“THE FREEBIRD”- A WEB-BASED APPLICATION WITH A JOB RECOMMENDATION SYSTEM

A CAPSTONE PROJECT REPORT

Submitted in partial fulfillment of the
requirement for the award of the
Degree of

BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE ENGINEERING WITH SPL. in DATA ANALYTICS

by

RASMIKA BILLA (18BCD7021)

Under the Guidance of

Dr. SANDIPAN MAITI

SCHOOL OF COMPUTERS ENGINEERING
VIT-AP UNIVERSITY
AMARAVATI- 522237

JANUARY 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

A freelance worker or Freelancer is a person who is self-employed and not necessarily committed to a particular employer long-term. Freelancers are sometimes represented by a company or a temporary agency that resells freelance labor to clients, others work independently by using professional associations or websites to get work[1].

The workers are usually forced into nonstandard work arrangements, for others, loosening attachment to organizations provides the opportunities to accommodate their wants, preferences, and individualized lifestyles[2]. Thus, the application will predict the skill levels for various workers/customers based on their skill set and project work. This prediction will be in a generalized form of getting into freelancing and establishing a network. The app would allow users to post, manage and delete their services or requests.

In another scenario, if the customer wishes to search for jobs and wants to work within a community then the skillset of the customer can further be implemented in the developed job recommendation system to draw suitable job postings. The project work is for dual purposes helping the workers in either way out.

The problems discussed above have led to a major concern especially during times of pandemic and an immediate solution for this problem has become necessary in order to sustain the economic development.

1.1 OBJECTIVES

The following are the objectives of this project:

- Creating an application to provide effective assistance to people in search of work.
- Client and freelancer registration and login to make the application secure.
- Applying crud operations such as a client or a freelancer can add, post or delete services or requests.

- On the other end, any user can view all the services or any service details he/she is interested in.
- Predicting the best three job recommendations and required postings, if the user wishes to work for a company based on their qualification, location and work experience[3].
- Predicting the percentage chance of getting selected by depicting the scale of vacancies, recommendation index and application index.

1.2 BACKGROUND AND LITERATURE SURVEY

India had (according to the data provided by Mahesh Vyas of the Centre for Monitoring Indian Economy) around 403.5 million employed people and around 35 million (or 3.5 crores) openly unemployed people in the country just before the Covid-19 crisis, at the end of the 2019-20 financial year.

The unemployment rate in India rose sharply and touched a six-month high. The fact that the unemployment rate has been rising since September 2020 despite faster recovery has raised concerns. India's unemployment rate rose sharply to 9.1 percent in December 2020 even as economic activity continues to increase which is one of the major concerns. The project aims to solve the major concerns that are listed and is motivated to decrease unemployment.

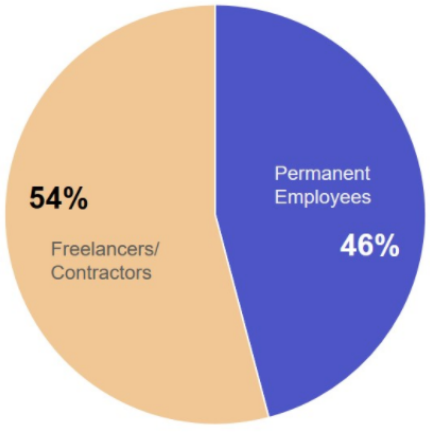

Figure 1. Google workforce of 220.000 workers

Google's workforce consisted of more than a lakh contractors/freelancers, which was more than the number of permanent employees on its payroll as of March 2019. The above figure depicts the workforce distribution at Google.

The rise in the adoption of remote teams has made freelancing a viable career path for skilled professionals in regions where domestic opportunities are low. The below figure amounts to the number of countries growing based on the freelancing markets. Among all these India ranks in the Top 10 countries with the fastest-growing earnings for freelancers.

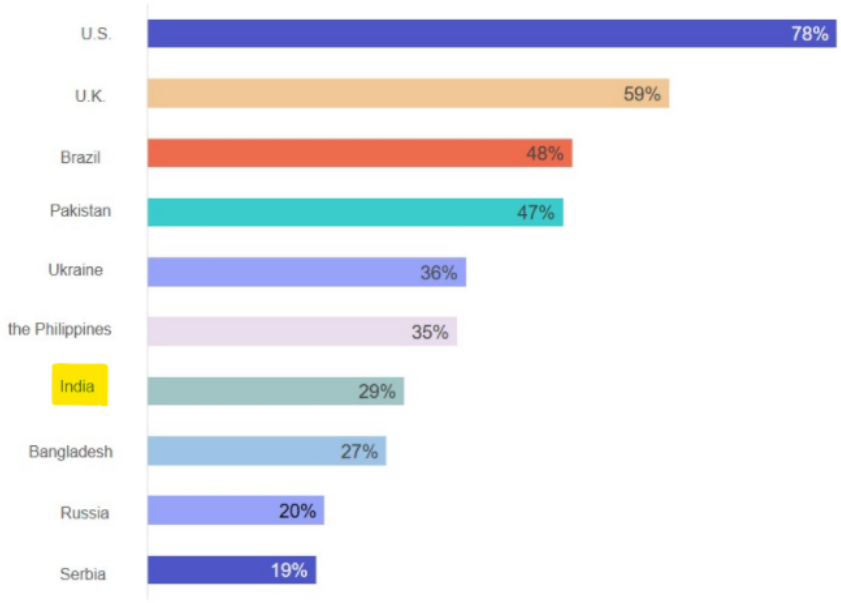

Figure 2. Fastest growing freelance markets

The data collected and analysed draws a decision that there is big concern among the markets in dual ways. The sectors are diverging into different kinds of markets where one is establishing productiving individually and the other within a community. To make balance between the both, the application developed provides a platform for the customers interested in freelancing as well as job postings[4]. A person who pursues individual productivity can go through freelancing and expertise in one to more fields whereas the other can look for best suitable job posts and apply for the ones where they have a higher chance of productivity.

