



AWS Whitepaper

Grundsätze der Zeitreihenprognose mit Amazon Forecast



Grundsätze der Zeitreihenprognose mit Amazon Forecast: AWS Whitepaper

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Die Marken und Handelsmarken von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, die geeignet ist, die Kunden zu verwirren oder Amazon in einer Weise herabzusetzen oder zu diskreditieren. Alle anderen Marken, die nicht Eigentum von Amazon sind, sind Eigentum ihrer jeweiligen Inhaber, die mit Amazon verbunden oder nicht verbunden oder von Amazon gesponsert oder nicht gesponsert sein können.

Table of Contents

Kurzbeschreibung und Übersicht	i
Übersicht	1
Sind Sie Well-Architected?	2
Über Prognosen	3
Prognosesystem	3
Wo treten Prognoseprobleme auf?	3
Zu bedenken, bevor versucht wird, ein Prognoseproblem zu lösen	4
Fallstudie: Problem bei der Prognose der Einzelhandelsnachfrage für ein E-Commerce-Geschäft	6
Schritt 1: Daten sammeln und aggregieren	10
Beispiel	12
Schritt 2: Daten vorbereiten	13
Umgang mit fehlenden Daten	13
Beispiel 1	13
Beispiel 2	16
Konzepte der Featurisierung und verwandte Zeitreihen	16
Beispiel 3	17
Schritt 3: Erstellen eines Predictors	19
Schritt 4: Predicators auswerten	21
Backtesting	21
Vorhersagequantile und Genauigkeitsmetriken	23
Gewichteter Quantilverlust (Weighted Quantile Loss, wQL)	23
Gewichteter absoluter prozentualer Fehler (Weighted Absolute Percentage Error, WAPE)	24
Quadratischer Mittelwertfehler (Root Mean Square Error, RMSE)	24
Probleme mit WAPE und RMSE	25
Schritt 5: Generieren und Verwenden von Prognosen für die Entscheidungsfindung	27
Probabilistische Prognosen	27
Visualisierung	28
Zusammenfassung zu Prognose-Workflows und APIs	30
Verwenden von Amazon Forecast für gängige Szenarien	31
Implementierung von Forecast in die Produktion	32
Fazit	34
Mitwirkende	36
Weitere Informationen	37

Anhang A: Häufig gestellte Fragen 38

Anhang B: Referenzen 42

Dokumentverlauf 43

Hinweise 44

AWS-Glossar 45

Grundsätze der Zeitreihenprognose mit Amazonas Forecast

Datum der Veröffentlichung: 1. September 2021 ([Dokumentverlauf](#))

Moderne Unternehmen verwenden alles von einfachen Spreadsheets bis zu komplexer Finanzplanungssoftware, um zu versuchen, zukünftige Geschäftsergebnisse, beispielsweise Produktbedarf, Ressourcenanforderungen oder Finanzperformance, präzise zu prognostizieren. Dieses Whitepaper bietet eine Einführung in die Prognoseerstellung, die verwendete Terminologie und die damit zusammenhängenden Herausforderungen und Anwendungsfälle. Anhand einer Fallstudie werden Prognosekonzepte und Forecast-Schritte beschrieben und es wird erläutert, wie [Amazon Forecast](#) dazu beitragen kann, zahlreiche praktische Herausforderungen zu meistern, die aus realen Prognoseproblemen resultieren.

Übersicht

Die Prognose ist die Wissenschaft der Zukunftsvorhersage. Historische Daten helfen Unternehmen dabei, Trends besser zu verstehen und abzuschätzen, was wann passieren könnte. Und sie können diese Daten in ihre Zukunftspläne integrieren – angefangen beim Produktbedarf über die Bestandsplanung bis hin zur Personalbesetzung.

Angesichts der Folgen, die Prognosen haben könnten, spielt die Genauigkeit hierbei eine wichtige Rolle. Bei einer Überprognose investieren Kunden möglicherweise zu viel in Produkte und Personal und verschwenden damit Investitionen. Bei einer Unterprognose investieren Kunden möglicherweise zu wenig, was zu einem Mangel an Rohstoffen und Lagerbeständen und folglich zu einer schlechten Kundenerfahrung führt.

Heutzutage versuchen Unternehmen, Prognosen sowohl mithilfe einfacher Tabellen als auch mithilfe komplexer Nachfrage-/Finanzplanungssoftware zu erstellen. Es ist jedoch nach wie vor schwer, dabei eine hohe Genauigkeit zu erreichen. Und dies aus zwei Gründen:

- Erstens tun sich traditionelle Prognosen schwer damit, historische Daten in großen Mengen zu integrieren und sind nicht in der Lage, wesentliche Signale aus der Vergangenheit im Datenrauschen zu erkennen.
- Zweitens berücksichtigen herkömmliche Prognosen selten verwandte, zugleich aber unabhängige Daten, die wichtige Zusammenhänge liefern können (wie Preise, Feiertage/Ereignisse, Fehlbestände, Marketingaktionen usw.). Da ihnen die komplette Historie und der breitere Kontext fehlt, sind die meisten Prognosen nicht in der Lage, genaue Zukunftsvorhersagen zu machen.

[Amazon Forecast](#) ist ein vollständig verwalteter Dienst, der diese Probleme löst. Amazon Forecast bietet die besten Algorithmen je nach Prognoseszenario. Es stützt sich auf modernes Machine Learning (ML) und, wenn erforderlich, auf Deep Learning, um hochgenaue Prognosen zu liefern. Amazon Forecast ist benutzerfreundlich und erfordert keine Vorkenntnisse im Bereich Machine Learning. Der Service stellt automatisch die erforderliche Infrastruktur bereit, verarbeitet Daten und erstellt benutzerdefinierte/private ML-Modelle, die in AWS gehostet werden und bereit sind, Vorhersagen zu machen. Amazon Forecast integriert stets die sich rasant weiterentwickelnden Techniken im Bereich Machine Learning, damit Kunden von den Verbesserungen in puncto Genauigkeit profitieren können, und das mit minimalem oder gar keinem zusätzlichen Aufwand ihrerseits.

Sind Sie Well-Architected?

Das [AWS-Well-Architected Framework](#) hilft Ihnen, die Vor- und Nachteile der Entscheidungen zu verstehen, die Sie beim Aufbau von Systemen in der Cloud treffen. Die sechs Säulen des Frameworks ermöglichen es Ihnen, architektonische bewährte Methoden für die Entwicklung und den Betrieb zuverlässiger, sicherer, effizienter, kostengünstiger und nachhaltiger Systeme zu erlernen. Mit dem [AWS Well-Architected Tool](#), das kostenlos in der [AWS-Managementkonsole](#) verfügbar ist, können Sie Ihre Workloads anhand dieser bewährten Methoden überprüfen, indem Sie für jede Säule eine Reihe von Fragen beantworten.

Der Schwerpunkt des [Brennpunkts „Machine Learning“](#) liegt auf der Konzeption, Bereitstellung und Architektur von Machine-Learning-Workloads in der AWS Cloud. Dieser Brennpunkt dient als Ergänzung der bewährten Methoden für das Well-Architected Framework.

Weitere Expertenleitungen und bewährte Methoden für Ihre Cloud-Architektur – Bereitstellungen von Referenzarchitekturen, Diagramme und Whitepapers – finden Sie im [AWS-Architekturzentrum](#).

Über Prognosen

In diesem Dokument beschreibt der Begriff „Prognose“ die Vorhersage zukünftiger Werte einer Zeitreihe: Die Eingabe- oder Ausgabedaten eines Problems sind zeitabhängig.

Prognosesystem

Ein Prognosesystem schließt eine Vielzahl von Benutzern ein:

- Endbenutzer, die die Prognose für ein bestimmtes Produkt wünschen und entscheiden, wie viele Einheiten gekauft werden sollen. Dabei kann es sich um eine Person oder ein automatisiertes System handeln.
- Geschäftsanalysten/Business Intelligence, die Endbenutzer unterstützen und aggregierte Berichte ausführen und organisieren.
- Datenwissenschaftler, die Nachfragemuster und kausale Effekte iterativ analysieren und neue Funktionen hinzufügen, um das Modell schrittweise zu verbessern oder das Prognosemodell zu verbessern.
- Ingenieure, die die Infrastruktur der Datenerfassung einrichten und sicherstellen, dass die Eingabedaten dem System zur Verfügung gestellt werden.

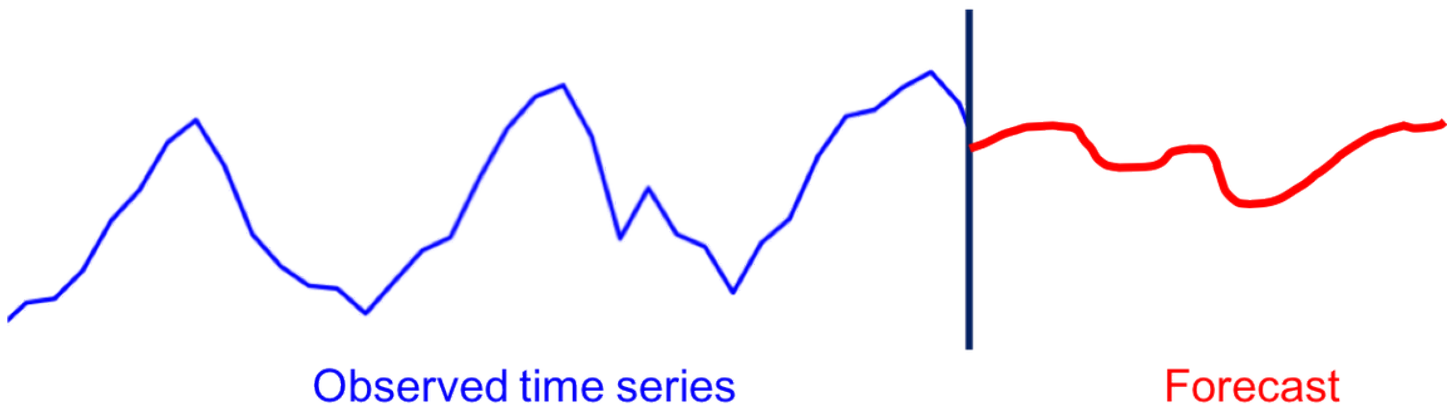
Amazon Forecast erleichtert Softwareingenieuren die Arbeit und ermöglicht Unternehmen, die nur begrenzt über Data-Science-Funktionen verfügen, die Nutzung modernster Prognosetechnologien. Für Unternehmen, die über Datenwissenschaftsfunktionen verfügen, bietet Amazon Forecast eine Reihe von Diagnosefunktionen, die dabei helfen, Prognoseprobleme effektiv zu lösen.

Wo treten Prognoseprobleme auf?

Prognoseprobleme treten in vielen Bereichen auf, in denen Zeitreihendaten auf natürliche Weise erzeugt werden. Dazu zählen Einzelhandelsumsätze, medizinische Analysen, Kapazitätsplanung, Überwachung von Sensornetzwerken, Finanzanalyse, Social Activity Mining und Datenbanksysteme. In den meisten Unternehmen spielen Prognosen bei der Automatisierung und Optimierung von Betriebsprozessen eine Schlüsselrolle, weil sie eine datengesteuerte Entscheidungsfindung ermöglichen. Prognosen zum Produktangebot und zur Produktnachfrage können zur Optimierung der Bestandsverwaltung, Personalplanung und Topologieplanung verwendet werden und

stellen im Allgemeinen eine wichtige Technologie im Hinblick auf die meisten Aspekte der Lieferkettenoptimierung dar.

Die folgende Abbildung bietet einen Überblick über das Prognoseproblem, wenn dieses auf einer beobachteten Zeitreihe basiert, die ein Muster aufweist (in diesem Beispiel ist es die Saisonalität), und es wird eine Prognose über einen bestimmten Zeitraum hinweg erstellt. Die horizontale Achse stellt die Zeit dar, die von der Vergangenheit (links) in die Zukunft (rechts) reicht. Die vertikale Achse stellt gemessene Einheiten dar. Die Vergangenheit (blau) erstreckt sich bis zur vertikalen schwarzen Linie, die Zukunft (rot) entspricht der Prognoseaufgabe.



Überblick über die Prognoseaufgabe

Zu bedenken, bevor versucht wird, ein Prognoseproblem zu lösen

Bevor versucht wird, Prognoseprobleme zu lösen, sollten folgende wichtige Fragen beantwortet werden:

- Müssen Sie ein Prognoseproblem lösen?
- Warum versuchen Sie das Prognoseproblem zu lösen?

Da Zeitreihendaten überall vorkommen, lassen sich Prognoseprobleme überall leicht auffinden. Es stellt sich jedoch die Kernfrage, ob es wirklich notwendig ist, ein Prognoseproblem zu lösen, oder ob darauf gänzlich verzichtet werden kann, ohne gleichzeitig Abstriche in puncto effizienter Entscheidungsfindung im Unternehmen machen zu müssen. Es ist sehr wichtig, sich diese Frage zu stellen, da Prognosen im Zusammenhang mit Machine Learning aus wissenschaftlicher Sicht das größte Problem darstellen.

Denken Sie beispielsweise an Produktempfehlungen für einen Online-Händler. Dieses Problem der Produktempfehlung kann als ein Prognoseproblem angesehen werden, bei dem Sie für jedes

Paar von Kundenlagerhaltungseinheiten (SKU) die Anzahl der Einheiten eines bestimmten Artikels prognostizieren, die dieser spezifische Kunde kaufen würde. Diese Problemformulierung bietet zahlreiche Vorteile. Ein Vorteil ist, dass dabei die Teilkomponente ausdrücklich berücksichtigt wird, sodass Sie Produkte entsprechend der Kaufmuster der Kunden empfehlen können.

Probleme im Zusammenhang mit Produktempfehlungen werden jedoch selten als ein Prognoseproblem formuliert, da es viel schwieriger ist, ein solches Prognoseproblem (z. B. spärliche Informationen auf Kunden-SKU-Ebene und das Ausmaß des Problems) zu lösen, als das Empfehlungsproblem direkt zu lösen. Wenn Sie also eine Prognoseanwendung auswählen, sollten Sie zunächst überlegen, wozu die Prognose später dienen wird, und Sie sollten überprüfen, ob sich das Problem mithilfe eines alternativen Ansatzes lösen lässt.

