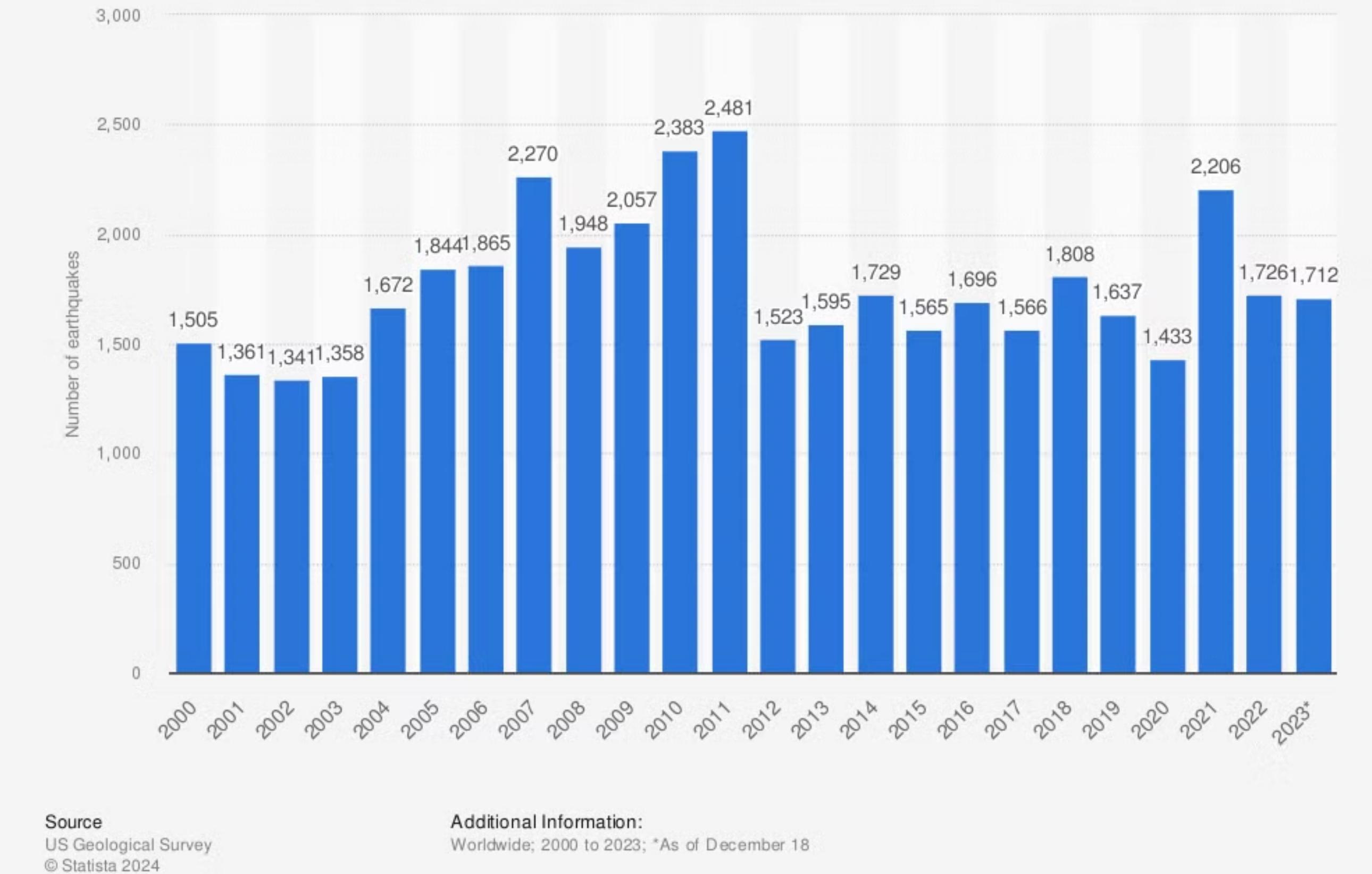






## Development of the number of earthquakes (M5+) worldwide from 2000 to 2023



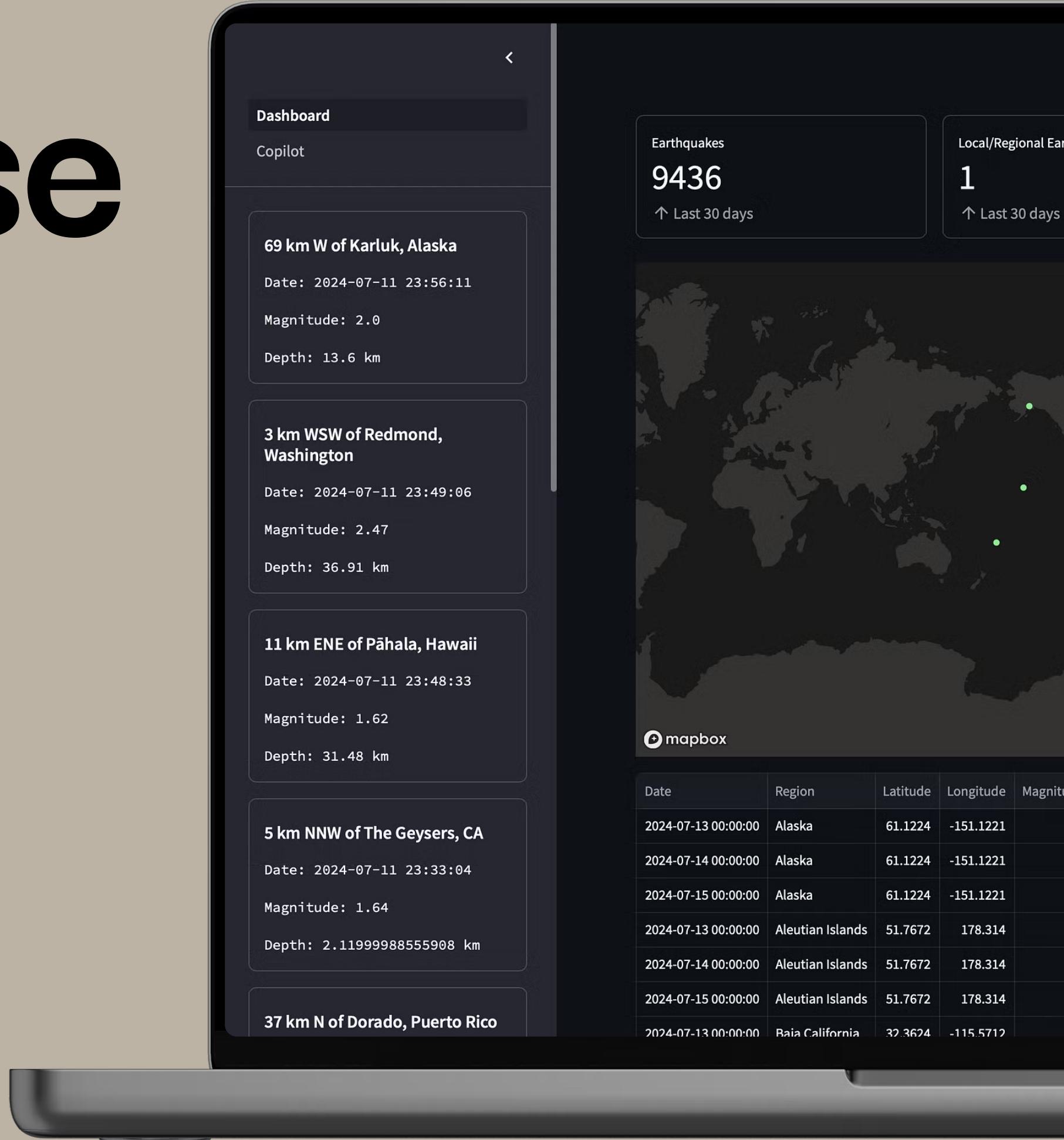
# SeismicSense

## Forecasting Earthquakes

Advanced Machine Learning