1.3 ORGANIZATION OF THE REPORT

The remaining chapters of the project report are described as follows:

- Chapter 2 contains the proposed system, methodology, hardware, and software details.
- Chapter 3 gives a detailed explanation and implementation of the project.
- Chapter 4 discusses the results obtained after the project was implemented.
- Chapter 5 concludes the report.
- Chapter 6 consists of codes.
- Chapter 7 gives references.

CHAPTER 2

THE FREEBIRD APPLICATION AND JOB RECOMMENDATION SYSTEM

This Chapter describes the proposed system, working methodology, the dataset used, the pre-processing techniques applied on the dataset, software and hardware details of the project work.

2.1 PROPOSED APPLICATION/SYSTEM

The application is intended to help the users in trial ways. The user registered as a freelancer can use the web-based system to predict their skill level, post their services, manage their services, ad lastly have person requests as well. The user registered as client can easily view the services posted, and additionally can generate a customised requests explaining his/her requirements. The user wanting to work within a community can use the job recommendation system[5] to get their best three skills acknowledged and apply for the job posts recommended with a given percentage change of getting placed in that particular job.

2.2 WORKING METHODOLOGY

The application is connected to SQL[6] server which helps secure and maintain the data. All the webpages connected remain active till the work is executed. The prediction models are implemented behind the loop with reference to Flask for its integration.

2.2.1 DATASETS

The datasets for skill level prediction and job recommendation system are different and connected to the aspect of employment. The dataset analysed for predicting skill level of the user is FIVER dataset. The dataset has about six columns and eight thousand rows. The below figure.3 depicts a snapshot of the dataset. The second dataset implemented for the project work is JOB_POSTS dataset. The dataset consists of about twenty-four columns and thirty thousand rows. The below figure.4 depicts a snapshot of the dataset. Both the datasets are taken from Kaggle resource.

	Category	Subcat	name	price	stars	votes
0	Programming & Tech	Data Analysis & Reports	build automated and insightful power bi report	137.76	5.0	10
1	Lifestyle	Greeting Cards & Videos	get kermit to personalize a video birthday gre...	4.44	5.0	1k+
2	Programming & Tech	Website Builders & CMS	build or redesign your existing squarespace we...	66.66	5.0	68
3	Lifestyle	Cooking Lessons	create original eastern food cooking video recipe	17.78	4.9	5
4	Writing & Translation	Legal Writing	write gdpr privacy policy and terms and condit...	8.89	5.0	2

Figure 3. FIVER Dataset

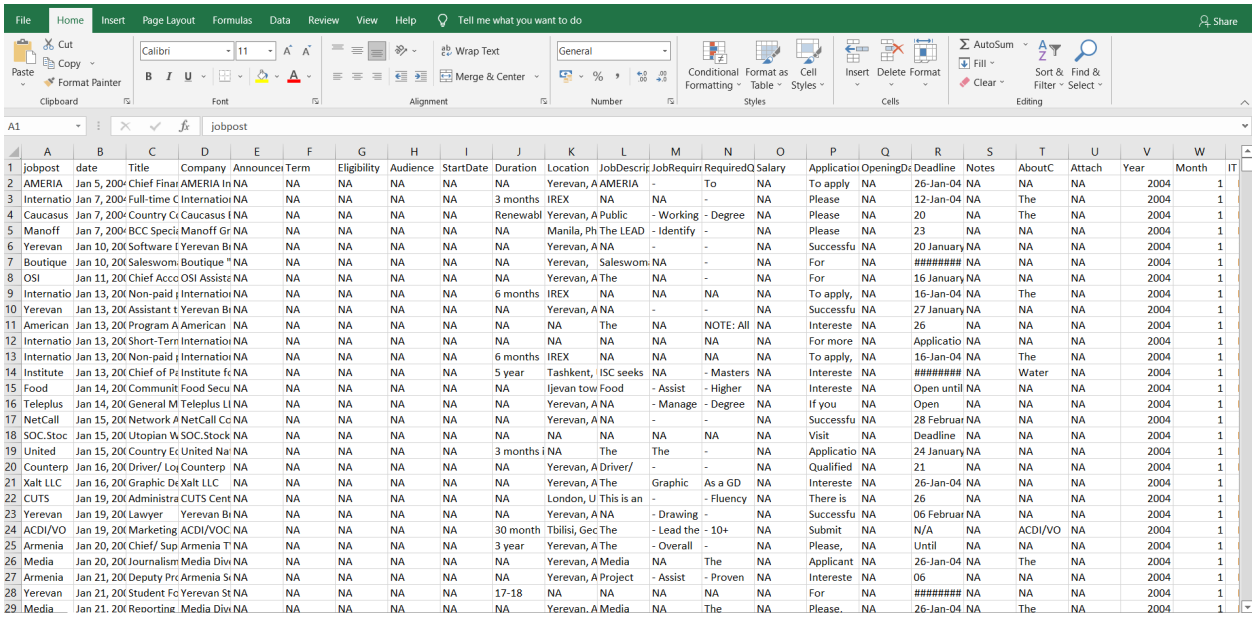

Figure 4. JOB_POSTS Dataset

2.2.2 PRE-PROCESSING TECHNIQUES

The technique used to modify the FIVER dataset is ordinal encoding[7] technique. The technique converts the unique categorical values to numerical values for better analysis. The figure below shows the implementation and output of ordinal encoding on the dataset.

```
from sklearn.preprocessing import OrdinalEncoder

ord_enc = OrdinalEncoder()

df['Category'] = ord_enc.fit_transform(df[['Category']])
```

```
df['Subcat'] = ord_enc.fit_transform(df[['Subcat']])
```

```
df
```

