

Zwischenpräsentation

Projekt Safe-Cycling

Lisa Rübel / Nicole Driebe / Sabine Matthies





Vorhersage des Gefahrenpotentials einer Strecke für Radfahrende in Berlin.



Welche Datensätze?

- ❑ Datensatz zu Fahrradunfällen in Berlin
- ❑ Datensatz zur Verkehrsdichte in Berlin
- ❑ SimRa-Daten zu Beinahe-Unfällen
- ❑ OSM-Daten (Straßenbeschaffenheit, Art Fahrradwege, etc.)



Fokus auf diese Datensätze:

- ☐ Datensatz zu Fahrradunfälle in Berlin
- ☐ Datensatz zur Verkehrsdichte in Berlin
- ☒ SimRa-Daten zu Beinahe-Unfällen
- ☒ OSM-Daten (Straßenbeschaffenheit, Art Fahrradwege, etc.)



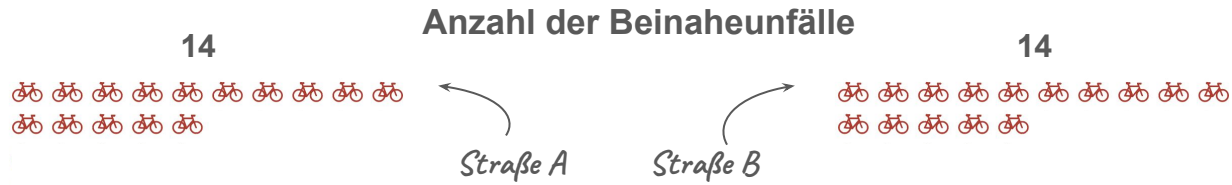


SimRa (Sicherheit im Radverkehr) - Beinahe-Unfälle

- ❑ Sammlung von Daten über Beinaheunfälle im Radverkehr und Fahrradrouten
- ❑ crowdsourcing-basiertes Projekt - Daten werden per SimRa-App erhoben
- ❑ Daten von 2019 bis jetzt
- ❑ unser erster Ansatz:
Datensatz ausschließlich mit Beinaheunfällen → verworfen



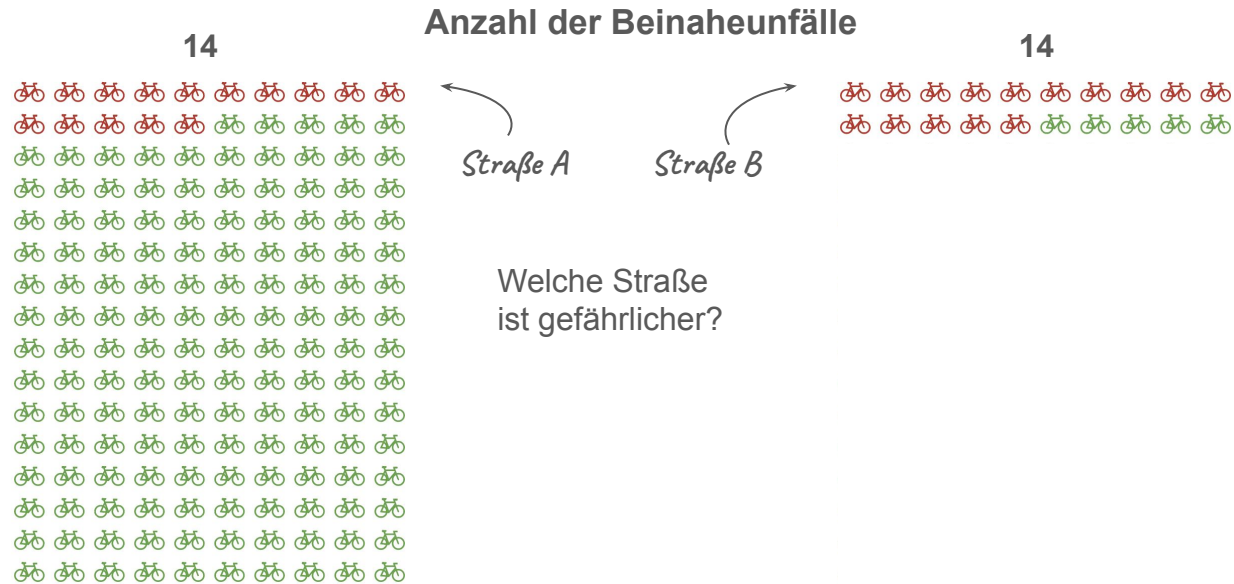
SimRa - Beinahe-Unfälle



Welche Straße
ist gefährlicher?



