

Explainable AI: Displaying Time-Series Data in Immersive Virtual Reality Applications

Projektgruppe „Praxis der Forschung“ – Sommersemester 2020

1 Hintergrund

In der KI-Forschung wird die Erklärbarkeit von Anwendungen im Bereich des maschinellen Lernens ein immer wichtiger werdender Aspekt. Mittel zu besitzen, um die Effekte sowie Nebeneffekte von Modellen besser zu verstehen und Veränderungen bei Feineinstellungen der Modellparameter direkt im dreidimensionalen Raum mitzerleben ist von großer Bedeutung. Darauf aufbauend wird die Beurteilung der Genauigkeit, sowie die Leistungsdarstellung von Modellen begünstigt.

Ein weiterer positiver Effekt besteht in der Stärkung des gesellschaftlichen Vertrauens in KI-Applikationen durch Erleben der Modelle, um dementsprechend die Anwendung in gesellschaftlich wichtigen Bereichen zu fördern. Des Weiteren wird im Bereich des Wissenstransfers das Lernen von maschinellem Lernen positiv beeinflusst. Es gibt unterschiedliche Ansätze. Zum einen werden aus sich heraus „verständliche“ Modelle wie Naive Bayes oder Entscheidungsbäume trainiert, die aber auch schnell schwer erfassbar werden oder es werden komplexe Modelle anhand einzelner Entscheidungen bewertet wie dies beispielsweise LIME tut. Beide zuvor genannten Ansätze stoßen bei hoch-dimensionalen Daten und komplex errechneten Merkmalen an ihre Grenzen.

In der Sichtung von hochdimensionalen Simulationsdaten ist eine visuelle dreidimensionale Darstellung bereits üblich, da hier die Daten von Natur aus räumlich interpretiert werden können. Des Weiteren wurden bei Clustering Analysen sowie interaktiver visueller Analysen hochdimensionaler Daten deutliche Erfolge erzielt.

2 Projektbeschreibung

Die Arbeit soll mögliche Interessante Forschungsaspekte und wegweisende Richtungen für die Analyse von gelernten Modellen innerhalb immersiver Räume im Bereich der virtuellen Realität untersuchen – inklusive der Nutzung von 3D Audio zur Signalisierung von markanten Merkmalen wie z. B. False Positives / Positive Negatives. Ziel ist die Implementierung nebst Nutzer-Evaluation von unterschiedlichen Methoden anhand von typischerweise schwerer verständlichen Eingangsdaten – wie z. B. Zeitserien – auf Grundlage bereits existierender Ansätze der dreidimensionalen Visualisierung sowie des erklärbaren Maschinellen Lernens.

3 Literatur

1. Francois Bertrand, *Project NEO: Virtual Reality Data Visualization for Machine Learning*, <https://www.youtube.com/watch?v=myI4P9C34A0&feature=youtu.be>
2. Benjamin Resnick, *Visualizing High Dimensional Data In Augmented Reality*, <https://www.kdnuggets.com/2017/09/ibm-visualizing-high-dimensional-data-augmented-reality.html>

4 Kontakt / Betreuung

Rainer Duda
Dr. Till Riedel

duda@kit.edu, Raum 209 (Geb. 07.07)
till.riedel@kit.edu, Raum 211 (Geb. 07.07)