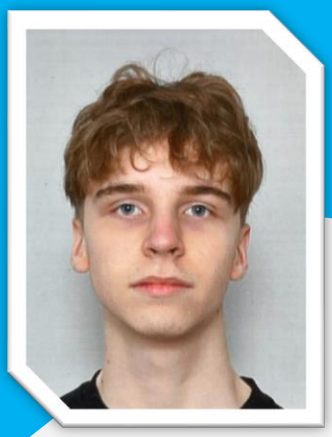

Robin (R.P.M.) Kras

Contact
Ypenburgbocht 9-A303
2316 WB
Leiden, Zuid-Holland
Netherlands

robkraseu@gmail.com
+31 6 27390907

Socials
robkras.com
GitHub: rbkrs
Kaggle: robkraseu

Languages
Dutch – Native
English – Bilingual
German – Basic
French – Basic

Certifications
English C1 Advanced
May 2018
Issued by Cambridge Assessment International Education

EDUCATION

Estimated finalisation Mid 2025

MSc Computer Science, Rijksuniversiteit Leiden

- Data Science & Artificial Intelligence specialization track

Dec 2023

BSc Computer Science, Vrije Universiteit Amsterdam

- Minor: Data Science

SKILLS

Programming: Python, C, C++, Assembly, SCALA

Technical skills: Keras, NumPy, Pandas, TensorFlow, PyTorch, (My)SQL, data manipulation, data visualization, machine learning, data science, reinforcement learning, GIT, web scraping, data mining, NLP, Hugging Face transformers, SciPy, Scikit-Learn, Librosa, CNN/RNN/GAN), NetworkX, spaCy

Personal skills: love of learning, time management, communication, adaptability

HOBBIES

Swimming, cooking, video games, working out

PROJECTS

Kaggle

Competition entries are regularly updated and stored on my domain (robkras.com).

Applied Skills & Techniques

Machine Learning

- Developed **supervised learning models** (XGBoost, kNN, Random Forest, Linear Regression) to predict rainfall and classify Titanic survival outcomes.
- **Optimized models** using GridSearchCV and KFold cross-validation, achieving improved accuracy and efficiency.
- Engineered new features, handled missing data, and applied **one-hot encoding & label encoding** for categorical variables.
- Trained and fine-tuned **deep neural networks** using TensorFlow/Keras.

Data Processing & Analysis

- Conducted **Exploratory Data Analysis (EDA)** using Seaborn & Matplotlib to identify trends and correlations.
- Applied Matplotlib and Seaborn to discover **variability and outliers** in numerical and categorical feature distributions.
- Cleaned and preprocessed datasets using **Pandas & NumPy**, ensuring high-quality input data.
- Scaled numerical features using **StandardScaler** to improve model convergence.

Model Evaluation & Interpretability

- Assessed models with **RMSE, R^2, MAE, accuracy, and ROC-AUC scores** for performance benchmarking.
- Applied **SHAP values** for explainability and feature importance analysis.
- Used **SMOTE** to balance imbalanced datasets, improving prediction robustness.

Master's Thesis

I am conducting research on cutting-edge VLMs such as Llama3.2 and Molmo to assess their capabilities in drawing cross-modal associations. This work contributes to practical applications, such as improving tools for individuals with disabilities.