Frankfurt School

Deutsche Bahn Delay

A Study of Official Statistical Bias and Empirical Modeling

Xiao Shen, Yuanye Xu

The Passenger Experience Gap

Punctuality

Development in the year under review

PUNCTUALITY / %	2024	2023	2022
DB Group (rail) in Germany	89.4	90.1	90.9
DB rail passenger transport in Germany	89.5	90.3	91.0
DB Long-Distance	62.5	64.0	65.2
DB Regional [1]	90.7	91.4	92.2
DB Cargo (Germany)	68.0	70.5	66.1
DB Regional (bus)	85.9	85.2	86.0
DB Cargo	68.2	69.7	66.3
Punctuality (whole journey) (DB Long-Distance)	67.4	68.9	69.3

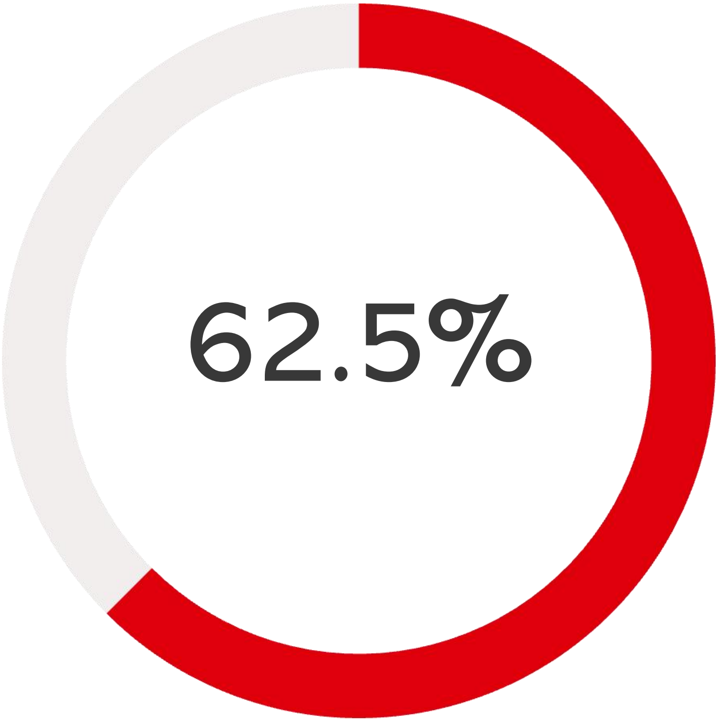
89.5%

Official Rate

DB's reported punctuality

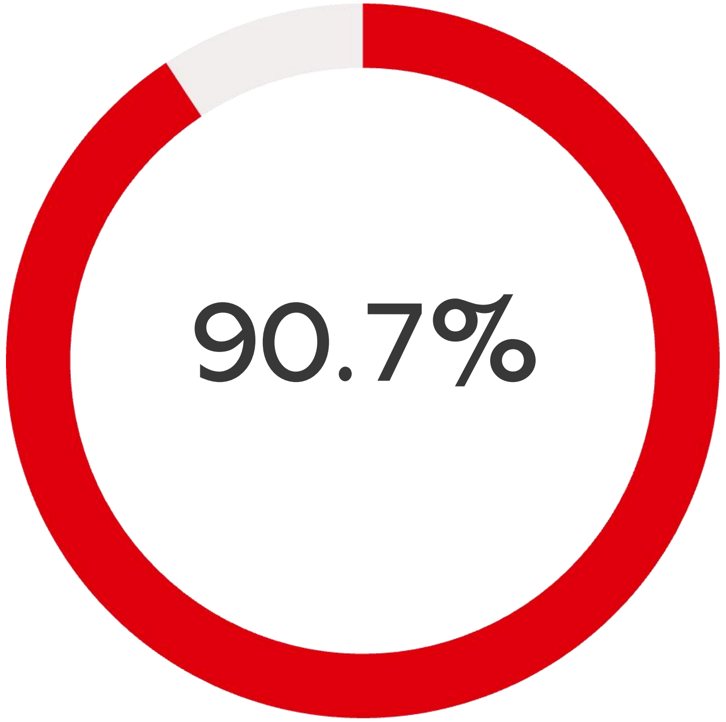
Official statistics told a different story than passenger reality.

The Statistical Illusion



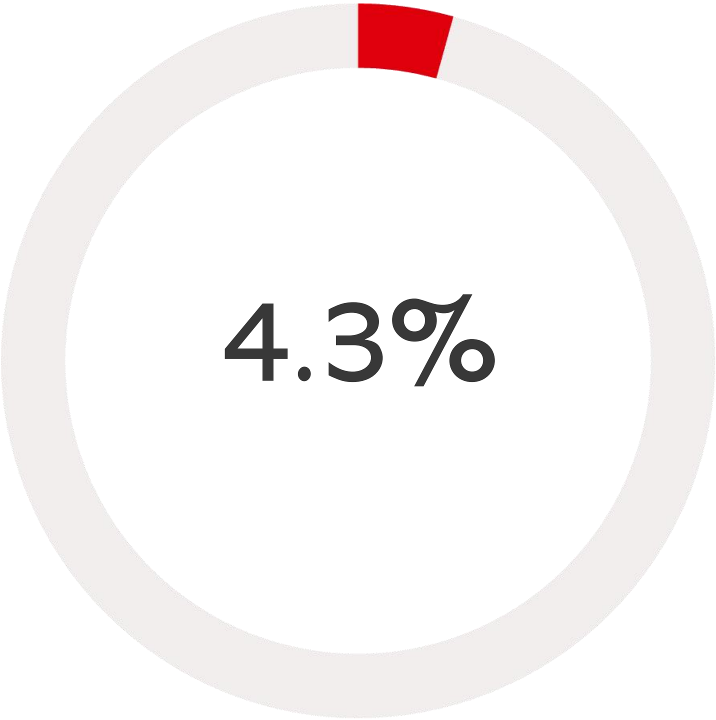
Long-Distance

ICE, IC trains punctuality



Regional

RE, RB, S-Bahn punctuality



Traffic Share

Long-distance proportion

Regional trains' high punctuality statistically "dilutes" poor long-distance performance, masking severity of delays.

RESEARCH MISSION

Reassess DB's reporting methodology and build a realistic, data-driven delay prediction model that exposes structural biases in official statistics.

Methodological Bias #1

Overly Lenient Definition

Punctuality

Share of on-time stops in relation to all stops along and at the end of lines in Germany. A stop is considered operationally punctual if the scheduled arrival time is exceeded by less than six minutes in passenger transport or by less than 16 minutes in freight transport. At DB Regional Rail, punctuality has been reported since 2024 on a third-weighted basis between S-Bahn (metro) alternating current, S-Bahn (metro) direct current and DB Regional Rail without S-Bahn (metro). At DB Regional Road, buses that leave more than one minute early have also been counted as unpunctual with retroactive effect since 2020. For the arrival time punctuality of passengers, see punctuality (whole journey).

