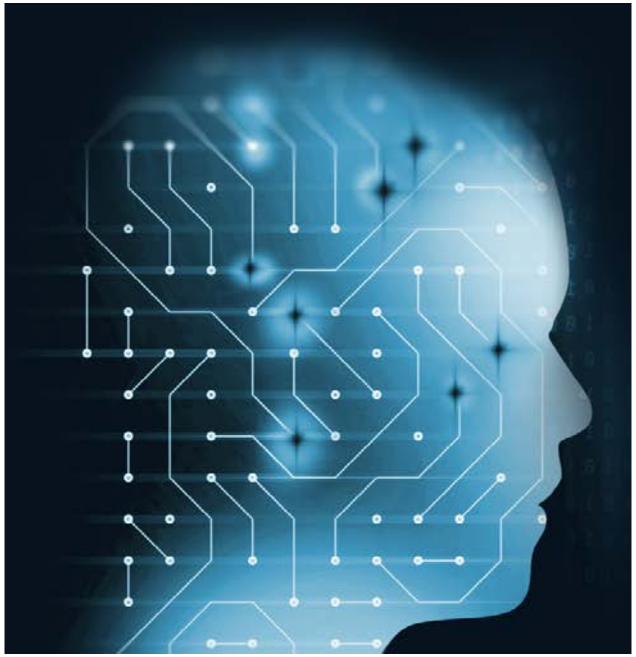


Explainable AI in Urban Applications

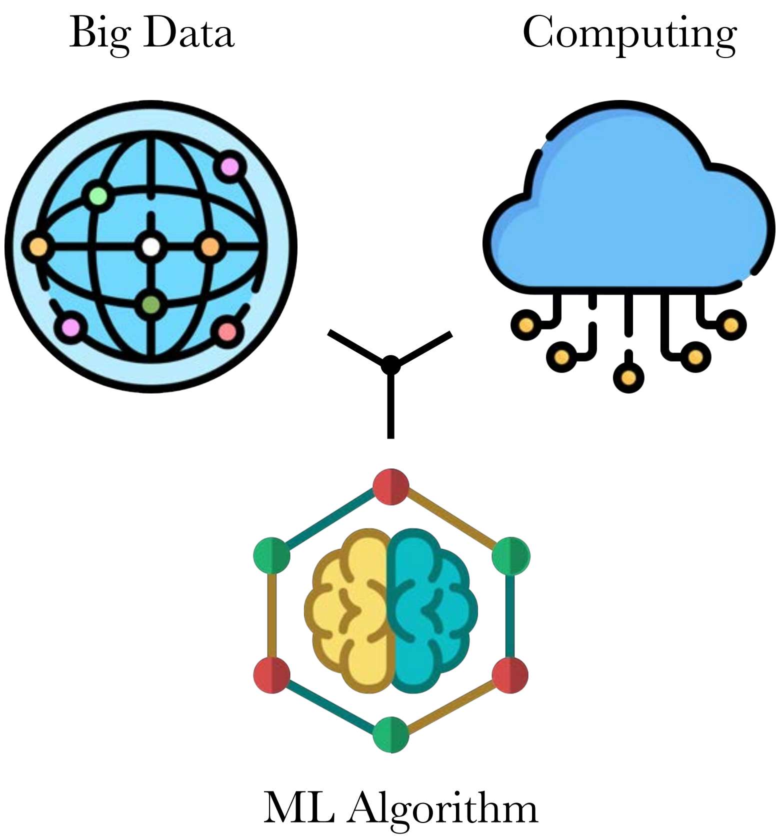
Ziqi Li

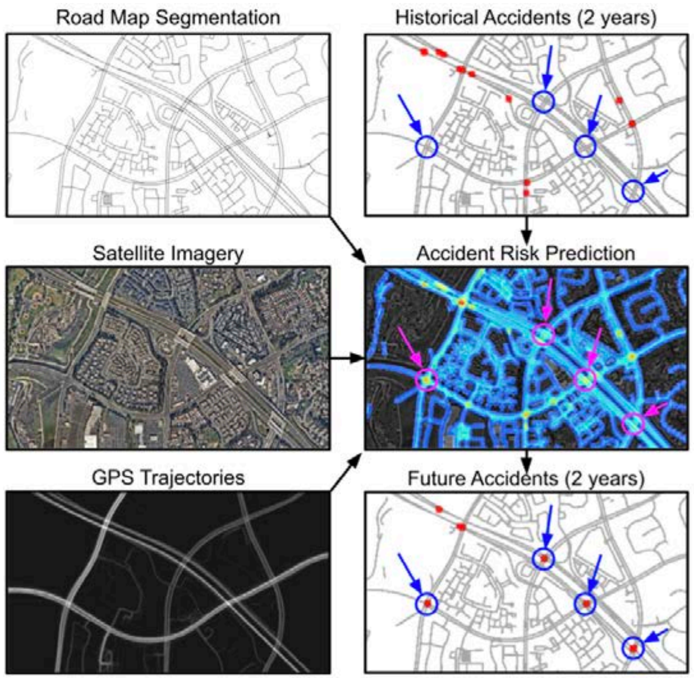
School of Geographical and Earth Sciences

Today

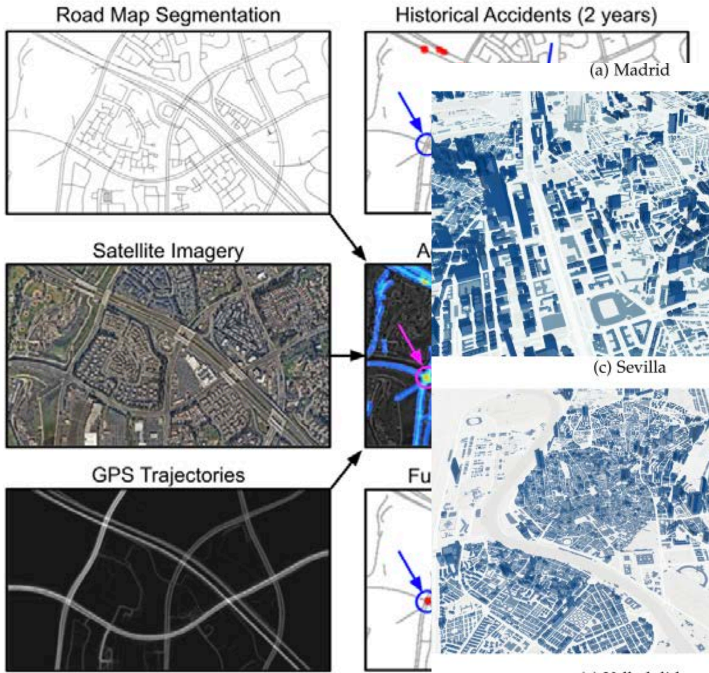
- What is eXplainable AI (XAI) and why does it matter?
- A case study of modelling ride-sharing preferences using XAI and big data.

Emergence of AI

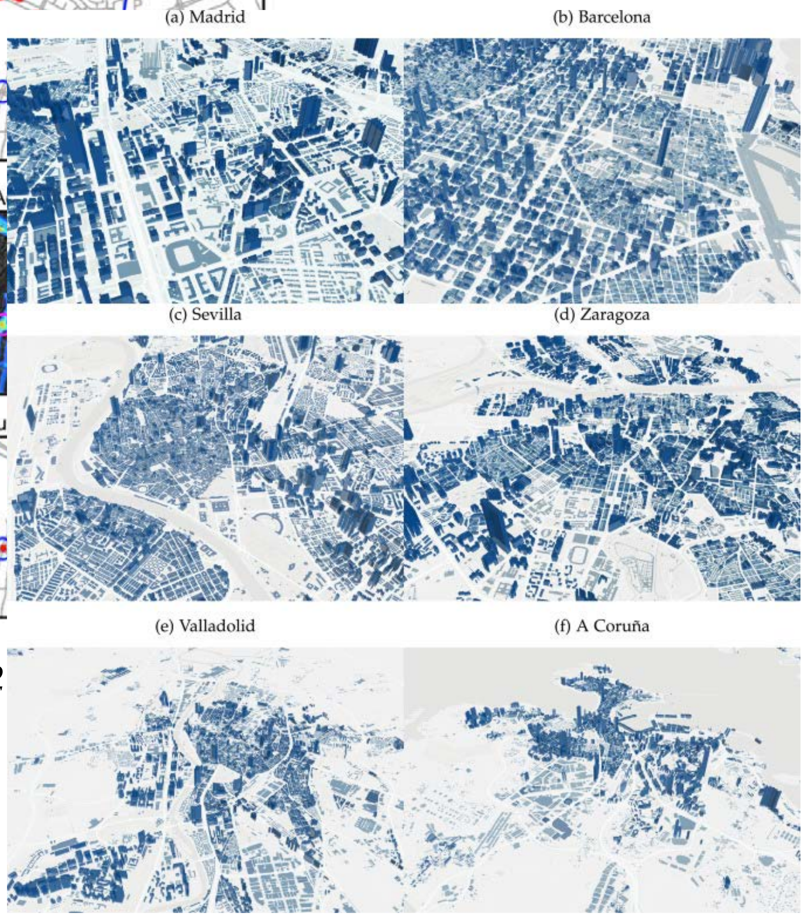



Road Map Segmentation
Historical Accidents (2 years)
Satellite Imagery
Accident Risk Prediction
GPS Trajectories
Future Accidents (2 years)

