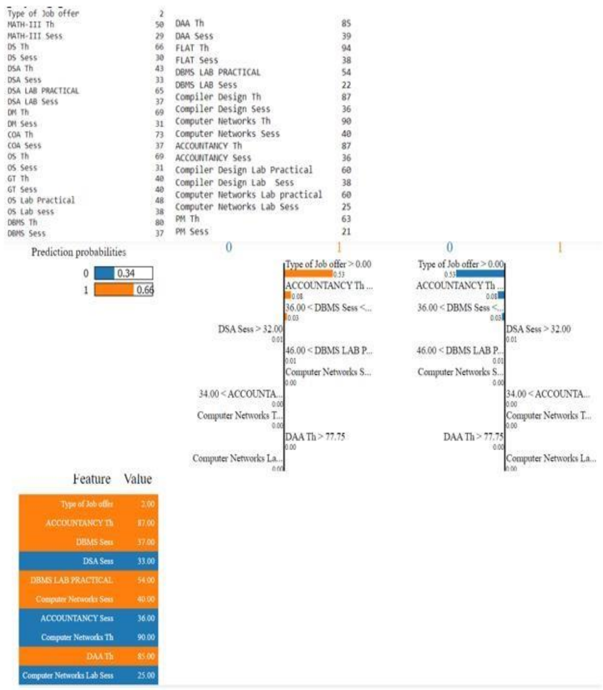

Fig. 10 LIME plot for the instance with Placement status: PLACED (Course category: **Professional core courses**, Type of Job Offer: High LPA)

The Fig. 10 describes the list of professional core courses used in the prediction process of another test instance. The individual has performed good in some of the important core subjects like – ACCOUNTANCY and DAA Theory. The internal marks obtained in Computer Network sessional is also quite impressive followed by DBMS sessional. However, he/she was not able to score good in the DBMS practical which is very much essential for the Data Base Developers and Data Base Administrator roles in IT industry.

3) Electives and employability

Students are offered Professional Elective (PE) courses related to their specialization in accordance with the AICTE model curriculum. Additionally, undergraduate engineering students can select Open Electives (OE) from emerging areas. The Fig. 11 illustrates how important are the optional courses in terms of students' employability.

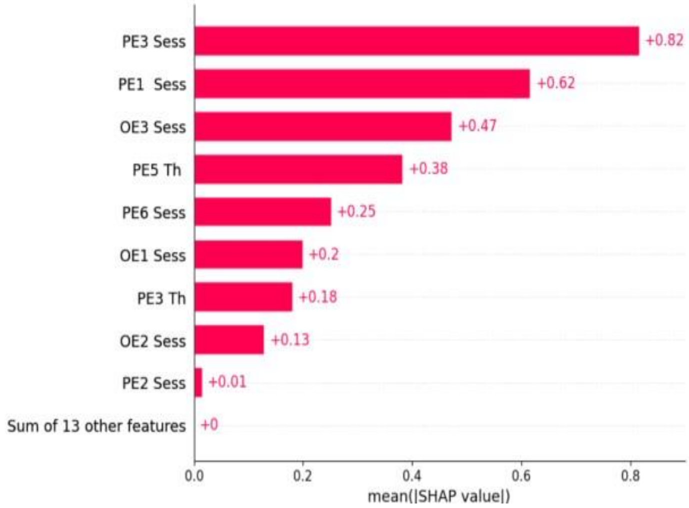

Fig. 11 Mean SHAP plot showing the contribution of important elective courses for employability

The Fig. 11 confirms that internal marks in PE3 course has the largest impact on employability. The other important predictors are - internal marks in PE1 and OE3, PE5 Th and so on.

.

The Fig. 12 shows the local interpretation provided by the LIME plot that indicates the performance of the students in this category and their placement status

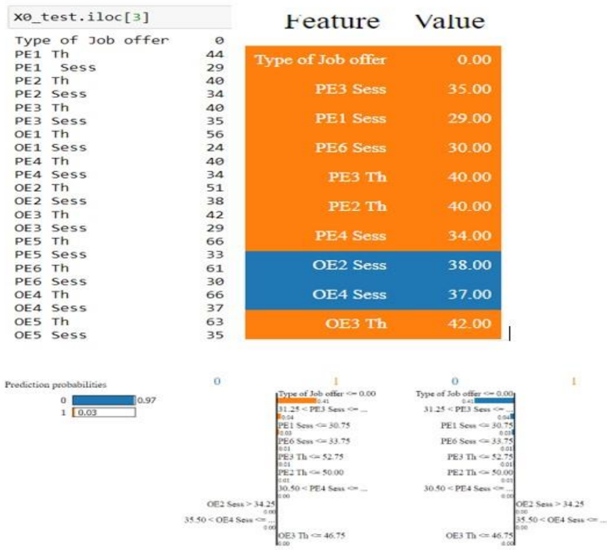

Fig. 12 LIME plot for the instance with Placement status:

