

Einführung in Data Science und maschinelles Lernen

ZEITREIHENANALYSEN

- Fragen
- Muster in Zeitreihenanalysen
- Non-Stationarity
- Baseline Modelle und Naïve Forecasting
- Projektpräsentation
- Undjetzt?

ADD-ON: DROPOUT LAYER FÜR NEURONALE NETZE

Video von Udacity zu Dropout Layern

```
the function len() returns the length of this list of variable names (i.e. the number of varial
    inputs = tf.keras.Input(shape=[len(r.training_features.keys())])
56
57
    # Normalization layer
58
    x = normalizer(inputs)
60
61
    # 1st hidden layer
    x = Dense(10, activation='relu')(x)
63
64
    # Dropout Layer
    x = Dropout(.2)(x)
65
66
    # 2nd hidden layer
67
    x = Dense(4, activation='relu')(x)
69
70 # Output layer
    output = tf.keras.layers.Dense(1)(x)
    # Model definition
73
    model = tf.keras.Model(inputs, output)
74
75
```

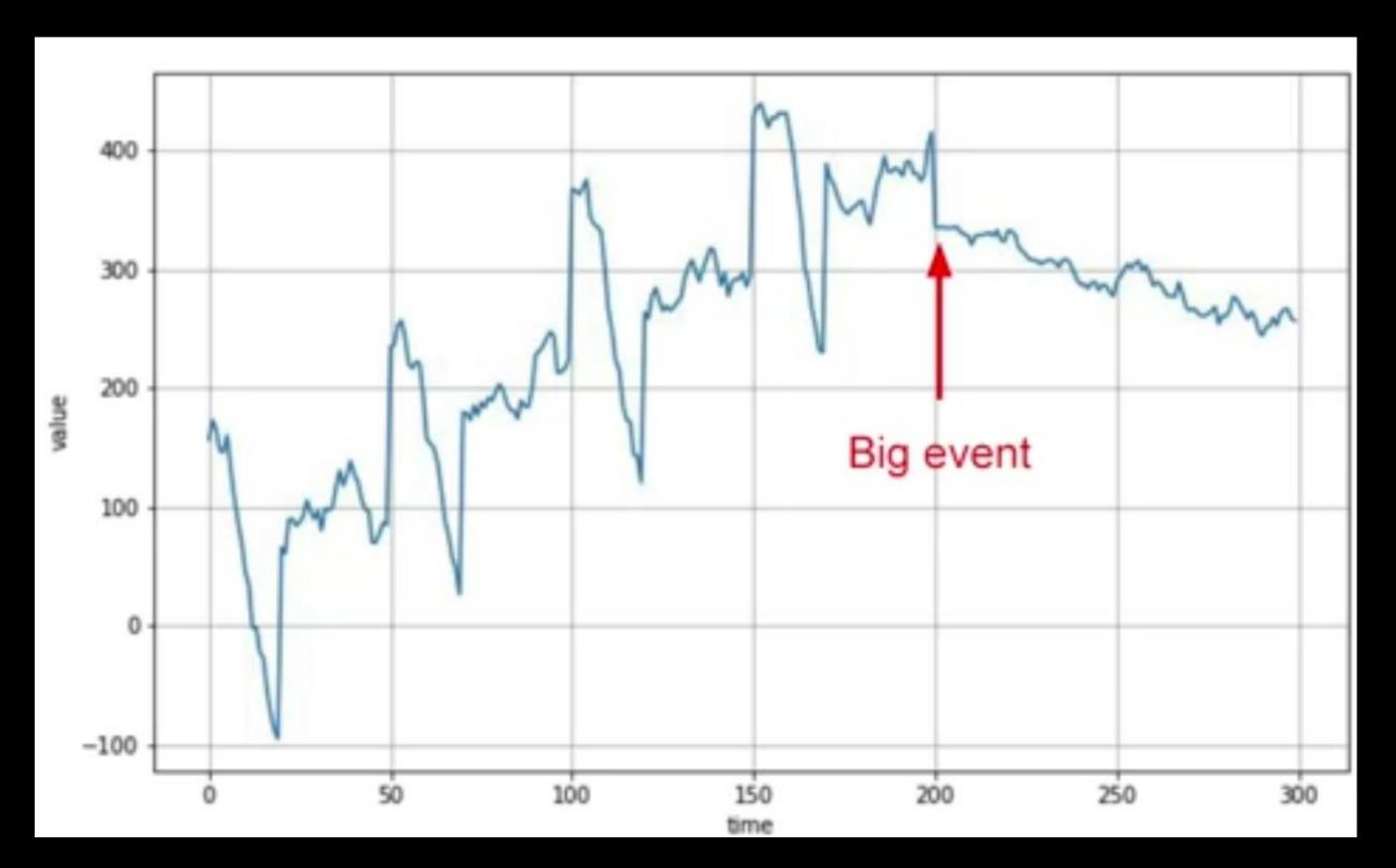
FRAGEN

ZUR LERNRATE DES OPTIMIZERS

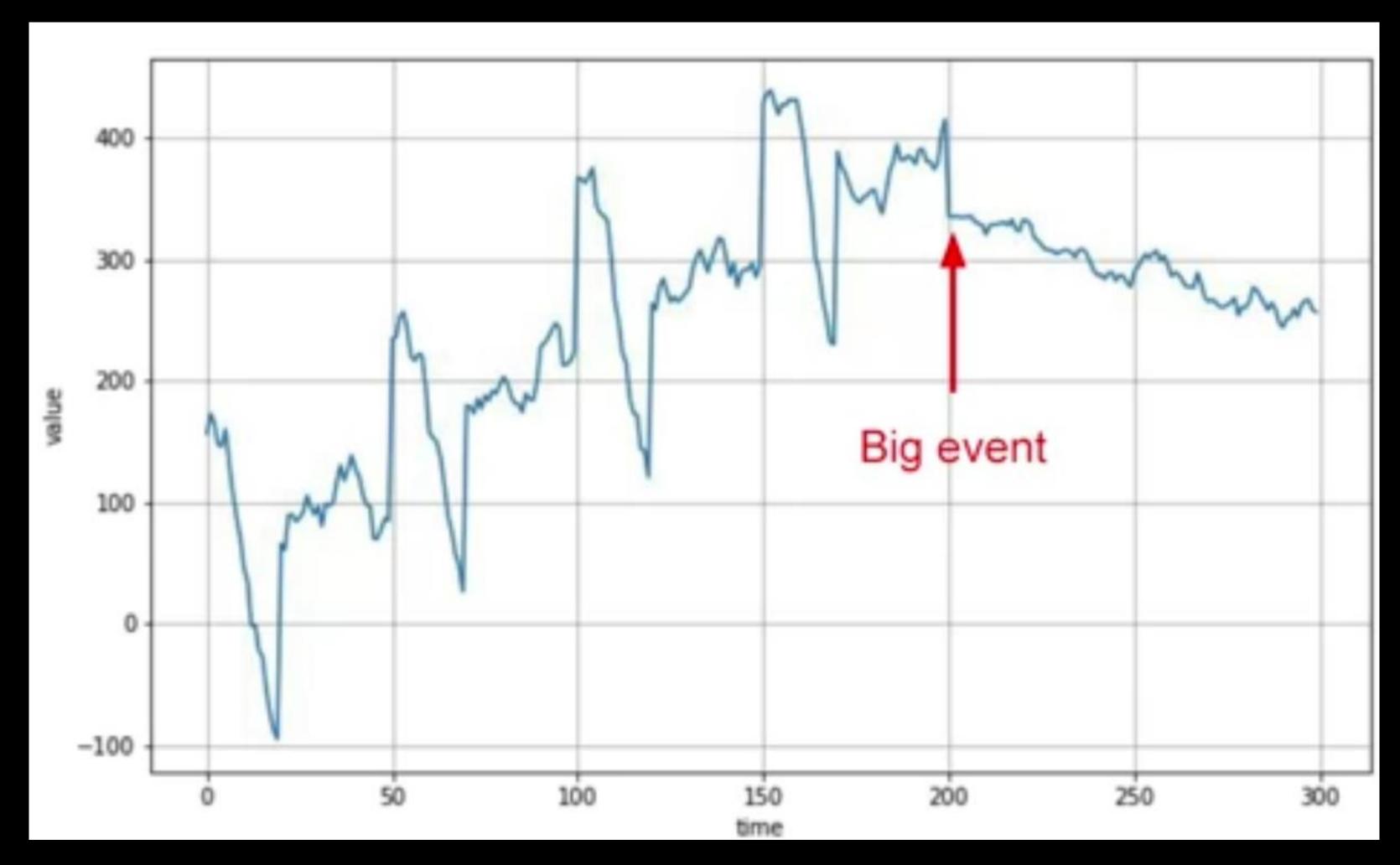
- Große Lernrate:
 Schnelle Optimierung, aber Gefahr, dass nicht das tatsächliche Minimum gefunden wird.
- Kleine Lernrate: Sehr langsame Optimierung, aber größere Sicherheit, dass man das tatsächliche Minimum findet.

