

16. Deutscher Nahverkehrstag
„30 Jahre Regionalisierung – Nah gedacht, weit gekommen?“

Echtzeit- Prognosemeldungen

Vertrauen ist gut, Analyse ist besser

Dr. Christine Langhanns & Dr. Kolja Thormann, 24.06.2026, Koblenz



Die rms GmbH – ein spezialisierter Dienstleister für den ÖV.

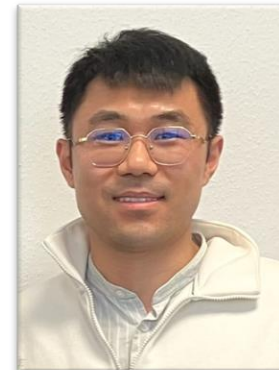
Data Analytics Team verwendet Daten u. a. aus dem ÖV in allen Formaten zur Beantwortung von Analysefragestellungen, zum Monitoring und zur Entwicklung von mathematischen oder KI-Modellen.



Dr. Christine Langhanns
Team- und Projektleitung



Dr. Kolja Thormann
Data Scientist



Weinan Li
Data Engineer

Artificial Intelligence And **MO**bility

Förderung und Konsortialpartner



Unterauftragnehmer



Partnerstädte

Stadt Landau in der Pfalz



Stadt Leipzig

Projekträger



Bundesanstalt
für Verwaltungsdienstleistungen

Gefördert durch:



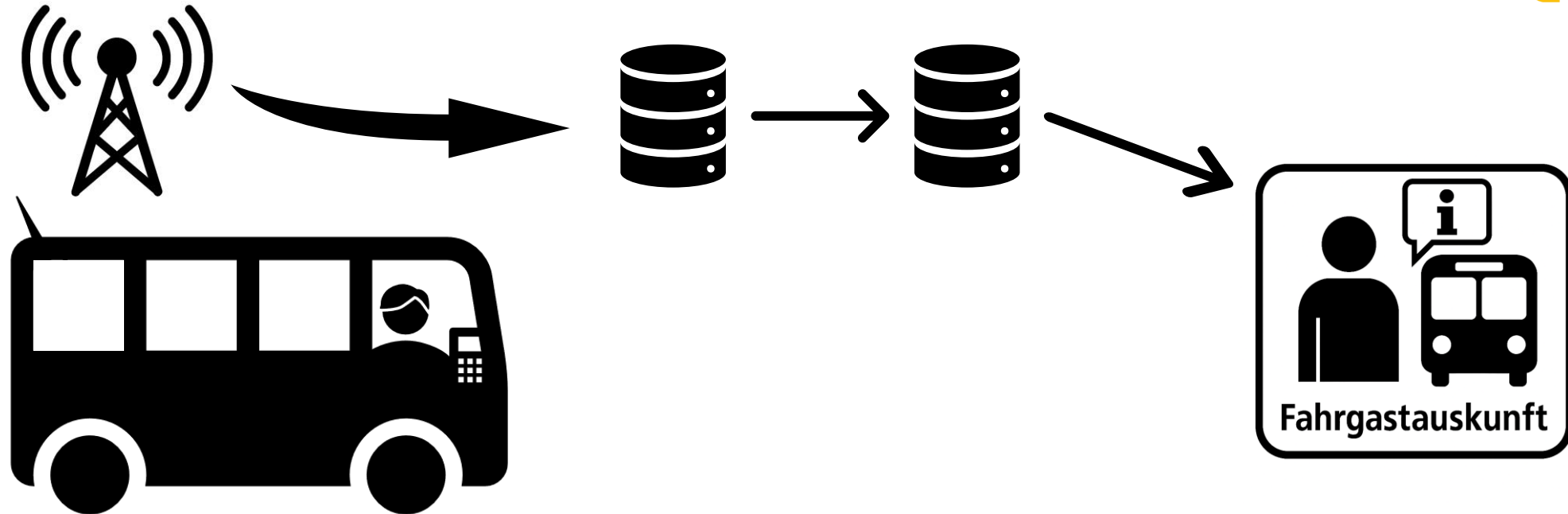
Bundesministerium
für Digitales und
Staatsmodernisierung

aufgrund eines Beschlusses
des Deutschen Bundestages

Echtzeit- Prognosemeldungen VDV-454 AUS: IST-Daten

Regio Cluster Süd, DELFI e. V.

