

Christina Erlwein-Sayer, Hochschule für Technik und Wirtschaft HTW Berlin

LSTM in wechselnden Zuständen: Regime Shifts und maschinelle Lernmodelle für das Finanzrisikomanagement

basierend auf einer Arbeit mit S. Grimm, A. Pieper und R. Alsaç

Highlights

Um unterschiedliche Marktbedingungen zu erfassen, entwickeln wir ein Modell, das ein rekurrentes neuronales Netz mit langem Kurzzeitgedächtnis (LSTM) mit einem Hidden Markov Modell (HMM) kombiniert.

- Kurze Einführung in Machine Learning (ML) im Finanzwesen
- Erkennung von verborgenen Zuständen in beobachteten Finanzzeitreihen
- Einbeziehung von Regimeinformationen in LSTM
- Studie zu Vorhersage von Kreditaufschlägen von Unternehmen in wechselnden Marktregimen

Übersicht

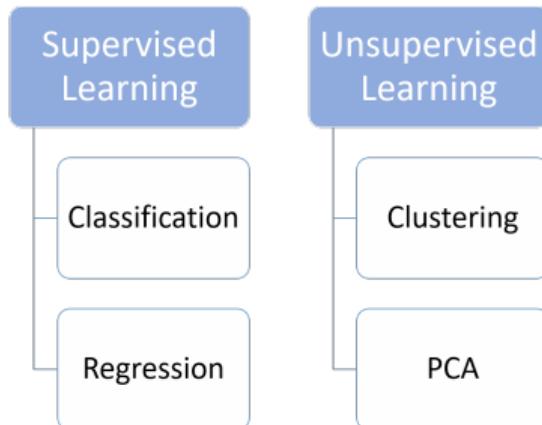
- Finanzmarkt und maschinelles Lernen
- Hidden Markov Modell auf Basis des OU-Prozesses
- Algorithmus in Batches
- Verknüpfung von Long Short Term Memory Model (LSTM) und HMM
- Vorhersage der Kreditaufschläge (Credit Spreads)
- Ausblick auf die Modellierung von Strompreisen
- Fazit

Modellierung von Credit Spreads

- Credit Spreads von Unternehmen reagieren auf Unternehmens-, aber auch auf allgemeine Marktbewegungen
- Turbulenzen am Finanzmarkt und ruhigere Marktphasen zeigen gemeinsame Bewegungen der Credit Spreads und ihre Eigenschaften \Rightarrow **wechselnde Marktregime**
- Die Credit Spreads weiten sich aus, wenn das finanzielle Risiko steigt: **makroökonomische Faktoren** (BIP, Inflation, Staatsverschuldung) beeinflussen die Kreditmärkte und werden mit zeitlich variierenden Koeffizienten berücksichtigt (z.B. [Bernoth et al 2012] oder [Maltritz (2012)])
- In der Kreditrisikomodellierung werden **Techniken des maschinellen Lernens** eingesetzt, um Kreditausfallwahrscheinlichkeiten zu modellieren (z. B. [Papouškova and Hajek (2019)]) oder Kreditrisikomodelle zu kalibrieren [Manzaf (2021)]

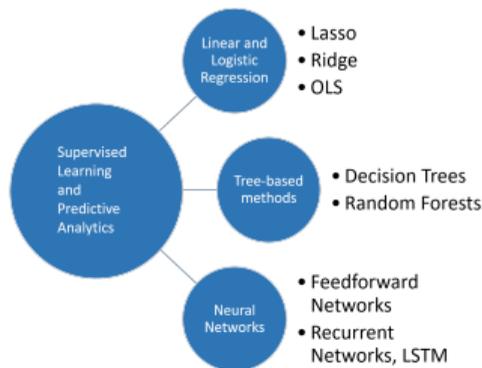
Machine Learning in Finance

- Methoden des maschinellen Lernens werden immer wichtiger für die Modellierung von Finanzzeitreihen
- Big Data - VOLUME, VELOCITY, VARIETY von Finanzdaten
- Auswahl, Einschränkungen und Verfügbarkeit von Daten: aufgeworfene Fragen bestimmen die Modellklasse



