

Karan Bhukar

MSCS STUDENT, NYU COURANT

✉ choudhary121karan@gmail.com | 🌐 karan121bhukar | 🏠 karan121bhukar.github.io | in karan-bhukar | 🏠 Karan Bhukar

Education

New York University Courant

MASTER OF SCIENCE IN COMPUTER SCIENCE

CGPA: **-/4.00**

September 2026 - Present

Indian Institute of Technology Hyderabad

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE & ENGG.

CGPA: **9.04/10.00**

Aug 2018 - June 2022

Key Coursework: Reinforcement Learning, Machine Learning, Deep Learning for Vision, Linear & Convex Optimization, Linear Algebra, Probability

Publications

C=Conference Paper, W=Workshop Paper, M=Manuscript. * denotes Equal Contribution

[C.5] Scalable and Efficient Large-Scale Log Analysis with LLMs: An IT Software Support Case Study 📄

Gupta P., **Bhukar K.**, Kumar H., Nagar S., Mohapatra P., Kar D.

(40th **AAAI** 2026)

[C.4] LogAn: An LLM-Based Log Analytics Tool with Causal Inferencing 📄

Gupta P., **Bhukar K.**, Kumar H., Nagar S., Mohapatra P., Kar D.

(25th **ICPE** 2025)

[C.3] Dynamic Alert Suppression Policy for Noise Reduction in AIOps 📄

Bhukar K., Kumar H., Mahindru R., Nagar S., Aggarwal P., Arora R., Paradkar A.

(46th **ICSE** 2024)

[W.1] Dynamic-X-Y: A Tool for Learning Dynamic Alert Suppression Policies in AIOps 📄

Bhukar K., Kumar H., Arora R., Mahindru R., Nagar S., Thornhill M., Manning I., Buggins J., Cook S., Paradkar A., Aggarwal P.

(16th **COMSNETS** 2024)

[C.2] Learning Representations on Logs for AIOps 📄

Gupta P., Kumar H., Kar D., **Bhukar K.**, Aggarwal P., Mohapatra P.

(16th **IEEE Cloud** 2023)

[C.1] End-to-End Deep Reinforcement Learning for Conversation Disentanglement 📄

Bhukar K., Kumar H., Raghu D., Gupta A.

(37th **AAAI** 2023)

[M.1] Target-Agnostic Adversarial Attacks on Language Understanding Models 📄

Chauhan J., **Bhukar K.**, Kaul M.

(arXiv preprint 2021)

Patents

P=Patent, S&M=Systems and Methods

[P.6] S&M for Logline Enrichment for Error Analysis

Bhukar K., Kumar H., Nagar S., Gupta P., Samanta S., Mohapatra p., Kar D., Aggarwal P.

U.S. Patent under filing 2025

[P.5] S&M for Explainability-Driven Control of an Alert Suppression Policy 📄

Bhukar K., Kumar H., Nagar S., Mahindru R., Paradkar A., Manning I., Hussey S.

U.S. Patent 2026

[P.4] S&M for Determining Log Templates of a Computing Device for Log Templatzation 📄

Bhukar K., Gupta P., Kumar H., Verma M.

U.S. Patent under filing 2024

[P.3] S&M for Dynamic Policy Management for Events and Alerts Noise Reduction in IT systems 📄

Bhukar K., Nagar S., Kumar H., Aggarwal P., Mahindru R., Paradkar A., Manning I., Thornhill M., Arora R., Hussey S., Forti F.

U.S. Patent 2023

[P.2] S&M To Recommend/Revoke Agents Guided by the AIOps Conversation Context 📄

Bhukar K., Kumar H., Azad A., Joshi S.

U.S. Patent 2023 & CN/JP Patent 2024

[P.1] S&M to Automatically Generate Close Notes from SRE Interactions for Downstream AIOps Tasks 📄

Bhukar K., Kumar H., Azad A., Joshi S.

U.S. Patent 2023

Relevant Industrial Research Experience

Amazon

Bengaluru, India

APPLIED SCIENTIST

Oct 2024 - Aug 2026

• Multilingual Search Query Understanding at Scale

- Architected a unified transformer model for spell-correction and music NER, replacing legacy edit-distance and BiLSTM-CRF pipelines.
- Shipped a highly robust multilingual model (US, CA, MX, BR) capable of handling complex English-Spanish-Portuguese code-switching.
- Generated ~5B training sequences spanning 100M+ entities by building an LLM and metadata-driven data pipeline.
- Scaled **distributed training** (hybrid FSDP) for **billion parameter scale models**, driving down training time and infrastructure costs.
- Boosted **core search metric (SRPR)** by **151 bps** in South America and **61 bps** in North America, directly improving user search discovery.

