

(408)-621-4007

kabir@manghnani.com

Kabir Manghnani

Machine Learning Scientist @ Tempus labs

linkedin.com/in/kabir-

manghnani

Impact: Track record of published and production-deployed work in novel Machine Learning methods, clinical decision support, and product POCs across academia & industry. A drive to enable new products and improved user experiences.

WORK EXPERIENCE

Senior Machine Learning Scientist – GenAI

Jun 2020 — Present (4.5 years)

Tempus Labs Inc

Chicago, IL

- Efficient Automated Structuring of Clinical Notes: Tech Lead. Developed and evaluated LLM agents for scalable data extraction including finetuned models using QLoRA/PeFT. Experience with LangChain/HF transformers /trl/peft/Microsoft guidance.
- Tempus Search: Major contributor. Developed RAG system for cohort identification from Tempus' 2 million+ clinical patient records – became Tempus' Flagship GenAI product. Considerations of medical domain NLU, RAG strategies, and user guardrails.
- GenAI Initiative: founding technical member. Scoped, built, and validated PoCs for clinical LLM products. Several PoCs graduated to product offerings. Submitted several Abstracts/Papers at top ML/oncology conferences/journals (NuerIPS, ASCO, AACR, JCO)
- RNA biomarker discovery: Contributor. TUO (Tumor of Unknown Origin), CMS (Consensus Molecular Subtypes), Brain Met.
- Lead data-modeling work for Tempus-Janssen COVID-19 clinical-data Collaboration

Consultant

Jan 2020 — May 2020 (6 Months)

Shoreline AI

Remote, CA

- Consulted on low-resource Machine Learning methods for 'on-device' anomaly detection (<50kb-memory).
- Resulted in Patent no. 20210133607 Methods for Self-learning IOT devices

RESEARCH & INTERNSHIPS

Work completed while associated with Koyejo Lab, University of Illinois | Jan17 – Jan20. Topics of interest include generative models, representation learning, model interoperability, inductive biases, & disentanglement.

Research Intern

May 2019 — Dec 2019 (8 Months)

Google AI Research (f.k.a Google Brain)

Mountain View, CA

- Explored extensions and implications of sub-group level and individual fairness constraints
- Developed novel PyTorch models for joint optimization of individual fairness and performance metrics. Evaluated effects on performance at the intersection of protected sub-groups.

Academic Collaborator

Nov 2018 — Mar 2019 (8 Months)

ScreenShop Inc (acquired by Snapchat)

Remote

- POCs on latent space disentanglement of images and contributions to Experimentation Framework

Computational Biology Intern*

May 2018 — Dec 2018 (8 Months)

Freenome Inc

South San Francisco, CA

- Developed methods for supervised dimensionality reduction to combat confounding in molecular datasets (cfDNA)
- Resulted in first author publications and posters at NeurIPS'18 and ISCB'18 workshops. (*not associated w/ Koyejo)

PUBLICATIONS & PATENTS

1. Manghnani, K. *et al.* Evaluation of Large Language Model (LLM)-based clinical abstraction of Electronic Health Records (EHRs) for Non-Small Cell Lung Cancer (NSCLC) Patients in American Society of Clinical Oncology 2025 (2025).
2. Manghnani, K. *et al.* METCC: METric learning for Confounder Control Making distance matter in high dimensional biological analysis. *CoRR* abs/1812.03188. arXiv: 1812.03188. <http://arxiv.org/abs/1812.03188> (2018).
3. Thompson, W. *et al.* Large Language Models with Retrieval-Augmented Generation for Zero-Shot Disease Phenotyping in Deep Generative Models for Health Workshop NeurIPS 2023 (2023). <https://openreview.net/forum?id=Ujqjn2q9Gi>.
4. Lourentzou, I., Manghnani, K. & Zhai, C. Adapting Sequence to Sequence models for Text Normalization in Social Media. *CoRR* abs/1904.06100. arXiv: 1904.06100. <http://arxiv.org/abs/1904.06100> (2019).
5. Michuda, J. *et al.* Validation of transcriptome-based assay for classifying cancers of unknown primary. *medRxiv* (2022).
4. Systems and Methods for Self-Learning IoT Devices | Co-Inventor – **Patent no. 20210133607** | Assignee Shoreline AI
5. Real-world data to enable large-scale assessment of WHO CNS5 glioma classification. **ASCO 2022 - Poster Discussion** | secondary author

EDUCATION

Bachelor of Science, Computer Science (Dean's List), University of Illinois Urbana-Champaign, GPA: 3.84/4.00 **Aug 2016 — Dec 2019**

Courses: Machine Learning (CS 446) • Machine Learning for Signal Processing (CS 598 PS) Algorithms (CS 374) • ML for Computational Biology (CS 598 JP) • Numerical Methods • Statistics & Probability • Linear Algebra • Text Information Systems