



AI-Powered Career Advisor: Leveraging Machine Learning and Chatbots for Personalized Guidance

Kandula PavaniKrishna

Asst. Prof., Information Technology
Sagi Rama Krishnam Raju Engineering
College (A), Bhimavaram
pavanikrishnakandula@gmail.com

Prathipati Madhubabu

Information Technology(CSBS)
Sagi Rama Krishnam Raju Engineering
College (A), Bhimavaram
madhubabuprathipati458@gmail.com

Bapatla Panvi Srinivas

Information Technology(CSBS)
Sagi Rama Krishnam Raju Engineering
College (A), Bhimavaram
bapatlapanvisrinivas@gmail.com

Gullipalli Jagadeeswara rao

Information Technology(CSBS)
Sagi Rama Krishnam Raju Engineering
College (A), Bhimavaram
jagadeeswararagullipalli64@gmail.com

Kusume Lek haz Samuel Raju

Information Technology(CSBS)
Sagi Rama Krishnam Raju Engineering
College (A), Bhimavaram
kusumelekhaz@gmail.com

Abstract— The proposed paper presents a Smart Career Advisor system that operates on machine learning methods to deliver recommendations on careers to individual students. The system combines two ML models, including XGBoost and CNN. XGBoost models that make use of label encoding are as high as 95.29 accuracy and the CNN model which involves both scaling and label encoding attain a high accuracy of 98.77. Besides the prediction models, the system has a Gemini AI-enhanced chatbot, which is capable of communicating with the users to answer essential questions on career roadmaps, skills needed, and job suitability. Students are also able to seek further specific guidance by posing more questions. The innovative system will support the problem of students choosing appropriate career options by offering customized recommendations based on the data. The project would be a revolution in the career counseling process, where machine learning is combined with conversational AI.

Keywords: *Smart Career Advisor, Machine Learning, XGBoost, Convolutional Neural Networks, Career Recommendation, Chatbot, Gemini AI, Personalized Career Guidance, Career Counseling.*

I. INTRODUCTION

The issue of the right career choosings can be taken as one of the most significant ways which students can make at the moment. However, the contemporary idea of career guidance does not tend to provide information and facts about the individual abilities and expectations. The possibility of improving career guidance process is unlimited with the presence of ML and AI.

The Smart Career Advisor system proposed will close this gap by proposing the use of more powerful ML model to give career suggestions. The system may make the appropriate career recommendations through the use of data-driven predictions by the profiling of a student such as the education history, talent and preferences. To enhance the user interaction, there is a Gemini AI-based chatbot that answers different questions such as career path, skill sets required in specific jobs, and job match based on inputs in the system made by the student.

The project will target the paradigm change of career counseling as well as offering a solution that is innovative,

accessible and scalable and will assist students to take informed decisions about their future career. The use of machine learning and conversations AI will ensure that the students will receive the correct, up-to-date and personal assistance based on their individual needs.

II. REVIEW OF LITERATURE

Recent research has increasingly focused on developing intelligent career guidance and recommendation systems using ML and AI, aiming to provide personalized suggestions and improve decision-making for students.

Career guidance is one of the areas where (ML) and (AI) have been applied to provide a high-scale response with personalized but not the classical response. The lack of precision and scalability of the traditional approach to career advising is clarified over the system of career guidance that incorporates the model of ensemble ML that involves the DT, RF, and NB to give the recommendations based on the profile and interests of the students revealing a more robust and effective implementation scheme crminated by the possibilities of ensemble learning models [1].

Career guidance system A career guidance system is an AI-driven career guidance system that specialized career guidance, and automated career advice by machine learning. The technique can be likened to the conventional psychometric frameworks which are normally inflexible and fail to adapt fast to the dynamic interest and abilities of the students. The concept is more dynamic and hence, better and proposes an efficient alternative to the conventional means of career counseling [2].

The key to remembering the importance of hybrid models which can place more than one algorithm as yet to enhance the accuracy and individualization of career advisor is the use of random forest, support vegetable machine (SVM), choice tree, and XGBoost with an NLP chatbot. The system will also advance the interaction with the users through the addition of chatbots, which are projected to be founded on the language models and that will offer real-time career guidance through the chat interface [3].