André Gilbert, Felix Noll, Marc Grün

Try Pitch

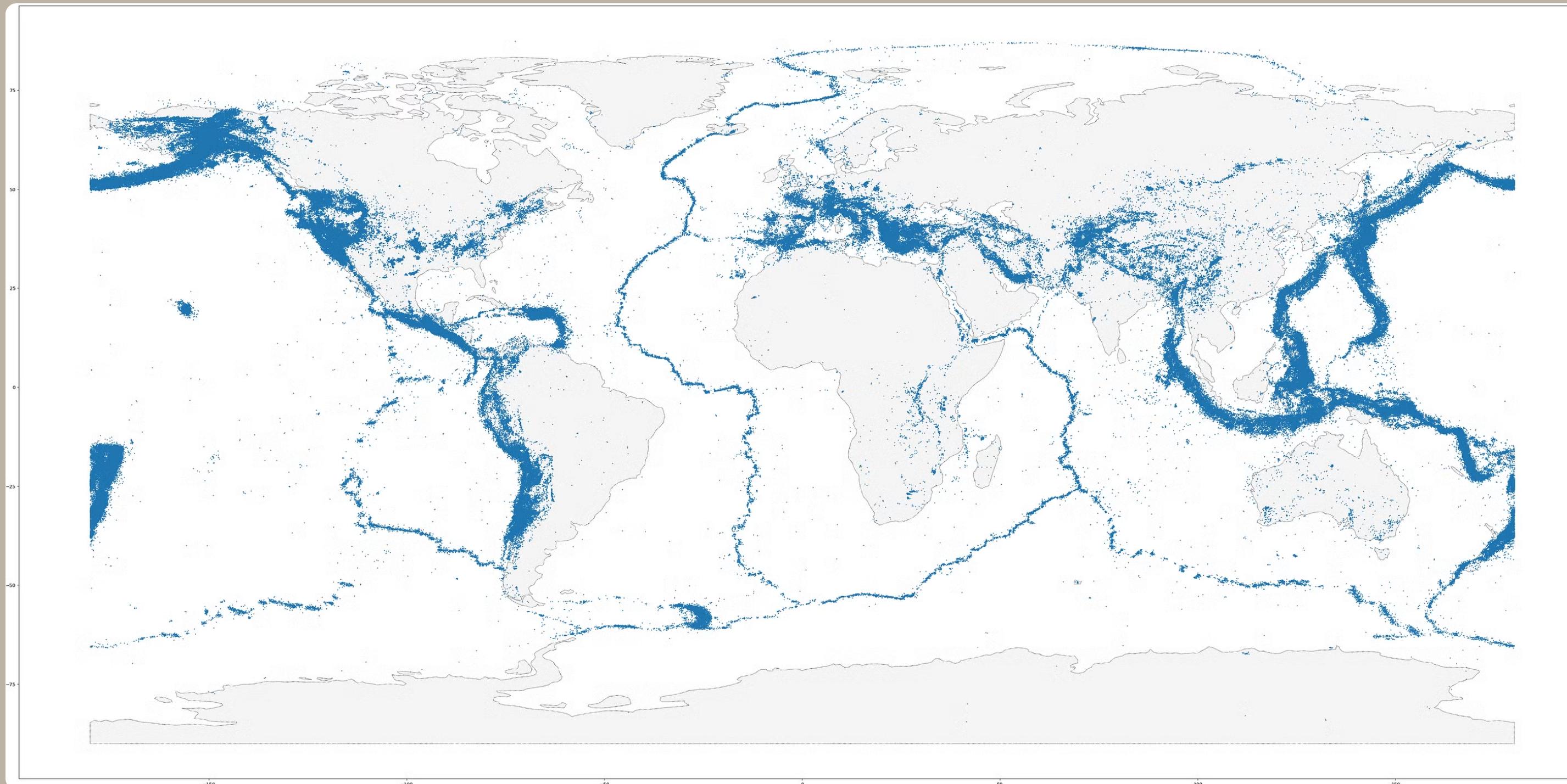


# Agenda

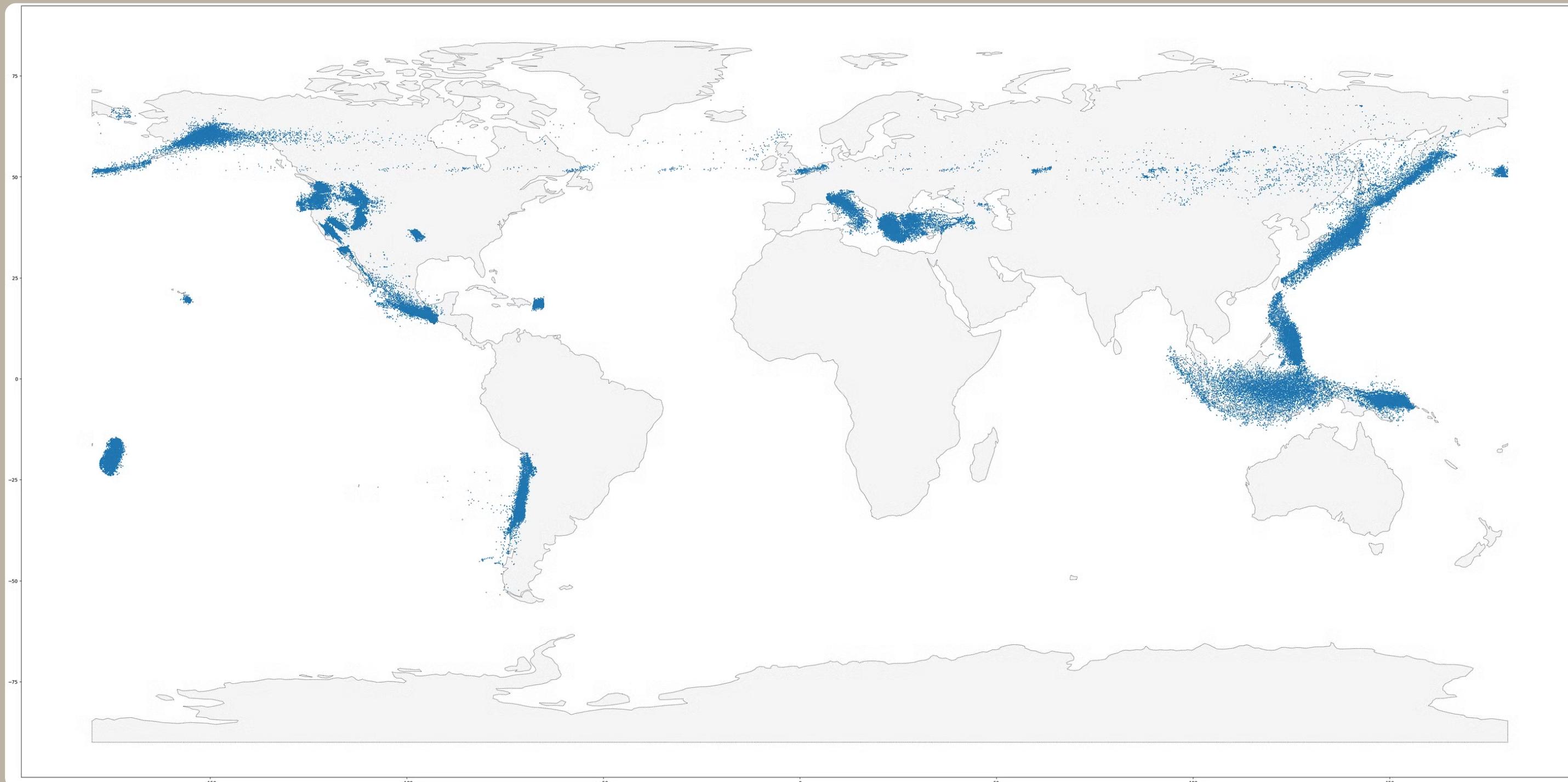
1. Daten
2. Modell
3. Resultat
4. AI Agent
5. Demo



