

Einsatz von intelligenten Agenten-Ensembles im Produktentstehungsprozess

Automatisches Zuordnen von Bauteilen zu Montagezonen bei der Prototypenmontage

Dr.-Ing. **A. Lommatzsch**, Technische Universität Berlin, Berlin

Dipl.-Inf. **T. Plumbaum**, Technische Universität Berlin, Berlin

C. Sauer, Technische Universität Berlin, Berlin

Prof. Dr. **S. Albayrak**, Technische Universität Berlin, Berlin

Kurzfassung

Die Entwicklung neuer Fahrzeuge wird in zunehmendem Maße durch rechnergestützte Verfahren (z. B. Simulationen) erleichtert und virtualisiert. Trotz der fortschreitenden Technisierung sind das Wissen und die Menge der benötigten Informationen im Produktentstehungsprozess (PEP) immens und erfordern weiterhin auch eine manuelle Bearbeitung von Daten. So verbraucht beispielsweise bei der virtuellen Montage von Prototypen das händische Zuordnen von Bauteilen in die zugehörigen Strukturen (Montagezonen) und das Prüfen der Merkmale (z. B. der Übereinstimmung mit aktuellen Normen) einen Großteil der Zeit, die eigentlich zur Überprüfung der Montierbarkeit der Fahrzeuge verwendet werden sollte. Dabei bieten sich im Produktentstehungsprozess vielfältige Möglichkeiten zur Effizienzsteigerung, z. B. indem die Zonenverantwortlichen durch den Einsatz von Methoden der Künstlichen Intelligenz unterstützt werden und das in bereits abgeschlossenen Projekten erworbene Wissen auf neue Fahrzeugprojekte übertragen wird.

Im Rahmen eines Projektes im Volkswagen Konzern wurde untersucht, wie durch die Kombination von Multi-Agententechnologien mit Verfahren zur automatischen Klassifikation von Bauteilen die Zuordnung von Bauteilen für ein neues Fahrzeugmodell zu den Montagezonen eines Fahrzeugs optimiert werden kann. Dabei wurden für die einzelnen Zonen spezialisierte Agenten erstellt, die jeweils einen auf die Anforderungen der jeweiligen Zone abgestimmten Klassifikationsalgorithmus einsetzen. Darüber hinaus wurde analysiert, wie die Ergebnisse mehrerer Agenten in einem Ensemble kombiniert werden können, um so die Klassifikationsgüte zu verbessern.

Die durchgeführte Evaluation zeigt, dass der gewählte Ansatz eine präzise Klassifikation von Bauteilen erlaubt und damit aus abgeschlossenen Projekten gewonnenes Wissen effektiv für die Entwicklung neuer Modelle herangezogen werden kann. Bauteile, für die aufgrund unzureichender Informationen keine eindeutige Zuordnung errechnet werden kann, werden markiert, um interaktiv mit dem Konstrukteur eine korrekte Klassifikation vorzunehmen. Dies ermöglicht es, den Produktentstehungsprozess effektiver zu gestalten. Durch das Berücksich-

tigen zusätzlicher abgeschlossener Projekte sowie das Einbeziehen des Feedbacks von Konstrukteuren, können die Agenten inkrementell nachtrainiert werden, um so die Klassifikationsgenauigkeit stetig zu verbessern.

Abstract

Nowadays the construction of new automotive models is supported by various computer based tools (e. g. Simulations). Nevertheless, the amount of information needed for the construction of new models still requires manual merging and arranging of data (e.g. inventory lists, hook lists). The manual assignment of components into predefined structures (zones) and the examination of features (e.g. compliance with current norms) consume a majority of the time, which actually should be used for designing the vehicle itself. Therefore, the product creation process offers various possibilities to support the technical designer, e.g. by the application of AI methods and by transferring knowledge acquired in previously finished projects to the product creation process of new automotive models.

In a project at the Volkswagen AG we researched how the combination of multi-agent technologies with methods of artificial intelligence can be used for assisting and optimizing the product creation process. The focus was put on the automatic assignment of components of new automotive models to predefined zones. For each zone, the developed system provides specialized agents deploying classification algorithm trained to meet the requirements of the respective zone. Beyond that, we analyzed, how the classification results of different agents, so called agent ensembles, can be combined, in order to improve the overall classification quality.

The evaluation shows, that the chosen approach provides a high precision. The data collected from previous projects provide a good basis for learning high quality classification strategies. Automotive components, for that a classification is impossible due to insufficient information, are marked, in order to allow a manual classification by the technical designer. This allows us to optimize the classification process by including additional projects as well as feedback from designers. Thus, the classification algorithms used by the agents can be constantly improved, considering additional data and feedback.

1. Motivation

Bei der Konstruktion neuer Fahrzeugmodelle besteht ein Großteil der Arbeit innerhalb des Produktenstehungsprozesses (PEP) aus der Suche nach Informationen, die zur Überprüfung der Montierbarkeit eines Fahrzeuges nötig sind. Das traditionelle Vorgehen erfordert die Suche nach Bauteilen in verschiedenen Dokumenten (Hakenliste, Baustandsliste, Normteillis-

ten, Baugruppen-Blätter, etc.) und das manuellen Zuordnen von Bauteilen zu den Montagezonen. Dieses Suchen und Zuordnen der richtigen Bauteile und der notwendigen Verbindungselemente für jeden Konstruktionsschritt ist nicht trivial und nimmt einen Großteil der Zeit in Anspruch. Die Entwicklung einer intelligenten Softwarelösung ermöglicht daher die Effizienz der virtuellen Montage zu erhöhen, die Mitarbeiter von Routineaufgaben zu entlasten und den Konstruktionsprozess dadurch erheblich zu beschleunigen.

