

# Alessandro Marrella

hello@alessandromarrella.com | linkedin.com/in/alessandromarrella

## Education

---

### Master of Science in **Computer Science**

Sep 2014 – Oct 2016

Major: Business Informatics

University of Pisa, Italy

Graduation date: 07/10/2016 | GPA: 29.33/30.0

### Bachelor's Degree in **Business Administration**

Sep 2010 – Sep 2014

University of Verona, Italy

Graduation date: 11/09/2014 | GPA: 27.3/30.0 | Final grade: 109/110

### Exchange year (**Erasmus** scholarship)

Sep 2013 – May 2014

University College Dublin, Ireland

## Work Experience

---

### Master Thesis Internship

Sep 2014 – Oct 2016

ABB AG, Corporate Research, Ladenburg, Germany

Title of the thesis: "*Data Anonymization for the Industrie 4.0*"

- Creation of a database of the relevant literature with a review
- Development of a tool to collect data anonymization requirements for data analytics
- Implementation of state-of-the-art anonymization algorithms in Python
- Anonymization of industrial data for frequent pattern mining on event logs and reliability analysis tasks
- Evaluation on the impact of anonymization on privacy and quality
- Improvement of existing models from centralized to distributed
- Presentation of the algorithms and the results to in-house data scientists
- Submission of a paper to the CLOUDSPD'16 workshop

## Projects

---

### Scalable spam detection

Jan 2016

Created a support vector machine classifier able to predict spam comments on a website (Reddit.com)

- Context analysis and feature extraction with Word2Vec and SentiWordNet, clustering with K-Means
- Processing and modeling with Google BigQuery, Hadoop, Spark and MLlib

### Mobility data mining

Jun 2015

Analysis of GPS traces from ~500 taxis in San Francisco

- Study of the moving points behavior both visually and with statistical analysis
- Trajectory clustering with DBSCAN to find the frequent access patterns to the city
- Processing and modeling with M-Atlas, PostgreSQL and PostGIS

### Other data mining projects

Jan 2015 – Jun 2015

- Classification of Serie A soccer players based on game statistics with *C4.5* and *random forests*
- Market basket analysis on transactions from a local supermarket with *Apriori* and *FP-Growth*
- Social Network Analysis and community discovery on data collected from Last.fm
- Processing and modeling with Python, Pandas, Numpy, NetworkX, KNIME and Weka

## Skills

---

**Computer skills:** Python, Java, C, PHP, R, Hadoop, Hive, Spark, Weka, KNIME, SQL, SSAS, SISS, SRSS, BigQuery, Data Mining, Machine Learning

**Languages:** *Italian:* native, *English:* fluent, *German:* basic

**Other skills:** Teamwork, Literature review, Reporting, Problem solving, Critical thinking

## Volunteer

---

**Communication Manager** - Erasmus Student Network, Pisa, Italy

Dec 2014 – Jul 2015

**Public Relations Officer** - Erasmus Student Network, Dublin, Ireland

Jan 2014 – May 2015

**Emergency Medical Technician** - Croce Verde, Verona, Italy

Dec 2012 – Sep 2013