[Amazon Personalize](#) kann Ihnen dabei helfen. Amazon Personalize ist ein Machine Learning-Service, der Entwicklern das Erstellen individualisierter Empfehlungen für Kunden mithilfe ihrer Anwendungen erleichtert.

Nachdem Sie festgelegt haben, dass Sie ein Prognoseproblem lösen müssen, sollten Sie sich als Nächstes die Frage stellen, warum Sie es lösen? In zahlreichen Geschäftsumgebungen dienen Prognosen in der Regel nur als Mittel zum Zweck. Im Einzelhandel kann die Prognose beispielsweise dazu dienen, Nachfrageprognosen zu erstellen, um auf deren Basis Entscheidungen zur Bestandsverwaltung zu treffen. Das Prognoseproblem bietet typischerweise Input für ein Entscheidungsproblem, das wiederum als Optimierungsproblem modelliert werden kann.

Zu solchen Problemen gehören beispielsweise Entscheidungen darüber, wie viele Einheiten erworben werden müssen oder welcher Ansatz sich am besten dazu eignet, den vorhandenen Lagerbestand zu verwalten. Andere Probleme in puncto Geschäftsprognose sind Prognosen zur Serverkapazität oder zur Nachfrage nach Rohstoffen/Teilen im Fertigungsbereich. Diese Prognosen können als Input für andere Prozesse verwendet werden: entweder für Entscheidungsprobleme (wie oben beschrieben) oder für Szenariosimulationen, die anschließend zur Planung ohne explizite Modelle genutzt werden. Es gibt Ausnahmen von der Regel, dass Prognosen nicht dem Selbstzweck dienen. In der Finanzprognose wird die Prognose beispielsweise direkt zum Anlegen finanzieller Reserven verwendet oder sie wird Anlegern präsentiert.

Um den Zweck von Prognosen zu verstehen, sollten Sie sich die folgenden Fragen stellen:

- Wie weit in die Zukunft sollte Ihre Prognose reichen?
- Wie oft müssen Sie Prognosen erstellen?
- Gibt es bestimmte Aspekte der Prognosen, mit denen Sie sich eingehend befassen sollten?

Fallstudie: Problem bei der Prognose der Einzelhandelsnachfrage für ein E-Commerce-Geschäft

Betrachten Sie den Fall eines E-Commerce-Geschäfts, das Produkte online verkauft, und verschaffen Sie sich auf diese Weise einen Überblick über Prognosekonzepte. Die Optimierung von Entscheidungen für Lieferketten (z. B. die Lagerverwaltung) ist für die Kernwettbewerbsfähigkeit dieses Unternehmens ausschlaggebend, da dadurch ermöglicht wird, an den entsprechenden Fulfillment-Standorten über die genaue Anzahl von Produkten zu verfügen. Dies bedeutet im Wesentlichen, dass das Unternehmen eine große Auswahl, kürzere Versandzeiten und wettbewerbsfähige Preise anbieten kann, was wiederum die Kundenzufriedenheit erhöht. Der wichtigste Input für das Lieferketten-Softwaresystem ist die Nachfrageprognose oder die Prognose des potenziellen Umsatzes, den die einzelnen Produkte aus dem Katalog generieren können. Diese Prognose ermöglicht wichtige nachgelagerte Entscheidungen, wobei folgende Entscheidungen zu den wichtigsten zählen:

- Planung auf Makroebene (strategische Prognose): Wie hoch ist das prognostizierte Wachstum in Bezug auf den Gesamtumsatz/Umsatz für ein Unternehmen? In welcher Region sollte das Unternehmen aktiv(er) sein? Wie sollten Arbeitskräfte verteilt sein?
- Nachfrage- (oder Bestands-)prognose: Wie viele Einheiten jedes Produkts werden voraussichtlich pro Standort verkauft?
- Werbeaktionen (taktische Prognosen): Wie sollten Werbeaktionen ausgeführt werden? Sollten Produkte liquidiert werden?

Der Rest der Fallstudie konzentriert sich auf das zweite Problem, das zur Familie der Probleme bei der operativen Prognose gehört (Januschowski & Kolassa, 2019). Dieses Dokument folgt den wichtigsten Anliegen: Daten, Modelle (Predictors), Inferenzen (Prognosen) und Produktion.

Bei dieser Fallstudie ist zu beachten, dass das Prognoseproblem ein Mittel zum Zweck ist. Obwohl Prognosen für das Unternehmen ausschlaggebend sind, spielen nachgelagerte Entscheidungen für die Lieferkette eine noch wichtigere Rolle. In unserer Fallstudie werden diese Entscheidungen von automatisierten Einkaufssystemen getroffen, die auf mathematischen Optimierungsmodellen aus dem Bereich der Betriebsforschung basieren. Diese Systeme versuchen, die voraussichtlichen Kosten für das Unternehmen zu minimieren.

Das Schlüsselwort lautet voraussichtlich, was bedeutet, dass die Prognosen nicht nur eine mögliche Zukunft, sondern alle möglichen Zukünfte abdecken sollten, wobei die entsprechende Gewichtung

gemäß der Wahrscheinlichkeit eines bestimmten Ergebnisses erfolgt. Zu diesem Zweck, und um eine nachgelagerte Entscheidungsfindung zu unterstützen, bedarf es nicht lediglich einer punktuellen Prognose sondern vielmehr einer kompletten Verteilung der Prognosewerte. Die folgende Abbildung zeigt eine probabilistische Prognose (auch als Dichteprognose bezeichnet). Beachten Sie, dass Sie aus dieser probabilistischen Prognose leicht eine einzelne Punktprognose (die wahrscheinlichste Zukunft) ableiten können. Es ist jedoch schwieriger, von einer Punktprognose zu einer probabilistischen Prognose zu wechseln.

Eine probabilistische Prognose liefert Ihnen verschiedene Statistiken und Sie können die Ergebnisse so anpassen, dass Sie sie bei Ihrer Entscheidungsfindung unterstützen. Das E-Commerce-Geschäft bietet möglicherweise zahlreiche wichtige Produkte an, die in seinem Lagerbestand niemals fehlen sollten. Verwenden Sie in diesem Fall ein hohes Quantil (z. B. die 90. Perzentile). Dies bedeutet, dass die Produkte zu 90 % der Zeit auf Lager sein werden. Für andere Produkte, z. B. Produkte, die sich leichter ersetzen lassen (z. B. Stifte), ist es wahrscheinlich besser, eine niedrigere Perzentile zu verwenden.

In Amazon Forecast können Sie aus der probabilistischen Prognose ganz einfach verschiedene Quantile abrufen.

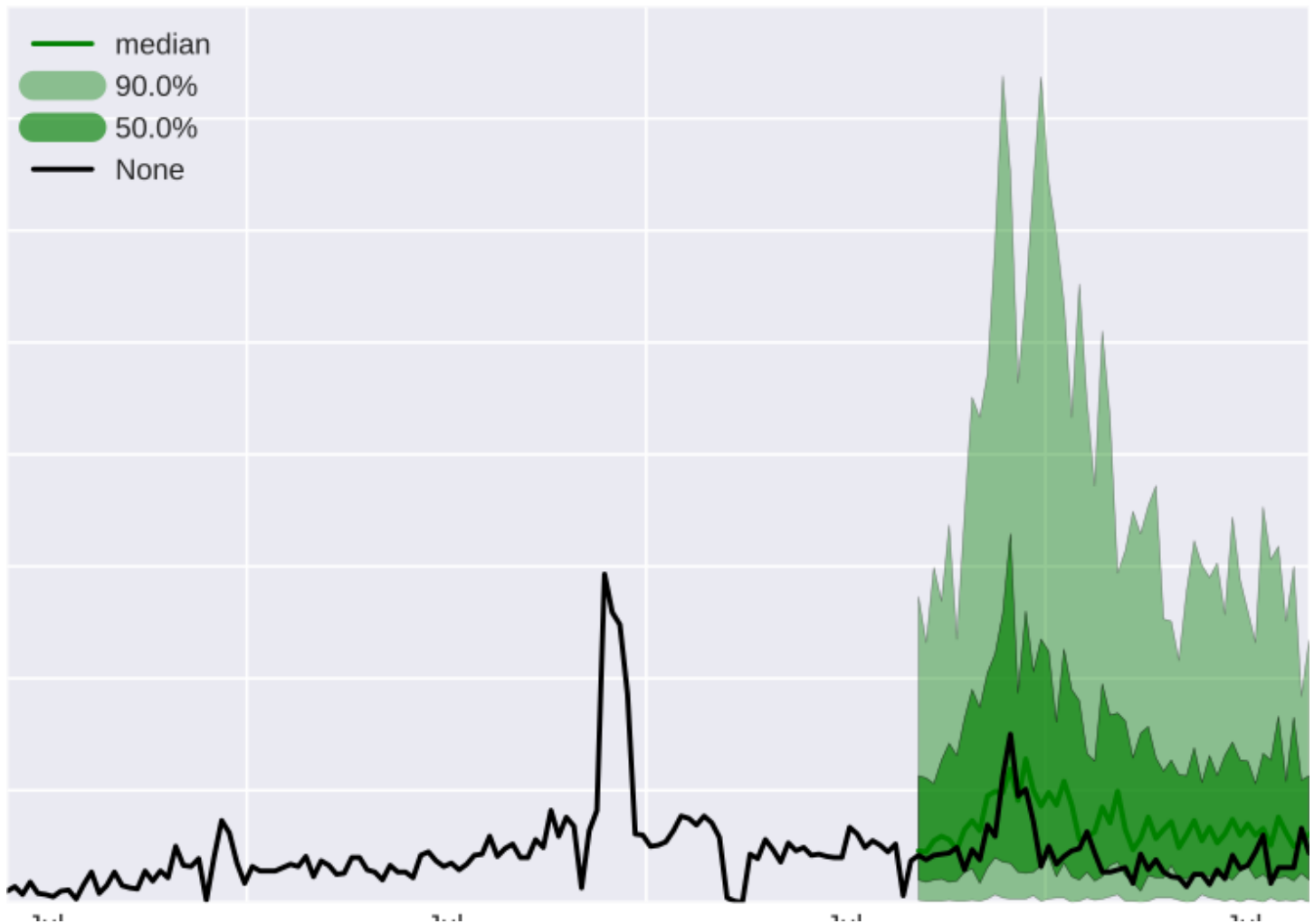


Abbildung: Probabilistische Prognose

In der vorhergehenden Abbildung entspricht die schwarze Linie den tatsächlichen Werten, die dunkelgrüne Linie ist der Median der Prognoseverteilung, der dunkelgrün schattierte Bereich stellt das Prognoseintervall dar, für das Sie erwarten, dass 50 % der Werte sinken werden und der hellgrüne Bereich stellt das Prognoseintervall dar, für das Sie erwarten, dass 90 % der tatsächlichen Werte sinken werden.

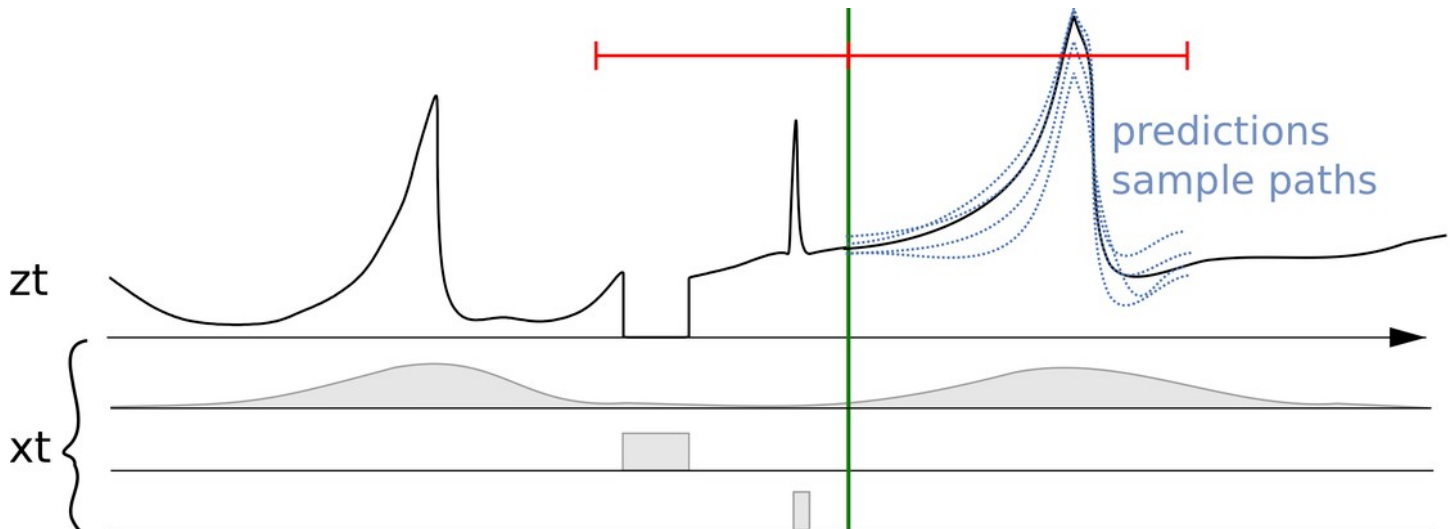
In den folgenden Abschnitten werden die Schritte behandelt, die zur Lösung des Prognoseproblems dieses Unternehmens beitragen, darunter:

- [Datenerfassung und Aggregation \(Schritt 1\)](#)
- [Datenaufbereitung \(Schritt 2\)](#)
- [Erstellen eines Predictors \(Schritt 3\)](#)
- [Bewertung von Predictors \(Schritt 4\)](#)

- [Automatisieren der Prognosegenerierung \(Schritt 5\)](#)

Schritt 1: Daten sammeln und aggregieren

Die folgende Abbildung zeigt ein mentales Modell für das Prognoseproblem. Ziel ist es, die Zeitreihe z_t in die Zukunft zu prognostizieren und dabei so viele relevante Informationen wie möglich zu verwenden, um eine möglichst präzise Prognose zu liefern. Deshalb besteht der erste und wichtigste Schritt darin, so viele richtige Daten wie möglich zu sammeln.



Eine Zeitreihe z_t mit dazugehörigen Merkmalen oder Kovariaten (x_t) und mehreren Prognosen

In der vorherigen Abbildung sind rechts neben der vertikalen Linie mehrere Prognosen dargestellt. Diese Prognosen sind Stichproben aus der probabilistischen Prognoseverteilung (oder können im Gegenteil zur Darstellung der probabilistischen Prognose verwendet werden).

