

Adithya Narayan

+918884275321

Bangalore, Karnataka

[Linkedin](#)

raoadi.8@gmail.com

<https://badirao.github.io>

EXPERIENCE

Research Assistant

Ramaiah Institute of Technology

July 2019 - June 2020

- Compared and Contrasted Classifier Encoder combinations for identifying user intent from Code mixed Hindi-English and Kannada-English queries.
- Published an **Open Source Rule Based Entity Extraction Module for Python**, built specifically for Hindi-English and Kannada-English queries
- Built an **autonomous, crowdsourced, data collection platform** using Cron,Telegram, AWS and MongoDB.
- Built an **API testing dashboard** using React, Flask-Restful, axios,and joblib to verify Response structures and Response times across all the interconnected modules of the Backend.
- Refactored the date utility in python 3 to support Hindi and Kannada date objects(day,month,year,time).
- Developed an algorithm to efficiently create a grammatically coherent code mixed hindi kannada dataset from monolingual hindi and monolingual kannada datasets using parse trees and CFG.

ML and Software Engineering Intern

Vahan Inc

Jan 2020 - June 2020

- Responsible for refactoring the User Validation and Verification vertical
- Used **Selenium, Hadoop MapReduce, and Tesseract** to build one of the largest repositories of Indian names on the web, that would then serve as a lookup table for identifying Name specific information in end-user chat messages. This helped the Operations team **easily identify verified, real-world end users, drastically reduce noise in internal user databases caused by spam and non-contributive end users/end user messages, and reduce investment in external 3rd party verification services.**
- **Refactored the chat validation module** used categorize and validate incoming end user chat messages into one of many different categories, making use of Syntactic and Morphological features. **Increased recall of the same from 85% to 95%.**
- Built an Aadhar verification module to exploit QR code encoded end-user info rather than Vanilla OCR, **decreasing reliability on 3rd party OCR engine by 33%.**

ML and Software Engineering Intern

Stride.AI

June 2019 - Aug 2019

- Built an **ETL pipeline** that extracted text and coordinate data from Unstructured legal documents/PDF's and loaded the same onto CSV files. **Used NLTK and Regex for preprocessing** obtained datasets. Used the same dataset to classify passages of text into one of 8 different topics **using sklearn and low resource data mitigation techniques.**
- Trained and **Deployed handwritten text localization and classification models on French KYC documents using TensorFlow and Protobuf.** Performed A/B testing to rank model architectures and end to end pipelines based on memory requirements, processing speeds and Testing accuracy.

SKILLS

Languages	Python, Java, JavaScript, Shellscript
Web Frameworks	React.JS, (HTML+CSS+JS), Flask, AWS, Bootstrap, Selenium
Tools	Jira, Confluence, Git
Data Frameworks	Hadoop, NLTK, OpenCV, TensorFlow, Joblib, sklearn, Numpy, Matplotlib
Databases	MySQL, MongoDB

PROJECTS

- Building a real time multimedia data collection module in python, that can be used to build datasets that need to source data from Youtube, Instagram and/or reddit. **Making use of Selenium and PyAutoGUI to automate Browser and Keystroke interactions, and bash to deploy as a CLI.**

EDUCATION

Ramaiah Institute of Technology, Bangalore, India
Bachelor of Engineering, Information Science

August 2016 - May 2020

