

Simone Viozzi

Deep Learning Engineer

✉ simoneviozzi97@gmail.com

☎ +39 3891575096

📍 Torino, Italy

🐙 [Github](#)

🌐 [LinkedIn](#)



Profile

Highly skilled and innovative Deep Learning Engineer with a strong background in computer vision and natural language processing. Equipped with a deep understanding of cutting-edge deep learning techniques and frameworks, such as PyTorch, ONNX, and Huggingface transformers. Passionate about developing state-of-the-art solutions to complex challenges, leveraging large language models and advanced computer vision architectures. Proven track record of success in training and deploying deep learning models, optimizing performance, and delivering exceptional results. Effective collaboration with international teams to successfully drive projects to completion. Committed to staying at the forefront of emerging technologies and continuously expanding expertise in deep learning and AI.

Professional Experience

Senior AI/ML Computational Science Analyst

Accenture [🔗](#)

12/2023 – present

Torino, Italy

Data Engineer

Software Inside SRL

09/2023 – 12/2023

Italy

In my role as a Data Engineer, I designed and implemented multiple Extract, Load, Transform (ELT) pipelines using Databricks notebooks for processing and Azure Data Factory pipelines for orchestration.

I employed Databricks, Python, PySpark, and SQL to manipulate and transform intricate datasets, ensuring they aligned with specific project requirements.

Additionally, I utilized Data Factory to coordinate the ingestion of data from various Linked Services, processed it through Databricks, and deposited the results into another Linked Service.

Deep Learning Engineer

Cloe Ai GmbH [🔗](#)

10/2021 – 06/2023

Germany

In my role as a Deep Learning Engineer, I conducted extensive research and leveraged advanced deep learning techniques to develop state-of-the-art products, effectively addressing complex challenges faced by third-party clients.

Key accomplishments and responsibilities include:

- Trained an Instance Segmentation model using the detectron2 framework.
- Co-developed a synthetic damage generator in a team setting, to simulate realistic damages on undamaged cars, thereby creating a synthetic dataset of car damages. The synthetic dataset was utilized for pre-training an Instance Segmentation model.

- Led and supervised a team of international annotators, overseeing an image classification and instance segmentation project. Implemented robust quality control measures, including comprehensive documentation and weekly review sessions with annotators, to ensure the accuracy and consistency of annotations.
- Designed, trained, and deployed a custom U-NET model as a CVAT (Computer Vision Annotation Tool) interactor, significantly improving image annotation workflows and achieving a five-fold increase in annotation speed. The model was developed using pure PyTorch and subsequently packaged into a Docker container with a standardized API interface for seamless integration with the CVAT instance.
- Independently designed, implemented, and deployed a high-performance Text Classification pipeline leveraging SetFit, a state-of-the-art Natural Language Processing library for few-shot text classification. Demonstrated exceptional accuracy, exceeding 70%, on a production dataset of less than 10,000 samples. Seamlessly integrated the pipeline into the company platform, empowering clients to efficiently perform text classification tasks and independently train and deploy models.
- Leveraged large language models, such as OpenAI completion models and ChatGPT, to develop a zero-shot text classification pipeline that achieves high-performance without the need for any training.

Technical Skills:

- Deep learning frameworks: Proficient in PyTorch and ONNX for developing and training deep learning models.
- Computer Vision: Experienced in utilizing OpenCV for image processing tasks. Familiar with Unet model architecture for semantic segmentation and Mask R-CNN model architecture for instance segmentation tasks.
- Natural Language Processing: Skilled in working with the Huggingface transformers library for various NLP tasks. Familiar with the OpenAI API structure for language model interactions.
- Model Deployment: Proficient in ONNX compression techniques for optimizing model size and performance. Experienced in using Docker and Docker-compose for containerization and deployment of machine learning models. Familiar with FastAPI for building efficient and scalable API interfaces.

Projects

Mask R-CNN training with docker containers on Sagemaker

05/2021 – 07/2021

Associated with Università Politecnica delle Marche

As part of a university project, I collaborated with a team of two to train a Mask R-CNN model using Docker containers on Amazon SageMaker. Our objective was to leverage the Matterport implementation of Mask R-CNN, proposed by Ross Girshick et al., with enhancements such as the Feature Pyramid Network and ResNet101 backbone.

