

Data Analytics in Scientific Research – An Introduction

Basisinformation:

Ausbilder: Prof. Dr. Dr. h. c. mult. Horst Wildemann /
Dr. Sebastian Eckert

Format: Bei diesem Seminar handelt es sich um ein 3-tägiges Intensivseminar. Für TUM Doktoranden werden mit diesem Seminar die Mindestanforderungen von 22,5 Vollzeit-Präsenzstunden erreicht. Im Rahmen des Seminars werden die Vorgehensweisen im Bereich von Data Analytics Analysen sowie die Funktionsweise und Zusammenhänge von Bewertungsmodellen mit Fokus auf wissenschaftliche Anwendungen (Dissertationsvorhaben) sowie deren Vor- und Nachteile geschult. Durch das Aufzeigen von spezifischen Fallstudien und deren Anwendungsfelder sollte dem Doktoranden ein verbessertes Verständnis des Vorgehens vermittelt werden. Mittels der Einführung zum Analyse-Tool Alteryx sollte die Grundlage für die Gruppenarbeit gelegt werden. Die eigenständige Bearbeitung eines spezifischen wissenschaftsnahen Problems sollte das Erlernte vertiefen. Ziel des Seminars ist, dass die Teilnehmer weitere Analyse-Tools im Bereich des Data Analytics erlernen und diese gezielt innerhalb ihrer Forschungsarbeiten einsetzen können.

Teilnehmer: Das Seminar richtet sich an Doktoranden der TUM School of Management, die ein besonderes Interesse im Forschungsbereich der methodischen und wissenschaftlichen Betriebswirtschaftslehre zeigen. Die Teilnehmerzahl für dieses Seminar ist zur Aufrechterhaltung qualitativ hochwertiger Diskussionen und wegen der Durchführung von exemplarischen Analysen mittels der Software Alteryx auf 20 Teilnehmer begrenzt.

- Vorbereitung:** Im Vorfeld des Seminars ist keine spezielle Vorbereitung notwendig. Relevante Literatur wird den Teilnehmern zur Einarbeitung in die Thematik vorab zur Verfügung gestellt.
- Anmeldung:** Alle interessierten Doktoranden müssen sich bis spätestens zwei Wochen vor Veranstaltungsbeginn per E-Mail anmelden (E-Mail: wisekretariate@wi.tum.de).
- Ort:** Detaillierte Informationen zum Veranstaltungsort werden den Teilnehmern rechtzeitig vor Veranstaltungsbeginn mitgeteilt.
- Benotung:** In diesem Seminar wird keine Benotung vorgenommen. Das Bestehen/nicht Bestehen des Seminars wird anhand der Anwesenheit der Doktoranden belegt.

Seminarziele und Lernziele:

Das Ziel des 3-tägigen Seminars besteht darin den Doktoranden ein interdisziplinäres Wissen zur Anwendung von Data Analytics Lösungen im wissenschaftlichen Umfeld zu vermitteln. Im Seminar wird den Doktoranden ein ganzheitlicher Einblick im Vorgehen, von der Datenaufbereitung bis hin zur Interpretation der Analyseergebnisse, gelehrt. Ergänzend erfolgt eine Einführung in die Softwarelösung Alteryx, welche im Rahmen der Gruppenarbeit benötigt wird. Die in den Gruppen zu bearbeitenden Problemstellungen sollte den Doktoranden ermöglichen erste Erfahrungen mit Alteryx zu erlangen, die Grundlagen aus dem Seminar anzuwenden sowie Vor- und Nachteile eines solchen Analyse -Tools zu erkennen.

