
EDUCATION

- **Indraprastha Institute of Information Technology** New Delhi, India
Bachelor of Technology, Computer Science and Engineering; CGPA: 8.88 *Aug. 2016 – May 2020*
- **Sanskriti School** New Delhi, India
High School; Percentage: 93.8% *2014 – 2016*

PUBLICATIONS

- Aditya Chetan*, **Brihi Joshi***, Hridoy Sankar Dutta, Tanmoy Chakraborty. CoReRank: Ranking to Detect Users Involved in Blackmarket-based Collusive Retweeting Activities. In *The 12th ACM International Conference on Web Search and Data Mining (WSDM 2019)*. (Acceptance Rate: 16%, CORE2018 A*)
- Hridoy Sankar Dutta, Aditya Chetan*, **Brihi Joshi***, Tanmoy Chakraborty. Retweet Us, We Will Retweet You: Spotting Collusive Retweeters Involved in Blackmarket Services. In *The 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2018)*. (Acceptance Rate: 15%) arXiv:1806.08979 [cs.SI]
- Nishtha Madaan, Gautam Singh, Sameep Mehta, Aditya Chetan*, **Brihi Joshi***. Generating Clues for Gender based Occupation De-biasing in Text arXiv:1804.03839 [cs.CL]

* Equal Contribution

SKILLS

- **Relevant Courses:** Machine Learning, Natural Language Processing, Linear Algebra, Data Mining, Probability and Statistics, Real Analysis, Analysis and Design of Algorithms, Discrete Mathematics, ODEs, Digital Audio
- **Languages:** Python, Java, C, Bash, JavaScript
- **Tools & Technologies:** Git, MATLAB, scikit-learn, nltk, spaCy, numpy

RESEARCH EXPERIENCE

- **Laboratory for Computational Social Systems** New Delhi, India
Undergraduate Researcher *Jan 2018 - Present*
 - **Understanding adversarial collusive activities in OSNs:**
Advised by - Dr. Tanmoy Chakraborty
 - * Worked on detecting collusive retweeters on Twitter, focussing on users of freemium services
 - * Curated an open dataset of manually annotated users from various freemium services
 - * Proposed a novel set of features for detection and tested its effectiveness on several machine learning models
 - * Developed an unsupervised and semi-supervised approach for detection of collusive retweeters
 - Keywords:** *Machine Learning, OSNs, Data Science*
- **IBM India Research Laboratory** New Delhi, India
Undergraduate Researcher *Nov 2017 - Present*
 - **Complex Answer Retrieval:**
Advised by: Dr. Sumit Bhatia
 - * Working on the problem of [Complex Answer Retrieval](#).
 - * Did a literature survey on existing approaches to artificial Q&A.
 - * Currently exploring meticulous IR approaches to help improve accuracies of downstream DL models in Q&A.
 - * **Keywords:** *NLP, Complex Answer Retrieval, IR, Deep Learning*
 - **Persuasive Style Transfer:**
Advised by - Dr. Deepak Padmanabhan (QU, Belfast)
 - * Detecting and removing “persuasiveness” quality in political speeches.

- * Modelled the detection and removal of “*persuasive*” quality in text as an optimisation problem.

Keywords: *NLP, Machine Learning, Deep Learning*

- **Occupational Debiasing [code][preprint]:**

Advised by - Dr. Sameep Mehta

- * Developed a system for detecting Occupational Gender Bias in text, taking into account the influence of demography and time-frame.
- * Developed and curated a dataset comprising of various occupation names and occupational evidences.
- * Proposed a pipeline that detects potential gender bias in occupations.
- * **Poster accepted at GHCI 2018**

