

Whitepaper

Effektives Forecasting:

Technische Methoden, profitable Anwendung & Herausforderungen im Unternehmensumfeld



Forecasting verwandelt Daten aus der Vergangenheit in belastbare Prognosen über die Zukunft. Erfahren Sie hier, wie Forecasting technisch implementiert werden kann und Unternehmen davon in der Praxis profitieren.

Über das Whitepaper und die Autoren

Finden Sie heraus, was Sie in diesem Whitepaper erwarten und welche Autoren hinter dem Whitepaper stecken.

In diesem Whitepaper erfahren Sie, was Forecasting ist und wie Unternehmen verschiedene Forecasting-Methoden technisch implementieren und gewinnbringend einsetzen. Wir stellen Use Cases aus der Praxis vor und diskutieren die Herausforderungen für Unternehmen bei der Implementierung.



Marlon Ziegler
Senior Data Scientist

Marlon Ziegler ist Senior Data Scientist spezialisiert auf Zeitreihen-Daten, Forecasting und Anomaly Detection. Er leitet die interne Arbeitsgruppe „Forecasting“, dessen Ziel es ist den Wissensaustausch und -aufbau zu fördern.



Daniel Lüttgau
Head of AI Development

Daniel Lüttgau verantwortet bei statworx den Bereich AI Development und ist Experte für die Anwendung von KI, um Mehrwerte für Unternehmen zu generieren.

statworx®

statworx GmbH
Hanauer Landstr. 150
60314 Frankfurt am Main

www.statworx.com
info@statworx.com
+49 (0)69 6783 067 - 51

Inhalt

4

-
- **Einführung**
Ein kompakter Einstieg in das Thema Forecasting

5

-
- **Definition & Überblick**
Die Bedeutung von Forecasting und die Ziele der Nutzung

8

-
- **Use Cases**
Drei Praxisbeispiele, in denen Forecasting Anwendung findet

12

-
- **Herausforderungen Part 1**
Aktuelle Herausforderungen für Organisationen

14

-
- **Technischer Überblick**
Methoden zur Erstellung von Forecasting-Modellen

21

-
- **Herausforderungen Part 2**
Typische Herausforderungen bei der Implementierung

26

-
- **Fazit & Ausblick**
Wie Forecasting Mehrwert auf mehreren Ebenen schafft

27

-
- **Über statworx**
Wie wir Unternehmen bei ihrer KI-Transformation unterstützen

Einführung

Einstieg

Von Modellen zur Kundenentwicklung, über Auftrags-eingangs- und Absatzvorhersagen für ein verbessertes Stockmanagement bis hin zur Prognose von Liquiditätszeitreihen: Die Erstellung von Prognosen (Forecasting) ermöglicht Unternehmen, sich auf Veränderungen vorzubereiten, Ressourcen effizient zu allokalieren und die Profitabilität zu optimieren.

Forecasting ist eine zentrale Aufgabe in fast allen Abteilungen eines Unternehmens. Typische Use Cases sind:



In allen aufgeführten Fällen kann eine datengetriebene Prognose einen erheblichen Mehrwert liefern. Dieser Mehrwert kann auf zwei Arten entstehen:

1) Entweder ist die Prognose selbst das Ziel und bietet finanziellen Nutzen.

2) Oder die Prognose ist ein unverzichtbarer Bestandteil eines größeren Prozesses, der zur Wertschöpfung beiträgt.

Dieses Whitepaper bietet Unternehmen einen Überblick über verschiedene Prognosemodelle und zeigt auf, welche Modelle sich für welche Herausforderungen eignen. Es stellt außerdem dar, wie die Integration von datengetriebenen Prognosen in verschiedenen Bereichen wie Vertrieb, Finanzen, Produktion und Logistik zu einer effizienteren Ressourcenallokation und gesteigerter Profitabilität führt.

Zudem geht das Whitepaper auf die unterschiedlichen methodischen Ansätze - von klassischen statistischen Methoden bis hin zu modernen Deep Learning- und Foundation-Modellen - sowie die Herausforderungen und Lösungen bei der Implementierung ein. Ziel ist es, eine umfassende Übersicht zu bieten, wie Unternehmen datengetriebene Prognosen effektiv nutzen können, um strategische Klarheit und operationale Exzellenz zu erreichen.

Weiter auf der nächsten Seite →

1

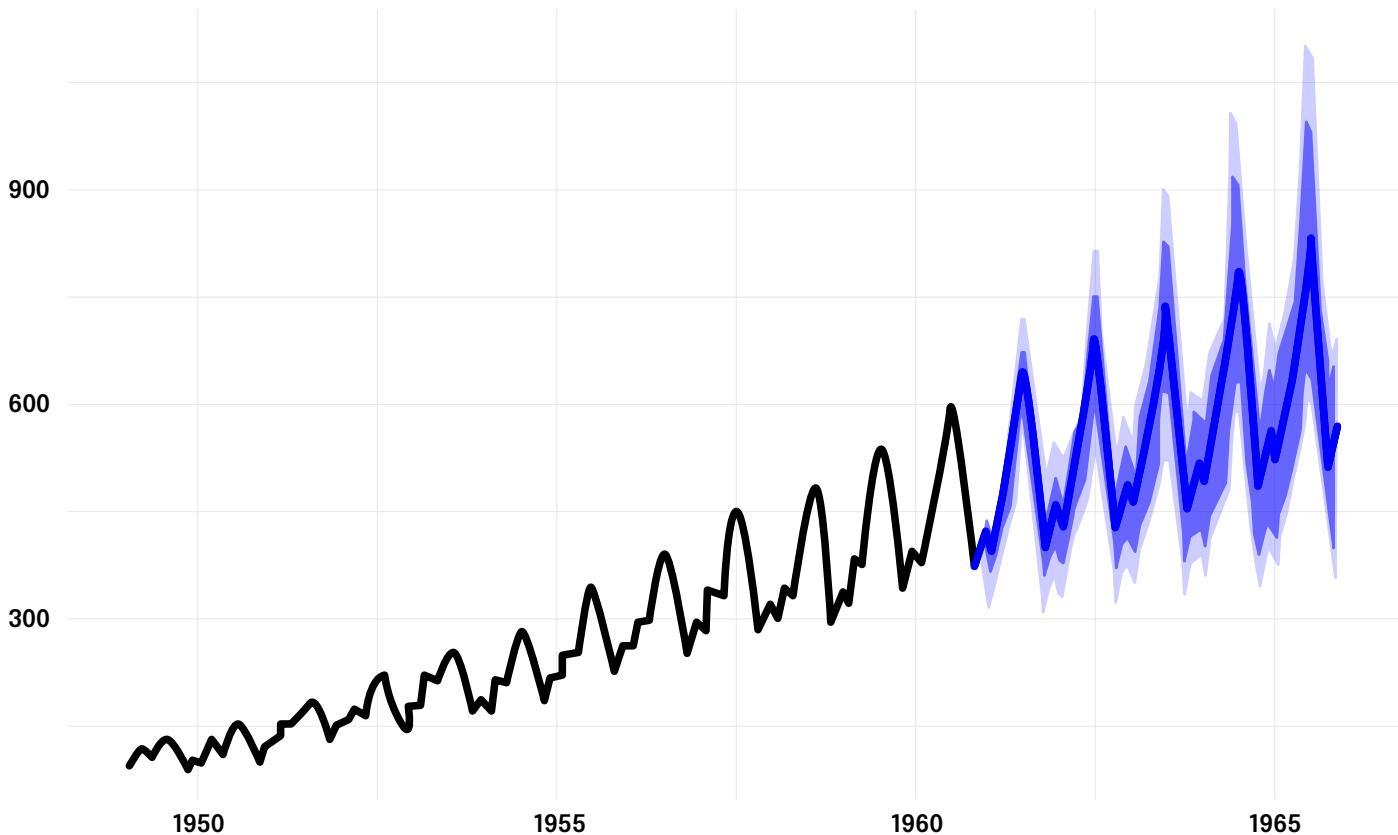
Definition: Forecasting

Definition & Überblick

Was ist Forecasting?

Der Begriff **Forecasting** bezeichnet den methodischen Prozess der Vorhersage zukünftiger Ereignisse, Trends oder Entwicklungen auf der Grundlage historischer Daten und entsprechender Modelle. Forecastings basieren meist auf der zeitlichen Abfolge von Daten, sie können aber auch auf Querschnittsdaten oder strukturellen Modellen beruhen. In diesem Whitepaper liegt der Fokus auf Time-Series-Forecasting, zu Deutsch: Zeitreihen-Prognose bzw. -Vorhersage. Es bezeichnet eine Methode, mit der zukünftige Ereignisse oder Werte auf Basis von bisher beobachteten Daten vorhergesagt werden. Ein Anwendungsbeispiel ist die Temperaturvorhersage: Aus einer Liste von täglich gemessenen Temperaturen lassen sich Muster erkennen und so Schätzungen über zukünftige Temperaturen erstellen. Es geht darum, aus der Vergangenheit zu lernen, um fundierte Vorhersagen für die Zukunft ableiten zu können.

Eine **Zeitreihe** ist eine chronologisch geordnete und in gleichbleibenden Abständen erhobene Sequenz von Beobachtungen einer numerischen Größe, wie z. B. monatlichen Absatzzahlen eines bestimmten Produkts. Zeitreihen können in verschiedene Komponenten zerlegt werden, wie Trend und Saisonalität. Eine Vorhersage ist eine Schätzung, wie sich die Sequenz der Beobachtungen in der Zukunft fortsetzen wird. Ebenso vielfältig wie die Anwendungsfälle von Forecasts ist der methodische Spielraum: Von klassisch-statistischen, univariaten Methoden über etablierte multivariate Machine-Learning-Modelle bis hin zu komplexen Deep-Learning-Ansätzen gibt die Forecasting-Toolbox einiges her. Die Wahl des geeigneten Ansatzes richtet sich nach den Eigenschaften der Input-Daten und den Anforderungen an den Forecast.