Machine Learning in Finance

- **Überwachtes Lernen** durch unterschiedliche Methoden möglich, wie z.B. Entscheidungsbäume, Random Forest oder Neuronale Netze
- Starke Vorhersageleistung durch **neuronale Netze**
- Interpretierbarkeit ist schwieriger als bei anderen Methoden
- Jüngere Entwicklungen umfassen die Modellierung von Finanzzeitreihen durch ein LSTM

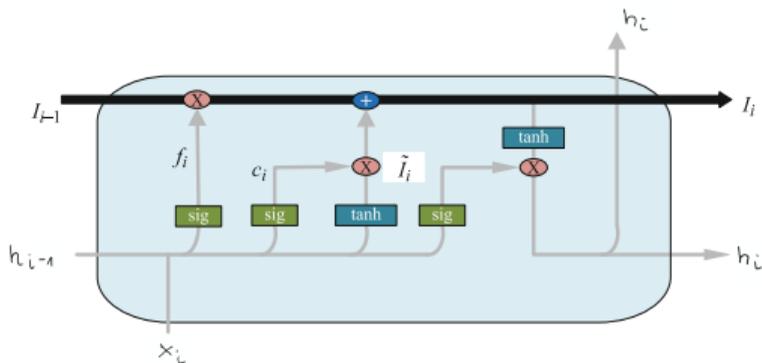


Machine Learning in Finance

- **Rekurrente neuronale Netze** eignen sich gut für die Modellierung und Vorhersage von Finanzzeitreihen (z.B. [Murphy (2022)])
- Probleme mit verschwindenden oder explodierenden Gradienten und der Entscheidung, welche Informationen weitergegeben werden sollen –> Verwendung von **LSTM für Finanzzeitreihen**
- Offen bleibt, ob rekurrente neuronale Netze (insbesondere LSTM-Modelle) klassische Zeitreihenmodelle (z.B. ARIMA, GARCH) für Finanzzeitreihen übertreffen, siehe z.B. [Siarni-Namini et al (2018)] oder [Kim et al (2021)].
- Wir untersuchen, ob und wie wechselnde Marktregime durch LSTM berücksichtigt werden können

Long-short term memory (LSTM) model

- LSTM-Modelle sind rekurrente neuronale Netze, die zur Modellierung von Zeitreihen an Popularität gewonnen haben [Hochreiter and Schmidhuber (1997)]
- Das Modell berücksichtigt vergangene Bewegungen der Zeitreihe
- Verwendung von Toren (Gates) zur Regulierung der bei jedem Zeitschritt zu behaltenden oder zu verwerfenden Informationen: Input-Gate, Forget-Gate und Output-Gate

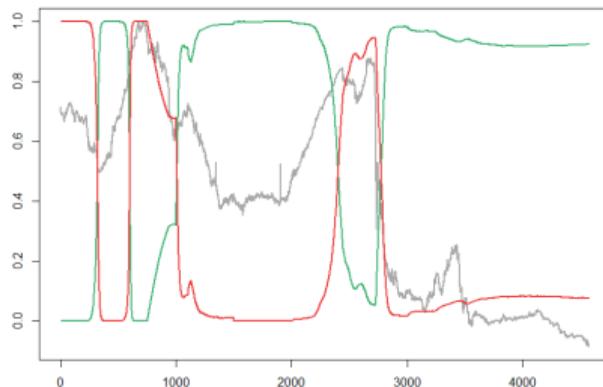


LSTM Modell

- Forget-Gate: welche langfristigen Informationen werden gespeichert und welche vergessen; Problem der langfristigen Informationen im Netz wird überwunden
- Feedback-Verbindungen helfen bei der Verarbeitung sequentieller Daten
- In [Xiong et al. (2016)] führte die Modellierung der Spreads von Unternehmensanleihen durch ein LSTM-Modell zu hoher Genauigkeit, die Random Forest und Bayesian Additive Regression Trees übertraf.

