Jainendra kumar

Senior Machine Learning Engineer
Jobiak.ai
https://www.jobiak.ai/
Hyderabad, India

+91-7598469637
jainendraiitgcs@gmail.com
jainendraiitg1991@gmail.com
https://github.com/vermadev54
https://www.linkedin.com/in/jainendra15

Education

- **Indian Institute of Technology, Guwahati** — Guwahati, India
  Master of Technology, Computer Science; CGPA: 7.02 — *Jun 2015 – Jul 2017*
- **Pondicherry University** — Pondicherry, India
  P.G.Diploma in Statistical & Research methods; Percentage: 62.10% — *Jun 2014 – Jul 2015*
- **Mizoram University** — Mizoram, India
  Bachelor of Technology, Information Technology; CGPA:7.54 — *Jun 2009 – Aug 2013*

Areas of Expertise

- **NLP**: Text Classification, Named Entity Recognition, Information Extraction, Conversational AI, Document Summarization, Sentiment Analysis, Machine Translation, Question Answering, Generative AI, and Large Language Models (LLM).
- **ML Application in HR Tech**: SEO Keyword Recommender, Resume Screening, Candidate Ranking, Resume and Job Description Matcher, SEO Article Generation based on LLM , SEO-Friendly Job Titles and Descriptions Generator, Job Graph and Similar jobs

Experience

- **Jobiak.ai** — Hyderabad, India
  Senior Machine Learning Engineer — *Oct 2019 - Present*
  - **Keyword Recommendation Ai Agent**:
    Ranking jobs in GFJ(Google for job) or Google organic need lots of keyword optimization on the job description page. Deciding which keyword to add on job description need lots of human power that too processing 1M job daily. Sources of keywords are SEO tool (Semrush, Google Console, Google Trend, Keyword Finder). Contributed towards building AI Agent to decide which keyword will get appended on job page based on job title, company, location.
    Technology/tool: MongoDB, LangGraph, LangSmith, Qdrant, Streamlit, ChatGPT 4o, RAG framework.
  - **Salary Estimation Ai Agent**:
    Ranking a jobs in GFJ (Google for job) or Google organic it is necessary to have complete details of the jobs. salary is one of the entity that Google likes to have in job description to Rank the job higher. In real life 50% of jobs don't have salary. contributed towards build the Ai Agent based on data we have collected over the time on salary to estimate the salary of job from input Job title and Location if originally not metion in jobs.
    Technology/tool: MongoDB, LangGraph, LangSmith, Qdrant, Streamlit, Llama 3.1, Vector Search, RAG framework.
  - **Job Title Cleaning LLM model**:
    Ranking a jobs in GFJ (Google for job) or Google organic, it is necessary to have a clean title to rank job higher. Fine-tuned Google T5 model from our manually cleaned 50K data.
    Technology/tool: LLM, PEFT, fine-tuning, LoRA (Low-Rank Adaptation), QLoRA.
  - **Automated Job Posting software** :
    Machine Learning Model Building and Optimization for Automated Job Posting software (Jobiak.ai Product).
    - Experience building REST APIs and containerizing for ML model using Flask.
    - Experience building ML based crawling technique to find the job listing page in the company website.
    - Experience building Name Entity extraction(NER) model to extract job entity example jobid, jobtitle, jobcompany, joblocation etc...
    - Experience building Recommender system to recommend keywords to improve job ranking in GFJ.
    - Experience doing data analysis and data visualization to improve the job position in GFJ.
  - **Job Description to Job Title Recommender** :
    Adding relevent keyword on job page makes SEO strong page for better job ranking. CNN model that take input as job description and give output as embedding in space of title embedding. nearest embedding point is the

recommended title for given job description. The cosine similarity between the job title embedding and the job description embedding was used as a scoring function
Graph embedding for title: Scraped jobs from the GFJ is the input to make Graph. Vertices in Graph is either searched Query job title in GFJ or Query result job title in GFJ and edge weight between Query job title to response title or vice versa is rank of response job title for particular query job title searched.
CNN model: Our model architecture consists of a convolutional neural network (CNN) that generates an embedding for a job description and a lookup table with job title embeddings.
Technology/tool: Python, scikit-learn, Graph Embedding,Annoy, node2vec, PecanPy.

- **Intelligent Job Scraping Tech** :
Scraping job from careers page of a company is not easy. It leads to explore each page and check is this careers page or job page and also follow the link intelligently if not do so ends up exploring every page. This is not feasible within time and cost. ML model has been developed to follow the link and check the page for careers or job page. Major ML model used are:-
Follow link: Discover which page to follow among all link on page.
Joblisting page finder: Binary classifier Model that say is this job listing page or not.
Job page: Binary classifier Model that say is this job page or not.
Technology/tool: Python, scikit-learn.
- **FAQ Generation for Jobs** :
Adding FAQ on job page has more impact on job ranking in GFJ found on our internal research. To generate FAQ from the job description and answer with help of Generative AI(LLM). A finetuned LLM model has been developed.
Technology/tool: LLM, PEFT.

