

# Kurzfrist-Prognose der Stromerzeugung von Photovoltaiksystemen mit Machine Learning Techniken

## 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: Dörr-Kling



Vorname: Lasse



Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 14.08.2020

---

# Inhaltsverzeichnis

<b>Titelblatt</b> .....	<b>i</b>
<b>Inhaltsverzeichnis</b> .....	<b>ii</b>
<b>Abbildungsverzeichnis</b> .....	<b>iv</b>
<b>Tabellenverzeichnis</b> .....	<b>v</b>
<b>Formelzeichen</b> .....	<b>vi</b>
<b>Abkürzungen</b> .....	<b>vii</b>
<b>Abstract</b> .....	<b>viii</b>
<b>1 Einleitung</b> .....	<b>1</b>
<b>2 Machine Learning</b> .....	<b>4</b>
2.1 Machine Learning Techniken .....	4
2.1.1 Überwachtes Lernen .....	5
2.1.2 Nicht überwachtes Lernen .....	5
2.1.3 Bestärkendes Lernen .....	6
2.2 Künstliche neuronale Netze .....	6
2.2.1 Allgemeine Funktionsweise .....	6
2.2.2 Perzeptron .....	7
2.2.3 Multidimensionale, mehrschichtige neuronale Netze .....	7
2.3 Lernvorgang und Optimierung .....	10
2.3.1 Fehlerfunktionen .....	11
2.3.2 Gradientenabstiegsverfahren .....	12
2.3.3 Fehlerrückführung .....	14
2.3.4 Momentum .....	14
2.3.5 Adam .....	15
2.3.6 Minibatch Training .....	15
2.4 Überanpassung und Unteranpassung .....	15
2.4.1 Überanpassung .....	16
2.4.2 Unteranpassung .....	17
2.5 Regulierungsmethoden für KNN .....	17
2.5.1 Weight Decay .....	18
2.5.2 Early Stopping .....	19
2.6 Weitere Modelle für die Vorhersage von Zeitreihen .....	20
2.6.1 Bayes-Netze .....	21
2.6.2 Markov-Modelle .....	21
2.6.3 Support Vector Machines .....	22
2.6.4 Hybridmodelle .....	24
2.6.5 Performance der Methoden .....	24
2.7 Unterscheidung der Prognosehorizonte .....	25
2.7.1 Kurzfristprognose .....	25
2.7.2 Mittelfristprognose .....	25
2.7.3 Langfristprognose .....	25
<b>3 Modellauswahl und Datenaufbereitung</b> .....	<b>26</b>
3.1 Modellauswahl und -aufbau .....	26
3.2 Auswahl der Programmiersprache .....	27

---

3.3 Datenauswahl und -aufbereitung.....	28
3.3.1 Inputdaten .....	29
3.3.2 Outputdaten .....	30
3.3.3 Datenaufbereitung .....	30
3.3.3.1 Fehlerbehebung .....	31
3.3.3.2 Angleichen der zeitlichen Auflösung der Datensätze .....	32
3.3.3.3 Auswahl des Zeitabschnitts der Daten .....	35
<b>4 Modellentwicklung zur Prognose der Stromerzeugung von Photovoltaiksystemen.....</b>	<b>36</b>
4.1 Implementierung .....	36
4.2 Training und iterative Verbesserung des Modells .....	39
4.3 Indikatoren der Performance.....	39
4.4 Optimierung .....	41
4.4.1 Anpassung der Modellparameter .....	41
4.4.2 Anwendung von Regulierungsmethoden.....	43
4.4.2.1 Weight Decay .....	43
4.4.2.2 Early Stopping .....	44
4.4.3 Schlussfolgerungen.....	45
<b>5 Analyse und Evaluierung.....</b>	<b>46</b>
5.1 Prognosen.....	46
5.2 Diskussion der Ergebnisse .....	48
5.2.1 Vergleiche zu anderen Modellen .....	48
5.2.2 Begründung der erzielten Performance .....	48
5.3 Schwächen und Limitationen des Modells .....	50
<b>6 Fazit und Ausblick .....</b>	<b>51</b>
<b>Literaturverzeichnis.....</b>	<b>52</b>
<b>Anhang .....</b>	<b>56</b>
<b>Ehrenwörtliche Erklärung .....</b>	<b>58</b>

