



## JOB POST AND SEARCH WEB PORTAL

Hritvik Gupta

Computer Science and Engineering

Geetanjali Institute of Technical Studies Udaipur, India

hritvik7654@gmail.com

Harshit Paliwal

Computer Science and Engineering

Geetanjali Institute of Technical Studies Udaipur, India

harshitpaliwal95@gmail.com

Jinkal Nagda

Computer Science and Engineering

Geetanjali Institute of Technical Studies Udaipur, India

nagdajinkle@gmail.com

Shoaib Khan

Computer Science and Engineering

Geetanjali Institute of Technical Studies Udaipur, India

shoaibk909@gmail.com

Aman Thakur

Computer Science and Engineering

Geetanjali Institute of Technical Studies Udaipur, India

aman06@gmail.com

Rakshit Kothari

Computer Science and Engineering

Geetanjali Institute of Technical Studies Udaipur, India

rakshit007kothari@gmail.com

Mayank Patel

Computer Science and Engineering

Geetanjali Institute of Technical Studies

Udaipur, India

mayank999\_udaipur@yahoo.com

**Abstract:** Job portals and search portal are sites that can search and can advertise jobs. A job search portal is a website that provides information about jobs, such as the company hiring, the type of position, and the salary. These portals also provide information on how to apply for jobs online. Job search portals are run by companies and organizations that want to promote their job openings. They may have partnerships with other websites that offer job postings. Job seekers can use these sites to find opportunities without having to spend time searching through individual company websites or social media pages. Therefore to resolve the overhead of writing out the job description or to filling out the jobs details we have used the AI which can extract every elements of the job from the document uploaded like pdf or document using Natural Language Processing and then further post it on portal

**Keywords:** Job search and posting, Natural Language Processing, Artificial Intelligence, Brain Tumor.

## I. INTRODUCTION

## Artificial Intelligence

[1] has altered society in the last decade by reducing corporate and consumer overhead by automating each and every one of their tasks with precision. This new type of job portal has introduced new means of posting and looking for suitable positions, hence boosting the odds of picking suitable applicants and jobs for both job seekers and employers. As a result, developing an AI-based portal is important due to the increasing number of candidates and, consequently, the proportionally increasing number of job postings for candidates. As a result, the overhead for the job poster or advertiser to write down everything about the job posting on the portal is reduced and also for the job seeker to look out appropriate jobs according to resume.

This research looks into various elements of natural language processing, particularly the Natural Language Understanding text extractor, which is based on neural network

and cosine similarity base functions. Text extraction using natural language processing entails extracting the required text from a document using various pre-trained and built models, such as the spaCy [2] web core models, which may then be trained on a specific dataset to extract the relevant data from the given dataset.

Therefore, this web-based portal is trained for both the job seeker and the job poster or advertiser according to their requirements. One more feature that has been included in this web portal is match in the resume skills with required job advertisement which allows job seekers to appropriately apply for the jobs on portal. As unstructured data grows at a breakneck speed, the enormous volume of data presents new hurdles for information extraction algorithms. Traditional information extraction algorithms couldn't handle the massive volume of unstructured big data. These Information Extraction technologies required to be improved due to the large volume and variety of data. Natural language processing techniques have allowed for the recognition and summary of

Extraction concerns thanksto recent technological advancements.

Natural Language Understanding (NLU)

[3] simulates a human's capacity to understand a natural language such as English, Spanish, or Chinese, allowing machines to "read" text (or other input such as speech). Natural Language Processing encompasses both Natural Language Understanding and Natural Language Generation, which simulates the ability of humans to write natural language text, such as to summarize [4] information or participate in a conversation. Natural language processing (NLP) has matured as a technology in the last ten years, with products like Siri, Alexa, and Google's voice search using it to interpret and reply to user requests. Medical research, risk management, customer service, insurance (fraud detection), and contextual advertising have all developed sophisticated text mining applications. The tough task for anyone attempting to analyze textual data is not discovering the correct texts, but finding the right information from these documents. Comprehending the relationships between entities, understanding how events have unfolded, or simply finding hidden pearls of information are all things that anybody is searching for when reading a text [5]. As a result, devising an automated method of extracting information from textual material and presenting it in a structured fashion will enable us to gain several benefits while drastically reducing the amount of time we must spend skimming through text documents. This is precisely what data extraction aims to accomplish.

