final paper

by Sarvjeet Utkarsh

Personalized Career Path Recommendation Using Deep Learning For Freshers

Mr. Sarvjeet Singh
Dept. of Computer Science
ABES Engineering College
Ghaziabad, India
sarvjeet2606@gmail.com

Mr. Utkarsh Jain
Dept. of Computer Science
ABES Engineering College
Ghaziabad, India
utkarshjain09091@gmail.com

Ms. Sapna Jain
Dept. of Computer Science
ABES Engineering College
Ghaziabad, India
sapna.jain@abes.ac.in

Mr. Vikas Pargai
Dept. of Computer Science
ABES Engineering College
Ghaziabad, India
pargai.vikas03@gmail.com

Abstract—The sole purpose of this project is to develop a method for providing personalized career guidance to a college fresher.A web application called ABES4U has been developed for ABES Engineering College in Ghaziabad. It provides comprehensive career guidance to freshers who have a lot to explore but very limited time.For this project, a survey was conducted in computer science and allied branches at the college, focusing on Computer Science Engineering students. The dataset contains 1000 data points along with 16 attributes.To provide career path recommendations, a deep learning model called Friends AI was developed. It takes both technical and personality attributes as input parameters and suggests a career path based on the experience of a senior student from the college. Additionally, ABES4U has other features to ensure the overall grooming of the college students. Ultimately, this research and ABES4U will help them become industry-ready by the time they graduate. It is suggested to implement this approach for core branches that are not related to computer science as well. Hence, ABES4U will help students to achieve more and perform better in their careers.

Index Terms—Machine Learning, Deep Learning, Neural Networks, Career Guidance.

I. INTRODUCTION

This research was initiated to develop a solution that can minimize the gap between the increasing demands of the tech industry and the skills possessed by recent computer science graduates. We have observed a significant decline in job opportunities within the tech industry, particularly for freshers. This situation was predicted during the Covid-19 pandemic.

After conducting an in-depth analysis of the factors that contribute to defining a student's skill set, we have identified several dominant factors. To assist students in their engineering journey, we have made the decision to incorporate all of these features into ABES4U. The aim is to enhance the holistic development of students and equip them with the necessary skills to thrive in the industry.

Fig 1 is showing the Use-case diagram of ABES4U application. In this section, only the features are described implementation details will be shared in the later section of this paper.

A. Friends AI

This is the main feature of ABES4U and the center of this research paper. It is a neural network model with an accuracy of 92%. The layer architecture is present in the later sections.
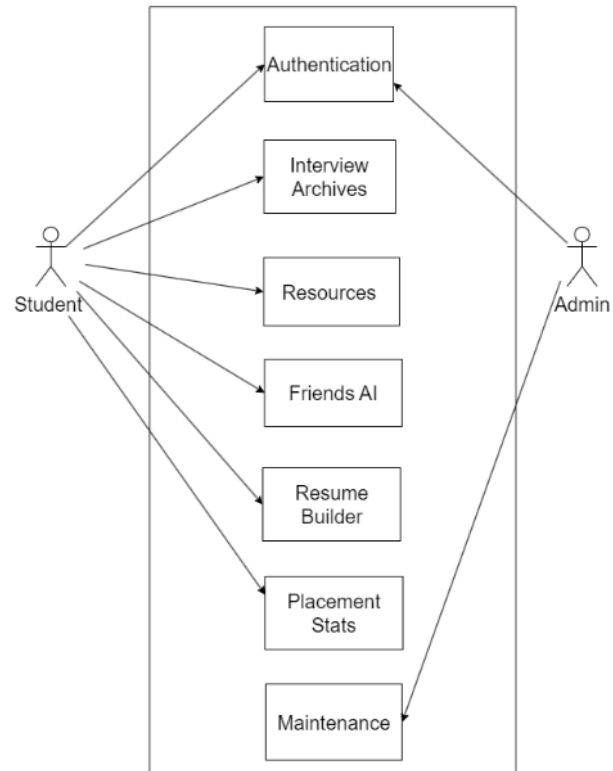

Fig. 1. Use Case Diagram of ABES4U.

It takes user's personality and technical input, then it match the user's features with a group of the seniors having same attributes, finally the most relevant senior's technical domain is recommended to the user.

B. Interview Archives

These archives enables students to analyze the interviews of their seniors and take important insights about the important topics and types of questions asked by a company in their recruitment process. It help students to be more aware about the industry and help them in preparing for the interviews.

C. Placement Stats

While the Interview archives tell about the interview experience of a particular senior. This feature provide a bird eye view to the students of their senior placements. It give insights regrading the overall placement scenario of previous batches. We are fetching this data from the placement cell of our college and this will be updated and maintained by the Placement officers in our college.

D. Resources

This section encompasses a collection of knowledge resources that have been passed down by seniors to juniors as a valuable legacy. It comprises a comprehensive range of materials related to computer science, including notes in PDF format and curated YouTube playlists covering various topics. Additionally, this section also provides a well-defined roadmap, containing detailed information about the materials available for each domain.

E. Resume Builder

A resume is one of the most crucial factors in securing students their dream opportunities. While many students possess the necessary skills, they often struggle with effectively showcasing their abilities. A well-crafted, concise single-page PDF resume can have a significant impact on their future prospects. Furthermore, it is vital to keep the resume up to date, as the job market evolves rapidly.

In today's competitive landscape, resumes should be easily parsed by various parsing tools. With the sheer volume of applications received, HR professionals are unable to manually review each resume. Consequently, automated tools are employed for initial shortlisting in many large companies. Recognizing this need, we have developed an additional feature on ABES4U that enables students to effortlessly create and maintain well-structured resumes online. This empowers them to increase their chances of being recognized and considered for desired positions.

F. Authentication

ABES4U is an exclusive platform designed specifically for students of ABES Engineering College. With the utmost priority given to privacy and security, ABES4U incorporates an authentication module that ensures only genuine students can access and utilize its applications. This stringent authentication process guarantees the confidentiality and protection of our students' information.

ABES4U will provide the necessary guidance and support at the right time so that students do not have to waste their crucial time in figuring out everything on their own. It will assist them and help them to shape their career in a better direction and will give them a head start and an fair advantage on their competition. Furthermore this research does not intend to limit a student's exploration but rather to help them prioritize domains better and save precious time.

II. LITERATURE SURVEY

In this research paper, authors have proposed a novel approach to recommend job and career path to the students. They have used text mining and collaborative filtering for the recommendation purpose. Their dataset contains only 1000 jobs and 100 users. They have split data into 70:30 training test ratio. They have used user based filtering for the recommendations to a new user. The authors have proposed using Hadoop clusters for more data handling. The authors have used two metrics one is precision and RMSE for evaluation of their recommendation system. [1]

In this research paper, author have suggested to choose pattern learning mode is chosen to train the algorithm rather than batch learning mode. The author have C4.5 Algorithm to predict the career path.This model focuses on the psychological parameters of the students and recommend them career domain of student. The lack of awareness among the students create the issue in choosing the career path and hence student fails to work upon their interest.The author improves the accuracy with statistical parameters applied to C4.5 algorithm and achieved 86% of it. [2]

The authors have implemented a unique skill graph generation model for career path prediction for a working professional, The authors have provided a three module based system for generating the skill graph of an user. They have used text mining techniques for extracting data from the users resume, typically that data is about education and about the skills, experience of the user. They have got recall and precision of 86.44% and 80.54% on their text extraction model. [3]

In this paper authors have recommended the career department by the score of a objective test given by a particular student. This makes a approach a very vague approach.They have applied K-Means Clustering algorithm for predicting the career department of the student.The authors have created a web based application to create a guidance system for the student who are unable to choose the department or stream after their senior secondary.The approach via objective test makes it very brutal and vague to recommend the career department. [4]