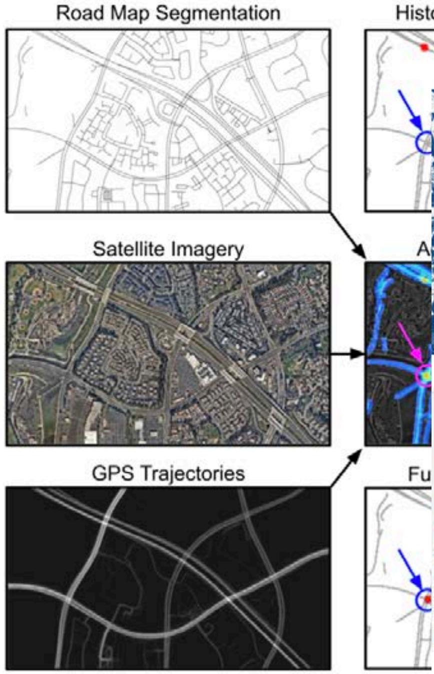
He et al., 2021



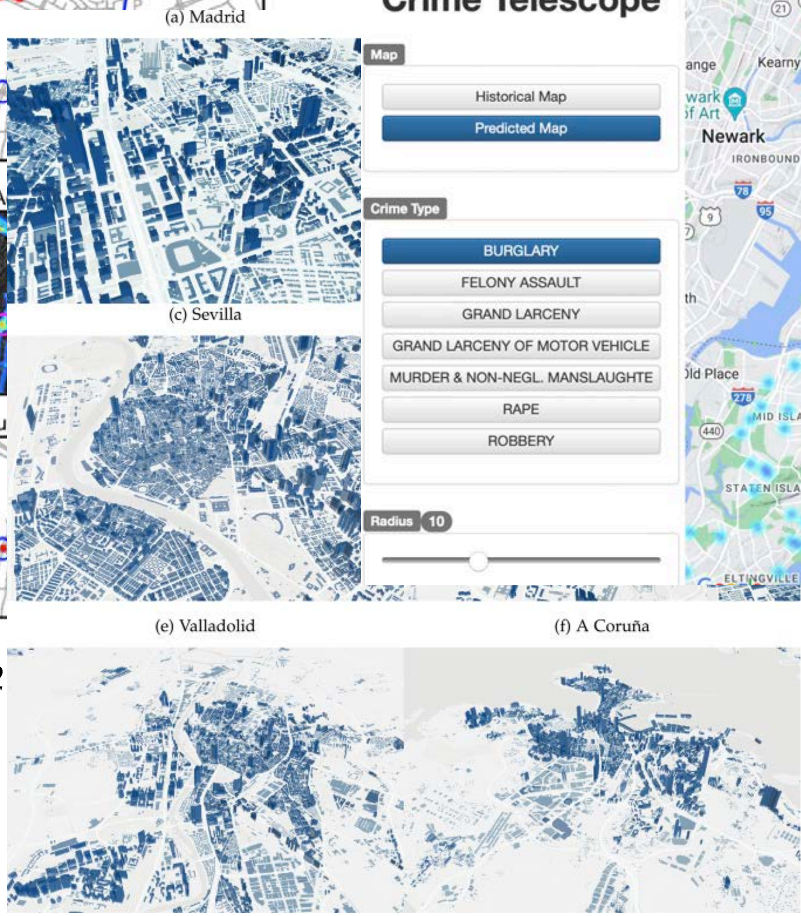
He et al., 202



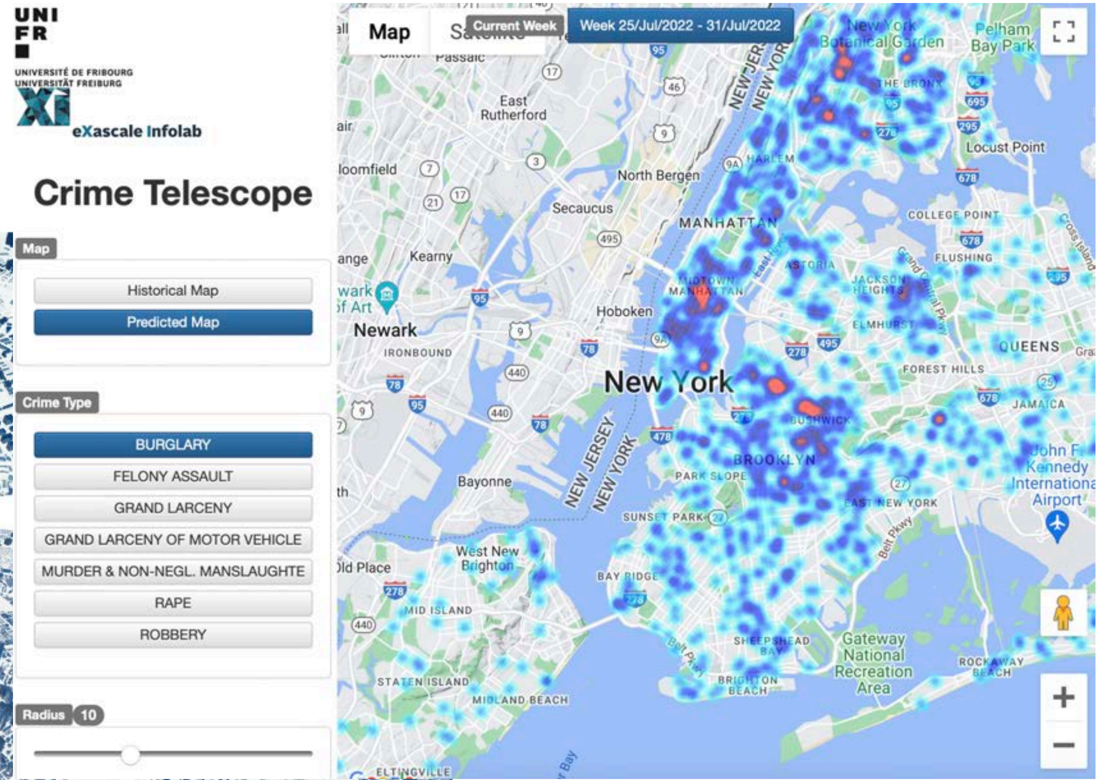
Arribas-Bel et al., 2021



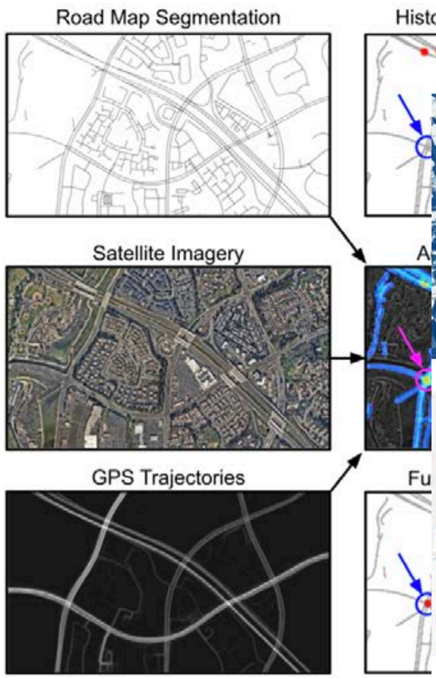
He et al., 202



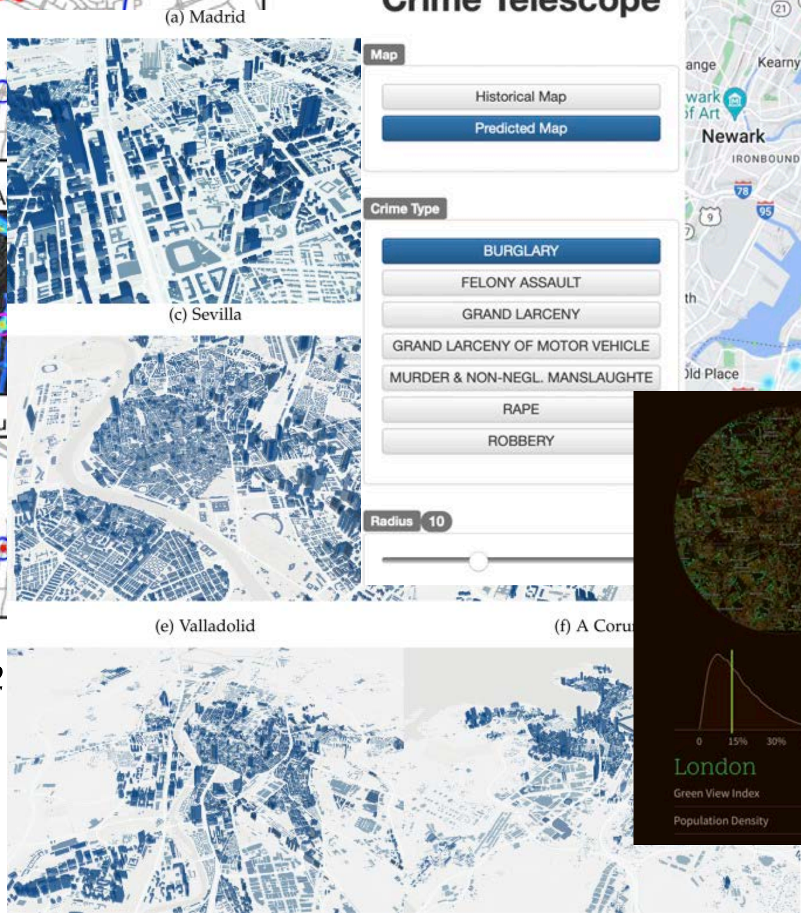
Arribas-Bel et al., 2021



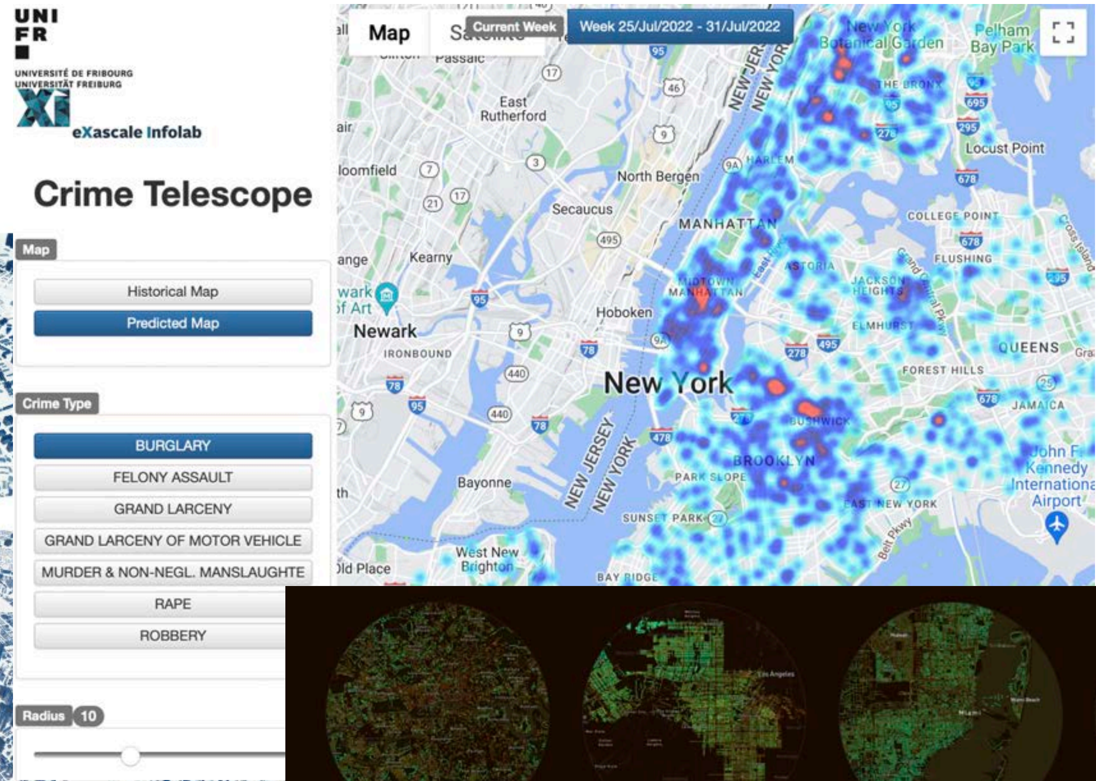
Yang et al., 2018



He et al., 202



Arribas-Bel et al., 2021