2. Problembeschreibung

Im Rahmen des Projektes wird ein Optimierungsproblem betrachtet. Der Produktentstehungsprozess soll effizienter gestaltet werden, indem die Menge von Bauteilen eines neuen Fahrzeuges den einzelnen Montagezonen automatisch zugeordnet werden. Als Montagezonen eines Fahrzeugs lassen sich beispielsweise Innenraum, Getriebe oder Fahrertür anführen.

Zum Erlernen einer Zuordnungsstrategie stehen den Daten bereits abgeschlossener Fahrzeugprojekte zur Verfügung. Dabei ist zu beachten, dass einige Fahrzeugprojekte nur ausgewählte Bereiche eines Fahrzeugs berücksichtigen (z. B. nur Karosserie oder nur den Innenraum), die auf Montierbarkeit überprüft werden sollten.

Die Bauteile eines Fahrzeugs werden jeweils durch eine Menge von Merkmalen beschrieben. Wichtige Bauteilattribute sind dabei die eindeutige Bauteilnummer (ID), der Name des Bauteils, die Art des Bauteils, die textuelle Beschreibung sowie die jeweiligen Abmessungen der einzelnen Teile.

Das Erlernen einer Strategie zum Zuordnen von Bauteilen zu den Zonen eines Fahrzeugs erfordert es, das Gesamtproblem in mehrere Teilprobleme zu zerlegen.

Erstellen einer geeigneten Trainingsmenge Es muss ermittelt werden, welche Trainingsdaten für die aktuelle Lernaufgabe besonders gut geeignet sind. Zu große Trainingsmengen erhöhen den Trainingsaufwand; nicht zur jeweiligen Klassifikationsaufgabe passende Bauteile in der Trainingsmenge führen zu einer schlechten Zuordnungsqualität, da das notwendige Wissen in der Trainingsdatenmenge nicht vorhanden ist.

Ermitteln der Merkmale, die für eine Zuordnungsentscheidung relevant sind Aus der Vielzahl der zur Bauteilbeschreibung verfügbaren Daten müssen die Bauteil-Attribute ausgewählt werden, die für das Ableiten einer Klassifikationsentscheidung besonders geeignet sind. Zu berücksichtigen sind dabei die unterschiedlichen Arten von Merkmalen (die Bauteilnummer ist eine Zahl, die Bauteil-Beschreibung ein längerer Text) sowie mögliche Vorverar-

beitungsschritte für Merkmalswerte, wie z. B. das Aufteilen von komplexen Merkmalen (z. B. Bauteilnummern) in mehrere kurze Bestandteile. Durch die Verwendung geeigneter Vorverarbeitungsschritte kann die Generalisierungsfähigkeit von erlernten Klassifikationsstrategien verbessert werden.

Auswahl der Lernalgorithmen und Finden geeigneter Konfigurationen Bei der Wahl eines der Lernalgorithmen muss berücksichtigt werden, dass der Klassifikationsalgorithmus zum Typ der gewählten Attribute passt und dass die zum Lernen einer Strategie notwendigen Ressourcenanforderungen akzeptabel sind. Darüber hinaus ist zu beachten, dass Normteile (z. B. Schrauben) mehreren Zonen zugeordnet werden können. Das bedeutet, dass Klassifikationsalgorithmen die Bauteile jeweils nur genau einer Zone zuordnen können, nicht direkt verwendet werden sollten, sondern für die Mehrklassenzuordnung erweitert werden müssen. Alternativ können Klassifikatoren (die jeweils für eine Klasse entscheiden) mit weiteren Klassifikatoren kombiniert werden ("Ensemble-Klassifikatoren"). Bei der Auswahl von Klassifikationsalgorithmen sollte zudem beachtet werden, dass die ausgewählten Algorithmen eine gute Generalisierungsfähigkeit aufweisen und möglichst Erklärungen für getroffene Klassifikationsentscheidungen bereitstellen sollten.

3. Ansatz

Im Rahmen des Projektes wird ein Software-Framework entwickelt, mit dem verschiedene Klassifikationsverfahren und Parameterkonfigurationen getestet und evaluiert werden können. Grundlage des Systems bildet eine Multi-Agentenplattform auf Basis des JIAC Frameworks [7], die es erlaubt, Klassifikationsalgorithmen als Agenten zu repräsentieren und die Ergebnisse mehrerer Agenten zu aggregieren. Es besteht die Möglichkeit, verschiedene Strategien zur Kombination von Klassifikatoren sowie die Güte unterschiedlicher Parameterkonfigurationen zu untersuchen. Die jeweils durch die Klassifikationsagenten erreichte Zuordnungsqualität wird anhand der aus abgeschlossenen Projekten verfügbaren Daten untersucht.

