Diogo Ferreira

Data Scientist/Software Engineer

Diogo Daniel Soares Ferreira
23 years old
Oliveira do Bairro
Portugal Driving License – Class B
Fluent in Portuguese and English

+351 961253338

diogodsferreira@gmail.com

diogodanielsoaresferreira.github.io

linkedin.com/in/diogo-ferreira

https://github.com/diogodanielsoaresferreira

Languages

Python ★★★★

Java ★★★

Elixir ★★

C/C++ ★★

Matlab ★★

Go ★

Hi! I'm Diogo Ferreira and I love to use technology to solve daily life problems. I have special interest in the areas of Machine Learning and Software Architectures, although I am ready to deal with any Software Engineering related problem.

I am also passionate about nature and music, being member of the Portuguese Scouts (CNE) and attending several music summer festivals in the past years.

Experience

September 2019 -

Data Scientist

Talkdesk – TDX Coimbra/Aveiro

Design, implementation and test of NLP machine-learning based techniques to help agents in contact centers to improve the customer experience. Working with sentence embeddings, document similarity, word extraction from text and other NLP topics.

March 2019 – August 2019

Student Researcher

Instituto de Telecomunicações – Aveiro

Study and development of machine learning models for time series prediction metrics in 5G slices, using Deep Learning techniques. Design and implementation of a distributed real-time architecture to forecast anomalies and act on the network to improve the quality of service. The project was developed in the Network Architectures and Protocols (NAP) group, having as mentor the Prof. Dr. Susana Sargento.

July 2018

Summer Internship

Critical Software – Coimbra

Summer Innovation Week – Developing a platform for monitoring the noise and the temperature in the Critical Software open-spaces.

April 2018 – September 2018

Undergraduate Student Researcher

Instituto de Telecomunicações – Aveiro

Study of intelligence mechanisms in the core of a 5G network for monitoring, optimization and security. The project was developed in the Network Architectures and Protocols (NAP) group, having as mentor the Prof. Dr. Susana Sargento.

October 2017 – March de 2018

Undergraduate Student Researcher

Instituto de Telecomunicações – Aveiro

Continuation of a project made in a third-year course ("Project in Informatics Engineering"), in partnership with *Altice Labs*, having as main goal to develop a platform with support for management, monitoring, alerts and actuation over the data of a city, in real-time (*smart city*).

The project was developed in the Network Architectures and Protocols group (NAP), having as mentor the Prof. Dr. Susana Sargento.

On the scope of the project, I had the opportunity to go to events as exhibitor, like *Techdays '17, Symposium 2017 CMU (Carnegie Mellon University) Portugal* or *Sharing Cities – Lisbon*.

July 2016 – August 2016

Summer Internship at Aptoide

Lisboa

R&D department - Design and build a software with the purpose of finding apps on the Aptoide Store with explicit content (nudity) using Machine Learning techniques.

July 2013

Summer Internship at PT Inovação S.A./Altice Labs

Aveiro

Market study and research about the state of the art of telemedicine platforms and comparison with the Medigraf architecture.

Education

2014 -2019 **Integrated MSc in Computer and Telematics Engineering**

Aveiro University

Finished with a grade of **18.21**, being the best student of the course finishing in 2018/2019.

2011 - 2014 **Science and Technology course (10th – 12th grades)**

Gafanha de Nazaré High School

Finished with an average grade of **17**.

Publications

Distributed Real-time Forecasting Framework for IoT Network and Service Management
Diogo Ferreira, Carlos Senna, Susana Sargento
IEEE/IFIP Symposium Network Operations and Management (NOMS), Budapest, Hungary, April, 2020

Root Cause Analysis of Reduced Acessibility in 4G Networks
Diogo Ferreira, Carlos Senna, Paulo Salvador, Luís Cortesão, Cristina Pires, Rui Pedro, Susana Sargento
Machine Learning for Networking. MLN 2019. Lecture Notes in Computer Science, vol 12081. Springer, Cham
DOI: 10.1007/978-3-030-45778-5_9

Breaking Text-Based CAPTCHA with Sparse Convolutional Neural Networks
Diogo Daniel Ferreira, Luís Leira, Petya Mihaylova, Petia Georgieva
IbPRIA 2019: Pattern Recognition and Image Analysis pp 404-415
DOI: 10.1007/978-3-030-31321-0_35

Relevant projects

Other projects can be seen in https://github.com/diogodanielsoaresferreira.

- **IoT City** – Project developed in a 3rd year course, with IT (Instituto de Telecomunicações) and in partnership with Altice Labs, having as goal to develop a platform to manage, monitor, alarm and actuate over the data of a city, in real-time (smart city). Developed in Python using the Django Framework, stores the values received from many sensors of several verticals in a city and shows them to the city manager. Besides, also allows to define alarms based on the received values and send values to actuators when alarms are triggered.
- **NoSoundAllowed** – Project developed in the Summer Innovation Week, promoted by Critical Software. The goal is to monitor meeting tables according to the produced noise level, as well as the environment temperature, for everyone in the open-space to feel comfortable.

  Project link: https://github.com/diogodanielsoaresferreira/NoSoundAllowed
- **Captcha Solver –** Project developed in the 5th year course "Machine Learning", where it was proposed to recognize the digits in *Captcha* images. It was used an artificial neural network with several convolutional layers, and it was achieved an accuracy of 93.9% in character recognition.

  Project link: https://github.com/diogodanielsoaresferreira/captcha_solver
- **Aptoide Images Detector** – Project developed in an internship at Aptoide, with the goal of detecting explicit content in application images on Aptoide Store. With the Machine Learning techniques used, it was achieved an error percentage of less than 6%.

