

Seminar Sommersemester 2021**Masterseminar Finance****Machine Learning Applications in Finance**

Ort der Präsentationstermine: tba
(abweichende Veranstaltungsorte sind unten angegeben)

Termin	Uhrzeit	Thema
01.02.2021 (Zoom)	16:30 Uhr	Kick-off Termin: Vorbesprechung mit Themenvergabe
12.04.2021 (Zoom)	10–12 Uhr (c.t.)	Einführung in Gliederung und inhaltliche Strukturierung der Seminararbeit
26.04.2021 (Zoom)	13–18 Uhr (s.t.)	Gliederungsbesprechung
20.05.2021 (LH-Gebäude)	bis 12:00 Uhr	Abgabe der Arbeiten
30.06.2021	10–18 Uhr (s.t.)	1. Resampling Methods 2. Regularization Methods 3. Regression Trees
01.07.2021	10–18 Uhr (s.t.)	4. Artificial Neural Networks 5. Sentiment Analysis 6. Text Readability and Similarity
01.07.2021	ab 18 Uhr	Gemeinsames Abendessen

Weitere Informationen finden Sie auf der Website des Lehrstuhls unter udue.de/finance.

Die verbindliche Anmeldung für das Seminar findet während der Vorbesprechung mit Themenvergabe am **01.02.2021** um **16:30 Uhr** statt. Jeder Seminarteilnehmer gibt drei Themenwünsche an (Erst-/Zweit-/Drittwunsch). Verwenden Sie bei der Anmeldung bitte zwingend die **stud.uni-due.de-E-Mail**-Adresse. Um Ihre Anmeldung abzuschließen, senden Sie **bis zum 02.02.2021** Ihren **Leistungsnachweis** an finance@uni-due.de.

Die Themenzuweisung erfolgt bis zum **11.02.2021**. Bis zum **14.02.2021** besteht die Möglichkeit, die Themen mit anderen Studierenden unter Verwendung des bereitgestellten Formblatts zu tauschen.

Am **12.04.2021** findet eine Einführung in die Erstellung einer Gliederung statt. Hier erhalten Sie Informationen über die grundlegenden Anforderungen an Ihre Gliederung. Die Gliederung (inkl. vorgesehener Seitenzahl und kurzer Inhaltsangabe je Kapitel) sowie ein Gantt-Diagramm sind bis zum **26.04.2021** fertigzustellen und an dem Tag mit dem Betreuer durchzusprechen.

Die Seminararbeiten sind spätestens am **20.05.2021** von **09-12 Uhr** in gedruckter Form (Abgabe bei Frau Fricke-Groenewold, LH 002) und zusätzlich per E-Mail an finance@uni-due.de einzureichen. Alle später eingereichten Arbeiten werden nicht mehr angenommen!

Von den Seminarteilnehmern wird Anwesenheit in allen Sitzungen erwartet, wobei aktive Mitarbeit zwingend erforderlich ist und einen wichtigen Bestandteil des Seminars darstellt.

Bearbeitungs- und Literaturhinweise:

Die folgenden Hinweise sollen den Einstieg in das Thema erleichtern und Schwerpunkte für die Bearbeitung erkennbar werden lassen. Eine Gliederung ist damit jedoch noch nicht vorgegeben. Zudem ist die angegebene Literatur nur als Basisliteratur zu verstehen.

Bitte achten Sie beim Verfassen auf die Berücksichtigung des neuen Leitfadens zur Anfertigung wissenschaftlicher Arbeiten sowie die lehrstuhleigenen Ergänzungen auf der Lehrstuhlhomepage. Beachten Sie bei der Anfertigung Ihrer Arbeit, dass der Umfang von **12 Seiten** nicht überschritten werden darf. Als Schriftart ist Times New Roman in Schriftgröße 12 Punkt mit 1,5-fachem Zeilenabstand zu verwenden. Das Nichteinhalten des Umfangs und der Bearbeitungshinweise wirkt sich negativ auf die Bewertung aus.

Thema 1: Resampling Methods

Unterscheiden Sie kurz die Resampling-Methoden Validation-Set-Approach, Leave-One-Out-Cross-Validation und K-Fold-Cross-Validation. Gehen Sie dabei auch auf deren Relevanz im Kontext von Machine Learning ein und erklären Sie das Problem des Overfitting. Erläutern Sie außerdem gängige Fehler-Metriken zur Evaluierung der Vorhersagegüte eines Modells. Zeigen Sie in weiterer Folge die empirischen Anwendungsmöglichkeiten der Resampling-Methoden auf und diskutieren Sie ausführlich die empirische Implementierung von Bianchi et al. (2020).

Berg, Tobias/Burg, Valentin/Gombović, Ana/Puri, Manju (2020): On the Rise of FinTechs: Credit Scoring Using Digital Footprints. In: *Review of Financial Studies*, 33: 2845–2897.

Bianchi, Daniele/Büchner, Matthias/Tamoni, Andrea (2020): Bond Risk Premiums with Machine Learning. In: *Review of Financial Studies*, demnächst.

Gürtler, Marc/Hibbeln, Martin/Usselmann, Piet (2018): Exposure at Default Modeling – A Theoretical and Empirical Assessment of Estimation Approaches and Parameter Choice. In: *Journal of Banking and Finance*, 91: 176–188.