NOT_PLACED_OR_OTHER_PLAN (Course category: Electives)

The internal marks obtained by the student in OE2, and OE4 are satisfactory as shown in the Fig. 12. However, the placement status for this instance is marked as 'NOT_PLACED_OR_OTHER_PLAN'.

The instance under study as explored in the Fig. 7 and Fig. 10, is examined again to determine his/her performance in elective courses leading to the placement status as shown in Fig. 13.

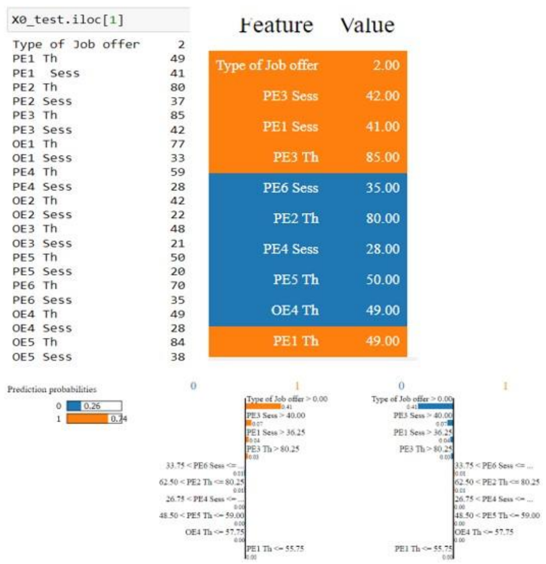

Fig. 13 LIME plot for the instance with Placement status:

PLACED (Course category: Electives, Type of Job Offer: High LPA)

Looking at the marks obtained by the students in various Elective courses, it is observed that student was able to score good marks in PE3, both in theory and internal. However, marks in PE1 theory were not good as shown in the Fig. 13.

4) Internship, projects and employability

The students gain real time industry exposure from internship opportunities, which cannot be replicated in the classroom environment, resulting in the production of qualified industry experts. Apart from the internship, the students are required to complete their final year project in two phases – phase I and phase II. Internship and project work measure their knowledge gained over the past academic years of BTech program **[24].**

The Fig. 14 illustrates how internships and projects help students to prepare for the workforce and increase their chances of landing a better job.

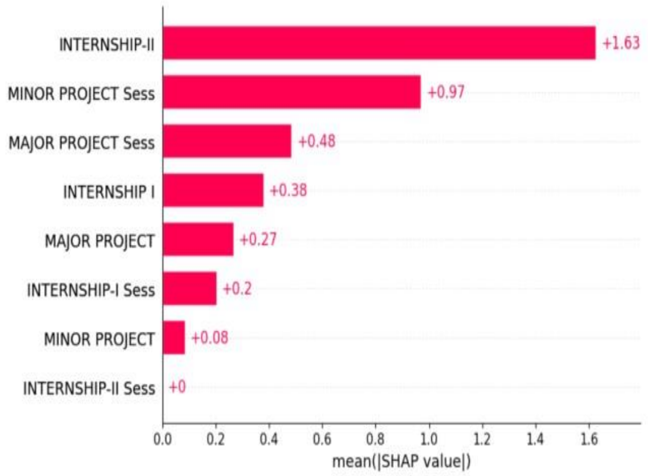

Fig. 14 Mean SHAP plot - Global interpretation of contribution of internship and project works on campus placement of LE students

Based on Fig. 14, it is evident that grades in Internship II have substantial influence on employability. Moreover, the internal marks in MINOR PROJECT and MAJOR PROJECT help the students to get placed in high paid jobs. The MAJOR PROJECT is more crucial than MINOR PROJECT as it is the Phase II of the Project cycle. The MAJOR PROJECT measures student's competency level in core courses, team building and communication skill.

The Fig. 15 shows the local interpretation provided by the LIME plot that indicates the performance of the students in this category and their placement status.