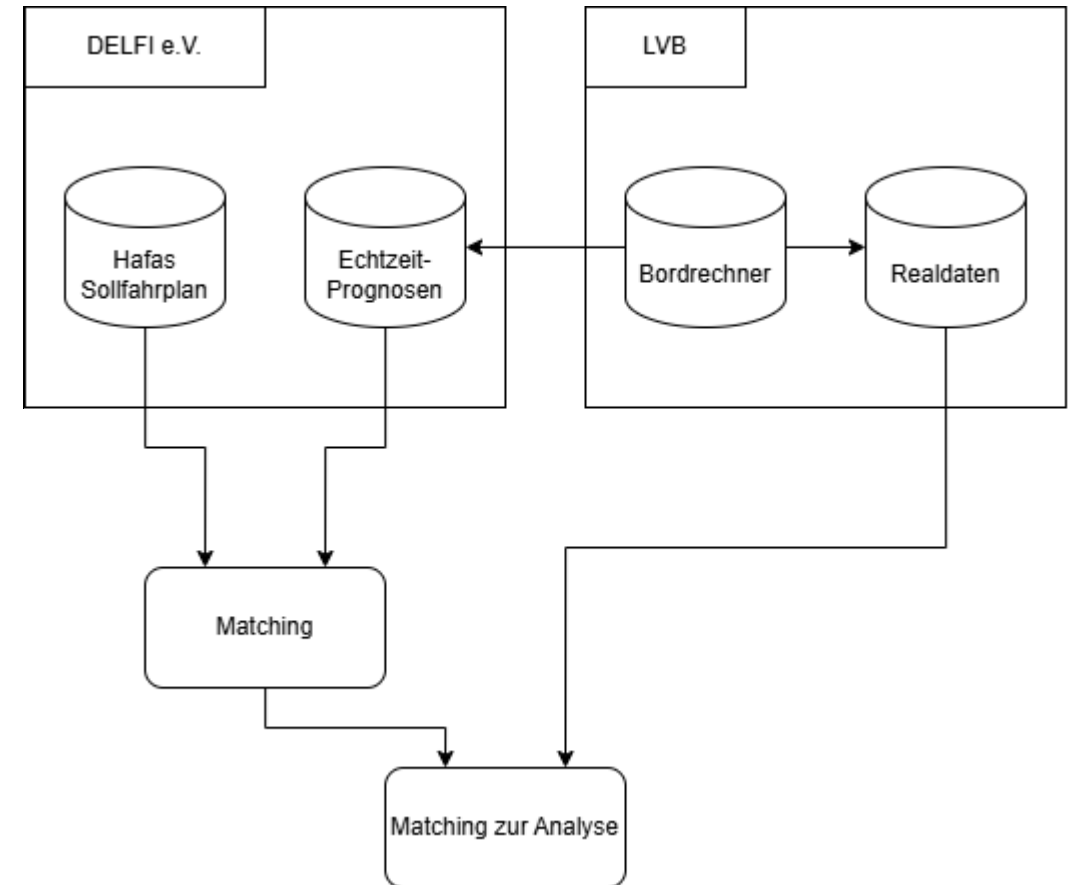
IST-Daten – Halten sie der Realität stand?



Untersuchung des Zusammenhangs von **Realdaten** und **Echtzeit-Prognosemeldung**

