

João Fonseca

RESEARCHER · DATA SCIENTIST · LECTURER

☎ (+351) 911 989 441 | ✉ jpfonseca@novaims.unl.pt | 🏠 joaopfonseca.github.io | 📧 joaopfonseca | 🌐 jpmrfonseca | 📺 live:jpmrfonseca | ORCID | Google Scholar

RESEARCH INTERESTS: Machine Learning, Synthetic Data Generation, Explainable AI

Relevant Experience

Tandon School of Engineering (Center for Responsible AI) — New York University

POSTDOCTORAL RESEARCHER

New York, USA

Oct 2024 — Sep 2025

VISITING RESEARCH SCHOLAR

Jul 2023 — Sep 2024

NOVA Information Management School — NOVA University Lisbon

INVITED ASSISTANT PROFESSOR

Lisbon, Portugal

Sep 2023 — Sep 2024

Tandon School of Engineering (Center for Responsible AI) — New York University

RESEARCH INTERN

New York, USA

Mar 2023 — Jun 2023

NOVA Information Management School — NOVA University Lisbon

PHD CANDIDATE

Lisbon, Portugal

Sep 2021 — Jul 2023

INVITED TEACHING ASSISTANT

Sep 2019 — Feb 2023

MACHINE LEARNING RESEARCHER

Apr 2019 — Aug 2021

NOVA School of Business and Economics — NOVA University Lisbon

TEACHING ASSISTANT

Carcavelos, Portugal

Sep 2018 — Aug 2019

JUNIOR RESEARCHER

Out 2017 — Mar 2019

Data Science for Social Good Europe — Summer Fellowship

INFRASTRUCTURE SUPPORT

Lisbon, Portugal

Jun 2018 — Aug 2018

Education

NOVA Information Management School — NOVA University Lisbon

Lisbon, Portugal

PHD IN INFORMATION MANAGEMENT

MIT Portugal Scholarship — GRANT NUMBER SFRH/BD/151473/2021

Approved with Distinction and Honor (HIGHEST CLASSIFICATION)

Sep 2020 — Oct 2023

Dissertation title: THE ROLE OF SYNTHETIC DATA IN IMPROVING SUPERVISED LEARNING METHODS: THE CASE OF LAND USE/LAND COVER CLASSIFICATION

MSc. IN INFORMATION MANAGEMENT

SPEC. IN BUSINESS INTELLIGENCE AND KNOWLEDGE MANAGEMENT (GPA: 17.07 / 20.00)

Sep 2017 — Jan 2019

NOVA School of Business and Economics — NOVA University Lisbon

Lisbon, Portugal

MSc. IN MANAGEMENT — SPEC. IN DIGITAL BUSINESS (GPA: 16.80 / 20.00)

Sep 2016 — Jan 2019

BSc. IN ECONOMICS (GPA: 13.49 / 20.00)

Sep 2013 — Jun 2016

Universidad de Granada — Facultad de Ciencias Económicas y Empresariales

Granada, Spain

ERASMUS PROGRAM (GPA: 8.28 / 10.00)

Sep 2015 — Feb 2016

Scholarships, Awards & Achievements

Best AI Track Paper Award ([link](#))

Boston, MA, USA

ACM CONFERENCE ON EQUITY AND ACCESS IN ALGORITHMS, MECHANISMS, AND OPTIMIZATION (EAAMO'23)

