Amey Agrawal
agrawalamey.github.io

Email: agrawalamey12@gmail.com
Github: /AgrawalAmey

Education

- **Birla Institute of Technology and Science** — Pilani, India
  Bachelor of Engineering in Computer Science; GPA: 8.13 / 10.0 — *Aug. 2014 – July. 2018*

Publications

- **Learning Digital Circuits: A Journey Through Weight Invariant Self-Pruning Neural Networks**
  ***Amey Agrawal**, and Rohit Karlupiya*
  Proceedings of New in ML Workshop, NeurIPS, 2019, Vancouver [paper]
- **Delog: A Privacy Preserving Log Filtering Framework for Online Compute Platforms**
  ***Amey Agrawal**, Abhishek Dixit, Namrata Shettar, Darshil Kapadia, Rohit Karlupiya, Vikram Agrawal, and Rajat Gupta*
  Proceedings of IEEE International Conference on Big Data, 2019, Los Angeles [paper]
- **Logan: A Distributed Online Log Parser**
  ***Amey Agrawal**, Rajat Gupta, and Rohit Karlupiya*
  Proceedings of IEEE International Conference on Data Engineering (ICDE), 2019, Macau [paper]

Experience

- **Qubole** — Bangalore, India
  Member of Technical Staff — *Jul 2018 - Present*
  - **Logan**: Developed a distributed log parsing algorithm that provides **39x** speed-up over the previous state-of-the-art. Integrated the parser with Qubole Spark for real-time anomalous log detection. Received **Spotlight** award for this effort.
  - **RStudio on QDS**: Contributed across Qubole's stack to integrate RStudio as a first-class citizen. Made open-source contributions to **Sparklyr** and **S3FS**.
  - **Zendesk Ticket Similarity Service**: Built an unsupervised learning model to identify similar Zendesk tickets.
- **BITS Pilani** — Pilani, India
  Teaching and Research Assistant — *Dec 2017 - May 2018*
  - **Research Assistant - Prof. Kamlesh Tiwari**: Developed an autoencoder-like architecture to perform controlled geometric transformations on images in order to learn translation invariant features for Iris images.
  - **Research Assistant - Prof. Surekha Bhanot**: Evaluated feasibility of reinforcement learning based meta-optimizer for learning rate control with quantized action space.
  - **Teaching Assistant - Neural Networks & Fuzzy Logic**: Developed **Callisto**, an Electron app which provides zero-setup Python environment for conducting evaluative assignments with Jupyter notebooks.
  - **Teaching Assistant - Machine Learning**: Designed and evaluated programming assignments in Python. Held boot-camp sessions on scientific Python ecosystem.
- **Qubole** — Bangalore, India
  Software Enginnering Intern — *Jul 2017 - Dec 2017*
  - **Deep Learning Cluster**: Developed beta version of Qubole's enterprise distributed deep learning platform. Received **Spotlight** award for cross-team collaboration.
- **Norah.ai** — Bangalore, India
  Machine Learning Research Intern — *May 2017 - Jun 2017*
  - **Text to Humanoid Animation**: Developed a sequence to sequence model for text to animation conversion using Keras.
- **Bharat Heavy Electrical Limited (BHEL)** — Trichy, India
  Software Engineering Intern — *May 2016 - Jul 2016*
  - **Sitemap-Draw**: Created a web-crawler and sitemap visualization tool in NodeJs.
  - **Log Warehouse**: Developed log storage and management tool using MongoDB and Python.

Selected Projects

- **Callisto** *Jan 2018*
  Guide: Prof. Surekha Bhanot [Blog] [Demo] [GitHub]
  - Developed cross-platform desktop application built with Electron and Express for neural networks course.
  - Automatically sets up uniform scientific python development environment independent of operating system.
  - Allows students to download and launch assignment Jupyter notebooks with a single click.
  - Evaluation mode enables programming contests using nbgrader.

