Montreal,QC - Canada
☎ +1 (780) 200-8010
✉ markosviggiato@gmail.com
markosviggiato.github.io/
linkedin.com/in/markosviggiato

Markos Viggiato

Highlights of Qualifications

- Proven record of collaboration with several research teams (papers [P2, P4]) and stakeholders
- Record of research projects using quantitative models with real-world data
- Technical expertise: NLP, statistical modelling, prediction and explanatory models, unstructured data analysis, cloud platforms (GCP, Snowflake), SQL

Education

Jan-2019–Present **PhD in Electrical and Computer Engineering**, *University of Alberta*, Edmonton, Canada
- Natural Language Processing/Understanding research. GPA: 4.0 (out of 4.0)

Mar-2017–Dec-2018 **MSc in Computer Science**, *Federal University of Minas Gerais*, Belo Horizonte, Brazil
- Machine Learning and Data Mining for Software Engineering. GPA: 9.0 (out of 10.0)

Mar-2011–Dec-2016 **Bachelor in Control and Automation Engineering**, *Federal University of Minas Gerais*, Belo Horizonte, Brazil. GPA: 7.6 (out of 10.0)

Industry Experience

Oct 2020–Present **NLP Intern**, *Prodigy Education*
I have investigated how to leverage ML/NLP techniques together with large-scale data to automatically extract knowledge from existing testing scenarios in natural language and to build a pipeline to automatically improve newly-created test descriptions **(Python, SQL)**

Jan-2016–Apr-2016 **Automation Engineering Intern**, *Centre Suisse d'Electronique et de Microtechnique - Brazil*
Autonomous system to collect and process solar energy-related data **(C, C++, Java)**

Research Experience

Jan-2019–Present **PhD Researcher**, *University of Alberta*
Research in applied Machine Learning and NLP **(Python, R, Java)**
- Investigated trade-off between statistical and neural language models to build a system that automatically analyzes test scenarios in natural language and provides improvement recommendations.
- Built a transformer-based (SBERT) technique to identify similarity in complex documents and achieved a performance of 83%.
- Modelled user behavior to build explainable win prediction models (XGBoost, Random Forest, Logistic Regression) for the Dota 2 game using the SHAP interpretability technique and achieved a performance of 86%.
- Implemented a sentiment analysis classification pipeline to analyze 12M of game reviews. Identified key problems that degrade the sentiment analysis performance, with a potential performance improvement of up to 11%.

Jan-2017–Dec-2018 **MSc Researcher**, *Federal University of Minas Gerais*
Research in machine learning and data mining for software engineering **(Python, Java, R)**

- Modelled the behaviour of developers using statistical and explanatory models and leveraged frequent itemset algorithms to identify co-evolution of changes in software.
- Built models to classify commits into maintenance activities using machine learning algorithms, which increased the state-of-the-art accuracy by 5%.
- Collaborated on a project to build explainable prediction models for software defects using XGBoost and SHAP values and improved the prediction accuracy by 15%.

Jan-2016–Dec-2016 **Undergraduate Researcher**, *Federal University of Minas Gerais*
Research in algorithm design, computational complexity and software quality **(Java, R, HTML, CSS)**

- Implemented efficient heuristics to configure software products.
- Investigated factors that affect software quality in e-commerce, health, and game industries.

Selected Publications

P1 **Using Natural Language Processing Techniques to Improve Manual Test Case Descriptions.** Markos Viggiato, Dale Paas, Chris Buzon, Cor-Paul Bezemer. *The 44th IEEE/ACM Conference on Software Engineering (ICSE 2022)*

P2 **Identifying Similar Test Cases That Are Specified in Natural Language.** Markos Viggiato, Dale Paas, Chris Buzon, Cor-Paul Bezemer. *arXiv* (Submitted to a journal)

P3 **Trouncing in Dota 2: An Investigation of Blowout Matches.** Markos Viggiato, Cor-Paul Bezemer. *The 16th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2020)*

P4 **What Causes Wrong Sentiment Classifications of Game Reviews?** Markos Viggiato, Dayi Lin, Abram Hindle, Cor-Paul Bezemer. *IEEE Transactions on Games*

Additional Information

Awards

- Alberta Innovates Graduate Student Scholarship (Jan 2020 - present). 3-year duration scholarship
- Alberta Graduate Excellence Scholarship (AGES) (Sep 2019)
- Early Career Researcher Award (Sep 2019) provided by the University of Alberta

Leadership positions

- Weekly seminar organizer in the Software and Machine Learning research laboratory during the master, 2017–2018
- Organizing member of the 6th Computer Science Summer School, *Federal University of Minas Gerais*, Brazil, 2017

Other Skills

- Experience with project management, git, SQL, bash script, Linux environment, Google Cloud Platform (GCP), Snowflake, Jupyter Notebook, sklearn, spacy, PyTorch, HuggingFace
- Experience with the following languages: Python, R, C/C++, Java