

RESUME CLASSIFIER & JOB PORTAL

24/9/2021

OVERVIEW

1. Project Topic

The project aims to work on Text classification; by providing working implementation of:

1. Resume classification and suggester.
2. Job Description suggester.

2. Motivation for selection

- This technique will allow candidate to select a vacancy he or she wishes to fill and the recruiter selects a candidate possessing the skills needed to accomplish this same job, a match between person and job needs. First all the CV are categorized using machine learning technique (eg. SVM), based on these categorized CV and job description provided the CV are ranked on bases of the job description.
- The person will provide the resume and the system will suggest those resumes.

3. User interfaces

The two interfaces that will be covered are:

1. Web Application
2. Desktop Application

4. GitHub Link

Github Link : <https://github.com/ABDURRAFAY360/RESUME-CLASSIFIER-JOB-PORTAL>

5. Middleware

The two interfaces that will be covered are:

1. Database: MS SQL Server
2. Web Service – ngrok

6. Methodology of machine learning model

The aim of this work is to find the right candidates resume from the pool of resumes. To achieve this objective, we have developed a machine learning based solution. The proposed model worked in mainly in two steps: i) Prepare and ii) Deploy and Inference.

7.1.1 Preprocessing:

In this process, the CVs being provided as input would be cleansed to remove special or any junk characters that are there in the CVs. In cleaning, all special characters, the numbers, and the single letter words are removed. We got the clean dataset after these steps having no special characters, numbers or single letter word. The dataset is split into the tokens using the NLTK tokenizes. The raw CV file was imported and the data in the resume field was cleansed to remove the numbers and the extra spaces in the date.

Stop words removal:

The stop words such as and, the, was, etc. are frequently appeared in the text and not helpful for prediction process, hence it is removed.

Stemming:

Stemming is the method of decreasing word inflection to its root forms such as mapping a group of words to the same stem even though the stem itself is not a valid term in the language.

Lemmatization:

Unlike Stemming, lemmatization decreases the inflected phrases to ensure that the root word belongs to the language correctly.

7.1.2 Feature Extraction:

On preprocessed dataset, we have extracted the features using the Tf-Idf. The cleansed data was imported and feature extraction was carried out using Tf-Idf. The machine learning based classification model or learning algorithms need a fixed size numerical vector as input to process it. ML based classifiers did not process the raw text having variable size in length. Therefore, the texts are converted to a required equal length of vector form during the preprocessing steps. Specifically, we have calculated Tf-Idf (term frequency, and inverse document frequency) for each term present in our dataset using the scikit learn library function: `sklearn.feature_extraction.text.TfidfVectorizer` to calculate a Tf-Idf vector.

7.1.3 Deployment and Inference:

In this process the tokenized CV data and the job descriptions (JD) would be compared based on the cosine similarity approach and the model would provide CVs relevant to the job description as an output.

7.1.4 Deployment as a Web service:

The service is deployed using a ngrok server and it parses json document and gives appropriate response.

7.1.5 Dataset Description:

The data was downloaded from the online portal(s) and from Kaggle. The data is in Excel format, with three column ID, Category, and Resume. ID - The sequence number of the resume, Category - Industry sector to which the resume belongs to, and Resume - The complete CV of the candidate. The table below show the 25 distinct categories of resume and the number of records belonging to each category.

Resume Categories	Number of records each category of resume
Java Developer	84
Testing	70
DevOps Engineer	55
Python Developer	48
Web Designing	45
HR	44
Hadoop	42
Data Science	40
Mechanical Engineer	40
Sales	40
Operations Manager	40
ETL Developer	40
Blockchain	40
Arts	36
Database	33
Health and fitness	30
Electrical Engineer	30
PMO	30
Business Analyst	28
DotNet Developer	28
Automation Testing	26
Network Security Engineer	25
Civil Engineer	24
SAP Developer	24
Advocate	20

7.1.6 Results and Discussion:

Two models have been built on the cleansed data: i) Classification - Based on the resume and category the model has been designed to categories the resume in the right category and ii) Recommendation -The model take a job description provided by the recruiter and give the list of most relevant resume based on the similarity between resume and jobs description.