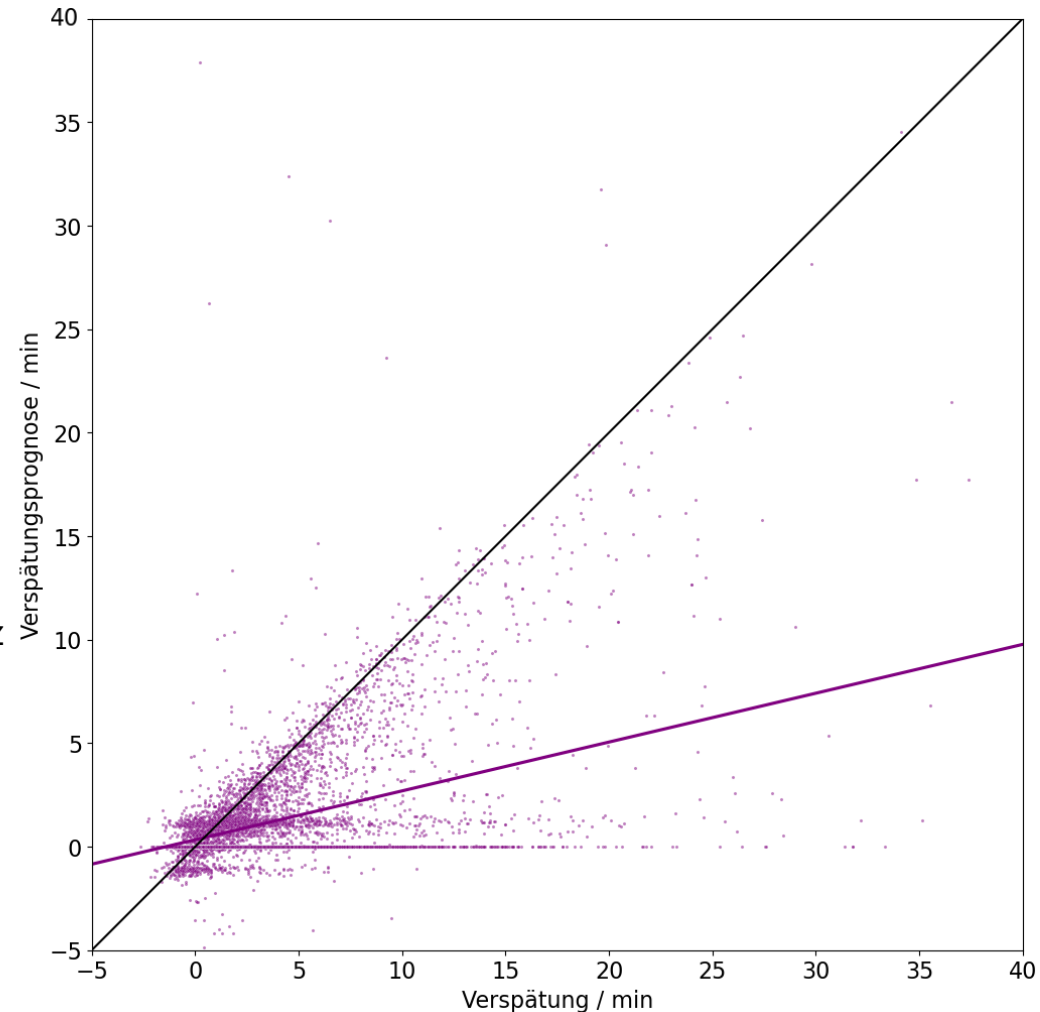
Datenanalyse zur Einschätzung der Prognosegüte

- ca. 6000 Fahrten pro Tag, ca. 1 Mio. Meldungen pro Tag
- Daten lassen sich nicht direkt prüfen
- Zuordnung der Prognosemeldungen zum Fahrplan
 - Herausforderung → fehlendes gemeinsames Element
 - fehlende Halte oder Zusatzhalte
 - abweichende Sollabfahrtszeiten
 - fehlende Mastschärfe der Haltestellen
- Ergebnis der Datenzuordnung
 - ca. 97 % der Prognosemeldungen lassen sich dem Sollfahrplan zuordnen

