

Jacobo Del Valle

linkedin://jjdelvalle • github://jjdelvalle • Aachen, Germany
https://jjdv.xyz • delvallejacob@gmail.com

Experience

- **Amazon Inc.** **Aachen, Germany**
July 2022 – Present
Applied Scientist Intern
 - Interned at the Alexa AI NLU team over the summer with the task of adapting large NLP models to more constrained scenarios without having to fine-tune fully or at all
 - Guide models in a constrained NLG scenario in the context of Alexa AI
 - **HealthCare.com** **Remote (US-based Company)**
Apr 2020 – May 2022
Data Scientist (Part-time since Sep 2021)
 - Developed an NLP model to predict the intent behind a search query and/or keywords using TF-IDF vectorization, stemming, and clustering techniques
 - Researched, designed, and implemented special vocabulary and tokenization techniques that allowed cutting the vocabulary size by ~70% without impacting performance
 - Helped develop a recommendation system and its A/B testing infrastructure by tracking events for each different experiment and saving them on S3
 - Optimized API endpoint by modifying queries and implementing Redis caching systems
 - API creation for model endpoints using FastAPI and Docker to deploy
 - **Plus Technologies & Innovations S.A.** **Guatemala, Guatemala**
Oct 2016 – Apr 2020
Backend Developer & Data Scientist
PlusTI is a leading software provider for financial institutions all around the globe
 - Developed, optimized (using parallelization and Cython) and maintained pre-processing, clustering and data exploration scripts in both R and python
 - Application and regularization of Gradient Boosting algorithms in R
 - Deep neural network design using Tensorflow and hyperparameter optimization using Bayesian methods
 - AWS EC2 and S3 resource management, cutting down costs by over 50%
-

Education

- **Charles University, Saarland University** **Czech Republic/Germany**
(Anticipated) Sep 2023
M.Sc. Computer Science (NLP Focus); EMLCT double degree scholarship recipient
 - **Galileo University** **Guatemala, Guatemala**
Nov 2019
B.Sc. Computer Science (Magna Cum Laude Honors)
-

Core Technical Skills

Languages: Proficient in Python, R, Java, SQL; acquainted with PHP

Applications: VSCode, Jupyter Notebooks, RStudio, tensorflow, pytorch, git, Airflow, Redis, VIM, L^AT_EX, svn

Open Source Contributions: ClosedXML, dolphin-emu, fastai, streamlit

Projects

LEFTy (Language Efficient Text Portray): Use fastai and pytorch to apply state of the art NLP transfer learning techniques to address the task of author profiling in multiple languages

GradStats: Built a web app using *streamlit* that allows student to explore historical graduate admissions data. Data was gathered by scraping TheGradCafe's website

Languages & Extracurricular Activities

Native Languages: Spanish, English (TOEFL score of 119/120)

Foreign Languages: German (TestDaF C1 Certificate)

GRE, and TOEFL tutor on online forums and through personal blog