The classification was done using Linear Support Vector Classifier (LSVM) and Logistic Regression.

Among the two-classification model the one which gives the best accuracy is Linear Support Vector Machine (LSVM) which is recorded as 1.0. So, we use LSVM trained model for resume classification.

7. Features

- The system shows the resume which job description the resume matches best.
- The system ranks the resumes on the basis of job description.
- Machine learning is used to create the model which predicts the labels.

- Employers can put up different job posts with their required job descriptions and applicants can view them.
- Users can apply for jobs on jobs posted by different employers.
- Employers can view the applicants which applied for the job.
- The application is available for both on the web and desktop.

8. Screenshots

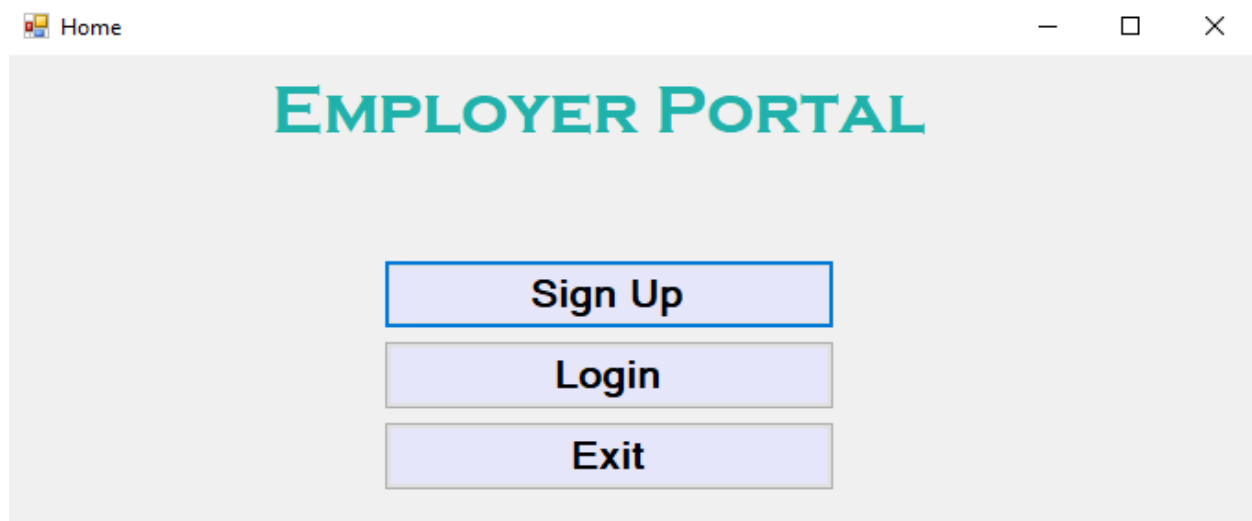


Figure 1. Desktop Main Page

EMPLOYER SIGN UP

First Name: Anas

Last Name: Bin Waqar

Email: k180198@nu.edu.pk

Password: *****

Current Education: B.Sc

Employment Status: Employed

Phone Number: 03061633994