The AI and machine learning can turn the competence of students into a match with the recruitment needs and, thereby, assist in building unique career proposal. The failure of the orthodox career guidance model is frowned upon, and guidance based on AI can be more precise and more customized to the suggestions positioned in terms of proposing career path to the students by studying a broader set of student data, and align it with the industry demands [4].

The consideration is given to the AI-based career guidance systems and the emphasis is put on the use of chatbots and ML models, such as RF, KNN, and SVM. In the research article, they describe an enhanced role of conversational agent in the provision of career advice whereby in such systems, the system would provide a more interactive and scalable response compared to the traditional systems whose systems cannot meet system requirements of large number of users [5].

The system of career path recommendation based on machine learning works on random forest, SVM and KNN. As shown in this paper, machine learning can be utilized to help make informed decisions, process information (academic and non-academic) to push decisions. The system is among the examples of the application of ML that can suggest career opportunities based on the aggregate profiles of the students [6].

Random forest and a SVM-based system do the same predictions of suggesting the right career option because it is one of the students in its capacity to propose their talent and personality developed by way of the administration of questionnaires. Although the study explains the implications of these approaches with limitations, it establishes the necessity of the existence of more adaptive approaches that take into account as much data as they can and base their career advice on the changing trends of career markets [7].

The overall examination of the case under ML-based career recommendation system, with such alternatives as resume analysis, identification of skills gaps, and job/course recommendations, highlights more detailed systems that are able to fully examine the profile of a student and deliver a balanced career advice, which cannot be limited to standard career services [8].

Predicting career paths employs the process of natural-language processing (NLP), and the process applies a multiplicity of classifiers to a resume and cover letter text. The research indicates that the use of NLP in career foretelling models is essential because they are utilized sequentially in a bid to encompass all the complexities of language in the data-driven procedure that overlooks it [9].

Relevance of machine learning career counseling chatbots has been addressed since the conversational agents were reported to be effective regarding the delivery of interactive career guidance in a personalized manner. According to the study, chatbots, compared to the traditional questionnaires, are more versatile and scalable since they provide a dynamic user experience, and methods towards providing advice in real time [10].

It is suggested to apply an integrated personalized career advice NLP and deep learning model, which has the

advantage that it generates more scalable personalized career advice. It takes heed to a vexed custom rule-grounded strategies whereby AI is more efficient to grasp superficial information on students and achieve more plausible decisions in the view of the current career markets [11].

Career path recommendations system is a mash-up between machine learning and NLP, the emphasis on the scale career recommendations. According to the research, since the available data is limited and only simple decision rules exist, one can exploit the existing systems only to a certain extent, whereas the alternative based on the use of the integrated model is much more powerful as it employs the complete variety of information that take place on a particular student [12].

The growing worry of the chatbots regarding career advice is explained and the way the conversational AI can be implemented to provide individual counseling. The paper has addressed the issue of making the chatbot systems more specialized and provide them with more specific and context-focused suggestions and points to the need to implement the systems based on deep learning models so that they become flexible [13].

Career advice systems that utilise various supervised ML models as RF, SVM, KNN and ANNs are explored. This suggests that the AI-based systems are able to offer individualized instructions because they can compute big and diverse volumes of information, and this is far beyond the capabilities of the conventional counseling services. Among the benefits that may be introduced into the field of career counseling through machine learning, the study refers to precision, scalability and personalization that would be accomplished [14].

The AI-based career guidance system offers career planning recommendations to the student across the profile, skills, and interests. It has been observed in the study that the classical career guidance models include the inability to keep up with dynamic changes in the student profile as well as the labor market and develops an idea that machine learning architectures can be more flexible and offer more precise solutions [15].

