

Analyse der Anwendbarkeit und Grenzen von KI-Frameworks für Energieprognosen

Bachelorarbeit

zur Erlangung des akademischen Grades „Bachelor of Science (B. Sc.)“
im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik,
Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät
der Leibniz Universität Hannover

vorgelegt von

Name: Arnold

Vorname: Florian

Geb. am:



in:



Prüfer: Prof. Dr. rer. nat. Michael. H. Breitner

Betreuerin: M. Sc. Jana Gerlach

Hannover, den 13.07.2022

Abstract

In dieser Arbeit werden die KI-Frameworks TensorFlow, PyTorch und Scikit-learn bezüglich ihrer Eignung zum Aufstellen von Energieprognosen miteinander verglichen. Es soll herausgefunden werden, welches der Frameworks sich besonders gut für Energieprognosen eignet. Zur Beantwortung dieser Frage werden die KI-Frameworks anhand verschiedener Kriterien beurteilt und verglichen. Es werden mit jedem Framework mehrere Neuronale Netze für zwei Anwendungsbeispiele trainiert und evaluiert. Zum einen für die Prognose der Stromerzeugung erneuerbarer Energien und zum anderen für die Prognose des Strompreises, beides auf Basis von Wetterdaten. Das Ergebnis dieser Arbeit ist, dass sich die Frameworks Scikit-learn und PyTorch von TensorFlow abheben. Scikit-learn überzeugt durch eine deutlich höhere Prognosegenauigkeit bei gleichzeitig einfacher Anwendbarkeit und kurzen Trainingszeiten für eine höhere Anzahl an Epochen. PyTorch bietet vor allem den Vorteil, dass sich dessen Funktionen leicht an die Anforderungen der Anwenderin oder des Anwenders anpassen bzw. erweitern lassen. Es ist wesentlich flexibler als Scikit-learn und auch als TensorFlow. Ein relevanter Nachteil von TensorFlow sind sehr lange Vorhersagezeiten, wodurch die sich an das Training anschließende Evaluation erheblich verlangsamt wird und die Anwendung dadurch deutlich weniger praktikabel ist. Zusammengefasst wird für einfache Anwendungen das KI-Framework Scikit-learn empfohlen, für komplexere und individueller zu gestaltende Anwendungen hingegen PyTorch.

Inhaltsverzeichnis

Abbildungsverzeichnis	IV
Tabellenverzeichnis	V
Anlagenverzeichnis.....	VII
Abkürzungsverzeichnis	VIII
1 Einleitung	1
2 Energieprognosen	3
2.1 Hintergrund.....	3
2.2 Anwendungsbereiche	5
2.2.1 Prognose der Stromerzeugung erneuerbarer Energien	5
2.2.2 Prognose des Strom- und Energieverbrauchs	7
2.2.3 Prognose des Strompreises.....	7
2.3 Zeithorizonte.....	8
3 Theoretischer Hintergrund zu Neuronalen Netzen.....	10
3.1 Künstliche Intelligenz und Maschinelles Lernen.....	10
3.2 Bestandteile Neuronaler Netze	11
3.2.1 Neuronen.....	11
3.2.2 Layer	13
3.2.3 Aktivierungsfunktionen.....	14
3.2.4 Verlustfunktion.....	16
3.2.5 Hyperparameter.....	16
3.3 Das Lernen Neuronaler Netzer	19
3.3.1 Forward Propagation	19
3.3.2 Backpropagation.....	20
3.3.3 Optimizer	20
3.4 Problematiken des Maschinellen Lernens.....	21
3.4.1 Underfitting und Overfitting	21
3.4.2 Exploding und Vanishing Gradients	22
4 Methodik.....	23
5 Genutzte Anwendungsplattform.....	25
6 Betrachtete KI-Frameworks	26
6.1 TensorFlow.....	26
6.2 PyTorch.....	28

6.3	Scikit-learn.....	30
7	Datengrundlage	31
7.1	Rohdaten.....	31
7.2	Datenaufbereitung	33
7.2.1	Wetterdaten.....	33
7.2.2	Stromerzeugung	35
7.2.3	Strompreis.....	36
7.3	Normalisierung der Daten.....	37
7.4	Trainings- und Testdatensatz	38
8	Analyse der KI-Frameworks.....	38
8.1	Kriterien	38
8.2	Anwendung der KI-Frameworks	40
8.3	Auswertung und Vergleich	40
9	Diskussion der Ergebnisse.....	51
10	Kritische Würdigung	54
10.1	Limitationen	54
10.2	Implikationen	55
10.3	Ausblick.....	57
11	Fazit	58
	Literaturverzeichnis	IX
	Anlagen.....	XVI
	Ehrenwörtliche Erklärung.....	LXV

