



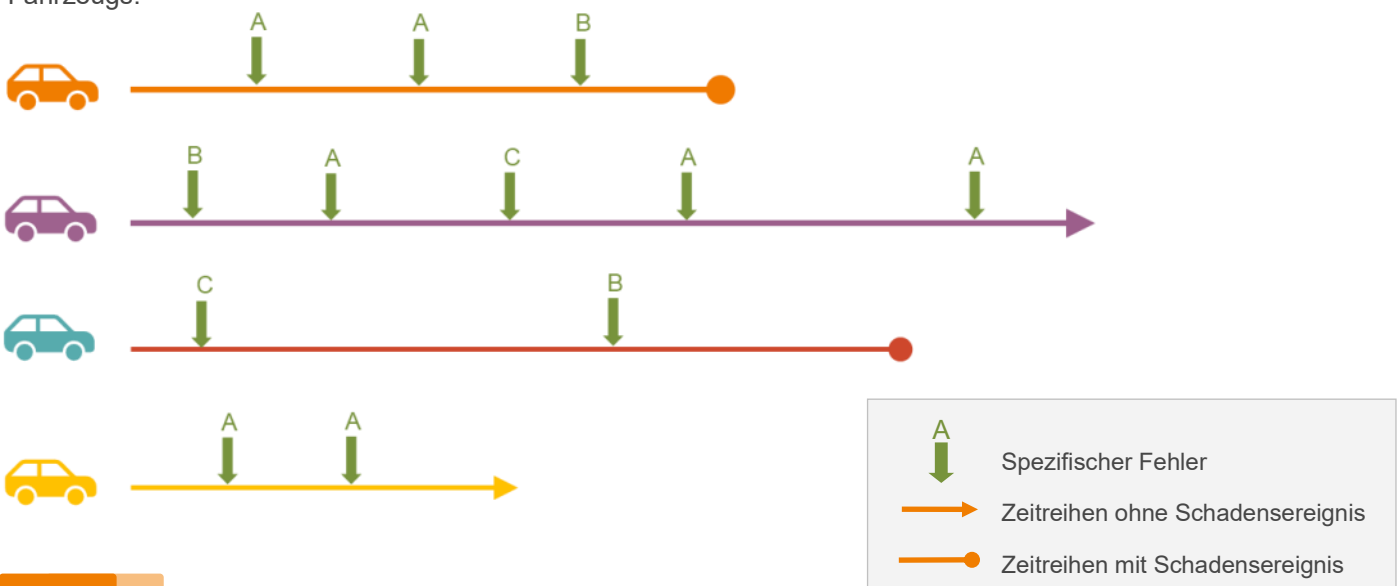
FEHLERPROGNOSE BEI INNOVATIVEN FAHRZEUGPROJEKTEN - WIE BIOSTATISTISCHE ANSÄTZE AUF AUTOMOTIVE-PROJEKTE ÜBERTRAGEN WERDEN KÖNNEN

PROJEKTZIEL

Die Entwicklung und Markteinführung innovativer Fahrzeuge, mit neuen Technologien, stellt eine besondere Herausforderungen für das Qualitätsmanagement dar. Es gibt keine historischen Informationen und der Prozess kann nicht auf vergangenen Erfahrungen basieren. Für das Qualitätsmanagement ist es wichtig, potenzielle Schadenstypen und deren Faktoren oder Muster zu identifizieren. Für neu entwickelte Fahrzeuge gibt es keine früheren Informationen, bei denen Muster von Fahrzeugkonfigurationen und Fehlermustern zu spezifischen Fahrzeugausfällen führen. Ziel dieses Projektes war es, einen Algorithmus zu entwickeln, der in der Lage ist, Cluster von Fehlermustern und Fahrzeugkonfigurationen zu identifizieren, welche zu bestimmten Fahrzeugausfällen führen. Wichtig ist, dass das Clustering berücksichtigen muss, dass sich die Fahrzeuge in unterschiedlichen Zuständen ihres Lebenszyklus befinden und dass der Algorithmus neu lernen muss, sobald neue Daten eintreffen. Die identifizierten Cluster wurden von Experten bewertet und gekennzeichnet, was das fortlaufende Anlernen eines Fehlertyp-Klassifikators ermöglicht. Um diese Herausforderung zu meistern wurde ein Ansatz aus der Biostatistik verwandt und in die Automobilindustrie übernommen.

