Arjun R. Rao

(+1) 720-323-8946 | arjun.rao@colorado.edu | arjun.fyi | arjun-rao | arjunra0

Education

University of Colorado, Boulder — *Boulder, US*

M.S. in Computer Science — *Aug. 2019 - May 2021*

- **Graduate Student Staff** for Intro to AI (CSCI3202)

Ramaiah Institute of Technology (RIT) — *Bangalore, India*

B.E. in Information Science and Engineering — *Aug. 2014 - June 2018*

- CGPA: **9.41/10**
- **Best Outgoing Student** (Batch of 2014-2018), Dept. of Information Science & Engineering

Industry Experience

Stride.ai Inc. (A TechStars Company) — *Bangalore, India*

NLP Engineer — *July 2018 - May 2019*

- **ACE**:
  - Worked on building Stride's flagship product, ACE, which supports document classification and information extraction using limited labelled data. I was responsible for the entire UI/UX design of the project in-addition to contributing to Backend APIs.
  - Built a **custom python library** that allows users to **train, evaluate ML models and perform inference** for document classification tasks in ACE using Tensorflow and Keras.
- **Backend and APIs**:
  - Designed database schemas and implemented RESTful APIs using Django and techniques like request caching, polling, and message queues (Redis-Queues).
  - Worked on defining API contracts between the Backend and Frontend teams and built client libraries for abstracting API calls.
- **Frontend Libraries and Frameworks**:
  - Standardized the design language and theme for all Stride products
  - Built Stride's **core frontend framework** with reusable components and custom styles using **Angular**.
  - Built a custom PDF Viewer library with Angular support using PDF.js that allowed user annotation functions, support for multiple monitors, intelligent document comparison, and cross-browser support. The viewer is being used in 10+ Projects and efficiently handles PDFs with over 500 pages.
- Technologies Used: **Python**, **Django**, **Typescript**, **Angular**, **Keras**, **Tensorflow**, D3.js, PDF.js, Internal tools

Research and Development Intern — *Oct. 2016 - Dec. 2016, Jan. 2018 - Apr. 2018*

- Worked on **optimizing Named Entity Recognition (NER) models** to improve performance by combining standard NER datasets with context specific data. This work was later extended and converted into an internal library for building custom NER models using augmented datasets.
- Built **scrapers** for automatically extracting information from multiple websites using Selenium and BeautifulSoup.
- Technologies Used: Python, Tensorflow, Keras, Gensim, Selenium

Google Inc. — *San Francisco, US*

Developer Programs Engineer Intern, Cloud Developer Relations — *June 2017 - Aug. 2017*

- Was an intern on the Cloud DPE team, and worked on 2 projects as an individual contributor responsible for both design and implementation.
- **GitHub Issue Tracker**:
  - Developed a web app that allows users to configure public GitHub repositories to receive **consolidated email notifications** for events like new issues, comments, etc.
  - The tool greatly reduced the number of email notifications received in a day from GitHub by summarizing all activity across configured repositories into a single digest email.
  - Link to open source release: github.com/GoogleCloudPlatform/issuetracker
- **GitHub Issues Mirror**:
  - This was a project to mirror issues and comments on configured GitHub repositories into the internal bug tracker to allow Googlers to triage and track issues in a centralised place.
  - I used internal libraries, and **App Engine** to coordinate cron jobs to synchronize internal bug entries with GitHub data in **BigQuery**.
- Technologies Used: **Go**, **Angular**, Google Cloud Datastore, Google BigQuery, Google App Engine

Skills

Programming Languages	**Python**, **TypeScript**, **Go**, **Java**, C/C++, PHP
Platforms & Frameworks	**Django**, **Tensorflow**, **Angular**, **Android**, **Google Cloud Platform**, Express
Languages	**English** (Native Language), Hindi, Kannada