> "Züge gelten als pünktlich, wenn sie mit weniger als 6 Minuten Verspätung ankommen."

Trains arriving within 5 minutes 59 seconds classified as "on time"

France SNCF

5 minutes

Japan JR

1 minute

Deutsche Bahn

6 minutes

The Passenger Perspective

1

1-2 Minute Delay

Disrupts tight transfers

2

Causes Anxiety

Uncertainty compounds

3

Erodes Trust

System reliability questioned

Six-minute tolerance fails to reflect actual traveler experience and obscures network inefficiencies.

Our Stricter Standard

1min

Delay Threshold

Trains delayed if ≥1 minute late

Purpose

- Enhance analysis sensitivity
- Capture micro-level deviations
- Reveal systemic inefficiencies
- Reflect passenger-centered reality

Methodological Bias #2

Exclusion of Cancelled Services

Ersatzverkehr

Substitute bus services excluded from statistics

Ausfälle

Cancelled trains removed from dataset

Our Approach

Cancellations = extreme delays, included in analysis

Methodological Bias #3

Oversimplified Classification

DB's System

- Long-distance (ICE, IC)
- Regional (RE, RB, S-Bahn)

S-Bahn's high punctuality artificially inflates regional average

Our Three-Tier System

Long-Distance	ICE, IC, EC, ECE, RJ, RJX, TGV, FLX, FEX, D, EN, NJ, UEX, ES, EST, WB
Regional	RE, RB, IRE, ALX RE, BRB RE, WFB RE, NWB, NWB RS, VIA, ME, MET, HLB RE, MEX, MEX a, MEX c, R, HBX, SVG
Urban	S, S X, SE

DATA COLLECTION & PROCESSING

Data Collection Challenge

The Problem

DB doesn't publicly release detailed delay data. Real-time boards refresh every few seconds, showing only short-term records.

The Solution

Third-party Kaggle dataset from independent developers' automated web scraping of DB's real-time boards.

Coverage

July 20–25, 2024 and September 1–2, 2024 across 20 major German cities.

	A	B	C	D	E	F	G	H	I
1	date	Hbf	scheduled_	expected_t	train_mode	route	platform	real_time_	has_delay
2	2024/9/1	Köln Hbf	13:00	13:00	RE 8 (	Rommerski	3	13:00	0
3	2024/9/1	Hannover H	13:00	13:00	S 2	Nienburg (W	1	13:00	0
4	2024/9/1	Münster (We	13:00	13:00	RB 89 (	Paderborn	3	13:00	0
5	2024/9/1	Hamburg Hb	13:01	13:01	S 5	Hamburg E	4Hamburg H	13:01	0
6	2024/9/1	Frankfurt	13:01	13:01	S 4	Kronberg (T	101Frankfu	13:01	0
7	2024/9/1	Hamburg Hb	13:02	13:02	RE 60 (	Westerland	11D-F	13:02	0
8	2024/9/1	Wuppertal	13:02	13:02	S 8	Mönchengla	5	13:02	0
9	2024/9/1	Berlin Hbf	13:03	13:03	FEX19832	Flughafen	6	13:03	0
10	2024/9/1	Hamburg Hb	13:03	13:03	ME RE3 (	Uelzen, 1	12A-B	13:03	0
11	2024/9/1	Hamburg Hb	13:03	13:03	S 1	Wedel (Hols	3Hamburg H	13:03	0
12	2024/9/1	Hamburg Hb	13:03	13:03	S 2	Hamburg-Al	4Hamburg H	13:03	0
13	2024/9/1	Hannover H	13:03	13:03	S 5	Paderborn	2	13:03	0
14	2024/9/1	Hannover H	13:03	13:03	WFB RE60 (	Braunschwe	12	13:03	0
15	2024/9/1	Frankfurt	13:04	13:04	VIA RB10 (	Neuwied,	23	13:04	0
16	2024/9/1	Nürnberg H	13:04	13:04	RE 32 (	Lichtenfel	16	13:04	0
17	2024/9/1	Hamburg Hb	13:05	13:05	S 3	Hamburg-Ne	1Hamburg H	13:05	0

Data Processing : remaining useful info

departure_city(H	route_cleaned	train_model_cleaned	scheduled_time	arrival_hour_bucket	expected_time	expected_delay	real_tim	has_delc	real_delay_min	causes
Berlin	Baruth(Mark)	RE	13:11	13	13:14	3	13:14	1	0	
Duisburg	K?ln/Bonn Flughafen	RE	13:11	13	13:14	3	13:14	1	0	
Hannover	Norddeich Mole	IC	13:13	13	13:14	1	13:14	1	0	
Bonn	Walporzheim	RB	13:13	13	13:14	1	13:14	1	0	
Hamburg	Stuttgart Hbf	ICE	13:14	13	13:14	0	13:31	1	17	Grund: Ve
Hannover	Koblenz Hbf	ICE	13:14	13	13:14	0	13:34	1	20	Grund: Te
Duisburg	Dresden Hbf	FLX	13:14	13	13:14	0	13:35	1	21	Grund: Vc
Düsseldorf	Kaarster See	S	13:14	13	13:14	0	Fahrt f?l	1		
Hannover	Koblenz Hbf	ICE	13:14	13	13:14	0	Fahrt f?l	1		
Köln	Berlin Ostbahnhof	ICE	13:11	13	13:15	4	13:15	1	0	
Münich	Herrsching	S	13:13	13	13:15	2	13:15	1	0	
Leipzig	Hannover Hbf	IC	13:14	13	13:15	1	13:15	1	0	Grund: Ve
Nürnberg	Regensburg Hbf	RE	13:15	13	13:15	0	13:24	1	9	Grund: Wa
Duisburg	Arnhem Centraal	VIA RE	13:15	13	13:15	0	13:23	1	8	

- Removing Numbers/Words
- Splitting Complex Phrases
- Create hour buckets with scheduled_time
- Calculating “real_delay_minutes” = “real_time_due_to_delay” - “expected_time”

