

29.07.2022 | Campus

Erfolgreiche Premiere beim Data Mining Cup

Große Datenmengen mit statistischen Methoden, computergestützten Verfahren und Algorithmen zu analysieren, um neue, nützliche Erkenntnisse zu gewinnen – das geschieht beim sogenannten Data Mining. Kevin Kröll, Leander Wernst und Moritz Uhlig, Studierende des Bachelor-Studiengangs Media Systems von der Fakultät Design, Medien, Information (DMI), haben ihre Fähigkeiten in diesem Bereich beim diesjährigen internationalen Data Mining Cup (DMC) unter Beweis gestellt: Bereits mit der ersten Teilnahme eines Teams der HAW Hamburg an dem studentischen Wettbewerb erreichten sie den 16. Platz und damit eine Platzierung im oberen Drittel.



Studierende des Bachelor-Studiengangs Media Systems, die erstmals am Data Mining Cup teilgenommen haben, zusammen mit ihrer Professorin (v.li.): Leander Wernst, Prof. Dr.-Ing. Sabine Schumann, Kevin Kröll und Moritz Uhlig

„Ein 16. Platz, das klingt vielleicht im ersten Moment nach keiner besonderen Leistung“, sagt Sabine Schumann, Professorin für Wirtschaftsinformatik am Department Medientechnik, „aber wenn man bedenkt, dass wir als HAW Hamburg das erste Mal mit einem studentischen Team an dem internationalen Wettbewerb teilgenommen haben und viele Hochschulen, die sich auch angemeldet hatten, am Ende gar keine Lösung einreichten oder nur null Punkte erzielten, dann ist das schon ein Grund, die Sektkorken knallen zu lassen!“ Prof. Dr.-Ing. Sabine Schumann hatte die Bachelor-Studierenden im Rahmen des projektbasierten Wahlpflichtmoduls Adaptive Systems and Artificial Intelligence (ASaAI) auf den Wettbewerb aufmerksam gemacht. Dieses Modul führt Frau Schumann gemeinsam mit Prof. Dr. Larissa Putzar durch, es steht den Studierenden beider Studiengänge (Media Systems und Medientechnik) des Departments offen.

Erst einmal blickten wir ziemlich ratlos auf die Tabellen mit den endlosen Zahlenkolonnen.

KEVIN KRÖLL

Bis zur Frustrationsgrenze

Die Studierenden, die sich in dem Wahlpflichtmodul das erste Mal mit dem Thema Künstlicher Intelligenz und Data Science beschäftigt hatten, standen dabei anfangs vor anscheinend unlösbaren Aufgaben. „Als wir am 12. April die Fragestellung mit den dazugehörigen Daten erhielten, blickten wir erst einmal ziemlich ratlos auf Tabellen mit den endlosen Zahlenkolonnen. Wir sollten aus einer anonymisierten Bestellhistorie eines Online-Händlers ein Modell entwickeln, das vorhersagt, ob und wann ein Produkt wiedergekauft wird.“, berichtet Kevin Kröll. „Nach den ersten Wochen war unsere Frustrationsgrenze schon fast erreicht“, pflichtet ihm sein Teamkollege Moritz Uhlig bei. „Wir hatten bereits versucht, mit unterschiedlichsten Verfahren neue Erkenntnisse aus den vorliegenden Daten zu gewinnen, aber kamen einfach zu keinem befriedigenden Ergebnis, das uns bei der Fragestellung voranbrachte.“

„Das waren erstmal nur numerische Werte für Kunden-, Produkt-IDs, Zeitpunkte und Eigenschaften und wir hatten lange Zeit Schwierigkeiten damit, diese Daten überhaupt zielführend zu transformieren.“, sagt auch Leander Wernst. „Zusätzlich umfassten die Daten nur den Zeitraum vom 1. Juni 2020 bis zum 31. Januar 2021. Das ist eine verhältnismäßig kurze Zeitspanne, um Voraussagen über das Kaufverhalten im Februar 2021 abzuleiten.“ Doch nach diversen Versuchen fanden die drei herangehenden Medieninformatiker etwa drei Wochen vor dem Abgabetermin am 28. Juni 2022 endlich einen Erfolg versprechenden Ansatz.

Wettbewerbe wie der DMC sind für Studierende extrem motivierend, da sie ihr Wissen direkt anwenden und vor allem mit echten Daten arbeiten können.

PROF. DR.-ING. SABINE SCHUMANN

Intensive Abschlussphase

„Danach haben wir die letzten zwei Wochen vor der Abgabe sehr intensiv, oft auch bis zu zehn Stunden täglich, an einer Lösung gearbeitet“, sagt Kevin Kröll, der sich vorstellen kann, diese Form der Datenanalyse später beruflich zu seinem Schwerpunkt zu machen. Umso größer war die Freude, am Ende tatsächlich ein Ergebnis einreichen zu können.

Insgesamt haben sich 78 Teams von 59 Universitäten aus 23 Ländern zu dem Wettbewerb angemeldet, 44 reichten eine Lösung ein, sechs davon wurden allerdings mit null Punkten bewertet. Den ersten und dritten Platz belegten Teams der Asia Pacific University aus Malaysia, auf Platz zwei lag eines der zwei Teams der Hochschule Zwickau. „Wir wollen auf jeden Fall auch noch einmal mit den besser platzierten Studierenden Kontakt aufnehmen“, sagt Kevin Kröll. „Einige davon haben nämlich Verfahren benutzt, die bei uns nicht die gewünschten Ergebnisse brachten. Es wäre jetzt natürlich interessant zu erfahren, was diese Teams anders gemacht haben und was wir beim nächsten Mal besser machen können.“

Prof. Dr.-Ing. Sabine Schumann zieht ein durchweg positives Resümee: „Wettbewerbe wie der DMC sind für Studierende extrem motivierend, da sie ihr Wissen direkt anwenden und vor allem mit echten Daten arbeiten können. Das bedeutet für die Studierenden zwar ein erhebliches zusätzliches Engagement zum Semesterstundenplan, aber als Lehrende versuchen wir die Studierenden zu motivieren, auch Freiräume zum eigenständigen Lernen zu nutzen. Wir danken unseren zahlreichen Unterstützer*innen, die diese gelungene Premiere ermöglicht haben.“

Text: Maren Borgerding

Data Mining Cup (DMC)

Der DMC wurde im Jahr 2000 von dem Unternehmen prudsys ins Leben gerufen und findet seitdem jährlich im Frühjahr statt. Pro Bildungseinrichtung dürfen (seit 2009) maximal zwei Teams teilnehmen. Die drei erstplatzierten Teams erhalten ein Preisgeld. Das Unternehmen prudsys wurde 1998 in Chemnitz durch Prof. Dr. Werner Dilger (TU Chemnitz-Zwickau), Prof. Dr. Andreas Ittner (seit 2007 Professor an der HS Mittweida) und weitere ehemalige Studierende/Beschäftigte der TU Chemnitz-Zwickau gegründet. Das Unternehmen ist Inhaber zahlreicher Patente und arbeitet eng mit vielen wissenschaftlichen Instituten zusammen. Der Hauptsitz der Firma ist mittlerweile in Berlin, weitere Büros sind in Chemnitz, Almere, Mailand und Istanbul.

Weitere Informationen:

[🔗 Data Mining Cup 2022](#)

Kontakt

Prof. Dr.-Ing. Sabine Schumann
Professorin für Wirtschaftsinformatik
Department Medientechnik

T +49 40 428 75-7619

✉ [sabine.schumann\(@\)haw-hamburg.de](mailto:sabine.schumann(@)haw-hamburg.de)