# Optimierte Leistungsprognosen an neuen Standorten durch induktives Multi-Task- und Transfer Learning von wetterabhängigen Merkmalen innerhalb eines PV-Portfolios

Gerrit Hein & Dominik Beinert Fraunhofer-Institut für Energiewirtschaft und Energiesystemtechnik IEE
Königstor 59, 34119 Kassel, Germany
Tel.: +49 (0) 561 / 7294 - 1725
E-Mail: gerrit.hein@iee.fraunhofer.de
www.iee.fraunhofer.de

## Einleitung

Um den steigenden Strombedarf zu decken und unsere Klimaziele zu erreichen, werden auch in Zukunft stetig neue PV Anlagen gebaut. Die derzeitigen politischen Entwicklungen wirken auf den Ausbau eher weiter beschleunigend. Für die wirtschaftliche Vermarktung der erzeugten Energie und um Planungssicherheit zu garantieren, ist die Güte der Netzeinspeisungsprognosen von großer Bedeutung. Die Netzeinspeisung, welche maßgeblich durch die PV Erzeugung bestimmt wird, ist unter anderem von einer Wechselwirkung des Anlagen Eigenverbrauchs abhängig. Die holistische Betrachtung dieses Zusammenspiels mehrerer Anlagen haben das Potential die Prognosen zu verbessern. Für die Modellierung des komplexen Zusammenspiels der Umgebungsbedingungen und der erzeugten und verbrauchten Energie werden immer häufiger datengetriebene Modelle herangezogenen. Der herkömmliche Weg zur Berechnung einer Lastflussprognose sieht vor, dass eine PV Anlagen unabhängig von anderen PV Anlagen, also einzeln trainiert wird.
Da bei vielen neuen Anlagen oftmals keine, bis wenig historische Messungen als Datengrundlage vorliegen, können die Leistungen erst nach einer gewissen Laufzeit verlässlich bewertet und Prognosen gemacht werden. Die Idee des Projekts ist, Wissen von Anlagen mit vorhandenen Zeitreihen auf Anlagen mit wenig bis keinen Daten zu übertragen. Die einzelne Lastflussprognose einer Anlage kann man auch als Task bezeichnen, für welche ein Modell trainiert wird, um die Zusammenhänge der Umgebungsbedingungen (Wetter, Nutzungsverhalten, ...) auf die PV-Produktion und den Gesamtverbrauch abzubilden.

Das Ziel dieser Arbeit ist, das vorhandene Wissen bzw. die Daten so auszunutzen, das neue Tasks (mit geringerer Datenmenge) von der vorhandenen Datengrundlage anderer Tasks profitieren und sich somit insgesamt die Vorhersage verbessern lässt.

Hierbei lässt sich unterscheiden zwischen den verschiedenen Vorgehensweisen des Single Task Learnings (STL), des Multi-Task Learnings (MTL) und des Transfer Learnings (TL).
Anders als beim STL, für das pro Task ein Modell einzeln trainiert wird, wird beim MTL ein Gesamtmodell gesucht, welches die PV Prognose über alle Anlagen optimiert. Im Gegensatz dazu steht das Transfer Learning, welches eine Gruppe von Anlagen als „Quelldaten“ zum Vortrainieren nutzt, um auf den „Zieldaten“ den eigentlichen Task zu trainieren. Die Ähnlichkeit zwischen Tasks ist die zentrale Grundvoraussetzung, damit dieser zweistufige Trainingsprozess erfolgreich abläuft.

