Job Recommendation System - B Tech Project

Akshat Jain (18ucs071)
Ashish Lawani (18ucs194)

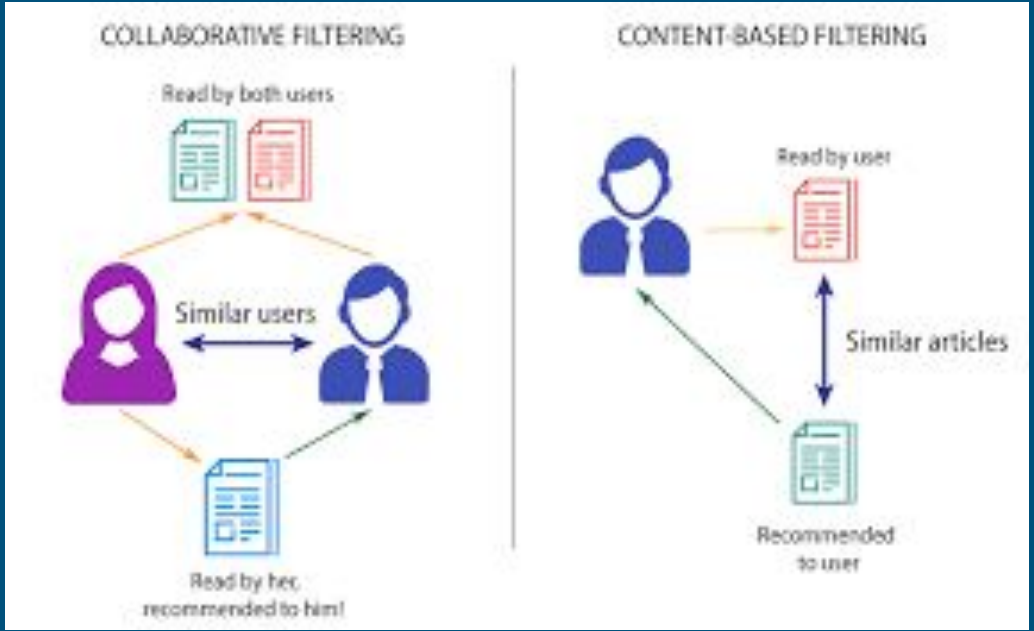
Problem Statement

Input:
Skills, Education, Experiences,Qualifications, Area of Interest, etc.

Output:
Most Suitable Job Titles with Matching Roles

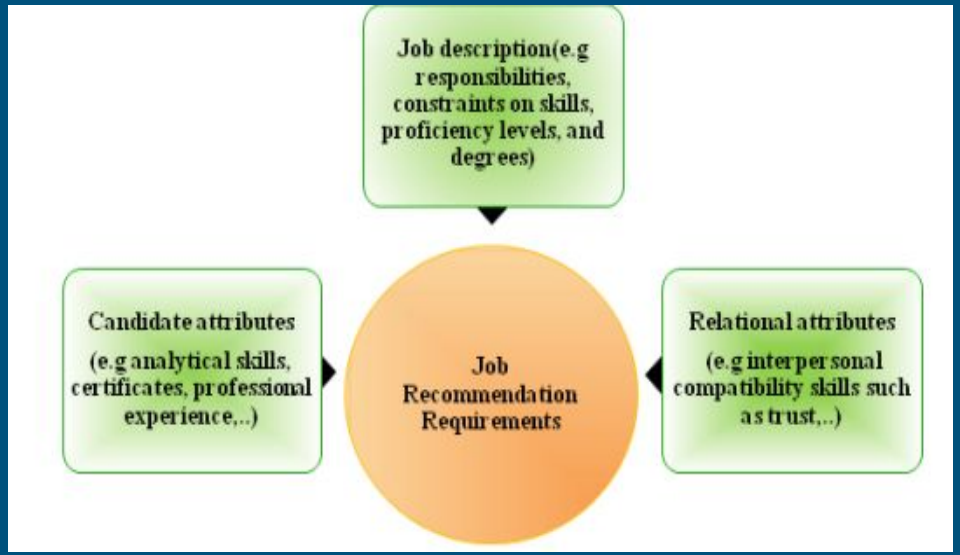
Quick Recap

- **Recommendation system:**
  Suggesting
  relevant items to users
  (movie to watch,text to
  read, products to buy.....)
- **Techniques (main):**
  1. Collaborative Filtering
  2. Content Based Filtering
  3. Hybrid Filtering



