

SAYAR GHOSH ROY

+91-7980704678

@ sayar.1997@gmail.com

sayarghoshroy

Sayar's Work

in sayarghoshroy

53/24 R.N Guha Road, Kolkata 700028

EDUCATION

International Institute of Information Technology, Hyderabad

Jun 2017 – Apr 2022

- Bachelor of Technology (Honours) in Computer Science
- Master of Science (by Research) in Computational Linguistics

CGPA: 9.75/10

RELEVANT RESPONSIBILITIES

System Administrator

Jan 2020 - Present

- Managing servers + other resources shared by Centre for Innovation and Entrepreneurship (CIE) & Information Retrieval and Extraction Lab (iREL)

Lecturer at 'Intro to Deep Learning Workshop'

May 2020

- Conducted four sessions for prospective candidates joining the Language Technologies Research Centre (LTRC), IIIT-H
- Created teaching material can be accessed [here](#)

Research Mentor

Jun 2020 - Present

- Mentoring three senior students in their research work in the areas of Text Coherence Analysis, Natural Language Generation from Tabular Data, and domain specific Text Summarization

HONORS AND AWARDS

- Dean's List, IIIT-H for securing the highest SGPA in every single academic semester, 2017-2020
- Endorsed by IIIT-H for the S.N Bose Scholars Program, 2018
- Selected for INSPIRE Scholarship, 2016 by Dept. of Science & Technology, Govt. of India
- Secured 1st School Rank, 8th State Rank in the Indian School Certificate (ISC) with an aggregate of 97.25% & a perfect score in Computer Science
- Award for General Proficiency, 2015, Salt Lake Point School, Kolkata
- Student of the Year 2014, St. Mary's Orphanage and Day School, Kolkata

PROGRAMMING

Languages

python c++ c java javascript bash racket haskell
matlab octave sql html css markdown latex

Libraries, Frameworks, & Tools

pytorch scikit-learn tensorflow spacy nltk corenlp numpy scipy
cvxpy matplotlib pandas opencv opengl webgl pygame
flask django bootstrap lucene(solr) neo4j riscV bluespec

RELEVANT COURSEWORK

Computer Programming Algorithms Data Structures Discrete Mathematics
Artificial Intelligence Statistical Methods in AI Optimization Methods
Probabilistic Graphical Models Principles of Programming Languages
Natural Language Processing (Intro + Advanced) Deep Learning: Theory & Practices
Computer Networks Intro to Databases Operating Systems Graphics
Linear Algebra Probability and Statistics Real & Complex Analysis
Linguistics (Intro + Social & Historical + Typology & Universals)
Computational Linguistics Neural & Cognitive Modeling Engineering Systems

Do check out my work [here](#).

EXPERIENCE

Undergraduate Researcher

Information Retrieval and Extraction Lab (iREL)

Aug 2019 – Present

Hyderabad Area, India

- Guided by Prof. Vasudeva Varma and Dr. Manish Gupta
- Created a system for predicting sentence-wise relevance of online news articles in answering Microsoft Bing search queries
- Currently working on Automatic Text Summarization, Fact Salience, and Scientific Discourse Analysis

Software Engineering Intern

Digitalee

May – Jul 2019

Hyderabad Area, India

- Built a pipeline for generating summaries of executive briefings
- Implemented a PDF to text converter adhering to the structure of large executive briefing documents
- Designed a summarizer based on state-of-the-art research in related areas which considers various sections and units of typical executive briefing documents
- Created an end-to-end architecture for accepting PDF documents & producing text summaries

Software Engineering Intern

Theodor

Sep – Nov 2018

Hyderabad Area, India

- Built a system facilitating conversion of hand-filled forms into a digital format
- Utilized CTPNs for text detection and CRNNs for text recognition
- Implemented a classifier for form fields such that scanned images of forms can be converted into JSON
- Created a web app using Django for users to upload scanned forms, receive digitized information, and have a dashboard of all activities

SELECTED PROJECTS

Acronym Sense Disambiguator | python • nltk

- Identifies acronyms in a text file and disambiguates possible expansions, where a guided disambiguator module runs queries on an indexed Wikipedia dump

Image classification formulations | python • scikit-learn • TensorFlow

- Performance of various deep learning architectures for face recognition and miscellaneous image classification tasks: insights and analysis

COVID-19 modeling for India | python • scipy

- Modeling the Coronavirus Statistics for India using the Susceptible-Infected-Recovered (SIR) Model

Detecting parse errors due to linguistic repair | python • spacy

- Detecting errors in dependency parsing of speech transcript primarily due to active linguistic repair - correlations between types of parse failures and the existence of repair in the utterance

Extractive summarization unit | python • scikit-learn • spacy

- A lightweight easy-to-train extractive summarization formulation for the CNN dataset

Machine Translation systems | python • PyTorch

- Statistical MT system for English → Hindi utilizing IBM models and an HMM formulation for relative token alignment
- Supervised neural MT system for English → Bengali using a seq2seq encoder-decoder architecture with an attention mechanism

C shell | c

- From scratch implementation of a user-defined interactive shell program similar to a bash shell capable of creating and managing new processes

Game Design

- Mario on Terminal: Simple version of the popular Mario game playable on your terminal | python
- Funride: A basic twist on the classic Jetpack Joyride | c++ • openGL
- Space Invaders Retro: Space Invaders recreated with the retro-feel | pygame
- Angry Paper Jet: Frugal Fighter Jet Simulation Game | c++ • openGL
- Green Lawn Surfers: Subway Surfers but on a lawn & with Angry Birds | javascript • webGL