

# Ioannis Prapas

Machine Learning, Big Data, Software Engineer

\* 27th of August 1992  
✉ io.prapas@gmail.com  
🌐 <https://iprapas.github.io>

## EXPERIENCE

Research Data Scientist (OCTOBER 2020 – NOW)  
*National Observatory of Athens, Greece*

**Deep Learning** on Satellite data for Fire Prediction.

Software Developer (SEPTEMBER 2017 – AUGUST 2018)  
*GEODE team, CERN Switzerland*

Refactor and enrich the application which is used to update and analyze positional measurements of the whole accelerator complex.

ML Research Associate (OCTOBER 2016 – JULY 2017)  
*Complex Event Recognition group, NCSR Greece*

Human Activity Reasoning with Computational Logic and **Deep Learning** (link). Grade 10/10

Software Engineer (MAY 2015 – JULY 2016)  
*EDMS team, CERN Switzerland*

Design and implement integration with pdf conversion engine. Refactor core script architecture. (SW design, Agile Scrum, Python, SQL, Unix Shell)

OS Lab Assistant (OCTOBER 2014 – FEBRUARY 2015)  
*National Technical University of Athens*

Helped university students understand basic OS concepts and complete the lab exercises.

## EDUCATION

2018 – 2020

M.Sc. in Big Data Management & Analytics  
*ULB (Brussels), UPC (Barcelona), TUB (Berlin)*

2-year joint Master program in **Big Data** technologies, Distributed Systems and **scalable Data Science**. GPA: 8.4/10

2010 – 2017

M.Eng. in Electrical & Computer Engineering  
*NTUA, Greece*

5-year studies specialized in Algorithms, Complexity, Computer Architecture, Operating Systems, AI, Networks & Security. GPA: 7.9/10

## SELECTED ONLINE COURSES

Computational thinking and **Data science** (MIT)  
**Scalable Machine Learning** with Spark (Berkeley)  
**Big Data** with Spark (Berkeley)  
**Machine Learning** (Stanford)  
**Deeplearning.ai** (Coursera)

## RECENT PROJECTS

Redeployment of **DL models** (2/20 - 9/20)

Efficiently **retrain and deploy Deep Learning** models in Pytorch and Tensorflow. (MSc Thesis)

Dynamic Resource Scaling (11/19 - 2/20)

Use of Reinforcement Learning to automatically scale a stream processing system (Apache Flink) that is fed data from Kafka sources.

Graph Recommendation System (2/2019 - 9/2019)

Define RDF ontology for scientific articles. Build publication graph on Neo4j and recommend reviewers based on their relevant impactful articles.

Kaggle Classification Competition (9/18 - 12/18)

Classify astronomical objects, based on their light curves (time-series data). Top 2% solution.

## SOFTWARE SKILLS

### PROGRAMMING LANGUAGES

Advanced: **Python, SQL**

Very Good: Java, Bash, JavaScript

Intermediate: C, R, Scala

### BIG DATA

Familiar with HDFS, HBase, MongoDB, Elasticsearch. Exposed to Graph Data Processing (Flink, Graphlab, Neo4j) and Knowledge Graphs (OWL, RDF).

Experienced in Batch and Stream Data Processing using Apache Spark, Flink, Kafka.

### DATA MINING - MACHINE LEARNING

Proficient in **Python ML libraries (scikit-learn, pandas, jupyter, numpy)**, **PyTorch** and **Tensorflow**.

### OTHER

Advanced in Git, Unix Shell, Oracle DB, Postgres.

Exposed to Agile Scrum, Kanban, Neo4j, Tableau, Flask, Cloud Computing (AWS, GCP), Docker.

## LANGUAGES

EA Native

EN Excellent (C2)

FR Very Good (B2)

DE,ES Basic (A2)