



# PROGNOSE- UND DATENANALYSEPLATTFORM FÜR GAS UND STROM

IMPLEMENTIERUNG EINER ZENTRALEN EUROPaweITEN FORECASTING- UND DATENANALYSE-LÖSUNG

## AUSGANGSSITUATION

Unser Kunde bietet Dienstleistungen in den Bereichen Energieversorgung, erneuerbare Energien und innovative Energielösungen, wobei der Fokus auf einem beschleunigten Übergang hin zu einer dezentralen und nachhaltigen Energiezukunft liegt. Das Unternehmen strebt danach, durch Kundenorientierung, fortschrittliche Energielösungen und effiziente Netzwerke die Energieversorgung der Zukunft sicher und umweltfreundlich zu gestalten. Vor der Migration verwendete unser Kunde eine Vielzahl von individuell pro Region entwickelten Prognosemodellen für kurzfristige Vorhersagen zum Gas-, Strom- und Wärmeverbrauch sowie zur Produktion von Wind- und Solarenergie im innereuropäischen Raum. Diese Modelle waren unabhängig voneinander implementiert und basierten größtenteils auf proprietären Black Box Modellen, die eine Fehlersuche bei falschen Vorhersagen erschwerten. Viele der Quelldaten kamen aus arbeitsintensiven und sich wiederholenden manuellen Aufgaben und nutzten unterschiedliche Werkzeuge, Zeitgranularitäten, Quelldaten, Datenqualitäten und Standards. Diese mussten den lokalen Anforderungen für gesetzliche Auflagen unter anderem in der Kundendatenanonymisierung entsprechen. Zudem verursachten die jährlichen Lizenzgebühren der bisherigen Prognosetools erhebliche Kosten. Um die Lizenz- und Nutzungsgebühren zu verringern, die Datentransparenz zu erhöhen und den Data-Governance-Ansprüchen zu entsprechen, war die schrittweise Migration der regionalen Einheiten hin zu einer zentral gewarteten und vereinheitlichten Datenplattform mit eigens im Hause entwickelten Modellen geplant. Unsere Berater\*innen haben von Anfang an bei der Entwicklung der ersten regionalen Vorhersagelösung und der zugehörigen ETL-Strecke mitgearbeitet. Auf dieser technischen Umsetzung basierten die Modelle der nachfolgenden Jahre maßgeblich.

## VORGEHEN

Teil der Anforderungen für den Projekterfolg war es, die unabhängig voneinander agierenden regionalen Einheiten unseres Kunden individuell von den Vorteilen der zentralen Plattform zu überzeugen. Nach erfolgreichen Pilotphasen beschlossen die ersten, ihre bestehenden Modelle auf die neue zentrale Plattform zu migrieren. Grundlegend wurde der Migrationsprozess dabei unter den folgenden vier Gesichtspunkten angegangen:

**Quelldaten-Migration:** Die grundlegenden Prozesse der Quelldatenbeschaffung wurden von manuellen Downloads und Excel-basierten Eintragungen sowie proprietären Drittanbieterlösungen auf automatisierte API-Calls und Echtzeitdatenabfragen umgestellt. Die Daten wurden von den alten, dezentralen Data-Warehouse-Lösungen auf eine zentrale Datenplattform migriert und kuratiert. Das durch das Projektteam geschulte Personal sorgte dafür, dass die Kontrolle und die Verantwortung über die Daten und Modelle gemäß den Vorgaben bei den jeweiligen regionalen Einheiten blieben.

**Transformation und Feature-Engineering:** Mittels Azure Databricks, Azure Datafactory, einer kuratierten Logik-Komponentenbibliothek in Azure Machine Learning und serverlosen Funk-

## ENERGIE

**Unsere Consultants legten gemeinsam mit den Entwickler\*innen unserer Kunden den Grundstein für eine europaweit genutzte und zentral kuratierte Datenplattform. Dank einer Lösung im Microsoft Azure Cloud Ökosystem wurde der Austausch zwischen unterschiedlichen regionalen Einheiten unseres europaweit agierenden Kunden vereinfacht. Neue Implementierungen können problemlos mittels eines einfach übertragbaren Schemas realisiert werden. Die Wartung und die Optimierung bereits bestehender Modelle wurden drastisch vereinfacht. Zudem wurden Abhängigkeiten und Kosten bei den dezentralen proprietären Vorhersagemodellen und Datenverarbeitungslösungen reduziert.**