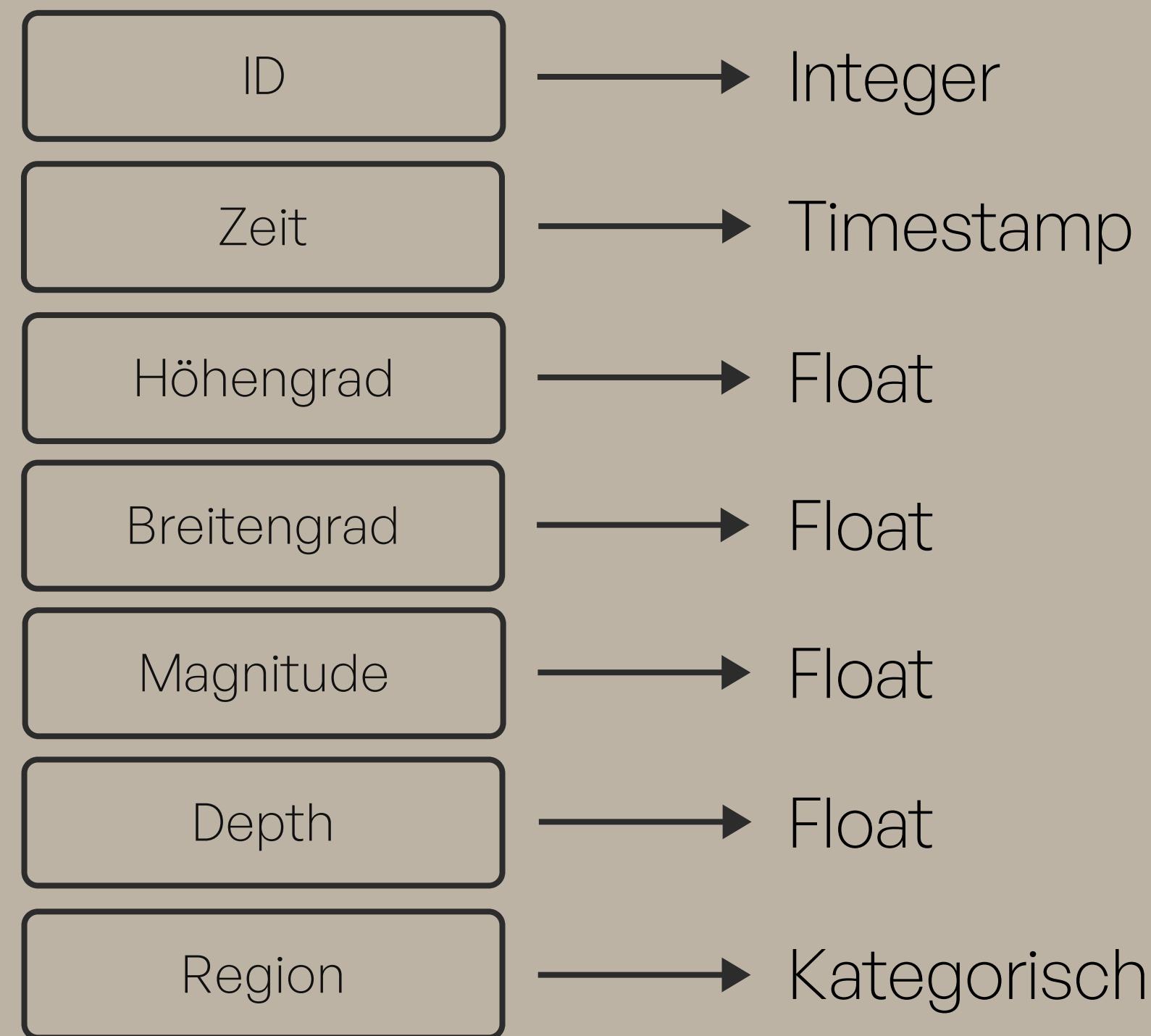
# Globale Verteilung von Erdbeben



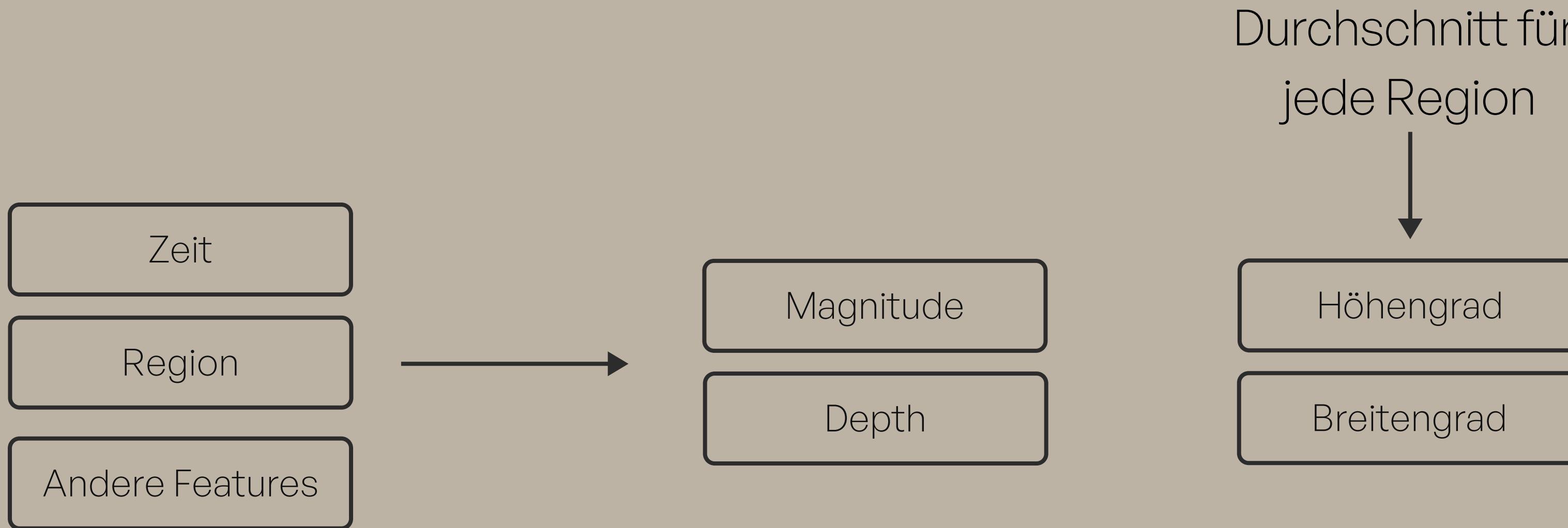
# Erdbeben in den Top 25 Regionen



# Datenstruktur der USGS API



# Modeling Problem



Vorhersage pro Region

# Feature Engineering

7 Lags



Timestamp  
aggregation

Timestamp	Feature_1	Lag
1	100	NaN
2	150	100
3	125	125

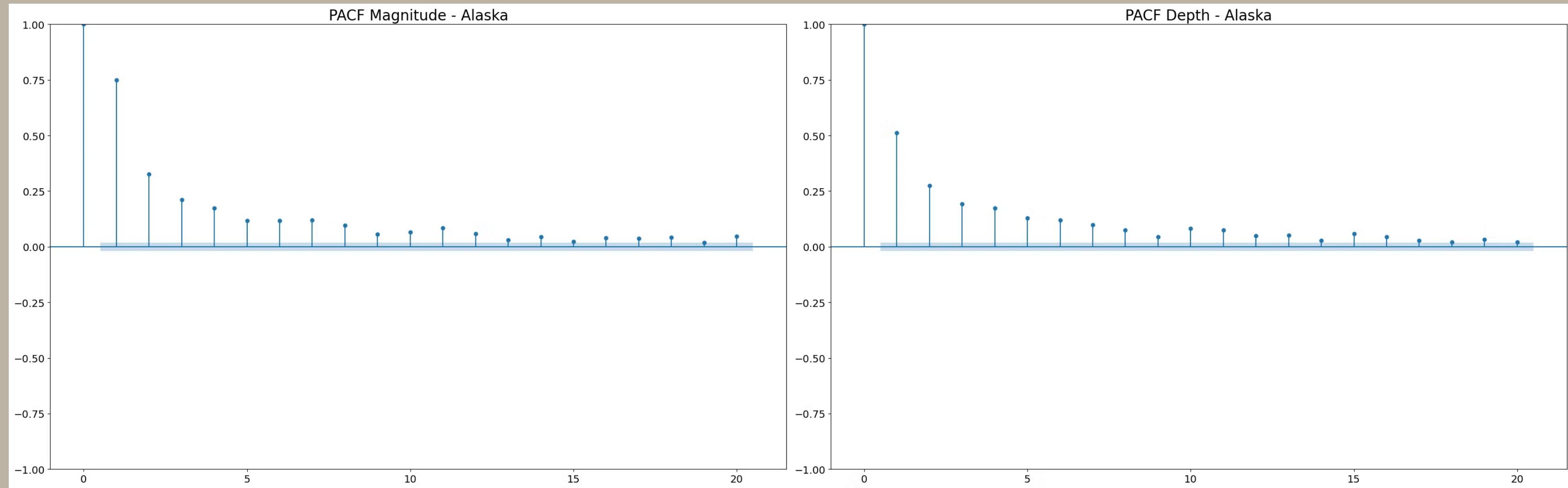
EWMA von 7



$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot S_{t-1}$$

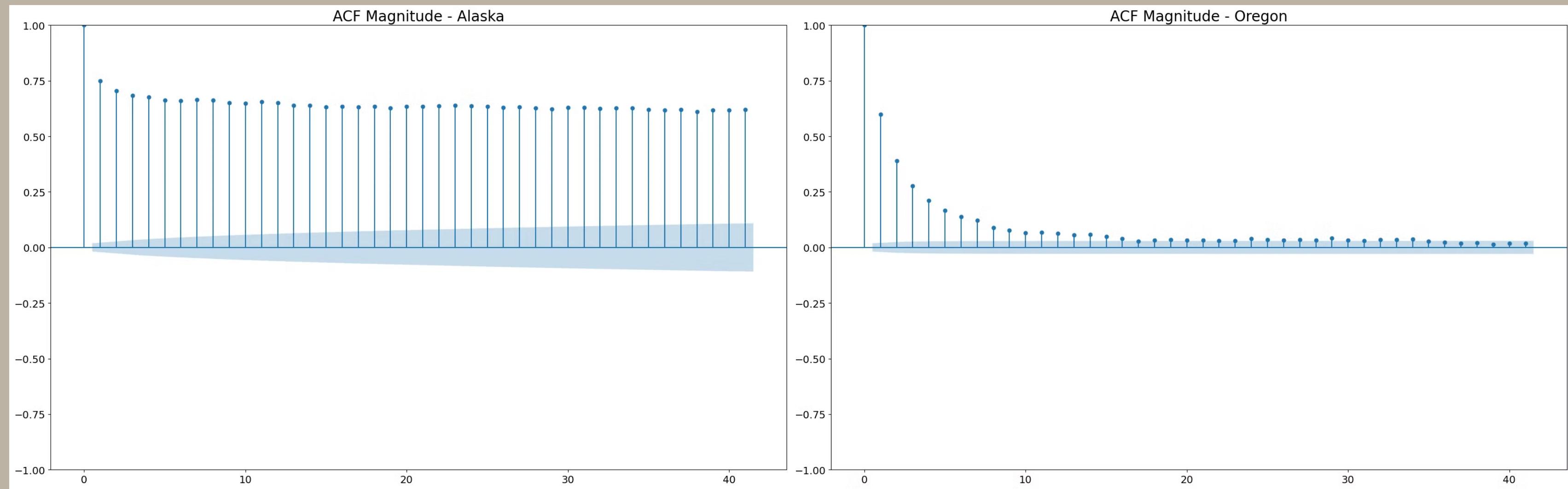
# Feature Engineering

## PACF Magnitude und Depth für Alaska



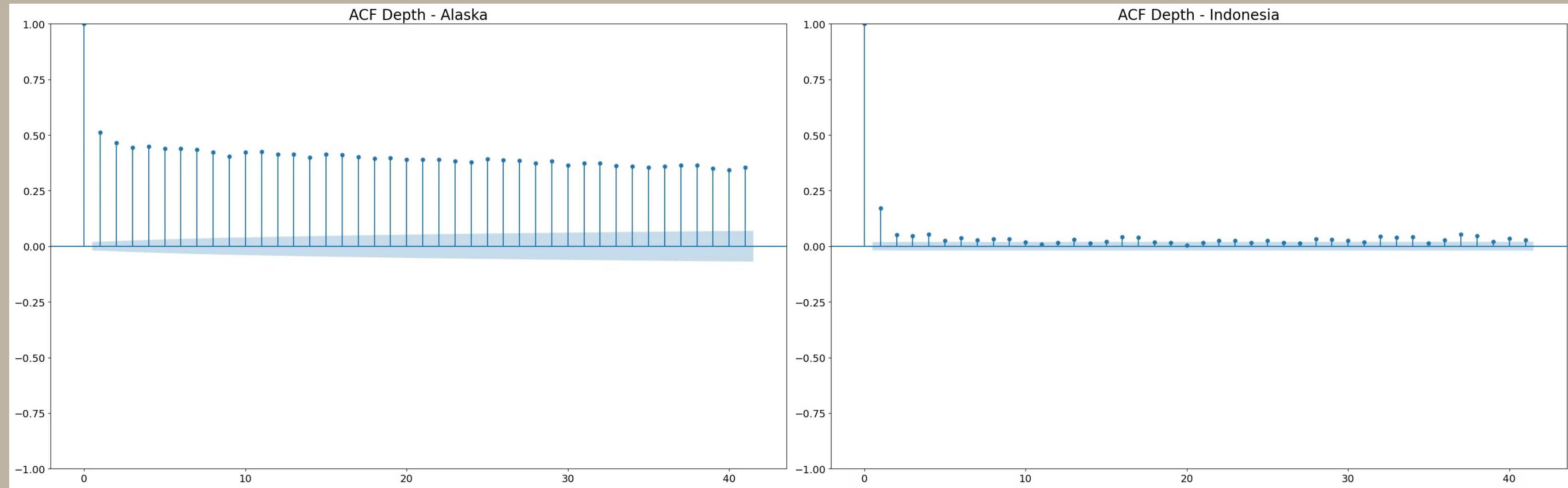
# Feature Engineering

## Autocorrelation Magnitude Alaska und Oregon



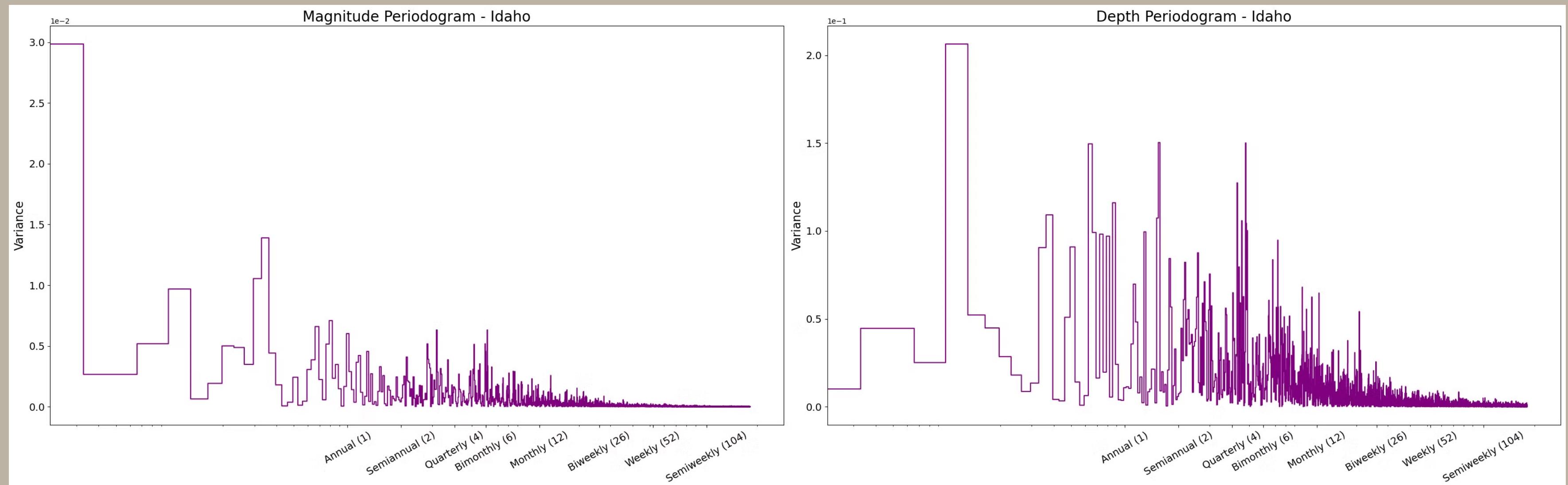