This paper provides a simple yet efficient approach to recommend a career to the users. They have identified three major problems in the traditional collaborative filtering approaches that are used by researchers commonly. These problems are privacy, trust and cold start. Cold start refers to the process of generating recommendations for a new user. The authors have used cosine similarity between the two vectors for finding the similarity in the career approaches and matched a graduate

student with a working professional. They have used python to develop their application and Flask framework for developing the back-end of the application.Their dataset only contains 20,000 entries and 17 columns. They are taking input from the user using a form and then after cleaning, storing data they have calculated a cosine similarity between the previous user preference and jobs available. Their application also have a feedback mechanism on which they are using Natural Language Processing to identify it as a good or bad review. [5]

This paper have a different approach for introducing a career prediction system.Specially the author have focused on the taiwan students, because according to the survey reports in Taiwan more than the half of the students lack knowledge about which path to choose as their career.To solve this problem the author has invented the solution with the help of TOPSIS and Fuzzy Cognitive Map (FCM) which uses typical and atypical factors of choosing the right career for students. Author thinks that this approach is more viable for the counselling as well as operational career prediction system. [6]

The authors have achieved a significantly improved accuracy then of the traditional SVM (Support Vector Machine). Their deep learning approach yield an accuracy of 95% while SVM models are only 62.5% accurate. They have use [38] Myers-Briggs Type Indicator (MBTI) and Five factor model for judging the personality of the user.They have used text normalization, feature extraction and Data pre-processing techniques for the classification of the texts into the personality types. They have used an open dataset for Kaggle which has 8675 samples classified into 16 classes. [7]

Dichotomy	Training Set	Test Set
E - I	1500	898
S - N	1600	698
T - F	3000	1700
J - P	2500	1642

Fig. 2. Authors dataset distribution.

The primary objective of this paper is to develop a career prediction system by emphasizing various crucial elements. These elements include personality traits, interests, and the individual's ability to pursue a particular course. The authors acknowledge the significant problem highlighted by the Council of Scientific and Industrial Research (CSIR), which states that approximately 40 percent of students struggle to select a suitable career path. Consequently, these students face difficulties in their future endeavors, leading to a negative impact on their overall performance. To address this issue, the authors have employed a decision tree-based classifier to

categorize different approaches based on the aforementioned key factors. [8]

In this research paper, authors have proposed a novel method to give career recommendations to the users and provide them their skill gap analysis. They have implemented a Monotonic Linear State Space (MLSS) model that is interpretable. They have feeded binary values which are non-decreasing in nature. They have extracted their data from the Indeed and LinkedIn profiles of the users which contain the basic information of the user. Their approach have sometime outperform the baselines models in certain scenarios. [9]

In this paper author has used data mining and linear classification models.The authors think that re-evaluation in careers is becoming common these days.They think that this re-evaluation can be prevented when the guidance is provide and hence they created a Intuitive Career System. They have previous student data to train models where where the students personalities along with aptitude knowledge is used to predict the career domain.They calculated the personality on basis of student social media accounts.They achieved approximately 77% of accuracy. [10]

The authors have introduced a new method to recommend the jobs to the graduates that do not have any kind of experience. This paper highlights the drawbacks of recommending career path using Collaborative based filtering and proposed P2CF (Personalized Preference Collaborative Filtering Recommendation Algorithm).They have taken account students family economic condition into account using regional economic index. They are also considering graduates preferences. [11] The authors have grouped graduates and their dataset comprise

---

Algorithm 1 Personalized preference collaborative filtering
Input: data, $\mathbf{F}, \mathbf{RFI}, \mathbf{REI}, \mathbf{D}, \lambda, \alpha, \eta$
Output: $\hat{\mathbf{R}}$
1: initialize $\mathbf{U}, \mathbf{V}, \mathbf{A}$ randomly
2: generate the triplet $\mathcal{R}_B$ from training data
3: **repeat**
4:     draw (u, i, j) from $\mathcal{R}_B$
5:     $\hat{R}_{ui} = \mathbf{U}_u \mathbf{V}_i^T + \mathbf{A}_u \mathbf{F}_i^T$
6:     $\hat{R}_{uj} = \mathbf{U}_u \mathbf{V}_j^T + \mathbf{A}_u \mathbf{F}_j^T$
7:     $\hat{R}_{uij} = \hat{R}_{ui} - \hat{R}_{uj}$
8:     **for** θ in $[\mathbf{U}, \mathbf{V}, \mathbf{A}]$ **do**
9:         $\theta = \theta - \alpha(-\frac{e^{-\hat{R}_{uij}}}{1+e^{-\hat{R}_{uij}}} \cdot \frac{\partial \hat{R}_{uij}}{\partial \theta} + \lambda \theta)$
10:     **end for**
11: **until** convergence
12: compute $\hat{R}_{ui}$ according to (13)
13: learn personalized preferences function f according to (15)
14: **for** graduate g in testing data **do**
15:     calculate regional preferences P_{gi} according to (14)
16:     compute $\hat{R}_{gi}$ according to (16)
17: **end for**
18: return $\hat{\mathbf{R}}$

Fig. 3. P2CF Algorithm.

of 1,389 college graduates.

This paper gives us a deep insight about how academic, interpersonal and technical factors of computer science graduated helps in deciding the career path for future. With the help of machine learning algorithm and data mining techniques authors have tried to predict the career path for students and also allocated students teacher who need special attention. Their model was well tested and had shown good performance in content based learning environment. Their prime focus was on Computer Science Engineering Students and how to help them at early stages of their graduation. [12]

In this research paper authors have described about the unique CNN (Convolution Neural Network) architecture to provide the predictions of a person's next job title and salary details. They have implemented this model for predicting the career paths of the person. [13]

| InputLayer | input: | (None, 16, 100) |
| | output: | (None, 16, 100) |

| Conv1D | input: | (None, 16, 100) |
| | output: | (None, 16, 48) |

| MaxPooling1D | input: | (None, 16, 48) |
| | output: | (None, 8, 48) |

| Flatten | input: | (None, 8, 48) |
| | output: | (None, 384) |

| Dropout | input: | (None, 384) |
| | output: | (None, 384) |

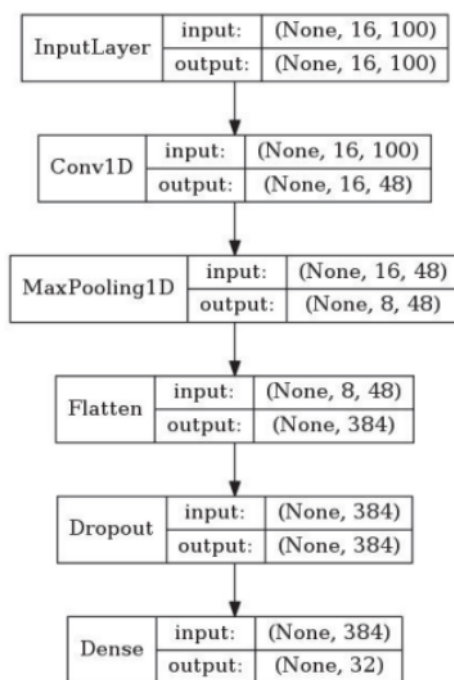| Dense | input: | (None, 384) |
| | output: | (None, 32) |

Fig. 4. Authors Neural Network Architecture

They have used ReLU activation function with a dropout rate of 40%. The authors have tested their model on a dataset of 70,000 resumes, out of which they have used 40,806 resumes. They have achieved an accuracy of 0.488 using position as label. This paper provides the opportunity of researching the use of convolutions neural networks in the field of career path suggestion.