Fig. 15 LIME plot for the instance with Placement status:

NOT_PLACED_OR_OTHER_PLAN

(Course category: Internship and Projects)

As shown in Fig. 15, even though the student has shown good performance in MINOR PROJECT and INTERNSHIP II, the placement status of the individual is marked as 'NOT_PLACED_OR_OTHER_PLAN'.

The Fig. 16 displays the interpretation result that indicate student's performance (taken from the same instance as displayed in the Fig. 7, Fig 10 and Fig. 13) in academic and industry internships along with the projects.

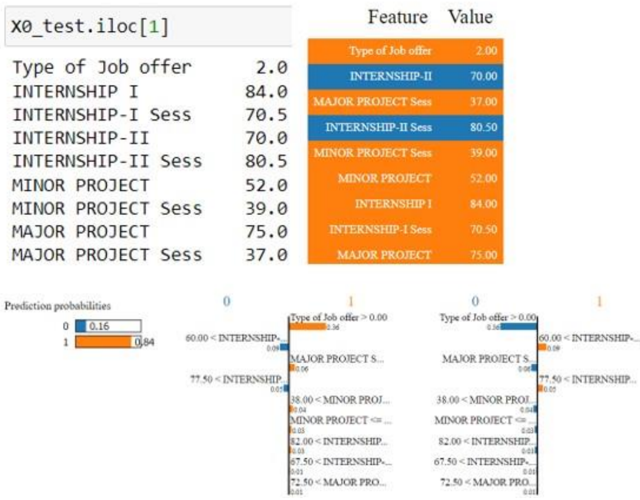

Fig. 16 LIME plot for the instance with Placement status:

PLACED (Course category: Internship and projects, Type of Job Offer: High LPA)

From the Fig. 16, it is observed that student was able to score good marks of above 70% in INTERNSHIP I and MAJOR PROJECT which helped to land in a high paid job.

C. Other influential factors on employability

In addition to their academic performance, the employability status of LE students is compared with the RE students with respect to the CEE marks. It helps to investigate the correlation between CEE marks and employability after graduation. Similarly, this study has explored the correlation between students' employability and accommodation during their study period (on campus or off campus).

1) CEE and employability

As an eligibility criterion RE students are required to have a good score in CEE which is exempted for LE students. The Fig. 17 and Fig. 18 show the employability status of LE and RE students in terms of CEE.

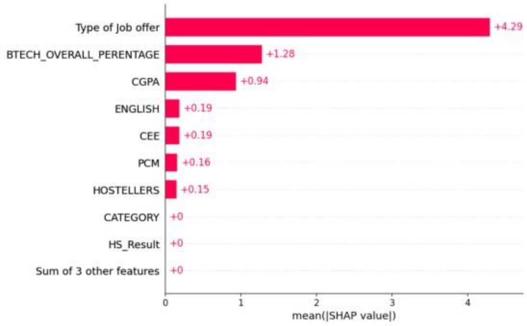

Fig. 17 Impact of CEE on Type of Job (RE students)

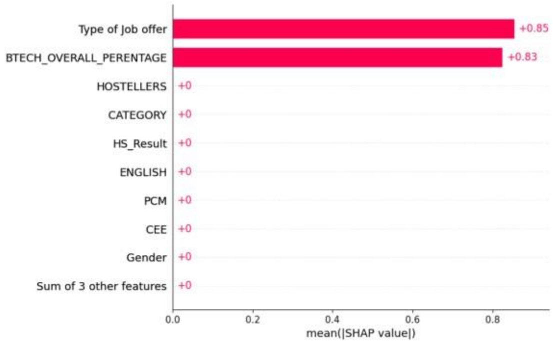

Fig. 18 Impact of CEE on Type of Job (LE students)

The employability of RE students is positively impacted by CEE marks that is indicated by the positive mean SHAP value of +0.19 as shown in the Fig. 17. However, as Fig. 18 illustrates, there is no correlation between employability and BTech admission criteria exercised by the HEIs for LE students and it is confirmed in the Fig. 18 showing the mean SHAP value of +0 for CEE in case of LE students.

2) Campus accommodation and employability

Many times, students prefer to stay in the university or college hostel for easy communication. It also works as a time saving system as the students do not have to spend hours in transportation for their regular classes. They acquire knowledge in extra-curricular activities being the residents of the same university or college where they pursue their BTech program. The students can improve their soft skills and get longer study hours compared to the other students who manages their own accommodation. The Fig. 19 illustrate how the hostel accommodation facility affects students' performance of LE students.