City: Karachi

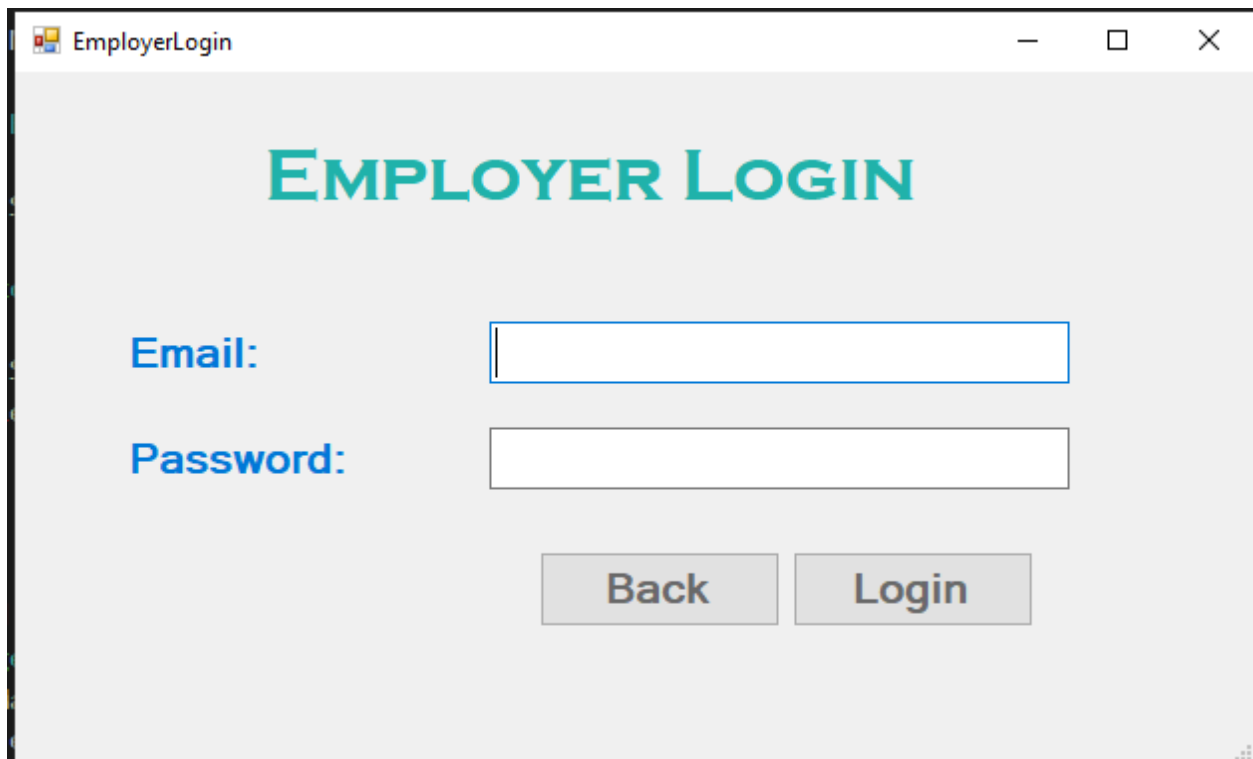
Country: Pakistan

Company Name: CFR

Company Contact Number: 03061633995

Back Sign Up

Figure 2. Desktop Signup form



The image shows a window titled "EmployerLogin". It has a light gray background. At the top center, the text "EMPLOYER LOGIN" is displayed in a large, bold, teal font. Below this, there are two labels in a blue font: "Email:" and "Password:". To the right of each label is a white rectangular input field with a thin blue border. At the bottom center, there are two gray buttons with black text: "Back" and "Login".

Figure 3. Employer Login



The image shows a window titled "Employer Page". It has a light gray background. At the top center, the text "EMPLOYER PORTAL" is displayed in a large, bold, teal font. Below this, the text "Hello, Abdur Rafay!" is displayed in a blue font. In the center, there are four light blue buttons with black text, stacked vertically: "Post Job", "View Applicants", "Logout", and "Exit".

Figure 4. Employer Form

Job Post

Job Designation:

Data Science

Job Description:

THE
YOUNG
FOUNDATION

DATA SCIENTIST & VISUALISATION OFFICER: JOB DESCRIPTION

Job title: Data Science Officer
Area: Research

Location: Bethnal Green, London
Line manager: Director of Research

Purpose of role

Back **Post Job**

Figure 5. Post Job as Employer on desktop

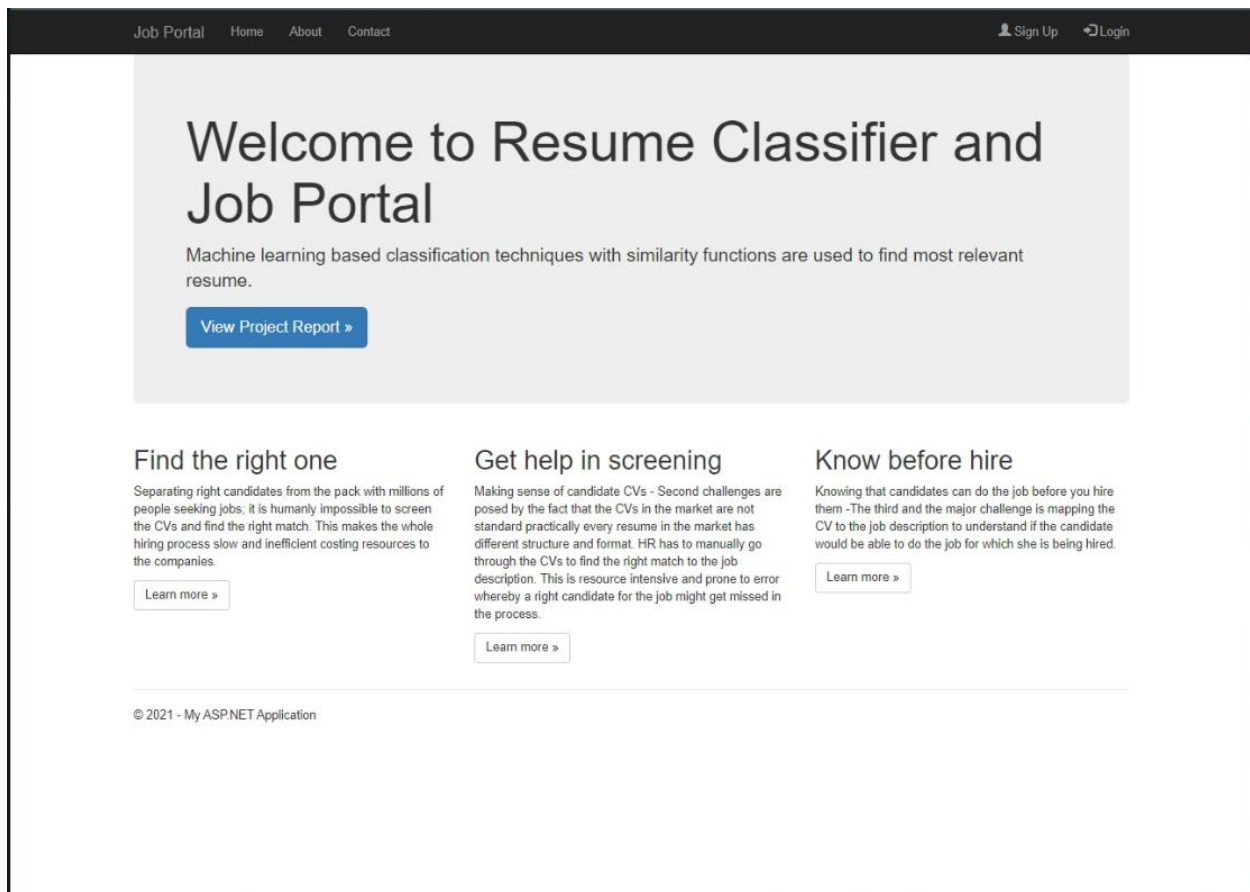


Figure 6. Website Homepage

[Job Portal](#) [Home](#) [About](#) [Contact](#) [Sign Up](#) [Login](#)

UserSignUp

User

First Name

Abdur

Last Name

Rafay

Email

abdurrafay360@gmail.com

Password

Current Education

Undergraduate

Employement Status

Unemployeed

Contact Number

03341872949

City

Karachi

Country

Pakistan

Create

- Want to find candidates? Signup as a Employer
- Already have a employer account? Login

© 2021 - My ASP.NET Application

Figure 7. User Sign Up form

[Job Portal](#) [Home](#) [About](#) [Contact](#) [Sign Up](#) [Login](#)

Login

User

Email address

abdurrafay360@gmail

We'll never share your email with anyone else.

Password

.....