1 Einleitung

Die zehn wärmsten Jahre seit Beginn der Wetteraufzeichnung im Jahr 1880 liegen allesamt im 21. Jahrhundert. Neun der zehn wärmsten Jahre waren zwischen den Jahren 2010 und 2020 (NCDC & NOAA, 2021). Diese Erhöhung der globalen Durchschnittstemperatur ist seit Jahren zu beobachten und lässt sich auf den Menschen zurückführen. Gründe dafür sind, dass sich u. a. durch das Verbrennen von fossilen Energieträgern und durch die Viehwirtschaft treibhauswirksame Gase in der Atmosphäre anreichern. Eines dieser treibhauswirksamen Gase ist Kohlenstoffdioxid (CO₂) (Umweltbundesamt, 2016). Der Trend der weltweiten CO₂-Emissionen ist steigend und bis ins Jahr 2019 waren die CO₂-Emissionen eines Jahres nahezu immer höher oder nahezu gleich wie im vorangegangenen Jahr. Dies ist in Abbildung 1 anschaulich dargestellt.

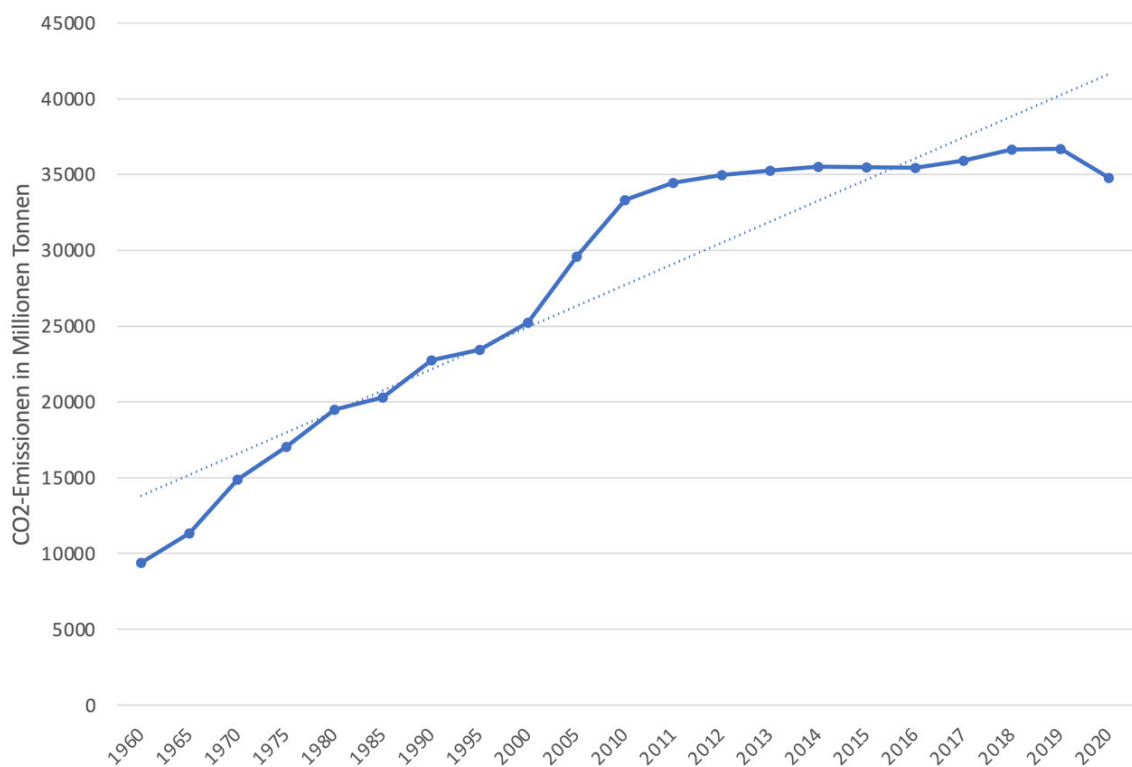


Abbildung 1: Weltweite CO₂-Emissionen in den Jahren 1960 bis 2020 in Millionen Tonnen, in Anlehnung an Global Carbon Project (2021)

