

Sachin Malhotra

• 1253.5 West 36th Pl, Los Angeles

• +1-213-822-2536

• sachinma@usc.edu

• Portfolio

• Github

• LinkedIn

"If you can't explain it simply, you don't understand it well enough" - Albert Einstein

Education

University of Southern California

Los Angeles, CA

M.S. in Computer Science

August 2017 - May 2019 (anticipated)

Coursework: CSCI 570 - Analysis of Algorithms, CSCI 544 - Applied Natural Language Processing, CSCI 561 - Artificial Intelligence

Work Experience

Information Sciences Institute

Marina del Rey, CA

Research Intern

May 2018 - Present

- Working under Prof. Premkumar Natarajan on the problem of Neural Abstract Generation.
- Introduced novel architectural changes to an existing model for abstract generation.
- Designing changes focussed around injecting additional context into the model for better guiding the generation process.

Information Sciences Institute

Marina del Rey, CA

Student Worker

Sept 2017 - Present

- Familiarized myself with the internal workings of a Feed Forward Network and learned about LSTMs and Sequence to Sequence models.
- Prepared a custom dataset consisting of 50,000 titles and abstracts for the task of Neural Abstract Generation.
- Trained custom word embeddings using text extracted and cleaned from the PDFs for 50,000 scientific papers.

Hike Messenger

Gurgaon, HR

Software Developer

June 2015 - July 2017

- Reduced the number of Redis cache calls from 72,000 to just 4000 per second by updating a user's activity time in batches.
- Achieved 2.4 million concurrent connections on a single HAProxy instance with 16 cores in a load testing effort. This helped reduce the size of the HAProxy's production cluster by a factor of 3.
- Reported and rectified the user's last seen time issue on the application, which was affecting the entire user base of 70 million users.
- Consolidated all the last seen related data in a single Redis cache thus removing multiple sources of truth and hence resolved the issue.

Hike Messenger

Gurgaon, HR

Software Developer Intern

Jan 2015 - June 2015

- Accelerated the time to ship a server component by a factor of 2 by setting up developer testing environments on Docker.
- Programmed a messaging proxy solution from scratch that would prevent dropping of user connections on server restarts. This improved end user experience and provided seamless deployments of server components.

Projects

LSTM based Sequential Short Text Classifier

Python | PyTorch

- Implemented an LSTM based sentence level classifier for annotating sentences of a given abstract as Introduction, Body or Conclusion.
- Trained the classifier on the PubMed dataset consisting of 20,000 abstracts and achieved a test-set accuracy of 90%.

Baby Lizards Solver - A fun variant of the classic N-Queens puzzle

Python

- Coded 4 different algorithms for solving the N-Queens puzzle.
- Implemented a fast Depth First Search algorithm to solve this variant of the N-Queens problem with obstacles.
- Achieved a remarkable result of placing 97,000 lizards on a 1000-by-1000 board in under 2 seconds with this approach.

Hotel Reviews Classifier

Python

- Developed a Vanilla perceptron and an Average perceptron for classifying hotel reviews as True/Fake and Positive/Negative.
- Achieved a test-set F1 score of 0.93 on the provided dataset with the Averaged perceptron.

Coderchef Kitchen

Python | C++

- Founder and maintainer at Coderchef Kitchen which is an actively growing repository of amazing programming problems with well commented solutions and beautiful diagrammatic explanations in the form of blog posts.

Hidden Markov Models based Part of Speech Tagger

Python

- Programmed a part-of-speech tagger using Trigram and Bigram HMMs, the Viterbi algorithm and Laplace smoothing.
- Obtained a test-set F1 score of 0.92 on the provided data set with Trigram implementation.
- Published a two part series for the same on the freeCodeCamp publication on Medium.

Skills

Programming Languages: Python, Java (Moderate), C++ (Moderate).

Tools and Frameworks: Jupyter Notebooks, NumPy, PyTorch, TensorboardX, Git.

Databases: Redis, MongoDB, MySQL.

Achievements

- Technical Writer for the largest tech publications on Medium, freeCodeCamp.
- Published a wide variety of articles with topics ranging from Data Structures and Algorithms to Natural Language Processing.
- Received over 130,000 views spread over 17 posts till date.