	Category	Subcat	stars
0	4.0	33.0	5.0
1	2.0	52.0	5.0
2	4.0	122.0	5.0
3	2.0	27.0	4.9
4	6.0	60.0	5.0

Figure 5. Ordinal Encoding

The JOB_POSTS dataset is the second data in which the technique used to preprocess is Customised Tokenizer[8] and Label Encoding. The technique is used to convert the textual data into simple format and extract the required skills and experience for predicting job roles for the user. The below figure is a code snippet implemented for converting textual data.

```
from nltk import word_tokenize
from nltk.stem import WordNetLemmatizer
class LemmaTokenizer(object):
    def __init__(self):
        # lemmatize text - convert to base form
        self.wnl = WordNetLemmatizer()
        # creating stopwords list, to ignore lemmatizing stopwords
        self.stopwords = stopwords.words('english')
    def __call__(self, doc):
        return [self.wnl.lemmatize(t) for t in word_tokenize(doc) if t not in self.stopwords]

# removing new line characters, and certain hypen patterns
df['RequiredQual']=df['RequiredQual'].apply(lambda x: x.replace('\n', ' ').replace('\r', '').replace('- ', ''). replace(' - ', ' to '))
```

Figure 6. Custom Tokenizer

2.3 SYSTEM DETAILS

This section describes the software and hardware details of the system:

2.3.1 SOFTWARE

- The frontend development for the application is completed with the help of HTML, CSS, and JavaScript as the programming languages.

- The backend development of the application is implemented using Php, MYSQL, and Flask for python integration.
- Google Colab is used for predicting and implementing various machine learning algorithms and Jupiter Notebook for developing the job recommendation system.
- Windows is used as the operating system.

2.3.2 HARDWARE

The programs were executed on a HP laptop with Intel core i5 8th generation CPU and 8 GB RAM with 256 GB SSD hard disk.

CHAPTER 3

DISCUSSION

The designs and pages are accomplished with HTML5[9] and CSS3[10]. The Index Page from which users can Login, Signup, Explore services, and go through contact and about web pages as well.

