

Analyse von neuronal network Architekturen im Bezug auf climate change temperature data

TORGE SCHWARK* and JOSCHUA QUOTSCHALLA*, Institute of Computer Science, University of Kiel, Germany

Abstract noch kürzer?

Dieses Paper untersucht verschiedene neuronale Netzwerkarchitekturen im Kontext der Analyse von Klimawandel-Daten, insbesondere der Erdoberflächentemperatur. Die Architekturen umfassen MLPs, 1D Convolutional Neural Networks, LSTMs und Transformer. Die Analyse konzentriert sich auf den *Climate Change: Earth Surface Temperature Data-Datensatz*. Dazu wird zunächst eine effiziente Data-Loader-Pipeline implementiert, Visualisierungen zum jeweiligen Datensatz als auch der Performance der Modelle erstellt und die Leistung der Architekturen qualitativ und quantitativ bewertet. Dabei weisen MLPs als auch 1D ConvNets ein sehr zuverlässiges und schnelles Training auf und bieten gleichzeitig die besten Ergebnisse. Die Studie beleuchtet auch Normalisierungseffekte, die Auswirkungen variabler Input-Sequenzen auf die Performance und liefert Temperaturvorhersagen auf bis zu 100 Jahre in die Zukunft, um weitere Schlüsse im Bezug auf den klimawandel ziehen zu können.

Additional Key Words and Phrases: neuronale Netzwerke, NN, MLP, 1D CONV, LSTM, Transformer, Klimawandel, Temperatur

ACM Reference Format:

Torge Schwark and Joschua Quotschalla. 2018. Analyse von neuronal network Architekturen im Bezug auf climate change temperature data. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 7 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

Die vorliegende Arbeit widmet sich der ganzheitlichen Entwicklung und Evaluierung von neuronalen Netzwerkarchitekturen für die Klimavorhersage unter Verwendung eines ausgewählten Datensatzes. Angesichts der wachsenden Bedeutung präziser und zuverlässiger Klimavorhersagen für zahlreiche Anwendungsgebiete, vom landwirtschaftlichen Management bis hin zu infrastrukturellen Planungsprozessen, ist es von entscheidender Bedeutung, leistungsstarke und effektive Modelle zu entwickeln. Das Ziel des vorliegenden Papers besteht darin, eine umfassende Analyse verschiedener Aspekte des Modelltrainings und der Architekturleistung durchzuführen, um fundierte Einblicke in die Vorhersagegenauigkeit und das Verhalten der Netzwerke zu gewinnen. Dabei liegt ein zentraler Schwerpunkt auf der Implementierung einer effizienten Data-Loader Pipeline, die darauf abzielt, Sequenzen von Datenpunkten effizient als Mini-Batches bereitzustellen. Durch die konsequente Anwendung bewährter Methoden zur strukturierten Vorverarbeitung der Daten sollen optimale Voraussetzungen für das Training der neuronalen Netzwerke geschaffen werden. Darüber hinaus werden Visualisierungen für die Analyse der Daten angestrebt, um detaillierte Einblicke in die Strukturen und Charakteristika des Datensatzes als auch der Modelle zu gewinnen. Im Zusammenhang dazu liegt ein weiterer zentraler Fokus dieser Arbeit auf der eingehenden Untersuchung der Leistung verschiedener Architekturen für neuronale Netzwerke, um die Frage zu beantworten, welche Architektur qualitativ und quantitativ die besten Vorhersageergebnisse erzielt. Dabei

Methoden nennen?

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

werden verschiedene Architekturen wie 1-dimensionale Convolutional Neural Networks (1D ConvNets), Multilayer-Perzeptronen (MLPs), Long Short-Term Memory Networks (LSTMs) und Transformer-Architekturen systematisch analysiert und verglichen. Darüber hinaus wird die Wirkung der Datennormalisierung auf die Vorhersagegenauigkeit untersucht, indem die Performance der Modelle bei Verwendung von normalisierten und nicht-normalisierten Daten verglichen wird. Dies ermöglicht es, die Auswirkungen der Datennormalisierung auf die Modellleistung zu bewerten und präzise Rückschlüsse auf deren Bedeutung für die Vorhersagegenauigkeit zu ziehen. Zusätzlich wird die Leistung verschiedener Architekturen in Abhängigkeit von der Länge der Input-Sequenzen von 8, 16, 32 und 64 analysiert, um die Auswirkungen der Input-Sequenzlänge auf die Modellperformance zu untersuchen und wichtige Erkenntnisse für die spätere Modellauswahl zu gewinnen. Insgesamt zielt diese Arbeit darauf ab, durch eine systematische und methodisch fundierte Vorgehensweise umfassende Einblicke in die Entwicklung und Leistungsfähigkeit von neuronalen Netzwerkarchitekturen für die Klimavorhersage zu gewinnen. Durch die Kombination von Vorverarbeitungstechniken, architektur-spezifischen Analysen und Vergleichen sowie der Untersuchung verschiedener Einflussfaktoren auf die Modellperformance soll ein umfassendes Verständnis für die Gestaltung effektiver und präziser Klimavorhersagemodelle erlangt werden.