In this research paper, authors have addressed the well known problems with content and collaborative filtering based methods for the job recommendation systems. The problem with these methods are they completely dependent on the users previous behaviour and results i.e. "Cold Start" problem. Authors have proposed a novel approach based on homogeneous Graph-Based Recommendation architecture (GBR). In there graph the nodes represents the jobs and edges represents the similarity. They have collaborated with CareerBuilder platform to utilize their data. They have utilized both content based features and behavioral features of the users. [14]

Email Spam Classification -Authors used four algorithms i.e., ID3, J48, Simple CART and ADTree to classify the email as spam or not. Then he calculated the accuracy over these four algorithms where the J48 classifier outperforms the ID3, CART and ADTree. Classification involves classifying each email as spam or non-spam. Dataset contains 58 attributes with each 4601. Moreover, Simple CART had also given similar results. Later in future he discussed extending its simulation in the WEKA environment in which he can improve his accuracy by keeping the number of instances the same in his dataset but decreasing the number of attributes. Reducing the dimension while keeping the instances the same will bring out the very useful insights in later research fields. [21]

Although Naïve Bayes is a widely recognized classification algorithm, its performance in terms of probability calculation is often considered inadequate. In this study, the author proposes utilizing Naïve Bayes for ranking customers based on their decision to purchase a product or not. However, the author acknowledges the limitations of Naïve Bayes and instead opts for a more advanced and commonly used decision tree learning algorithm called C4.4. The author measures the ranking performance of customers using the area under the receiver operating characteristic curve (AUC). Through empirical and theoretical analysis, the author demonstrates why Naïve Bayes outperforms C4.4 in this specific context. The study focuses on addressing two main challenges: conjunctive literals and 6-of-n concepts. These challenges are then utilized to evaluate the classification performance of Naïve Bayes. [22]

The author of this study explores the concept of querying an individual profile database, which is maintained through a social networking service provider. The goal is to identify similar profiles to the one queried by the user. Throughout the research, the author emphasizes the importance of user privacy in today's digital age. To address this concern, the author proposes three solutions. One solution involves privacy-preserving user profile matching using homomorphic encryption techniques and multiple servers. This approach allows individuals to find matching profiles with the assistance of multiple servers without compromising the privacy of the

query or the profiles being queried. The author thoroughly examines and implements this solution to address privacy issues and evaluates its overall performance. [23]

This paper focuses on addressing the issue of the substantial area required for training and implementing large neural networks. The computational power necessary for such networks is often both costly and inaccessible to many. To mitigate this problem, the authors propose a novel hardware-based approach that reduces the area requirements for neural networks. By employing this approach, the authors successfully decrease the number of physical hidden layers from N to N/2, all while maintaining the accuracy of the neural network. This achievement is made possible through the utilization of multiplexing techniques for the input and output layers. The proposed approach is implemented within the Tensorflow framework and yields a remarkable 42% reduction in required area. [24]

The authors of this study have focused their work on ImageNet detection and ImageNet localization. However, instead of solely aiming to achieve state-of-the-art results, this paper primarily investigates the behaviors of extremely deep networks. It is widely known that training deeper neural networks can be more challenging. To address this issue, the authors propose a residual learning framework designed to facilitate the training of significantly deeper networks compared to those used previously. This is achieved by explicitly reformulating the layers to learn residual features relative to the inputs of each layer. The authors present comprehensive empirical evidence in this study, demonstrating that these residual networks can enhance accuracy while also being comparatively easier to optimize. [25]

This study concentrates on a group of efficient architectures known as MobileNets, which are suitable for mobile and embedded vision applications. MobileNets employ a simplified architecture that utilizes intensity-sensitive separable convolutions to create lightweight deep neural networks. The paper delves into hyperparameters like Latency and accuracy. It also explains the MobileNet network structure and concludes with a discussion of the two model shrinking hyperparameters, width multiplier and resolution multiplier. The study highlights the use of MobileNets in various applications such as recognition in Stanford dogs, large scale geolocalization, compressing large systems with unknown or complex training procedures, and object detection. [26]

In this paper, the authors have introduced an intrinsic taxonomy of contemporary neural network architectures. Their focus is primarily on the versatile applications of deep convolutional neural networks in fields such as computer vision, natural language processing, and speech recognition. The authors have categorized these architectures into seven distinct

classes based on various factors including multi-path structures, width of networks, feature-map exploitation, channel boosting techniques, spatial exploitation, depth of networks, and attention mechanisms. In addition, the authors provide a concise overview of recent architectures, outlining their components, challenges, and applications of deep convolutional layers in each case. [27]

In this paper, the authors focus on introducing a simpler mobile architecture called MobileNetV2, which aims to enhance the performance of mobile models across various tasks and benchmarks, while also accommodating different model sizes. The authors also demonstrate the construction of mobile semantic segmentation models using a condensed version of DeepLabv3. Notably, the paper highlights the significance of eliminating non-linearities in narrow layers to preserve representational power. The authors thoroughly evaluate the trade-offs between accuracy, multiply-add operations (MAdd) as a measure of computational load, actual latency, and the number of parameters. Their evaluation is based on the ImageNet dataset and the MobileNetV2 architecture, resulting in performance improvements across a wide range of aspects. The proposed architecture exhibits superior accuracy compared to real-time detectors and demonstrates reduced complexity in object detection tasks on the COCO dataset. [28]

In this review paper, authors have discussed the various computational challenges that traditional systems face while doing complex tasks like image and video processing, speech recognition, pattern recognition. They are crucial tasks for the development of the more useful and complex intelligent systems. Authors have further discussed the feedforward and signal flow based architectures of the neural networks, additionally they have mentioned the characteristics of the neural networks that are fault tolerance, network structures, distributed memory, learning ability, collective solution, parallel processing. [29]

In this research paper, the authors aim to enhance the performance of neural networks by investigating the concept of model scaling and achieving a balanced approach across network depth, width, and resolution. To achieve this, they propose a novel scaling method that uniformly scales all dimensions (depth, width, and resolution) using a compound coefficient that is effective, simple, and efficient. The authors implement this scaling method specifically for MobileNets and ResNet, effectively scaling up a family of neural networks that outperform traditional convolutional neural networks. They refer to this architecture as EfficientNets. The authors emphasize the significance of this scaling method by demonstrating a notable 2.5% improvement in accuracy, which cannot be achieved through single dimension scaling methods alone. [30]

The author has described the neural network based approach for classification purpose and he has provided a comparison between conventional approach and neural network based approach. They have examined the misclassification costs and mentioned a way of feature variable selection. Author has also focused on the issues of posterior probability estimation. This paper is intended to provide a summary for the recent classification approaches that are generally neural network based, we use different architectures of deep neural networks to classify our data into different classes. [31]

In this research paper, the authors have proposed a new method or approach for object detection tasks. This method is named as YOLO in this research paper. YOLO method uses end to end strategy to detect the probability of class using regression methods, which predicts the bounding boxes and separates class probability that is associated with that bounding box. The YOLO method is extremely fast and it can be easily optimized without involving 20 much computational power. The YOLO method processes images in real-time at 45 frames per second while a more optimized and smaller network called Fast YOLO can process up to 155 frames per second. YOLO is an effective and efficient method that can be used as mobile nets,when usually a lot of computational power is not available but better performance is expected. The YOLO method is prone to errors while detecting objects but it doesn't predict much false positives. [32]

In this paper, authors have provided a comparative and comprehensive study on the different architectures of deep neural networks specifically for multi class pattern classification. The Author has implemented single and multiple neural network systems with different modelling parameter classes using one-against-all, one-against-one, one-against higher-order, and P-against-Q. They have tested 6 different neural network architectures on well known benchmark datasets. They have implemented K-class pattern classification using a single neural network system and K-class pattern classification using multiple neural networks. They have also discussed the time complexity for different neural network systems as well as the learning capabilities of: [33]

- Imbalance datasets.
- Small vs large datasets.
- Large number of pattern classes

In this paper the author has answered some difficult questions like which neural network architecture to choose ? What is the best learning algorithm ? He has described how MLP (multilayer perceptron) architecture is less powerful than BMLP (bridged multilayer perceptron) architecture. Some conclusions are following: [34]

- BMLP is more powerful than conventional MLP architecture because it not only helps in overfitting but also

TABLE 1–NUMBER OF NEURONS/WEIGHTS REQUIRED FOR DIFFERENT PARITY PROBLEMS USING NEURAL NETWORK ARCHITECTURES.