Erlerntes Wissen/Fähigkeiten:

Die Doktoranden erhalten mit dem Seminar die Grundlagen von Data Analytics in der Wissenschaft sowie eine Einführung in der Anwendung von der Softwarelösung Alteryx. Die Teilnehmer werden in den folgenden Aspekten geschult:

- Vorgehensweisen im Bereich von Data Analytics Analysen im wissenschaftlichen Umfeld und innerhalb des Dissertationsvorhabens: Data Cleaning, Data Manipulation, Data Layout und Data Enrichment
- Funktionsweise und Zusammenhänge von Bewertungsmodelle mit Fokus auf wissenschaftliche Anwendungen sowie deren Vor- und Nachteile und spezifische Anwendungsfelder
- Gezielte Anwendungen von Modellevaluierungskriterien zur Abschätzung der Modellqualität bei der Anwendung von Data Analytics Analyse-Tools
- Erste Erfahrungen durch die eigenständige Anwendung von Alteryx bei spezifischen wissenschaftsnahen Problemstellungen
- Ausweiten der wissenschaftlichen Arbeitsweise durch das Erlernen neuer Methoden neben den traditionellen statistischen Analysemöglichkeiten
- Erfahrungsaustausch zwischen den Doktoranden zur Anwendung spezifischer Data Analytic-Tools im Bereich des wissenschaftlichen Arbeitens

Seminarverlauf:

Tag 1

bis 13.30 Uhr	Eintreffen der Teilnehmer
14.00 Uhr – 15.00 Uhr	Übergeordnete Vorstellung und allgemeine Einführung in Data Analytics im wissenschaftlichen Umfeld
15.00 Uhr – 18.00 Uhr	Vorstellung der Data Culture (Datentypen, Datenstrukturen, Datenherkunft sowie der Datenqualität) und Erläuterung der Analytics Grundlagen mit Fokus auf Prozessabläufe und Anwendungsbereiche des Machine Learning. Abgrenzung von <i>Supervised Learning</i> , <i>Unsupervised Learning</i> und <i>Reinforcement Learning</i> .
18.00 Uhr – 19.00 Uhr	Gemeinsames Abendessen und Diskussionen in der Gruppe
19.00 Uhr – 21.00 Uhr	Darstellung von Regressionsproblemen sowie Einführung in die Evaluierung verschiedener wissenschaftlicher Modelle (Ablauf der Validierung, Unterschied Trainings- und Testdatensatz, Unterschied Varianz und Bias)
21.00 Uhr – 22.00 Uhr	Vorstellung unterschiedlicher Softwarelösungen aus dem Bereich Data Analytics (Alteryx, Rapidminer, DATAIKU, PowerBI)

Tag 2

08.00 Uhr – 10.00 Uhr	Aufbau, Inhalt und Funktionsweise des Data Cleanings (Soll-Ist-Vergleich von Datensätzen, Datenbewertung) und der Data Manipulation (Filter, Stack / Split, Aggregation). Visualisierung anhand von Fallbeispielen.
10.00 Uhr – 12.00 Uhr	Aufbau, Inhalt und Funktionsweise des Data Layouts (Transpose, Fold / Unfold, Pivot) und des Data Enrichments (Joins & Duplicates, Top N Recipe, Sliding Windows). Visualisierung anhand von Fallbeispielen.
12.00 Uhr – 13.00 Uhr	Gemeinsames Mittagessen und Diskussionen in der Gruppe

13.00 Uhr – 19.00 Uhr	Einführung in die Software Alteryx als Basis des Hackathons an Tag 3. Fokus auf die Softwarebereiche <i>Data Preparation</i> , <i>Data Investigation</i> , <i>Forecasting</i> sowie <i>Prediction</i> . Aufzeigen der Vor- und Nachteile im Vergleich zur manuellen Datenaufbereitung und Datenauswertung. Ausblick für den Einsatz in der Forschung durch Präsentation von Beispielanwendungen
19.00 Uhr – 19.30 Uhr	Einteilung der Gruppen und Zuweisung spezifischer Problemstellungen für den Alteryx-Hackathon an Tag 3
ab 19.30 Uhr	Gemeinsames Abendessen und Diskussionen in der Gruppe