Regimewechsel im Kreditmodell

- Der Zustand des Marktes ist meist nicht beobachtbar und wird als latente Variable behandelt –> Hidden Markov Modell



Regimewechsel im Kreditmodell

- Hauptproblem: wie kann man **Informationen der Regimewechsel** filtern und in **LSTM** einbeziehen.
- **Unser Aufbau:** Spreads von Unternehmensanleihen werden mit einem HMM modelliert, das auf einem diskretisierten Ornstein-Uhlenbeck-Prozess basiert
- Terme des **Niveaus und der Geschwindigkeit der Rückkehr zum Mittelwert sowie die Volatilität** werden von einer **versteckten Markovkette** in diskreter Zeit geleitet
- Vorhersage von Credit Spread-Zeitreihen auf der Grundlage eines HMM
- Aktueller gefilterter Zustand der Markovkette wird als eingehende Variable im LSTM verwendet

HMM - Finanzmarkt und Modellstruktur

Sei (Ω, \mathcal{F}, P) der zugrunde liegende Wahrscheinlichkeitsraum der homogenen Markovkette \mathbf{x} mit endlichem N -dimensionalen Zustandsraum in diskreter Zeit.

- Unter P folgt die Markovkette der Dynamik

$$\mathbf{x}_t = \Pi^\top \mathbf{x}_{t-1} + \mathbf{v}_t,$$

wobei \mathbf{v}_t ein Vektor von Martingalzuwächsen ist, $\pi^{i,j} = P(\mathbf{x}_t = \mathbf{e}_j \mid \mathbf{x}_{t-1} = \mathbf{e}_i)$ ist die Wahrscheinlichkeit von Zustand i zu Zustand j zu wechseln, $i, j \in \{1, \dots, N\}$ und $\Pi = (\pi^{i,j}) \in \mathbb{R}^{N \times N}$ ist die Übergangswahrscheinlichkeitsmatrix.

- Der Credit Spread c folgt dem stochastischen Prozess

$$dc_t = a(\mathbf{x}_t)[\beta(\mathbf{x}_t) - c_t] dt + \xi(\mathbf{x}_t) dW_t \quad (1)$$

für $c_0 \geq 0$ mit $a(\mathbf{x}_t) = \langle \mathbf{a}, \mathbf{x}_t \rangle$, $\beta(\mathbf{x}_t) = \langle \boldsymbol{\beta}, \mathbf{x}_t \rangle$ und $\xi(\mathbf{x}_t) = \langle \boldsymbol{\xi}, \mathbf{x}_t \rangle$, wobei $\langle \cdot, \cdot \rangle$ das euklidische Skalarprodukt bezeichnet.

Finanzmarkt und Modellstruktur

- Der Beobachtungsprozess $\{y_{k+1}, 0 \leq k \leq K-1\}$ ist eine diskrete Darstellung des Credit Spreads und basiert auf der diskretisierten Lösung von (1).

$$y_{k+1} = \alpha(\mathbf{x}_k)y_k + \gamma(\mathbf{x}_k) + \eta(\mathbf{x}_k)z_{k+1} \quad (2)$$

wobei $\alpha(\mathbf{x}_k) = e^{-a(\mathbf{x}_k)\Delta t_{k+1}}$, $\gamma(\mathbf{x}_k) = \beta(\mathbf{x}_k)(1 - e^{-a(\mathbf{x}_k)\Delta t_{k+1}})$, $\eta(\mathbf{x}_k) = \xi(\mathbf{x}_k)\sqrt{\frac{1 - e^{-2a(\mathbf{x}_k)\Delta t_{k+1}}}{2a(\mathbf{x}_k)}}$
 and $\Delta t_{k+1} = t_{k+1} - t_k$.

- $\{z_k\}$ is eine Folge von u.i.v. standard-normalverteilten Zufallsvariablen
- die Modellparameter α , γ und η sind von den Zustandswahrscheinlichkeiten der Markovkette abhängig.

HMM für Credit Spreads

- Die Schätzung optimaler Parameter wird mit Hilfe eines filterbasierten Erwartungsmaximierungsalgorithmus (EM) durchgeführt (erstmal entwickelt für HMM von *[Elliott(1994)]*).
- Adaptive Filter für Zustände der Markovkette, Sprung- und Verweildauerprozesse (siehe *[Erlwein and Mamon (2009)]* für Herleitung und weitere Details) werden für die Parameterschätzungen aufgestellt zum Filtern unbeobachtbarer Informationen

