

# e|m|w

Energie. Markt. Wettbewerb.

**e|m|w.trends**

## Stromhandel mit maschinellen Agenten

Von **Dr. Ludwig Rauch**, Machine Learning Engineer,  
**David Goebel**, Manager Data Science, EXXETA AG  
und **Julius Kittler**, Masterand, Linköping University, EXXETA AG

# Stromhandel mit maschinellen Agenten

## Algorithmischer Stromhandel mit Reinforcement Learning im EPEX Intraday-Markt

Automatisierter Handel beherrscht mittlerweile den Intraday-Markt der EPEX. Indem von Menschen entwickelte Handelsstrategien durch Algorithmen abgebildet werden, können Computer Strom ohne menschliche Interaktion handeln. Den nächsten evolutionären Schritt für derartige Algorithmen stellen maschinelle Agenten dar, die selbstständig Strategien durch Reinforcement Learning entwickeln.

✦ Von **Dr. Ludwig Rauch**, Machine Learning Engineer, **David Goebel**, Manager Data Science, EXXETA AG und **Julius Kittler**, Masterand, Linköping University, EXXETA AG

Das Potenzial, unternehmerische Prozesse durch maschinelles Lernen zu optimieren, ist in den vergangenen Jahren zu einem globalen Trend geworden. Neue Disziplinen wie „Data Science“ oder „Machine Learning“ widmen sich der Optimierung von Prozessen durch datengetriebene Analysen und maschinelles Lernen. Die Ergebnisse beeinflussen schon heute unseren Alltag, was sich an Beispielen von autonomem Fahren, Bilderkennung oder sprachgesteuerter Software belegen lässt.

In diesem Artikel wird eine neue, zukunftsweisende Technik des maschinellen Lernens für den Intraday-Stromhandel vorgestellt. Dabei entwickeln die Autoren intelligente Agenten, die autonom Entscheidungen treffen können. Diese Technik wurde bereits von Google in AlphaGo erfolgreich eingesetzt, um den derzeitigen Großmeister im chinesischen Brettspiel Go durch selbstständig erlernte und dem Menschen bisher unbekannte Strategien zu besiegen. Können diese maschinellen Agenten auch neue, profitablere Strategien im Intraday-Stromhandel erlernen und diesen damit optimieren?

### Stromhandel im Intraday-Markt

Der Intraday-Markt der EPEX bietet Händlern die Möglichkeit, Strom kontinuierlich über einen Zeitraum von einem Tag vor der physischen Lieferung zu handeln. Das Ziel der Händler ist es, kurzfristige

Änderungen in der Stromproduktion und im Stromverbrauch möglichst genau zu kompensieren, um die Netzstabilität in Deutschland sicherzustellen. Die Auslieferung des Stroms erfolgt in viertelstündlichen, halbstündlichen oder stündlichen Zeitintervallen, welche die drei Hauptprodukte dieses Strommarktsegments definieren. Somit ist es Stromproduzenten sowie gewerblichen Verbrauchern möglich, die Energiebilanz von Erzeugung und Verbrauch möglichst gering zu halten, um auf diese Weise gesetzliche Ausgleichszahlungen bei Abweichungen zu vermeiden.

Unter diesen Bedingungen kann ein algorithmisch gesteuerter Handel großes Potenzial freisetzen, da sich der Strompreis hauptsächlich durch Angebot und Nachfrage bestimmt. So kann zum Beispiel das zukünftige Stromangebot durch fortlaufend aktualisierte Wetterprognosen vorhergesagt werden. Der zu erwartende Stromverbrauch ist mithilfe historischer Verbrauchsdaten modellierbar. Durch Parameter, die das zu erwartende Angebot und die zu erwartende Nachfrage abbilden, kann der Strompreis frühzeitig mithilfe von Algorithmen vorhergesagt werden.

Computer können dabei, im Gegensatz zu menschlichen Händlern, große Datenmengen in Sekundenschnelle verarbeiten und einen Informationsvorteil generieren. Im Vergleich zu herkömmlichen Lösungen im algorithmischen Handel kann der

unternehmerische Handelserfolg durch die in Echtzeit berücksichtigte Vielzahl an Parametern und die Anpassung der Algorithmen an individuelle Portfolios der Unternehmen gesteigert werden.