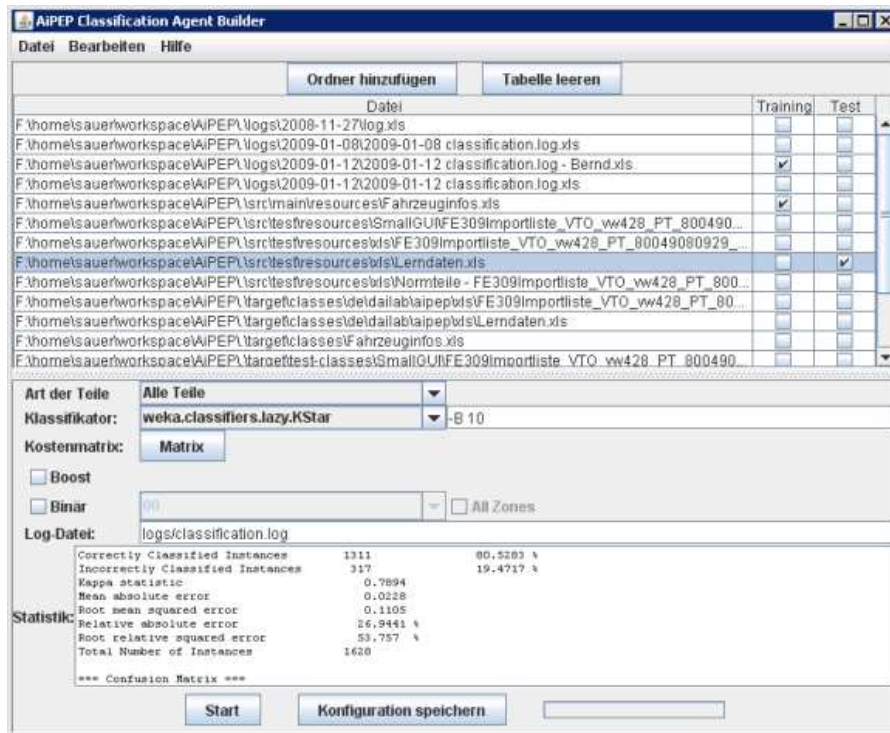


Abbildung 3.1: Das Menü zum Zusammenstellen der Trainings Szenarien und zum Evaluieren der gelernten Klassifikationsstrategien

Unter Verwendung des entwickelten Frameworks wurden die für das Lernen einer guten Zuordnungsstrategie relevanten Fragestellungen untersucht. Die durchgeführten Analysen werden in den folgenden Abschnitten detailliert erläutert, dabei wird auf die in Kapitel 2 genannten Teilprobleme detailliert eingegangen.

Ermitteln einer geeigneten Trainingsmenge Es wird analysiert, welchen Einfluss die zum Trainieren von Zuordnungsstrategien verwendete Menge von Bauteilen auf die Zuordnungsqualität hat. Eine zentrale Fragestellung ist dabei, ob jeweils eine auf ein bestimmtes Lernszenario optimierte Trainingsmenge oder eine umfassende, allgemeine Trainingsmenge verwendet werden sollte.

Wahl des Klassifikationsalgorithmus Es werden unterschiedliche Arten von Klassifikationsalgorithmen untersucht. Wichtige Kriterien für die Auswahl der Algorithmen ist die erreichte Zuordnungsqualität, der Bedarf an Ressourcen sowie die Nachvollziehbarkeit (Erklärung) von Zuordnungsentscheidungen. Im Rahmen des Projektes wurden vier Klassifikationsalgorithmen analysiert, die jeweils unterschiedliche Ansätze verfolgen: (1) Der Algorithmus KStar [1] basiert auf dem Prinzip der Analyse der k nächsten Nachbarn (kNN). (2) Der Algo-

rithmus C4.5 [2] erstellt anhand einer Trainingsmenge einen Entscheidungsbaum. Die Auswahl der Attribute ("Features"), die für die Erstellung von Knoten herangezogen werden, erfolgt auf Basis eines entropiebasierten Maßes ("Transinformation"). Zur Vereinfachung der gebildeten Entscheidungsbäume und zur Reduzierung von Overfitting können Teilbereiche des Entscheidungsbaumes zu Blättern zusammengefasst werden ("Error Pruning"). Entscheidungsbäume bieten eine akzeptable Klassifikationsgenauigkeit bei guter Nachvollziehbarkeit der gelernten Zuordnungsstrategie. (3) Als probabilistischer Klassifikationsalgorithmus wird der Bayes-Net-Algorithmus [3] berücksichtigt. Der Algorithmus verwendet Bayesianische Netze und trifft die Klassifikationsentscheidungen anhand der Verknüpfung von bedingten Wahrscheinlichkeiten. Der Algorithmus ermöglicht auch bei einer begrenzten Ressourcenverfügbarkeit präzise Zuordnungsergebnisse. (4) Es wird eine Support-Vektor-Maschine [4,5] zum Lernen einer Klassifikationsstrategie getestet. Support Vector Maschinen unterteilen eine Menge von Objekten so in Klassen, dass um die Klassengrenzen herum ein möglichst breiter Bereich frei von Objekten bleibt ("Large Margin Classifier"). Die Support Vector Maschine passt dabei in den Raum der Trainingsvektoren eine Hyperebene (die als Trennfläche fungiert) so an, dass der Abstand der Vektoren, die der Hyperebene am nächsten liegen, maximiert wird. Der breite Rand ("Large Margin") soll sicherstellen, dass auch neue Objekte möglichst zuverlässig klassifiziert werden. Support-Vektor-Maschinen ermöglichen in vielen Einsatzszenarien eine sehr gute Klassifikationspräzision, stellen aber oft hohe Anforderungen an Ressourcen.

