

Gustavo Aguilar

gustavoaguilar91@gmail.com

linkedin.com/in/gustavoag91 ◇ github.com/gaguilar ◇ gustavoaguilar.io

EDUCATION

Ph.D. in Computer Science

University of Houston

Dissertation: Neural Sequence Labeling on Social Media Text

GPA: 3.9 / 4.0

Spring 2017 - Fall 2020

Houston, TX

Advisor: Dr. Thamar Solorio

Bachelor of Science in Computer Science

Universidad Centroamericana José Simeón Cañas

Thesis: VR Algorithm Using Kinect Camera for Amputees Undergoing Rehabilitation

GPA: 8.31 / 10.0 — Top 3 in class of 2015

Mar 2010 - Oct 2015

La Libertad - El Salvador

EXPERIENCE

Applied Scientist II

Amazon, Alexa AI

- Built large-scale KGs from customer-Alexa interactions and trained GNNs for query rewriting (QR)
- Mentored intern for a multilingual vocabulary-free neural tokenizer (accepted at RepL4NLP 2022)
- Deployed a graphical Markov model for QR on Hindi data

March 2021 - Present

Seattle, WA (remotely)

Research Assistant

University of Houston, [RiTUAL lab](#), Dr. Thamar Solorio

- Named entity recognition on user-generated text ([NSF grant #1910192](#))
- English-Spanish code-switching on social media data

September 2016 - December 2020

Houston, TX

Deep Learning Research Scientist (Intern)

Salesforce Research, Einstein AI, MS Bryan McCann & Dr. Nitish Keskar

- Investigated robust transformer-based language models for noisy text
- Extended the subword embedding space using robust character compositionality ([paper](#))

May 2020 - August 2020

Palo Alto, CA (remotely)

Applied Scientist (Intern)

Amazon, Alexa AI, Dr. Benjamin Yao

- Conducted research on knowledge distillation for QA relevance
- Proposed a new method for internal knowledge distillation of transformers ([paper](#))

May 2019 - August 2019

Seattle, WA

Applied Scientist (Intern)

Amazon, Alexa Speech, Dr. Chao Wang

- Conducted research on emotion recognition with multimodal data
- Proposed multi-view learning to induce semantics into acoustic models ([paper](#))

May 2018 - August 2018

Cambridge, MA

Teaching Assistant

Introduction to Natural Language Processing ([COSC 6336](#))

- Elaborated and graded in-class assignments, homework, and practical exams
- Held weekly office hours and covered a few lectures when needed

Jan 2018 - May 2018

Houston, TX

Game Developer

Beyond Games Inc. (based in San Francisco, CA)

- Programmed in different areas of game development (Gameplay, Graphics and AI)
- Co-led 14 engineers, reviewed the team's code, and presented updates to senior management

November 2013 - July 2016

San Salvador, El Salvador

GRADUATE COURSEWORK

- As MS student (one semester) — **Fall 2016:** Artificial Intelligence, Natural Language Processing, Visualization;
- As PhD student — **Spring 2017:** Fundamentals of Machine Learning (at Rice University); Advanced Numerical Analysis; **Fall 2017:** Advanced Natural Language Processing; Digital Image Processing; **Spring 2018:** Computer Architecture, Research Methods; **Fall 2018:** Operating Systems

HONORS AND AWARDS

- [2020 Outstanding PhD Student Award](#) by the CS department, University of Houston
- [2019 Snap Inc. Research Fellowship](#) (\$10,000 + research internship offer)
- [2019 Facebook PyTorch Scholarship](#) for the Udacity [Deep Learning](#) and [Deep RL](#) Nanodegrees
- Scholarships to attend [2019 SACNAS National Diversity in STEM](#) (Waikiki, HI), [2019 ACM Richard Tapia](#) (San Diego, CA), and the 2019 CRA Grad Cohort Workshop for URMD (Honolulu, HI)
- [2018 Best Junior Ph.D. Student Award](#) by the CS department, University of Houston
- Univ. of Houston Alumni Association and Natural Sciences and Mathematics Scholarship, 2018
- First place on [Noisy User-generated Text \(W-NUT\) shared task](#) in all the categories, EMNLP 2017
- Dean's List, 2010-2014, Universidad Centroamericana José Simeón Cañas (UCA)
- [National chess champion of El Salvador in 2012](#) and governmental award for chess excellence in 2008

RESEARCH ACTIVITIES AND SERVICES

- Guest lecture for Intro. to NLP (COSC 6336, Spring 2020): Sequence to Sequence Models ([slides](#), and [code](#))
- Reviewer at EACL-21, NAACL-21, EMNLP-20, CALCS-20, W-NUT-19-20, ACL-20, FLAIRS-20.
- Co-organizer of the SemEval-2020 Task 9 on [Sentiment Analysis for Code-Mixed Social Media Text](#)
- Speaker at Data Science El Salvador community, [December 2018](#), and [December 2019](#) ([slides](#), [code1](#), [code2](#))
- Speaker at the [Computer Science Focus on Research \(FOR\)](#) seminars at the Univ. of Houston, Fall 2018
- Co-organizer of the 3rd workshop on [Computational Approaches to Linguistic Code-Switching \(CALCS\)](#)

PUBLICATIONS

Selected publications (full list available in my [Google Scholar profile](#)):

- [Char2Subword: Extending the Subword Emb. Space with Robust Character Compositionality](#). **EMNLP '20**. **Gustavo Aguilar**, Bryan McCann, Tong Niu, Nazneen Rajani, Nitish Keskar, and Tamar Solorio.
- [From English to Code-Switching: Transfer Learning with Strong Morphological Clues](#). **ACL 2020**. **Gustavo Aguilar** and Tamar Solorio.
- [SemEval-2020 Task 9: Overview of Sentimnet Analysis of Code-Mixed Tweets](#). **SemEval 2020**. P Patwa*, **G Aguilar***, S Kar, S Pandey, S Pykl, D Garrette, G Björn, T Chakraborty, T Solorio, A Das.
- [LinCE: A Centralized Benchmark for Linguistic Code-switching Evaluation](#). **LREC 2020**. **Gustavo Aguilar**, Sudipta Kar, and Tamar Solorio.
- [Knowledge Distillation from Internal Representations](#). **AAAI 2020**. **Gustavo Aguilar**, Yuan Ling, Yu Zhang, Benjamin Yao, Xing Fan, and Chenlei Guo.
- [Multimodal and Multi-view Models for Emotion Recognition](#). **ACL 2019**. **Gustavo Aguilar**, Viktor Rozgić, Weiran Wang and Chao Wang.
- [Dependency-Aware Named Entity Recognition with Global and Relative Attentions](#). **arXiv 2019**. **Gustavo Aguilar** and Tamar Solorio.

- [NER on Code-switched Data: Overview of the CALCS 2018 Shared Task](#). *3rd CALCS Workshop*, **ACL 2018**. **Gustavo Aguilar**, Fahad AlGhamdi, Victor Soto, Mona Diab, Julia Hirschberg and Thamar Solorio.
- [Modeling Noisiness to Recognize Named Entities using Multitask NNs on Social Media](#). **NAACL 2018**. **Gustavo Aguilar**, A. Pastor López, Fabio A. González and Thamar Solorio.
- [A Multi-task Approach for Named Entity Recognition on Social Media Data](#). *3rd W-NUT*, **EMNLP 2017**. **Gustavo Aguilar**, Suraj Maharjan, A. Pastor López, and Thamar Solorio.