Lehrbuchbeispiel für eine Zeitreihe von Box & Jenkins: Monatliche Gesamtzahlen der Passagiere im internationalen Flugverkehr, 1949 bis 1960. Vorhersage über 5 Jahre.

Definition & Überblick

Das Ziel von Forecasting ist nicht, die Vorhersagen von Expert:innen mit Domänenwissen vollkommen aus der Gleichung auszuschließen und durch datenbasierte Erkenntnisse zu ersetzen. Im Gegenteil: Domänenwissen und Bauchgefühl bzw. Instinkt haben nach wie vor ihren verdienten Stellenwert. In manchen Fällen bieten sie sogar einen entscheidenden Vorteil gegenüber strukturiert verfügbaren Daten, denn Modelle können nicht alle potenziellen Einflussfaktoren berücksichtigen, die ein Mensch mit viel Erfahrung in seine Prognose einfließen lassen kann. Vielmehr geht es beim Forecasting darum, Expert:innen neben ihren Erfahrungswerten einen weiteren Wissens-Input bereitzustellen, der auf datengetriebener Mustererkennung basiert.

Die Integration von menschlichem Wissen und maschinellen Forecasting-Lösungen stellt für Organisationen eine Herausforderung dar. Denn Expert:innen müssen genau verstehen, welche Informationen in den Forecasting-Modellen enthalten sind und welche nicht. Erst diese zusätzliche Information befähigt sie dazu, Vorhersagen anzupassen und zu verfeinern. Zum Beispiel können Expert:innen Vorhersagen korrigieren, wenn sie wissen, dass aktuelle Lieferkettenengpässe nicht durch die vorhandenen Daten erfasst werden und daher in den Vorhersagen nicht berücksichtigt sind. Diese zusätzliche menschliche Information ergänzt die maschinengenerierten Vorhersagen und trägt zur Erhöhung der Genauigkeit bei.

Um eine solche Synergie zwischen Mensch und Maschine zu erreichen, ist ein effektives Change-Management unerlässlich. Organisationen müssen sicherstellen, dass ihre Mitarbeiter:innen die nötigen Schulungen und Werkzeuge erhalten, um die neuen Technologien effizient nutzen zu können. Dies umfasst nicht nur technische Schulungen, sondern auch die Entwicklung eines Verständnisses für die Grenzen und Möglichkeiten der eingesetzten Forecasting-Modelle. Darüber hinaus spielt die Kommunikation eine wesentliche Rolle. Es muss ein kontinuierlicher Dialog zwischen den Entwickler:innen der Forecasting-Lösungen und den Anwender:innen stattfinden. Nur so kann sichergestellt werden, dass die Modelle ständig verbessert und an die sich ändernden Anforderungen der Organisation angepasst werden.



2

Forecasting Use Cases

Use Cases

Praxisbeispiele von Forecasting

Forecasting nimmt viele Formen an und wird in verschiedenen Branchen eingesetzt, um zukünftige Entwicklungen vorherzusagen. Dieses Kapitel präsentiert verschiedene Praxisbeispiele, die die Vielseitigkeit und den praktischen Nutzen von Forecasting in der Geschäftswelt aufzeigen. Von Verkaufsprognosen in der Modeindustrie über Preis- und Rabattanpassungen bis hin zur Liquiditätsprognose – die Beispiele zeigen, wie präzise Vorhersagen strategische Entscheidungen unterstützen und Unternehmen einen Mehrwert bieten können.

Praxisbeispiel: Forecasting in der Fast Fashion Industrie

HERAUSFORDERUNG

Unser Kunde, ein führendes Unternehmen für Sportartikel im europäischen Markt, stand vor der Herausforderung, die Verkaufsprognosen für ein umfangreiches und komplexes Produktpotfolio zu verbessern. Wir entwickelten eine automatisierte Forecasting-Engine, die in der Lage ist, erklärbare und probabilistische Vorhersagen für zukünftige Produktverkäufe zu generieren. Diese Vorhersagen erstrecken sich über Zeiträume von bis zu 18 Monaten.

LÖSUNG

Die Forecasting-Engine basiert auf einem fortschrittlichen, erklärbaren Deep-Learning-Modell, das in der Lage ist, große Mengen an Daten zu verarbeiten und detaillierte Vorhersagen zu liefern.

ERGEBNIS

Die Engine zeichnet sich durch ihre hohe Flexibilität und Modularität aus. Dadurch können zukünftige Algorithmen einfacher integriert werden. Die Forecasting-Engine liefert folgende Vorteile:

Erhöhte Genaugigkeit: Manuelle Vorhersagen wurden im Durchschnitt um 10 Prozentpunkte übertroffen. Dadurch können unnötige Lagerkosten reduziert und fehlende Kapazitäten verringert werden.

Zeitersparnis: Die monatliche Planungszeit wurde von Wochen auf Stunden reduziert.

Breite Anwendbarkeit: Die Engine kann Vorhersagen für über 500 Produkte in den Kanälen Online, Einzelhandel und Großhandel im gesamten europäischen Markt erstellen.

Unterstützung für Cold-Start-Produkte: Neue Produkte, für die keine historischen Daten vorliegen, lassen sich auf Basis von Produktbildern und Stammdaten existierender Produkte prognostizieren.

Skalierbarkeit: Dank einer vollständig skalierbaren Cloud-Implementierung und Datenpipelines, die verteiltes Rechnen nutzen, kann die Lösung auf bis zu 20.000 Produkte erweitert werden.



Use Cases

Praxisbeispiel: Forecasting für einen optimalen Kapitaleinsatz

HERAUSFORDERUNG

Unser Kunde, ein großes Retail-Unternehmen, stand vor der Herausforderung, ein definiertes Budget für Rabatte optimal zu nutzen, um gleichzeitig Gewinne zu maximieren, Lagerbestände zu optimieren und die Absatzziele zu erreichen. Wir entwickelten eine Lösung, die auf einem Machine-Learning-Ansatz basiert, um die Preiselastizitäten der Produkte zu berechnen und zukünftige Absätze präzise vorherzusagen. Dadurch kann das Budget gezielt und effizient verteilt werden, selbst bei Produkten ohne historische Daten.

LÖSUNG

Wir entwickelten eine umfassende Lösung, die aus drei zentralen Komponenten besteht:

Preiselastizitäten: Basierend auf einem Machine-Learning Modell für die Modellierung der Wahrscheinlichkeit einer Bestellung werden die Preiselastizitäten ermittelt. Hierdurch kann ermittelt werden, wie stark der Absatz eines jeden Produktes auf Preisänderungen reagiert. Dies hilft, Produkte mit geringer Preissensitivität zu identifizieren, um dort höhere Preise oder niedrigere Rabatte anzusetzen, während preissensitive Produkte stärkere Rabatte erhalten.

Forecast: Eine zuverlässige Basis für die Optimierung wird durch die Prognose zukünftiger Absätze geschaffen. Ein Machine-Learning-Modell erstellt genaue Vorhersagen unter unveränderten Bedingungen, auch für neue Produkte ohne historische Daten, indem es auf Daten ähnlicher Vorgängerprodukte sowie Saisonalitäts- und Trendmuster zurückgreift.

Optimierung: Auf Grundlage der berechneten Preiselastizitäten und der prognostizierten Absätze wird das Budget optimal auf die Produkte verteilt. Überschreitungen des Budgets sind nicht erlaubt, während eine Unterschreitung angestrebt wird, um die gesetzten Ziele oft mit geringerem Budget zu erreichen.

ERGEBNIS

Die Lösung ermöglichte eine flexible Anpassung der Preise und Rabatte, die sowohl die Margen als auch den Absatz und die Lagerbestände deutlich verbesserte. Für die Optimierung wurden mehrere Ziele festgelegt, die je nach Bedarf eingesetzt werden können: Margen erhöhen, Absatzziele erreichen oder eine Mischung aus beiden. Die Lösung führt dabei zu mehreren Vorteilen:

Erhöhte Gewinnmargen: Gezielte Preisstrategien verbesserten die Marge bei Produkten mit geringer Preissensitivität, während der Absatz preissensitiver Produkte durch größere Rabatte signifikant gesteigert wurde, was zusammengenommen den Gesamtgewinn erhöhte.

Optimierte Lagerhaltung: Präzise Absatzprognosen ermöglichen eine bessere Lagerplanung, die Überbestände und Fehlmengen minimierte und die Kapitalbindung verbessert.

Effiziente Nutzung des Budgets: Durch die gezielte Verteilung der Rabatte konnte das Unternehmen seine Absatzziele erreichen und gleichzeitig das Budget effizienter nutzen.

Skalierbarkeit: Die Lösung ist vollständig in der Cloud implementiert und nutzt verteiltes Rechnen, um auf eine größere Anzahl von Produkten skaliert zu werden. Sie unterstützt auch neue Produkte, die keine historischen Daten haben, durch Analysen von Produktbildern und Stammdaten bestehender Produkte. Die technische Herausforderung, neue Produkte ohne historische Daten in die Prognosen einzubeziehen, gelang durch die Verwendung von Daten ähnlicher Vorgängerprodukte sowie Saisonalitäts- und Trendanalysen.



Use Cases

Praxisbeispiel: Liquiditätsprognose zur effizienten Zinsreduktion

HERAUSFORDERUNG

Ziel des Projekts war es, eine möglichst genaue Vorhersage der Liquidität (Ein- und Auszahlungen) auf verschiedenen Hierarchieebenen für mehrere Unternehmen, die zu einer Holding gehören, zu erstellen. Die Hypothese des Kunden war, dass menschliche Planungsexpert:innen bei der Schätzung des Cashflows einen gewissen Bias haben und tendenziell zu konservativ sind. Das bedeutet, dass sie den Inflow eher unterschätzen und den Outflow eher überschätzen. Ein datengetriebener und damit unvoreingenommener Machine-Learning-Ansatz sollte die tatsächlichen Werte genauer vorhersagen können.