### Reinforcement Learning

Der Begriff „Reinforcement Learning“ (dt. Bestärkendes Lernen) bezeichnet eine Methode, Algorithmen durch maschinelles Lernen (siehe Infokasten: Maschinelles Lernen) zu entwickeln. Die Grundlage bilden Daten, in denen ausreichend Informationen enthalten sind, um relevante Entscheidungen treffen zu können. Im Intraday-Strommarkt ergeben sich die relevanten Variablen aus den Daten der gehandelten Produkte wie Volumen und Preis. Die einzelnen Transaktionszeitpunkte bilden die unterschiedlichen Stromproduktionsperioden und das Stromverbrauchsverhalten ab. Zudem benötigen die maschinellen Agenten eine Vorgabe der erlaubten Handlungen, unter denen Sie abwägen müssen.

Da es sich um eine Handelsumgebung handelt, sind die möglichen Handlungen auf Kaufen, Verkaufen und das Halten einer Position beschränkt. Abschließend muss ein abstraktes Ziel definiert werden, das der Agent versucht zu erreichen. Beispiele sind ein rein profitorientierter Handel durch Kauf, Verkauf und Halten von Positionen oder auch nur der optimale Verkauf einer vordefinierten Strommenge



Foto: © istockphoto/MF3d

eines Stromproduzenten. Der maschinelle Agent versucht in einer sogenannten Trainingsphase, Strategien anhand von Beobachtungen in historischen Handelsdaten zu entwickeln und zu optimieren, um sein Ziel möglichst optimal zu erreichen (siehe Abb. 1). In der Trainingsphase erhält der maschinelle Agent (links) ein abstraktes Ziel, etwa den rein profitorientierten Handel über Kauf, Verkauf und Halten von Positionen. Anschließend führt der Agent Aktionen an der Börse (rechts) aus, evaluiert das Resultat und passt seine nächsten Entscheidungen beziehungsweise Aktionen an. Das wiederholt der Agent solange, bis er keine besseren Entscheidungen mehr findet. Gewünschte Entscheidungen (Aktionen) in Bezug auf das definierte Ziel

werden belohnt, unerwünschte werden bestraft. Zeitpunkt und Höhe der Belohnung müssen vorab definiert werden und beeinflussen die Entwicklung des maschinellen Agenten maßgeblich.

So können die Algorithmen zu tendenziell risikobereiteren oder konservativeren Handelsakteuren ausgebildet werden. Diese Belohnungsfunktion ist vom Entwickler mit großer Sorgfalt zu wählen, damit unerwünschte Strategien und Handlungen vermieden werden. Durch eine große Anzahl an Iterationen von historischen Handelssituationen, getroffenen Entscheidungen, Beobachtungen und zugeteilten Belohnungen wird die Lernfähigkeit der maschinellen Agenten bestärkt.

### Maschinelle Agenten im EPEX Intraday-Stromhandel

Die beschriebene Methodik des Reinforcement Learning wurde in der hier entwickelten Software auf historische Daten der Intraday-Handelsbörse EPEX angewandt, wobei ausschließlich Stundenprodukte betrachtet wurden. Die Methodik kann ohne größere Anpassungen auf viertel- oder halbstündige Stromprodukte erweitert werden.

#### Trainingsphase der maschinellen Agenten

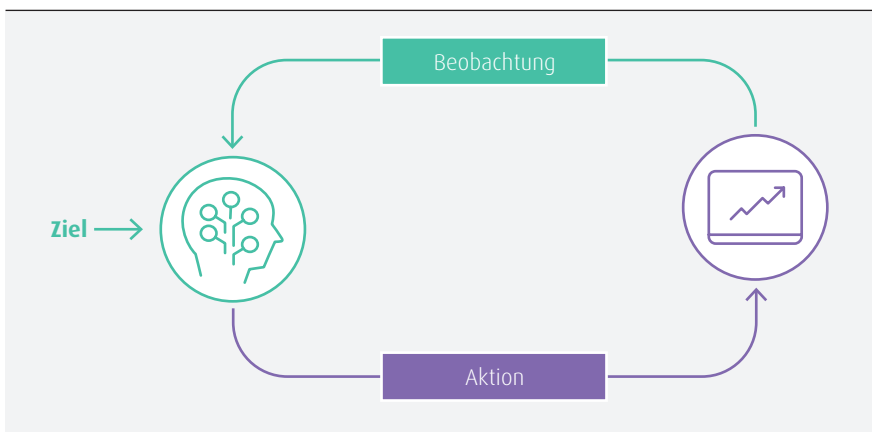
Die über mehrere Jahre gesammelten Handelsdaten werden in die jeweiligen Auslieferungzeitpunkte (Kontrakte) aufgeteilt und jeweils in kürzere Zeitintervalle unterteilt. In diesen diskreten Zeitschritten werden Marktinformationen aggregiert und die für den Handel relevanten Parameter abgeleitet.

