

Ruan Chaves Rodrigues

 ruanchaves93@gmail.com

 github.com/ruanchaves

 linkedin.com/in/ruanchaves

Education

- **University of the Basque Country**
Erasmus Mundus Master in Language and Communication Technologies
Donostia, Spain
September 2021 – Present
 - **Federal University of Goiás**
Bachelor's degree, Computer Science
Goiânia, Brazil
January 2017 – June 2021
-

Experience

Recognai

Data Science Intern

Madrid, Spain

November 2021 – Present

- Contributed to improve the support for data-centric weak supervised learning on [Rubrix](#), an open source NLP library for data annotation and management.

NeuralMind

Data Scientist

Campinas, Brazil

February 2021 – May 2021

- Developed domain-specific search engines by combining Transformer models with data augmentation and hyperparameter optimization techniques.
- Achieved performance improvements over previous state-of-the-art technologies, resulting in publications in the field of Information Retrieval.

AI Center of Excellence (CExIA) & Deep Learning Brasil

Researcher

Goiânia, Brazil

August 2020 – June 2021

- Investigated how recent research findings in NLP could be applied to chatbots and other sorts of conversational agents in the Portuguese language.
 - Applied techniques such as data augmentation, domain adaptation and multi-task learning to deal with data scarcity in low-resource domains across several NLP tasks.
 - Contributed to improve the internal conversational systems of one of the top 10 energy trading companies in Brazil.
-

Accomplishments

- ◆ **1st place** at two international Textual Entailment competitions: [COLIEE 2021](#) (NeuralMind team) [[GitHub](#)] and [ASSIN 2](#) (Deep Learning Brasil team) [[GitHub](#)].
 - ◆ **Erasmus Mundus Scholarship** associated with the Erasmus Mundus Master in Language and Communication Technologies (EMLCT).
 - ◆ **Outstanding Student Award** by the Institute of Informatics (INF/UFG) and a **Certificate of Honor** by the University Council of the **Federal University of Goiás** (UFG) in recognition of my academic achievements.
-

Projects

- ◆ Created the [Hashformers](#) framework for hashtag segmentation, which has been [independently audited](#) and confirmed to consistently outperform the previous state-of-the-art.
 - ◆ [Fixed a severe bug](#) on the evaluation procedure of widely utilized Portuguese word embeddings and [published my findings](#) at the main conference in the field of Portuguese NLP, [PROPOR 2020](#).
-

Skills

Python, PyTorch, Tensorflow, Google Cloud Platform (GCP), Natural Language Processing libraries (NLTK, Gensim, spaCy, transformers), Data Science libraries (Pandas, NumPy, scikit-learn), Docker, Linux, Git, Bash