LÖSUNG

Die Lösung bestand darin, einen Machine-Learning-basierten Ansatz zur Vorhersage der Liquidität zu entwickeln.

Datenklassifizierung und -modellierung: Die einzelnen Zeitreihen wurden aufgrund ihrer Heterogenität in Bezug auf Historie, Saisonalität und Varianz in verschiedene Kategorien klassifiziert. Anschließend wurden sie mit unterschiedlichen Modellierungsansätzen vorhergesagt. Bei kurzen Zeitreihen kamen naive Vorhersagen oder einfache statistische Modelle zum Einsatz. Bei ausreichender Datengrundlage wurde ein LightGBM-Modell verwendet, das mit exogenen Informationen angereichert wurde.

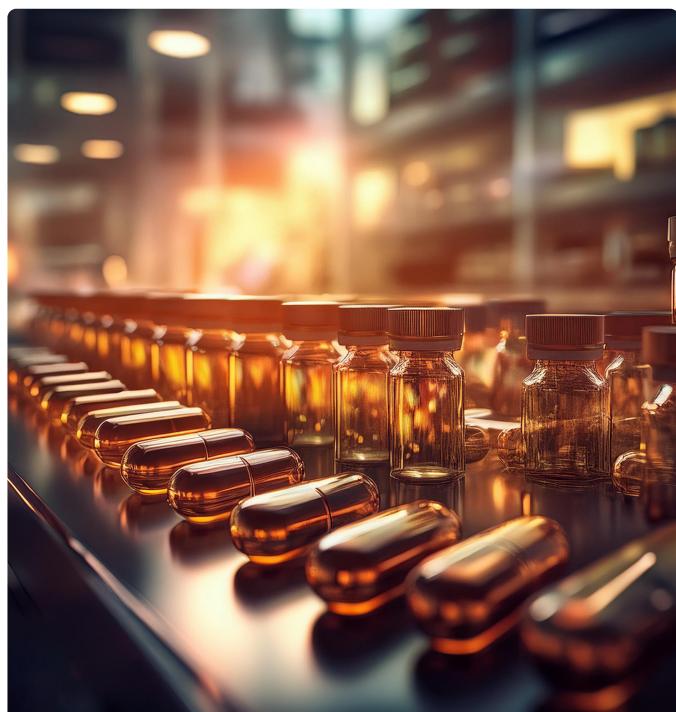
Hierarchische Konsistenz: Da die Zeitreihen hierarchisch strukturiert waren, war es wichtig, Konsistenz über die verschiedenen Aggregationsebenen hinweg zu gewährleisten. Dafür wurde der „Optimal Reconciliation“-Ansatz von Hyndman und Athanasopoulos (2014) genutzt. Bei diesem Ansatz werden alle Zeitreihen zunächst separat vorhergesagt und anschließend optimal harmonisiert, um die besten Ergebnisse zu erzielen.

ERGEBNIS

Die Evaluation zeigte, dass das Machine-Learning-Modell in den Unternehmen, in denen keine aktive Planung durch Expert:innen stattfand, regelmäßig bessere Vorhersagen lieferte als menschliche Expert:innen. In Unternehmen mit aktiver Expert:innenplanung schneidet das Modell hingegen schlechter ab. Diese Ergebnisse verdeutlichen, dass datengetriebene Lösungen nicht zwangsläufig die bestehenden menschlichen Ressourcen ersetzen sollen. Vielmehr sollen die Expert:innen durch empirisch fundierte Informationen unterstützt werden. Gleichzeitig kann der Mehrwert einer ML-basierten Vorhersage auf Unternehmen skaliert werden, in denen es keine spezialisierten Planungsexpert:innen gibt.

ROI UND ZINSKOSTENVERMEIDUNG

Der ROI des ML-basierten Ansatzes lässt sich durch die Vermeidung von Zinskosten ermitteln. Jeder Euro an vermeintlich notwendiger Liquidität, der durch die Expert:innenplanung bereitgestellt, aber nicht benötigt wurde, führte zu Zinsverlusten. Jeder Euro, der durch die genauere datenbasierte Vorhersage nicht bereitgestellt und letztlich auch nicht gebraucht wurde, resultierte in einer Vermeidung von Zinskosten. Diese Vermeidung wurde als KPI definiert und ermöglichte eine klare Bewertung des finanziellen Vorteils des neuen Ansatzes.



3

Herausforderungen: Organisationen

Herausforderungen

Aktuelle Herausforderungen für Organisationen

Viele Unternehmen bearbeiten Forecasting Use Cases isoliert in verschiedenen Abteilungen. Das führt zu sogenannten Silos, also einem System, das Mitarbeiter:innen, die zusammenarbeiten und kommunizieren sollten, strukturell voneinander trennt. Silos stellen Barrieren dar, die der teamübergreifenden Zusammenarbeit und Kommunikation im Weg stehen, die Effizienz mindern und den Informationsfluss behindern.

Ein typisches Beispiel: Absatzprognosen werden sowohl vom Vertrieb als auch von der Produktion erstellt. Diese redundante Arbeit ist ineffizient, verursacht inkonsistente Ergebnisse und erschwert den Gesamtüberblick. Wenn jede Abteilung eigene Technologien und Frameworks nutzt, erhöht sich zudem der Wartungsaufwand, was IT-Ressourcen unnötig belastet. Die durch Silos entstandene Fragmentierung erschwert insgesamt die Zusammenarbeit und führt

zu ineffizienten Prozessen und höheren Kosten. Es fehlt eine zentrale Datenbasis, was das Risiko von Dateninkonsistenzen und Fehlentscheidungen erhöht.

Unterschiedliche Datenquellen und deren unterschiedliche Interpretationen in den Abteilungen führen zu ungenauen Prognosen, die die strategische Planung beeinträchtigen. Unterschiedliche Technologien und Frameworks verursachen Kompatibilitätsprobleme und erschweren den Informationsaustausch. Der erhöhte Schulungsaufwand für Mitarbeiter:innen, die sich ständig an neue Systeme anpassen müssen, ist ein weiteres Problem. Deshalb sollten Unternehmen eine harmonisierte und zentralisierte Dateninfrastruktur schaffen, die Prognosen konsolidiert und eine einheitliche Datenübersicht bietet. Nur durch enge Zusammenarbeit und standardisierte Prozesse können Effizienz und Prognosequalität nachhaltig verbessert werden.

Weiter auf der nächsten Seite →

4

Forecasting-Modelle: Technik

Forecasting-Modelle

Technischer Überblick

Die Methoden, die zur Erstellung von Forecasting-Modellen eingesetzt werden können, sind vielfältig und reichen von klassischen statistischen Modellen über maschinelle Lernverfahren bis hin zu hochkomplexen Deep-Learning-Architekturen und den aktuell häufig diskutierten Foundation-Models. Jede dieser Methoden hat Stärken und Schwächen. Die Wahl der richtigen Methode hängt stark von den spezifischen Anforderungen des Anwendungsfalls und den verfügbaren Daten ab.

Zeitreihen Modelle

Klassische Zeitreihenmodelle wie ARIMA, SARIMA und ETS sind seit Langem bewährte Werkzeuge für die Vorhersage von univariaten Zeitreihen mit ausgeprägten saisonalen und trendbezogenen Mustern. Diese Modelle sind besonders wertvoll für Benchmarking und Anwendungen, bei denen die Vorhersage auf einer einzigen Variablen basiert. Ihre Stärken liegen in der Einfachheit und der Fähigkeit, klare und verständliche Ergebnisse zu liefern. Nachfolgend sind verschiedene Zeitreihenmodelle mit ihren Stärken aufgeführt. Hyndman und Athanasopoulos (2021) bieten für diese und weitere Zeitreihenmodelle detaillierte Ausführungen.

(S)ARIMA (AutoRegressive Integrated Moving Average):

ARIMA-Modelle kombinieren Autoregression (AR), Differenzierung (I) zur Stationarisierung und gleitende Durchschnitte (MA), um zeitliche Abhängigkeiten und Trends zu modellieren. SARIMA erweitert diese Methodik durch die Integration saisonaler Komponenten, um periodische Muster zu berücksichtigen.

ETS (Error, Trend, Seasonal):

ETS-Modelle umfassen eine Familie von Modellen, die Fehler, Trends und Saisonalität berücksichtigen. Diese

Methode verwendet eine additive oder multiplikative Kombination dieser drei Komponenten, um zeitliche Muster zu erfassen. Sie ist besonders nützlich für Daten mit klaren saisonalen und trendbezogenen Strukturen und ermöglicht eine flexible Anpassung an unterschiedliche Zeitreihencharakteristiken.

Neben den zwei klassischen Zeitreihenmodellen gibt es Zeitreihenmodelle, die eine höhere Komplexität bieten.

TBATS (Trigonometric Seasonal Components, Box-Cox Transformation, ARIMA Errors, Trend, Seasonal Components):

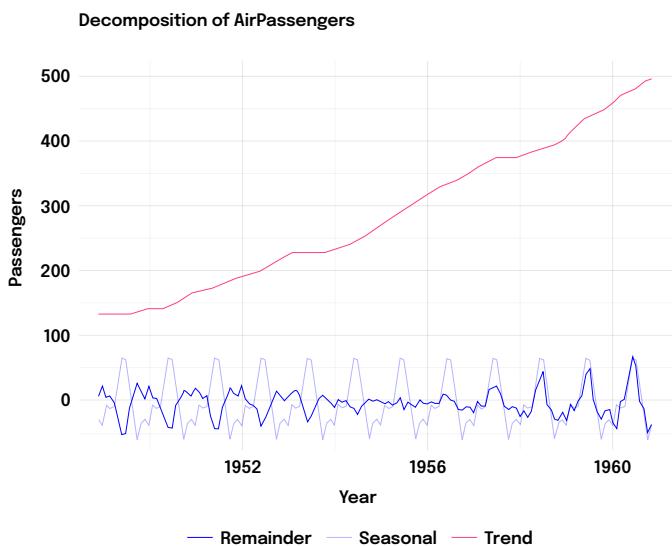
TBATS-Modelle sind besonders geeignet für komplexe saisonale Muster und lange saisonale Zyklen, die durch klassische Modelle schwer zu erfassen sind. Sie kombinieren mehrere Techniken: Box-Cox-Transformation zur Stabilisierung der Varianz, trigonometrische Funktionen zur Erfassung saisonaler Muster, ARMA-Modelle zur Modellierung von Kurzzeitabhängigkeiten und Trendkomponenten zur Berücksichtigung langfristiger Trends. TBATS bietet sich besonders für Zeitreihen an, die eine komplexe Saisonalität in Form multipler Saisonalitäten aufweisen.

