Breno L. S. de Almeida

🌐 Website ꧁ Google Scholar 🐙 brenoslivio 💼 brenoslivio ✉ brenoslivio@pm.me

SKILLS

PROGRAMMING
Python, R, Octave, Java, C, C++, SQL, Bash

PROGRAMMING TOOLS
Scikit-learn, Tensorflow, Pandas, Polars, Numpy, Scipy, NetworkX, Streamlit, Matplotlib, Plotly

SOFTWARE TOOLS
Jupyter Notebook, R Markdown, PostgreSQL, Git, Linux, LaTeX, Biopython, EMBOSS, Infernal, seqkit, CD-HIT

LANGUAGES
Brazilian Portuguese (Native),
English (IELTS: 6.5, exam taken on 23 March, 2024),
German (Elementary proficiency)

EDUCATION

PH.D. COMPUTER SCIENCE AND COMPUTATIONAL MATHEMATICS
UNIVERSITY OF SÃO PAULO

July 2024 - Present | São Carlos, Brazil
Research area: Artificial Intelligence - Machine Learning and Data Mining. Double Degree PhD Program with the University of Leipzig. Funded by FAPESP, Process Number #2024/10958-1.

B.SC. COMPUTER SCIENCE
UNIVERSITY OF SÃO PAULO

Mar 2020 - Dec 2023 | São Carlos, Brazil
Concluded certificate programs in data science and data engineering at the Institute of Mathematics and Computer Science. GPA: 9.4/10 (Rank: 7/116). Funded by multiple FAPESP scholarships, Process Number #2021/08561-8, #2023/00264-0 and #2024/03304-5.

ENGINEER'S DEGREE, COMPUTER ENGINEERING
UNIVERSITY OF SÃO PAULO

Mar 2017 - Dec 2019 | São Carlos, Brazil
Transitioned to Computer Science, applying credits from prior coursework.

COMPLEMENTARY EDUCATION

- Applied Data Science with Python Specialization - Coursera (Online). University of Michigan.
- Data Science Specialization - Coursera (Online). Johns Hopkins University.
- Optional class in Bioinformatics at the University of São Paulo.

EXPERIENCE

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ | RESEARCH ASSISTANT
July 2024 - Present | Leipzig, Germany

- Collaborating with UFZ since 2021. Officially assumed the role of Research Assistant in the Microbial Data Science group at the Department of Computational Biology and Chemistry upon starting PhD. In-person research activities will commence in July 2025.

FUDAN UNIVERSITY | RESEARCH INTERN
July 2022 – Sep 2022 | Remote

- Research Internship at the Big Data Biology Lab, Fudan University in Shanghai, headed by Dr. Luis Pedro Coelho. Worked with large data sets for the rarefaction of multiple environments regarding the richness of unique smORFs. The work can be consulted here: Global Microbial Small ORFs Catalog - Gene rarefaction for smORFs.

INSTITUTE OF MATHEMATICS AND COMPUTER SCIENCES (ICMC) | TEACHING ASSISTANT
May 2021 – Nov 2022 | São Carlos, Brazil

- Assisted in four undergraduate courses: Object-Oriented Analysis and Design, Statistics, Fundamentals of Data Science, and Computational Visualization. Responsibilities included grading assignments, preparing exams, and delivering practical lectures to support student learning.

EFESTUS | CO-FOUNDER AND COORDINATOR
June 2018 – Feb 2021 | São Carlos, Brazil

- Student organization focused on computer science and data science for social good, especially on debating the ethical implications of technological advances. Fully funded participation in events such as TechShift Summit, Chapel Hill, NC (USA).

AWARDS AND RECOGNITION

- **J. F. Marar Artificial Intelligence Prize for Undergraduate Research**: Winner project "BioDeepFuse: Empowering Researchers in Life Sciences with Deep Learning". Project was selected winner considering criteria such as academic and social relevance, adherence to the United Nations' Sustainable Development Goals (SDGs), and academic performance. More details in Portuguese here: News shared by the University of São Paulo.
- **Prototypes for Humanity**: BioAutoML was selected by Prototypes for Humanity as one of the top 100 projects from nearly 3,000 applications submitted by graduates from over 100 countries. The project received funding for participation and was presented at the international festival in Dubai. More information is available at: BioAutoML on the event website.

