Parth Parikh

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

Education

Rajiv Gandhi Institute of Technology (University of Mumbai) *August 2017 - June 2021*
Bachelor in Engineering - Department of Computer Engineering *CGPA - 9.09*

Publications

Spectral Bloom Filters for Client Side Search *Final Draft*
***Parth Parikh**, Mrunank Mistry, Dhruvam Kothari, Sunil Khachane*
11^{th} IEEE Annual IEMCON, November 2020, Vancouver, Canada

Technical Experience

Research Intern - Indian Institute of Information Technology, Allahabad *May 2020 - July 2020*
Affective Analysis of Project Gutenberg's corpus *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. Incorporated multiple models such as KNN, Adaboost, and Random Forest 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.

MLH Fellow (Major League Hacking, New York) *September 2020 - December 2020*
BentoML - Framework for serving, managing, and deploying ML models. (*commiter*)

- Added support for URL prefix to allow users to run YataiService behind a reverse proxy server.

Authored sections of a handbook to help fellows navigate large codebases of any open-source project.

Projects

LuaNLP - Natural Language Processing Library for Lua *February 2021 - Present*
Implemented modules for Penn Treebank word tokenizer, sentence tokenization using Naïve Bayes classifier, Wordnet-based lemmatization, Porter Stemmer, Parts-of-speech tagger using averaged perceptron, VADER algorithm for sentiment analysis, RAKE algorithm for keyword extraction, and TextTeaser for text summarization.

Sthir - Spectral Bloom Filters for Client-Side Search *June 2020 - October 2020*
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.

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.
This project was submitted to the Indian Space Research Organization for Smart India Hackathon, 2020.

Anaphora Resolution *February 2020 - April 2020*
Designed a BERT model to obtain the contextual word embeddings, and encode the mentions of interest. Implemented a logistic regression-based model to predict if a mention-pair creates an anaphoric reference.

Crossword Solver to solve mini New York Times' crosswords *December 2019*
Designed a class to guess clues using databases like Wordnet, Moby's thesaurus and using word2vec models like GloVe and positioned the guesses on the crossword-board using Z3 Theorem Prover (SMT solver).

Indian 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. Curated *The Indian Movie Database(TIMDB)*, **currently the largest database available for Indian movies**.

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

Branch Target Buffer *April 2019*

Designed it with a 2-bit saturating counter using doubly linked-lists and LRU replacement policy.

Scaled-down Search Engine *July 2018*

Designed a web-crawler using Scrapy, back end using Flask, and used an Inverted Index for mapping. Added auto-completion using Tries, ranking using TF-IDF weighting scheme and *did you mean* feature using Levenshtein distance algorithm.

Technical Skills

Programming Languages	Proficient in Python, Lua, C; Prior experience in Java, Bash, GNU Octave
Libraries/Frameworks	Django, Flask, Scikit-learn, Pandas, Numpy, NLTK, Scrapy, Tkinter
Software Skills	Linux Shell utilities, MySQL, SQLite, Markdown, Git, LaTeX, AutoCAD
Web Technologies	Javascript, HTML, CSS
Operating Systems	Ubuntu, Debian, Tiny Core Linux, CentOS, MacOS, Windows

Courses Undertaken

Structured Programming Approach, Digital Logic Design and Analysis, Discrete Structures, Data Structures, Analysis of Algorithms, Advanced Algorithms, Computer Graphics, Microprocessor, Database Management System, Theory of Computer Science, Machine Learning, System Programming and Compiler Construction, Software Engineering, Data Warehousing and Mining

Obtained a full grade in the above relevant courses
Guest lectured on the topic of Support Vector Machines as part of the Machine Learning course

Last Updated: March 13^{th} 2021