Data Processing for modelling

Raw Dataset

30+ train categories including bus services

Cleaning

Removed all Ersatzverkehr bus records

Reclassification

Three operational categories by service characteristics

Final Attributes

Date, stations, category, time, punctuality status

Delay Definition

Train classified as **delayed** if arrival delay exceeded **one minute**

Cancelled trains for any reason assigned delay status

Binary variable: `has_delay`

- 0 = on-time
- 1 = delayed

0/1

Binary Classification

Final Data for Modeling

1	date	Hbf	arrive_station	train_category	depart_hour_bucket	has_delay
2	2024/7/20	Hannover Hbf	Wien Hbf	long-distance	20	1
3	2024/7/25	Nürnberg Hbf	München Hbf	long-distance	8	1
4	2024/7/23	Hamburg Hbf	Karlsruhe Hbf	long-distance	11	1
5	2024/7/22	Berlin Hbf	München Hbf	long-distance	12	0
6	2024/9/2	Berlin Hbf	Wünsdorf-Waldstadt	long-distance	8	1
7	2024/7/20	Hamburg Hbf	Leipzig Hbf	long-distance	9	0
8	2024/9/1	Berlin Hbf	München Hbf	long-distance	17	0
9	2024/7/23	Hannover Hbf	München Hbf	long-distance	15	1
10	2024/7/21	Nürnberg Hbf	Karlsruhe Hbf	long-distance	20	0
11	2024/7/20	Bremen Hbf	Norddeich	long-distance	8	0
12	2024/7/21	Hannover Hbf	Hamburg Hbf	long-distance	11	1
13	2024/9/2	Dortmund Hbf	Hamburg-Altona	long-distance	8	1
14	2024/7/25	Stuttgart Hbf	Karlsruhe Hbf	long-distance	12	0
15	2024/7/24	Dortmund Hbf	Münster(Westf)Hbf	long-distance	7	1
16	2024/9/2	Leipzig Hbf	Berlin Hbf	long-distance	8	0
17	2024/9/2	Stuttgart Hbf	Nürnberg Hbf	long-distance	7	0
18	2024/7/24	Berlin Hbf	Aachen Hbf	long-distance	14	1
19	2024/7/20	Hannover Hbf	Berlin Ostbahnhof	long-distance	17	0
20	2024/7/21	Köln Hbf	Berlin Ostbahnhof	long-distance	1	1
21	2024/7/25	Bielefeld Hbf	Berlin Ostbahnhof	long-distance	9	1
22	2024/7/20	Nürnberg Hbf	München Hbf	long-distance	7	0
23	2024/9/1	Duisburg Hbf	Frankfurt(Main)Hbf	long-distance	22	1
24	2024/7/20	Hamburg Hbf	Lübeck Hbf	long-distance	8	1
25	2024/7/23	Hannover Hbf	Wien Hbf	long-distance	20	1
26	2024/7/24	Leipzig Hbf	Hannover Hbf	long-distance	8	1
27	2024/7/21	Köln Hbf	Amsterdam Centraal	long-distance	19	1
28	2024/7/20	Dresden Hbf	Praha hl.n.	long-distance	22	1
29	2024/7/20	Leipzig Hbf	Norddeich Mole	long-distance	22	1
30	2024/7/23	Hannover Hbf	Hamburg Hbf	long-distance	15	1

OBSERVATION
& VISUALIZATION

Preliminary Observation

- The total average value of actual delay time was 4.6 minutes. Among the trains that experienced delays, the average delay duration increased to 8.84 minutes, with the maximum delay reaching over 400 minutes.

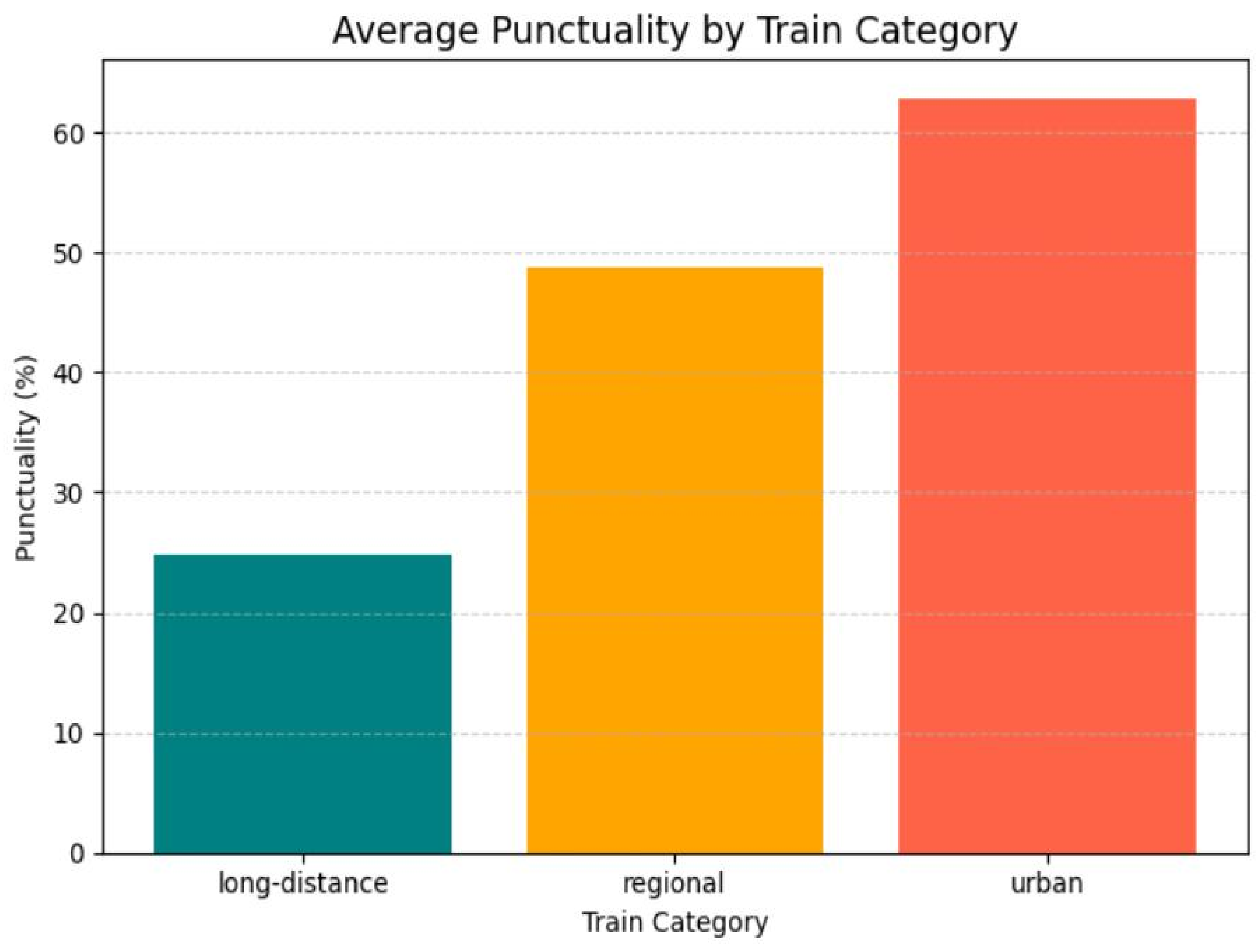
- Regarding the causes of delays, track changes (Gleisweckseln) accounted for 3.8%, delays caused by preceding trains represented 13.3%, construction work (Bauarbeiten) contributed 1.6%, and technical issues such as maintenance depot problems made up 7.31%. Among all delayed services, 1.3% were recorded as “Verkehrt” (wrong routing), and 5.9% were marked as “Fällt aus” (cancelled).