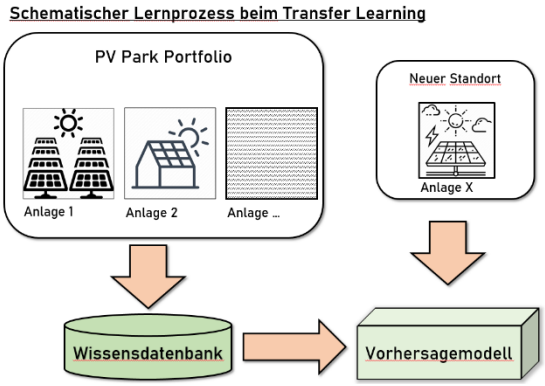


*Abbildung 1: Schematischer Lernzprozess beim Transfer Learning*

So kann zum Beispiel ein ähnlicher Task dazu beitragen, die „Datenlücken“ eines anderen Tasks teilweise zu schließen. Übertragen auf den Anwendungsfall der PV Prognosen bedeutet dies, dass Anlagen mit ähnlichen Eigenschaften (zum Beispiel Größe der Anlage, Ausrichtungswinkel oder ähnlichen Wetterbedingungen) zu einem positiven Wissenstransfer beitragen können. Hat man es jedoch mit komplett verschiedenen Systemen und Übertragungen zu tun, wird man eher einen negativen Transfer bekommen, so dass das kollektive Training zu einem schlechteren Ergebnis führt, als wenn beide Tasks unabhängig voneinander betrachtet werden.

## Daten

Zur Verfügung standen ein Jahr (2019) normierte Messdaten in 15-minütiger Auflösung von SMA. Es handelte sich dabei um 100 Gitterkoordinaten mit einer Auflösung von ca. 10 km x 10 km. Diese Punkte stellten aggregierte PV-Anlagen verschiedenere Kategorien dar: Gewerbeanlagen mit mehr als 30 kW und Heimanlagen mit weniger als 30 kW, jeweils mit oder ohne Batterie. Als Zielgrößen wurden am Ende die normierten Größen von PV-Produktion, Gesamtverbrauch oder der direkte Eigenverbrauch gewählt.

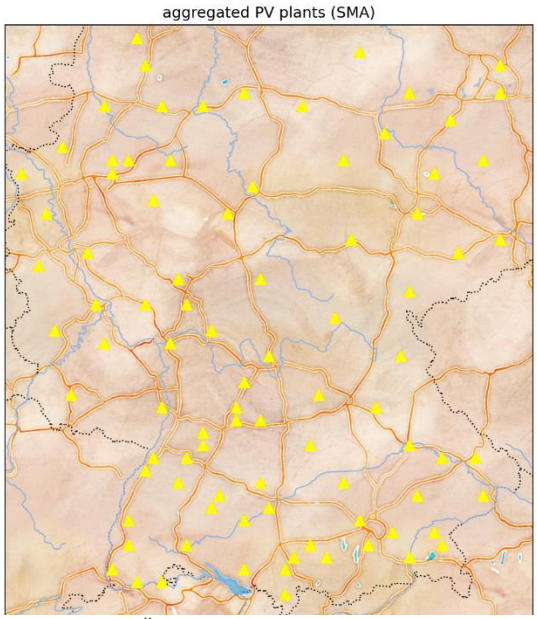


*Abbildung 2: Übersichtskarte der im Datensatz zur Verfügung gestellten anonymisierten PV Anlagen*

Als Input wurden numerische Wetterprognosen (Folgetags-Prognosen) verwendet. Aus einer Reihe von verfügbaren Größen wurden letztendlich die Parameter Temperatur, Luftdruck, Taupunkt, Einstrahlung und die Windgeschwindigkeit gewählt. Außerdem wurden kalendarische Variablen (Wochenenden oder Feiertage) und die Lastprognose für ganz Deutschland als Eingangsdaten für die Prognose berücksichtigt.

