

KI-gestützte Verbrauchsprognosen auf Basis von Smart-Meter-Daten

Motivation

Im Stromnetz der Zukunft wird es elementar sein, den Stromverbrauch von Haushalten, Gewerbebetrieben, Ladestationen für Elektroautos und weiteren Verbrauchern akkurat zu prognostizieren. Nur so kann auch bei schwankender Einspeisung ein stabiler Netzzustand garantiert werden. Durch die Einführung von Smart Metern nimmt die verfügbare Datenmenge kontinuierlich zu und könnte in Zukunft eine gute Basis zur Lastprognose in Verteilnetzen bieten. Besonders die Granularität der Daten verspricht hohe Erkenntnisgewinne, welche in Kombination mit Datenanalyse-Methoden als Grundlage für eine weitergehende Automatisierung im Netzbetrieb dienen könnte.

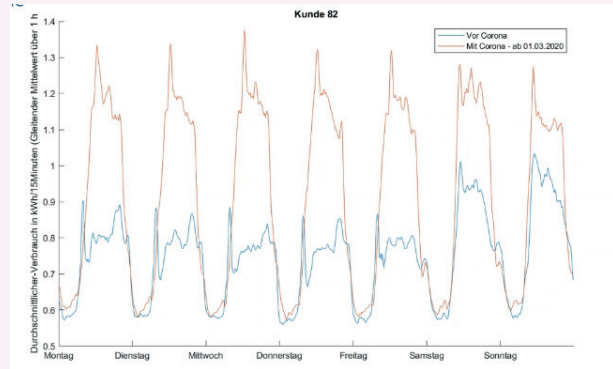
Ziele

Für diese Challenge war es die Aufgabe der Teams Smart-Meter-Datensätze von Netzbetreibern zu analysieren und daraus Verbrauchsprognosen abzuleiten. In einem ersten Schritt sollten Cluster aus den zur Verfügung gestellten Daten gebildet werden, die sich in ihrem Verbrauch ähnelnde Messungen beinhalten.

Auf dieser Grundlage wurden in einem zweiten Schritt 1-Tages-Prognosen für einzelne Smart-Meter, die gebildeten Cluster und das Gesamtportfolio berechnet. Als Teil der Data-Science-Challenge wurden die Teams anschließend aufgrund der Abweichungen ihrer Prognose zu den realen Werten aus dem Testdatensatz bewertet.

Daten

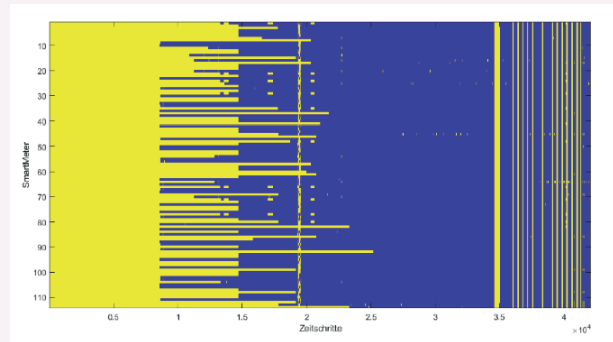
Für diese Challenge konnten vom Stromnetz Hamburg und enercity Netz Verbrauchsdaten zur Prognoseerstellung verwendet werden. Insgesamt erhielten die Teams Informationen von 438 Smart-Metern inklusive einer teilweisen Einteilung in die BDEW-Cluster. Einige Tage des Datensatzes wurden den Teams vorenthalten, um eine Evaluierung der entwickelten Modelle anhand dieser realen Testdaten vornehmen zu können. In der unteren Abbildung ist der Datensatz von enercity visualisiert, wobei ein senkrechter Strich einen Tag an Smart-Meter-Messungen repräsentiert. Während der linke Bereich die unterschiedlichen Inbetriebnahme Zeitpunkte der Geräte zeigt, stellen die gelben Striche auf der rechten Seite die Tage zur Evaluation der Modelle dar, welche den Teams vorenthalten wurden.



Vergleich des Verbrauchs vor und nach Beginn der Corona-Pandemie

Ergebnisse

Bei der Wahl geeigneter Cluster orientierten sich die Teams vom Fraunhofer IEE und NAECO Blue an Jahreszeiten und Wochentagen. Das Team von ifesca hingegen nahm eine Zuordnung der einzelnen Zeitreihen zu den bereits bekannten Standardlastprofilen des BDEW vor. Zwecks einheitlicher Bewertung verständigten sich die Teams und Paten für den zweiten Schritt, der Verbrauchsprognose, auf die Cluster nach letzterer Methode.



Verfügbarkeit der Smart-Meter-Messungen (enercity)

In der anfangs von allen Teams durchgeführten Datenanalyse konnte gezeigt werden, welche Aussagekraft Smart-Meter-Daten besitzen. Als Beispiel sei der Vergleich des aufgezeichneten Verbrauchs einer Woche vor und nach Beginn der Corona Pandemie genannt (s. Abb. oben). Es ist klar erkennbar, dass der private Stromverbrauch in Zeiten des Home-Offices gestiegen ist (rote Linie). Solche Veränderungen, insbesondere die Lastspitzen, sind wichtige Informationen für Verteilnetzbetreiber, welche erst durch den Einsatz von Smart-Metern sichtbar werden.

