Anfrage zur schriftlichen Beantwortung E-005390/2020

an die Kommission

Artikel 138 der Geschäftsordnung

Patrick Breyer (Verts/ALE)

Betrifft: Erklärungsversuche von „Black Box“-Modellen des maschinellen Lernens von KI-Programmen und Verwendung interpretierbarer Modelle

Es ist Kritik[[1]](#footnote-0) daran lautgeworden, dass es Anreize für die Privatwirtschaft gibt, intransparente „Black Box“-Algorithmen zu entwerfen und zu vermarkten. Einfache interpretierbare Modelle mit gleichem oder höherem Leistungsniveau wären für jeden frei nutzbar.

1. Was hält die Kommission von dem Vorschlag[[2]](#footnote-1), dass für bestimmte wichtige Entscheidungen von großer Tragweite keine „Black Box“-Modelle zugelassen werden sollten, wenn es bereits ein interpretierbares Modell mit gleichem Leistungsniveau gibt?

2. Was hält die Kommission von dem Vorschlag[[3]](#footnote-2), dass Organisationen, die „Black Box“-Modelle einführen, auch verpflichtet werden sollten, über die Genauigkeit interpretierbarer Modellierungsmethoden Bericht zu erstatten?

1. Rudin, Cynthia, „Hören Sie damit auf, ‚Black Box‘-Modelle des maschinellen Lernens zu erklären zu versuchen, die bei wichtigen Entscheidungen von großer Tragweite eingesetzt werden, und verwenden Sie stattdessen interpretierbare Modelle“, Nature Machine Intelligence [↑](#footnote-ref-0)
2. Siehe oben. [↑](#footnote-ref-1)
3. Siehe oben. [↑](#footnote-ref-2)