Thema 2: Regularization Methods

Erläutern Sie kurz die Regularisierungsmethoden Least-Absolute-Shrinkage-and-Selection-Operator (LASSO), Ridge-Regression und Elastic-Net. Gehen Sie außerdem auf die Motivation für die Nutzung von Regularisierungsmethoden und den Bias-Variance-Tradeoff ein. Diskutieren Sie in weiterer Folge die ökonomischen Herausforderungen bei der Implementierung und die spezifische Anwendung von Regularisierungsmethoden auf Basis von Chinco et al. (2019).

Agarwal, Arvind/Gupta, Aparna/Kumar, Arun/Tamilselvam, Srikanth G. (2019): Learning Risk Culture of Banks Using News Analytics. In: *European Journal of Operational Research*, 277: 770–783.

Chinco, Alex/Clark-Joseph, Adam D./Ye, Mao (2019): Sparse Signals in the Cross-Section of Returns. In: *Journal of Finance*, 74: 449–492.

Feng, Guan Hao/Giglio, Stefano/Xiu, Dacheng (2020): Taming the Factor Zoo: A Test of New Factors. In: *Journal of Finance*, 75: 1327–1370.

Thema 3: Regression Trees

Erklären Sie kurz die Funktionsweise von Regressionsbäumen als Methode nichtparametrischer Vorhersage. Verdeutlichen Sie das hierbei entstehende Problem des Overfitting, und skizzieren Sie dessen Lösungsmöglichkeiten durch Boosting und Bagging. Fokussieren Sie Ihre Ausführungen zum Bagging auf die Methode des Random Forest. Fassen Sie anschließend mögliche Anwendungen von Regression Trees im Kontext Finance in einen Literaturüberblick zusammen, und analysieren Sie ausführlich die empirische Implementierung von Easley et al. (2020).

Easley, David/de Prado, Marcos L./O’Hara, Maureen/Zhang, Zhibai (2020): Microstructure in the Machine Age. In: *Review of Financial Studies*, demnächst.

Gu, Shihao/Kelly, Bryan/Xiu, Dacheng (2020): Empirical Asset Pricing via Machine Learning. In: *Review of Financial Studies*, 33: 2223–2273.

Mercadier, Mathieu/Lardy, Jean-Pierre (2019): Credit Spread Approximation and Improvement Using Random Forest Regression. In: *European Journal of Operational Research*, 277: 351–365.

Thema 4: Artificial Neural Networks

Beschreiben Sie kurz den grundlegenden Aufbau eines künstlichen neuronalen Netzes und fokussieren Sie Ihre Ausführungen auf die einfachste Form – klassische Feed-Forward-Netze. Gehen Sie dabei auch kurz auf die Vor- und Nachteile von künstlichen neuronalen Netzen gegenüber konventionellen statistischen Methoden ein. Erläutern Sie in weiterer Folge die Anwendungsmöglichkeiten im Bereich Finance in einem Literaturüberblick und diskutieren Sie ausführlich die empirische Implementierung eines neuronalen Netzes von Gu et al. (2020).

Bianchi, Daniele/Büchner, Matthias/Tamoni, Andrea (2020): Bond Risk Premiums with Machine Learning. In: Review of Financial Studies, demnächst.

Gu, Shihao/Kelly, Bryan/Xiu, Dacheng (2020): Empirical Asset Pricing via Machine Learning. In: Review of Financial Studies, 33: 2223–2273.

Wu, Wenbo/Chen, Jiaqi/Zhibin, Yang/Tindall, Michael L. (2020): A Cross-Sectional Machine Learning Approach for Hedge Fund Return Prediction and Selection. In: Management Science, demnächst.

Thema 5: Sentiment Analysis

Erläutern Sie kurz, wie die Textanalyse zur Untersuchung des Sentiments von finanziellen Texten verwendet werden kann. Beschreiben Sie kurz die Vorgehensweise bei der Implementierung der Textanalyse. Zeigen Sie deren Anwendungsmöglichkeiten in einem ausführlichen Literaturüberblick unter Einbeziehung der angegebenen Literatur auf. Gehen Sie dabei insbesondere auf die ökonomischen Herausforderungen bei Ammann/Schaub (2020) ein.

Ammann, Manuel/Schaub, Nic (2020): Do Individual Investors Trade on Investment-Related Internet Postings? In: Management Science, demnächst.

Hillert, Alexander/Jacobs, Heiko/Müller, Sebastian (2014): Media Makes Momentum. In: Review of Financial Studies, 27: 3467–3501.

Hillert, Alexander/Jacobs, Heiko (2016): Alphabetic Bias, Investor Recognition, and Trading Behavior. In: Review of Finance, 20: 693–723.

Thema 6: Text Readability and Similarity

Beschreiben Sie kurz, welche Maße genutzt werden können, um Lesbarkeit und Ähnlichkeit finanzieller Texte zu quantifizieren. Beziehen Sie in Ihre Ausführungen zur Lesbarkeit auch den Fog-Index ein, und diskutieren Sie dessen Vor- und Nachteile gegenüber weiteren gebräuchlichen Maßen. Analysieren Sie im Anschluss ausführlich die Anwendung von Lesbarkeits- und Ähnlichkeitsmaßen im Bereich Finance und weisen Sie auf die empirischen Herausforderungen hin.

Bonsall IV, Samuel B./Leone, Andrew J./Miller, Brian P./Rennekamp, Kristina (2017): A Plain English Measure of Financial Reporting Readability. In: Journal of Accounting and Economics, 63: 329–357.

Cohen, Lauren/Malloy, Christopher/Nguyen, Quoc (2020): Lazy Prices. In: Journal of Finance, 75: 1371–1415.

Loughran, Tim/McDonald, Bill (2014): Measuring Readability in Financial Disclosures. In: Journal of Finance, 69: 1643–1671.