Fig. 19 Impact of Hostel in the final performance of LE students

As shown in the Fig. 19, there is no correlation between Hostel accommodation and employability in case of LE students. The results obtained from IR clarifies that both the instances under study have availed Hostel accommodation. Although one of these two instances was able to get a high paid job, while the other was not placed.

The IR approach is also applied for RE students to verify the correlation between Hostel accommodation and employability status as depicted in the Fig. 20.

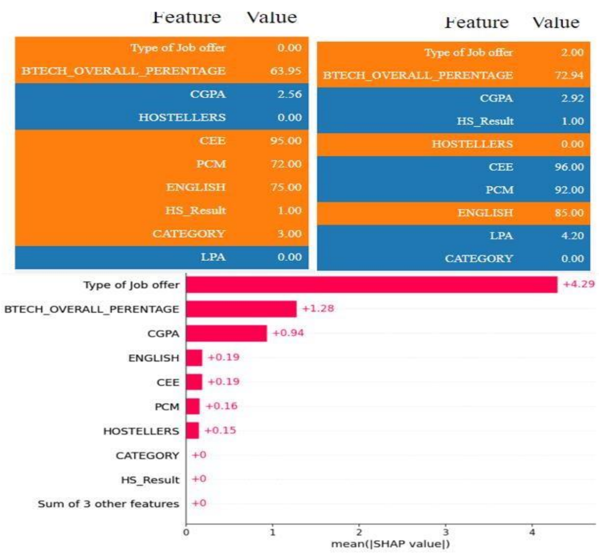

Fig. 20 Impact of Hostel in the final performance of RE students

There are some students who do not like the restriction imposed by the hostel wardens to its resident students. Because of this fear, such students prefer to stay outside the college or university campus. The RE students as shown in the Fig. 20 staying outside campus as non-hostellers are able to show better performance by getting placed in high paid jobs. However, there are some students who are also non-hostellers, but with
employability status as –

'NOT_PLACED_OR_OTHER_PLAN'.
From this analysis it is hard to draw a conclusion that "employability and accommodation" is a topic of IR where it is difficult for us to give a generic conclusion about the facts.

VI. Conclusion And Future Work

Student's employability prediction at their early academic days is one of the important responsibilities of country's HEIs. The LE students having their diploma in engineering can join any branch of engineering as per the New Education Policy (NEP) to gear up their technical knowledge. However, this study has verified that there is a large gap of inequality in terms of employability between RE and LE students. The proposed framework for employability prediction has leveraged the power of CNN and GRU with numerous optimization techniques that helps students pursuing BTech in CSE to predict their final academic performance in terms of employability. Consequently, the students would be motivated and focused more on the subjects that are in real sense trending in the current job market. The projected model is able to provide favourable prediction performance compared to the base line models under this study. The analytical results indicate that programming, internship and major projects have significant importance in employability. Being in the engineering discipline, the practical skills are equally important for the students to get selected through placement drives. Moreover, soft skills have greater impact on employability that could be gained by improving the internal marks associated with each course. The internal marks evaluate student's sincerity through attendance mechanism and accountability through timely submission of assignments. The consistency level is evaluated based on their continuous improvement in class tests or midterm examination conducted every semester. The IR results claim that student's employability has no corelation with either CEE marks or on campus or off campus accommodation. By implementing the best practices prescribed by the framework, LE students together with their fellow RE mates could shape their professional career leading to great satisfactory and comfortable job with respect to salary packages and work culture. As a future scope, we continue to work on some advanced optimization techniques and deep explainable AI tools to achieve better results in the field of Educational Data Mining.

Acknowledgment

The authors greatly acknowledge Assam Science and Technology University for giving the opportunity to pursue research in the area of Education Data Mining using AI and Deep learning.

References

[1] https://timesofindia.indiatimes.com/education/news/lateral-entry-foradditional-btech-degree-to-enhance-job-prospects/articleshow/85650880.cms, last accessed 2024/02/28.

[2] G. S. Nandakumar, L. M. Palanivelu, "A Research on the Performance of the Diploma Students in the Lateral Entry Scheme of Engineering Degree Program", International Journal of Engineering and Advanced Technology (IJEAT), Vol. 8, 2019.

[3] R. Costa-Mendes, T. Oliveira, M. Castelli, C. J. Frederico, "A machine learning approximation of the 2015 Portuguese high school student grades: A hybrid approach", Educ Inf Technol, vol. 26, pp. 1527–1547, 2021.

[4] S. Alturki, I. Hulpuș, H. Stuckenschmidt, "Predicting Academic Outcomes: A Survey from 2007 Till 2018", TechKnow Learn, vol. 27, pp. 275–307, 2022.