ARCHITECTURE	PARITY-3	PARITY-7	PARITY-15	PARITY-31	PARITY-63
MLP	4/16	8/64	16/256	32/1024	64/4096
BMLP	3/14	5/44	9/152	17/560	33/2144
FCC	2/9	3/27	4/70	5/170	6/399

Fig. 5. Neurons required with different architecture for Parity-N Problem.

produces better results.
- FCC (Fully Connected Cascade) architecture can solve the parity-N problem with the least number of neurons.
- For optimal performance, the number of neurons should be minimum, so that they can perform better on new patterns, rather than overfitting on the trained dataset.
- Traditional EBP algorithm is very slow and may be able to perform better with fewer neurons.
- LM and NBM algorithms are not efficient in working with large neural network sizes.

This paper presents a series of case-based studies aimed at selecting and implementing appropriate neural network architectures. The authors provide a comprehensive analysis of the advantages, disadvantages, and applications of different neural network architectures. To validate their approaches, the authors develop three distinct models: a multilayer perceptron neural network-based (MPNN) solar irradiance forecasting model, an improved backpropagation neural network-based (IBNN) rainfall forecasting model, and an Elman neural network-based (ENN) temperature forecasting model. Each of these neural networks is trained using datasets spanning 5 years for MPNN and IBNN, and 7 years for ENN. It should be noted that this study has limitations in terms of dataset size, making forecasting challenging for larger datasets. Additionally, the authors highlight several interesting hindrances to the growth of artificial neural networks. [35]

- Human rights violations, privacy concerns.
- Scarcity of job opportunities.
- Possibility to endanger humans and environment.

In this survey paper, authors have presented their extensive survey work in the deep learning field. They have discussed the importance of deep learning, various types of neural network architecture and learning algorithms, merits and demerits of these architecture, application of deep learning, different hardware requirements for the training of deep neural networks.

They further added an evaluation matrix and benchmark datasets. They have described various CNN (convolutional neural network) architectures as well as RNN (recurrent neural network) architecture.
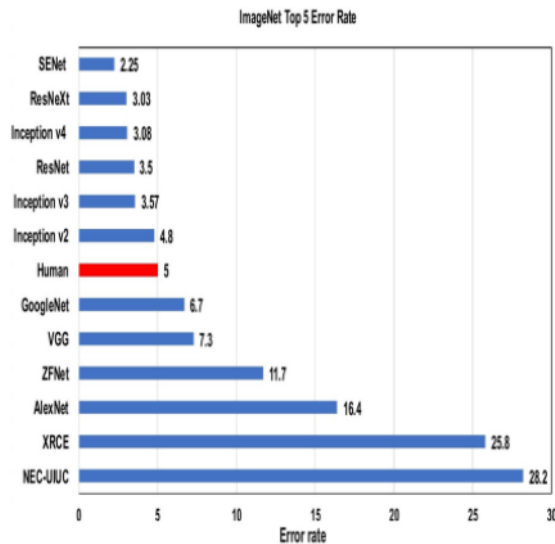

Fig. 6. Comparison of neural network architecture with human performance.

III. ANALYSIS

In [12] research we have found that most researchers have used content based and collaborative filtering methods for career recommendation. These type of models have their own disadvantages such as they are not very flexible, they depend upon the previous history and performance in a specific test organized by the research team. They also suffer from cold start problem that means that their models cannot provide accurate recommendations to a new user.
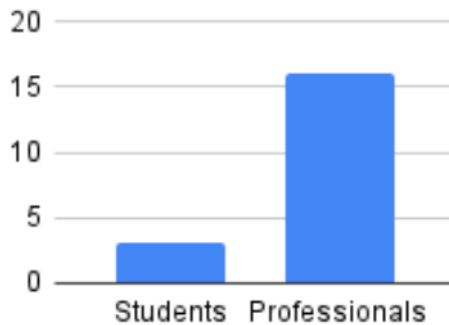

Fig. 7. Students VS Professionals.

Limited research is conducted for early graduate students, most of the research was focused on either working professionals or already graduated students. Some researchers have also implemented their systems or models to suggest a career path to high school and senior secondary students. Mostly researchers have taken similar input features like Personal Information, academic performance and personality aspects.
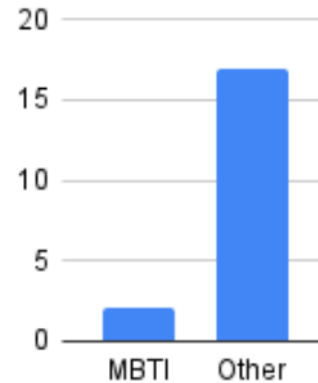

Fig. 8. MBTI vs Other Methods.

For judging the personality few researchers have used [38] [37]Myers Briggs personality test to classify the students on the basis of their personality. While other researchers have used different social media platforms of the students to classify the personality of the students.It is evident from our research that every strategy has taken personality as a key aspect in their model for career suggestions.

Out of 19 papers in this domain only one have implemented Neural networks for this task.Researchers have applied various machine learning algorithms to predict the career for a student. While some of them have introduced their own algorithms and modified methods, it was observed that Neural Networks and Deep learning is still an unexplored domain in this problem statement.

We have also studied various papers on Neural Network architectures and methods to improve the performance of the neural networks for both prediction and classification tasks. Some papers also methods optimized algorithms that do not require so much computational power.

IV. RESEARCH METHODOLOGY

We followed a linear methodology to develop ABES4U, which served as our Final Year Engineering project. The entire development process took us over a year to complete, and we are proud to share that we received two awards and appreciations during this rewarding journey.

Figure 11 illustrates the methodology employed in the development of ABES4U, consisting of the following main steps:

- Data Collection for Friends AI model.
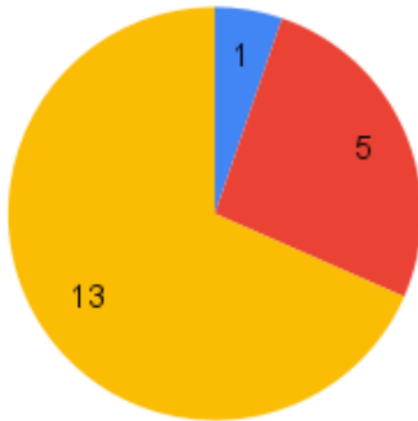- Data analysis and cleaning.

Fig. 9. Neural Networks vs Other Methods.

Fig. 10. Research Methodology of ABES4U.

- Friends AI Training.
- ABES4U Implementation.