Submit

- Want to find candidates? Signup as a Employer
- Already have a employer account? Login

© 2021 - My ASP.NET Application

Figure 8 Login form

[Job Portal](#) [Home](#) [About](#) [Contact](#) [Upload Resume](#) [View Applied Places](#) [Apply for Job](#) [Logout](#)

Upload your Resume

Resume

Choose File No file chosen

Create

© 2021 - My ASP.NET Application

Figure 9. Resume Uploading Form

[Job Portal](#)
[Home](#)
[About](#)
[Contact](#)
[Add New Jobs](#)
[View Job Applicants](#)
[Resume Suggester](#)
[Logout](#)

Applied Jobs

Job ID	Job Designation	Job Description	Applicant
16	Automation Testing	<p>POSITION TITLE DEPARTMENT LOCATION REPORTING TO OBJECTIVE OF THE ROLE JOB DESCRIPTION AND SPECIFICATION Software Tester - Manual & Automation Testing Technology & Software Development Mumbai To the Project Manager/Tech Lead Software Tester – Manual & Automation Testing - KEY RESPONSIBILITIES - - COMPETENCIES/SKILLS ? ? ? ? ? Understand the business/customer requirements thoroughly and be a part of team meetings where requirements are being conveyed to the technical team Thoroughly test all products/modules assigned to self and report issues/bugs Maintain strong documentation of all test cases and use cases from technical as well as functional perspective to ensure thorough and end to end testing Familiarity with AGILE development methodologies Perform all kinds of testing like stress testing, regression testing, performance testing, functional testing, etc. Write and execute test scripts Run manual as well as automated tests for best results Testing on web as well as mobile environments Ensuring good performance of the system from end user perspective Ensuring the system is robust as well as user friendly Review documentation made by technical team as well as business team for clarity and point out errors, if any Quality assurance and timely feedback Staying up-to-date on the latest technologies and using the same to improve efficiency Process orientation Working on multiple projects simultaneously Delivering tasks as per desired and decided timelines Excellent verbal and written communication skills in English Strong technical background, work experience and qualification Previous experience in similar role highly preferred Knowledge of Logistics and Supply Chain sectors will be an added advantage Openness and curiosity to constantly keep learning and improving Performance and goal oriented and highly focused individual ? ? ? ? ? Problem Solving Skills Strong competence on MS Word, Excel and Technical Automated Testing Tools Ability to adapt and work in a high performance, dynamic environment Strong Documentation Skills Attention to Detail and Desire to deliver perfection to end customers QUALIFICATION Essential: Graduate(B.E/B.Tech preferred) YEARS OF EXPERIENCE COMPANY WEBSITE Essential: 2 yrs Desired: 2-5yrs www.elixia.tech CORPORATE VIDEOS www.youtube.com/watch?v=OX3R4CABJ0 www.youtube.com/watch?v=cZ4WGkly3hl</p>	View Here

© 2021 - My ASP.NET Application

Figure 10. Created Jobs

[Job Portal](#)
[Home](#)
[About](#)
[Contact](#)
[Add New Jobs](#)
[View Job Applicants](#)
[Resume Suggester](#)
[Logout](#)

Suggested Resume on Basis of Ranking

Rank No	Ranking Score	Resume
1	0.0859205795543046	View Resume

© 2021 - My ASP.NET Application

Figure 11. Ranking Scores

The top panel shows a Python script in `app.py` with the following code:

```

19
20 lemmatizer = WordNetLemmatizer()
21 stemming = PorterStemmer()
22 resumeText = resumeText.lower()
23 resumeText = re.sub('http\S+\S*', ' ', resumeText) # remove URLs
24 resumeText = re.sub('RT|cc', ' ', resumeText) # remove RT and cc
25 resumeText = re.sub('#\S+', ' ', resumeText) # remove hashtags
26 resumeText = re.sub('@\S+', ' ', resumeText) # remove mentions
27 resumeText = re.sub('[%s]' % re.escape("""!"#$%&'()*+,-./:;<=>?@[\]^_`{|}~"""), ' ', resumeText) # remove punctuation
28 resumeText = re.sub(r'[\x00-\x7f]', ' ', resumeText)
29 resumeText = re.sub(r'[0-9]', ' ', resumeText)
30 resumeText = re.sub('\s+', ' ', resumeText) # remove extra whitespace
31 word_tokens = word_tokenize(resumeText)
32 resumeText = [w for w in word_tokens if not w in stopwords.words("english")]
33 resumeText = ' '.join(map(str, resumeText))
34 resumeText = stemming.stem(resumeText)
35 resumeText = lemmatizer.lemmatize(resumeText)
36 return resumeText
37
38 #This function will calculate the tf_idf of job description
39 def job_description_tfidf_vector(job_description):
40     global vectorizer
41     query_vector = []
42     query_vector_res = []
43     job_description = cleanResume(job_description)
44     query_vector.append(job_description)
45     query_vector_res = vectorizer.transform(query_vector)
46     return query_vector_res
47

