



SAS erweitert Trustworthy AI-Angebot für mehr Transparenz und Fairness

Model Cards zur Überwachung der Modellqualität, einschließlich Unterstützung für Open Source --- AI Governance Advisory als Service hilft Kunden beim sicheren und ethischen KI-Einsatz

Heidelberg, 29. April 2024 -- SAS, einer der weltweit führenden Anbieter von Lösungen für Daten und künstliche Intelligenz (KI), erweitert seine Produktpalette um umfassende Funktionen für einen verantwortungsvollen Einsatz von KI (" Trustworthy AI "). Model Cards und AI Governance Advisory Services sollen Unternehmen dabei unterstützen, sich im komplexen KI-Umfeld zurechtzufinden, Risiken zu minimieren und ihre Ziele besser umzusetzen - immer im Einklang mit den Richtlinien für einen ethisch vertretbaren Einsatz der Technologie und mit den Verbrauchererwartungen.

"Unsere Kunden sind von den Möglichkeiten, die KI bietet, begeistert - gleichzeitig sind sie aber unsicher, wie sie diese am besten nutzen", erklärt Reggie Townsend, Vice President Data Ethics Practice bei SAS. "Sie stellen die richtigen Fragen in Bezug auf verantwortungsvolle und ethische KI. Unsere Aufgabe ist es, ihnen anhand jahrzehntelanger Erfahrung die Tools und die Beratung zu geben, damit sie KI so integrieren können, dass sie ihre Profitabilität steigern und gleichzeitig Schaden vermeiden."

Model Cards: Der "Nutri-Score" für verantwortungsvolle KI

Als neues Feature werden automatisch generierte Model Cards in SAS Viya integriert, die künftig dafür sorgen, dass alle Beteiligten - vom Entwickler bis zum Vorstandsvorsitzenden - auf dem gleichen Kenntnisstand hinsichtlich der Performance und Zuverlässigkeit von KI-Modellen sind.

Die Model Cards sollen im Laufe des Jahres verfügbar sein und funktionieren ähnlich wie "Nutri-Scores" bei Lebensmitteln: Bewertet werden KI-Modelle im Hinblick auf Genauigkeit, Fairness und Abweichungen, die sich im Laufe der Zeit ergeben (beispielsweise nachlassende Performance). Weiterhin umfasst die Beurteilung Details zu Governance, darunter Zeitpunkt und Verantwortlichkeiten bei der Erstellung, sowie zur Nutzung, einschließlich Use Cases außerhalb des vorgegebenen Rahmens. Auch Nutzungseinschränkungen werden abgebildet - das wird spätestens relevant, wenn Transparenz und Modellprüfung Bestandteil regulatorischer Vorgaben werden. Das Tool generiert die Model Cards automatisch und unterstützt auch mit Open-Source-Programmiersprachen erstellte Modelle, zunächst für Python.

AI Governance Advisory: Orientierung im KI-Dschungel

Die Verbreitung von KI ist nicht mehr aufzuhalten. Gleichzeitig nehmen die Bedenken der Unternehmen zu, wie eine produktive und gleichzeitig sichere Nutzung ihrer Daten nachhaltig möglich ist. SAS bietet seinen Kunden hier mit dem AI Governance Advisory einen neuen Service an, der Klarheit schafft.

Die Beratung beginnt mit einem kurzen Meeting, bei dem SAS zusammen mit dem Unternehmen definiert, was AI Governance speziell für die jeweilige Organisation heißt. Mehrere Unternehmen haben das AI Governance Advisory schon in Pilotversuchen getestet - mit sehr positiven Ergebnissen: höhere Produktivität, besseres Vertrauen in die Datennutzung, größere Chancen für Gewinnung und Bindung von Top-Talenten und ein Boost für das Markenimage. Frühzeitig allen regulatorische Vorgaben zu entsprechen, schafft zudem Wettbewerbsvorteile und erlaubt es Unternehmen, neue Märkte agiler zu erschließen.

Steven Tiell wird den neuen Geschäftsbereich als Global Head of AI Governance bei SAS leiten. Er verfügt über umfassende Expertise für ethische KI, nicht zuletzt als Verantwortlicher für Global Data Ethics and Responsible Innovation Practice bei Accenture.

Zuverlässiges Risikomanagement für KI-Modelle

Darüber hinaus hat SAS anhand von Empfehlungen des US National Institute of Standards and Technology (NIST) einen Trustworthy AI Life Cycle Workflow entwickelt, der Unternehmen das Risikomanagement in Verbindung mit KI erleichtern soll. Sie bekommen ein Produktionsmodell an die Hand, das die Dokumentation ihrer Sorgfaltspflicht umfasst. Damit sind sie in der Lage zu belegen, dass sie faire Modelle verwenden und dass ihre Prozesse keinen Schaden verursachen, beispielsweise indem sie ausschließlich repräsentative Trainingsdaten verwenden, die keine Personengruppen benachteiligen.

Der SAS Trustworthy AI Life Cycle Workflow ist ab sofort über den SAS Model Manager Resources Github verfügbar.

circa 4.100 Zeichen

Pressekontakt

Dr. Haffa & Partner GmbH

Frau Anja Klauck
Karlstraße 42
80333 München

haffapartner.de
postbox@haffapartner.de

Firmenkontakt

SAS Institute GmbH

Herr Thomas Maier
In der Neckarhelle 162
69118 Heidelberg

https://sas.com/de_de/home.html
thomas.maier@sas.com

SAS ist ein weltweit führender Anbieter von Daten und künstlicher Intelligenz (KI), die auch als spezifische Lösungen für verschiedenste Branchen verfügbar sind. SAS macht aus Daten zuverlässige Informationen, mit denen Unternehmen schneller sichere Entscheidungen treffen können. So verschafft SAS seinen Kunden seit 1976 THE POWER TO KNOW.

Firmensitz der US-amerikanischen Muttergesellschaft ist Cary, North Carolina. SAS Deutschland hat seine Zentrale in Heidelberg. Weitere Informationen unter http://www.sas.com/de_de/company-information.html.

Anlage: Bild

SAS Model Manager - Manage Models

< Models Opened Items

Stent Failure Risk Project name: Stent (1.1) Publish

Model Card Files Variables Properties Versions

Stent Failure Risk Version 1.1 ⚠

The purpose of this model is to classify whether patients are at high risk for stent placement failure.

Model Health ⚠ Caution Training Performance Monitoring

Accuracy ⓘ
MCE Pass AUC Pass TPR

Fairness ⓘ
Age Fail Ethnic group

Model Drift ⓘ
Hours active Pass Stent age Pass Event probability

Model modified: August 21, 2023 09:10 AM

Influential Variables ⓘ
The most influential variables ranked by relative impact:

- Stent Material
- GeographicMiss
- Hours Active
- Cell Design
- Stent Age

Fairness Variables ⓘ
Model variables with information privacy classifications: ⓘ

- Age
- Gender
- Ethnic Group

Model variables that have strong correlations with information privacy variables: ⓘ

- Stent material
- Zip code
- Diabetic

Tags: Health Stent Risk Score

Contributors: Jim Box Allie DeLonay

Responsible party: Smith HealthTech Inc.

Project Status: **Deployed** | Limitations documented: **Yes** | Fairness assessment available: **Yes** | Training code type: **SAS**