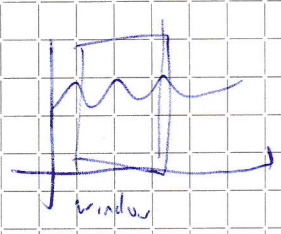
Montag, 18.12.2023

GAN ~ 2014

Recurrent Neural Network

Zeitreihenanalyse – Time Series Analysis

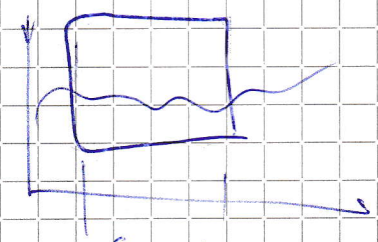


x (30, 20, 1) y = (30,)

30 Punkte

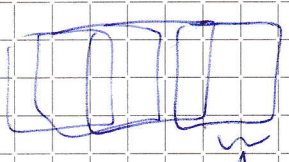
sequence length.

(Datenpunkte, Sequenzlänge, Merkmale)



(50, 1)

(29, 20, 1) →



~~the~~ 20 sequences means only need 29 Data points, if the window 50.

← ~~one~~ shift 1 Step

Dienstag, 19.12.2023

Recurrent Neural Network

Zeitreihen in einzelne Sequenzen aufteilen

↳ moving window

~~Projektor~~ LSTM & GRU

Bayesian Neural Network

tensorflow-probability ← be careful with version during pip install

↓

ca. 0.22 ↑

Mittwoch, 20.12.2023

Tensor Flow Probability Distribution

Text Analyse

- Tokenizer:

"Aku makan nasi" ⟶ ["Aku", "makan", "nasi"]

- Text in Zahlen konvertieren
  - ↳ Dictionary
  - ↳ count-Vectorizer (Bag of words)
  - ↳ Buchstabenbasiertes Tokenizing ⟶ character to ASCII
  - ↳ One hot encoding
  - ↳ Word-vektoren / Word-vector

Donnerstag, 21.12.2023

Recurrent Neural Network ~ Natural Language Proces,
~ Text Analysis

Tokenizer
↳ from sentence(s) to word/syllables/character

~~Warum~~ Why Tokenize only train data?
↳ model size

Was macht die Funktion pad_sequence? ← keras.
↳ Token-count length per sequence

Was ist die optimale Länge der Sätze (Sequenzen)?
↳ 20? 500?

Was sind Wort-Vektoren?

Was macht eine Embedding Schicht? ← Embedding Layer in keras
↳ Token into → Word-vector

Warum werden für die Textanalyse rekurrente Schichten verwendet?
↳ Wörter hinter einander.
↳ Wörter hat andere Bedeutung

BLEU-Metrik / BLEU-Metrics : 0 (bad) – 1 (good)

Übersetzung / Translation ~ Encoder-Decoder-Architecture

Freitag, 22.12.2023

Projekt?

Bidirectional RNN
(Recurrent Neural Network)

Beam Search

- Attention layer
  ↳ no trainable parameter/weights

output = attention_layer(input_a, input_b)

output = attention_layer(vector_a, vector_a)

self attention, because 2 inputs are the same