# Vishal Panjeta

+91 6239279040 | vishalpanjeta@gmail.com
LinkedIn: vishal-panjeta | Github: VPanjeta

## Experience

**Senior Member of Technical Staff** Nov 2023 – Current
*Thoughtspot India Pvt Ltd* *Bengaluru, KA*

- Architected a multi-agent system, streamlining agent development and enabling dynamic expert agent registration.
- Developed and implemented an AutoML system for business data, utilizing advanced algorithms, including fine-tuned and grammar specific self-hosted transformer models, for anomaly detection, forecasting, and tailored data analysis.
- Developed a multi-agent insight system to answer complex business intelligence questions.
- Implemented high-performance in-house caching systems optimised for time series and language data analysis, reducing average API latency by 30%.
- Developed a system to answer ”why” questions by providing grounded results and data analysis, which includes a mini-agent system for the analysis.

**Machine Learning Engineer - 3** May 2019 – Nov 2023
*Wellthy Therapeutics* *Bengaluru, KA*

- Implemented the digital patient insight system to serve 2X concurrent patients using the same resources.
- Architected backend platform from scratch using Typescript, NodeJS and Express which caters to 60%+ API hits.
- Implemented chat platform on modern protocols like MQTT and XMPP and throughput increased by 80%
- Utilised AmazonMQ, AWS SQS to increase platform reliability and increase platform SLA to over 99.5%.
- Implemented AWS lambda + API Gateway to build microservice architecture to serve 2X requests.
- Architected the platform's SDK that can integrate 3rd party services to Wellthy's ecosystem without data leak.

**Software Engineer** March 2018 – April 2019
*Bankbuddy* *Bengaluru, KA*

- Founding engineer of the company and developed product from the scratch.
- Built the NLU engines using Rasa NLU and custom lightweight language models.
- Built the API engine on Flask + multithreaded pool to maintain a throughput of over 120 requests/minute.
- Developed one click deploy frameworks with docker containers to deploy the entire project stack with one command using on-premises portainer instances.
- Fast inference models to cater to mission critical algorithms of investment banks.
- Helped investment banks automate onboarding of patients to reduce data filling time of users by half.
- Developed aggregate python script to increase data visualization in Superset by 30% by building data ahead of time.

**Software Engineer Intern** January 2018 – June 2018
*Intel Corporation* *Bengaluru, KA*

- Wrote Jenkins CI/CD pipelines for microservices using multiple EC2 instances to reduce build times by 20%.
- Wrote a code flow analysis tool to analyze code and generate documentation and postman collections that helped developers reduce documentation time by almost 50%.

## Projects

**What If? - Alternate History Generation Based On Events** | ***Python, Transformers, LLM*** April 2022

- Used Large Language Models (transformer models) to generate alternate history timelines based on some events.
- Questions can be asked like: ”What if Roman Empire never fell?”

**LLaMA CPU: LLM on CPU** | ***Python, Transformers, CUDA*** February 2023

- Harnessing the power of LLM quantization to make LLM run on CPU.
- Enhanced performance of LLaMA models on CPU to get over 10x speed up on CPU inference.

**ModiScript** | ***Python, Compiler*** November 2016 – December 2016

- Used abstract syntax trees in Python to build a turing complete lexer-parser kit interpreter.
- Script accepts code in the languages of PM Narendra Modi and executes the program.

## Education

**Ramaiah Institute of Technology** Bangalore, KA
*Bachelor of Engineering, Computer Science and Engineering, CGPA: 9.2* *June 2014 – June 2018*