B	D	F	J	K	L	M	N	O	P	Q
date	departure_city(	route_cleaned	train_model_cleaned	scheduled_time	arrival_hour_bucket	expected_time	expected_delay	real_tir	has_del	real_delay_min
2024/7/23	Berlin	Budapest-Nyugati	EC	18:43	18	18:43	0	1:35	1	412
2024/7/24	Berlin	Budapest-Nyugati	EC	18:43	18	18:43	0	1:37	1	414
2024/7/24	Berlin	Budapest-Nyugati	EC	18:43	18	18:43	0	1:40	1	417
2024/7/23	Dresden	Budapest-Nyugati	EC	16:50	16	16:50	0	23:48	1	418

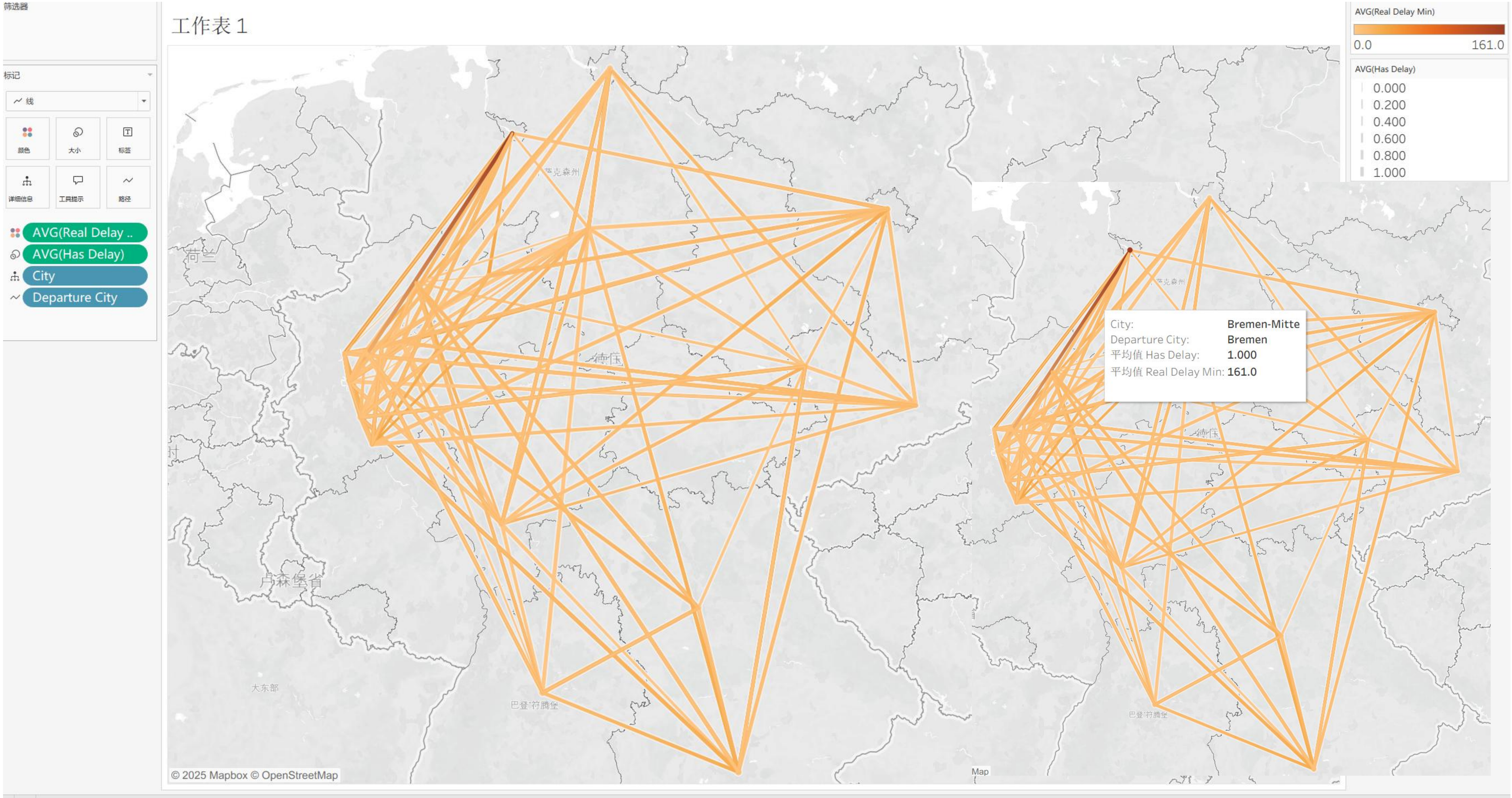
Punctuality by Common Train Model

Train Type	Delay Rate (%)	Average Delay (min)
S-Bahn	36.9	1.95
ICE	77.3	16.91
RE	57.1	6.69
IC	65.3	13.60
RB	45.4	3.74
EC	78.1	41.57

