YASSER GONZALEZ

Website – http://yassergonzalez.com
Email – contact@yassergonzalez.com

SUMMARY

- Passionate about solving challenging problems that require developing scalable machine learning models and drawing conclusions from large datasets.
- Interested in areas of artificial intelligence and machine learning such as deep learning, natural language processing, reinforcement learning, optimization, evolutionary computation, and recommender systems.
- Bachelor's degree in computer science and a master's degree in information systems & technology with an emphasis on mathematical optimization.

EXPERIENCE

2018–present   **Research Scientist.** Amazon, Canada.

2017–2018   **Senior Data Scientist.** Canopy Labs, Canada.

- Led the development of personalization algorithms such as recommender systems and propensity models for marketing using technologies such as Spark, PyTorch, Keras, TensorFlow, scikit-learn, xgboost, Python and Scala.
- Researched the applications of deep learning and recurrent neural networks for predictive customer analytics.

2015–2017   **Data Scientist.** Canopy Labs, Canada.

- Developed large-scale recommender systems for diverse application domains – participating in the definition of the requirements and business metrics as well as the design, implementation and evaluation of the systems in production.
- Employed technologies such as Spark with Python and Scala running on Mesos and YARN, Amazon EMR & S3 and MongoDB.

2013–2015   **Research & Teaching Assistant.** York University, Canada.

- Designed and implemented reinforcement learning techniques to optimize the user interaction in a configuration process.
- Developed methodologies to improve the performance of search heuristics on multimodal optimization problems.
- Implemented scientific software in Python (using NumPy, pandas and SciPy), R and MATLAB/Octave.

2011–2013   **Research Assistant**. Institute of Cybernetics, Mathematics and Physics, Cuba.

- Developed new estimation of distribution algorithms (EDAs) using copulas and vines to model the probability distributions.
- Implemented a group of R packages available on CRAN and a C library for dependence modeling using vines.

EDUCATION

2013–2015   **Master's degree, Information Systems & Technology.**
York University, Canada.

- *Thesis:* Efficient Calculation of Optimal Configuration Processes.

- *Selected Courses:* Mining of Massive Datasets, Advanced Information Retrieval Systems, Introduction to Computational Linguistics.

2014–2015	**Data Science Specialization.** A non-credit series offered by Johns Hopkins University through Coursera.

- *Selected Courses:* R Programming, Getting and Cleaning Data, Exploratory Data Analysis, Statistical Inference, Regression Models, Practical Machine Learning.

2006–2011	**Bachelor's degree, Computer Science.** University of Havana, Cuba.

- *Thesis:* Estimation of Distribution Algorithms Based on Copulas and Vines.
- *Selected Courses:* Linear Algebra, Calculus, Probability & Statistics, Design & Analysis of Algorithms, Operating Systems, Computer Networks, Database Systems, Artificial Intelligence, Information Retrieval Systems.

SELECTED OPEN-SOURCE SOFTWARE

- Code used in Kaggle competitions. https://github.com/yasserglez/kaggle.
- Python package implementing different reinforcement learning methods for optimizing the user interaction in a configuration process. https://github.com/yasserglez/configurator.
- R package for implementing and studying estimation of distribution algorithms (EDAs) based on copulas. https://github.com/yasserglez/copulaedas.
- R implementation of the vine graphical models for representing high-dimensional probability distributions. https://github.com/yasserglez/vines.

For more information, please see http://yassergonzalez.com/software.

SELECTED PUBLICATIONS

- Y. Gonzalez-Fernandez, S. Chen. (2015). Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information. In *IEEE Congress on Evolutionary Computation*, 776–783. IEEE. http://dx.doi.org/10.1109/CEC.2015.7256970.
- Y. Gonzalez-Fernandez, M. Soto. (2014). copulaedas: An R Package for Estimation of Distribution Algorithms Based on Copulas. *Journal of Statistical Software*, 58(9), 1–34. http://www.jstatsoft.org/v58/i09.
- Y. Gonzalez-Fernandez, S. Chen. (2014). Identifying and Exploiting the Scale of a Search Space in Particle Swarm Optimization. In *Conference on Genetic and Evolutionary Computation*, 17–24. ACM. http://doi.acm.org/10.1145/2576768.2598280.
- M. Soto, A. Ochoa, Y. Gonzalez-Fernandez, Y. Milanés, A. Álvarez, D. Carrera and E. Moreno. (2012). Vine Estimation of Distribution Algorithms with Application to Molecular Docking. In S. Shakya and R. Santana (eds.), *Markov Networks in Evolutionary Computation*, 209–225. Springer. http://link.springer.com/chapter/10.1007/978-3-642-28900-2_13.

For more information, please see http://yassergonzalez.com/publications.