Anpassen / erläutern, warum wir andere Werte haben (840, 300)

2 EINFÜHRUNG UND DATENVERSTÄNDNIS

Die Grundlage dieser Studie bildet der "Climate Change: Earth Surface Temperature" Datensatz. Ziel ist es, die Besonderheiten der Daten zu verstehen und eine strukturierte Organisation für den Einsatz in neuronalen Netzwerken zu gewährleisten. Die Temperaturen werden dabei auf Float-Werte mit zwei Nachkommastellen gerundet und nach geografischen Kategorien wie Land, Staat, Stadt und Hauptstadt in separate Textdateien sortiert. Eine zunächst vermutete Optimierung bestand darin, die Daten beim Einlesen in ein Array zu speichern, um den Zugriff während des Trainings zu beschleunigen. Nachdem dieser Ansatz aber keine messbaren Veränderungen mit sich brachte, wurden weitere Optimierungen bezüglich einer Parallelisierung auf der CPU und GPU untersucht, welche eine deutliche Beschleunigung des Trainingsprozesses von bis zu 50% ermöglichten. Hierzu wird die Klasse Pool des Paketes multiprocessing verwendet, welche eine einfache Parallelisierung von Python-Code ermöglicht, indem eine Funktion mit mehreren Inputs auf mehrere Prozesse verteilt wird.

3 DATA-LOADER PIPELINE

Die Data-Loader Pipeline wurde implementiert, um Mini-Batches von Sequenzen von Datenpunkten gemäß den Anforderungen der Architekturen bereitzustellen. Dieser Prozess wurde optimiert, um die Effizienz des Trainings zu steigern. Zur Analyse wurden Visualisierungen wie Histogramme für geografische Merkmale, Datenverteilungen und Kartenplots verwendet. T-SNE-Plots ermöglichen eine multidimensionale Darstellung der Datenpunkte für eine bessere Analyse.

4 VISUALISIERUNGEN FÜR DATENANALYSE UND PREDICTIONS

Die Implementierung verschiedener Visualisierungen spielt eine entscheidende Rolle bei der Analyse von Daten und Vorhersagen. Hierbei wurden Histogramme für den Längen- und Breitengrad, die Verteilung von Datenlücken, Temperaturverteilungen und Unsicherheiten erstellt. Zusätzlich wurde ein Kartenplot der Datenpunkte generiert. Diese Visualisierungen bieten Einblicke in die geografische Verteilung der Daten und ermöglichen eine detaillierte Analyse von Städten mit identischen Werten.

