Mohammad Haghjoo



Data Science & Analytics

Werkstudent · Verfügbar ab Dez. 2026 · bis zu 20 Std./Woche · Nürnberg, Erlangen, München oder hybrid

Nürnberg · mohammad.haghjoo@fau.de · mhmmdhaghjoo.github.io · github.com/mhmmdhaghjoo · linkedin.com/in/mohammad-haghjoo-b241823b0

KURZPROFIL

Data Science & Analytics Professional mit fünf Jahren Vollzeit- (sowie sechs Jahren Teilzeit-) Berufserfahrung im Betrieb von PostgreSQL-Datenbanken, automatisierten ETL-Reporting-Pipelines, internen Python-Systemen und angewandtem ML. Schwerpunkt auf Datenqualitäts-Audits, statistischer Analyse und klarem Reporting mit SQL, Tableau Public, Streamlit und R Shiny. Aktuell im M.Sc. International Information Systems an der FAU Erlangen-Nürnberg (B.Sc. Computer Engineering); verfügbar als Werkstudent ab Dez. 2026.

TECHNISCHE KENNTNISSE

Programmierung & SQL	Python (pandas, NumPy, SciPy), SQL, PostgreSQL, R (tidyverse, ggplot2), Kimball-Sternschema
Statistik & Analytik	EDA, Hypothesentests, bayesianische Inferenz, A/B-Testing, RFM- & Kohortenanalyse
BI & Daten-Apps	Tableau Public, Executive Dashboards, Data Storytelling, Streamlit, R Shiny, Plotly
ML & Engineering	scikit-learn, Kalibrierung, Leakage- & Fehleraudits, FastAPI, pytest, mypy, ruff, Git, CI

BERUFSERFAHRUNG

08/2020 – 10/2025

Information Systems Specialist (Datenbanken, ETL, Python-Systeme & ML)
Shiraz Petrochemical Co. · Teheran, Iran · Vollzeit (5 Jahre)

- **Produktivbetrieb & Datenbank-Zuverlässigkeit:** Verantwortete fünf Jahre lang Design und Betrieb von PostgreSQL-Datenbanken (normalisierte Schemata, Query-Tuning, rollenbasierte Zugriffsrechte, geplante Backups) für abteilungsübergreifende Systeme.
- **Automatisierte ETL- & Reporting-Pipelines:** Entwickelte SQL-Integrationsstrecken und Datenschnittstellen, die operative Tabellen in regelmäßige Berichte für interne Fachabteilungen überführten.
- **Python-Unternehmenssoftware & Automatisierung:** Entwickelte und wartete interne Python-Anwendungen (pandas, NumPy) sowie Skripte, die manuelle Datenimporte, Berichterstellung und Backup-Prüfungen ablösten.
- **Angewandtes ML auf Produktionsdaten:** Entwickelte Machine-Learning-Modelle in Python (pandas, scikit-learn) auf petrochemischen Produktionsdaten zur Unterstützung operativer Analysen.

01/2014 – 06/2020

Information Technology Specialist (Datenbanken, SQL-Reporting & Python-Tools)
Tankerranan Petroleum Transport Company · Rafsanjan, Iran · Teilzeit (6 Jahre)

- **Datenbankbetrieb & Stabilität:** Wartete und optimierte über sechs Jahre operative PostgreSQL-Datenbanken (Updates, Indizierung, regelmäßige Backups) für stabilen Zugriff auf Betriebsdaten.
- **Regelmäßiges SQL-Betriebsreporting:** Lieferte strukturierte SQL-Abfragen und wiederkehrende Betriebsberichte als Datengrundlage für laufende Geschäftsentscheidungen.
- **Interne Python-Tools:** Entwickelte interne Python-Skripte und Daten-Tools (pandas, NumPy) zur Automatisierung wiederkehrender Auswertungen neben dem Bachelorstudium.

AUSGEWÄHLTE OPEN-SOURCE-PROJEKTE

Retail Analytics Warehouse & Executive BI `SQL, Kimball-Sternschema & Tableau Public · v0.1.0`
`github.com/mhmmdhaghjoo/retail-analytics-warehouse`

Kimball-Sternschema in SQL und 10-Punkte-Finanzabstimmung auf **541.909 europäischen Einzelhandelstransaktionen** (UCI Online Retail). Entfernt 5.268 Duplikate, trennt 9.251 Retouren und 132.186 Gastbestellungen, stimmt **£10,64 Mio. Brutto- / £9,75 Mio. Nettoumsatz** centgenau ab (**10/10 DQ-Checks bestanden**) und veröffentlicht ein interaktives **Tableau-Public-Dashboard** (RFM, Karte, Kohorten).

Marketing Campaign Decision & Calibration Lab `Tabular ML, FastAPI & Streamlit · v0.1.0`
`github.com/mhmmdhaghjoo/marketing-campaign-decision-lab`

Leakage-geprüftes Tabular ML, Kalibrierung, FastAPI-Scoring und interaktiver **Streamlit**-Entscheidungssimulator auf **45.211 echten Marketing-Anrufen** (UCI Bank Marketing). Deckt `duration`-Leakage auf (0,91 scheinbar → **0,81 ehrlicher ROC-AUC, 0,008 ECE**), prüft Drift und verwandelt **−29,8 T€ Verlust** in **+40,2 T€ Nettogewinn (+164,5 % ROI)**.

Experimentation & Bayesian Inference Workbench (R + Shiny) `Statistik, R & Shiny · v0.1.0`
`github.com/mhmmdhaghjoo/ab-experimentation-workbench-r`

A/B-Test-Prüfung und Entscheidungs-Workbench in **R und Shiny** auf einem reproduzierbaren 20.000-Session-Benchmark. Automatisiert Sample-Ratio-Mismatch-Tests (`χ²`), frequentistische vs. exakte **bayesianische Beta-Binomial**-Posteriors (`P(B>A)=90,7 %`), Power-Planung und **Holm-Bonferroni**-Subgruppenkorrektur gegen Fehlalarme (p=0,026 → 0,102).

AUSBILDUNG

`10/2025 – 03/2028`
`voraussichtlich`

M.Sc. International Information Systems
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Relevante Kurse: Data Science: Foundations, Tools, Applications in Socio-Economics & Marketing; People Analytics; Experimental Behavioral Research in Data Science; Knowledge Discovery in Databases; Pattern Recognition; AI Assistance in the Data Science Value Chain (Seminar); Business & Global Operations Strategy; Introduction to Software Engineering; Free & Open Source Software.

`10/2013 – 07/2020`

B.Sc. Computer Engineering (Software Engineering)
Islamic Azad University, Central Tehran Branch

Relevante Kurse: Grundlagen des Datenbankentwurfs; Datenbanklabor; Ingenieurstatistik & Wahrscheinlichkeitsrechnung; Künstliche Intelligenz; Datenstrukturen; Algorithmenentwurf; Software Engineering I & II.

SPRACHEN

Englisch C1 (IELTS Academic 7,5) · **Deutsch** B1 · **Persisch** Muttersprache · **Türkisch** Gute Kenntnisse