

Wie Data Science Prognosen im hochvolatilen Strommarkt ermöglicht

Advanced Analytics – Prognosen auch in volatiltem Umfeld

„Künstliche Intelligenz“ (KI) ist heutzutage in aller Munde und verspricht sowohl enorme Effizienzsteigerungen als auch völlig neuartige Anwendungen. So haben schon sehr viele Unternehmen ein KI-Projekt durchgeführt, in der Hoffnung, von diesem Trend zu profitieren. Doch leider scheitern allzu viele dieser Projekte – meist nicht an der technischen Expertise, sondern an unklaren Zielsetzungen, übertriebenen Erwartungen und dem Wunsch, unbedingt „KI“ einzusetzen, wo es eventuell auch ein einfacher Mittelwert mit Standardabweichung tut. Man spricht hier auch von Advanced Analytics. Sinnvoll ist es deshalb, ein Vorgehensmodell zu verwenden, das klare Zielsetzungen bedingt und sich technisch nicht nur auf Fancy-Analysemöglichkeiten beschränkt. Ein solches Modell stellen wir am Beispiel eines Preis-Forecasts im Strommarkt vor.

Strommärkte sind sehr komplexe Konstrukte, die im exakten Gleichgewicht zwischen Energieerzeugung und -verbrauch gehalten werden müssen. Denn Strom ist nicht lagerbar! Das klingt zunächst nach einem physikalischen Detail, hat aber enorme Auswirkungen auf den Strommarkt: Marktteilnehmer *müssen* teilweise kaufen oder verkaufen, da sie sonst einen Blackout riskieren.

Advanced Analytics

Unter Advanced Analytics versteht man autonome oder halbautonome Analyseverfahren von Daten oder Inhalten. Hierbei werden unterschiedlichste Methoden und Technologien genutzt, zum Beispiel Verfahren der sogenannten Künstlichen Intelligenz, wie Machine Learning oder Neuronale Netze. Aber auch mit weniger komplexen Ansätzen aus der Statistik und der Stochastik können in vielen Advanced-Analytics-Projekten hervorragende Ergebnisse erzielt werden. Insofern ist das Spektrum der genutzten Methoden und Technologien relativ umfangreich und muss – Case-basiert – auf die jeweilige Zielsetzung hin angepasst werden. Dabei gilt der Grundsatz: vom Einfachen hin zum Komplexen.

Beispielsweise kann die Prognose an der Strombörse durch Neuronale Netze erfolgen, wobei sie auf einer Vielzahl von verschiedenen Datenpunkten aufsetzt. Erfahrungsgemäß ist es aber schneller und einfacher, zunächst mit einfachen Regressionsmodellen zu starten, um dann – wenn nötig – komplexere Verfahren zu testen.

Infolge dieses Verkauf-/Kaufzwangs kommt es zu enormen Preisschwankungen von bis zu 100 Prozent. Auch negative Verkaufspreise sind keine Seltenheit im Strommarkt. In diesem hochvolatilen Strommarkt versuchen die Marktakteure natürlich, ihren eigenen Strom möglichst gewinnbringend zu verkaufen bzw. möglichst günstig fremden Strom einzukaufen und so ihren Gewinn zu erhöhen bzw. Verlustrisiken zu minimieren. Entsprechend kann eine automatische Preisvorhersage bzw. -tendenz hier zu enormen Gewinnen führen und Verluste minimieren. Da Analytics-Projekte hochiterative und dynamische Projekte sind, ist ein exploratives Vorgehen unumgänglich. Nur so kann man sich einem so komplexen Thema Schritt für Schritt nähern. Abbildung 1 zeigt schematisch das Vorgehensmodell, das sich am CRISP-DM- und ASUM-DM-Prozess orientiert.

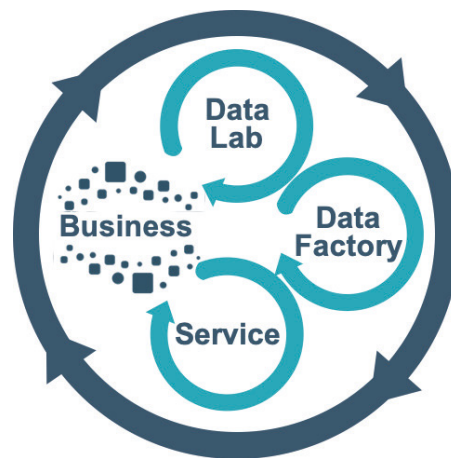


Abb. 1: Das infologistix Vorgehensmodell

Business

In diesem ersten Prozessschritt gilt es die Idee eines Forecasts im Strommarkt zu konkretisieren: Welche(s) Stromprodukt(e) soll(en) prognostiziert werden? Über welchen Zeitraum? Reicht eine einfache Preistendenz oder muss es ein konkreter Preis sein?