Prophet:

Entwickelt von Meta, bietet Prophet eine robuste Lösung für Daten mit starken saisonalen Effekten und fehlenden Werten. Es verwendet eine additive Modellierung, bei der separate Komponenten für den Trend sowie für die saisonale und Feiertagseffekte definiert werden. Dies macht es besonders benutzerfreundlich und anpassungsfähig. Zudem können externe Features zusätzlich eingebunden werden. Das stellt einen Vorteil gegenüber den meisten anderen Zeitreihenmodellen dar, die Faktoren wie Wochentage, Feiertage oder Marketingkampagnen nicht berücksichtigen können.

Forecasting-Modelle

Bei statworx nutzen wir zu Beginn eines Projektes, in dem Zeitreihenprognosen eine Rolle spielen, meist klassische Zeitreihenmodelle und naive Methoden wie den Mittelwert oder die letzte bekannte Observation. Insbesondere in den frühen Phasen eines Projekts ermöglichen diese einfachen Ansätze eine rasche Einschätzung der Daten und liefern teils bereits überraschend gute Ergebnisse. Durch ihre Einfachheit und geringen Rechenanforderungen eignen sie sich hervorragend als Ausgangspunkt, bevor komplexere und rechenintensivere Modelle zum Einsatz kommen. So können wir schnell erste Vorhersagen treffen und die Modellierungsstrategie flexibel anpassen, ohne unnötig Zeit und Ressourcen zu investieren. Darüber hinaus integrieren wir diese klassischen Modelle teilweise in Ensemble-Modelle, um ihre Stärken optimal auszuschöpfen. Besonders bei Zeitreihen mit ausgeprägter Saisonalität bieten sie einen erheblichen Mehrwert: Für einen Teil der Modellierung bereinigen wir die Zeitreihen um die saisonale Komponente und modellieren die Saisonalität explizit mit einem Zeitreihenmodell.



Tree Based Modelle (ML)

Baumbasierte Modelle wie LightGBM, XGBoost und CatBoost sind seit vielen Jahren aufgrund ihrer Effizienz beliebt. Sie sind in der Lage, große Mengen an exogenen Features effizient zu verarbeiten und komplexe nicht-lineare Beziehungen zu modellieren. Diese Methoden sind besonders nützlich, wenn es darum geht, eine Vielzahl von Einflussfaktoren in die Vorhersage einzubeziehen. Zudem

bieten sie durch Techniken wie SHAP (Shapley-Werte) eine hohe Transparenz und Erklärbarkeit der Modellentscheidungen, was für das Vertrauen und die Akzeptanz der Modelle in der Praxis ein wichtiges Kriterium darstellt.

Grundkonzept eines baumbasierten Modells: Baumbasierte Modelle basieren auf Entscheidungsbäumen, bei denen die Daten in rekursiven Schritten in immer kleinere Teilmengen aufgeteilt werden. Jede Verzweigung („Node“) im Baum stellt eine Entscheidung auf Basis eines Merkmals dar, die zu einer weiteren Verzweigung führt, bis eine Endbedingung (Leaf Node) erreicht wird. Die Endknoten repräsentieren die Vorhersagewerte. Diese Struktur ermöglicht es den Modellen, komplexe nicht-lineare Beziehungen in den Daten abzubilden.

Gradient Boosting und Ensemble

- **Gradient Boosting:** Gradient Boosting-Algorithmen wie LightGBM und CatBoost bauen sequenziell mehrere Bäume auf, wobei jeder neue Baum die Fehler der vorhergehenden Bäume zu korrigieren versucht. Dies führt zu einem starken, zusammengesetzten Modell, das präzise Vorhersagen liefern kann.
- **Ensemble:** Im Gegensatz dazu erzeugt der Random-Forest-Algorithmus viele Entscheidungsbäume parallel, indem er zufällige Stichproben der Daten und Features verwendet. Die finale Vorhersage ergibt sich aus dem Durchschnitt oder der Mehrheit der individuellen Prognosen der einzelnen Bäume. Durch die zufällige Auswahl der Merkmale und das Sampling der Trainingsdaten entstehen unabhängige Fehlerquellen, die dazu beitragen, Overfitting zu vermeiden.

Bei der Entwicklung der Modelle haben Algorithmen basierend auf Gradient Boost diverse Vorteile: Gradient Boosting-Methoden wie LightGBM, XGBoost und CatBoost bieten im Vergleich zu Random Forest meist eine höhere Effizienz und bessere Performance, insbesondere wenn die Hyperparameter getuned werden. Die modernen GBM-Algorithmen integrieren fortschrittliche Funktionen, die ihre Skalierbarkeit und Effizienz verbessern. Zudem bietet CatBoost im Speziellen den Vorteil, dass kategoriale Features sehr effizient und nativ in die Modellierung eingebunden werden können.

Forecasting-Modelle

Beim Einsatz von baumbasierten Modellen ist allerdings Vorsicht geboten. Diese Modelle treffen Vorhersagen basierend auf Splits in den Daten, die durch Input-Features bestimmt werden. Das schränkt ihre Fähigkeit ein, Trends außerhalb der Trainingsdaten vorherzusagen. Um eine Extrapolation für baumbasierte Modelle dennoch zu gewährleisten, können folgende Strategien genutzt werden:

- **De-Trending:**

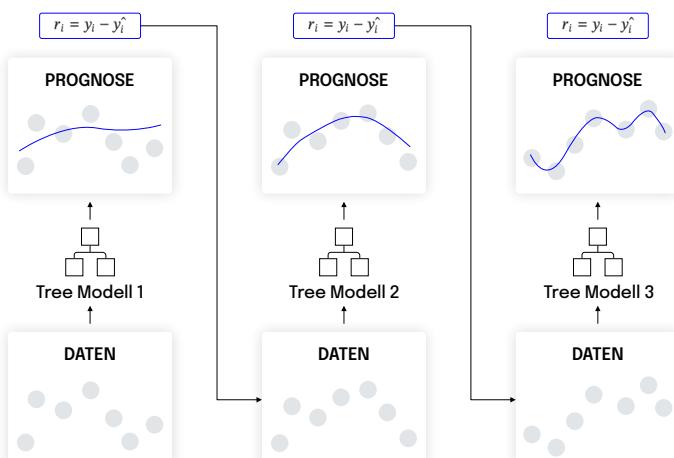
Das Entfernen des Trends aus der Time Series ermöglicht es dem Modell, sich auf zyklische und saisonale Muster zu konzentrieren. Der Trend wird separat geschätzt und nach den Vorhersagen des Modells wieder hinzugefügt. Dieser Ansatz ist leicht umsetzbar, hängt jedoch stark von der Güte der Trend-Schätzung ab.

- **Differencing:**

Hierbei wird die Differenz zwischen aufeinanderfolgenden Beobachtungen berechnet, um die Zeitreihe zu stabilisieren. Obwohl diese Methode simpel ist, kann übermäßiges Differencing zu Informationsverlust führen und eine Modellierung der Zeitreihe erschweren.

- **Verwendung der Linear Tree Method:**

Die Integration einer linearen Komponente in die Baumstruktur erlaubt es dem Modell, sowohl lineare als auch nichtlineare Muster zu erfassen und damit Extrapolationen des Trends zu ermöglichen, ohne die zu prognostizierende Zeitreihe zu transformieren. Dieser Ansatz ist modellspezifisch und kann die Komplexität erhöhen.



Deep Learning

Deep-Learning-Modelle haben das Potenzial, sehr komplexe und nicht-lineare Muster in den Daten zu erkennen und zu nutzen. Diese Modelle erfordern in der Regel große Datenmengen und eine aufwändige Implementierung, bieten jedoch erhebliche Vorteile, wenn es um die Vorhersage in hochdynamischen und komplexen Umgebungen geht. Die Modellarchitektur spielt hierbei eine entscheidende Rolle, und die Optimierung dieser Modelle erfordert spezialisierte Kenntnisse und Erfahrung.

Der Einsatz von Deep Learning in Forecasting ist dabei keineswegs neu. Erste Modelle wie Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN) und Long Short-Term Memory (LSTM) Modelle wurden bereits in den 1990er-Jahren entwickelt. In den letzten Jahren allerdings erfährt Deep Learning auch im Forecasting-Bereich immer mehr Aufmerksamkeit, was insbesondere auf die sogenannten Transformer-Modelle zurückzuführen ist. Durch den Einsatz einer speziellen Architektur, der sogenannten Attention, können die Modelle kontextuelle Abhängigkeiten in den Daten besser erfassen und verarbeiten.