- **Disentanglement learning for iris image indexing** *Jan 2018 - May 2018*
  Guide: Prof. Kamlesh Tiwari [Blog] [Google Colab]
  - While working at the AI Lab in BITS Pilani, designed an autoencoder architecture to learn horizontal translation-invariant representations of normalized iris images.
  - Established proof of concept on MNIST, Fashion-MNIST and CIFAR10 datasets.

- **Deep Q-learning for autonomous warehouse robots** *Jan 2017 - Apr 2017*
  Guide: Prof. Surekha Bhanot [GitHub]
  - Implemented a Deep Q-learning algorithm to make warehouse robots which can learn to navigate autonomously.
  - Developed 2D simulations using pybox2D and 3D simulations using V-rep.
  - Implemented the neural network using keras.

- **CNN visualization toolkit** *Apr 2017*
  Guide: Prof. Sundaresan Raman [GitHub]
  - Integrated a collection of popular CNN visualization techniques into a single framework which can take any Keras CNN model as input.
  - Developed backend web server using Flask.

- **NIRF Ranking Analysis** *May 2018*
  Open Source Initiative [Blog] [GitHub]
  - Performed extensive analysis of ranking metrics and proposed an alternative metric to measure research outcome of universities.
  - Scrapped data used to calculate rankings of Indian engineering schools from National Institutional Ranking Framework's (NIRF) website.
  - Parsed PDFs by creating a finite state machine using TextFSM.

- **Automated news-in-shorts** *Nov 2016*
  Guide: Prof. Poonam Goyal [GitHub]
  - Latest posts from RSS feeds of multiple news agencies are clustered using K-Means with TF-IDF vectorization.
  - Trending topics on twitter are mapped to clusters of news articles to identify trending news.
  - All the articles on a given trending topic are summarized using extractive text summarization using sumy.
  - Developed web clients and REST API using MongoDB and Express.
  - Implemented text processing pipeline using NLTK, Scikit-learn and Gensim.

- **Pokemon MMORPG** *Mar 2017 - Apr 2017*
  Guide: Prof. Rahul Banerjee [GitHub]
  - Designed and developed fully distributed multi-player online game entirely in C.
  - Used Redis as message broker and MongoDB as database.

- **Arxiv-Sanity v2.0** *Feb 2017*
  Open Source Initiative [GitHub]
  - Revamped Andrej Karpathy's Arxiv-Sanity for improved functionality and a better UI.
  - Built upon the existing flask web server and migrated database to MongoDB for scalability.

- **Real-time classification of network traffic** *Mar 2017 - Apr 2017*
  Guide: Prof. Amit Dua
  - Developed random forest and multilayer perceptron models to perform real-time classification of network traffic.
  - Implemented models using Scikit-learn and Keras and captured live packets using PyShark.

- **Predicting Election Results using Twitter** *Apr 2015*
  Guide: Prof. Vandana Agrawal
  - Built a CNN model with word embeddings to perform sentiment analysis on tweets pertaining to US presidential elections.
  - Implemented CNNs from the ground up in vanilla python.

Technical Skills

- **Languages**: Proficient in Python, Scala, JavaScript, Bash, Familiar with Java, Ruby, R, C/C++
- **Data Science**: PyTorch, Keras, Spark, TensorFlow, Scikit-learn
- **Web & UI**: Express, Flask, Jetty, Ruby on Rails, Electron, JQuery
- **Databases**: MySQL, Redis, Hive, MongoDB
- **Misc**: Git, Markdown, LaTeX, Jenkins, CorelDraw

Selected Coursework

- **Mathematics**: Multivariable Calculus, Linear Algebra, Probability & Statistics, Differential Equations
- **Software Engineering**: Object Oriented Programming, Computer Networks
- **Systems**: Database Management Systems, Database Systems
- **Data Science**: Machine Learning, Information Retrieval, Neural Networks and Fuzzy Logic, Cognitive Computing

Extracurricular Activities

- Designed and developed websites for BITS-Pilani's cultural and technical festivals while working at students' union's department of visual media.
- Worked with English Press Club at BITS-Pilani to design vector graphics and illustrations for institute's annual students' magazine Cactus Flower.
- Organized Concepticon, an event for awareness about career options in STEM. The event saw participation from more than one thousand high school students.