Beispielhaft werden hier Eingangs- und Ausgangsdaten eines Standortes gegenübergestellt, mittels welcher die Modelle trainiert wurden:

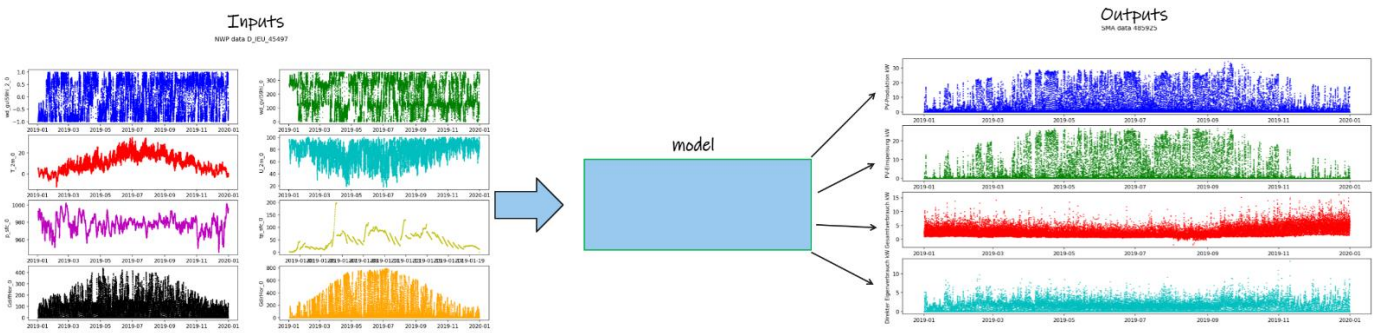
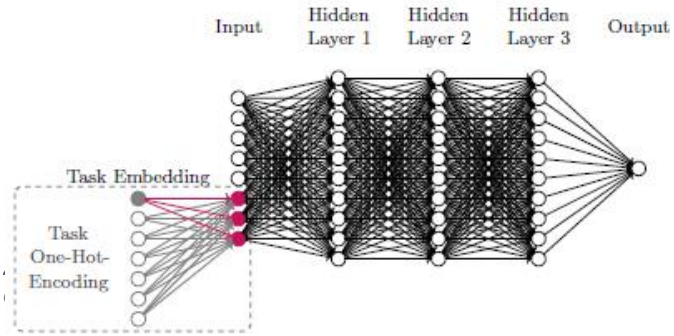



*Abbildung 3: Beispielhafte Zeitreihen der meteorologischen Inputs und der Outputs: PV-Produktion, PV-Einspeisung, Gesamtverbrauch und Eigenverbrauch (von oben) welche die verschiedenen Modelle durchlaufen.*

## Experimente

### MULTITASK LEARNING SZENARIO




Für das Multi-Task Learning wurden verschiedene Netzwerkarchitekturen untersucht und bezüglich ihrer Prognosegüte für die verschiedenen Zielgrößen verglichen. Multi-Task Learning mittels Bayes Embedding schaltet einem Neuronalem Netz eine Embedding Schicht vor, welche genutzt werden kann, um ähnliche Tasks miteinander zu verknüpfen. Hier wird eine Wahrscheinlichkeitsfunktion gelernt anstatt wie beim „Deterministic Embedding“ ein diskreter „Schaltwert“. Dem gegenüber gestellt wurde wieder das Single Task Learning (mittels Gradient Boosted Regression Trees) und ein „Unified Task Learning“, welches die Embedding Schicht weglässt und ein allgemeines Modell trainiert, dass von allen Inputs auf alle Outputs schließt.

**TRANSFER LEARNING SZENARIO**

In diesem Szenario wurde eine Verkürzung „neuer Aufgaben“ simuliert. Das bedeutet, dass die im zweiten Schritt zu trainierenden Target Tasks nur zwischen 10 % bis 90 % der anfänglichen Trainingslänge zur Verfügung hatten. Es wurden 50 PV Standorte zufällig als Source Tasks gewählt und die restlichen als Source Tasks verwendet.