III. DATA COLLECTION

The data of Smart Career Advisor system includes the characteristics that determine the extent of the skills held by the students in various fields such as Database Fundamentals, Cyber Security, Software Development, AI and ML, Project Management and others. The separation of features into certain levels of proficiency is done to the following levels of proficiency (in this order): Average, Beginner, Excellent, Intermediate, Not Interested, Poor, and Professional. This assists the system in ranking the student in regards to his or her strength and dynamism in different fields.

Job opportunities, which will be targeted, will be Database Administrator, Software developer, project manager, Data scientist, and Cyber Security Specialist among others. This data set performs the role of determining the right career paths with the help of suggestions based on the talents and interests



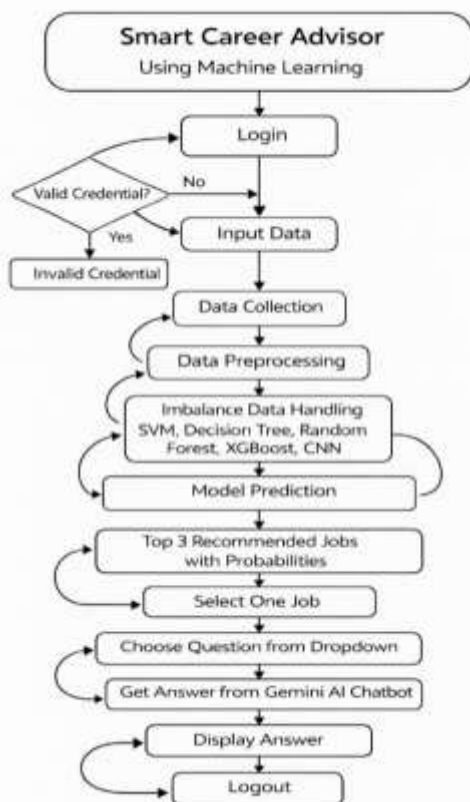
of the students that provides a tailor-made recommendation to the students in good career choices.

IV. PROPOSED SYSTEM

The proposed Smart Career Advisor is an ML-based system, permitting to provide the students with a personal career advice, basing on their abilities, academic results, and career choices. The system will consist of three primary elements, including a machine learning career prediction model, chatbot interface with the help of Gemini AI, and recommendation engine. Convolutional Neural Networks (CNN) and XGBoost are machine learning models, which can be used to examine proficiency of the customer, on a per-area basis, including technicalism, such as Database fundamentals, Cyber Security, Software development and soft skills, such as Communication and Project Management. In this analysis the system predicts the most adequate career positions among an existing set of pre-determined positions based on Software Developer, Cyber Security Specialist, and Data Scientist.

The chatbot (Gemini) is an AI-based tool aimed at the users being provided with tailored guidance on the careers available, skills required to be employed in the job opportunity, and overall suitability to employment. The students will also be able to pose other questions and the chatbot will reply to it with the relevant information as per the inputs. The system will consist of an accessible, scalable, and data-driven system of career counseling to facilitate the closeness between the skills of the students and the available career opportunities at hand. Smart Career Advisor helps students to take rational decisions on their careers by applying AI and advanced machine learning algorithms on their professional future. As shown as Fig. 1.

Fig. 1. Flowchart of the proposed system



A. Dataset Preparation and Exploratory Data Analysis(EDA)

The first stage of the discourse of developing the Smart Career Advisor system is the data preparation and exploratory data analysis (EDA). The data has been designed in such a manner that it contains multiple attributes that describe the students competence in the majority of the fields such as Database Fundamentals, Cyber Security, Software Development among many others. The EDA stage is applied to explore the values of the data set on the completeness, structure and distribution. This incorporates making sure that data is not missed or no outliers or that there are data imbalances of the data. This could be known by visualization of the data in terms of the histogram of data, box plot, and correlation matrices. EDA helps in identifying trends and hints that will be taken into account to create the next steps of preprocessing and model choice.

