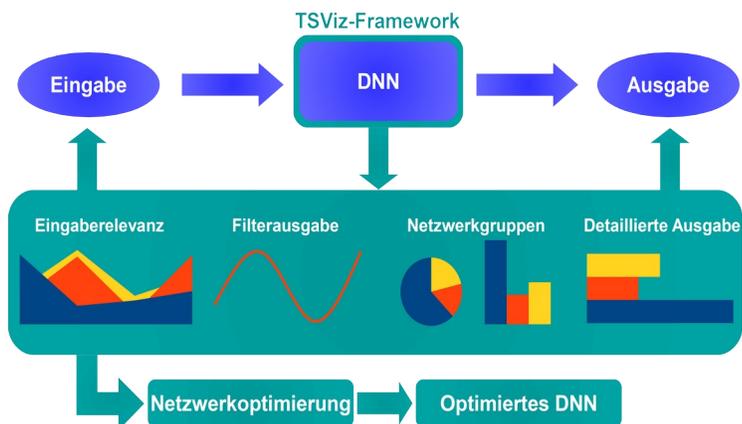


Time Series Visualization Framework – TSViz

Deep neural networks are being deployed in a range of different applications, which poses a great question on their reliability and robustness. In order to tackle the goal of interpretability of DNNs, DFKI developed a framework that acts as a bridge between the researchers, developers, and users to provide suited information about the decisions of neural networks for time-series analysis.



The framework enables various user groups to understand the network's prediction as a whole as well as the specific parts of the model which contributed towards it. Therefore, only the information that is aligned with the background of the user group is presented. The goal of the framework is to understand the behavior of the network to enable to further optimization based on the gained insights.

Nowadays DNNs are being used various tasks these days, and sometimes the end-user does not even know about them, e.g. for search, recommendations, and automated analysis. For some safety critical domains, together with accuracy, trustworthiness is of extreme importance. Therefore, it is necessary to have an accurate system which can precisely explain the decision i.e. how and why a particular decision was taken, especially to establish trust for its failure cases.

Finding an explanation for a given network decision is a challenging task that involves various different factors such as the domain, target group, and data format. Especially when it comes to domain, most of the existing explanations are designed for images or language as these modalities are directly intelligible for humans. Conversely, in the case of time-series, this becomes a significant challenge, as the data in time-series modality is not directly intelligible as images and text, which TSViz attempts to solve.

Covering the time-series domain:

The framework acts as an instance-based post-hoc method and can handle any convolutional neural network architecture. Furthermore, it is not limited to a specific task and can be used with classification, forecasting, as well as for regression networks.

User-specific explanation enables the broad applicability:

Due to an interactive interface, it is possible to hide information that is not aligned with the background of the target group. Therefore, the system can be used by developers and users without limitation to a specific group.

Kontakt

Prof. Dr. Andreas Dengel
Leiter Forschungsbereich Smarte
Daten & Wissensdienste

Andreas.Dengel@dfki.de

+49 631 20575 1000

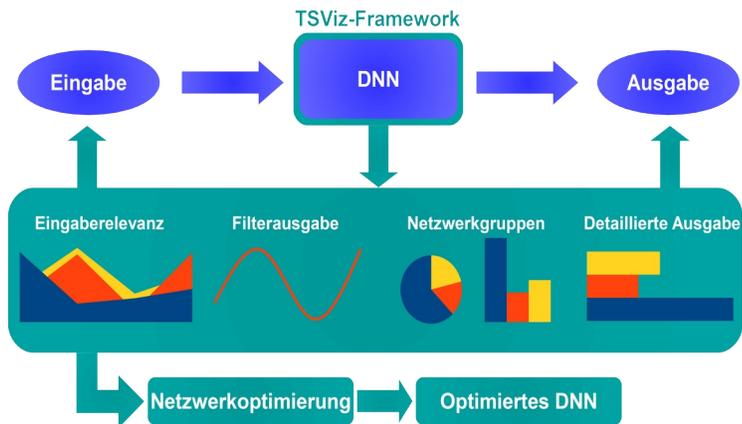
Dr. Sheraz Ahmed
Dominique Mercier
Shoaib Ahmed Siddique
Forschungsbereich Smarte Daten
& Wissensdienste