## II. RESEARCH METHODOLOGY

A literature review on the major issues of the web-based job portal is conducted, which include improper matching of the required job with the offered job, and second, because different amounts of jobs are posted every day, it is not appropriate to write details of each and every job on the portal; instead, upload a text document such as pdf or document, and the remaining work is done by an AI-based NLP model that is trained for these types of tasks.

### I. RESEARCH PROCEDURE

"How can we use the usability of the Natural Language Processing Spacy to extract the job elements from the supplied document, while there was no pre-defined dataset for this task?" is the major question answered here.

The sub-questions indicated below should be answered in order to fully explain the main themes in this paper:

- 1) Why are jobs elements that are needed to be extracted from the uploaded pdf or docx?
- 2) What are those elements?
- 3) What kind of predefined good accuracy model we

can use to deploy this task?

- 4) What kind of dataset we can use to train our model?
- 5) How can we create such kind of dataset and feed to model?

In this paper the focus was also on developing an algorithm which included the NLP architecture, which further trained on the self-defined dataset. The following steps were used in order to develop the NLP architecture.

- 1) Understanding how the job elements can be extracted from the pre-defined document
- 2) Create the dataset according to requirement. And create such amount of data so that it can be used to train the model up to approximately 90% of the accuracy
- 3) Developed a more sophisticated algorithm that is being used to train the model on the developed dataset and provide the output.

A computational learning system that employs a network of functions to interpret and transform a data input in one form into a desired output, generally in another form, is known as artificial neural networks. Human biology and the way neurons in the human brain work together to interpret inputs from human senses inspired the artificial neural network concept. In the actual world, ANNs have already found a wide spectrum of uses. Their ability to identify and recognize a pattern has enticed researchers to use them to solve a wide range of therapeutic issues. As we become more aware that diagnosis, treatment, and outcome prediction in many clinical situations are dependent on a complex interaction of many clinical, biological, and pathological variables, there is a growing demand for analytical tools such as artificial neural networks (ANNs) that can exploit the intricate relationships between these variables. This could be done by perceptron. Many versions of the fundamental Perceptron network have been developed, but the multilayer feed forward Perceptron (fig. 2) has proven to be the most popular. These networks are made up of layers of neurons, typically an input layer, one or more middle or hidden layers and an output layer, each of which are fully connected to another layer [6]. Based on the questions answered in the literature review section this research is carried forward in a step by step manner where, first we defined the different methodology and mathematical equations that are used in creating the algorithm for the project. It's tough to process raw text effectively since most words are uncommon, and it's typical for words that appear to be entirely different to imply almost the same thing. The same words in a different sequence might have entirely distinct meanings. In many languages, even dividing text into useable word-like units can be challenging. While certain issues may be solved using simply raw letters, it is typically preferable to employ linguistic understanding to provide important information. That is precisely what spaCy is intended to accomplish. When paired with additional

predictions like named entities, the dependency parse may be a valuable tool for information extraction. The model used in these projects extracts money and currency values, i.e. entities tagged as MONEY, and then utilizes the dependency parsing to locate the noun phrase they are referring to. In this model we normally include one or more trained pipeline packages into your continuous integration workflow and build process if your application relies on them. While spaCy includes a number of helpful tools for obtaining and loading pipeline packages, the core functionality is solely relied on native Python packaging. This enables your application to manage a spaCy pipeline in the same way that it would any other package requirement. Standard expressions (called REs, or regexes, or regex designs) are basically a little, exceedingly specialized programming dialect inserted interior Python and made accessible through the re module. Utilizing this small dialect, you indicate the rules for the set of conceivable strings that you just need to coordinate [7] this set might contain English sentences, or email addresses, or TeX commands or anything you like. The application that we have developed in other words also can be called a web portal is deployed on Heroku backend architecture and hosted on the Netlify platform. In this web portal there are 3 main html web pages which are supported by java script and connected to each other using the hyperlink the first page is home page which are used to redirect user from home page to either registration portal or to application portal which is either used for the job seeker or the for the job poster [8]. The registration page includes the first time registration user either job seeker or any job provider.