```

The bottom panel shows the terminal output:

```

https://scikit-learn.org/stable/modules/model_persistence.html#security-maintainability-limitations
warnings.warn(
C:\Users\zah\AppData\Local\Programs\Python\Python310\lib\site-packages\sklearn\base.py:324: UserWarning: Trying to unpickle estimator SVC from version 0.24.2 when using version 1.0.1. This might lead to breaking code or invalid results. Use at your own risk. For more info please refer to: https://scikit-learn.org/stable/modules/model_persistence.html#security-maintainability-limitations
warnings.warn(
127.0.0.1 - - [19/Dec/2021 22:38:11] "POST / HTTP/1.1" 200 -
<class 'dict'>
127.0.0.1 - - [19/Dec/2021 22:49:20] "POST /JobSuggestion HTTP/1.1" 200 -
127.0.0.1 - - [19/Dec/2021 22:52:40] "POST / HTTP/1.1" 200 -
<class 'dict'>
127.0.0.1 - - [19/Dec/2021 22:54:44] "POST /JobSuggestion HTTP/1.1" 200 -
<class 'dict'>
127.0.0.1 - - [19/Dec/2021 22:55:06] "POST /JobSuggestion HTTP/1.1" 200 -

```

The screenshot shows the Ngrok web interface with the following details:

- ngrok by @inconshreveable** (Ctrl+C to quit)
- Session Status:** online
- Account:** Asim Imtiaz (Plan: Basic)
- Version:** 2.3.40
- Region:** Asia Pacific (ap)
- Web Interface:** http://127.0.0.1:4040
- Forwarding:** http://rafay.ap.ngrok.io -> http://localhost:80
- Forwarding:** https://rafay.ap.ngrok.io -> http://localhost:80
- Connections:**

t1	opn	rt1	rt5	p50	p90
68	0	0.01	0.01	0.42	1.39
- HTTP Requests:** -----

Figure 12. Ngrok Server accepting requests

```

/* Response.Write(read_resumes[0][0]);
Response.Write(read_resumes[1][0]);
Response.Write(read_resumes[2][0]);
Response.Write(read_resumes[3][0]);
Response.Write(read_resumes[4][0]);
Response.End();*/
var user = new Dictionary<string, List<string>>
{
    { "resumes", read_resumes},
    { "job_description", read_jd}
};
var json = JsonConvert.SerializeObject(user);
var data = new StringContent(json, Encoding.UTF8, "application/json");
var client = new HttpClient();
var response = await client.PostAsync("https://rafay.ap.ngrok.io/JobSuggestion", data);
string x = await response.Content.ReadAsStringAsync();
Dictionary<string, double> suggestion_dic = new Dictionary<string, double>();
Dictionary<string, string> htmlAttributes =
JsonConvert.DeserializeObject<Dictionary<string, string>>(x);
int index = 0;
foreach (KeyValuePair<string, string> kvp in htmlAttributes)
{
    //textBox3.Text += ("Key = {0}, Value = {1}", kvp.Key, kvp.Value);
    suggestion_dic.Add(file_paths[index].ToString(), Convert.ToDouble(kvp.Value));
    index++;
}
var sortedDict = suggestion_dic.OrderByDescending(v => v.Value).ToDictionary(v => v.Key, v => v.Value);
ViewBag.Resumes = sortedDict;
return View("SuggestedResume");
}