Danach werden dem maschinellen Agenten die aufbereiteten Handelsinformationen in allen in der Vergangenheit liegenden Zeitintervallen des Kontrakts als Grundlage für seine Entscheidung zur Verfügung gestellt. Jedes Zeitintervall und jede Historie wird fortlaufend dem maschinellen Agenten als Information gegeben, um so die historischen Handelsverläufe zu simulieren. Dieser Prozess wird solange wiederholt, bis der Agent aufhört, nach neuen Strategien zu suchen und die Summe der Belohnung zu einem konstant hohen Wert konvergiert.

#### Handelsaktionen der Agenten

Nachdem die Trainingsphase des Agenten abgeschlossen ist, werden seine erlernten Handelsstrategien an für ihn bisher unbekanntem Handelstagen evaluiert.

01 Lernmechanismus der Agenten im Reinforcement Learning





## ► Maschinelles Lernen

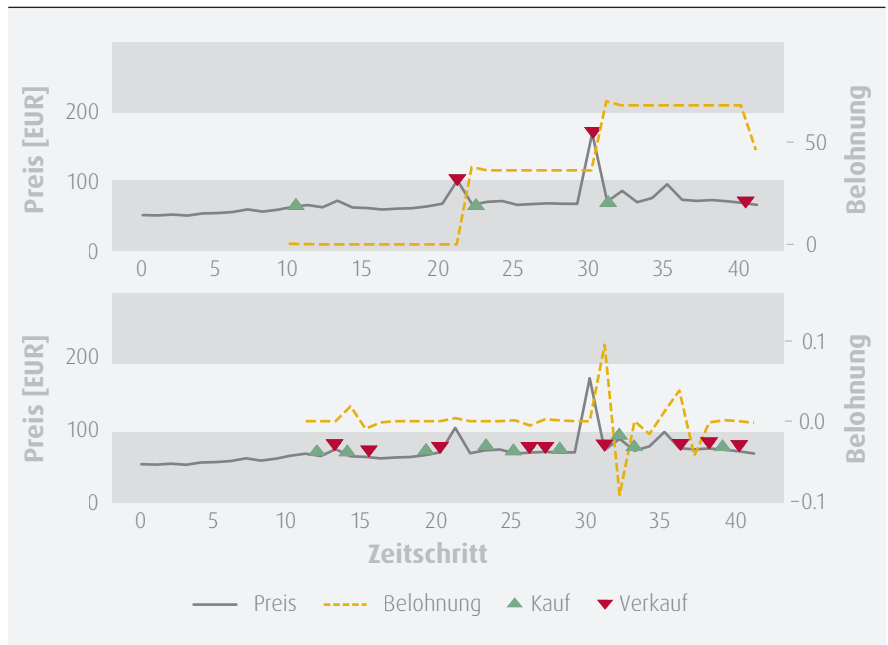
Bei maschinellem Lernen werden Algorithmen eingesetzt, um komplexe Informationen innerhalb von Datensätzen zu erlernen und entweder einzelne Elemente zu klassifizieren oder bestimmte Zielparameter vorherzusagen. Im Allgemeinen werden Techniken des Supervised (dt. überwacht), Unsupervised (dt. unbewacht) und Reinforcement Learning unterschieden. Beim Supervised Learning sind für einen ähnlichen Datensatz die Zielgrößen bereits bekannt und der Algorithmus erlernt aus den Eingabedaten, die Zielgröße zu bestimmen.

Das komplexere Unsupervised Learning versucht, selbstständig und ohne vordefinierte Zielgröße, einen Datensatz zu verstehen. Mit dem erlernten Wissen, wie zum Beispiel ein Sprachverständnis, kann der Algorithmus verwandte sprachliche Themen lösen. Beim Reinforcement Learning werden neben dem Bereitstellen der Daten lediglich erlaubte Handlungen sowie ein abstraktes Ziel vordefiniert. Lösungsstrategien werden dann von dem Algorithmus selbstständig gesucht und optimiert.