Die wichtigsten Daten, die ein Einzelhandelsunternehmen erfassen muss, sind:

- Transaktionsdaten: Zum Beispiel Stock Keeping Unit (SKU), Standort, Zeitstempel und verkaufte Einheiten.
- Detaillierte SKU-Daten: Metadaten eines Artikels. Z. B. Farbe, Abteilung, Größe usw.
- Preisdaten: Die Preiszeitreihe jedes Artikels mit Zeitstempeln.
- Daten zu Werbeaktionen: Verschiedene Arten von Werbeaktionen, entweder für eine Sammlung an Artikeln (Kategorie) oder für einzelne Artikel mit Zeitstempeln.
- Lagerinformationen: Information, die für jede Zeiteinheit angibt, ob eine SKU vorrätig war oder noch gekauft werden konnte oder im Gegenteil nicht mehr vorrätig war.
- Standortdaten: Der Standort eines Artikels oder Verkaufs zu einem bestimmten Zeitpunkt kann als Zeichenfolge `location_idstore_id` oder als tatsächliche Geolokalisierung dargestellt

werden. Für Geolokalisierungen kann eine Ländervorwahlnummer und eine fünfstellige Postleitzahl verwendet werden oder die `latitude_longitude`-Koordinate. Der Standort drückt die „Dimension“ der Transaktion aus.

In [Amazon Forecast](#) werden die historischen Daten der zu prognostizierenden Menge als Zielzeitreihe (Target Time Series, TTS) bezeichnet. Im Einzelhandelsgeschäft entsprechen Transaktionsdaten der TTS. Andere historische Daten zu Transaktionen, die zeitgleich mit den anderen Transaktionen stattgefunden haben, werden als verwandte Zeitreihen (Related Time Series, RTS) bezeichnet. Für das Einzelhandelsgeschäft würde die verwandte Zeitreihe die Variablen Preis, Werbeaktion und Lagerbestand beinhalten.

Hinweis: Lagerbestandsdaten sind wichtig, da dieses Problem die Nachfrage-Prognose betrifft und nicht die Umsatz-Prognose. Das Unternehmen speichert aber lediglich die Verkaufsdaten. Wenn eine SKU nicht mehr vorrätig ist, ist die Anzahl der Verkäufe geringer als die potenzielle Nachfrage. Daher ist es wichtig zu wissen und festzuhalten, wann solche Fälle auftreten.

Weitere Datensätze, die berücksichtigt werden sollten, sind die Anzahl der Webseitenbesuche, Informationen zu Suchbegriffen, soziale Medien und Wetterinformationen. Oft ist es wichtig, über Vergangenheitsdaten und Zukunftsdaten zu verfügen, um sie in Modellen verwenden zu können. Dies ist eine Voraussetzung für viele Prognosemodelle und für das Backtesting (siehe Abschnitt [Schritt 4: Prognoseauswertung](#)).

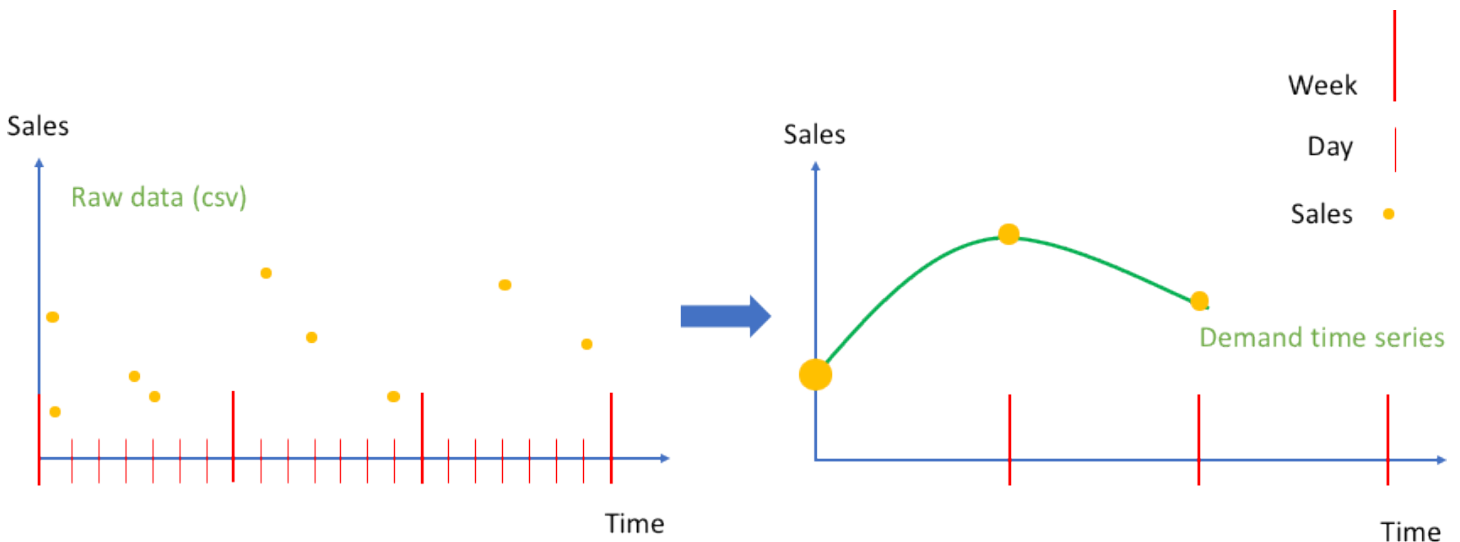
Bei einigen Prognoseproblemen stimmt die Häufigkeit der Rohdaten perfekt mit der Häufigkeit des Prognoseproblems überein. Ein Beispiel hierfür ist die Anforderung an das Servervolumen, das minutenweise abgetastet wird, wenn Sie eine Prognose im Minutentakt erstellen möchten.

Oft werden für die Aufzeichnung von Daten innerhalb eines Zeitabschnitts eine feinere Frequenz oder schlichtweg zufällige Zeitstempel verwendet, für das Prognoseproblem wird jedoch eine gröbere Granularität festgelegt. Dies kommt häufig in Fallstudien aus dem Bereich Einzelhandel vor. Hier werden die Verkaufsdaten normalerweise als Transaktionsdaten aufgezeichnet. Wenn z. B. Verkäufe getätigt werden, besteht das Format aus einem Zeitstempel mit einer feinen Granularität. Im Anwendungsfall für Prognosen ist eine geringe Granularität wahrscheinlich nicht erforderlich und es ist wahrscheinlich sinnvoller, diese Daten im Stunden- oder Tagesformat zu aggregieren. Hier entspricht die Aggregationsebene dem nachgelagerten Problem, z. B. der Bestandsverwaltung oder der Ressourcenplanung.

Beispiel

In der folgenden Abbildung zeigt das linke Diagramm ein Beispiel für die Rohdaten der Kundenverkäufe, die in Amazon Forecast als CSV-Datei (durch Kommas getrennte Werte) eingegeben werden können. In diesem Beispiel wurde für die Verkaufsdaten ein feineres Tagesraster eingestellt und das Problem besteht darin, die zukünftige wöchentliche Nachfrage auf Basis eines gröberen Zeitrasters vorherzusagen. Amazon Forecast aggregiert die Tageswerte in einer bestimmten Woche im `imcreate_predictor` API-Aufruf.

Das Ergebnis wandelt die Rohdaten in eine Sammlung an wohlgeformten Zeitreihen um, für die eine wöchentliche Frequenz eingestellt wurde. Das rechte Diagramm veranschaulicht diese Aggregation für die Zielzeitreihe, wobei für die Aggregation als Summe die Standardmethode verwendet wird. Andere Aggregationsmethoden unterstützen Berechnungen von Mittelwert, Maximum, Minimum oder die Auswahl eines einzelnen (z. B. des ersten) Punktes. Die Aggregationsgranularität und die Methode müssen so gewählt werden, dass sie genau zu dem Anwendungsfall der Daten passen. In diesem Beispiel entspricht der aggregierte Wert der wöchentlichen Aggregation. Benutzer können andere Aggregationsmethoden mithilfe des `FeaturizationMethodParameters`-Schlüssels des `FeaturizationConfig`-Parameters des `create_predictor`-API einstellen.



Aggregation von Rohverkaufsdaten zu Ereignissen (links), eine Zeitreihe mit gleichen Intervallen (rechts)

Schritt 2: Daten vorbereiten

Sobald Ihnen Rohdaten zur Verfügung stehen, müssen Sie sich mit Komplikationen wie fehlenden Daten befassen und sicherstellen, dass Sie für die Prognosemodelle Daten vorbereiten, die die beabsichtigte Interpretation am besten erfassen.

Umgang mit fehlenden Daten

Bei realen Prognoseproblemen kommt es häufig vor, dass in den Rohdaten Werte fehlen. Fehlt in einer Zeitreihe ein Wert, so bedeutet das, dass der wahre entsprechende Wert zu jedem Zeitpunkt mit der angegebenen Häufigkeit für die weitere Verarbeitung nicht verfügbar ist. Werte können aus mehreren Gründen als fehlend markiert werden.

Es kann vorkommen, dass Werte fehlen, weil es keine Transaktionen gab, oder weil bei der Messung Fehler gemacht wurden (z. B. weil ein Dienst, der bestimmte Daten überwacht hat, nicht richtig funktioniert hat oder weil die Messung nicht richtig durchgeführt werden konnte). Das Hauptbeispiel für einen solchen Fall in der Einzelhandelsfallstudie ist ein Fehlbestand in der Nachfrageprognose, denn das bedeutet, dass die Nachfrage an diesem Tag nicht dem Umsatz entspricht.

Ähnliches kann in Cloud-Computing-Szenarien vorkommen, wenn ein Dienst ein Limit erreicht hat (z. B. alle [Amazon EC2](#)-Instanzen sind in einer bestimmten [AWS-Region](#) ausgelastet). Als weiteres Beispiel für fehlende Werte kann ein Produkt oder eine Dienstleistung erwähnt werden, die noch nicht auf den Markt gebracht wurde oder deren Produktion eingestellt wird.

Fehlende Werte können auch mithilfe von Komponenten zur Bearbeitung von Features eingefügt werden. Auf diese Weise kann sichergestellt werden, dass die Zeitreihe genauso lang ist wie das Padding. Bei ausreichender Prävalenz können fehlende Werte die Genauigkeit eines Modells signifikant beeinflussen.

Beispiel 1

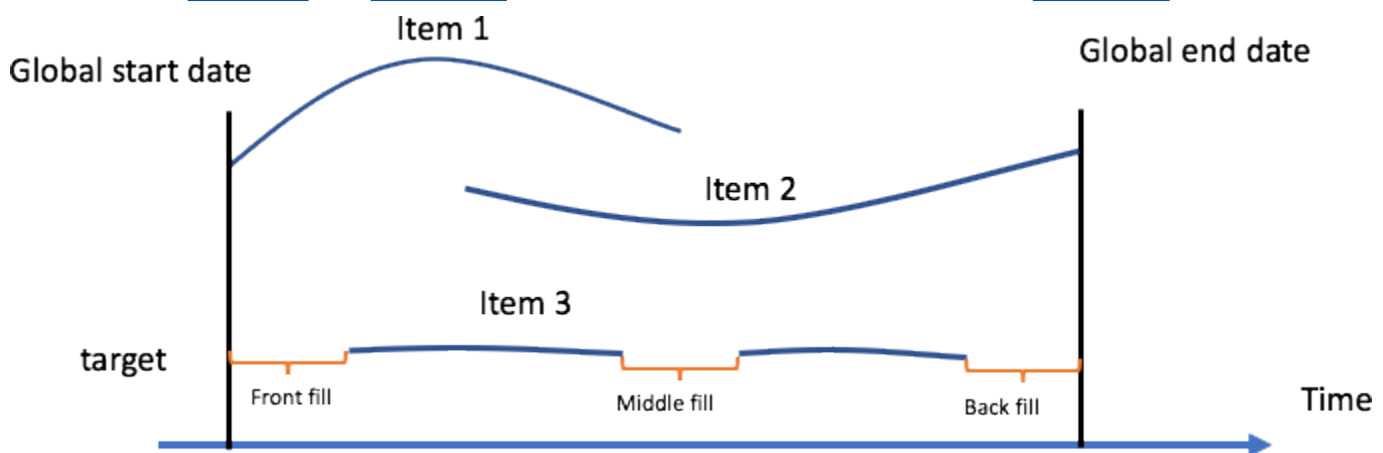
Füllen ist der Prozess des Hinzufügens standardisierter Werte zu fehlenden Einträgen in Ihrem Dataset. Die folgende Abbildung veranschaulicht eine Reihe von Methoden, die Amazon Forecast zum Behandeln fehlender Werte bietet: Vorfüllen, mittlere Füllung, Nachfüllen und zukünftige Füllung – für Element 2 in einem Dataset mit drei Elementen.

Amazon Forecast unterstützt Füllmethoden für Zielzeitreihen und verwandte Zeitreihen. Das globale Startdatum ist das früheste Startdatum der Startdaten aller in Ihrem Dataset enthaltenen Elemente.

Im folgenden Beispiel wird das globale Startdatum für Element 1 angezeigt. Ähnlich wird das globale Enddatum als das späteste Enddatum der Zeitreihe für alle Elemente festgelegt. Das gilt für Element 2.

Beim Vorfüllen werden alle Werte vom Beginn der jeweiligen Zeitreihe bis zum globalen Startdatum ausgefüllt. Zum Zeitpunkt der Veröffentlichung dieses Dokuments aktiviert Amazon Forecast kein Vorfüllen und erlaubt, dass alle Zeitreihen zu unterschiedlichen Zeitpunkten beginnen. Mittlere Füllung bezeichnet Werte, die in der Mitte der Zeitreihe ausgefüllt wurden (z. B. zwischen dem Start- und dem Enddatum der Elemente). Beim Nachfüllen werden Werte vom letzten Datum dieser Zeitreihe bis zum globalen Enddatum ausgefüllt.

Für die Zielzeitreihe befolgt die Methode der mittleren Füllung und die Nachfüllmethode eine Standardfülllogik NULL. Die zukünftige Füllung (nur für verwandte Zeitreihen) füllt alle fehlenden Werte zwischen dem globalen Enddatum der Artikel und dem vom Kunden angegebenen Prognosehorizont aus. Zukünftige Werte sind für die Verwendung des verwandten Zeitreihen-Datasets mit [Prophet](#) und [DeePar+](#) erforderlich. Für die Verwendung mit [CNN-QR](#) sind sie optional.



