**Karthik V. Uppulury, PhD**

<https://www.linkedin.com/in/karthik-uppulury-phd-0a819558/>

103 Myrtle Ave, Millburn, NJ, 07041 [karthik.uppulury@gmail.com](mailto:karthik.uppulury@gmail.com)

https://github.com/uppulury +1-713-614-6549

**SUMMARY**

Experienced full-stack ML engineer with academic research background. Published 8 peer-reviewed publications in scientific journals. I’m passionate about building ML/AI data pipelines at scale.

**TECHNICAL SKILLS**

**Cloud**: K8s, Vertex AI, Docker, Jenkins/CI-CD, Kubeflow, Gitlab, Confluence

**ML Frameworks**: PyTorch, Keras, TensorFlow

**Programming**: Python (NumPy, SciPy, Scikit-learn, Pandas, Dask), SQL (Big Query, Big SQL, Redshift), Shell/Unix, NLP (NLTK, Stanford CoreNLP, spaCy)

**Big Data:** Apache Spark (PySpark), Ray

**EXPERIENCE**

Sr ML Engineer, ***Macy’s Inc*** (New York, NY) *Feb 2021 – Current*

-**Developed** an LLM-pipeline to generate ‘*Store Manager Insights*’. The pipelines *summarized* store data from several dashboards into interpretable text summary of the relevant business KPIs. **GPT**/**Gemini** suite of models were leveraged to generate ***autopilot code*** via prompt engineering to process the datasets for text summarization.

-**Designed** & **implemented** customer facing ML features such as *browse ranking* (using **CatBoost),** *customer* *price-preference* (using **Matrix Factorization),** *Cross-sell* (using **NLP**, **Association Rule Mining** techniques such as entity extraction, embeddings (via BERT/ResNet/LLMs) & *complimentary recommendations* (using **Graph Neural Nets**). **Served** recommendations to **1 million**+ users. **Scaled** up recommendation features using ***PySpark***.

-**Achieved** ~80% accuracy on test data sets.

-**Deployed** customer & supply-chain models to production using ***Kubeflow*** *&* ***Vertex AI*** pipelines (GCP).

-**Built** & **managed** Email & Teams notifications for Vertex pipelines.

Sr Data Scientist, ***ADP*** (New York, NY) *March 2020 – Jan 2021*

-**Developed** classification models (using *H2O.ai*) to identify prospective companies buying payroll products.

-**Achieved** 80% model accuracy (with 20% misclassification rate) using ensemble methods (*XGBoost*, *LightGBM*).

-**Deployed** models on Amazon EKS and served results on Salesforce UI.

Sr Data Scientist, ***IBM*** (New York, NY) *July 2019 – Feb 2020*

-**Developed & Deployed** **microservices (DevOps)** pipelines on ***Kubernetes*** clusters for data catalog dashboards.

-**Improved** metadata quality & graph connectedness for the *metadata* ***Knowledge Graph*** (REST API).

-**Monitored** dashboard for metadata completeness.

Research Scientist, ***Ingenus Pharmaceuticals*** (Denville, NJ)*August 2017 – July 2019*

**-Designed & implemented** standard test procedures to detect trace level (*parts per billion*) elements for (~18) generic cancer drugs. **Achieved** FDA approvals (per *ICH Q3D*) for drug release to market.

Research Scientist, ***Children’s Hospital*** (Columbus, OH) M*arch 2017 – August 2017*

-**Deciphered** mechanisms of *Respiratory Syncytial Virus* (**RSV**) lifecycle in cotton rats using **parameter optimizations** using *profile likelihood* estimates. **Elucidated** the role of *adaptive immune system* in suppressing RSV. The insights will aid in the *design* & *testing* of new vaccine candidates for RSV.

Postdoctoral Associate, ***Texas Tech Univ.*** (Lubbock, TX**)** *June 2015 – December 2016*

-**Elucidated** new mechanisms of DNA damage reactions for *cancer therapy*. **Achieved** accelerated code performance (2 months wall time to 1 month wall time) using Open-MPI framework.

Postdoctoral Fellow, ***Emory University*** (Atlanta, GA) *June 2013 - May 2015*

-**Characterized** gel bilayers using *molecular simulation*. **Discovered** the importance of the role of *glycerol-backbone disorder* on the final properties of gel bilayers that are not accessible from any SOTA measurement technology.

Graduate Researcher, ***Rice University*** (Houston, TX). *August 2009 - May 2013*

-**Developed** *Monte-Carlo* algorithms & *Markov* models to describe *emergent properties* in biological cells. **Validated** SOTA measurements using *probabilistic* models.

**EDUCATION**

**Ph.D.** in *Chemistry* (2013) RICE UNIVERSITY, Houston, TX

***Thesis***: “*Theoretical Investigation of Biological Transport Processes using Discrete State Stochastic*

*Models and Simulations*”

***Courses***: Statistical Mechanics, Quantum Mechanics, Classical Mechanics, Advanced Differential Equations

**M.Sc.** in *Chemistry* (2008) – INDIAN INSTITUTE OF TECHNOLOGY, Kanpur, India

***Thesis***: “*Fermi Resonant Hamiltonians: Bifurcations, Level velocities and Entanglements*”

***Minor Electives***: Data Structures & Algorithms, Differential Equations, Advanced Calculus

**PUBLICATIONS**

1. Electron Nuclear Dynamics Simulations of Proton-Cancer Therapy Reactions: Water Radiolysis and Proton- and Electron-Induced DNA Damage in Computational Prototypes (**K. Uppulury**, *et al*), Cancers, *10*, 136 (2018)

2. Channel-Facilitated Molecular Transport: The Role of Strength and Spatial Distribution of

Interactions, (**K. Uppulury**, A.B. Kolomeisky), Chemical Physics, 481, 34-41 (2016)

3. Molecular Simulation of the DPPE Lipid Bilayer Gel Phase: Coupling between molecular packing order and tail tilt angle, (**K. Uppulury**, *et al*), J.Phys.Chem B, 119, 8725-8733 (2015).

**232 Citations H-index: 7**

<https://scholar.google.com/citations?user=maJc4_QAAAAJ&hl=en>

**VOLUNTARY ACTIVITY**

Peer-reviewed several top tier journal articles (IEEE Transactions, PCCP, MDPI)

**HOBBIES**

Classical Guitar, Piano