Jorge Saldivar

Dr. César López Moreira 444
Asunción, Paraguay
Email: jorgesaldivar@gmail.com
Webpage: jorgesaldivargalli.com
Github: github.com/joausaga
+595981534660

DATA SCIENTIST, RESEARCHER, AND SOFTWARE ENGINEER

EDUCATION

PhD Information and Communication Technologies, University of Trento, Italy *2017*

Informatics Engineering, Catholic University of Asuncion, Paraguay *2010*

DATA SCIENCE PROJECTS

- **Motivation Factors and Participation Patterns in Crowdsourcing (∼2k records).** Applied **exploratory data analysis (EDA)** to study the profile of the participants of crowdsourced civic participation processes. Employed **non-parametric statistical tests** (Wilcox, Spearman, Chi-square, Kruskal-Wallis, Friedman) to examine change in the motivation factors that drive people to crowdsourced civic participation processes. Used **Logistic Regression** to predict the odds of participants to stay engaged in crowdsourcing processes. (Tools: R, dplyr, ggplot2)[1]

- **The effectiveness of Social Sharing Practices (∼35k records).** Used **Multivariate Linear Regression** and **parametric statistical tests** (Pearson correlation, T-test) to study the effectiveness of the ubiquitous social media sharing buttons in increasing participation in online social communities (Tools: R, ggplot2, dplyr, reshape)[2]

- **Collective and Individual Behavior in Online Communities (∼300k records).** Used **K-means algorithm** and **non-parametric statistical tests** (Chi-square, Kruskal-Wallis) to discover patterns in the collective behavior of online innovation communities (Tools: R, ggplot2, dplyr)[3]

- **Politic Bots (∼200k records).** Employed **exploratory data analysis (EDA)** techniques (summary statistics and visualizations) to understand how social media bots and fake accounts are used to promote and manipulate information during electoral periods (Tools: Python, Pandas, Numpy, Matplot, MongoDB)[4]

- **Online Popularity of Movies (∼10k records).** Used **exploratory data analysis (EDA)** methods (summary statistics, visualizations, and correlation) to understand the factors that contribute to the online popularity of movies (Tools: Python, Pandas, Numpy, Matplot)[5]

LANGUAGES

Spanish (Native), English (Full professional proficiency), Italian (Limited working proficiency)

PROFESSIONAL EXPERIENCES

Research Fellow
Catholic University of Asuncion - DEI, Asunción, Paraguay *2016-Present*

- Coordinated the implementation of an analytics tool that uses machine learning (ML) and natural language processing (NLP) techniques (**classification**, **clustering**, **concept extraction**, **sentiment analysis**) to automate the process of crowdsourced textual data
- Implemented a **K-means** algorithm to cluster similar documents
- Applied **K-means** algorithm to discover patterns in the collective behavior of online innovation communities
- Employed **exploratory data analysis (EDA)** to study the use of social media bots for political propaganda
- **Published six papers in top conferences such as ACM CHI (2) and ACM CSCW (1)**

---

[1] Repo: https://github.com/joausaga/finland-experiment-study
[2] Repo: https://github.com/joausaga/ims-sn-study
[3] Repo: https://github.com/joausaga/collective-behavior-im-communities
[4] Repo: https://github.com/ParticipaPY/politic-bots
[5] Repo: https://github.com/joausaga/moviesonlinepopularity

PhD Researcher
University of Trento - DISI, Trento, Italy *2011-2015*

- Employed **non-parametric statistical tests** to examine the motivation factors that drive citizens to participate in crowdsourced civic participation processes
- Used **exploratory data analysis (EDA)** to study the profile of the participants of crowdsourced civic participation processes
- Applied **multivariate linear regression analysis** to study effectiveness of current social sharing practices
- **Published three articles in IEEE and ACM journals and two conference papers**

Visiting Scholar
University of California, Berkeley - CITRIS, California, USA *2014*

- Modeled, designed, and implemented a Twitter app that allows citizens of California to propose suggestions on issues that merit the attention of the government
- Contributed to the execution of a real-case process of crowdsourced policymaking in the state of California (participation > 10,000 residents)
- **Published a paper in an ACM Conference**

Visiting Researcher
Analytics Lab, Hewlett Packard, Palo Alto, California, USA *2013*

- Designed and implemented a web-based interactive analytic tool for conducting real-time visualization of large-scale parallel executed SQL queries
- **Filled the patent: Real-time monitoring and analysis of query execution (PCT/US2014/032133)**

Software Engineer
OLPC project, NGO Paraguay Educa, Asunción, Paraguay *2008-2011*

- Contributed to the developed of open-source educational software, such as Poll and Labyrinth
- Developed the educational game Club de Othello XO (30,000 downloads)
- Developed an open-source fingerprint attendance system (9,000 downloads)

Front-End Developer
AISA, Asunción, Paraguay *2007*

- Collaborated in the implementation of the front-end of an ERP system that controls the selling, billings, transportations, and clients of the company

---

TECHNICAL SKILLS

Data Analytics and Machine Learning. Exploratory data analysis (correlation, location and variability estimation, visualization), statistical experiments and significance testing (hypothesis tests, t-tests, confidence intervals, ANOVA, chi-square, wilcox, kruskal-wallis), regression (simple linear regression, multiple linear regression), classification (logistic regression), clustering (k-means)

Software and Programming Languages. Python, SQL, R, iPython (pandas, scikit-learn, numpy), Javascript, HTML, CSS, Java

Data Visualization. ggplot2, jQuery Sparklines, JavaScript InfoVis Toolkit, Matplotlib

Scripting. Linux/Unix

PUBLICATIONS

Published **12 peer-reviewed papers** in relevant international conferences and journals in the following topics: Crowdsourcing, Collective Intelligence, E-Participation, Human-Computer Interaction, Virtual Reality, and Business Process Management (please refer to http://me.jorgesaldivargalli.com/publications for the complete list of my publications)