Check for updates

Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling

Pratiyush Guleria[1] · Manu Sood[2]

Abstract

Machine Learning concept learns from experiences, inferences and conceives complex queries. Machine learning techniques can be used to develop the educational framework which understands the inputs from students, parents and with intelligence generates the result. The framework integrates the features of Machine Learning (ML), Explainable AI (XAI) to analyze the educational factors which are helpful to students in achieving career placements and help students to opt for the right decision for their career growth. It is supposed to work like an expert system with decision support to figure out the problems, the way humans solve the problems by understanding, analyzing, and remembering. In this paper, the authors have proposed a framework for career counseling of students using ML and AI techniques. ML-based White and Black Box models analyze the educational dataset comprising of academic and employability attributes that are important for the job placements and skilling of the students. In the proposed framework, White Box and Black Box models get trained over an educational dataset taken in the study. The Recall and F-Measure score achieved by the Naive Bayes for performing predictions is 91.2% and 90.7% that is best compared to the score of Logistic Regression, Decision Tree, SVM, KNN, and Ensemble models taken in the study.

Keywords Artificial intelligence · Career counseling · Educational data mining · Explainable AI · Machine learning

✉ Pratiyush Guleria
pratiyushguleria@gmail.com

1   National Institute of Electronics and Information Technology (NIELIT), Shimla, Himachal Pradesh, India

2   Department of Computer Science, Himachal Pradesh University, Shimla, Himachal Pradesh, India

 Springer

1 Introduction

The right career direction boosts up the performance of students and increases their motivation level (The Importance of Motivation in an Educational Environment, 2012). The career counseling-based systems are most important to judge the skill of students and accordingly, recruiters assign the right job role to them (Viheräkoski, 2020). Career counseling is required to solve the problems which come into the mind of parents for their children where guidance cell will play an important role to guide parents to opt for the right course for their children. Here, Machine Learning will be used to carry out the computational intelligence task in a smart manner (Yang et al., 2020).

ML techniques help solves students', parent's queries related to the right career choice. It uses data mining techniques and algorithms to learn from experiences. The ML-based system uses Artificial Intelligence Techniques, Deep Learning, Neural Network, Natural Language Processing, etc. for effective decision making related to the career guidance of students. Career counseling helps students to take the right decision to choose the right career approach to leverage their interests and capabilities. Moreover, due to the increase in educational institutions and the competitive environment, it becomes a tedious task for parents as well as children to choose the right career direction. The Machine Learning enriched system recommends the students to choose the right course curriculum as per the capability of the incumbent. The ML-enabled systems help in guiding students according to the capability of the students. The system requires input data from students which consists of attributes as follows: a) subject interest of students in senior secondary (Non-Medical, Medical, Arts, Accounts), b) specialization, c) marks obtained in the employability tests, d) work experience, e) status of placements, f) marks obtained in the post-graduation, etc. The data comprising of these traits may be collected from elementary education, higher education, employment offices, and through local administration using qualitative, quantitative methods like online conventions, communities, surveys, questionnaires, face-to-face interaction programs with students in collaboration with educational institutions and colleges.

The standard approach of career counseling followed without incorporating intelligent tools and techniques is shown in Fig. 1. In this approach, career counselors get input from a student about his interests, aptitude skills, goals, etc. Similarly, data about parents' educational background, profession, income, etc. are collected. Subsequently, the counselor tries to discover the hidden insights from the responses of parents, students and suggests the appropriate career guidance or field chosen by children like arts, finance, engineering, medicine, etc. The standard approach of counseling has a limited scope and the subject domain area is also limited. There is an absence of intelligent techniques, knowledge base, inference engine, and database for storing the data gathered. In absence of data, ML techniques cannot be applied to infer the results and perform predictive analytics.

Compared to the standard counseling approach, researchers have proposed intelligent educational frameworks using AI, Machine learning techniques for
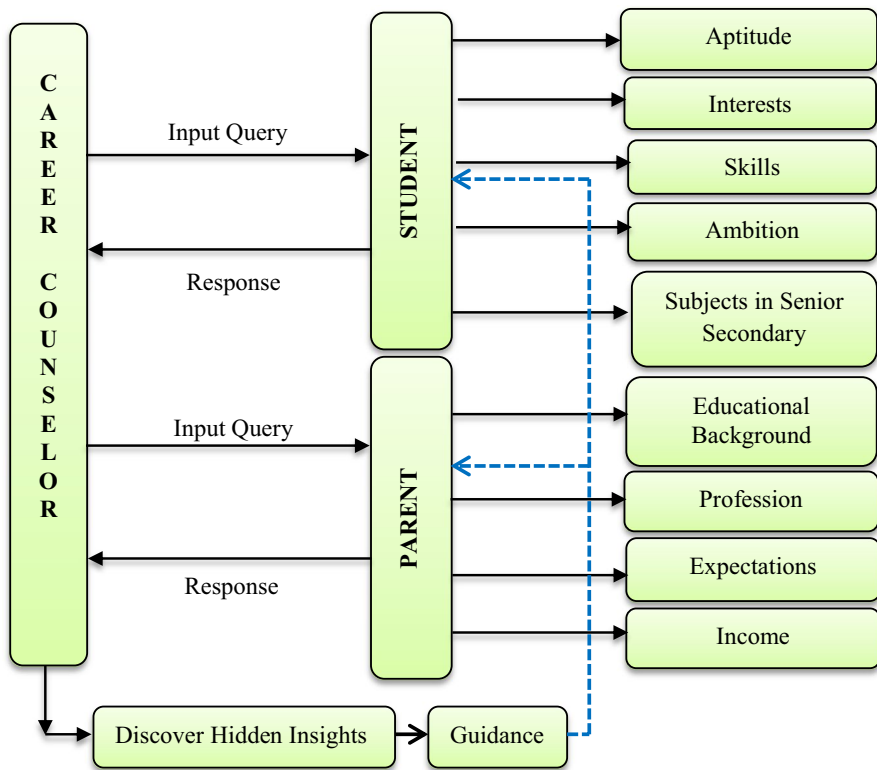
Fig. 1 Standard Approach of Career Counseling

effective learning outcomes. (Khan et al., 2021) have proposed a conceptual framework for attribute selection and predicting student performance using ML models. ML is implemented in the automatic assessment of students learning using responses, simulations, educational assessment, etc. (Zhai et al., 2021). Iatrellis et al. (2021) have implemented the K-Means clustering algorithm and discussed the potential of the clustering-aided approach for predicting student outcomes in higher education. During COVID-19 crises, remote learning has facilitated the students in achieving higher education and ML techniques have been implemented for predicting student satisfaction with the online learning models (Ho et al., 2021). In pandemic crises time, learning management systems integrated with technologies like AI are helpful in data analysis, improved learning, and effective resource management etc. (Villegas-Ch et al., 2020). ML-based frameworks along with simulation modeling are proposed by authors (Iatrellis et al., 2020) for designing learning pathways and evaluating quality assurance indicators in higher education. ML approaches like artificial neural networks are effective in the classification of various educational outcomes e.g. grade point average of students, academic retention and degree completion outcomes etc. (Musso et al., 2020).

1.1 Explainable Artificial Intelligence (XAI) and Black Box models

Explainable AI (XAI) is a term coined by Artificial Intelligence. XAI is one of the emerging fields which are trying to make the ML models especially Black Box models in a more understandable form for the users. XAI is important and focuses on ML models' transparency, interpretability, accountability, and final explanation of models' decisions and results. In machine learning, the models are divided into the White Box and Black Box models. The White-Box models can be interpreted as follows the approaches like inductive logic programming, rule learners, etc. The Black Box models are opaque especially the deep neural networks. A White Box Model inner working and steps are easily interpretable and are clear compared to the Black Box Models. However, the accuracy level of Black Box Models is more compared to the White-Box Models but the results of the Black Box Models are hard to interpret. Example of White-Box models includes simple Decision Trees, Linear Regression Models, Bayesian Networks etc. (Pintelas et al., 2020). The ML models' predictive results are divided into 2 subclasses i.e. global interpretability and local interpretability. In global interpretability, there is an understanding of the whole logic of the model and follows the entire reasoning of the possible outcomes (Linsley et al., 2018; Seo et al., 2017) whereas, in local interpretability, reasons are explained for a specific decision or single prediction that is interpreted locally (Molnar, 2021).

A Recursive Partitioning technique for global model interpretation is proposed by authors (Yang et al., 2018). Valenzuela-Escárcega et al. (2018) have proposed a supervised learning approach for information extraction whereas the activation maximization technique is implemented by (Erhan et al., 2010). The technique for local interpretable model agnostic explainability for local approximation of black box model is proposed by Ribeiro et al. (2016) whereas another technique is called Leave-One-Covariate-Out (LOCO) is implemented by Lei et al. (2018) for generating local explanation models.

XAI and ML are working closely with each other to develop the systems which properly explain the results and give answers to questions mainly asked by users: Why, Whom, and How?

The Black Box models are complex and the results obtained from the Black Box Models like neural networks are not easily understandable and interpretable. As current AI systems are ML-centric, the ML models are opaque, non-intuitive, and difficult to understand.

Explainable AI and ML are essential for customers to trust and accept AI-based applications. The major difference between Black Box AI and XAI is that, in the case of Black Box AI, there are decisions, recommendations but a lot of confusion to users whereas in the case of XAI, the decisions generated by ML models have proper explanations and feedback from users are also collected.

The Black Box Models like deep neural networks are attaining popularity, as in certain tasks, these models are performing better than human beings but they have disadvantages that they lack explicit declarative knowledge and decisions are not understandable, explainable to humans (Holzinger, 2018). ML model explainability is the primary and the foremost requirement as it builds trust among the users and as a result, users adopt AI-based applications (Gade et al., 2020).

XAI and ML have potential in the educational sector. AI-enabled Adaptive learning systems are increasingly utilized in educational systems (How, 2019). Educational data mining is a branch of AI and is used to analyze educational data. EDM helps to extract the hidden information from an available dataset and construct the model to predict the trends related to educational data, student outcomes. XAI systems need to be developed such that, humans understand the predictive results in a better way to better help the students and help them in effective decision making for students' carrier (Alonso & Casalino, 2019).

The XAI-based system must leverage the explanation for improving the performance of the ML model which includes improvement through feature selection, feature engineering, modifying the models' architecture, hyperparameters are taken, increasing the iterations, running the trained model on test data (Rai, 2020). The improvisation in the ML model also involves cross-validation, splitting the ratio of data into training and test data. The ML algorithms like decision trees, bayesian classifiers, linear models, etc. are inherently interpretable models compared to deep learning models which are having complicated structures, learning mechanisms and are not easily understandable to users (Hall & Gill, 2019). The results obtained from algorithms like Decision trees, LR models, Bayesian networks, etc. are traceable and transparent in their decision making. These algorithms use a restricted number of internal components like paths, rules, or features. The deep learning algorithms are being employed in areas where high-dimensional input size is there, speech recognition, object detection using convolutional neural network, face recognition system, natural language processing (NLP), etc. but the deep learning algorithm results cannot be easily interpretable to humans and neither transparent. Authors (Hall & Gill, 2019; Ribeiro et al., 2016), have proposed classification of XAI techniques for deep learning models. The techniques are classified by them into two dimensions which are described as:

- whether the technique is model specific or model agnostic?
- whether the technique is designed to explain, that is global in scope to the model or one that is local in scope to predictions?