• Popularity Bias in Music Search Ranker [Ongoing Research] (with Prof. Arpit Agarwal, IIT Bombay)

- Consulting **Lead Scientist** for the **Amazon-IIT Bombay** collaboration, investigating fairness in SOTA ranking algorithms.
- Analyzing the long-term effects of popularity bias on user retention, focusing on the churn rate of users with niche content preferences.
- Guiding the development of a novel theoretical framework (Coupled Dynamical System Analysis and Two-Time Scale Stochastic Approximation) by reconciling constraints and insights from large-scale production environments.
- Leading the empirical validation of the theoretical models using extensive production data to quantify bias impact and substantiate findings for a high-impact publication.

IBM Research

RESEARCH ENGINEER MENTORED BY HARSHIT KUMAR, DINESH RAGHU

Gurgaon, India

June 2022 - Oct 2024

- **BERTOps, Pretrain-to-Production Log Analytics (Demo: 🎥)** (Oral at IEEE CLOUD '23, ICPE '25 & IAAI '26) [C.2] [C.4] [C.5] [P.6]
 - Built BERTOps, a domain-specialized masked-LM for semi-structured IT logs, to capture template regularities that general LLMs miss.
 - Released labeled datasets and a unified benchmark for Log Format Detection, Golden Signal Classification, and Fault Category Prediction including the first public labels for GSC/FCP.
 - Productionized models into a log-analytics tool; optimized for CPU-only model inference for millions of logs via log templating.
 - At production scale, saved 8K engineer-minutes by processing 1.04B log lines (877 GB) across 100+ internal software support teams.
- **False Alert Suppression on Observability Platform (Demo: 🎥)** (Oral at ICSE '24 and Workshop at COMSNET '24) [C.3] [W.1] [P.3] [P.5]
 - Designed *Dynamic-X-Y*: an unsupervised alert-suppression policy learner using moving-average envelopes over historical alert streams; applies per-signal dynamic thresholds at runtime.
 - On log and metric anomaly datasets, improved accuracy over baselines by **+7.39%** and **+35.7%**, respectively.
 - Built a train-to-deploy pipeline that learns per-signal X-in-Y suppression thresholds from historical alerts and pushes them as runtime policies via AI Model Management.
 - In internal deployments, cut non-actionable incident notifications by ~90% while preserving critical alerting.

IBM Research

RESEARCH INTERN MENTORED BY HARSHIT KUMAR, AJAY GUPTA, DINESH RAGHU

Gurgaon, India

July 2021 - Sept 2021

- **Reinforcement Learning for Conversation Disentanglement(CD)** (Oral & Poster at AAAI '23; Patents in US, CN & JP) [C.1] [P.1] [P.2]
 - Framed CD as end-to-end RL problem, optimizing a global conversation-level objective over local link predictions.
 - Designed a global clustering-based reward and a policy that sequentially assigns messages to threads.
 - Proposed a novel hybrid reward balancing the global disentanglement metric with local link-level accuracy.
 - Translated the research into applied systems, resulting in two patent applications one pursued internationally (US/CN/JP) and one US-only.

ServiceNow

MACHINE LEARNING INTERN

Gurgaon, India

May 2021 - July 2021

- **Multilingual Lexical Normalization for Noisy IT Incidents**
 - Identified incident-ticket classifier's brittleness due to lexical noise; scoped pre-classification normalization problem in the pipeline.
 - Implemented a probabilistic spell-correction + word-segmentation module based on Norvig's noisy-channel approach and segmentation scoring, adapted to IT incident jargon.
 - Built multilingual normalization (EN/FR/ES/DE) via language-specific frequency models and dictionaries
 - Integrated normalizer into production; improved downstream classification accuracy by +5%.

Research Experience

Lab 1055, The Machine Learning and Vision Lab, IIT Hyderabad

Hyderabad, India

UNDERGRADUATE RESEARCHER MENTORED BY PROF. VINEETH N. BALASUBRAMANIAN, PROF. C.V JAWAHAR

May 2020 - May 2021

- **AutoNUE'21 - Domain Adaptation & Semantic Segmentation Benchmarks for IDD (CVPR Workshop)**
 - Defined and baselined five benchmarks: Supervised, Semi-supervised, Weakly-supervised, Unsupervised Domain Adaptation, and Semantic Segmentation for workshop challenge.
 - Set a clear cross-domain setup: train on Mapillary, Cityscapes (fine), BDD, and GTA; evaluate on IDD (Level-3, 26 classes) to measure adaptation to Indian roads.
 - Released public baseline code & data-prep pipelines for reproducible training, evaluation, and leaderboard submission. 🌐