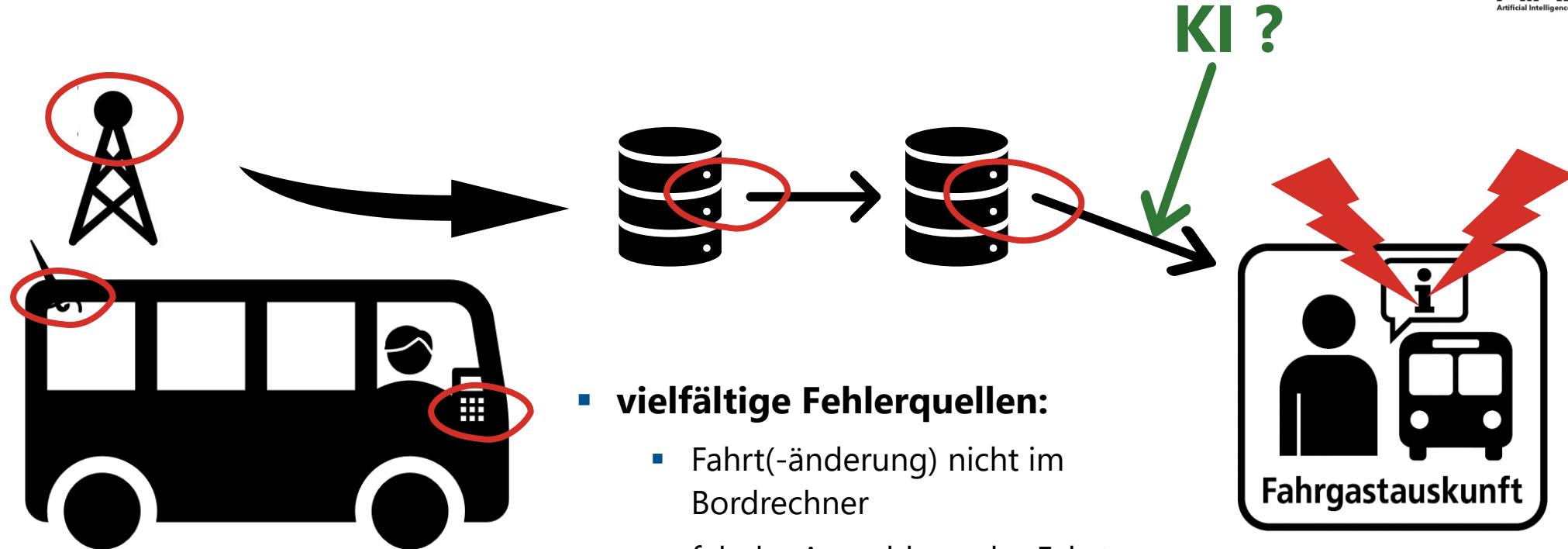


Abfahrtsprognose = tatsächliche Abfahrt ?

- Zeitraum: 03. – 09.11.2025
- Grafik: Datenausschnitt: 10.000 von 12 Millionen Datensätzen
- Differenz zwischen Prognose und tatsächlicher Abfahrt
 - Wertebereich: -130 min bis 165 min
 - Median: 0.90 min
 - Mittelwert: 1.99 min $\pm$ 5.32 min
- ca. 9 % der Prognosen melden Pünktlichkeit (Sollfahrplan) trotz Abweichung bzw. Verspätung
- Analyse des Zusammenhangs (Korrelation)
 - $r = 0.40$ (schwach-mittel)



IST-Daten – Halten sie der Realität stand?



■ vielfältige Fehlerquellen:

- Fahrt(-änderung) nicht im Bordrechner
- falsche Anmeldung der Fahrt
- „Black Box“ Bordrechner
- Funkantenne defekt
- Fahrzeug im Funkloch
- potenziell: Datendrehscheiben