1.2 Stages of Explainable AI (XAI) and framework

There are 3 stages of the XAI i.e. Pre-Modelling, Explainable Modelling, and Post Modelling stage. The detailed stages and XAI process is shown in Fig. 2.

Pre-modelling In the Premodelling stage, the dataset is explored and its prior understanding is crucial. Therefore, visualization of the dataset is also important to apply the statistical techniques and get the information like Mean, Standard Deviation, etc. and help in making the vigorous model.

Explainable modelling It is a stage where an explanation is a part of their inherent property. In this stage, the model generates the predictive results after analyzing
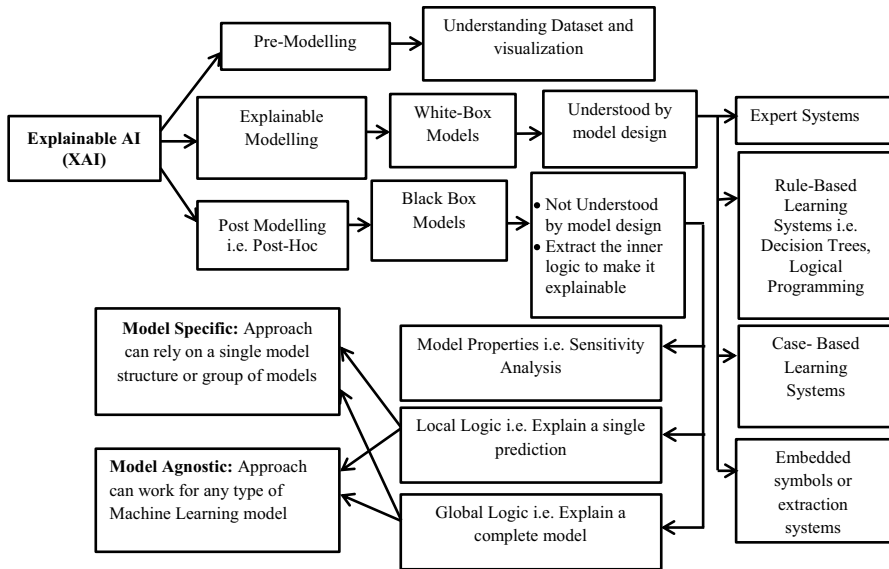
Fig. 2 Explainable AI (XAI) White Box and Black Box Model stages

the dataset and the model design is itself understood. The results are also trusted by the users and this stage consists of White Box Models.

Post-modelling stage In this stage, post justification of the predictive results by the model is given. In this, the results are not self-explanatory and are not understood by the model design. The White Box model design includes 4 main classes i.e. expert systems, rule-based learning system (Guidotti et al., 2018), case-based reasoning, embedded symbols and extraction (Mao et al., 2019). Black-box models are not explainable by themselves and therefore, techniques like model properties, local logic, global logic, etc. are adopted to make the black-box model explainable from the internal logic or the output of the model.

The Post-Hoc i.e. Post modelling stage is divided into local and global techniques followed by model agnostic and model specific approaches. In the model agnostic approach, there is an explanation for every type of machine learning model. In global logic techniques, there is an explanation of a complete model whereas, in the local logic technique, there is an explanation of a single prediction. An example of global logic is decision tree which approximates the prediction function $f(x)$ and for the local methods, there is a XAI framework called LIME (Local Interpretable Model Agnostic Explanations). The greedy function approximation i.e. a gradient boosting of regression trees (Friedman, 2001), individual conditional expectation (ICE) for visualizing statistical learning with plots of an individual conditional expectation (Goldstein et al., 2015), classical partial dependence plots (PDP), ALE (Accumulated Local Effects) plots shows the interaction between features and effect of a feature selection on prediction function $f(x)$ of a machine learning model (Apley

& Zhu, 2020) etc., follows the global techniques whereas local techniques rely on frameworks like Shapley values using game theory to explain the prediction of a model (Shapley, 1953) and SHAP (Shapely Addictive Explanations), an approach to interpret model predictions and extract local explanations of model instance predictions (Lundberg & Lee, 2017).

Apart from it, there is an another XAI framework i.e. known as what-if tool which is an interface for understanding the classification and regression problems related to Black Box Models. A toolkit known as AIX360 i.e. AI explainability 360 developed by IBM and supports model interpretability.

1.3 Research questions

A framework for career counseling of students using ML and AI techniques is proposed in the study that assists them in college-to-work transformation. There are following research questions related to XAI which are developed based on the empirical research literature and are harnessed to guide this research:

RQ1: Why ML model generates such a decision and why not something else?
RQ2: How much iteration is performed by the ML model to get success and come to the desired decision?
RQ3: What is the supportive evidence which helps the user to trust the results?
RQ4: Are all models in all defined-to-be-interpretable model classes equally interpretable (Doshi-Velez & Kim, 2017)?
RQ5: Are explanations always important (Bunt et al., 2012)?
RQ6: Why does accuracy generally require more complex prediction methods (Breiman et al., 2001)?
RQ7: Whether predictions are more important rather than their explanations (Yarkoni & Westfall, 2017)?

These questions are mainly targeting the Post- Modelling stage, where the inner logic of the black box models needs to be explored to know the explainability of the output predictive results.

1.4 Motivation

In the competitive environment, unfortunately, every student after completing their prerequisite degree doesn't get the right placement or career outcome. Therefore, career guidance has always been important for students because career counseling becomes a decision-making step from which they can give shape to their career and also helps them to stress those factors which may lead them to opt the right decision for enhancing their chances of employment and get skilled. In such a scenario, the career counseling objective is to make the right decision for students and increase skill-oriented approach among the students by opting for the right course curriculum. The right career counseling results in the dissemination of new academic trends and fills the gap between formal and non-formal sectors.

2 Related Work

The facts and searching are used in the Expert systems for effective decision making. A web-enabled career counseling system for students of Nigerian is developed to seek guidance for taking admission in any course (AlaoKazeem & IbamOnwuka, 2017). In expert systems, the knowledge base is used for logic and case-based reasoning systems (Aamodt & Plaza, 1994). Academic evaluation and students' academic adaptation are necessary for career counseling (Alexitch & Page, 1997). Career planning is important for students undergoing senior secondary education (Bardick et al., 2004; Witko et al., 2005). The outcome indicators for successful career development require a) career learning, b) development, and advancement (Bimrose et al., 2005). The right career choice is the foundation for success in life. Effective career counseling is a cycle that involves parents, their wards, teachers, counselors, and students, etc. The right career depends on the right career choice and also involves decisions from schooling itself. In maximum cases, students and parents come in an impression of oral marketing and choose the wrong career without judging their children's capabilities and interests.

With a blend of artificial intelligence and machine learning, the proposed work helps students make well-informed career decisions. Students are engaged with regular career updates and well-researched content. Advanced career guidance increases students' awareness of diverse career paths and universities. As a result, they pursue new-age careers and dynamic courses, across a more diverse range of institutes. The objective of the right career counseling is to motivate the students to enhance their skills and work upon their scholastic, familiar, passionate, and personal development. Career counseling makes educational opportunities available and accessible to everyone by rationalizing them. Practice exercises are helpful to students in an online education system. (Huang et al., 2019) proposed a deep reinforcement learning framework for recommending suitable learning exercises to students and increasing interaction by receiving students' performance feedback. ML and Educational data mining techniques are effective ways of predicting learner performance. A model is proposed using three ML algorithms i.e. decision tree, naive bayes, and neural network for predicting students' performance (Mimis et al., 2019). Qazdar et al. (2019) conducted a case study on students of high school in Morocco and presented a framework for predicting the students' performance using ML techniques.

Data mining algorithms namely logistic regression, random forest, and artificial neural network are implemented for early detection of potential difficulties faced by University students (Hoffait & Schyns, 2017). Goga et al. (2015) have designed a framework of an intelligent recommender system for improving the students' performance and predicted the first-year academic performance.

Similarly, intelligent tutoring system is devised using machine learning techniques for identifying the students who face problems when attempting homework exercises (Abidi et al., 2019).

The author Bilon (2013) has discussed current trends in research and theory for career counseling. Fritz (1997) has defined the AI system as a computer

programmed system that performs just like a human brain but it differs in working. According to Fee and Holland-Minkley (2010), there should be good pedagogy so that students improve their ability to analyze and solve complex computational problems independently. Gorad et al. (2017) have discussed the perspective of data mining for career counseling whereas Hendahewa et al. (2006) have mentioned an artificial intelligence approach for effective career guidance. An AI-based system is used as a carrier guidance cell which helps students in selecting the right course for bachelor's degree courses after senior secondary. Technology has an impact on business change and its significance for career selection (Hoyt, 1987). Authors have worked upon the expert system and correlated the relationship between problem space and decision-maker (Jackson, 1998; Russell & Norvig, 2010). The problem space means a lack of career guidance to students whereas the decision-maker is the student itself who is deciding the right career choice. A study has been conducted by authors for a private university in Thailand and based on the study, developed an intelligent recommendation system. The system discovers the fact that student results are correlated with student history (Kongsakun et al., 2010a, 2010b). Author (Kongsakun et al., 2010a, 2010b) has proposed an intelligent recommendation system framework for managing students' data and derives useful relations from it. Agent-based systems help develop an expert system whereas Kjellin and Boman (1994) have stressed the need for machine learning systems and knowledge acquisition algorithms for vocational training. Authors (Loan & Van, 2015) have proposed a system where data is collected, user modeling is done, and then using an intelligent system, information is derived which includes course information and useful training courses. AI uses computer programs to train and make intelligent machines. It uses data mining and machine learning models to make a machine intelligent (McCarthy, 2019). A typical example of an artificial intelligence system is an expert system (ES). The intelligent system can be referred to as any system with artificial intelligence. Career counseling is important for students completing the final semester of their degree course and students completing senior secondary face problems related to the right career choice (Morgan & Ness, 2003). Mihaela and Cristina (2014) have researched educational counseling and career guidance in Romania. Norasiah et al. (2003) have designed the software to manage the students' academic activities using an intelligent student information system. Another important aspect of career counseling is that it predicts the latest trends of industry from which both teachers and parents both are unaware. The student's interest, passion, and ability to learn are also important for enhancing quality education. The author has described a study conducted in Australia among 800 14-year-old adolescents and their purpose was to investigate the value orientations behind students' occupational choices, influences that will affect students' choices, and the attitudes of students toward vocational guidance (Poole & Juchnowski, 1974). Authors have proposed a model for student career counseling using artificial intelligence. The random forest and linear regression machine learning algorithms are used for the proposed work (Pujari et al., 2019). To achieve effective career prospects, course selection is the most important and decisive factor for students undergoing higher education (Saraswathi et al., 2014). Sodhi et al. (2016) have discussed

the efficacy of machine learning techniques and artificial neural networks to assist in taking decisions related to career counseling. Sun and Yuen (2012) have discussed career counseling projects in Chinese universities. According to them, suggestions and feedback are very important factors for further improvement and research on career counseling platforms. Clustering is also an effective approach to extract meaningful information from the data related to career counseling. A model based on classification techniques using past cases from the student database to predict the likely student's grade point average of the prospective student and current students. Watts (1986) has examined the transformation of traditional career counseling with the help of computerized tools, architectures, AI techniques, and algorithms whereas Watkins (n.d.) has proposed an architecture that assists students in the right course selection for the prospective future.