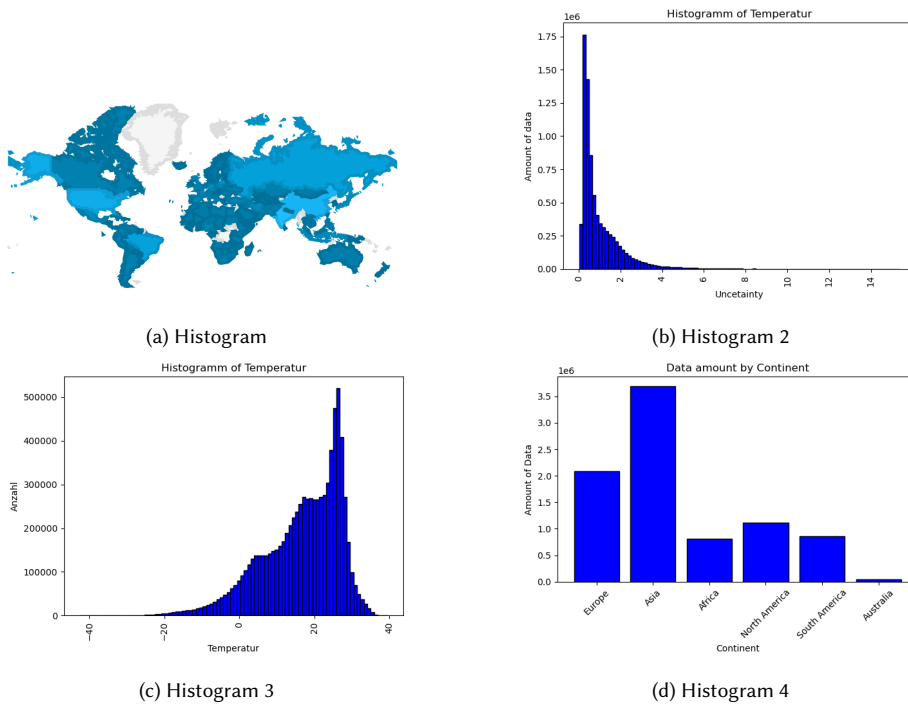


Fig. 1. Four histograms

Table 1. Network Architekturen

Netzwerk	Anzahl Parameter	Anzahl Layern
MLP	TODO:	meisten Parameter aufgrund von dense layern
CONV 1D	TODO:	geringere Anzahl durch feature extraction convolution layer vor den dense layern
LSTM	TODO:	TODO:
Transformer	TODO:	(zweit) meisten Parameter aufgrund von mehreren inneren MLPs

5 TRAINIEREN MIT VERSCHIEDENEN NETZWERKARCHITEKTUREN

Die Netzwerkarchitekturen MLP, 1D Conv Net, LSTM und Transformer wurden auf den vorbereiteten Datensätzen trainiert. Im ersten Schritt wurde eine Grid Search durchgeführt, um die optimalen Modelle und deren Hyperparameter zu bestimmen. Nach dieser automatisierten Suchstrategie wurde ein händisches Fine-Tuning durchgeführt, um die Leistung der Modelle weiter zu verbessern. Während des gesamten Prozesses wurde darauf geachtet, wie sich jede Architektur an die spezifischen Charakteristiken des Datensatzes anpasst und wie gut sie die Temperaturvorhersagen durchführen kann. Die sorgfältige Auswahl und Anpassung der Hyperparameter und Aktivierungsfunktionen spielte eine entscheidende Rolle bei der Optimierung der Modelleistung.

Rewe GPT: Optimierung von Architekturen und Parametern Die Auswahl und Optimierung der geeigneten Parameter für neuronale Netzwerkarchitekturen spielt eine entscheidende Rolle bei der Leistungsfähigkeit und Zuverlässigkeit

der Klimavorhersagemodelle. Im Rahmen der Untersuchung wurden verschiedene Schlüsselparameter betrachtet und optimiert, um die bestmöglichen Ergebnisse zu erzielen.

MLP (Multilayer-Perzeptronen) Im Kontext von Multilayer-Perzeptronen (MLPs) spielen insbesondere die Anzahl der Hidden Layers, die Anzahl der Neuronen pro Layer, die Wahl der Aktivierungsfunktionen und die Lernrate eine bedeutsame Rolle bei der Optimierung. Die Anpassung dieser Parameter beeinflusst maßgeblich die Fähigkeit des MLP-Modells, komplexe nichtlineare Zusammenhänge zu erfassen und präzise Vorhersagen zu generieren. Unsere Experimente haben gezeigt, dass die Auswahl und Feinabstimmung dieser Parameter entscheidend sind, um eine ausgewogene Modellkapazität zu gewährleisten und Überanpassung/Overfitting zu vermeiden. Die systematische Optimierung der genannten Architekturen und Parameter bildet die Grundlage für die Evaluierung ihrer Leistungsfähigkeit im Kontext der Klimavorhersage. Es ist wichtig, die spezifischen Verhaltensweisen und Interaktionen der Optimierungsergebnisse zu analysieren, um fundierte Schlussfolgerungen hinsichtlich der Vorhersagequalität und der Quantifizierung des Modellverhaltens ziehen zu können.

LSTMs (Long Short-Term Memory Networks) Ein zentraler Parameter bei LSTMs ist die Anzahl der Memory Units oder sogenannten Neuronen in der Zelle. Durch gezielte Anpassung dieser Anzahl kann die Netzwerkleistung verbessert und die Fähigkeit zur Erfassung von langfristigen Abhängigkeiten gesteuert werden. In dem Experimenten wurde festgestellt, dass eine moderate Anzahl von Memory Units zu einer ausgewogenen Leistung führt, während zu wenige Units die Kapazität des Netzwerks begrenzen und zu schlechterer Modellleistung führen können.

