

SEJAL DUA

Tufts University Class of 2021
Data Science & Biomedical Engineering

WORK EXPERIENCE

IBM Research

Machine Learning Research Intern

Jun 2020 - Sep 2020, Yorktown Heights, NY

- ❖ Engineered primary NLP classification model for IBM Drug Repurposing for Cancer pipeline by performing scientific document corpus filtering through heuristic rules and statistic modeling.
- ❖ Infused domain expertise into Snorkel binary classifier and achieved 83% accuracy when applying the model on 127,000 unlabeled PubMed articles and a validation set of 1,400 labeled articles.

Textbook Exchange Network

Director of Data Analytics

Jan 2019 - Present, Medford, MA

- ❖ Gathered data-driven insights from 6000+ API exchanges representing textbook transactions that have saved students \$500,000 compared to campus bookstore prices.
- ❖ Calculated Key Performance Indicators (KPIs) via SQL queries and Python statistical methods to measure the health and wealth of the organization and market estimated impact on [website](#).
- ❖ Managed a team of computer science students through numerous cycles of exploratory data analysis, integration, visualization, and presentation at monthly showcases.

LEGO Education Engineering Lab

Software Development Intern

Jun 2019 - Aug 2019, Boston, MA

- ❖ Developed a Python software package which enables bidirectional communication between a LEGO SPIKE Prime hub and various microprocessors, thus making it compatible with existing products.
- ❖ Improved LEGO image recognition functionality by implementing a smile detection neural network demonstration using the sklearn package.

Oregon Center for Aging & Technology

Software Engineering Intern

May 2018 - Aug 2018, Portland, OR

- ❖ Facilitated and executed a research validation experiment designed to obtain gait speed data from an elderly cohort of participants with Alzheimer's or dementia; engineered laser trip beams and 3D-printed sensor enclosures for passive infrared sensing virtual hallway.
- ❖ Wrangled and visualized 20 years of passive infrared sensing data, investigating the gradual degradation of the circadian rhythm as a result of aging.

Tufts University Computer Science Department

Teaching Fellow

Jan 2019 - Present, Medford, MA

- ❖ Hosted office hours, facilitated labs, and improved grading infrastructure for 120+ person classes; Served as liaison between professor and 30+ person teaching staff.

CO-CURRICULARS

Tufts Women in Technology Conference

Speaker Coordinator

Apr 2020 - Present

Networked with women in executive positions at top tier tech companies; Recruited 8 speakers to inspire and empower college students at virtual women in technology conference.

Towards Data Science

Freelance Data Journalist

Nov 2019 - Present

Performed sentiment analysis on song lyrics using the Spotify API, Python, and Tableau; Wrote 6 Medium articles reaching over 26K readers in the data journalism space.

HACKATHONS

TechTogether 2020

IBM: Drug Repurposing for Cancer | Dell Technologies: Driving Human Progress

Jan 2020

Built an NLP-powered, user-facing annotation tool that sifts through abstracts from therapeutic cancer intervention studies and identifies the drug type, cancer type, therapeutic association, and study type.

URx InternHacks

Best Technical Problem

Jun 2020 - Jul 2020

Integrated a Flask web application that facilitates intellectual conversation between users with a K-Means Clustering algorithm, to help mitigate social media echo chambers.

CONTACT



(503) 505-0595



sejal.dua@tufts.edu



sejaldua



www.sejaldua.com



[sejal-dua-761371164](https://www.linkedin.com/in/sejal-dua-761371164)

EDUCATION

Tufts University (Medford, MA)

GPA: 3.77 Dean's List

Graduation: May 2021 (expected)

Jesuit High School (Portland, OR)

GPA: 4.26 weighted, 3.96 unweighted

Graduation: June 2017

RELEVANT COURSES

- Data Structures
- Algorithms
- Machine Learning
- Data Visualization
- Neural Networks and Deep Learning
- Big Data
- Artificial Intelligence
- Cybersecurity

TECHNICAL SKILLS

Coding Languages

Python, C, C++, Java, R, SQL, HTML, CSS, Javascript, PHP, MATLAB, Bash

Frameworks & Technologies

React, Node.js, Flask, Tableau, MongoDB, Postgres, AWS, Docker, Makefile, Git

Packages & Libraries

matplotlib, numpy, scipy, spacy, pandas, sklearn, D3, Keras, Tensorflow, PyTorch, Snorkel

Productivity Tools

VSCode, Jupyter Notebooks, Xcode, LaTeX, Microsoft Office Suite, Trello

Microprocessors & Fabrication

Raspberry Pi, Arduino, micro:bit, OpenMV Camera, PyBoard, SolidWorks, 3D Printing, Laser Cutting

INTERESTS

Fitness Tech, Mandarin Chinese, Sports Analytics, Music, Sushi Connoisseur, Travel, Netflix Documentaries, Yoga & Meditation