Temporal Fusion Transformer (TFT): Der TFT ist eine spezialisierte Deep-Learning-Architektur, die entwickelt wurde, um Zeitreihendaten effizient zu verarbeiten und präzise Vorhersagen zu treffen, insbesondere in Szenarien mit komplexen und dynamischen Umgebungen (Lim et al. 2021). Die Architektur des TFT kombiniert mehrere Konzepte, unter anderem „Attention“-Module, um sowohl die zeitlichen Abhängigkeiten als auch die Beziehungen zwischen verschiedenen Features zu erfassen. Sie haben folgende technische Charakteristiken:

- **Multi-Head Attention:**

Um auch langfristige Abhängigkeiten zu erfassen, verwendet der TFT ein Multi-Head Attention-Modul. Dieses Modul ermöglicht es dem Modell, verschiedene Zeitpunkte in der Vergangenheit zu beachten und gleichzeitig die Wichtigkeit dieser Zeitpunkte für die aktuelle Vorhersage zu bestimmen. Multi-Head Attention bedeutet, dass das Modell mehrere Heads (parallele Aufmerksamkeitsebenen) nutzt, um unterschiedliche Aspekte der Daten zu analysieren. Diese parallele Betrachtung ermöglicht es, sowohl detaillierte lokale Muster als auch breitere, globale Abhängigkeiten über längere Zeiträume hinweg zu erfassen.

Forecasting-Modelle

- **Gating-Mechanismen:**

Der TFT verwendet spezialisierte Gating-Mechanismen, um den Informationsfluss durch das Netzwerk zu steuern. Einer der Schlüsselmechanismen ist das „Variable Selection Network“-Gate, das in jedem Schritt bestimmt, welche Eingabeveriablen für die Vorhersage relevant sind. Dies sorgt dafür, dass das Modell nicht von irrelevanten oder rauschenden Features überwältigt wird, sondern sich auf die wesentlichen Informationen konzentrieren kann.

- **LSTM Layers:**

Ein zentraler Bestandteil der TFT-Architektur sind die Long Short-Term Memory (LSTM) Layers, die dazu dienen, zeitliche Abhängigkeiten in den Daten zu erfassen. LSTMs sind besonders gut darin, langfristige Abhängigkeiten zu lernen, indem sie Informationen über vorherige Zeitschritte speichern und für zukünftige Vorhersagen nutzen. Im TFT werden diese LSTM Layers verwendet, um die Dynamik in den Zeitreihendaten zu modellieren, sowohl für vergangene als auch für zukünftige Zeitpunkte.

- **Static Covariate Encoders:**

Statische Merkmale, oder Static Covariates, sind Merkmale, die sich über die Zeit hinweg nicht verändern, wie zum Beispiel die geografische Lage, manche demografische Informationen oder bestimmte Eigenschaften eines Produkts. Diese Merkmale sind oft entscheidend für die Modellierung, da sie wichtige kontextuelle Informationen bereitstellen, die die dynamischen Prozesse in den Zeitreihen beeinflussen können. Im TFT werden diese statischen Merkmale durch sogenannte Static Covariate Encoders in das Modell integriert. Diese Encoder transformieren die statischen Merkmale in sogenannte Kontextvektoren, die das gesamte Netzwerk beeinflussen. Diese Kontextvektoren dienen als Bedingung für die zeitlichen Dynamiken im Modell.

Inzwischen gibt es viele weitere Transformer-Modelle, die sich die Eigenschaften der Architektur zu Nutze machen, um mit den Schwierigkeiten im Forecasting umzugehen:

- **Informer** wurde für die effiziente Vorhersage langer Sequenzen und Zeitreihen entwickelt und führt eine spezielle Form der Self-Attention ein, um die Rechenkomplexität zu reduzieren. Durch Self-Attention-Distillation und ein generatives Decoder-Modul, können lange

Sequenzen in einem einzigen Durchgang vorhergesagt werden (Zhou et al., 2021).

- **TimesNet** behandelt Zeitreihen als eine Kombination aus saisonalen Mustern und nutzt die „frequency-temporal Attention“, um die Zeitreihen zu zerlegen und Abhängigkeiten auf mehreren Ebenen zu erfassen (Wu et al., 2022).
- **PatchTST** adaptiert Vision-Transformer-Konzepte für die Zeitreihenvorhersage, indem es die Eingabesequenzen in nicht überlappende Bereiche (Patches) unterteilt. Komplexere multivariate Daten werden pro Kanal unabhängig verarbeitet. Zusätzlich kann das Modell durch selbstüberwachtes Pretraining trainiert werden (Nie et al., 2022).
- **Crossformer** ermöglicht die Erfassung kurzfristiger und langfristiger Abhängigkeiten in Zeitreihendaten durch segmentweise Self- und Cross-Scale Attention. Durch einen Multiscale-Encoder werden Eingaben auf verschiedenen zeitlichen Ebenen verarbeitet, was eine effiziente Verarbeitung sowohl lokaler als auch globaler Muster ermöglicht (Wang et al., 2021).
- **iTransformer** umfasst die Nutzung von Attention und Feed-Forward-Networks auf invertierte Dimensionen. Dabei werden die Zeitpunkte einzelner Zeitreihen in Variate-Tokens embedded, die vom Attention-Mechanismus genutzt werden, um multivariate Korrelationen zu erfassen. Gleichzeitig wird das Feed-Forward Network auf jeden Variate-Token angewendet, um nichtlineare Repräsentationen der Zeitreihen zu lernen (Liu et al., 2024).

Darüber hinaus bieten Modelle wie DLinear, N-BEATS und NHITS Architekturen mit geringerer Komplexität, die sowohl flexibel als auch effizient sind. Besonders bei weniger komplexen Daten können sich diese Modelle als geeignete Alternativen zu klassischen Zeitreihenmodellen sowie zu den zuvor genannten Architekturen erweisen.

Ein Vorteil von Deep-Learning-Methoden ist ihre inhärente Fähigkeit, Features automatisch zu extrahieren und zu transformieren. Dies reduziert den Bedarf an manuellem Feature-Engineering, das bei traditionellen ML-Methoden oft notwendig ist. Durch den Einsatz von Techniken wie

Forecasting-Modelle

Batch Normalization, Dropout und anderen Regularisierungsmethoden können diese Modelle zudem Überanpassungen (Overfitting) effektiv entgegenwirken und eine bessere Generalisierung auf unbekannte Daten erreichen.

Deep-Learning-Modelle sind in der Regel komplexer und erfordern eine weitaus größere Rechenleistung sowie längere Trainingszeiten. Diese Komplexität erhöht nicht nur den Aufwand bei der Implementierung, sondern auch bei der Modellwartung und -optimierung. Zudem sind Deep-Learning-Modelle oft empfindlicher gegenüber Overfitting, insbesondere wenn sie auf kleinen oder unzureichend repräsentativen Datensätzen trainiert werden. Während oben genannte Techniken wie Dropout und Regularisierung helfen können, bleibt die Gefahr bestehen, dass das Modell die Trainingsdaten zu stark anpasst und dadurch auf neuen, unbekannten Daten schlechter performt.

In der Praxis können Deep-Learning-Modelle vor allem dann genutzt werden, wenn die Komplexität der Datenstruktur traditionelle baumbasierte Ansätze überfordert oder wenn die Vorhersagegenauigkeit bei volatilen und dynamischen Datensätzen optimiert werden muss. Trotz der jüngsten Erfolge von Transformer-basierten Modellen im Bereich von Zeitreihen-Forecasts, zeigt die Forschung, dass sogar einfache Architekturen in Form eines One-Layer Linear-Models komplexe Transformer-Modelle outperformen können. Diese überraschenden Ergebnisse von Zeng et al. (2023) legen nahe, dass die zeitlichen Beziehun-

gen in den Daten in manchen Fällen möglicherweise besser von weniger komplexen Modellen erfasst werden können. Insbesondere in Kombination mit Transfer-Learning und fortschrittlichen Optimierungsalgorithmen können diese einfacheren Modelle maßgeschneiderte Lösungen für eine Vielzahl von Anwendungsfällen bieten. Aus diesem Grund sollte die Entscheidung für Deep-Learning-Modelle sorgfältig getroffen werden, da der tendenziell höhere Aufwand in der Implementierung nicht zwangsläufig zu einer besseren Lösung führen muss.



Weiter auf der nächsten Seite →

Forecasting-Modelle

Foundation Models

Foundation-Modelle stellen einen relativ neuen Fortschritt im maschinellen Lernen dar. Derzeit gelten Sie als eines der intensivsten Forschungsgebiete im Bereich der Vorhersage. Ein prominentes Beispiel ist ChatGPT, das wahrscheinlich bekannteste große Sprachmodell (LLM). Diese Large-Time-Series-Models (LTSMS) zeichnen sich durch ihre enorme Leistungsfähigkeit aus, da sie auf riesigen Datenmengen trainiert werden und somit eine Vielzahl von Aufgaben bewältigen können. Im Bereich der Vorhersage bieten sie den Vorteil, dass sie sowohl für die Absatzprognose von Autos als auch für die Auslastungsprognose in der Produktion gleichzeitig eingesetzt werden können. Der wesentliche Unterschied zu anderen Modellklassen besteht darin, dass sie direkt genutzt werden können, ohne dass sie separat und oft sehr zeitaufwendig trainiert werden müssen.

Die zugrundeliegende Modellarchitektur basiert aktuell meist auf Transformern, einer Deep-Learning-Technologie, ähnlich wie bei den großen Sprachmodellen. Das Potenzial dieser Modellklasse ist gewaltig: Unternehmen können in Sekunden eigene Prognosen erstellen, ohne dafür Features zu entwickeln, Modelle zu trainieren, zu evaluieren und auszuwählen. Diese Effizienz und Vielseitigkeit machen Foundation-Modelle zu einer vielversprechenden Technologie. Ein zusätzlicher Vorteil ist, dass Unternehmen oft im Kontext von Zeitreihen nur über kleine Datenmengen verfügen, sodass die Vorteile von Deep-Learning-Forecasting-Modellen häufig ungenutzt bleiben und sie auf andere Modellklassen zurückgreifen müssen.