SimRa - Beinahe-Unfälle



SimRa - Beinahe-Unfälle

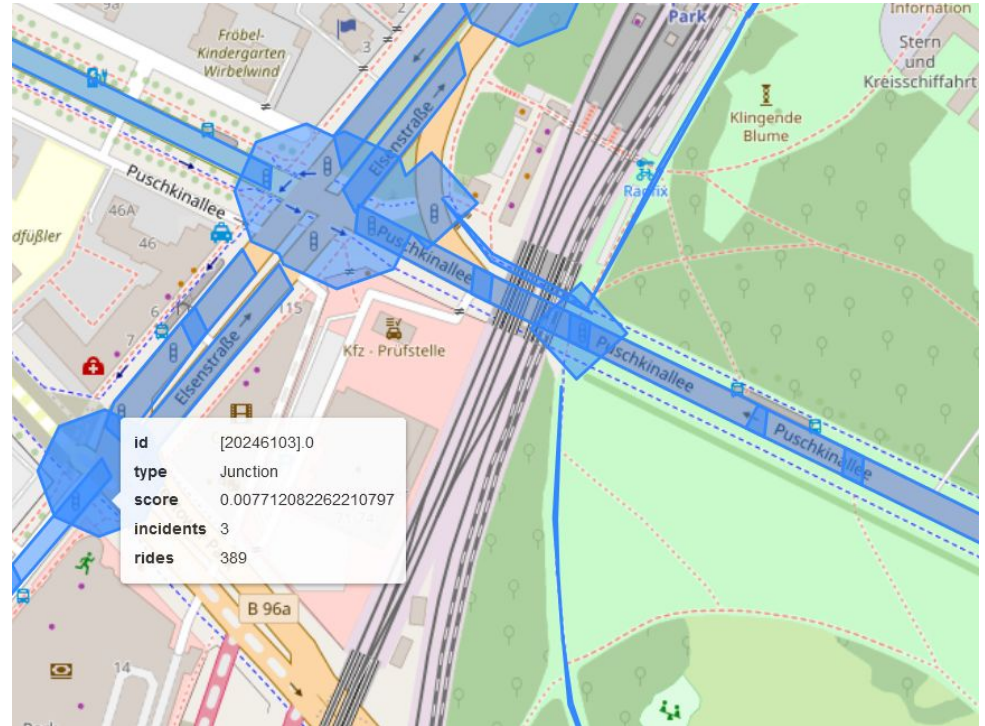
$$\text{Score} = \frac{\# \text{ Beinaheunfälle}}{\# \text{ Fahrten}}$$

