Guideline und Methoden für KI in sicherheitskritischen Anwendungen

Künstliche Intelligenz (KI) ist in unserem Alltag weit verbreitet und kommt auch immer mehr in komplexen, sicherheitskritischen Anwendungen zum Einsatz. Wo Menschenleben, Umwelt und hohe Sachwerte auf dem Spiel stehen, kann Fehlverhalten nicht toleriert werden. Die Komplexität Neuronaler Netzwerke und ihre schwer durchschaubaren Entscheidungsprozesse stellen eine der größten Hürden für den Durchbruch der KI in sicherheitskritischen Anwendungen dar.

Dabei ist es möglich, KI sicher einzusetzen. Es bedarf jedoch klarer Richtlinien und Sicherheitsmaßnahmen, um Risiken zu minimieren. Aus diesem Grund hat FEV.io die „Guideline for AI in safety-critical applications“ entwickelt. Diese Guideline fasst wesentliche Grundlagen zusammen, die für eine sichere Anwendung von KI in allen Mobilitätslösungen erforderlich sind.

Die Guideline enthält 15 Kernaussagen, welche sich auf die Bereiche „Functional specification“, „Architectural design“, „Implementation methods“ sowie „Training, testing and validation“ konzentrieren und so den kompletten Entwicklungszyklus technischer Systeme abdecken. Jeder einzelne Punkt mag auf den ersten Blick einfach erscheinen, doch gerade das Zusammenspiel der einzelnen Faktoren ist entscheidend für ein sicheres Systemverhalten. Jede Kernaussage wird mit einer kurzen Erläuterung und wissenschaftlichen Quellen ergänzt.

Die Guideline bietet Ingenieuren und Entwicklern Sicherheit und Orientierung und kann als Checkliste für die Entwicklung sicherer KI-Anwendungen genutzt werden. Mit diesem Ansatz möchte FEV.io die Entwicklung intelligenter Systeme in sicherheitskritischen Bereichen fördern.

Die Guideline ist öffentlich verfügbar unter: https://tinyurl.com/FEVetamax

Auf Basis der Guideline hat FEV.io Methodiken und Konzepte erarbeitet, die eine Entwicklung von Systemen mit KI-Komponenten sowie den Nachweis von sicherem Systemverhalten ermöglichen. Entscheidendes Umdenken ist dabei, dass Sicherheitsanforderungen in viel stärkerem Maße das Systemdesign und die Architektur bestimmen, als dies bei konventionellen Systemen der Fall ist. Ebenso entscheidend ist die gesamtheitliche Betrachtung des Systemverhaltens im Zusammenwirken von konventionellen und intelligenten Komponenten.

Im Vortrag werden mehrere dieser zum Patent angemeldeten Konzepte vorgestellt und anhand von Beispielen veranschaulicht.