

# David P. Huber

2833 NE Davis St  
Portland, OR 97232  
(973) 714-0735

dhuber1@gmail.com  
<https://www.linkedin.com/in/davidhuberpx>

## OVERVIEW:

Recently completed a master's degree in information and data science, concentrating on machine learning and natural language processing techniques. Former software engineer with an obsessive desire to solve problems elegantly balanced with the ability to deliver results quickly. I am excited to apply my experience and passion at an industry-leading organization with a commitment to innovation and advancement.

## TECHNICAL SKILLS:

- Python, SQL, R, C/C++, Pandas, Numpy, Linux, Windows, MapReduce, Spark, GIT, Machine Learning, Natural Language Processing, Deep Learning, AWS, GCP, Docker

## PROJECTS:

### Reality Check: *Choose Your News* - [www.realitycheck.works](http://www.realitycheck.works) (Synthetic Capstone Design)

- Full-term project exploring the problem of fake news and a potential remedy.
- Fine-tuned a transformer network pretrained on a language model to classify news articles into 10 categories of news and combined those and other results into a unified ranking score.
  - Python, GCP with GPU, AWS EC2 with GPU, Neo4J, Postgres

### Kaggle Competition: *Facial Keypoint Detection* (Introduction to Machine Learning)

- Explored neural network architectures, activation functions and optimizations to train a classifier to detect 15 facial keypoints on input facial images.
  - Keras, Theano, Python, Amazon EC2 with GPU acceleration

### Twitter Flu Shot Experiment (Experiments and Causality)

- Ran an experiment to measure the effect of informational reply tweets on retweet count and engagement rate of tweets with a negative flu shot sentiment.
  - Amazon Lambda, Amazon DynamoDB, Keras, Python

## EDUCATION:

Dec 2018	<i>University of California, Berkeley, Berkeley, CA</i> <b>Master of Information and Data Science, School of Information</b> <b>Courses:</b> Statistics For Data Science      Experiments and Causality Introduction to Machine Learning      Scaling Up! Really Big Data NLP with Deep Learning      Machine Learning at Scale
May 2006	<i>Rutgers University, New Brunswick, NJ, with honors</i> <b>BS, Electrical and Computer Engineering, Economics minor, College of Engineering</b> <b>BS, Computer Science, Physics double major, Rutgers College</b>

## **PROFESSIONAL EXPERIENCE:**

**Conversica, Inc.,** Seattle, WA

6/18 – 9/18

---

### **Data Science Intern**

- Applied state-of-the-art NLP techniques in a fast-paced startup culture.
  - Built intent model training set through a combination of active learning, data wrangling, and human annotation.
  - Used LSTMs for data-refinement and intent modeling.
  - Investigated sentence clustering techniques.
  - Improved email section detection decision tree model.
- Transitioned models for GDPR compliance.
- Communicated key insights to project stakeholders.
- Participated in a bi-weekly scrum cycle.

**UC Berkeley, School of Information,** online

1/18 – 12/18

---

### **Teacher's Assistant for W261: Machine Learning at Scale**

- Assist course instructors in a graduate-level course that explores scalable algorithms and how to implement them with Hadoop MapReduce and Apache Spark.
  - Provide student support for course concepts.
  - Grade assignments and produce comprehensive feedback.

**CyberExtruder.com, Inc.,** Newark, NJ

6/05 – 1/06 & 7/09 – 4/15

---

### **Senior Research and Development Engineer**

- Responsible for several C++ applications in a multithreaded, Windows environment.
- Built DLL interfaces to stream video from local and remote devices for real-time processing.
- Redesigned existing code for improved memory management and multithreaded efficiency by utilizing the Windows API.
- Created interface libraries to abstract third party facial recognition and IP camera SDKs.
- Implemented computer vision techniques for human detection and object tracking.

**ImageWare Systems, Inc.,** San Diego, CA, Portland, OR

7/06 – 7/09

---

### **Software Engineer**

- Maintained and enhanced a distributed, multi-modal, biometric template matching engine.
- Improved query efficiency by redesigning memory management in the biometric template cache and utilizing multiple processors via the Windows API.
- Developed bulk loading utilities for data ingestion.
- Provided all-purpose support and training to customers, partners, and employees.