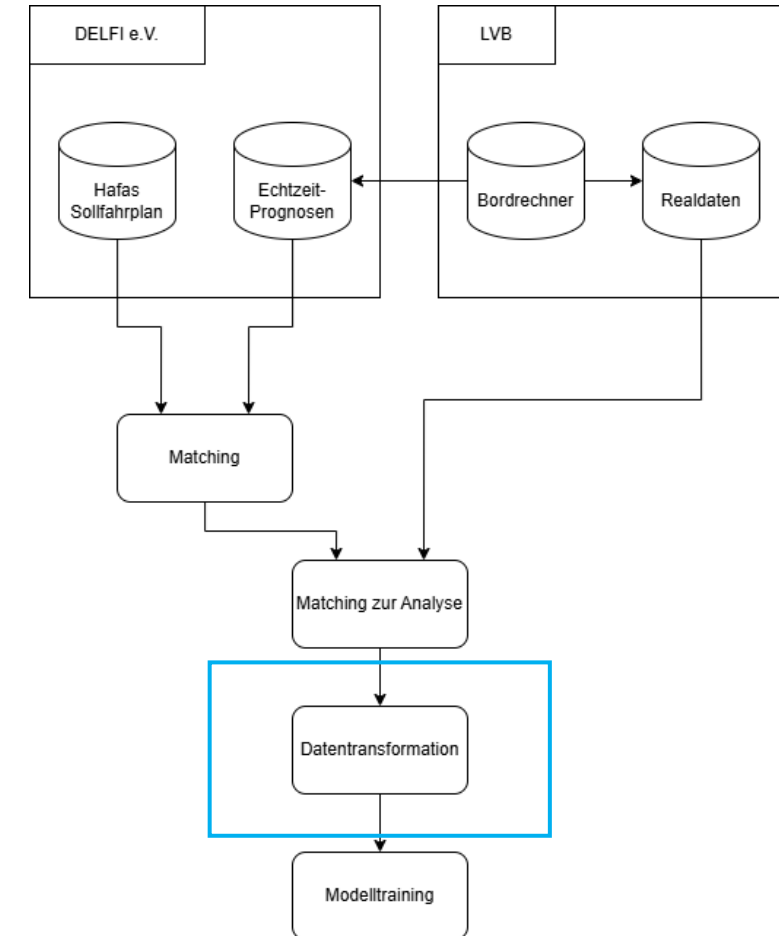
KI-Analysen

Voraussetzung, Schritte und Ergebnisse der KI-Modellierungen

Allgemeines Vorgehen – Entscheidungen, die zu treffen sind

- Datenaufbereitung
 - Aktuelle Meldung regelmäßig neu evaluieren
 - Meldungen vor Ist-Halt und ab 10 Minuten vor Soll-Abfahrt
- Kriterien für Modellwahl
 - Tabellarische Daten und teils kategorische Features
 - Robustes Modell für Machbarkeit
 - Erklärbarkeit der Modellentscheidungen

→ Iterativ erstellter Wald aus Entscheidungsbäumen (XGB)
- Was hatten wir nicht
 - Linienfahrwege, Besetztgrad, Abschließende „real“ Meldungen
 - Live GPS-Koordinaten