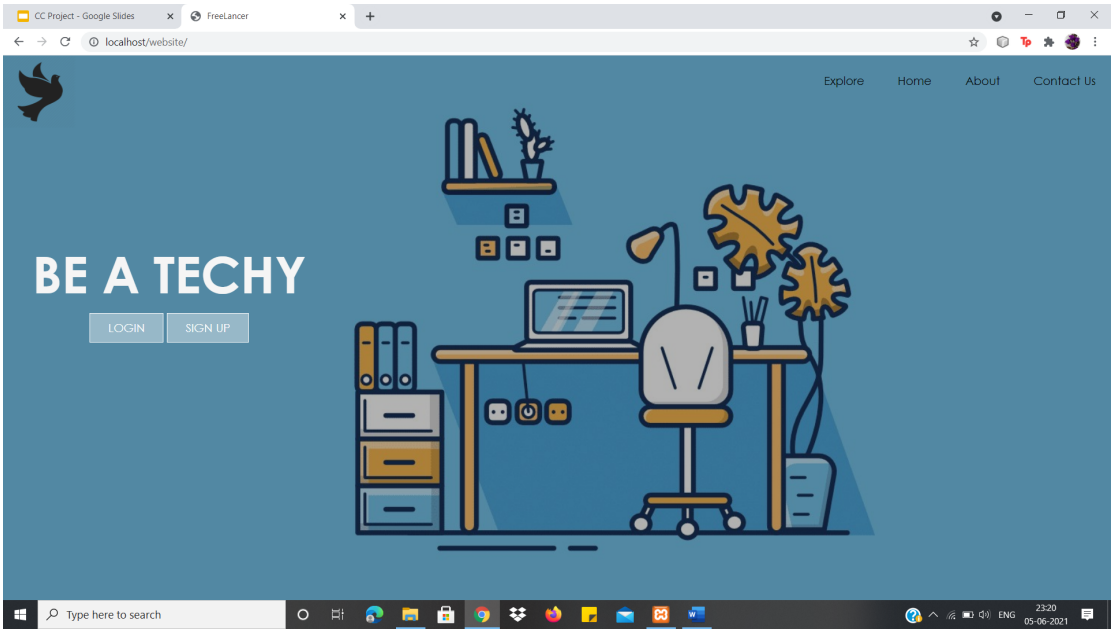

Figure 7. Index page

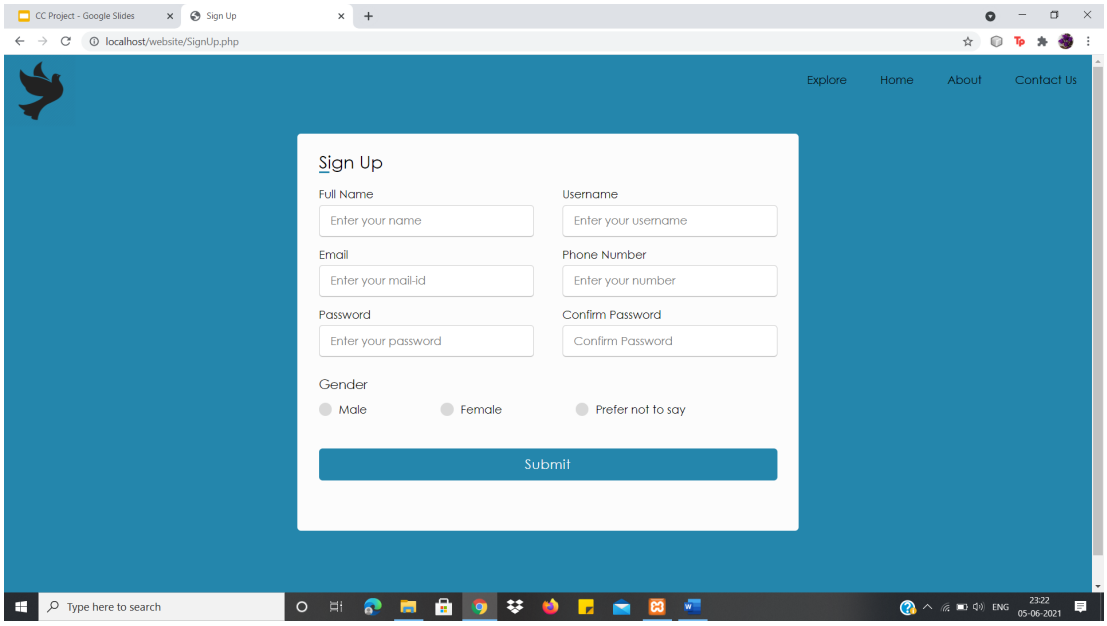

Figure 8. Signup Page

The user has to re-enter if the password and the confirm password is not matched. Also, if the username or password is incorrect the user is notified during the login process.

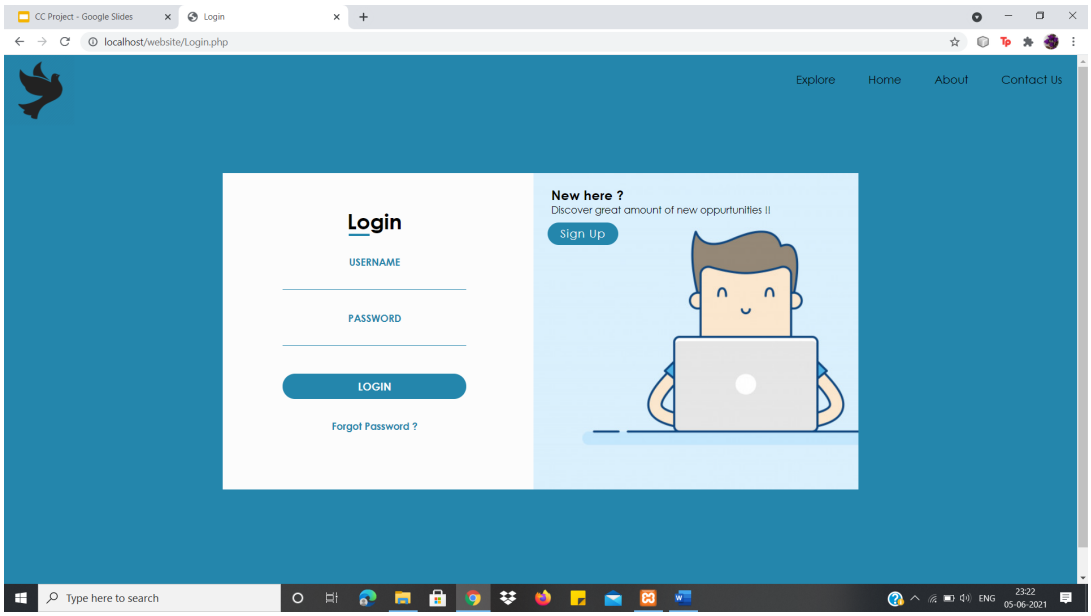

Figure 9. Login Page

The Home page where the user can either become a buyer or a seller depending on their necessities. The user can register for both as well.

