Andrés **Cabero Busto**

MACHINE LEARNING ENGINEER · SOFTWARE ENGINEER

Stuttgart, Germany

☎ +4917684048732 | ✉ cabero96@protonmail.com | ⬡ cbr9 | 💼 andres-cabero-busto

Skills

Data Science, AI & ML	Pandas, Plotly, Matplotlib, SQL, SLURM, W&B, PyTorch, PyTorch Lightning, LangChain, HuggingFace, Captum, SpaCy, NLTK
Programming	Python, Rust, JavaScript, TypeScript, HTML, CSS, LaTeX
Languages	Spanish, English, Italian, German
Other	Docker, Linux, NixOS, Nix, Bash, CLI, Git

Experience

Sony Research   *Stuttgart, Germany*
MASTER THESIS STUDENT   *Apr. 2023 - Oct. 2023*

- Performed domain-adaptation on an open-source, self-supervised (SSL) speech recognition model.
- Applied regularization-based and replay-based continual learning methods to avoid incurring into catastrophic forgetting on the old Librispeech ASR task.

Sony Research   *Stuttgart, Germany*
WORKING STUDENT   *Jul. 2022 - Mar. 2023*

- Developed a natural language grammar based on weighted finite-state transducers (WFSTs) and the Thrax framework, similar to NVIDIA NeMo grammars, that enabled a 10% WER improvement for the pre-existing Italian ASR system.
- Time-aligned and transcribed speech utterances using Label Studio for use in text-to-speech (TTS) model training.

University of Stuttgart   *Stuttgart, Germany*
PYTHON DEVELOPER   *Jul. 2022 - Dec. 2022*

- Developed the core of a toolkit for academic NLP research in lexical semantic change detection (LSCD), openly available at `https://github.com/ChangeIsKey/LSCDBenchmark`

University of Stuttgart   *Stuttgart, Germany*
LINUX SYSTEM ADMINISTRATOR   *Jul. 2022 - Mar. 2022*

- Administration and maintenance of Linux servers and users.

University of Stuttgart   *Stuttgart, Germany*
DATA ANNOTATOR   *Nov. 2021 - Jan. 2022*

- Data annotator for the first Shared Task of Lexical Semantic Change Discovery in Spanish (`https://fdzr.github.io/lscdiscovery/`). The process involved annotating the degree of difference between uses of the same word based on the surrounding context, using a Likert scale.
- Identified some problems in the annotation platform, and developed a small JavaScript solution to help my teammates and I work more efficiently.

Projects

Model Compression using Explainable AI (XAI)   *Stuttgart, Germany*
INSTITUT FÜR MASCHINELLE SPRACHVERARBEITUNG (UNIVERSITY OF STUTTGART)

- Compressed a Bi-LSTM trained for a speech emotion recognition task.
- Using XAI methods, we realized that more than half of the input features were completely irrelevant, thus allowing us to reduce the number of parameters of the network.
- Because we were able to reduce the parameters in this way, this also allowed us to experiment with other hyperparameters more efficiently.
- The final model preserved the exact same accuracy of the original model, but with more than half of its parameters.

Education

M.Sc. Computational Linguistics
Institut für Maschinelle Sprachverarbeitung (University of Stuttgart)

Apr. 2021 - Oct. 2023

Stuttgart, Germany

- Achieved an in-depth understanding of modern deep learning (DL) and natural language processing (NLP) techniques.

B.A. Linguistics
University of Cádiz

Sept. 2016 - Dec. 2020

Cádiz, Spain

- Took several courses related to programming, natural language processing (NLP) and computer science (CS), achieving honors in several of them.

Certifications

PyTorch For Deep Learning
HTTPS://WWW.UDEMY.COM/CERTIFICATE/UC-C8142260-A945-4F41-91F5-D213671B8493/

Apr. 2022

Online

Mathematics For Machine Learning
HTTPS://WWW.COURSERA.ORG/ACCOUNT/ACCOMPLISHMENTS/SPECIALIZATION/CERTIFICATE/DJPXWT9X9KVS

Mar. 2022

Online

Neural Networks and Deep Learning
HTTPS://WWW.COURSERA.ORG/ACCOUNT/ACCOMPLISHMENTS/CERTIFICATE/JXLQNFEH3H9S

Dec. 2021

Online

Linear Algebra
HTTPS://WWW.UDEMY.COM/CERTIFICATE/UC-0D00A3F9-2A0C-4881-A50E-4341FA30B055/

Mar. 2020

Online

Probability & Statistics
HTTPS://WWW.UDEMY.COM/CERTIFICATE/UC-BHG9O9SJ/

Mar. 2020

Online