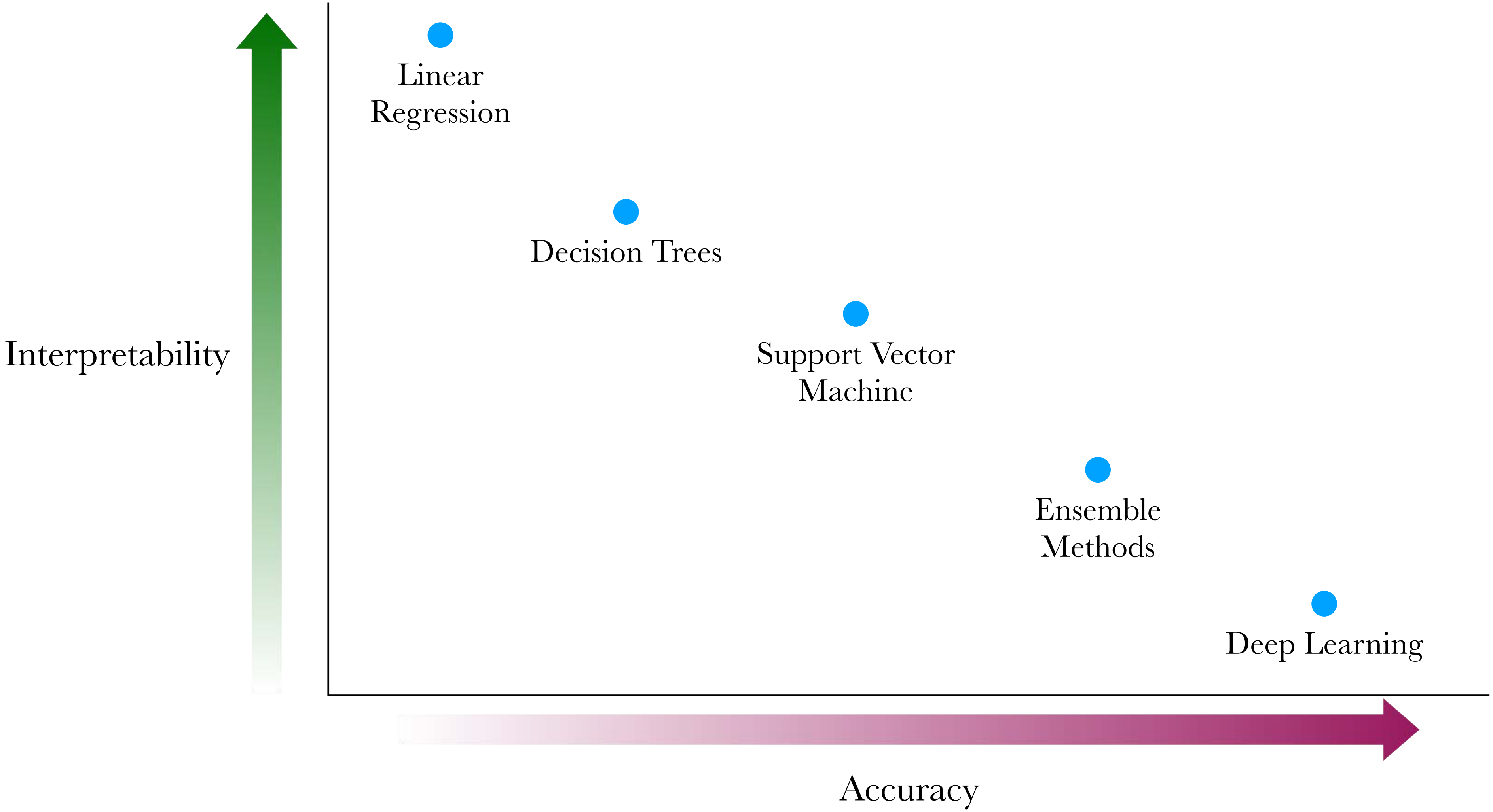


MIT Senseable City Lab

Why AI/ML

- Unstructured data (e.g. image, video, text, speech, GPS, point clouds, etc.)
- Scalability for big data
- Fewer assumptions (distribution, relationship)
- Automated model selection
- Superior predictive accuracy

Tradeoff between accuracy and interpretability



Black box of AI

- AI models are intrinsically hard to interpret due to a huge number of parameters being estimated.

