Shashank Kirtania

Research Fellow, PROSE Team
Microsoft

shashankkirtania123@gmail.com
LinkedIn | GitHub |
Website | Google Scholar

Education

Thapar Institue of Engineering and Technology 2019 - 2023
Computer Engineering

Research Interests

Reasoning and Planning with foundational Models, Multi-Agent systems, AI4Code, Neuro-Symbolic reasoning

Experience

Microsoft Research Sep. 2023 – Present
Research Fellow (Predoctoral) with the PROSE group *Remote / Bengaluru*

- Working on automated fixing of GitHub issues over python repositories.
- Created pipeline to localise Terrascan policy violation using static tools that self-improve with developer feedback.
- Created a MultiAgent System for automated knowledge base improvement for RAG pipeline for code migration.
- Worked on automatic code repair based on dependabot flag for component governance in C# repositories.

Wadhwani AI Jan. 2023 - July 2023
Machine Learning Research Intern *Remote / Bengaluru*

- Worked on identifying Lexical reading complexity using statistical analysis using Dirichlet's prior along with other Seq2Seq models for improved understanding of phonetic complexity of text in low resource languages.
- Worked in developing solution for Oral Reading Fluency assessment in public schools of Gujarat for grade 3 to 8.
- Collaborated with HCI team to improve shortcomings of existing solutions to assess reading capabilities.
- Tool was used by more than 50,000 teachers to conduct more than 10,00,000 unique assessments by end of 2023.

Google Summer of Code, PyMC May 2022 - Sep 2022
Open Source Contributor *Remote*

- Built APIs to improve PyMC deployment pipeline for probabilistic machine learning models.
- Increased test coverage of PyMC & PyMC-experimental package from 86% - 91% & 83.8% - 88.1% respectively.
- Contributed to the docker image and built CI/CD pipelines to automate the updates on dockerhub.

Indian Institute of Technology, Delhi June 2022 - July 2022
Research Intern *Remote*

- Curated dataset for time series forecasting for day ahead bidding at Indian Energy Exchange (IEX).
- Used wavelet transform to extrapolate on non-stationary data with 90% prediction interval for week ahead pricing.
- Benchmarked statistical models and SOTA neural networks to predict day ahead pricing.

Indian Institute of Information Technology June 2021 - May 2022
Research Intern *Remote*

- Implemented High-Throughput JPEG2000 (lossless) compression for images, reducing image size by over 40%.
- Analyzed the impact of DWT coefficient bands (HH and HL) on classification accuracy, enhancing interpretability.
- Finetuned U-net to segment images using DWT coefficients in a semi-compressed domain.

Publications

- Priyanshu Gupta*, **Shashank Kirtania***, Ananya Singha, Sumit Gulwani, Arjun Radhakrishna, Gustavo Soares. *MetaReflection:Learning Instructions for Language Agents using Past Self-Reflections.* preprint ***under review in EMNLP 2024***

- **Shashank Kirtania**, Priyanshu Gupta, Arjun Radhakrishna. *Logic-LM++: Progressive Refinement in Semantic Comprehension can Improve Logic Enhanced LM Systems* Natural Language Reasoning and Structured Explanations Workshop @ **ACL 2024**

- **Shashank Kirtania**, Shalini Batra. *NoTouch: Contactless Fingerprint Verification.* **Best Capstone Research Project** ***2023***

- Tejasvee Bisen, M Javed, **Shashank Kirtania**, P Nagabhushan. *DWT-CompCNN: Deep Image Classification Network for High Throughput JPEG 2000 Compressed Documents*. Pattern Analysis and Applications, Springer, June, 2023. *link*

Technical Skills

Languages: Python, C++, C#, TeX, Shell, MATLAB*, Julia* * = Elementary Proficiency
Software and Tools: PyTorch, AutoGen, Langchain, Fairseq*, Flask, Git, Docker, Postman

Selected Projects

Autofill for data manipulation in Julia Aug 2023

Julia, Programming by Example

- Implemented PBE based approach same as FlashFill, to generate programs from input-output examples.
- Developed DAGs with iterative intersection and unification functions for data process optimization.

Sticky-Stuff incident analysis with PyMC Dec 2022

Statistical Modeling, PyMC

- Implemented a change-point model from scratch using PyMC to interpret raw data from MLB 2021 stats of pitcher's swing rates to find the exact date of new regulations on tampering.
- Presented work at PyData Delhi in 2023.

Achievements and Position of Responsibilities

- Positioned 6th nationally among 5000 + teams in Amazon ML challenge 2023.
- Bronze Medal in American Express Time Series forecasting Challenge 2022.
- PyData Delhi talk on Introduction to Probabilistic Programming.
- Joint Secretary of Creative Computing Society at TIET in 2022.
- Head of Photography at Fine Arts and Photography Society at TIET in 2021.
- Top 20 percentile in Facebook Hackercup 2020 in Round-1.