B. Preprocessings:

Preliminary processing is a critical step in the ML pipeline that will be created to process the data to be utilized to train the model. The data then fall under a series of manipulations to be used with the ML algorithms. The model is not biased with redundant data being removed by the process of eliminating the duplication row. Second, nominal attributes, e.g. level of proficiency acquired and skills are changed into numbers through Label Encoding. This would be transformed in a way that machine learning algorithms will work with the data in an effective manner. Encoding entails assigning every category a special numeric nomenclature and this is so in order to make categorical variables get viewed out of the boxes by the models. There are also other models where feature scaling is applied to normalise the data, to some models like (CNN) all features have equal roles in model training. This helps in the rapid convergence and also the better performance of CNNs that is sensitive to the scale of the input.

C. Model Building and Training Strategies

The second process is to build and train ML models after data has been preprocessed. The data will be categorised into 2 parts: training the model and the other part is applied to evaluate the performance of the model. The model will be then trained using the given data that will identify the relationship among the features with the target variable that is the career Role which the system will present. It has training based on a large number of ML algorithms, such as XGBoost, and (CNN). Each of these models is trained using a small part of the data and the objective is to train such models on how all the different features can be correlated to create the best career roles. After the training process, the model is checked

k=1



which the associations among features may be multifaceted and unnoticeable. The XGBoost is a speedy and predictive gradient boosting algorithm that is known due to its rapidity and forecast abilities. It can also be applied to the tabular data and it may be used in classification and regression. A type of ensemble learning that is referred to as random forest is also in use because it is robust and simple to apply. It achieves this by mixing the forecast of different decision tree in order to increase accuracy and reduce over fitting. The reason behind implementing these three approaches in the system is that they forecast the best job positions to the user based on on unknown data to check its generalization and prediction accuracy. The train-test split approach whereby the data is divided to form training and testing set in 80/20 or 70/30 ratio are used to ensure that the model will be evaluated without any bias.

D. Performance Evaluation and Comparative Analysis:

Once the models have been made, their performance is evaluated through the performance measurement that takes into consideration the different measures of performance to establish the accuracy and the ability of the models to make the correct predictions. The former is accuracy which was the only key measure as the rest of the evaluation metrics such as the confusion matrix are a more informative indicator of the model performance in terms of what types of classifications are made and what ones are not. One can use the confusion matrix to calculate other important metrics including the precision, recall and F1 score that is essential in understanding the ability of the model to handle the different classes and more importantly in imbalanced collections of data. The different models are compared i.e.: Random Forest, XGBoost and CNN and the most appropriate model to the problem explained was chosen. The comparison of the accuracy scores, training time and the computational efficiency is involved in this analysis and this allows one to make a viable decision on the model that is going to be applied in the actual implementation.

E. Methods:

Three major machine learning algorithms are employed in the system; Convolutional Neural Networks (CNN), XGBoost, and. CNN is chosen due to its ability to learn the hierarchical features representation and can be used in a problem where the feature extraction process is not easily resolved. It is effective in particular with structured data in

his/her expertise when each of the models plays a unique role in the overall effectiveness of the prediction.

1) XGBoosting

XGBoost is a useful gradient boosting algorithm that continuously adds decision trees to correct the previous errors. It also fewer custom objective function which is a loss function plus a regularization term, to remove overfitting. XGBoost is the most preferable in the case of classification projections by the fact that it is expeditious and it can survive on vast data and as well as organized data.

$$Obj(\theta) = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum^K \Omega(f_k) \quad (1)$$

2) CNN

The deep learning models also referred to as (CNN) are designed to manipulate grid-like information and images specifically. The CNNs have convolutional layers and pooling layer and fully connected layer, which automatically extract hierarchical aspects between the available raw data. Convolutional layers apply filters to the input data and the pooling layers reduce the dimension. The CNNs are also effective in classifying images because they can acquire spatial hierarchies of features without the need to work on features.

$$O(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n) \quad (2)$$

V. RESULTS AND DISCUSSIONS

A. Results of Models

The training accuracy performance is the percentage of the correct predictions the model had made of the training data and the testing accuracy performance is the percentage of the correct predictions which the model had made of the unknown testing data. The CNN model possesses the remarkably high training and testing accuracy statistics of 99 and 98.77, respectively, implying that the model is very good, as far as it is used with the training and testing data sets. The accuracy of XGBoost is greatly low at 96 percent during its training and 95.29 percent during its testing that is, it is a good performer but not as accurate as CNN. Random Forest, has training and testing accuracy of 94 percent and 94.58 respectively i.e. it has been able to recognize well on the testing data even though it slightly overfitted on the training data as compared to CNN model.