Besonders beängstigend: Faktor 4,4



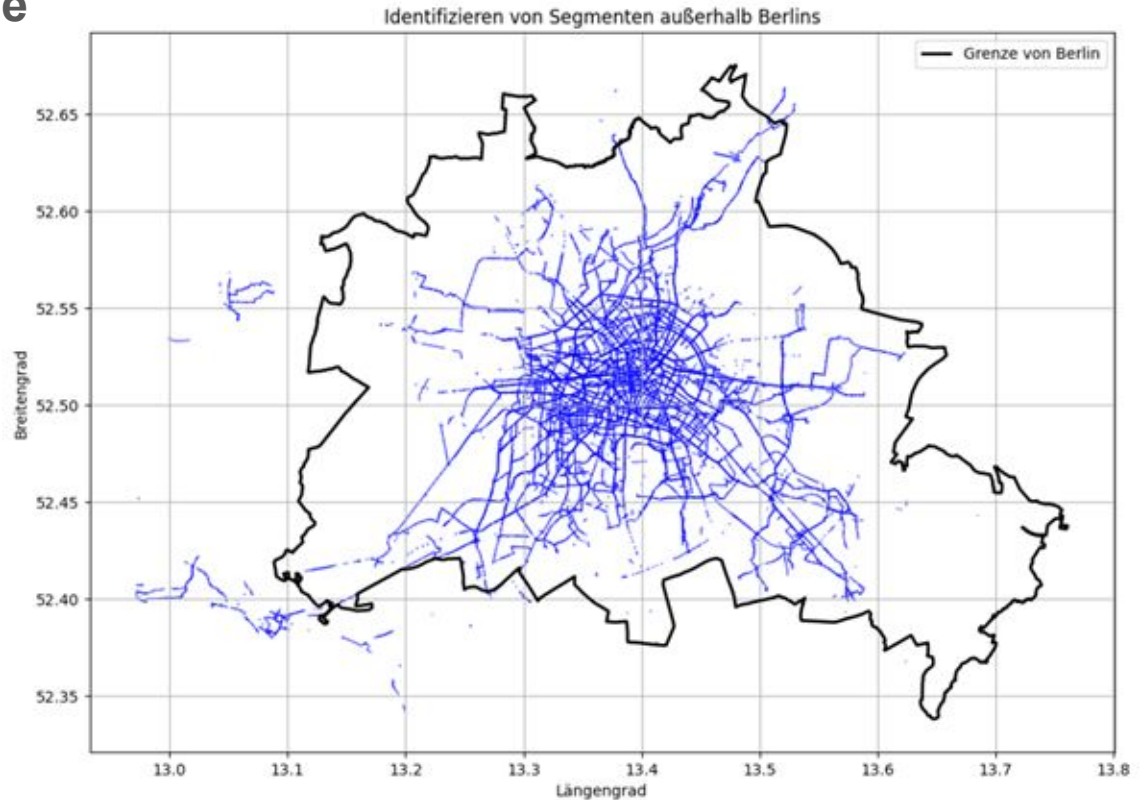
SimRa - Beinahe-Unfälle

- ❑ Geometrien = Polygone
- ❑ nach Fahrten / Beinaheunfällen zusammengefasst
- ❑ 19 Features
- ❑ jedes Polygon:
 - Score
 - incidents
 - rides



SimRa - Beinahe-Unfälle

- ❑ Datenanalyse
Duplikate? Fehlende Werte? Outlier?
- ❑ Räumliche Analyse:
alle Geometrien außerhalb Berlins entfernt
- ❑ Datensatz:
 - vorher: 16819 Einträge
 - nachher: 16294 Einträge



SimRa - Beinahe-Unfälle

	id	type	score	incidents	rides	length	rides south west	incidents south west	score south west	rides north east	...	nlorh	saho	ssho	tailgating	nd	near- dooring	dao	other	markers	geometry			
0	[79310].0	Street	0.000000		0	57	20.129336	57.0		0.0	0.000000	0.0	...	0	0.0	NaN	0	0.0	NaN	0	0	[]	POLYGON ((13.37410 52.53031, 13.37421 52.53020...	
1	[196724641, 196725586, 866264912].0	Junction	0.000649		1	1541	NaN	NaN	NaN	NaN	NaN	...		0	NaN	0.0	0	NaN	0.0	0	0	13.417860660000001, 52.514469009999999]...	[[]]	POLYGON ((13.41751 52.51461, 13.41779 52.51442...
2	[112051].0	Street	0.000000		0	190	18.022919	190.0		0.0	0.000000	0.0	...	0	0.0	NaN	0	0.0	NaN	0	0	[]	POLYGON ((13.50914 52.45236, 13.50899 52.45249...	

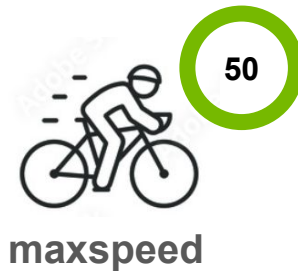


	id	type	score	incidents	rides	geometry
0	[79310].0	Street	0.000000	0	57	POLYGON ((13.37410 52.53031, 13.37421 52.53020...
1	[196724641, 196725586, 866264912].0	Junction	0.000649	1	1541	POLYGON ((13.41751 52.51461, 13.41779 52.51442...
2	[112051].0	Street	0.000000	0	190	POLYGON ((13.50914 52.45236, 13.50899 52.45249...



