



ENTWICKLUNG EINES AUF DATA SCIENCE ANALYSEN UND NEURONALEN NETZEN BASIERENDEN SOFTWARE SYSTEMS ZUR MENGENBEDARFSBESTIMMUNG VON ENTWICKLUNGSFAHRZEUGEN IN DER FAHRZEUGENTWICKLUNG

TOBIAS DJERMESTER (MASTERSTUDIUM INFORMATIK)

Betreuer: Prof. Dr. Jochen Schmidt, Prof. Dr. Gerd Beneken

Einleitung

Die Fahrzeugentwicklung ist risikoreich. Auf Fehler in der Entwicklung folgen oft hohe Kosten. Entscheidend für den Entwicklungserfolg und zugleich grundlegender Bestandteil jedes Fahrzeugentwicklungsprojektes ist die Absicherung sowie die initial vorgenommene Absicherungsplanung. Die Absicherung des Entwicklungsprozesses wird zyklisch durch die Erprobung und den Test von Bauteilen und des Gesamtfahrzeuges vorgenommen. Dabei werden Ressourcen benötigt, um insbesondere die Anforderungen der Bauteile separat als auch die Integration der Bauteile innerhalb des Gesamtfahrzeuges zu validieren. Den größten Anteil des in der Absicherung benötigten Ressourcenbedarfs stellen Entwicklungsfahrzeuge dar.

Die herkömmliche Absicherung beruht auf der manuellen Planung und Prognose des für die Absicherung benötigten Ressourcenbedarfs. Eine Übernahme oder Unterstützung des Planungsprozesses und damit eine Optimierung resultierender Planungen ist durch den Einsatz von Künstlicher Intelligenz auf Grundlage vergangener Entwicklungsprojektdaten denkbar.

Im Rahmen der Arbeit wird ein auf Künstlicher Intelligenz basierendes Softwaresystem zur Prognose des Entwicklungsfahrzeugbedarfs für neue Fahrzeug- und Antriebsprojekte entwickelt. Dabei wird die Möglichkeit als auch das Potential der Verwendung Neuronaler Netze in der Bedarfsplanung untersucht.

Vorgehensweise

Die Auswahl und Aufbereitung von Daten ist Grundlage und fundamentaler Bestandteil in der Entwicklung des Planungssystems. Um eine Reduktion der

Merkmalsmenge im Verhältnis zur verfügbaren Datenmenge zu erzielen, wird die Auswahl und Bewertung verschiedener Merkmalsvektoren vorgenommen. Anschließend werden Zielwerte der KI auf die benötigten Ergebnisse der Bedarfsplanung ausgelegt und entsprechend einer detaillierten Verteilung des Fahrzeugbedarfs auf Teilbereiche der Entwicklung aufgebaut. Durch die Adaption von Hyperparametern und die Optimierung der Topologie der KI wird die Leistung angewandter Algorithmen weiter optimiert. Damit die Prognose des Fahrzeugbedarfs durch Fachkräfte der Bedarfsplanung durchgeführt werden kann, wird die Entwicklung einer durch eine graphische Benutzerschnittstelle geführte Anwendungssoftware vorgenommen. Änderungen und Erweiterungen des Systems werden durch die Entwicklung eines Erzeugungs- und Bereitstellungsprozesses neuer Softwareversionen der Anwendungssoftware auf Grundlage aktualisierter Daten- und Trainingszustände der KI umgesetzt. In Vorbereitung für zukünftige Untersuchungen neuer Datenstände des Softwaresystems wird eine Analysekomponente integriert, welche Analyseergebnisse anschaulich aufbereitet und wiederkehrende Analysedurchläufe der zugrunde liegenden Datenbasis ermöglicht.

Analyse

Bei der allgemeinen Problemstellung handelt es sich um ein Regressionsproblem. Aus einem mehrdimensionalen Merkmalsvektor werden stetige quantitative Ergebniswerte abgeleitet. Im Rahmen der Analyse werden strukturiert mehrere Schritte durchlaufen.

