Pinkesh Badjatiya

pinkeshbadjatiya.github.io
pinkeshbadjatiya@gmail.com | pinkeshbadjatiya | pinkeshbadjatiya | +91-720-774-6433

Applied Scientist 2
Sept '21 - Current (3 years)

Working in the **Microsoft Copilot Squad @Bing, STCI** where I focus on **Building, Productionizing and Scaling Machine Learning systems for Bing and Microsoft Copilot** users while maximizing **DAU & User Experience**, with **Web-ready latency (<10ms)**.

Core Impact:

- Part of the **Copilot Squad, leading the SLM finetuning and modeling for Suggestion Chips**. Responsibilities include developing the complete life-cycle including data collection, model finetuning, model quantization & optimization and deployment at Million scale at <300ms latency. Leading the **SLM Finetuning** & Optimization efforts in the org as well.
- Proposed, Trained, Optimized & Distilled the Relevance Model, then quantized on FPGA architecture for <10ms latency. In production, ONLY Relevance Model for PeopleAlsoAsk@Bing, Shipped with **100% Bing Coverage**
- Productionized various Quality models – **FPGA Relevance Model**, **Rel-Rank Model**, **QnA Generation Model (offline)**, **Video QnA Generation**, **Realtime QnA Retrieval – aggregating to Bing DAU gains of 0.1+%**
- Shipped the ranker "**Dynamic Ranker**" – Improved **engagement by 10+%**, Provides quality data for all the LLM modelling efforts; Provides **50% of the online ranking signal**; Shipped w/ **60% Bing Coverage**
- Shipped new experience "**Video QnA**" – Created Model to generate QnA from Video content, Collaborated with IIT Madras, Shipped to production w/ **2% Bing coverage**
- Shipped new experience "**Build Your Own Prompt**" for Copilot on Bing – **Manually Reformulated Queries reduced by -0.24%, Lowest bad transfers % to Copilot, ensuring user-success, Filed a patent, Picked up by SEO Tabloids**
- Built a **real-time QnA Generation SLM Model (25ms @6 tok latency)**. Optimized model using distillation, quantization, architecture pruning, and compression.
- Helped with the backend for other new experiences like Health QnA, QnA Rewrite, Relevance v2, etc.
- **Top performer every year, Huge Impact on all KPIs – DAU, Coverage, Engagement, Experience etc.**

Team Impact:

- Team Champ for POC & New Idea Execution from 0 to 1. Performed POCs for experiences like People Also Watch, Copilot on Whatsapp, Copilot Quality for Indic Languages, etc..
- Leading a team of 2 applied scientists and 3 engineers, vested in their success.
- Additional Responsibilities include – Admin for STCI-wide GPU cluster, Regular talks as part of SLM v-Team etc..
- Internal talks – SLMs, Prompt Engineering, Modelling & Distillation, etc
- External talk & workshop – *Prompt Engineering in AI Applications* **@TakeLessons**, *World of LLMs & Prompt Engineering* **@BITS GOA**, *Prompt-a-thon L200 Workshop* **@Microsoft IDC**
- Mentored **3+ internship projects**

Past Experience

Service

External Reviewer / PC Member

- **Reviewer:** AI-ML Systems 2024, ICCV 2021, CVPR 2021, AAAI 2021, EMNLP 2018, ICWSM 2019, CODS-COMAD 2021, WOAH-EMNLP 2020, ICON 2019, ALW 2019
- **PC Member:** AAAI 2021, WOAH-EMNLP 2021, CODS-COMAD 2022, ICON 2019

Machine Learning Researcher
June '19 - Sept '21 (2 yrs 4 months)