Punctuality by Train Category

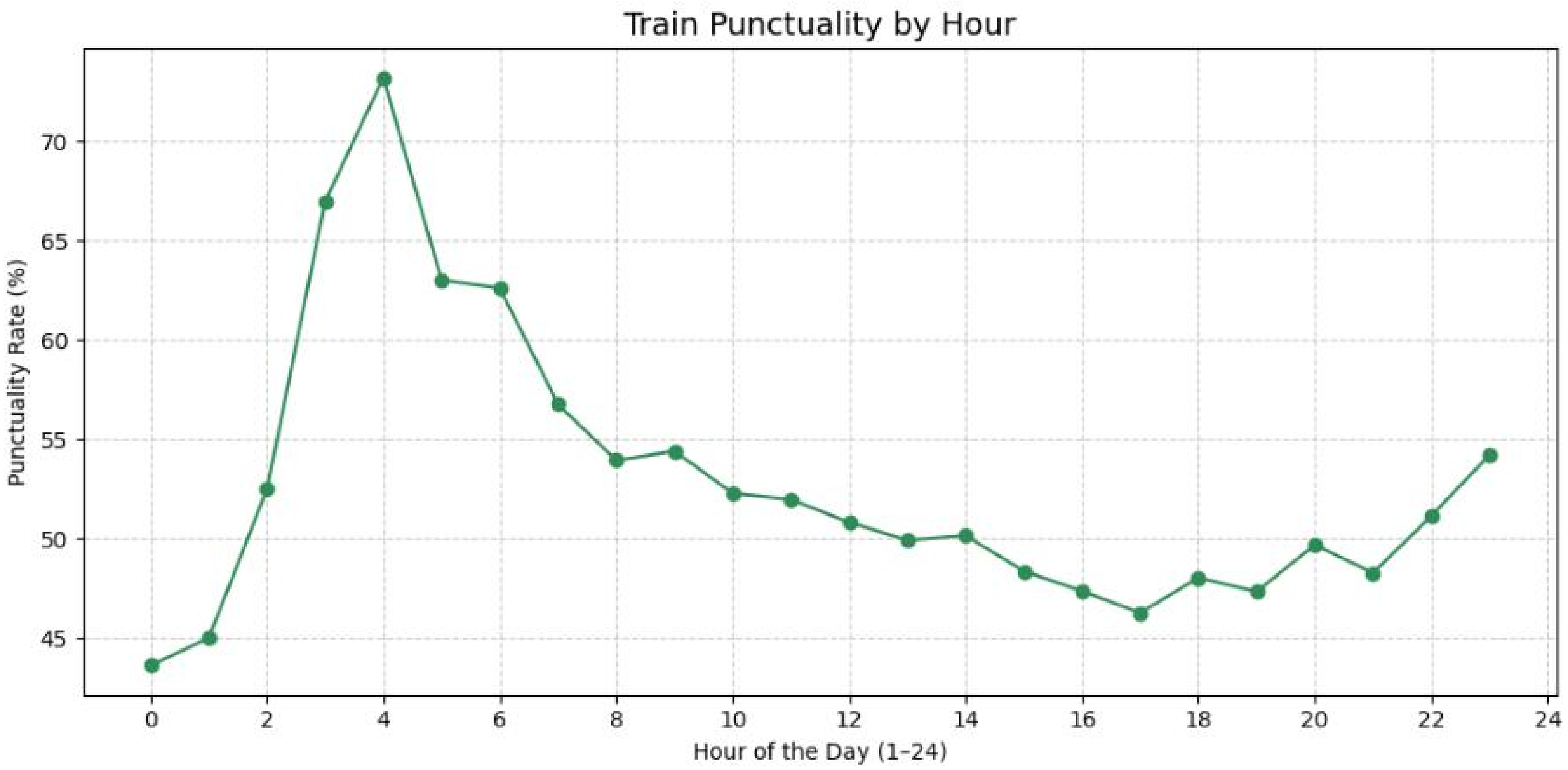


Long-distance trains achieved an on-time rate of only **24.84%**, meaning roughly one in four arrived as scheduled. Regional trains reached **48.74%**, showing moderate stability. Urban trains (S-Bahn) performed best at **62.90%,** reflecting the resilience of short-distance commuter operations.


筛选器
标记
线
颜色
大小
标签
详细信息
工具提示
路径
AVG(Real Delay ..
AVG(Has Delay)
City
Departure City
工作表 1
AVG(Real Delay Min)
0.0
161.0
AVG(Has Delay)
0.000
0.200
0.400
0.600
0.800
1.000
荷兰
德国
卢森堡省
巴登-符腾堡
© 2025 Mapbox © OpenStreetMap
City: Bremen-Mitte
Departure City: Bremen
平均值 Has Delay: 1.000
平均值 Real Delay Min: 161.0
Map

Temporal Patterns

Inverted-U Throughout the Day



highest between 03:00–04:00 (67–74%),

relatively low between 15:00–21:00 (46–50%),

slightly recover in the late evening,

sink to the bottom at mid-night .

Traffic density and infrastructure load strongly affect on-time performance.

Overall: Below 50%

Under stricter one-minute threshold, DB's actual punctuality substantially lower than official 89.5%

MODELLING APPOACH

Machine Learning Approach

Data Split

Merged July and September samples

- 80% training
- 20% testing

Balanced distributions for robust generalization

Four Models

- Logistic Regression
- Random Forest
- LightGBM
- XGBoost

Model 1: Logistic Regression

Baseline Model

$$\text{logit}(p) = \ln \frac{p}{1-p} = \beta_0 + \beta_1(\text{departure station}) + \beta_2(\text{arrival station}) + \beta_3(\text{train category}) + \beta_4(\text{departure hour})$$

$$p = \frac{1}{1 + e^{-(\beta_0 + \sum_i \beta_i x_i)}}.$$

Characteristics

- Linear model with interpretability
- OneHotEncoder for categorical variables
- Recall: 0.80 on-time, 0.75 delayed
- Stable, transparent baseline