The ML techniques and approaches followed by the different authors are concluded in Table 1.

From the academic studies and the work conducted by the different researchers, the prediction of the students' performance, the dropout rate from the course or the retention rate in a course, etc. is done. The factors like employability, career placements, subjects undertaken in higher, senior secondary are not determined and predicted in the majority of the studies. the predictive results are shown using ml techniques but the inner logic, results are not interpreted.

The different educational attributes are taken by researchers in the datasets but the feature extraction methods are not discussed in the results in most of the studies. The absence of feature extraction results in downgrading the performance of the machine learning models (Cai et al., 2018). There is insufficient information related to the XAI and Black Box Models in the research studies. The predictive results obtained on the different educational datasets are not interpreted for the Black and White Box Models. The correct predictions are discussed but the incorrect predictions, overfitting; underfitting factors, learners, splits, iterations, etc. are missing which lacks the trust of users on the results generated. There are some unanswered questions from the research studies like ML techniques implemented by authors are model agnostic or model specific, the predictive results obtained are globally interpretable or locally interpretable etc.

3 Proposed framework using machine learning techniques

The proposed career counseling-based educational framework is integrated with ML techniques that can help in finding the career interests of students. The primary requirement of the framework is the dataset over which the ML models get trained. The input data desired from students and parents can be obtained in collaboration with educational institutions, academic organizations, employment offices, or local administrative bodies. The data collected from students is the major component because some students having an interest in maths might not select science as a stream rather they may choose commerce, arts as a future career option. The framework proposed in Fig. 3 acts as an advisor to both parents, students and guides them in the right selection of courses. The AI-based multi-agent systems find a correlation

Table 1 Machine Learning Techniques and Approaches Followed by Authors related to Career Counseling in Educational Sector

Reference	ML Techniques/Approach Followed	Description of the Work Done
Abidi et al., 2019	Machine Learning Techniques	Intelligent tutoring system for identifying the students who face problems when attempting homework exercises
AlaoKazeem & IbamOnwuka, 2017	Expert systems	Web-enabled career counseling system for students of Nigeria to seek guidance for taking admission in any course
Al-Sudani and Palaniappan, 2019	Artificial Neural Network	Finding low-performing students at an early stage of the semester
Goga et al., 2015	ML Framework	An intelligent recommender system for improving the students' performance and predicted the first-year academic performance
Hendahewa et al., 2006	Artificial intelligence	Artificial intelligence approach for effective career guidance
Ho et al., 2021	ML Model	Investigated the satisfaction level of undergraduate students using remote learning in higher education
Hoffait & Schyns, 2017	Logistic Regression, Random Forest, and Artificial Neural Network	Early detection of potential difficulties faced by university students
Huang et al., 2019	Deep Reinforcement Learning Framework	Suitable learning exercises to students and increasing interaction by receiving students' performance feedback
Hussain et al., 2018	Decision Tree	Observing the students who express low-engagement during assessment activities
Iatrellis et al., 2021	K-Means Clustering Algorithm	Clustering-aided approach for predicting student outcomes in higher education
Jishan et al., 2015	Naïve Bayes, Decision Tree, and Artificial Neural Networks	Forecast students' final result before the final exam
Kausar et al., 2020	Ensemble techniques	Examine the relationship between students' semester course and final results
Khan et al., 2021	Artificial Neural Networks	Proposed a conceptual framework for attribute selection and predicting student performance
Mimis et al., 2019	Decision Tree, Naive Bayes, and Neural Network	Predicting students' performance
Musso et al., 2020	Artificial Neural Networks	Grade point average of students, academic retention and degree completion outcomes
Pujari et al., 2019	Random Forest and Linear Regression	Student career counseling using artificial intelligence

Table 1 (continued)

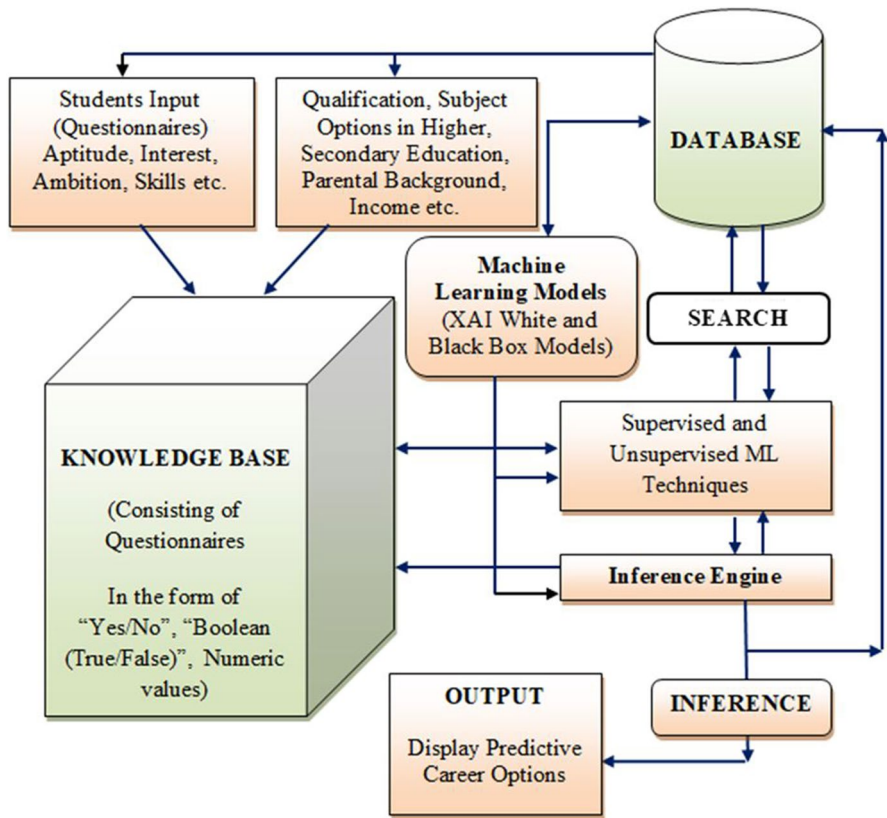| Reference	ML Techniques/Approach Followed	Description of the Work Done
Villegas-Ch et al., 2020	Artificial Intelligence and Data Analysis	Integrating machine learning and data analysis in learning management system
Zhai et al., 2021	ML Framework	An automatic assessment of students learning using responses, simulations, educational assessment

Fig. 3 Machine learning framework for career counseling

between student interests, his/her aptitude and accordingly suggest them for the course to pursue. The ML techniques fit and predict the mean accuracy of classification models. The framework is supposed to work like an expert system involving the combination of ML and AI for problem-solving. The system implements the knowledge engine and generates a combination of many results derived from the historical data. The expert system like MYCIN has been proved successful in areas of healthcare for patient diagnostics, fault-tolerance, data analysis, and measurement (Bratko, 2001). The expert system helps in generating valuable results with less time, better accuracy, and fewer economic measures. In Fig. 3, the inputs in the form of attributes are fetched from the dataset and the Machine Learning techniques shown in Table 2 are being followed to develop a framework. The input of students and parents in the form of questions and their answers are stored in a database as a knowledge base and the machine will learn the system, generate the outcome after applying the multiple combinations to generate the right output.

The ML models taken in the study are trained over the educational dataset shown in Table 3. They learn from the experiences and derive the results from

Table 2 Mathematical description of ML algorithms taken in a study

SrNo	Classifier Name	Description
1	K-Nearest Neighbor	K-Nearest Neighbor is implemented for classification and regression. Here, the new point is classified based on the nearest distance to the point. It calculates the point based on the two measures i.e. similarity and distance. KNN calculates the similarity based on the Euclidean and Manhattan distance functions as shown in Eqs. 1 and 2 Euclidean Distance: $d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$ (1) Manhattan Distance: $d(x, y) = \sum_{i=1}^{n}|x_i - y_i|$ (2)
2	SVM	SVM is known as a Support vector machine. It is also a supervised Machine Learning algorithm as it works on labeled training data. In SVM, the separating Hyperplane approach is followed and it categorizes the plane into two parts. The separating Hyperplane divides the attribute into two classes i.e. one class to one side of Hyperplane and second class to another side of Hyperplane. In the case of Linear SVM, for a given training dataset of n points of the form as shown in Eq. 3, $x_1, y_1, \ldots \ldots, x_n, y_n$ (3) where the y_i are either 1 or −1, each indicating the class to which the point x_i belongs. Each x_i is a p dimensional real vector. Any hyperplane can be written as the set of points x satisfying $w^t x - b = 0$, (4) where wi s the normal vector to the hyperplane. The Hyperplane can be described by Eqs. 5 and 6 $w^t x - b = 1$, (5) which means anything on or above this boundary is one class whereas anything on or below this boundary is another class $w^t x - b = -1$(6)
3	Naïve Bayes	Naïve Bayes is based on the probability of events and performs predictions. The Naïve Bayes is used to predict the probability of data belonging to an input class based on prior knowledge i.e. given data. It equates a Posterior Probability as shown in Eq. 7. $P(i)$ is the prior probability i.e. known from the past events and $P(X)$ is the predictor prior probability, that X will be observed. In the equation, $P(X|i)$ is the prior knowledge and $P(i|X)$ is the posterior probability of i, for given value X $P(i|X) = \frac{P(X|i)P(i)}{P(X)}$ (7)

Table 2 (continued)

SrNo	Classifier Name	Description
4	J48	J48 is a class for generating a pruned or unpruned C4.5 decision tree, first proposed by Quinlan (1993). It is a decision tree and if–then conditions are used to predict the target variables values based on the other independent variables on different nodes of a tree. It is also a supervised Machine Learning model which makes predictions on hidden data. In the Decision Tree, the attribute having the highest information gain is considered as the root node and the dataset will split on this attribute to perform predictions. The other nodes of the tree are the decision nodes and if there is no further split on any node, such nodes are the leaf nodes. It calculates the impurity using Gini Index and Entropy. The Gini Impurity Index and Entropy, both perform the classification task. Equation 8 shows the Gini Index (GI) $$GI = \sum_{i=1}^{m} -x_i\,(1 - x_i),\ (8)$$ where x_i is the frequency of the attribute i at a node and n is the number of unique attributes and, the Entropy (E_t) is calculated as shown in Eq. 9 $$E_t = \sum_{i=1}^{n} -f_i \log(f_i)\ (9)$$ The pruned Decision Tree is shown in Table 9 and the Decision Tree is visualized in Fig. 10