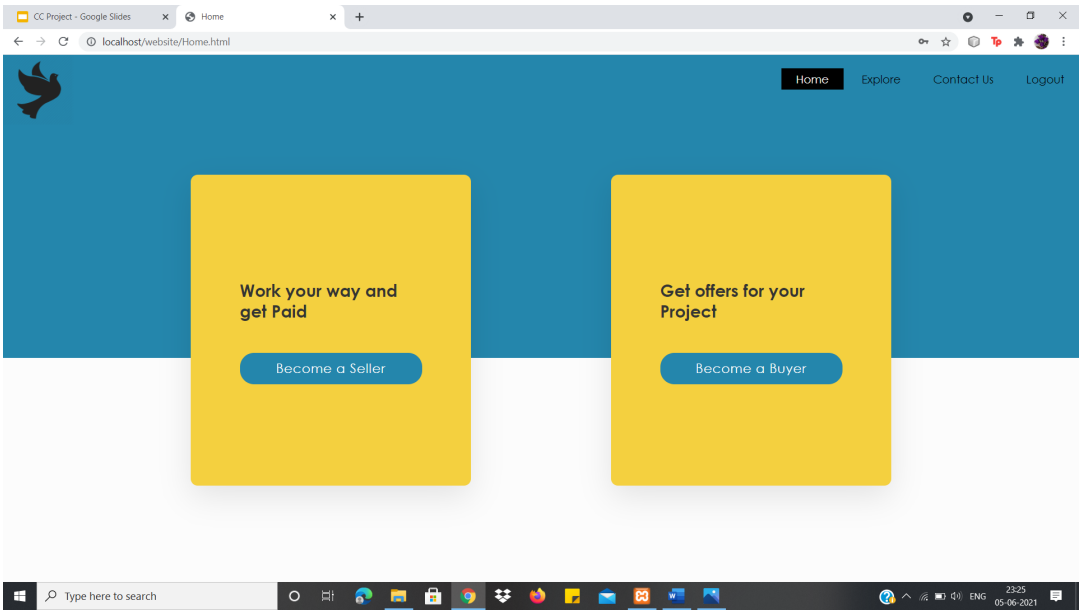

Figure 10. Home Page

The user registered as a Seller can go through their skill level, Publish/Post a service, manage their posts as well as manage the client requests.