Aktuell gibt es einige Anbieter dieser Modelle. [Nixtla's TimeGPT-1](#) (NIXTLA, 2024) ist dabei das erste kostenpflichtige Forecasting Foundation Model. Allerdings gibt es auch open-source Alternativen wie bspw. [TimesFM](#) von Google (Das et al., 2023) oder [MOMENT](#) der Kooperation von der University of Pennsylvania und der Carnegie Mellon University (Goswami et al., 2024). Die Anwendung dieser Modelle ist simpel, Datensätze können in einer gewissen Struktur samt Einstellungsparameter an eine Schnittstelle geschickt werden, die daraufhin eine Vorhersage und zum Teil sogar Unsicherheitsintervalle als Antwort ausgibt. Bei einigen Modellen ist der Vorhersagehorizont frei wählbar, sprich es kann beliebig weit in die Zukunft vorhergesagt werden. Natürlich sinkt die Vorhersagequalität bei längeren Horizonten – so wie bei allen anderen Modellklassen auch.

Ähnlich wie LLMs lassen sich auch LTSMS finetunen. Das heißt, das Modell kann nicht nur angewendet werden, sondern auch auf die eigenen Daten angepasst und trainiert werden. LLMs können z. B. auf die firmeneigene Kommunikation (z. B. Corporate Language) und Fachterminologie angepasst werden. Im Zusammengang mit LTSMS bezieht sich dies auf spezifische Merkmale der Daten eines Unternehmens, wie zum Beispiel die unterschiedlichen saisonalen Verkaufszyklen in der Modebranche. Anders als bei großen Sprachmodellen ist das Finetuning kostengünstiger, da LTSMS deutlich weniger komplex sind als LLMs. Garza und [Mergenthaler-Canseco \(2023\)](#) haben gezeigt, dass die Vorhersagegüte bei einem Finetuning im Vergleich zur reinen Anwendung der LTSMSs enorm gesteigert werden kann.

Zeitreihen enthalten oft äußerst sensible Unternehmensdaten, weshalb die Nutzung von Foundation-Modellen intern abgestimmt werden muss. Eine sichere Nutzung dieser Modelle kann durch Hosting in der eigenen Cloud oder on-premise gewährleistet werden. So bleiben die Daten stets in der vertrauten Arbeitsumgebung und werden sicher innerhalb der eigenen Infrastruktur verarbeitet.

Foundation-Modelle für die Zeitreihenvorhersage stellen eine benutzerfreundliche und vielversprechende Alternative zu traditionellen Modellklassen dar. Durch die zunehmende Anzahl von Anbietern werden mehr Foundation-Modelle auf den Markt kommen, was in der Zukunft zu einer fortlaufenden Verbesserung der Qualität und Leistung führen wird. LTSMS eignen sich als Referenzmodelle, da diese mit geringem Aufwand implementiert werden können und eine hohe Vorhersagequalität liefern. Allerdings ist es wichtig zu betonen, dass sie keineswegs ein Allheilmittel für alle Anwendungsfälle im Bereich der Zeitreihenvorhersage darstellen. Sie sollten stets mit klassischen statistischen Methoden verglichen werden, da diese ebenfalls oft mit geringem Aufwand implementiert werden können.

Trotz ihrer Vielseitigkeit und Leistungsfähigkeit erfordert der Einsatz von LTSMS eine sorgfältige Datenvorbereitung. Die Qualität und Konsistenz der Daten sind und bleiben die entscheidenden Faktoren für die Güte der Vorhersagen. Zudem sollte man die Interpretierbarkeit der Modelle nicht außer Acht lassen. Während LTSMS leistungsstarke Werkzeuge sind, bleibt die Transparenz der Entscheidungsprozesse oft begrenzt. Daher kann es sinnvoll sein, ergänzend erklärbare Modelle einzusetzen, um Einblicke in die zugrunde liegenden Muster und Beziehungen innerhalb der Daten zu gewinnen.

5

Herausforderungen: Implementierung

Herausforderungen

Typische Herausforderungen bei der Implementierung

Der Bedarf nach servicebasierten Forecasting-Lösungen, die Nutzer:innen auf Knopfdruck Vorhersagen auf Basis eines simplen Dateninputs liefern, wächst. Viele etablierte Dienstleister bieten dafür Standardlösungen. Doch je nach Branche und Projekt übersteigt die Komplexität des Forecasting-Bedarfs die Fähigkeiten solcher Lösungen. Damit Vorhersagen ihren vollen Mehrwert entfalten können, müssen komplexe Business-Prozesse mitberücksichtigt werden: Beispielsweise müssen Lieferzeiten und Beschaffungsprozesse bei der Absatzprognose einbezogen werden. Ein prognostizierter Absatz sollte realistisch im Hinblick auf die Lieferfähigkeit sein, um Engpässe oder Überbestände zu vermeiden. Das bedeutet, dass Vorhersagen mit Produktions- und Bestellzyklen abgestimmt sein müssen, um sicherzustellen, dass die benötigten Produkte rechtzeitig verfügbar sind. Auch die Art des Produktes hat entscheidenden Einfluss auf das Forecasting: Wie erstellt man eine Vorhersage für ein komplett neues Produkt ohne Historie („Coldstart“)? Oder im umgekehrten Fall: Welche Auswirkung hat ein Produktionsstopp und eine daraus resultierende kontinuierliche Reduzierung des Restbestands auf die Absatzentwicklung dieses Produkts und inwiefern müssen „rohe“ Modell-Prognosen dahingehend korrigiert werden?

Das sind nur einige der vielen Fragen, die eine hochwertige, nachhaltige Forecasting-Lösung berücksichtigen muss. Deshalb bieten Dienstleistungen, die auf numerische Vorhersagen spezialisiert sind, allenfalls einen Teil der Lösung. Forecasting hat grundsätzlich zwei Anwendungarten. Entweder ist der Forecast selbst die Lösung,

die in einen der oben skizzierten Business-Prozesse eingebettet ist und einen eigenen Mehrwert liefert. Oder der Forecast ist ein Mittel zum Zweck in einem anderen datengetriebenen Use Case. Typische Beispiele hierfür sind die Erreichung von Absatzzielen oder die Optimierung von Lagerbeständen über gezielte Preisanpassungen (Intelligent Pricing).

Bei der Implementierung einer umfassenden Lösung, die einen spürbaren Mehrwert liefert, stehen Unternehmen oft vor spezifischen Herausforderungen. Diese können sowohl durch externe Faktoren als auch durch die Anforderungen der Lösung selbst bedingt sein. Typische Herausforderungen treten in Projekten immer wieder auf. Im Folgenden beleuchten wir die häufigsten Probleme und wie diese gelöst werden können.

Exogene Schocks wie Pandemien und Lieferkettenprobleme

Kernproblematik: Unvorhersehbare externe Ereignisse wie die COVID-19-Pandemie oder Störungen in den Lieferketten können die Genauigkeit von Forecasting-Modellen erheblich beeinträchtigen. Derartige Schocks sind durch die verfügbaren Daten selten prognostizierbar und führen zugleich zu abrupten und oft drastischen Veränderungen in den Daten. Diese Strukturumbrüche, die nicht dem gewöhnlichen Marktumfeld entsprechen, können zu erheblichen Verzerrungen bei den trainierten Modellen sorgen, wenn sie nicht entsprechend berücksichtigt werden.

Herausforderungen

Lösungsansatz: Das Problem exogener Schocks kann mittels der Modellauswahl sowie der Datenverarbeitung selbst adressiert werden, um negative Auswirkungen auf die Performance bestmöglich zu kompensieren.

- **Datenaufbereitung:** Bei der Datenverarbeitung ist es zunächst erforderlich, dass alle Anomalien und Zeitabschnitte, die mit exogenen Schocks in Verbindung stehen, identifiziert sind. Das kann einerseits durch Expert:innen und qualitativen Ansätzen erfolgen. Insbesondere im Kontext von Corona sind die Zeiträume in den Daten meist klar erkennbar. Andererseits können Algorithmen genutzt werden, um Anomalien und Strukturumbrüche zu identifizieren. Nach der Identifikation müssen die betroffenen Daten entweder korrigiert oder aus dem Trainingssatz entfernt werden. Eine Möglichkeit besteht darin, diese Daten durch simulierte oder geschätzte Werte zu ersetzen, die die normalen Bedingungen widerspiegeln.
- **Modellauswahl:** Der Einsatz robuster Modelle, die gegenüber Ausreißern und Anomalien widerstandsfähig sind, ist besonders hilfreich, um die Auswirkungen exogener Schocks zu minimieren. Baumbasierte Modelle wie LightGBM und Random Forest bieten Mechanismen, um mit Anomalien umzugehen und die Modellgenauigkeit zu verbessern. LightGBM verwendet einen Leaf-wise Growth Algorithmus und umfangreiche Regularisierungsoptionen, die Überanpassung und die Auswirkungen von Ausreißern kontrollieren. Random Forest reduziert die Empfindlichkeit gegenüber Ausreißern durch Bootstrap Aggregation und Feature Randomness, indem es mehrere Entscheidungsbäume auf zufälligen Datenproben trainiert und deren Ergebnisse aggregiert.

Produkteinführungen

In vielen Branchen werden regelmäßig neue Produkte eingeführt, für die Prognosen erstellt werden sollen. Da für neue Produkte natürlich noch keine Historie vorliegt, können diese Produkte nicht explizit beim Modelltraining berücksichtigt werden. Um dennoch eine Prognose für derartige Produkte zu gewährleisten, bieten sich die folgenden Ansätze besonders gut an:

- **Naive Ansätze:** Insbesondere für Bereiche, in denen es Vorgängerprodukte gibt, bieten sich naive Ansätze

an. Auf Basis der Vorgängerinformation kann so auf das Absatzniveau des künftig eingeführten Produktes geschlossen werden. Saisonale Komponenten und die Absatzentwicklung verganglicher Ramp-Ups können zusätzlich einen inhaltlichen Mehrwert für die Prognose bieten.