Auswahl der für die Klassifikation verwendeten Merkmale Die aus abgeschlossenen Projekten zur Verfügung stehenden Daten werden in ein für die Klassifikatoren geeignetes Format gebracht. Die den Zonen des Fahrzeugs zuzuordnenden Bauteile werden jeweils durch einen Merkmalsvektor beschrieben. Dabei werden komplexe Bauteil-Merkmale in mehrere einfache Attribute zerlegt, um den unterschiedlichen Informationsgehalt komplexer Merkmale bestimmen zu können. Die Auswahl der jeweils am besten für die Klassifikation geeigneten Attribute wird durch die einzelnen Klassifikationsalgorithmen durchgeführt (z. B. anhand der Entropie der auftretenden Merkmalswerte).

Modellierung der Klassifikationsaufgabe Zum Lernen der Klassifikationsstrategie werden verschiedene Ensemble-Agenten analysiert, wobei ein Agent jeweils eine Klassifikationsstrategie repräsentiert.

Zunächst wird der Fall analysiert, dass nur ein einzelner Klassifikationsagent zum Einsatz kommt. Es wird dabei für jedes Bauteil eine eindeutige Zuordnung zu einer Fahrzeugzone

gelernt. Bauteile, die in mehreren Zonen verwendet werden, ordnet man dabei jeweils der Zone zu, in der sie anteilmäßig am häufigsten auftreten.

Um das Problem der Zugehörigkeit eines Bauteils zu mehreren Zonen des Autos zu behandeln, wurde untersucht, welche Klassifikationsgüte erreicht wird, wenn ein Ensemble aus binären Klassifikations-Agenten eingesetzt wird [6,7]. Dazu wird für jede Zone des Fahrzeugs ein Agent trainiert, der jeweils auf "seine" Zone optimiert ist und jeweils entscheidet, ob ein Bauteil zu einer Klasse gehört oder nicht. Da die Agenten unabhängig voneinander ihre Entscheidungen treffen, kann ein Bauteil mehreren Zonen zugeordnet werden. Zudem können unterschiedliche Algorithmen eingesetzt werden, um die Besonderheiten der jeweiligen Zonen des Fahrzeugs zu berücksichtigen.

Um die Klassifikationspräzision zu verbessern, wird ein Agent zur Koordination der Klassifikationsagenten eingesetzt [8]. Dieser analysiert die Klassifikationsentscheidungen und die von den einzelnen Agenten ermittelten Daten über die Zuverlässigkeit der Entscheidungen, um durch Zusammenführen der Daten eine hochqualitative Entscheidung abzuleiten. Bauteile, die in der Menge zu zuzuordnenden Teile mehrfach (n -fach) vorkommen, können so geeignet mehreren (m) Fahrzeugzonen zugeordnet werden (wobei $m \leq n$).

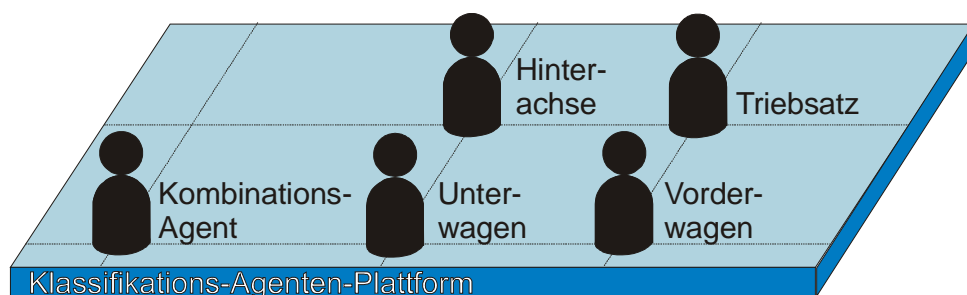


Abbildung 3.2: Ein Ensemble von auf verschiedene Zonen des Fahrzeugs spezialisierten Agenten

Durchgeführten Experimente und eingesetzte Hardware Es wurden im Rahmen des Projektes vielfältige Szenarien und Konfigurationen (z. B. unterschiedliche Strategien zum Zusammenstellen der Trainingsmenge, verschiedene Ansätze zum Definieren der Attributwerte der Bauteile, Beschränkung auf eine oder mehrere Zonen eines Fahrzeugs) analysiert, um die Leistungsfähigkeit der betrachteten Algorithmen zu evaluieren. Die Experimente wurden auf einem Standard-Desktop-Computer durchgeführt (Pentium D, 3 GHz, 4 GB RAM). Die Konfiguration wurde jeweils so gewählt, dass mit den begrenzten Ressourcen in akzeptabler Zeit Ergebnisse ermittelt werden können.

4. Evaluation

Die Evaluation der getesteten Verfahren zeigt, dass mit den eingesetzten Verfahren des Maschinellen Lernens eine hohe Zuordnungsgenauigkeit von Bauteilen erreicht werden kann. Die für ein neues Fahrzeugmodell erreichte Klassifikationsgüte hängt dabei vor allem davon ab, wie umfangreich die zum Trainieren des Klassifikators genutzte Trainingsmenge von Bauteilen ist. Aus einer umfangreichen Trainingsmenge extrahieren die verwendeten Klassifikationsalgorithmen die Informationen, die für die Zuordnungsaufgabe benötigt werden.