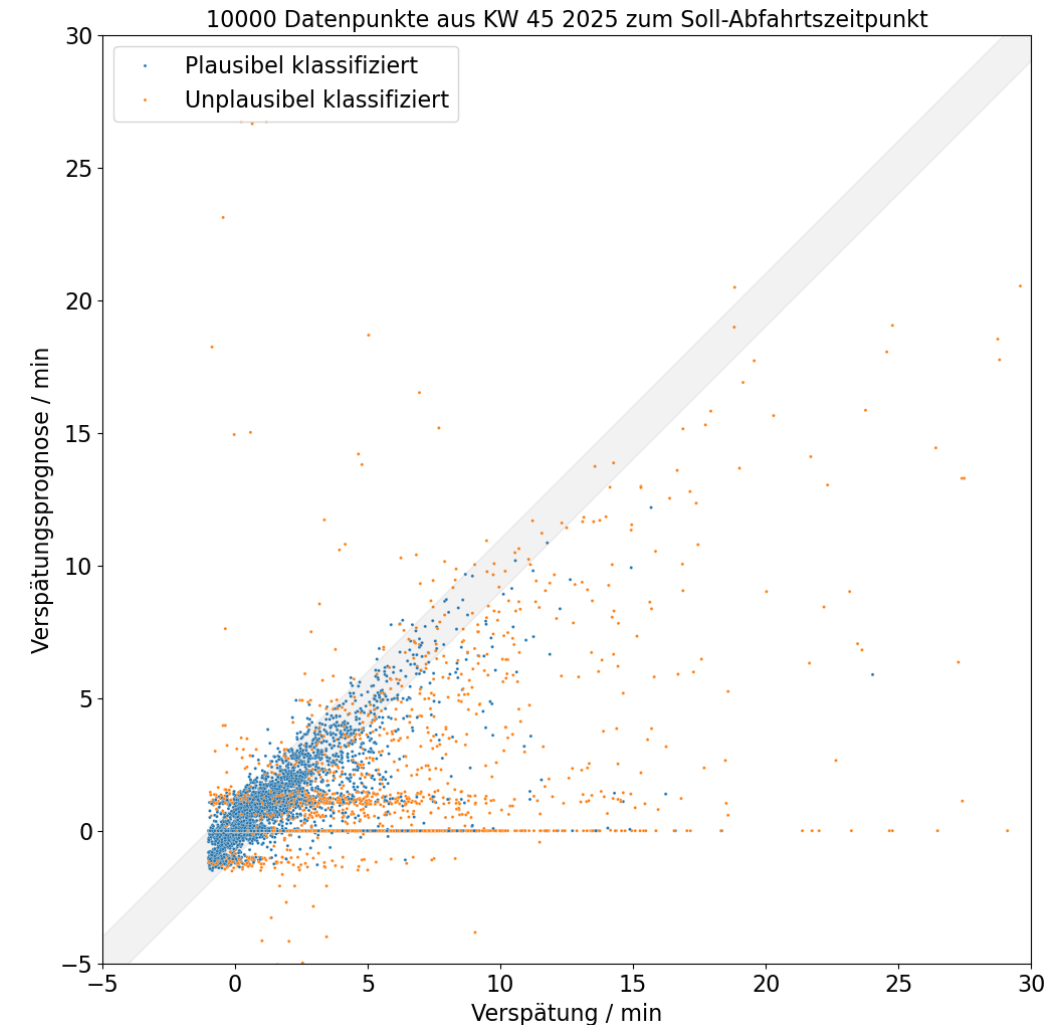


Klassifikation der Echtzeit-Prognosen

- Klassen
 - Abweichung innerhalb ± 1 min → plausibel (grau)
 - Abweichung außerhalb ± 1 min → unplausibel (weiß)
- Klassifikation von Plausibilität:
 - Modell: XGBClassifier, trainiert mit Daten vom 03. – 17.10.2025
 - Balanced Accuracy, getestet an Daten vom 03. – 09.11.2025: **0.71** 😊

	Plausibel	Unplausibel
Plausibel klassifiziert	36.5%	18%
Unplausibel klassifiziert	11.5%	34%

- Folgefragen:
 - Kann das Modell weiter verbessert werden?
 - Können unplausible Meldungen direkt verbessert werden?