Requirements for Job Recommendation

- Skills and Abilities
- Candidate Attributes
- Relational Attributes
- Uniqueness of candidate



Previous Work

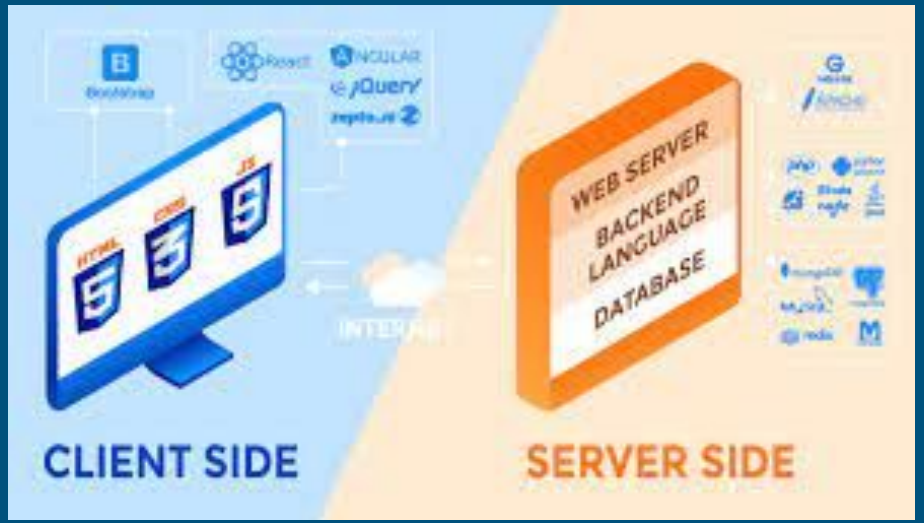
- **Simple Movie Recommendation System:**
  - Collaborative Filtering:
  - Pearson Correlation
  - Cosine Similarity
- **Implementation and Evaluation of Supervised Learning Algorithms:**
  - K-Fold Cross Validation
  - Confusion matrix, Classification table

Data Set (Actual Project)

- Alternate Titles
- Educational Experience
- Knowledge
- Occupation

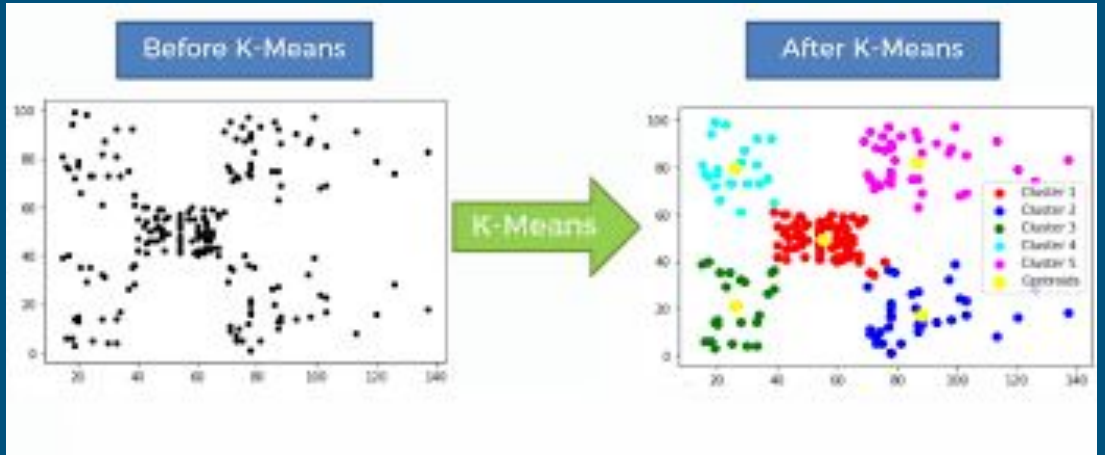
Tech Stack (Backend)

- Language: Python
- IDE: Google Collab
- Libraries: Pandas, Numpy, Pickle, Sklearn
- Algorithms: K-Means Clustering, Random Forest