77.4%	0.77
Accuracy	F1-Score

Model 2: Random Forest

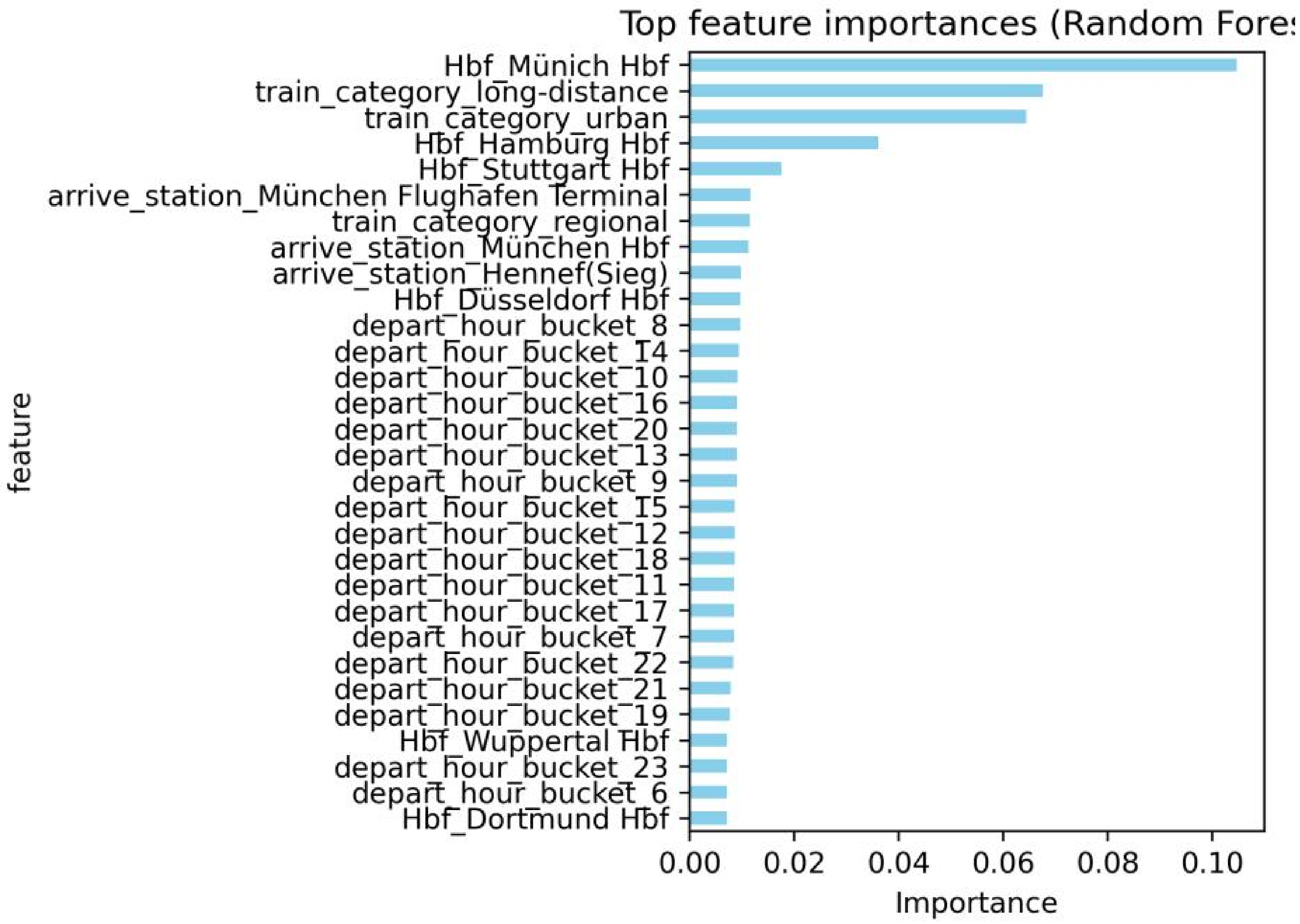
Advantage

Captures nonlinear relationships

Better on borderline cases

78.2%	0.78
Accuracy	F1-Score

Feature Importance Revealed



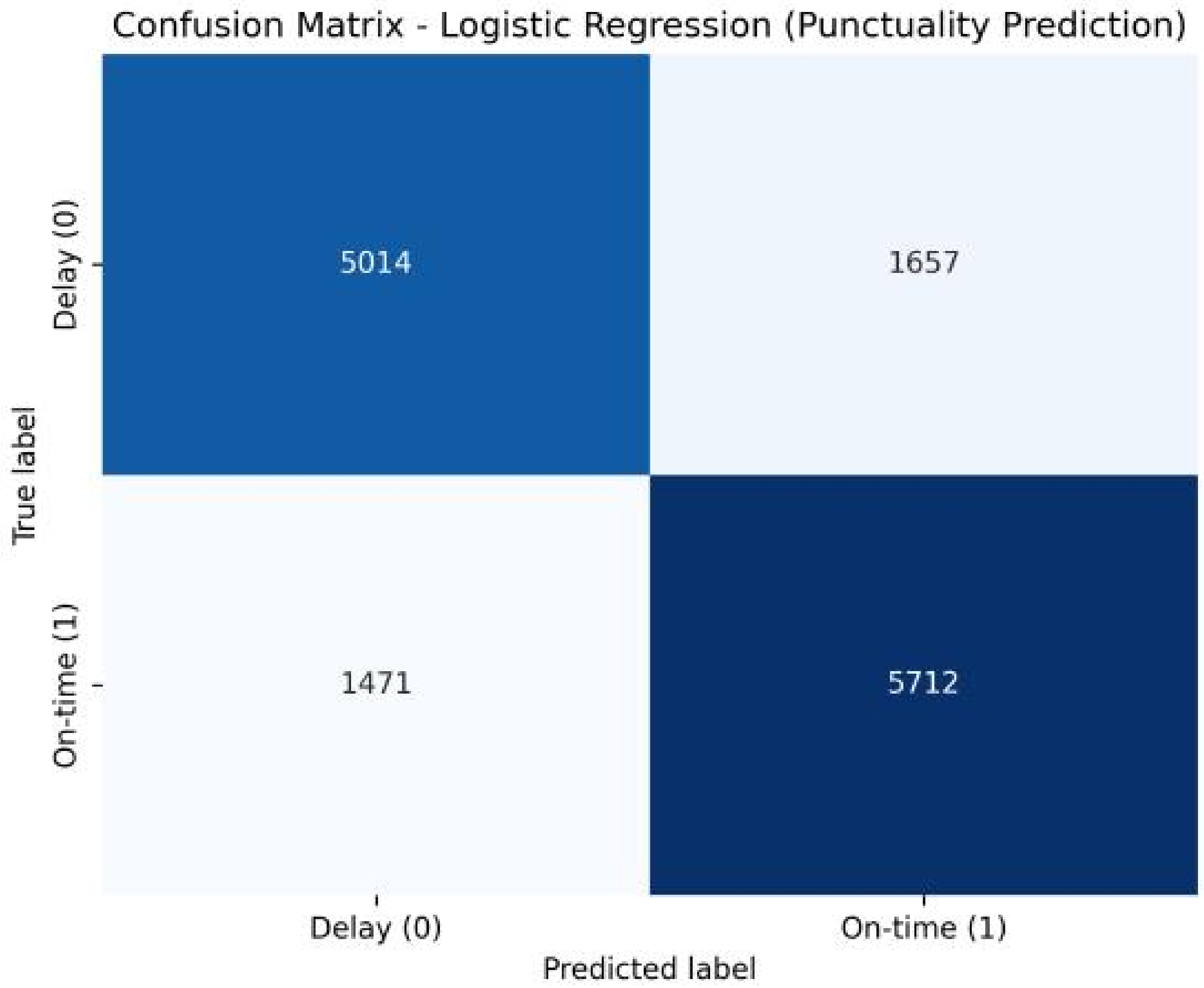
Departure and arrival stations

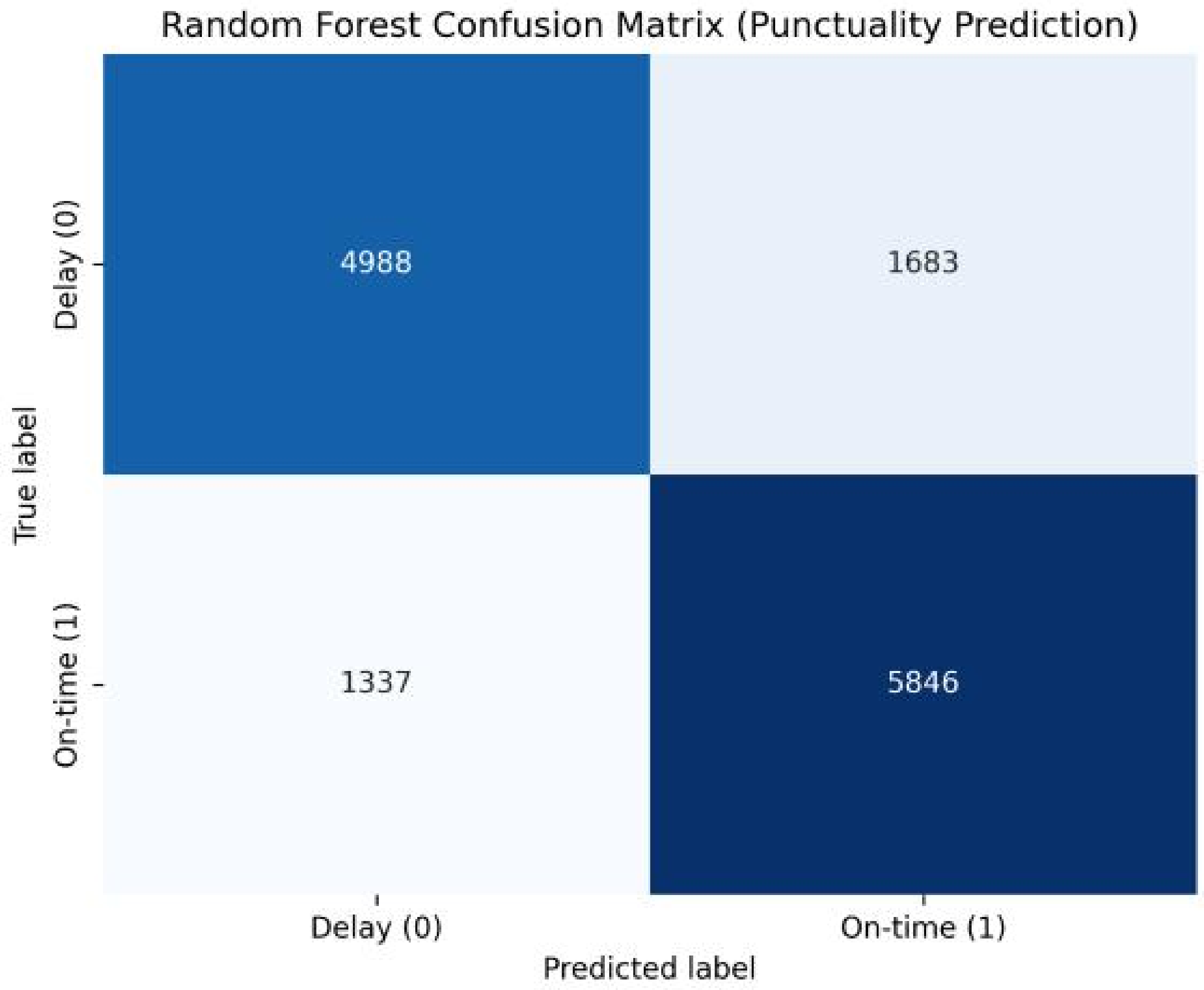
most influential.

Major hubs

—Munich, Hamburg, Stuttgart—

dominate rankings.





Improvement

More on-time trains identified (5846 vs 5712)

