Bhiman Kumar Baghel

+1-412-844-1085   ✉ bkb45@pitt.edu   🌐 bhimanbaghel.github.io   in bhiman-kumar-baghel   🎓 Google Scholar

📍 130 N Bellefield Ave, Pittsburgh, PA 15260, USA

RESEARCH OVERVIEW

My research focuses on improving the internal representations of large language models (LLMs) through direct model editing. I formalized key limitations in locate-and-edit algorithms, specifically UnderEdit and OverEdit, and developed methods to address them. This work led to a 38 percentage point improvement over the previous state-of-the-art. More recently, I've begun exploring editing techniques to enhance the legal reasoning capabilities of LLMs, with the goal of making them more accurate, interpretable, and trustworthy.

EDUCATION

- **PhD IN COMPUTER SCIENCE** — AUG. 2023 - PRESENT
  University Of Pittsburgh, PA, USA
  Advisor: DR. XIANG (LORRAINE) LI

- **M.TECH IN COMPUTER SCIENCE** — JUL. 2017 - MAY 2019
  Indian Institute of Technology (IIT), Kharagpur, India
  Thesis: DCLL - A Deep Learning Model for Travel Time and Traffic Congestion Prediction

- **B.TECH IN COMPUTER SCIENCE** — JUN. 2013 - JUN. 2017
  National Institute Of Technology (NIT), Jalandhar, India

PUBLICATIONS

[5] **Bhiman Kumar Baghel**, Scott M. Jordan, Zheyuan Ryan Shi, Xiang Lorraine.
"Resolving UnderEdit & OverEdit with Iterative & Neighbor-Assisted Model Editing."
arXiv, 2025. [PDF]

[4] **Bhiman Kumar Baghel**, Arun Balajiee Lekshmi Narayanan, Michael Miller Yoder.
"A Fairness Analysis of Human and AI-Generated Student Reflection Summaries."
ACL GeBNLP Workshop, 2024. [PDF] [Talk]

[3] Yang Zhong, **Bhiman Kumar Baghel**.
"Multimodal Understanding of Memes with Fair Explanations"
CVPR MULA Workshop, 2024. [PDF] [Talk]

[2] Niraj Kumar, **Bhiman Kumar Baghel**.
"Intent Focused Semantic Parsing and Zero-Shot Learning for Out-of-Domain Detection in Spoken Language Understanding"
IEEE Access Journal, 2021. [PDF]

[1] Niraj Kumar, **Bhiman Kumar Baghel**.
"Smart Stacking of Deep Learning Models for Granular Joint Intent-Slot Extraction for Multi-Intent SLU"
IEEE Access Journal, 2021. [PDF]

PATENTS

[4] Sourabh Tiwari, **Bhiman Kumar Baghel**, Jalaj Sharma, Manish Chauhan, Boddu Venkata Krishna Vinay, Syed Khaja Moinuddin.
"Method and system for time based personalization management in multi-device environment"
Filed with Samsung Electronics Co Ltd, 2024. Patent Application No. WO2025018568A1. [Link]

[3] Venkata Krishna Boddu Vinay, **Bhiman Kumar Baghel**, Gorang Maniar, Syed Khaja Moinuddin, Sudhansu Ranjan Acharya.
"Methods and systems for enabling seamless indirect interactions."
Filed with Samsung Electronics Co Ltd, 2023. Patent Application No. US 18517995. [Link]

[2] Niraj Kumar, **Bhiman Kumar Baghel**.
"Method and system for mitigating physical risks in an iot environment."
Filed with Samsung Electronics Co Ltd, 2023. Patent Application No. US 18202687. [Link]

[1] Niraj Kumar, **Bhiman Kumar Baghel**.
"Methods and systems for determining missing slots associated with a voice command for an advanced voice interaction."
Filed with Samsung Electronics Co Ltd, 2023. Patent Application No. US 17835387. [Link]

WORK EXPERIENCE

• **GRADUATE RESEARCH ASSISTANT**   AUG. 2024 - PRESENT
Department of Computer Science, University of Pittsburgh, USA

- Improved model editing performance by **38 percentage points** over the **state-of-the-art** in LLaMA-3, LLaMA-2, and GPT-J, enabling efficient knowledge updates without full fine-tuning.
- Applied memory editing and LoRA to boost legal reasoning in LLMs (LLaMA-3, OPT), achieving **100% accuracy** across **100+ tasks** and improving interpretability.
- Analyzed gender bias in GPT-3.5 and BART-generated student reflection summaries, identifying a **10% male skew** using Jensen-Shannon Divergence.
- Created a **2.9K-meme** dataset and found **40%** of LLaVA/MiniGPT-4 explanations biased; traced sources to visual cues, named entities, and text-image imbalances.

