Andres Cremisini

Data Scientist

Miami, FL, 33126

1-954-612-5378

acremisini@gmail.com

https://www.linkedin.com/in/andrescremisini/

Data Scientist and published researcher. Industry and research experience with the entire Data Science pipeline, significant background in Natural Language Processing (NLP). Nuanced, formal understanding of statistical analysis and machine learning techniques, always curious to learn more. Strong humanities background, highly adept at creative thinking and explaining complex topics to non-experts. Graduating with an M.S. in Data Science in Dec. 2019, available Jan. 2020. Authorized to work in the E.U. and the U.S.

Work History

2017-08 - Current

Graduate Researcher: NLP And Data Science

Cognition, Narrative And Culture Lab, Miami, FL

- Implemented state-of-the-art machine learning language model to detects when different descriptions of events refer to the same occurrence (event coreference resolution)
- Extensive literature review of academic papers
- Two academic publications (see *Publications & Personal Projects* section)

2019-06 - 2019-08

Data Scientist Intern

Bureau Of Labor Statistics, Washington, DC

- Developed a natural language (NLP) classifier that achieved > 97% accuracy, providing fundamental technology to automate a procedure that costs millions of dollars
- Designed a statistical framework to account for classifier uncertainty
- Developed an intuitive, automated dashboard that allows non-technical staff to train new classifiers and interpret results
- Delivered highly successful talks about machine learning to non-technical managers
- Collaborated with internal stakeholders to understand a complex coding system and develop features for the classifier

2019-06 - 2019-08

Data Science Fellow

Coding It Forward, Washington, DC

Selected as a Data Scientist for the Bureau of Labor Statistics, the principal fact

finding agency of the U.S. Government. *Coding it Forward* is a competitive 10 week internship program that places students in civic tech positions at federal agencies.

2017-06 - 2017-08

Computer Scientist Intern

Sandia National Laboratories, Albuquerque, NM

- Designed and implemented a wind model using a sophisticated statistical technique and real-time data from NOAA
- Picked up C++ quickly from no previous experiencce due to strong programming fundamentals, writing well-designed code that is still in use
- Significantly improved existing weather model, allowing the lab to use a more realistic simulation environment

2016-08 - 2017-08

Computer Science Fellow

Applied Research Center At Florida International University, Miami, FL

- Analyzed requirements and designed, developed and implemented the middle-tier for a mobile app
- Evaluated potential software products based on new and existing system development and migration requirements

2013-08 - 2014-05

5th Grade And 12th Grade Teacher

Colegio Integral El Avila, Caracas, Venezuela

- Designed and taught original curriculum in English and creative development for two 5th grade classrooms
- Designed and taught original curriculum in the philosophy of music for three 12th grade classrooms

2010-01 - 2016-01

Composer

Freelance

- Music for TV, Film and Advertising. Successfully courted clients and translated often convoluted requirements into bespoke music
- Commission from the Miami Symphony Orchestra for original orchestral piece
- Performances across the United States and Italy

Publications & Personal Projects

Extracting and Aligning Timelines
*Finlayson, M., **Cremisini, A.**, Ocal, M.,*
Events and Stories in the News, Cambridge University Press (forthcoming)
Book chapter exploring the computational extraction of narratives and timelines from a corpus of documents.

A Challenging Dataset for Bias Detection: The Case of the Crisis in the Ukraine
***Cremisini, A.**, Aguilar, D., Finlayson, M.,*
2019 SBP-BRiMS Conference, Springer Lecture Notes in Computer Science
Conference paper studying the suitability of document embeddings for bias

detection in the news.

Neural Language Embeddings for Comment Clustering
Implemented a state of the art deep learning network (Siamese-LSTM) in pytorch as part of a sophisticated user comment clustering application.

Miami MetroRail Audit
Statistical analysis and data collection/cleaning for an audit of the Miami Metrorail. The project led to the first advocacy-led redesign of a bus system in the United States. View the audit here: https://transitalliance.miami/campaigns/transit-audit

🎓 Education

2017-08 - Current

Master of Science: Data Science
Florida International University - Miami, FL
Grade 3.91 / 4
Graduate Research Fellow
Research in NLP. Courses in Data Science with a focus on Machine Learning and NLP.

2016-01 - 2017-05

Bridge To Graduate Study: Computer Science
Florida International University - Miami, FL
Grade 3.95 / 4
Undergraduate courses in preparation for graduate study including Programming (OO), Data Structures, Databases and mathematics up to Linear Algebra, Probability and Statistics.

2011-08 - 2013-05

Master of Arts: Music Composition
University of California, Berkeley - Berkeley, CA
Grade 3.91 / 4
Eugene Cota-Robles Fellow
Research in Composition, Music Cognition and Computer Music

2006-08 - 2010-12

Bachelor of Arts: Music Composition
University of Miami - Coral Gables, FL
Grade 3.83 / 4
Magna Cum Laude, "Outstanding Graduate" in Composition

🖥 Software

Python, Java, C#, C++, SQL, R