Table 2 (continued)

SrNo	Classifier Name	Description
5	Ensembling Methods	AdaBoost is the Ensembling technique to boost up the performance of the decision trees and is used mainly for classification. Random Forest first proposed by Ho (1995) are ensembling methods and perform classification, regression tasks on decision trees during run time and output the class, that is either the mode of the classes i.e. classification or average prediction of individual trees i.e. Regression(Ho, 1998). Random Forest overcomes the problem of overfitting of training data as in the case of decision trees (Hastie et. al., 2008). The bagging algorithm proposed by Breiman (1996) is also known as bagging predictors or Bootstrap Aggregating and it calculates the aggregate of multiple versions of a predicted model. In the case of regression, an average is taken over all the outputs predicted by the individual learners as shown in Eq. 10 (https://blog.paperspace.com/bagging-ensemble-methods/amp/), whereas in classification, either the most voted class is accepted i.e. hard-voting or the highest average of all the class probabilities is taken as the output i.e. called soft-voting or aggregation. Bagging reduces variance and avoids overfitting problems. Given a training set for two classes, $S = \{(x_1, y_1), \ldots, \ldots, x_n, y_n)\}$. A Machine is trained on each S_i, i = 1 …to T samples and obtains a sequence of T outputs $f_1(x), \ldots, f_T(x)$ $\bar{f}_{bag} = \bar{f}_1(x) + \bar{f}_2(x) + \cdots + \bar{f}_b(x)$ (10) Here, $\bar{f}_{bag}$ is the bagged prediction and $\bar{f}_1(x), \ldots, +\bar{f}_b(x)$ are the individual learners The final aggregate classifier for regression is shown in Eq. 11 $\bar{f}_x = \sum_{i=1}^{T} f_i(x)$, (11) Here, x is the point and the average of $\bar{f}_i$ for i = 1…….T; The final aggregate classifier for classification is shown in Eq. 12 $f(x) = sign(\sum_{i=1}^{T} f_i(x))$ (12)
6	Association Rule Mining	Association Rule Mining is a rule based machine learning and derives the interesting relations between attributes in databases (Piatetsky-Shapiro, 1991). In Association Rule Mining, support and confidence for the frequently occurring item sets are derived from the dataset and displayed in the form of Key-Value pair

Table 3 Sample Dataset

sl_no	gender	ssc_p (Secondary Education %)	ssc_b (Educational Board Secondary)	hsc_p (Higher Secondary %)	hsc_b (Educational Board Higher Secondary)	hsc_s (Subjects in Higher Secondary)	degree_p (% of marks in degree)	degree_t (Subjects in Under Graduate Degree)	Workex (Work Experience)	etest_p (Employability Test)	Specialisation Post-Graduation(MBA)	mba_p % obtained	Status Placement Status	Salary Salary offered
1	M	67	Others	91	Others	Commerce	58	Sci&Tech	No	55	Mkt&HR	58.8	Placed	270,000
2	M	79.33	Central	78.33	Others	Science	77.48	Sci&Tech	Yes	86.5	Mkt&Fin	66.28	Placed	200,000
3	M	65	Central	68	Central	Arts	64	Comm&Mgmt	No	75	Mkt&Fin	57.8	Placed	250,000
4	M	56	Central	52	Central	Science	52	Sci&Tech	No	66	Mkt&HR	59.43	Not Placed	NaN
5	M	85.8	Central	73.6	Central	Commerce	73.3	Comm&Mgmt	No	96.8	Mkt&Fin	55.5	Placed	425,000

the Inference Engine. The results obtained from the ML models are classified and their comparative metrics are shown in Table 8. The data mining techniques like association rule mining helps in discovering the relationships from the dataset and accordingly perform predictions. In the ML Process, which consists of supervised and unsupervised learning techniques, the XAI, Black Box Models get trained over the dataset. The advantage of the XAI model is that it is capable to explain the predictions performed by the model whereas in Black Box models, it becomes complex even for its designers to explain the results obtained (Sample, 2017).

The knowledge base shown in the proposed framework is the storage of information gathered through students and parents whereas inference engine is a component that applies logic rules to the knowledge base to infer new information. The framework could request further information concerning the student, as well as, the predictive results help students, to arrive at a probable decision, after which it would recommend a course of action for choosing the right course and career. If requested, the framework using Explainable AI (XAI) would explain the reasoning that led to its suggestion and recommendation based on the production rules. The results obtained using white box and black box models are interpreted in Tables 4 and 5. Artificial intelligence methods like Expert Systems (ES) can help and save time in this domain because an ES can provide a piece of fast expert advice based on the knowledge from its knowledge base component. The right career selection boosts the performance of the student, becomes skilled, and takes a keen interest to achieve excellence in that subject.

3.1 Machine learning techniques

The predictive analytics is performed using machine learning techniques. Machine learning techniques work in three ways. The techniques are: a) unsupervised, b) supervised and c) semi-supervised learning.

In Fig. 3, for developing a proposed framework, Machine learning techniques are implemented in a phased manner. In supervised learning, the majority of algorithms, the machine is trained using well-labeled data, inputs and outputs are matched. The process is to map function and takes inputs and matches to output, creating a target function. The subtypes of supervised learning are classification, regression. Example: Linear Regression, Random Forest, SVM etc. In unsupervised learning, there are unlabeled data (inputs only) is analyzed. Here, learning happens without supervision. The inputs are used to create a model of the data. The subtypes of unsupervised learning are clustering, association, etc. Example: K-Means, Hierarchical clustering. In semi-supervised learning, some data is labeled and some are unlabeled. The goal is to obtain better results from the labeled data and is good for real-world data. Example: self-training, mixture models, semi-supervised SVM. In the proposed work, experimentation is performed on supervised machine learning models. Machine learning techniques taken for experimentation are summarized along with their mathematical description in Table 2.

Table 4 Decision Tree, SVM, and Logistic Regression

Model Type	Kernel Function/ Ensemble Method	Accuracy	Prediction Speed	Training Time	Feature Selection	Enable Principal Component Analysis(PCA) or Not	Categorical Predictors Used or Not
Decision Tree	Fine Tree Maximum number of splits: 100 Split criterion: Gini's diversity index	63.7%	~ 2000 obs/sec	6.2652 s	All features used in the model, before PCA	Numeric predictors are used	All 7 categorical predictors are used in the model. PCA is not applied to categoricals
Linear SVM	Gaussian Kernel Function Multi-Class Method: One-vs-One	65.6%	~ 2200 obs/sec	2.4079 s	✓	✓	✓
Logistic Regression		68.8%	~ 1700 obs/sec	4.9942 s	✓	✓	✓

Table 5 Ensemblers and narrow neural network

Model Type	Kernel Function/Ensemble Method	Accuracy	Prediction Speed	Training Time	Feature Selection	Principal Component Analysis(PCA)	Predictors
Boosted Trees	Ensemble method: AdaBoost Learner Type: Decision Tree Maximum number of splits: 20 Number of learners: 30 Learning rate: 0.1	64.2%	~740 obs/sec	2.9058 s	✓	Numeric predictors are used	All 7 categorical predictors are used in the model. PCA is not applied to categoricals
Bagged Trees	Ensemble method: Bag Learner Type: Decision Tree Maximum number of splits: 214 Number of learners: 30	65.6%	~830 obs/sec	2.9635 s	✓	✓	✓
RUS Boosted Trees	Ensemble method: RUSBoost Learner Type: Decision Tree Maximum number of splits: 20 Number of learners: 30 Learning rate: 0.1 Split criterion: Gini's diversity index	58.1%	~920 obs/sec	2.8604 s	✓	✓	✓
Neural Network	Narrow Neural Network Number of fully connected layers: 1 First layer size: 10 Activation: ReLU Iteration limit: 1000	68.8%	~2800 obs/sec	2.841 s	✓	✓	✓

Table 6 Confusion matrix of White Box models

Model type	Class	Outcome variable	Predicted Class		TPR	FNR	PPV	FDR
			Not Placed	Placed				
Decision Tree	True Class	Not Placed	20	47	29.9%	70.1%	39.2%	60.8%
		Placed	31	117	79.1%	20.9%	71.3%	28.7%
SVM	True Class	Not Placed	0	67	0	100%	0	100%
		Placed	1	147	99.3%	0.7%	68.7%	31.3%
Logistic Regression	True Class	Not Placed	0	67	0	100%	0	0
		Placed	0	148	100%	0	68.8%	31.2%

TPR = True Positive Rate; FNR = False Negative Rate; PPV = Positive Predicted Value; FDR = False Discovery Rate

3.2 Experimentation performed

In the proposed work, the experimentation is performed using MATLAB R2021 Machine learning toolbox and Weka 3.8.5 software. The system configuration used for the study is Intel(R) Core(TM) i3-7100 CPU @ 3.90 GHz and the operating system is Windows 10 Pro. Supervised machine learning techniques are implemented on the sample dataset shown in Table 2 for performing academic and employability factor placement predictions as a part of the proposed career counseling model. Apart from it, the performance evaluation of the white and black box ML models is done on MATLAB and Weka. The information disseminated by both Models is summarized in Table 6 and Table 7 to make the results easily interpretable and explainable to the user. The "predictor" and "response" variables are selected from the dataset.

Table 7 Confusion matrix of Black Box models

Model type	Class	Outcome variable	Predicted Class		TPR	FNR	PPV	FDR
			Not Placed	Placed				
Boosted Trees	True Class	Not Placed	17	50	25.4%	74.6%	38.6%	61.4%
		Placed	27	121	81.8%	18.2%	70.8%	29.2%
Bagged Trees	True Class	Not Placed	17	50	25.4%	74.6%	39.5%	60.5%
		Placed	26	122	82.4%	17.6%	70.9%	29.1%
RUSBoosted Trees	True Class	Not Placed	37	30	55.2%	44.8%	38.5%	61.5%
		Placed	59	89	60.1%	39.9%	74.8%	25.2%
Neural Network	True Class	Not Placed	0	67	0%	100%	0%	0%
		Placed	0	148	100%	0%	68.8%	31.2%

3.2.1 Dataset

The dataset is available online and retrieved from the Kaggle website (https://www.kaggle.com/benroshan/factors-affecting-campus-placement). The dataset comprises both boolean and categorical data. For the experiment, datasets related to academia are in-depth analyzed, and accordingly, dataset is preprocessed by structuring it into a supervised ML format which improvised the performance of ML algorithms and precision of the prediction process. The total number of instances in a dataset is 215 and total 15 attributes are there. The sample dataset is shown in Table 3.

