Giovanni Martucci

Section "CONTACT INFORMATION"
- email: giovannimar19@gmail.com;
- birthday: 1996;
- Nationality: Italiana;
- Web-site: https://giovanni-martucci.github.io
- Linkedin: https://www.linkedin.com/in/giovanni-martucci-08855b206
- Github: https://github.com/Giovanni-Martucci

Section "PROFESSIONS"
Data Scientist/Engineer, Machine/Deep learning Engineer, AI Engineer, Software developer, Full-stack developer, Cloud system developer, IT Security, Penetration tester.

Section "EXPERIENCE"

1. Experience 1
Where: Roma (Jun 2023 - Present)
Professions: Data Scientist/Engineer - AI Engineer
Company: Target Reply
Key point of company: Machine Learning, Data Science, Generative AI, Python, Deep Learning, NLP, LLM, RAG, LangChain, Amazon Web Services (AWS), Azure, Big-Data, Spark, Hue, Scripting, Automatization, SQL, Cloudera, Hive

About:
I carried out various activities of Data Science, MLOps, Data Engineering and AI Engineer and Automation for this
company:

2. Experience 2
Where: Catania (Jan 2023 - May 2023)
Companies: Samsung Innovation Campus and University of Catania.

The course, which selected the top 25 students, focuses on topics of Internet of Things applied to the
Consumer Electronics market, Machine Learning, Deep Learning, Artificial Intelligence, Computer Vision,
Cybersecurity, Cloud and Big Data.
Link: https://www.samsung.com/it/campaign/innovation-campus/

3. Experience 3:
Where: Catania (Jan 2023 - Present)
Professions: ComputerVision Engineer / Research Assistant
Company: University of Catania

About work:
Research activities with IPLab, ARSLab and STMicroelectronics.

In this research project, I developed the vision algorithm to provide a vision for the robot that has the
ultimate goal of playing a game of table football. Specifically the technologies used were: Pytorch,
OpenCV, Yolov8, RNN, Morphological operations.

4. Experience 4:
Where: Hatfield, England, United Kingdom (Jan 2022 - Jul 2022)
Erasmus: master thesis project
Company: University of Hertfordshire

About:
Erasmus focused on developing the master thesis project titled: AI and Computer Vision techniques
applied to problem of underestimating distance perceptions in VR environments.
Specifically, in this project we made use of the methodology called "Triangle Similarity" used for a first
level of inference and to generate a dataset labeled to train a CNN (ResNet50) to infer the distances of
the objects or people from the user who is immersed in a virtual environment.

5. Experience 5:
Where: Catania (Jul 2020 - Feb 2022)
Profession: Research Assistant
Company: University of Catania

About:
This research project, called "Saturn", approved by MiSE, involves STMicroelectronics and the research
team from the University of Catania, which aims to find a solution to an NP-Hard scheduling problem.
Several strategies are employed in this project to find the optimal solution, including Greedy
methodologies, operational programming, machine learning strategies, reinforcement learning, and
CSP.

6. Experience 6:
Where: Catania (Dec 2020 - May 2021)
Profession: Internship
Company: iCTLab

About:
Research and development of deep learning strategies for the identification of DeepFake.
Using a forensic approach, we tried to extrapolate latent patterns using the Expectation and
Maximization (E-M) algorithm. These hidden patterns are incorporated into images during generation by
generative neural networks (e.g., STARGAN). Next, through a clustering algorithm (DBSCAN), we tried to

place these latent patterns in the correct cluster. Each cluster represents a generative neural network
model of deep fake creation.

7. Experience 7:
Where: Catania (Jul 2019 - Dec 2019)
Profession: Internship
Company: TIM

About:
Internship carried out at Tim's Joint Open Lab for the development of Re-Identification systems in the
smart city context, obtaining a system for searching for hidden or missing individuals using advanced
Computer Vision and Machine Learning techniques.
The key technologies employed included PyTorch, OpenCV, and Convolutional Neural Networks (CNN),
specifically ResNet18 and ResNet50.
For user interaction, I designed an interactive web platform. Users could upload photos, crop individuals
from camera frames, or even retrieve images from the internet. This platform allowed for the
identification of both ordinary and/or famous people. The system aimed to enhance safety and security
in urban environments by aiding in the search for missing persons.

