

# Arnau Puche

<https://github.com/puchee99>

Email : [arnaupuchevila@gmail.com](mailto:arnaupuchevila@gmail.com)

Mobile : (+34)-673-821-695

## EDUCATION

---

- **Universitat Autònoma de Barcelona**  
*Bachelor's degree of Data Engineering (240 ECTS); Avg. Grade: 7.13*  
Bellaterra, Spain  
*Sep. 2018 – Present*
  - **TFG:**
- **Coursera**  
*Data Science*  
Sep. 2020 – Present
  - **IBM AI Enterprise Workflow (4/6):** connecting business priorities to technical implementations.
  - **Neural Networks and Deep Learning:** foundational concept by Andrew Ng.
- **Chartered Financial Analyst - CFA**  
*Investment management*
  - **Level 1:** studying to take the exam.

## EXPERIENCE

---

- **CaixaBank Tech**  
*Data Scientist*  
Barcelona, Spain  
August 2022 -  
**Tools:** Python, SQL, Bash
  - **Detecting and preventing frauds:** monitoring and analyzing consumers's financial actions to detect any suspicious or harmful behavior.
  - **Predictive and real-time analysis:** run various analytical algorithms and generate important information for forecasting future occurrences.
- **UAB**  
*Research Assistant*  
Bellaterra, Spain  
Sept 2021 - Feb 2022  
**Tools:** Python, SQL, Bash
  - **Database:** Create a SQL database from biomedical device data.
- **UVE Solutions**  
*Intern Data Scientist*  
Manresa, Spain  
Feb 2021 - July 2021  
**Tools:** Python, SQL, Bash, Streamlit, Pandas, Bitbucket, Confluence, Jira
  - **Data ingestion:** Build a data ingestion tool (Streamlit App) to perform ETL on different data formats. ML techniques were used to calculate similarity between words.
  - **BI:** Analyze the distribution of the hostelry sector in the Spanish territory.

## HACKATHONS

---

- **NUWE JOBarcelona 2022 Final (7 h)**  
*Data (teams of 4) - 1st*  
Camp Nou, Barcelona, Spain  
May 2022  
**Tools:** Python, Github (**Repo link**), Bash, Pytorch, Sckit-Learn, Streamlit, Pandas, XGBoost  
**Challenge:** Predict the fair market value of football players.
  - **Feature Engineering:** Through different techniques and combinations with the columns, we were able to add new variables with a higher correlation.
  - **Multivariant Regression:** We use different ML models with Grid Search to find the best parameters. We also tested with neural networks.
- **NUWE CaixaBank 2022 (14 h)**  
*Data -*  
online  
May 2022  
**Tools:** Python, Github (**Repo link**) , Bash, Pytorch, Sckit-Learn, Pandas, XGBoost  
**Challenge:** Modeling a predictive algorithm that allows knowing the future dynamics of the market from the historical data of the value of the IBEX35 and some Tweets.
  - **Sentiment analysis:** Some new columns were added in the market data. One was made by tweet's daily sentiment score. Other columns with popularity information from Google Trends.
  - **Time Series Forecasting:** Tested different Time Series and Classification models (up - down) .

- **NUWE Talent Squad 2022 (1 week)**

*Data - 1st*

online  
*July 2022*

**Tools:** Python, Github (**Repo link**) , Bash, Pytorch, Sckit-Learn, Pandas

**Challenge:** The dataset was data created by sensors from the company SpaceX. The goal was to classify the data into 5 types of weather conditions in order to manage the flight.

- **Multivariant Classification:** Neural networks were used.

- **NUWE JOBarcelona 2022 Classification (1 week)**

*Data - Qualified for the final*

online  
*April 2022*

**Tools:** Python, Github (**Repo link**) , Bash, Pytorch, Sckit-Learn, Pandas

**Challenge:** Classify insects according to the value of different sensors.

- **Multivariant Classification:** Neural networks were used.

## PROJECTS

---

- **Flask App:** create a functional app with different databases in the cloud. (Github link)

## PROGRAMMING SKILLS

---

- **Languages:** Python, Javascript, HTML, SQL, R, Bash, C
- **Technologies:** GCP, Docker, Tableau, AWS, Airflow, MongoDB, Pyspark, Flask

## LANGUAGES

---

**Catalan:** Native

**Spanish:** Native

**English:** Professional (B2 Certificate)