DISKUSSION

 Welche verschiedenen Arten von Mustern kann man in der dargestellten Abbildung erkennen?



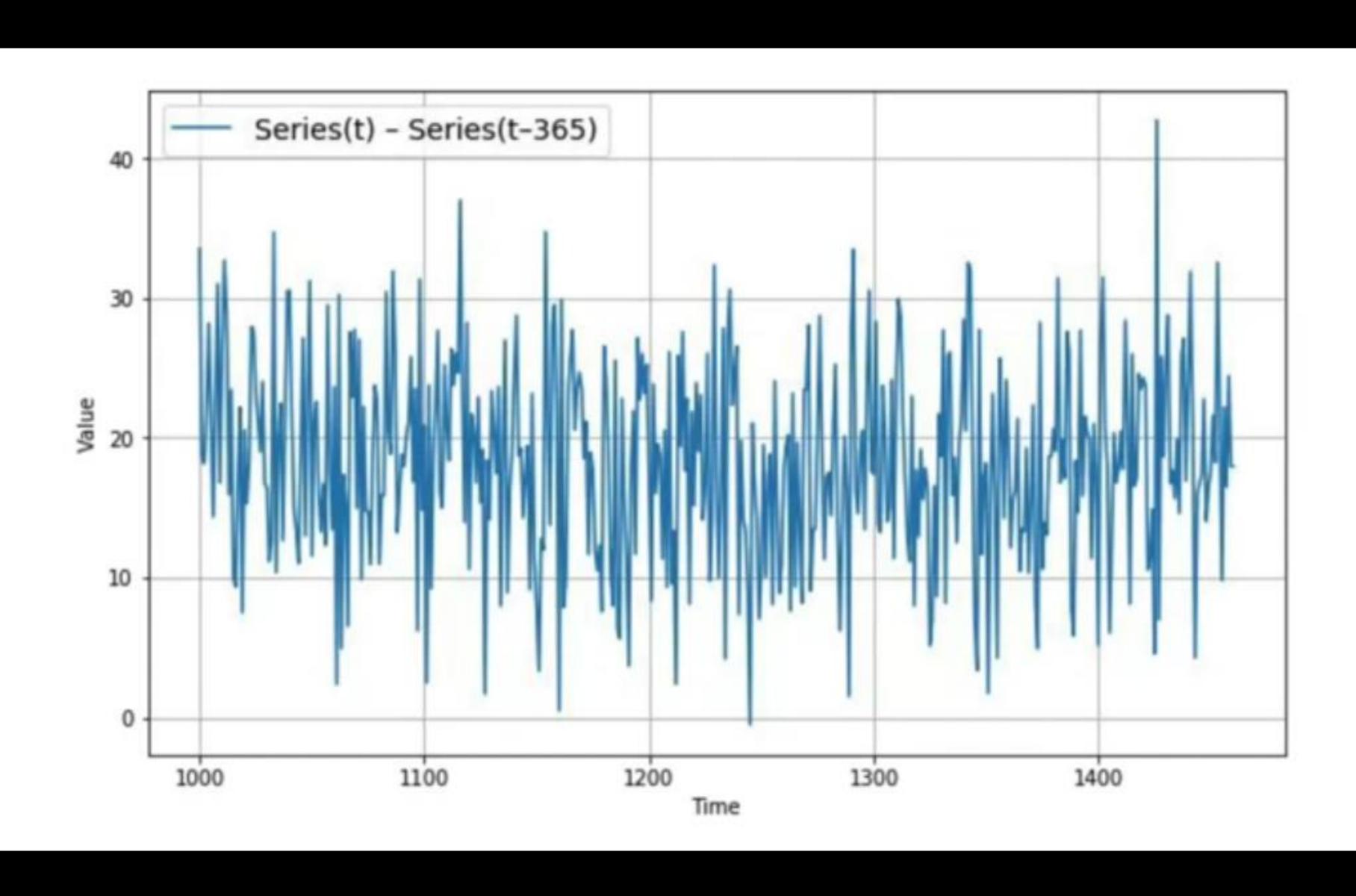
NON-STATIONARITY



MUSTER IN ZEITREIHENANALYSEN

- Trends
- "Jahresgang" (Saisonality)
- Rauschen (Noise)
- Autokorrelation

DIFFERENCING



DIFFERENCING

 Methode um die Zeitabhängigkeit von Zeitreihen zu entfernen bzw. zu mindern.

Subtraktion der vorherigen Beobachtung von der aktuellen Beobachtung

LAG-DIFFERENZ

 Die Differenz zwischen aufeinanderfolgenden Beobachtungen wird als Lag-1-Differenz bezeichnet.

 Die Lag-Differenz kann an die spezifische zeitliche Struktur angepasst werden.

 Bei Zeitreihen mit einer saisonalen Komponente kann man davon ausgehen, dass die Verzögerung der Periode (Breite) der Saisonalität entspricht.