Amazon-Forecast-Methoden für den Umgang mit Werten

In der vorherigen Abbildung bezeichnet das globale Startdatum das früheste Startdatum für die Startdaten aller Elemente, und das globale Enddatum bezeichnet das späteste Enddatum für die Enddaten aller Elemente. Der Prognosehorizont ist der Zeitraum, in dem Forecast Prognosen für den Zielwert bereitstellt.

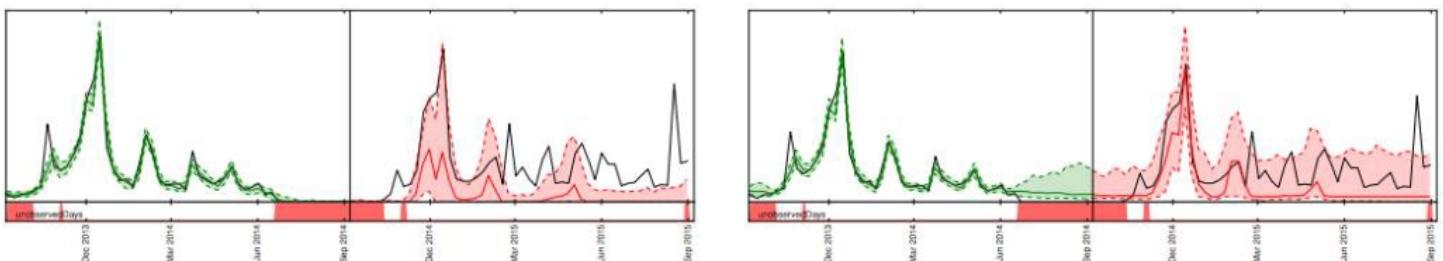
Dies ist ein gängiges Einzelhandelsszenario, das für Transaktionsdaten für verfügbare Artikel einen Null-Umsatz darstellt. Diese Werte werden als echte Nullen behandelt und in der Komponente zur Kennzahlbewertung verwendet. Amazon Forecast ermöglicht es dem Benutzer, tatsächlich fehlende Werte zu identifizieren und sie als NaN (Not a Number) zu codieren, damit die Algorithmen sie

verarbeiten können. In diesem Whitepaper wird als Nächstes untersucht, worin sich diese beiden Fälle unterscheiden und wann sie jeweils nützlich sind.

In der Einzelhandelsfallstudie unterscheidet sich die Information, dass ein Einzelhändler null Einheiten eines verfügbaren Artikels verkauft hat, von der Information, dass null Einheiten eines nicht verfügbaren Artikels verkauft wurden, und zwar entweder in den Zeiträumen, in denen er nicht existiert hat (z. B. vor seiner Markteinführung oder nach seiner Abschaffung), oder in Zeiträumen, in denen er existiert hat (z. B. als er teilweise vergriffen war oder als für diesen Zeitraum keine Verkaufsdaten aufgezeichnet wurden). Im ersten Fall wird die Standardfüllung NULL angewandt. Im zweiten Fall werden für den als fehlend markierten Wert zusätzliche Informationen übermittelt, obwohl der entsprechende Zielwert typischerweise Null ist. Es empfiehlt sich, die Information darüber, dass Daten gefehlt haben, aufzubewahren und diese nicht zu löschen. Im folgenden Beispiel wird veranschaulicht, warum es wichtig ist, Informationen aufzubewahren.

Amazon Forecast unterstützt zusätzliche Fülllogiken für Wert, Mittelwert, Median, Minimum und Maximum. Für die verwandte Zeitreihe (z. B. Preis oder Werbeaktion) wurden für die mittlere Füllung, das Nachfüllen oder die zukünftige Füllung keine Standardfülllogiken angegeben, da die korrekte Logik für fehlende Werte je nach Attributtyp und Anwendungsfall variiert. Die Fülllogik, die für verwandte Zeitreihen unterstützt wird, umfasst Null, Wert, Mittelwert, Median, Minimum und Maximum.

Um fehlende Werte auszufüllen, geben Sie die zu implementierenden Füllarten an, wenn Sie den Vorgang [CreatePredictor](#) aufrufen. Die Fülllogik wird in [FeaturizationMethod](#)-Objekten definiert. Um beispielsweise einen Wert zu codieren, der nicht den Null-Umsatz eines nicht verfügbaren Produkts in der Zielzeitreihe darstellt, markieren Sie einen Wert als wirklich fehlend, indem Sie den Fülltyp auf NaN setzen. Im Gegensatz zur Null-Füllung werden die als NaN kodierten Werte als wirklich fehlend behandelt und in der Komponente zur Kennzahlbewertung nicht verwendet.



Die Auswirkung der Füllung mit 0 im Vergleich zur Füllung mit NaN auf die Prognosen für denselben Artikel

In der vorherigen Abbildung (linke Grafik) sind die Werte links von der vertikalen schwarzen Linie mit Nullen gefüllt, was zu einer Unterprognose führt (rechts von der vertikalen schwarzen Linie). In der rechten Grafik sind diese Werte als NaN gekennzeichnet, was zu einer richtigen Prognose führt.

Beispiel 2

Die vorherige Abbildung veranschaulicht, wie wichtig es ist, fehlende Werte bei einem linearen Zustandsraummodell wie [ARIMA oder ETS](#) korrekt zu behandeln. Es stellt die Nachfrageprognose für einen Artikel dar, der teilweise vergriffen ist. Die Trainingsregion ist in der linken Grafik grün markiert, der Vorhersagebereich ist im rechten Feld rot markiert und das echte Ziel ist schwarz markiert. Die roten Linien bzw. der schattierte Bereich repräsentieren die Median-, p10- und p90-Prognosen. Die vergriffenen Artikel (80 % der Daten) sind unten in Rot dargestellt. Im linken Diagramm wurden die Werte für vergriffene Artikel ignoriert und mit 0 ausgefüllt.

Infolgedessen gehen die Prognosemodelle davon aus, dass viele Nullen vorhergesagt werden müssen und die Prognosen daher zu niedrig sind. Im rechten Diagramm werden die Fehlbestandsbereiche als tatsächlich fehlend gekennzeichnet, weshalb die Nachfrage in der Region, in der die Artikel vergriffen sind, ungewiss wird. Da die fehlenden Werte für die vergriffenen Artikel entsprechend mit NaN ausgefüllt wurden, sehen Sie in diesem Diagramm im Prognosebereich keine Unterprognose. Amazon Forecast füllt diese Datenlücken aus und erleichtert Ihnen den korrekten Umgang mit fehlenden Daten, ohne dass Sie alle Eingabedaten explizit ändern müssen.

Konzepte der Featurisierung und verwandte Zeitreihen

Amazon Forecast ermöglicht es Benutzern, verwandte Daten einzugeben, um die Genauigkeit bestimmter unterstützter Prognosemodelle zu verbessern. Bei diesen Daten kann es sich um zwei Typen von Daten handeln: verwandte Zeitreihen oder statische Elementmetadaten.

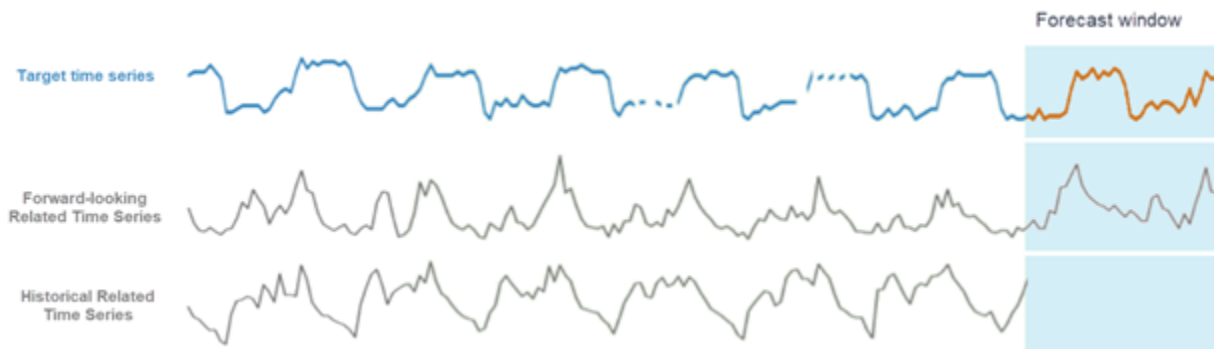
Note

Metadaten und verwandte Daten werden im Bereich des Machine Learning als Features und in der Statistik als Kovariaten bezeichnet.

Verwandte Zeitreihen sind Zeitreihen, die in gewissem Maße mit dem Zielwert korrelieren und die der Prognose des Zielwerts eine gewisse statistische Aussagekraft verleihen sollten, da sie eine intuitive Erklärung bieten (ein Beispiel finden Sie unter [Amazon Forecast: Vorhersage von Zeitreihen](#)).

in großem Maßstab). Verwandte Zeitreihen sind im Gegensatz zu Zielzeitreihen bekannte Werte in der Vergangenheit, die sich auf die Zielzeitreihe auswirken und möglicherweise bekannte Werte in der Zukunft haben.

In Amazon Forecast können Sie zwei Arten verwandter Zeitreihen hinzufügen: historische Zeitreihen und zukunftsgerichtete Zeitreihen. Historisch verwandte Zeitreihen enthalten Datenpunkte bis zum Prognosehorizont und keine zukünftigen Datenpunkte innerhalb des Prognosehorizonts. Zukunftsgerichtete Zeitreihen enthalten Datenpunkte bis zum und innerhalb des Prognosehorizonts.



Verschiedene Ansätze zur Verwendung verwandter Zeitreihen mit Amazon Forecast

Beispiel 3

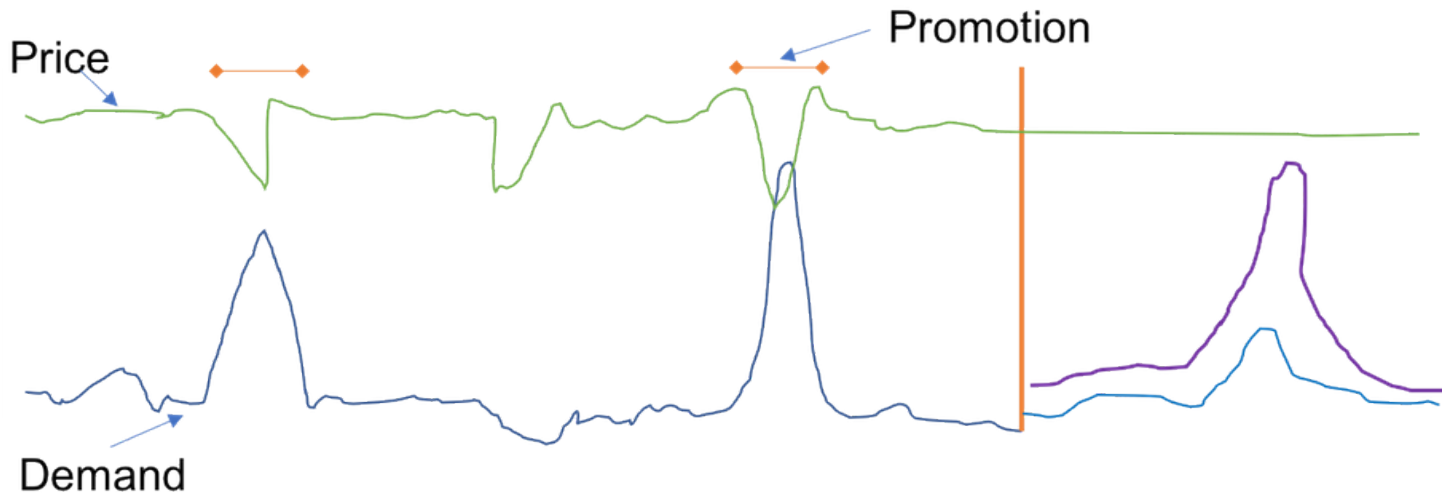
Die folgende Abbildung zeigt ein Beispiel dafür, wie verwandte Zeitreihen verwendet werden können, um die zukünftige Nachfrage nach einem beliebten Buch vorherzusagen. Die blaue Linie stellt den Bedarf in der Zielzeitreihe dar. Der Preis wird als grüne Linie dargestellt. Die vertikale Linie stellt das Startdatum der Prognose dar und die Prognosen für die beiden Quantile werden rechts von der vertikalen Linie angezeigt.

In diesem Beispiel wird eine zukunftsgerichtete Zeitreihe verwendet, die mit der Zielzeitreihe in puncto Prognosegranularität übereinstimmt und zu allen (oder den meisten) Zeitpunkten in der Zukunft bekannt ist, und zwar im Bereich vom Startdatum der Prognose bis zum Startdatum der Prognose, inkrementiert um den Prognosehorizont (Prognoseenddatum).

Die folgende Abbildung zeigt außerdem, dass der Preis ein passendes Feature ist, da Sie Korrelationen zwischen einem Preisrückgang und eine Umsatzsteigerung bei dem Produkt erkennen können. Verwandte Zeitreihen können Amazon Forecast über eine separate CSV-Datei zur Verfügung gestellt werden, die die Artikel-SKU, den Zeitstempel und die verwandten Zeitreihenwerte (in diesem Fall den Preis) enthält.

Amazon Forecast unterstützt Aggregationsmethoden wie Durchschnitt und Summierung für Zielzeitreihen, nicht jedoch für verwandte Zeitreihen. Zum Beispiel macht es wenig Sinn, einen Tagespreis mit einem Wochenpreis zu summieren. Dasselbe gilt für tägliche Werbeaktionen.

Amazon Forecast kann automatisch Informationen zu [Wetter](#) und [Feiertagen](#) in ein Modell integrieren, indem es integrierte Feature-Datasets einbezieht (siehe [SupplementaryFeature](#)). Wetterinformationen und Feiertage können die Nachfrage im Einzelhandel erheblich beeinflussen.



