

ADITYA SRIVASTAVA

Links : [LinkedIn](#), [GitHub](#)

+91-836-908-8672 ◇ aditya.srivastava@research.iiit.ac.in ◇ Indian Citizen

WORK AND RESEARCH EXPERIENCE

SentiSum

October 2021 - Present

Member of Technical Staff - NLP and Machine Learning

Working on the research and development of ML pipelines for text classification, textual anomaly detection and unsupervised information extraction. [ElasticSearch, AWS, Kubernetes, PyTorch, Docker]

Language Technologies Research Center (LTRC) - IIIT Hyderabad

August 2017 - December 2021

MS by Research and Research Assistant

Previously working with Dr. Dipti Misra Sharma (Head, LTRC) on sentiment extraction and identification of political bias in news; now working on generation and sentiment analysis of code mixed language, by exploiting techniques such as grounding as solutions to low resource NLP.

Rakshak Project - AI for Medicine

August 2020 - February 2021

Research Assistant

Built a platform that facilitates ML and data science on medical datasets under Prof. Ramesh Loganathan and Prof. Vikram Pudi, in collaboration with Intel Labs India and funded by the Govt. of India. I worked specifically on the database and platform design, including the data ingestion, indexing and retrieval systems, and more sporadically on the APIs and other backend systems. [FastAPI, ElasticSearch]

Institute for High Performance Computing and Networking (ICAR CNR, Italy)

May 2019 - March 2020

Research Intern

Worked with Dr. Giuseppe Manco (Senior Researcher, National Council for Research, Italy) and team, on the application of deep neural networks in recommendation systems, more specifically, the application of memory networks, graph neural networks and capsule networks for collaborative filtering.

International Institute of Information Technology, Hyderabad

August 2019 - December 2019

Teaching Assistant - CL1.101 Introduction to Linguistics

Responsibilities included conducting tutorial classes, and setting/grading assignments and test papers.

Qalaxia (A Carrot Inc.)

June 2018 - October 2018

Machine Learning Intern

Worked at Qalaxia with Dr. Manish Shrivastava, on an ML based NLP pipeline for automated word problem solving, by mapping them to handcrafted solution frames that extract entities from the problems and then solve them algorithmically.

Nebulae Digital

August 2016 - October 2016

Intern

Wrote a library for in-house scraping, cleaning and indexing of textual data for use in downstream NLP tasks.

Tata Institute of Fundamental Research (TIFR)

May 2016 - July 2016

Web-Development Intern

Developed multiple web-portals for use within various departments at TIFR.

OMitra Journey Solution Pvt. Ltd.

Jan 2016 - April 2016

Intern

Worked on backend API development and database design.

PUBLICATIONS

Pending submission still in anonymity period, awaiting review.

Due in 2022

TLDR: Transformers for Learning Dictionaries and their Representations

Due in 2022

- Submitted to the [CODWOE shared task in the SemEval Workshop 2022, ACL](#), where we achieved **first place** on the leaderboard for the definition modeling subtask and competitive scores on the embedding generation subtask.
- We used transformer models for both subtasks, experimenting with techniques like unsupervised pretraining, multitask learning and contrastive learning, and will be publishing our results.

HCMS for Sentiment Classification in Code-Mixed Language

December 2020

- Accepted at the SemEval Workshop 2020, ACL.
- Devised a multi-level neural architecture, employing CNNs and self-attention to perform sentiment analysis in a low-resource, code-mixed language setting.

PROJECTS UNDERTAKEN

Collaborative Memory Nets. for Recommendation Systems (Pytorch port)

June 2019

- Created a PyTorch port for the architecture detailed in the aforementioned paper.

Sentiment Analysis on Code-Mixed Conversations

November 2018

- Developed a neural pipeline inspired by Siamese networks for context-aware code-mixed sentiment analysis.
- Employed transfer learning through the use of sentiment specific word embeddings and GloVe embeddings, alongside word level feature engineering.
- The data and problem statement were part of SemEval 2019 - our system placed within the top 5 percentile of all participants.

Wikipedia Search Engine

August 2018

- Designed custom indexing scheme to preprocess the whole dataset to for faster searches.
- Created a search engine to parse custom queries and search over the index to score and retrieve relevant documents from the Wikipedia dataset.