## TECHNOLOGIEN & METHODEN

- Python
- pandas
- Apache Spark
- Scikit-learn
- TensorFlow
- Prophet
- Azure Data Lake Storage Gen2
- Azure Machine Learning
- Azure Data Factory
- Azure Databricks
- Azure Functions App
- Snowflake

tionen in Azure Functions haben wir eine modulare, baukastenartige Lösung konzipiert. Sie ist - je nach notwendiger Ausführungsgeschwindigkeit, Komplexität und Häufigkeit der Prozesse - wartbarer, anpassbarer sowie einfacher verständlich und transparenter als die vorangegangenen Lösungen.

**Vorhersagemodellierung:** Den regionalen Anforderungen entsprechend, wurden teils existierende Modelle aus anderen technologischen Frameworks mittels Reverse-Engineerings in das eigene Ökosystem übertragen oder bei Bedarf gänzlich neu implementiert. Diese komplexen und teils aus mehreren hundert Features bestehenden Modelle wurden unter Einhaltung datenwissenschaftlicher Standards mit Modeling-Experimenten und explorativen Datenanalysen zur Feature-Entwicklung und Hyperparameter-Optimierung analysiert. Die Komplexität der Vorhersagemodelle reichte dabei je nach Bedarf und Datenlage von linearen Regressoren bis zu neuronalen Netzen mit mehreren hundert Features.

**Distribution:** Die Vorhersagedaten landeten nach dem Post-Processing in der gemeinsamen Datenplattform. Wöchentliche Schulungen bereiteten die betroffenen Analysten-Teams gründlich auf die Nutzung, die Wartung und eventuelle Erweiterung der Modelle vor.

## ERGEBNIS

Geopolitische Entwicklungen haben zunehmend spürbare Auswirkungen auf viele wirtschaftliche Bereiche, insbesondere auf den Gas- und Energiemarkt. Eine gesteigerte Reaktionsfähigkeit des Unternehmens auf unvorhersehbare Ereignisse liefert somit einen enormen Wettbewerbsvorteil und den größten Motivator für eine Umstellung auf eine gemeinsame Plattform mit erklärbaren Vorhersagemodellen.

Die zentralisierten und kuratierten Daten in der Plattform fördern immens die Flexibilität der regionalen Analyst\*innen und ihren Austausch über Ländergrenzen hinweg. Sie sorgen zugleich für klare Verantwortlichkeiten im Falle von Defekten und Ausreißern in den Daten. Regelmäßige Mentoring-Sessions durch die implementierenden Teams sorgten für einen effizienten gegenseitigen Austausch zu technischen Anforderungen und ermöglichten ein schnelles und nachhaltiges Einarbeiten in die neue Plattform. Die kuratierten AML-Komponenten entsprechen dem Low-Code-Ansatz und sind wahlweise vollständig ohne Coding-Kenntnisse nutzbar, bzw. mittels der AML Python SDK einfach anpassbar. Beide Ansätze wurden - je nach technischem Stand und Bedarf der Analysten-Teams - im Rahmen der Mentoring-Sessions gecoach, um die nachhaltige Nutzbarkeit der neuen Plattform zu gewährleisten.

Die verbesserte Modell-Performance und der Go-live der ersten von drei aktuell involvierten regionalen Einheiten haben weitere Regionen von der Effizienz der neuen zentralisierten Lösung und dem damit verbundenen Coaching-Ansatz überzeugt. Im Laufe der letzten drei Jahre wuchs das implementierende Team von rund 5 auf inzwischen fast 30 Mitarbeitende an, um der wachsenden Nachfrage im Unternehmen gerecht zu werden.

Aufgrund des Projekterfolgs ist auch von einer vierten und fünften Region die Migration zur zentralen Datenplattform geplant. Infolge ihrer zunehmenden Relevanz ist die Implementierung eines zentralen Datenmodells für Daten jenseits des Forecastings in Arbeit, um einen einheitlichen Datenstandard für alle europäischen Geschäftsregionen zu etablieren.