Reduced false negatives (1337 vs 1471)

Model 3 & 4: Gradient Boosting

LightGBM and XGBoost

LightGBM

Accuracy: 0.7925

F1: 0.8019

AUC: 0.8727

Best performance

XGBoost

Accuracy: 0.7864

F1: 0.7969

AUC: 0.8689

Slightly lower but more stable

Model Comparison: ROC Curves

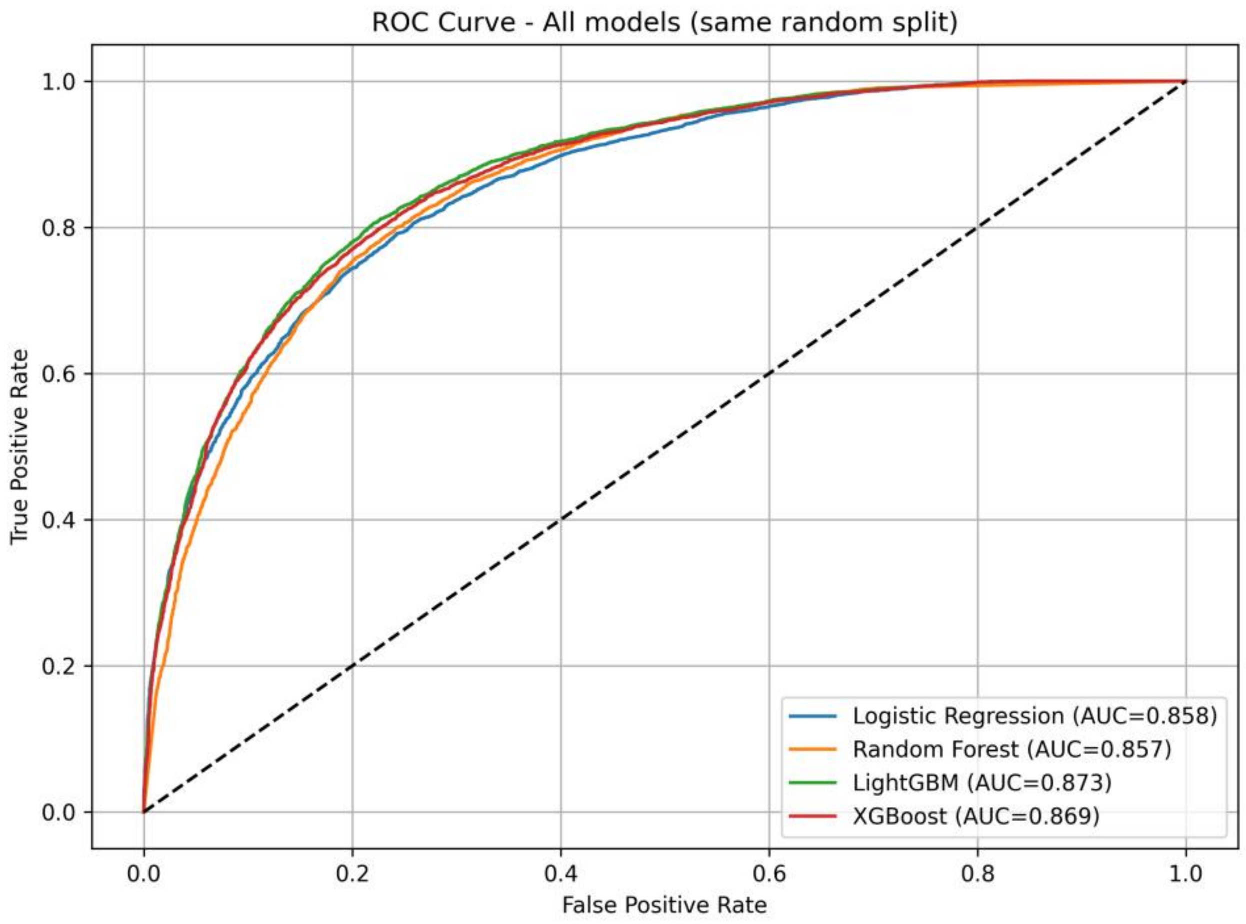
77-79%

Accuracy Range

All four models

Similar overall performance indicates available features nearly exhaust dataset's explanatory power