3.2.2 Methodology Adopted

The methodology adopted for applying ML techniques in the proposed framework works in a phased manner. The first step is to fetch the training data in the form of.csv,.arff or.xlsx files etc. After fetching the data, the preprocessing of data is done to remove the outliers. The dataset without preprocessing contains blank, missing, and NaN (Not a Number) values. The outliers are removed to boost up the accuracy of Machine Learning models.

Feature selection is an important phase to create a model. The predictor variables are selected so as to improve the prediction performance of the model (Miao & Niu, 2016; Xue et al., 2015). Neighborhood Component Analysis (NCA) feature selection is implemented in MatLab environment for classification and the supported data type are continuous features. The feature importance is estimated for distance based supervised models that pair's distance between observations to forecast the feedback. Supervised learning is used to train the model on the educational dataset shown in Table 3. ML model evaluation and the prediction results are discussed in Section 4. The Machine learning workflow for training the model on the dataset and performing predictions is shown in Fig. 4.
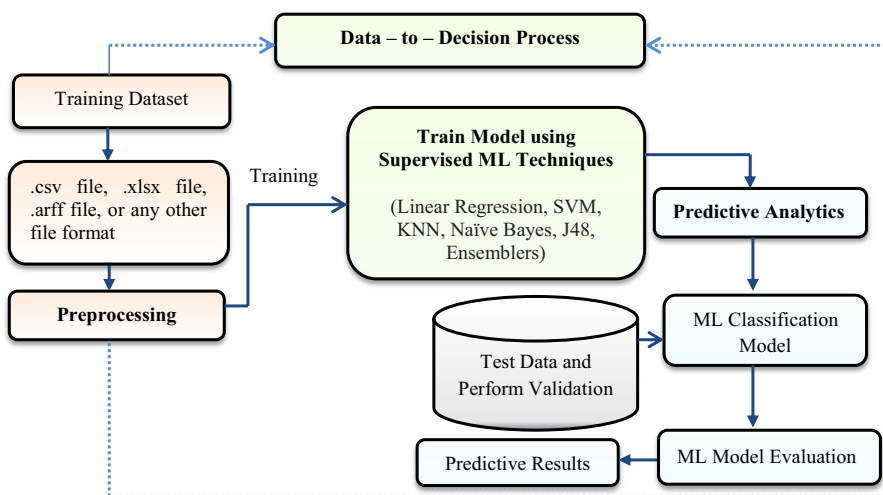

Fig. 4 Machine learning workflow for performing predictions

4 Results and Discussions

The Machine learning classifiers are trained on the sample dataset shown in Table 3. The 'PlacementStatus' attribute is considered as the response variable and other attributes are the predictor variables. The tenfold cross-validation is applied to protect against overfitting by partitioning the dataset into folds and estimating accuracy on each fold. The resampling technique i.e. K-fold cross-validation is used to estimate the skill of a model on the new data. The ML model once gets trained on the educational dataset, it will be capable enough to perform the predictions on the test data. Cross-validation is required when there are chances of overfitting. Machine Learning algorithms also deal with underfitting and overfitting problems to avoid the poor performance of the ML algorithms. The overfitting means, the ML algorithm learns the training data extraordinarily well and the machine learning model can't generalize or fit well on the unseen dataset. A clear sign of machine learning overfitting is if its error on the testing or validation dataset is much greater than the error on the training dataset (https://datascience.foundation/sciencewhitepaper/underfitting-and-overfitting-in-machine-learning). The under fitting means, the ML models perform poorly on the training data. In overfitting, the performance on training data increases, as it learns the irrelevant data. In both overfitting and underfitting, the output learned by the ML models is applied poorly on the new data. The performance is normally being the same on training and test data.

In the experiment, White box models i.e. Naive Bayes, Linear SVM are trained over the dataset. A naive bayes classifier uses the Gaussian distribution for numeric predictors and multivariate multinomial distribution (MVMN) distribution for categorical predictors whereas a Linear SVM makes a simple linear separation between classes using the linear kernel. In the ensembling technique, the model creates an ensemble of medium decision trees and takes less time, memory for computations.

The simpler interpretation of the Decision Tree, Linear SVM, and Logistic Regression results in MATLAB is shown in Table 4 and the results of Ensemblers, Neural Networks are shown in Table 5. The campus placement results of the students are displayed using Decision Trees and Bagging Ensembling methods only because, it is derived from the confusion matrix results of ML models displayed in Table 6 that, SVM and Logistic Regression White-box models have shown zero observations for 'Not Placed' predicted class and resulted into overfitting. On the other side, the Neural Network Model has also shown 100% True Positive Rate (TPR) for the 'Placed' predicted class in Confusion Matrix shown in Table 7.

The predictors taken for performing predictions related to the campus placement of students using ML trained models are 'ssc_p(Secondary Education)', 'hsc_p(Higher Secondary)', 'etest_p(Employability Test)', 'degree_t(Subjects in Under Graduate Degree)'.

The correct and incorrect predictions performed by Decision Tree model, on taking the attributes 'ssc_p', 'hsc_p' and 'etest_p', 'degree_t' are shown in Figs. 5(a), (b) and 6(a), (b) and the predictions performed by the Bagging ensembling techniques for the same attributes i.e. 'ssc_p', 'hsc_p' and 'etest_p', 'degree_t' is shown in Figs. 7(a), (b) and 8(a), (b). The findings show that the students who have secured above 60% in their Higher and Senior Secondary Examinations have shown good placement results.
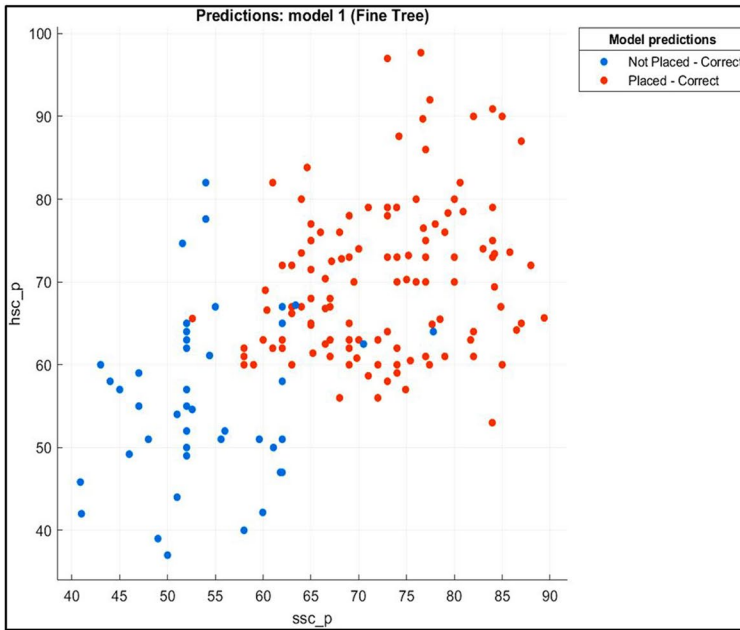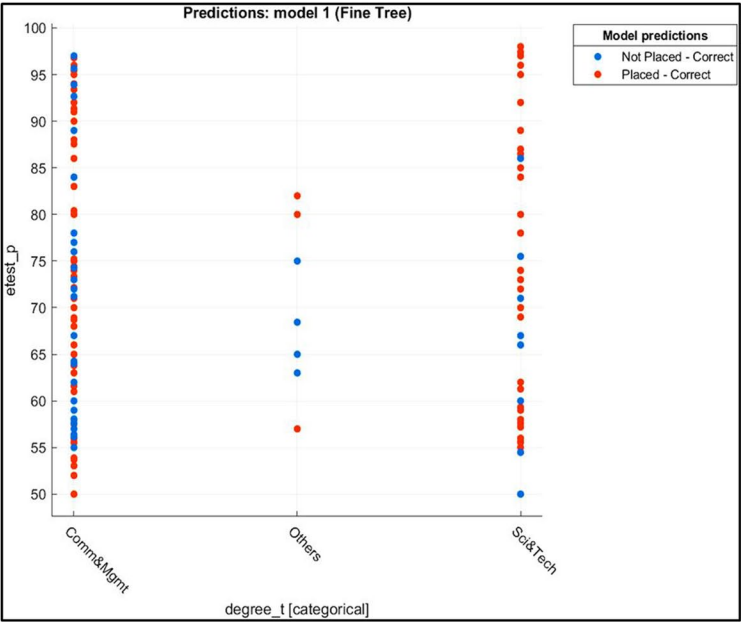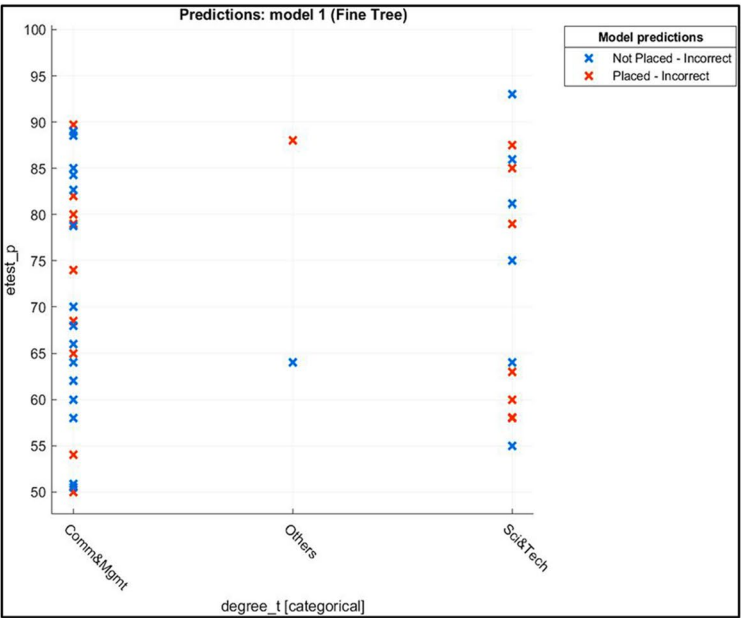
(a)
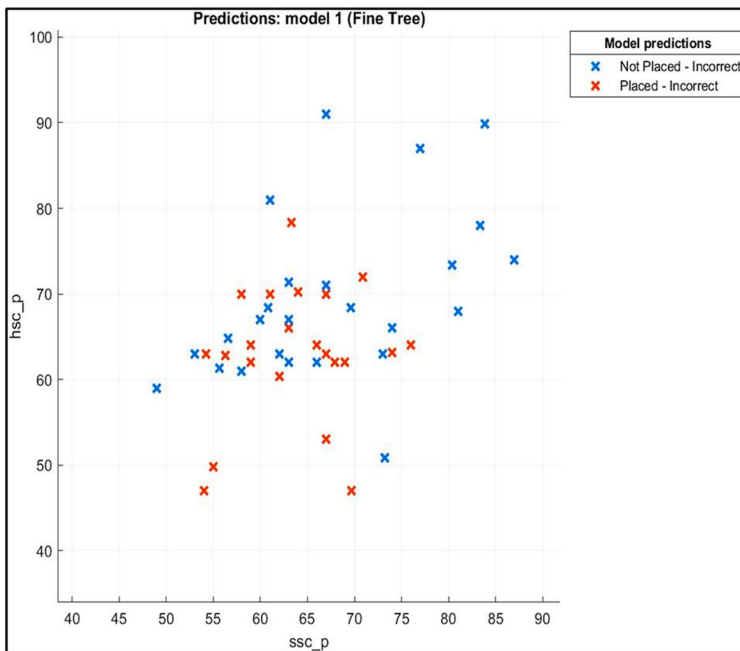
(b)

