

Rithesh Kumar

📍 222 Columbia St, Santa Cruz, CA, 95060 ✉ rkumar45@ucsc.edu ☎ 831-266-7201 🔗 [linkedin.com/in/rithesh-kumar-91332a156](https://www.linkedin.com/in/rithesh-kumar-91332a156) 🐙 github.com/Rithesh17 🌐 <https://rithesh17.github.io/>

SUMMARY

I am an AI Researcher with over 3 years of experience in Natural Language Processing and Computer Vision, focusing on Generative AI. My industrial expertise lies in developing and deploying analytical models to drive credit decisioning in my 3 years at Goldman Sachs. I am actively seeking an intern position in Artificial Intelligence research to broaden my knowledge horizon and make substantial contributions to the world.

EXPERIENCE

Decision Scientist - Goldman Sachs

08/2020 - 07/2023, Bangalore, India

- Developed 2 models based on SVM and Knowledge-Tree graphs for customer delinquency and risk scoring for personal loans and further assisted in its deployment. The addition of the new models resulted in a 6% decrease in the overall delinquency rate and an increase in customer satisfaction by 8%
- Built 2 regression models to assist in customer acquisition risk scoring and personalize the loan offers for each customer. Deployment of these models led to an increase in the loans offered by 5%
- Spearheaded the research on changing market conditions during the COVID period by leveraging the tools of LSTM-based time series analysis over the years. The findings of the research were then used to refine the existing policies, leading to an increase in new customers by 10%
- Supervised a team in building a comprehensive end-to-end pipeline for the credit decision and loan offer process on Provenir, leading to a decrease in the latency of loan processing by 15%, an increase in load handling by 13% and reducing the cost of maintenance of the pipeline by 20% compared to the previous pipeline deployed on a legacy platform, FICO.

Machine Learning Research Intern - Sprinklr India

05/2019 - 07/2019, Gurgaon, India

- Developed and deployed a high-performing LSTM model for spam classification and product review analysis on the Twitter database, resulting in a 10% reduction in spam detection errors.
- Conducted in-depth research on model compression algorithms for LSTMs in mobile applications and proposed a teacher-student model that achieved a 60% reduction in model size with minimal loss in accuracy.

SKILLS

Areas of Expertise: Generative AI, Artificial Intelligence, Natural Language Processing, Computer Vision, Reinforcement Learning, Deep Learning, Model Optimization, Image Processing, AI Model Deployment, Machine Learning Model Evaluation, Data Mining and Visualisation, Pattern Recognition.

Languages: Python, R, C, C++, Javascript, SQL, Git.

Tools: TensorFlow, Keras, PyTorch, spaCy, Scikit Learn, openCV, Pandas, Numpy, Scipy, MySQL, MongoDB, Github.

Soft Skills: Software development, Team collaboration, Leadership, Mentoring and training, Problem-solving, Critical thinking, Communication skills, Programming, Project management, Research presentation, Research publication.

EDUCATION

Master of Science in Computer Science • University of California Santa Cruz

Santa Cruz, CA • 09/2023 - 05/2025 • 4 GPA

- Working as a **Teaching Assistant** from Jan 2024

Bachelor of Technology in Computer Science • National Institute of Technology Karnataka

Surathkal, India • 08/2016 - 05/2020 • 3.81 GPA

PROJECTS

Prostate Cancer Grading using Multistage Deep Neural Networks • [Publication](#)

MIND 2021 • 07/2021 - 12/2021 • Springer

- Developed a novel multi-stage deep learning framework for automated Gleason system grading (GSG) and grade group (GG) classification of prostate cancer cells. This approach differs from existing methods by treating Gleason pattern (GP) classification as a classification problem combined with segmentation.
- The system achieved an overall diagnostic accuracy exceeding 90% f1-score for each CNN, demonstrating its effectiveness in GSG and GG classification. Additionally, the paper reports precision and recall values for each GP, both exceeding 90%, beating the state-of-the-art techniques.
- This DL-based system holds promise for enhancing objectivity and efficiency in prostate cancer diagnosis, potentially leading to early detection and cure.

Network Anomaly Detection using Artificial Neural Networks Optimised with PSO-DE Hybrid • [Publication](#)

SSCC 2018 • 01/2018 - 05/2018 • Springer

- Proposed a hybrid PSO-DE algorithm combining Particle Swarm Optimization and Differential Evolution to optimize ANNs for network anomaly detection. The hybrid PSO-DE algorithm leverages the complementary advantages of both techniques to achieve better exploration and exploitation capabilities.
- The optimized ANNs significantly improved anomaly detection accuracy compared to traditional methods. The paper reports an accuracy of 98.7%, which significantly improves the accuracy of traditional ANN-based methods.
- The paper used the NSL-KDD dataset, a standard benchmark dataset for network intrusion detection. The results show that the proposed approach can effectively detect anomalies in real-world network traffic.

Academic Video Summarizer and Enhancer using Retrieval Augmented Generation with LLMs

01/2024 - Present

- Built an LLM based RAG pipeline that takes in the video transcripts and generates a context-based summary enhanced with the concepts from the reading material. The proposed pipeline should generate a summary that is more student-friendly and informative than the naive summary generators.
- Employed few-shot learning and prompt-based fine-tuning techniques to improve performance over the corresponding pre-trained summarizers.
- Preliminary analysis shows promising results that are outperforming the conventional state-of-the-art summary generators. Currently, the project is in the stage of collecting more feedback through a survey and human evaluation of the results.

Relaxation of Differential Privacy in Machine Learning

08/2019 - 05/2020

- Analyzed the importance of Differential Privacy (DP) in Machine Learning, zoning in on the convex bounds for privacy and accuracy in the models.
- Conducted a comprehensive literature survey and investigated the convex properties of relaxed privacy guarantees, with a particular focus on Renyi DP. This research lays the groundwork for the application of these methods in machine learning model security and privacy.
- Proposed a Renyi-DP-inspired Gradient Descent algorithm for Neural Networks, demonstrating its potential to balance individual privacy with strong model performance.