# 1 Einleitung

Der steigende weltweite Energiebedarf, bedingt durch Wirtschafts- und Bevölkerungswachstum und die sich gleichzeitig verschlechternde Umweltsituation aufgrund des Klimawandels, zwingt Wirtschaft und Politik immer mehr zu einer Energiereform (Das et al., 2018). Die Bedeutung von nachhaltigen, regenerativen Energieformen wie der Solarenergie bzw. der Photovoltaik (PV) ist seit den 1990er Jahren stark gestiegen und tut dies in Zeiten der Energiewende weiterhin. Dieser Umstand wird in Abbildung 1.1 verdeutlicht (Bundesministerium für Wirtschaft und Energie (BMWi), 2020). Zu beobachten ist ein exponentieller Anstieg des von PV-Anlagen erzeugten Stroms zwischen den Jahren 2000 und 2012. Danach flacht die Kurve leicht ab. Die derzeitige Tendenz weist jedoch wieder einen Anstieg der Stromerzeugung durch PV-Anlagen auf. Es ist zu erwarten, dass sich dieser positive Trend aufgrund der Umweltsituation und politischer Vorgaben und Subventionen weiter deutlich verstärkt. So definierte die EU bereits einige Klimaziele, welche eine Senkung der Treibhausgasemissionen bis 2030 um 40% gegenüber 1990 vorsieht (Europäische Kommission, 2020). Auch der Anteil der Primärenergie aus erneuerbaren Quellen soll bis 2030 auf 30% steigen (Das et al., 2018). Weltweit betrug dieser 2017 etwa 13,6% (Bundeszentrale für politische Bildung, 2019). In der EU lag er im Jahr 2018 bei 18% (Eurostat, 2020). Die Bundesrepublik Deutschland visiert derzeit einen Anteil nachhaltiger Energieformen von 65% bis 2030 und 80% bis 2050 an (Eising et al., 2020).

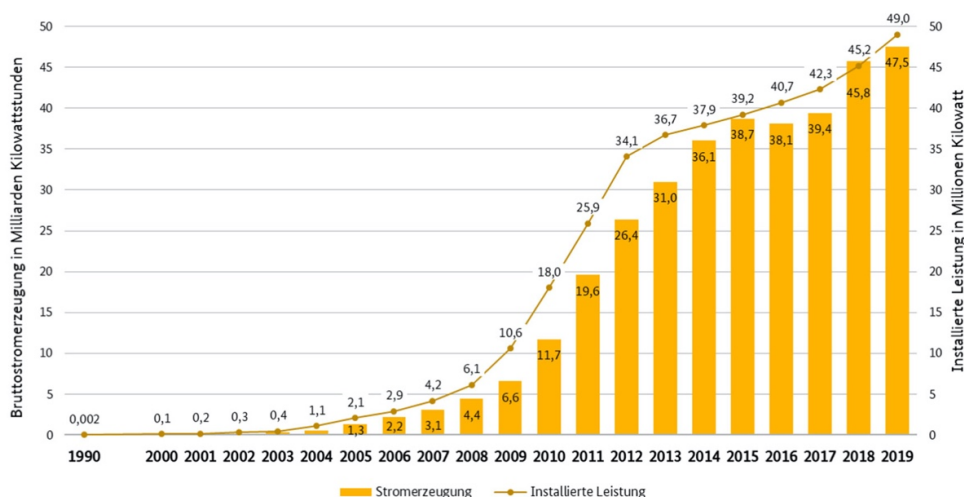


Abbildung 1.1: Entwicklung der Stromerzeugung und der installierten Leistung von Photovoltaikanlagen seit 1990 (Quelle: Bundesministerium für Wirtschaft und Energie (BMWi), 2020)