Fig. 5 **(a)** Decision tree correct predictions 'ssc_p' vs 'hsc_p'. **(b).** Decision tree incorrect predictions 'ssc_p' vs 'hsc_p'
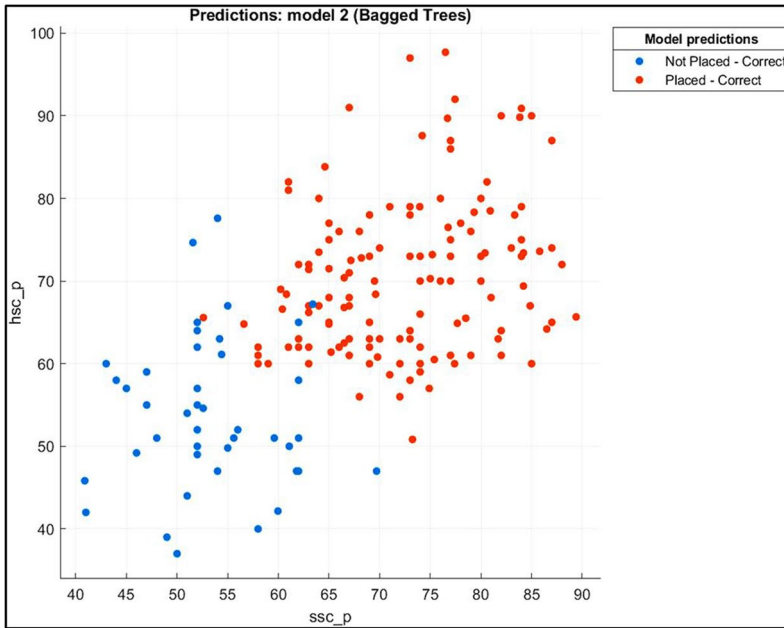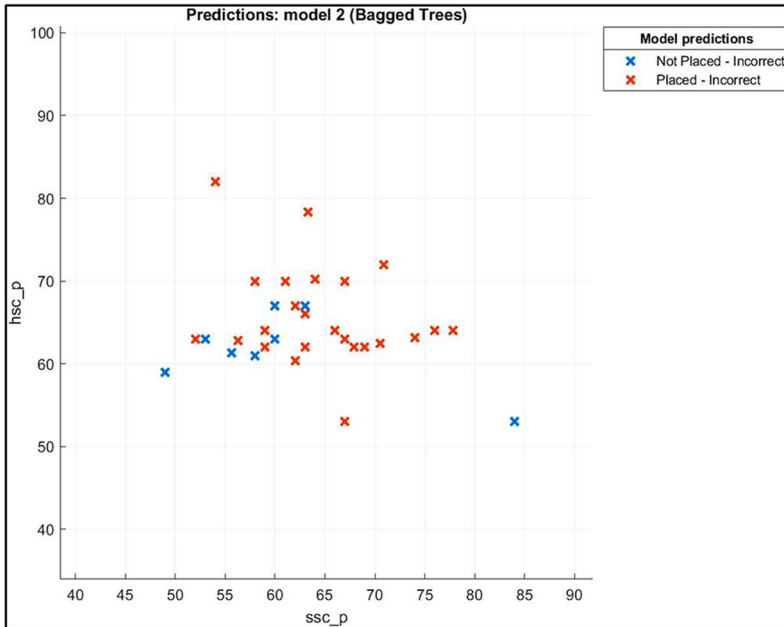
(a)

(b)

Fig. 6 **(a)** Decision tree correct predictions 'etest_p' vs 'degree_t'. **(b).** Decision tree incorrect predictions 'etest_p' vs 'degree_t'
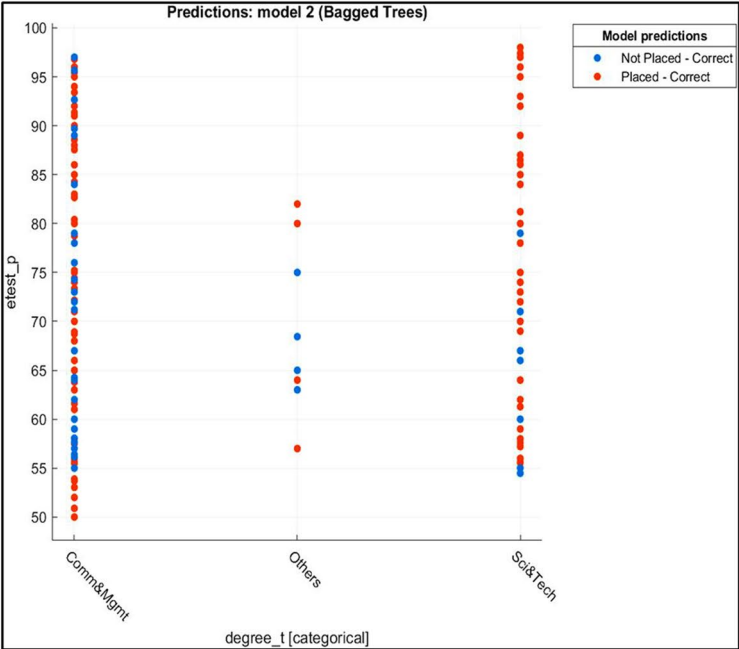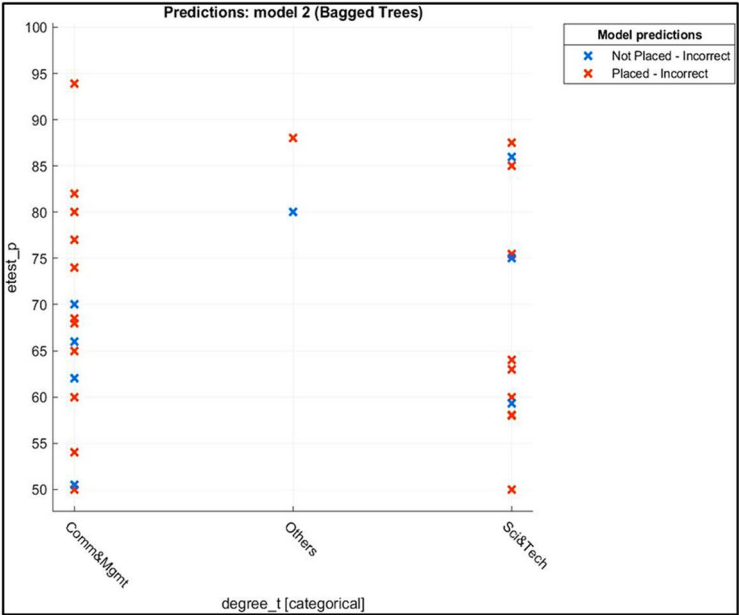
(a)

(b)

Fig. 7 **(a)** Bagging correct predictions 'ssc_p' vs 'hsc_p'. **(b).** Bagging incorrect predictions 'ssc_p' vs 'hsc_p'

(a)

(b)

Fig. 8 (a) Bagging correct predictions 'etest_p' vs 'degree_t'. (b). Bagging incorrect predictions 'etest_p' vs 'degree_t'

The results obtained from the ML model for attributes 'etest_p' i.e. Employability Test and 'degree_t' i.e. Subjects in undergraduate degree shows that the students having subject 'Science and Technology, Commerce and Management have got higher placements compared to students having another subject in Undergraduate degree. On training the models, the accuracy achieved by Fine Tree is 63.7% whereas the accuracy achieved by Linear SVM and Logistic Regression is 65.6%, 68.8%. The accuracy achieved by SVM and Logistic Regression is higher than the Decision Tree but both SVM, Logistic Regression have resulted in zero observations in the predicted class category. On the other side, the accuracy achieved by Boosted trees, Bagged, RUSBoosted Trees, and Narrow Neural Network is 64.2%, 65.6%, 58.1%, and 68.8%.

The accuracy obtained by Neural Network is higher than Ensemblers, but Narrow Neural Network has shown zero observations for the "Placed', "Not Placed' Predicted classes and the True Positive Rate value for 'Placed' class is 100% which shows that the model has resulted into the overfitting. The confusion matrix obtained in MATLAB R2021a of both the White Box and Black Box models featuring true positive rates, false negative rates, and the positive predictive values, false discovery rates are shown in Table 6 and Table 7 whereas, the ML Model metrics are compared in WEKA 3.8.5 and the results are shown in Table 8. The performance evaluation of the ML Models is shown in Fig. 9. It is derived from the Fig. 9, that Naïve Bayes has correctly classified the highest number of instances and the F-Measure, Recall values are highest in it.

It is derived from the results shown in Table 8 that the maximum number of instances is correctly classified by the Naïve Bayes. The root mean squared error of the Naive Bayes model is 0.2541 which measures the predictive performance of the model. The predictive results are globally interpreted using supervised machine learning techniques. The Decision Tree shown in Fig. 10 has considered 'hsc_p' i.e. percentage in higher education as its root node and the node having the highest information gain.

TPR is the true positive rate and is also known as the sensitivity or recall metric of the ML Model. It measures the actual positive samples from the dataset which are correctly identified. TP rate means true positive which means samples are correctly classified and fall in positive class. The TP rate formula is shown in Eq. 13.

$$True\ Positive\ Rate\ (tpr)\ or\ Sensitivity\ or\ tp\ rate = \frac{tp}{tp + fn} \qquad (13)$$

FN means false negatives which show the samples are incorrectly classified as negative. FP means a false positive which shows that the samples are incorrectly classified as positive whereas TN means true negative which shows that the samples are correctly classified but fall into negative class. FPR is false positive rate and is also known as specificity (Wang and Zheng, 2013). The false positive rate is shown in Eq. 14.

$$False\ Positive\ Rate\ (fpr)\ or\ Specificity\ or\ fp\ rate = \frac{fp}{fp + tn} \qquad (14)$$

In Table 6 and Table 7, PPV and NPV are the metrics to show the percentage of predictive values. PPV means positive predictive value and NPV means negative predictive value. PPV and NPV are shown in Eq. 15 and Eq. 16.