[5] A. J Baruah, S. Baruah, "PFCMVO: Political fractional competitive multiverse optimization enabled deep neuro fuzzy network for student performance estimation in spark environment", International Journal of Modeling, Simulation, and

Scientific Computing, vol. 13, pp. 2250050:1-2250050:25, 2022.
[6] M. Bora, R. Baruah, "Sustainable Development of Employability and Skill Development in Outcome-based Education: A Machine Learning Approach" in International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, pp. 394-400, 2023.
[7] M. Elmasry, "Applying Deep Learning Neural Networks in Predicting Students' Cumulative Grades", International Journal of Engineering and Advanced Technology, vol. 10, pp. 293-297, 2020.
[8] N. Kumar, H. Krishna, S. Shubham, P.P. Rout, "Automated categorization of stu-dent's query", International Journal of Computers and Applications, vol. 44, pp. 1056-1064, 2022.
[9] W. Tianchang, "Online student achievement prediction model based on Data Envel-opment Analysis (DEA) for multidimensional educational internet of data mining", International Journal of Computers and Applications, vol. 43, pp. 315-319, 2021.
[10] B. Sekeroglu, K. Dimililer, K.Tuncal, "Student Performance Prediction and Classi-fication Using Machine Learning Algorithms", ICEIT 2019: Proceedings of the 2019 8th International Conference on Education-al and Information Technology, pp. 7-11, 2019.
[11] A. Çetinkaya, O. Baykan, "Prediction of middle school students' programming talent using artificial neural networks", Engineering Science and Technology, an Inter-national Journal, vol. 23, 2020.
[12] M. Tadayon, G. J. Pottie, "Predicting Student Performance in an Educational Game Using a Hidden Markov Model", IEEE Transactions on Education, vol. 63, pp. 299-304, 2020.
[13] T.T. Dien, T. -N. Nguyen, H. Nguyen, S. Luu, "Deep Learning with Data Transformation and Factor Analysis for Student Performance Prediction", International Journal of Advanced Computer Science and Applications, vol. 11, pp. 711-721, 2020.
[14] S. Poudyal, M. Mohammadi-Aragh, J. Ball, "Prediction of Student Academic Per-formance Using a Hybrid 2D CNN Model, Electronics, vol. 11, 2022.
[15] S. Jung, J. Moon, S. Park, E. Hwang, "An Attention-Based Multilayer GRU Model for Multistep-Ahead Short-Term Load Forecasting", Sensors, vol. 21, 2021.
[16] A. Aljaloud, D. Uliyan, A. Alkhalil, M. Elrhman, A. Alogali, Y. Altameemi, M. Al-tamimi, P. Kwan, "A deep learning model to predict students' learning outcomes in LMS using CNN and LSTM", IEEE Access, vol. 10, 2022.
[17] K. Biswas, S. Kumar, S. Banerjee, A., Pandey, "TanhSoft - a family of activation functions combining Tanh and Softplus".
[18] L. Trottier, P. Gigu, B. Chaibdraa, "Parametric exponential linear unit for deep convolutional neural networks, in: Machine Learning and Applications (ICMLA)", in 16th IEEE International Conference on, IEEE, pp. 207–214, 2017.
[19] https://www.datacamp.com/tutorial/introduction-to-shap-values-machinelearning-interpretability, last accessed 2024/03/29.
[20] O. Saidani, L. J. Menzli, A. Ksibi, N. Alturki, A. S. Alluhaidan, " Predicting Student Employability Through the Internship Context Using Gradient Boosting Models", IEEE Access, vol. 10, pp. 46472-46489, 2022.
[21] P. Guleria, M. Sood, "Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling", Educ Inf Technol, vol. 28, pp. 1081–1116, 2023.
[22] M. Cilínio, M. Pereira, D. Duarte, L. Mata, P. Vieira, "Explainable Fault Analysis in Mobile Networks: A SHAP-Based Supervised Clustering Approach", in 16th International Conference on Signal Processing and Communication System (ICSPCS), Bydgoszcz, Poland, pp. 1-9, 2023.
[23] https://www.v7labs.com/blog/neural-networks-activationfunctions(last accesses, 2024/03/28)
[24] https://medium.com/@mailforengineering/what-is-the-importance-of-final-year-engineering-projects-for-students-1730ff44557e#:~:text=By%20doing%20a%20project%2C%20the,and%20they%20reinforce%20their%20learning.(last accesses, 2024/03/28)