1) CNN Model:

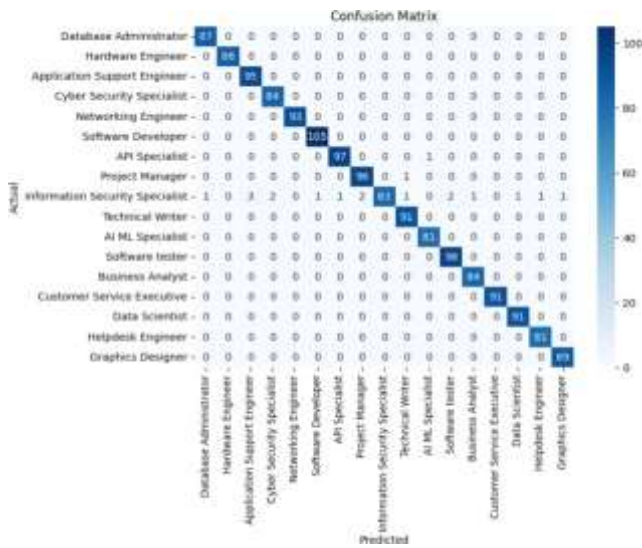


Fig. 2. Confusion matrix of CNN

The confusion matrix offers a good understanding of the working of the classification model, especially on the correct study of job roles. As referred in both Table 1 and Fig. 1, the diagonal values are the job positions that are correctly classified e.g., Software Developer, Hardware Engineer, and Application Support Engineer with the high degree of accuracy. Nevertheless, the model also has certain misclassifications, with the position such as the Information Security Specialist being wrongly merged with Technical Writer and AI/ML Specialist. In spite of such misclassifications, the model has a high overall accuracy in job role prediction, and very few erroneous predictions. The accuracy of 0.98 shows that 98 percent of the positive forecasts are accurate which means that the model reduces false positives. On the same note, the Recall of 0.98 indicates that the model can correctly identify 98 percent of the true positives and still miss only a few of these True positives amounting to 2 percent of the true positives. A good balance of precision and recall is represented by the F1-Score of 0.98, implying that the model is not skewed in the false positive and is very effective in recognizing job roles correctly.

2) XGBoost Model:

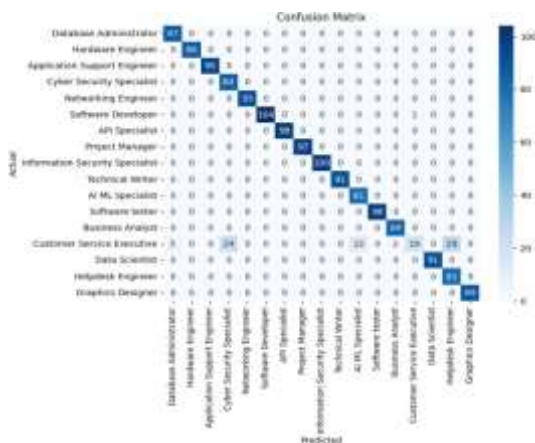


Fig. 3. Confusion matrix of XGBOOST

The confusion graph shows that XGBoost classification model predicts job roles well as there is no huge error between the actual and predicted job roles. The diagonal values are proper predictions with the greatest ones depicting perfect job role allocations, Software Developer and Information security specialist. The classification is however messed up, particularly in the role of Customer Service Executive where multiple job classifications are not correctly foretold. The Precision of 0.96 shown in fig. 2, that is when the model gives a positive prediction, there is a probability of it being right 96 percent of the time and, though a little less than the CNN model will give it good accuracy in positive predictions. The Recall of 0.95 implies that XGBoost errs only 5% of true positives but it misses a few true positives to identify the remaining 95 percent. The F1-Score of 0.94 indicates a balanced performance and indicates the fact that the model is performing well in terms of addressing the trade-off between precision and recall with a slight gap between the two, which can still be considered as appropriate due to the nature of the issue.