BASELINE MODELLE

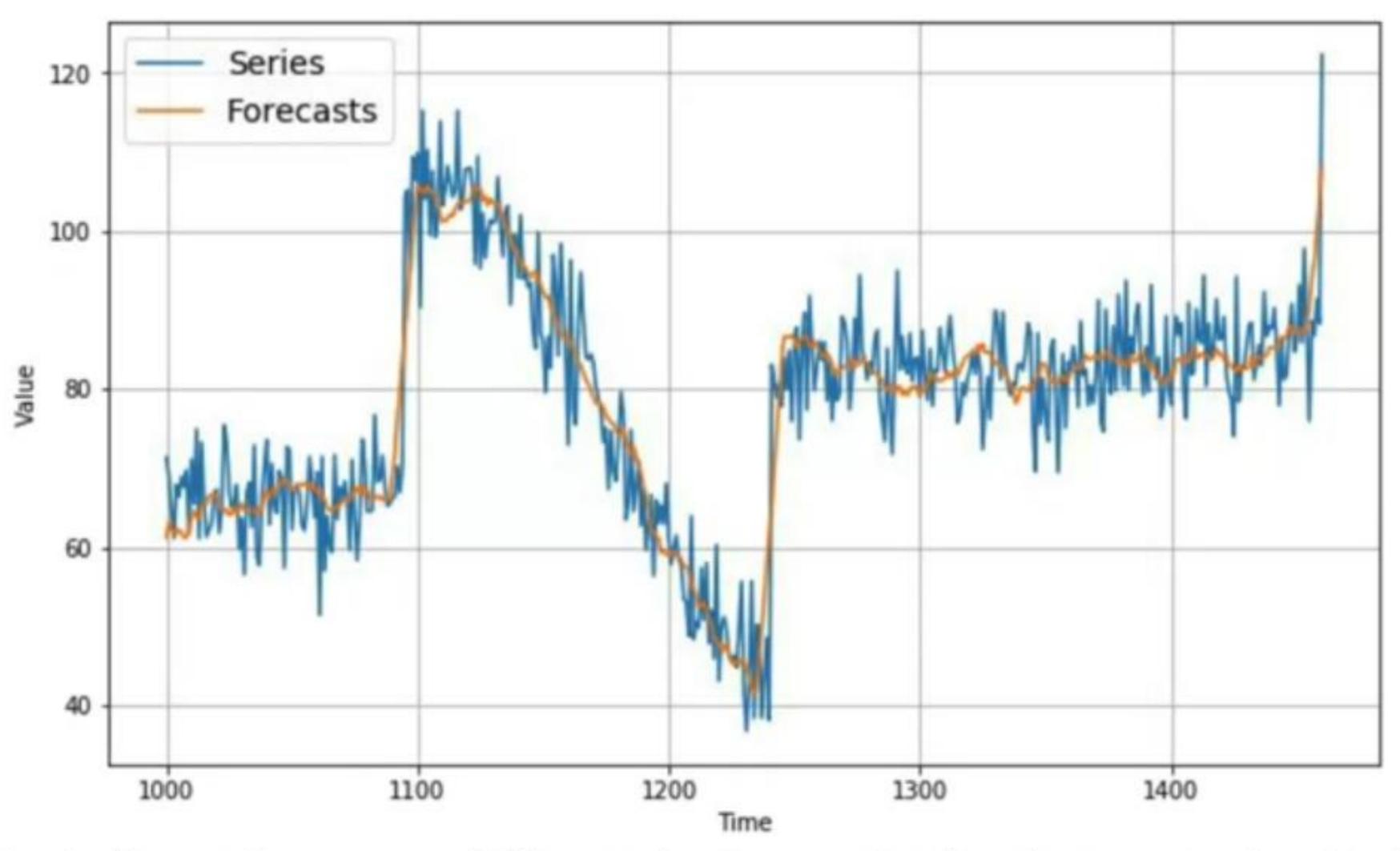
- Allein an Metriken wie MAPE oder RMSE kann man häufig schlecht abschätzen, wie viel das eigene Modell gelernt hat.
- Es ist daher wichtig, die Ergebnisse anderer Modelle als "Baseline" bzw. Referenz zu nutzen.

Mögliche Baselines:

- Ergebnisse bisher genutzter Modelle für den gleichen Datensatz
- Ergebnisse von Modellen auf artverwandten Datensätzen
- Speziell bei Zeitreihen: Ergebnisse basierend auf Naïve Forecasting

NAIVE FORECASTING

- Mögliches Baseline Modell für Zeitreihenanalysen
- Vorhersage entspricht dem jeweils letzten beobachteten Wert
- Saisonal Naïve Forecasting:
 Vorhersage entspricht dem letzten Wert mit der gleichen Saisonalität.



