Mayank Goel

BTECH CS + MS, COMPUTATIONAL LINGUISTICS

Phone Number
+917890545454

Website
mayankgoel28.github.io

Email Address
themayankgoel28@gmail.com

Github
mayankgoel28

AWARDS

Best Project (Theme track)
Natural Language API hackathon, expert.ai
June 2021

1st Prize, Equity Track
HackUMBC, University of Maryland
November 2020

Best Web Application
HackAtHome, Brown University
October 2020

3rd Prize, Walmart Track
TAMU Datathon, Texas A&M University
October 2020

Best Technical Product
SummerUp 2020, Ecell IIIT Hyderabad
June 2020

Top 30, India
Panini Linguistics Olympiad
March 2019

LANGUAGES AND TOOLS

Python, Javascript, MERN, Bash, C++

Nutch, Solr, Pytorch, Tensorflow, Plotly

COURSEWORK

Computer Systems Organization · Computational Linguistics (I and II) · Algorithm Analysis and Design · Database and Applications · Linear Algebra · Probability and Statistics · Design and Analysis of Software Systems · NLP · Machine Data and Learning

INTERESTS

Stand-up Comedy

Debate

Philosophy

EDUCATION

IIIT Hyderabad
BTech + MS Computer Science and Computational Linguistics
(July 2019 - Present)
8.51

WORK EXPERIENCE

Trivedi Centre for Political Data, Ashoka University
Research Associate
(May 2021 - June 2021)

Made data visualizations using Plotly, displaying them on a dashboard made using Streamlit, with interactive features and parameters to view the data, for an internal diagnostics tool. Used numpy and pandas to develop an extensive set of features.

Proxzar
Intern
(April 2021 - May 2021)

Crawled and extracted data from E-Commerce websites, using Apache Nutch and Apache Solr. Used BS4 and various libraries for a tool for generalized data extraction, automating the process through Bash and Python. Solved challenges such as login barriers and different layouts for each site.

Panini Linguistics Olympiad
Member, Organizing Committee
(January 2019 - Present)

Organization Committee member for the Indian chapter of the International Olympiad of Linguistics. Involved with handling email, paper correction, and miscellaneous duties.

PUBLICATIONS

Towards Conversational Humor Analysis and Design
11th Humor Research Conference
January 2021

Worked on identifying and creating jokes, based on core theories of Humor. Used NLP tools to show challenges involved with jokes, along with a novel approach of template-extraction and filling using BERT.

PROJECTS

Echodex
https://github.com/chaudhary1337/Echodex

Wrote a generalized tool to fetch data from a news API, Twitter and Reddit. Worked on fetching and cleaning the data to a frontend made using streamlit, after using NLP tools for analysis such as sentiment analysis and document classification.

Domain Specific Search Engine
https://github.com/MayankGoel28/DSSE

Made a Knowledge-based ontological Search engine, automating most part of the cleaning process and using Semantic Similarity to relay data from over 20k Walmart products.

Name Entity Recognition System
https://github.com/DomainNER/NamedEntityRecognizer

Automatically tagged a large corpus of 500k sentences with Named Entities to create training data, and used multiple approaches for NER.

Maze Algorithms
https://github.com/Defectors-IIIT/Puzzles

Worked on randomized Maze Generation, based on an Object Oriented Template developed by team. Did Complexity Analysis and analysis of bias with generation, and also worked on an Adversarial Search Algorithms for a PacMan-like game on the mazes generated.

Summarizer-Bot
https://github.com/MayankGoel28/Summariser-Bot

Uses cosine similarity on sentence vectors to summarize long texts, along with keyword extraction for a discord bot. Encountered and solved problems such as code mixing and data processing using a predominant language agnostic approach.