Felix Brunner

Data Scientist / Machine Learning Specialist

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

Machine learning specialist with experience in commercial and academic projects. Main expertise lays in implementing AI projects in computer vision, time series, and NLP with the help of deep learning methods.

PROFESSIONAL EXPERIENCE

Machine Learning Scientist

dida Machine Learning — *05/2021 – 09/2023*

- Conducted deep learning research projects in computer vision and time series
- Implemented commercial machine learning software in a manufacturing context
- Developed data processing pipelines for the deployment into live production
- Conceptualized application of large language models for question answering
- Represented and organized a five-person project team
- Prepared analyses and visualizations of large amounts of unstructured data

Teaching Assistant

Universidade Nova de Lisboa — *09/2017 – 02/2022*

- Supported teaching and grading of Master's courses in financial data analysis
- Generated course materials, assignments and exams

Quantitative Strategies Intern

risklab / Allianz Global Investors — *07/2015 – 11/2015*

- Provided data-driven research on fund performance and portfolio optimization
- Prototyped tools for investment strategies in Matlab and VBA for Excel
- Modeled time series with regularized regressions and scenario simulations

Project Management Intern

zeb Consulting — *03/2014 – 08/2014*

- Prepared senior management presentations to guide an IT project in risk management

DEGREES & CERTIFICATIONS

PhD, Econometrics

Nova SBE — *09/2016 – present*

- Carried out research in applied econometrics and empirical data analysis
- Worked on the high-dimensional estimation of network graphs and applications
- Employed data science tools such as python, Matlab, SQL in the implementation
- Submitted dissertation „Connecting the dots: An exploration of spillover networks in economics and finance" in late 2023, defense expected in April 2024

Data Science Bootcamp

Lisbon Data Science Academy — *06/2020 – 02/2021*

- Completed weekly modules on NLP, recommender systems, model deployment, etc.
- Participated in weekend-long bootcamp and competed in monthly hackathons

MSc, Finance

Nova SBE — *09/2014 – 05/2016*

BSc, International Business Administration

WU Vienna — *10/2010 – 05/2014*

SKILLS

python
machine learning
data science
artificial intelligence (AI)
deep learning
neural networks
data wrangling
time series
git
bash / CLI
linux
docker
computer vision
NLP / LLMs
PyTorch
TensorFlow
scikit-learn
huggingface
MatLab
SQL
econometrics
statistics
causal inference
OOP
MS Office
API
R
Excel / VBA

LANGUAGES

English
Full working proficiency

German
Native proficiency

Portuguese
Limited working proficiency

French
Limited working proficiency

Felix Brunner

www.felixbrunner.github.io
/felixbrunner
/@fbrunner

More details and linked materials about the listed projects are available at www.felixbrunner.github.io/projects.

REFERENCE PROJECTS

Machine status detection in industrial 3D printing based on infrared image data

Computer vision *Manufacturing Sector* *10/2021 – 09/2023*

Baseline: Client has infrared cameras installed inside production machines for manual monitoring.

Goal: Automate the process monitoring to identify irregularities from live images.

Activities: I contributed to this project as a *machine learning developer* and *project coordinator*:

- Guided systematic data collection and pre-processing for the machine learning algorithms
- Defined the labeling process and implemented an interface to annotate the datasets
- Programmed a visual deep learning algorithm to detect machine pollution in live production
- Implemented data augmentation techniques to deal with machine heterogeneity
- Supply a containerized model with API endpoints for deployment to the production machines
- Coordinated and represented a five-person project team, prepared presentations and reports

Outcome: Provided a production-ready system to inform operators of potential faults in real time.

PyTorch
Convolutional neural networks
docker
git
OpenCV
FastAPI
computer vision
pytorch-lightning
ipywidgets

Quality prediction based on time series data in a manufacturing context

Time series *Manufacturing Sector* *10/2021 – 09/2023*

Baseline: Multitude of sensors collects time series data from production machines at high frequency.

Goal: Identify patterns in the data linked to final product quality.

Activities: I worked in this project as a *machine learning developer* and *project coordinator*:

- Analyzed and visualized large amounts of time series data
- Explored, filtered and connected various data sources to construct consistent datasets
- Conceptualized the data pipelines and modeling approaches
- Implemented a machine learning system for quality prediction in a manufacturing context
- Tested various predictive algorithms including statistical models and deep neural networks

Outcome: Derived qualitative insights to guide process engineers in optimizing production.

PyTorch
CNNs / ROCKET model
data wrangling
git
time series classification
explainable AI

Automated question answering via retrieval of internal documents

NLP *IT Services Industry* *05/2021 – 08/2023*

Baseline: Company employees manually search internal documentation for information.

Goal: Integrate chatbot for question answering (QA) into company communication platform.

Activities: I supported this project as a *machine learning scientist*:

- Created a demo for extractive QA based on a provided document
- Developed proof-of-concept for automatic question answering (QA) with semantic search
- Expanded processing pipeline to larger quantities of internal documents via automatic document retrieval and generative QA with large language models (LLMs)
- Informed the developers about recent advancements in natural language processing (NLP)

Outcome: Made responsive chatbot for document-based extractive QA available to all employees.

transformers
haystack
LLMs
document retrieval
semantic search
question answering
BERT
git
beautifulsoup4

Estimation and analysis of variance spillover networks for academic research

Time series *Academic Research* *05/2020 – 12/2023*

Baseline: Existing network estimation methodology only works reliably on small datasets.

Goal: Apply machine learning methods to analyze large datasets of economic and financial data.

Activities: I realized this project as an *econometrics researcher*:

- Explored and compared statistical learning algorithms for multivariate forecasting
- Acquired datasets from SQL databases and set up an automated pre-processing pipeline
- Wrote object-oriented code to run cross-validated regularized machine learning algorithms
- Conducted extensive statistical and econometric analyses to empirically analyze the results
- Authored research papers that present the results at academic standard

Outcome: Published acclaimed paper in *Quantitative Economics*, follow-up papers under review.

pandas / numpy
time series forecasting
scikit-learn
python
SQL
glmnet
econometrics
vector auto-regression
networkx

STACK