LEI MA

Data Scientist & Ph.D. in Physics

@ leima137@gmail.com | Cologne, Germany | leima.is | emptymalei

EXPERIENCES AND PROJECTS

Lead Data Scientist

Orion Engineered Carbons

April 2024 – Now | Cologne, Germany

- Established and leading a team on BI and data science.
- Initiating and managing data science projects across multiple functional areas inside the company, bringing the power of data science into manufacturing.
- Building a data science platform to enable more employees to use the power of data science.

Applied Scientist in Pricing and Forecasting

Zalando SE

Dec 2021 – April 2024 | Berlin, Germany

- Lead a project building a time series forecasting deep learning framework, enabling the team to experiment fast in deep learning time series forecasting models. The team has been using the framework for publications as well as improving our production demand forecasting model.
- Built new forecasting models and evaluation toolkit for a diverse of sales events from scratch, providing solid foundation for optimal pricing during sales events.
- Prototyping and applying new forecasting and evaluation algorithms to our problems, e.g., initiated a graph neural network seminar, where we experiment with applications of graph neural networks in our forecasting and optimization systems.
- Fine-tuned a llama 2 based large language model for demand forecasting, providing potential generic applications for ad-hoc forecasts in the company.

Data Scientist

Saloodo! GmbH, DPDHL Group

Nov 2019 – Sep 2021 | Cologne, Germany

- Built an automated road freight cost model with a complete ETL, training, and deployment machine learning pipeline for DHL Freight, providing insights and predictions to road freight pricing.
- Built a pricing system for the bidding process on our road freight marketplace, improving the fairness of the marketplace for all carriers and shippers while preserving privacy.
- Built classification models for our imbalanced financial data to help the finance team save 30 hours every month.
- Supply and Demand models of the road freight market to help the management and operations make decisions based on data.

Data Scientist

Homelike Internet GmbH

Aug 2018 – Nov 2019 | Cologne, Germany

- Built ETL pipelines and maintained data warehouses for all the data homelike has, ensuring stable and clean data for our data science projects.

THEORY & TECH

Deep Learning **Statistical Learning**

Neural Networks | Neural ODE | Transformer | Graph Neural Networks | Random Forest | Gradient-boosted Trees | Spiking Neural Networks

Time Series **Dynamical Systems**

Forecasting | Time Series Analysis

Physics

Theoretical Physics | Complex Networks

Python **Mathematica**

SQL **C/C++**

git | pandas | Spark | PyTorch | sklearn

Data Visualization **Dashboard Building**

matplotlib | plotly | PowerBI | streamlit

Data Engineering **Data Scraping**

ETL | BigQuery | Serverless | Data Wrangling | CI/CD | Big Data | API

LLM | Gen AI | Pricing | Logistics | Scientific Computing

LANGUAGES

Chinese **English**

German(A1)

EDUCATION

Ph.D. in Physics

University of New Mexico

2013.09 – 2018.07 | Albuquerque, USA

B.Sc. in Physics

Shandong University
National Science Talents Training Base

2006.09 – 2010.07 | Jinan, China

- Designed and built data products on geo-location-based analysis and forecasting for customers on the supply side, helping them with data-driven decisions.
- Built an automated data-driven marketing system with our CIO to automatically market on google Ads and boost our performance marketing efficiency.
- Built user profiling algorithms and API to assist the pre-sale team and boosted their efficiency in lead management.

Ph.D. in Physics

University of New Mexico

Aug 2013 – May 2018 Albuquerque, USA

- Developed analytical and numerical methods to compute and interpret neutrino oscillations in different environments.
- Proposed a new interpretation for a neutrino oscillation phenomenon, the MSW effect, in the presence of matter inhomogeneity.
- Developed artificial neural network methods to solve neutrino oscillations under the supervised learning scheme.
- Wrote e-books of statistical physics and neutrino physics and helped people to dive deep into modern statistical physics and neutrino physics.

Open Source, Content Creation, and Community

- I create science and data-related software, e.g., some python packages listed under KausalFlow and DataHerb.
- I write about deep learning and time series models in a book Time Series Forecasting with Deep Learning.
- Neuronstar is a community I am running since 2016, establishing connections between Neuroscience, Physics, Complex Networks and Machine Learning through a series of online seminars and reading groups, on topics such as spiking neural networks, energy-based learning, biological neural networks.

RECENT PUBLICATIONS

Book

[1] Graeme Davidson and Lei Ma. *Time Series with PyTorch*. Packt. ISBN: 9781805120421. URL: https://www.packtpub.com/en-ec/product/time-series-with-pytorch-9781805120421.

Journal Articles (non-physics Topics)

[2] Manuel Kunz, Stefan Birr, Mones Raslan, Lei Ma, Zhen Li, Adele Gouttes, Mateusz Koren, Tofigh Naghibi, Johannes Stephan, Mariia Bulycheva, et al. “Deep Learning based Forecasting: a case study from the online fashion industry”. In: *arXiv preprint arXiv:2305.14406* (2023).