Our project commenced with the development of the Friends AI model. Initially, we adopted a clustering-based matching approach, leveraging unsupervised machine learning techniques. However, this approach failed to provide a satisfactory solution. Recognizing the limitations, we made a strategic shift towards a deep learning-based solution.By transitioning to a deep learning-based approach, we aimed to enhance the performance and capabilities of our Friends AI model. This new methodology allowed us to harness the power of neural networks and leverage their ability to learn complex patterns and relationships within the data. As a result, we were able to develop a more robust and effective solution for our project.Throughout the development process, we dedicated significant effort to data analysis and cleaning. This step involved thoroughly examining the collected data, identifying any inconsistencies or anomalies, and applying appropriate techniques to preprocess and prepare the data for training and implementation.With the Friends AI model trained and optimized, we proceeded to integrate it into the ABES4U application. The implementation phase involved developing the necessary software components and user interfaces to create a functional and user-friendly system.Overall, the journey to develop ABES4U has been challenging yet rewarding. We have overcome obstacles, learned valuable lessons, and ultimately achieved a successful outcome by adopting a deep learning-based solution for our Friends AI model.

A. Data Collection

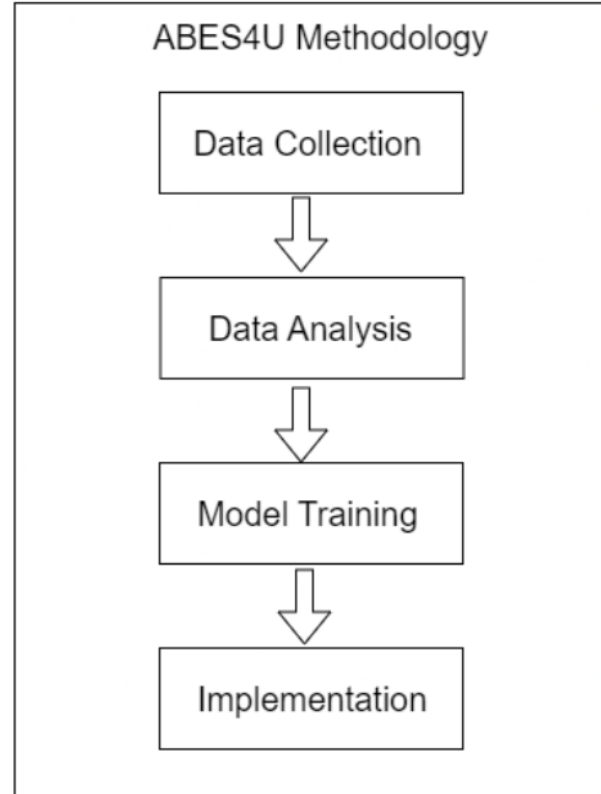We have collected data twice for Friends AI model. Initially we had collected unlabeled data but after shifting towards the neural networks based solution, we have again collected labeled data.Final datset consists of 600 instances and 19 attributes. We have collected data from all computer science and Information technology related branches in our college. It took us 3 months to prepare the dataset.

1) Features Description:

In the development of the Friends AI model, we carefully considered 16 features that encompassed various indicators to quantify the technical, aptitude, and personality aspects of students. Since Friends AI is specifically designed for fresher students who are in the initial phase of their graduation, we selected features that are commonly applicable to both seniors and freshers, allowing us to assess their technical and aptitude capabilities. The following is the list of features we incorporated:

- 10th percentage: This feature captures the academic performance of the student in their 10th grade.
- 11th percentage: Similarly, this feature reflects the academic performance in the 11th grade.
- 12th percentage: This feature represents the academic performance in the 12th grade, which is a crucial factor for admission into higher education.
- Did the student have any guidance?: This feature assesses

whether the student received any guidance or counseling during their academic journey.

- Extrovert VS Introvert 1: This feature aims to measure the student's personality trait on the extroversion-introversion spectrum.
- Extrovert VS Introvert 2: Another measure to capture the student's extroversion or introversion tendencies.
- Extrovert VS Introvert 3: The third feature in this series helps refine the assessment of the student's extroversion or introversion.
- Sensing VS Intuition 1: This feature assesses the student's preference for sensing or intuition, which relates to how they gather information and perceive the world.
- Sensing VS Intuition 2: A complementary feature to further determine the student's inclination towards sensing or intuition.
- Sensing VS Intuition 3: The third feature in this category provides additional insights into the student's sensing or intuition preferences.
- Thinking VS Feeling 1: This feature evaluates whether the student tends to make decisions based on logic and rationality or emotions and personal values.
- Thinking VS Feeling 2: Another measure to refine the assessment of the student's decision-making tendencies.
- Thinking VS Feeling 3: The third feature in this series helps capture the nuances of the student's thinking or feeling approach to decision-making.
- Judging VS Perceiving 1: This feature measures the student's preference for structure and organization versus flexibility and spontaneity.
- Judging VS Perceiving 2: A complementary feature to further determine the student's inclination towards judging or perceiving.
- Judging VS Perceiving 3: The third feature in this category provides additional insights into the student's judging or perceiving tendencies.

By incorporating these 16 features, we aimed to create a comprehensive assessment tool that considers academic performance, guidance, and various personality traits relevant to the students' early stages of higher education. This holistic approach allows the Friends AI model to provide valuable insights and guidance to fresher students in their academic and career journeys.

2) How we are judging personality ?:

The utilization of the Myers Briggs Personality Test (MBTI) has been employed in this study to gain insights into the personalities of the students [38]. The Myers-Briggs Personality Test is a well-known and widely utilized assessment tool designed to categorize individuals based on their unique personality traits. Developed by Katharine Cook Briggs and her daughter Isabel Briggs Myers, this test draws inspiration from the theories put forth by the esteemed psychologist Carl Jung. The assessment evaluates individuals across four primary dimensions of personality: extraversion (E) versus introversion (I), sensing (S) versus intuition (N), thinking (T) versus feeling (F), and judging (J) versus perceiving (P).

Through a series of questions, participants are assigned a four-letter code representing their specific personality type, such as ISTJ (Introverted-Sensing-Thinking-Judging) or ENFP (Extraverted-Intuitive-Feeling-Perceiving). The Myers-Briggs Personality Test offers valuable insights into an individual's preferences, tendencies, and potential career paths, fostering a deeper understanding of oneself and facilitating personal growth and development. However, it is important to acknowledge that the test is subjective and should be regarded as a tool for self-reflection rather than an absolute measure of personality.

B. Data Analysis

After collecting data from the users we have cleaned the data and transformed the data. After removing outliers, mainly we have identified the personality class of the student on the basis of 12 questions that were asked to the user. There are 4 classes in the MBTI test and we have defined 3 questions for each of these 4 classes. This identified class is then passed to the Friends AI model along with the other 4 features.

Secondly we have labeled our output classes. After a thorough research we have identified 18 domains that can be pursued by a fresher. These domains are identified on the basis of off campus internships grabbed by our seniors. The domains list is following:

- Frontend using HTML, CSS and JavaScript.
- Frontend using React.
- Backend using JavaScript frameworks.
- Backend using Python Framework.
- Backend using Java Frameworks.
- Competitive Programming using C++.
- Competitive Programming using Python.
- Competitive Programming using Java.
- Full Stack (JavaScript Based Tech Stack).
- Full Stack (Python Based Tech Stack).
- Full Stack (Java Based Tech Stack).
- Data Science / Artificial Intelligence / Machine Learning.
- Blockchain.
- Cyber Security.
- Application Development using Kotlin.
- Application Development using Java.
- Application Development using Swift.
- Application Development using Flutter.

C. Model (Friends AI) Training

Figure 12 describes the architecture of Friends AI neural Network. We have considered Adam Optimizer and sparse_categorical_crossentropy as our loss metric. It was trained for 500 epochs. It has achieved has the highest accuracy of 92% on the test dataset. We have split our dataset into 90% train datset and 10% test datset.