Deshalb besteht noch viel Handlungsbedarf, woraus sich gute Wachstumsaussichten für den Photovoltaiksektor ergeben. Aufgrund der dadurch zunehmenden Bedeutung werden sowohl technische Parameter, wie der Wirkungsgrad einer Photovoltaikanlage als auch wirtschaftliche Aspekte, wie Stromgestehungskosten und die optimale Verwendung mit dem dadurch erzeugten Strom, immer wichtiger. Dabei spielt die Information über die zeitliche Entwicklung

der Stromerzeugung aus Photovoltaikanlagen bei der Energieversorgung eine wichtige Rolle. Dies kann durch Vorhersagemodelle mit hoher Genauigkeit erreicht werden, wodurch der Energieversorger frühzeitig Informationen über die zukünftige Energieproduktion erhalten kann. Zudem werden Abschätzungen über den Anteil der Einspeisung von Solarenergie in das Stromversorgungsnetz ermöglicht und darüber, wieviel dieser Energie verkauft, gespeichert oder in chemische Energie (z.B. Wasserstoff) umgewandelt werden sollte. Daher können entsprechende Vorhersagemodelle beim Umgang mit den genannten Problemen helfen und zu Kostensenkungen bei der Stromerzeugung führen (Espinár et al., 2013).

Fluktuierende Spannung und Stabilitätsprobleme bei Solarenergie sind weitere Gründe für die Notwendigkeit von Vorhersagen in diesem Sektor (Voyant et al., 2017). Die starken Schwankungen bei der Stromerzeugung durch wechselnde Sonneneinstrahlung je nach Bewölkung, weiteren Umwelteinflüssen wie Temperatur oder Wind und durch den Wechsel von Tag und Nacht sind typisch bei Photovoltaikanlagen. Derartige Faktoren erschweren das Stromnetzmanagement zusätzlich. Übersteigt der Strombedarf aufgrund von ungünstigen Wetterverhältnissen die Stromproduktion, muss zusätzliche Energie ins System eingespeist werden. Dies kann durch konventionelle Kraftwerke oder gespeicherte Energiereserven geschehen. Dieser Regelungsvorgang muss jedoch vorbereitet und eingeleitet werden, der je nach Stromaggregat zwischen einigen Minuten (z.B. Wasserkraftwerk) und mehreren Stunden (z.B. Kohlekraftwerk) variieren kann. Dies verdeutlicht weiterhin die Relevanz von genauen Vorhersagemodellen zur Verbesserung dieser Regelvorgänge.

Derartige Vorhersagen sind insbesondere für isolierte Stromnetzsysteme von großem Interesse. Ein solches isoliertes Stromnetzsystem kann oftmals auf Inseln vorgefunden werden, welche nicht auf eine zuverlässige und gut steuerbare Stromversorgung von konventionellen Großkraftwerken auf dem Festland zurückgreifen können. Daher werden solche Systeme als Inselnetze bezeichnet. Typisch für Inselnetze sind eine geringere Ausfallsicherheit, größere Schwankungen im Spannungs- und Frequenzbereich sowie hohe Kosten für Stromreserven, verglichen mit gewöhnlichen Verbundnetzen. In diesen Fällen ist ein geschickter Umgang mit dezentralen, regenerativen Energieformen von besonderer Bedeutung (Steinhart et al., 2017). Hinzu kommt der Aspekt, dass eine präzise Vorhersage bessere Voraussetzungen für den Kauf von Energiereserven am Markt ermöglicht.

Die genannten Aspekte bilden die Grundlage der Motivation und Zielsetzung dieser Arbeit. Das Ziel ist die Kurzfrist-Prognose der Stromerzeugung von Photovoltaikanlagen. Als Kurzfrist wird hier der Zeithorizont von 24 Stunden bezeichnet. Datengrundlage bilden dabei Wetterdaten. Die Umsetzung soll mit Machine Learning Techniken erfolgen. Auf die verschiedenen Varianten von Machine Learning Techniken, die Methoden und Funktionen, sowie die technischen und mathematischen Hintergründe wird in den nachfolgenden Kapiteln eingegangen.