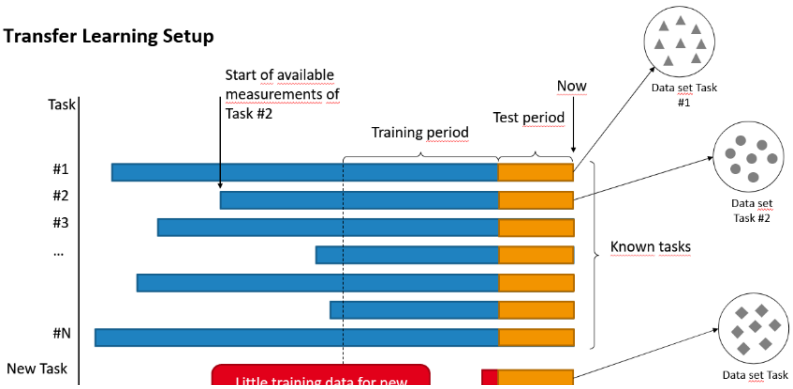



*Abbildung 5: Das Transfer Learning Setup: Quelltasks mit langer Historie, Source task mit kurzer Historie.*

Wir haben die Modelle für die drei Zielgrößen PV-Erzeugung, Gesamtverbrauch und Eigenverbrauch berechnet und jeweils auf den verkürzten Target Tasks nachtrainiert.

## Ergebnisse

Bei den Multi-Task Learning Experimenten zeigten die vier verschiedenen Modelle eine ähnliche Performance auf dem Testdatensatz (Abbildung 6). Das Bayes Embedding neigte am ehesten zum Overfitting der Trainingsdaten.