# Feature Engineering

## Autocorrelation Depth Alaska und Indonesia



# Feature Engineering

## Periodogram Magnitude und Depth in Idaho



# Model



## Yandex CatBoost

Automatische  
Umwandlung der  
kategorischen Features

Bessere Möglichkeiten der  
Interpretation

Schnell und wenig  
Overfitting

# Konvertierung kategorischer Features



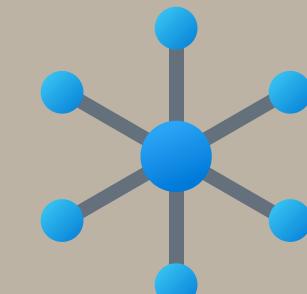
$$ctr_i = \frac{countInClass + prior}{totalCount + 1}$$

**countInClass:** Wie oft der Labelwert bei Objekten mit dem aktuellen Eigenschaftswert i überschritten wurde

**totalCount:** Anzahl der Objekte mit einem gleichen Feature Value

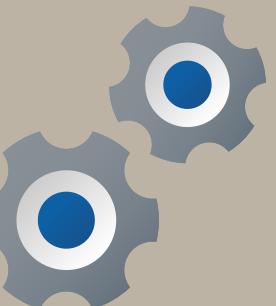
**prior:** Konstante, die als Parameter definiert ist z.B. 0.05