Vergleich der Klassifikationsalgorithmen Für die Klassifikation von Bauteilen wurden verschiedene Klassifikationsverfahren untersucht.

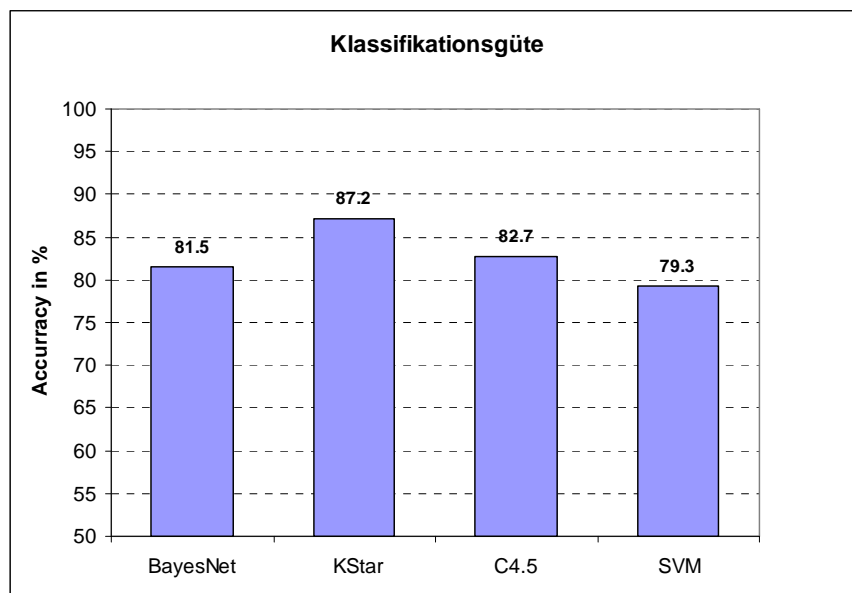


Abbildung 4.1: Vergleich der mit verschiedenen Klassifikationsalgorithmen erreichten Klassifikationsgüte (getestetes Fahrzeug: VW Sharan)

Die Evaluation zeigt, dass die besten Ergebnisse mit einem Nearest-Neighbour-Klassifikator (KStar) erreicht werden. Die Algorithmen Bayes-Net und C4.5 liefern eine geringfügig geringere Klassifikationsgüte, wobei insbesondere C4.5 weniger stabil erscheint (d. h. die erreichte Klassifikationsgüte variiert bei Versuchen mit ähnlichen Bedingungen). SVMs wurden aufgrund des hohen Bedarfs an Ressourcen und der zu langsamen Erstellung von Klassifikationsstrategien nicht weiter betrachtet (die SVM benötigte bei gleicher Aufgabe 10000 mal so lange wie der Klassifikationsalgorithmus C4.5).

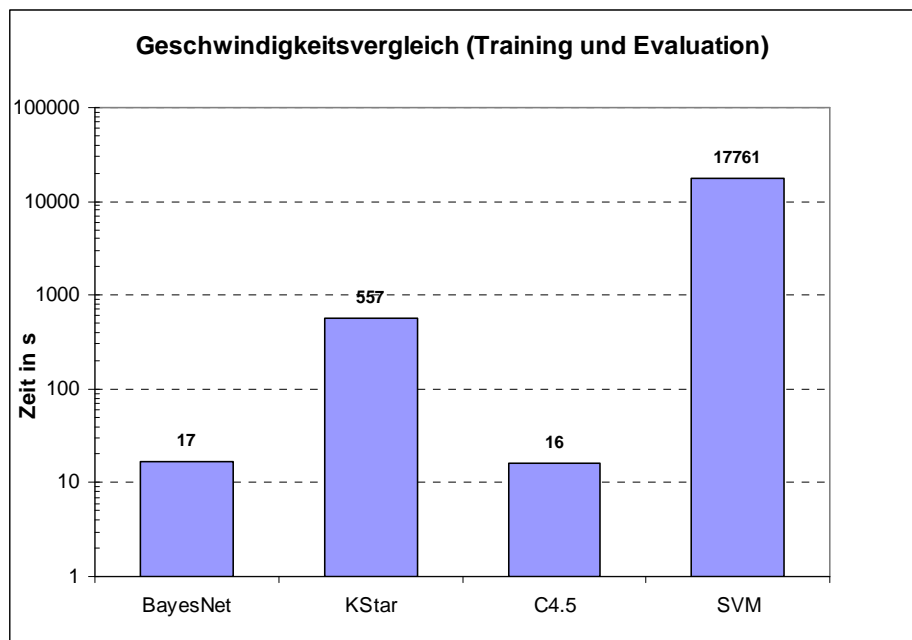


Abbildung 4.2: Vergleich des zeitlichen Aufwandes für Training und Evaluation der Klassifikatoren (getestetes Fahrzeug: VW Sharan)

Bei den betrachteten Klassifikationsalgorithmen müssen die Parameter jeweils so gewählt werden, dass eine geeignete Generalisierung gelernt werden kann. Es muss sichergestellt werden, dass keine zu starke Anpassung an die Trainingsmenge erfolgt ("Overfitting"); aber trotzdem eine Strategie zum Unterscheiden der Zielklassen gelernt wird. Vergleicht man die mit dem KStar-Algorithmus erreichte Klassifikationsgüte mit der, die durch reines Nachschlagen der Bauteile in der Trainingsmenge (Auswendiglernen) erreicht wird, so stellt man fest, dass der Klassifikationsalgorithmus im Allgemeinen eine zwischen 10% und 20% höhere Zuordnungsqualität erreicht.