Krama Lab, IIT Hyderabad

Hyderabad, India

UNDERGRADUATE RESEARCHER MENTORED BY PROF. MANOHAR KAUL

Aug 2019 - May 2020

- **Target Agnostic Black-Box Adversarial Attacks for NLU** (arXiv preprint) [M.1]
 - Proposed a target-model-agnostic black-box attack for NLU under strict query budgets with strong cross-model transferability.
 - Built an offline adversarial-generation pipeline using enhanced ON-LSTM + multi-head attention with synonym-constrained substitutions that preserve semantics.
 - Demonstrated impact: on MNLI, BERT accuracy drops from 82.7% → 61.6% with budget $K=20$, exceeding baseline degradation by $> 15\times$.
 - Achieved practical efficiency: batched attacks run $9-12\times$ faster than baselines; a staggered strategy reuses adversaries across models.

Pertinent Skills

Languages Proficient in: Python, C++; Familiar with: SQL, Java, JavaScript, Haskell, Prolog, HTML, CSS, Eiffel, MATLAB

Software & Tools JAX, PyTorch, TensorFlow, Transformers, Pyspark, Spark, Django, AWS, $\LaTeX$, Git, Vim

Talks & Presentations

- 2024 **46th International Conference on Software Engineering**, @ Centro Cultural de Belém, Lisbon, Portugal
- 2024 **IBM Research Industry Expo Booth**, @ 37th NeurIPS, New Orleans, USA
- 2023 **37th AAAI Conference on Artificial Intelligence**, @ Walter E. Washington Convention Center, Washington, D.C., USA

Selected Scholastic Achievements

- 2024 **Outstanding Technical Achievement Award (OTAA)**, IBM Research's highest distinction for technical excellence, recognizing innovative work that enabled multi-million-dollar business growth.
- 2024 **Entrepreneurship Award**, awarded at IBM for end-to-end ownership of a project that matured into a major internal product.
- '22-24 **6 Inventor Award**, acknowledged at IBM for successful filing of multiple patents in USPTO
- 2021 **Among 50**, from **10,000+** applicants selected for the prestigious **IBM Research Blue Scholar Internship Program**.
- 2018 **All India Rank of 681**, among **0.15 million** candidates in the **JEE Advanced** conducted by the Joint Admission Board.
- 2018 **All India Rank of 358**, among **1.5 million** candidates in the **JEE Main**, conducted by CBSE.
- 2018 **KVPY Fellowship**, by the Dept. of Science, Government of India, among top **2000** out of **100,000+** applicants.
- 2016 **NTSE Scholarship, Ranked 5**, in Rajasthan state conducted by NCERT, Government of India, among **100,000+** applicants.

Aposite Volunteering & Community Service

- 2024 **Reviewer @ AAAI '24**, for submissions pertaining to AAAI '23 research track
- 2024 **Internship Mentor**, for Shivangi Shreya, CS Masters student at IIT Madras
- 2020-21 **Volunteer @ COLING '20 & NeurIPS '21**, assisted in the online conference organization during covid
- Fall '21 **Teaching Assistant @ Dept. of CSE, IIT-H**, Foundations of ML, lectured by **Prof. Vineeth N. Balasubramanian**
- Fall '20 **Teaching Assistant @ Dept. of CSE, IIT-H**, Data Structures, lectured by **Prof. Karteek Sreenivasaiah**
- Spring '19 **Teaching Assistant @ Dept. of CSE, IIT-H**, Artificial Intelligence, lectured by **Prof. P. K. Srijith**

Notable Positions Of Responsibility and Extracurriculars

- **Core @ ACM-IITH Student Chapter**
 - POR: **Treasurer**, for academic year (2021)
 - POR: **Moderator**, for Theoretical Computer Science group (2020)
- **Co-ordinator @ Sports Council, IIT Hyderabad**: Table Tennis ('19-'21)
- **Overall Champion** representing the Charaka block @ Milan, **The General Championship**, IIT-H, Mar 2022 and Runners Up in 2020, via podiums in events like Table Tennis and E-sports (CSGO, Valorant).
- Represented IIT Hyderabad at The **Inter IIT Sports Meet** for Table Tennis at IIT Bhubaneswar, Dec 2018.
- Represented IIT Hyderabad at The **Inter IIT Sports Meet** for Table Tennis at IIT Guwahati, Dec 2018.
- Represented Rajasthan (Home State) at **National Table Tennis Championship** organized by TTFI at Gandhidham, Gujarat, 2013.
- Represented Rajasthan (Home State) at **National Table Tennis Championship** organized by SGFI at RAJKOT, Gujarat, 2012.