Nonlinear models offer stronger flexibility but limited improvement without dynamic features



Why LightGBM Wins

Histogram-Based Splits

Efficient handling of sparse data

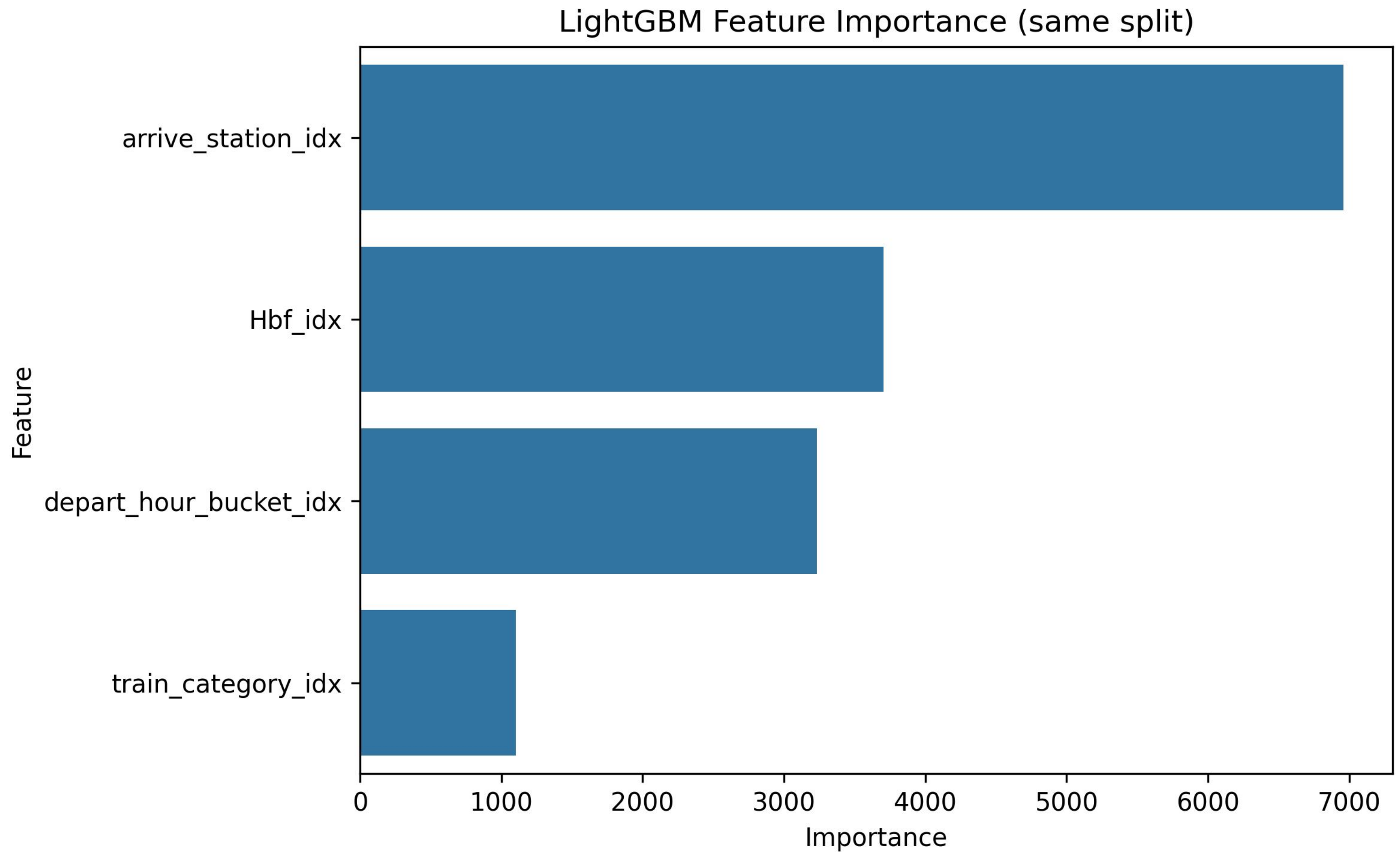
Leaf-Wise Growth

Enhanced efficiency in high-dimensional spaces

Feature Interactions

Captures complex nonlinear relationships

Key Finding: Nonlinear Interactions



Key points:

- Punctuality is strongly influenced by nonlinear interactions
- Departure–arrival station combinations contribute over 60% of total feature importance
- Strong interactions among train type, departure time, and departure station

Practical Implications

For Travelers

High-Risk Corridors

Journeys between Frankfurt (Main), Munich(München), Cologne(Köln) carry significantly higher delay risk

Buffer Time

Allow extra time for transfers on major hub routes

Timing Strategy

Avoid peak hours (15:00–21:00) and holidays when possible

Practical Implications

For Deutsche Bahn

Critical Areas

High-risk routes need capacity management priority

Scheduling Optimization

Focus on densest traffic corridors

Operational Discipline

Enhance reliability through systemic improvements

LIMITATION

Limitations

1

Feature Ceiling

Categorical variables (stations, types, time) mostly linear relationships. Limited room for improvement.

2

Missing Dynamics

No weather, congestion, maintenance, or construction data etc in current dataset.

3

Time Scope

Only two time windows (July, September 2024). Seasonal patterns not captured.

Future Research Directions

- Weather Integration
- Maintenance Schedules
- Holiday Patterns
- Historical Delays
- Temporal Dynamics

Expanding data scope will enable models to capture broader dynamics and achieve higher predictive precision.

THANK YOU
FOR YOUR LISTENING !

References

1.https://www.kaggle.com/datasets/santiagoravotti/trains-and-delays-deutsche-bahn?resource=download
2.https://ibir.deutschebahn.com/2024/en/combined-management-report/product-quality-and-digitalization/the-customer-is-at-the-center-of-our-actions/punctuality/?utm_source=chatgpt.com
3.https://ibir.deutschebahn.com/2024/en/glossary/
4.https://github.com/FabreXUYY/Deutsche-Bahn-Delay