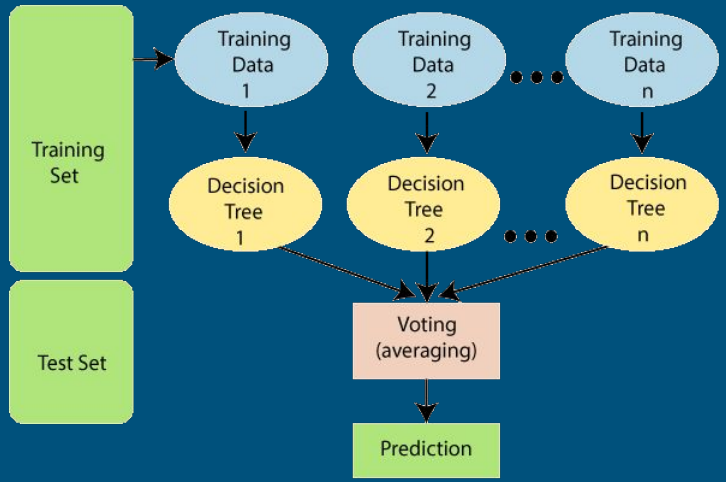
K Means Clustering

- Unsupervised Learning Algorithm
- Items ⇒ k groups
- Items of same group are similar
- Similarity measures: Eucledian Distance, Cosine Similarity, etc.

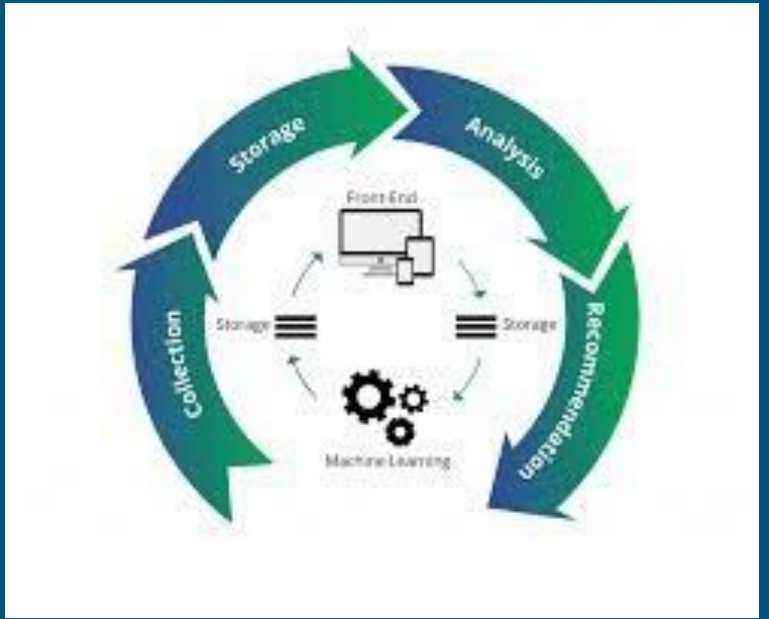


Random Forest

- Supervised Learning Algorithm
- Extension Of Decision Tree
- Multiple Trees, different subset of dataset for each
- Avg of all the results to increase Accuracy



Workflow



- Questionnaire from user
- Input ⇒ Database (MongoDB)
- Prediction of Job Titles
  using skills
- User's Score of education
  and experience
  corresponding to predicted titles
- Sorting of titles (based on above step)
- Additional Information from dataset (JD,core tasks…)
- Output

Simulation and Results

```
while True:
    try:
        print("*****Education Level**************")
        print ("1 - Less than High School Diploma")
        print ("2 - High School Diploma")
        print ("3 - Post Secondary Certificate")
        print ("4 - Some College Course")
        print ("5 - Associate's Degree")
        print ("6 - Bachelors's Degree")
        print ("7 - Post-Baccalaureate Certificate")
        print ("8 - Master's Degree")
        print ("9 - Post-Master's Certificate")
        print ("10 - First Professional Degree")
        print ("11 - Doctoral Degree")
        print ("12 - Post-Doctoral Training")
        print("******End of Education Levels*******")
        ed_level = int(input("Please enter suitable Education Level (1-12):"))
    except ValueError:
        print("Sorry, I didn't understand that.")

        continue
    else:
        break

print(f"Education Level: {ed_level}")
```

```
*****Education Level**************
1 - Less than High School Diploma
2 - High School Diploma
3 - Post Secondary Certificate
4 - Some College Course
5 - Associate's Degree
6 - Bachelors's Degree
7 - Post-Baccalaureate Certificate
8 - Master's Degree
9 - Post-Master's Certificate
10 - First Professional Degree
11 - Doctoral Degree
12 - Post-Doctoral Training
******End of Education Levels*******
Please enter suitable Education Level (1-12):8
Education Level: 8
```

