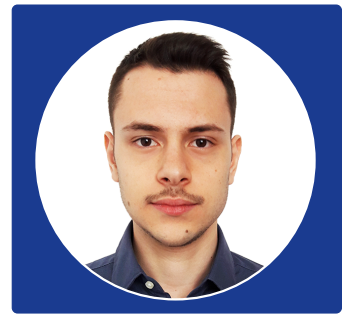


# Claudio Scalzo

DATA SCIENCE AND SOFTWARE ENGINEERING STUDENT

☎ (+39) 346 24 55 116 • ✉ claudio.scalzo@outlook.com • 🌐 github.com/claudioscalzo • 🔗 linkedin.com/in/claudioscalzo

*Born in Italy in 1994. Passionate about data science and software development. Strong analytical thinking and team-working skills. Successful academic results due to the passion for these topics and the high precision in on-time fulfillment of projects, tasks and assignments.*



## Education

### Master's Degree in Data Science and Engineering

EURECOM / TELECOM PARISTECH

*Sophia Antipolis, France*

*Sep. 2017 - Mar. 2019*

- Double Degree programme between Telecom ParisTech and Politecnico di Torino

### Master's Degree in Software Engineering

POLITECNICO DI TORINO

*Turin, Italy*

*Sep. 2016 - Mar. 2019*

### Bachelor's Degree in Software Engineering

POLITECNICO DI TORINO

*Turin, Italy*

*Sep. 2013 - Jul. 2016*

- with the highest grade, 110/110

## Experience

### SAP France

DATA SCIENTIST / SOFTWARE ENGINEER INTERN

*Paris, France*

*Jul. 2018 - Dec. 2018*

- Provisioned...
- Built...
- Implemented...

## Projects

### Team leader for an optimization project for TIM / SWARM Joint Open Lab

GITHUB.COM/CLAUDIOSCALZO/COLOTE-OMA-PROJECT

- Solving of a VRP (Vehicle Routing Problem) optimization problem proposed by TIM and the SWARM Joint Open Lab, for an IOT project named *ColoTe*. Multistart tabu-search approach, written in Java (with the *OpenTS* Java library).
- Achieved first position in the final ranking. Secured great comprehension of metaheuristics and, overall, improved logical/algorithmic competences and team-working skills.

### Virtual Assistant for answering music related questions

GITHUB.COM/D2KLAB/MUSIC-CHATBOT

- Developing of a virtual assistant (in the chatbot and vocal assistant forms), capable of answering music related questions and providing detailed graphical results. Informations extracted from the *DOREMUS* knowledge base, queried using the *SPARQL* language.
- Built using *Node.js* code with the *BotKit* framework. Used and trained Google's *Dialogflow* as NLP. Interfaced with the *Facebook Messenger* and *Slack* clients, using the respective APIs.

### House prices Kaggle challenge: predicting sales prices with advanced regression techniques

GITHUB.COM/CLAUDIOSCALZO/AML/TREE/MASTER/CHALLENGE

- Solution of one of the most famous Kaggle challenges, developed during the *AML (Algorithmic Machine Learning)* course at *EURECOM*.
- Top x% on the global ranking, thanks to smart preprocessing techniques (like *PCA* and *DBSCAN* for outlier removal), and thanks to advanced stacked regressors models.
- Written in *Python*, using *Pandas DataFrames* for the data structures and *scikit-learn* for the modeling phase.

### Challenge on the Fashion-MNIST and CIFAR-10 datasets: Naive Bayes and Bayesian Linear Regression

GITHUB.COM/CLAUDIOSCALZO/ASI-CHALLENGE

- Solution of the *ASI (Advanced Statistical Inference)* course. Implemented (from scratch) the Naive Bayes Classifier and the Bayesian Linear Regression. Used in the classification tasks of the Fashion-MNIST and CIFAR-10 images datasets.
- Written in *Python* using *NumPy* and *Pandas*. Achieved extremely satisfying results in terms of accuracy and computational efficiency.

## Skills

<b>Programming</b>	Python (+ NumPy, Pandas, scikit-learn, TensorFlow, PySpark, Spark MLlib) • C • Java • SQL
<b>Big Data</b>	Apache Spark • MapReduce & HDFS • NoSQL Architectures • Data processing (cleansing, analysis, mining)
<b>Machine Learning</b>	Deep Learning • Probabilistic Machine Learning
<b>OS &amp; Tools</b>	GNU/Linux (+ Bash, AWK, Sed) • Git • Jupyter Notebooks • Dialogflow (+BotKit) • MATLAB • LaTeX