- **Transfer Learning:** Auch wenn ein Training mit expliziter Berücksichtigung zukünftiger Produkte nicht möglich ist, können bestehende Modelle genutzt werden, um Prognosen für neue Produkte zu erstellen. Dabei greift man auf Modelle und Features zurück, die für ähnliche Produkte oder Vorgängern genutzt werden.

In beiden Ansätzen wird davon ausgegangen, dass Rückschlüsse auf neue Produkte durch die Analyse verwandter oder ähnlicher Produkte gezogen werden können. Es ist jedoch wichtig zu beachten, dass Prognosen für neue Produkte grundsätzlich mit höherer Unsicherheit verbunden sind als für bereits etablierte Produkte mit umfangreicher Historie. Ein wesentlicher Erfolgsfaktor für diese Methoden ist die rechtzeitige Verfügbarkeit von Informationen über bevorstehende Produkteinführungen.

Erklärbarkeit

In vielen Projekten ist es entscheidend, dass Prognosen nachvollziehbar und verständlich sind. Da komplexe Modelle oft als Black Box wahrgenommen werden, besteht das Risiko, dass ihre Ergebnisse auf geringe Akzeptanz stoßen. Um die Akzeptanz und Nachvollziehbarkeit der Prognosen zu erhöhen, ist es daher wichtig, Wege zur „Erklärbarkeit“ der Modelle und ihren Vorhersagen zu finden. Für dieses Ziel gibt es zwei mögliche Ansätze. Erstens kann das Modell allgemein erklärt werden. Dabei wird untersucht, welche Informationen für das Modell wichtig sind und wie diese Informationen die Vorhersagen beeinflussen. Zweitens können einzelne Vorhersagen erklärt werden. Dabei wird gezeigt, wie jede Information die jeweilige Vorhersage beeinflusst hat. Im Folgenden werden einige der gängigen Methoden genauer beschrieben:

- **SHAP (SHapley Additive exPlanations):** SHAP ist eine Methode, die auf den Shapley-Werten aus der Spieltheorie basiert (Lundberg & Lee 2017). Sie berechnet den Beitrag jedes Features zu einer Vorhersage, indem sie immer wieder einzelne Merkmale verschleiert und analysiert, wie das Modell da-

Herausforderungen

rauf reagiert. Dabei bietet SHAP eine detaillierte Erklärung sowohl auf individueller (local) als auch auf aggregierter (global) Ebene. Zudem berücksichtigt SHAP Interaktionen zwischen Features, wodurch es zu stabilen Erklärungen kommt. Insbesondere für baumbasierte Modelle wie LightGBM und CatBoost hat sich SHAP als effizienter Standard etabliert.

- **LIME (Local Interpretable Model-agnostic Explanations):**

LIME erstellt lokale Erklärungen, indem es ein einfaches, interpretiertes Modell (z. B. lineare Regression) um die Vorhersage eines komplexen Modells herum anpasst. Das hilft, die Vorhersage für individuelle Datenpunkte verständlicher zu machen. LIME kann auf eine Vielzahl von Modellen angewendet werden, einschließlich komplexer Modelle wie Deep-Learning-Netzwerke und baumbasierte Modelle. Im Gegensatz zu SHAP ist LIME schneller und weniger rechenintensiv, jedoch ist strittig, ob LIME zuverlässig wahre Erklärungen produziert, da es weniger mathematisch fundiert ist als SHAP.

Permutation Feature Importance: Die Permutation Feature Importance ist eine Technik, die die Bedeutung eines Features für ein Modell anhand eines Scores abbildet. Dabei wird die Bedeutung eines Features durch die Veränderung der Modellleistung gemessen, wenn die Werte dieses Features zufällig permutiert werden. Diese Methode bietet eine globale Erklärung des Modells, indem sie zeigt, welche Features den größten Einfluss auf die Vorhersagen haben.

Die Permutation Feature Importance ist sehr effizient, da sie ohne tiefgreifende Änderungen am Modell durchgeführt werden kann. Allerdings basiert diese Methode auf der Annahme, dass die Features unabhängig voneinander sind, was zu Fehleinschätzungen der tatsächlichen Bedeutung eines Features führen kann. Wenn Interaktionen zwischen Features bestehen oder eine hohe Korrelation vorliegt, kann die ermittelte Feature Importance irreführend sein. Daher sollte sie als erste Indikation verwendet werden, während für fundierte Aussagen Methoden herangezogen werden sollten, die Abhängigkeiten zwischen den Features berücksichtigen.

Partial Dependence Plots (PDPs): PDPs zeigen den Einfluss eines oder mehrerer Features auf die Vorhersage,

während andere Features konstant gehalten werden. Der Vorteil von PDPs besteht darin, dass eine globale Sicht darauf geboten wird, wie sich Änderungen eines Features auf die Modellvorhersagen auswirken. Es ist jedoch wichtig zu beachten, dass die Grundlage der PDPs eigentlich Individual Conditional Expectation (ICE) Plots sind, die eine lokale Sichtweise bieten. Beide Arten von Erklärungen – global durch PDPs und lokal durch ICE Plots – sind möglich. Zudem sind oft Accumulated Local Effects (ALE) den PDPs vorzuziehen, da sie Feature-Interaktionen besser berücksichtigen.



Local Explainability:

Diese Erklärungen werden für einzelne Vorhersagen oder Datenpunkte erstellt. Hierzu gehören Methoden wie SHAP und LIME, die detaillierte Einblicke geben, wie einzelne Merkmale die Vorhersage für einen spezifischen Fall beeinflussen. Diese Erklärungen sind besonders hilfreich, wenn es darum geht, einzelne Entscheidungen des Modells zu verstehen.

Global Explainability:

Diese Erklärungen beschreiben die gesamte Funktionsweise des Modells. Das umfasst allgemeine Einsichten darüber, welche Merkmale im Modell wichtig sind (z. B. Permutation Feature Importance) und wie sich Merkmale im Durchschnitt auf die Vorhersagen auswirken (PDPs). Diese Erklärungen sind hilfreich, um ein allgemeines Verständnis des Modells zu erhalten.

Hierarchische Forecasts

Hierarchische Forecasts sind Prognosen, die in einer hierarchischen Struktur organisiert sind, bei der Vorhersagen für verschiedene Ebenen einer Hierarchie, wie z. B. geografische Regionen oder Produktkategorien, durchgeführt werden. Diese Art von Forecasts ist besonders relevant für Organisationen, die ihre Prognosen auf verschiedenen Aggregationsebenen benötigen, z. B. natio-

Herausforderungen

nale Verkäufe, regionale Verkäufe und lokale Verkäufe. Die Herausforderung besteht darin, sicherzustellen, dass die Vorhersagen auf allen Ebenen der Hierarchie konsistent und kohärent sind. Es gibt verschiedene Ansätze, um diese hierarchischen Forecasts zu erstellen, von klassischen Ansätzen wie Bottom-Up- und Top-Down bis hin zu Ansätzen wie Optimal Reconciliation.

- **Bottom-Up:** Der Bottom-Up-Ansatz beginnt mit der Vorhersage auf der untersten Ebene der Hierarchie und aggregiert diese Vorhersagen, um Vorhersagen für höhere Ebenen zu erzeugen. Dies ermöglicht eine detaillierte Betrachtung der unteren Ebenen, kann jedoch zu inkonsistenten Vorhersagen auf höheren Ebenen führen.
- **Top-Down:** Im Gegensatz zum Bottom-Up-Ansatz startet der Top-Down-Ansatz auf der obersten Ebene der Hierarchie und verteilt die Vorhersagen auf die unteren Ebenen basierend auf historischen Anteilen oder anderen Methoden. Dies sorgt für Konsistenz auf den oberen Ebenen, kann jedoch die Genauigkeit auf den unteren Ebenen beeinträchtigen.
- **Middle-Out:** Dieser Ansatz ist ein hybrider Ansatz, der Bottom-Up- und Top-Down kombiniert. Hierbei wird eine mittlere Ebene der Hierarchie als Anker gewählt. Vorhersagen über dieser Ebene werden mithilfe der Bottom-Up-Methode berechnet, während die unteren Ebenen die Top-Down-Methode verwenden. Diese Methode ist besonders nützlich, wenn die mittlere Ebene eine gute Balance zwischen detaillierten und aggregierten Daten bietet.
- **Optimal Reconciliation:** Der Ansatz strebt die bestmögliche Konsistenz zwischen den verschiedenen Hierarchieebenen an. Diese Methode, auch als Optimal Combination bekannt, verwendet einen verallgemeinerten „generalized least squares estimator“ unter Berücksichtigung der Kovarianzmatrix der Kohärenzfehler. Dies gewährleistet, dass die endgültigen Vorhersagen sowohl auf unteren als auch auf höheren Ebenen kohärent und optimal kombiniert werden. Dieser Ansatz wurde von Hyndman et al. (2011) vorgeschlagen und wird als eine der effektivsten Methoden zur Konsolidierung hierarchischer Forecasts angesehen.

Unsicherheit von Prognosen

Punktschätzungen, bei denen ein einzelner Wert als Vorhersage angegeben wird, können irreführend sein, da sie die Unsicherheit in den Daten und Modellen nicht berücksichtigen. Dies ist besonders problematisch in dynamischen und komplexen Umgebungen, in denen viele unbekannte Faktoren die Vorhersage beeinflussen können. Ein einzelner Punktwert kann zu einer falschen Sicherheit führen, was zu suboptimalen Entscheidungen führt. Die Berücksichtigung der Unsicherheit in Vorhersagen bietet erhebliche Vorteile. Anstatt nur einen einzigen Wert zu prognostizieren, liefern Intervallprognosen eine Bandbreite möglicher Ergebnisse. Dies ermöglicht es Entscheidungsträgern, das Risiko besser zu managen und fundiertere Entscheidungen zu treffen, insbesondere in Szenarien mit hoher Volatilität oder Unsicherheit.

