

By this all the system will recommend the personalized roadmap for the unique user's profile.

Related Work

The research, Career Guidance System Using Machine Learning, having objective of build a career guidance system which predicts suitable careers for students based on **personality traits and learning styles**. It defines the web-based platform recommending career paths based on predicted **personality type and learning style**. In this paper the approach was to Uses **Big Five Personality Test** and **VAK learning style (Visual, Auditory, Kinesthetics)** data, Data sourced from **Kaggle (Big Five)** and **Google Dataset Search (VAK)** and preprocessing involved cleaning, encoding, normalization, and splitting into train/test sets. It mainly focuses on **personality and learning styles** for career prediction.[1]

The paper, A Machine Learning-based Career Recommendation that design a **data-driven career recommendation system** that predicts career paths based on **academic performance and extracurricular factors** using dataset of 1,258 student records and have 17 career categories (Software Engineer, Doctor, Data Scientist, Teacher, etc.). It mainly focuses on **academic performance and extracurriculars** with stronger **data analysis and model comparisons**.[2]

The research article, A personalized M.L. based system to evaluate student's skillset and analyse the gap between academic and industry for engineering students which extracts **job skills** from job advertisements and compares student-acquired skills with required skills for predicts placement outcome and suitable job role with the help of **Naive Bayes, K-Nearest Neighbour (K.N.N.), Support Vector Machines (S.V.M.), and Artificial Neural Networks (A.N.N.)**.[3]

Methodology

The proposed system will work in four main stages:

A. Data Collection

Student's data will be in term of 10th marks (especially English and Maths) ,12th marks (PCM),technical interest, extra-curricular interest and job .This data for the roadmap formation to the particular job. On other side data of the alumni to Guide the students.

B. Data Preprocessing

Some students cannot fill their full information ,so to fill that empty cells and make the prediction smooth we use mean, median and mode .Also most of the inputs are text-based , so we use label

encoding to convert them in numeric form so the model can process them.

C. Model Training

We have first used the algorithms like Decision tree, Random tree forest and XGBoost algorithm .Finally, we XGBoost , which builds multiple tree sequentially . Each tree corrects the error of previous tree ,making it smarter and more accurate.

D. Roadmap Generation

Once the model is train then it's predict the job for the students. Along with them it also create the perfect roadmap for the job.

Algorithms

M.L. is a subset of A.I. that enables computer systems to learn from data and enhance their performance over time without being explicitly programmed for every instruction. It consists of algorithms to find the pattern in dataset ,to predict the result on the basis of the data .

E. Random Forest

Random forest is the machine learning algorithm in which it developes make decision trees for the better prediction of the dataset. Each tree looks at the random part of the dataset and then vote for the classification and average for the regression. This reduces error and improves accuracy of model.

$$Y^{RF} = \frac{1}{T} \sum_{T=1}^T y^T$$

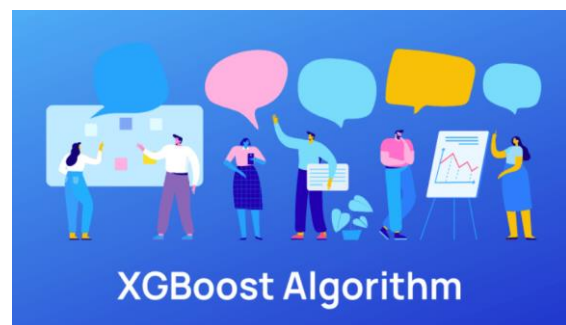
Where,

Y^t = prediction of yth tree

T = number of trees

F. XGBoost algorithm

Earlier, in traditional era we used decision tree and random forest algorithm to predict and it is easy to do that . But we had to face problems like accuracy and complex dataset .XGBoost which is an advance M.L. algorithm designed to enhance the speed, efficiency and high performance.