Convolutional Neural Networks (ConvNets) Für ConvNets sind insbesondere die Größe und Anordnung der Kernel/Filter sowie deren Anzahl in den Convolutional Layern von großer Bedeutung. Während größere Kernel/Filter und eine höhere Anzahl von Filtern zu einer komplexeren Modellarchitektur und möglicherweise zu einer besseren Erfassung räumlicher Merkmale führen können, besteht die Herausforderung darin, ein ausgewogenes Verhältnis zu finden, das Overfitting vermeidet und effektivere Modelle ermöglicht.

Transformer-Architekturen Bei der Optimierung von Transformer-Modellen, insbesondere in Bezug auf Klimavorhersagen, sind die Anzahl der Layer, die Dimensionalität der Embeddings, die Anzahl der Köpfe in den Multi-Head Attention Mechanismen und die Lernrate wichtige Parameter. Unsere Untersuchungen haben gezeigt, dass die sorgfältige Abstimmung dieser Parameter in einem komplexen Zusammenspiel entscheidend ist, um eine optimale Leistungsfähigkeit der Transformer-Architekturen zu erreichen. Es ist wichtig zu betonen, dass Optimierungsansätze stark von den spezifischen Merkmalen des betrachteten Datensatzes und den Zielsetzungen der Klimavorhersagen abhängen. Daher ist eine systematische und gewissenhafte Evaluierung der Optimierungsergebnisse unerlässlich, um fundierte Schlussfolgerungen hinsichtlich der Performanz der unterschiedlichen Architekturen zu ziehen.

6 ARCHITEKTURVERGLEICH UND NORMALISIERUNGEN

Die Leistung der Architekturen wurde sowohl qualitativ als auch quantitativ bewertet. Aktivierungsfunktionen wie Selu und Relu wurden für MLP analysiert. Bei 1D Conv Net wurde die Auswirkung von Global Average Pooling und Flatten vor den Dense Layern untersucht. LSTM und Transformer wurden auf ihre spezifischen Eigenschaften und Fehlerquellen analysiert. Der Vergleich erfolgte durch Bewertung der qualitativen Anpassung und quantitativen Metriken wie dem Mean Absolute Error (MAE).

7 NORMALISIERUNG UND VARIATION DER INPUT-SEQUENZEN

Die Daten wurden vor dem Training normalisiert, und die Auswirkungen auf die Vorhersagegenauigkeit wurden verglichen. Zusätzlich wurde die Performance der Modelle für variable Input-Sequenzen von 8, 16, 32 und 64 untersucht. Dies ermöglichte eine tiefgreifende Analyse, wie die Länge der Input-Sequenzen die Modellleistung beeinflusst.

Insgesamt bietet diese Studie eine umfassende Analyse der Climate Change-Daten unter Verwendung verschiedener Netzwerkarchitekturen. Die gewonnenen Erkenntnisse ermöglichen nicht nur eine bessere Anpassung der Modelle an die Daten, sondern auch Optimierungsmöglichkeiten für zukünftige Forschungen im Bereich des Klimawandels.

8 CITATIONS AND BIBLIOGRAPHIES

The use of \LaTeX for the preparation and formatting of one's references is strongly recommended. Authors' names should be complete — use full first names (“Donald E. Knuth”) not initials (“D. E. Knuth”) — and the salient identifying features of a reference should be included: title, year, volume, number, pages, article DOI, etc.

The bibliography is included in your source document with these two commands, placed just before the `\end{document}` command:

```
\bibliographystyle{ACM-Reference-Format}
\bibliography{bibfile}
```

where “bibfile” is the name, without the “.bib” suffix, of the \LaTeX file.

