

SAMBBHAV GARG

+91 836 851 6631 • sambbhavgarg@gmail.com

EXPERIENCE

Dubverse.ai (Lensmatic Solutions Pvt. Ltd.), Gurugram, India

January 2024 - Present

Machine Learning Engineer - II

- Developed and optimised a Langchain-like LLM-first translation / AI-assisted text-editing module **boosting acceptance of machine-generated translations by 30%** and **driving down the need for editing translations by 40%**
- Leveraged vector databases and RAG in LLMs to **personalise translations** of newly created projects
- Developed a feature to **automatically generate consumable content** such as titles, descriptions, hashtags, and social media captions from user projects using LLMs
- Drafted a detailed translations dashboard giving visibility into **how translations are being used** - understanding language usage, translation editing, acceptance / rejection of AI-assisted editing, time spent on editing projects - generating key insights for planning new features and planning a roadmap
- Published a [blog](#) on the current SOTA LLM and encoder-decoder models **evaluating and benchmarking** the models for the translation task

Koo (Bombinate Technologies Pvt. Ltd.), Bengaluru, India

December 2021 - January 2024

Machine Learning Engineer - I

- Led the development of the multilingual 'Koo Explore' feature, **increasing user engagement by over 25%** through personalized trending topics and hashtags
- Achieved significant cost savings by optimizing language translation processes using IndicTrans, reducing expenses by over **USD 100K annually** for **5+ Indic languages** while maintaining **80ms latency and 99.99% uptime**
- Successfully migrated over 10 critical systems from AWS to GCP, enhancing system efficiency and reliability
- Enhanced Named Entity Recognition (NER) accuracy by approximately **30%**, streamlining content enrichment processes
- Pioneered the expansion of the topics feature to over **10 Indic languages**, setting a new standard in the microblogging industry
- Automated named-entity discovery and incorporation, reducing process time **from 1.5 days to less than half a day**, significantly increasing content timeliness
- Contributed to system design reviews, authored articles for the Koo technology blog, and advocated for best software practices

Machine Learning Engineer Intern

December 2021 - March 2022

- Developed a heuristic-based algorithm for real-time named entity standardization, enabling the system to quickly adapt to trending topics and enhance user engagement

SKILLS

- Development** - Python (Numpy / Pandas / SciPy / Pytorch / TF / Seaborn / Huggingface), SQL (Postgres, Presto, BigQuery), Docker, Flask/Django/BentoML, Apache Spark / PySpark
- Tools** - Prompt Engineering, Langchain/LlamaIndex, Git, MilvusDB/ChromaDB/FaissDB, AWS/GCP, Apache Kafka, Metabase, Prometheus, Elasticsearch, Grafana, New Relic, Kibana, Visual Studio, bash, cron, Postman, Amplitude, various data labeling software, Google Suite, K8s
- Data Understanding** - Exploratory Data Analytics (Quantitative / Qualitative / Distribution Analysis), Data Visualization
- Data Modeling** - Classical Machine Learning, Deep Learning / Transformer-based NLP solutions
- Cloud Services** - AWS S3/EC2/Lambda/Athena/Sagemaker, Google Compute Engine/Composer/ Dataproc/Cloud Storage/Vertex AI

EDUCATION

UPES, Dehradun • GPA: 7.81

2017-2021

B.Tech. Computer Science and Engineering, specialization in Business Analytics and Optimization

Delhi Public School, R.K. Puram • 88%

2003-2017

All India Senior School Certificate Examination

PUBLICATIONS

Schlegelmilch, Bodo B., Sharma, Kirti & **Garg, Sambbhav**. (2021). Employing Machine Learning for Capturing COVID-19 Consumer Sentiments from Six Countries: A Methodological Illustration. *International Marketing Review*. [10.1108/IMR-06-2021-0194](#)

ACADEMIC PROJECTS

Hand Gesture Recognition for the Indian Sign Language (ISL)

August 2020 - December 2020

Mentored by: Dr. Shamik Tiwari, Sr. Associate Professor, UPES

Using Deep Learning to convert the ISL to Text: assisting hearing impaired individuals for smooth communication over online interactions, especially during the COVID-19 pandemic