



Abschlusspräsentation: Gruppe "Schiffe versenken"

Willkommen an Bord



- 1. Live-Demo des Programms
- 2. Datenanalyse
- 3. Architektur
- 4. Machine Learning





Live-Demo des Programms

Das Beste zu Beginn



Präsentation des Programms





Datenbereinigung



Routen:

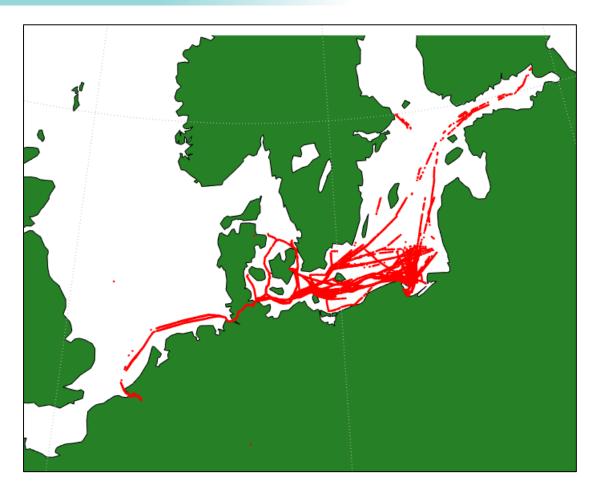
- Bremerhaven Hamburg
- Kiel Gdynia

Umwandlung von .arff zu .csv Import in MySQL-Datenbank

Auswahl zu löschender Datensätze in Matlab Berechnung zusätzlicher Feature in MySQL Trennung in Test- und Trainingsda ten

Alle Wege führen nach Gdynia...

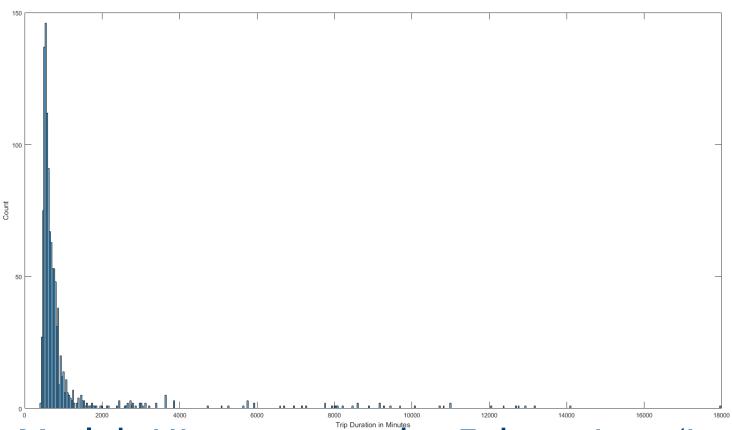




Matlab-Plot aller Koordinatenpaare im Kiel-Gdynia-Datensatz. Bereinigung mit Bounding Box.

Bereinigung über Fahrtzeiten





Matlab-Histogramm der Fahrtzeiten (in Minuten) zwischen Bremerhaven und Hamburg. Bereinigung über Perzentil.

Bereicherung des Datensatzes



TimeDate	WeekdayInt	WeekOfYearInt	HourInt
2016-07-02 10:22:00	5	26	10
2016-07-02 11:31:00	5	26	11
2016-07-02 11:34:00	5	26	11

Ziel: Nutzen der Aussagekraft des Messzeitpunkts *TimeDate* zur Reisezeit.

Methode: Extrahieren von drei eindeutigen Features (*Wochentag*, *Tageszeit*, *Kalenderwoche*), für die ein Einfluss auf die Reisezeit des Schiffes vermutet wird.