JOURNAL PAPERS

- Bonidia, R. P., Santos, A. P. A., **de Almeida, B. L.**, Stadler, P. F., da Rocha, U. N., Sanches, D. S., & de Carvalho, A. C. (2022). BioAutoML: automated feature engineering and metalearning to predict noncoding RNAs in bacteria. Briefings in Bioinformatics, 23(4), bbac218.
- Bonidia, R. P., Avila Santos, A. P., **de Almeida, B. L.**, Stadler, P. F., Nunes da Rocha, U., Sanches, D. S., & De Carvalho, A. C. (2022). Information theory for biological sequence classification: A novel feature extraction technique based on Tsallis entropy. Entropy, 24(10), 1398.
- Duan, Y., Santos-Júnior, C. D., Schmidt, T. S., Fullam, A., **de Almeida, B. L.**, Zhu, C., ... & Coelho, L. P. (2024). A catalog of small proteins from the global microbiome. Nature Communications, 15(1), 7563.
- Avila Santos, A.P., **de Almeida, B.L.**, Bonidia, R.P., Stadler, P.F., Stefanic, P., Mandic-Mulec, I., Rocha, U., Sanches, D.S. and de Carvalho, A.C., 2024. BioDeepfuse: a hybrid deep learning approach with integrated feature extraction techniques for enhanced non-coding RNA classification. RNA biology, 21(1), pp.1-12. **(Shared first authorship)**

CONFERENCE PAPERS

- **de Almeida, B. L. S.**, Queiroz, A. P., Santos, A. P. A., Bonidia, R. P., da Rocha, U. N., Sanches, D. S., & de Carvalho, A. C. P. D. L. F. (2021). Feature importance analysis of non-coding dna/rna sequences based on machine learning approaches. In Advances in Bioinformatics and Computational Biology: 14th Brazilian Symposium on Bioinformatics, BSB 2021, Virtual Event, November 22–26, 2021, Proceedings 14 (pp. 81-92). Springer International Publishing.

CONFERENCE ABSTRACTS

- Santos, A. P. A., **de Almeida, B. L.**, Bonidia, R. P., Stadler, P. F., da Rocha, U. N., Sanches, D. S., & de Carvalho, A. C.. Feature Extraction Techniques based on Deep Learning Framework for Enhanced Classification of Non-coding RNA. In: 16th symposium on Genetics and Bacterial Ecology 2023 (BAGECO 2023), 2023, Copenhagen. New approaches/technologies in microbial ecology, 2023.
- Bonidia, R. P., Santos, A. P. A., **de Almeida, B. L.**, Stadler, P. F., da Rocha, U. N., Sanches, D. S., & de Carvalho, A. C.. BioAutoML: End-to-End Machine Learning Package for Life Sciences. In: 10th FEMS Congress of European Microbiologists, 2023, Hamburg. 10th FEMS Congress of European Microbiologists, 2023.
- **de Almeida, B. L.**, Kamath S., Santos, A. P. A., Bonidia, R. P., Stadler, P. F., Wierckx N., Borchert E., Rother D., de Oliveira A., Sanches, D. S., de Carvalho, A. C., & Ulisses Rocha. HiPfam: Expanding the universe of known protein using large language models. In: Helmholtz AI Conference, 2024, Düsseldorf. Helmholtz AI Conference, 2024.
- Santos, A. P. A., **de Almeida, B. L.**, Bonidia, R. P., Stadler, P. F., Sanches, D. S., de Carvalho, A. C., & da Rocha, U. N.. BioMeta-Explorer: democratizing the analysis of non-Coding RNA with multi-omics and data-centric AI. In: Helmholtz AI Conference, 2024, Düsseldorf. Helmholtz AI Conference, 2024.