Folgen der erhöhten Klimaänderung sind das Abschmelzen von Gebirgsgletschern und Schneebedeckungen sowie das gehäufte Auftreten von Extremwetterereignissen. Aus diesem Grund beschäftigen sich Institutionen, wie z. B. der Weltklimarat, mit Maßnahmen zur Reduktion des Ausstoßes von Treibhausgasen. (Umweltbundesamt, 2016)

Eine Möglichkeit zur Reduktion der Treibhausgase ist der Einsatz von erneuerbaren Energiequellen zur Stromerzeugung. Zu den erneuerbaren Energiequellen zählen u. a. Photovoltaik und Windkraft (Aslam et al., 2021). Die in Deutschland vorhandene Nenn-

leistung, d. h. die elektrische Leistung aller Kraftwerke dieser beiden Energiequellen, ist in den letzten beiden Jahrzehnten stetig gewachsen (BMWK, 2022; windmonitor.de, 2019). Durch die Zunahme des Anteils von erneuerbaren Energiequellen an der Stromerzeugung sind akkurate Energieprognosen essenziell. Die Gründe hierfür werden in Kapitel 2 genauer betrachtet.

Im Bereich der Künstlichen Intelligenz (KI) ziehen Ansätze des Maschinellen Lernens (ML) und des Deep Learnings (DL) in den verschiedensten Anwendungsbereichen die Aufmerksamkeit von immer mehr Wissenschaftlerinnen und Wissenschaftlern auf sich. Grund dafür ist ihre hohe Leistungsfähigkeit, nichtlineare Zusammenhänge zu modellieren (Alkhayat & Mehmood, 2021). Neuronale Netze (NN) als Teil von ML und DL eignen sich dazu, Prognosen auf Basis von Inputdaten aufzustellen. Dabei lässt sich beweisen, dass sich jede kontinuierliche Funktion mithilfe eines NN beliebig genau annähern lässt, sofern die Zahl der Neuronen des NN groß genug ist (Bishop, 1994; Choo, Greplova, Fischer & Neupert, 2020).

NNs lassen sich mithilfe von KI-Frameworks entwickeln. Sie stellen eine Art Baukasten zur Verfügung, die die wichtigsten Komponenten zum Aufbau von NNs beinhalten. Gleichzeitig ist die komplexe Mathematik hinter den NNs bereits in den KI-Frameworks implementiert, sodass sich ausschließlich auf die Entwicklung der Struktur eines NN konzentriert werden kann. (Ketkar & Moolayil, 2021)

Da es nicht das eine, beste KI-Framework für alle Anwendungsfälle gibt (Devopedia, 2020), soll in der vorliegenden Arbeit folgende Frage geklärt werden: „Welches der betrachteten KI-Frameworks eignet sich am besten für Energieprognosen?“. Hierzu werden in dieser Arbeit drei verschiedene KI-Frameworks unter verschiedenen Gesichtspunkten miteinander verglichen. Diese drei KI-Frameworks sind TensorFlow (TF), PyTorch (PT) und Scikit-learn (SK). Um die direkte Nutzung der KI-Frameworks mit in den Vergleich einzubeziehen, werden NNs für zwei unterschiedliche Anwendungsbeispiele erstellt, trainiert und evaluiert. Das erste NN soll die Stromerzeugung erneuerbarer Energiequellen auf Basis von Wetterdaten prognostizieren. Das zweite NN soll, ebenfalls auf Basis von Wetterdaten, lernen, den Strompreis vorherzusagen.

Um die Forschungsfrage beantworten zu können ist die Arbeit wie folgt strukturiert. In den Kapiteln 2 und 3 wird auf theoretische Grundlagen zu Energieprognosen und NNs eingegangen. Im Anschluss wird die Methodik bzw. die Vorgehensweise bei der Erstellung dieser Arbeit in Kapitel 4 vorgestellt, wonach in Kapitel 5 auf die genutzte Anwendungsplattform und in Kapitel 6 auf die getesteten KI-Frameworks eingegangen wird. Kapitel 7 beschäftigt sich mit der Auswahl und Aufbereitung der Datensätze für die NNs und in Kapitel 8 werden die KI-Frameworks anhand verschiedener Kriterien beurteilt. Daran anschließend werden die Ergebnisse in Kapitel 9 diskutiert. In Kapitel 10 findet die kritische Würdigung dieser Arbeit statt und abschließend folgt das Fazit in Kapitel 11.

dazu die notwendigen Funktionen bereit, NNs zu konstruieren, zu trainieren, zu evaluieren und letztendlich zu verwenden. Durch Weiterentwicklung der Frameworks könnten hierbei weitere Fortschritte, ein größerer Funktionsumfang oder eine noch einfachere Anwendbarkeit sein.

