PARTH PARIKH

(+91) 9082600410 ◇ parthparikh1999p@gmail.com ◇ https://github.com/pncnmnp

EDUCATION

Rajiv Gandhi Institute of Technology, Mumbai *August 2017 - June 2021*
Bachelor in Engineering
Department of Computer Engineering *CGPA - 8.78*

COURSES UNDERTAKEN

Structured Programming Approach, Digital Logic Design and Analysis, Discrete Structures, Data Structures, Analysis of Algorithms, Advanced Algorithms, Operating System, Computer Organization and Architecture, Computer Networks, Database Management System, Theory of Computation

RESEARCH EXPERIENCE

Indian Institute of Information Technology, Allahabad *May 2020 - July 2020*
Topic: Affective Analysis of Project Gutenberg's corpus
Guide: Prof. Uma Shanker Tiwary and Mr. Punit Singh

- Worked with JeMaS and Doc2Vec models to predict the valence, arousal, and dominance space of each section in a book.
- Experimented with KNN, Adaboost, and Random Forest models to classify the emotions present in Emobank and Gutenberg corpora.
- Implemented a variant of the Discrete Emotions Questionnaire to estimate the mood of any reader and recommend books to improve their mood.

PROJECTS

Sthir - Spectral Bloom Filters for Client-Side Search *June 2020 - Present*
Innovated a memory-efficient library to perform client-side searching using the probabilistic data structure - Spectral Bloom Filters. Implemented Okapi BM25 for ranking the documents. Based on two performance benchmarks: size comparison (gzip-based) and search result comparison (using Kendall's Tau and Spearman's Footrule), this library produces rankings comparable to Lunr.js with less memory footprint.

Detecting air pollution hotspots and identifying source trajectories *Jan 2020 - Feb 2020*
Developed a model to detect air-pollution hotspots and predict their forward/backward source trajectories using data obtained from satellites such as ERA5 and Sentinel-5P. Explored the use of DBSCAN, a clustering technique, to detect air-pollution from a non-flat geometry. Designed a geographically weighted regression model to convert satellite data to ground concentrations for validation.
This project was submitted to ISRO for Smart India Hackathon, 2020.

Popup Encyclopedia *May 2019*
A browser extension aimed to provide word meanings on double-clicking a word. Optimized it to perform faster than state-of-the-art software like Google Dictionary using offline indexing techniques. Improved search accuracy by implementing word-lemmatization.

Crossword-Solver *December 2019*
Designed a class to guess clues using databases like Wordnet, Moby's thesaurus, and Gensim's glove-wiki-gigaword-100. An implementation based on Z3 Theorem Prover (SMT solver) was developed to position the guesses on the crossword-board.

Movie Recommendation System *September 2019 - November 2019*
Developed a movie recommendation engine using content-based and collaborative filtering approaches

with a hybrid recommender system optimizing their recommendations. Implemented machine learning approaches like KNNBaseline and Cosine Similarity. Curated *The Indian Movie Database(TIMDB)*, the largest database available for Indian movies, with over 4500 titles released between 1950 and 2019.

Pizza Delivery Chatbot *March 2019*
Conceptualized and developed a parser with support for universals, generic responses, sentiment analysis, and curse-word detection. Analyzed customer priority using KNN, hand-curated core datasets and added support for various customer-analytics using SQLite.
This project secured an Honourable mention in DBIT hackathon.

Search Engine *July 2018*
Designed a web-crawler using Scrapy, performed word refining using NLTK, and indexing using an Inverted Index. Added auto-completion using Tries and *did you mean* feature by analyzing Levenshtein distance between words. Implemented retrieval and error detection in Flask and ranking using Tf-Idf.

MINOR PROJECTS

- Designed a **Branch Target Buffer** for Computer Organization and Arch. Lab. *April 2019*
- Developed a **progressive-web-app/website** to provide news in rural areas. With low request sizes and fast loading time, the site was optimized for low bandwidth connections. *July 2019*
- Developed a feature-rich **URL shortening service** with support for duplicate link verification, auto expiry of URL, custom URLs, and spam detection. Designed a spam detection library using mmh3 and bitarrays. *February 2019*
- Designed a news-aggregator based toolkit in Django which parsed RSS feed, found breaking news from them (both categorically and non categorically), performed summary scraping and retrieved stocks from BSE, NASDAQ, and NSE. *November 2018 - June 2019*
- Created an AI-based Tic-Tac-Toe game using the Minimax algorithm. *February 2018*

TECHNICAL SKILLS

Programming Languages	Proficient in Python, C; Prior experience in Java, Bash, Octave
Libraries	Django, Flask, Scikit-learn, Pandas, Numpy, NLTK, Scrapy, Tkinter
Software Skills	MySQL, SQLite, Octave, Markdown, Git, LaTeX, AutoCAD
Web Technologies	Javascript, HTML, CSS

Last Updated: 18 August 2020