Google AI Blog

The latest from Google Research

NVIDIA. DEVELOPER HOME BLOG FORUMS DOCS

Over 500 GTC sessions now

TECHNICAL BLOG

NEWS

OpenAI Presents GPT-3, a 175 Billion Parameters Language Model

By Nefi Alarcon

Discuss (0) Share 0 Like

Tags: featured, Machine Learning & Artificial Intelligence, News, Speech & Audio Processing, Supercomp

Pathways Language Model (PaLM): Scaling to 540 Billion Parameters for Breakthrough Performance

Monday, April 4, 2022

Posted by Sharan Narang and Aakanksha Chowdhery, So

In recent years, large neural networks trained for lang
achieved impressive results across a wide range of ta
models (LLMs) can be used for *few-shot learning* and
scale task-specific data collection or model paramete
LaMDA, Gopher, and Megatron-Turing NLG, achieved
tasks by scaling model size, using sparsely activated
more diverse sources. Yet much work remains in und
few-shot learning as we push the limits of model scal

Alberto Romero

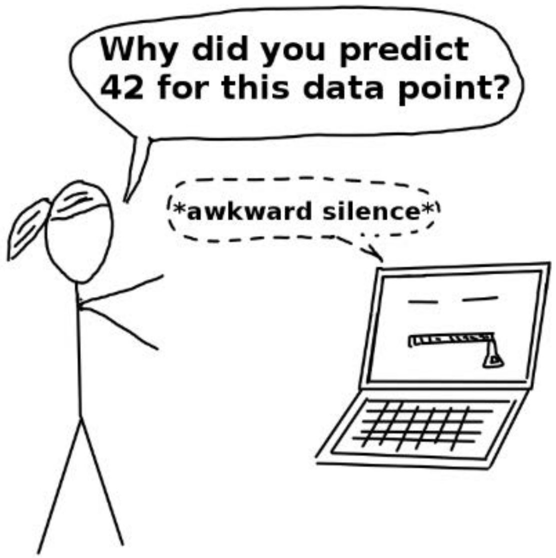
Sep 11, 2021 · 6 min read ★ ·

ARTIFICIAL INTELLIGENCE | NEWS

GPT-4 Will Have 100 Trillion Parameters — 500x the Size of GPT-3

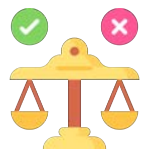
Are there any limits to large neural networks?

Black box of AI

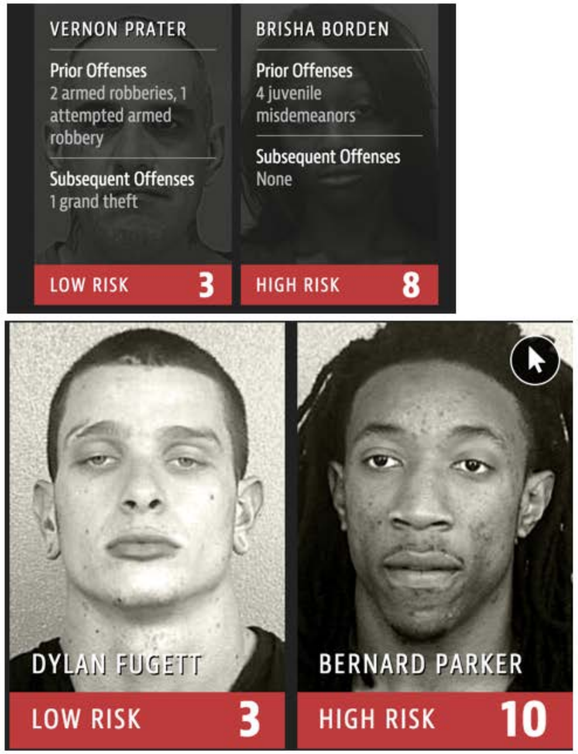


Trust issues with black-box AI

- Technical:
  - Why a certain decision is made; when does the system work/fail, how to correct the error and improve the model?
- Ethical:
  - Critical decisions are made by AI: healthcare, finance, security, etc.
  - Discrimination and biases



Racial bias in COMPAS



Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Black offenders were seen almost twice as likely as white offenders to be labeled a higher risk but not actually re-offend.

Angwin et al. (2016)

nature

Explore content | About the journal | Publish with us | Subscribe

nature > news > article

NEWS | 24 October 2019 | Update 26 October 2019

Millions of black people affected by racial bias in health-care algorithms

Study reveals rampant racism in decision-making software used by US hospitals — and highlights ways to correct it.

Heidi Ledford

Science

Current Issue | First release papers | Archive | About | Submit manus

RESEARCH ARTICLE

Dissecting racial bias in an algorithm used to manage the health of populations

ZIAD OBERMEYER, BRIAN POWERS, CHRISTINE VOGELI, AND, SENDHIL MULLAINATHAN Authors Info & Affiliations

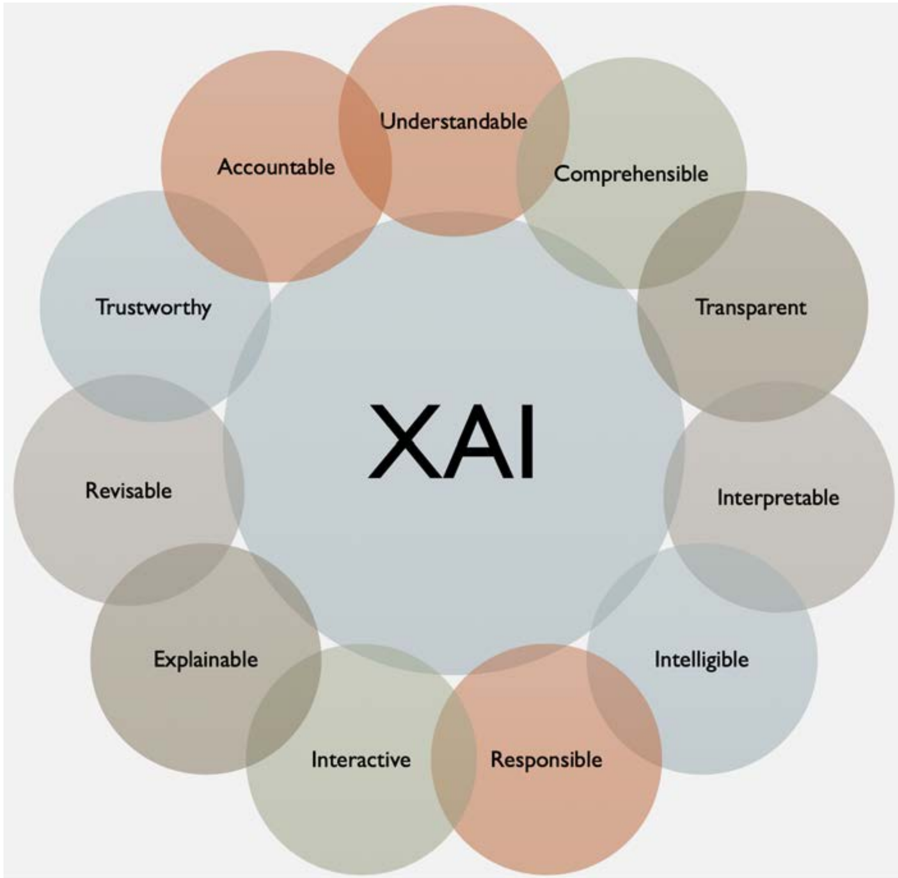
SCIENCE • 25 Oct 2019 • Vol 366, Issue 6464 • pp. 447-453 • DOI: 10.1126/science.aax2342