BEREITGESTELLTE DATEN

Jeder Datenpunkt bestand aus Informationen über die Konfiguration des Fahrzeugs, Daten aus dem Produktionsprozess, dem normierten Protokoll der Fehlermeldungen aus verschiedenen Steuermodulen, aufbereiteten Daten vom Kunden, Problembeschreibungen, Reparaturberichten und der aktuellen Laufzeit des Fahrzeugs.



HERAUSFORDERUNGEN

Für die Entwicklung des Clustering-Algorithmus gab es eine Reihe von Herausforderungen.

Zunächst einmal waren die meisten der bereitgestellten Features nominale Daten mit verschiedenen Kategorien. Hierfür war ein Feature Engineering erforderlich, da die meisten Clustering-Algorithmen nur mit numerischen Daten arbeiten.

Viele der bereitgestellten Daten aus den Fahrzeugen wurden zu verschiedenen Zeitpunkten des Lebenszyklus aufgenommen und waren zensiert. Zusätzlich waren in den Datensätzen kaum Fehlfunktionen aufgetreten.

Es gab mehrere Fehlerquellen, die die Genauigkeit des Algorithmus beeinträchtigt haben könnten. Vererbte Fehlermeldungen oder Reparaturen, welche den Fehler nicht beseitigt haben, könnten zu irreführenden Schlussfolgerungen über den tatsächlichen Fehler führen.

Damit die Cluster nach ihrem Ausfallmuster bewerten werden können, mussten die Cluster sich selbst an den eingehenden Datenstrom anpassen und entsprechend ihrer Haupteinflussfaktoren interpretiert werden können.

ANGEWANDTE METHODEN

Um das Problem der zensierten Daten zu lösen, haben wir einen Feature-Engineering-Ansatz implementiert, der die verschiedenen Phasen des Lebenszyklus berücksichtigt. Die Fehlerereignisvariablen wurden überarbeitet, so dass sie die Relevanz eines bestimmten Fehlerereignisses für ein einzelnes Fahrzeug mit der entsprechenden Relevanz des Fehlerereignisses für alle Fahrzeuge in Beziehung gesetzt wurde. Diese Variablen wurden anhand der Lebenszeit der Fahrzeuge gewichtet. Auf diese Weise konnte eine Reihe von Fehlerprotokollen nach ihrer Relevanz gruppiert und über die Zeit verglichen werden. Dies ermöglichte die Bildung von Clustern von Ausfallarten für Fahrzeuge, die sich in unterschiedlichen Phasen ihres Lebenszyklus befinden.

Um die Herausforderung mit den nominalen Daten zu bewältigen, wurden die bereitgestellten Features in den numerischen Raum projiziert, indem die Multiple Correspondence Analyse angewendet wurde, die gleichzeitig die Dimensionalität reduziert und Rechenleistung gespart hat.

Die überarbeiteten und projizierten Features wurden dann für die Clustering-Algorithmen verwendet. Der beste Algorithmus war der DBSCAN, der durch den Silhouetten Koeffizienten ausgewählt wurde.

Basierend auf den identifizierten Clustern der Hauptkomponenten wurden die Haupteinflussfaktoren der Cluster ermittelt. Für jeden Cluster wurde eine individuelle Überlebensanalyse durchgeführt, um die Zeit bis zu einem Schadensereignis zu bestimmen. Ein Co-Occurrence Score wurde weiter berechnet, um Fehlerprotokollreihen zu identifizieren, die zu den gleichen Schadensmustern führten.

PROJEKTERGEBNIS

Der entwickelte Algorithmus kann zur Identifizierung von Schadensmustern verwendet werden, welche über verschiedene Fahrzeugkonfigurationen verteilt waren.

Fahrzeuge in verschiedenen Phasen ihres Lebenszyklus können nun in das Clustering einbezogen werden.

Neue Fahrzeugvarianten können in das dominierende Schadensmuster eingeordnet werden. Dies ermöglicht die Entwicklung einer geeigneten Qualitätsmanagementstrategie.