Test- und Trainingsdaten

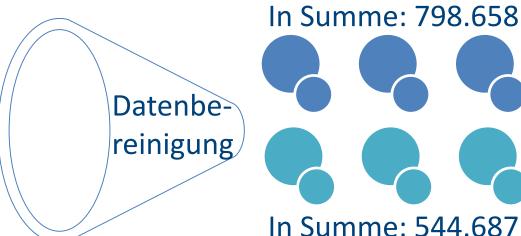


Aufteilung des Datensatzes:

- Eigene Datensätze je Agent (drei Agenten je Route)
- Split: Trainingsset (80%) und Testset (20%)
- Wichtig: Trennung anhand vollständiger Trips

Bremerhaven – Hamburg 958.350

> Kiel - Gdynia 752.980







Architektur



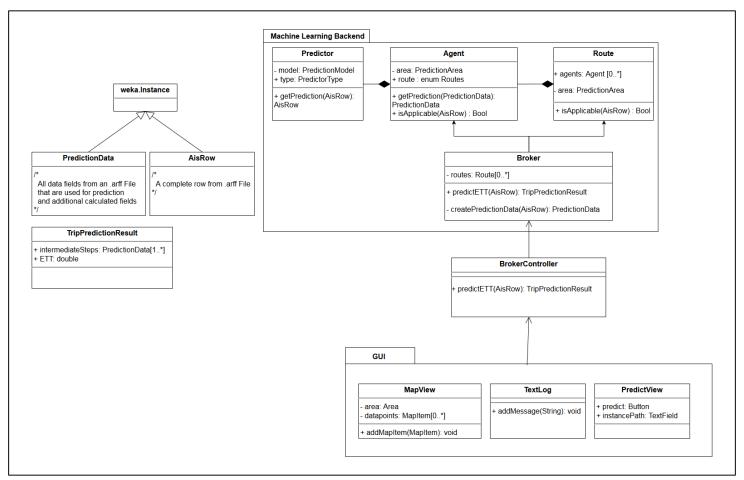
Verteilt: GUI und Broker werden getrennt ausgeführt

Betriebssystemunabhängig: Beide Applikationen sind in "reinem" Java verfasst

Skalierbar: Es ist ohne Komplexität möglich neue Routen und Agenten hinzuzufügen oder den Broker zu wechseln

UML-Klassendiagramm

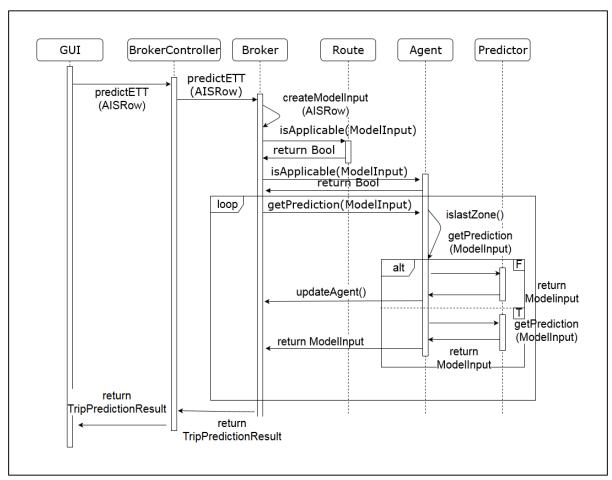




Verteilung auf zwei Systeme: Die GUI und den Broker

UML-Sequenzdiagramm





Verschachtelte Aufrufe bis zur eigentlichen Prediction

Tech Stack



Verwendete Technologien und Hilfsmittel:





Tech Stack



Verwendete Technologien und Hilfsmittel:











Machine Learning

Verwendete Modelle

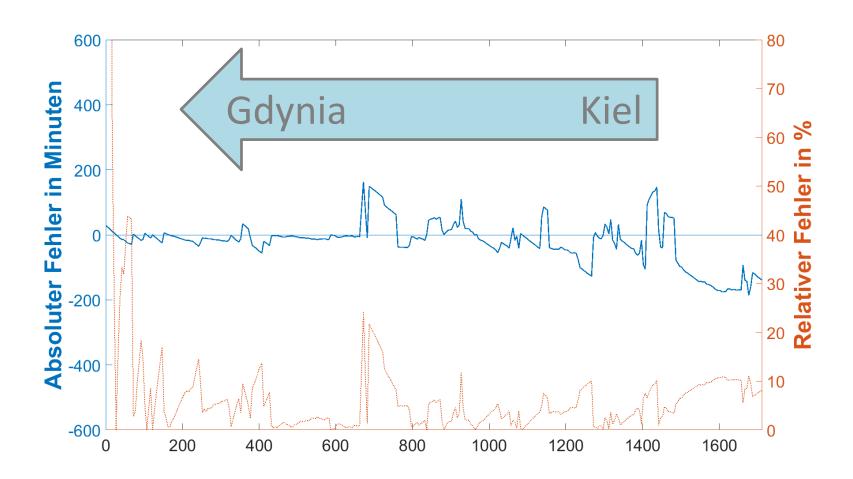


Für das MultiAgent-Verfahren werden neben der RemainingTravelTime noch weitere Features vorhergesagt:

Viol nach		Agent	RTT	LAT	SOG	COG	
	Multi	1	M5P, 50.0	M5P, 50.0	M5P, 4.0	M5P, 4.0	
	Agenten	2	KNN, 100	M5P, 50.0	M5P, 4.0	KNN, 100	
		3	M5P, 100.0	-	-	-	
	Single Agent	Random Forest					
		Agent	RTT	LAT	SOG	COG	
Bremer-	Multi	Agent 1	RTT RT, 50.0	LAT M5P, 50.0	SOG M5P, 4.0	COG M5P, 4.0	
haven	Multi Agenten						
		1	RT, 50.0	M5P, 50.0	M5P, 4.0	M5P, 4.0	

