Samriddhi Sinha

http://samriddhisinha.com
samriddhidjokestersinha@gmail.com
Github:// djokester LinkedIn:// samriddhisinha Quora:// Samriddhi Sinha

Last Updated on 8th June 2020

EDUCATION

IIT KHARAGPUR
BTech: Civil Engineering, 2019
Cum. GPA: 7.19

DELHI PUBLIC SCHOOL
Grad. May 2014 | Ruby Park, Kolkata

POSITIONS OF RESPONSIBILITY

TECHNOLOGY STUDENT'S GYMKHANA
| Technology Coordinator
July, 2017 – Present | IIT Kharagpur

ACM-ICPC 2017 ASIA KHARAGPUR REGIONAL
| Organizing Team Head
September, 2017 - December, 2017 | IIT Kharagpur

SKILLS

LANGUAGES
Python, Julia, R
Matlab and C/C++
Javascript, HTML/CSS

PACKAGES, LIBRARIES AND TOOLS
tensorflow, pytorch, keras
scikit-learn, scipy, numpy, pandas
NLTK, spacy

CLOUD AND DEVOPS
Docker, Kubernetes
AWS, Google Cloud, Azure
Continuous Integration tools (Travis/Jenkins)

MENTORSHIP

KHARAGPUR WINTER OF CODE
December, 2017, IIT Kharagpur

GIRLSCRIPT SUMMER OF CODE
2018

STUDENT WELFARE PROGRAM
2017-19, IIT Kharagpur

WORK EXPERIENCE

CUDDLE.AI, FRACTAL ANALYTICS
| Data Scientist
May 2019 – Present | Mumbai, India
Data Scientist working on the NLP backend of an Artifical Business Analyst. Primarily involved in the development and deployment of algorithms for entity detection and semantic disambiguation in a question answering system that maintains a strict F1 Score above 95 percent.

INTERNSHIPS

GOOGLE SUMMER OF CODE
| Student Software Developer
May-September, 2017

- Developed a natural language processing toolkit in Python for dealing with Indian languages under the mentorship of Portland State University.
- Engineered models for basic NLP functionalities like tokenization, gender tagging, lemmatization and POS tagging.
- Created a model based on Recurrent Neural Network with LSTM units for Part of Speech tagging that utilized word embeddings as features.

FRACTAL ANALYTICS
| Natural Language Processing Intern
May 2018 – July 2018 | Mumbai, India

- Worked as a Natural Language Processing intern with Cuddle.ai. Cuddle.ai puts an AI-powered personal analyst in the hand of the user, delivering mission-critical business insight in a timely manner.
- Developed a BiLSTM based Artificial Named Entity Classifier (ANEC) aimed at restricted business domain systems. The ANEC outperformed a previously established Stanford Constituency Parser based NER in terms of both accuracy as well as search time.

FREELANCE EXPERIENCE

PROLOGIC FRIST
| Natural Language Processing Consultant
Oct 2018 – Mar 2019

Built a Natural Language interface aimed at the Hospitality industry which would interact with AI Assistants like Amazon's Echo and Google Home by extracting structured information from the end-user's commands/requests and in response initiating a certain activity based on the user's commands as well as generating a natural language response providing an update/feedback on the user's command.

AWAIRE
| Machine Learning Engineer
Dec 2017 – Jan 2018

Project based on using deep learning models to perform classification on captured heart data in an attempt to detect abnormal heart sound, heart valvular disease, and even predict possible signs of congestive heart failures with the help of features extracted from heart sound using signal processing techniques.

PROJECTS

AUTOMATIC CONCEPT MAP EXTRACTION | Graduation Project
July 2018 - May 2019 | IIT Kharagpur
The objectives of the project is to create a pipeline that is capable of:

- Extract text from a large collection of documents presented in Portable Document Format(PDF) in a reliable, and widely reproducible manner.
- Extract Important Concepts and Sub-Concepts from the text and link them together with the help of short textual summaries.
- Represent the concept map visually in an effective and easy to consume manner.

COMPETITIVE STRENGTH PREDICTIONS OF ATM VENDORS IN CALIFORNIA
| Silver Technology General Championships
Jan 2017 | IIT Kharagpur

- Part of a 15 member team to analyse the competitive strength of three major ATM vendors from the demographic data of California.
- Scraped data from http://www.unitedstateszipcodes.org/ for demographic data based on PIN codes
- Visualised feature importance with Tableau and clustered ATM locations based on K-Means approach
- Combined per county demographic model with per state demographic model to generate final revenue generation of ATM Location.