Racial bias in health algorithms

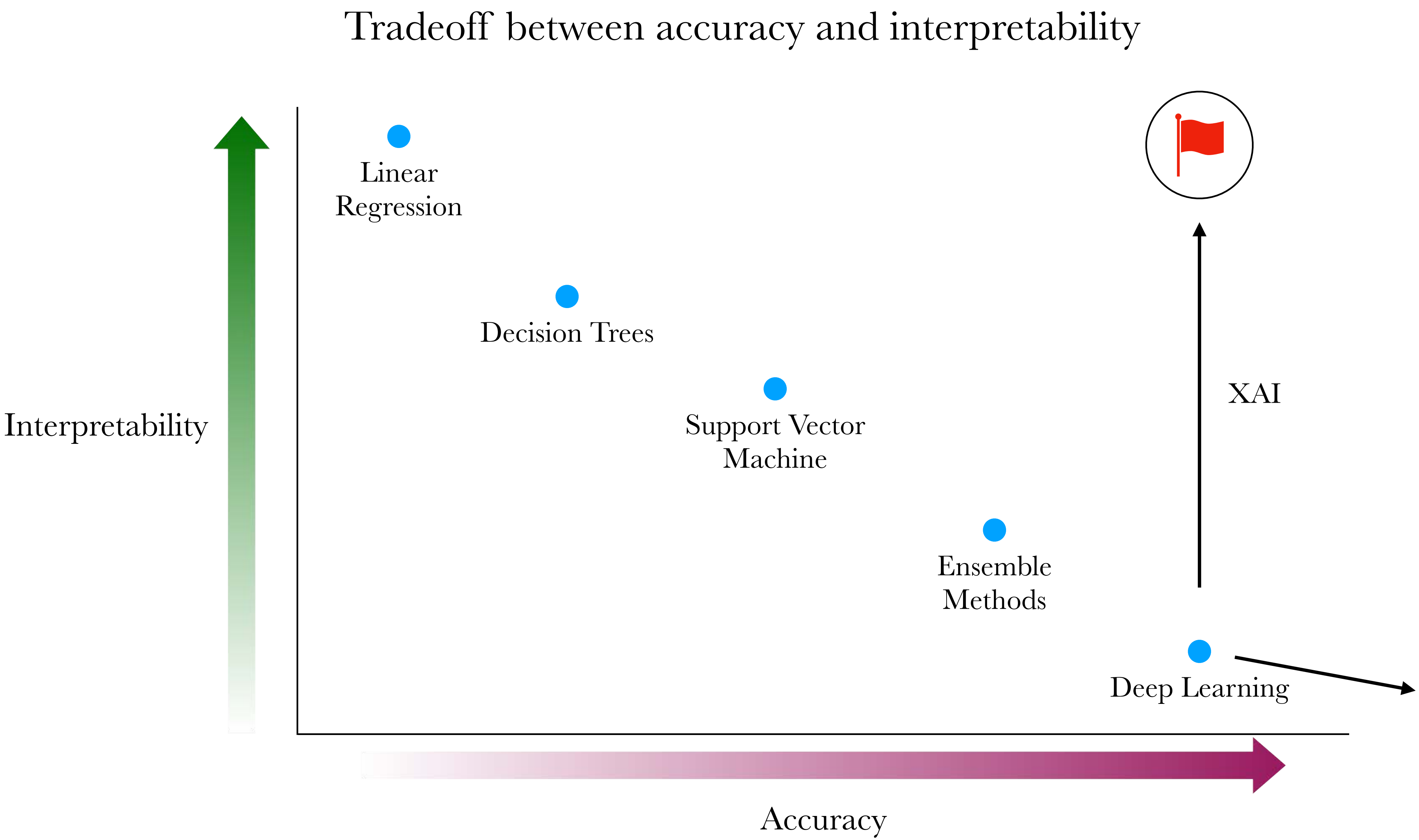
The U.S. health care system uses commercial algorithms to guide health decisions. Obermeyer *et al.* find evidence of racial bias in one widely used algorithm, such that Black patients assigned the same level of risk by the algorithm are sicker than White patients (see the Perspective by Benjamin). The authors estimated that this racial bias reduces the number of Black patients identified for extra care by more than half. Bias occurs because the algorithm uses health costs as a proxy for health needs. Less money is spent on Black patients who have the same level of need, and

Explainable AI (XAI)

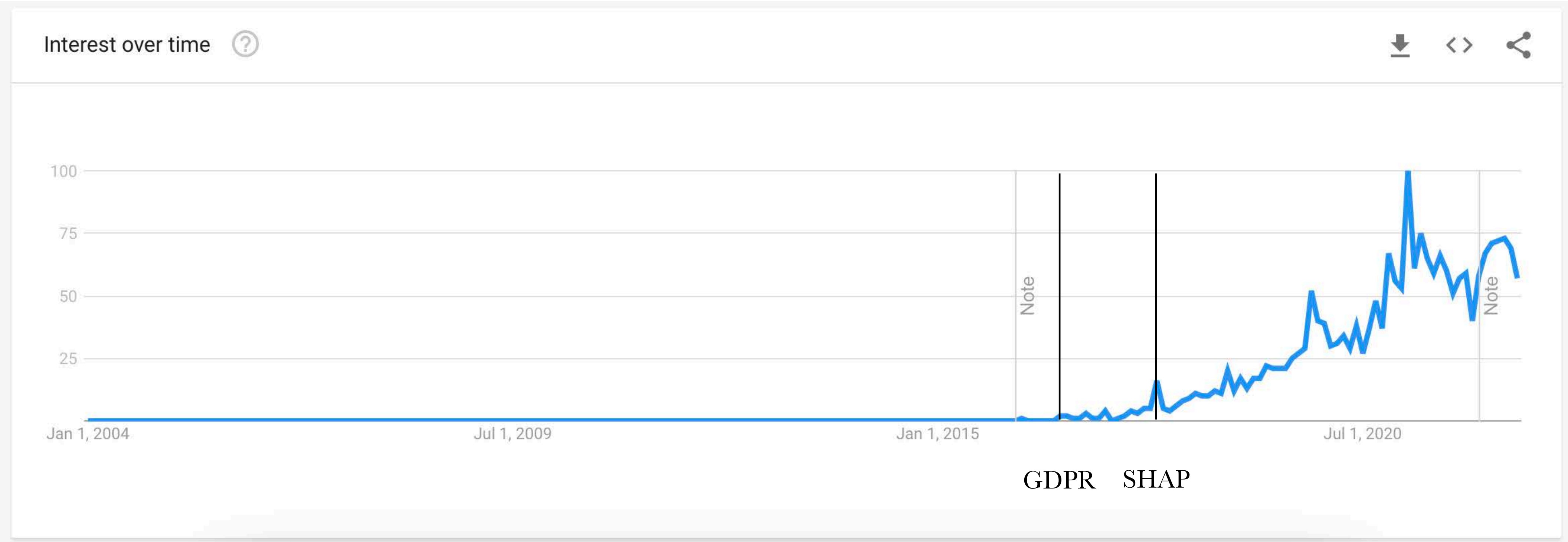
- Explainable artificial intelligence (XAI) is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms.
  - Improve understanding of the underlying decision processes.
  - Provide credibility and confidence of the model parameters and outcome.



https://miro.medium.com/max/1400/1*-vNI5N7f1GUBWFZUfwfd8w.png


Tradeoff between accuracy and interpretability
Interpretability
Linear Regression
Decision Trees
Support Vector Machine
Ensemble Methods
Deep Learning
XAI
Accuracy

Google trends searching "XAI"



XAI in urban applications

ELSEVIER

Computers, Environment and Urban Systems

Volume 94, June 2022, 101789

Interpretable machine learning models for crime prediction

Xu Zhang [a, b], Lin Liu [b, c], Minxuan Lan [d], Guangwen Song [b], Luzi Xiao [b], Jianguo Chen [b]

Show more

ELSEVIER

Science of The Total Environment

Volume 761, 20 March 2021, 144057

Predicting stream water quality under different urban development pattern scenarios with an interpretable machine learning approach

Runzi Wang [a], Jun-Hyun Kim [b], Ming-Han Li [b]

ELSEVIER

Transportation Research Part C: Emerging Technologies

Volume 124, March 2021, 102962

Decoding pedestrian and automated vehicle interactions using immersive virtual reality and interpretable deep learning