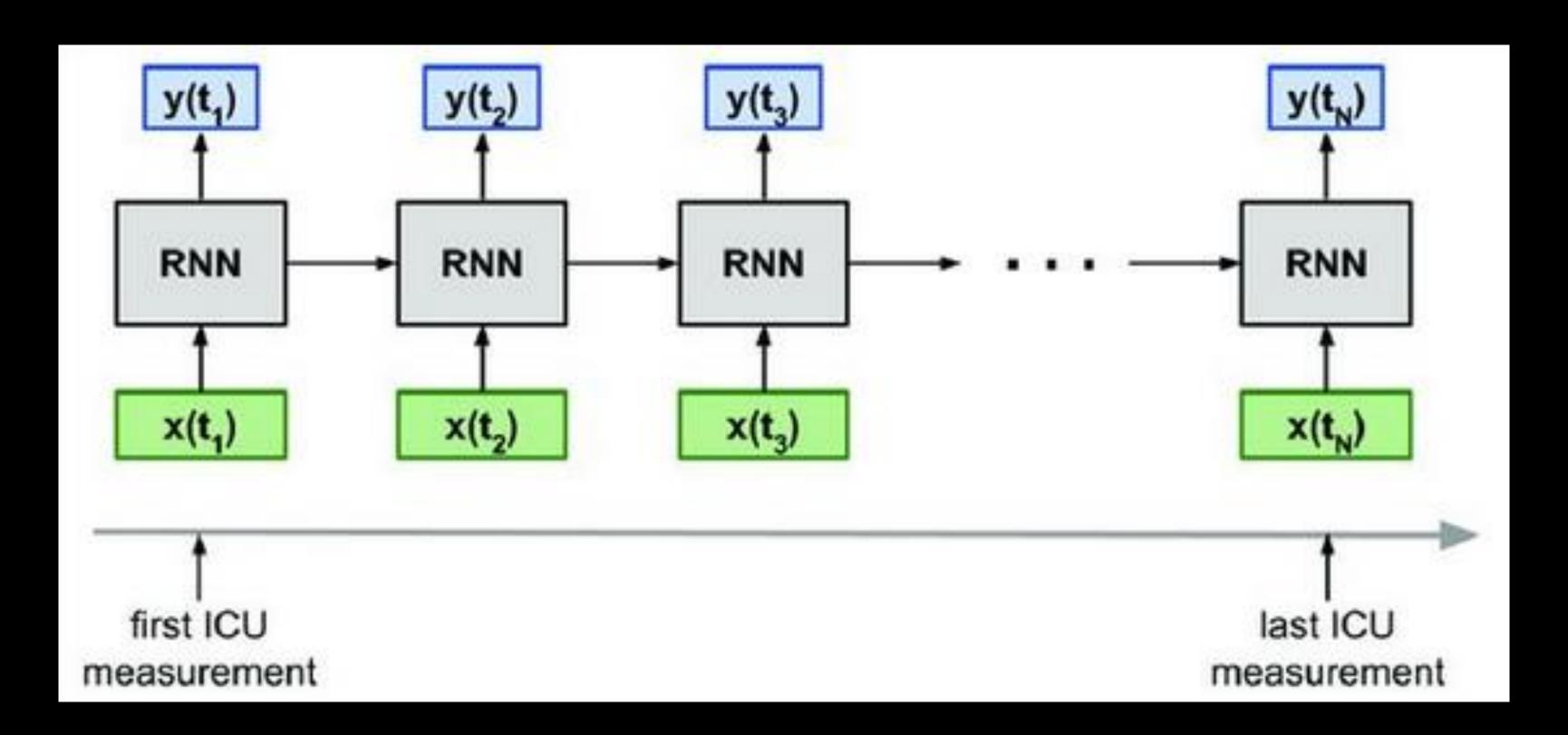
Forecasts = trailing moving average of differenced series + centered moving average of past series (t - 365)

SEQUENCE BIAS

- Generell:
 Die Reihenfolge der Dinge beeinflusst die Auswahl.
- Im maschinellen Lernen:
 Die Reihenfolge der Daten beeinflusst die Schätzung der Parameter.

RECURRENT NEURAL NETS

 Neurales Netz bei dem die Knoten einer Schicht untereinander verknüpft sind.



INHALT DER PROJEKTPRÄSENTATION

Dauer: ca. 8 Minuten pro Team

- Eure Namen
- Auflistung und kurze Beschreibung der selbst erstellten Variablen
- Balkendiagramme mit Konfidenzintervallen für zwei selbst erstellte Variablen
- Optimierung eines neuronalen Netzes:
 - Source Code
 - MAPE für einen Testdatensatz mit 10% der Ursprungsdaten
 - Warengruppenumsätze für den ersten nicht im Datensatz enthaltenen Tag

TIPPS

- Präsentation des R-Markdown Notebooks als HTML
- Formatierung des Codes mithilfe von styler





R news and tutorials contributed by hundreds of R bloggers

ABOUT RSS ADD YOUR BLOG!