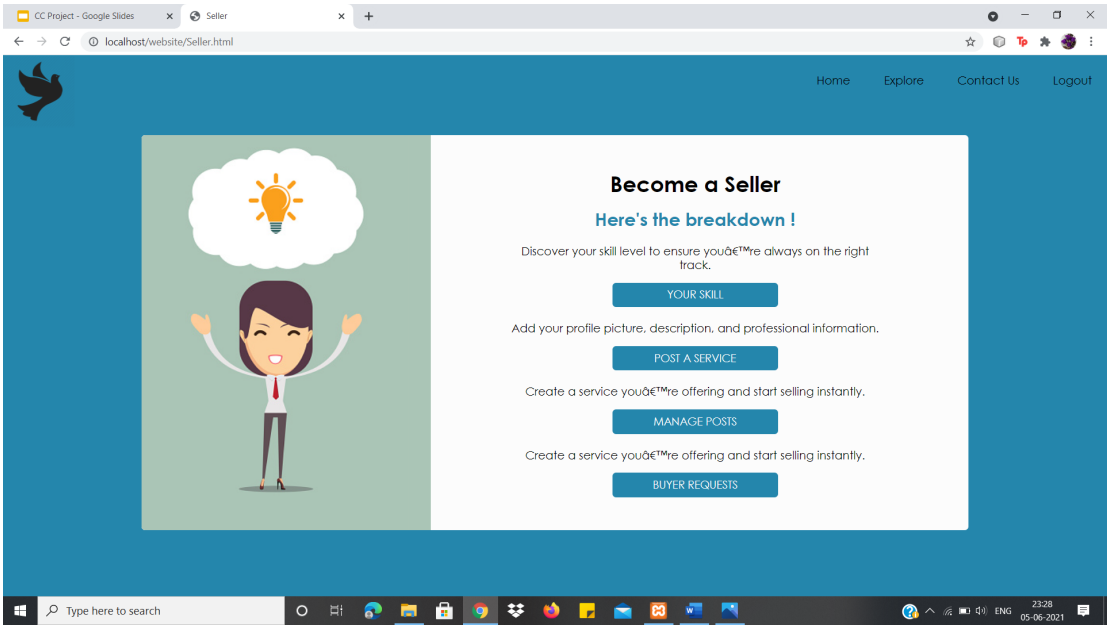

Figure 11. Freelancer main Page

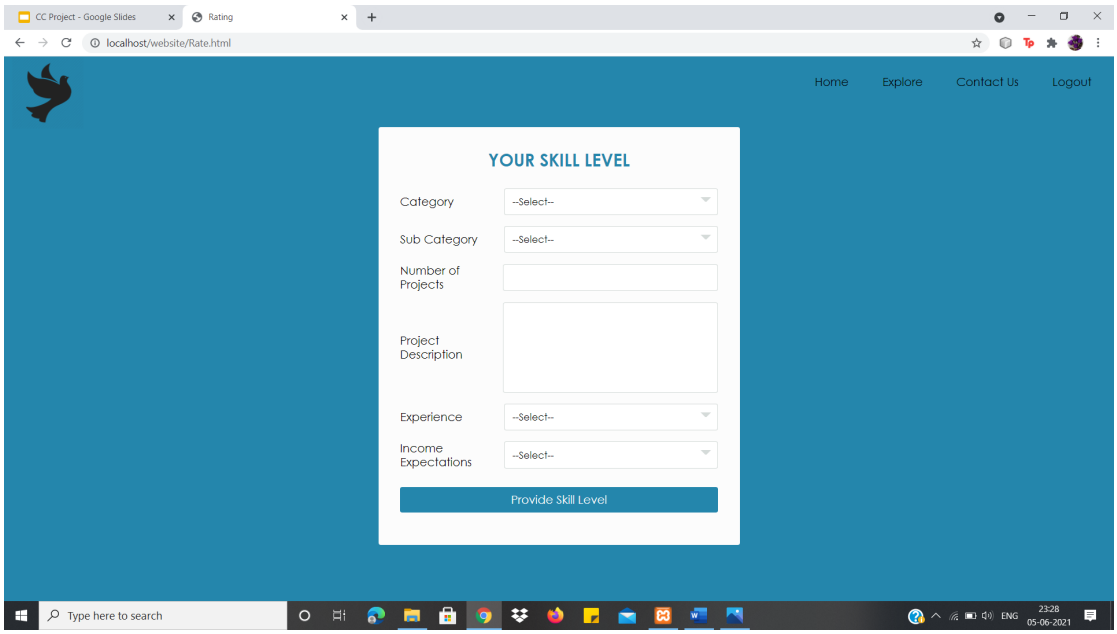

Figure 12. Predicting Skill Level

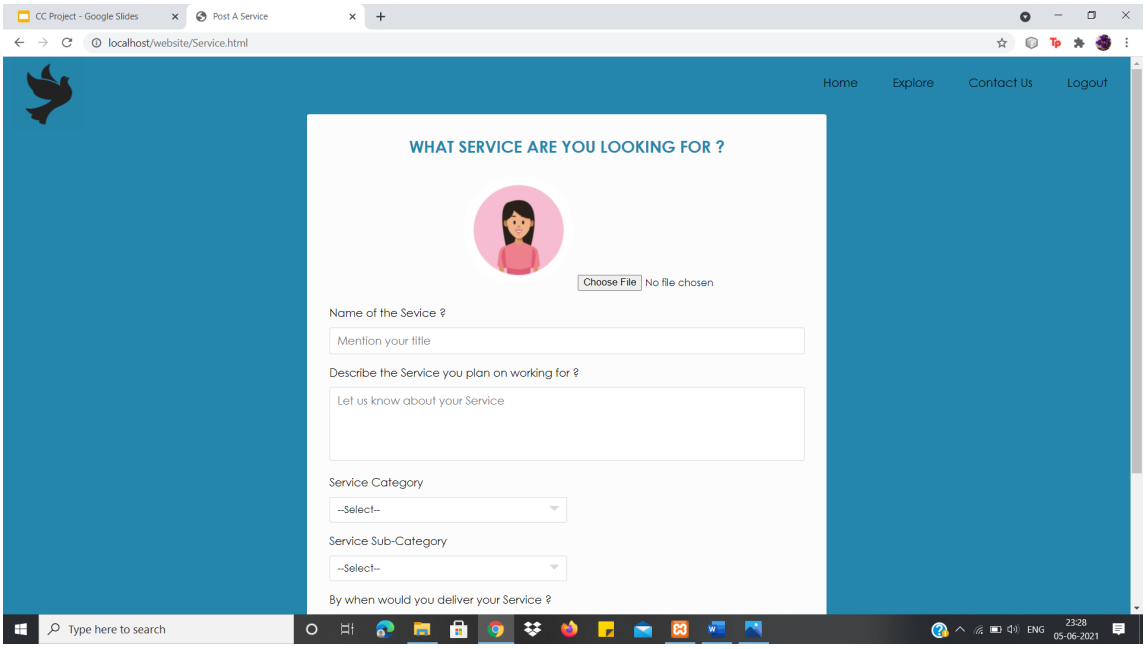

Figure 13. Posting services

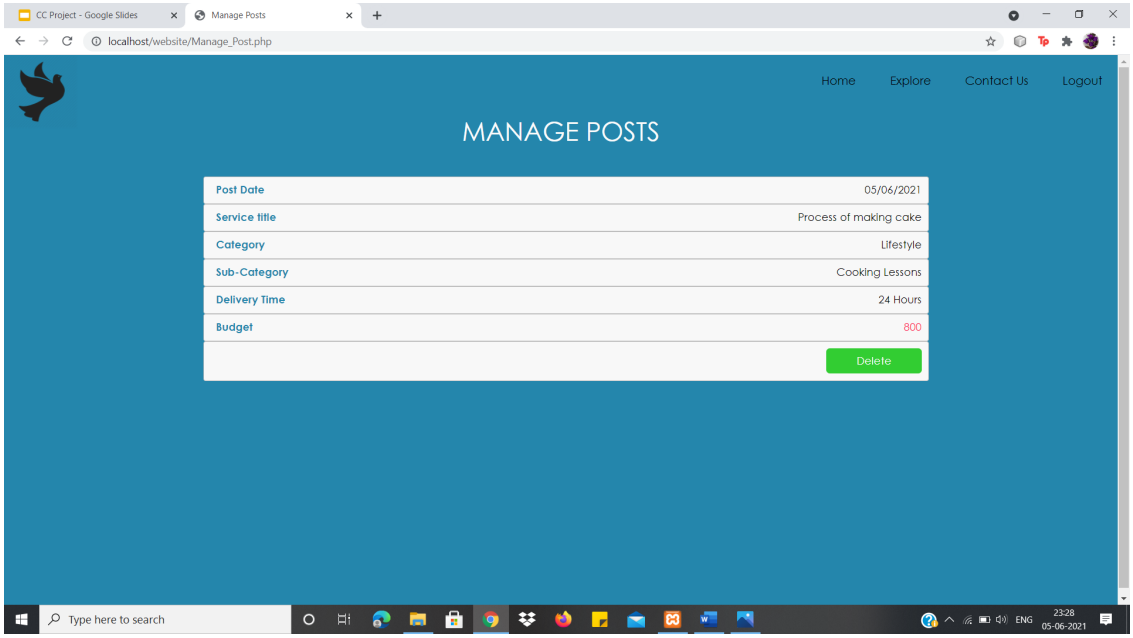

Figure 14. Managing the services posted

The user registered as a Buyer can publish/post a service request, manage their request posts as well as search and explore various services.