HMM für Credit Spreads

Lemma

Sei \hat{J} , \hat{O} und $\hat{T}(f)$ die gefilterten Werte der Prozesse der Markovkette J , O und $T(f)$. Die optimalen Parameterschätzer basierend auf dem adaptiven EM-Algorithmus $\hat{\pi}_{ji}$, $\hat{\alpha}_i$, $\hat{\gamma}_i$, $\hat{\eta}_i$ sind gegeben durch

$$\hat{\pi}_{ji} = \frac{\hat{J}_k^{ji}}{\hat{O}_k^i}$$

$$\hat{\alpha}_i = \frac{\hat{T}_k^i(y_{k+1}, y_k) - \hat{T}_k^i(y) \hat{\gamma}_i}{\hat{T}_k^i(y^2)}$$

$$\hat{\gamma}_i = \frac{\hat{T}_{k+1}^i(y) - \hat{T}_k^i(y) \hat{\alpha}_i}{\hat{O}_k^i}$$

$$\hat{\eta}_i = \sqrt{\frac{\hat{T}_{k+1}^i(y^2) + \hat{\alpha}_i^2 \hat{T}_k^i(y^2) + \hat{\gamma}_i^2 \hat{O}_k^i - 2\hat{\alpha}_i \hat{T}_k^i(y_{k+1}, y_k) - \frac{2\hat{\gamma}_i \hat{T}_{k+1}^i(y) + 2\hat{\alpha}_i \hat{\gamma}_i \hat{T}_k^i(y)}{\hat{O}_k^i}}{\hat{O}_k^i}}$$

HMM für Credit Spreads

Algorithm 1: Filterbasierter EM-Algorithmus in Abschnitten

```

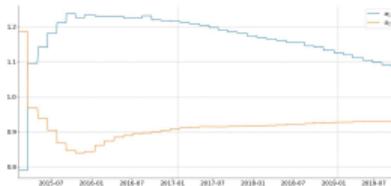
Setze Anfangswerte für  $\theta_0 = (\Pi, \alpha(\mathbf{e}_i), \gamma(\mathbf{e}_i), \eta(\mathbf{e}_i), i = 1, \dots, N)$ ;
for  $k$  in 1 to  $(\text{Length of time series})/(\text{p} \cdot \text{Length of Batch})$  do
  Setze Anfangswerte für Filter  $J_k^{dc}, O_k^c, T_k^c(f)$ ;
  for  $i$  in 1 to  $p$  do
    Verwende  $\theta_{k-1}$  als Anfangswert;
    for  $j$  in 1 to  $\text{Length of Batch}$  do
      (Maß) Bestimme  $\text{RNG}(j)$  für den Maßwechsel;
      (FILTER) Berechne den Filter  $j$  für Prozesse der Markovkette;
      (MC) Erhalte Schätzung  $\hat{\mathbf{x}}_j$  der Wahrscheinlichkeit des Zustandes
    end for
    (ML) Erhalte ML-Schätzung  $\theta_i$  von Lemma
  end for
end for

```

Pseudo-Algorithmus für die Parameterschätzung; läuft über beobachtete Daten in Abschnitten, Parameterschätzer werden on-line am Ende jedes Zeitabschnittes ermittelt.

HMM für Credit Spreads

Adaptive Parameterschätzung in einem 2-Zustands-HMM für Zeitreihe von Credit Spreads aus Spanien

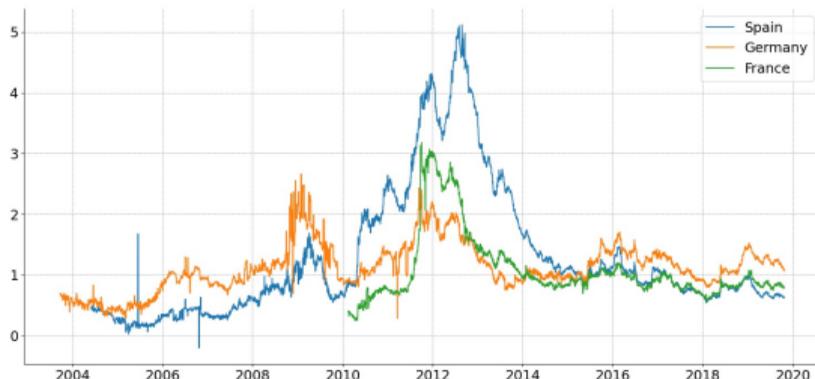


HMM-LSTM

- Ziel: **Zusammenführung** der aus dem **HMM** gewonnenen Informationen mit der Prognoseleistung des **LSTM**
- LSTM Basismodell ist in der Lage, Zeitreihen für Credit Spreads gut vorherzusagen
- **Gefilterte Zustandsschätzung** der versteckten Markov-Kette wird nun als erklärende Variable in das LSTM aufgenommen, um die Vorhersage zu verbessern