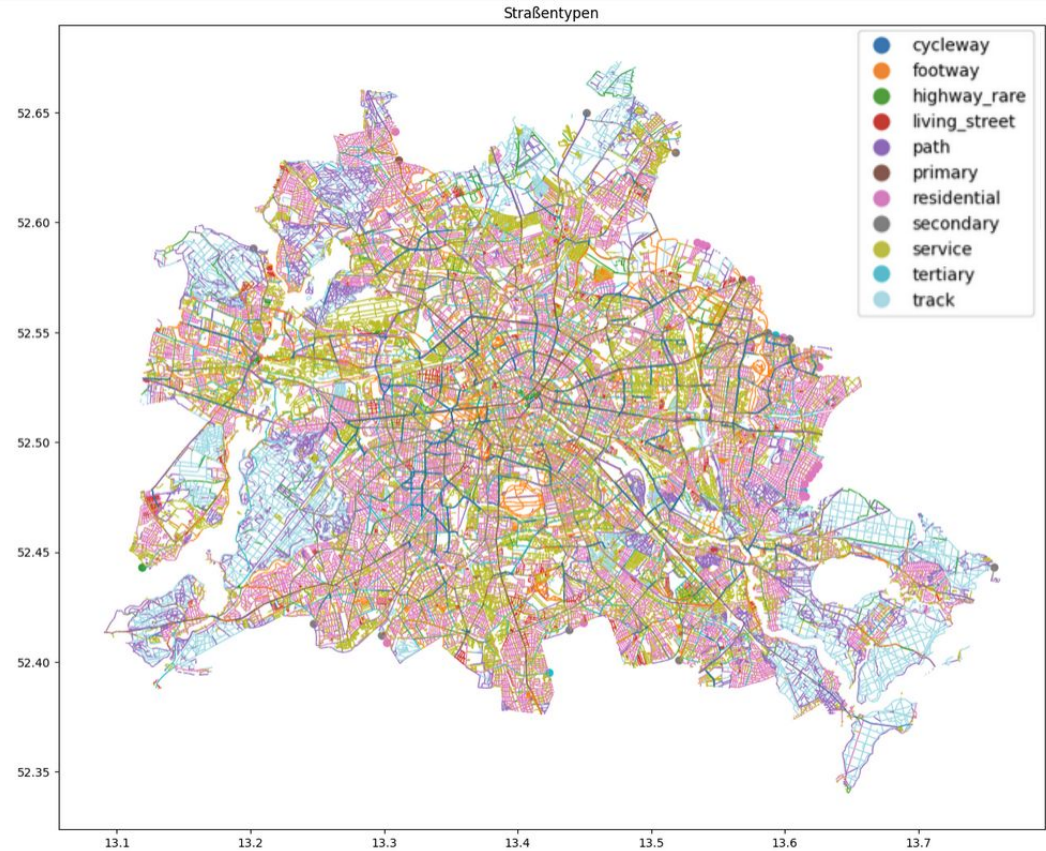
OSM - OpenStreetMap

- ❑ Open Source, durch Community gepflegt
- ❑ offen zugängliche Datenbank aller beigetragenen Geoinformationen
- ❑ Wir nutzen `network_type="cycling"` (Fahrradnetzwerk)
alle Straßen, die mit dem Fahrrad befahren werden können/dürfen
- ❑ enthält 41 Features
- ❑ Features, die wir nutzen wollen:



