

Vishal Singh Yadav

LinkedIn: [linkedin.com/in/tokudai/](https://www.linkedin.com/in/tokudai/)

Github: github.com/tokudai-vs

Email: vishal.singh@alumni.iith.ac.in

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
 - * Develop and implement Natural Language Processing (NLP) models to streamline the operations of the underwriting team.
 - * Design automation tools leveraging NLP models to enhance business processes.
 - * Utilize 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.
 - * Utilize advanced machine learning techniques to optimize the underwriting process, reducing costs and improving accuracy.
 - * Responsible for creating LLM-based tools specifically designed for the Underwriting Teams.
 - * Analyze and formulate algorithms to craft policies tailored for Underwriters.
 - * Conduct research and develop algorithms to identify High-Cost Claimants.
 - * Design machine learning models to predict potential high-cost consumers, 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.
 - * Engineered a Document Management System (DMS) capable of providing documents and performing approximate matching on a large volume of documents in real time.
 - * Collaborated in creating algorithms and products to identify quantum vulnerabilities in existing codes and binaries.
 - * Developed frameworks and efficient knowledge bases to enhance the performance of conversational chatbots.
 - * Deployed machine learning models and chatbots into production, ensuring seamless integration and functionality.
 - * Guided 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 the utilization of 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)

PUBLICATIONS

- **Context Aware Question Routing in Community Question Answering Sites:**
Vishal Singh Yadav and Manish Singh - May 2023
Proposed a recommendation system for online QA 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.

PROJECTS

- **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 were provided with 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.
- **Home automation for IoT Devices (C++, Networking, Socket Programming):** We developed this project in collaboration with Samsung. We implemented back-end and front-end automation of the entire home to control every possible IoT device, such as lights, fans, HVAC, doors, etc. Features included manual control of devices by the user, automatic scenes based on time of day, automatic settings based on specific user's presence, and automatic controls when no one is present.
- **Tweet Popularity Detection (Machine Learning, NLP):** We implemented a machine learning model to predict the future popularity of a tweet from its text. The model only uses the text in the tweet to predict the tweet's popularity at different times in the future, such as 1 hour, 4 hours, 12 hours, etc., up to 48 hours. This model allows us to find the probability of a tweet being viral and to filter harmful content or misinformation.
- **Music Genre Classification using Deep Learning (Machine Learning):** We implemented a deep learning model to classify music into ten genres of music as defined in the GTZAN data set. Features like mel-cepstral coefficients, zero-crossing rate, power, loudness, etc., are extracted from audio files using the librosa library, which feeds the data into the ML models to classify data. Mel Spectrogram images are formed from audio files, which are powerful features to discriminate the properties of two audio files. These images are then fed into the CNN model for classification of music into different genres.

ACHIEVEMENTS

- **Carelon Global Solutions:** Recognized with the "Go Above 250 Points 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, showcasing our collective skills and teamwork. This accomplishment underscores my ability to perform under pressure and exceed expectations.
- **Mentorship:** Offered mentorship to emerging professionals in the field of AI and advanced technologies through a course facilitated by NSE Talentsprint and IIT Hyderabad.
- **Machina Doctorina:** Secured 2nd position in Machina Doctrina Machine Learning Competition organised by IITH.
- **IndoML:** Secured 3rd position in IndoML 2022 Datathon organised by IITGN.

SKILLS SUMMARY

- **Languages:** Python, C/C++, Bash
- **Frameworks:** Langchain, Scikit-learn, NLTK, SpaCy, Pytorch, Pytorch Geometric, HuggingFace
- **Tools:** LLMs, Docker, GIT, Matlab, \LaTeX
- **Skills:** NLP(Transformers, BERT, GPT), MLOps, Deploying ML in production
- **Platforms:** Linux, Web, Windows, Arduino, Raspberry, AWS, GCP