

KSHITISH GHATE

<https://kshitishghate.github.io/> • [Google Scholar](#) • [linkedin.com/in/kshitish-ghate/](#) • kghate@andrew.cmu.edu • (878)-834-9301

EDUCATION

CARNEGIE MELLON UNIVERSITY - SCHOOL OF COMPUTER SCIENCE Pittsburgh, PA
Master of Science in Intelligent Information Systems, LANGUAGE TECHNOLOGIES INSTITUTE May 2025
Relevant Courses : Advanced Natural Language Processing, Introduction to Machine Learning (graduate level), Multilingual NLP

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI Goa, India
Bachelor of Engineering in Computer Science and Master of Science in Economics | CGPA : 8.99/10.00 Mar. 2023
Relevant Courses : Artificial Intelligence, Data Structures and Algorithms, Database Management Systems, Foundations of Data Science, Game Theory, Machine Learning, Object Oriented Programming, Probability and Statistics

EXPERIENCE

Carnegie Mellon University Pittsburgh, PA
Graduate Research Assistant Sept. 2023 - Present

- Assessing incorporation of personal information in language models with the goal of developing a decoding-time intervention mechanism that restricts the generation of personal data, under the supervision of Prof. [Mona Diab](#).

Implicit Social Cognition Lab, Harvard University Cambridge, MA
Research Assistant Mar. 2022 - Aug. 2023

- Collaborated on artificial intelligence fairness research with Prof. [Mahzarin Banaji](#) at Harvard, Prof. [Aylin Caliskan](#) at the University of Washington and Prof. [Tessa Charlesworth](#) at the Kellogg School of Management.
- Measured impact of individual social groups in propagating intersectional biases in large language models. [\[paper in preparation\]](#)
- Demonstrated propagation of social biases from intrinsic settings to downstream zero-shot settings in language-vision models such as CLIP, BLIP-2 and FLAVA. [\[paper in preparation\]](#)

Amazon Alexa – AI Bangalore, India
Applied Scientist Intern Aug. 2022 - Dec. 2022

- Implemented a novel training methodology and model architecture drawing from Curriculum Learning literature to address problem of classifying long tail data in natural language understanding tasks.
- Achieved 5% improvement in Slot F1 and Intent Classification accuracy by applying a holistic sample difficulty metric in training.

Cambridge Centre for Alternative Finance, Cambridge Judge Business School Cambridge, UK
Research Intern Jun. 2021 - Jun. 2022

- Automated the data extraction process and streamlined data analysis workflow for the Global Benchmarking Research team.
- Published the [Global COVID-19 Fintech Market and Impact Resilience](#) Study detailing impact of COVID-19 on Fintech.

SELECTED PUBLICATIONS

- [Kshitish Ghate*](#), Tessa E. S. Charlesworth*, Aylin Caliskan, Mahzarin R. Banaji, (in preparation), "Identifying and validating flexible methods for the study of intersectional stereotypes in language".
- Aswini Kumar Mishra and [Kshitish Ghate](#), (2022), "Dynamic connectedness in non-ferrous commodity markets: Evidence from India using TVP-VAR and DCC-GARCH approaches". Resources Policy, 76, 102572. [\[paper\]](#)
- Kakade Kshitij, Aswini Kumar Mishra, [Kshitish Ghate](#), and Shivang Gupta, (2022), "Forecasting Commodity Market Returns Volatility: A Hybrid Ensemble Learning GARCH-LSTM based Approach." Intelligent Systems in Acc., Fin. and Management. [\[paper\]](#)

SKILLS

Languages Python, R, STATA, C/C++, Matlab, Java (intermediate), SQL (intermediate)
Toolkits PyTorch, TensorFlow, Keras, Scikit-Learn, Pandas, NumPy, Matplotlib, PySpark (intermediate)

SELECTED PROJECTS

Self-Supervised Learning for Tabular Data Annotation BITS Pilani | Jul. 2021

- Developed a self-supervised learning algorithm using an innovative reconstruction loss to classify large proportions of unlabelled data. [\[under review at IEEE Transactions on Affective Computing\]](#)
- Accomplished significant reduction in labelling overhead of continuous emotion annotation based tabular data.

COVID-19 Pandemic Sentiment and Stock Market Behavior BITS Pilani | Jun. 2020

- Determined impact of COVID-19 pandemic sentiment on stock market behavior by constructing a novel text-based attention index to serve as a proxy for the public's attention to the pandemic. [\[paper\]](#)