Miguel Â. Simões Valente

http://www.miguelvalente.xyz

Email : miguelvalente@protonmail.com
Mobile : +31 681 022 805

Education

- **Utrecht University** — Utrecht, Netherlands
  Master of Science in Artificial Intelligence — *Sep. 2019 - Feb. 2022*
  - **Thesis**: Generative Based Zero-Shot Learning: Classifying Images from Text
- **Instituto Politécnico de Castelo Branco** — Castelo Branco, Portugal
  Bachelor of Computer Engineering — *Aug. 2016 – Jun. 2019*
  - **Thesis**: Compendium of three publications in applied object detection. *Extra information and relevant links to publications available on my personal website.*

Experience

- **Oddity.ai** — Utrecht, Netherlands
  Deep Learning Developer — *May 2022 - December 2022*
  - **Deep Learning Modeling**: Architected custom models in PyTorch and Tensorflow using CNNs, RNNs, and self-attention as building blocks. Leveraged pre-trained weights for different styles of transfer learning from CNN and transformer-like architectures. Researched and applied novel approaches for action recognition.
  - **Model Training & Experimentation**: Introduced automatic experiments to decrease timesinks on hyperparameter searches while providing an overview of results. Introduced experimentation best practices to reduce friction between colleagues' workflow and ensure model lineage.
  - **Computer Vision & Image Data**: Created datasets for action recognition. Experimented with different approaches for data augmentation, on the image and video level.
  - **Model Optimization/Testing**: Exported models to ONNX and optimized them with TensorRT. TensorRT models were combined with a RUST backend to speed up inference time.
- **TNO(Netherlands Organization for Applied Scientific Research)** — The Hague, Netherlands
  Research Intern — *Feb 2021 - Oct 2021*
  - **Research**: Zero-shot image classification through text embeddings.
  - **Nutshell**: Encoding of noisy text descriptions into a per-class embedding representation. Trained Normalizing Flows to generate images conditioned on the class embeddings. Generated images for classes without visual examples. Train a classifier on the generated images and real images to obtain zero-shot capacities.
  - **Methodology**: Work was done on vector space with the sole intent of classification. Images were encoded with pre-trained networks from the timm library. The text was encoded with TF-IDF and recent models from the HuggingFace library. The image data comprised of ImageNet and the text was sourced from Wikipedia.

Projects & Certifications

- **Whisperer**: Automatic text-audio dataset maker with Audio Deep Learning Models. Does: Transcription, diarization, and ids speakers across audio files with hierarchical clustering
- **MLOps Specialization**: Machine Learning Engineering for Production (MLOps) Specialization
- **CUDA NVIDIA Specialization**: Fundamentals of Accelerated Computing with CUDA Python
- **Greenhouse**: Application of data science methods on time series for weather prediction and next-day electricity consumption forecasting.

Programming Skills

- **Languages:** Python, SQL, Latex, Rust(entry-level)
- **Technologies:** PyTorch, TensorFlow, ONNX, TensorRT, HuggingFace, Weights&Biases, Scikit-Learn, Pandas, Scipy, Docker, GitHub Actions, GNU/Linux, Typer, FastAPI, Streamlit
- **Machine Learning Domains:** Predominantly worked with unstructured data, but familiar with structured data problem solving