Silvio Pavanetto

silviopavanetto.com

Email : silvio.pavanetto@gmail.com
Github : github.com/reidemeister94

EXPERIENCE

Jobtome   Stabio, Switzerland
Software Engineer   *11/2021 - Present*

Developed and maintained **search APIs** written in **Python** that query **Elasticsearch**, capable of handling **1000 requests per second**.

Created and managed pipelines using **Apache Beam** for ingesting and processing **millions of records per day**, from **Google Pub/Sub to BigQuery**.

Developed a classification model using **NLP** and **deep learning** techniques (**BERT, CNN**) and deployed it on **Google Vertex**. The model processes approximately **1 million jobs per day with an accuracy of 95%.**

Created a software architecture in **Python** for a **job offers crawler** consisting of over **500** crawlers for different web pages. Scheduled and executed in production using **Docker, Airflow, and Kubernetes**, capable of collecting **100k jobs per day**.

Polytechnic University of Milan   Milan, Italy
Research Fellow   *10/2019 - 10/2021*

Monitored **3 millions of Italian conversations around vaccines on social media** (Twitter, Facebook) and understand the interplay between online public discourse and vaccine hesitancy/uptake rates.

Developed a **data pipeline** for ingesting, analyzing and producing insights from **hundreds of online reviews and social media posts per day**, from **100 Italian museums**.

Deloitte Digital   Milan, Italy
Software Developer   *02/2019 - 08/2019*

Developer role on international cloud projects using **Salesforce** technology, which includes several programming languages such as **Java** and **Javascript**.

EDUCATION

Polytechnic University of Milan   Milan, Italy
M.Sc. in Computer Science and Engineering; Final grade: 110 with honors/110   *09/2016 – 04/2019*

Polytechnic University of Milan   Milan, Italy
Bachelor Engineering of Computing Systems   *09/2013 – 09/2016*

SELECTED PUBLICATIONS

- **ICWSM 2021 - Online**: *A Content-based Approach for the Analysis and Classification of Vaccine-related Stances on Twitter: the Italian Scenario;* Collection and analysis of 3 millions of Italian conversations about COVID-19 vaccines on Twitter, investigating the geographical, temporal and lexical distribution of data. Trained a binary classifier that predicts the stance of tweets towards vaccines. **Link**
- **ICWE 2020 - Helsinki, Finland**: *Generation of Realistic Navigation Paths for Web Site Testing using Recurrent Neural Networks and Generative Adversarial Neural Networks.* Method for generating high-quality weblog data using deep learning techniques, involving comparison of results with a suite of data mining algorithms as a baseline. **Link**

SKILLS

- **Languages:** Python, Java, Javascript, C++

- **Technologies:** Google Cloud Platform (GCP), Docker, Kubernetes, Git, Airflow, FastAPI, Elasticsearch, Pandas, Numpy, Matplotlib, Scikit-Learn, Tensorflow, Keras.

- **Concepts:** Data engineering, data pipelines, machine learning, NLP, distributed systems, microservices, cloud computing, agile development, test-driven development (TDD), continuous integration/continuous delivery (CI/CD).