8. Experience 8:
Where: Catania (Apr 2021 - Apr 2021)
FameLab21: Scientific dissemination on DeepFake theme.

About:
In this scientific popularization meeting, I exposed and presented the problem related to Deep Fake and
Deep-Porn, exposing all the various problems that occur in cascade.
Apr 2018 Web Developer/ Developer
Development of websites, web apps, and applications for various clients. The technologies used are
Angular, React Native, Bootstrap, Materialize CSS, Firebase, MongoDB, Node.js, Express.js, Flask,
Javascript, HTML, Selenium, Telegram-bot, Scripting, Automatization Scripting, REST API.

Section "EDUCATION"

1. I have a master degree in Data science, Artificial Intelligence and Cybersecurity, with an experimental thesis entitled "AI and Computer Vision techniques applied to the problem of underestimating distance perceptions in VR environments”.
Grade: 110/110 cum Laude.
This is the list of subjects that I studied during my studies:
•Algorithms and complexity

- Analytical-forensic methods for scientific investigations (means)
- Artificial intelligence and laboratory
- Big data
- Cloud systems and IoT (AWS)
- Computer security
- Computer vision
- Distributed systems engineering
- English
- Enterprise Startup and Business Models
- Fundamentals of data analysis
- Machine/Deep learning
- Optimization
- Vulnerability Assessment and Penetration Testing (VAPT)
- Internship in ICTLab

2. I have a bachelor degree in Computer Science, with an experimental thesis entitled "Development of Re-Identication systems in the Smart City environment".

Grade: 105/110.

This is a list of subjects that I studied during my studies:

- Advanced Programming Laboratory 2
- Algorithms
- Architecture of Elaborators
- Competitive and Distributed Programming Techniques
- Computer Forensics
- Computer Networks
- Data Mining
- Databases
- Discrete Mathematics
- Elements of Mathematical Analysis
- English
- Fundamentals of Informatics
- Interaction and Multimedia
- Internet Security
- Mathematical and Statistical Methods
- Operating Systems
- Physics
- Programming 1
- Programming 2
- Software Engineering
- Internship in TIM Wave

Section "SKILLS"

1. General skills: Machine Learning, Deep Learning Problem solving, Vulnerability Assessment e Penetration Test Team working.

2. Programming-languages skills: C++, Python, C, C#, Java, SQL/HQL, Assembly, React Native, JavaScript, Processing.

3. Framework skills: AWS, Pytorch, Cloudera, Amazon, Azure AI, Docker, LLM, RAG, LangChain, Selenium, Appium, Angula, Firebas, MongoDB, Flask, Boostrap, Node.js, Express.js, Telegram Bot API.

4. Languages skills: Italian, English (IELTS English Certification level B2).

Section "PROJECTS"

1. Deep learning project: In this project it's possible to see an implementation of ResNet-18 pre-trained to resolve a task of classification of medical images using the framework Pytorch Lightning.

2. Market analysis - Data analysis: The subject of this project is the study and analysis of customer personality, aiming for a detailed analysis of ideal customers for a company. This analysis helps to better understand customers needs, preferences, and behaviors, allowing the company to modify its products according to the specific requirements of different types of customers and to direct its marketing activities towards customer
groups more likely to purchase the offered products. This project illustrates and uses the main methodologies of data analysis.
In detail, we will see:
- Data Cleaning and Pre-processing;
- Density analysis;
- Data normalization;
- Dimensionality reduction: through PCA and LDA;
- Linear regression;
- Clustering: K-means;
- Classification using: Logistic Regression, Naive Bayes, LDA, QDA, SVM, decision trees, tree forests, KNN, neural network;
- Ensemble methodologies;
- Cross-validation of generated models.

3. Neural network for classification (CNN): The task of this project is to "classify the different touches of the various knobs in a domestic oven".
Specifically, given an oven, our algorithm, through a specially trained neural network, must be able to
recognize if no knob is being touched, or if one is being touched, which one specifically.
In this project, we had to work with images, used as input for our network. This is why we used a Convolutional Neural Network (CNN) which allows neural networks to be applied to image processing, able to scale large images and large datasets of images.
This problem of classifying touch/no-touch in a domestic oven, can actually be extended to industrial machinery by considering actions that result from specific choices.