Academic Experience

Ramaiah Institute of Technology

Bangalore, India

Senior Project - The Milo IDE (miloide.github.io) *Sept. 2017 - April 2018*

- Built a web-based IDE to help students with no prior programming experience learn Machine Learning and Linear Algebra.
- Customized **Google's blockly project**, and designed a visual programming language that supports data science operations.
- Implemented a data explorer with built-in datasets along with support for using custom numeric, image and textual datasets.
- Implemented common ML algorithms using **Tensorflow.js** as blocks and used **D3.js and Plotly.js** for interactive visualizations.
- Presented and published a paper based on a user study with the IDE at IEEE VLHCC 2018 (See Publications).
- Technologies Used: **Node**, **Javascript**

Teaching Assistant, Applied CS with Android Course designed by Google *Jan. 2017 - May 2017*

- Handled **25 two-hour tutorial sessions** for a class of **52 students**.
- **Developed codelabs, internal sites, auto grading systems** for assignments.
- Based on my work, RIT was selected as **one of the 5 universities worldwide, to be featured in a case study** for University Adoption of the program. Link: https://appliedcsskills.withgoogle.com/case-studies/AppliedCS_RIT.pdf
- Technologies Used: **Android**, **Java**, Go, Javascript

Workshops & Talks

Cloud Study Jam

Bangalore india

Facilitator, Google Developer Groups Bangalore *Feb. 2019*

- Conducted a day long hands-on workshop on ML and Google Cloud.
- Topics covered: Intro to Kubernetes, Cloud ML APIs, Cloud ML Engine.
- Used the QwikLabs platform to conduct hands-on sessions
- Link to Slides: https://goo.gl/wPFBsZ

Learning AI Development and UX

Bangalore india

Speaker for Women Tech Makers Bangalore's session on Learning AI Development *Dec. 2018*

- Conducted a day long hands-on workshop on AI development.
- Topics covered: Text classification, Word Embeddings, Feature Engineering, Serving ML Models in Production.
- Used Google's Colab environment to design hands on exercises using Keras and Scikit-Learn.
- Link to Slides: https://goo.gl/x7Etym

Apps Script Study Jam

Bangalore, India

Facilitator for GDG Bangalore's Workshop on Google Apps Script *Feb. 2018*

- Conducted a day long hands-on workshop on Google Apps Script for Google Developer Groups Bangalore (GDG Bangalore).
- Topics: Google Apps Script, Automating workflows with Apps Script, Google Clasp, Integrating third-party services with Apps Script.
- Link to Slides: https://goo.gl/ZH8WQK

The Power of Google Apps Script

Bangalore, India

Speaker for GDG Bangalore's DevFest *Oct. 2017*

- Introduced Google Apps Script with use cases, workflows and demos at Google Developer Groups Bangalore's annual developer conference.
- Link to Slides: https://goo.gl/9gfUUq

Select Publications

Milo: A visual programming environment for Data Science Education

Lisbon, Portugal

Arjun R Rao, Ayush Bihani, Mydhili K Nair *Oct. 2018*

- Proceedings of 2018 IEEE Symposium on Visual Languages and Human-Centric Computing **(VL/HCC'18)**, pp. 211-215. [PDF]
- Designed and implemented a novel **visual programming environment** to help novice students and non-programmers learn **Data Science and ML concepts** using block based programming.
- DOI: 10.1109/VLHCC.2018.8506504

Context Based Approach for Second Language Acquisition

New Orleans, USA

Nihal V Nayak, **Arjun R Rao** *June 2018*

- System paper for Duolingo's shared task on Second Language Acquisition Modelling **(SLAM 2018)**. [PDF]
- Published in the Proceedings of the **NAACL-HLT Workshop** on Innovative Use of NLP for Building Educational Applications **(BEA at NAACL 2018)**.
- Trained a logistic regression model to predict the likelihood of a student making a mistake while answering an exercise on Duolingo. Made use of features inspired by research in **code-mixed language learning** where context plays an important role.
- Result: **AUROC scores for English/Spanish = 0.821**, Spanish/English = 0.790 and French/English = 0.812. **2nd best linear model**, finished 9th overall in SLAM 2018