Arash Kalatian, Bilal Farooq

Show more

ELSEVIER

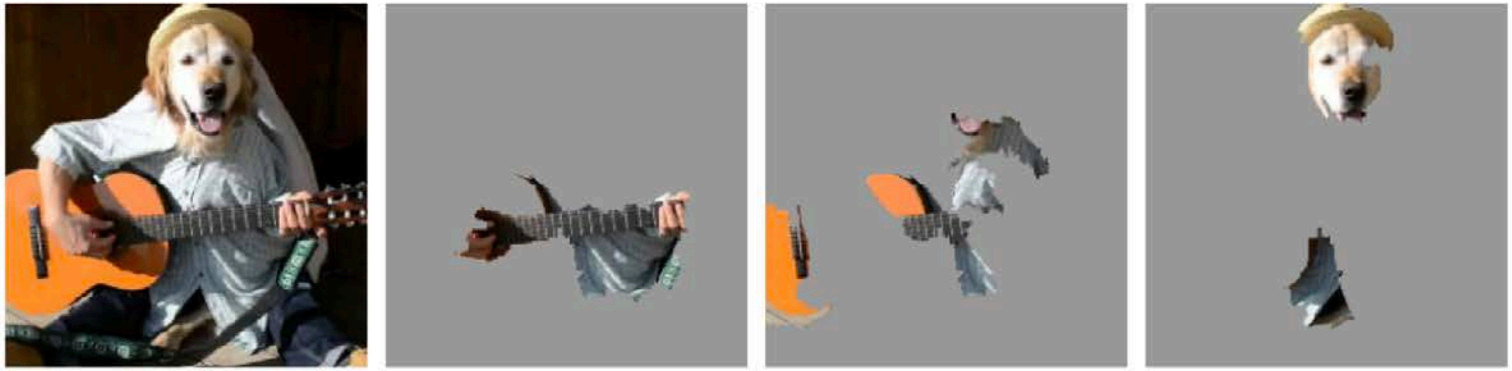
Neurocomputing

Volume 468, 11 January 2022, Pages 123-136

Hybrid interpretable predictive machine learning model for air pollution prediction

Yuanlin Gu [a], Baihua Li [b], Qinggang Meng [c]

XAI for images

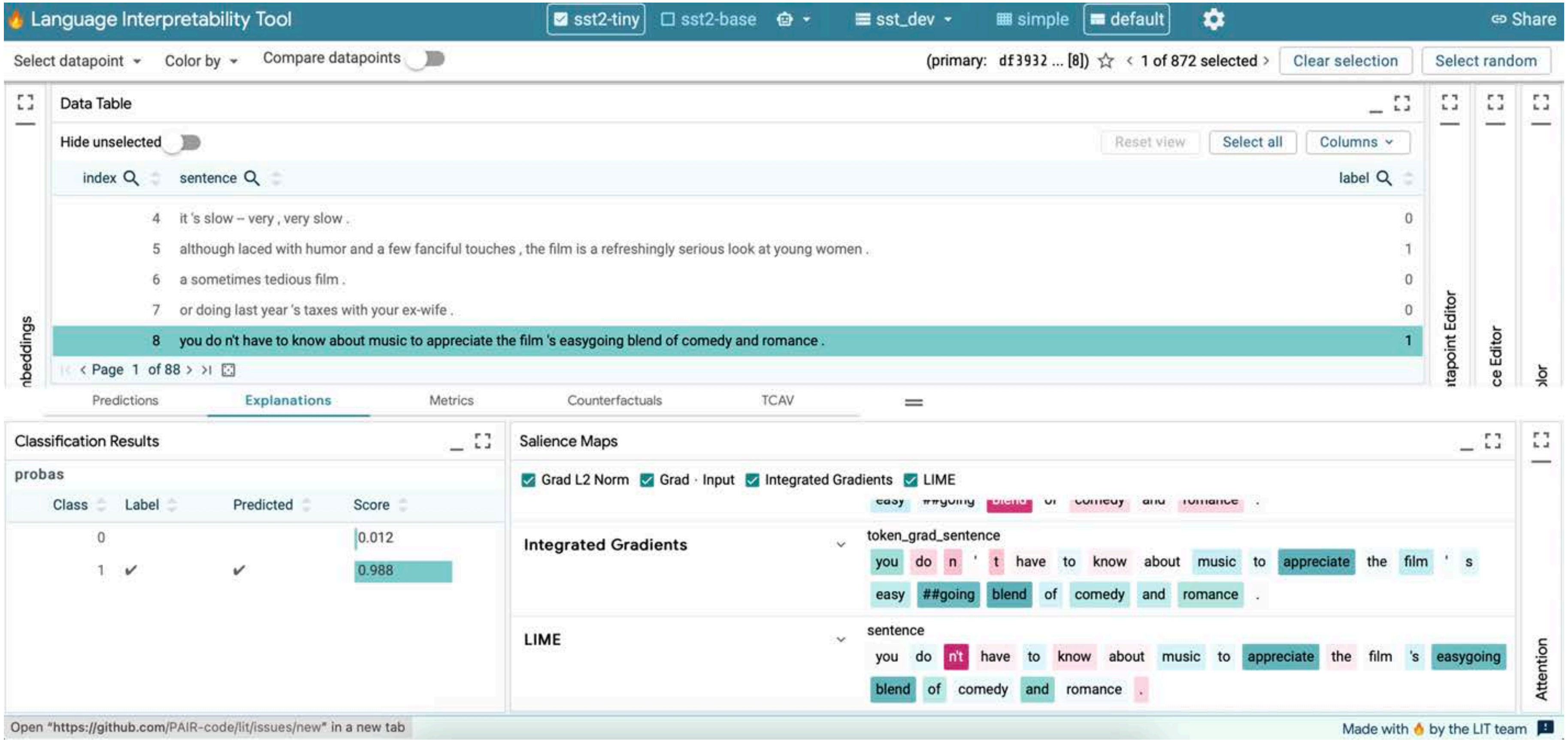


(a) Original Image (b) Explaining Electric guitar (c) Explaining Acoustic guitar (d) Explaining Labrador

Fig. 9. Local explanations of an image classification prediction described using LIME [65]. Here, top three classes are "electric guitar" $(p = 0.32)$, "acoustic guitar" $(p = 0.24)$ and "labrador" $(p = 0.21)$. By selecting a group of 'superpixels' from the input image, the classifier provides visual explanations to the top predicted labels.