Working of the XGBoost :-

1. Start from base learner x
2. Calculate error
3. Train the next tree

- Repeat the process
- Combine the prediction

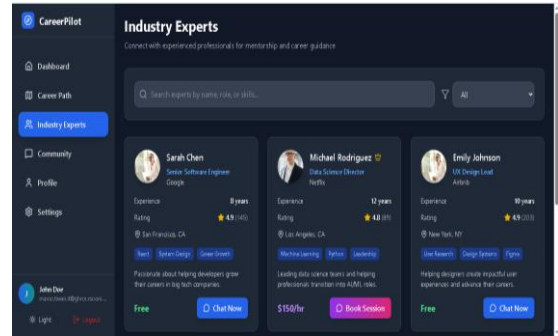
$$Y^{\wedge}i = \sum_{k=1}^K f_k(x_i)$$

Where,

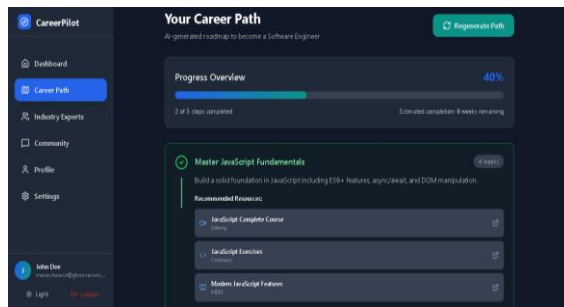
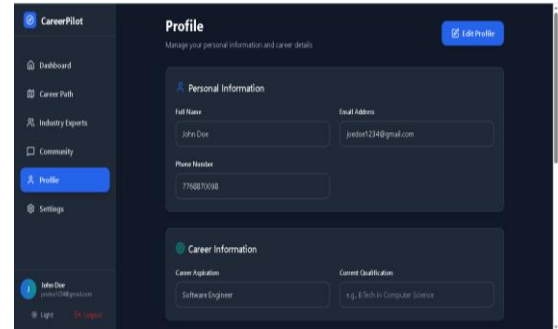
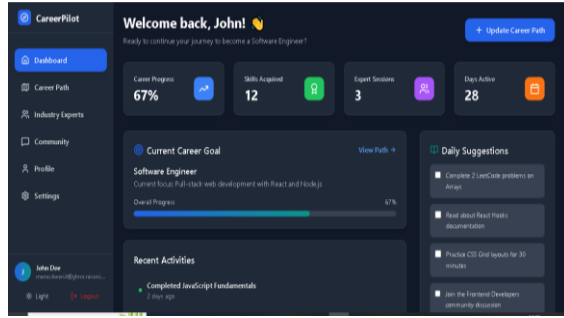
$Y^{\wedge}i$ = Final Predicted value

K = number of tree

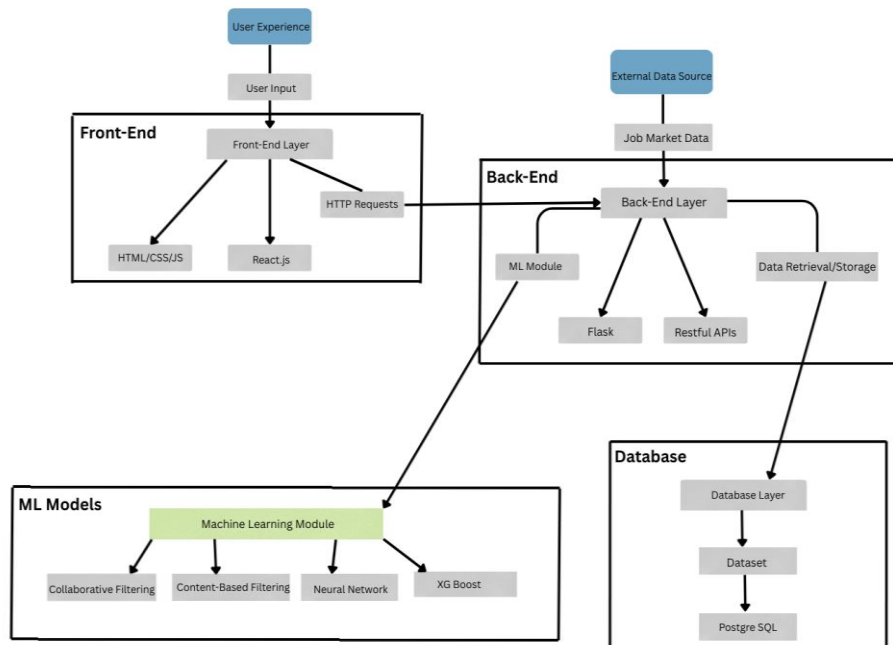
$f_k(x_i)$ = prediction of Kth tree for the ith data point