Simulation and Results

```
while True:
    try:
        print("*******Experience Level***********")
        print ("1 - No experience")
        print ("2 - Upto 1 month experience")
        print ("3 - (1-3 months) experience")
        print ("4 - (3-6 months) experience")
        print ("5 - (6 months - 1 year) experience")
        print ("6 - (1-2 years) experience")
        print ("7 - (2-4 years) experience")
        print ("8 - (4-6 years) experience")
        print ("9 - (6-8 years) experience")
        print ("10 - (8-10 years) experience")
        print ("11 - (> 10 years) experience")
        print("*******End of Experience Level******")
        exp_level = int(input("Please enter suitable Experience Level (1-11):"))
    except ValueError:
        print("Sorry, I didn't understand that.")

        continue
    else:
        break

print(f"Experience Level: {exp_level}")
```

```
*******Experience Level***********
1 - No experience
2 - Upto 1 month experience
3 - (1-3 months) experience
4 - (3-6 months) experience
5 - (6 months - 1 year) experience
6 - (1-2 years) experience
7 - (2-4 years) experience
8 - (4-6 years) experience
9 - (6-8 years) experience
10 - (8-10 years) experience
11 - (> 10 years) experience
*******End of Experience Level******
Please enter suitable Experience Level (1-11):11
Experience Level: 11
```

Simulation and Results

```
[ ]  result = loaded_model.predict(test_data)
     print("The test data belongs to Class: ", result[0])
```

```
The test data belongs to Class:  95
```

```
selected_title_group =cluster_group_df.loc[cluster_group_df.index==result[0]]
print("The jobs are:",selected_title_group.values)

selected_title_group.head()
```

```
The jobs are: [['Online Merchants']
 ['Business Intelligence Analysts']
 ['Real Estate Brokers']
 ['Marketing Managers']
 ['Sales Managers']
 ['Management Analysts']
 ['Distance Learning Coordinators']
 ['First-Line Supervisors of Non-Retail Sales Workers']]
```

Class	Title
95	Online Merchants
95	Business Intelligence Analysts
95	Real Estate Brokers
95	Marketing Managers
95	Sales Managers

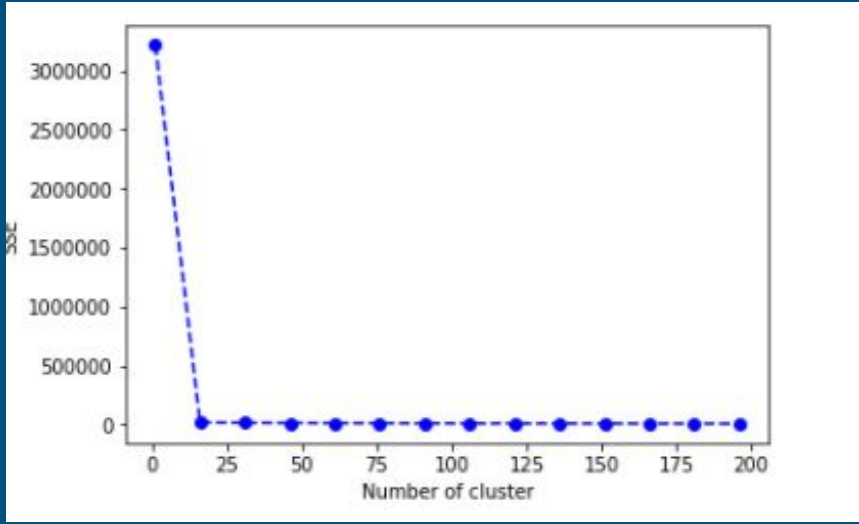
Simulation and Results

```
#from sklearn import tree
from sklearn.ensemble import RandomForestClassifier
rf = RandomForestClassifier(n_estimators=200)
test=DF[:1]   # DF->Importance dataset
test_target=test["Title"]
test=test.drop("Title",axis=1)
test["Active Listening"]=0
test["Mathematics"]=0
test["Writing"]=3.25
test["Reading Comprehension"]=3.62
test["Critical Thinking"]=0
test["Science"]=0
rf.predict(test)
```

```
array(['Chief Executives'], dtype=object)
```

Evaluation Metrics

Elbow Method:



Silhouette Analysis:

```
For n_clusters=25, The Silhouette Coefficient is 0.25819784539767043
For n_clusters=50, The Silhouette Coefficient is 0.2849557725377086
For n_clusters=75, The Silhouette Coefficient is 0.30417968181741606
For n_clusters=100, The Silhouette Coefficient is 0.3307248059253102
For n_clusters=125, The Silhouette Coefficient is 0.3500809376536469
For n_clusters=150, The Silhouette Coefficient is 0.3532717073774063
For n_clusters=175, The Silhouette Coefficient is 0.3460653058030282
```

