

# **ABOUT ME**



June 8, 1993



Groningen, Netherlands



+31 6 53228437



phil.m.pappas@gmail.com



www.philpappas.com



philippospappas



phil-pappas



English, Greek

# **INTERESTS**



Online Games



Physical Activities



Heavy Lifting
Travelling



Coffee



Books

# FILIPPOS PAPPAS

DATA SCIENTIST

# **SUMMARY**

I am a passionate Data Scientist who loves to get his hands dirty with messy data and identify hidden business insights. However, from my point of view, data are not only meant for discovering insights, but also to extract knowledge and discover alternative ideas than can convert the business challenges to cutting edge solutions.

# **EDUCATION**

#### **BACHELOR DEGREE**

BSc in Informatics and Telecommunications at University of Peloponnese

#### Certifications:

Certification of Data Science at National & Kapodistrian University of Athens

#### **MASTER DEGREE**

MSc in Computer Science at University of Peloponnese

Thesis: Extracting emotions from songs using Machine Learning algorithms

# **EXPERIENCE**

### **DATA SCIENTIST**

Feb 2019 to present | IBM

## **European Union**

- Pre-processing in Python all the received data from EU in order to optimise the search results of Watson Discovery Service
- Tested out-of-the-box techniques such as N-Grams, TF/iDF and Words Vectors for keywords generation that will improve the search accuracy of the delivered solution.
- Deployed the solution in IBM Cloud

### ABN AMRO

- Responsible for designing a state-of-the-art unsupervised learning solution for the bank's chatbot intent classification. Solution was deployed in AWS and it is been used bi-weekly to discover new client intents.
- Hyper-parameter tuning among clustering algorithms such as Gaussian Mixture Models,
   LDA and NMF to determine the optimal for discovering new user intents
- Exploratory data analysis of the user logs in order to understand better the data before feeding them to the algorithms

## NATO

- Data wrangling and analysis of the received data
- Experienced spaCy library for natural language processing
- Used NLP techniques such as stemming and tokenization for data quality improvement

# **EXPERTISE**

Supervised Learning	••••
Unsupervised Learning	••••
NLP	••••
Data Cleaning	••••
Data Visualization	••••

# **SKILLS**

Python	••••
Pandas / NumPy	••••
Matplotlib / Plotly	••••
SQL	••••
Git	••••
Cloud Computing	••••
Java	••••