- **Cognizant** Hyderabad, India
Machine Learning Engineer *Jul 2017 - Oct 2019*
  - **Chatbot Development, Deployment and Integration for Creatives2Code (Cognizant GTO product)—Duration April 2019 - Aug 2019**:
  Product Description : Creatives2Code used to convert hand drawn/tool designed sketch prototyping of product to code (HTML/Angular..etc). Here chatbot used to add behavioral component to control in interactive way(Chat with ChatBot). Example: loading country JSON to country dropdown.
  led product Development, Deployment and Integration.
  Model Dataset Generation/Training: Incrementally added the validated conversation dialog-flow dataset to the model and training.
  ChatBot Deployment: deployed to Cognizant internal sever (Window Server) and server as Http Api to Creatives2Code.
  Integration Webchat Widget to Creatives2Code: Filtering ChatBot response(JSON) and modifying the Creatives2Code response(JSON) for given Element-ID captured during chatbot user conversation.
  Technology/tool: Python, RASA NLU/Core, Liferay, JavaScript, Html, Webchat Widget.
  - **Ultra Fast Video Surveillance system—Duration Dec 2018 - Mar 2019**: Developed Video Surveillance system to recognize person identity in real time streaming video. technology/method: Python, YOLO, Object detection
  - **Analysis of the effective Incremental learning algorithm for image classification—Duration June 2018 - Dec 2018** : Developed a techniques to incrementally train the Machine learning model without using previous dataset only previous model weights and new dataset required while training the model to reduce time and space of system. for experiment considered image datasets CIFAR-100.
  Technology/tool/Method: Python, Tensorflow, CovNet.
  - **WIKI Question Answering System —Duration Jan 2018 - May 2018**: Developed QA system that able to search answer for a given question in a potentially very large corpus of unstructured documents. Thus the system has to combine the challenges of document retrieval (finding the relevant documents) with that of machine comprehension of text and answer the question correctly.
  Technology/tool/Method: Python, NLTK, CoreNLP, Tensorflow.
  - **Speaker Identification application for mobile—Duration Aug 2017 - Dec 2017**: Developed Mobile Api for security. Allow mobile owner to set voice characteristics based password. Considered voice characteristics Feature are MFCC, DELTA and Spectrogram.
  Technology/tool/Method: Python, python-speech-features, scikit-learn, coolEdit.

- **Indian Institute of Technology, Guwahati** Guwahati, India
Master of Technology, Computer science, Research and Teaching Assistant *Jun 2015 - Jul 2017*
  - **Energy Efficient Migration Aware Proportional Fair Scheduling on Multiprocessor**: This project includes Design, Implementation Analysis of Fault-Tolerant Multiprocessor Scheduling Strategy for Real-Time Embedded System. Technology: C++.

- **Portfolio Optimization In Stocks**: This project includes optimization of the stock portolio so that it will give a high profit at given risk using convex optimization.Technologies: Python
- **Speech Processing: Voice Controlled Robot**: This project included the development of a speech recognition system using a Hidden Markov Model which can be trained using any language and can be used as a word recognizer in real time. Tools/Technologies: Visual Studio, C++.

Patent

- **MACHINE LEARNING APPLICATIONS TO IMPROVE ONLINE JOB LISTINGS**: [Paper]

SKILLS

Python, REST APIs development, Ipython Notebook, HTML, JavaScript, SQL, NoSQL, Statistics analysis, Natural Language Processing(NLP), Artificial Intelligence(AI), Machine learning(ML), Deep Learning(DL), Data Visualization, Web crawling, Generative AI (Gen AI), LLMs, finetuning, Prompt Engineering, Reinforcement Learning, Kubernetes, tensorflow serving.

Generative AI Frameworks and Tools

LangChain, LlamaIndex, LangSmith, Hugging Face, FAISS, Pinecone, Qdrent, Chroma, Gemini,OpenAi, ChatGPT, Mistral, Claude, Llama, Vertex AI, Ollama, Amazon Bedrock, Hugging Face Transformers.

TOOLS

Django, Pandas, Numpy, Matplotlib, Tableau, Scikit-learn,Tensorflow, Keras, NLTK, PySpark, Docker, AWS, Beautiful Soup, MongoDB, kserve, Amazon EKS, Docker, bentoML, MLOps MLflow, Model Management, Prefect Dataflow Automation, AWS Kinesis, AWS Lambda, Grafana, Evidently, github action, Terraform, Flask.

Certification

- **DataTalksClub Machine Learning Zoomcamp**: certificate
- **DataTalksClub MLOPS ZOOMCAMP**: certificate
- **Generative AI with Large Language Models**: certificate

ACHIEVEMENTS

- **GATE**: Qualified in GATE 2015, 2016, 2017.
- **UGC-NET**: Qualified for Assistant Professor by UGC-NET in July 2018.

POSITIONS OF RESPONSIBILITY

- **Worked as Teaching Assistant in Programming lab (C, Java)** : Carried out my duties of conducting labs, checking codes of enrolled students and clarifying their lab related doubts
- **Worked as Teaching Assistant for Developing and Designing Departmental Automation System**: Department of CSE, IIT Guwahati.
- **Volunteered for Placements (2015 - 2016).** : Centre for Career Development, IIT Guwahati

KEY COURSES TAKEN

Data Structures and Algorithms, Artificial Intellengence, Machine Learning, Statistics, Speech Processing, Logic in Computer Science, Mathematics for Computer Science, Theory of Computation, Computer Systems.