Tag 3

08:00 Uhr – 11:00 Uhr	Hackathon zur Bearbeitung spezifischer wissenschaftlicher Problemstellungen mit der Softwarelösung Alteryx
11:00 Uhr – 11:30 Uhr	Vorstellung und gemeinschaftliche Diskussion der erarbeiteten Umsetzungswege und Ergebnisse der Gruppenarbeiten
12:30 Uhr – 13:30 Uhr	Lessons Learned und allgemeine Rezension der Seminarergebnisse und Einsatzmöglichkeiten von Data Analytics im Rahmen des Promotionsvorhabens
13:30 Uhr – 14:00 Uhr	Gemeinsamer Mittagsimbiss und Diskussionen in der Gruppe
14:00 Uhr	Ende der Veranstaltung und Heimreise

Seminar-begleitende Literatur:

Den Teilnehmern ist freigestellt sich im Vorfeld des Seminars in die entsprechende Literatur einzulesen.

Folgende Publikationen und Bücher werden zur Vorbereitung des Seminars empfohlen:

- Aggarwal, Charu C. (2015): Data Mining. Cham: Springer International Publishing.
- Albright, S. Christian; Winston, Wayne L. (2017): Business analytics. Data analysis and decision making. 6th edition. Boston, Mass.: Cengage Publishing.
- Bange, C.; Grosser, T.; Janoschek, N. (2015): Big data use cases.
- Bartlett, Randall Kenyon (2013): A practitioner's guide to business analytics. Using data analysis tools to improve your organization's decision making and strategy. New York NY u.a.: McGraw Hill.
- Bose, R. (2009): Advanced analytics: opportunities and challenges. In: Industrial Management & Data Systems (109(2)), S. 155–172.
- Chahal, H.; Jvoti, J.; Wirtz, J. (2019): Understanding the Role of Business Analytics. Some applications. [S.I.]: SPRINGER.
- Chamoni, Peter (2011): BI-Strategie zum Ausgleich von Technologie-Push und Business Pull. In: IMDM (193), S. 13–22.
- Chen; Chiang; Storey (2012): Business Intelligence and Analytics: From Big Data to Big Impact. In: MIS Quarterly 36 (4), S. 1165. DOI: 10.2307/41703503.
- Cleve, Jürgen; Lämmel, Uwe (2014): Data Mining. [Elektronische Resource]. München: De Gruyter Oldenbourg.
- Davenport, Thomas H.; Harris, J. G. (2007): Competing on Analytics: The new science of winning. In: Harvard Business School Press.
- Ereth, J.; Kemper, H.-G. (2016): Business Analytics und Business Intelligence. In: Controlling 28 (8-9), S. 458–464.
- Gluchowski, Peter (2016): Business Analytics – Grundlagen, Methoden und Einsatzpotenziale. In: HMD 53 (3), S. 273–286. DOI: 10.1365/s40702-015-0206-5.
- Gronau, Norbert; Thim, Christof; Fohrholz, Corinna (2016): Business Analytics in der deutschen Praxis. In: Controlling 28 (8-9), S. 472–479. DOI: 10.15358/0935-0381-2016-8-9-472.

- Seiter, Mischa (2017): Business Analytics. Effektive Nutzung fortschrittlicher Algorithmen in der Unternehmenssteuerung. München: Franz Vahlen.
- Wildemann, Horst (2017): Neue Geschäftsfelder, Geschäftsmodelle, Technologien. Müssen sich Unternehmen neu erfinden? München: TCW Transfer-Zentrum GmbH & Co KG (TCW-Report, Nr. 69).
- Wildemann, Horst (2018c): Produktivität durch Industrie 4.0. 1. Auflage. München: TCW-Verlag (TCW, 34).
- Wildemann, Horst (2020): Neue Geschäftsmodelle in der Industrie 4.0. Leitfaden zur Entwicklung neuer Geschäftsmodelle in der Industrie 4.0. 4. Auflage. München: TCW-Verlag (Leitfaden / TCW Transfer-Centrum für Produktions-Logistik und Technologie-Management, 147).
- Zott, Christoph; Amit, Raphael; Massa, Lorenzo (2011): The Business Model: Recent Developments and Future Research. In: Journal of Management 37 (4), S. 1019–1042. DOI: 10.1177/0149206311406265.