• **LEAD NLP ENGINEER**   JUN. 2019 - AUG. 2023
Samsung Research Institute, Bangalore

- Led development of **CoSMIC**, a multi-intent BERT-based NLU system for Smart Home, achieving **96% accuracy** and **reducing errors by 67%** in real user interactions.
- Scaled CoSMIC to the Korean market by guiding a parallel team at **Samsung HQ**, resolving tokenization bottlenecks and **boosting performance by 25%**.
- Designed deep learning models for Conversational AI (Alexa-style), improving multi-intent recognition (**91%**) and OOD detection (**90%**) across production pipelines.

• **MACHINE LEARNING RESEARCH INTERN**   MAY 2018 - JUL. 2018
IBM, Bangalore, India

- Developed a self-learning LSTM-based system for predicting and auto-resolving IT infrastructure failures across 33 metrics, achieving **97% accuracy**.

PROJECTS

• **Chat-Enabled AI Web Agent for complex flight search**   *PhD Term Project, CMU*
Tools: BrowserGym, Gradio, OpenAI GPT-4o, Pytorch   [▶]

- Designed modular prompting strategies enabling the agent to reason over multi-step flight search actions based on dynamic browser observations and user goals, enhancing the agent's temporal and spatial reasoning capabilities.

• **Automatic Concept Map generation from Learning Material**   *Term Project, IIT Kharagpur*
Tools: PySpotlight, FastText, Stanford CoreNLP   [○]

- Built a pipeline to generate and visualize concept maps from Wikipedia by extracting entities and semantic relations using entity linking, word embeddings, and syntactic parsing.

• **Twitter Sentiment Analysis**   *B.Tech Final Project, NIT Jalandhar*
Tools: Tweepy, NumPy, Scikit-learn, Flask   [○]

- Developed a web application that fetches real-time tweets based on user queries and classifies their sentiment (positive, negative, neutral) using a Naive Bayes classifier.

Honors and Awards

• **Samsung High Performance Bonus (Three times)** *2023*
Samsung Research Institute, Bangalore, India

  ◦ Spearheading the development of CoSMIC, a Smart Home AI Assistant, achieving 96% accuracy with 67% error reduction in real user interactions.

  ◦ Pioneering Conversational AI research, resulting in 4 US A1 patents and 2 papers.

• **Excellence Award (Five times)** *2023*
Samsung Research Institute, Bangalore, India

  ◦ Pioneering Conversational AI research, resulting in 4 US A1 patents and 2 papers.

  ◦ Samsung CLab Finalist for innovative solution to indirect communication in smart home.

• **2nd Runner-Up - Audience Poll Category** *2018*
IBM, Bangalore, India

  ◦ IBM Extreme Blue Expo outstanding project.

Teaching Experience

• **GRADUATE TEACHING ASSISTANT** AUG. 2023 - JUL. 2024
Department of Computer Science, University of Pittsburgh, USA
Courses: Intro to Natural Language Processing, Operating Systems, Programming, Discrete Mathematics

• **GRADUATE TEACHING ASSISTANT** JUL. 2017 - JUL. 2018
Indian Institute of Technology (IIT), Kharagpur, India
Courses: Intro to Programming and Data Structures, Intro to Database Management Systems

Mentorship Experience

• Avyukth R. Nilajagi FEB. 2025 - PRESENT
BS Computer Science, University of Pittsburgh, USA
Project: Preventing model collapse in LLMs due to model editing

• Yuelong Xu SEP. 2024 - DEC. 2024
MS Computer Science, University of Pittsburgh, USA
Project: Tracing commonsense knowledge in LLMs

• Samsung PRISM - PReparing and Inspiring Student Minds MAY. 2023 - AUG. 2023
Samsung Research Institute, Bangalore, India
Project: Multi-intent spoken language understanding in code-mixed language

Leadership Experience

• **EVENT MANAGER** APR. 2025 - PRESENT
ANKUR, International Graduate Student Organization Pittsburgh, USA

Review Service

• Multimodal Learning and Applications (MULA) Workshop, CVPR 2025

• Gender Bias in Natural Language Processing (GeBNLP) Workshop, ACL 2025

• Representation Learning for NLP (RepL4NLP) Workshop, NAACL 2025

SKILLS

- **Deep Learning Frameworks:** PyTorch, TensorFlow, PEFT, Hugging Face, vLLM, Lightning, Keras, MLFlow
- **Programming Languages:** Python, C, C++, JavaScript
- **Web Technologies:** HTML, CSS, PHP, CakePHP, React, Bootstrap, Flask, Django
- **Databases:** SQL, MongoDB, Neo4j

REFERENCES

DR. XIANG (LORRAINE) LI
Assistant Professor, Department of Computer Science, University of Pittsburgh, USA
🌐 https://lorraine333.github.io/   ✉ xiangli@pitt.edu

DR. MICHAEL MILLER YODER
Teaching Assistant Professor, School of Computing and Information, University of Pittsburgh, USA
🌐 https://michaelmilleryoder.github.io/   ✉ mmyoder@pitt.edu

DR. NIRAJ KUMAR
Research Manager, Fujitsu Research, Bangalore, India
🌐 https://www.nirajai.com/   ✉ nirajrkumar@gmail.com