Pinkesh Badjatiya

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

Machine Learning Researcher & Engineer 2 — June '19 - Current

Currently working with the **Media and Data Science Research (MDSR) Lab in Adobe** as a *AI Researcher* where I primarily work with NLP and Computer Vision research. I also work on productionizing ML algorithms.

- Created "**Related Conversations**" feature on https://community.adobe.com that provides recommendation to thousands of users everyday. Increased interaction CTR by **XX%** and reduced Jarvis help-ticket's CTR by **XX%**.
- Presented **1 center-stage breakout talk**, **2 posters** and **1 tutorial** in the Adobe Tech Summit.
- Submitted **4 papers** and **3 patents** in US and other countries.
- Published a blog on Adobe Tech Blog: http://bit.ly/magix-adobe-tech

Awards, Grants, Honors and Achievements

2021 | **3 patents** being filed in US and other countries — Adobe

2021 | Promoted to **ML Engineer 2** at Adobe (in 12 months) — Adobe, India

2020 | Promoted to **SDE-2** at Adobe (in 6 months) — Adobe, India

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

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
26th International World Wide Web Conference, 2017
Received award for publication *Deep Learning For Hate Speech Detection in Tweets* at the WWW'17 Conference, Perth, Australia.

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.

2015 | Ranked 2nd in *Microsoft Code.Fun.Do* Hackathon — Microsoft, Hyderabad

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)**

Past Experience

Service — **External Reviewer / PC Chair**
- **Reviewer:** ICCV 2021, CVPR 2021, AAAI 2021, CODS-COMAD 2021, WOAH-EMNLP 2020, ICWSM 2019, ICON 2019, ALW 2019, EMNLP 2018
- **PC Chair:** CODS-COMAD 2022, WOAH-EMNLP 2021, AAAI 2021, ICON 2019

Machine Learning Research Intern — Jan '19 - April '19
Working towards solving social dilemmas using cooperation for Multi-agent setting using Deep Reinforcement Learning using Model-based Value functions.

Research Assistant — Sept '18 - Sept '19
Currently working 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
Designed an algorithmic solution using constraint optimization in **java** and **scala**. Used TDD with OOP concepts and Design patterns. Performed error analysis and deployed to production.

Data Analytics/Site Reliability Engineer Intern — Sep '17 - Mar '18
- 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 using **Google Analytics** and **Python**

Teaching Assistant — Aug '16 - Dec '18

- 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

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.

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/

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

PUBLICATIONS

APRIL 2021 — **Combating Online Hate Speech Roles of Content, Networks, Psychology, User Behavior and Others**
Sarah Masud, **Pinkesh Badjatiya**, Dr. Amitava Das, Dr. Manish Gupta, Dr. Vasudeva Varma, Dr. Tanmoy Chakraborty
Tutorial Accepted at ECML/PKDD 2021

SEPTEMBER 2020 — **TRACE: Transform Aggregate and Compose Visiolinguistic Representations for Image Search with Text Feedback** *Under Review*
Surgan Jandial*, Ayush Chopra*, **Pinkesh Badjatiya***, Pranit Chawla, Mausoom Sarkar, Balaji Krishnamurthy

AUGUST 2020 — **Inducing Cooperative behaviour in Sequential-Social dilemmas through Multi-Agent Reinforcement Learning using Status-Quo Loss** *Under Review*
Pinkesh Badjatiya, Mausoom Sarkar, Abhishek Sinha, Siddharth Singh, Nikaash Puri, Jayakumar Subramanian, Balaji Krishnamurthy (Adobe, India)
Under review in conference. 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, 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

* denotes equal contribution

Major Projects

Wiki Search Engine	**Search Engine for indexing & querying entire Wikipedia dump** • Built efficient positional index for 64GB of Wikipedia dump for supporting multi-word queries. • Field queries with execution time **<5 secs**. Ranking results based on relevance, field and term scores.
d.TF-IDF	**Distributed TF-IDF for documents** Developed a master-slave distributed architecture using **C++** using **MPI interface** for computing TF-IDF scores for a set of documents to equally distribute the load across slaves.
Tic-Tac-Toe AI	**A game bot for modified version of tic-tac-toe game** • Developed a game bot for Ultimate Tic-Toe-Game using **Minimax Algorithm** & **Alpha-Beta pruning**.
PeerNet	**P2P file sync over TCP/UDP protocols with integrity checks** • A peer-to-peer file sharing application capable of sharing files using TCP/UDP protocols along with integrity check using SHA256 checksum. Supports remote directory listing along with complete directory sync. • Built in C using socket programming and UNIX system calls.
Nex	**Mini multi-threaded proxy web server + HTTP web server** • Capable of serving html/images/audio/videos and directory listing or used as a proxy server. • Other features: Template injection, banned hosts, conditional header processing, HTTP Basic Authentication.
Xv6 Priority Scheduler	**A scheduler for xv6 operating system based on priority of processes** • Created a priority based scheduler in place of round robin scheduler. • Added a new system call **set_priority()** to change the priority of processes.
C Shell	**A shell built in c using POSIX standards and thread programming** • A shell in C using knowledge of threads, forking, signals, process groups, system calls and other OS concepts. • Features include piping, file redirection, process job management, foreground/background processes, tracking background processes, signal handling, variable assigning, and built-in commands.

Technical Skills

•	Proficiency: **Advanced**, Intermediate, Basic
Programming Languages	**Python**, **C++ (STL)**, **C**, java, Scala
OS	**GNU/Linux** ♡, **Microsoft Windows**
Scripting	**Python**, **Bash**
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, POSIX, pthreads, sockets, OpenGL
Interests	Machine Learning, Natural Language Processing, Deep Learning, Reinforcement Learning, Algorithms, Operating Systems