4. Big Data Projects and Methodologies: These projects make use of the primary Big Data methodologies, including:

- Hadoop: Hadoop is a framework for distributed storage and processing of large datasets across clusters of computers. It allows for parallel processing of data and fault tolerance, making it a popular choice for big data applications.
- Pyspark/ Spark: Spark is an open-source distributed computing system that is designed to be fast, flexible, and easy to use. It is built on top of the Hadoop Distributed File System (HDFS) and allows for in-memory processing of large datasets.
- LSH: Locality-Sensitive Hashing (LSH) is a technique used in data mining and machine learning to approximate similarity between data points. It is commonly used in applications such as nearest neighbor search and recommendation systems.

Together, these technologies form a powerful ecosystem for processing and analyzing big data, allowing for efficient storage, processing, and analysis of massive datasets.

5. Minimum weight vertex cover problem with Iterated Local Search (ILS): The aim of the project was to study and implement the Iterated Local Search to solve the problem of the Minimum Weight Vertex Cover belonging to the group of problems NP-Complete, therefore unsolvable with deterministic algorithms. A local search starts from a valid solution and iteratively builds a new solution trying to improve it until reaching a global optimal. During the execution of the algorithm the solution may end up in a very good local. Thanks to the perturbation operators and the acceptance criteria, the solution of local optimal is then removed by exploring new areas of the research space.

6. ComputerVision Projects:

- Camera calibration;
- Creation of a stereoscopic system;
- Measurement of real-world objects;
- Texture classification
- Deep learning

7. Cloud project - AWS + Docker: This project aims to develop a web-app in a distributed manner on AWS (using different AWS services offered by Amazon) with a dockerization system (Docker). The application contains a control dashboard for real-time monitoring of data extracted from microcontrollers on Arduino. Specifically, from the web interface, it is possible to retrieve, through a dedicated box, data generated by microcontrollers for the probability of rain, the probability of sunny weather, and temperature. Data extraction occurs through micro-services implemented in nodejs, which are invoked by the back-end connected to the web interface. They intercept the last message sent on the channel by Arduino, through a specific configuration of the MQTT communication protocol, in which Arduino is involved as a publisher, therefore publishing the values to a channel, while the microservices act as subscribers to the channel.

Once the requested values are intercepted, they are sent to the web interface that triggered all the cascade calls.

8. Bioinformatics project: This project aims to identify patients with similar tumors by analyzing the biological processes of different patients and finding similar patterns among them, using data-mining and big-data techniques.
The software and algorithms used are:

- Mithrill algorithm, which was used to map expression values to corresponding genes in the network and assigned a perturbation value to each gene of each patient;
- Multimotif: a software that can calculate motifs for each patient's meta-pathway, finding the most recurring labeled subgraphs in the network

9. Vulnerability-Assessment-and-Penetration-Testing: The project aims to penetrate a server by gaining access to the website hosted on the server. Once access to the server is obtained, using the web application as an attack vector, different privilege escalation techniques will be employed to obtain a reverse shell with root privileges (obtaining remote control of server with administrator privileges).

10. Security project: Study and exploitation of a known vulnerability in IBM Q-Radar software.

11. Automated extraction: In this project, you can observe various automation techniques, including:

- Web-scraping
- Selenium
- Appium (Android)
- Scripting
- REST API

12. Remote Design patterns:

- Implementation of communication through RabbitMQ: Routing and RPC.
- Design and development of the following design patterns to increase the level of security and performance in the application developed during the cloud project on AWS:
  - Reference Monitor;
  - Timeout;
  - Cache system.

13. Insurance app: In this project an application for an insurance agency has been developed. In particular, two versions have been developed, one for customers and one for insurers. The customer version allows you to immediately send photos, location and other details of a road accident to your appraiser. The version for insurers will provide a list of all road claims made by customers. The applications were developed using

React-Native with the addition of Firebase.

14. IoT project: The system developed aims to detect the presence of combustible gases and smoke through the use of
a gas sensor, in order to increase safety and prevent fires and casualties in homes, offices, etc.