Um eine Reduktion der Datenkomplexität und eine Optimierung des Verhältnisses von Datenmenge zur berücksichtigten Merkmalsmenge zu schaffen,

wird die Untersuchung verschiedener Merkmalskombinationen vorgenommen. Dabei wird im ersten Schritt der Analyse ein Merkmalsvektor der Größe 31 ausgewählt. Im Rahmen einer Korrelationsanalyse wird die Beziehung der Merkmalswerte untereinander als auch ihre Auswirkung auf Zielwerte genauer untersucht. Abbildung 1 zeigt Korrelationen mittels einer Korrelationsmatrix, erzeugt über Merkmale und Zielwerte der Problemstellung, graphisch auf.

Für die Ermittlung der Relevanz individueller Merkmale wird das Random Forest-Klassifikationsverfahren eingesetzt. Dabei wird die Relevanz der Merkmale unter Bestimmung der Gini

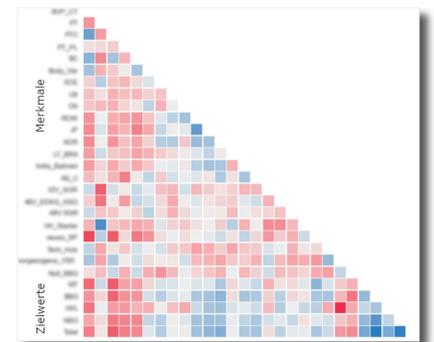


Abbildung 1: Korrelationsmatrix über Merkmale und Zielwerte

Importance (Mean Decrease in Impurity) sowie der Permutation Feature Importance (Mean Decrease in Accuracy) auf Basis berechneter Entscheidungsbäume bestimmt. Das Ergebnis individueller Merkmale wird zur Ermittlung potentieller Kombinationsmöglichkeiten des Merkmalsvektors herangezogen. Im Rahmen dessen werden Merkmalsvektoren unterschiedlicher Größe und Zusammensetzung erstellt und entsprechend ihrer Auswirkung auf den relativen Fehler korrespondierender Neuronaler Netzwerke untersucht. Der

ROSENHEIMER INFORMATIKPREIS INF-MASTER

Merkmalsvektor wird durch Reduktion um Merkmale einer geringen Relevanz und entsprechend einer Optimierung der Ergebnisgenauigkeit Neuronaler Netzwerke auf eine Größe von 20 Merkmalen reduziert.

Ausgaben der Algorithmen der Künstlichen Intelligenz können durch unterschiedliche Ergebnisvektoren repräsentiert werden. Abhängig von dem Anwendungsfall können Ergebnisvektoren beliebig groß aufgebaut sein. Nachfolgend wird der Einfluss unterschiedlicher Variationen des Ergebnisvektors auf die Ergebnisgenauigkeit untersucht. Dabei wird die Gruppierung von Zielwerten in kategorischen Einheiten von Fachbereichen und Entwicklungsphasen vorgenommen. Ein Optimum der Konstellation des Ergebnisvektors wird, unter Beurteilung von dessen Eignung in der Problemlösung und der Ergebnisgenauigkeit Neuronaler Netzwerke, ermittelt. Dabei stellt sich eine Granularität von vier Entwicklungsphasen und neun Fachbereichen je Entwicklungsphase als Optimum heraus. Die Ergebnisgenauigkeit kann darüber hinaus durch Anwendung einer relativen Verteilung des Fahrzeugbedarfes je Fachbereich gegenüber einer absoluten Verteilung weiter gesteigert werden.

Hinsichtlich der Analyse und Selektion einer Künstlichen Intelligenz werden Neuronale Netzwerke bezüglich ihrer Topologie und Hyperparameter untersucht. Neuronale Netzwerke können durch eine unterschiedliche Ausprägung und Anordnung von Schichten und Neuronen gekennzeichnet sein. Dabei können Neuronale Netzwerke verschiedene Formen und Strukturen annehmen. Im Verlauf der Analyse wird sowohl die Anzahl der Schichten als auch die Größe jeder Schicht durch die automatisierte Erweiterung um neue Neuronen analysiert. Dabei wird eine Reduktion oder ein Anstieg des Fehlers der Ausgaben Neuronaler Netzwerke gemessen. Das Ergebnis legt nahe, dass der minimal zu erzielende Fehler unter den getesteten Netzwerktopo-

logien bei zwei verdeckten Schichten erzielt werden kann. Neben der Topologie des Netzwerkes werden Optimierungen an Trainings- und Netzwerkparametern vorgenommen, um den Lernerfolg weiter zu steigern.

