

## **PERSÖNLICHE DATEN**

**Name:** Tom Roßmann

**Stadt:** Bonn, Deutschland

**E-Mail:** [rossmann.tom99@gmail.com](mailto:rossmann.tom99@gmail.com)

**Website:** <https://itxtx.dev/>

**Github:** <https://github.com/itxtx>

**Linkedin:** <https://www.linkedin.com/in/tom-rossmann/>

## **PROFIL**

**Applied Data Scientist** mit starkem mathematischem Hintergrund, derzeit im M.Sc.-Studium der Mathematik und mit praktischer Erfahrung aus einem ML-fokussierten Praktikum. Erfahrung im Aufbau, der Evaluation und Weiterentwicklung von Machine-Learning-Modellen auf realen Datensätzen, insbesondere in den Bereichen NLP, Zeitreihenanalyse und Anomalieerkennung. Versiert in der strukturierten Analyse komplexer Fragestellungen sowie in der nachvollziehbaren Kommunikation und prototypischen Umsetzung praxisnaher Lösungen in Python. Besonders interessiert an Rollen mit Fokus auf Modellqualität, Transparenz und Skalierbarkeit.

## **BERUFSERFAHRUNG**

### **08.2025 – 01.2026 Praktikum Data Science / Machine Learning – Jobmonitor**

*Bertelsmann Stiftung*

- Evaluierte moderne Embedding-Modelle (u. a. Jina v3, multilingual-e5) zur semantischen Analyse von Stellenanzeigen und bereitete Ergebnisse visuell und erklärbar auf (UMAP, SHAP).
- Entwickelte und evaluierte einen hierarchischen Klassifikationsansatz mit über 1.000 Klassen auf Basis von PyTorch und Hugging Face (PEFT / LoRA).
- Prototypisierte mathematische Modelle zur Abbildung beruflicher Übergänge und implementierte diese performant auf Embedding-Distanzen.
- Nutzte Databricks und R Shiny zur Analyse und Visualisierung arbeitsmarktbezogener Daten.
- Dokumentierte Modellarchitekturen und Evaluationen strukturiert und präsentierte Ergebnisse für technische und nicht-technische Stakeholder.

## **AUSBILDUNG**

2025 - 2027

**Master of Science in Mathematik**

FernUniversität in Hagen

Schwerpunkte: Parametrische Statistik, Nichtlineare Optimierung.

2020 - 2024

**Bachelor of Science in Mathematik**

Linnaeus University, Växjö (Schweden)

### **Bachelorarbeit: Kernel Methods for Regression, 2023**

- Vergleich von Kernel-Regression und Ridge-Regression hinsichtlich Generalisierungsleistung und Trainingsaufwand.
- Analyse des Einflusses verschiedener Kernelfunktionen auf Genauigkeit und Skalierbarkeit.  
URN: urn:nbn:se:lnu:diva-126158

## **PROJEKTE (AUSWAHL)**

### **Erkennung von Kreditkartenbetrug, 2025**

- Entwickelte überwachte und unüberwachte Modelle zur Betrugserkennung unter extremer Klassenungleichheit.
- Implementierte Resampling-Strategien und evaluierte Modelle mit prázisionsorientierten Metriken.

### **Projekt zur Solarprognose, 2025**

- Entwickelte ein LSTM-basiertes Zeitreihenmodell zur Vorhersage von Solarerträgen.
- Führt systematische Hyperparameteroptimierung sowie robuste Fehleranalyse und Unsicherheitsbewertung durch.

### **End-to-End-Projekt: Cloud-Natives Churn-Modell mit MLOps-Pipeline, 2025**

- Entwickelte ein End-to-End-Churn-Modell inklusive API und Deployment.
- Implementierte CI/CD-Pipelines für Training, Tests und Deployment.
- Etablierte Experiment-Tracking und reproduzierbare Modellartefakte.

## **KENNTNISSE UND FÄHIGKEITEN**

- **Programmiersprachen & Datenbanken:** Python (sehr gut), R, SQL, Matlab (Grundlagen), SQLite, MongoDB
- **Data Science & Machine Learning:** Scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, SciPy, XGBoost, Optuna, Polars

- **Deep Learning & NLP:** PyTorch, Hugging Face (Transformers, Sentence-Transformers)
- **Methoden & Statistik:** Regression & Klassifikation, Kernel-Methoden, Ensemble Learning, Clustering, Dimensionsreduktion, Zeitreihenanalyse, Anomalieerkennung, Statistische Modellierung, Hypothesentests
- **MLOps & Cloud:** Docker, MLflow, DVC, FastAPI, Databricks, CI/CD, ML-Pipelines, API-Deployment, AWS, Azure (Grundlagen)
- **Tools & Visualisierung:** Git/GitHub, Jupyter, RStudio, R Shiny

## **SPRACHEN**

Deutsch (Muttersprache) • Englisch (fließend)

Stand: Januar 2026