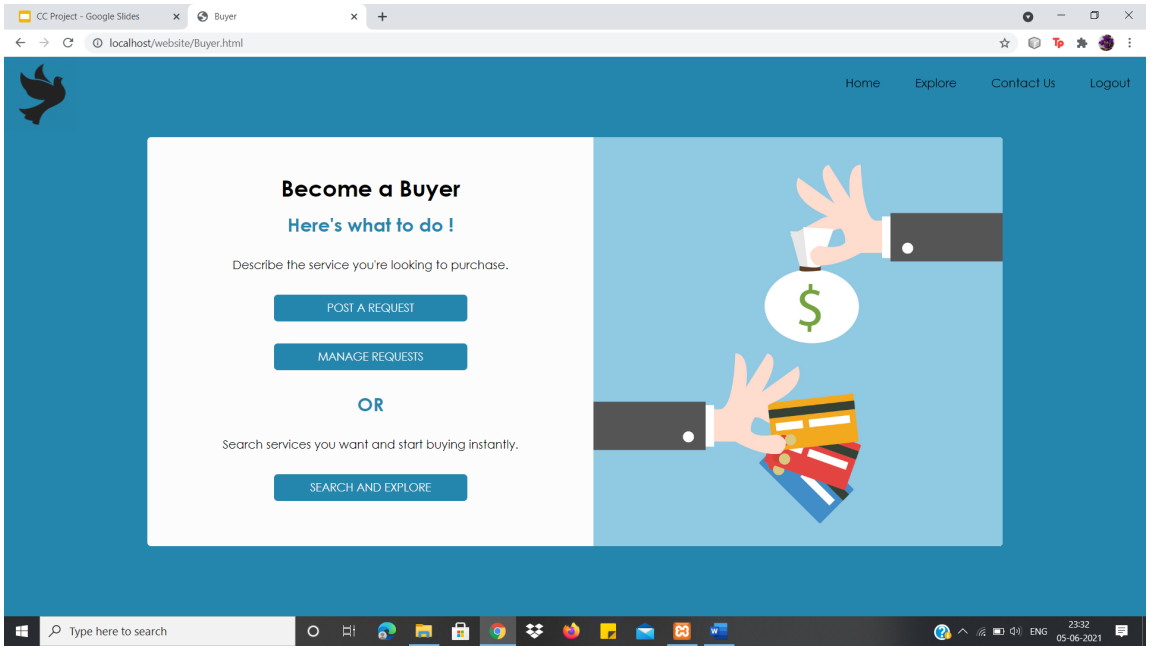

Figure 15. Clients main Page

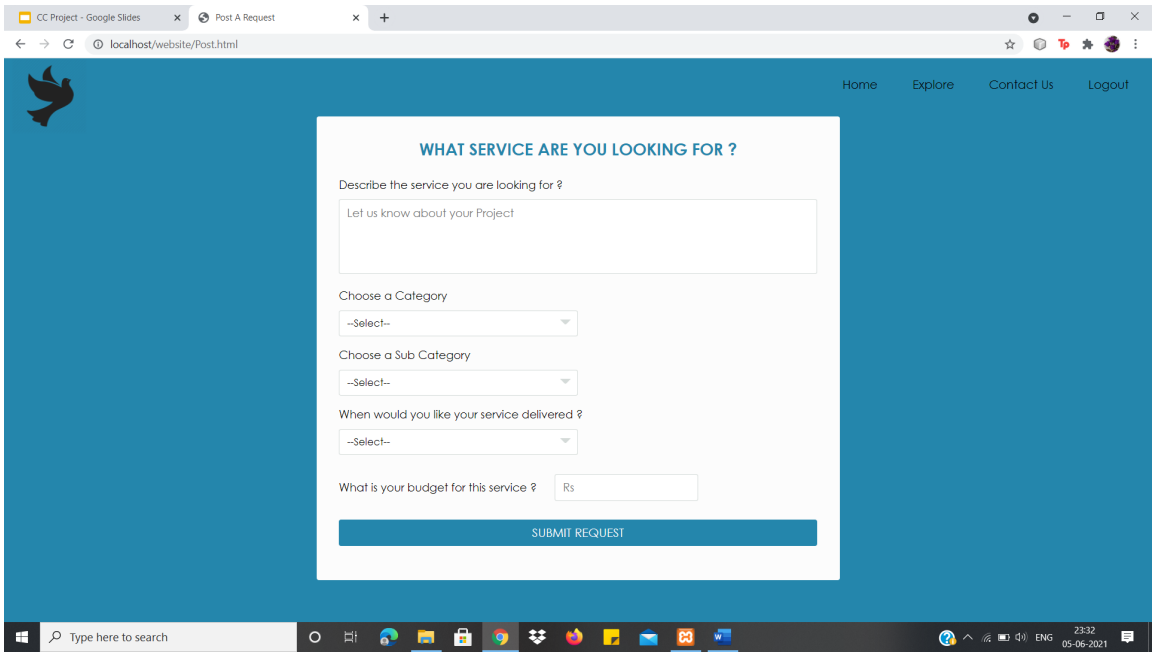

Figure 16. Client requests

The clients and freelancers can easily communicate in this way through the application developed and gain from the resources available. The data of the users is maintained in a secured manner as shown in the below figure.