Table 8 Comparative metrics of ML models

| Sr.No	ML Model Type	Correctly Classified Instances	Incorrectly Classified Instances	TP Rate	FP Rate	Recall	F-Measure	Root mean squared error	Accuracy	Time is taken to build the model (seconds)
[1]	J48(C4.5)	178	37	0.828	0.249	0.828	0.826	0.3923	82.790%	0.03
[2]	Bagging	180	35	0.837	0.262	0.837	0.833	0.3368	83.720%	0.03
[3]	KNN	159	56	0.740	0.559	0.740	0.676	0.5078	73.953%	0
[4]	Logistic Regression	188	27	0.874	0.204	0.874	0.872	0.3088	87.441%	0.29
[5]	Naïve Bayes	196	19	0.912	0.195	0.912	0.907	0.2541	91.162%	0
[6]	SVM	186	29	0.865	0.208	0.865	0.863	0.3673	86.511%	0.03
[7]	Random Forest	184	31	0.856	0.261	0.856	0.849	0.3258	85.581%	0.12
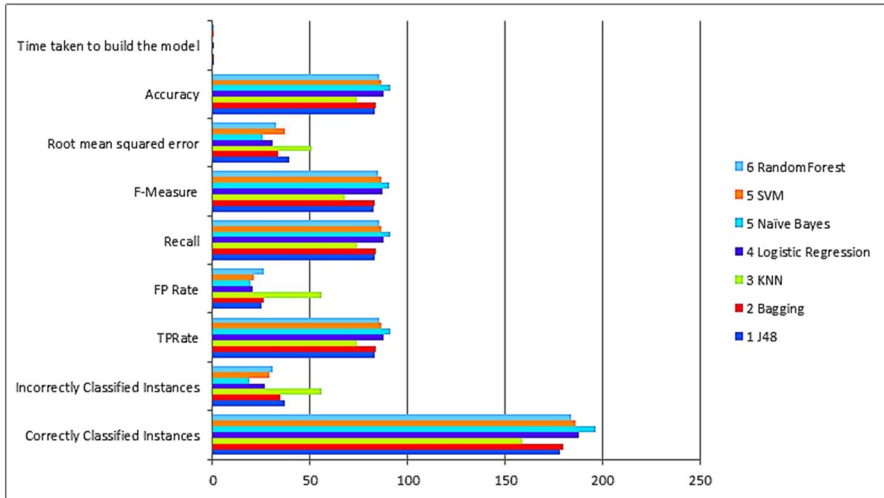
Fig. 9 Comparative evaluation of the ML classifiers
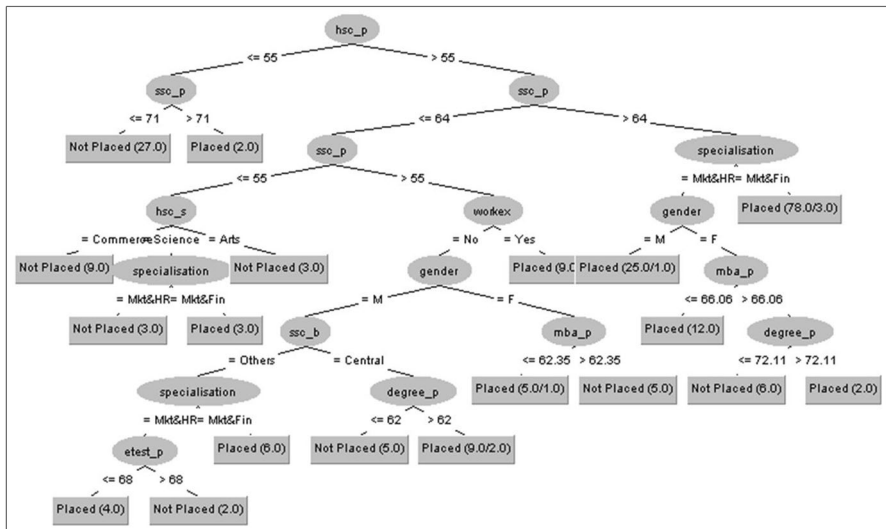
Fig. 10 Decision Tree

$$Positive\ Predictive\ Value\ (ppv) = \frac{tp}{tp + fp} \tag{15}$$

$$Negative\ Predictive\ Value\ (npv) = \frac{tn}{tn + fn} \tag{16}$$

The metrics true negative rate (TNR) and false negative rate (FNR) are shown in Eq. 17 and Eq. 18. TNR represents the samples that are correctly predicted to the negative class and in FNR the samples belonging to the positive class are predicted as negative.

$$True\ Negative\ Rate\ (tnr) = \frac{tn}{tn + fp} \qquad (17)$$

$$False\ Negative\ Rate\ (fnr) = \frac{fn}{fn + tp} \qquad (18)$$

FDR is the false discovery rate i.e. also known as type 1 errors in testing null hypothesis while performing multiple comparisons (Benjamini & Hochberg, 1995). FDR is the ratio of false positives to the sum of false positives and true positives. False discovery rate is shown in Eq. 19.

$$False\ Discovery\ Rate\ (fdr) = \frac{fp}{fp + tp} \qquad (19)$$

5 Conclusion and future scope

In the present competitive environment, the right career counseling is the need of an hour. The right career choice is very important for those students who are pursuing their education and are in their final semester courses. Career counseling helps students in analyzing their expectations, aptitude, skills, and educational interests for appropriate course selection. Effective career guidance tries to advise pupils for effective decision-making for the right course orientation and skill-based learning. In the majority of cases, students take enrollment in courses without proper career counseling and parents also come in the wrong impression of oral marketing. The students take admissions into courses without understanding their capabilities, interests, skills, and goals. Apart from it, career counseling is also important for recruiters to choose the right skill. Although there are portals for career counseling, machine learning-based career counseling applications are still least explored. Machine learning is the application of artificial intelligence where the machine will access the data and learn from them. In the proposed work, predictive analytics is performed using the white box and black box machine learning techniques. Apart from it, in-depth analysis of the White and Black box model techniques is explored for proper explainability and interpreting AI results on a Career counseling-based educational dataset. The accuracy of different machine learning algorithms is compared to perform predictions. The best machine learning algorithm is implemented in the proposed career counseling framework for effective decision making.

Limitations of this study include that the dataset taken contains a limited scope of attributes and sample size. We intend to conduct a similar study with more traits, and socio-economic factors related to students counseling.

Machine learning models need to be trained on the educational datasets collected from the multiple educational institutes with more attributes for assessment of the students and their placement factors need to be considered for employability enhancement and skill development. The proposed work can be further extended using unsupervised learning techniques and deep learning frameworks for training the models on educational datasets with more instances and attributes (Table 9).

Table 9 J48 pruned tree

```
hsc_p <= 55
|   ssc_p <= 71: Not Placed (27.0)
|   ssc_p > 71: Placed (2.0)
hsc_p > 55
|   ssc_p <= 64
|   |   ssc_p <= 55
|   |   |   hsc_s = Commerce: Not Placed (9.0)
|   |   |   hsc_s = Science
|   |   |   |   specialisation = Mkt&HR: Not Placed (3.0)
|   |   |   |   specialisation = Mkt&Fin: Placed (3.0)
|   |   |   hsc_s = Arts: Not Placed (3.0)
|   |   ssc_p > 55
|   |   |   workex = No
|   |   |   |   gender = M
|   |   |   |   |   ssc_b = Others
|   |   |   |   |   |   specialisation = Mkt&HR
|   |   |   |   |   |   |   etest_p <= 68: Placed (4.0)
|   |   |   |   |   |   |   etest_p > 68: Not Placed (2.0)
|   |   |   |   |   |   specialisation = Mkt&Fin: Placed (6.0)
|   |   |   |   |   ssc_b = Central
|   |   |   |   |   |   degree_p <= 62: Not Placed (5.0)
|   |   |   |   |   |   degree_p > 62: Placed (9.0/2.0)
|   |   |   |   gender = F
|   |   |   |   |   mba_p <= 62.35: Placed (5.0/1.0)
|   |   |   |   |   mba_p > 62.35: Not Placed (5.0)
|   |   |   workex = Yes: Placed (9.0)
|   ssc_p > 64
|   |   specialisation = Mkt&HR
|   |   |   gender = M: Placed (25.0/1.0)
|   |   |   gender = F
|   |   |   |   mba_p <= 66.06: Placed (12.0)
|   |   |   |   mba_p > 66.06
|   |   |   |   |   degree_p <= 72.11: Not Placed (6.0)
|   |   |   |   |   degree_p > 72.11: Placed (2.0)
|   |   specialisation = Mkt&Fin: Placed (78.0/3.0)


Number of Leaves:      19
Size of the tree       :      36
```

Declarations

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical approval This article does not contain any studies with human subjects and/or animals performed by any of the author.

Informed consent This article does not contain any studies with human participants.

References

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications, 7*(1), 39–59.

Abidi, S. M. R., Hussain, M., Xu, Y., & Zhang, W. (2019). Prediction of confusion attempting algebra homework in an intelligent tutoring system through machine learning techniques for educational sustainable development. *Sustainability, 11*(1), 105.

Al-Sudani, S., & Palaniappan, R. (2019). Predicting students' final degree classification using an extended profile. *Education and Information Technologies, 24*(4), 2357–2369.

AlaoKazeem, A., & IbamOnwuka, E. (2017). Development of a web-based intelligent career guidance system for pre-tertiary science students in Nigeria.

Alexitch, L. R., & Page, S. (1997). Evaluation of academic and career counseling information and its relation to students' educational orientation. *Canadian Journal of Counseling, 31*(3), 205–218.

Alonso, J. M., & Casalino, G. (2019). Explainable artificial intelligence for human-centric data analysis in virtual learning environments. *International workshop on higher education learning methodologies and technologies online* (pp. 125–138). Springer.

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (statistical Methodology), 82*(4), 1059–1086.

Bardick, A. D., Bernes, K. B., Magnusson, K. C., & Witko, K. D. (2004). Junior high career planning: What students want. *Canadian Journal of Counseling and Psychotherapy*, *38*(2).

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (methodological), 57*(1), 289–300.

Bilon, A. (2013). Career Counseling: Current Trends in Research and Theory.

Bimrose, J., Brown, J., & Barnes, S. A. (2005). A systematic literature review of research into career-related interventions for higher education.

Bunt, A., Lount, M., & Lauzon, C. (2012). Are explanations always important? A study of deployed, low-cost intelligent interactive systems. In Proceedings of the 2012 ACM international conference on Intelligent User Interfaces, (pp. 169–178).

Bratko, I. (2001). *Prolog programming for artificial intelligence*. Pearson education.

Breiman, L., Cox, D., & Breiman, L. (2001). Comment-statistical modeling: The two cultures. *Statistical Science, 16*(3), 199–231.

Breiman, L. (1996). Bagging predictors. *Machine Learning, 24*(2), 123–140.

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing, 300*, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

Erhan, D., Courville, A., & Bengio, Y. (2010). Understanding representations learned in deep architectures. *Department dInformatique et Recherche Operationnelle, University of Montreal, QC, Canada, Tech. Rep, 1355*(1).