```

Figure 13. Suggestion Post request function

Job Portal Home About Contact Upload Resume View Applied Places Apply for Job Logout	
The places you have applied	
Job Designation	Job Description
Mechanical Engineer	<p>MECHANICAL ENGINEER Key Responsibilities: ? Gather information from inspection reports and plan for maintenance of mechanical static based on department maintenance integrated planning and integrated operation planning to sustain the integrity of the equipment. ? Attend to any troubleshooting, breakdown or failures related to mechanical static and provide support and solution using the safest, reliable and cost efficient approach and compliance with the world industrial standard. ? Response to any technical query regarding any issues related to mechanical static based on latest standard engineering practice ? Ensure the maintenance and improvement related to mechanical static is in line with the department objectives ? Prioritize maintenance and improvement activity to minimize production interruption. ? Develop scope of works based on site inputs and inspection reports and prepare budgetary estimates for any materials and service requisition ? Review and study the design of existing assets for further improvement whenever necessary ? Develop master plan for any non-shutdown or shutdown related activities and coordinate with HSE, Logistic and other disciplines prior execution ? Develop site specific procedures, workpack and personnel mobilization plan for unit shutdown and total shutdown windows. ? Provide technical input for tender plan, technical evaluation and presentation of topside and underwater maintenance contracts related to mechanical static. ? Monitor, coordinate and execute EPCC projects, improvement projects or like to like replacement from onshore fabrication up to offshore installation. ? Ensure all P&ID, piping plan, isometric drawings, assembly drawings and equipment layout are up to date and properly documented and have easy access for references. ? Ensure proper procedures in place and in line with the standard engineering and statutory requirements prior any fabrication, testing, lifting, transportation, installation and hook up commissioning activities. ? Monitor all activities and responsible to make sure that any maintenance and fabrication is strictly compliance with the company standard working procedures (HAMS), in line with ASME, API, NACE, TEMA, SSPC and etc. and company HSE policy. Key Requirements: ? Bachelor of Engineering (Hons) Mechanical Engineering or Manufacturing System. ? At least five (5) years working experience in related fields in downstream and upstream oil, gas and petrochemical industries. ? Good knowledge in design, fabrication, installation and maintenance of mechanical static equipment, pressure vessel and shell and tube heat exchangers, process skid, process columns, piping, plant erection, valves and etc. ? Well verse in the applicable codes and standard for oil, gas and petrochemical industries such as ASME, ASTM, NACE, ANSI, TEMA, API, NPFA, AWS and etc. ? Knowledge in design and project planning softwares such as Autocad, PDMS, MS Projects, Primavera and etc. and other related software for maintenance and project such as Maximo, SAP and etc. ? Good written and spoken communication skills in English and able to work with all levels of management and multinational staff. ? Proficient in Microsoft Office applications i.e., Word, Excel, PowerPoint and Project. ? Familiarity with the national oil & gas company's established policies and procedures is an added advantage. ? Applicants should be Malaysian citizens or hold relevant residence status. How to Apply? Those interested may submit their application directly to hresource@noex.com.my. Please include a passport sized photograph and information on your current and expected salary in your application. </p>
Data Science	<p>THE YOUNG FOUNDATION DATA SCIENTIST & VISUALISATION OFFICER: JOB DESCRIPTION Job title: Data Science Officer Area: Research Location: Bethnal Green, London Line manager: Director of Research Purpose of role The purpose of the role is to develop and embed a new data science and visualisation function within the Young Foundation Research Team, while providing strong analytical input into our work. The post holder will work closely with the Research Team, other Young Foundation colleagues and external partners to achieve this. Duties of the role e Leading on data analytics across Young Foundation research projects. This is primarily expected to include working with large publicly available data sets but may also include analysis of secondary research data, our own survey data or bespoke data sets provided by our partners. e Leading on the creation of data visualisations (both static and interactive). e Prototype, develop, test & execute data analytics & data visualisation tools — including exploring how we can make greater use of large volumes of unstructured data e = Identification of appropriate tools and software to support the delivery and growth of data analytics and visualisation within The Young Foundation e Supporting colleagues across The Young Foundation to learn new data analysis skills e Working closely with others as part of a team and with stakeholders and local partners to manage relationships and deliver work. e Producing high quality and varied research outputs for a range of audiences including analysis, report writing and presentations where necessary. e Taking responsibility for data and project management and administration as it relates to delivering work. e Any other duties as required by the line manager in line with the grade of the post. PERSON SPECIFICATION Prior experience - essential e Experience in data analytics using large data sets, gained either through a postgraduate degree in a relevant subject area or comparable experience gained through work in the public, private or third sectors. e Experience of combining multiple data sets to create new insights e Experience in producing creative and insightful data visualisations, particularly with a geographic focus e Experience of creating interactive data tools that allow the interrogation of bespoke data sets e Evidence of education or training in a relevant area (e.g. data science, statistics, mathematics, etc.), ideally to postgraduate level e Experience of cleaning datasets, conducting data linkage exercises and other data preparation work to facilitate high quality data analysis e = Advanced skills in Excel, analysis packages such as R (or similar), and relevant programming languages (e.g. Python) e Evidence of ability to grasp complex concepts and debates quickly and to translate these into appropriate questions for data analysis e Evidence of ability to work across multiple</p>