Citations and references are numbered by default. A small number of ACM publications have citations and references formatted in the “author year” style; for these exceptions, please include this command in the **preamble** (before the command “`\begin{document}`”) of your \LaTeX source:

```
\citestyle{acmauthoryear}
```

Some examples. A paginated journal article [?], an enumerated journal article [?], a reference to an entire issue [?], a monograph (whole book) [?], a monograph/whole book in a series (see 2a in spec. document) [?], a divisible-book such as an anthology or compilation [?] followed by the same example, however we only output the series if the volume number is given [?] (so Editor00a's series should NOT be present since it has no vol. no.), a chapter in a divisible book [?], a chapter in a divisible book in a series [?], a multi-volume work as book [?], a couple of articles in a proceedings (of a conference, symposium, workshop for example) (paginated proceedings article) [? ?], a proceedings article with all possible elements [?], an example of an enumerated proceedings article [?], an informally published work [?], a couple of preprints [? ?], a doctoral dissertation [?], a master's thesis: [?], an online document / world wide web resource [? ? ?], a video game (Case 1) [?] and (Case 2) [?] and [?] and (Case 3) a patent [?], work accepted for publication [?], 'YYYYb'-test for prolific author [?] and [?]. Other cites might contain 'duplicate' DOI and URLs (some SIAM articles) [?]. Boris / Barbara Beeton: multi-volume works as books [?] and [?]. A couple of citations with DOIs: [? ?]. Online citations: [? ? ?]. Artifacts: [?] and [?].

9 ACKNOWLEDGMENTS

Identification of funding sources and other support, and thanks to individuals and groups that assisted in the research and the preparation of the work should be included in an acknowledgment section, which is placed just before the reference section in your document.

This section has a special environment:

`\begin{acks}`

`...`

`\end{acks}`

so that the information contained therein can be more easily collected during the article metadata extraction phase, and to ensure consistency in the spelling of the section heading.

Authors should not prepare this section as a numbered or unnumbered `\section`; please use the “acks” environment.

10 APPENDICES

If your work needs an appendix, add it before the “`\end{document}`” command at the conclusion of your source document.

Start the appendix with the “appendix” command:

`\appendix`

and note that in the appendix, sections are lettered, not numbered. This document has two appendices, demonstrating the section and subsection identification method.

11 MULTI-LANGUAGE PAPERS

Papers may be written in languages other than English or include titles, subtitles, keywords and abstracts in different languages (as a rule, a paper in a language other than English should include an English title and an English abstract). Use `language=...` for every language used in the paper. The last language indicated is the main language of the paper. For example, a French paper with additional titles and abstracts in English and German may start with the following command

```
\documentclass[sigconf, language=english, language=german,
               language=french]{acmart}
```

The title, subtitle, keywords and abstract will be typeset in the main language of the paper. The commands `\translatedXXX`, `XXX` begin title, subtitle and keywords, can be used to set these elements in the other languages. The environment `translatedabstract` is used to set the translation of the abstract. These commands and environment have a mandatory first argument: the language of the second argument. See `sample-sigconf-i13n.tex` file for examples of their usage.

12 SIGCHI EXTENDED ABSTRACTS

The “sigchi-a” template style (available only in \LaTeX and not in Word) produces a landscape-orientation formatted article, with a wide left margin. Three environments are available for use with the “sigchi-a” template style, and produce formatted output in the margin:

- `sidebar`: Place formatted text in the margin.
- `marginfigure`: Place a figure in the margin.
- `marginfigure`: Place a table in the margin.

ACKNOWLEDGMENTS

To Robert, for the bagels and explaining CMYK and color spaces.

A RESEARCH METHODS

A.1 Part One

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi malesuada, quam in pulvinar varius, metus nunc fermentum urna, id sollicitudin purus odio sit amet enim. Aliquam ullamcorper eu ipsum vel mollis. Curabitur quis dictum nisl. Phasellus vel semper risus, et lacinia dolor. Integer ultricies commodo sem nec semper.

A.2 Part Two

Etiam commodo feugiat nisl pulvinar pellentesque. Etiam auctor sodales ligula, non varius nibh pulvinar semper. Suspendisse nec lectus non ipsum convallis congue hendrerit vitae sapien. Donec at laoreet eros. Vivamus non purus placerat, scelerisque diam eu, cursus ante. Etiam aliquam tortor auctor efficitur mattis.

B ONLINE RESOURCES

Nam id fermentum dui. Suspendisse sagittis tortor a nulla mollis, in pulvinar ex pretium. Sed interdum orci quis metus euismod, et sagittis enim maximus. Vestibulum gravida massa ut felis suscipit congue. Quisque mattis elit a risus ultrices commodo venenatis eget dui. Etiam sagittis eleifend elementum.

Nam interdum magna at lectus dignissim, ac dignissim lorem rhoncus. Maecenas eu arcu ac neque placerat aliquam. Nunc pulvinar massa et mattis lacinia.