ANA VALDIVIA

Fairness in Machine Learning, Computational Social Science, Natural Language Processing

@ anna.valdivia.garcia@gmail.com   github.com/anavaldi   @ana_valdi
valdilab.wordpress.com

EXPERIENCE

Data Scientist
Trilateral Research

October 2019 to currently   London, United Kingdom

- Design and develop software, machine learning algorithms and other data analysis and visualisation tools for threat and vulnerability analysis to support decision-making, stabilisation efforts, and humanitarian response (`www.trilateralresearch.com/project/hamoc`).
- Lead the data science team to understand the relationship between mental health and criminality on adolescents living in UK. Partnering with a charity.
- Perform research work related to current projects, write reports or sections of reports and develop other deliverables.
- Participate in the preparation of research proposals for funding (H2020).

Data Science for Social Good Fellow
University of Chicago

May 2018 to August 2018   Chicago, United States

- Identify factors that contribute to students dropping out in El Salvador.
- Build a database to ensure secure, ethical and privacy-considerate processes.
- Envision the Ministry of Education of El Salvador will use these factors to better decide how to allocate their social program budget as well as drive decisions regarding programs to be implemented in schools
- Advisors: Rayid Ghani (Director of Data Science for Public Policy Center of University of Chicago, Ex-Chief Scientist of Obama For America)
- `github.com/dssg/el-salvador-mined-public`

Research Assistant
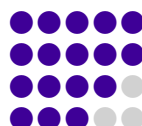**IESE Business School**

2014 to 2015   Barcelona, Spain

- Build a model to study and manage the impact of exhibitions on the number of museum visitors.
- Estimate the model using data collected from two large museums in Barcelona (MNAC and CaixaForum).
- Design an optimization framework in which the museum operator can choose the best duration and synchronization of exhibitions.
- Advisor: Víctor Martínez de Albéniz (PhD at the Operations Research Center of MIT)

SKILLS

Python/R
Tableau/Gephi
SQL/AWS
C++

EDUCATION

PhD in Machine Learning & Natural Language Processing with Honors
Universidad de Granada

2016 to 2019   Granada, Spain

- Title: Sentiment Analysis-Based Methods for Cultural Monuments.
- Advisor: Francisco Herrera (highly cited author)
- `github.com/anavaldi/phd`

MsC in Data Science and Computer Engineering
Universidad de Granada

2015 to 2016   Granada, Spain

BsC in Mathematics
Universitat Politècnica de Catalunya - BarcelonaTech

2008 to 2014   Barcelona, Spain

- Erasmus Program at University of Groningen, The Netherlands.

HONORS & AWARDS

🏆 **The most innovative paper presented at HAIS 2017**
Our work *Sentiment Analysis on TripAdvisor: Are There Inconsistencies in User Reviews?* was awarded and considered for publication at Neurocomputing.

🏆 **UniversityHack CAJAMAR 2017 - Special Mention**
The jury rewarded our data visualizations on the proposed problem.

🏆 **Datathon CETAQUA 2016 - Winner Team**
Top position at the first datathon organized by the University of Granada, CETAQUA, and Emasagra.

PUBLICATIONS

📄 Journal Articles

- Valdivia Ana, Casillas Jorge and Sanchez Javier (2020). "How fair can we go? Assessing the boundaries of fairness in decision tree algorithms". In: *working paper*.
- Valdivia Ana, et al. (2020). "Governance in online citizen participation: Auditing the urban planning process in Decidim Barcelona". In: *working paper*.
- López, Miguel et al. (2019). "E2SAM: Evolutionary ensemble of sentiment analysis methods for domain adaptation". In: *Information Sciences* 480, pp. 273–286.
- Valdivia, Ana et al. (2019). "Inconsistencies on TripAdvisor reviews: A unified index between users and Sentiment Analysis Methods". In: *Neurocomputing* 353, pp. 3–16.
- Zuheros, Cristina et al. (2019). "Deep recurrent neural network for geographical entities disambiguation on social media data". In: *Knowledge-Based Systems*. link.
- Martınez-de-Albéniz, Victor and Ana Valdivia (2018). "Measuring and exploiting the impact of exhibition scheduling on museum attendance". In: *Manufacturing & Service Operations Management*.
- Valdivia, Ana et al. (2018). "What do people think about this monument? Understanding negative reviews via deep learning, clustering and descriptive rules". In: *Journal of Ambient Intelligence and Humanized Computing*. link, pp. 1–14.
- Valdivia, Ana, M Victoria Luzón, Erik Cambria, et al. (2018). "Consensus vote models for detecting and filtering neutrality in sentiment analysis". In: *Information Fusion* 44, pp. 126–135.
- Valdivia, Ana et al. (2017b). "Sentiment analysis in tripadvisor". In: *IEEE Intelligent Systems* 32.4. link, pp. 72–77.

👥 Conference Proceedings

- Valdivia, Ana et al. (2020). "Auditing the urban planning process in Decidim Barcelona". In: *Data Justice Conference*. Data Justice Lab.
- – (2017a). "Neutrality in the sentiment analysis problem based on fuzzy majority". In: *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, pp. 1–6.
- Valdivia, Ana, M Victoria Luzón, and Francisco Herrera (2017). "Sentiment Analysis on TripAdvisor: Are There Inconsistencies in User Reviews?" In: *International Conference on Hybrid Artificial Intelligence Systems*. Springer, pp. 15–25.

VOLUNTEERING

- ✏️ **Teacher at Tibetan Monastery, Kathmandu, Nepal 2019**
  I taught Maths and English to Tibetan monks between the ages of 4 and 12. For three weeks I lived with them in the monastery, soaking up the wisdom and philosophy of Buddhism.

- 🎓 **Teacher at Alpargate, Granada, Spain 2017-currently**
  I teach Math to teenagers which live in marginal neighbourhoods and motivate them so that someday they can neighbourhood their dreams.

- ♀ **Geek & Tech Girls, Granada, Spain 2017-currently**
  Geek& Tech Girls aims to visibilize and empower women and girls in the field of technology.

RECENT PROJECTS

Social Media Analysis Workshop - Main Organizer
- Teaching students from law and sociology degrees to analyze data from Twitter.
- github.com/anavaldi/Taller_Sociologia

Toxic Comment Classification Challenge
- Tutoring Master's degree students to build a multi-headed model that's capable of detecting different types of of toxicity like threats, obscenity, insults, and identity-based hate.

Data Ethics and Social Good Workshop - Main Organizer
- Contributing to a social debate around Artificial Intelligence applications to know how they affect us on our daily lives.

RECENT TALKS

Big Data for Social Good - 18 July, 2019
Invited by The Fluxions Podcast
medialab.ugr.es/evento/the-fluxions-1x05-ana-valdivia-big-data-for-social-good

Reducing Early School Dropouts in El Salvador- 23 March, 2019
Medialab Prado, Madrid, Spain
medialab-prado.es/actividades/datos-etica-y-bien-comun

NLP for Linguistics & Social Sciences - 21 March, 2019
Invited by University of Salamanca, Salamanca, Spain
github.com/anavaldi/pyladiesNLP/upload/master

Big Data for Optimistics- 13 March, 2019
Invited by Centro Mediterráneo, Granada, Spain
canal.ugr.es/evento/conferencia-una-mirada-optimista-del-bigdata

How Big Data is Shaping our Society - 12 March, 2019
Invited by Medialab UGR, Granada, Spain

NLP for Mathematicians - 05 March, 2019
Invited by the Young Mathematical Researchers Group, Granada, Spain

Big Data for Social Good - 19 February, 2019
Invited by the Geek & Tech Girls, Granada, Spain