disy/BMBF

Umweltdaten mit KI erschließen

[15.12.2023] Das BMBF-Forschungsprojekt Simplex4Learning will KI- und Machine-Learning-Ansätze praxistauglicher machen. Das Projektkonsortium untersucht, wie sich heterogene Datenbestände in Umweltbehörden für intelligente Datenanalysen besser nutzen lassen.

Das Bundesministerium für Bildung und Forschung (BMBF) will die heterogenen Datenbestände der Umweltbehörden mit KI- und Machine Learning-Methoden besser für intelligente Datenanalysen erschließen und hat deswegen das Forschungsprojekt Simplex4Learning gestartet. Dies berichtet das auf Datenanalytik spezialisierte Unternehmen disy Informationssysteme, das die Leitung des Projektkonsortiums übernommen hat.
Methoden der Künstlichen Intelligenz (KI) wie Machine Learning (ML) können dazu beitragen, ökologische Phänomene und ihr komplexes Zusammenspiel – etwa bei der Ursachen-Wirkung-Forschung in Waldökosystemen – möglichst gut zu verstehen und zu überwachen. Mit ihnen lassen sich Datenbestände aus verschiedenen Quellen auswerten, fehlende Daten mit Vervollständigungsmethoden generieren, Umweltphänomene prognostizieren oder räumliche Verteilungen besser verstehen. Deren Anwendung erfordert jedoch vertiefte KI-Expertise, die in Umweltbehörden in der Regel nicht verfügbar ist. Hier setzt das im Oktober 2023 angelaufene und auf 30 Monate ausgelegte Forschungsprojekt Simplex4Learning an. Zum Projektkonsortium gehören die Unternehmen disy Informationssysteme, Simplex4Data und die Hochschule für Technik und Wirtschaft (HTW) Berlin. Als assoziierte Partner sind das Landesamt für Natur-, Umwelt- und Verbraucherschutz Nordrhein-Westfalen, die Landesanstalt für Umwelt Baden-Württemberg und der Landesbetrieb Forst Brandenburg eingebunden. Sie unterstützen das Forschungsvorhaben mit Umweltdaten und Anwendungsfällen.

Maschinelles Lernen im Zusammenspiel mit disy Cadenza

Dabei soll für die Analyse-Software disy Cadenza, die in der öffentlichen Verwaltung bereits auf zehntausenden Arbeitsplätzen im Einsatz ist, ein praxistauglicher Ansatz entwickelt werden. Die Idee ist, dass Anwendende aus disy Cadenza heraus über die bestehende Analyse-Erweiterung, die als Schnittstelle zu KI-Verfahren fungiert, Daten, Lern- und Analyseaufgaben senden. Diese wiederum kann auf trainierte KI-Modelle zugreifen, die in einem ML-Repository abgelegt sind. Für das Training der KI-Modelle wird im Projekt ein ML-Framework aufgebaut, das die automatische Erstellung von ML-Modellen mit unterschiedlichen Algorithmen unterstützt. Die ML-Resultate und generierten Ergebnisse werden über die Analyse-Erweiterung an disy Cadenza übergeben und können von den Anwendenden in ihrer Analyseumgebung visualisiert oder in weitere Analyseprozesse eingebunden werden.
Um Umweltdaten für das maschinelle Lernen bereit stellen zu können, hat die Datenhaltung eine Schüsselrolle. Deswegen ist es im Rahmen des Projekts auch vorgesehen, den Simplex-Ansatz des Projektpartners Simplex4Data weiter zu entwickeln. Dies ermöglicht die Datenhaltung in einem einheitlich strukturierten Daten-Pool, der unabhängig von zweckgebundenen Anwendungsfällen ist. Die so strukturierten Umweltdaten können in vorhandene (Geodaten-)Infrastrukturen integriert werden, sodass disy Cadenza darauf zugreifen und sie für die KI-Verfahren bereitstellen kann.

(sib)

https://www.disy.net
https://www.simplex4data.de
https://www.bmbf.de

Stichwörter: Geodaten-Management, BMBF, Disy Informationssysteme, KI, Machine Learning