Key Contributions and Accomplishments:

- Prepared and curated a custom dataset for training the Mask R-CNN model, ensuring its suitability for the specific application domain.
- Developed functions to adapt the dataset annotations from the Supervisely structure to the input format required by the Matterport implementation.
- Created a script that facilitated the training process by launching a Docker container on Amazon SageMaker, effectively utilizing the infrastructure capabilities of the cloud service.

- Collaborated with the team to optimize the training workflow, considering the specific requirements and limitations of SageMaker.
- Actively engaged in troubleshooting and debugging to overcome challenges encountered during the training process.
- Open-sourced the project, making it accessible to the wider deep learning community, and associated it with the university as part of an exam.

This project allowed us to gain hands-on experience with training advanced computer vision models, containerization using Docker, and utilizing cloud infrastructure for scalable and efficient model training. By successfully training the Mask R-CNN model, we demonstrated our ability to apply cutting-edge deep learning techniques to address real-world challenges. The project also highlighted our collaboration and problem-solving skills within a team environment.

Education

Bachelor's degree, Computer and automation Engineering 2016 – 2022

Università Politecnica delle Marche [↗](#)

Grade: 106/110

Diploma Istituto Tecnico e Professionale, Elettronica 2011 – 2016

Istituto Tecnico Tecnologico Statale "G. e M. Montani" [↗](#)

Grade: 86/100

Courses with Certification

Deep Learning Specialization [↗](#) 07/2020

Coursera

The Deep Learning Specialization is a comprehensive program that equips you with the theoretical knowledge and practical skills to understand and contribute to the cutting-edge field of deep learning, covering various neural network architectures and their applications in industries such as speech recognition, chatbots, and natural language processing.

Natural Language Processing Specialization [↗](#) 09/2020

Coursera

This Natural Language Processing (NLP) Specialization focuses on the technical aspects of using algorithms to interpret and manipulate human language, enabling the design of NLP applications for question-answering, sentiment analysis, language translation, text summarization, and chatbot development.

Generative Adversarial Networks (GANs) Specialization [↗](#) 03/2021

Coursera

This GANs Specialization course provides a comprehensive and hands-on introduction to generating realistic images, videos, and voice outputs using Generative Adversarial Networks (GANs), while also covering technical topics such as cybersecurity, data anonymization, image processing, and more.

DeepLearning.AI TensorFlow Developer Professional Certificate [↗](#)

08/2020

Coursera

This technical course focuses on building and training neural networks using TensorFlow, improving network performance with convolutions, natural language processing, and text processing.

Reinforcement Learning Specialization [↗](#)

09/2020

Coursera

The Reinforcement Learning Specialization explores the power of adaptive learning systems and AI, teaching learners to implement complete RL solutions and apply them to various real-world problems such as game development, customer interaction, and finance.

Google Cloud Computing Foundations [↗](#)

12/2022

Google Cloud Skills Boost

The Google Cloud Computing Foundations course provides a comprehensive introduction to Google Cloud, covering cloud computing fundamentals, infrastructure, networking and security, and data, machine learning, and AI in Google Cloud.

The Complete SQL Bootcamp: Go from Zero to Hero [↗](#)

09/2023

Udemy

This course covers comprehensive SQL skills, including PostgreSQL, SQL syntax, aggregate functions, JOIN commands, and database creation with constraints, enhancing data analysis and Python integration.

Languages

Italian



English



Skills

Python

Pandas, Matplotlib, Numpy



Deep Learning

PyTorch, TensorFlow, ONNX, Convolutional Neural Networks (CNN), Huggingface transformers, Generative Adversarial Networks (GANs)



Computer Vision

OpenCV, UNet, Mask R-CNN, Instance Segmentation, Semantic Segmentation, Image Processing



Natural Language Processing (NLP)

Huggingface transformers, OpenAI API, Text Classification, Large Language Model Interactions



System Administration

Bash scripting, File system management, SSH tunnels



Model Deployment

ONNX Compression, Docker, Docker-compose, FastAPI, API Interfaces



GitHub

Pull requests, PR review, Semantic versioning, Semantic commits



Data Engineer

Databricks, SQL, PySpark



DevOps

Amazon SageMaker, Cloud Infrastructure, Containerization with Docker



GDPR Consent

I authorize the processing of personal data contained in my curriculum vitae based on art. 13 GDPR 679/16.

Simone Viozzi