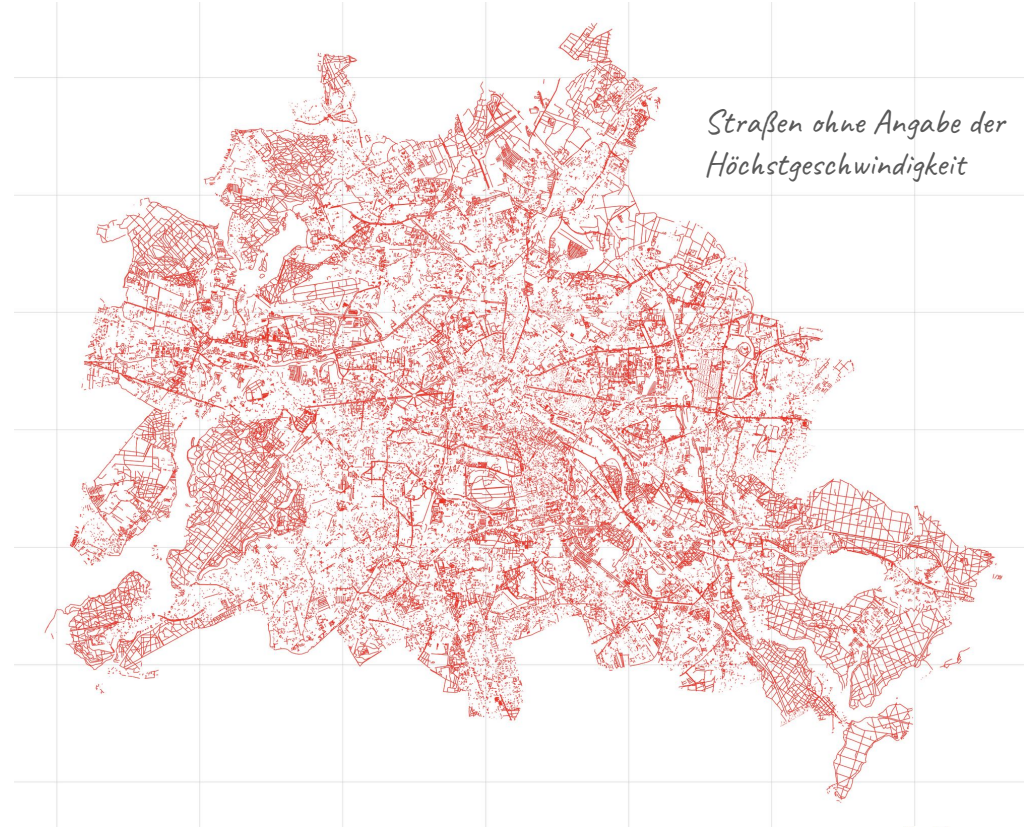
OSM - highway

- ❑ 29 verschiedene Straßentypen in OSM
- ❑ Zusammenfassen seltener Typen:
< 1% → *highway_rare*
- ❑ bearbeiteter Datensatz enthält 11 Straßentypen



OSM - maxspeed

- ❑ von 177.849 Straßen,
119.212 **ohne maxspeed -> 67%!**
- ❑ 18 verschiedene maxspeed-Typen
numerisch u. nicht numerisch
- ❑ fehlende Werte je nach
Straßentyp gesetzt
- ❑ Kategorien gebildet
z.B.: 5 km/h: walk, 5, 6, 7
- ❑ nach Bearbeitung:
7 verschiedene maxspeed-Typen
keine NaN-Werte