Oct. 30th 2023

MIT Portugal PhD Grant Recipient

Lisbon, Portugal

MPP2030-FCT PHD GRANT, FOCUS ON DATA SCIENCE APPROACHES APPLIED TO EARTH SYSTEMS: OCEANS TO NEAR SPACE

Dec. 11th 2020

Publications

* - Equal contribution

16. Fonseca, J., Bhandari, A., Gilad, A., Kimelfeld, B., Roy, S., & Stoyanovich, J. (2025). **Understanding Fairness/Performance Trade-Offs in Synthetic Data Generation Using Integrity Constraints**. Working paper.
15. Hwang, D., Fonseca, J., Bell, A., Wang, S., & Stoyanovich, J. (2025). **Data Representation Attacks on SHAP-based Explanations**. (Under Submission)
14. Fonseca, J.*, Bell, A.*, & Stoyanovich, J. (2025). **Safeguarding Large Language Models in Real-time with Tunable Safety-Performance Trade-offs**. arXiv preprint arXiv:2501.02018. (Under Submission)
13. Bell, A.*, & Fonseca, J.* (2025). **Output Scouting: Auditing Large Language Models for Catastrophic Responses**. arXiv preprint arXiv:2410.05305. (Under Submission)
12. Bell, A.*, Fonseca, J.*, Abrate, C., Bonchi, F., & Stoyanovich, J. (2024). **Fairness in Algorithmic Recourse Through the Lens of Substantive Equality of Opportunity**. arXiv preprint arXiv:2401.16088. (Under Submission) [\(link\)](#)
11. Pliatsika, V., Fonseca, J., Wang, T., Stoyanovich, J. (2024). **ShaRP: Explaining Rankings with Shapley Values**. arXiv preprint arXiv:2401.16744. (Under Submission) [\(link\)](#)
10. Bell, A.*, Fonseca, J.*, & Stoyanovich, J. (2024) **The Game Of Recourse: Simulating Algorithmic Recourse over Time to Improve Its Reliability and Fairness**. In SIGMOD Conference Companion (pp. 464-467). [\(link\)](#)
9. Fonseca, J.*, Bell, A.*, Abrate, C., Bonchi, F., Stoyanovich, J. (2023). **Setting the Right Expectations: Algorithmic Recourse Over Time**. In Equity and Access in Algorithms, Mechanisms, and Optimization (pp. 1-11). **[Best AI Track Paper]** [\(link\)](#)
8. Fonseca, J., & Bacao, F. (2023). **Geometric SMOTE for Imbalanced Datasets with Nominal and Continuous Features**. Expert Systems with Applications, 234, 121053. [\(link\)](#)
7. Fonseca, J., & Bacao, F. (2023). **Tabular and latent space synthetic data generation: a literature review**. Journal of Big Data, 10(1), 115. [\(link\)](#)
6. Fonseca, J., & Bacao, F. (2023). **Improving Active Learning Performance Through the Use of Data Augmentation**. International Journal of Intelligent Systems, 2023. [\(link\)](#)
5. Fonseca, J., & Bacao, F. (2022). **Research Trends and Applications of Data Augmentation Algorithms**. arXiv preprint arXiv:2207.08817. [\(link\)](#)
4. Fonseca, J., Douzas, G., & Bacao, F. (2021). **Increasing the Effectiveness of Active Learning: Introducing Artificial Data Generation in Active Learning for Land Use/Land Cover Classification**. Remote Sensing, 13(13), 2619. [\(link\)](#)
3. Fonseca, J., Douzas, G., & Bacao, F. (2021). **Improving Imbalanced Land Cover Classification with K-Means SMOTE: Detecting and Oversampling Distinctive Minority Spectral Signatures..** Information, 12(7), 266. [\(link\)](#)
2. Crayton, A., Fonseca, J., Mehra, K., Ng, M., Ross, J., Sandoval-Castañeda, M., von Gnecht, R. (2021). **Narratives and Needs: Analyzing Experiences of Cyclone Amphan Using Twitter Discourse**, in Tackling Climate Change with Machine Learning Workshop at NeurIPS 2020. [\(link\)](#)
1. Douzas, G., Bacao, F., Fonseca, J.*, & Khudinyan, M.* (2019). **Imbalanced Learning in Land Cover Classification: Improving Minority Classes' Prediction Accuracy Using the Geometric SMOTE Algorithm**. Remote Sensing, 11(24), 3040. [\(link\)](#)

Talks

2024 ACM SIGMOD/PODS Conference [\(link\)](#)

BELL, A.*, FONSECA, J.*, STOYANOVICH, J. "THE GAME OF RECOURSE: SIMULATING ALGORITHMIC RECOURSE OVER TIME TO IMPROVE ITS RELIABILITY AND FAIRNESS" — DEMO PAPER PRESENTATION

Santiago, Chile

Jun. 11th 2024

ACM conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO'23) [\(link\)](#)

FONSECA, J.*, BELL, A.*, STOYANOVICH, J. "SETTING THE RIGHT EXPECTATIONS: ALGORITHMIC RECOURSE OVER TIME" — PAPER PRESENTATION [AI TRACK PAPER AWARD]

