

David Paule, PhD

COMPUTING SCIENTIST AND ENGINEER

☎ (+34) 686 119 418 | ✉ j davidpaule@gmail.com | 🏠 davidpaule.es | 📄 github.com/Salias | 📺 davidpaule

Experience

Data Scientist/Engineer

Barcelona, Spain

JOBANDTALENT

July 2019 - Current

- Work on Matching/Recommendation Systems
- Define and improve metrics of main business objectives
- Perform descriptive and predictive analysis
- Design of statistical models or applied machine learning
- Involved in the full data science stack (from prototyping to production)
- Building ETL (Extract, Transform and Load) Pipelines using airflow
- AWS Stack (S3, ECS, EC2, Lambda, Kinesis, CloudWatch, Redshift)
- Production data services/API in Scala

PhD Researcher/Research Assistant

Glasgow, Scotland

INFORMATION RETRIEVAL GROUP/URBAN BIG DATA CENTRE - UNIVERSITY OF GLASGOW

Apr. 2014 - Feb. 2019

- PhD Thesis (IR/NLP)
- Research Assistant of researcher in other areas
- Collection and Pre-processing and cleaning of Big Datasets (Cloudera Manager, Hadoop, Spark).
- Identify research problems and address them with critical thinking and scientific rigour.
- Design of research experiments.
- Adequately interpret and communicate research findings.

Undergraduate Tutoring

Glasgow, Scotland

UNIVERSITY OF GLASGOW

Sep. 2015 - Jun. 2017

- Help and guide many groups of students through their studies.
- Programming (Python).
- Computer systems & databases.
- Human-computer interaction
- Algorithm foundations.

Junior Backend Engineer

Glasgow, Scotland

TWIG WORLD - TWIG-WORLD.COM

Jan. 2014 - Apr. 2014

- Successfully developed, maintained and tested complex database-driven web applications.
- Adopted and proposed new software development methodologies and experimented with new tools and technologies.
- Python/Django.
- Test Driven Development (TDD).
- Continuous Integration (CI).
- CVS (GIT).
- MySQL and Object-Relational Mapping (ORM).

Software Engineer & Data Analyst

Tenerife, Spain

CANARIAS.COM

May. 2013 - Aug. 2013

- Developed an email marketing web application integrated into the company's software ecosystem.
- PHP/Javascript (Frontend and Backend).
- Microsoft SQL Server.
- Data mining for the Marketing Department (Weka).

Software Engineer

Tenerife, Spain

FREELANCE

Sep. 2012 - Sep. 2013

- Analysis, design and implementation of a press-oriented CMS that provide local and small clients with a tool that helps them to automate and simplify the management of an online newspaper.
- Full Stack (Frontend and Backend).
- Python/Django.
- Test Driven Development (TDD).
- Continuous Integration (CI).
- CVS (GIT).
- MySQL and Object-Relational Mapping (ORM).

Education

PhD in Computing Science (Information Retrieval and Machine Learning)

Glasgow, Scotland

UNIVERSITY OF GLASGOW

Apr. 2014 - Nov. 2018

PhD thesis title: "Inferring the Geolocation of Tweets at a Fine-Grained Level" (<http://theses.gla.ac.uk/41007/>).

Bachelor's Degree in Computer Engineering

Tenerife, Spain

UNIVERSITY OF LA LAGUNA

2007 - 2013

I obtained my Bachelor degree at the University of La Laguna (Spain), which has a focus on Information Systems and Software Engineering.

Research

RESEARCH PUBLICATIONS

During my PhD, I had the opportunity of conducting interdisciplinary research and published multiple research papers that have contributed to the fields of Computing Sciences, as well as Social Sciences and Geographic Information Sciences.

PhD Thesis - Inferring the Geolocation of Tweets at a Fine-Grained Level

2019

JORGE DAVID GONZALEZ PAULE

University of Glasgow

Learning to Geolocalise Tweets at a Fine-Grained Level

2018

JORGE DAVID GONZALEZ PAULE, YASHAR MOSHFEGHI, CRAIG MACDONALD AND IADH OUNIS.

In Proc. of the 27th International Conference on Information and Knowledge Management (CIKM'18)

On Fine-Grained Geolocalisation of Tweets and Real-Time Traffic Incident Detection

2018

JORGE DAVID GONZALEZ PAULE, YERAN SUN AND YASHAR MOSHFEGHI.

In Information Processing & Management, Elsevier.

Beyond Geo-Tagged Tweets: Exploring the Geo-Localisation of Tweets for Transportation Applications

2018

JORGE DAVID GONZALEZ PAULE, YERAN SUN AND PIYUSHIMITA THAKURIAH.

In Transportation Analytics in the Era of Big Data, Springer.

On Fine-Grained Geolocalisation of Tweets

2017

JORGE DAVID GONZALEZ PAULE, YASHAR MOSHFEGHI, JOEMON JOSE AND PIYUSHIMITA THAKURIAH.

In Proc. of the 3rd International Conference on the Theory of Information Retrieval (ICTIR'17). Amsterdam, Netherlands.

Spatial analysis of users-generated ratings of yelp venues

2017

YERAN SUN AND JORGE DAVID GONZALEZ PAULE.

In Open Geospatial Data, Software and Standards

Sensing spatiotemporal patterns in urban areas: analytics and visualizations using the integrated multimedia city data platform

2016

PIYUSHIMITA THAKURIAH, KATARZYNA SILA-NOWICKA AND JORGE DAVID GONZALEZ PAULE

In Built Environment 42(3), pages 415-429.

Skills Summary

DATA SCIENCE

Theoretical Background	Information Retrieval; Machine Learning; Natural Language Processing; Data (Text) Mining;
Information Retrieval	Lucene (Java); Terrier (Java)
Machine Learning	Gensim; Scikit-learn; Weka
Analysis/Querying	Strong SQL; Excel;
Data Analysis	SQL; Pandas; Numpy; SciPy; Jupyter Notebook
NLP	NLTK (Python); SpaCy (Python)

DATA ENGINEERING

Data Services/API	Scala (Finatra); Python
ETL Pipeline	Airflow (python)
Cloud	AWS (Redshift; S3; ECS; EC2; Kinesis; Lambda; Cloudwatch)

SOFTWARE ENGINEERING

Programming Languages	Python; Scala
Software Development	Continuous Deployment/Integration; Agile; Testing; Code Reviews
Dev Tools	Git; Docker
DBMS	MySQL; PostgreSQL;

OTHERS

Languages	Spanish (Native); English (Professional working proficiency)
Documents	Spanish Passport; Eligible to work in the EU
References	Under request