Author Identification on Emails using BiLSTMs

August 2017

- Introduced a novel setup using bidirectional LSTMs, which generate representations for all the data in the email.
- Feature engineering on the data revealed multiple stylometric patterns that were key to identifying individuals as the original authors of the emails.

Vector Representations for Twitter Users

August 2017

- Explored the viability of representing twitter users as vector representations of their tweets.
- Used these representations to recommend user to follow and cluster similar users.
- Created a custom preprocessing toolkit to help with processing tweets.

BrickBreaker, Bloxorz & 3D Aquarium

February 2017

- Brickbreaker and Bloxorz are 2D and 3D games respectively implemented from scratch in OpenGL and C++ without using any physics/game engines. 3D Aquarium is an aquarium simulation implemented in WebGL.

Tree2DB

January 2016

- Created a utility to convert Hindi/Urdu treebanks into SQL databases, with a database schema that supports a diverse range of queries.

SCHOLARSHIPS

The Dean's List of Meritorious Students - IIIT Hyderabad

August 2016

Recipient of the Dean's Merit List scholarship of students with exemplary academic performance (top 20% of the batch) in the academic year 2015-2016.

ACHIEVEMENTS

Winner of Megathon 2018

September 2018

Finished at 1st position in the security domain at Megathon, India's biggest student-run hackathon. Our submission was a system to detect malicious software by spawning multiple virtual machines to run the software in, and then checking the VMs for changes made to them.

Winner of Howzhack 2019

September 2019

Finished at 1st position in the security domain at Howzhack, a national, university-level hackathon. Our submission consisted of a system which used mobile-app usage telemetry to make predictions in the future, and then compare actual usage to predicted usage to flag anomalies. We also developed a web-UI to showcase this.

First Runner Up at Megathon 2019

November 2019

Our submission was a system for automated tagging and indexing of math problems based on the concepts used, to generate balanced tests for school students.

Secured rank 8 worldwide in the Labor challenge at AICrowd.

January 2021

Secured rank 11 worldwide in the SCRBL challenge at AICrowd.

January 2021

AICrowd is a Kaggle-like platform that hosts international ML contests. These two challenges were part of the [AICrowd Blitz #2](#) tournament in which my overall ranking was 43rd worldwide.

EXTRA-CURRICULAR ACTIVITIES

1st Place Inter-House Music Competition

January 2020

2nd Place Inter-House Basketball Tournament

December 2019

1st Place Inter-House Music Competition

August 2015

Coordinator of the IIIT Music Club

August 2018 - May 2019

Organizer for various events conducted by the music club in the academic year 2018-19.

VOLUNTEER AND SOCIAL WORK

Tutor at AshaKiran

January 2017 - April 2017

AshaKiran is a tuition where volunteering IIIT Hyderabad students tutor middle-schoolers and high-schoolers from under-privileged backgrounds, free of cost, to give them a leg-up at school and academics.

Member of the Gender Sensitization Committee (GSC)

August 2019 - April 2019

The GSC aims to make the IIIT campus a safe space for people of all genders, with special focus on women and the LGBTQ community. It tries to achieve this by organizing talks and open discussions about gender issues and by deliberating on and instituting progressive campus policies. It also serves as a place for campus members to confidentially report issues that they may be facing.

Miscellaneous

... - August 2015

Volunteer at multiple blood donation drives, and at the [Tamana NGO](#) which trains differently-abled children to give them the skills to be independent adults.

EDUCATION

International Institute of Information Technology, Hyderabad

August 2015 - Present

M.S. by Research in Computational Linguistics

CGPA 8.25

B.Tech. in Computer Science & Engineering

CGPA 7.45

TECHNICAL SKILLS

Programming Languages

Python, C/C++, JS, BASH

Libraries/Technologies

ElasticSearch, AWS, Kubernetes, Docker, Tensorflow, PyTorch

Miscellaneous

GNU-Linux, Windows, Git, Arduino

RELEVANT COURSES

Data Structures & Algorithms, Artificial Intelligence & Statistical Mechanics in AI, Computational Socio-Pragmatics, Natural Language Processing & Applications, Computational Methods in Linguistics, Distributed Computing, Information Retrieval & Extraction, Language Typology and Universals, Linear Algebra, Speech Technology, Game Theory