# CatBoost Setup



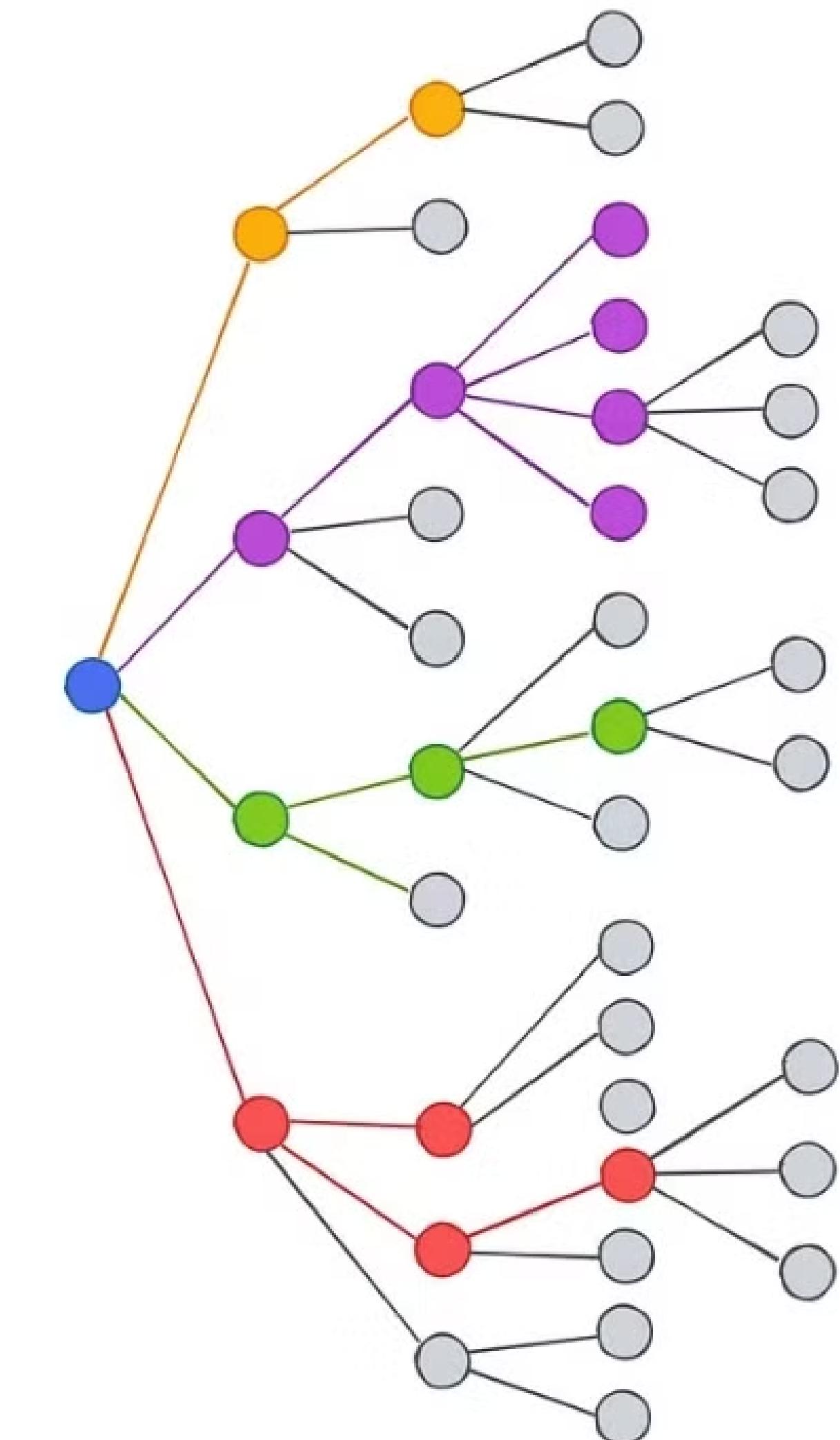
## Genutzte Techniken

- GridSearch
- Early Stopping
- MultiRMSE



## Parameter

- Tiefe : 10
- Iterations : 1000
- Lernrate : 0.1
- L2 Leaf Regularization : 5



# Resultat

2,9344