Accuracy for Random Forest Classifier: 75%

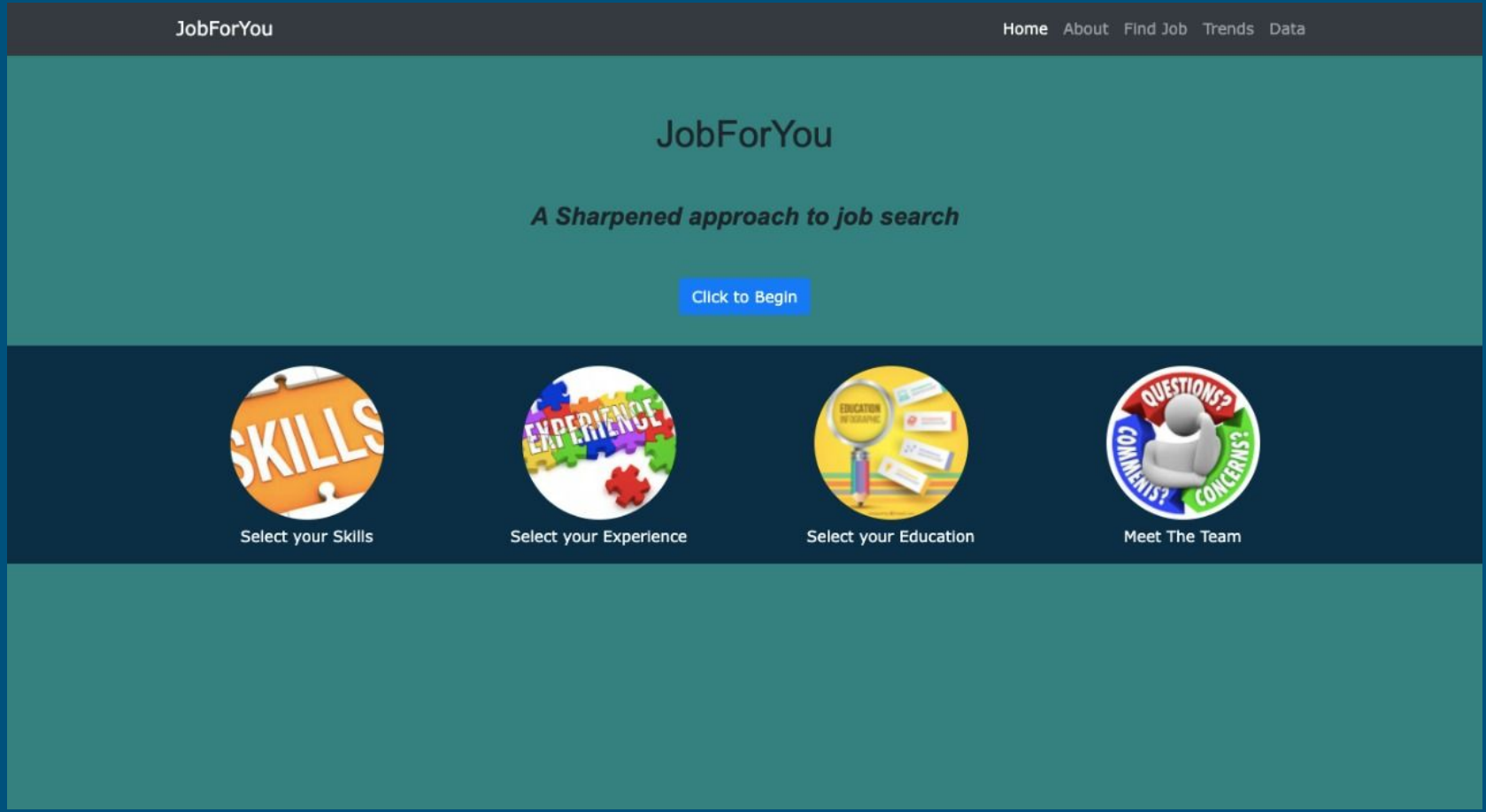
Front End

- **Home:** Links to all other pages
- **Find Jobs:** Taking Inputs, Making Recommendations
- **About:** Overview of the Site
- **Trends:** Trending Jobs
- **Data:** Statistical Details