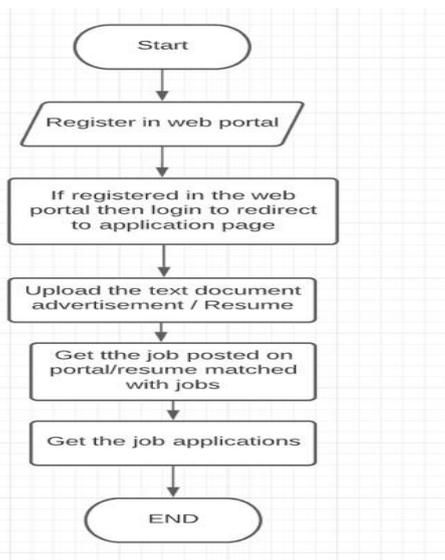


Figure 1. Application Flowchart

### III. SCOPE

The basic functionality delivered by the application features is basically two main functionality the first is to

post the job advertisement and the second is to upload the resume and get the matching skills jobs. Basic changes were made as follows:

1. Used In place of text boxes to write each and every thing down to post request or to search for the job.
2. The Functionality allows job seeker and provider to directly upload pdf or docx and get the required job elements extracted.

### IV. RESULT

The process starts when a user login to the web portal. Job posting and search web portal is created to give user reliable experience and decrease the overhead of search and post the job by removing the factor of writing each and everything down. Rather provide a considerable good solution using AI which is built using is sub-category Natural language processing, Natural Language processing which again uses its methodologies and pre-defined algorithms and models in order to create a well-defined model which can extract the job elements from the uploaded document. Now, the main focus of this web portal is the model training and the dataset creation part where we have used custom or pre-build dataset to train pre-defined model of spaCy library of NLP. The model is trained on the dataset to achieve a remarkable amount of accuracy.

This Application web portal maintains a 3way view of the web pages where each page is connected to each other using the hyperlink and other each page redirected using the click buttons of the html. The user data is stored using the json format which is the user registration data is stored using the json format and can be accessed using the same while matching that data with the user login data. The user data on server stored using the SQL where the format of storage remains the same and storage is directly connected to the Heroku backend [9]. The study found that by incorporating AI into every part of the job search and posting process, it can be made significantly easier [10]. Although it took some time to put in place, the pre-built algorithms perform well on the unique dataset, which was created specifically to extract job information such as jobs skills [11], location, income provided, number of opportunities, and matching CV skills with appropriate jobs [12]. The user required to first registered if not so that to create the login and password in order to move forward to application portal the registration of the user requires:

1. Organization Name/User Name.
2. Location of organization/user
3. Field of work they are currently in.
4. Number of employees for the organization login and for job seeker not necessary to fill it.
5. Then set and confirm password
6. After submitting the button user redirected again to the login page where they have to login to proceed to the application. Each

HTML Page is connected to the each other using the hyperlink which is redirected using click button [13].

7. The mobile application receives JSON response and will convert it into HTML and render it. Further the response is stored using SQL server-based dataset which is connected to the Heroku server [14].
8. Nextatthemainapplicationpage userhas touplodeither job advertisement or the resume to matched it withthe appropriate job according to his/her skills and resultsarepostedonthesamepagewithoutbeinglo adedagain.

## V. CONCLUSION

The job posting and search online portal was built to provideusers with a dependable experience and to reduce the overheadof searching for and posting jobs by eliminating the need towriteeverythingdown. Rather, giveasignificantgoods olutionutilizingAI, whichisdevelopedusingthesub-categoryNaturallanguageprocessing, whichutilizesitsm ethods and pre-defined algorithms and models to produce awell-defined model that can extract the job aspects from thesubmitteddocument. Themodeltraininganddatasetge neration sections of this online portal are now the majoremphasis, withbespokeorpre-built datasetsbeingusedtotrainpresetmodelsofthespaCylibra ryofNLP. Toattaina goodresult, themodelistrained onthedataset. To turn this study into a real-world application that can be used bya broader population, work must be done on the backend, whereyou must build up the application utilizing cloud services such asAmazonwebservices, Azure clouds, and others. Because thismodelwastrainedwithaverymodestquantityofdata, r e-enforcement learning will be required in the future so that it canservea bigger population.

