

Amartya Hatua

Boston, MA | 601.307.8198 | amaryahatua@gamil.com

[Senior Member IEEE](#) | [GitHub](#) | [Google Scholar](#)

Summary

- A solution-oriented data scientist working for Fidelity Investments for the last 3 years.
- Technically proficient software engineer with 3 years of industry experience playing a pivotal role in the development of LG Smart android phones G2, G3, G4, G5.
- An innovator filed 2 patents on Smart TV and 1 patent on AR compatible devices.

Work Experience

Senior Data Scientist

March 2021-Present

Fidelity Investments

- **Lifetime Value Calculation (LTV):** Employed various machine learning methods to forecast the Lifetime Value (LTV) for both current and potential clients. Utilized a combination of seven distinct data sources, including internal data from Fidelity and publicly available information. Emphasis was placed on enhancing prediction explainability and benchmarking data as key components of this initiative.
- **Personal Identifiable Information (PII) Document Extraction:** Designed a pipeline empowered by Large Language Models (LLM) to streamline the process of extracting data from Personally Identifiable Information (PII) documents, including Driver Licenses, Birth Certificates, Death Certificates, and Marriage Registrations, thereby minimizing manual effort. The pipeline achieved an extraction accuracy of 85%. Evaluation of the results involved various approaches, including manual assessment, statistical analysis, and transformer-based models.
- **PDF Data Extraction:** Developed and deployed an AI-based automated pipeline to download, process, extract, and generate information pertaining to target funds and income funds from PDF files submitted to the Department of Labor (DoL) by both current and prospective clients. This innovative pipeline enhances overall process efficiency by 80% and significantly contributes to the calculation of Lifetime Value (LTV).

Affiliate

July 2021-Present

SETI Institute

- **Cameras for All-sky Meteor Surveillance (CAMS):** Designed to monitor the whole sky for signs of meteors, which can be traced back to their cometary origins. Implemented artificial intelligence pipeline to reduce the amount of work required of human operators. To visualize the meteors showers NASA's Meteor Observation Portal is implemented. <https://meteorshowers.seti.org>

Post-Doctoral Fellow

November 2020-March 2021

University of Houston

- **Fact Extraction and Verification:** Proposed different synthetic data generation techniques using GAN and BERT to solve the problem of fact extraction and verification. The proposed techniques are applied on a publicly available dataset, FEVER.

Research Assistant

August 2016-August 2020

The University of Southern Mississippi

- **Social Bot Analysis and Information Diffusion in Social Network:** Keen, detail-oriented research and development of skills on Social Network Analysis. Identified 3 major features for Information Diffusion and 7 distinguishable characteristics of intelligent social bots on Twitter.
- **Chat-bot and Automatic dialogue generation system:** Improved the performance of GoalOriented chatbot by 10-15% using Transfer Learning and Attention Mechanism. Usage of Deep Reinforcement Learning along with the various concepts derived from cognitive science is the goal of this research.
- **Blockchain in Swarm Robotics:** Accelerated the consensus time by 30% and reduced the falsepositive to 0% in collective decision-making problem for Swarm Robotics using Blockchain technology. www.youtube.com/watch?v=RT-HBghq_Jw
- **Detection of Diabetic Retinopathy:** Explored and identified signs of diabetic retinopathy in eye images using Hadoop, MapReduce, Mahout framework and KNN classifier in a distributed environment, to classify the retina images to different classes. The proposed model produced an improved result of 5-10 % compared to previous models.

Senior Software Engineer

LG Soft India

August 2013-June2016

- **Road direction in camera preview of android smartphone:** Pioneered a novel feature in the Camera application of LG Android mobile phone. It eases the user to navigate by displaying the real time direction on the camera application. Designed and integrated Google Map APIs and LG Camera application. Formalized the algorithm as part of the 5-member team. www.youtube.com/watch?v=loGSTczQJbo
- **Camera flash reduction:** Reengineered the algorithm to reduce the light effects in the image taken by a mobile phone's camera. Redesigned the algorithm and implemented using Poisson blending, thereby enhancing the efficiency of the algorithm by 10-15%.
- **Structured data retrieval in file manager application of android smartphone:** Incorporated a feature in the file manager application of LG Android mobile phone. Modeled and merged Google Map APIs and LG file manager application. Developed the algorithm and implemented it as part of the 5-member team.

Education

Ph.D (Computational Science), University of Southern Mississippi **2016-2020** **M.Tech, Information Technology**, National Institute of Technology Karnataka **2011-2013**
B.Tech, Information Technology, Calcutta Institute of Engineering and Management **2006-2010**

Techniques, Software & Research Interest

Programing Language C, Java, Python, R, Matlab

Research Interest Machine Learning, Deep Learning, Data Mining, Natural Language Processing (NLP), Recurrent Neural Network (RNN), Convolution Neural Network (CNN), Long Short-Term Memory (LSTM), Generative Adversarial Network (GAN), Blockchain, Swarm Robotics, Transfer Learning, Attention Mechanism, Reinforcement Learning

Current Research Areas Chatbot, Social Bot, Artificially Dialogue Generation, Social Network Analysis

Libraries Tensorflow, Keras, Scikit Learn, Pandas, Numpy

Patents

- IMAGE DISPLAY: [KR20180035045A](#)
- DISPLAY DEVICE AND CONTROLLING METHOD THEREOF: [KR20170012998A](#)
- APPARATUS FOR CONTROLLING IMAGE DISPLAY AND METHOD THEREOF: [KR20160008893A](#)

Awards

- Awarded the best paper in Natural Language Processing category at IEEE-UEMCON 2023.
- Selected among Top 1.2% of students in the Graduate Aptitude Test in Engineering (GATE)-2011 in Computer Science and Information Technology from 136027 students.
- Recipient of a full scholarship for Bachelor of technology in Calcutta Institute of Engineering and Management (2006-2010).
- Recipient of Graduate Aptitude Test (GATE) in Engineering Fellowship from Ministry of Human Resource Development, Government of India for Master of Technology (2011-2013).
- Recipient of Graduate Travel Award (Fall-2019) from the College of Arts and Sciences, The University of Southern Mississippi.

