

+49 163 9078036
Karlsruhe | Vienna
roshanprizak@gmail.com
roshanprizak.com

ROSHAN PRIZAK

DATA SCIENTIST

 <https://www.linkedin.com/in/roshan-prizak-05a185254/>

About me

Skilled data scientist with 9+ years experience in data analysis, statistical analysis, data modelling, data visualisation, research paper writing and collaborative research projects. Strong quantitative background in pure and applied mathematics, statistics, data analysis, and computer science. Driven to learn new tools and concepts, actively building an independent data science portfolio on real-world problems.

General Skills

Data scraping

Data cleaning

Data mining

Data modeling

Data visualization

Hypothesis testing

Statistical analysis

Clustering

NLP

PCA

Regression analysis

Markov Chain Monte Carlo

Deep learning

Machine learning

Technical Skills

Programming/libraries:

Python, R Programming, R Markdown, R Shiny, MATLAB, C++, Bash, SQL, D3.js, HTML, CSS, Javascript, Vue.js, Quasar, scikit-learn, Tensorflow, matplotlib, ggplot2, Plotly, Leaflet

Other software/tools:

Firebase, Node.js, AWS, Canva, GitHub, git, LaTeX, conda, Docker

Soft Skills

- Collaboration
- Communication
- Multi-tasking
- Leadership
- Decision-making
- Critical thinking
- Quick learner
- Out-of-box thinking

Data Science Projects

How accessible is Vienna's public transport?

- Using GTFS feeds, database of addresses and a local Docker-based installation of openrouteservice, built a dataset of closest public transport
- Visualized results in a R Shiny interactive app, and deployed it on AWS.
- Next, using routing and real-time info, quantify first-mile and last-mile connectivity, robustness to cancellations, redundancy and other metrics

NLP analysis of the resolutions of the World Health Assembly

- Using NLP techniques and network analysis, uncovered topics and patterns behind the resolutions adopted by the sessions of the WHA
- Created a clean dataset of WHA resolutions from all WHA sessions by obtaining PDFs, converting to text using PyPDF2 and cleaning it

Age gaps between lead male & female actors in Indian cinema

- Extracted and clean data from Wikipedia and other public sources
- Visualized the age gaps using matplotlib

Looking at cricket through numbers

- Extract ball-by-ball details - batter, bowler, fielding, events, commentary - from ESPNcricinfo, ICC and other sources and prepare a clean dataset
- Translate qualitative questions about player strategies, performances, umpire biases, team predictions etc. into quantitative insights

Positions



Karlsruhe Institute of Technology

Postdoctoral scientist 

Mar 2019 – Present

Key responsibilities:

- **Image and data analysis** of complex microscopy datasets
- Bioinformatics analysis of **large-scale genomic data**
- **Physical modeling** and polymer simulations of chromatin, followed by **data analysis of simulation output data** in Python
- Worked as part of multiple research teams, contributed to various projects and **co-wrote published research papers**
- Supervised students, **taught R and Python** to Master and students



Austria Model World Health Organization

Vice-President

Aug 2022 – Present

Key responsibilities:

- **Co-organised** simulations of the World Health Assembly
- **Designed and deployed website** at austriamodelwho.org
- **Co-wrote theme guides** for Austria Model WHO simulations
 - Energy Security and Health (October 2022)
 - Climate Change and Health (April 2023)
- Admin roles - **delegate training** during simulations, **dissemination material design**

+49 163 9078036
Karlsruhe | Vienna
roshanprizak@gmail.com
roshanprizak.com

ROSHAN PRIZAK

DATA SCIENTIST

 <https://www.linkedin.com/in/roshan-prizak-05a185254/>

Teaching

Data Analysis, Statistics and R

as part of Chemomics and Phenomics
Instructor (MFOR-3209, KIT, 2022)
Instructor (MFOR-3209, KIT, 2021)

Making graphs with Python

as part of Python Applications
Instructor (BIF-IGS course, KIT, 2023)
Instructor (BIF-IGS course, KIT, 2022)

Selected Topics in Biophysics

Teaching Assistant (IST Austria, 2014)

Network Theory (EE 224)

Teaching Assistant (IIT Bombay, 2012)

Courses

- Data analysis
- Statistical physics
- Statistics
- Biophysics
- Bioinformatics
- Core electrical engineering
- Pure and applied math
(number theory, real and complex analysis, linear algebra, DDEs, PDEs, probability theory, optimization, Markov processes, Markov chains)
- Machine learning and AI
- Systems and networks

Languages

Fluent/native

English Telugu Hindi

A2+

German

Tamil

Malayalam

A1+

Spanish

Kannada

Education Background



Institute of
Science and
Technology
Austria

Institute of Science and Technology Austria

PhD student

Sept 2013 – Dec 2018

Bioinformatics, biophysics, evolutionary theory

Deep learning

Machine learning

Data modeling

Physical modeling

Statistical analysis

Hypothesis testing

Clustering

R Programming

MATLAB

Markov chains

Markov processes

Markov Chain Monte Carlo



Indian Institute of Technology Bombay (GPA 8.63/10)

Master of Technology

Bachelor of Technology

July 2008 – Aug 2013

Electrical Engineering (Communication & Signal Processing)

Data modeling

Physical modeling

Markov chains

Markov processes

Statistical analysis

Probability theory

Control theory

Systems theory

Network theory

Markov Chain Monte Carlo

Fourier transforms

Internships

Institute of Science and Technology Austria

May 2012 – July 2012

University of Surrey, UK

May 2011 – July 2011

Jawaharlal Nehru Centre for Advanced Scientific Research, India

May 2010 – July 2010

Awards and honours

Heinz Scholarship for graduate education (IST Austria, 2013 – 2016)

Rajiv Gandhi Science Talent Research Fellow (Bangalore, 2010 – 2011)

CBSE Merit Scholarship for undergraduate education (India, 2008 – 2010)

Ranked 125 out of 300,000 applicants in **IIT-JEE** (India, 2008)

Gold Medal, top 25 at **camp for International Chemistry Olympiad** (India, 2008)

Among **top 42** in the **Indian National Chemistry Olympiad** (India, 2008)

Top 1% in the **National Science Examination in Physics** (India, 2008)