Credit Spreads: Datengrundlage

- analysierte Datensatz umfasst 150 Zeitreihen von Credit Spreads von Unternehmen aus drei europäischen Ländern über einen Zeitraum von zehn Jahren
- Länder: Spanien, Frankreich und Deutschland
- Untersuchung von Gruppen von Credit Spreads für Unternehmen: Gruppenklassifizierung auf der Grundlage der HMM-Prognoseleistung von Spreads im Zeitverlauf



Credit Spreads: Modelle

Für die 150 Spread-Zeitreihen führen wir eine Ein-Schritt-Vorhersage mit den entwickelten Modellen durch:

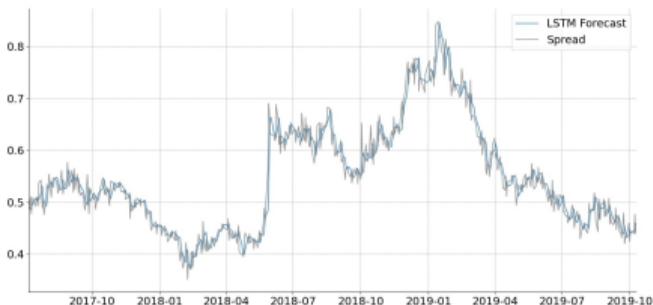
- i) das OU-HMM in einem 2-Zustandsmodell wird für jede Spread-Zeitreihe berechnet und die Vorhersagegenauigkeit wird analysiert
- ii) das LSTM wird zur Vorhersage der Credit Spreads verwendet und
- iii) der dritte Ansatz kombiniert beide Modelle: eine Zustandsvorhersage aus dem HMM für jede Credit Spread-Zeitreihe wird als zusätzliche erklärende Variable in das LSTM aufgenommen.

LSTM Vorhersage

LSTM wird in Python implementiert mit dem Paket `keras`

Hyperparameter im Modell:

- units per hidden layer = 4
- epochs = 1000
- early stopping patience = 120
- training/test = 0.5/0.5
- ts steps in, mc input, ts steps out = 3, 1, 1
- activation function: sigmoid
- optimizer = 'adam', loss function = mse



LSTM Vorhersage - Vergleich

Beispiel einer Zeitreihe aus Spanien



Einbindung der HMM Information

Beispiel eines Spreads einer französischen Unternehmensanleihe
LSTM mit HMM Information



Vorhersage: Gütemaße

- Die Vorhersage wird mit verschiedenen Gütemaßen bewertet: MSE und MdRAE für die Modellierung von Credit Spreads
- LSTM zeigt gute Resultate - kann mit verschiedenen Merkmalen der Zeitreihen umgehen; die Ein-Schritt-Vorhersage für Credit Spreads ist angemessen und liegt innerhalb einer niedrigen Fehlergrenze
- Ensemble-Modell zwischen LSTM und HMM führt zu Verbesserung des Fehlers über den gesamten Datensatz hinweg
- Untersuchung einzelner Länder zeigt höchste Verbesserung bei Modellierung von Zeitreihen in volatilen Märkten

Fehlermaße der Vorhersagen aller Modelle

| | HMM | LSTM Basismodell | LSTM mit MK | LSTM mit MK und mittlerer MK |
|------------|--------|---------------------|----------------|---------------------------------|
| mean MSE | 0.227 | 0.061 | 0.049 | 0.054 |
| min MSE | 0.0003 | 0.0003 | 0.0004 | 0.0003 |
| max MSE | 10.10 | 5.80 | 3.92 | 4.74 |
| mean MdRAE | 12.37 | 4.12 | 4.06 | 3.69 |
| min MdRAE | 0.995 | 0.873 | 0.860 | 0.866 |
| max MdRAE | 160.2 | 190.3 | 155.8 | 171.8 |

Tabelle 1. Fehlermaße über alle 150 Credit Spreads

LSTM-HMM Vorhersage gruppiert nach Ländern

| | HMM | LSTM Basismodell | LSTM mit MK | LSTM mit MK und mittlerer MK |
|----------|--------|---------------------|----------------|---------------------------------|
| mean MSE | 0.4937 | 0.1255 | 0.0866 | 0.1083 |
| min MSE | 0.0006 | 0.0005 | 0.0005 | 0.0005 |
| max MSE | 10.10 | 5.8002 | 3.9172 | 4.7393 |