Das and Rad, 2020

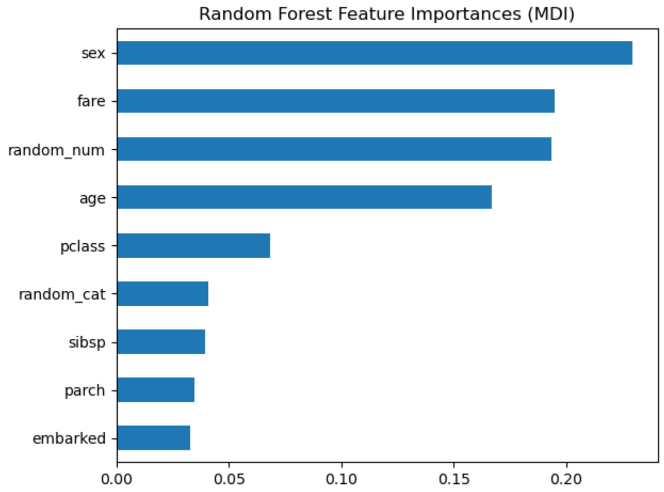
XAI for text

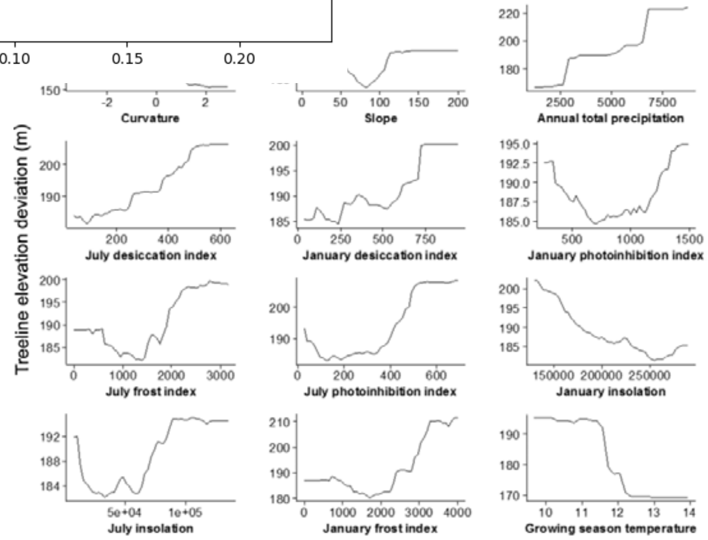


https://pair-code.github.io/lit/demos/

XAI for tabular data

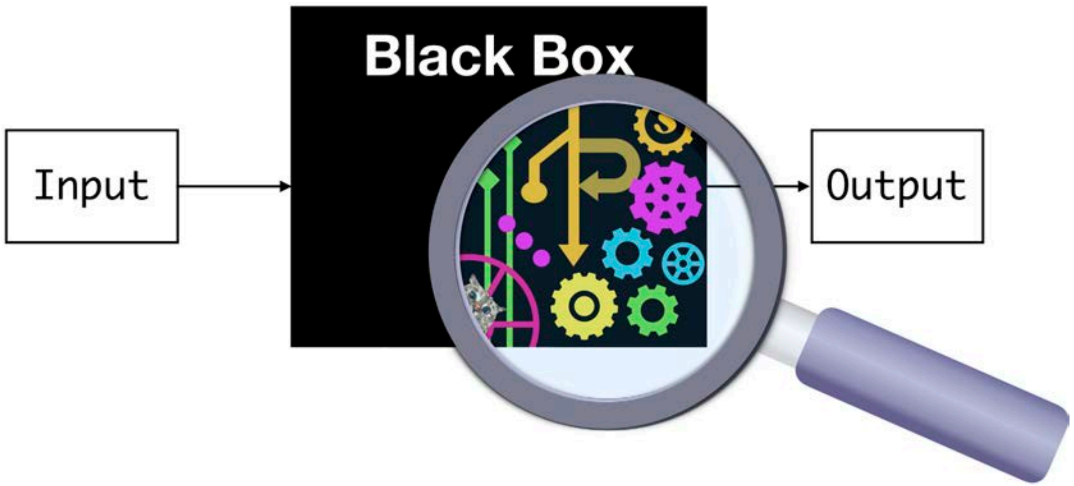
- Feature “importance”
    - What are the major contributors to the model.
- Partial dependence plot
    - Relationships between X and y.





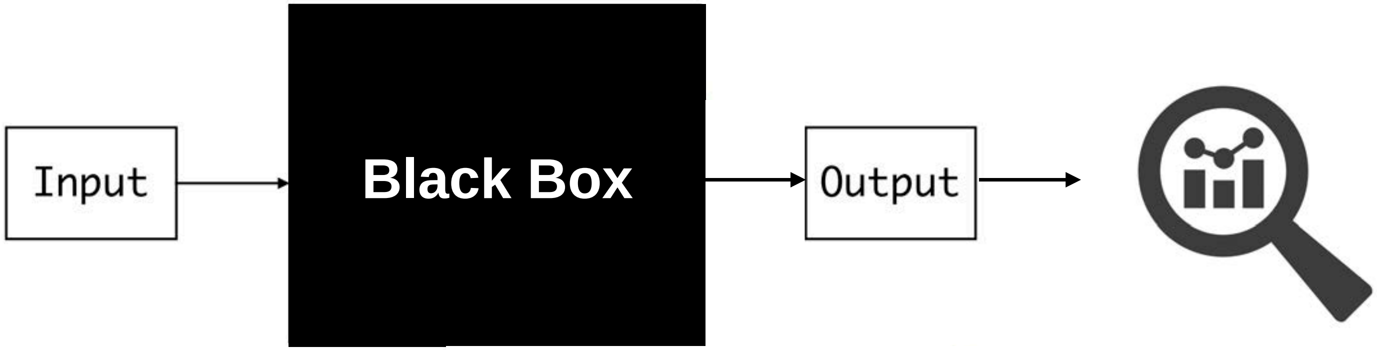
Two types of model explanations

Model-based



Looking at internal structure of a model
Model-specific
ML Expert

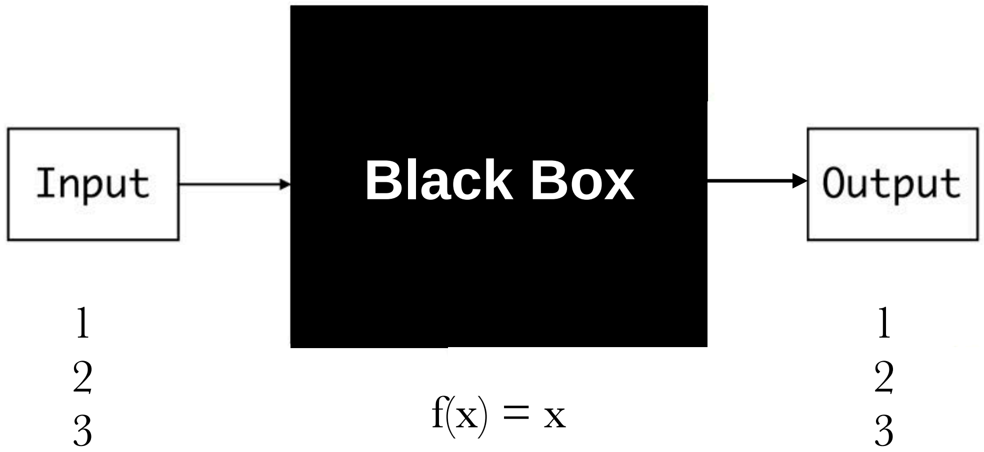
Post-hoc

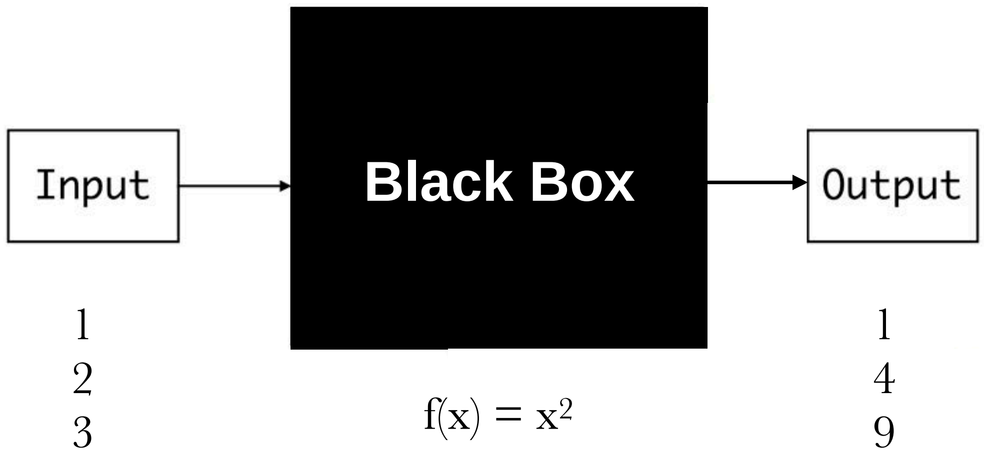


Explaining predictions from a model
Model-agnostic
Analysts and non-experts

Murdoch et al., 2019. PNAS

Post-hoc explanation

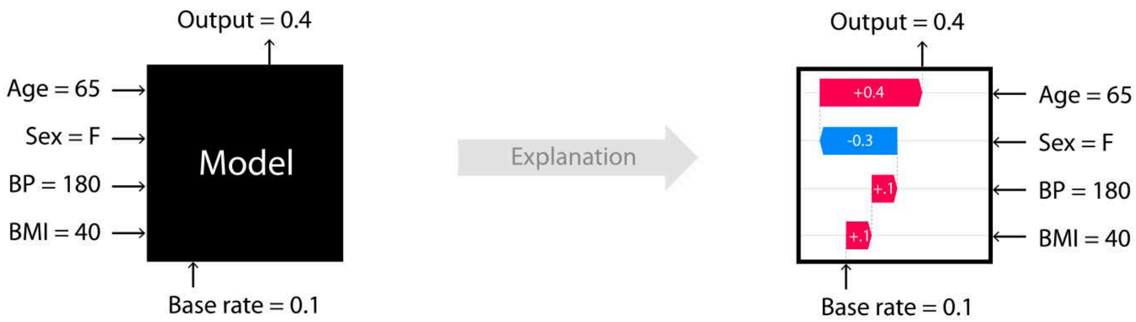




SHAP

- SHAP (Shapley Additive Explanations): is a game theoretic approach to quantify the contribution of each feature in the model that collectively makes the prediction.

