

ErklärKI RL

Erklärbare künstliche Intelligenz für die Lieferkette

Programm / Ausschreibung	IWI, IWI, Basisprogramm Ausschreibung 2022	Status	abgeschlossen
Projektstart	01.06.2022	Projektende	31.07.2023
Zeitraum	2022 - 2023	Projektaufzeit	14 Monate
Keywords			

Projektbeschreibung

Ziel dieses Projektes ist es, die Vorhersageplattform von Circly weiterzuentwickeln und dabei erklärbare KI (Explainable AI (XAI)) zu integrieren. XAI ist eine künstliche Intelligenz, bei der die Ergebnisse der Lösung für den Menschen nachvollziehbar sind und dadurch Vertrauen in KI-Modelle schafft. Dabei werden XAI-Methoden analysiert und genutzt, um LSTM und Gradient-Boost-KI-Modelle (die Circly bereits entwickelt hat) erklärbar zu gestalten und anschließend die gewonnenen Erklärungsdaten wie Feature Importance, Partial Dependence Plots, etc. in graphische Darstellungen zu überführen, die für die Nutzer der Plattform einfach zu verstehen sind. Diese Darstellungen werden prototypisch in die Plattform von Circly integriert und zu Projektende in einem Workshop mit Nutzern evaluiert. Beispielweise soll die Entwicklung Nutzern im Handel ermöglichen, die Auswirkungen von Marketingaktionen besser zu verstehen und die Gründe der Entscheidung der KI zu verstehen.

Es sollen sowohl modellagnostische XAI-Konzepte wie LIME oder SHAP als auch modellbasierte (für Gradient Boost) evaluiert werden, die aus der Grundlagenforschung stammen und für unsere Zwecke angewandt werden sollen. Auch ist die Recherche über aktuelle Entwicklungen Teil des Projektes.

Herausforderungen des Projektes werden der Zielkonflikt zwischen Erklärbarkeit und Genauigkeit von KI-Modellen, die Aussagekräftigkeit der Erklärungen und das Potential für falsches Verständnis, die globale Erklärbarkeit von Modellen und die Darstellung der Erklärungsdaten selbst sein.

Projektpartner

- Circly GmbH