Fee, S. B., & Holland-Minkley, A. M. (2010). Teaching computer science through problems, not solutions. *Computer Science Education, 20*(2), 129–144.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189–1232.

Fritz, W. (1997). Intelligent systems and their societies. *A free e-book in* http://www.intelligent-systems.com. *ar/intsyst/index.htm*.

Gade, K., Geyik, S., Kenthapadi, K., Mithal, V., & Taly, A. (2020). Explainable AI in Industry: Practical Challenges and Lessons Learned. In Companion Proceedings of the Web Conference 2020, (pp. 303–304).

Goga, M., Kuyoro, S., & Goga, N. (2015). A recommender for improving the student academic performance. *Procedia-Social and Behavioral Sciences, 180*, 1481–1488.

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics, 24*(1), 44–65.

Gorad, N., Zalte, I., Nandi, A., & Nayak, D. (2017). Career counseling using data mining. *International Journal of Innovative Research in Computer and Communication Engineering.*

Guidotti, R., Monreale, A., Ruggieri, S., Pedreschi, D., Turini, F., & Giannotti, F. (2018). Local rule-based explanations of black box decision systems. *arXiv preprint* arXiv:1805.10820.

Hall, P., & Gill, N. (2019). *An introduction to machine learning interpretability*. O'Reilly Media, Incorporated.

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning*. Springer. ISBN 0-387-95284-5.

Hendahewa, C., Dissanayake, M., Samaraweera, S., Wijayawickrama, N., Ruwanpathirana, A., & Karunananda, A. S. (2006). Artificial intelligence approach to effective career guidance. *Sri Lanka Association for Artificial Intelligence.*

Ho, I. M. K., Cheong, K. Y., & Weldon, A. (2021). Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques. *PLoS ONE, 16*(4), e0249423.

Ho, T. K. (1995). *Random decision forests*. In Proceedings of 3rd international conference on document analysis and recognition, vol. 1, pp. 278–282, IEEE.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(8), 832–844.

Hoffait, A.S, Schyns, M. (2017). Early detection of university students with potential difficulties. *Decision Support Systems*, vol 101, (pp. 1–11), ISSN 0167–9236. https://doi.org/10.1016/j.dss.2017.05.003.

Holzinger, A. (2018). From machine learning to explainable AI. In 2018 world symposium on digital intelligence for systems and machines (DISA), (pp. 55–66). IEEE.

How, M. L. (2019). Future-ready strategic oversight of multiple artificial superintelligence-enabled adaptive learning systems via human-centric explainable AI-empowered predictive optimizations of educational outcomes. *Big Data and Cognitive Computing, 3*(3), 46.

Hoyt, K. B. (1987). The impact of technology on occupational change: Implications for career guidance. *The career development quarterly.*

Huang, Z., Liu, Q., Zhai, C., Yin, Y., Chen, E., Gao, W., & Hu, G. (2019). Exploring multi-objective exercise recommendations in online education systems. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, (pp. 1261–1270).

Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. (2018). Student engagement predictions in an e-learning system and their impact on student course assessment scores. *Computational intelligence and neuroscience.*

Iatrellis, O., Savvas, I. K, Fitsilis, P., & Gerogiannis, V. C. (2021). A two-phase machine learning approach for predicting student outcomes. *Education and Information Technologies, 26*(1), 69–88.

Iatrellis, O., Savvas, I. K., Kameas, A., & Fitsilis, P. (2020). Integrated learning pathways in higher education: A framework enhanced with machine learning and semantics. *Education and Information Technologies, 25*(4), 3109–3129.

Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. *Decision Analytics, 2*(1), 1–25.

Jackson, P. (1998). *Introduction to expert systems*. Addison-Wesley Longman Publishing Co., Inc.

Kausar, S., Oyelere, S., Salal, Y., Hussain, S., Cifci, M., Hilcenko, S., & Huahu, X. (2020). Mining smart learning analytics data using ensemble classifiers. *International Journal of Emerging Technologies in Learning (iJET), 15*(12), 81–102.

Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). A conceptual framework to aid attribute selection in machine learning student performance prediction models. *International Journal of Interactive Mobile Technologies*, 15(15).

Kjellin, H., & Boman, M. (1994). Fielded machine learning system for vocational counseling. *Applied Artificial Intelligence an International Journal, 8*(4), 543–563.

Kongsakun, K., Fung, C. C., & Chanakul, T. (2010a). Developing an intelligent recommendation system for a private university in Thailand. *Issues in Information Systems, 11*(1), 467–472.

Kongsakun, K., Fung, C. C., Borirug, S., & Philuek, W. (2010b). An intelligent recommendation system framework for student relationship management.

Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R. J., & Wasserman, L. (2018). Distribution-free predictive inference for regression. *Journal of the American Statistical Association, 113*(523), 1094–1111.

Linsley, D., Scheibler, D., Eberhardt, S., & Serre, T. (2018). Global-and-local attention networks for visual recognition. arXiv preprint arXiv:1805.08819.

Loan, D. T., & Van, N. (2015). Career guidance in secondary schools–a literature review and strategic solutions for Vietnamese rural areas. *American International Journal of Social Science, 4*(5), 135–142.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

Mao, J., Gan, C., Kohli, P., Tenenbaum, J. B., & Wu, J. (2019). The neuro-symbolic concept learner: Interpreting scenes, words, and sentences from natural supervision. *arXiv preprint* arXiv:1904.12584.

McCarthy, J. (2019). What is artificial intelligence (2007). *URL* http://www-formal.stanford.Edu/jmc/whatisai/whatisai.html. Accessed 01.11.2021.

Mihaela, N. A., & Cristina, I. G. (2014). A research on the educational counseling and career guidance in Romania. In 2nd International Scientific Forum, *ISF 2014*, vol. 130, 170, (pp. 28).

Morgan, T., & Ness, D. (2003). Career decision-making difficulties of first-year students. *The Canadian Journal of Career Development, 2*(1), 33–39.

Miao, J., & Niu, L. (2016). A survey on feature selection. *Procedia Computer Science, 91*, 919–926.

Mimis, M., El Hajji, M., Es-Saady, Y., Guejdi, A. O., Douzi, H., & Mammass, D. (2019). A framework for smart academic guidance using educational data mining. *Education and Information Technologies, 24*(2), 1379–1393.

Molnar, C. (2021). Interpretable machine learning. A Guide for Making Black Box Models Explainable, 2019.

Musso, M. F., Rodríguez Hernández, C. F., & Cascallar, E. C. (2020). Predicting key educational outcomes in academic trajectories: a machine-learning approach.

Norasiah, M. A., & Norhayati, A. (2003, January). Intelligent student information system. In 4th National Conference of Telecommunication Technology, 2003. NCTT 2003 Proceedings, (pp. 212–215). IEEE.

Piatetsky-Shapiro, G. (1991). Discovery, analysis, and presentation of strong rules. *Knowledge discovery in databases*, 229–238.

Pintelas, E., Livieris, I. E., & Pintelas, P. (2020). A grey-box ensemble model exploiting black-box accuracy and white-box intrinsic interpretability. *Algorithms, 13*(1), 17.

Poole, M. E., & Juchnowski, M. V. (1974). Career Guidance in Schools.

Pujari, A. M., Dalvi, R. M., Gawde, K. S., & Tatwadarshi, P. N. (2019). SCCAI-A Student Career Counseling Artificial Intelligence. *VIVA-Tech International Journal for Research and Innovation, 1*(2), 1–6.

Qazdar, A., Er-Raha, B., Cherkaoui, C., & Mammass, D. (2019). A machine learning algorithm framework for predicting students' performance: A case study of baccalaureate students in Morocco. *Education and Information Technologies, 24*(6), 3577–3589.

Quinlan, R. C. (1993). 4.5: Programs for machine learning morgan kaufmann publishers inc. *San Francisco, USA*.

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48*(1), 137–141.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). "Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, (pp. 1135–1144).

Russell, S. J., & Norvig, P. (2010). Artificial Intelligence: A Modern.

Saraswathi, S., Reddy, M. H. K., Kumar, S. U., Suraj, M., & Shafi, S. K. (2014). Design of an online expert system for career guidance. *The International Journal of Research in Engineering and Technology*, *3*.

Sample, I. (2017). Computer says no: Why making AIs fair, accountable and transparent is crucial. *The Guardian, 5*, 1–15.

Seo, S., Huang, J., Yang, H., & Liu, Y. (2017). Interpretable convolutional neural networks with dual local and global attention for review rating prediction. In Proceedings of the eleventh ACM conference on recommender systems, (pp. 297–305).

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games, 2*, 307–317.

Sodhi, J. S., Dutta, M., & Aggarwal, N. (2016). Efficacy of Artificial neural network based decision support system for career counseling. *Indian Journal of Science and Technology, 9*, 32.

Sun, V. J., & Yuen, M. (2012). Career guidance and counseling for university students in China. *International Journal for the Advancement of Counseling, 34*(3), 202–210.

The Importance of Motivation in an Educational Environment (2012). Retrieved from https://study.com/academy/lesson/the-importance-of-motivation-in-an-educational-environment.html. Accessed 01.11.2021.

Valenzuela-Escárcega, M. A., Nagesh, A., & Surdeanu, M. (2018). Lightly-supervised representation learning with global interpretability. arXiv preprint arXiv:1805.11545.

Villegas-Ch, W., Román-Cañizares, M., & Palacios-Pacheco, X. (2020). Improvement of an online education model with the integration of machine learning and data analysis in an LMS. *Applied Sciences, 10*(15), 5371.

Viheräkoski, J. (2020). Strengths in career development: Modeling strength-based career counseling through reflecting customer experience.

Watkins, K. A. N. O. K. W. A. N. An improved recommendation models on grade point average prediction and postgraduate identification using data mining. *Advances in Neural Networks, Fuzzy Systems and Artificial Intelligence*, 186–194.

Watts, A. G. (1986). The role of the computer in careers guidance. *International Journal for the Advancement of Counseling, 9*(2), 145–158.

Witko, K., Bernes, K. B., Magnusson, K., & Bardick, A. D. (2005). Senior high school career planning: What students want. *The Journal of Educational Enquiry, 6*(1).

Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2015). A survey on evolutionary computation approaches to feature selection. *IEEE Transactions on Evolutionary Computation, 20*(4), 606–626.

Yang, C., Rangarajan, A., & Ranka, S. (2018, June). Global model interpretation via recursive partitioning. In 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), (pp. 1563–1570). IEEE.

Yang, C., Huan, S., & Yang, Y. (2020). A practical teaching mode for colleges supported by artificial intelligence. *International Journal of Emerging Technologies in Learning (IJET), 15*(17), 195–206.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science, 12*(6), 1100–1122.

Zhai, X., Krajcik, J., & Pellegrino, J. W. (2021). On the validity of machine learning-based Next Generation Science Assessments: A validity inferential network. *Journal of Science Education and Technology, 30*(2), 298–312.