Mean Absolute Error

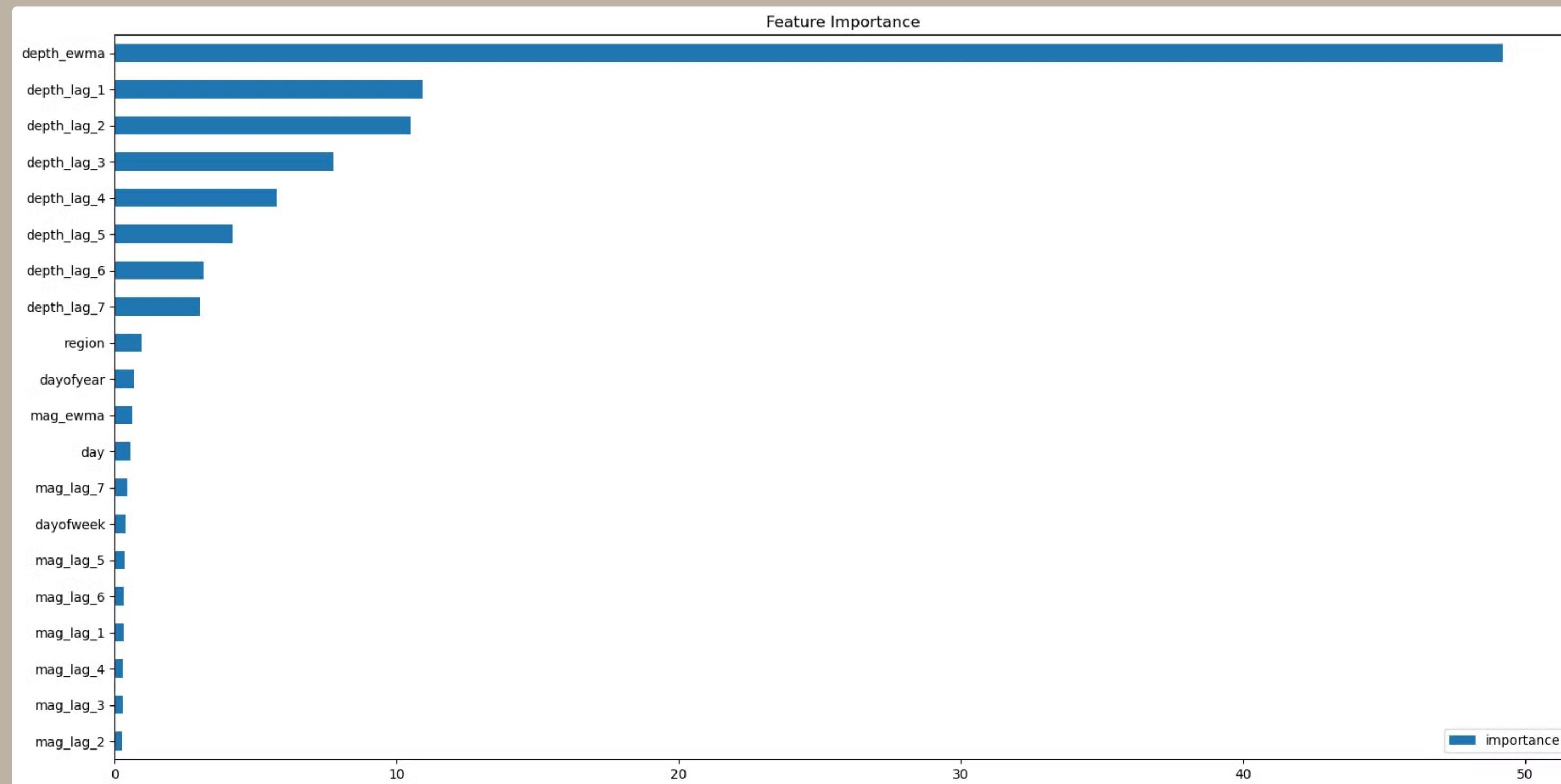
10,3577

Root Mean Squared Error

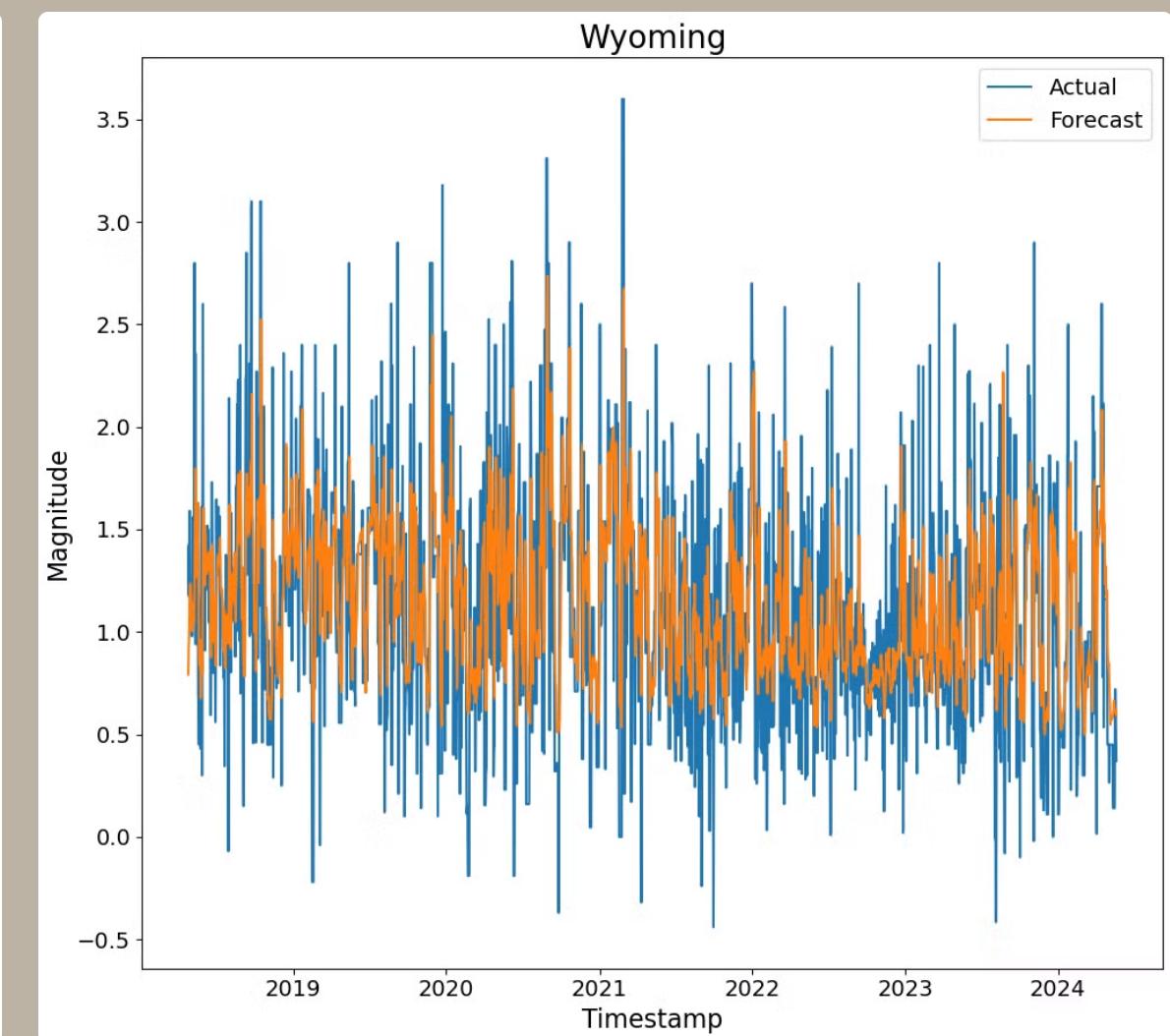
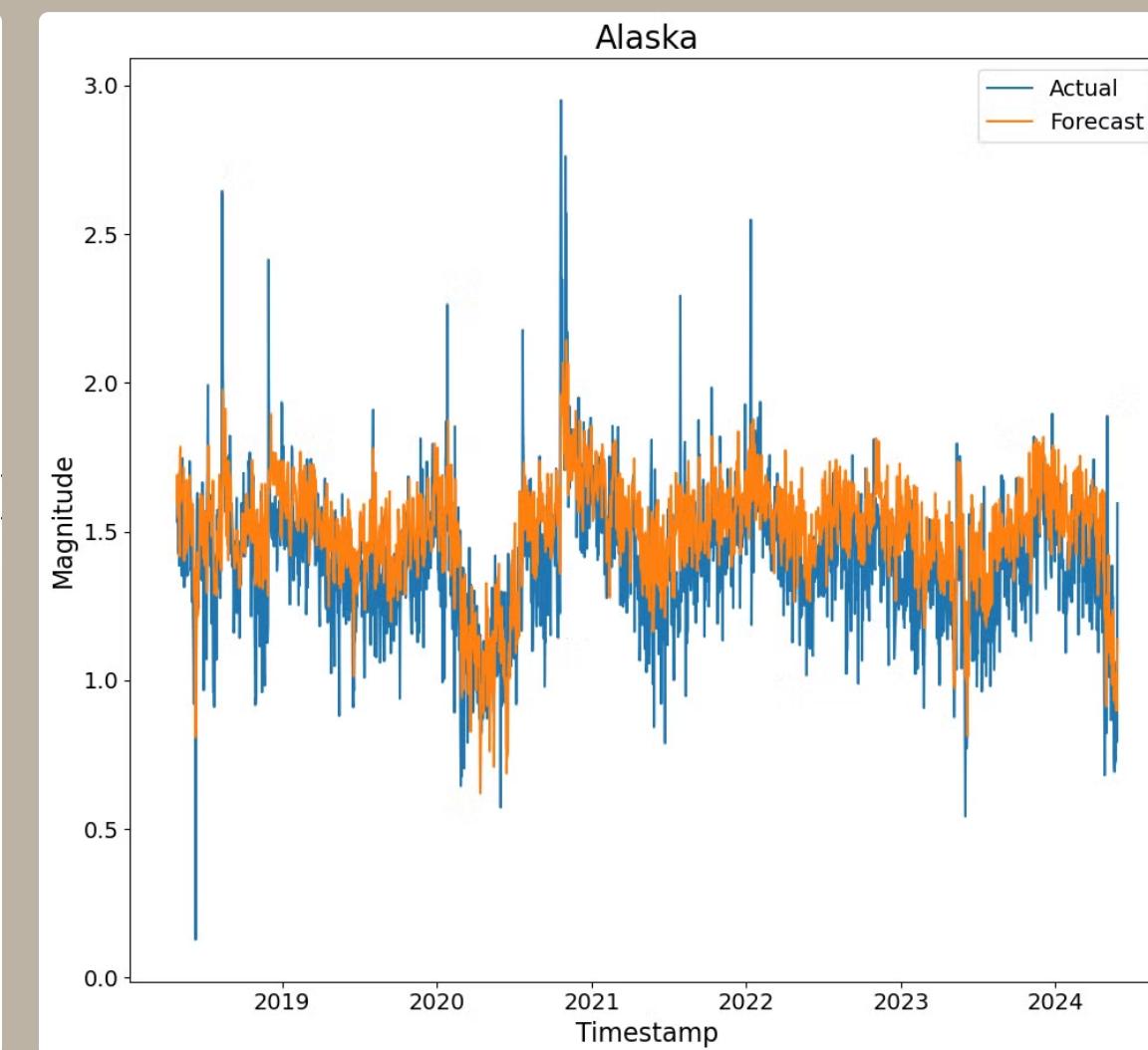
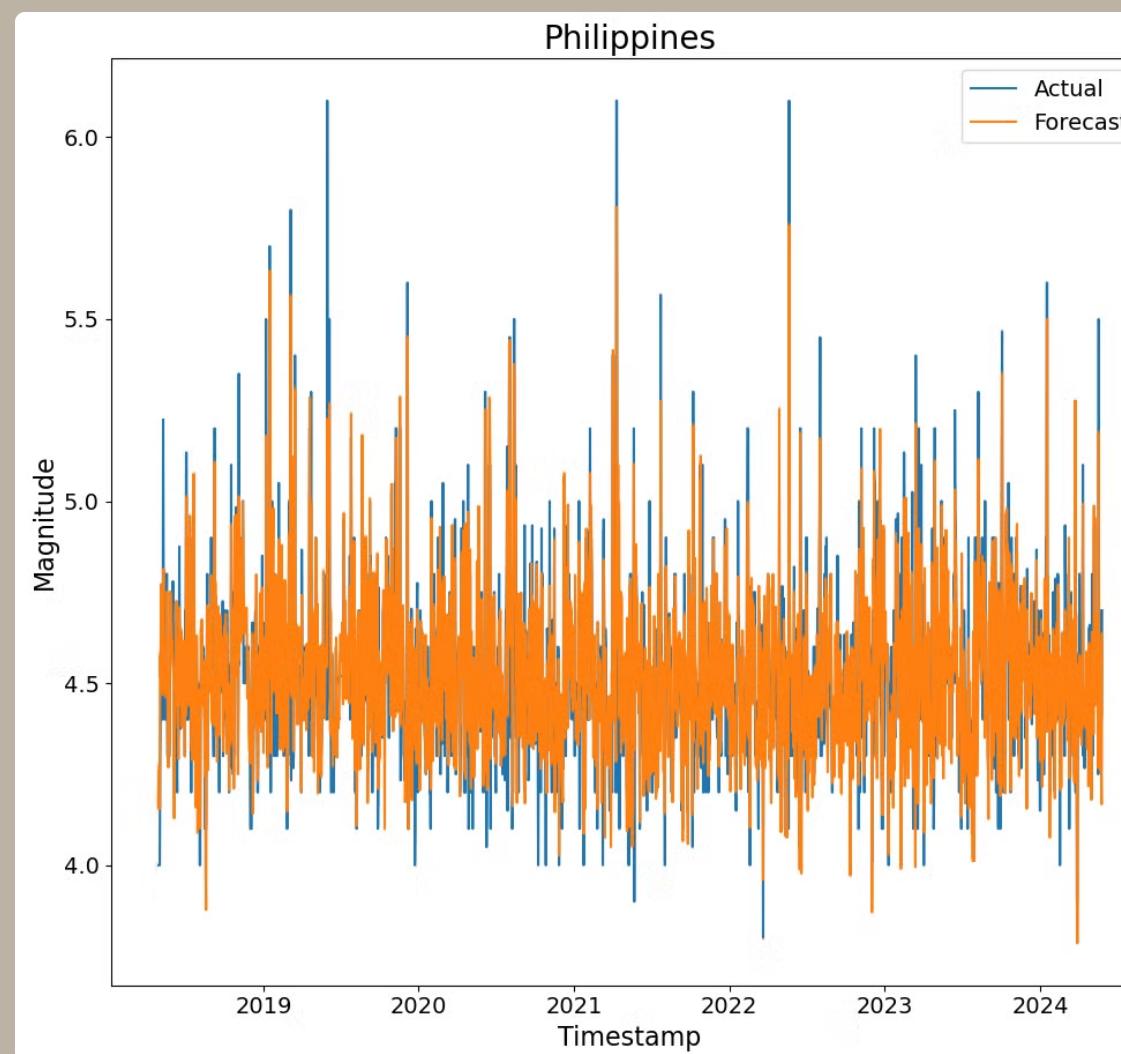
0,9495