D. Implementation

We have successfully incorporated all our research findings into the development of ABES4U, an innovative web application that effectively utilizes our extensive research work.
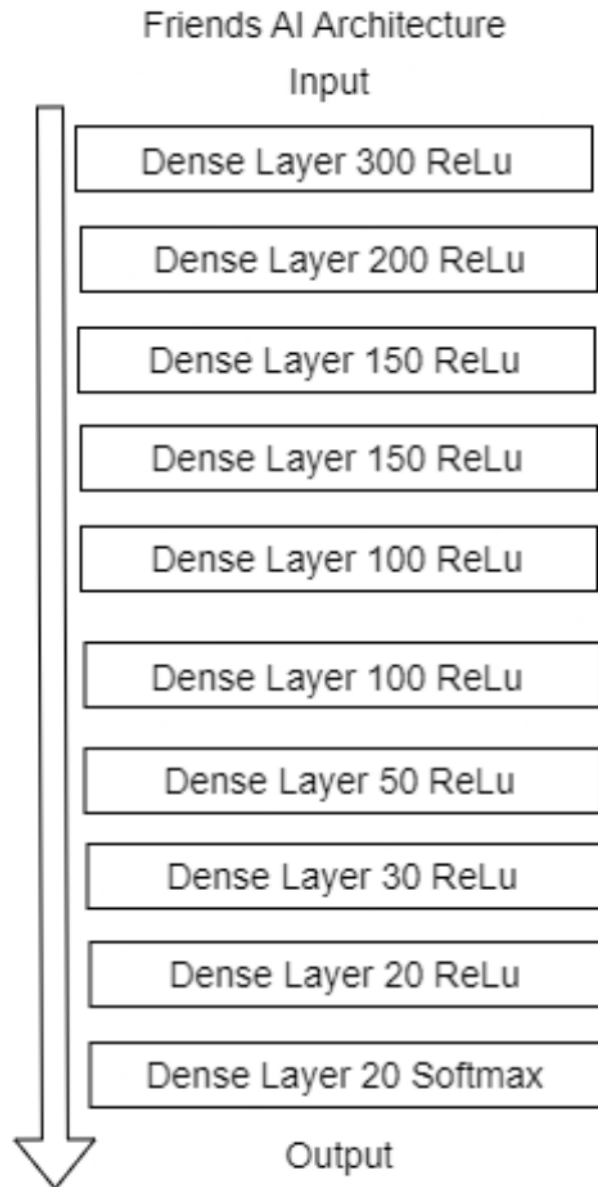
Fig. 11. Friends AI Architecture.

store and manage the application's data effectively.

- HTML5, CSS3, Javascript: Leveraged for frontend development, enabling us to craft an engaging and interactive user interface.
- Jupyter Lab Tool: Used for comprehensive data analysis, facilitating the exploration and interpretation of complex datasets.
- TensorFlow 2: Employed to develop the Friends AI component, utilizing the power of deep learning for enhanced functionality.

Throughout the development process, we extensively utilized Class-Based Views, an inherent capability provided by Django 4, which greatly facilitated the implementation of crucial CRUD (create, read, update, delete) operations within ABES4U. Additionally, Django forms were employed to seamlessly capture user input for the Friends AI model, ensuring a smooth user experience. To prioritize security, we implemented Custom User Authentication specifically tailored for ABES4U, ensuring the protection of user data and access controls.

ABES4U serves as a responsive full-stack deep learning-enabled dashboard, offering a range of features we previously discussed. To bring this application to life, we utilized a carefully selected tech stack, which includes:

- Django 4.1.4: Employed for backend development, providing a robust framework for creating dynamic web applications.
- MySQL: Utilized as the database management system to

The implementation phase of ABES4U spanned a period of six months, during which we meticulously designed, developed, and refined the application to meet our desired standards. This allowed us to create a feature-rich and robust web application that leverages our research findings to provide users with an exceptional experience.
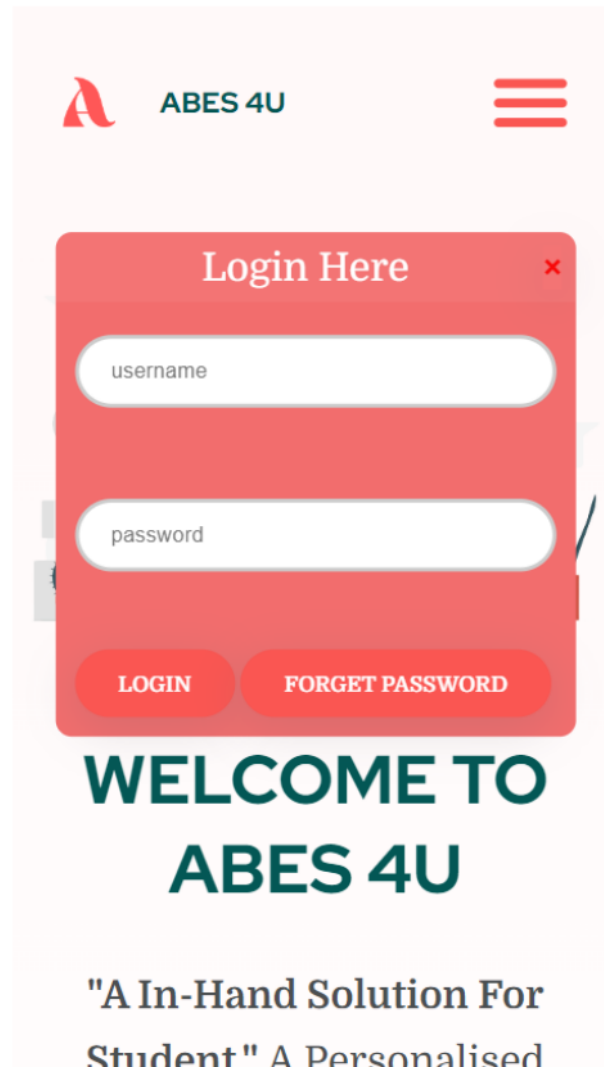
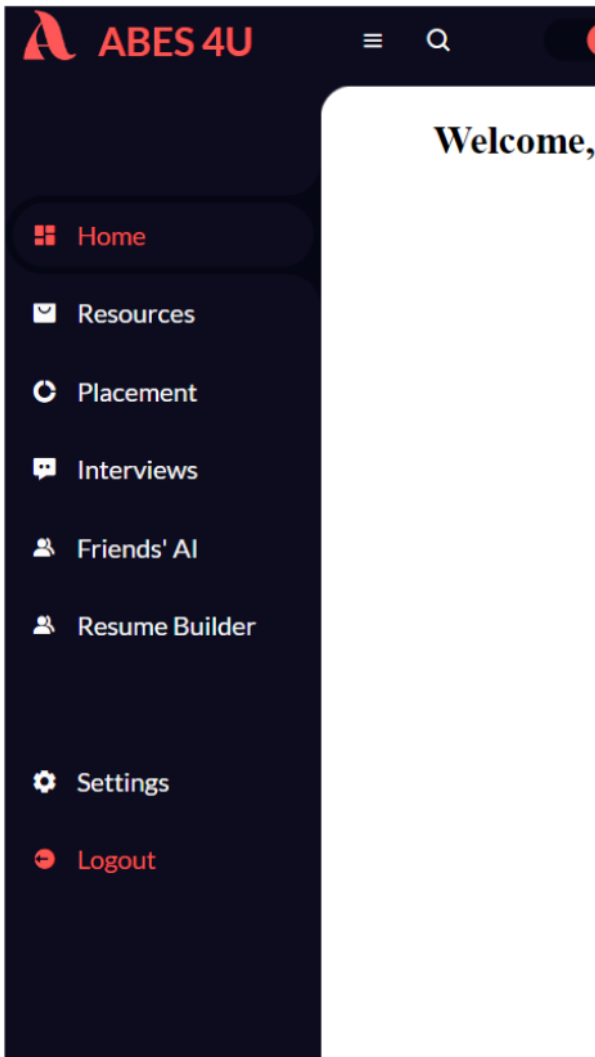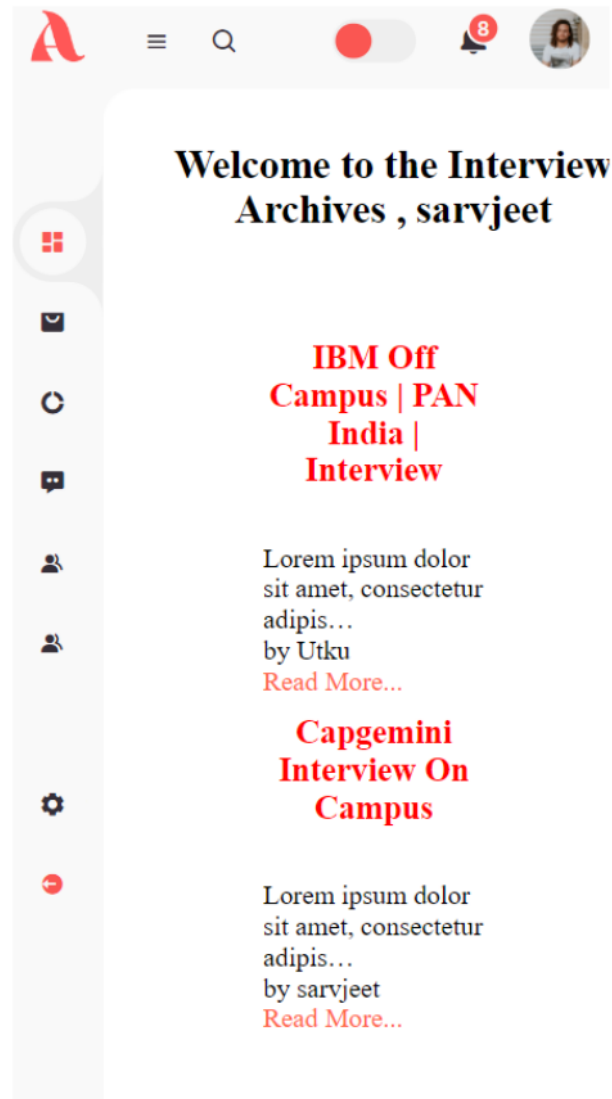Fig. 12. Landing Page.