Basierend auf diesen neuen Informationen und der erlernten Strategie aus historischen Daten entscheidet sich der Agent an jedem neuen Zeitintervall für eine der drei möglichen Aktionen Kaufen, Verkaufen oder Halten. Die Handelsstrategie kann dann anhand der Handelsperformance bewertet und mit alternativen Algorithmen verglichen werden (Backtests).

In Abb. 2 wird ein solcher Verlauf exemplarisch über einen Kontrakt dargestellt. Der durchschnittliche Strompreis ist über den zeitlichen Verlauf durch die graue Linie gekennzeichnet. Während des Handels sind die Handelsentscheidungen des maschinellen Agenten beim Kauf mit grünen Dreiecken und beim Verkauf mit roten Dreiecken gekennzeichnet. Die positive Belohnung bei profitabilem Handel ist durch die orange Linie illustriert. In diesem Beispiel erkennt man die erlernte Strategie, da der Akteur bei kurzfristigen Kurssprüngen seine Positionen mit Gewinn veräußert. Die Performance der Agenten kann durch sogenannte Backtests mit anderen Algorithmen verglichen werden.

## 02 Beispiel einer Handelssimulation eines gehandelten Stundenproduktes.



Weitere Strategien können durch das Anpassen der Belohnungsfunktion gefunden werden. So kann der Agent zu mehr oder weniger Risiko oder auch zu verstärktem Handel angetrieben werden. In einem zweiten Beispiel (Abb. 2 unten) wurde der Agent dazu angetrieben, nur in Zeiten von geringen Fluktuationen im Markt zu agieren. Zu beobachten ist, dass der Agent häufiger handelt, bei starken Kursänderungen aber passiv agiert. Der maximale beobachtete Verlust ist in diesem Fall geringer, jedoch fällt auch der im Durchschnitt erzielte Gewinn geringer aus.

Der Stromhandel mit maschinellen Agenten steht erst am Anfang der Entwicklung. Reinforcement Learning bietet die Möglichkeit, relevante Parameter für den Stromhandel in Echtzeit zu verarbeiten und selbstständig Handelsstrategien unter Berücksichtigung individueller Randbedingungen zu entwickeln. Die so trainierten maschinellen Agenten sind in der Lage, auf kurzfristige Änderungen und unter der Berücksichtigung vorab definierter Ziele zu reagieren. Die Entwicklung dieser maschinellen Agenten kann nicht nur dazu beitragen, die aktuellen Herausforderungen am Strommarkt zu überwinden, sondern zum Beispiel auch zur Optimierung komplexer Entscheidungen im Bilanzkreismanagement eingesetzt werden. ◀

### DR. LUDWIG RAUCH

#### Jahrgang 1987

- 2014–2017: Promotion in Physik, Max-Planck-Institut für Kernphysik, Heidelberg
- 2017–2019: Postdoctoral Researcher, Astroteilchenphysik, DESY, Berlin
- seit 2019: Machine Learning Engineer, EXXETA AG
- ✉ Ludwig.Rauch@EXXETA.com

### DAVID GOEBEL

#### Jahrgang 1988

- 2013–2016: M.Sc. Energy Science and Engineering, Darmstadt
- 2016–2019: Data Scientist, EXXETA AG
- seit 2019: Data Science Manager, EXXETA AG
- ✉ David.Goebel@EXXETA.com

### JULIUS KITTLER

#### Jahrgang 1994

- 2018–2020: M.Sc. Statistics and Machine Learning, Linköping
- seit 2020: Masterand, EXXETA AG
- ✉ Julius.Kittler@EXXETA.com

# e | m | w

Energie. Markt. Wettbewerb.

energate gmbh

Norbertstraße 3-5

D-45131 Essen

Tel.: +49 (0) 201.1022.500

Fax: +49 (0) 201.1022.555

[www.energate.de](http://www.energate.de)

[www.emw-online.com](http://www.emw-online.com)

Bestellen Sie jetzt Ihre persönliche Ausgabe!

[www.emw-online.com/bestellen](http://www.emw-online.com/bestellen)