Für die Analyse der Leistungsfähigkeit der Klassifikatoren analysiert man die Konfusionsmatrizen (vgl. Abbildung 4.3). Diese Matrizen zeigen für jede Zone an, wie viele Teile jeweils korrekt zugeordnet werden und zwischen welchen Zonen es Klassifikationsfehler gibt.

	Karosserie:Plattform	Karosserie:Hut	Kabelbaeume	Cockpit	Unterrwagen	Vorderwagen	Hinterachse	Vorderachse	Triebsatz	Hinterwagen	Innenwagen	Frontend	Klappen	Frontklappe	Tür:vorne.links	Tür:vorne.rechts	Tür:hinten.links	Tür:hinten.rechts	Heckklappe	Stopfen	unstrukturiert	Karosserie:Bolzen			
143	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<-- classified as	
3	136	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Karosserie:Plattform	
0	0	32	0	2	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Karosserie:Hut	
4	10	0	30	0	0	0	1	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	Kabelbaeume	
0	0	0	0	173	2	3	1	0	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	Cockpit	
0	2	1	0	115	2	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	1	0	Unterrwagen	
0	0	0	0	2	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Vorderwagen	
0	0	0	0	0	0	0	111	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Hinterachse	
0	0	0	0	1	0	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Vorderachse	
0	0	0	2	2	0	0	0	0	47	2	0	0	0	0	4	1	0	0	0	0	0	0	0	Triebsatz	
0	0	1	1	3	0	0	0	0	0	200	4	0	1	1	0	0	0	1	0	0	0	0	0	Hinterwagen	
1	1	0	0	2	0	0	0	0	0	1	101	0	0	0	0	0	0	0	0	0	0	0	0	Innenwagen	
0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Frontend	
0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	Klappen
1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	63	11	0	0	0	0	0	0	0	0	Frontklappe
0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	40	35	0	0	0	0	0	0	0	0	Tür:vorne.links
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	6	30	5	0	0	0	0	0	0	Tür:vorne.rechts
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	6	30	5	0	0	0	0	0	0	Tür:hinten.links
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	8	5	30	0	0	0	0	0	0	Tür:hinten.rechts
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	0	0	0	0	0	Heckklappe
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86	0	0	0	0	Stopfen
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	_unstrukturiert
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	Karosserie:Bolzen

Abbildung 4.3: Ein Beispiel für eine Konfusionsmatrix des KStar-Klassifikators (getestetes Fahrzeug: VW Polo)

Ensemble Klassifikatoren Man analysiert, wie stark sich die Klassifikationsgüte durch die Verwendung von Agentenensembles verbessern lässt. Die Evaluation zeigt, dass die mit Ensembles erreichte Klassifikationsgenauigkeit meist geringfügig über der des KStar Algorithmus liegt. Abbildung 4.4 zeigt die Klassifikationsgüte von zwei verschiedenen ensemblebasierten Strategien im Vergleich zum KStar-Klassifikator.

Die Strategie "**Ensemble of {KStar, C4.5, BayesNet}**" sendet die Anfrage jeweils an drei Klassifikationsstrategien. Es wird jeweils der Klassifikator ausgewählt, der seine Klassifikationsentscheidung als besonders sicher einschätzt.

Die Ensemble-Strategie "**Ensemble of Binary**" erstellt für jede Zone einen Binären Klassifikator. Ein Bauteil wird jeweils der Zone zugeordnet, für die die höchste Verlässlichkeit der Zuordnung ermittelt wurde.

Im betrachteten Szenario führte der Einsatz von Agentenensembles nur zu einer geringfügigen Verbesserung der Zuordnungsgenauigkeit. Agentenensembles sind aber hilfreich, um Bauteile mehreren Zonen eines Fahrzeugs zuordnen zu können.