Fig. 13. Login Page.

Fig. 14. Dashboard.

ABES 4U

Welcome,

- Home
- Resources
- Placement
- Interviews
- Friends' AI
- Resume Builder
- Settings
- Logout

Welcome to the Interview Archives , sarvjeet

IBM Off Campus | PAN India | Interview

Lorem ipsum dolor sit amet, consectetur adipis…
by Utku
Read More...

Capgemini Interview On Campus

Lorem ipsum dolor sit amet, consectetur adipis…
by sarvjeet
Read More...

Fig. 15. Interview Archives.

Fig. 16. Friends AI.

V. CONCLUSION

We have full confidence that ABES4U will give students a competitive edge over their peers, allowing them to excel in off-campus competitions and other academic endeavors. Moreover, the benefits of ABES4U are not limited to computer science students alone. This approach can be extended to students from various other branches of study, enabling them to explore practical aspects related to their respective fields. By leveraging the application's features, students from different domains can expand their knowledge, gain hands-on experience, and sharpen their skills, ultimately boosting their overall performance and success.Furthermore, the approach utilized in ABES4U can be effectively employed in scenarios where matching between two groups of people is required based on specific, well-defined features. This could be particularly valuable in various contexts, such as forming project teams, organizing collaborative research efforts, or creating diverse workgroups. By leveraging the application's capabilities, users can streamline the process of finding suitable matches and ensure optimal group composition.

REFERENCES

[1] Patel, B., Kakuste, V., Eirinaki, M. (2017). CaPaR: A Career Path Recommendation Framework. 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService).

[2] L. S. Katore, B. S. Ratnaparkhi and J. S. Umale, "Novel professional career prediction and recommendation method for individual through analytics on personal traits using C4.5 algorithm," 2015 Global Conference on Communication Technologies (GCCT), Thuckalay, India, 2015

[3] Gugnani, A., Reddy Kasireddy, V. K., Ponnalagu, K. (2018). Generating Unified Candidate Skill Graph for Career Path Recommendation. 2018 IEEE International Conference on Data Mining Workshops (ICDMW).

[4] S. Vignesh, C. Shivani Priyanka, H. Shree Manju and K. Mythili, "An Intelligent Career Guidance System using Machine Learning," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021

[5] Yadalam, T. V., Gowda, V. M., Kumar, V. S., Girish, D., M, N. (2020). Career Recommendation Systems using Content based Filtering. 2020 5th International Conference on Communication and Electronics Systems (ICCES).

[6] Chen, Y.-T., Peng, W.-C., Yu, H.-Y. (2018). Identify Key Factors for Career Choice by Using TOPSIS and Fuzzy Cognitive Map. 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS).

[7] Harrouk, A. I., Barbar, A. M. (2018). A Psycholinguistic Approach to Career Selection Using NLP with Deep Neural Network Classifiers. 2018 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET).

[8] Shankhdhar, A., Agrawal, A., Sharma, D., Chaturvedi, S., Pushkarna, M. (2020). Intelligent Decision Support System Using Decision Tree Method for Student Career. 2020 International Conference on Power Electronics IoT Applications in Renewable Energy and Its Control (PARC).

[9] Ghosh, A., Woolf, B., Zilberstein, S., Lan, A. (2020). Skill-based Career Path Modeling and Recommendation. 2020 IEEE International Conference on Big Data (Big Data).

[10] Rangnekar, R. H., Suratwala, K. P., Krishna, S., Dhage, S. (2018). Career Prediction Model Using Data Mining and Linear Classification. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[11] Q. Zhou, F. Liao, L. Ge and J. Sun, "Personalized Preference Collaborative Filtering: Job Recommendation for Graduates," 2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, 2019, pp. 1055-1062, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00203.

[12] Arafath, M. Y., Saifuzzaman, M., Ahmed, S., Hossain, S. A. (2018). Predicting Career Using Data Mining. 2018 International Conference on Computing, Power and Communication Technologies (GUCON).

[13] He, M., Shen, D., Zhu, Y., He, R., Wang, T., Zhang, Z. (2019). Career Trajectory Prediction based on CNN. 2019 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI).

[14] Shalaby, W., AlAila, B., Korayem, M., Pournajaf, L., AlJadda, K., Quinn, S., Zadrozny, W. (2017). Help me find a job: A graph-based approach for job recommendation at scale. 2017 IEEE International Conference on Big Data (Big Data).

[15] Qamhieh, M., Sammaneh, H., Demaidi, M. N. (2020). PCRS: Personalized Career-Path Recommender System for Engineering Students. IEEE Access, 8, 214039–214049.

[16] Maurya, A., Telang, R. (2017). Bayesian multi-view models for member-job matching and personalized skill recommendations. 2017 IEEE International Conference on Big Data (Big Data).