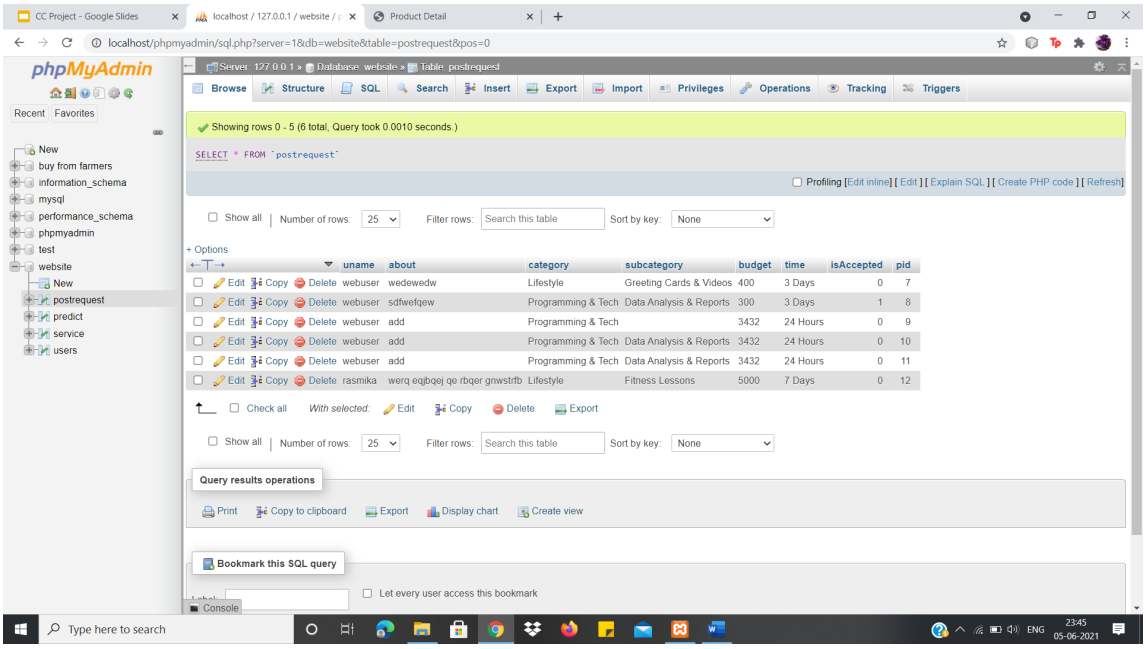

Figure 17. MySQL Server[11]

The user intending to join a community to earn a living can use the job recommendation system to get to his/her best three suitable job roles and their respective job postings that is developed by analysing the textual data with the help of natural language processing[12]. The analysis is based on the three narrowed prospects i.e required job qualification, skillset, experience, location, and vacancies. The below figure shows the recommended job titles for their respective work and experience with the specified job posting as per localities.

	Skill and Experience	1st Recommendation	2nd Recommendation	3rd Recommendation
0	Good knowledge of OOP and OOD; Experience in p...	Java Developer	Software Developer	Web Developer
1	Two years of experience in Java web programmin...	Java Developer	Software Developer	Web Developer
2	BS in Computer Science or any related technica...	.NET Developer	Web Developer	Software Developer
3	Excellent knowledge of PHP, MySQL's SQL, JavaS...	Web Developer	Java Developer	Software Developer
4	Minimum 2 years experience in relevant field; ...	Web Developer	Software Developer	Software Engineer
...	...	...	...	...
74	Positive attitude with a "can do" mentality; A...	iOS Developer	Web Developer	.NET Developer
75	Bachelor of Science degree in Computer Science...	Java Developer	Software Engineer	Software Developer
76	Bachelor's/Master's degree in CS or related di...	Software Developer	Software Engineer	Web Developer
77	Proficient in ASP.NET, C#, Java Script; And/or...	Programmer	Software Developer	Web Developer
78	Degree in Computer Science, Information Techno...	Java Developer	Software Developer	Software Engineer

79 rows × 4 columns

Figure 18. Job role recommendations

	jobpost	date	Title	Company
2158	Essence Development LLC\nTITLE: Tester/ Quali...	Feb 16, 2006	Tester/ Quality Assurance Engineer	Essence Development LLC
3138	Essence Development LLC\nTITLE: Tester/ Quali...	Jan 8, 2007	Tester/ Quality Assurance Engineer	Essence Development LLC
3295	Essence Development LLC\nTITLE: Tester/ Quali...	Feb 12, 2007	Tester/ Quality Assurance Engineer	Essence Development LLC
3518	Essence Development LLC\nTITLE: Tester/ Quali...	Apr 2, 2007	Tester/ Quality Assurance Engineer	Essence Development LLC

Figure 19. Job posts based on role, locality and vacancies[13]

3.1 ML ANALYSIS

The accuracy scores for predicting the skill level is shown in the below table.

Model	Accuracy Score
1. K-nearest neighbours	76.21%
2. Support Vector Machine[14]	72.89%
3. Decision Tree	92.94%
4. Random Forest	96.78%
5. Gradient Boosting	84.78%
6. XGB Boost	81.29%
7. Naive Bayes Gaussian Classifier[15]	86.34%
8. Naive Bayes Multinomial Classifier	76.21%

Table 1. Models implemented and their accuracies(1)

The accuracy scores for implementing the job recommendation system is shown in the below table.

Model	Accuracy Score
1. K-nearest neighbours[16]	67.29%
2. Support Vector Machine	77.21%
3. Decision Tree	78.83%
4. Random Forest	81.09%
5. Gradient Boosting	46.45%
6. XGB Boost	56.89%
7. Naive Bayes Gaussian Classifier	62.54%
8. Naive Bayes Multinomial Classifier	69.92%
9. Logistic Regression	95.67%
10. AdaBoost + Random Forest	83.54%
11. Single Classifier	91.29%
12. Meta Classifier	80.20%

Table 2. Models implemented and their accuracies (2)

The highest accuracy score for skill level prediction is given by Random Forest Algorithm[17], whereas for the job recommendation system, Logistic Regression[18] shows an eminent progress. Even though the complex algorithms such as bagging and boosting[19] are applied there is no improvement in the accuracy. The reason behind the situation is that the data provided is textual and requires NLP techniques with Logistic regression[20]. Also, while predicting skill level the data is converted from categorical to numerical making the Random Forest algorithm more dominant.