Vishal Singh Yadav

LinkedIn: linkedin.com/in/tokudai/   Email: vishal.singh@alumni.iith.ac.in
Github: github.com/tokudai-vs   Mobile: +91-9555756235

EDUCATION

Indian Institute of Technology   Hyderabad, India
Masters of Technology (by Research) - *Computer Science and Engineering; GPA: 8.38*   *Jan 2020 - Dec 2022*
Thesis: *Context Aware Question routing in CQA sites*
Relevant Courses: *Foundations of Machine learning, Deep Learning for Vision, Natural Language Processing, Computational Topology, Advanced-Data structures and algorithms design*

GGS Indraprastha University   New Delhi, India
Bachelors of Technology - *Computer Science and Engineering; Percentage: 71.2*   *Jul 2014 - May 2018*
Relevant Courses: *Artificial Intelligence, Machine Learning, Applied Mathematics, Algorithm Design, Data structure*

EXPERIENCE

Carelon Global Solutions   Hyderabad, India
Associate Data Scientist   *Jul 2023 - Present*
- ∗ Awarded by the Director of Data Science for accelerated new product launch.
- ∗ Awarded with the Go Above Award 2 times for exceptional and accelerated delivery of products with strict deadlines and changing requirements.
- ∗ Develop and implement Natural Language Processing (NLP) models to streamline the operations of the underwriting team, which included utilizing Language Model (LLM) technology to develop conversational AI and other tools that support business operations.
- ∗ Collaborate with cross-functional teams to integrate machine learning models into the company's workflow, enhancing efficiency and decision-making. I worked with other teams to deploy models locally and on AWS and Azure.
- ∗ Responsible for creating LLM-based tools specifically designed for the Underwriting Teams.
- ∗ Conduct research and Design machine learning models to identify High-Cost Claimants and Special medical conditions, providing valuable predictive analysis.

Qulabs Software India Pvt Ltd   Hyderabad, India
Machine Learning Engineer   *Mar 2023 - Jul 2023*
- ∗ Spearheaded the development of a conversational AI chatbot and its underlying frameworks from the ground up. Additionally, I engineered a Document Management System (DMS) capable of providing documents and performing approximate matching on a large volume of documents in real-time. Helped the team create models using existing models from Huggingface as a base and creating from scratch.
- ∗ Collaborated in creating algorithms and products to identify quantum vulnerabilities in existing codes and binaries. Also developed frameworks and efficient knowledge bases to enhance the performance of conversational chatbots.
- ∗ Deployed machine learning models and chatbots into production(on AWS and in-house), ensuring seamless integration and functionality while guiding fellow developers on code development, data management, and deployment.

Krama Lab, IIT Hyderabad   Hyderabad, India
Research Assistant   *Jan 2020 - Dec 2022*
- ∗ Served as a Research Assistant under the sponsorship of GreatFour Systems, contributing to various research initiatives.
- ∗ Engaged in multiple research projects within Krama Lab, focusing on areas such as point cloud segmentation and classification, knowledge graph completion, graph representation learning, and developing few-shot methods for these tasks.

Bosch Global Software Technologies   Bangalore, India
Machine Learning Research Intern   *Jan 2022 - May 2022*
- ∗ Graph Neural Networks - Conducted in-depth research on using Graph Neural Networks (GNNs) within the automotive sector.
- ∗ Pointcloud/LIDAR data for autonomous vehicles - Executed segmentation and classification tasks on point cloud data derived from LIDAR technology for applications in autonomous vehicles.

CERTIFICATIONS

- **Carelon Global Solutions & Prizmato**: Advanced Certification in Generative AI
- **nvidia**:
  - Applications of AI for anomaly detection
  - Fundamentals of Deep Learning, Building Transformer-based NLP Applications
  - Fundamentals of Accelerated Data Science, Fundamentals of Accelerated Computing, and Accelerating Data Engineering Pipelines
- **University of Washington**: Machine Learning Specialization (through Coursera)
- **University of Michigan**: Applied Machine Learning (through Coursera)
- **deeplearning.ai**: Deep Learning Specialization (through Coursera)

PROJECTS

- **ARROW - In-house underwriting risk assessment tool (Machine Learning, NLP)**: Part of the team that developed ARROW, an in-house tool for Elevance Health, an AI/ML-based risk assessment tool, to streamline the underwriting process by analyzing financial, clinical, and demographic data. The tool incorporates risk scores into our rating system, enabling quicker, more informed decisions on insurance rates. ARROW utilizes AI/ML in underwriting tools, making data-driven, tech-smart operations that improve our processes and enhance our market competitiveness.
- **Medical-LLM (Machine Learning, NLP, LLM)**: Contributed to the development of Medical-LLM, a tool for Elevance Health. This tool uses Language Model (LLM) technology to process medical data while ensuring HIPAA compliance and user privacy. This initiative has significantly improved the handling of medical data at Elevance Health while allowing other teams to utilize LLMs for their tasks.
- **Medical Compliance (Machine Learning, NLP, LLM)**: Contributed to the development of 'Compliance', a comprehensive suite of tools designed to automate medical compliance in the US Healthcare sector for Qulabs Software. I spearheaded the creation of Machine Learning tools, which encompassed the development of Language Model (LLM) powered chatbots, document matching algorithms, and medical compliance checks, among other functionalities. This initiative significantly enhanced the efficiency and accuracy of compliance processes.
- **Understand Code in Python Notebooks (Machine Learning, NLP)**: We did this project for Google's AI4Code challenge. The task was to understand the relationship between code and comments in Python notebooks and reconstruct the order of markdown cells in a given notebook based on the order of the code cells, demonstrating comprehension of which natural language cell references which code.
- **Multilingual toxic/abusive comment classification and detection (Machine Learning, NLP)**: We did this project for Datathon 2021, organized by IndoML. The task was to classify and detect toxic comments on social media platforms such as ShareChat, Instagram, Moj, etc., in 10+ Indian languages. We used semi-structured raw data in multiple Indian languages. We cleaned the data, created proper data pipelines, and then created NLP models for toxicity detection, which work for various Indian languages.

PUBLICATIONS

- **Context Aware Question Routing in Community Question Answering Sites**:
  Vishal Singh Yadav and Manish Singh - May 2023
  Proposed a recommendation system for online Q&A platforms, improving question routing to suitable answerers. The system uses various techniques to enhance the relevance and diversity of recommendations. The system's performance showed significant improvement in various evaluation metrics.

ACHIEVEMENTS

- **Carelon Global Solutions**:
  - Awarded by the director of Data Science for an accelerated product launch.
  - Recognized with the "Go Above Award" for demonstrating exceptional dedication and a proactive attitude. Led the team to deliver a critical member dashboard within a challenging timeframe of 24 hours.
  - Recognized with the "Go Above Award" for accelerated delivery of ARROW while accommodating business changes in a strict timeline.
- **Mentorship**: Offered mentorship to emerging professionals in the field of AI and advanced technologies through a course facilitated by NSE Talentsprint and IIT Hyderabad.

SKILLS SUMMARY

- **Languages**: Python, C/C++, Bash
- **Frameworks**: Pytorch, Pytorch Geometric, Tensorflow, Langchain, Scikit-learn, NLTK, SpaCy, HuggingFace
- **Tools**: LLMs, OpenAI, AzureAI, Docker, Git, Matlab, LaTeX
- **Skills**: NLP(LLM, Transformers, BERT, GPT), Statistics, Data Preparation, MLOps(Deploying ML models)
- **Platforms**: Linux, Web, Windows, Arduino, Raspberry, AWS, GCP