[17] Jain, H., Kakkar, M. (2019). Job Recommendation System based on Machine Learning and Data Mining Techniques using RESTful API and Android IDE. 2019 9th International Conference on Cloud Computing, Data Science Engineering (Confluence).

[18] Peng Yi, Yang, C., Chen Li, Zhang, Y. (2016). A job recommendation method optimized by position descriptions and resume information. 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC).

[19] Almalis, N. D., Tsihrintzis, G. A., Karagiannis, N., Strati, A. D. (2015). FoDRA — A new content-based job recommendation algorithm for job seeking and recruiting. 2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA).

[20] Katiyar, S., Walia, H., Kumar, S. (2020). Personality Classification System using Data Mining. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO).

[21] Kumar, Aman Assistant, Sharma Sahni, Suruchi. (2011). A Comparative Study of Classification Algorithms for Spam Email Data Analysis. International Journal on Computer Science and Engineering 3.

[22] Zhang H. and Su J., "Naive bayesian classifiers for ranking", 15th European Conference on Machine Learning, ECML 2004. Accessed online http://www.cs.unb.ca/profs/hzhang/publications/NBRanking.pdf

[23] Yi, Xun Bertino, Elisa Rao, Fang-Yu Bouguettaya, Athman. (2016). Practical privacy-preserving user profile matching in social networks. 373-384. 10.1109/ICDE.2016.7498255.

[24] Khalil, Kasem Eddash, Omar Kumar, Ashok Bayoumi, Magdy. (2018). An Efficient Approach for Neural Network Architecture. 745-748. 10.1109/ICECS.2018.8617887.

[25] K. He, X. Zhang, S. Ren, J. Sun, 'Deep Residual Learning for Image Recognition', CoRR, . abs/1512.03385, 2015.

[26] A. G. Howard .., 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications', CoRR, . abs/1704.04861, 2017.

[27] Khan, A., Sohail, A., Zahoora, U. et al. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53, 5455–5516 (2020). https://doi.org/10.1007/s10462-020-09825-6

[28] Mark Sandler, Andrew G. Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen: Inverted Residuals and Linear Bottlenecks: Mobile Networks for Classification, Detection and Segmentation. CoRR abs/1801.04381 (2018)

[29] K. I. Ahamed, "A Study on Neural Network Architectures," The International Institute for Science, Technology and Education (IISTE), Sep. 2016.

[30] Mingxing Tan, Quoc V. Le: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. CoRR abs/1905.11946 (2019).

[31] Zhang, Peter. (2000). Neural Networks for Classification: A Survey. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 30. 451 - 462. 10.1109/5326.897072.

[32] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, Ali Farhadi:

[33] You Only Look Once: Unified, Real-Time Object Detection. CoRR abs/1506.02640 (2015)

[34] Ou, G., Murphey, Y.L. (2007). Multi-class pattern classification using neural networks. Pattern Recognit., 40, 4-18.

[35] M. Madhiarasan M. Louzazni, 'Analysis of Artificial Neural Network: Architecture, Types, and Forecasting Applications', Journal of Electrical and Computer Engineering, . 2022, . 5416722, 2022.

[36] L. Alzubaidi .., 'Review of deep learning: concepts, CNN architectures, challenges, applications, future directions', Journal of Big Data, . 8, . 1, . 53, 2021.

[37] https://www.thoughtco.com/myers-briggs-personality-types-4686022

[38] https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/original-research.htm

final paper

PRIMARY SOURCES

1	www.hindawi.com Internet Source	1%
2	Rishabh Bajpai, Rahul Kumar Sharma, Rohit Vashisht. "Identification of peak hour and active member on WhatsApp group using Python", 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022 Publication	1%
3	Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen. "MobileNetV2: Inverted Residuals and Linear Bottlenecks", 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018 Publication	<1%
4	www.enggjournals.com Internet Source	<1%
5	Kasem Khalil, Omar Eldash, Ashok Kumar, Magdy Bayoumi. "An Efficient Approach for Neural Network Architecture", 2018 25th IEEE	<1%

International Conference on Electronics, Circuits and Systems (ICECS), 2018
Publication

6   Lecture Notes in Computer Science, 2004.   <1 %
    Publication

7   Tiffany Maldonado, Dusya Vera, William D.   <1 %
    Spangler. "Unpacking humility: Leader
    humility, leader personality, and why they
    matter", Business Horizons, 2021
    Publication

8   "Proceedings of International Conference on   <1 %
    Trends in Computational and Cognitive
    Engineering", Springer Science and Business
    Media LLC, 2021
    Publication

9   Submitted to Liverpool John Moores   <1 %
    University
    Student Paper

10 Reza Eslamloueyan. "Designing a hierarchical   <1 %
    neural network based on fuzzy clustering for
    fault diagnosis of the Tennessee–Eastman
    process", Applied Soft Computing, 2011
    Publication

11 dokumen.pub   <1 %
    Internet Source

12 Atharva Patil, Divyansh Suwalka, Aryan   <1 %
    Kumar, Gaurav Rai, Jayita Saha. "A Survey on

Artificial Intelligence (AI) based Job Recommendation Systems", 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2023
Publication

13 sportold.ubbcluj.ro
Internet Source   <1%

14 www.bartleby.com
Internet Source   <1%

15 1library.co
Internet Source   <1%

16 Guobin Ou, Yi Lu Murphey. "Multi-class pattern classification using neural networks", Pattern Recognition, 2007
Publication   <1%

17 Submitted to ABES Engineering College
Student Paper   <1%

18 www.mdpi.com
Internet Source   <1%

19 Prem N. Ramkumar, Sergio M. Navarro, Heather S. Haeberle, Jaret M. Karnuta et al. "Development and Validation of A Machine-Learning Algorithm After Primary Total Hip Arthroplasty: Applications to Length of Stay   <1%

and Payment Models", The Journal of Arthroplasty, 2018
Publication

20   Submitted to SUNY, Binghamton   <1%
     Student Paper

21   Omid Moztarzadeh, Mohammad (Behdad)   <1%
     Jamshidi, Saleh Sargolzaei, Fatemeh Keikhaee
     et al. "Metaverse and Medical Diagnosis: A
     Blockchain-Based Digital Twinning Approach
     Based on MobileNetV2 Algorithm for Cervical
     Vertebral Maturation", Diagnostics, 2023
     Publication

22   arxiv.org   <1%
     Internet Source

23   intellipaat.com   <1%
     Internet Source

24   www.semanticscholar.org   <1%
     Internet Source

25   "Data Driven Approach Towards Disruptive   <1%
     Technologies", Springer Science and Business
     Media LLC, 2021
     Publication

26   "Handbook of Computational Social Science   <1%
     for Policy", Springer Science and Business
     Media LLC, 2023
     Publication

27 "ICDSMLA 2019", Springer Science and Business Media LLC, 2020
Publication

<1 %

28 "Operations Research and Enterprise Systems", Springer Science and Business Media LLC, 2020
Publication

<1 %

29 Fabio Cisne Ribeiro, Raphael Torres Santos Carvalho, Paulo Cesar Cortez, Victor Hugo C. De Albuquerque et al. "Binary Neural Networks for Classification of Voice Commands From Throat Microphone", IEEE Access, 2018
Publication

<1 %

30 advancesincontinuousanddiscretemodels.springeropen
Internet Source

<1 %

31 pubs.sciepub.com
Internet Source

<1 %

32 pure.rug.nl
Internet Source

<1 %

33 repositori.uji.es
Internet Source

<1 %

34 www.spiedigitallibrary.org
Internet Source

<1 %

35 Mahdi Nabiyouni, Delasa Aghamirzaie. "A Highly Parallel Multi-class Pattern

<1 %

Classification on GPU", 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012), 2012

Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		