

Selina Frank

Data Scientist

✉ selina.frank@example.de
📍 Berlin, Deutschland
🌐 linkedin.com/in/selinafrank

☎ +49 30 4567 8901
🌐 selinafrank.dev
📄 github.com/selinafrank



Profil

Data Scientist mit 4+ Jahren Erfahrung im Marketplace- und Logistik-Bereich bei HelloFresh SE und Delivery Hero SE. Versende-ETA-Modelle, Pricing-Experimente, MLOps. Fluent in DoubleML, EconML, MLflow 2.x, Feast, Snowflake und Statsig. Treibe Geschäftsergebnisse in zweistelliger Millionen-EUR-Höhe mit kausal validierten Pricing- und Logistik-Hebeln.

Berufserfahrung

Data Scientist

HelloFresh SE • Berlin, Deutschland • 06/2023 - heute

Marketplace-Data-Science im Fulfillment- und Pricing-Team

- Entwicklung eines LightGBM + Quantil-Regressions-ETA-Modells, MAE von 4,1 auf 2,7 Minuten auf 9 Mio. tägliche Bestellungen reduziert
- Treiber von +22 Mio. EUR jährlichem GOV durch Tip-Prompt-Redesign, validiert über 28 GrowthBook-A/B-Tests ($p=0,003$)
- Rebuild der Airflow + dbt + Snowflake-Pipelines, Daten-Freshness-SLA von 6h auf 35 min auf 14 Legacy-Jobs verkürzt
- Aufbau eines Feast-Feature-Stores mit p99-Latenz 25 ms für 18 Online-Modelle
- Mentoring von 2 Junior-Data-Scientists, monatliche Modell-Review-Cadence eingeführt

Data Scientist

Delivery Hero SE • Berlin, Deutschland • 01/2021 - 05/2023

Pricing-Analytics im Marketplace-Team

- Aufbau eines XGBoost-Preiselastizitätsmodells (R^2 0,84) auf 38 Mio. Bestellungen, 6 Mio. EUR jährlicher Kontributionsmarge-Hebel
- Statsig-Experimentation-Plattform mit CUPED + sequential testing in 9 Monaten ausgerollt, 42 Tests/Quartal
- Migration von 9 Legacy-Modellen auf MLflow 2.x, Deploy-Lead-Time von 5 Tagen auf 8 Stunden verkürzt

Ausbildung

M.Sc. Statistik

Ludwig-Maximilians-Universität München (LMU)

München, Deutschland

10/2018 - 09/2020

Statistik • 1,2

B.Sc. Mathematik

Universität Heidelberg • Heidelberg, Deutschland • 10/2015 - 09/2018

Mathematik • GPA: 1,4

Fähigkeiten

- Python (PyTorch, scikit-learn)
- XGBoost & LightGBM
- DoubleML & EconML
- MLflow 2.x & KServe
- Feast Feature Store
- Snowflake & dbt • Airflow
- Statsig & GrowthBook

Sprachen

Deutsch • Muttersprache

Englisch • C2

Spanisch • B2

Stärken

Kausales Denken

Wähle zwischen RCT, DoubleML und IPW bewusst auf Basis der Daten- und Stakeholder-Lage

Lieferorientierung

Bringe Modelle in Produktion (MLflow + KServe) statt in Notebooks zu enden

Stakeholder-Übersetzung

Erkläre ATE/ITE für Produktmanager mit klaren Konfidenzintervallen und Risikohinweisen

Projekte

ETA-Modell-Refactoring

- 03/2024 - 08/2024

Migration von Random Forest zu LightGBM + Quantilregression, Latenz p99 von 95 ms auf 38 ms

GrowthBook-Migration

- 10/2022 - 02/2023

Ablösung von hausinternem A/B-Test-Tool durch GrowthBook + CUPED, 4-fache Test-Velocity

Zertifikate

AWS Certified Machine Learning Specialty

- 10/2025

Databricks Certified Machine Learning Professional

- 05/2024

Google Cloud Professional Machine Learning Engineer

- 11/2023