R2 Score

# Feature Importance



# Magnitude Forecast

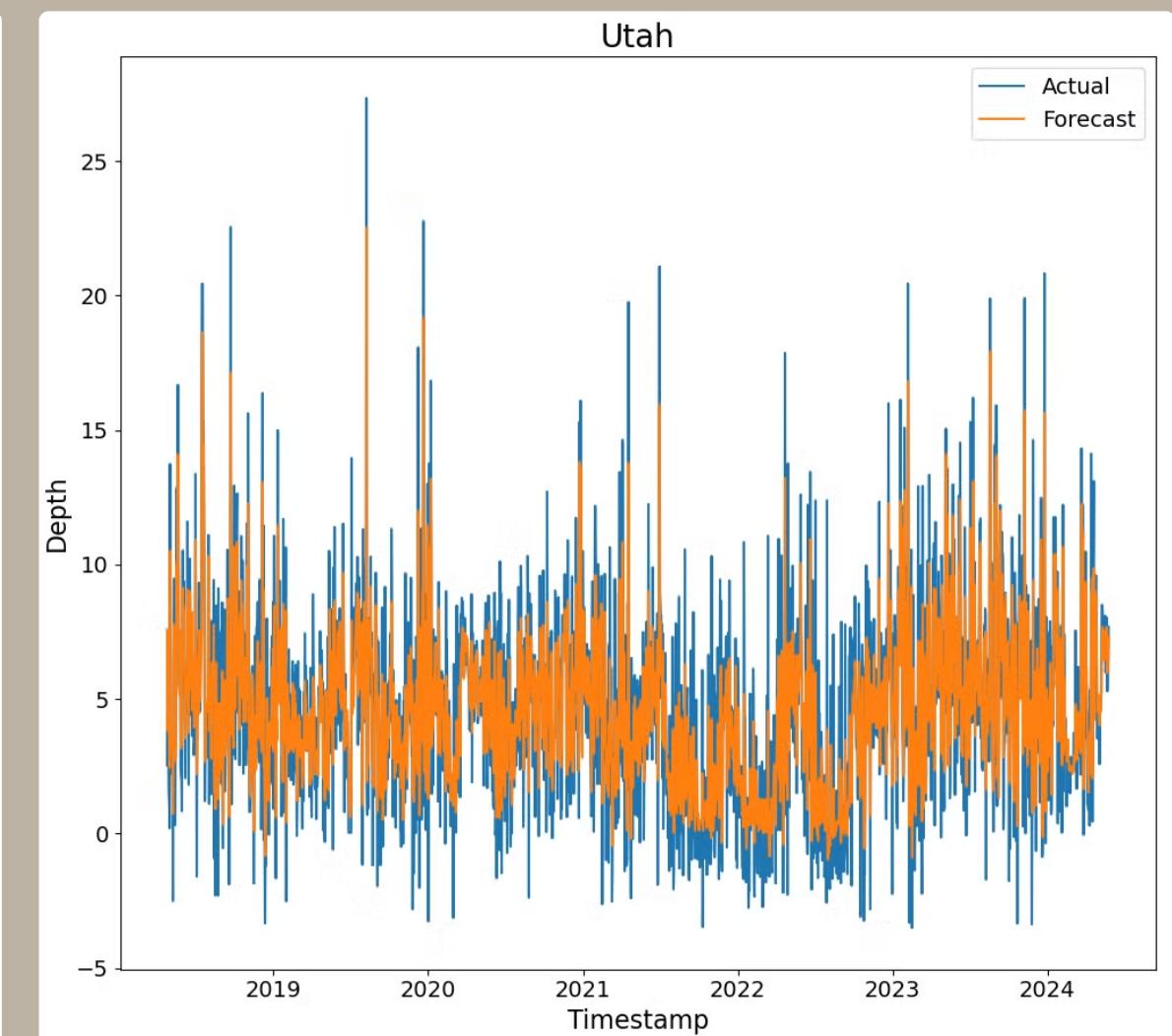
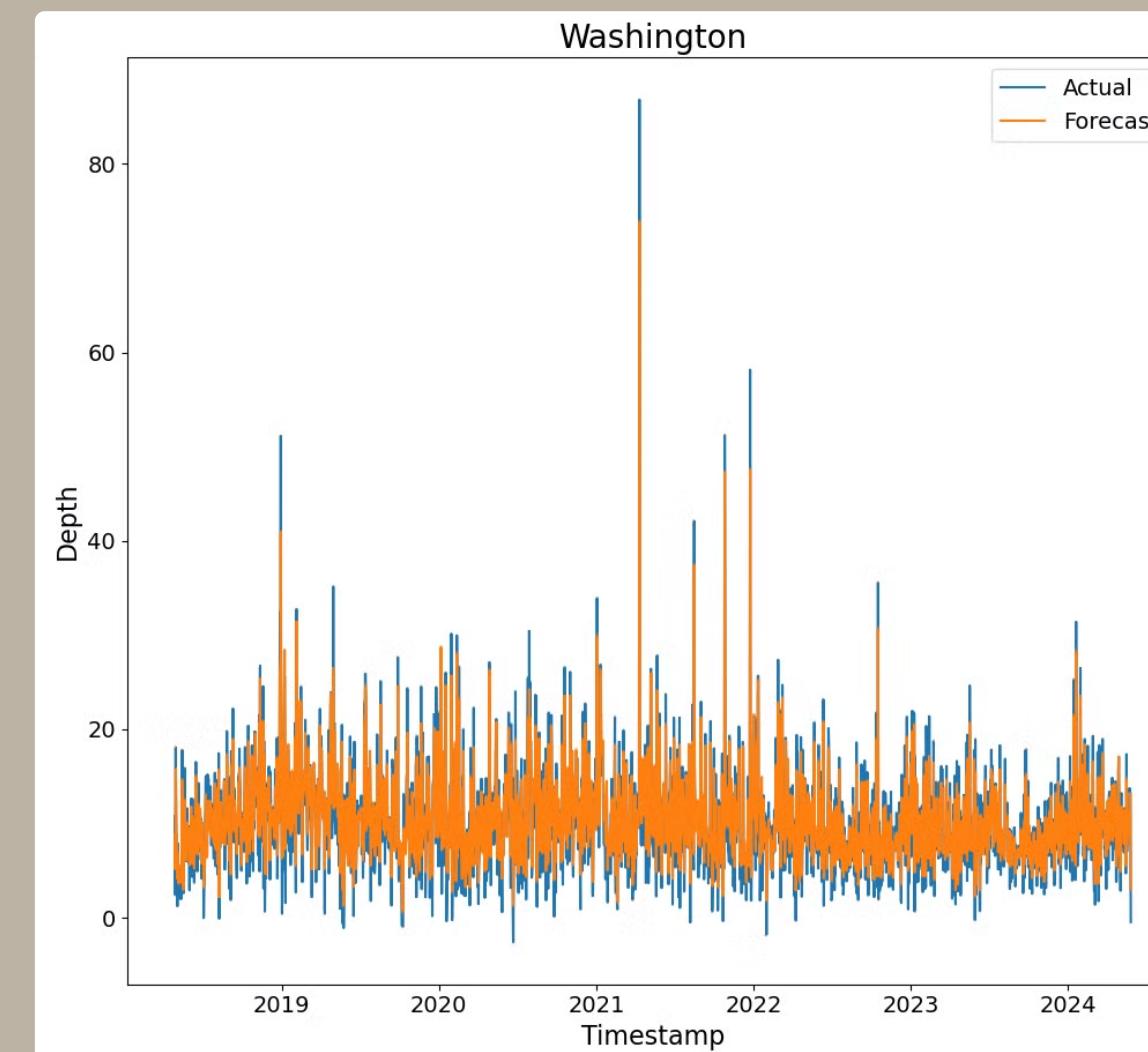
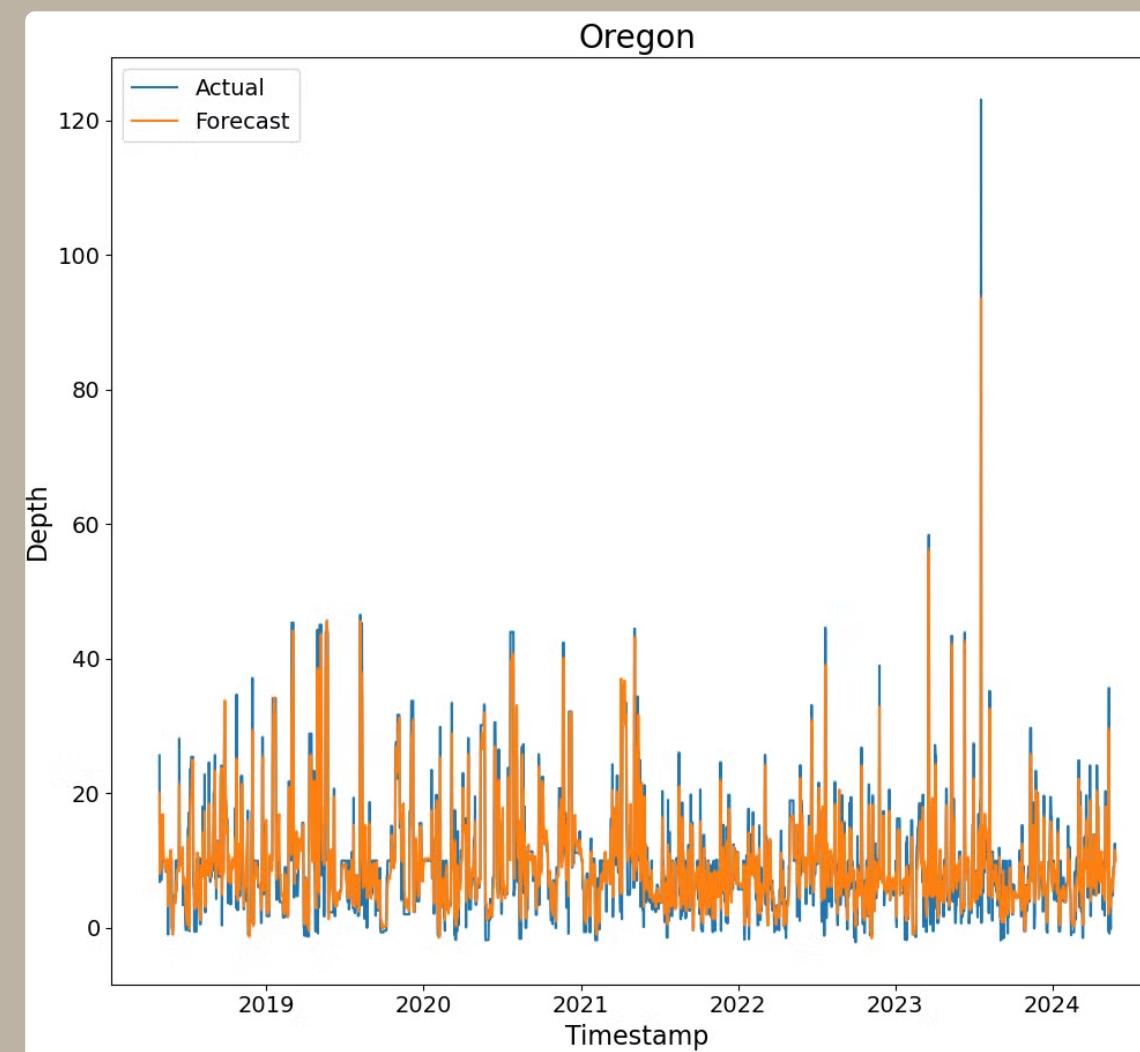