  Project link: https://contribsoft.caixamagica.pt/wiki/AptoideImagesDetector
- **SmartFarm** – Project developed in a 4th year course, where it was proposed a medium complexity project using DevOps techniques. The project was developed using SpringBoot Framework (Java) and has as objective to gather data from sensors in a farm, show them in real-time to the farmer and to add alerts related to the received values. Several technologies were used, such as Docker, Kafka, Kafka Streams, ReactJS, MongoDB, PostgreSQL, Jenkins, Cucumber, TICK Stack, ELK Stack, among others.

  Project link: https://github.com/diogodanielsoaresferreira/SmartFarm
- **Electric Cars Charging System** – Project developed in two 5th year courses, where it was built an electric car charging system, to ease the current electric charging system. The system contained an app that allowed the users to scan the QR code of the station and enabled the user to monitor the charging of his car. The notification service (built with by Elixir and Phoenix) would notify the user that his car was ready, being possible to pay the charging with the app. The communication between the station and the server was made with LoRa technology.

  Project link: https://github.com/diogodanielsoaresferreira/notification-service
- **Automated Detection of Cyber-trolls** – Project developed in the 5th year course "Machine Learning", with the goal of recognizing offensive tweets. Machine Learning techniques were used for text classification and an accuracy of 95.6% was achieved.

  Project link: https://github.com/diogodanielsoaresferreira/OffensiveTweetsDetection
- **IoT network monitoring** – Project developed in a 4th year course, where it was proposed to monitor the attacks in a network done by IoT devices. Using Machine Learning techniques, several scenarios were tested, and all of them had an accuracy of over 99%.

  Project link: https://github.com/diogodanielsoaresferreira/IoTNetworkMonitoring
- **DoS attack monitoring in real-time with a Bloom Filter, using an FPGA** – Project developed in a 4th year course, where the goal was to detect a DoS attack in a 5G network and block the attacker's IP, using an FPGA with a set of hardware-programmed Bloom Filters.

  Project link: https://github.com/diogodanielsoaresferreira/DosAttackMonitoringUsingFPGA
- **Search Engine Construction** – Project developed in a 5th year course with the goal of creating a search engine. Developed in Java, some modules created for the search engine were a tokenizer, a document indexer and a ranking system, based on TF-IDF weighting.

  Project link: https://github.com/diogodanielsoaresferreira/SearchEngine

Major relevant skills

- Experience in the implementation and testing of NLP algorithms, such as TF-IDF, LDA, word embeddings, sentence embeddings, document similarity and word extraction
- Experience in training an attention-based language model, such as BERT, GPT, RoBERTA or Transformer-XL, to perform a specific NLP-related task, such as intent recognition, entity extraction and sentence embeddings
- Experience in NLP frameworks, such as Spacy, GenSim, Duckling, Magnitude, Rasa and NLTK
- Experience in problem resolution using Machine Learning techniques and state-of-the-art Deep Learning techniques, with frameworks like Keras or Scikit Learn
- Experience in data analysis problems using data manipulation and visualization tools such as NumPy, SciPy, Matplotlib, Jupyter, Pandas and similar
- Experience in small and medium project management in a multidisciplinary team using Agile techniques
- Academic experience in DevOps and associated tools, such as Docker, ELK Stack, TICK Stack, Cucumber (for integration tests), Nginx, Gunicorn and similar
- Academic experience in event-driven architectures using micro-services and associated tools, such as Kafka or RabbitMQ (message queues), distributed databases and similar tools
- Experience in usage of web frameworks, such as Django, Spring Boot, Phoenix/Elixir or Falcon
- Experience in managing relational and non-relational (NoSQL) databases
- Experience in client-server protocols, such as REST or SOAP
- Experience in the problem domain in Smart Cities projects, IoT projects, and similar
- Knowledge of elaboration of technical reports using LATEX
- Knowledge about the construction and optimization of linear and non-linear Machine Learning models, particularly Deep Learning models

Social skills

Beyond informatics, I have always had other hobbies, where I played an active role in the community. I attended catechesis from 1st to 12th year and I was part of Corpo Nacional de Escutas for 12 years, where I was guide of my team and secretary/treasurer.

I have enrolled in several activities from Aveiro diocese, namely “Convívios Fraternos”, Taizé pilgrimages, Santiago de Compostela pilgrimages and other activities.

I also did sports since I was 5 years old, playing Football/Futsal in several teams until I was 18 years old (S.C.Beira-Mar, Atómicos S.C., Clube P.T., Grupo Desportivo da Gafanha e Oliveira do Bairro S.C.), being team captain on several occasions.

During the summer holidays I attended in several holiday camps, namely “Academia Viva’12” at the Aveiro University (in the courses “The communication networks are fun!” and “Electronics and Informatics”) and “Oliveira Viva” at Oliveira do Bairro sports park.

Currently, I am part of the juvenile choral group at the Oliveira do Bairro parish and I am a candidate to scout leader of Corpo Nacional de Escutas. All these experiences brought me bigger sense of responsibility, leadership, team spirit, punctuality and assiduousness.

Interests

On informatics I have great interest by machine learning, programming and data science areas. I have done online courses to learn more about machine learning, such as the Machine Learning course by Andrew Ng in Coursera, the Deep Learning Specialization by Deep Learning.ai, or CS224n: Natural Language Processing with Deep Learning by Stanford University.

Beyond the digital world, I am also passionate about music, having attended several music summer festivals in the past years. I also enjoy sports, mainly football.

In my free time I also enjoy reading books and attend events related to informatics, technology and programing, like JNation 2020, DSPT Day 2019, National Meeting of Informatic Engineering Students (ENEI 2019 and ENEI 2018), TechDays (2017) and Talk a Bit (2017), DSPT meetups among others.