Es gibt verschiedene Methoden, um Unsicherheit in Vorhersagen zu integrieren:

- **Konfidenzintervalle:** Bei der Nutzung von Zeitreihenmodellen wie ARIMA oder ETS stehen Konfidenzintervalle standardmäßig zur Verfügung. Diese Intervalle quantifizieren die Unsicherheit der Vorhersage, indem sie einen Bereich definieren, in dem der zukünftige Wert mit einer bestimmten Wahrscheinlichkeit (z. B. 95%) liegen wird.
- **Quantilregression:** Bei baumbasierten Modellen wie XGBoost kann die Unsicherheit durch die Anwendung von Quantilregressionen modelliert werden. Anstatt nur eine Punktprognose (Median) zu liefern, berechnet das Modell Vorhersagen für verschiedene Quantile, z. B. das 5%- und 95%-Quantil. Diese Quantile geben dann die Grenzen eines Intervalls an, das die Unsicherheit der Prognose widerspiegelt.
- **Monte Carlo Dropout:** Für Deep-Learning-Modelle kann eine Intervallprognose durch die Nutzung von Monte Carlo Dropout erzielt werden. Hierbei wird während des Vorhersageprozesses zufällig ein Teil der Neuronen deaktiviert (Dropout), was mehrere verschiedene Vorhersagen ermöglicht. Die Verteilung dieser Vorhersagen gibt dann Aufschluss über die Unsicherheit. Das Modell liefert somit eine Verteilung von Ergebnissen anstelle eines einzelnen Punktwerts.

Fazit & Ausblick

Mehrwert auf mehreren Ebenen

Forecasting bietet Unternehmen ein enormes Potenzial, strategische Vorteile zu realisieren und operative Effizienz zu steigern. In einer zunehmend datengetriebenen Welt, in der die Verfügbarkeit von Informationen exponentiell wächst, wird die Fähigkeit, präzise Vorhersagen zu treffen, zu einem entscheidenden Wettbewerbsfaktor. Unternehmen, die Forecasting effektiv einsetzen, können ihre Ressourcen optimal nutzen, Risiken besser managen und ihre Entscheidungsfindung auf eine fundierte, datengestützte Basis stellen.

Der wahre Wert von Forecasting liegt jedoch nicht nur in der Fähigkeit, zukünftige Entwicklungen vorherzusagen, sondern auch in der Flexibilität, die es Unternehmen bietet. Egal, ob es darum geht, Liquiditätsengpässe frühzeitig zu erkennen, Absatzprognosen für eine gezielte Marketingstrategie zu erstellen oder die gesamte Supply Chain effizienter zu gestalten – Forecasting kann in nahezu jedem Bereich eines Unternehmens signifikanten Mehrwert schaffen. Es ist ein Werkzeug, das sowohl direkten als auch indirekten Nutzen generieren kann, sei es durch unmittelbare Einsparungen oder durch langfristige Prozessoptimierungen.

Jedoch ist es wichtig zu verstehen, dass es keine „One-Size-Fits-All“-Lösung im Forecasting gibt. Jede Organisation ist einzigartig und hat einzigartige Anforderungen, Ziele und Herausforderungen. Die Wahl der richtigen Forecasting-Methode hängt von einer Vielzahl von Faktoren ab, wie der Verfügbarkeit und Qualität der Daten, den spezifischen Geschäftsanforderungen und den zu erwartenden externen Einflüssen. Standardlösungen können oft nicht die notwendige Präzision und Anpassungsfähigkeit bieten, um in komplexen und dynamischen Umgebungen erfolgreich zu sein.

Die in dem Whitepaper thematisierten Anforderungen und Herausforderungen wie Erklärbarkeit, externe Schocks und Produktlebenszyklen sind grundsätzlich zu bedenken bei der Umsetzung von Forecasting Use Cases. Darüber hinaus gibt es viele weitere Anforderungen, die vor allem durch unternehmens- oder branchenspezifische Business-Logiken geprägt sind. Solange sich diese Anforderungen quantifizieren lassen und die dafür notwendigen Daten verfügbar sind, können auch Lösungskonzepte für eine Implementierung entwickelt und implementiert werden.

Auch heute werden Forecasts in vielen Unternehmen noch von Expert:innen mit hohem manuellen Aufwand erstellt. Das ist ineffizient und stößt bei hoher Produktanzahl oder komplexen Eigenschaften schnell an Grenzen. Datengetriebene Forecasts können die Effizienz signifikant erhöhen und komplexe Zusammenhänge objektiv berücksichtigen. Expert:innen können dabei weiterhin wichtige Bestandteile des Prozesses bleiben, um mit ihrem Fachwissen eine Qualitätssicherung der Forecastings zu gewährleisten.

Durch die Berücksichtigung individueller Anforderungen und die Anwendung maßgeschneiderter Lösungen kann Forecasting nicht nur operative Herausforderungen bewältigen, sondern auch langfristige, strategische Vorteile sichern. Unternehmen, die diese Potenziale erkennen und nutzen, sind besser gerüstet, um in einer sich ständig verändernden Welt erfolgreich zu agieren.

Creating value from data & AI.

statworx ist eines der führenden Beratungs- und Entwicklungsunternehmen für Daten und KI in der DACH-Region.

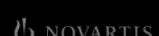
→ Über uns

Als führendes Beratungs- und Entwicklungsunternehmen für Daten und KI unterstützen wir Sie bei allen Aspekten der digitalen Transformation - von strategischer KI-Beratung über gezielte KI-Weiterbildung bis zur Entwicklung modernster KI-Lösungen.

→ Zahlen und Fakten



→ Einige unserer Kunden



Starten Sie Ihre Reise in die Zukunft mit Daten und KI

Unser **AI Potential Workshop** ist der optimale Startpunkt für Ihre KI-Transformation – verständlich und praxisorientiert.

AI Potential Workshop

Eine der großen Herausforderungen bei der Einführung von KI in Unternehmen ist die Identifikation von **werthaltigen Anwendungsfällen**. Mit unserem AI Potential Workshop bieten wir Unternehmen die Möglichkeit, schnell und effektiv spannende Use Cases für KI-Technologie zu finden und diese zu verproben.

[TERMIN BUCHEN](#)[MEHR INFOS](#)

Ablauf

- **Step 1: Assessment**
 - Abstimmungsgespräch zur Bewertung der Situation und Herausforderungen
- **Step 2: Preparation**
 - Use Cases mit KI-Potenzialen werden vorbereitet
 - Beinhaltet die Bewertung von Daten und deren Nutzen zur Effizienzsteigerung
- **Step 3: Workshop**
 - Eintägiger Workshop
 - Evaluierung und Entwicklung von agilen und interaktiven Strategien für DS, ML und Analytics Use Cases
 - Fokus auf Erkennung und Umsetzung der größten Chancen
- **Step 4: Prototype**
 - Entwicklung eines Prototypen
 - Entwicklungsbudget: 10 Personentage

Ihr Ansprechpartner



Daniel Lüttgau
Head of AI Development

Daniel Lüttgau verantwortet bei statworx den Bereich AI Development und ist Experte für die Anwendung von KI, um Mehrwerte für Unternehmen zu generieren.

Quellen

Quellenangaben und weiterführende Informationen zu den Inhalten aus diesem Whitepaper.

Das, A., Kong, W., Sen, R., & Zhou, Y. (2023). A decoder-only foundation model for time-series forecasting. arXiv preprint arXiv:2310.10688.

Garza, A., & Mergenthaler-Canseco, M. (2023). TimeGPT-1. arXiv preprint arXiv:2310.03589.

Goswami, M., Szafer, K., Choudhry, A., Cai, Y., Li, S., & Dubrawski, A. (2024). Moment: A family of open time-series foundation models. arXiv preprint arXiv:2402.03885.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

Hyndman, R.J., & Athanasopoulos, G. (2021) Forecasting: principles and practice, 3rd edition, OTexts: Melbourne, Australia. OTexts.com/fpp3

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. Computational statistics & data analysis, 55(9), 2579-2589.

Hyndman, R. J., & Athanasopoulos, G. (2014). Optimally reconciling forecasts in a hierarchy. Foresight: The International Journal of Applied Forecasting, (35), 42-48.

Kong, Y., Wang, Z., Nie, Y., Zhou, T., Zohren, S., Liang, Y., ... & Wen, Q. (2024). Unlocking the Power of LSTM for Long Term Time Series Forecasting. arXiv preprint arXiv:2408.10006.

Lim, B., Arik, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748-1764.

Liu, Y., Hu, T., Zhang, H., Wu, H., Wang, S., Ma, L., & Long, M. (2023). itransformer: Inverted transformers are effective for time series forecasting. arXiv preprint arXiv:2310.06625.

Nie, Y., Nguyen, N. H., Sinthong, P., & Kalagnanam, J. (2022). A time series is worth 64 words: Long-term forecasting with transformers. arXiv preprint arXiv:2211.14730.

NIXTLA. (2024). Meet TimeGPT. <https://www.nixtla.io/timegpt>

Wang, W., Yao, L., Chen, L., Lin, B., Cai, D., He, X., & Liu, W. (2021). CrossFormer: A versatile vision transformer hinging on cross-scale attention. arXiv 2021. arXiv preprint arXiv:2108.00154.

Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., & Long, M. (2022). Timesnet: Temporal 2d-variation modeling for general time series analysis. arXiv preprint arXiv:2210.02186.

Zeng, A., Chen, M., Zhang, L., & Xu, Q. (2023). Are transformers effective for time series forecasting?. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 9, pp. 11121-11128).

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021, May). Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of the AAAI conference on artificial intelligence (Vol. 35, No. 12, pp. 11106-11115).

statworx®

statworx GmbH
Hanauer Landstr. 150
60314 Frankfurt am Main

www.statworx.com
info@statworx.com
+49 (0)69 6783 067 - 51

AdobeStock
714890654, 901678106,
963719726, 966896174, 981988733