11 Fazit

Im Rahmen der vorliegenden Arbeit wurde einleitend die Notwendigkeit von Energieprognosen aufgezeigt, um im Anschluss die hierfür einsetzbare Methodik von NNs zu beschreiben. Es wurden drei verschiedene KI-Frameworks betrachtet, die sich zur Konstruktion von NNs eignen. Hierbei stellte sich die Frage „Welches der betrachteten KI-Frameworks eignet sich am besten für Energieprognosen?“.

Getestet wurden die drei KI-Frameworks PT, SK und TF. Anhand verschiedener Kriterien wurden diese in Kapitel 8 miteinander verglichen. Neben allgemeinen Kriterien zu den Frameworks wurden diese außerdem anhand von zwei Anwendungsbeispielen, der Prognose der Stromerzeugung erneuerbarer Energiequellen und der Prognose des Strompreises, jeweils auf Basis von Wetterdaten, in ihrer Anwendung gegenübergestellt. Dabei wurde insbesondere die Trainingszeit, die Vorhersagezeit und die Prognosegenauigkeit der NNs evaluiert.

Das Ergebnis dieser Arbeit ist, dass sich SK und PT unter den gewählten Randbedingungen als die besten KI-Frameworks herausgestellt haben. SK überzeugt durch die leichte Anwendbarkeit bei der gleichzeitig besten Prognosegenauigkeit unter den drei Frameworks. Auch bei den meisten anderen Kriterien konnte SK überzeugen. Der hauptsächliche Nachteil von SK besteht in der geringen Anpassbarkeit und Flexibilität. Dementsprechend bietet sich in Anwendungsfällen, bei denen genau dies erforderlich ist, die Benutzung von PT an. PT ist das flexibelste der drei Frameworks. Es bietet viele Möglichkeiten, Funktionen und Prozesse für die eigenen Anforderungen zu erweitern bzw. anzupassen. TF befindet sich in diesen beiden Bereichen jeweils im Mittelfeld. Der größte Nachteil von TF ist zudem die extrem lange Vorhersagezeit, die die Evaluationsprozesse von NNs auf Basis von großen Datensätzen entscheidend verlangsamt.

Sich an diese Arbeit anschließende Untersuchungen können die Analyse weiterer Funktionen der verwendeten KI-Frameworks sein oder andere KI-Frameworks in den Vergleich mit einzubeziehen. Weiterhin lassen sich die NNs beider Anwendungsbeispiele weiter verbessern, um dann die Leistungsfähigkeit der Frameworks neu zu beurteilen.

Abschließend bietet Tabelle 17 einen zusammenfassenden Überblick zur Auswertung der in dieser Arbeit betrachteten KI-Frameworks. Die für die abschließende Bewertung relevantesten Kriterien aus Kapitel 8 werden hierzu vergleichend gegenübergestellt. In der Bewertung wird unterschieden von sehr gut (+ +), bewertet mit + 2, über durchschnittlich (o), bewertet mit 0, bis sehr schlecht (- -), bewertet mit - 2. Die Bewertungen wurden zudem

unterschiedlich gewichtet. Für den GPU-Support wurde der Faktor 1, für die Flexibilität, die Trainingszeit und die Vorhersagezeit der Faktor 2 und für die Vorhersagegenauigkeit der Faktor 3 gewählt. Das resultierende Endergebnis für jedes KI-Framework findet sich in der letzten Spalte der Tabelle.

Tabelle 17: Abschließender Vergleich der betrachteten KI-Frameworks

	GPU-Support	Flexibilität	Trainingszeit	Vorhersagezeit	Vorhersagegenauigkeit	Σ
Gewichtung	1	2	2	2	3	
PyTorch	+	++	o	++	-	6
Scikit-learn	-	-	++	++	+	8
TensorFlow	+	+	+	--	o	1

Der Vergleich zeigt, dass keines der KI-Frameworks in allen Kriterien besser als die anderen beiden Frameworks ist. Insgesamt werden die Ergebnisse aus Kapitel 9 sehr gut widerspiegelt. Auf Basis der betrachteten, bewerteten und gewichteten Kriterien verdeutlicht das resultierende Endergebnis, dass SK gegenüber PT und TF das am besten geeignete KI-Framework für Energieprognosen ist.