Der Umsatz eines bestimmten Artikels (blau, links neben der vertikalen roten Linie)

Artikelmetadaten, auch als kategoriale Variablen bekannt, sind weitere hilfreiche Features, die in Amazon Forecast eingegeben werden können (ein Beispiel finden Sie unter [Amazon Forecast: Vorhersage von Zeitreihen in großem Maßstab](#)). Der Hauptunterschied zwischen kategorialen Variablen und verwandten Zeitreihen besteht darin, dass kategoriale Variablen statisch sind, d. h. sie ändern sich nicht im Laufe der Zeit. Zu den gängigen Beispielen aus dem Einzelhandels gehören die Artikelfarben, Bücherkategorien und binäre Indikatoren, die angeben, ob es sich bei einem Fernseher um einen Smart-TV handelt oder nicht. Diese Informationen können durch Deep-Learning-Algorithmen erfasst werden, um daraus zu erfahren, welche Ähnlichkeiten SKUs aufweisen. Dabei gilt aber die Voraussetzung, dass ähnliche SKUs ähnliche Umsätze erzielen. Da diese Metadaten keine Zeitabhängigkeit haben, besteht jede Zeile in der CSV-Datei ausschließlich aus der Artikel-SKU und der zugehörigen Kategoriebezeichnung oder Beschreibung.

Schritt 3: Erstellen eines Predictors

Ein Predictor kann auf zwei Arten erstellt werden: durch Ausführen von [AutoML](#) oder durch manuelles Auswählen eines der sechs Algorithmen, die in Amazon Forecast integriert sind. Beim Ausführen von AutoML testet Amazon Forecast automatisch die sechs integrierten Algorithmen und wählt den Algorithmus mit den niedrigsten durchschnittlichen Quantilverlusten beim 10., 50. (Median) und 90. Quantil aus.

Amazon Forecast bietet vier lokale Modelle an:

- Autoregressive Integrated Moving Average ([ARIMA](#))
- Exponential Smoothing ([ETS](#))
- Non-Parametric Time Series ([NPTS](#))
- [Prophet](#)

Lokale Modelle sind Prognosemethoden, bei denen ein einzelnes Modell an jede einzelne Zeitreihe (oder eine bestimmte Kombination aus Element und Dimension) angepasst wird. Anschließend werden diese Modelle dazu verwendet, um die Zeitreihe in die Zukunft zu extrapolieren.

ARIMA und ETS sind skalierbare Versionen beliebter lokaler Modelle aus dem R-Forecast-Paket. NPTS ist eine lokale Methode, die bei Amazon entwickelt wurde und die im Vergleich zu den anderen lokalen Modellen einen wesentlichen Unterschied aufweist. Im Gegensatz zu einfachen Tools für saisonale Prognosen, die Punktprognosen liefern, indem sie den letzten Wert oder einen saisonabhängigen Wert wiederholen, erstellt NPTS probabilistische Prognosen. NPTS verwendet einen festen Zeitindex, wobei der vorherige Index ($T - 1$) oder die vergangene Saison ($T - \text{tau}$) die Vorhersage für den Zeitschritt T darstellt. Der Algorithmus tastet nach dem Zufallsprinzip einen Zeitindex (t) aus der Menge $\{0, \dots, T - 1\}$ ab, um eine Stichprobe für den aktuellen Zeitschritt T zu generieren. NPTS ist besonders nützlich, wenn mit intermittierenden (manchmal auch spärlichen) Zeitreihen mit vielen Nullen gearbeitet wird. Forecast schließt auch die Python-Implementierung von Prophet ein, ein Bayes'sches strukturelles Zeitreihenmodell.

Amazon Forecast bietet zwei globale Deep-Learning-Algorithmen:

- [DeepAR+](#)
- [CNN-QR](#)

Globale Modelle trainieren ein einzelnes Modell über alle Zeitreihen in einem Datensatz hinweg. Dies ist insbesondere dann nützlich, wenn ein Satz an Querschnittseinheiten ähnliche Zeitreihen enthält. Zum Beispiel Nachfrage-Zeitreihengruppen für unterschiedliche Produkte, Serverlasten und Anforderungen für Webseiten.

Im Allgemeinen steigt bei einer zunehmenden Anzahl der Zeitreihen auch die Effizienz von CNN-QR und DeePAR+. Bei lokalen Modellen ist dies nicht immer der Fall. Deep-Learning-Modelle können auch zum Erstellen von Prognosen für neue SKUs mit wenigen bis gar keinen historischen Verkaufsdaten verwendet werden. Dies wird als [Cold-Start-Prognose](#) bezeichnet.

	Neural Networks		Flexible Local Algorithms	Baseline Algorithms		
	CNN-QR	DeePAR+	Prophet	NPTS	ARIMA	ETS
Computationally intensive training process	High	High	Medium	Low	Low	Low
Accepts historical related time series*	✓	✗	✗	✗	✗	✗
Accepts forward-looking related time series*	✓	✓	✓	✗	✗	✗
Accepts item metadata (product color, brand, etc)	✓	✓	✗	✗	✗	✗
Suitable for sparse datasets	✓	✓	✗	✓	✗	✗
Performs Hyperparameter Optimization (HPO)	✓	✓	✗	✗	✗	✗
Allows overriding default hyperparameter values	✓	✓	✗	✓	✗	✗
Suitable for What-if analysis	✓	✓	✓	✗	✗	✗
Suitable for Cold Start scenarios (forecasting with little to no historical data)	✓	✓	✗	✗	✗	✗

Vergleichen Sie die in Amazon Forecast verfügbaren Algorithmen

Weitere Informationen zu verwandten Zeitreihen finden Sie unter [Verwandte Zeitreihen](#).

Schritt 4: Predicators auswerten

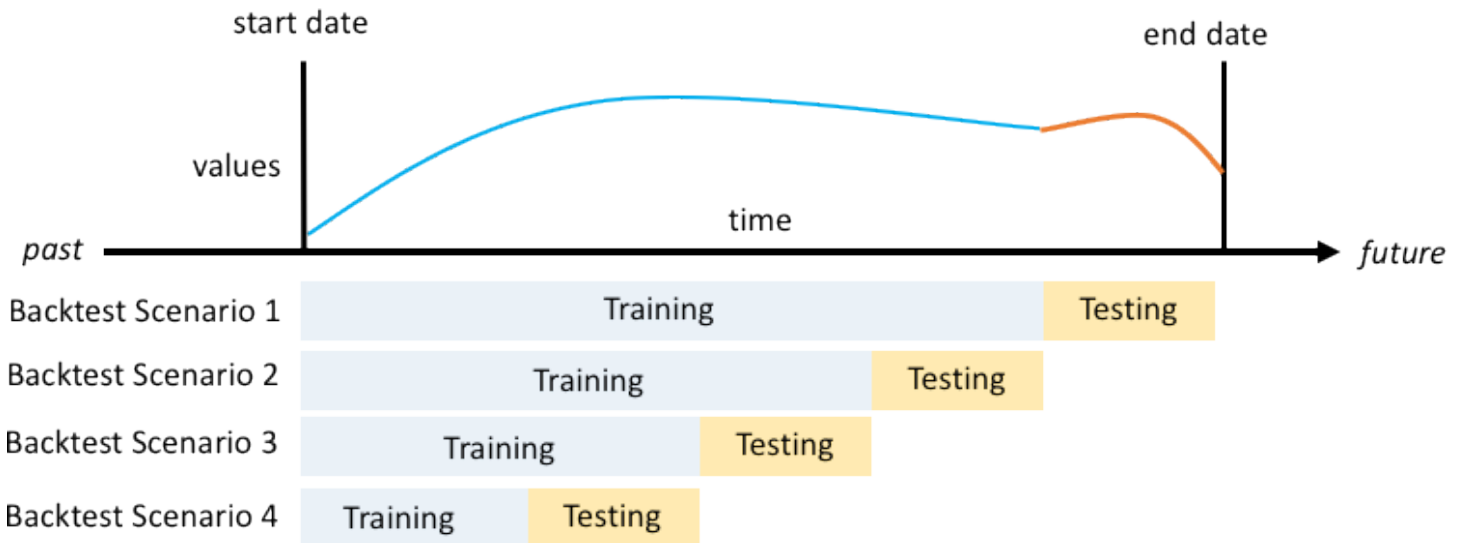
Ein typischer Machine-Learning-Workflow besteht darin, eine Reihe von Modellen oder eine Kombination von Modellen mithilfe eines Trainingssatzes zu trainieren und deren Genauigkeit anhand eines Hold-Out-Datensatzes zu bewerten. In diesem Abschnitt wird erläutert, wie historische Daten aufgeteilt und welche Metriken in der Zeitreihenprognose verwendet werden, um Modelle zu bewerten. Die Technik des Backtesting dient als Hauptinstrument zur Bewertung der Prognosegenauigkeit.

Backtesting

Ein adäquates Evaluierungs- und Backtesting-Framework stellt eines der zentralen Elemente bei der Erstellung einer erfolgreichen Machine-Learning-Anwendung dar. Ein erfolgreiches Backtesting Ihrer Modelle ist die Grundlage dafür, um Vertrauen in die zukünftige Vorhersagekraft der Modelle zu gewinnen. Darüber hinaus können Sie die Hyperparameteroptimierung (HPO) nutzen, um Modelle zu optimieren, zu lernen, wie Modelle miteinander kombiniert werden, und um Meta-Learning und AutoML zu aktivieren.

Im Vergleich zu anderen Bereichen des angewandten Machine Learning stellt das Merkmal der Zeitreihenprognose Zeit ein wesentliches Unterscheidungsmerkmal in Bezug auf die Bewertungs- und Backtesting-Methode dar. Um einen Prognosefehler in einem Backtest zu evaluieren, teilen Sie bei ML-Aufgaben in der Regel einen Datensatz nach Elementen auf. Für die Kreuzvalidierung im Rahmen bildrelevanter Aufgaben trainieren Sie beispielsweise anhand eines bestimmten Prozentsatzes der Bilder. Anschließend verwenden Sie zum Testen und Validieren andere Teile. Bei Prognosen müssen Sie in erster Linie nach dem Faktor Zeit (und in geringerem Maße nach Elementen) aufteilen, um sicherzustellen, dass in den Test- oder Validierungssatz keine Informationen aus dem Trainingssatz gelangen und dass Sie den Produktionsfall so genau wie möglich simulieren.

Die Aufteilung nach Zeit muss sorgfältig erfolgen, da Sie nicht einen einzelnen Zeitpunkt sondern mehrere Zeitpunkte auswählen wollen. Andernfalls hängt die Genauigkeit zu stark vom Prognosestartdatum ab, das durch den Trennpunkt definiert wird. Eine fortlaufende Prognoseauswertung, bei der Sie mehrere Aufteilungen über mehrere Zeitpunkte hinweg durchführen und das Durchschnittsergebnis ausgeben, liefert robustere und zuverlässigere Backtest-Ergebnisse. Die folgende Abbildung zeigt vier verschiedene Backtest-Aufteilungen.



Abbildung, die vier verschiedene Backtesting-Szenarien mit zunehmender Trainingsatzgröße, jedoch konstanter Testgröße darstellt

In der vorherigen Abbildung stehen für alle Backtesting-Szenarien alle Daten zur Verfügung. Das ermöglicht das Abgleichen der prognostizierten Werte mit den tatsächlichen Werten.

Da die meisten Zeitreihen in der realen Welt normalerweise nicht stationär sind, bedarf es mehrerer Backtest-Fenster. Das E-Commerce-Geschäft aus der Fallstudie hat seinen Sitz in Nordamerika. Es verzeichnet die höchste Produktnachfrage im vierten Quartal, mit besonderen Bedarfsspitzen um Thanksgiving herum und vor Weihnachten. In der Einkaufssaison im vierten Quartal ist die Variabilität der Zeitreihen höher als während des restlichen Jahres. Mithilfe mehrerer Backtest-Fenster können Sie Prognosemodelle in einer ausgewogeneren Umgebung evaluieren.

Die folgende Abbildung zeigt für jedes Backtest-Szenario die grundlegenden Elemente der Terminologie von Amazon Forecast. Amazon Forecast teilt Daten automatisch in Trainings- und Test-Datasets auf. Amazon Forecast entscheidet, wie die Eingabedaten aufgeteilt werden. Hierzu verwendet das Tool den `BackTestWindowOffset`-Parameter, der als Parameter in der `create_predictor`-API angegeben ist oder seinen `ForecastHorizon`-Standardwert.

In der folgenden Abbildung sehen Sie den ersten allgemeineren Fall, in dem die `ForecastHorizon`-Parameter und die `BackTestWindowOffset`-Parameter nicht identisch sind.

Der `BackTestWindowOffset`-Parameter definiert das Startdatum einer virtuellen Prognose, das in der folgenden Abbildung als gestrichelte vertikale Linie dargestellt wird. Es kann dazu verwendet werden, um die folgende hypothetische Frage zu beantworten: Wenn das Modell an diesem Tag eingesetzt würde, wie würde dann die Prognose aussehen? Der `ForecastHorizon` definiert die Anzahl der vorherzusagenden Zeitschritte vom Startdatum der virtuellen Prognose.

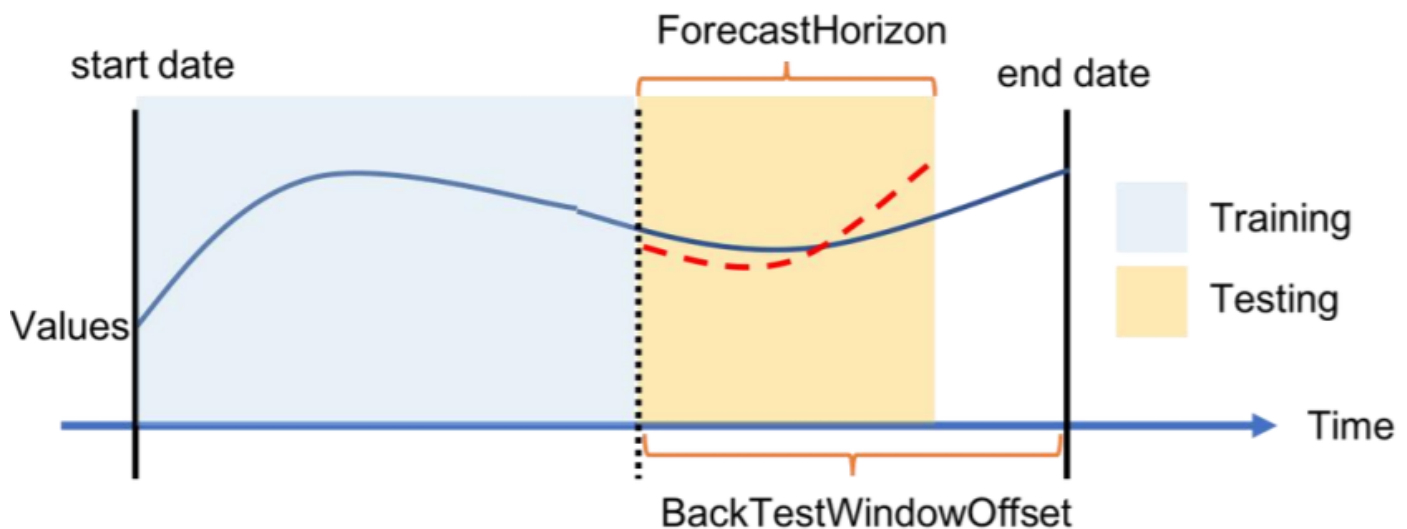


Abbildung eines einzelnen Backtest-Szenarios und seiner Konfiguration in Amazon Forecast

Amazon Forecast kann die prognostizierten Werte und Genauigkeitsmetriken, die beim Backtesting generiert wurden, exportieren. Die exportierten Daten können zur Auswertung bestimmter Elemente zu bestimmten Zeitpunkten und Quantilen verwendet werden.

