




# DEVANSH BATRA

 [github.com/devanshbatra04](https://github.com/devanshbatra04)  [devanshbatra.com](https://devanshbatra.com)  [devanshbatra51@gmail.com](mailto:devanshbatra51@gmail.com)

## EDUCATION

---

**Netaji Subhas Institute of Technology, New Delhi**  
*Bachelor in Engineering, Information Technology*

*Aug 2017 - May 2021*  
**CGPA – 8.64**

## TECHNICAL SKILLS

---

**Broad Interests:** Artificial Intelligence, Natural Language Processing, Machine Learning  
**Programming Languages:** C++, Python, Javascript, Haskell, C, Ruby  
**Relevant Frameworks:** **Full Stack Development:** ReactJS, NodeJS, Flask, Django  
**Deep Learning:** Tensorflow, Keras, Pytorch

## WORK EXPERIENCE

---

**Member Technical, D.E. Shaw India** *May 2021 - Present*  
- Primarily responsible for generating proprietary synthetic datasets for quantitative research using market data.

**Research Intern, MIDAS-IIIT Delhi** *May 2020 - Present*  
- Developed LiFi, an SoTA video frame interpolation model with linguistic information beating for speech videos  
- Developed first unsupervised approach for lip reading using adversarial learning of viseme probability distribution  
- Developing cross-lingual ASR for indigenous languages, accounting for phonetic similarities between languages

**Research Intern, Nanyang Technological University** *May 2020 - Dec 2020*  
- Created Inverse Text Normalization algorithms to improve Speech Recognition during air traffic control.  
- Wrote Finite State Transducers to develop millions of variations of colloquial ways to refer to addresses in English.  
- Created a realtime GUI with support for the Inverse Text Normalization Module along with Named Entity tagging.

**Summer Software Engineer Internship, Sprinklr** *May 2020 - Jun 2020*  
- Developed distributed, scalable and extensible microservice based architecture for monitoring live servers.  
- Added config-driven Slack Integration to quickly add new param based devops services with realtime reports.

**Google Summer of Code, OpenCV** *May 2020 - Aug 2020*  
- Created the Point Cloud (PtCloud) module in OpenCV allowing basic Point Cloud manipulation with OpenCV.  
- Implemented sample consensus based object fitting and segmentation methods using RANSAC and LMeDS.

**Research Intern, Complex Systems Lab, IIIT Delhi** *May 2019 - July 2019*  
- Summer research intern at Complex Systems Lab under the mentorship of Dr. Ganesh Bagler at IIIT-D.  
- Projects drove original efforts in the field of Computational Gastronomy pioneered by Dr. Ganesh Bagler.  
- Created RecipeDB - the single largest peer reviewed repertoire of 148k cooking recipes with nutritional content.  
- Integrated RecipeDB with FlavorDB and DietRx, thus adding flavor molecules and disease associations to all recipes.  
- Developed SOTA language models for novel recipe generation, beating the SOTA by 33% in terms of perplexity.

**Javascript and Machine Learning Intern, Taiyo** *July 2018 - June 2020*  
- Created modular visualization libraries, along with UX design of visualisation client for economic and finance data.  
- Integrated Elasticsearch with Grafana widgets for analysis from several machine learning models developed.  
- Created statistical Machine Learning models for prediction of cryptocurrency rates and for consumer sentiments.

**Software and AI Department Lead, ARES Robotics** *July 2018 - July 2020*  
- Incorporated Visual Inertial Odometry, fused GPS localization, and Path Planning in the ATOM Mars Rover.  
- Top 30 teams worldwide in European Rover Challenge 2018 and 2019, 2nd at Indian Rover Design Competition.

## PROJECTS IN PRODUCTION

---

### **Watchdog - Server Log Monitoring and Malware Analysis Dashboard**

**Tech Stack:** React, Python, Shell, Elasticsearch

- Cross platform app for detecting and eliminate multiple malicious remote connections in Windows and Unix servers.
- Real Time IP address scan and SAFE scores for intuitive remote connection health metric.
- Tunneled websocket traffic as TCP packets into the linux socket using websocketify for robust real time monitoring
- Incorporated all Sentinel rootkit detection and offline IP localization and ownership features.
- **Won complex problem statement category** at Smart India Hackathon 2020 under BEML Ltd.
- Assisted in onsite deployment at BEML India's datacenters.