- Researcher **@Media and Data Science Research (MDSR) Lab - a BU embedded research group in Digital Marketing** at **Adobe**
- **Few-Shot Learning** to **1-shot learning** for banner tagging for websites; (**abstract**, **human**, **non-human**) for personalizing the banner-ad on website; Obtained $\sim 94\%$ Top-1 accuracy and deployed the model to production; Project for **Samsung**.
- **Image Editing w/ GAN:** Image editing & interpolation using GANs; worked on GAN framework (extended StyleGAN-2) for generating interpolations across attributes like gender, age, smile, and eyeglasses; proposed a new interpolation strategy; collaborated with core Adobe Photoshop team to replace the existing StyleGan-2 Model; Filed a patent.
- **Major Research Projects:** Multi-modal Image Retrieval, Semantic Face Editing, Action Spaces in RL, Supervised/Unsupervised Representation Learning, GANs, Multi-modal Bundle Recommendation, Adversarial NAS, Semantic Similarity, Multi-Agent Reinforcement Learning, Policy Optimization, Logo Detector..
- Accepted **1 center-stage breakout talk**, **2 posters** and **1 tutorial** in the Adobe Tech Summit (Virtual due to Covid-19)
- Submitted **6 papers** and **10 patents** in US and other countries.
- Mentored **12+ internship projects** with **17+ student interns**.
- Built recommendation model for "**Related Conversations**" feature on https://community.adobe.com. Increased CTR by **38%**, Reduced Help Forum visits by **30%** (99% t-test confidence). Lead a team of 2 engineers + Collaborated with the respective Product Team Engineers.
- Published a tech-blog on Adobe Tech Blog http://bit.ly/magix-adobe-tech as part of Adobe India Hackathon.

Machine Learning Research Intern
Jan '19 - April '19 (4 months)

Working towards solving social dilemmas using cooperation for Multi-agent setting using Deep Reinforcement Learning using Model-based Value functions. Extended into a full-time offer.

Research Assistant
Sept '18 - Sept '19 (12 months)

Worked on **Identification of Discriminatory Content on Social Media** with Prof. Vasudeva Varma (IIIT-Hyderabad) and Prof. Manish Gupta (Principal Applied Scientist, Microsoft India R&D) at Information Retrieval & Extraction Lab (iREL), IIIT Hyderabad.

Software Engineer Intern
May '18 - July '18 (3 months)

- Designed an algorithmic solution using constraint optimization to minimize a KPI calculation based on few regulatory constraints.
- Resulted in equivalent monetary benefit of **X Billion $**. Extended into a full-time offer.
- Developed using **java** & **scala** with TDD, OOP concepts and design patterns. Performed error analysis & deployed to production.

Data Analytics/Site Reliability Engineer Intern — Sep '17 - Mar '18 (7 months)

- Developed analytics from the reports produced by ∼280 co-working startups incubated at T-Hub.
- Analyzed the user traffic flows to improve the site experience and generate weekly statistics for internal-analysis.

Teaching Assistant — Aug '16 - Dec '18 (2 years 6 months)

- Information Retrieval & Extraction *(Monsoon 2018)*
- NLP Applications *(Spring 2018)*
- Statistical Methods in Artificial Intelligence *(Monsoon 2017)*
- Data Structures *(Spring 2017)*
- ITWS-1 *(Monsoon 2016)*

Google Code-In '16 Mentor — Nov '16 - Jan '17 (3 months)

Mentored students working on the *ListenBrainz* project. Tasks involved creating tasks of varying difficulty levels and evaluating their submissions.

System Administrator and Organizer — Jul '16 - Jun '17 (11 months)

Felicity Threads is the annual technical fest of IIIT Hyderabad. I was incharge of the various servers on which different contests were hosted and maintain the AWS & email-servers, and perform tasks like creating websites, user-management, Linux instance maintenance, and maintaining docker, kafka, mysql and reddis servers.