Vorhersagequantile und Genauigkeitsmetriken

Prognosequantile können eine Ober- und Untergrenze für Prognosen angeben. Wenn Sie beispielsweise die Prognosetypen 0,1 (P10), 0,5 (P50) und 0,9 (P90) verwenden, entsteht um die P50-Prognose herum ein Wertebereich, der als 80-%-Konfidenzintervall bezeichnet wird. Wenn Sie Vorhersagen für P10, P50 und P90 generieren, können Sie davon ausgehen, dass der wahre Wert in 80 % der Fälle zwischen diesen Grenzen liegt.

Quantile werden in diesem Whitepaper unter [Schritt 5](#) näher erörtert.

Amazon Forecast verwendet die Genauigkeitsmetriken Weighted Quantile Loss (wQL), Root Mean Square Error (RMSE) und Weighted Absolute Percentage Error (WAPE), um beim Backtesting Predictors zu bewerten.

Gewichteter Quantilverlust (Weighted Quantile Loss, wQL)

Die Fehlermetrik Gewichteter Quantilverlust (wQL) misst die Genauigkeit der Prognose eines Modells bei einem bestimmten Quantil. Dies ist besonders nützlich, wenn für Unter- und Überprognosen unterschiedliche Kosten anfallen. Indem Sie die Gewichtung (τ) der wQL-Funktion festlegen, können Sie automatisch unterschiedliche Strafen für Zu- und Unterprognosen einbauen.

$$wQL[\tau] = 2 \frac{\sum_{i,t} [\tau \max(y_{i,t} - q_{i,t}^{(\tau)}, 0) + (1 - \tau) \max(q_{i,t}^{(\tau)} - y_{i,t}, 0)]}{\sum_{i,t} |y_{i,t}|}$$

wQL-Funktion

Wobei gilt:

- τ – ein Quantil in der Menge $\{0,01, 0,02, \dots, 0,99\}$
- $q_{i,t}(\tau)$ – das vom Modell vorhergesagte τ -Quantil.
- $y_{i,t}$ – der beobachtete Wert am Punkt (i, t)

Gewichteter absoluter prozentualer Fehler (Weighted Absolute Percentage Error, WAPE)

Der gewichtete absolute Prozentfehler (WAPE) ist eine häufig verwendete Metrik zur Messung der Modellgenauigkeit. Es misst die Gesamtabweichung der prognostizierten Werte von den beobachteten Werten.

$$WAPE = \frac{\sum_{i,t} |y_{i,t} - \hat{y}_{i,t}|}{\sum_{i,t} |y_{i,t}|}$$

WAPE

Wobei gilt:

- $y_{i,t}$ – der beobachtete Wert am Punkt (i, t)
- $\hat{y}_{i,t}$ – der vorhergesagte Wert am Punkt (i, t)

Die Prognose verwendet den Prognosemittelwert als prognostizierten Wert, $\hat{y}_{i,t}$.

Quadratischer Mittelwertfehler (Root Mean Square Error, RMSE)

$$RMSE = \sqrt{\frac{1}{nT} \sum_{i,t} (\hat{y}_{i,t} - y_{i,t})^2}$$

Der quadratische Mittelwertfehler (RMSE) ist eine häufig verwendete Metrik zur Messung der Modellgenauigkeit. Ebenso wie der WAPE misst der RMSE die Gesamtabweichung der Schätzungen von den beobachteten Werten.

Wobei gilt:

- $y_{i,t}$ – der beobachtete Wert am Punkt (i, t)
- $\hat{y}_{i,t}$ – der vorhergesagte Wert am Punkt (i, t)
- nT – Die Anzahl der Datenpunkte in einem Testsatz

Die Prognose verwendet den Prognosemittelwert als prognostizierten Wert, $\hat{y}_{i,t}$. Bei der Berechnung von Predictor-Metriken ist nT die Anzahl der Datenpunkte in einem Backtest-Fenster.

Probleme mit WAPE und RMSE

Generell sollten die Punktprognosen, die intern oder mit anderen Prognose-Tools generiert werden können, mit den p50-Quantil- oder Mittelwertprognosen übereinstimmen. Sowohl für WAPE als auch für RMSE verwendet Amazon Forecast die Mittelwertprognose, um den prognostizierten Wert ($\hat{y}$) darzustellen.

Für die Gleichung $\tau = 0,5$ in der wQL [τ] sind beide Gewichtungen gleich, und der wQL [0,5] reduziert sich auf den allgemein verwendeten gewichteten absoluten Prozentfehler (WAPE) für Punktprognosen:

$$wQL[0.5] = 2 \frac{\sum_{i,t} 0.5 [\max(y_{i,t} - q_{i,t}^{(0.5)}, 0) + \max(q_{i,t}^{(0.5)} - y_{i,t}, 0)]}{\sum_{i,t} |y_{i,t}|} = \frac{\sum_{i,t} |y_{i,t} - q_{i,t}^{(0.5)}|}{\sum_{i,t} |y_{i,t}|}$$

Wobei $\hat{y} = q(0,5)$ die Berechnungsprognose ist. In der wQL-Formel wird ein Skalierungsfaktor von 2 verwendet, um den Faktor 0,5 zu annullieren und den exakten WAPE [Median]-Ausdruck zu erhalten.

Beachten Sie, dass sich die obige Definition von WAPE von einer gängigen Interpretation des mittleren absoluten prozentualen Fehlers [MAPE](#) (Mean Absolute Percentage Error) unterscheidet. Der Unterschied liegt im Nenner. Die obige Definition von WAPE umgeht das Problem der Division durch 0. Dieses Problem tritt häufig in realen Szenarien wie der E-Commerce-Fallstudie auf, in der das Geschäft an einem bestimmten Tag oft 0 Einheiten einer bestimmten SKU verkauft.

Anders als bei der Metrik Gewichteter Quantilverlust für τ , die nicht gleich 0,5 ist, kann die inhärente Verzerrung jedes Quantils nicht durch eine Berechnung wie WAPE erfasst werden, bei der die Gewichtungen gleich sind. Die weiteren Nachteile des WAPE sind, dass er nicht symmetrisch ist, eine zu hohe prozentuale Fehlerquote für kleine Zahlen aufweist und lediglich eine punktuelle Metrik darstellt.

Der RMSE ist das Quadrat des Fehlerbegriffs im WAPE und eine häufige Fehlermetrik in anderen ML-Anwendungen. Die RMSE-Metrik zieht ein Modell vor, bei dem die einzelnen Fehler eine konsistente Größenordnung aufweisen, da große Fehlervariationen den RMSE überproportional ansteigen lassen. Aufgrund des quadratischen Fehlers können einige schlecht vorhergesagte Werte den RMSE in einer ansonsten guten Prognose erhöhen. Außerdem haben Terme mit kleineren Fehlern aufgrund der quadrierten Terme im RMSE ein geringeres Gewicht als im WAPE.

Genauigkeitsmetriken ermöglichen eine quantitative Prognosebewertung. Insbesondere für groß angelegte Vergleiche („Ist Methode A insgesamt besser als Methode B?“) sind diese entscheidend. Es ist jedoch oft wichtig, zu solchen Vergleichen Bildmaterial für die einzelnen SKUs hinzuzufügen.

Schritt 5: Generieren und Verwenden von Prognosen für die Entscheidungsfindung

Sobald Sie über ein Modell verfügen, das den für Ihren spezifischen Anwendungsfall erforderlichen Genauigkeitsschwellenwert (der durch Backtesting ermittelt wurde) erreicht hat, besteht der letzte Schritt darin, das Modell bereitzustellen und Prognosen zu erstellen. Um ein Modell in Amazon Forecast bereitzustellen, müssen Sie die `Create_Forecast`-API ausführen. Diese Aktion hostet ein Modell, das durch Trainieren mit dem gesamten historischen Datensatz erstellt wurde (im Gegensatz zur Aktion `Create_Predictor`, bei der die Daten in einen Trainings- und einen Testsatz aufgeteilt werden). Die über den Prognosehorizont hinweg generierten Modellprognosen können dann auf zwei Weisen verwendet werden:

- Sie können die Prognosen für einen bestimmten Artikel abfragen (indem Sie den Artikel oder die Kombination Artikel/Dimension angeben). Verwenden Sie hierzu die `Query_Forecast`-API in der [AWS CLI](#) oder greifen Sie darauf direkt über die [AWS-Managementkonsole](#) zu.
- Mithilfe der `Create_Forecast_Export_Job`-API können Sie über alle Quantile hinweg Prognosen für alle Kombinationen von Elementen und Dimensionen generieren. Diese API generiert eine CSV-Datei, die sicher an einem [Amazon Simple Storage Service](#) (Amazon S3)-Standort Ihrer Wahl gespeichert ist. Daraufhin können Sie die Daten aus der CSV-Datei verwenden und in Ihre nachgelagerten Systeme integrieren, die für die Entscheidungsfindung verwendet werden. Ihre vorhandenen Lieferkettensysteme können die Ergebnisse von Amazon Forecast beispielsweise direkt aufnehmen, um fundierte Entscheidungen zur Herstellung bestimmter SKUs zu treffen.

Probabilistische Prognosen

Amazon Forecast kann Prognosen mit unterschiedlichen Quantilen erstellen. Dies ist besonders nützlich, wenn sich die Kosten einer Unterprognose von den Kosten einer Überprognose unterscheiden. Ähnlich wie beim Training von Predictors können probabilistische Prognosen für Quantile im Bereich von p1 bis p99 generiert werden.

Standardmäßig generiert Amazon Forecast Prognosen mit den Quantilen, die auch beim Training von Predictors verwendet wurden. Wenn beim Predictor-Training keine Quantile angegeben werden, werden standardmäßig p10-, p50- und p90-Prognosen generiert.

Für die p10-Prognose wird erwartet, dass der wahre Wert in 10 % der Fälle unter dem prognostizierten Wert liegt. Zur Beurteilung der Genauigkeit kann die wQL [0,1]-Metrik verwendet werden. Das bedeutet, dass die P10-Prognose in 90 % der Fälle unter dem prognostizierten Wert liegt. Würde sie für den Lagerbestand verwendet, wäre der Artikel in 90 % der Fälle ausverkauft. Die P10-Prognose könnte bei mangelndem Lagerplatz nützlich sein oder wenn die Kosten des investierten Kapitals zu hoch sind.

Note

Die formale Definition einer Quantilprognose lautet $\Pr(\text{Istwert} \leq \text{Prognose bei Quantil } q) = q$. Aus technischer Sicht ist ein Quantil ein Perzentil/100. Statistiker neigen dazu, „P90-Quantilniveau“ zu sagen, da das einfacher auszusprechen ist als „Quantil 0,9“. Beispielsweise kann bei einer Prognose auf Quantilniveau P90 davon ausgegangen werden, dass der tatsächliche Wert in 90 % der Fälle unter dem prognostizierten Wert liegen wird. Insbesondere wenn der prognostizierte Wert bei einer Zeit t_1 und einem Quantilniveau 0,9 30 ist, bedeutet das, dass der tatsächliche Wert zum Zeitpunkt t_1 , wenn Sie 1.000 Simulationen durchgeführt haben, für 900 Simulationen voraussichtlich unter 30 liegt, und bei 100 Simulationen wird davon ausgegangen, dass der tatsächliche Wert über 30 liegt.

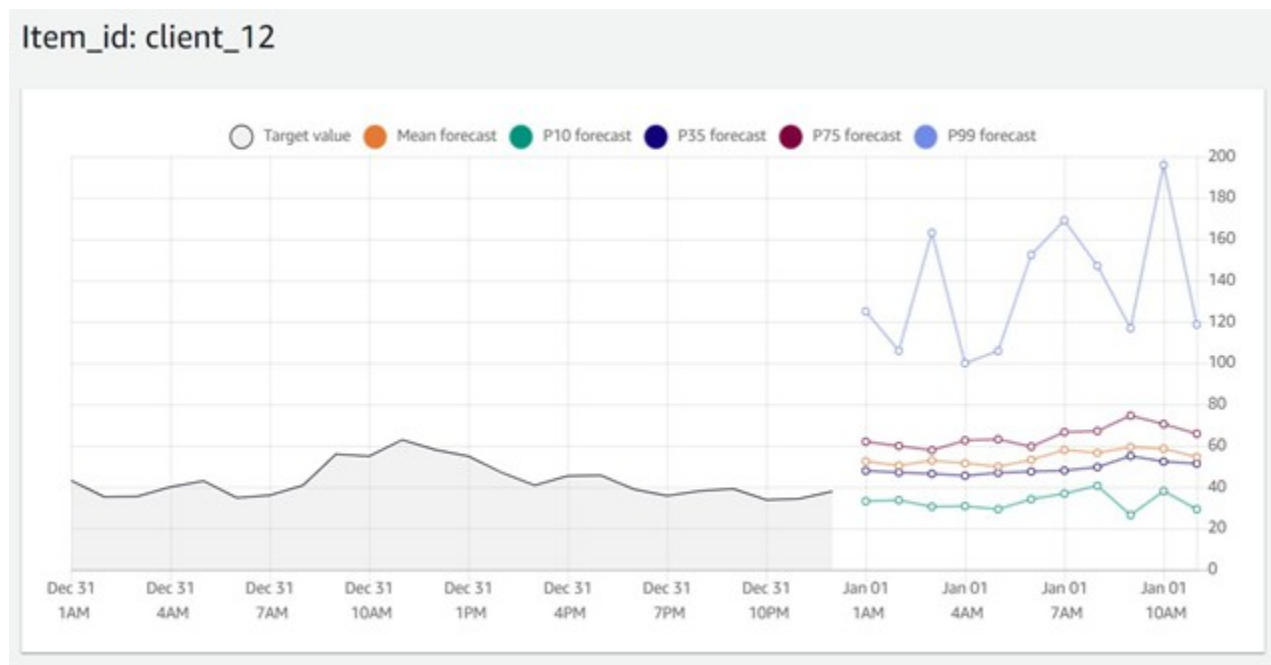
Andererseits ist die P90-Prognose in 90 % der Fälle eine Überprognose, und sie ist dann nützlich, wenn die Opportunitätskosten, die aus dem Nichtverkauf eines Artikels entstehen, extrem hoch sind, oder die Kosten des investierten Kapitals niedrig sind. Ein Lebensmittelgeschäft könnte die P90-Prognose für Produkte wie Milch oder Toilettenpapier verwenden, das es nicht will, dass ihm diese Produkte ausgehen und da es nichts ausmacht, wenn auf den Regalen immer etwas Milch oder Toilettenpapier übrig bleibt.