LEARN R R JOBS CONTACT US

How to perform Eta Squared in R

January 3, 2022 | finnstats

The post How to perform Eta Squared in R appeared first on finnstats. If you want to read the original article, click here How to perform Eta Squared in R. Eta Squared in R, Eta squared is a commonly-used effect size metric in ANOVA models. It is calculated as follows: ... [Read more...]

Uncovered Interest Rate Parity and F-test on Regression Parameters using R

January 2, 2022 | sang-heon lee

This post explains how to perform the F-test of joint parameter restrictions on a linear regression model. As an example, we use the data in Chen and Tsang (2013), who introduce so called relative Nelson-Siegel factor model to predict exchange r...

[Read more...]



Most viewed posts (weekly)

Using databases with Shiny Examining College Football Conference

Search R-bloggers..

Subscribe

52793 readers

Go

DuckDuckGo

Your e-mail here

DuckDuckGo blocked this Facebook Page

We blocked Facebook from tracking you when the page loaded. If you unblock this page, Facebook will know your activity.

Learn More

Unblock Page

How to perform the Sobel test in R

January 1, 2022 | finnstats

The post How to perform the Sobel test in R appeared first on finnstats. If you want to

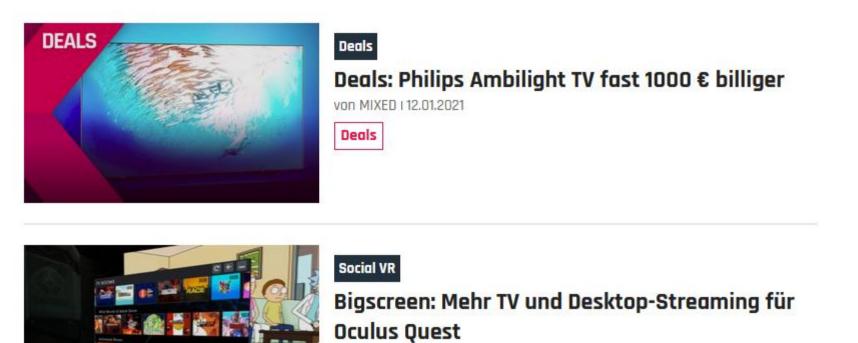






News.

News.



✓ Mixed-Reality-News per E-Mail
✓ VR, AR, KI
✓ Jederzeit kündbar
✓ Kostenlos

E-Mail-Adresse *

Abonnieren

Datenschutz

Machine Learning



Start writing

Latest Best

Georgia Deaconu in Towards Data Science

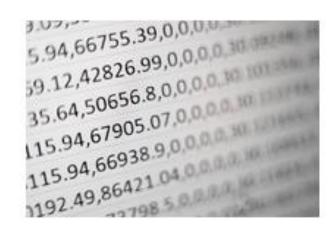


5 ways to deal with large datasets in Python



As a data scientist, I find myself more and more having to deal with "big data". What I abusively call big data...

2 days ago . 3 min read





 Q_2



Andrew Lombarti in Towards Data Science







Mistakes I've made on my journey and how you can avoid being like me when starting out

19 hours ago . 5 min read



70K 177K Stories Writers



Related Topics

Artificial Intelligence Data Science Deep Learning Python Technology Programming Neural Networks NLP

Top Writers



Terence Shin







Susan Li

Changing the world, one post at a time. Sr Data Scientist, Toront...







