



Trendinterview: KI-Algorithmen auch in lokalen Controllern?

Künstliche Intelligenz, Edge und Cloud sind getrennt zu betrachten

Viele Anwender gerade im Maschinenbau setzen bei der Einführung von Industrie4.0-Konzepten auf die Auswertung ihrer Daten auf der Maschinenebene. Mit skalierbarer Rechenleistung in der Steuerung oder im Edge-Device stehen ausreichend Ressourcen für die Datenauswertung direkt vor Ort an der Linie zur Verfügung. Doch wieviel Intelligenz an der Grenze zur Cloud macht Sinn und wie lassen sich die Aufgaben effektiv zwischen Edge und Cloud teilen? Diese Fragen beantwortet Dr. Josef Papenfort im nachfolgenden Trendinterview mit Andreas Gees, stellvertretender Chefredakteur der Elektro Automation.

Elektro Automation: Welche Gründe sprechen für die Integration von KI-Algorithmen in lokale Anwendungen bzw. Steuerungen oder Edge-Devices? Welche Vorteile ergeben sich gegenüber cloudbasierten Lösungen und welche Rolle spielt das Thema Daten-Sicherheit?

Dr. Josef Papenfort: In welchen Bereichen KI-Algorithmen ausgeführt werden, hängt im Wesentlichen von der vertretbaren Latenz und den maximal zulässigen Kosten ab. Ein KI-Algorithmus lebt von Daten. Zu unterscheiden sind hier zum einen die Daten, die beim Trainingsprozess (einmalig) benötigt werden, und zum

anderen die Daten, die kontinuierlich bei der Prädiktion mit einem gelernten Netz (Inferenz-Betrieb) benötigt werden. Sollen Lernen und Inferenz in einer Cloud ausgeführt werden, müssen entsprechend hohe Datenraten zur Verfügung stehen. Bei cloudbasierten Systemen kann dies dann zu hohen Kosten und möglicherweise zu Einschränkungen in der Funktionalität aufgrund zu geringer Bandbreite der Internetverbindung führen. Die Inferenz auf einem Edge-Device erfordert einen einmaligen Invest in die Hardware, kann dann aber in einem lokalen Netzwerk mit hohen Datenraten kommunizieren. Der Lernvorgang hingegen kann sinnvoll auch auf Cloud-Systemen ausgeführt werden. Das Lernen muss nicht kontinuierlich erfolgen, sodass ein einmaliger Upload von Daten genügt. Der Lernprozess erfordert



„Da der Lernprozess ggf. große Rechenressourcen erfordert, die nur temporär benötigt werden, ist dafür das Pay-as-you-go-Bezahlmodell in einer Cloud passend.“

Dr. Josef Papenfort,
Produktmanager TwinCAT bei Beckhoff

aber ggf. große Rechenressourcen, welche nur temporär benötigt werden. Dafür ist das Pay-as-you-go-Bezahlmodell in einer Cloud passend.

Elektro Automation: Während Steuerungsaufgaben meist in harter Echtzeit ablaufen, erfordern KI-Algorithmen nicht zwingend Echtzeit – wie lassen sich diese beiden Anforderungen lokal verbinden?

Dr. Josef Papenfort: Die Art der Verbindung hängt von der Applikation ab, welche durch einen KI-Algorithmus realisiert oder unterstützt werden soll. Klar ist, dass der KI-Algorithmus Daten aus dem Echtzeitprozess benötigt, also ein Datenstrom von der Echtzeit in die Nicht-Echtzeit-Welt realisiert werden muss. Im Wesentlichen muss aber dann die Frage gestellt werden, ob der Nicht-Echtzeit-Algorithmus einen Rückeingriff in den Echtzeitprozess benötigt. Darauf verzichtet werden kann beispielsweise dann, wenn der KI-Algorithmus im Bereich der prädiktiven Wartung eingesetzt wird, beispielsweise eine 'rest of useful life'-Prädiktion ausgeführt werden soll. Hier ist ein direkter Eingriff zurück in den Echtzeitprozess in der Regel nicht nötig. Soll hingegen eine geschlossene Optimierungsschleife realisiert werden, muss der KI-Algorithmus die Resultate wieder mit dem Echtzeitprozess austauschen.

Elektro Automation: Wann stoßen lokale KI-Anwendungen aufgrund der Rechenleistung an ihre Grenzen und welche Anwendungen sind dafür geeignet?

Dr. Josef Papenfort: Diese Frage kann kaum zielführend beantwortet werden. Es hängt von der verwendeten Hardware ab, ob ein Edge-Device die Rechenleistung mitbringt, ein bestimmtes Deep Neural Network in geeigneter Zeit zu berechnen. In PC-basierten Steuerungssystemen kann die Hardware bedarfsgerecht skaliert werden.

Elektro Automation: Gibt es aus Ihrer Sicht auch Anwendungsszenarien, die sich nur in der Cloud sinnvoll abbilden lassen, beispielsweise standortübergreifende Analysen?

Dr. Josef Papenfort: Auch hier ist wiederum die Auftrennung in Lernphase und Inferenz zu betrachten. Das Lernen basiert auf einer Sammlung von Daten. Wird an einem einzelnen Standort nicht genügend Datenvolumen erzeugt, so kann das Koppeln mehrerer Standorte helfen, die den gleichen Prozess abbilden. Dazu ist die Cloud ein geeignetes Bindemittel. Zudem stellt sie die Rechenleistung zur Verfügung, um zu einem bestimmten Zeitpunkt den Lernvorgang in Gang zu setzen. Ob die Inferenz dann auch in der Cloud stattfinden soll, hängt – wie oben erwähnt – von der konkreten Applikation, der Latenz und der Datenrate ab.