Für die p50-Prognose (oft auch als Median-Prognose bezeichnet) wird erwartet, dass der wahre Wert in 50 % der Fälle unter dem prognostizierten Wert liegt. Zur Beurteilung der Genauigkeit kann die wQL [0,5]-Metrik verwendet werden. Die p50-Quantilprognose kann nützlich sein, wenn es nichts ausmacht, Überbestände zu haben und wenn für einen bestimmten Artikel eine mittelgroße Nachfrage verzeichnet wird.

Visualisierung

Amazon Forecast unterstützt in der AWS-Managementkonsole das native Plotten von Prognosen. Darüber hinaus können Sie den vollständigen Python-Data-Science-Stack nutzen (siehe [Amazon](#)

[Forecast-Beispiele](#)). Amazon Forecast mithilfe der ExportForecastJob-API das Exportieren von Prognosen als CSV-Datei. Auf diese Weise können Benutzer die Prognose im Analyse-Tool ihrer Wahl visualisieren.



In der Amazon Forecast-Konsole bereitgestellte Visualisierung bei unterschiedlichen Quantilen

Zusammenfassung zu Prognose-Workflows und APIs

In der folgenden Tabelle werden die einzelnen Schritte des Prognose-Workflows der entsprechenden Amazon Forecast-API zugeordnet.

Tabelle 1: Prognoseschritte und Amazon Forecast-APIs

Schritt	API	API-Funktionen
Schritt 1: Daten sammeln und aggregieren Schritt 2: Daten vorbereiten	Create_Dataset_Group , Create_Dataset , Create_Dataset_Import_Job	1. Legen Sie die High-Level-Domäne (Einzelhandel, Metriken usw.) für das Problem fest. 2. Definieren Sie das Schema für die verschiedenen Datasets (Zielfdaten, verwandte Daten, Artikel-Metadaten). 3. Importieren Sie Daten aus Amazon S3 in Amazon Forecast.
Schritt 3: Erstellen eines Predictors Schritt 4: Auswerten von Predictors	Create_Predictor	1. Führt die ETL aus. 2. Teilt Daten in Trainings-/Testsätze auf und trainiert das Modell. 3. Verwenden Sie die Option Create_predictor_backtest_Export_job , um Backtest-Ergebnisse nach CSV zu exportieren und Metriken auf Elementniveau zu berechnen.

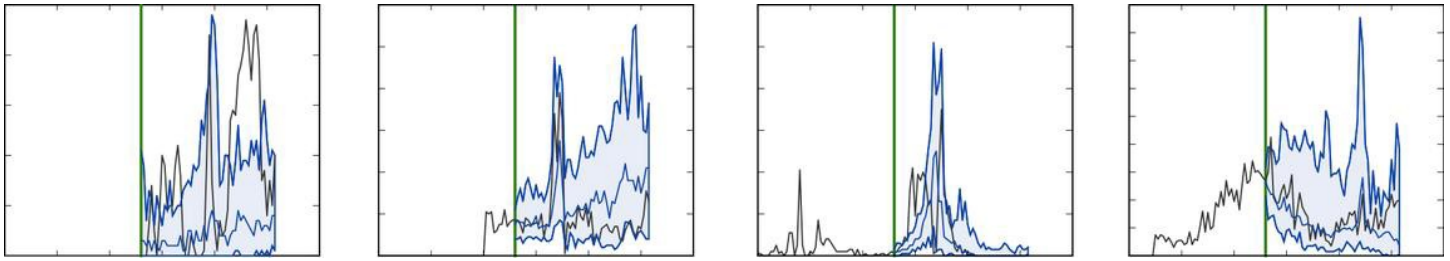
Schritt	API	API-Funktionen
<u>Schritt 5: Prognosen generieren und für die Entscheidungsfindung verwenden</u>	Create_Forecast	1. Trainiert/hostet das Modell. 2. Generiert Vorhersagen für ein bestimmtes interessantes Quantil (z. B. eine ganze Zahl von 1 bis 99 und den Mittelwert) über den Prognosehorizont hinweg.
	Query_Forecast	Ermöglicht es Ihnen, von Create_Forecast erstellte Prognosen zu verrechnen

Verwenden von Amazon Forecast für gängige Szenarien

Sie können auch Was-wäre-wenn-Analysen durchführen, indem Sie verschiedene Prognosen auf Grundlage sich ändernder externer Variablen (z. B. Preise oder Werbeaktionen) erstellen. In der E-Commerce-Fallstudie können Sie beispielsweise verschiedene Prognosen erstellen, die auf Werbeaktionen basieren, die Sie möglicherweise planen. Sie können die Nachfrage für ein Produkt prognostizieren, für das zunächst ein Rabatt von 10 % und anschließend ein Rabatt von 20 % gewährt wird, um daraus zu ermitteln, welche Menge von diesem Produkt Sie auf Lager haben müssten, um die Nachfrage zu decken. Dies können Sie tun, indem Sie einzelne Dataset-Gruppen einrichten und die entsprechenden Zeitreihen auf Grundlage des Szenarios, das Sie interessiert, in jeder Dataset-Gruppe aktualisieren.

Darüber hinaus können Sie Prognosen für Artikel erstellen, für die keine Historie vorliegt. Ein solcher Fall wird manchmal auch als „Kaltstartproblem“ bezeichnet. Um Prognosen für den neuen Artikel zu generieren, muss bei diesem Ansatz mithilfe von DeePar+ oder CNN-QR und auf Basis von Metadaten (z. B. ein Artikel-Metadaten-Dataset) ein Predictor erstellt werden.

Die folgende Abbildung zeigt Beispiele für reale betriebliche Prognoseprobleme anhand vier verschiedener SKUs.



Beispiele für reale betriebliche Prognoseprobleme anhand vier verschiedener SKUs.

In der obigen Abbildung befinden sich links von der vertikalen Linie die tatsächlich beobachteten Werte. Rechts von der vertikalen Linie sind in Blau die Prognosen zu sehen, im Vergleich zu den tatsächlichen Werten, die in Schwarz dargestellt sind. Beachten Sie, dass die Historie der einzelnen SKUs links von der vertikalen Linie kein Indikator für die rechts neben der grünen Linie dargestellte Entwicklung ist.

Implementierung von Forecast in die Produktion

Nach Beendigung des gesamten Amazon Forecast-Workflows ist es wichtig, die Hauptunterschiede zwischen der `Create_Forecast` -API und der `Create_Predictor`-API zu identifizieren und zu bestimmen, wann die jeweilige API verwendet werden sollten.

Die erste API wird hauptsächlich für Machbarkeitsstudien zur Bewertung der Modellgenauigkeit/Metriken verwendet. Die zweite API wird hingegen zur Erstellung von Prognosen in einer Produktionsumgebung angewandt.

Sobald sich `Create_Predictor` im Einsatz befindet, muss sie nicht jedes Mal, wenn eine Prognose generiert wird, ausgeführt werden, sondern nur dann, wenn das Modell wegen Datenänderungen oder im Rahmen eines festgelegten Rhythmus (z. B. alle zwei Wochen oder jeden Monat) neu trainiert werden muss. Da die Datasets mit neuen Daten aktualisiert werden, muss nur `Create_Forecast` ausgeführt werden, um Prognosen für den neuen Prognosehorizont zu erstellen.

Außerdem müssen Sie Ihre Dataset-Importe und Prognosevorgänge automatisieren, um fortlaufend neue Prognosen zu erstellen. Hierzu können heute Cron-Jobs verwendet werden, die Sie mithilfe einer Kombination aus [Amazon CloudWatch Events](#)-Protokollen, [AWS Step Functions](#) und [AWS Lambda](#)-Funktionen einrichten können. Durch das Einrichten von [Cron](#)-Jobs werden wiederum Amazon Forecast-API-Aufrufe für Import/Neutraining oder die Generierung von Prognosen automatisiert. Schließlich ist es wichtig, Ihre Ressourcen zu verwalten und sie in regelmäßigen Abständen zu löschen, damit Sie die vom Dienst vorgeschriebenen [Systemgrenzen](#) nicht

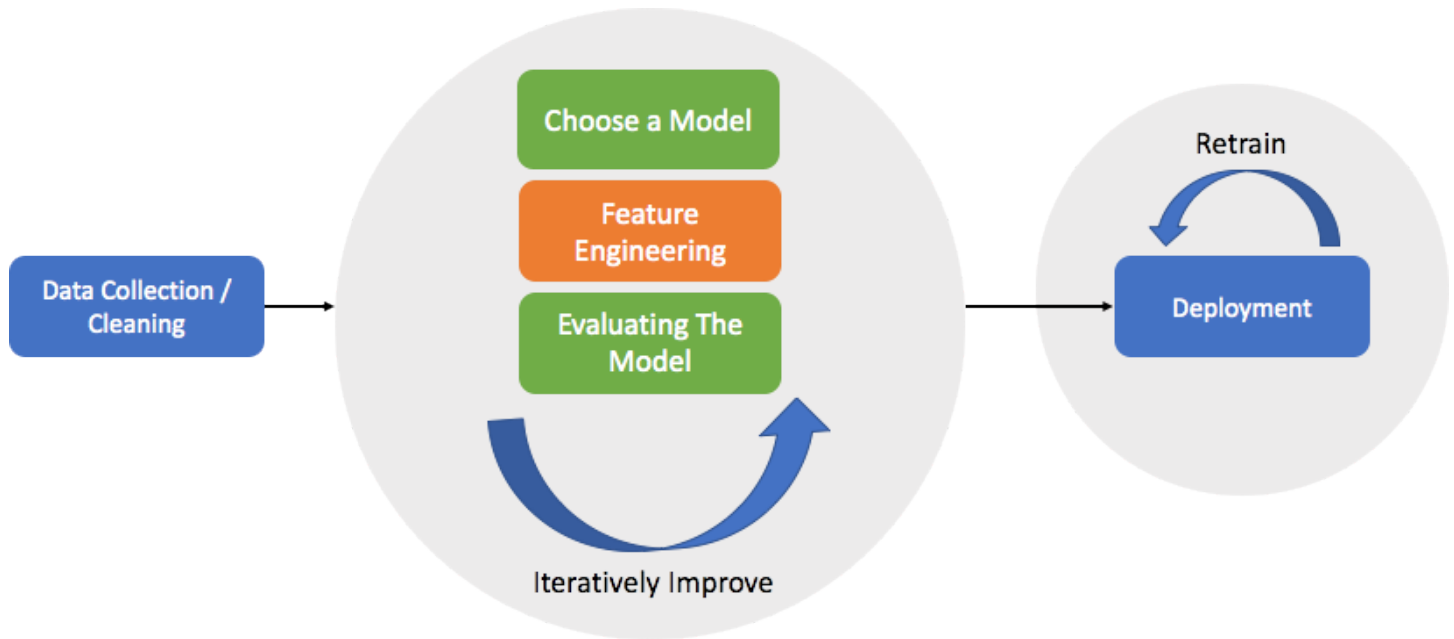
überschreiten. Lesen Sie diesen [Blog-Beitrag](#) sowie den Artikel zu [Amazon Redshift](#), um mehr über das Einrichten von geplanten Jobs zu erfahren.

Fazit

[Januschowski und Kolassa \(2019\)](#) bieten eine Klassifizierung der Prognoseprobleme, die auf die Entscheidungen ausgerichtet sind, die Unternehmen treffen müssen, einschließlich strategischer, taktischer und operativer Entscheidungen. Jeder Entscheidungsebene werden entsprechende Prognoseaufgaben zugewiesen.

Operative und taktische Prognosen sind deshalb problematisch, da sie großen Datenmengen und in der Regel einen hohen Automatisierungsgrad erfordern. Diese Probleme lassen sich mithilfe verschiedener Prognosemethoden lösen. Lokale Prognosemethoden eignen sich in der Regel gut zum Lösen strategischer Prognoseprobleme, auf Deep Learning basierende Methoden sind für operative Prognoseprobleme geeignet, und alle dazwischen liegenden Probleme können durch ein wenig Experimentieren gelöst werden. Obwohl in diesem Whitepaper operative Prognoseprobleme erörtert werden, vertritt Amazon Forecast keinerlei Meinung zu den vorgeschlagenen Modellen und bietet Modelle, mithilfe derer sowohl strategische als auch operative und taktische Prognoseprobleme gelöst werden können.

Der Prozess zum Lösen von operativen Prognoseproblemen lässt sich in grundlegende Schritte unterteilen, angefangen bei der Datenerfassung und -aufbereitung bis hin zum Erstellen und Bereitstellen von Modellen. Im Allgemeinen ist es am sinnvollsten, den Prozess eher als einen iterativen und nicht als einen linearen Prozess zu betrachten. Da Modelle und Anwendungsfälle beispielsweise zu einem besseren Verständnis beitragen, kann es sinnvoll sein, sich noch einmal die Phase der Datenerfassung anzusehen. Die Modellentwicklung ist an sich ebenfalls ein hochiterativer Prozess.



Ein vereinfachter Prozess zur Entwicklung eines produktionsreifen Prognosemodells.

