Anshul Singh

anshulsinghchambial@gmail.com
anshulsc.live

Education

- **Panjab University** — Chandigarh, India
  Bachelor of Engineering in Information Technology; CGPA: 9.17/10.00 (Rank: 04/120) — *Sep 2021 - June 2025*

Research Experience and Internships

Research Associate at IACV Lab — **IISc Bangalore, India**
Under Prof. Soma Biswas — *Aug 2025 – Present*

- **Multimodal Analysis of Open Source Information**
  Developing a data-efficient active learning framework for multimodal fake news detection using a hybrid method based on entropy-based uncertainty and LLM-guided disagreement, achieving improved sample efficiency over baseline approaches.

Research Intern at LT Research Group — **University of Hamburg, Germany**
Under Prof. Chris Biemann and Jan Strich — *Jan 2025 – June 2025*

- **Multi-Hop Reasoning for Multi-tabular Data using Vision Language Models**
  Worked on vision-language models for multi-tabular reasoning. Implemented post-training techniques like GRPO to handle multiple table images for multi-hop reasoning. Created the MTabVQA benchmark to evaluate VLMs on this task and built a synthetic QA generation pipeline producing the MTabVQA-Instruct dataset.

Visiting Research Intern at Dalhousie University — **Nova Scotia, Canada**
Under Prof. Ghader Manafiazar as MITACS Globalink Research Intern — *June 2024 – Sep 2024*

- **Machine Learning Applications for Dairy Cattle Vocal Pattern Analysis**
  Analyzed animal vocalizations to predict calving time and assess pre- and post-weaning behavior using MFCC spectrograms and hybrid ViT-ResNet models. Applied Attention-Guided CAM to identify key spectro-temporal features driving classifications, achieving 83% accuracy through factorial analysis of vocal changes.

Machine Learning Intern at Virtual Labs — **IIT Roorkee, India**
Under Prof. R.S. Anand and Dr. Rajeev Kumar — *June 2023 – July 2023*

- **Vibration Signal Analysis for Induction Motor Fault Detection**
  Designed a high-performance ML pipeline for fault detection in induction motors using advanced feature extraction and dimensionality reduction techniques. Developed analysis tools and visualization platforms that accelerated research, supporting three related publications.

Undergraduate Research Assistant at *Dept. of Information Technology* — **Panjab University, India**
Under Prof. Veenu Mangat — *Jan 2023 – May 2023*

- **Comparative Analysis of Classification Algorithms for Network Intrusion Detection**
  Conducted comparative analysis of machine learning classifiers for network intrusion detection, evaluating statistical feature selection and genetic algorithms. Explored provenance-based detection approaches using graph neural networks to enhance detection accuracy

Publications

- **A. Singh**, C Biemann, J Strich, *"MTabVQA: Evaluating Multi-Tabular Reasoning of Language Models in Visual Space"*, Findings of EMNLP, 2025.
- P. Kumari, V. Mangat, and **A. Singh**, *"Comparative Analysis of State-of- the-Art Attack Detection Models"*, 14th International Conference on Computing Communication and Networking Technologies, 2023, doi: 10.1109/ICCCNT56998.2023.10306428.
- S. K. Tah, L. Gupta, P. Katari, **A. Singh**, et al., *"HybridNet: LLM-Guided Active Learning for Multimodal Fake News Detection"*, 2025. [**Under Review**]

Projects

Cross-lingual Embedding Alignment for Indic Languages *Python, Pytorch, FastText, SciPy*
Project-II (PWIT 651)
Developed cross-lingual alignment pipeline using FastText with custom embeddings trained on Wikipedia dumps, achieving competitive Precision@1 scores (0.3464 vs. 0.3513 for pre-trained models). Implemented generative adversarial training for improved unsupervised alignment in low-resource language settings.

Aurelius: LLM For APIs *Python, PyTorch, Transformers, Peft, BitsandBytes*
Project-I (PWIT 552)
Fine-tuned suite of LLMs for API call generation, fine-tuned on LLaMA-7B and Mistral-7B with adapter-based techniques and quantization for efficient inference. Built highly relevant context retrieval system using ColBERT's token-level embeddings for enhanced API generation accuracy.

Metal-FL: Cross-Platform Federated Learning *Python, Kafka, gRPC, Socket.IO, Protobuf, PyTorch*
Developed decentralized federated learning architecture with Kafka and gRPC for real-time communication, utilizing Socket.IO and asynchronous programming for seamless client-server connections. Designed cross-platform model aggregation mechanism integrating weight updates from multiple heterogeneous machine nodes for distributed ML training.

Relevant Courses

• **Coursework:** Deep Learning, Machine Learning, Artificial Intelligence, Information Theory, Discrete Mathematics, Linear Algebra, Probability, Databases, Operating Systems, Data Structures.

• **Certifications:** Deep Learning Specialization (Coursera), Machine Learning by Stanford (Coursera).

Technical Skills

Languages: Python, C++, Java, JavaScript, CUDA, SQL, HTML/CSS, LaTeX

ML/DL Frameworks: PyTorch, TensorFlow, vLLM, veRL, DSPy, LangChain, Ray, MLflow

Cloud & Infrastructure: AWS, Kubernetes, Docker, SLURM, PySpark

Web & Databases: FastAPI, Streamlit, Node.js, MongoDB, MySQL, Pinecone

Awards & Leadership Experience

- **MITACS Globalink Research Internship Award** Awarded CAD $9,000 competitive research fellowship for summer research internship at Dalhousie University, Canada (2024).
- **Co-founder & Director** Uniqus Edutech Solutions, startup incubated at RUSA Innovation Cell, Panjab University.
- **ML/AI Lead** Google Developer Students Club (GDSC), Panjab University - Led AI/ML initiatives and workshops.
- **Executive Core Member** IEEE Student Branch, Served as Content Writing Head & Webmaster.
- **Workshop Instructor** Conducted IEEE Workshop on Basics of Machine Learning for 100+ college students.
- **Organizing Committee Member** IEEE National Conference of Women In Engineering, PEC Chandigarh (2022).

Languages

• **English:** Full Professional Proficiency **(IELTS C1)** • **Hindi:** Native/Bilingual Proficiency