Figure 14. Applied Jobs list shown to user

2021121939-1.pdf
1 / 1
100%

Abdur Rafay

A Passionate Undergraduate Student

A Computer Science student seeking opportunities in a organization to utilize acquired academic knowledge and skills. An energetic, passionate and skilled individual who believes in adding values to a company by achieving targets and developing the best performance in the organization.

abdurrafay360@gmail.com
 +923341872949

59-L block 2 PECHS , Karachi, Pakistan
 linkedin.com/in/abdur-rafay-aab7791a7

EDUCATION

Bachelors Degree in Computer Science
National University of Computer and Emerging Sciences (NUCES)
08/2018 - 06/2022 CGPA : 3.33

A-Level
Beaconhouse College Campus PECHS(BCCP)
08/2016 - 08/2020

O-Level
St. Peter's High School
08/2003 - 08/2016

ACADEMIC PROJECTS

Sentiment Analysis (04/2021 - 05/2021)

- Implemented on – **Python**.
- Features - An **application** is made using a **Scikit-learn** and **NLTK** library that predicts the **sentiment** of the sentence whether it is positive or negative. It was implemented by two **machine learning** approaches **Logistic Regression** and **Naive Bayes Classifier**. Each of these models is trained on the **IMDB review** dataset from **Kaggle** and **tweets dataset** of **NLTK** library respectively.

Automated Resume Recommender (04/2021 - 05/2021)

- Static website implemented on – **Python**.
- Features – An **application** is made using a **Scikit-learn** and **NLTK** library that first categories all the CV's given in the corpus using **machine learning** model trained on **SVM classifier**, the model is trained on a dataset from **Kaggle** next it takes **job description** as input and suggests the most relevant resumes on the bases of the job description.

Dynamic Problem Solver (01/2021 - 01/2021)

- Implemented on – **HTML, CSS, Javascript, JQuery, and PHP**.
- Features – Our website provides an **Interactive User**

SKILLS

View Applicants

Posted Jobs: Data Science

Candidates: 6

[View Resume](#)

[Back](#)

Gold medalist (08/2018)
Achieved gold medal from Beaconhouse College Campus PECHS for scoring exceptional grades in A-level CIE.

Scholarship (08/2016 - 08/2018)
Received 75% academic scholarship in AS-level for scoring A grades in O-level CIE and 100% academic scholarship in A-level for scoring straight A grades in AS-level CIE.

Best Actor Award (11/2017)
Performed as a sitcom character and won best actor award in Curtain Call hosted by Beaconhouse College Campus PECHS.

EXTRA-CURRICULAR ACTIVITIES

Think n Create(Socirty) Co-head Media and Promotion (10/2020 - 06/2021)
Society of the Fast National University of Computer and Emerging

Figure 15. Viewing job applicants result connected to Database shared by Website and desktop