F1 of 0.94 means that XGBoost is indeed doing well when losing a small portion of disparity of precision and recall with CNN. There is a slight difference between r-call and precision that are also important depending on nature of problem in this model.

On the whole, XGBoost is characterized by good performance with 96 percent accuracy, 95 percent recall, and 94 percent F1-score, whereas CNN is also better than XGBoost with its 98 percent accuracy, 98 percent recalls, and 98 percent F1-score, which depicts a high performance ratio. Which is shown in the Table 1.

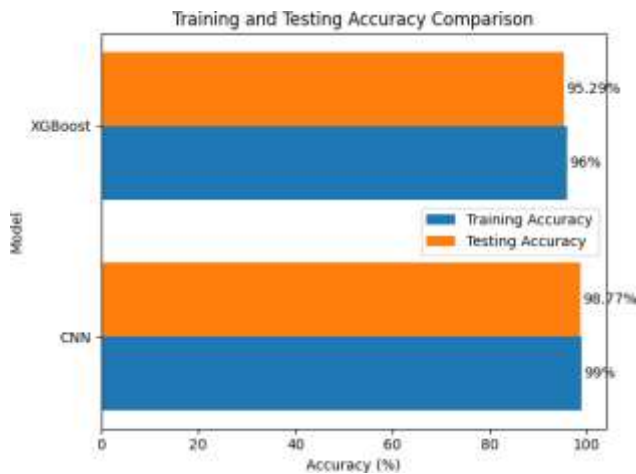


Fig. 4. Model Comparison

Model	Training Accuracy	Testing Accuracy	Average Precision	Average Recall	Average F1-Score
CNN	0.99	98.77	0.98	0.98	0.98
XGBoost	0.96	95.29	0.96	0.95	0.94

TABLE I. EVALUATION MATRIX

VI. CONCLUSION

The Smart Career Advisor system is also an efficient technique as it is based on the machine learning algorithm such as the use of XGBoost and Convolutional Neural Net (CNN) to infer the most desirable career paths that students will undertake based on their capabilities. The best model was XGBoost because it had a test accuracy of 95.29 percent which rendered it to be effective in large datasets. It has yet to be enhanced but is very accurate when applied in the aspect of career roles forecasting of the situation. This structured data did not go well with CNNs despite their high accuracy in training, only to demonstrate that CNNs are very effective where unstructured data are involved, but are not necessarily the best choice where predicting requires one to be structured.

It is also connected to the chatbot driven by the Gemini AI that enables the students to have questions about their career-related issues and get specific answers. This will also personalize and the interactive quality to advisor system that will ultimately boost the user experience.

To summarize, one can observe that the XGBoost model is the most suitable to implement in the situation with career prediction, yet CNNs and the application of the feature that provides the AI with the ability to make conversations are also impactful in terms of system value. Smart Career Advisor system informs about what machine learning and AI might introduce in the future to assist students in a personalized, data-driven career advising so that they could make better and more informed career-related decisions.

The Smart Career Advisor system was found to have impressive performance with the XGBoost registering outstanding performances in reactions in which the accuracy topped 95.29 percent in both. This was the most successful model. XGBoost having the testing accuracy, is speedy, scalable and hence, it can be applied to a bigger dataset. It, nonetheless, requires additional adjusting, and, therefore, it becomes a little more complex to use. Convolutional Neural Network (CNN) network with the test accuracy of 98.77 per cent was found to be one of the models that can pick up complicated patterns. However, CNNs are in a better position to work with unstructured information and whilst they were as effective in the structured career prediction data, they were not more effective in prediction than the XGBoost model. A more complex dataset is more likely to be addressed by CNN, but in the case provided, an overall model of XGBoost is the best to use, as it has reached 95.29% accuracy in the test, readability and is quite easy to operate compared to CNN.