Publications

1. Building a Learning Machine Classifier with Inadequate Data for Crime Prediction, in *Journal of Advances in Information Technology Vol 8 (2), 2017*.
2. Blockchain Approach to Solve Collective Decision Making Problems for Swarm Robotics, in *Blockchain and applications: International congress, 118-125, 2020*.

3. Information Diffusion on Twitter: Pattern Recognition and Prediction of Volume, Sentiment, and Influence, in *Proceedings of the Fourth IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, 157-167, 2017.
4. Dialogue generation using self-attention generative adversarial network, in *2019 IEEE International Conference on Conversational Data & Knowledge Engineering (CDKE)*, 33-38, 2019.
5. Early detection of diabetic retinopathy from big data in Hadoop framework, in *Displays, Elsevier, Vol 70*, 102061, 2021.
6. How the tables have turned: Studying the new wave of social bots on Twitter using complex network analysis techniques, in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 501-508, 2019.
7. Cumulative Training and Transfer Learning for Multi-Robots Collision-Free Navigation Problems, in *IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 0305-0311, 2019.
8. On the Feasibility of Using GANs for Claim Verification-Experiments and Analysis, in *Proceedings of the 2021 workshop on Reducing Online Misinformation through Credible Information Retrieval*, 2021.
9. Goal-Oriented Conversational System Using Transfer Learning and Attention Mechanism, in *IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 0099-0104, 2019.
10. AI-Enhanced Data Processing and Discovery Crowd Sourcing for Meteor Shower Mapping, in *arXiv preprint arXiv:2308.02664*, 2023.
11. Claim Verification using a Multi-GAN based Model, in *Recent Advances in Natural Language Processing*, 494–503, 2021.
12. User Level Multi-feed Weighted Topic Embeddings for Studying Network Interaction in Twitter, in *Big Data–BigData 2019: 8th International Congress, Held as Part of the Services Conference Federation, SCF, Proceedings 8*, 2019.
13. Influence Modeling, Volume Prediction and Sentiment Analysis of Short Texts on Twitter, in *31st International Conference on Computer Applications in Industry and Engineering*, 2018.
14. Optimal Feature Selection from VMware ESXi 5.1 Feature Set, in *arXiv preprint arXiv:1410.5784*
15. Active Learning for the Cameras for All-Sky Meteor Surveillance (CAMS) Meteor Orbit Survey, in *LPI Contributions 2678*, 2713, Vol 2678, 2713, 2022.
16. How to Detect AI-Generated Texts?, in *IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 2023. **(Awarded the Best Paper in Natural Language Processing Category)**

Talks

1. Machine Unlearning using a Multi-GAN based Model, at ETLTC-ICETM2024: The 6th International Conference on ICT Integration in Technical Education, The University of Aizu, January 2024, Aizuwakamatsu, Japan.
2. Important Mathematical Techniques for Machine Learning, at CANTOR SECT, Department of Mathematical Sciences, The University of Memphis, Memphis, TN 3815, October 2021.
3. Important Mathematical Techniques for Machine Learning, at Department of Computer Science and Engineering & Information Technology with Institution's innovation council of Bengal Institute of Technology, Kolkata, Tech Town, Basanti Highway, Bantala (Bus-stop: 1 No. Colony), West Bengal, India, Kolkata-700150, March 2022.
4. Claim Verification using a Multi-GAN based Model, RANLP: Recent Advances in Natural Language Processing, September 2021 (Online event).
5. On the Feasibility of Using GANs for Claim Verification-Experiments and Analysis, Workshop on Reducing Online Misinformation through Credible Information Retrieval, April 2021, Tuscany, Italy (Online event).
6. Dialogue Generation Using Self-Attention Generative Adversarial Network, First IEEE International Conference on Conversational Data & Knowledge Engineering (CDKE 2019), October 2019, San Diego, USA.
7. Collective Decision Making by Swarm Robotics Using Blockchain, International Congress on Blockchain and Applications, June 2019, Avila, Spain.
8. Building a learning machine classifier with inadequate data for crime prediction, International Conference of Information and Knowledge Engineering, July 2017, Las Vegas, Nevada, USA.

Review

Journals:

1. MDPI Journal: Applied Sciences
2. MDPI Journal: Applied Math
3. MDPI Journal: Computer Sciences & Mathematics Forum
4. MDPI Journal: Electronics
5. Natural Language Processing, Cambridge University Press
6. Natural Language Processing Journal

Conference:

1. 13th annual computing and communication workshop and conference (CCWC), Las Vegas, USA.
2. International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan.

Other:

1. Member of Application Review Team for the Research Experiences for Undergraduates (REU), 2024 at the SETI Institute.