Santiago, Chile

Jun. 11th 2024

NeurIPS 2020 Workshop on Tackling Climate Change with Machine Learning ([link](#))

CRAYTON, A., FONSECA, J., MEHRA, K., NG, M., ROSS, J., SANDOVAL-CASTAÑEDA, M., VON GNECHT, R., "NARRATIVES AND NEEDS: ANALYZING EXPERIENCES OF CYCLONE AMPHAN USING TWITTER DISCOURSE"

Online

Dec. 11th 2020

DSSG Summit 2020 ([link](#))

FONSECA, J., DAVID, S., WINNING COMMUNITY CHALLENGE SUBMISSION PRESENTATION: "KEYBOARD LAYOUT OPTIMIZATION FOR ALS PATIENTS COMPETITION"

Online

Oct. 21st 2020

Solve for Good - DataFest 2020 ([link](#))

CRAYTON, A., FONSECA, J., MEHRA, K., NG, M., ROSS, J., SANDOVAL-CASTAÑEDA, M., VON GNECHT, R. "AMPHAN: ANALYZING EXPERIENCES OF EXTREME WEATHER EVENTS USING ONLINE DATA"

Online

Sep. 17th 2020

Data Science for Hospitality and Tourism conference ([link](#))

"UNLEASHING THE POWER OF BIG DATA", DEBATE PANEL MEMBER, RECTORATE OF NOVA UNIVERSITY OF LISBON

Lisbon, Portugal

Dec. 4th 2018

NOVA School of Business and Economics' Inauguration Ceremony ([link](#))

FONSECA, J., MESTRINHO, L., ZEJNILOVIC, L. "DATA-DRIVEN PLANNING FOR SUSTAINABLE TOURISM IN PORTUGAL", INVITED SPEAKER, NOVA SCHOOL OF BUSINESS AND ECONOMICS

Carcavelos, Portugal

Sep. 29th 2018

Research Projects

Tandon School of Engineering — New York University

New York, USA

AUDITING AND SAFEGUARDING LARGE LANGUAGE MODELS

Sep 2024 — Present

- Developed Output-Scouting, an open-source library designed to search for dangerous/unsafe outputs within Large Language Models' possible output space.
- Developed SafeNudge, a real-time safeguarding method against red teaming Large Language Models.

RECOURSE-GAME — MULTI-AGENT ALGORITHMIC RECOURSE OVER TIME ([LINK](#))

Mar 2023 — Present

FUNDED BY "MIT PORTUGAL", GRANT NUMBER SFRH/BD/151473/2021

- Recourse-Game contains the software implementation of the Multi-agent Algorithmic Recourse framework proposed in the paper "Setting the Right Expectations: Algorithmic Recourse Over Time."

EXPLAINABILITY IN RANKING

Mar 2023 — Present

- Proposed ShaRP (Shapley for Rankings and Preferences), a framework that explains the contributions of features to different aspects of a ranked outcome, and is based on Shapley values.
- Developing a methodology for comparison and benchmarking of explainability methods.

NOVA Information Management School — NOVA University Lisbon

Lisbon, Portugal

ML-RESEARCH — AN OPEN SOURCE LIBRARY FOR MACHINE LEARNING RESEARCH ([LINK](#))

Sep 2020 — Present

FUNDED BY "MIT PORTUGAL", GRANT NUMBER SFRH/BD/151473/2021

- ML-Research contains the software implementation of most algorithms used or developed in my research. It contains Active Learning and Over-sampling implementations (among others), as well as Datasets and several utilities to conduct research.
- Continuously developed throughout my research activities.

MAPINTEL — INTERACTIVE VISUAL ANALYTICS PLATFORM FOR COMPETITIVE INTELLIGENCE ([LINK](#))

Fev 2021 — Aug 2021

FUNDED BY "FUNDAÇÃO PARA A CIÊNCIA E TECNOLOGIA", GRANT NUMBER DSAIPA/DS/0116/2019

- Project developed with Professor Fernando Bação, PhD and Georgios Douzas, PhD. This work focused on the development of an interactive visual analytics platform for competitive intelligence.
- Usage of various Natural Language Processing and visualization techniques.