In Kapitel 2 erfolgt ein Überblick über den heutigen Stand der Technik und Methodik auf dem Gebiet des Machine Learning. Es werden grundlegende Funktionsweisen und verschiedene Umsetzungsmöglichkeiten vorgestellt. Daran anschließend findet in Kapitel 3 die Modellauswahl und die Aufbereitung der zugrundeliegenden Daten statt. In Kapitel 4 erfolgt die Entwicklung und anschließende Optimierung des Modells. Schließlich wird in Kapitel 5 die Analyse und Evaluierung des Modells und seiner Performance durchgeführt. Kapitel 6 resümiert diese Arbeit in einem Fazit und beinhaltet einen Ausblick.

## 6 Fazit und Ausblick

Diese Arbeit beschäftigte sich mit der Prognose der PV-Stromerzeugung in der kurzen Frist. Die Motivation dieses Themas bildet die aktuelle und zukünftig voraussichtlich weiter steigende Relevanz von PV-Anlagen. Der damit in unmittelbarem Zusammenhang stehende bestmögliche, effiziente Umgang des durch die Anlagen gewonnenen Stroms, kann mit derartigen Prognosemodellen erleichtert und optimiert werden. Umgesetzt wurde dies anhand von Machine Learning Techniken. Dazu notwendige technische Grundlagen, Zusammenhänge, Umsetzungsmöglichkeiten, Probleme, Optimierungsmöglichkeiten und Modelle wurden im Grundlagenteil dargestellt. Für die Umsetzung wurden die für den Standort Konstanz bezogenen Wetterdaten und PV-Stromerzeugungsdaten aufbereitet. Schließlich wurde in der Programmiersprache Python, mithilfe der auf TensorFlow basierenden Programmbibliothek Keras ein einfaches FFNN erstellt und trainiert. Daran anschließend erfolgte ein Vergleich unterschiedlicher Programm-Setups, um die bestmögliche Konfiguration des Modells zu finden. Anhand dieses Setups wurden weitere Optimierungsversuche durch das Anwenden von Regulierungsmethoden unternommen. Dabei bewirkte die Methode Weight Decay jedoch keinerlei Verbesserungen, Early Stopping hingegen steigerte die Generalisierungseigenschaft des Setups geringfügig. Auf Grundlage des besten Setups konnten schließlich Prognosen erstellt und analysiert werden. Die daraus erhaltenen Ergebnisse wurden diskutiert und mögliche Gründe für das Abschneiden analysiert. Zudem konnten Schwächen des Modells herausgearbeitet werden.

Insgesamt kann bilanziert werden, dass das grundsätzliche Ziel der Arbeit erreicht wurde. Es wurde zudem gezeigt, dass ML-Techniken für PV-Prognosen geeignet sind. Jedoch fallen die Ergebnisse, bezugnehmend auf die Performance des Modells, mit den erzielten Werten von  $MAPE_{mean} \approx 21,83\%$ ,  $Acc_{test} = 50,28\%$  und  $MSE_{test} = 0,0029$  im Vergleich zu Modellen anderer Autoren deutlich schwächer aus. Zudem fehlt dem Programm, wie unter 5.3 bereits erwähnt, eine nutzerfreundliche Einbettung, die geeignetes Einlesen und Vorverarbeiten der Daten als auch die spätere Ausgabe ermöglichen. Eine derartige Erweiterung würde die Verwendung des Modells erleichtern.

Als weiterführende Arbeit empfiehlt sich daher das Erstellen solcher Funktionen sowie das Anwenden anderer Modellansätze für KNN, wie RBFNN oder RNN. Auch grundlegend andere Modellansätze wie Hybridmodelle, Bayes-Netze, Markov-Modelle oder SVR sollten in Betracht gezogen und nach objektiven Kriterien gegeneinander abgewogen werden.