Hier sollten unbedingt die Data Scientists mit einbezogen werden, damit die Ziele möglichst klar und realistisch formuliert werden und keine überzogenen Erwartungen entstehen. Für den Strommarkt wurde mit einem simplen

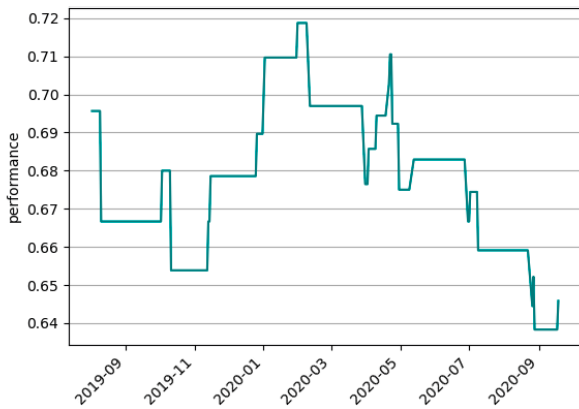


Abb. 2: Performance der eingesetzten Forecast-Verfahren

Preistrend (Kaufen/Verkaufen) begonnen, und das auch nur für ein einzelnes Produkt – den Day-Ahead-Markt, weil er einen recht kurzen Prognosezeitraum hat. Erst wenn man für diesen simplen Use-Case einen funktionierenden Service hat, beginnt man weitere Produkte zu erschließen.

Data Lab

Das Data Lab ist mit Abstand die dynamischste Phase des gesamten Projekts. Hier wird explorativ untersucht, welche Daten für einen Forecast relevant sind, welche Methoden sich am besten eignen und wie sie optimal umzusetzen sind. Dabei gilt immer die Devise: „fail fast, fail cheap“.

Für den Day-Ahead-Markt wurden zunächst historische Daten der vergangenen Jahre zusammengestellt sowie Preisdaten anderer Strommarktprodukte. Mit Hilfe einer (Auto-)Korrelationsanalyse konnten die interessantesten Daten für eine Prognose identifiziert werden.

Diese Daten dienen dann als Features für Prognose-Algorithmen. Die Auswahl an möglichen Algorithmen ist hier sehr groß – die Erfahrung hat aber gezeigt, dass es sinnvoll ist, mit einfachen zu beginnen und erst, wenn diese keine gute Performance zeigen, zu komplexeren und damit aufwendigeren überzugehen. Einfache Regressionsverfahren zeigten im Fall des Day-Ahead-Marktes schon gute Werte. Komplexe Verfahren aus dem Machine-Learning-Bereich wie zum Beispiel KNN, Random Forests oder SVMs konnten diese Performance leider nicht verbessern.

Am Ende erzielte eine Kombination verschiedener Verfahren das beste Ergebnis. Die Vorhersagegenauigkeit (Performance) konnte konstant auf über 65 Prozent erhöht werden (siehe Abbildung 2).

Für die Analyse wurde Python mit den NumPy-, pandas- sowie Scikit-learn- und statsmodels-Bibliotheken ver-

wendet. Wichtig war hier ein offenes System, das schnelles Programmieren und Änderungen erlaubt.

Data Factory

Mit dem PoC aus dem Data Lab wurde ein System aufgesetzt, das tagesaktuell Preisdaten von der deutschen Strombörse in Leipzig bezieht. Diese Daten werden dann jeden Vormittag für einen aktuellen Forecast auf den Day-Ahead-Markt verwendet. Hierzu wurde die Python-Umgebung aus dem Data Lab um einen Klickbot/Webcrawler erweitert sowie um eine InfluxDB, um die Daten abzuspeichern.

Service

Das Handelssignal – das eigentliche Ergebnis des gesamten Prozesses – kann dann entweder per E-Mail-Benachrichtigung zugestellt oder per Online-Dashboard visualisiert werden.

Fazit

Auch in einem so volatilen und schwer zu prognostizierenden Umfeld wie dem Strommarkt können durch Advanced Analytics zuverlässige Prognosen entwickelt werden. Hierzu sind weniger spezielle Hard- oder Softwarekombinationen notwendig, sondern eine zielgerichtete Vorgehensweise.

Grundlage ist immer ein klar definierter Business-Case (Ziel) – in unserem Beispiel ein Forecast für das Day-Ahead Trading. Ferner muss man akzeptieren, dass sich die Ergebnisse täglich ändern können und man unter Umständen in eine Sackgasse läuft. Analytics-Projekte sind eben iterativ und hochdynamisch. Und schließlich gilt das Gebot, mit einfachen Analysen zu beginnen und erst, wenn das nicht mehr reicht, weitere, komplexere Ansätze zu nutzen – bis hin zum Neuronalen Netz.

Kontakt

infologistix GmbH

Dr. Alexander Wunderle
 Gutenbergstraße 7
 D-85748 Garching

Telefon: +49 (0)89 81 88 59 79
 E-Mail: AdvancedAnalytics@infologistix.de
 Web: www.infologistix.de