## VI. REFERENCES

- [1] H. Gupta and M. Patel, "Method Of Text Summarization Using Lsa And Sentence Based Topic Modelling With Bert," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 511-517.
- [2] P. Patil, S. Dalmia, S. Abu Ayub Ansari, T. Aul and V. Bhatnagar, "Automatic text summarizer," 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), New Delhi, 2014, pp. 1530-1534
- [3] Moratanch, N. & Gopalan, Chitrakala. (2017). A survey on extractive text summarization. pp 1-6
- [4] Kothari, R., Choudhary, N. and Jain, K., 2021. CP-ABE Scheme with Decryption Keys of Constant Size Using ECC with Expressive Threshold Access Structure. In *Emerging Trends in Data Driven Computing and Communications* (pp. 15-36). Springer, Singapore.
- [5] B. V. Barde and A. M. Bainwad, "An overview of topic modeling methods and tools," 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, 2017, pp. 745-750, doi: 10.1109/ICCONS.2017.8250563.
- [6] K. Nokkaew and R. Kongkachandra, "Keyphrase Extraction as Topic Identification Using Term Frequency and Synonymous Term Grouping," 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP), Pattaya, Thailand, 2018, pp. 1-6, doi: 10.1109/iSAI-NLP.2018.8693001.
- [7] BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding. Jacob devlin, Ming -wei chane, Kenton lee, Kristina Tautanova
- [8] Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. *J Mach Learn Res* 3(Jan):993–102
- [9] Wang, Dingding & Zhu, Shenghuo & Li, Tao & Gong, Yihong. (2009). Multi-Document Summarization using Sentence-based Topic Models.. *ACL-IJCNLP*. 297-300.
- [10] Tong, Zhou & Zhang, Haiyi. (2016). A Text Mining Research Based on LDA Topic Modelling. *Computer Science & Information Technology*. 6. 201-210. 10.5121/csit.2016.60616.
- [11] Guo-Hua Wu and Yu-Tian Guo, "An enhanced LSA-based approach for update summarization," 2015 12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, 2015, pp. 493-497
- [12] Menaria, H.K., Nagar, P., Patel, M. (2020). Tweet Sentiment Classification by Semantic and Frequency Base Features Using Hybrid Classifier. In: Luhach, A., Kosa, J., Poonia, R., Gao, XZ., Singh, D. (eds) *First International Conference on Sustainable Technologies for Computational Intelligence*. *Advances in Intelligent Systems and Computing*, vol 1045. Springer, Singapore. [https://doi.org/10.1007/978-981-15-0029-9\\_9](https://doi.org/10.1007/978-981-15-0029-9_9)
- [13] Ameta, U., Patel, M., Sharma, A.K. (2021). Scrum Framework Based on Agile Methodology in Software Development and Management. In: Mathur, R., Gupta, C.P., Katewa, V., Jat, D.S., Yadav, N. (eds) *Emerging Trends in Data Driven Computing and Communications*. *Studies in Autonomic, Data-driven and Industrial Computing*. Springer, Singapore. [https://doi.org/10.1007/978-981-16-3915-9\\_28](https://doi.org/10.1007/978-981-16-3915-9_28)
- [14] Cherukuri, Aswani Kumar & Srinivas, S. (2006). Latent semantic indexing using eigenvalue analysis for efficient information retrieval. *Int. J. Appl. Math. Comput. Sci*. 16. 551-558.
- [15] Li, Juanzi & Fan, Qi'na & Zhang, Kuo. (2007). Keyword extraction based on tf/idf for Chinese news document. *Wuhan University Journal of Natural Sciences*. 12. 917-921 10.1007/s11859-007-0038-4
- [16] Alghamdi, Rubayyi & Alfalqi, Khalid. (2015). A Survey of Topic Modeling in Text Mining. *International Journal of Advanced Computer Science and Applications*. 10.14569/IJACSA.2015.060121.