Fraunhofer IEE

Für die Prognosen nutzte das Team des Fraunhofer IEE verschiedene Machine-Learning-Verfahren wie z.B. Entscheidungsbäume oder Neuronale Netze. Die finale Prognose basiert auf einer Kombination zweier Modelle, um Lücken in den Messwerten kompensieren zu können. Als besonders geeignet die komplexen Zusammenhänge zu erkennen, wird das Verfahren mithilfe eines Neuronalen Netzes genannt.

ifesca

Das Team von ifesca wählte als Kern einen Gradient Boosting Regressor, der eine geringe Trainingszeit besitzt und wenig anfällig für eine Überanpassung an Trainingsdaten ist, was den Einsatz in der Praxis erleichtert. Als finales Modell werden mehrere Stufen dieser Regressoren verwendet, wobei jede Stufe einen anderen Prognosefokus besitzt.

NAECO Blue

Auch NAECO Blue trainierte diverse Modelle, darunter Neuronale Netze, statistische Methoden und Regressoren. Dabei unterschieden sich die Ergebnisse je nach Datensatz. Bei der geringeren Datenmenge konnte das statistische Modell die Machine-Learning-Verfahren überbieten, ansonsten wurden die Boosting-Regressoren als bestes Modell gewählt.

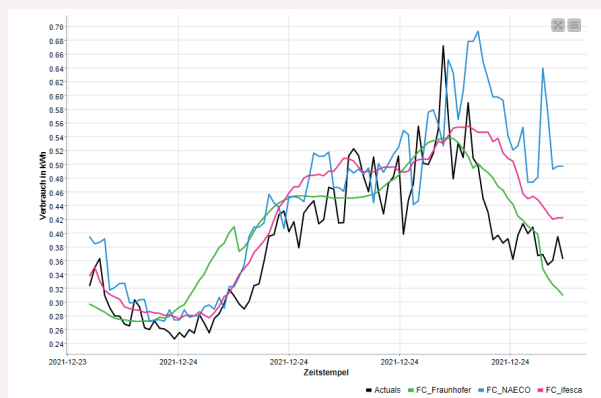
Beim Vergleich der Prognosen der Teams untereinander ist erkennbar, dass es zwei mögliche Modellierungsansätze gibt (s. Abb. rechts). Die Vorhersage ermittelter Werte oder die Nachbildung der realen Verbräuche inklusive kurzzeitiger Schwankungen. Mit Blick auf die gemessenen Abweichungen der Prognosen im Testdatensatz, scheint der gemittelte Ansatz bessere Ergebnisse zu erzielen, allerdings gaben die Netzbetreiber an, dass besonders die Spitzenlasten für die Netzberechnung ein wichtiger Faktor seien und daher auch der zweite Ansatz Vorteile bietet.

Erkenntnisse

Auch wenn die Prognosen der Teams bereits gute Ergebnisse mit geringen Abweichungen zu den realen Daten hervorgebracht haben, ist die Erweiterung des Datensatzes ein Kernelement für eine Optimierung der Modelle. Im Zuge des Smart-Meter-Rollouts bleibt daher zu beobachten, ob durch die steigende Datenmenge auch die schwer vorhersagbaren Spitzenlasten von KI-Modellen erkannt werden können.

Exogene Einflüsse konnten von den Teams nahezu nicht untersucht werden. Diese versprechen jedoch hohe Potenziale in Bezug auf die Vorhersagbarkeit der Verbräuche. Eine Kombination aus Smart-Meter-Daten und weiteren Verbraucherangaben sollte aus datenschutzrechtlicher Sicht allerdings kritisch analysiert werden, da die Anonymität der Nutzer erhalten bleiben muss.

In Hinblick auf die Modellierung der KI-Modelle ist in dieser Challenge verdeutlicht worden, dass die Wahl der Evaluationsmetrik ein relevanter Faktor zur Zielerreichung ist. Der gewählte Root-Mean-Squared-Error bietet zwar die Vergleichbarkeit der Ergebnisse, allerdings belohnt dieser auch das Mitteln der Verbräuche, wohingegen Netzbetreiber an den Spitzenlasten interessiert sind. Dieser Fall zeigt beispielhaft die Herausforderungen, mit denen bei der Umsetzung von KI-Projekten zu rechnen ist.

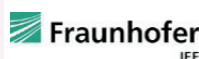


Verbrauchsprognose im Tagesverlauf in kWh

Ausblick

Mit Blick auf die Zunahme von Smart-Metern im Stromnetz sind zwei Faktoren für Netzbetreiber von Relevanz. Einerseits sollten die notwendigen datenbasierten Systeme so ausgelegt sein, dass eine Skalierung auf große Datenmengen erfolgen kann. Besonders die Rechenzeit muss bei einer Operationalisierung in Betracht gezogen werden. Andererseits können sich die Netzbetreiber die Frage stellen, inwieweit die Standardlastprofile verändert oder erweitert werden müssen, da Auswertungen der Teams belegen, dass bereits Verschiebungen der Spitzenlasten oder auch neue Peaks im Vergleich zu den Standardlastprofilen auftreten.

1. Platz



2. Platz



3. Platz



Platzierungen

Die Challenge hat gezeigt, dass die Verwendung von Smart Metern erhebliche Potenziale für Netzbetreiber bietet und daher verbundene Anwendungsfälle frühzeitig implementiert werden sollten. Neben Datenschutzrechtlichen Aspekten sollten aber auch die Anforderungen an klare Zieldefinitionen vor der Umsetzung beachtet werden.