Tabelle 2. 50 Credit Spreads aus Spanien

| | HMM | LSTM Basismodell | LSTM mit MK | LSTM mit MK und mittlerer MK |
|----------|--------|---------------------|----------------|---------------------------------|
| mean MSE | 0.0459 | 0.0335 | 0.0356 | 0.0322 |
| min MSE | 0.0003 | 0.0003 | 0.0004 | 0.0003 |
| max MSE | 0.4669 | 0.6585 | 0.6407 | 0.6230 |

Tabelle 3. 50 Credit Spreads aus Deutschland

Ausblick auf LSTM-HMM für Spotpreise

- HMM wird für deseasonalisierte Spotpreise auf dem deutschen Energiemarkt aufgebaut
- Wir erstellen ein Mixture of Experts (MoE) Modell, in dem in Untermengen LSTMs trainiert werden. Ein Gating-Modell entscheidet, welchem Experten zu vertrauen ist.
- Modellaufbau:
 - Aufgabe wird in Teilintervalle unterteilt
 - Für jede Untermenge wird ein LSTM entwickelt
 - Verwendung eines Gating-Modells, um zu entscheiden, welcher Experte verwendet werden soll
 - Vorhersagen und Gating-Funktionen werden für die endgültige Vorhersage kombiniert

LSTM-HMM für Spotpreise

Mixture of Experts

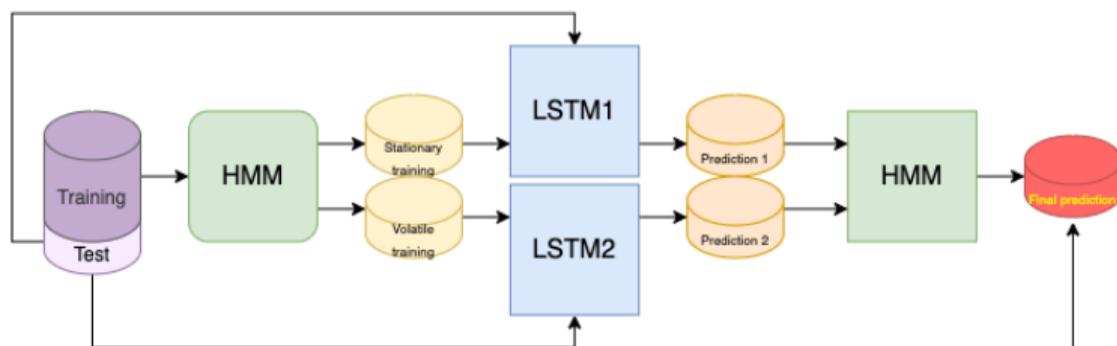


Abbildung 1. Modellaufbau

LSTM-HMM für Spotpreise

- Gating wird durch eine Markovkette mit zwei Zuständen übernommen, abhängig von geschätzten Übergangswahrscheinlichkeiten
- Schätzung mit Hilfe des Forward-Backward Algorithmus und dem Viterbi-Algorithmus zur Schätzung der optimalen Zustandssequenz

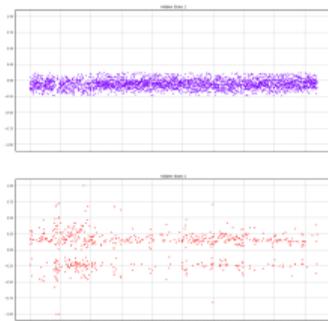


Abbildung 2. Beobachtungen aufgeteilt nach Zugehörigkeit zu geschätzten MK-Zuständen