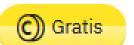
















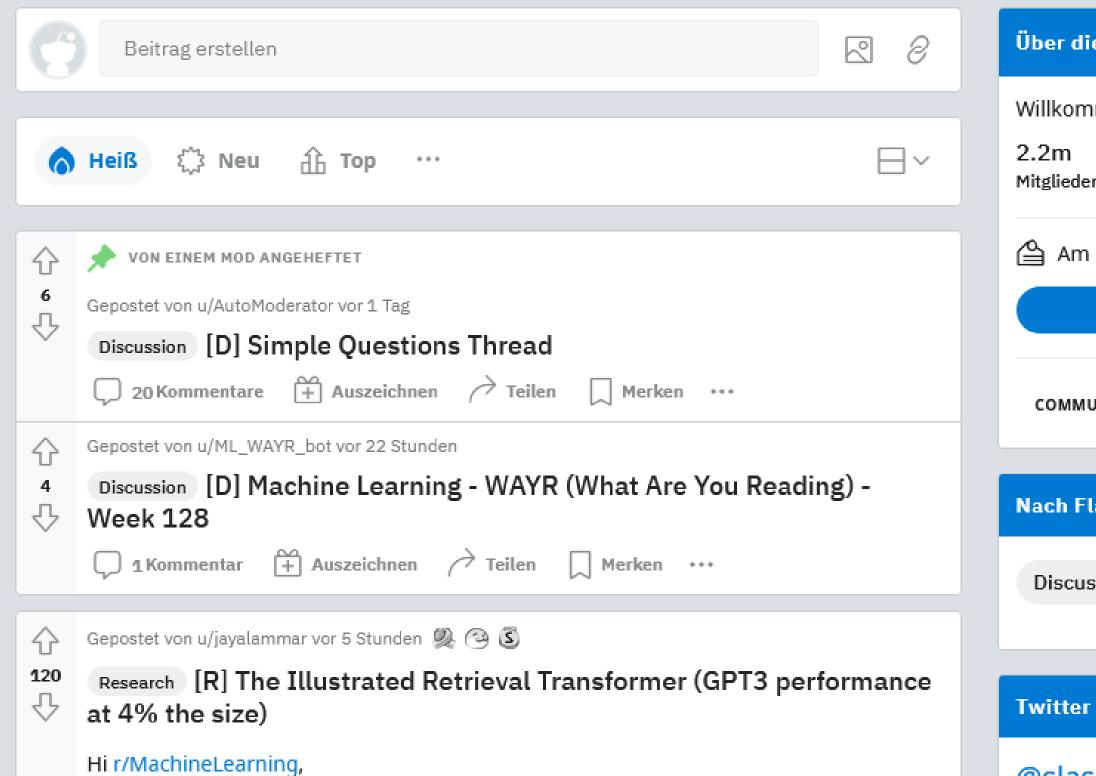
Machine Learning

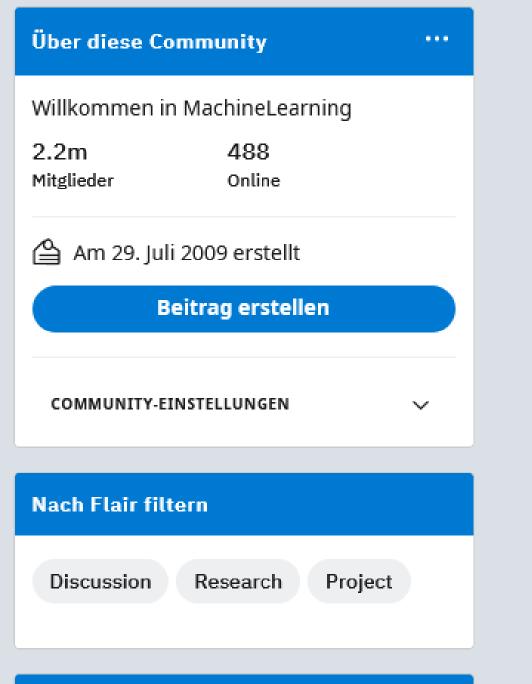
Beigetreten 🔔



r/MachineLearning

Beiträge







COURSEIC (UDACITY

100 Udemy









:: KI-Campus

Die Lernplattform für Künstliche Intelligenz

OPENCAMPUS.SH MACHINE LEARNING DEGREE

Machine Learning with TensorFlow or Deep Learning