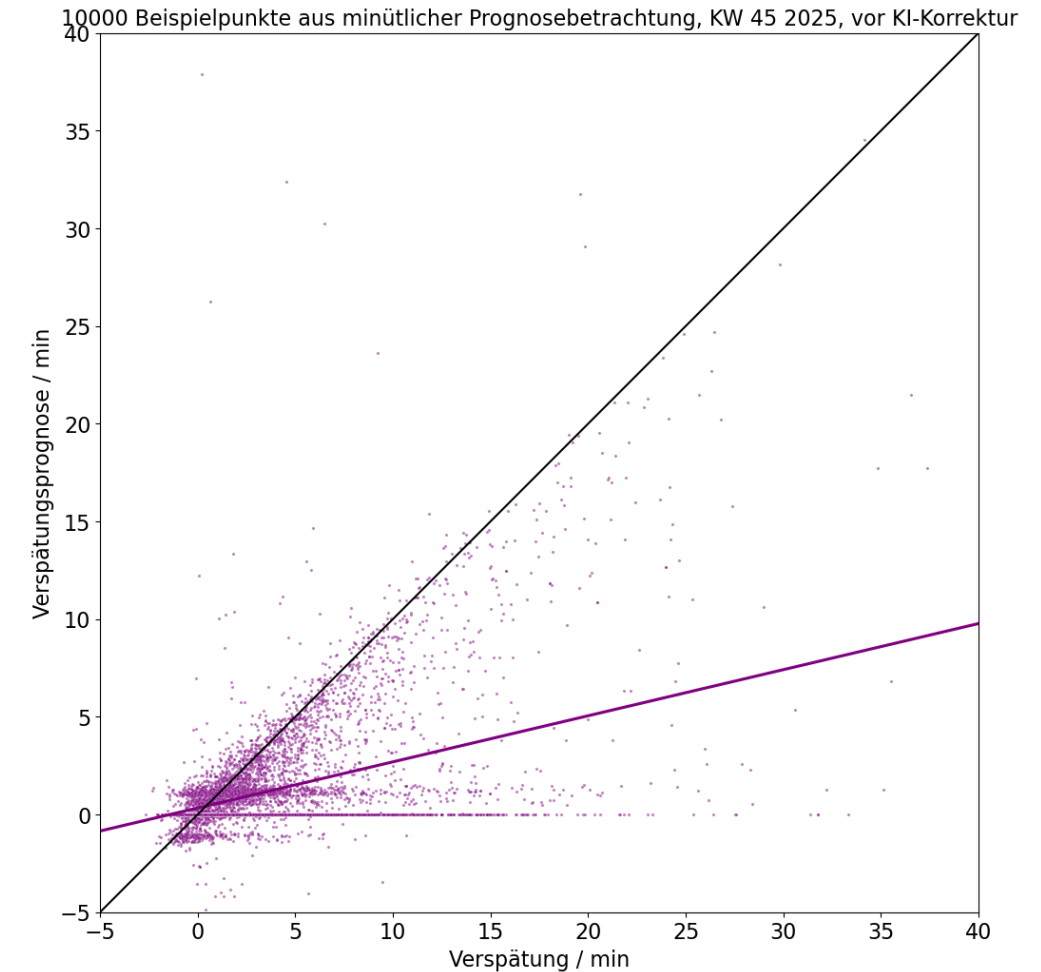


Schätzung der Abfahrtszeit mittels Regression

- Ziel ist, die wahre Abfahrtszeit zu bestimmen
 - Regressionsmodellierung
 - Selbe Datenbasis wie für Klassifikation
- Regressionsergebnisse
 - Modell: XGBRegressor
 - Einfluss auf die Prognosen in Novemberwoche (03. – 09.11.2025):

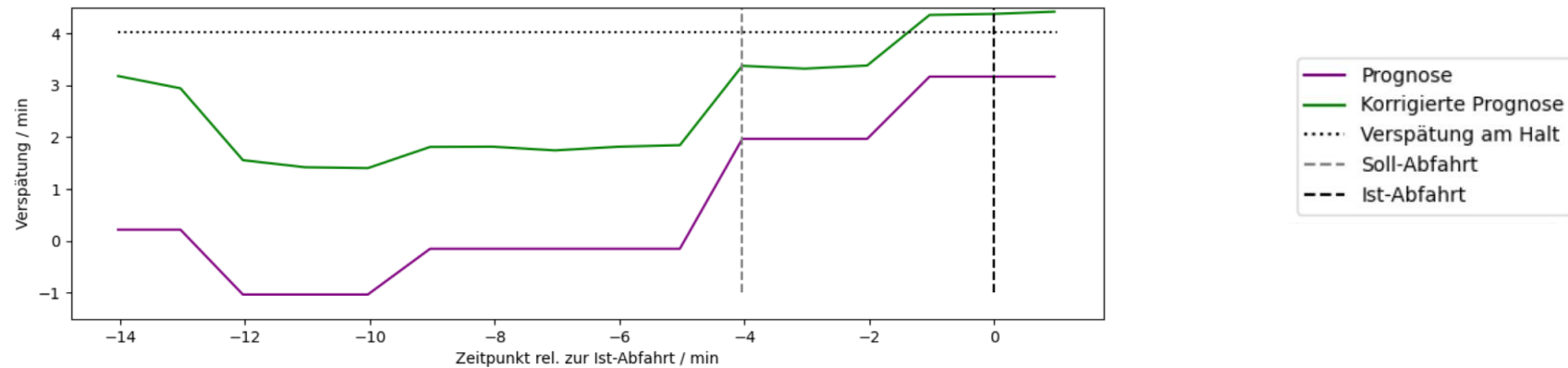
	RMSE	Verzerrung	Stabw.	<i>r</i>
Vor KI-Korrektur	5:41 min	2:00 min	±5:18 min	0.40
Nach KI-Korrektur	3:42 min	0:09 min	±3:42 min	0.77

RMSE: Wurzel des mittleren quadratischen Fehlers, Stabw.: Standardabweichung, *r*: Korrelationskoeffizient

