

# SEJAL DUA

Tufts University Class of 2021  
Data Science & Biomedical Engineering

## EXPERIENCE

### IBM Research

#### Machine Learning Research Intern • Jun 2020 – Sep 2020

- ❖ Built initial model within Drug Repurposing for Cancer NLP pipeline by performing scientific document corpus filtering through heuristic rules and statistic modeling.
- ❖ Applied Snorkel binary classifier on 127,000 clinical PubMed articles and a labeled test set of 1,400 documents; Achieved 83% accuracy, sufficient for scalability to other domains.

#### LEGO Education Research & Development Lab [Tufts CEEO]

#### Software Development Intern • Jun 2019 – Aug 2019

- ❖ Developed a Python software package which enables bidirectional communication between a LEGO SPIKE Prime hub and various microprocessors.
- ❖ 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

- ❖ Facilitated and executed a research validation experiment designed to obtain gait speed data from an elderly cohort of participants with mild cognitive impairment, often in the form of Alzheimer's or dementia.
- ❖ Sifted through 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

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

## PROJECTS

### Beyond the lyrics: the intersection of music and data visualization

#### Towards Data Science Article • Nov 2019

Performed sentiment analysis on song lyrics using the Spotify API, Python, and Tableau; wrote a Medium article using data to show how artists evolve and how genres differ.

### Intellivo

#### Best Technical Problem (InternHacks) • 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.

## CLUBS AND CO-CURRICULARS

### Textbook Exchange Network

#### Director of Data Science • Jan 2020 – Present

Overseeing a data analytics team working to gain insights from 4000+ API exchanges representing textbook transactions that have saved students \$300,000 compared to prices found at the campus bookstore.

### Tufts Women in Technology Conference

#### Speaker Coordinator • Apr 2020 – Present

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

## CONTACT



(503) 505-0595  
sejal.dua@tufts.edu  
sejaldua  
www.sejaldua.com  
sejal-dua-761371164

## EDUCATION

### Tufts University (Medford, MA)

GPA: 3.77 *Dean's List*  
Graduation Date: May 2021

### Jesuit High School (Portland, OR)

GPA: 4.26 weighted, 3.96 unweighted  
Graduation Date: June 2017

## RELEVANT COURSES

- Data Structures
- Machine Struct. & Assembly Lang.
- Algorithms
- Machine Learning
- Data Visualization
- Neural Networks and Deep Learning
- Big Data
- Artificial Intelligence

## TECHNICAL SKILLS

### Coding Languages:

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

### Packages & Libraries:

matplotlib, numpy, scipy, pandas, d3, keras, tensorflow, pytorch, sklearn

### Operating Systems:

Linux, UNIX, Windows

### Software Development Tools:

Github, VSCode, Sublime Text, Jupyter Notebooks, Xcode, LaTeX, Tableau

### Technologies & Microprocessors:

Raspberry Pi, Arduino, micro:bit, OpenMV Camera, PyBoard

### IoT Platforms:

Thingworx, Vuforia