IPSTERS — IPSENTINEL TERRESTRIAL ENHANCED RECOGNITION SYSTEM. ([LINK](#))

Apr 2019 — Jan 2021

FUNDED BY "FUNDAÇÃO PARA A CIÊNCIA E TECNOLOGIA", GRANT NUMBER DSAIPA/AI/0100/2018

- Worked with Professor Fernando Bação, PhD. Exploration of applications and limitations of artificial intelligence algorithms to perform automatic updates of Land Use/Land Cover maps for the Portuguese mainland using satellite imagery.
- Exploration of new data filtering/anomaly detection algorithms, as well as dimensionality reduction techniques and different types of classification algorithms, including ensemble methods and deep learning classifiers.

DSSG Solve for Good — Data Science for Social Good Foundation

Online

AMPHAN - ANALYZING EXPERIENCES OF EXTREME WEATHER EVENTS USING ONLINE DATA ([LINK](#))

Jul 2020 — Nov 2020

VOLUNTEER RESEARCH PROJECT

- Integrated a team of four data scientists to research and develop a method to analyze social media discourse around natural disasters in developing countries in partnership with the International Water Management Institute (IWMI).
- Collaborative effort to develop data collection, preprocessing, feature extraction and analysis methods. Most of the tasks done were centered around the exploration of state-of-the-art Natural Language Processing algorithms.

NOVA School of Business and Economics — NOVA University Lisbon

Carcavelos, Portugal

RESEARCH + MASTER'S THESIS (FINAL GRADE: 19 / 20) ([LINK](#))

Oct 2017 — Mar 2019

- Worked with Professor Leid Zejniliovic, PhD. Used digital traces created by tourists to understand and improve the efficiency of tourism flows in Portugal, particularly specifically in the region of Lisbon.
- Studied the potential of Big data to inform destination management organizations using three sources of Big data: Telecom, Social media and Airbnb data.
- Developed a Social Media Crawler to fetch, store and visualize data of three major social media channels (Facebook, Twitter and Instagram).

Teaching

NOVA Information Management School — NOVA University Lisbon

INVITED ASSISTANT PROFESSOR

Sep 2023 — Present

- **Computation II** ([link](#)): Course coordinator (spring semester). Undergraduate course taught to Technology and Information Systems students.
- **Computation III** ([link](#)): Course coordinator (fall semester). Undergraduate course taught to Technology and Information Systems students.
- **Data Mining** ([link](#)): Practical sessions instructor (fall semester). Masters course taught to Information Management students.

NOVA Information Management School — NOVA University Lisbon

INVITED TEACHING ASSISTANT

Sep 2019 — Feb 2023

- **Business Cases with Data Science** ([link](#)): Instructed students in using a range of data science techniques to analyze and solve complex business challenges.
- **Data Mining** ([link](#)): Instructed students in exploratory analysis, data preprocessing (e.g., outlier detection, data imputation, and dimensionality reduction), data visualization, and clustering methods in Python through hands-on projects.
- **Descriptive Methods of Data Mining** ([link](#)): Instructed students in exploratory analysis, data preprocessing (e.g., outlier detection, data imputation, and dimensionality reduction), and clustering methods.
- **Web Analytics** ([link](#)): Instructed students in generating web analytics reports using Google Analytics, develop customer segmentation, and digital advertising.

NOVA School of Business and Economics — NOVA University Lisbon

TEACHING ASSISTANT

Sep 2018 — Aug 2019

- **Introduction to Programming** ([link](#)): Instructed students in Python programming through weekly practical classes, exams, office hours, and group project grading.

Skills

Python , Pandas, Geopandas, Numpy, Matplotlib, Plotly, Seaborn, PyTorch, Tensorflow, Keras, Scikit-Learn	<i>Advanced</i>
Bash , several CLI tools including git	<i>Advanced</i>
LaTeX	<i>Advanced</i>
Hadoop , PySpark, Impala, Hive	<i>Intermediate</i>
SQL	<i>Intermediate</i>
Dashboarding , MS Power BI, Qlik Sense, Apache Superset	<i>Intermediate</i>
Statistical Software , SPSS, SAS Enterprise Miner, Gephi, RapidMiner Studio, Orange	<i>Intermediate</i>
Web Design , HTML, CSS, Javascript	<i>Basic</i>

Personal

Citizenship: Portugal

Languages: Fluent in Portuguese, English and Spanish.

Interests: Guitar (Jazz, Flamenco, Blues and Rock), Surfing, Padel, Bouldering, Cycling