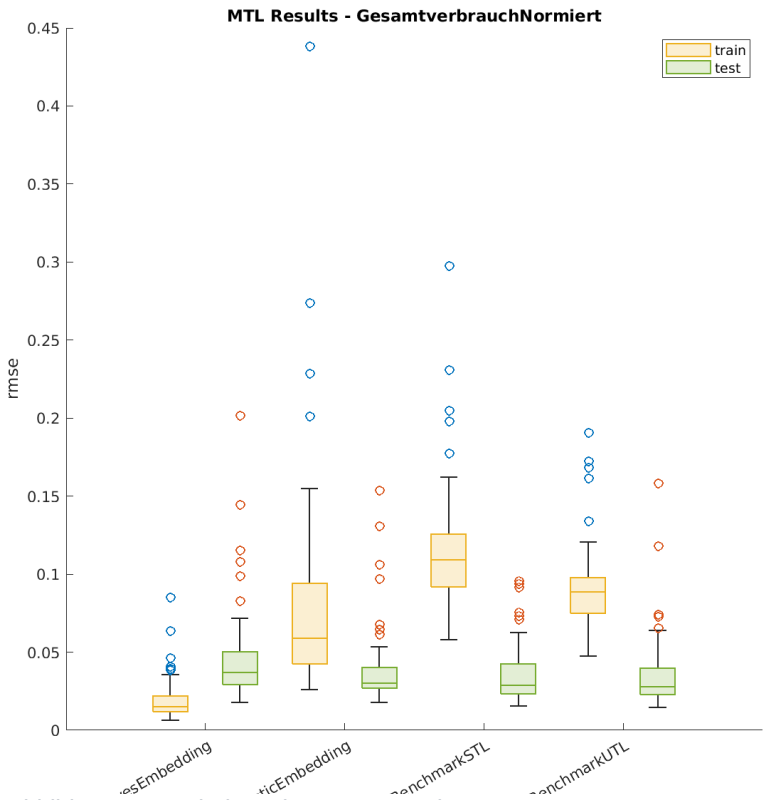



*Abbildung 6: Ergebnisse des MTL Experiments*

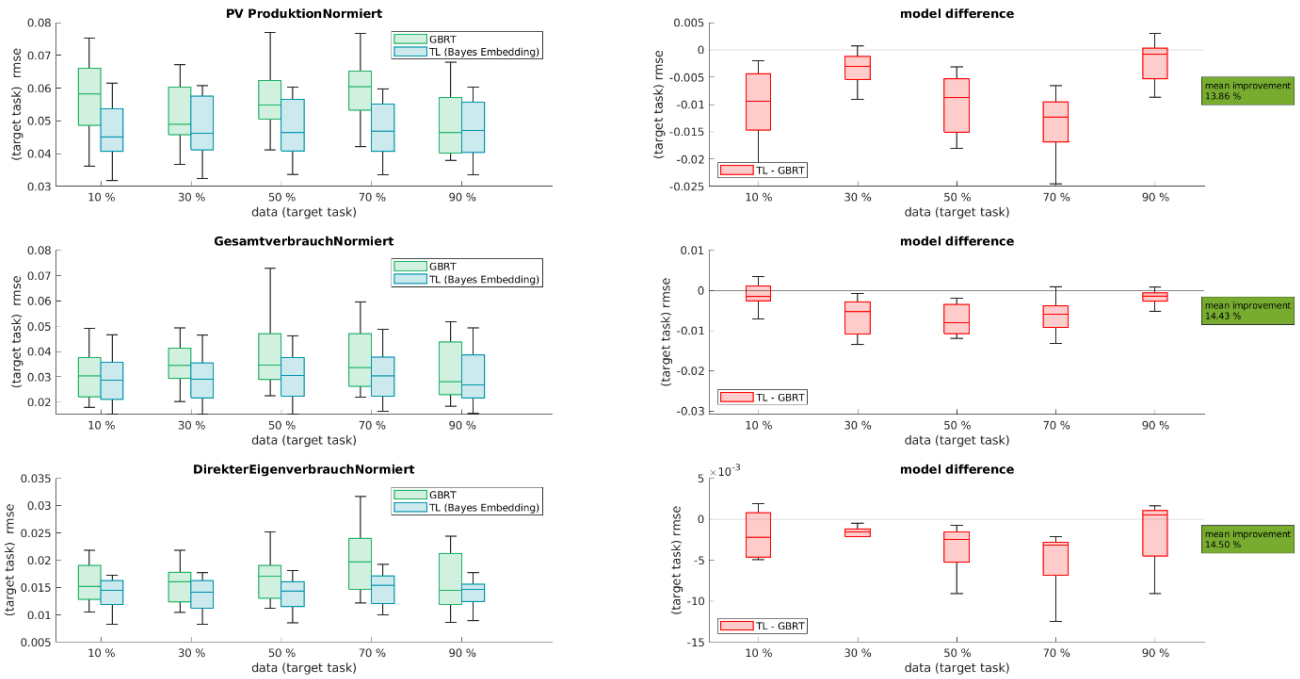



*Abbildung 7: Ergebnisse des Transfer Learning Experiments.*

Für das Transfer Szenario wurden zwei Modelle miteinander verglichen (Abbildung 7). Links sind die Testergebnisse der verschiedenen Target Tasks für beide Modelle als Balkenplot gezeigt, in Grün der GBRT-Benchmark und in blau das Transfer-Lernmodell. Auf der rechten Seite ist die Differenz zwischen den Benchmark und der Transfer-Learning Ergebnisse gezeigt. Hier erkennt man, dass der positive Effekt des Transfer Learnings bei einer nicht zu geringen, aber auch nicht zu hohen Datenmenge (zwischen 50 % -70 %) am stärksten ist.

## FAZIT

Der Einsatz von Multitask und Transfer Learning führt in den meisten Fällen zur Verbesserung der PV Leistungsprognosen. Ein weiterer Vorteil gegenüber der herkömmlichen Einzelberechnung ist die Zeitersparnis, da die Transfer Learning Modelle, wesentlich recheneffizienter sind als herkömmliche STL-Modelle zu trainieren.

Diese Arbeit wurde ermöglicht durch die BMBF-Förderung des Projekts TRANSFER – Förderkennzeichen 01IS20020A.