

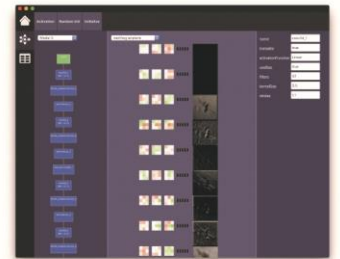
Visual Analytics for Explainable Deep Learning

Diese Thesis analysiert die Anwendungsgebiete und den Stand der Technik von visuellen Analyseverfahren, die eingesetzt werden, um Deep Learning-Modelle zu erklären. Sie behandelt verschiedene Ansätze systematisch und diskutiert deren Voraussetzungen, Potentiale und ungelöste Herausforderungen. Ein besonderer Fokus wird auf Systeme gelegt, die einen Schwerpunkt auf visuelle Analyseverfahren setzen. Basierend auf den gewonnenen Erkenntnissen wurden zwei System-Prototypen entwickelt, um ein besseres Verständnis für die Herausforderungen zu gewinnen und Lösungsansätze zu testen. Der erste Prototyp ermöglicht es, das Modellverhalten basierend auf benutzerdefinierten Datengruppen zu erforschen. Er demonstriert dabei die technischen Grenzen für Analyse-Systeme, die versuchen, beliebig komplexe Modelle mit traditionellen Verfahren zu visualisieren. Zusätzlich wird aufgezeigt, welche Herausforderung verschiedene Formate von Datensätzen für modellunabhängige Systeme darstellen. Zuletzt wird bekräftigt, dass die hauptsächliche Herausforderung in der Komplexitätsreduzierung von Deep Learning-Modellen liegt, damit diese von Menschen verstanden werden können, ohne zu viele Informationen zu verlieren. Dieses Problem wird im zweiten Prototyp weiter untersucht. Dieser verwendet verschiedene Kennzahlen, die Elemente (Neuronen, Filter, etc.) als einzelne Skalarwerte beschreiben. Basierend auf diesen Kennzahlen ermöglicht das System, Gruppen von Elementen zu bilden, die sich auffällig verhalten, wenn mit den visualisierten Daten interagiert wird. Durchgeführte Fallstudien deuten an, dass das System mithilfe einiger der gefundenen Kennzahlen als Leitsystem dienen kann, um relevante Bereiche im Modell zu finden, wo weitere Erklärungsstrategien angewendet werden können. Das wird exemplarisch mit der Untersuchung von Datensätzen demonstriert, die eine starke Interaktion mit den gefundenen Elementgruppen aufweisen.

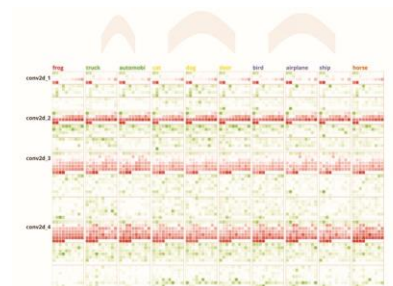


Diplomand/in
Yanick Lukic

Dozent
Philipp Ackermann



Der erste Prototyp ermöglicht es Benutzern, Deep Learning-Modelle zu laden, Gruppen von Datensätzen zu definieren und Interaktionen zwischen diesen Komponenten zu untersuchen.



Mit dem zweiten Prototyp können Kennzahlen von Modellelementen analysiert werden. Er verfügt hierfür über diverse interaktive Funktionen, z.B. Clustering basierend auf benutzerdefinierten Kennzahlgewichtungen.