Google Summer Of Code (GSoC) '16 Intern — Apr '16 - Aug '16 (4 months)

- Created a proxy submission **Flask** API compatible with Last.fm scrobblers.
- Added scrobbling support for desktop clients with support for tracking currently playing song in **Redis**.
- Blog: https://blog.musicbrainz.org/2016/08/23/gsoc-16-listenbrainz-fun/

Awards, Grants, Honors and Achievements

2023 | **Patent on BURP** being filed in US — Microsoft, USA

2022 | **10 patents** being filed in US and other countries — Adobe, USA

2021 | Promoted to **SDE 2** @Adobe (in 6 months); **ML Engineer 2** @Adobe (in 12 months.) — Adobe, India

2020 | **Project ANGEL | Google AI/ML Research Award 2020**
Insights from my undergraduate research work subsequently resulted in the formation of Project ANGEL at IIIT Hyderabad which later garnered the Google AI/ML Research Award 2020 to my advisor.

2019 | **TheWebConf Student travel grant** for attending TheWebConf '19 conference in US. — San Francisco, USA

2019 | **Microsoft Research (MSR) Travel Grant** for attending TheWebConf '19 conference in US. — San Francisco, USA

2018 | **Dean's Research Award** — IIIT Hyderabad
Awarded in semesters *Monsoon 17* & *Monsoon 18* for obtaining research publications during undergraduate studies.

2018 | **Dean's Merit List for academic excellence** — IIIT Hyderabad
Selected in Dean's List in semesters *Monsoon 15*, *Monsoon 17* & *Spring 18*. Awarded to Top 5% of the batch.

2018 | **ECIR-Grenoble Travel Grant** for attending ECIR'18 conference in France. — Grenoble, France

2017 | Ranked 3rd in the **Big Data Challenge**
Department of Higher & Technical Education Government of Rajasthan, India
1741 teams participated from all across India. Awarded prize by the Vasundhara Raje, Chief Minister of Rajasthan

2017 | **Best Poster Presentation Award** — WWW'17, Perth, Australia
Received award for publication *Deep Learning For Hate Speech Detection in Tweets* at the 26th International WWW Conference 2017

2016 | My team *SourceCode* ranked **80th** in the 2016 **ACM-ICPC** Asia Chennai Regional Contest, **67th** in the Asia-Chennai '16 Online Round and **107th** in the Asia Amritapuri '16 Online Round.

2014 | **AIR 3313/1.4 million (All India Rank) in JEE-Advance**

2014 | **AIR 7044/1.4 million (All India Rank) in JEE-Mains (State Rank-258)**

Notable Publications

2022 | **SAC: Semantic Attention Composition for Text-Conditioned Image Retrieval**
Surgan Jandial*, **Pinkesh Badjatiya***, Pranit Chawla*, Ayush Chopra*, Mausoom Sarkar, Balaji Krishnamurthy*
* denotes equal contribution
Accepted at WACV, 2022

2022 | **SHERLock: Self-Supervised Hierarchical Event Representation Learning**
Sumegh Roychowdhury, Sumedh A Sontakke, Mausoom Sarkar, **Pinkesh Badjatiya**, Milan Aggarwal, Nikaash Puri, Balaji Krishnamurthy, Laurent Itti
Published at International Conference on Pattern Recognition (ICPR) 2022

September 2020 | **TRACE: Transform Aggregate and Compose Visiolinguistic Representations for Image Search with Text Feedback** — *Under Review*

Surgan Jandial*, Ayush Chopra* (MIT Media Labs), **Pinkesh Badjatiya***, Pranit Chawla, Mausoom Sarkar, Balaji Krishnamurthy

August 2020 | **Status-quo policy gradient in Multi-Agent Reinforcement Learning** *Under Review*
Pinkesh Badjatiya, Mausoom Sarkar, Abhishek Sinha (Standford, USA), Siddharth Singh, Nikaash Puri, Jayakumar Subramanian, Balaji Krishnamurthy (Adobe, India)
Published at DeepRL Workshop NeurIPS 2021, Got accepted as Extended Abstract in AAMAS 2020