Keywords: *NLP, Machine Learning, Flask*

WORK EXPERIENCE

- **dunnhumby** New Delhi, India
Associate Analyst Intern *May 2017 - July 2017*
 - **Worked with the North America Data Team::**
 - * Studied the various aspects and applications of Data Science in the Business Dimension.
 - * Worked on creating a large-scale implementation demo of MLlib in Apache Spark to predict user purchase patterns for various clients of dunnhumby focusing on those based in North America
 - * Studied the Automation of ETL processes using Apache Airflow on Google Cloud Platform and worked in transferring data from SAS Archives to Hadoop Datalake.
 - * **Keywords:** *Data Analytics, Apache Spark, Hadoop, Business Intelligence*
- **Rails Girls Summer of Code** New Delhi, India
Student Scholar *May 2017 - July 2017*
 - **Worked on Tessel - an IoT and Robotics development platform:**
 - * Part of one of the selected 16 sponsored teams out of 192 teams worldwide.
 - * Developed tutorials - Prepared Documentation, code snippets, Fritzing diagrams and a working demonstration.
 - * Developed a model for a Humanoid Arm Project - Understood the idea behind product design and development.
 - * Studied and worked on Reach - an ESP32 (low power BLE module), that is supposed to be Tessel’s new launch in the domain of Low power boards.
 - * **Keywords:** *Open Source, IoT, Project Design*

AWARDS

- **Microsoft Research India Travel Grant:** Awarded travel support of 50000 INR for visiting WSDM 2019
- **ACM-W Scholarship:** Awarded travel support of 1200 USD for visiting WSDM 2019
- **Google Women Techmakers Scholarship, 2018:** Awarded to students who work for diversity and inclusion in the field of Computer Science.
- **Best Technical Poster Runner-up at GHCI 2018:** Received for the project, “Generating Clues for Gender based Occupation De-biasing in Text”
- **Dean’s Award for Innovation R&D:** Awarded to students who work on Research projects beyond coursework. Awarded for the academic years 2016-17 and 2017-18.
- **Dean’s List of Academic Affairs:** Awarded to students who demonstrate excellence in an academic year. Awarded for the academic year 2016-17.
- **Grace Hoppers Celebration India (GHCI) Scholarship, 2018:** Awarded travel grant and scholarship to attend the GHCI conference.

PROJECTS

- **img2L^AT_EX[demo]:**
 - * An end-to-end model for converting handwritten mathematical expressions to compilable L^AT_EX
 - * Made use of image segmentation, supervise classifiers such as CNNs, SVM, etc. and heuristics for code formation
 - * Developed as a course project for Machine Learning (CSE343) at IIT Delhi in Winter 2018
 - * **Keywords:** *Machine Learning, Deep Learning, Image Processing*

- **SemEval19 Task 3 : EmoContext:**
 - * Worked on the task of contextual emotion detection in text. This was a shared task for a workshop at ACL 2019.
 - * Used deep learning models like LSTMs, DeepMoji and ELMo for trying to increase accuracy.
 - * **Keywords:** *NLP, Contextual emotion detection, Deep Learning*
- **iDabba** [\[code\]](#)[\[poster\]](#):
 - * This project aims to improve the food storage and large scale handling of food silos in granaies
 - * Uses Computer Vision techniques with SIFT, Haar classifier and Microsoft Vision API in Python and checks the environmental conditions around it using Humidity, Temperature and Weight Sensors.
 - * Was amongst the top 10 projects in the first-year batch and received a mention in the Director's blog [\[link\]](#).
 - * **Keywords:** *Computer Vision, IoT, Flask*
- **Chain Reaction** [\[code\]](#):
 - * A JavaFX-based MacOS application based on the famous Android-based game
 - * Includes features like undo, resume and multiplayer mode
 - * **Keywords:** *JavaFX, GUI, MacOS, Java, Object-oriented programming*

CO-CURRICULAR ACTIVITIES

- **Talks**
 - * *Intro to Computer Vision:* Introductory talk on the self-taught topic, demonstrating basics of Images, colour, processing and basic CV algorithms. Delivered at WWCode Delhi. [\[slides\]](#)
 - * *Intro to Mathematics of ML:* Introductory talk on the self-taught topic, demonstrating basic ML algorithms from a mathematical point of view. Delivered at WWCode Delhi. [\[GitHub\]](#)
 - * *Intro to Web connectivity with Tessel:* An introductory talk about Tessel development board and its features. Delivered at LinuxChix India. [\[coverage\]](#)
- **Volunteering**
 - * *Summer School, IIITD:* Summer School is where middle-school children from government schools are taught basic life skills and beyond the subjects taught in school. I was a part of the Personality Development and Communication Skills team, where I taught students the importance of physical gestures like posture, speaking and listening skills and basics of english. I was also involved in the cultural development of students in Music and Dance. [\[coverage\]](#)