GIOVANNI BRUNETTI

Big Data and Software Engineer

SUMMARY

I'm a computer scientist based in Italy focused on working in the Big Data context. I love writing code and solving problems with it!
I have been a Big Data Engineer since 2020 across multiple domains: **energy**, **telecommunications** and **insurance**.

SKILLS

Technicals: GCP (BigQuery, Composer, DataFlow, BigTable, ecc), Cloudera Ecosystem, Scala, SQL, Python, Scala, Bash, Docker, Kubernetes, Spark, Hadoop, Hive, Jenkins, Airflow, Kafka, Sonar-Qube, Git, Maven, Sbt, Functional Programming, HDFS, Team Tools and methodologies (Jira, Confluence, Scrum)

Soft Skills: Problem Solving, Proactive, Team Working, Passion, Time Management, Challenge Seeker

EXPERIENCE

10/2022 – NOW   **Big Data and Software Engineer**   **Agile Lab**

- Consultant for a major Italian **insurance** company:
  - Data modelling and data partitioning in a Cloudera Platform ecosystem.
  - Defining and developing Spark job using Scala for Ingestion and processing real-time and batch data.
  - Development of NiFi pipeline to orchestrate jobs.
    MAIN TECHNOLOGIES: **SCALA, SPARK, HIVE, IMPALA, SQL, POSTGRESSQL, HBASE, HDFS, BASH, SSH, JENKINS, KAFKA**

01/2020 – 10/2023   **Big Data and Software Engineer**   **VT Solutions & Consulting**

- Consultant for a major German **telecommunications** company:
  - Data modeling and data partitioning in a Google Cloud Platform ecosystem.
  - Defining and developing DataFlow job using Scala for Ingestion and processing real-time and batch data.
  - Development of Airflow DAGs using Python to orchestrate ETL operations on data for a DWH.
    MAIN TECHNOLOGIES: **SCALA, PYTHON, BIG QUERY, COMPOSER, SQL, POSTGRESSQL, BIG TABLE, GCS, BASH, JENKINS, KAFKA, PUB SUB, AIRFLOW, SONARQUBE**
- Consultant for a major Italian **telecommunications** company. The main task was to write some guideline documents about the main approaches and best practices in a Big Data platform and its reference environment
  - Software component structure and code documentation.
  - Git integration Strategy (with particular focus on branching models like Git-Flow and GitHub-Flow, code and environment promotion).
  - Data Lake principles (data ingestion, data preparation, data exporting, testing).
  - Lambda architecture, data modeling, data preparation, partitioning.
  - Analysis for migration of Cloudera Ecosystem to Google Cloud Platform.
- Consultant for a major Italian **energy** company.
  - Development of Apache Spark batch jobs using Scala for ingestion and processing of Avro file landed in HDFS.
  - Design and development of Data Quality Spark Processing Jobs for reports based on Hive and Impala.
  - Data Lake principles (data ingestion, data preparation, data exporting, testing).
  - Data preparation, data modelling and data partitioning.
    MAIN TECHNOLOGIES: **SCALA, JAVA, PYTHON, SQL, POSTGRESSQL, BASH, JENKINS, SONARQUBE**

CERTIFICATIONS

07/2022   **GCP Professional Data Engineer**   **Google Cloud**
certified link

02/2022   **Astronomer Certification for Apache Airflow Fundamentals and DAG Authoring**   **Astronomer**
certified link Fundamentals
certified link DAG Authoring

10/2023   **PET B1**   **University of Cambridge**
certified link

10/2019   **Best Student Machine Learning**   **Università della Calabria - Department of Mathematics and Computer Science**
certified link

09/2012   **Attendance course at ETC International College**   **ETC International College, Bournemouth, England**
certified link

EDUCATION

2021	**Master Degree in Computer Science**	**University of Calabria**
	Università della Calabria - Department of Mathematics and Computer Science	
2017	**Bachelor Degree in Computer Science**	**University of Calabria**
	Università della Calabria - Department of Mathematics and Computer Science	

LANGUAGES

English - B1+, **Italian** - native

DRIVING LICENCE

Driving Licence: A1, B