

# Alon Albalak

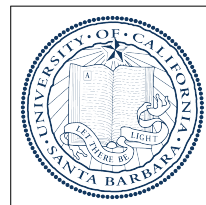
PhD Candidate, Computer Science  
University of California, Santa Barbara

✉ [alonalbalak@gmail.com](mailto:alonalbalak@gmail.com)

🌐 [Personal Webpage](#)

🐙 [GitHub](#) [in LinkedIn](#)

🐦 [Twitter](#) [Scholar](#)



## About Me

I am a Ph.D student in the Computer Science Department at the University of California, Santa Barbara, co-advised by William Yang Wang and Xifeng Yan. My work lies at the intersection of natural language processing, machine learning, and logic. I mainly focus on data efficiency in NLP through transfer learning, data augmentation and neuro-symbolic methods.

## Education

2018–present **Ph.D, Computer Science**, *University of California, Santa Barbara*.

[UCSB NLP Group](#)

Advisors: [William Yang Wang](#) and [Xifeng Yan](#)

2016–2018 **B.S., Mathematics**, *Wayne State University*.

## Publications and Patents

- 2022 **FETA: A Benchmark for Few-Sample Task Transfer in Open-Domain Dialogue**.  
[Alon Albalak](#), Yi-Lin Tuan, Pegah Jandaghi, Connor Pryor, Luke Yoffe, Deepak Ramachandran, Lise Getoor, Jay Pujara, William Yang Wang.  
[Preprint](#) [\[code\]](#)
- 2022 **Addressing Issues of Cross-Linguality in Open-Retrieval Question Answering Systems For Emergent Domains**.  
[Alon Albalak](#), Sharon Levy, William Yang Wang.  
[Preprint](#) [\[code\]](#)
- 2022 **D-REX: Dialogue Relation Extraction with Explanations**.  
[Alon Albalak](#), Varun Embar, Yi-Lin Tuan, Lise Getoor, William Yang Wang.  
NLP for Conversational AI Workshop, ACL 2022. [Paper](#) [\[code\]](#)
- 2021 **Systems and methods for determining and using semantic relatedness to classify segments of text**.  
Rohit Jain, Devin H. Redmond, Richard B. Sutton, [Alon Albalak](#), Sharon Huffner.  
US Patent US20210279420A1
- 2021 **Emotion Recognition in Conversation using Probabilistic Soft Logic**.  
Eriq Augustine, [Alon Albalak](#), Anurag Prakash, Connor Pryor, William Yang Wang, Lise Getoor.  
[Preprint](#)
- 2021 **Modeling Disclosive Transparency in NLP Application Descriptions**.  
Michael Saxon, Sharon Levy, [Alon Albalak](#), Xinyi Wang, William Yang Wang  
EMNLP 2021 Main Conference. [Paper](#)
- 2019 **The Sounds of Seizures: Audio-Triggered Detection by Convolutional Neural Network**.  
Maysaa Basha, [Alon Albalak](#), Hani Alhourani, Aashit Shah, Ming Dong  
Neurology. [Paper](#)
- 2017 **Decoding User's Intention from Surface EEG Signals Using Machine Learning Algorithms**.  
[Alon Albalak](#), Guanghua Xu, Chaoyang Chen, Ming Li, John M Cavanaugh  
Annual Meeting of BMES

---

## Selected Projects

- February 2021 – **Recommender Dialogue Systems**, *in collaboration with UCSC, USC, Google.*  
present
  - o Actively collaborating with researchers across institutions to solve problems in dialogue systems such as explainability, information extraction, and zero- or few-shot dialogue classification tasks
  - o Led a team in the design, development, and publication of state-of-the-art work on the explainability of relation extraction methods in dialogue
  - o Applied Probabilistic Soft Logic (PSL) to the task of emotion recognition in conversation to achieve state-of-the-art results
- Advisors : Industry - [William W. Cohen](#) and [Tania Bedrax-Weiss](#)  
Academic - [William Yang Wang \(UCSB\)](#), [Lise Getoor \(UCSC\)](#), and [Jay Pujara \(USC\)](#)
- June 2021 – **Alexa Prize Taskbot Challenge**, *Team Lead.*  
present
  - o 8% acceptance rate
  - o Led and advised UCSB's "Team GauchoBot" in developing an agent that assists real Alexa customers to complete cooking and do-it-yourself projects that require multiple steps and complex decision making
  - o Designed algorithms for intent classification and question answering as well as the communication architecture between modules
  - o Extensively used AWS services including, but not limited to, EC2, ECS, ECR, S3, CloudWatch, DynamoDB, and Lambda
- May 2021 – **COVID(ATAK)**, *in collaboration with IARPA and Peraton Labs.*
- October 2021
  - o Mentored a visiting undergraduate researcher
  - o Built a multilingual open-retrieval question answering system for COVID-related journal articles and a clinical trials database
  - o Designed and implemented:
    - a multilingual deep semantic indexing method to retrieve relevant documents
    - a multilingual reading comprehension system to find answers within a document
  - o Online demo will be available shortly

---

## Professional Experience

- June 2019 – **Research Associate**, *Theta Lake.*
- September 2020
  - o Built classifiers for automated risk detection in regulated industries through the use of natural language processing and other machine learning techniques
  - o Took multiple projects from inception to production, and developed 2 patent pending methods along the way
- December 2017 – **Machine Learning Research Associate**, *Machine Vision and Pattern Recognition Lab, Wayne State University.*
- September 2018
  - o Research funded by the Epilepsy Foundation, titled "The Sound of Seizures"
  - o Built computer vision based CNN-LSTM model predicting the onset of seizures with 91% accuracy
  - o Optimized neural network in Keras/TensorFlow for portability to mobile devices
- July 2016 – **Research Assistant**, *Robotic Rehabilitation and Neurophysiology Lab, Wayne State University.*
- July 2017
  - o Custom built an arduino controlled lower limb exoskeleton
  - o Implemented neural network on motor cortex EEG data to control exoskeleton using python and matlab
  - o Built system to control a motorized wheelchair through EEG and EMG with a deep neural network

---

## Fellowships & Awards

- 2018 **Integrative Graduate Education and Research Traineeship (IGERT) Fellow**, *University of California, Santa Barbara.*
- 2018 **Academic Excellence Fellowship**, *University of California, Santa Barbara.*
- 2018 **Chia Kuei Tsao Award**, *Wayne State University.*  
For outstanding academic achievement in the undergraduate mathematics program

## Technical skills

Tools Python, C++, Shell, AWS, Azure  
Packages PyTorch, TensorFlow, Keras, NumPy, SciPy  
Machine Learning Natural language processing (NLP), computer vision (CV), transformers, sequence to sequence models, statistical analysis, regression, clustering

## Teaching Experience

Spring, 2020 **CS 165a: Artificial Intelligence - Lead TA.**  
Fall 2020 – **CS 9: Object Oriented Programming.**  
Spring 2021

## Military Experience

2012 – 2015 **Reconnaissance Sabotage Unit, Israel Defense Forces.**

- o Engineering, demolitions, and reconnaissance specialty training
- o Battalion lead navigator