

KUSHAL KUMAR

Senior Applied Scientist @ Amazon, New York

Email: kushlku@amazon.com

Mobile: +1 (929) 669 7636

ACADEMIC DETAILS

Year	Education	College	Grade
2014-2018	B.S. in Mathematics and Scientific Computing [#]	Indian Institute of Technology, Kanpur	8.9/10.0

[#] With a **Minor** in Machine Learning and Applications within Computer Science department with GPA of 10.0/10.0.

KEY PUBLICATIONS

- Unsupervised Multi-Modal Representation Learning for Retrieval of Similar Products at E-commerce Scale, CIKM 2023
- PAVE: Lazy-MDP based Ensemble to Improve Recall of Product Attribute Extraction Models, CIKM 2022
- Solving Price Per Unit Problem: Formulating Fact Extraction as Question Answering, KDD TrueFact 2021

APPLIED SCIENTIST III @ AMAZON.COM

(May 2020 - Present)

- **Demeter: End-to-End GenAI Entity Resolution Workflow** (Oct'23 - Present)
 - Ideated a scalable pay-as-you-go entity resolution workflow, called Demeter, to curate product variation linkages.
 - Developed product representation and matching models that can grow seed density clusters without blocking.
 - The workflow achieved 90% recall @ 99% precision with 5× reduction in audits and efficient GenAI inference.
- **Mergen: Multi-Modal Product Retrieval Service** (Mentor: Ismail Tutar | Sep'22 - Oct'23)
 - Enhanced CLIP in-batch negatives using product taxonomy and optimized batch size for task-agnostic retrieval.
 - Studied ANNS parameters for finding the pareto-optimal configurations based on user specifications such as cost.
 - Achieved 18% lower reduction ratio and 73% cheaper search @ 95% recall. This work was accepted in CIKM'23.
- **PAVE: Product Attribute Value Ensemble** (Mentor: Anoop S V K K Saladi | Mar'21 - Sep'22)
 - Proposed a novel RL agent that improves recall of traditional attribute extraction models without dropping PR.
 - Trained PAVE agent with KB and novel rewards to get 10% absolute lift in RE. This work was accepted in CIKM'22.
- **Solving Price Per Unit Problem in Consumables** (Mentor: Tarik Arici | May'20 - Mar'21)
 - Built a two-step QnA framework for fact extraction task to extract price per unit (PPU) using product description.
 - Achieved < 50ms latency due to lightweight multi-span model. This work was accepted in KDD workshop '21.

SALIENT PROJECTS

- **Advancements in Approximate Nearest Neighbor Search** (Mentor: Hakan Ferhatosmanoglu | Oct'23 - Present)
 - Predictive modeling of search accuracy for efficient ANNS. This work is currently under review in SIGMOD'26.
 - Improving query performance through selective replication of important vectors in partitioning-based indices.
- **Automated Attribution of Value at Risk (VaR)** (Senior Risk Analyst @ Goldman Sachs | Jun'18-Apr'20)
 - Automated portfolio VaR changes report using correlation-based clusters on multivariate Risk Factor vectors.
 - Optimized clustering with PCA and created explanations using decision trees to reduce manual workload by 20%.
- **Predicting Enhanced Satellite Images** (Research Intern @ IBM Research | Mentor: Jagabondhu Hazra | May'17 - Sep'17)
 - Predicted high resolution Landsat-like weekly image from coarse Modis Daily image using prior information.
 - Implemented IDL based work on FSDAF algorithm in Python, making it independent of NASA's ENVI tool.
- **Short Story Endings** (IITK | Mentor: Prof. Harish Karnick | Jan'18-Apr'18)
 - Built a binary classifier that predicts one of the two endings for the given four sentences of a story.
 - Used ROCStories Corpora as the training set, and achieved an accuracy of 63% using feature ensemble methods.
- **Predict Word from Brain Activity** (IITK | Mentor: Prof. Piyush Rai | Aug'16-Nov'16)
 - Predicted the word from a new brain activity query given 300 fMRI brain scan images of 5 persons.
 - Tested several models like Multiclass Regression, SVM, Naive Bayes, LDA & LASSO to get best accuracy of 57/60.

TECHNICAL SKILLS

Languages: Python • SQL • PySpark • C++ **AI/ML:** AWS • Spark • HuggingFace • Pytorch • Tensorflow • PyCharm

AWARDS AND RECOGNITION

- **General Proficiency Medal** for graduating with the highest CGPA in Mathematics & Scientific Computing department.
- Recipient of the **Academic Excellence Award** on two occasions for achieving the highest semester GPA in the batch.
- Achieved an **All India Rank - 933** in IIT-JEE Advanced 2014 examination, competing against 1.35 million candidates.
- Recipient of fellowships such as **National Talent Search Exam**, **Kishore Vaigyanik Protsahan Yojana** and **INSPIRE**.

CO-CURRICULAR ACHIEVEMENTS

- Received the **Special Appreciation Award** for contributions to IIT Kanpur cricket team as the captain.
- **Won bronze medal** as the captain in Inter-IIT 2017 among 21 participating colleges and **gold medal** as a player in 2015.
- Served as a **Student Guide in Counseling Service**, mentoring nine freshmen to foster personal and academic growth.
- Successfully completed a year long **National Cadet Corps (NCC) camp** in IIT Kanpur between 2014-2015.