### **Sentinel - Remote Connection And Rootkit Scanner**

**Tech Stack:** QT Framework, Python, Shell, Several Python Libraries

- Cross platform application for detecting and eliminate multiple rootkits from Windows and UNIX systems.
- Incorporated intuitive GUI features as well as outsourced databases for the purpose of identifying IP addresses.
- Included system binary signatures for 79 UNIX rootkits gathered through virustotal archives.
- Included utility functions to deal with TCP remote connections like offline location scanning, manipulating iptables
- Real-time checking of processes for malware using multithreaded hashing of suspicious files for VirusTotal matches
- **Won complex problem statement category** at Smart India Hackathon 2019 under Ministry of External Affairs.
- Software aquired and used by Ministry of External Affairs across their data center.

### **Atom - Mars Rover Prototype**

**Tech Stack:** ROS, Gazebo, Bumblebee, Jetson Nano, Arduino, OpenCV

- Wrote scripts in ROS for Science Cache task to autonomously control the 5 Degree of Freedom Robotic Arm.
- Implemented ORB SLAM, GPS assisted Localization and Obstacle avoidance.
- The autonomous systems received particularly high score with jury commendation at University Rover Championship
- Ranked among the 30 top teams at European Rover Championship 2019, to represent India at the same.
- Demo available on YouTube - [www.youtube.com/watch?v=uGqkPknVmwI](http://www.youtube.com/watch?v=uGqkPknVmwI)

### **RecipeDB - A resource for exploring recipes, their culinary, nutritional, flavor and health correlates**

**Tech Stack:** Spacy, Tensorflow, Flask, SQL, Javascript, HTML, CSS

- Data curation and analysis of recipes, flavor molecules and disease associations and web app development.
- Used NLP Techniques to extract and find relationships between various characteristics of recipe ingredients.
- Developed models to infer nutritional information about, taking into account ingredients as well as cooking processes.
- Developed a method to find similarity between recipes based on the above exercises
- Developed entire backend and GUI. Deployed the same on institute's servers
- Available openly at <https://cosylab.iiitd.edu.in/recipeadb/>

## RESEARCH PROJECTS

---

### **RecipeDB and Allied Projects, 2019 to 2020**

- RecipeDB is a structured compilation of recipes, ingredients, and nutrition profiles.
- Worked on data compilation and cleaning, followed by nutritional estimation from parsed natural language text.
- Followup work included SoTA method for generating novel recipes.
- Publications at ICDE Workshops, COLING as oral paper, and OUP Database journal.

### **Visual Speech related projects, 2020 to present**

- Developed datasets and error metrics for fair evaluation of algorithms for generation of frames in speech videos.
- Evaluated previous SOTA models and presented novel ones on the aforementioned datasets and metrics.
- Published at ICASSP 2021.
- Currently working on unsupervised visual speech recognition, initial results submitted to AAAI 2022.

## ACHIEVEMENTS

---

- June 2017 - 99.5th percentile in JEE Advanced, with AIR 6575
- 2018 - 2019 - Won Multiple Hackathons, including Smart India Hackathon and MS Hacks (at Microsoft Gurgaon)
- Aug 2020 - Won Smart India Hackathon for the second time, research paper accepted at Oxford's Database
- Aug 2020 - Led ARES Robotics to 2nd place finish at Indian Rover Design Competition among 28 teams, 8 countries.

## PUBLICATIONS

---

- Yash Agarwal\*, **Devansh Batra\***, and Ganesh Bagler. “Building Hierarchically Disentangled Language Models for Text Generation with Named Entities” In 2020 28th International Conference on Computational Linguistics (COLING)
- **Batra**, et al. “RecipeDB: A Resource for Exploring Recipes.” Oxford University Publishing’s Database Journal (2020) (**I.F. 3.86**)
- Aradhya Neeraj Mathur\*, **Devansh Batra\***, Yaman Kumar\*, Rajiv Ratn Shah, Roger Zimmermann, and Amanda Stent. “LIFI: Towards Linguistically Informed Frame Interpolation.” ICASSP 2021.
- Jushaan Kalra\*, **Devansh Batra\***, Nirav Diwan, and Ganesh Bagler. “Nutritional Profile Estimation in Cooking Recipes.” In 2020 IEEE 36th International Conference on Data Engineering Workshops (ICDEW)
- Nirav Diwan\*, **Devansh Batra\***, and Ganesh Bagler. “A Named Entity Based Approach to Model Recipes.” In 2020 IEEE 36th International Conference on Data Engineering Workshops (ICDEW)