



# Rakesh Bal

 rakeshbal99@gmail.com  (424) 440-9976  rakeshbal99  Rakesh Bal  Rakesh Bal

## Education

### University of California Los Angeles (UCLA)

Master of Science (M.S.) in Computer Science | GPA - 3.81/4.0

Los Angeles, CA

Sep'21 - Jun'23

### Indian Institute of Technology (IIT) Kharagpur

Bachelor of Technology (B.Tech.) in Computer Science and Engineering | GPA - 9.04/10.0

Kharagpur, India

Jul'16 - May'20

## Work Experience

### Stovell AI Systems

Machine Learning Operations Engineer | Supervisor: [Kedar Waghmode](#)

Role: End-to-end management of data pipelines, model training & deployment, tracking of model accuracy and business ROI impact

- Trained custom ML architectures involving Transformers, LSTMs, CNNs and their hybrids to predict fuel price and volume demands
- Achieved 15–30% lower MAE through architecture experiments, advanced time series feature engineering and hyperparameter tuning
- Designed several *causal* models to predict *price-volume elasticity*; devised metrics to evaluate model performance & business impact
- Orchestrated pipelines for data ingestion, processing and model inference; ensuring **99.89%** uptime for all pipelines across customers
- Built monitoring framework w/ automated alerting in *DataDog* for data sanctity and tracking model accuracy across geographies

San Francisco, US

Jan'24 - Current

### Kanini Software Solutions

DevOps Engineer | Supervisor: [Rajendra Aduma](#)

- Automated provisioning of a healthcare platform with Terraform, improving scalability and reliability using AWS S3, Lambda & RDS

Los Angeles, US

Aug'23 - Jan'24

### Amazon Web Services (AWS)

Applied Scientist Intern | Advisors: [Erika Pelaez Coyotl](#), [Marc van Oudheusden](#)

- Benchmarked Protein LLMs (ProtBERT/T5/GPT) and integrated Diffusion Docking on Drug Target Interaction (DTI) prediction
- Employed data parallel distributed training framework w/ P6.48xlarge EC2 instance for training leading to 2x memory cost reductions
- Enhanced baseline performance by 20% and 12% on two datasets with ProtBERT; project showcased at Amazon ML Conference

San Francisco, US

Jun'22 - Sep'22

### Goldman Sachs

Software Engineer | Supervisor: [Nitin Garg](#)

- Integrated two internal bug & issue tracking softwares as a full-stack development project and worked on their migration to AWS
- Architected and built scalable RESTful APIs using Java w/ Spring Boot and intuitive UI with TypeScript-Angular and Redux
- Coordinated with end-users for the entire SDLC; deployed to production and handled adoption by over **50000** daily active users

Bangalore, India

Aug'20 - Sep'21

### Accenture Technology Labs

Research Intern | Advisor: [K. M. Annervaz](#)

- Engineered a GCN-based stock price prediction model integrating news and knowledge graphs, improving MSE by 5% over baseline

Bangalore, India

May'19 - Aug'19

### University of California Los Angeles

Graduate Teaching Associate | Instructors: [Huiling Shao](#), [Abby Kavner](#), [David Smallberg](#)

- Managed and led discussions/OH for 500 undergrads in CS32 (Intro to CS II) & Chem 20A, totaling over 500 hours in 5 quarters

Los Angeles, US

Sep'21 - Jun'23

## Publications

### Analysing the Extent of Misinformation in Cancer Related Tweets | Rakesh Bal *et al.*

[Paper]

- 14th International AAAI Conference on Web and Social Media (ICWSM 2020) | **42 citations**

- Curated a new dataset and developed an attention-based BiLSTM-CRF model to detect/analyze cancer misinformation on Twitter

### Two-Sided Fairness in Non-Personalised Recommendations | Rakesh Bal *et al.*

[Paper]

- 35th AAAI Conference on AI (AAAI 2021) Student Abstract and Poster Program | **9 citations**

### PGraphDTA: Improving DTI Prediction using Protein Language Models & Contact Maps | Rakesh Bal *et al.*

[Paper]

- 37th Advances in Neural Information Processing Systems (NeurIPS 2023) Workshop on AI for Science | **5 citations**

## Relevant Projects

### DeepSub: Fine-tuning LLMs for multi-source subtitle translation

[Code]

- Fine-tuned Gemma & LLaMA using QLoRA to synthesize context from 2 source languages and translate subtitle to a target language
- Evaluated a triplet translation pipeline and experimented with 6 *Indic* & 3 *European* languages with BLEU performance tracking

### CLIP for Visual Question Answering (VQA)

[Code]

- Harnessed OpenAI CLIP in VQA models like MCAN and Pythia in both zero-shot and finetuned settings with 2% upgrade
- Added Language Driven Semantic Segmentation (LSeg) to pipeline for answering number-based questions in VQA2.0 dataset

### Autonomous Underwater Vehicle (AUV), IIT Kharagpur

[Link]

- Implemented real-time underwater buoy detection using MobileNet and made inference of the model lightweight by 20%
- Used ROS packages like Actionlib and Smach to implement the Mission Planner Stack, participated in SAVe at NIOT, Chennai

## Skills

**Languages:** Python, Java, C, C++, SQL, JavaScript, TypeScript, Matlab, Go

**ML:** PyTorch, TensorFlow, Keras, Pandas, Sklearn, KubeFlow, PT-Lightning

**DevOps:** Docker, K8s, GCP, AWS, Azure, Terraform, Jenkins, Snowflake, DBT

**Web:** Angular, React, Flask, Node.js, MySQL, GraphQL, SpringBoot, MongoDB

## Achievements

– IIT-JEE Advanced: All India Rank **265**/200K

– IIT-JEE Mains: All India Rank **187**/13M

– KVPY Fellowship: All India Rank **191**/100K

– Puri District Chess Cup: **5th**/1000 players