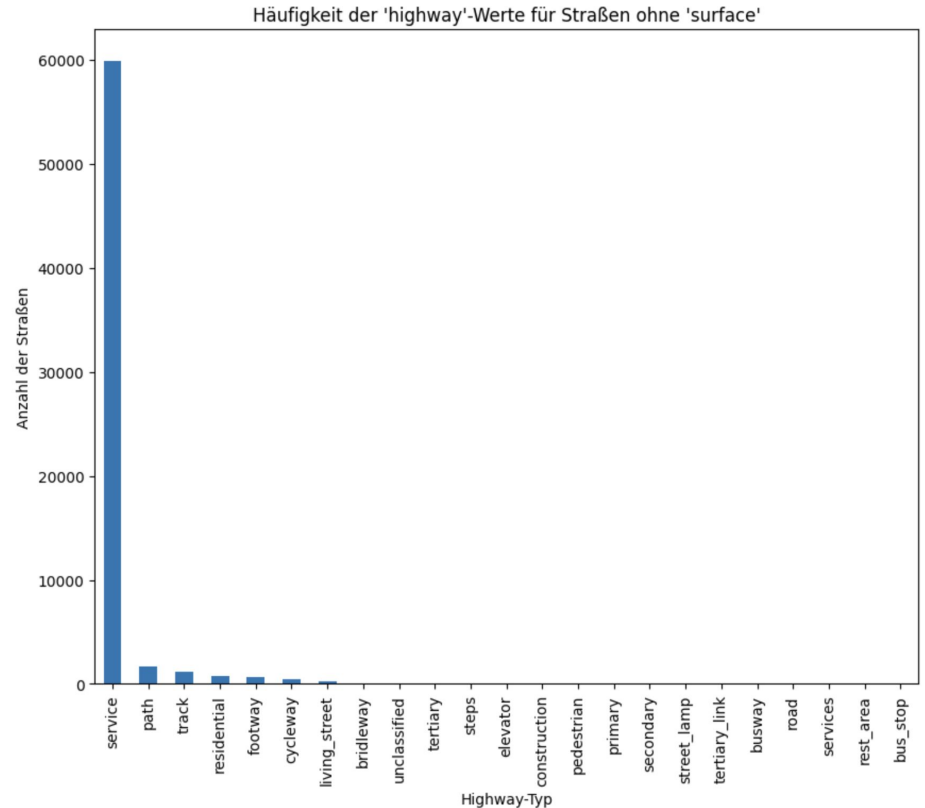


OSM - surface

- ❑ 36% ohne surface (65001 Straßen)
- ❑ 36 verschiedene surface-Typen
- ❑ teilweise doppelt gesetzte Werte
- ❑ fehlende Werte je nach Straßentyp gesetzt
- ❑ Zusammenfassung in Kategorien

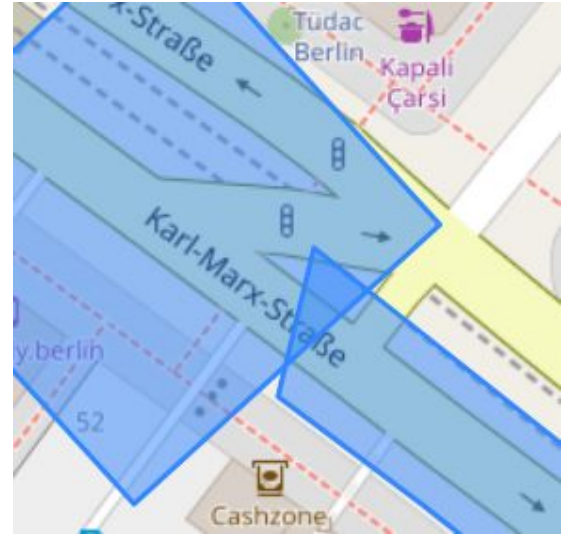
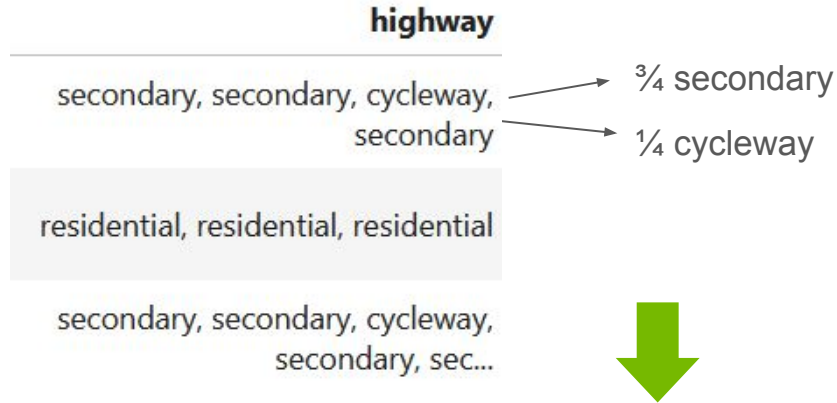
Die Gesamtzahl der Straßen ist: 177849
Die Anzahl der Straßen ohne 'surface'-Wert ist: 0
Prozentualer Anteil der Straßen ohne 'surface'-Wert: 0.00%

surface_category	Anzahl Strassen
0 paving_stone	62651
1 asphalt	57637
2 unpaved	47378
3 concrete	9330
4 sett	853



OSM + SimRa Räumliche Verknüpfung

Dummy-Variablen für die **highway**-Spalte



	cycleway	footway	highway_rare	living_street	path	primary	residential	secondary	service	tertiary	track
...	0.250000	0.0	0.0	0.0	0.0	0.0	0.000000	0.750000	0.0	0.000000	0.0





Erkenntnisse und Ausblick

- ❑ Datensätze bereinigt
- ❑ zusammengeführt
- ❑ Vorhersage des Gefahrenpotentials basierend auf den OSM-Merkmalen: Straßentyp, Geschwindigkeit & Straßenbelag
- ❑ Training des Modells steht an





 **Danke!**