September 2020 | **Leveraging Style and Content features for Text Conditioned Image Retrieval** *CVPR-W 2021*
Pranit Chawla, Surgan Jandial, **Pinkesh Badjatiya**, Ayush Chopra (MIT Media Labs), Mausoom Sarkar, Balaji Krishnamurthy
Published at CVFAD - CVPR 2021 (long paper)

August 2020 | **MixBoost: Synthetic Oversampling with Boosted Mixup for Handling Extreme Imbalance** *ICDM 2020*
Anubha Kabra, Ayush Chopra, Nikaash Puri, **Pinkesh Badjatiya**, Sukriti Verma, Piyush Gupta, Balaji Krishnamurthy (Adobe, India)
Published in Internation Conference on Data Mining (ICDM) 2020 (long paper)

November 2019 | **Multi-label Categorization of Accounts of Sexism using a Neural Framework** *EMNLP 2019*
Pulkit Parikh, Harika Abburi, **Pinkesh Badjatiya**, Radhika Krishnan, Niyati Chhaya (Adobe Research, India), Manish Gupta (Microsoft, India), Vasudeva Varma
Published in Empirical Methods in Natural Language Processing (EMNLP), 2019 (long paper)

October 2018 | **Stereotypical bias removal using Knowledge-based Generalization for Abuse Detection Task**
Pinkesh Badjatiya, Manish Gupta (Microsoft, India), Vasudeva Varma (IIIT-H)
Published in The World Wide Web Conference (TheWebConf) 2019 (long paper)

December 2017 | **Attention-based Neural Text Segmentation** *ECIR 2018*
Pinkesh Badjatiya, Litton J Kurisinkel, Manish Gupta (Microsoft, India), Vasudeva Varma (IIIT-H)
In proceedings of European Conference on Information Retrieval (ECIR), 2018 (long paper)

April 2017 | **Deep Learning for Hate Speech Detection in Tweets** *WWW 2017*
Pinkesh Badjatiya*, Shashank Gupta*, Manish Gupta (Microsoft, India), Vasudeva Varma (IIIT-H)
Publishd in International World Wide Web (WWW) Conference, 2017
- Received **Best Poster Presentation** Award at **WWW'17, Perth, Australia**
- Our work received significant print media coverage: https://bit.ly/hatespeech-iiith-best-poster

For complete list of publications, and patents, please visit pinkeshbadjatiya.github.io

Education

2018 - 2019 | MS by RESEARCH in Computer Science and Engineering
International Institute of Information Technology, Hyderabad (IIIT Hyderabad)
- **Specialization:** Information Retrieval, Natural Language Processing, Machine Learning, Deep Learning
- Best Poster Award (WWW'17), ECIR Student Travel Grant, Microsoft Research Travel Grant, WWW Student Travel Grant, Dean Research Award
- **Advisor:** Dr. Vasudeva Varma (IIIT Hyderabad); Dr. Manish Gupta (Microsoft India R&D)

2014 - 2018 | B.Tech in Computer Science and Engineering
International Institute of Information Technology, Hyderabad (IIIT Hyderabad)
Overall CGPA: 8.71 / 10

Technical Skills

- Proficiency: **Advanced**, Intermediate, Basic

Programming Languages	**Python**, **C++ (STL)**, **C**, java, Scala
OS	**GNU/Linux** ♡, **Microsoft Windows**
Scripting	**Python**, **Bash**, **Scala**
Web Technologies	**HTML**, **CSS**, **Flask**, JavaScript, Web2py, **Django**
Machine/Deep Learning	**TensorFlow**, **PyTorch**, **numpy**, **sklearn**, **pandas**, theano
Others	**MySQL**, **Git(VCS)**, **Markdown**, **apache2**, **nginx**, LaTeX, AWS, Redis, Cassandra, Docker, POSIX, pthreads, sockets, Grafana
Interests	Machine Learning, Natural Language Processing, Deep Learning, Reinforcement Learning, Algorithms, Operating Systems, Disatributed Systems

* denotes equal contribution