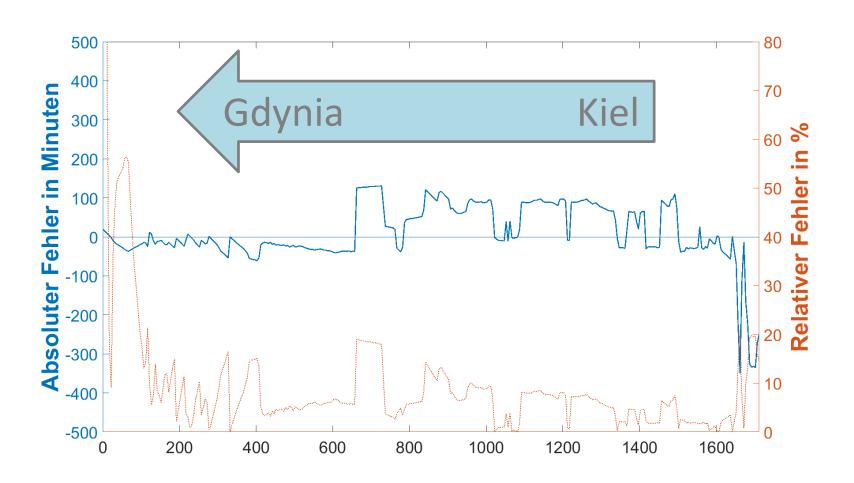
Kiel - Gdynia: Single Agent





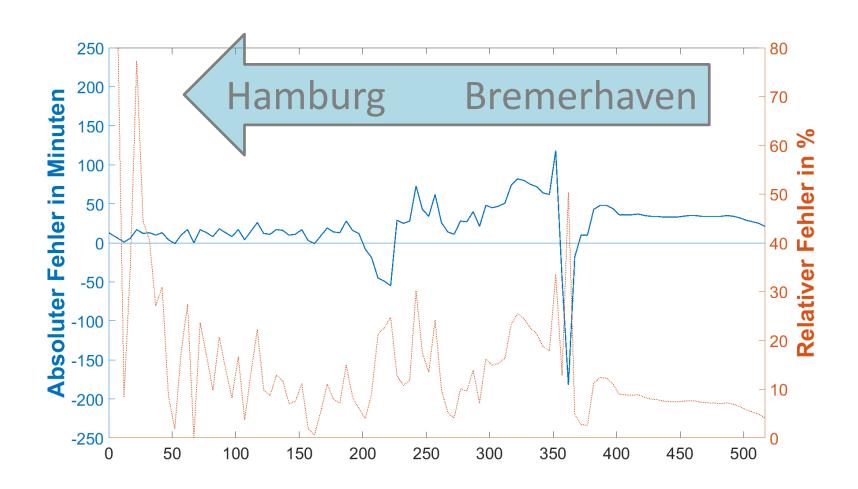
Kiel - Gdynia: Multi Agent





Bremerhaven - HH: Multi Agent





Erkenntnisse



Multi-Agent schlägt Single-Agent: Die Aufteilung des Gebiets in mehrere Zonen führt zu einer präziseren Gesamtvorhersage.

Mischung: Kein Modell sagt alle Feature am besten voraus, daher verschiedene Modelle verwenden.

"Hafen-Problem": Wenn das Schiff zum Ende eines Trips im Hafen warten muss, verändern sich die Features bis auf die Restzeit kaum noch. Schwierige Vorhersage.

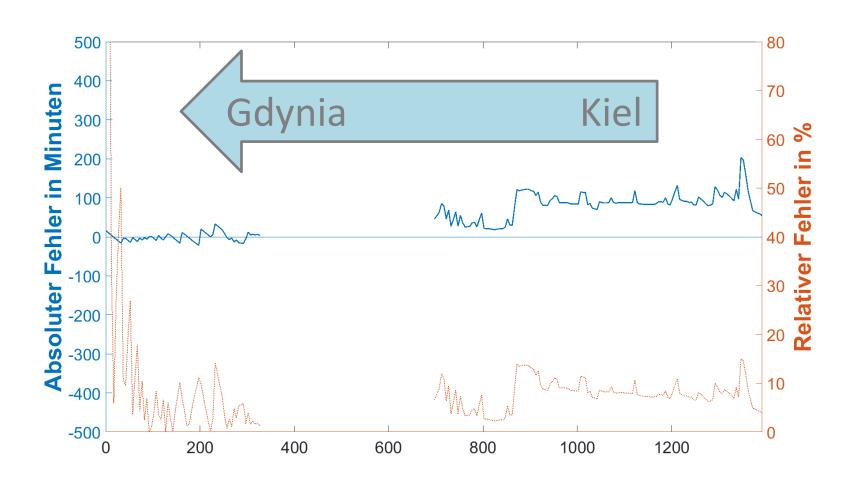




Wir lichten die Anker!

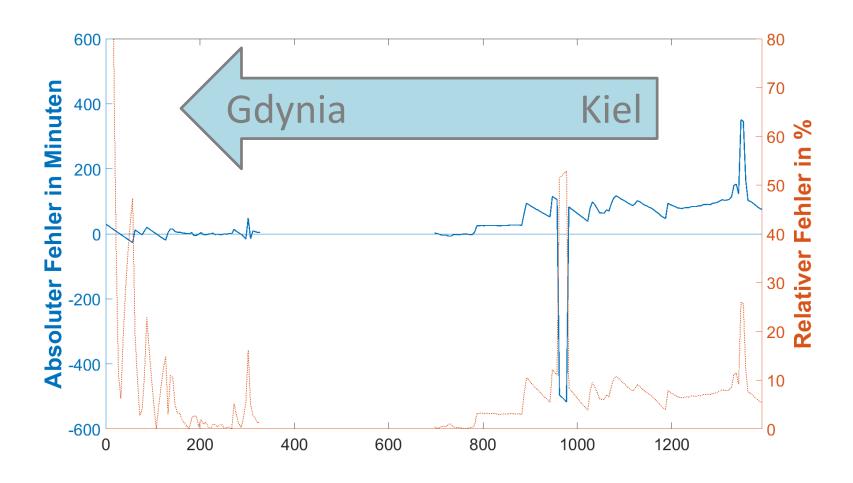
Kiel - Gdynia: Multi Agent





Kiel - Gdynia: Single Agent





Bremerhaven - HH: Single Agent



Matlab einfügen