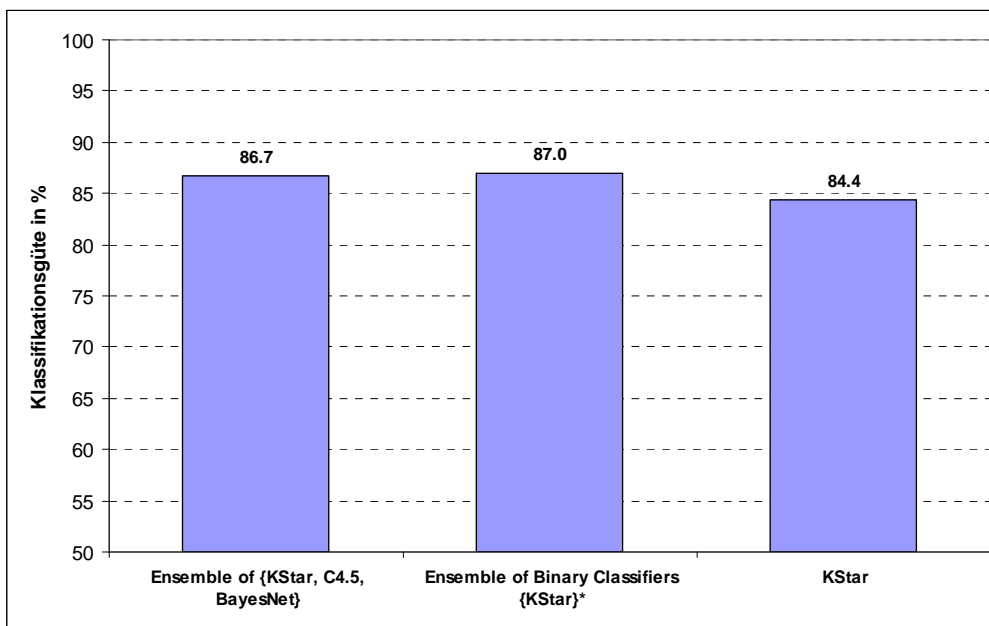


Abbildung 4.4: Vergleich der von Ensemble-Klassifikatoren erreichten Klassifikationsgüte im Vergleich zur Güte des KStar-Klassifikators (getestetes Fahrzeug: VW Polo)

Einfluss der Trainingsmenge Es wird untersucht, wie die Zusammenstellung der Trainingsmenge die Klassifikationsgüte beeinflusst. Es wird am Beispiel eines VW Polos mit 5-Gang Getriebe analysiert, welchen Einfluss unterschiedliche Trainingsmengen auf die erreichte Klassifikationsgüte haben. Es werden dabei die drei Fälle betrachtet:

- (1) Es werden beim Training alle 5-Gang-Polo-Modelle (außer das zu klassifizierende Modell) verwendet.
- (2) Es werden nur die 7-Gang-Polo-Modelle verwendet.
- (3) Es werden alle verfügbaren Polo-Modelle verwendet (außer das zu testende Modell).

Die Evaluation zeigt, dass das Hinzufügen von Bauteilen zur Trainingsmenge zu einem erhöhten Bedarf an Ressourcen führt, im Allgemeinen aber die Klassifikationsgüte erhöht. D. h. im betrachteten Szenario führt das Hinzufügen von nur indirekt relevanten Bauteilen zu einer Verbesserung der Klassifikationsgüte. D. h. auch die 7-Gang-Polo-Modelle enthalten Informationen, die für das Lernen einer Zuordnungsstrategie hilfreich sind. Durch das Berücksichtigen aller Bauteile gegenüber nur den (scheinbar relevanten) 5-Gang-Modellen verbessert sich die Klassifikationsgüte (in Abhängigkeit vom verwendeten Klassifikator) um 3 - 7%.

Meist ist es daher sinnvoll, eine möglichst umfangreiche Trainingsmenge zu verwenden. Bauteile, die für die Klassifikationsaufgabe irrelevant sind, führen nicht zu einer Verschlechterung der Zuordnungsgenauigkeit.

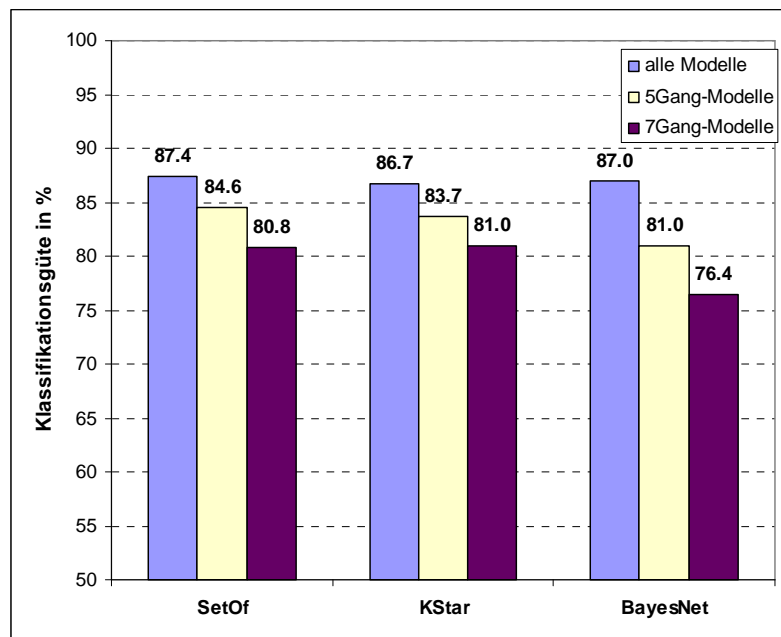


Abbildung 4.5: Vergleich der Klassifikationsgüte in Abhängigkeit von der verwendeten Trainingsmenge (alle Polomodelle, Modelle mit 5 Gang-Getriebe, Modelle mit 7 Gang-Getriebe); getestetes Modell 5 Gang VW Polo

5. Auswertung und Ausblick

Im Rahmen des Projektes wurde ein Framework zum Erstellen und Evaluieren von Agentenensembles erstellt. Die gewählte agentenorientierte Architektur des Frameworks erlaubt es, effizient vielfältige Klassifikationsalgorithmen einzubeziehen und Algorithmen zur Kombination mehrerer Klassifikationsalgorithmen zu testen. Das Framework ist offen für neue Klassifikationsalgorithmen und erlaubt es, Parameterkonfigurationen gezielt in Hinsicht auf ein ausgewähltes Szenario zu optimieren. Die integrierten Funktionen zur Messung der Klassifikationsgüte und zur Analyse der gelernten Klassifikationsstrategien ermöglichen die Evaluation der jeweils gelernten Zuordnungsstrategien. Für die analysierten Szenarien konnte eine sehr gute Klassifikationspräzision erreicht werden. Das entwickelte Framework ermöglicht es daher, die Teilezuordnung effektiv vorzunehmen und dadurch den Produktentstehungsprozess effizienter zu gestalten.