Gut

Durchschnittlich

Suboptimal

# Depth Forecast



Gut

Durchschnittlich

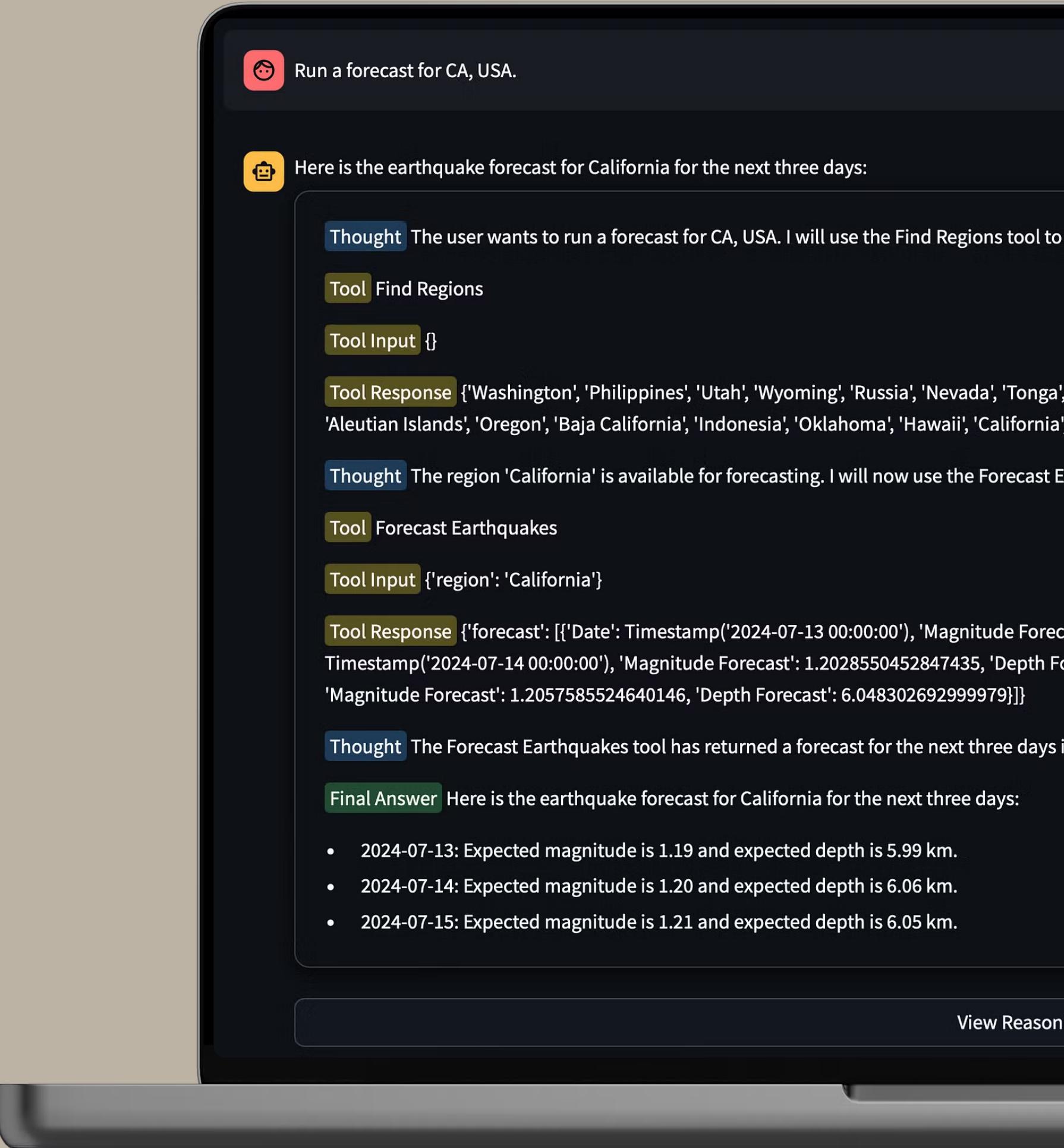
Suboptimal

# LLM Agent

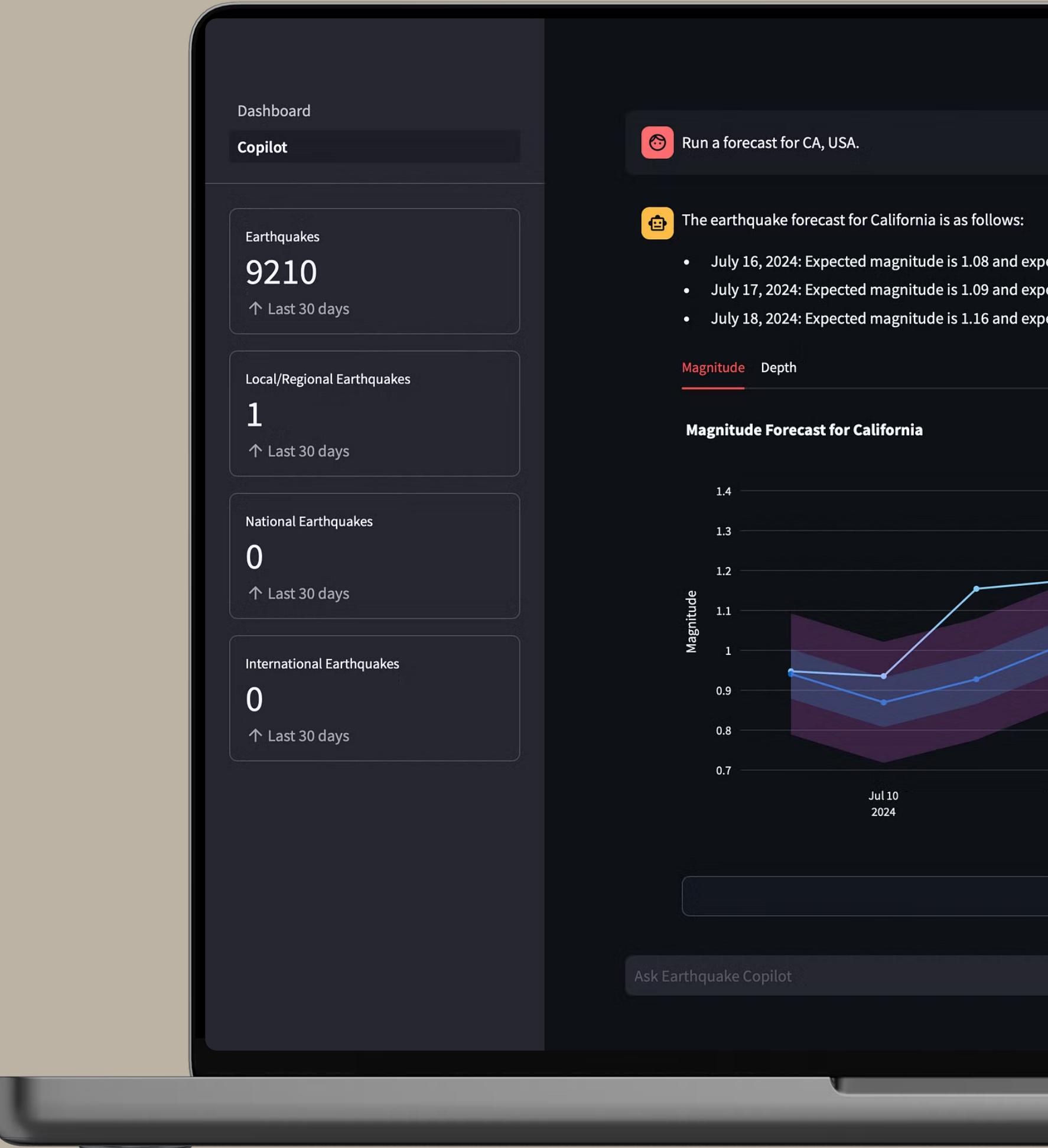
## ① Einfache Nutzung

## ② Besser zu verstehen

## ③ Weitere Informationen



# Demo



# Quellenverzeichnis

Spektrum.de (2024) ‘In jeder Generation gibt es ein Erdbeben in Deutschland’, Verfügbar unter:  
<https://www.spektrum.de/news/in-jeder-generation-gibt-es-ein-erdbeben-in-deutschland/1648960> (Zugriff am 23. Juli 2024)

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K. and Cao, Y., 2022. React: Synergizing reasoning and acting in language models. *arXiv preprint arXiv:2210.03629*.

USGS Earthquake Hazards Program (2024) ‘API Documentation - Earthquake Catalog’, Verfügbar unter:  
<https://earthquake.usgs.gov/fdsnws/event/1/> (Zugriff am 23. Juli 2024)

CatBoost.ai (2024) ‘Transforming categorical features to numerical features’, Verfügbar unter:  
[https://catboost.ai/en/docs/concepts/algorithm-main-stages\\_cat-to-numeric](https://catboost.ai/en/docs/concepts/algorithm-main-stages_cat-to-numeric) (Zugriff am 23. Juli 2024)