LSTM-HMM für Spotpreise

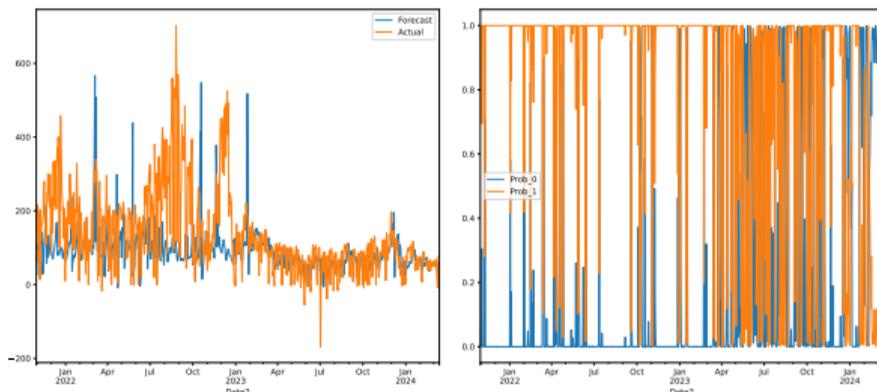


Abbildung 3. Vorhersage und Wahrscheinlichkeiten

Zusammenfassung

- Zusammenführung von HMM- und LSTM-Modellen, um Credit Spreads von Unternehmen zu prognostizieren
- Filter für die versteckte Markov-Kette und optimale Parameterschätzungen für OU-HMM
- Gute Ein-Schritt-Vorhersage mit reinem HMM und reinem LSTM
- Verbesserung der Vorhersage durch Einbeziehung der Schätzung der Markov-Kette als Merkmal im LSTM für Zeitreihen
- in volatilen Märkten bei Credit Spreads von Unternehmen größte Verbesserung möglich
- Ausblick: Robustheit des HMM, Optimierung der Hyperparameter des LSTM, LSTM für unterschiedliche Regimes in Energiemärkten und zur Handelsvolumenprädiktion

Literature

-  Bernoth, K., & Erdogan, B. (2012). Sovereign bond yield spreads: A time-varying coefficient approach. *Journal of International Money and Finance*, 31(3), 639-656.
-  Elliott, R.J. Exact adaptive filters for Markov chains observed in Gaussian noise. *Automatica* 30(9) 1399–1408, 1994.
-  Erlwein, C., & Mamon, R. (2009) An online estimation scheme for a Hull–White model with HMM-driven parameters. *Statistical Methods and Applications*, 18(1), 87-107.
-  Erlwein-Sayer, C., Grimm, S., Pieper, A., & Alsaç, R. (2023). Forecasting corporate credit spreads: Regime-switching in LSTM. *Econometrics and Statistics*.
-  Hochreiter, S. & Schmidhuber, J. (1997), Long short-term memory, *Neural computation* 9, 1735–80.
-  Kim, J.M, Kim, D.H., Jung, H. (2021) Applications of machine learning for corporate bond yield spread forecasting *North American Journal of Economics and Finance* 58 101540.
-  Maltritz, D. (2012) Determinants of sovereign yield spreads in the Eurozone: A Bayesian approach. *Journal of International Money and Finance*, 31(3), 657-672.

Literature

-  Manzo, G & Qiao, X (2021) Deep Learning Credit Risk Modeling, *Journal of Fixed Income*, vol. 31, no. 2, pp. 101-127. <https://doi.org/10.3905/jfi.2021.1.121>
-  Murphy, K (2022) *Probabilistic Machine Learning: An introduction*, MIT Press.
-  Siami-Namini, S., Tavakoli, N. und Siami Namin, A. (2018) A comparison of ARIMA and LSTM in forecasting time series. 17th IEEE international conference on machine learning and applications (ICMLA). IEEE, 2018
-  Papouskova M., Hajek P. (2019) Two-stage consumer credit risk modelling using heterogeneous ensemble learning *Decision Support Systems*, 118 , 33-45.
-  Yu, S.L. and Li, Z. (2018) Forecasting Stock Price Index Volatility with LSTM Deep Neural Network. *In: Tavana, M. and Patnaik, S. (Eds.) Recent Developments in Data Science and Business Analytics*, Springer International Publishing
-  Xiong, R. Y., Cai, H., Diego-Guerra, I., Lu, Y., Xu, X., & Yin, Y. (2016) Forecasting Credit Spreads: A Machine Learning Approach. *International Association for Quantitative Finance*, available at: https://www.iaqf.org/uploads/file_.pdf

DAV/DGVFM
Herbsttagung
2024

**Vielen Dank für
Ihre Aufmerksamkeit.**

*Christina Erlwein-Sayer, Hochschule für Technik und Wirtschaft HTW Berlin
erlwein@htw-berlin.de*