VII. REFERENCE

- [1] C. U. Betrand *et al.*, "Career Guidance System Using Decision Tree, Random Forest, and Naïve Bayes Algorithm," *International Journal of Science, Technology and Society* 2025, Volume 13, Page 35, vol. 13, no. 2, pp. 35–42, Mar. 2025, doi: 10.11648/j.ijsts.20251302.11.
- [2] L. Xu *et al.*, "Improving Question Embeddings With Cognitive Representation Optimization for Knowledge Tracing," *IEEE Trans. Cybern.*, vol. 56, no. 1, pp. 235–248, 2026, doi: 10.1109/TCYB.2025.3612399.
- [3] G. Danwani and N. Ahmed, "AspireNextGen: An Explainable AI-Powered Personalized Career and Skill Recommendation System," 2025, doi: 10.64388/IREV916-1713212.
- [4] I. Francis Joseph and U. Zannah Bukar, "AI-Powered Career Guidance Systems: Foundations, Opportunities, and Challenges," *International Journal of Education and Evaluation*, doi: 10.56201/ijee.v11.no2.2025.pg133.147.
- [5] Dr. K. B. Monika D S, "A Review on AI-Based Chatbots for Personalized Career Guidance," *International Journal of Advanced Research in Computer Science & Technology(IJARCST)*, vol. 8, no. 5, pp. 12833–12838, Oct. 2025, doi: 10.15662/IJARCST.2025.0805011.
- [6] V. Nayak and N. Vora, "NonCommercial International (CC BY-NC 4.0) License A Machine Learning-based Career Recommendation," *Journal of Trends in Computer Science and Smart Technology*, vol. 6, no. 4, pp. 374–390, 2024, doi: 10.36548/jtcsst.2024.4.004.
- [7] E. Padma, P. Sowdharshini, P. Shanmugapriya, K. M. Reshmaa, and C. N. Srimathi, "Career guidance system for students using machine learning," *Challenges in Information, Communication and Computing Technology*, pp. 666–671, Nov. 2024, doi: 10.1201/9781003559092-115.
- [8] K. Talekar, U. Tiwari, S. Vishwakarma, H. Gupta, and D. Singh, "CareerCraft AI: A Review on Intelligent Career Advisory Systems Using Machine Learning and Natural Language Processing," *International Journal for Research Trends and Innovation*, vol. 10, pp. 2456–3315, 2025, Accessed: Feb. 11, 2026. [Online]. Available: www.ijrti.org



- [9] R. Ade and P. R. Deshmukh, "Efficient knowledge transformation system using pair of classifiers for prediction of students career choice," *Procedia Comput. Sci.*, vol. 46, pp. 176–183, 2015, doi: 10.1016/j.procs.2015.02.009.
- [10] R. Goyal, N. Chaudhary, and M. Singh, "Machine Learning based Intelligent Career Counselling Chatbot (ICCC)," *2023 International Conference on Computer Communication and Informatics, ICCCI 2023*, 2023, doi: 10.1109/ICCCI56745.2023.10128305.
- [11] S. G. Satpute and H. Sambhaji Garudkar, "AI-Powered career guidance and skill development platforms for students," *International Journal of Advanced Scientific Research*, 2025, Accessed: Feb. 11, 2026. [Online]. Available: www.allscientificjournal.com
- [12] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.
- [13] V. Dalvi, C. Choudhary, V. Kokitkar, S. Kulkarni, and G. Potdar, "Chatbot for Academic / Career Guidance".
- [14] A. Balasubramanian, "PERSONALIZED CAREER PATHWAY: A HYBRID MACHINE LEARNING APPROACH FOR DYNAMIC RECOMMENDATIONS," *Journal of Artificial Intelligence, Machine Learning and Data Science*, vol. 1, no. 1, pp. 1999–2003, Nov. 2022, doi: 10.51219/jaimld/abhinav-balasubramanian/440.
- [15] M. Govind *et al.*, "AI-BASED CAREER GUIDANCE SYSTEM," *International Journal of Creative Research Thoughts*, vol. 13, pp. 2320–2882, 2025, Accessed: Mar. 02, 2026. [Online]. Available: www.ijcrt.org