

# DAVID CAÑONES

davidadrian.cc - david.canones.castellano@gmail.com - david.canones@whiteboxml.com

## ABOUT

---

Data Science Leader with a passion for AI and Machine Learning. I've been thriving in the field since 2015, before it was a trending topic.

My journey started at Servinform, crafting automation solutions that saved significant resources. I further honed my skills at Kernel Analytics and Pragsis Bidoop, focusing on predictive analytics and energy prediction models.

In 2019, I co-founded WhiteBox, aiming to deliver top-tier Data & Analytics projects in Spain, the UK, and the US. We've grown significantly, maintaining high-quality standards and attracting top talent. In 2021, we branched out to become one of the first DaaS companies in Spain, launching DataMarket, a platform providing quality datasets for our clients.

My goal is to continually innovate, add value through data, and promote technical excellence within my team.

## EXPERIENCE

---

### WhiteBox

October 2019 - Present

*Lead Data Scientist & Partner*

· Responsibilities:

- Leading the Data Science team, mentoring and overseeing their work, providing technical guidance and advice.
- Developing complex components of our projects and boosting the productivity of my team members.
- Collaborating closely with clients to understand their needs and identifying impactful data use cases.

· Achievements:

- Finalist in the international edition of the AWS DeepRacer in Las Vegas, 2019.
- Led the development of the Bricomart demand forecasting model in 2020, optimizing inventory for over 900 references across 30 stores.
- Developed the Vodafone causal inference framework in 2021, enabling the establishment of cause-effect relationships in churn, network quality, and customer service.

### DataMarket

January 2021 - Present

*Chief Data Officer (CDO) & Co-founder*

· Responsibilities:

- Steering the company's strategic vision and fostering a data-centric culture.
- Overseeing data sourcing, curation, and distribution, ensuring high quality and relevance for our customers.
- Building and maintaining client relationships, understanding their unique data requirements, and delivering tailored data solutions.

· Achievements:

- Proud to serve a growing roster of clients including Burger King, Adevinta, and Iryo, among others.
- Developed a robust platform offering instant access to a wide range of datasets, helping businesses complement their internal data and gain comprehensive insights.
- Positioned DataMarket as one of the first and foremost DaaS providers in Spain, playing a pivotal role in the country's data revolution.

## **Ironhack**

July 2019 - September 2021

*Lead Instructor Data & Analytics*

- Responsibilities:
  - Crafting a comprehensive, cutting-edge curriculum for the Data Analytics course, focused on the practical application of Machine Learning, Data Visualization, and Data Engineering concepts.
  - Guiding students through complex data projects, promoting a hands-on learning approach and fostering critical thinking.
  - Providing personalized mentorship to students, helping them overcome technical challenges and guiding them in their career development.
- Achievements:
  - Successfully mentored over 100 students during 2 years, many of whom secured roles in top tech companies post-graduation.
  - Developed a unique teaching methodology that simplified complex data concepts, earning a 95% satisfaction rating in student feedback consistently.
  - Actively participated in local Data Science meetups and events, representing Ironhack and inspiring a broader audience to explore the world of data analytics.

## **Pragsis Bidoop (now part of Accenture)**

July 2018 - October 2019

*Senior Data Scientist*

- Responsibilities:
  - Building robust Machine Learning models leveraging traditional (scikit-learn, XGBoost, lightGBM) and Deep Learning (TensorFlow, Keras) frameworks.
  - Scaling Machine Learning models from concept to deployment using distributed and parallel computing (Spark, Dask, Celery) and orchestration tools like Apache Airflow.
  - Developing cutting-edge Computer Vision solutions using Google Edge TPUs, Nvidia GPUs, TensorFlow, and OpenCV.
- Achievements:
  - Enhanced power production forecasting for a 1GW cluster of 13 wind farms in Washington, USA. This led to a 10% reduction in forecasting error, resulting in substantial cost savings for our client.
  - Contributed to the winning team in the Amazon Web Services DeepRacer League by developing a Reinforcement Learning algorithm.
  - Created a real-time tracking system for person detection using Computer Vision techniques optimized for low-consumption hardware (Google Edge TPU, Raspberry Pi).

## **Kernel Analytics (now part of Boston Consulting Group)**

October 2017 - July 2018

*Data Scientist*

- Responsibilities:

- Leveraging traditional (scikit-learn) and Big Data (Spark MLlib) frameworks to develop machine learning models.
- Developing interactive dashboards with Plotly Dash and Microsoft PowerBI for effective data exploration and visualization.
- Achievements:
  - Successfully designed and developed a Customer Experience Management (CEM) framework for a leading Mobile Operator. The pipeline ingested 3G/4G antenna data and created a model to relate Customer Experience with Churn and Complaints (contact center calls), resulting in the client gaining actionable insights about its mobile network infrastructure impact on customer experience.
  - Developed a predictive model for a renowned Mobile Operator that could predict user complaints (contact center calls) based on consumption patterns and user personal profile. This led to automation of a part of support process for our client.

## **Grupo Servinform**

February 2015 - October 2017

*Data Scientist*

- Responsibilities:
  - Designed and implemented automated data pipelines capable of cleaning and ingesting data from diverse inputs into relational databases.
  - Developed an innovative data product to simplify data exploration for pharmaceutical and healthcare researchers, enabling them to identify complex relationships between different data sources.
- Achievements:
  - Successfully developed an NLP model that could interpret user queries written in natural language and translate them into database queries (SQL).
  - Conceptualized and developed an automation framework that integrated seamlessly with existing back-office processes, paving the way for a new business line (RPA) and the ability to automate previously unfeasible processes.

## **EDUCATION**

---

### **EOI Business School**

MBA, Corporate Finance

*September 2014 - September 2015*

### **Universidad de Sevilla**

MS Industrial Engineering, Energy

*September 2007 - September 2014*

## **HONORS & AWARDS**

---

### **DeepRacer League, 3rd position**

*Amazon Web Services (AWS)*

Machine Learning competition organized by AWS which consisted on developing a Reinforcement Learning model for an autonomous car. I got the 3rd position in Spanish competition and was member of the team that made the top 3 positions (Gold, Silver and Bronze)

## **COURSES AND CERTIFICATIONS**

---

- **Machine Learning**  
*Stanford University*
- **Deep Learning Specialization**  
*deeplearning.ai*

- **TensorFlow Developer Certificate**  
*Google*
- **TensorFlow in Practice Specialization**  
*deeplearning.ai*
- **AI for Medicine Specialization**  
*deeplearning.ai*
- **Computational Thinking and Data Science Specialization**  
*MIT*
- **Advanced Machine Learning Specialization**  
*Higher School of Economics*
- **Learning From Data**  
*Caltech*
- **The Analytics Edge**  
*MIT*
- **Data Manipulation at Scale: Systems and Algorithms**  
*U. Washington*

## TECHNICAL STACK

---

<b>Machine Learning</b>	scikit-learn, XGBoost, lightGBM, H2O, MLlib, MLflow
<b>Deep Learning</b>	TensorFlow, Keras, PyTorch
<b>Computer Vision</b>	OpenCV
<b>NLP</b>	NLTK, spaCy, Fasttext, Transformers
<b>In Memory</b>	pandas, polars, NumPy, Apache Arrow
<b>Relational</b>	PostgreSQL, SQLite, Oracle
<b>Big Data</b>	Apache Spark, Hive, Impala, ClickHouse
<b>Cloud Computing</b>	AWS (Amazon Web Services), Microsoft Azure
<b>Orchestration</b>	Apache Airflow
<b>Languages</b>	Python, R, SQL