Für eine weitere Verbesserung des Frameworks ergeben sich verschiedene Ansatzpunkte.

Einbeziehung weiterer Daten und Auswertung von Feedback Die durchgeführten Experimente zeigen, dass durch umfangreichere Trainingsmengen die Klassifikationsgüte erhöht werden kann. Aus diesem Grunde sollten weitere Datenquellen beim Lernen von Klassifikationsstrategien einbezogen werden. Darüber hinaus erscheint es aussichtsreich, Feedback der Anwender einzubeziehen, um Klassifikatoren nachzutrainieren.

Da im betrachteten Szenario das Zuordnen eines Bauteils zu einer falschen Zone des Fahrzeugs zu höheren Kosten führt, als wenn ein Bauteil keiner Zone automatisch zugeordnet wird, sollte das System den Anwender beim manuellen Zuordnen von Bauteilen unterstützen, für die automatisch keine zuverlässige Zuordnung ermittelt werden kann.

Visualisierung der Ergebnisse Die Ergebnisse von Klassifikationsalgorithmen werden oft als abstrakt und als schwer nachvollziehbar angesehen. Um die Akzeptanz der Software bei den Anwendern zu erhöhen, sollte die graphische Visualisierung der Ergebnisse verbessert werden. Dabei sollte dargestellt werden, wie zuverlässig die Zuordnung eines Bauteils zu einer Fahrzeugzone ist und zu welchen verschiedenen Zonen ein Bauteil zugeordnet werden sollte (jeweils mit einer Erklärung).

Verfeinerung der Fahrzeugzonen In den durchgeführten Experimenten wurde jeweils mit 22 Fahrzeugzonen gearbeitet. Eine weitere Effizienzverbesserung ließe sich erreichen, indem die Fahrzeugzonen weiter unterteilt werden. Damit könnte eine detaillierte Reihenfolge von Bauteilen bei der Konstruktion eines Fahrzeugs generiert werden. Dabei könnten die derzeit verwendeten Klassifikationsalgorithmen durch Strategien zum Lernen von binären Präferenzen ergänzt werden [10,11].

Danksagung

Die präsentierten Ergebnisse entstanden im Rahmen eines Projektes in Kooperation mit der Volkswagen AG, Wolfsburg. Besonders danken wir Dr. Henry P. Bensler und Dr. Bernd Werther für ihr Engagement und die umfassende Unterstützung, ohne die dieses Projekt nicht möglich gewesen wäre.

Literaturverzeichnis

- [1] John G. Cleary, Leonard E. Trigg: K*: An Instance-based Learner Using an Entropic Distance Measure. In: 12th International Conference on Machine Learning, 108-114, 1995.
- [2] Ross Quinlan (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.
- [3] George H. John, Pat Langley: Estimating Continuous Distributions in Bayesian Classifiers. In: Eleventh Conference on Uncertainty in Artificial Intelligence, San Mateo, 338-345, 1995.
- [4] J. Platt: Machines using Sequential Minimal Optimization. In B. Schoelkopf and C. Burges and A. Smola, editors, Advances in Kernel Methods - Support Vector Learning, 1998.

- [5] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy (2001). Improvements to Platt's SMO Algorithm for SVM Classifier Design. *Neural Computation*. 13(3):637-649.
- [6] A. Lommatzsch, S. Albayrak: Using multi-agent systems for learning optimal policies for complex problems In *Proceedings of the 45th annual southeast regional conference*, 244-249, 2007, Winston-Salem, NC
- [7] Stefan Fricke, K. Bsufka, J. Keiser, T. Schmidt, R. Sessler und S. Albayrak: Agent-based telematic services and telecom applications In *Communications of the ACM*, vol. 44, no. 4, 43-48, 2001
- [8] R. Polikar: Ensemble based systems in decision making In *Circuits and Systems Magazine, IEEE*, 2006, vol. 6, no. 3, pp. 21 - 45
- [9] Wu, T., Lin, C., and Weng, R. C. 2004. Probability Estimates for Multi-class Classification by Pairwise Coupling. *J. Mach. Learn. Res.* 5 (Dec. 2004), 975-1005.
- [10] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, H. Li: Learning to rank: from pairwise approach to listwise approach In *ICML '07: Proceedings of the 24th international conference on Machine learning*, Corvallis, Oregon, USA , 2007, pp. 129 -136, ISBN 978-1-59593-793-3
- [11] C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, G. Hullender: Learning to rank using gradient descent In *ICML '05: Proceedings of the 22nd international conference on Machine learning*, Bonn, Germany, 2005, pp. 89 -96, ISBN 1-59593-180-5