Website



Architecture



Result And analysis

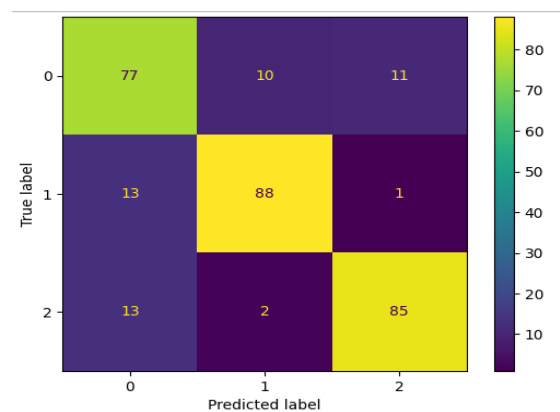
As we demonstrate the website with the different types of algorithm ,in which we get the best accuracy from the XGBoost algorithm.

Algorithms	Accuracy
Random Forest	79.9%
Decision tree	67.3%
XGBoost	81.7%

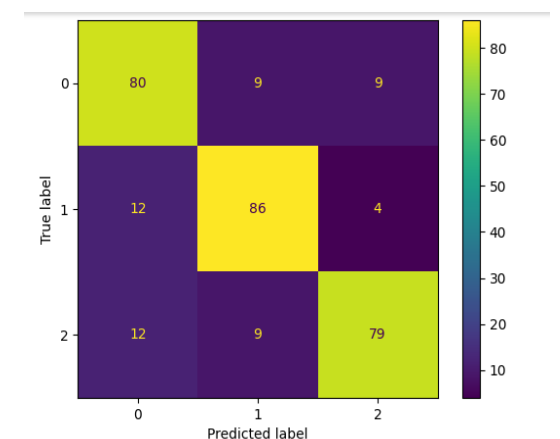
XGBoost algorithm is much more accurate then the decision tree and random forest. It can handle the large dataset and also handle the empty data of the dataset. That’s why we use XGBoost for this model.

Confusion Matrix

It's a N*N matrix which is used for summarizing the performance of the model. It’s helps in understanding where the model is making error and overall perform.



Confusion matrix of Random forest



Confusion matrix of XGBoost

Limitations

- Data dependency is the major issue ,if the student cannot fill the all information or

any incorrect information ,then it predict incorrect result.

- We have use limited data for the training ,so it may not cover all the possible career paths.
- It does not have feature of the emotional and psychological consideration.

These limitations does not reduces the usefulness of the system , but highlights the area of improvements for betterment of the system.

Conclusion

In this model , we focuses on the suggesting the job with the consideration of the student’s interest ,skills and marks. In this project , we are not only predict the job , we also provide the roadmap to reach the goal .Along with this we also provides the alumni connects for the good guidance .

To train this model we used various M.L. algorithms i.e. Random_forest , Decision_tree and XGBoost. But the XGBoost is the most accurate and suitable algorithm for this model.

This project is very important for the students , not to only get the job suggestion but also give the step-by-step roadmap to the students to reach their successfully. Moving forward , the focus will be on expanding it’s capacities to remain it as a trusted career guidance website.

Future Scope

- Currently it is mainly focused on engineering , further it can be expand to other profession also.
- The implementation of emotional feature can be help students to motivate them for their professional journey.
- Gamified version can also increase the interest and usability of the system
- Developing a Mobile app can increase the accessibility of the system.

References

S.Thavasi, T.Revathi [2024], “A personalized machine learning- based system to evaluate students skillset and analyze the gap between academia and industry for engineering students” [Online] <https://doi.org/10.1108/K-01-2024-0053>

Vaishnavi Nayak, Neha Vora [2024], “A Machine Learning-based Career Recommendation” , [Online] <https://irojournals.com/tcsst/article/pdf/6/4/4>

Subiddhya Panthee, Sudikshya Rajkarnikar, Rabiya Begum [2023], “Career Guidance System Using Machine Learning” [Online]

<https://www.nepjol.info/index.php/JACEM/article/view/55947/41813>

Harsh Kumar, Upendra Pandey, Pankaj Kumar [2022] "AI-Based Career Path Recommendation System"[Online]

[https://amity.edu/UserFiles/aijem/293Harsh1.0%20\(AJCS\).pdf](https://amity.edu/UserFiles/aijem/293Harsh1.0%20(AJCS).pdf)

Tianqi Chen , Carlos Guestrin[2016], "XGBoost: A Scalable Tree Boosting System" [Online]

<https://dl.acm.org/doi/pdf/10.1145/2939672.2939785>