

Hossam Zaki

Education

(401) 855 5567 • hossam_zaki@brown.edu • hossamzaki.com • github.com/hossam-zaki

Brown University (3.9/4.0)

Bachelor of Science in Computer Science and Biology,

Providence, RI

Expected Graduation: May 2022

- **Related Coursework:** Computational Biology, Computer Vision, Deep Learning, Genetics, Organic Chemistry, Physiology, Linear Algebra, Graduate Mathematical Statistics, Honors Statistical Inference.
- **Programming Languages:** Python, Java, R, Dart, Bash, Javascript, SQL, HTML/CSS, C++.
- **Libraries/Frameworks:** Tensorflow, Sklearn, Numpy, Pytorch, Pandas, Node.js, Angular.js, Flutter, Android Studio.
- **Scholarships:** Jackie Robinson Foundation Scholar, Merck ACS Scholar.

Work Experience

Computational Biology (CSCI 1810)

Undergraduate Teaching Assistant

Providence, RI

September 2020 - Present

- Iterated upon course content, graded assignments, conducted TA hours, and set up occasional labs.
- Worked with the Ethics TA program to develop new CS ethics content for both this course and the department in general.

Wessel Lab at Brown University

Undergraduate Researcher

Providence, RI

July 2017 - Present

- Building Deep Learning model to build regulatory networks of germ-line genes.
- Developed pipeline to search and find new genetic motifs in the RNA of germ-line cells using Python and the Knuth-Morris-Pratt algorithm.
- Co-authored an abstract presented in Developmental Biology of the Sea Urchin Conference titled: "Bisphenol A exposure differentially affects echinoderm embryogenesis. Developmental Biology of the Sea Urchin".

Broad Institute of MIT and Harvard

Undergraduate Researcher

Boston, MA

June 2020 - August 2020

- Researched effect of mutations in DNA repair genes on whole genome structural variants.
- Co-Authoring and presented abstract at SACNAS and ABRCMS titled "Characterizing Structural Variants in RAD51B"
- Developed algorithm using Python and R and analyzed data using various statistical tests.
- Combined data from several databases including ICGC, PCAWG, Uniprot, Ensembl, and TCGA.

Google

SPS Program Participant

San Francisco, CA

June 2020 - August 2020

- Worked with team to develop a real-time voting app using Java and HTML/CSS.
- Tested app with several Googlers and gave presentation to several people in the Translate team.

Protein Data Bank of Europe

Software Engineering and Bioinformatics Intern

Cambridge, UK

June 2019 - August 2019

- Developed a pipeline to superpose protein structures in a UniprotKB Accession, cluster the proteins based on structure similarity, and show them visually on PyMol. Now deployed on PDBE-KB protein pages.
- Designed Algorithm using Python and SQL and analyzed data using Hierarchical Clustering.
- Compiled the algorithm on over 47,000 Uniprot Accessions which included over 350,000 PDBE entries.
- Co-authored and published a paper to Nucleic Acids Research titled "PDBE: Improved findability of Macromolecular structure data in the PDB".

Software Projects

Apollo | A Speech-to-Text Note-Taking App for Doctors

<https://github.com/hossam-zaki/apollo> | <http://apollohealth.herokuapp.com/>

- Utilized Google's Speech Recognition Library to capture appointment transcript, and used Knuth-Morris-Pratt to search for relevant details.
- Encrypted patient data and appointment transcript using Google Tink.
- User-tested with medical students at Brown Medical School, and doctors at Rhode Island Hospital.

Computational Analysis of Sea Urchin Transcriptome

https://github.com/hossam-zaki/Sea_Urchin_Transcriptome_Analysis

- Developed pipeline to search for known genetic motifs in the transcriptome of the Sea Urchin, as well as discovering new motifs.
- Implemented Knuth Morris Pratt Algorithm to linear-time pattern recognition as well as a Suffix Trie to search for the longest most common substring.

X-Ray Classification With Bone-Shadow Suppression

<https://github.com/hossam-zaki/X-Ray-Classification>

- Constructed an auto-encoder to remove bone shadow on NIH X-Ray Images for classification
- Built a neural network to classify X-Ray images in NIH Dataset.
- Achieved an average of 78% True Positive Rate.
- Utilized Tensorflow and Keras frameworks.

Protein Structure Clustering through Superposition

PDBE News Article

- Built a tool to cluster proteins based on structural similarity, now used 25,000+ times by labs all over the world.
- Using protein superposition tool GESAMT and hierarchical clustering, protein structures are grouped by structural similarity.