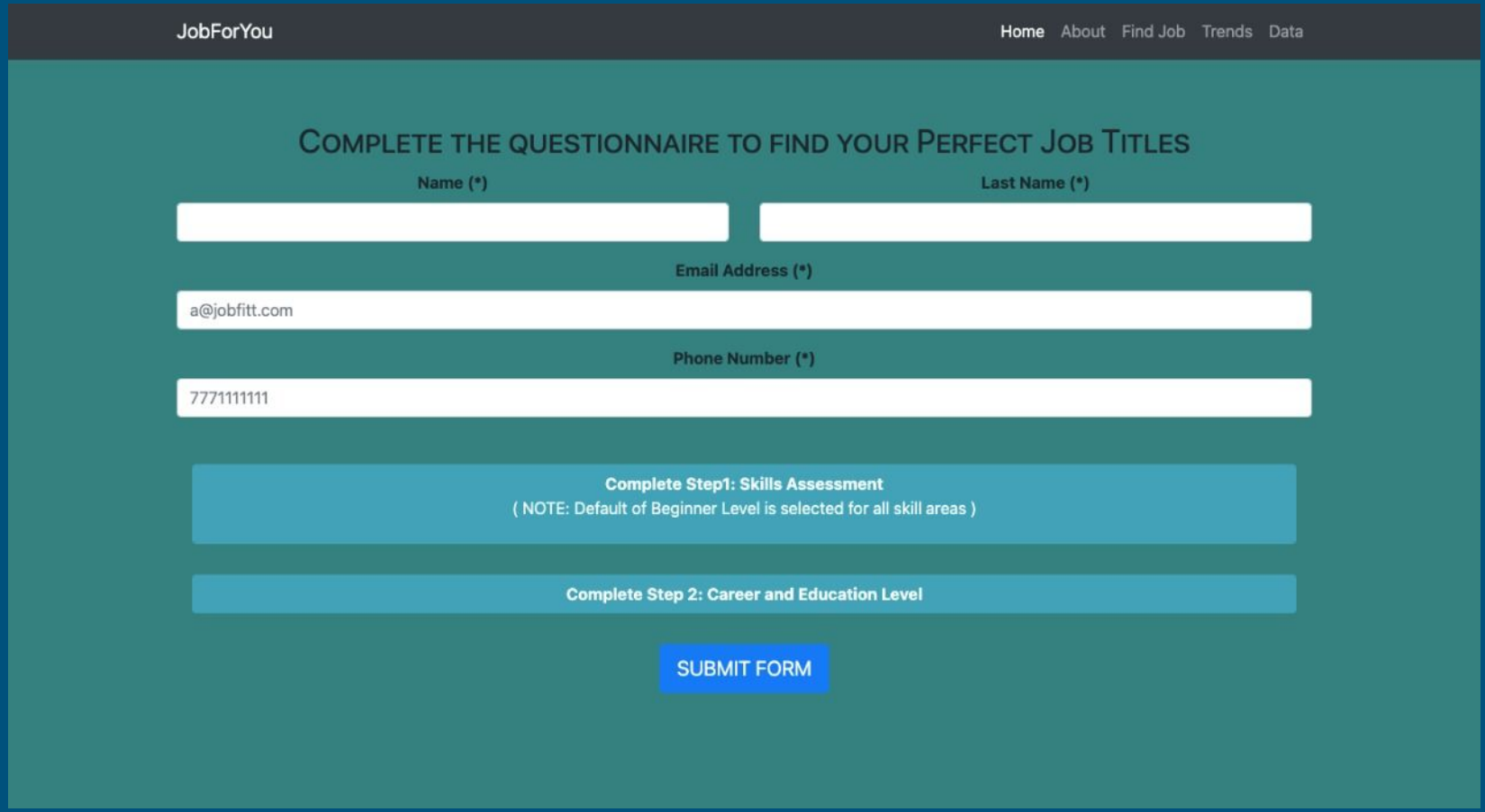
Front End - Tech Stack

- Languages: HTML, CSS, JavaScript
- IDE: Visual Studio Code
- JQuery AJAX(Pass User Input From Client side To Server side and get response from server)
- Bootstrap

Home Page


JobForYou
Home
About
Find Job
Trends
Data
JobForYou
A Sharpened approach to job search
Click to Begin
SKILLS
Select your Skills
EXPERIENCE
Select your Experience
EDUCATION
Select your Education
QUESTIONS?
COMMENTS?
CONCERNS?
Meet The Team

Find Jobs Page



Thanks!

Mentored by:
Dr. Suvidha Tripathi