AKSHAY JOSHI

akshay@blinkin.io | +49 15758132164 | https://github.com/akshayjoshii | Munich, Germany

Education

ETH Zurich, Switzerland — Feb 2022 - May 2023
Master of Science Research Thesis — *Advisor: Prof. Mrinmaya Sachan*
Research Area: Vision & Language, Visual Reasoning

University of Saarland, Germany — Oct 2019 - Jul 2023
Master of Science in Data Science & Artificial Intelligence — GPA: 1.8/5.0 (Best: 1.0)
Research Area: Natural Language Understanding — *Advisor: Prof. Josef van Genabith*

Visvesvaraya Technological University, India — Aug 2013 - Jun 2017
Bachelor of Engineering in Computer Science & Engineering — GPA: 8.03/10 (Best: 10.0)
Research Area: Distributed Computing, Security — *Advisor: Prof. Shashikumar D R*

Publications

- **Akshay Joshi**, Ankit Agrawal, Sushmita Nair, *"Art Style Classification with Self-Trained Ensemble of AutoEncoding Transformations"*, arXiv:2012.03377, 2020.
- **Akshay Joshi**, Peter Pfeiffer, Annalena Kohnert, Philip Hake, Peter Fettke, Peter Loos, *"PubMedSMBERT: A Pretrained Biomedical Language Model for Medical Smart Vigilance"*, 2023.
- Annalena Kohnert, **Akshay Joshi**, Peter Pfeiffer, *"Evaluating Medical LM Pretraining"*, 2023.

Experience

BlinkIn — Aug 2023 - Present
Senior Applied AI/LLM Scientist — *Munich, Germany*

- Lead researcher & manager of Multimodal Vision-Language Model (VLM) research and development team.
- Research on learning techniques to train Sparse Mixture-of-Domain-specific Expert VLMs with robust cross-modal semantic grounding, multi-hop reasoning capabilities, reduced hallucinations & low-cost inference.
- Co-leading the development of Text-to-Video Spatio-Temporal Semantic Video Chunker model as part of Multimodal Search & Retrieval system. Vision & Speech modalities are utilized for chunk identification.

German Research Center for Artificial Intelligence — Nov 2020 - Jul 2023
Deep Learning Researcher (HiWi) — *Advisors: Prof. Peter Loos, Peter Pfeiffer*

- Developed novel & explainable Self-supervised Transformer Language Model (LLM) architectures for Biomedical Natural Language Understanding. The models are pre-trained on ~40 million clinical research reports.
- Achieved ~2% relative improvement over the current state-of-the-art methods in PubMedQA, EBM PICO, BC5-disease & BIOSSES downstream tasks in the Microsoft BLURB Biomedical NLP benchmark.
- Built highly parallel & computationally efficient Semantic Search (retrieves & ranks ~2 million documents in <3 sec) & Recommendation System for Smart Vigilance in Medical Product Research & Development.

Google Research — Feb 2022 - May 2023
M. Sc. Research Thesis — *Advisors: Dr. Alessio Tonioni, Dr. Henry Rebecq*

- Worked with the Machine Perception - AR/VR team to develop novel self-supervised Multimodal Transformer architectures, and large-scale vision-language pretraining objectives for representation learning of densely structured scientific diagrams.
- Investigated structural parsing, semantic interpretation & multi-hop reasoning in the context of Diagram Question Answering & Cross-Modal Retrieval.
- Setup a Mechanical Turk task to congregate a 1,000,000 large open-source diagram dataset which would be released to the research community for large-scale self-supervised Vision-Language model pretraining.

German Research Center for Artificial Intelligence — Feb 2022 - May 2023
M. Sc. Research Thesis — *Advisors: Dr. Cristina E. Bonet, Yasser Hamidullah*

- Research on large-scale Language Modeling & Hybrid Multimodal Fusion in the 'Machine Translation' group at the 'Multilinguality and Language Technology' research department.
- Devising novel methods to efficiently encode extremely long diagram descriptions/captions & multimodal fusion strategies to mitigate strong priors imposed by language/text modality.

Algorithmic Business & Production Group - UdS Oct 2020 - Mar 2021
Graduate Teaching Assistant *Advisor: Prof. Jana Koehler*
- Tutored & graded the 'Architectures for Intelligent Systems' course, which had a cohort of ~45 M. Sc. students from Computer/Data Science, Embedded Systems, Visual Computing, and Bioinformatics majors.
- Performed 2 iterations of development of a reference architecture & corresponding architectural design documents for a cloud-powered Conversational Question Answering Smart Digital Assistant in ~4 months.

AMD Research & Development Aug 2018 - Sep 2019
Software Engineer *Bangalore, India*
- Implemented platform initialization routines of the off-chip phase of Platform Security Processor (PSP) firmware for Ryzen 3000 (Matisse & Castle Peak architecture) processors in ~5 months.
- Extended & validated the support for Microsoft PlayReady DRM protection technology in PSP firmware for Ryzen Pinnacle & Raven Ridge family of desktop & mobile x86 processors in 3 months.

AMD Research & Development Feb 2018 - Aug 2018
Software Engineering Intern *Bangalore, India*
- Developed & validated the AMD Ryzen Master Software Development Kit for CPU/Memory Overclocking (Frequency, Voltage, Timing), Core Parking, and Simultaneous Multithreading utilities.
- In a span of 6 months, delivered ~70 multi-platform Windows SDK APIs. Further, established & documented >100 unit test cases & a detailed test plan for Ryzen Master CLI tool validation.

Alstom Jan 2017 - Mar 2017
Engineering Intern *Bangalore, India*
- Developed an end-to-end PySpark data preprocessing pipeline which comprises of Data Cleansing, Transformation, Statistical Analysis & Feature Selection. Reduced the computation time of older pipeline by 12%.
- Built & evaluated Multivariate Regression & ARIMA Forecast Models for Predictive Analytics on Alstom EMEA employee compensation data using A/B Tests. Mean Average % Error of the best model: ~17%.

Skills

- **Languages**: Python, C++, SQL, Qdrant VectorDB
- **Libraries**: PyTorch, NumPy, Langchain, JAX, Scipy, OpenCV, Pandas, FastAPI, Spark, Kafka, LlamaIndex
- **Tools**: Git, Docker, Kubernetes, Jira, WandB, OpenAI API, Nginx, CMake, LaTeX, AWS, CI/CD, GCP

Projects

- Exploiting Point-level Correspondences for Unsupervised 3D Point Cloud Understanding.
- 3D Pose and Shape Estimation with Stitched Puppet Model & Max Product Belief Propagation.
- Open-domain Question Answering over Knowledge Graphs & Text Passages with Memory Networks.
- Real-time Event-driven Multi-lane Vehicle Detection & Tracking for Adversarial Situation Monitoring.
- RSNA-MICCAI Radiogenomic Classification of 3D fMRI scans to detect MGMT Methylation in Brain tissue.

Seminars

- Hybrid Machine Learning Seminar on 'Learning like Humans with Deep Symbolic Networks', University of Saarland, Germany. Nov 2020 - Feb 2021.
- Research Poster on 'Advanced Cryptographic Standards & Security', Computer Society of India. April 2017.

Achievements

- Merit certificate for achieving a rank in the **top 5**% of the graduating class of Bachelors in Computer Science & Engineering. Total number of graduating students: **106**.