

Workshop „KI selber programmieren“

Der TAO-SFZ-Workshop „KI selber programmieren“ fand vom 29. bis 31. Oktober 2018 statt. Insgesamt sechs interessierte Schüler fanden sich ein, um in praktischen Aufgaben zu erfahren, wie man Daten analysiert, aufbereitet und neues Wissen aus ihnen extrahiert. Am ersten Tag wurden die Schüler mit der Plattform „Rapidminer“ vertraut gemacht. Ein realer Datensatz, der Attribute von mehreren Tausend Pilzen beinhaltet, wurde zunächst auf Basis von mehreren Statistiken visualisiert. So wurde zum Beispiel entdeckt, dass mehr essbare als giftige Pilze im Datensatz vorkommen. Die nächste Aufgabe, vor der die Schüler standen war es, einen Entscheidungsbaum zu lernen, der auf Grundlage seiner Attributwerte vorhersagt, ob ein gegebener Pilz giftig oder ungiftig ist. Rapidminer erlaubte es mit wenigen Mausklicks den Entscheidungsbaum zu generieren. Das brachte zum Vorschein, dass man schon allein anhand des Geruchs eine gute Vorhersage treffen kann. Ist man sich allerdings unsicher, muss man noch weitere Attribute wie die Hutform oder die Lamellenfarbe berücksichtigen. Am Nachmittag wurde dann am Datensatz der Überlebenden des Titanic-Desasters erklärt, wie man Rapidminer dazu verwenden kann, die Parameter des Entscheidungsbaum-Algorithmus so einzustellen, dass das gelernte Modell auf einem separaten Test-Set besonders gut abschneidet.

Der nächste Tag stand ganz im Zeichen von Neuronalen Netzen und Deep Learning. Die Schüler lernten, dass man solche Verfahren vor allem dann nutzen kann, wenn die Daten in Form von Zahlen vorliegen. Der erste Datensatz beinhaltete medizinische Messwerte von Frauen aus dem Pima-Indianer-Stamm. Werte wie Alter, Blutdruck und Insulinwert wurden gemessen. Die relevante Klasse war eine Binärentscheidung, ob die Frauen Anzeichen von Diabetes aufwiesen. Diesmal wurde das Framework „Keras“ verwendet, mit dessen Hilfe die Schüler ein Neuronales Netz definieren, trainieren und evaluieren konnten. Nach mehreren Epochen konnte schon eine hohe Accuracy und ein niedriger Loss auf einem Test-Set gemessen werden, was ein Indikator für ein gutes Modell ist. Die nächste Station waren Convolutional Neural Networks (CNN); eine spezielle Form der Neuronalen Netze zum Erkennen von Objekten auf Bildern. Der Task war nun, auf einem Bild eine handschriftlich geschriebene Ziffer zu erkennen. Als Datensatz konnte nun MNIST verwendet werden; eine Sammlung von 60,000 handgeschriebenen Ziffern. Nach der Definition eines CNNs konnten die Schüler das Training starten. Diesmal dauerte es länger, bis das Modell zufriedenstellende Werte lieferte. Am Ende wurde aber eine von den Schülern gezeichnete Ziffer mit hoher Konfidenz korrekt erkannt. Am Abend wurde dann der Spielfilm Ex Machina angeschaut, der die Themen KI und Ethik aus einem philosophischen Blickwinkel betrachtet. Die Verköstigung kam natürlich nicht zu kurz.

Am Mittwoch, dem letzten Tag des TAO-Workshops, lernten die Schüler wie man mithilfe des Shiny R Frameworks Analysen visualisiert und interaktiv macht. Die Schüler erhielten zunächst eine Einführung in die Grundlagen der Programmiersprache R. Anschließend bauten sie sich selbst ein eigenes Dashboard, in dem sie beliebige Daten hochladen, anzeigen und analysieren konnten. Sie integrierten beispielsweise Auswahlformulare, Tabellen und Histogramme in die Nutzeroberfläche. Der Kreativität waren keine Grenzen gesetzt. Im Anschluss lernten die Schüler einen Vertreter maschineller Lernverfahren kennen: die k-Nearest Neighbour Methode. Auch hier standen sie vor der Herausforderung, dieses Verfahren in ihr eigenes Dashboard zu integrieren und auf dem hochgeladenen Datensatz anzuwenden. Doch das war zum Glück kein Problem! Zum Abschluss lernten die Schüler mithilfe einer in Shiny R implementierten Simulation, dass eine hohe Korrekturklassifizierungsrate (Accuracy) nicht unbedingt etwas über die Performanz eines Lernverfahrens aussagt und welche alternativen Maße man stattdessen heranziehen kann. Alles in allem war der TAO-Workshop eine abwechslungsreiche Rundreise durch die Welt der Datenanalyse und des maschinellen Lernens.

Der Workshop wurde von der Gruppe Kognitive Systeme durchgeführt.

Verantwortlicher Leiter: Johannes Rabold

Beiträge: Bettina Finzel, Mark Gromowski

Einführungsvortrag: Ute Schmid

TAO-Lehrkraft: Lutz Reuter

www.uni-bamberg.de/cogsys