Mitwirkende

An diesem Dokument haben folgende Personen mitgewirkt:

- Yuyang Wang, Senior Machine Learning Scientist, AI Vertical Services
- Danielle Robinson, Applied Scientist, AI Vertical Services
- Tim Januschowski, Manager, ML Applied Science
- Namita Das, Senior Product Manager, AI Vertical Services
- Christy Bergman, Senior AI/ML Specialist Solutions Architect
- Kris Tonthat, Technischer Redakteur, KI-/ML-Dokumentation

Weitere Informationen

Weitere Informationen zu Zeitreihenprognosen und Deep-Learning-Methoden finden Sie unter:

- [Amazon Forecast-Dokumentation](#)
- [Blog zur allgemeinen Verfügbarkeit von Amazon Forecast](#)
- [Jetzt in Amazon SageMaker verfügbar: DeepAR-Algorithmus für mehr Genauigkeit](#)
- [Amazon SageMaker DeepAR unterstützt nun fehlende Werte, kategorische Funktionen und Zeitreihenfunktionen sowie generalisierte Häufigkeiten](#)
- [Amazon Forecast kann jetzt Convolutional Neural Networks \(CNNs\) verwenden, um Prognosemodelle bis zu zweimal schneller und mit bis zu 30 % mehr Genauigkeit zu trainieren](#)
- [Amazon Forecast unterstützt nun Genauigkeitsmessungen für einzelne Artikel](#)
- [Mit Amazon Forecast können Sie die Genauigkeit des Prognosemodells messen, um Ihre Geschäftsziele zu optimieren](#)
- [Amazon Forecast Weather Index - Lokales Wetter automatisch einfügen, um die Genauigkeit Ihres Vorhersagemodells zu verbessern](#)
- [Wissenschaftliche Arbeiten zu Zeitreihenprognosemodellen](#)
- [GitHub-Seite mit Amazon Forecast-Beispielen](#)
- [AWS -Architekturzentrum](#)

Anhang A: Häufig gestellte Fragen

F: Welche ersten Schritte sind für die Benutzung von Amazon Forecast notwendig?

1. Sie benötigen zunächst ein AWS-Konto.
2. Öffnen Sie als Nächstes in der [AWS-Managementkonsole](#) den Forecast-Service, erstellen Sie eine Datensatzgruppe und importieren Sie eine .csv-Datei in den Zielzeitreihen-Datensatz (erforderlich). Um zu beginnen, benötigen Sie zumindest historische Daten für die Menge, die Sie vorhersagen wollen, z. B. Strom pro Zeitstempel pro Haushalt.
3. Erstellen Sie abschließend ein Modell, indem Sie [CreatePredictor](#) ausführen, und generieren Sie Ergebnisse, indem Sie [CreateForecast](#) ausführen. Weitere Informationen finden Sie auf der Dokumentationsseite [Erste Schritte](#).

Lesen Sie auch den [Leitfaden für die Einführung in GitHub und für bewährte Methoden](#).

F: Ist Amazon Forecast die richtige Wahl für mich?

Nicht alle Probleme, die mit dem maschinellen Lernen zusammenhängen, sind Prognoseprobleme. Die erste Frage, die Sie sich stellen sollten, lautet: „Beinhaltet mein Geschäftsproblem in seiner Aussage Zeitreihen?“ Benötigen Sie beispielsweise einen bestimmten Wert für eine bestimmte Uhrzeit und ein bestimmtes Datum, das in der Zukunft liegt? Prognosen eignen sich nicht für allgemeine, statische Probleme, bei denen ein bestimmtes Datum/eine bestimmte Uhrzeit keine Rolle spielt (wie z. B. bei der Erkennung von Betrug oder wenn es darum geht, einem Benutzer einen Filmtitel zu empfehlen. Für statische Probleme gibt es viel schnellere Lösungen.

Neben der Tatsache, dass Sie Zeitreihendaten haben sollten, sollten die Daten selbst „dicht“ beieinander liegen und eine lange Historie aufweisen. Dies ist in der folgenden Tabelle zusammengefasst:

Tabelle 2 – Kriterien und Amazon Forecast-Algorithmusklassen

Kriterien	Amazon-Forecast-Algorithmusklasse
Großer Datensatz mit bis zu fünf Millionen Zeitreihen mit ähnlichen zugrundeliegenden Mustern + saisonalen Effekten + verwandten Daten. Wenn versucht wird, jährliche Ereigniss	Proprietäres Deep Learning DeepAR+, CNN-QR von Amazon Forecast

Kriterien	Amazon-Forecast-Algorithmuskategorie
e zu erfassen, sollte jede Zeitreihe eine lange Historie (idealerweise mehr als zwei Jahre) haben und jede Zeitreihe sollte mehr als 300 Datenpunkte (idealerweise mindestens 1.000) aufweisen.	
Kleiner Datensatz mit Zeitreihen von 1–100, wobei die meisten Zeitreihen mehr als 300 Datenpunkte + saisonale Effekte + verwandte Daten aufweisen.	Prophet
Kleiner Datensatz mit Zeitreihen von 1–10, wobei die meisten Zeitreihen mehr als 300 Datenpunkte und saisonale Effekte aufweisen.	ETS, ARIMA
Intermittierend (spärlich, enthält viele 0) mit Zeitreihen von 1–10, wobei die meisten Zeitreihen mehr als 300 Datenpunkte aufweisen.	Proprietäres NPTS von Amazon Forecast
Kleiner Datensatz (regulär oder spärlich) mit Zeitreihen von 1–10, wobei die meisten Zeitreihen weniger als 300 Datenpunkte aufweisen.	Die Daten sind für Amazon Forecast zu klein. Probieren Sie stattdessen ETS in Excel oder die traditionellen statistischen Modelle ARIMA und Prophet aus.

Am besten ist es, das erste Mal mit Ihren Daten im AutoML-Modus in Ihrem Predictor zu trainieren. AutoML durchläuft automatisch alle Algorithmen (die DL-Algorithmen werden bei aktivierter HPO durchlaufen), um festzustellen, welcher Algorithmus für Ihre Daten am besten geeignet ist.

F: Was muss ich bei fehlenden Daten bedenken? Ab wann lassen sich keine vernünftigen Prognosen erstellen?

Es kann vorkommen, dass es Probleme mit der Erfassung von Daten gibt oder dass das Aggregationsniveau der Daten zu niedrig oder zu hoch ist. Allgemein gilt, dass die Prognoselänge nicht länger sein darf als 1/3 der Trainingsdaten.

Neben der Menge der fehlenden Daten ist außerdem die Imputation der fehlenden Daten zu berücksichtigen. Sie können alle 0-Werte in Nullen umwandeln und Amazon Forecast die aufwendige Aufgabe überlassen, fehlende Werte automatisch zu imputieren. Amazon Forecast erkennt automatisch, ob die fehlenden Werte auf die Einführung neuer Produkte (Cold-Starts) oder auf Altprodukte zurückzuführen sind. Sie können für fehlende Werte mehrere Logiken verwenden, darunter Wert, Median, Minimum, Maximum, Null, Durchschnitt und NaN (nur für Zielzeitreihen). In der [Dokumentation finden Sie Informationen zur Syntax zum Ausfüllen von fehlenden Werten](#).

- „Frontfill“ (nur TTS) bezieht sich auf neue Produkte oder Kaltstarts und darauf, wie Sie mit Nullen umgehen möchten, bevor das Produkt eine Historie aufweist
- „Middlefill“ bezieht sich auf Nullen in der Mitte der Zeitreihenwerte
- „Backfill“ bezieht sich auf Altprodukte und darauf, wie Sie mit Nullen umgehen möchten, nachdem ein Produkt aus dem Umlauf gezogen wurde
- „Futurefill“ (nur RTS) bezieht sich auf Nullen, die nach dem Ende der Trainingsdaten auftreten

F: Meine historischen Eingabedaten enthalten keine negativen Werte, in der Nachfrageprognose sehe ich aber negative Werte? Wie ist das möglich? Wie kann ich dies vermeiden?

Für alle Modelle außer NPTS (trainiert auf den nicht negativen Daten) und DeepAR (mit negativer Binomial-Wahrscheinlichkeitsfunktion) gibt es keine Garantie dafür, dass positive Zahlen generiert werden. Die Lösung besteht darin, zu einem der oben genannten Modelle zu wechseln oder die Prognosewerte auf nicht negative Werte zu kürzen.

F: Warum variieren Genauigkeitsmetriken bei Quantilen? Da es sich um dasselbe Modell handelt, müsste der Fehler doch derselbe sein?

Weitere Erläuterungen zur Abhängigkeit der Gewichtung vom Quantil finden Sie unter [Weighted Quantile Loss \(wQL\)](#).

Stellen Sie sich vor, Sie haben alle Ihre Prognosen für drei verschiedene Quantile erstellt: p10, p50, p90. Die drei Vorhersagen sind selbst Zufallsvariablen. Die Genauigkeiten werden separat auf jeder Quantilebene zwischen tatsächlichen Werten und Prognosen berechnet. Sie sehen möglicherweise eine Tabelle mit „wQL“ (Weighted Quantile Losses), wie unten dargestellt. Die wQL-Werte stehen nicht in einer deterministischen Beziehung zueinander. (Verlust wegen Rückruf bedeutet Fehler und ist daher ungeordnet; Quantilprognosen sind jedoch geordnet). Es gibt also keinen Grund dafür, dass p90 wQL größer sein muss als beispielsweise p50 wQL.

Tabelle 3 – Beispiel für Prognosequantile

	A	B	C
1	P10 wQL	P50 wQL	P90 wQL
2	0.18647	0.50879	0.30428

F: Wie kann ich die Prognosegenauigkeit verbessern?

Die Prognosegenauigkeit hängt davon ab, ob die richtigen Daten in der richtigen Menge vorliegen und der erwarteten Qualität entsprechen. Wenn die Genauigkeit unbefriedigend ist, kann es sinnvoll sein, sich die Frage zu stellen, wie vorhersehbar das Problem ist (oder wie zufällig/verrauscht/stationär die Daten sind). Weitere Faktoren, die berücksichtigt werden sollten, sind die Bewertung verschiedener Modelle, Hyperparametereinstellungen und die Integration zusätzlicher Funktionen mithilfe der zugehörigen Zeitreihen und Elementmetadaten-Datasets. Konkrete Vorschläge finden Sie in diesem Best Practices-Dokument zu GitHub unter <https://github.com/aws-samples/amazon-forecast-samples/blob/master/ForecastCheatSheet.md#iteratebp>.

F: Ich habe einen Algorithmus, der bei meinem Anwendungsfall sehr gut funktioniert und der in Amazon Forecast nicht angeboten wird. Was soll ich tun?

Das Amazon-Forecast-Team hilft Ihnen gerne bei diesem Anwendungsfall. Wenden Sie sich per E-Mail an das Service-Team von Amazon Forecast <amazonforecast-poc@amazon.com>.

Anhang B: Referenzen

[Januschowski, Tim and Kolassa, Stephan. A classification of business forecasting problems. Foresight: The International Journal of Applied Forecasting. 2019](#)

[Salinas, David, Flunkert, Valentin, Gasthaus, Jan and Januschowski, Tim. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting. 2019](#)

[Gasthaus, Jan, Benidis, Konstantinos, Wang, Yuyang, Rangapuram, Syama Sundar, Salinas, David, Flunkert, Valentin and Januschowski, Tim. {Probabilistic Forecasting with Spline Quantile Function RNNs. The 22nd International Conference on Artificial Intelligence and Statistics. 2019](#)

[Januschowski, Tim, Gasthaus, Jan, Wang, Yuyang, Salinas, David, Flunkert, Valentin, Bohlke-Schneider, Michael and Callot, Laurent. Criteria for classifying forecasting methods. International Journal of Forecasting. 2019 \(Anmeldung erforderlich\)](#)

[Januschowski, Tim, Gasthaus, Jan, Wang, Yuyang, Rangapuram, Syama and Callot, Laurent. Deep Learning for Forecasting. Foresight: The International Journal of Applied Forecasting. 2018](#)

[Januschowski, Tim, Gasthaus, Jan, Wang, Yuyang, Rangapuram, Syama Sundar and Callot, Laurent. Deep Learning for Forecasting: Current Trends and Challenges. Foresight: The International Journal of Applied Forecasting. 2018](#)

[Bose, Joos-Hendrik, Flunkert, Valentin, Gasthaus, Jan, Januschowski, Tim, Lange, Dustin, Salinas, David, Schelter, Sebastian, Seeger, Matthias and Wang, Yuyang. Probabilistic demand forecasting at scale. Proceedings of the VLDB Endowment. 2017](#)

Dokumentverlauf

Abonnieren Sie den RSS-Feed, um über Aktualisierungen des Whitepapers benachrichtigt zu werden.

Änderung	Beschreibung	Datum
Whitepaper wurde aktualisiert	Aktualisierungen.	1. September 2021
Erstveröffentlichung	Erstveröffentlichung des Whitepapers	4. Februar 2020

Note

Um RSS-Aktualisierungen zu abonnieren, muss für den von Ihnen verwendeten Browser ein RSS-Plug-in aktiviert sein.

Hinweise

Kunden sind eigenverantwortlich für die unabhängige Bewertung der Informationen in diesem Dokument zuständig. Dieses Dokument: (a) dient rein zu Informationszwecken, (b) spiegelt die aktuellen AWS-Produktangebote und Verfahren wider, die sich ohne vorherige Mitteilung ändern können, und (c) impliziert keinerlei Verpflichtungen oder Zusicherungen seitens AWS und dessen Tochtergesellschaften, Lieferanten oder Lizenzgebern. AWS-Produkte oder -Services werden im vorliegenden Zustand und ohne ausdrückliche oder stillschweigende Gewährleistungen, Zusicherungen oder Bedingungen bereitgestellt. Die Verantwortung und Haftung von AWS gegenüber seinen Kunden wird durch AWS-Vereinbarungen geregelt. Dieses Dokument ist weder ganz noch teilweise Teil der Vereinbarungen zwischen AWS und seinen Kunden und ändert diese Vereinbarung auch nicht.

© 2021 Amazon Web Services Inc. bzw. Tochtergesellschaften des Unternehmens. Alle Rechte vorbehalten.

AWS-Glossar

Die neueste AWS-Terminologie finden Sie im [AWS-Glossar](#) im AWS-Referenzhandbuch.