Entwicklung

Die Anwendung und Bedienung der Künstlichen Intelligenz im Umfeld der Bedarfsplanung erfordert die Bereitstellung einer abstrakten Funktionalität für die Ausführung von Bedarfsprognosen durch Fachkräfte. Im Rahmen dessen wird eine Anwenderkomponente inklusive graphischer Benutzerschnittstelle zur Ausführung von Funktionen der Künstlichen Intelligenz im Arbeitsumfeld erstellt. Abbildung 2 stellt die implementierte Benutzerschnittstelle der Anwenderkomponente dar.

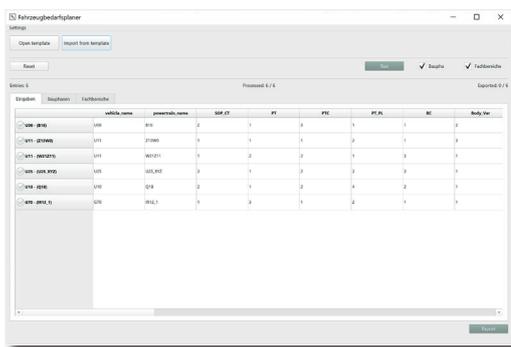


Abbildung 2: Anwenderkomponente für die Fahrzeugbedarfsplanung

Das Tool verfügt über Funktionen der Eingabe, Verarbeitung und Ausgabe von Bedarfsplanungen und Prognoseergebnissen. Die Prognose wird durch die trainierte Instanz der Künstlichen Intelligenz vorgenommen. Dabei wird ein Wrapper implementiert, welcher primäre Funktionen der Bedienung der Künstlichen Intelligenz abbildet.

Für eine stetige Erweiterung und Optimierung des Systems werden Erzeugungs- und Bereitstellungsprozesse der Anwenderkomponente entwickelt. Bereitstellungsprozesse sind in der Lage, neue Versionen der Anwenderkomponente in Relation zu neuen Trainings- und Datenzuständen der KI automatisiert zu erzeugen sowie die Benutzerschnittstellen und interne

Verarbeitungsprozesse auf Veränderungen in den Daten- und Trainingszuständen auszulegen. Prozesse werden unter Linux auf dem Server des Systems ausgeführt.

In Vorbereitung für zukünftige Untersuchungen neuer Datenstände wird eine Analysekomponente entwickelt. Implementierte Analyseskripte greifen auf gemeinsam genutzte Funktionen der Komponente zu und stellen verschiedene Analysevorgänge zur Aufbereitung und Selektion von Daten der KI bereit.

Zukunftsausblick

Entscheidend in der Beurteilung, ob zukünftige Daten für die Entscheidungsfindung geeignet sind, ist ihre Vorhersagekraft, nicht nur ihre Fähigkeit, die Vergangenheit darzustellen. So können Innovationen, wie neue Antriebsformen oder das autonome Fahren, auf Merkmalen und Merkmalskombinationen basieren, welche vom aktuellen Stand des Systems nicht abgebildet werden. Das Auslegen der Merkmale und Zielwerte des Systems auf die Bedarfsprognose zukünftiger Technologien ist entscheidend für die zukünftige Eignung des Systems.

Hinsichtlich der Erweiterungen der Datenbasis kann die Verwendung des Systems im Kontext der Datengewinnung bedacht werden. Die Überprüfung und Adaption der Planungsergebnisse sollten vorgenommen werden oder die Nutzung des Systems für den Erhalt neuer Daten eingeschränkt werden.

Bedarfsplanungen werden vor Entwicklungsbeginn durchgeführt, um den vermeintlich optimalen Fahrzeugbedarf (Soll-Wert) für die Abwicklung des Fahrzeugprojektes vorherzusagen. Der tatsächlich genutzte Fahrzeugbedarf bleibt dabei unberücksichtigt. Durch zukünftige Entwicklungen kann versucht werden, Plandaten zu adaptieren und somit langfristig eine Auslegung der Datenbasis auf den genutzten Fahrzeugbedarf (Ist-Wert) zu realisieren.