

# Vahideh Rasekhi, PhD

## DATA SCIENTIST

---

vrasekhi@gmail.com | (631)599-8516 | linkedin.com/in/vahidehrasekhi/ | github.com/vahidehrasekhi/

---

### PROFESSIONAL SUMMARY

Data Scientist with experience in data mining, data analysis, data visualization, natural language processing, machine learning, and data storytelling. Coming from academia, I have 10+ years of experience conducting research, discovering insights from data, presenting important findings to diverse audiences, and publishing results in journals. As a life-long learner with a strong drive to learn and master new technologies and techniques, I am passionate about combining my academic background and data science skills to solve problems and deliver actionable insights.

---

### SKILLS

- **Python** – data manipulation and exploratory data analysis (EDA) using *pandas*, *numpy*, *regex*, *matplotlib*, *seaborn*, machine learning using *scikit-learn*, *Keras* and *TensorFlow*, natural language processing (NLP) using *SpaCy*, *NLTK* and *VADER*, web scraping and html data manipulation with *beautifulsoup* and *requests*
- **SQL** – database import and manipulation (incl. complex joins), Common Table Expressions (CTE) and analysis
- **R** – data manipulation and EDA with *tidyverse*, *dplyr*, *ggplot2*, *plotly*, interactive app development using R Shiny
- **Git/GitHub** – Version control

**Other tools:** Microsoft 365 Office, Praat, LaTeX

---

### PROFESSIONAL EXPERIENCE

#### DATA SCIENTIST

Nashville Software School | Nashville, TN

09/2021 – 06/2022

*9-month intensive data science training. The program emphasized practical application of Python, R and SQL technologies for addressing real-world business problems by organizing and cleaning data, exploratory data analysis, performing machine learning, preparing data visualizations and data storytelling.*

#### CAPSTONE PROJECT – OFFENSIVE SPEECH DETECTION ON TWITTER

Utilized NLTK, Scikit-learn libraries to detect offensive language on Twitter. Cleaned and analyzed about 50,000 tweets in Python. Ran different machine learning models including Naïve Bayes, Logistics Regression, Random Forest, Decision Trees, LSTM/CNN, Topic modeling to identify the overall best performing model.

<https://github.com/vahidehrasekhi>

#### MID-COURSE CAPSTONE PROJECT – SENTIMENT ANALYSIS OF COVID VACCINATION TWEETS

Created RShiny app to analyze sentiment of people around the world towards COVID vaccination. Streamed tweets using Twitter API. Used Python for cleaning and exploratory data analysis, before diving into VADER Sentiment Analysis for 5,000,000 tweets and visualizing the output with R Shiny.

[https://vahidehrasekhi.shinyapps.io/Covid Twitter Sentiment Analysis](https://vahidehrasekhi.shinyapps.io/Covid%20Twitter%20Sentiment%20Analysis)

*Other projects included:*

- Automated extraction of information from 10K forms in HTML format, using SpaCy
  - Tokenized top 50 most downloaded books from “Project Gutenberg” using NLTK. Found their Jaccard and Cosine similarity, applied different N-gram probability models to predict the next word in a sentence.
  - Analyzed Medicare data, generated suggestions on increasing patient referrals to Vanderbilt University Medical Center using Python and SQL
  - Scraped Home Mortgage Disclosure Act website and created an R Shiny dashboard to visualize how peer institutions compare regarding their approval/denial rates of applications based on applicants’ demographics.
-

---

## PROFESSIONAL EXPERIENCE (CONT'D)

### POSTDOCTORAL RESEARCHER AND LECTURER

University of California, Los Angeles (UCLA) | Los Angeles, CA

09/2018 – 06/2021

- Designed/conducted research studies on syntax, language processing, phonology/phonetics, sociolinguistics.
- Collected data, conducted qualitative and quantitative data analysis in R.
- Presented findings at 20+ conferences, published them, have been cited by researchers in the field 65+ times.
- Created curriculum, taught several undergraduate courses, designed assessment/evaluation materials.

### DATA ANALYTICS & ALUMNI ENGAGEMENT COORDINATOR

Stony Brook University | New York, NY

08/2017 – 08/2018

- Coordinated programs and events that engage more than 200,000 alumni worldwide.
- Designed a survey using Qualtrics to assess alumni's experience, conducted quantitative / qualitative data analysis using Excel, in collaboration with the Advancement team.
- Communicated the results to the Vice President for Advancement Strategy, Executive Director of Alumni Relations, and alumni in a report, brought back 10% more alumni to the campus the following year.
- Updated and managed portfolios of prospect donors, started the Grateful Patient fundraising program, identified 500+ prospects and raised more than \$1.22M in major gifts and \$185K in annual giving.
- Revised/edited the Alumni Association's Board of Directors' Manual and Alumni Relations' Employee Manual.

### MANAGER OF LANGUAGE LEARNING AND RESEARCH CENTER

Stony Brook University | New York, NY

06/2015 – 08/2017

- Managed foreign language placement and proficiency exams required for undergraduate students.
- Wrote quarterly reports to the Director of the Center and Dean of College of Arts and Sciences.
- Liaised with the Orientation office the times and dates of the exams for freshman students.
- Hired and trained graduate students to proctor the exams following each exam language's rules and guidelines.
- Taught 25 graduate and undergraduate courses (e.g. LIN 541 & 542, LIN 307, LIN 355, ESL 192, PER 101)

### ELECTED PRESIDENT OF GRADUATE STUDENT ORGANIZATION

Stony Brook University | New York, NY

06/2015 – 06/2017

- Represented graduate student body (8500 students) from different disciplines at local and national levels.
- Successfully led the efforts at the university to increase graduate students' annual stipend by 10%.
- Liaised with the state and federal elected officials including Senator Schumer and former VP Joe Biden.
- Proposed initiatives on higher education, increasing federal research funding, DACA, immigration reform.
- Managed \$500,000 annual budget, supported students' professional development, social/cultural activities.
- Organized talks by raising \$100,000 from offices on campus, congregated 3000 university stakeholders.
- Supervised 4 board members, 3 staff, 80 graduate student representatives from different departments.

---

## EDUCATION

### NASHVILLE SOFTWARE SCHOOL

Data Science Bootcamp

09/2021 – 06/2022

Statistics for Data Science

07/2021 – 08/2021

### STONY BROOK UNIVERSITY

PhD Linguistics

08/2010 – 05/2018

### CALIFORNIA STATE UNIVERSITY, FRESNO

MA Linguistics

09/2008 – 12/2009