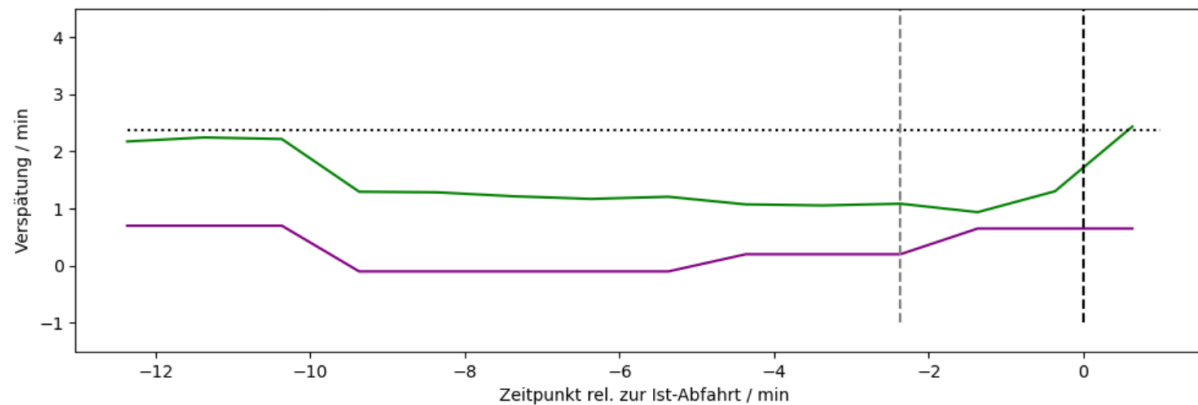


Beispiele von Korrekturen für einzelne Halte

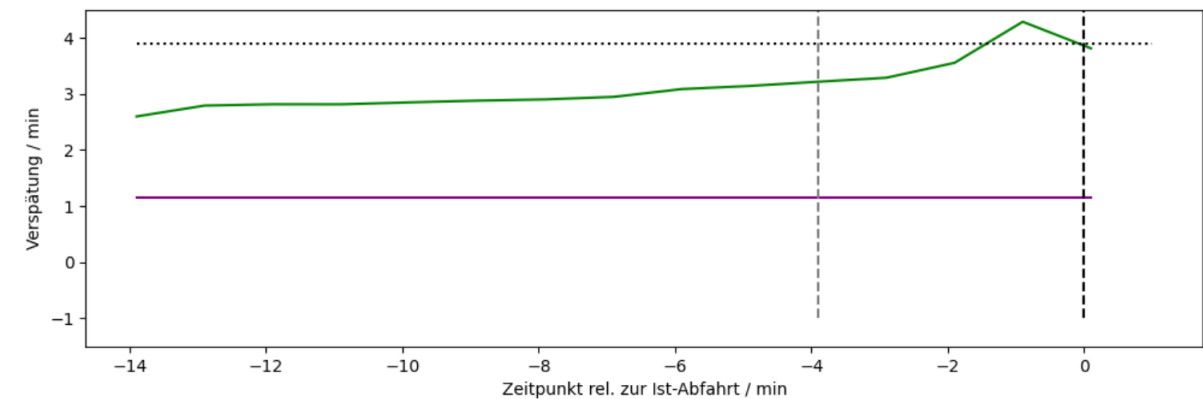
Leipzig, Einertstr., Linie 3, 2025-11-06 15:01:00



Leipzig, Antonien-/Gießerstr., Linie 1, 2025-11-04 13:11:00



Leipzig, Richard-Lehmann-Str./HTWK, Linie 11, 2025-11-07 10:00:00



Zusammenfassung und Ausblick



1. Echtzeit-Prognosemeldungen sind nicht häufig IST-Daten
 2. Fehlerquellen sind zahlreich und deren Behebung ist nicht nur aufwändig sondern auch kostenintensiv
 3. Positionsdaten der Fahrzeuge bedeutet einen einmaligen finanziellen Aufwand: In Funklöchern gibt es auch diese Daten nicht
 4. Maschinelles Lernen ermöglicht eine Bewertung der Meldungen und eine signifikante Schärfung der Prognosequalität
- Modellverbesserung:
Informationen zur Straßenverkehrssituation,
Voraussetzung: Veröffentlichung der Fahrwege
 - Anwendung der KI-Lösung in Echtzeit



Vielen Dank für Ihr Interesse!

Dr. Christine Langhanns

Teamleitung Data Analytics

Christine.Langhanns@rms-consult.de

Dr. Kolja Thormann

Data Scientist

Kolja.Thormann@rms-consult.de



rms GmbH

(Rhein-Main-Verkehrsverbund
Servicegesellschaft)

www.rms-consult.de
info@rms-consult.de

Hauptsitz Frankfurt

Am Hauptbahnhof 6
60329 Frankfurt am Main

Büro Berlin

Am Hamburger Bahnhof 4
10557 Berlin

Büro Hamburg

Burchardstraße 17
20095 Hamburg

Büro Dresden

Prager Straße 8a
01069 Dresden