



FELIX BRUNNER,

PHD

DATA SCIENTIST

📍 Berlin, Germany
✉ brunner.felix@gmail.com
📞 +49 174 2479404
🌐 www.felixbrunner.github.io
🌐 /brunnerfelix
🌐 /felixbrunner
🌐 /@fbrunner

CORE COMPETENCIES

DATA ANALYTICS

- time series
- pandas
- numpy / scipy
- matplotlib
- causal inference
- econometrics / statistics
- text mining / NLP
- SQL

PREDICTIVE MODELING

- machine learning
- deep learning / neural networks
- PyTorch / TensorFlow
- scikit-learn
- huggingface
- forecasting
- computer vision
- feature engineering

A.I. SOFTWARE ENGINEERING

- git / CI / CD
- linux / bash / CLI
- docker
- langchain / LLMs
- FastAPI
- MLLow

OTHER SKILLS

PROGRAMMING & TOOLS

- python
- jupyter notebooks
- VS Code
- SQL
- MATLAB
- R
- Spreadsheets / Excel / VBA
- MS Office / Google Suite

SOFT SKILLS

- proactive problem-solving
- creative / innovation-driven
- stakeholder communication
- adaptability / continuous learning
- project management / organization

LANGUAGES

- English - Full working proficiency
- German - Native proficiency
- Portuguese - Intermediary
- French - Intermediary

PROFILE SUMMARY

Data science and machine learning specialist with 3+ years of experience in delivering commercial and research projects. Main expertise lies in developing **predictive models**, **data analytics solutions**, and implementing **AI-driven applications** in Python to drive actionable insights and business value.



PROFESSIONAL EXPERIENCE

Machine Learning Scientist

dida Machine Learning

05/2021 – 09/2023

- **Led deep learning research** in computer vision and time series, improving model accuracy and performance for real-world applications.
- **Implemented machine learning solutions** in client projects, automating processes, increasing operational efficiency, and enabling data-driven decision-making.
- **Developed data pipelines** for model deployment into live production.
- **Conceptualized and prototyped large language models (LLMs)** for question-answering systems, automating workflows and enhancing NLP capabilities.
- **Managed a five-person project team**, coordinating tasks and ensuring timely delivery of milestones and successful project outcomes.
- **Prepared in-depth analyses and visualizations** of large-scale unstructured data, delivering actionable insights through data-driven storytelling.

Teaching Assistant

Universidade Nova de Lisboa

09/2017 – 02/2022

- **Assisted in teaching and grading graduate-level courses** on financial data analysis, mentoring 300+ students on complex data-driven methodologies.
- **Developed course materials, assignments, and exams**, promoting applications of financial modeling, statistical analysis, and data visualization techniques.

EDUCATION & MORE

PhD, Economics

Nova SBE, New University of Lisbon

05/2024

- **Conducted research in applied econometrics and data analysis**, with a focus on the estimation of network graphs from large datasets with machine learning.
- **Utilized data science tools** such as Python, SQL, and Matlab to implement complex models, analyze large-scale datasets, and make significant contributions to the field.
- **Successfully defended dissertation** in May 2024 with distinction and honors.

Data Science Bootcamp

Lisbon Data Science Academy

07/2020

- Completed a 9-month program in **NLP, recommender systems, model deployment and more** with hands-on projects in python.
- **Competed in hackathons** and weekend sprints, applying advanced machine learning techniques in collaborative environments.

Career break

Parenting and PhD completion

2024

- Took a planned career break to be a family man while completing my PhD and enhancing my skills in data science, software development and freelance business.

REFERENCE PROJECTS

Visual Machine Status Detection in Industrial 3D Printing

Machine Learning Engineer / Project Manager

Manufacturing, Computer Vision

09/2023

- o **Led project management in a five-person team**, ensuring smooth collaboration, preparing project reports, and delivering key stakeholder presentations.
- o **Guided data collection and pre-processing** of infrared image data from production machines.
- o **Designed and implemented a custom annotation interface** for efficient labeling of image datasets.
- o **Developed a deep learning model** using convolutional neural networks (CNN) to detect machine pollution.
- o **Implemented data augmentation techniques** to improve model performance and generalization across machines.
- o **Deployed the model via a containerized API** using Docker and FastAPI, enabling integration into live production.
- o **Result:** Supplied a **real-time monitoring system** that automatically detected irregularities, reducing manual interventions and improving production efficiency.

Technologies: Python, PyTorch, CNNs, Docker, Git, OpenCV, FastAPI, PyTorch Lightning, ipywidgets, ResNet.

Links: [Article](#) on custom annotation tool with code [repository](#).

Quality Prediction Based on Sensor Time Series Recordings

Machine Learning Researcher / Project Manager

Manufacturing, Time Series

09/2023

- o **Analyzed and visualized high-frequency sensor data** to identify patterns related to product quality.
- o **Integrated data from multiple sources**, processing, filtering, cleaning, and combining datasets for consistency.
- o **Designed data pipelines and modeling approaches** to automate data preparation and streamline data modeling.
- o **Developed a quality prediction system** using advanced machine learning techniques, testing algorithms from traditional statistical models to cutting-edge deep neural networks.
- o **Conducted experiments with explainable AI** to derive actionable insights for process engineers.
- o **Result:** Provided **analytic insights for production optimization**, empowering process engineers to make informed adjustments, ultimately improving product consistency and reducing defects.

Technologies: Python, PyTorch, CNNs, ROCKET, Time Series Classification, Data Wrangling, Git, Explainable AI, matplotlib, numpy.

Links: [Article series \(#1, #2, #3\)](#) on the refined ROCKET model with code [repository](#).

Automated Question Answering via Document Retrieval

Data Scientist / AI Engineer

IT Services, NLP

02/2023

- o **Developed an extractive QA demo** using NLP models to automate question answering based on internal documents.
- o **Created a proof-of-concept for semantic search and automatic document retrieval**, improving search accuracy.
- o **Expanded the retrieval pipeline with more documents**, enabling generative QA with large language models (LLMs).
- o **Updated developers on the latest NLP advancements** and applications, ensuring cutting-edge technology adoption.
- o **Result:** Deployed a **responsive chatbot** for document-based QA, reducing manual search times and streamlining employee access to information.

Technologies: Huggingface Transformers, Haystack, LLMs, BERT, Semantic Search, Document Retrieval, Git, BeautifulSoup4, web scraping.

Estimation and Analysis of Variance Spillover Networks

Data Scientist / Researcher

Academia, Time Series Forecasting

11/2023

- o **Explored and evaluated various statistical learning algorithms for multivariate forecasting**, identifying optimal methods for large-scale datasets.
- o **Acquired data from SQL databases** and set up an automated pre-processing pipeline for efficient data handling.
- o **Implemented object-oriented code** for running cross-validated algorithms to estimate variance spillover networks.
- o **Conducted rigorous statistical and econometric analyses**, producing valuable insights into variance spillovers in economic and financial networks.
- o **Authored research papers** targeting high-impact journals, presenting novel methodology and findings.
- o **Result:** Published a **well-received paper** in *Quantitative Economics*, with follow-up research papers under review, contributing significantly to the field of econometrics and financial research.

Technologies: Python, pandas, numpy, scikit-learn, PostgreSQL, glmnet, Vector Auto-Regression, Multivariate Time Series Forecasting, Networkx, MS Azure.

Links: Published [paper](#) with code [repository](#).