Sheraz.Ahmed@dfki.de

+49 631 20575 4818

Time Series Visualization Framework – TSViz

Tiefe neuronale Netze werden in einer Reihe von verschiedenen Anwendungen eingesetzt, was eine große Frage nach ihrer Zuverlässigkeit und Robustheit aufwirft. Um das Ziel der Interpretierbarkeit von DNNs in Angriff zu nehmen, hat das DFki einen Rahmen entwickelt, der als Brücke zwischen den Forschern, Entwicklern und Nutzern fungiert, um geeignete Informationen über die Entscheidungen neuronaler Netze für die Zeitreihenanalyse bereitzustellen.



Das Framework ermöglicht es verschiedenen Benutzergruppen, die Vorhersage des Netzwerks als Ganzes sowie die spezifischen Teile des Modells, die dazu beigetragen haben, zu verstehen. Daher werden nur die Informationen präsentiert, die auf den Hintergrund der Benutzergruppe abgestimmt sind. Das Ziel des Frameworks ist es, das Verhalten des Netzwerks zu verstehen, um auf der Grundlage der gewonnenen Erkenntnisse weitere Optimierungen zu ermöglichen.

Heutzutage werden DNNs für verschiedene Aufgaben verwendet und manchmal weiß der Endbenutzer nicht einmal etwas davon, z.B. für die Suche, Empfehlungen und automatisierte Analysen. Für einige sicherheitskritische Bereiche ist neben der Genauigkeit auch die Vertrauenswürdigkeit von extremer Bedeutung. Daher ist es notwendig, ein genaues System zu haben, das die Entscheidung genau erklären kann, d.h. wie und warum eine bestimmte Entscheidung getroffen wurde, insbesondere um Vertrauen für die Fehlerfälle zu schaffen.

Eine Erklärung für eine bestimmte Netzwerkentscheidung zu finden, ist eine anspruchsvolle Aufgabe, die verschiedene Faktoren wie die Domäne, die Zielgruppe und das Datenformat umfasst. Besonders wenn es um die Domäne geht, sind die meisten der vorhandenen Erklärungen für Bilder oder Sprache ausgelegt, da diese Modalitäten für Menschen direkt verständlich sind. Umgekehrt wird dies im Fall von Zeitreihen zu einer bedeutenden Herausforderung, da die Daten in der Zeitreihenmodalität nicht direkt als Bilder und Text verständlich sind, was TSViz zu lösen versucht.

Abdeckung der Zeitreihen-Domäne:

Das Framework agiert als instanzbasiertes Post-Hoc-Verfahren und kann mit jeder beliebigen CNN-Architektur eines neuronalen Netzes umgehen. Darüber hinaus ist es nicht auf eine bestimmte Aufgabe beschränkt und kann sowohl bei der Klassifikation, der Vorhersage als auch für Regression eingesetzt werden.

Benutzerspezifische Erläuterungen ermöglichen die breite Anwendbarkeit:

Durch eine interaktive Schnittstelle ist es möglich, Informationen auszublenken, die nicht auf den Hintergrund der Zielgruppe ausgerichtet sind. Daher kann das System von Entwicklern und Anwendern ohne Einschränkung auf eine bestimmte Gruppe verwendet werden.

Kontakt

Prof. Dr. Andreas Dengel
Leiter Forschungsbereich
Smarte Daten &
Wissensdienste

Andreas.Dengel@dfki.de

+49 631 20575 1000

Dr. Sheraz Ahmed
Forschungsbereich Smarte
Daten & Wissensdienste

Sheraz.Ahmed@dfki.de

+49 631 20575 4818