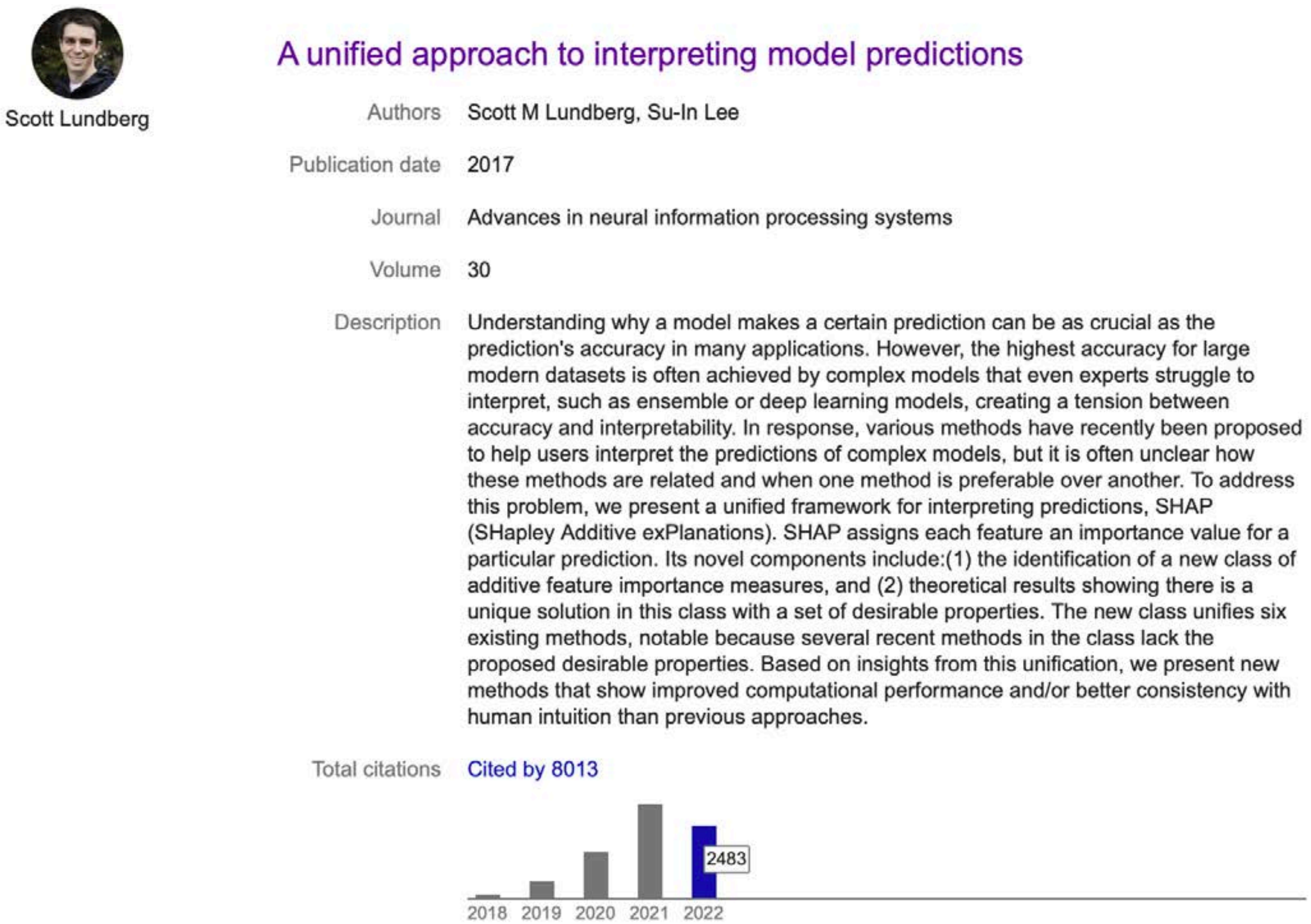
$$\hat{y}_i = shap_0 + shap(X_{1i}) + shap(X_{2i}) + \ldots + shap(X_{pi})$$



https://github.com/slundberg/shap (Lundberg and Lee, 2017)

SHAP

- SHAP (Sh uantify the contributi liction.



https://github.com/slundberg/shap (Lundberg and Lee, 2017)

Shapley

Lloyd Shapley (1923-2016)
Nobel Prize in Economics (2012)

$$\varphi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \overbrace{\frac{|S|!\;(n-|S|-1)!}{n!}}^{\text{Possible permutations}} \underbrace{(v(S \cup \{i\}) - v(S))}_{\text{Marginal contribution}}$$

Shapley value measures -
the average of marginal contribution of a player in a game over all possible different permutations in which the coalition can be formed.

A Shapley value example

- Null: 0
- Ziqi: 5
- Qunshan: 10
- Nick: 100

- {Qunshan + Ziqi}: 5
- {Nick + Ziqi}: 120
- {Nick + Qunshan}: 140
- {Nick + Qunshan + Ziqi}: 150

- So, what is the contribution of each us?

A Shapley value example

- Null: 0
- Ziqi: 5
- Qunshan: 10
- Nick: 100
- {Qunshan + Ziqi}: 5
- {Nick + Ziqi}: 120
- {Nick + Qunshan}: 140
- {Nick + Qunshan + Ziqi}: 150

- So, what is the contribution of each us?

Shapley(Ziqi) = {v(Ziqi) - v(Null)}/3 +
{v(Ziqi + Qunshan) - v(Qunshan)}/6 +
{v(Ziqi + Nick) - v(Nick)}/6 +
{v(Ziqi + Qunshan + Nick) - v(Qunshan + Nick)}/3
= (5)/3 + (-5)/6 + (20)/6 + (10)/3 = 45/6 = **7.5**

A Shapley value example

- Null: 0
- Ziqi: 5
- Qunshan: 10
- Nick: 100
- {Qunshan + Ziqi}: 5
- {Nick + Ziqi}: 120
- {Nick + Qunshan}: 140
- {Nick + Qunshan + Ziqi}: 150

- So, what is the contribution of each us?

Shapley(Ziqi) = {v(Ziqi) - v(Null)}/3 +
{v(Ziqi + Qunshan) - v(Qunshan)}/6 +
{v(Ziqi + Nick) - v(Nick)}/6 +
{v(Ziqi + Qunshan + Nick) - v(Qunshan + Nick)}/3
= (5)/3 + (-5)/6 + (20)/6 + (10)/3 = 45/6 = **7.5**

- Shapley(Qunshan): **20**
- Shapley(Nick): **122.5**

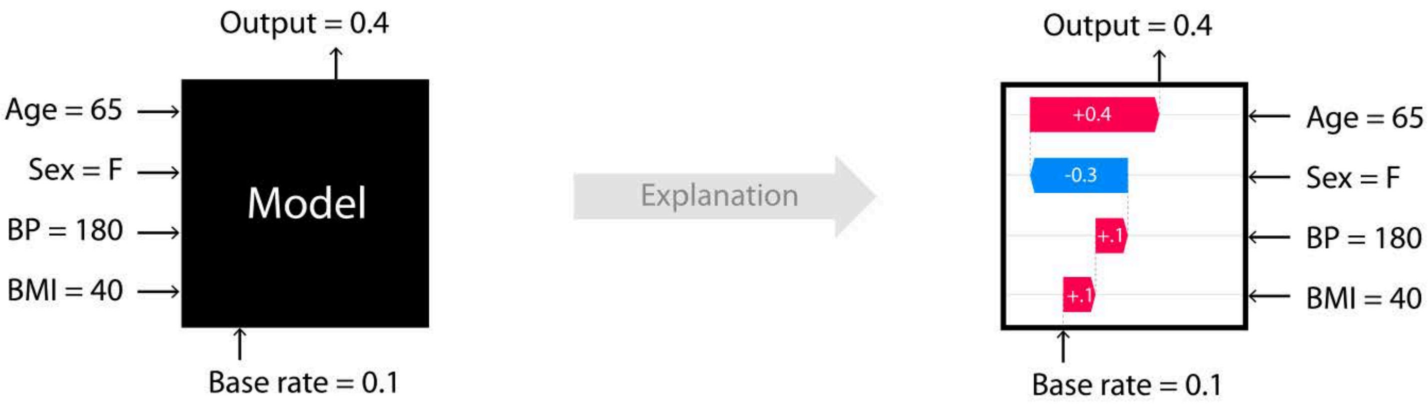
Shapley

- Shapley value is the unique solution for a fair distribution that has certain properties:
    - Null, additivity, efficiency, symmetry properties
- Shapley in ML:
    - Game-> model, player -> feature; outcome->prediction
    - Shapley value of a feature is the feature's contribution to the model prediction

Shapley

- The computation of Shapley value is NP-hard.
  - SHAP provides different approximation methods to estimate Shapley value.
  - Sampling based, Kernel based, Tree based, etc.
- SHAP also unifies some other XAI methods for text and image explanations.
- SHAP has been integrated into industry XAI software such as Amazon SageMaker and Google Vertex Explainable AI.

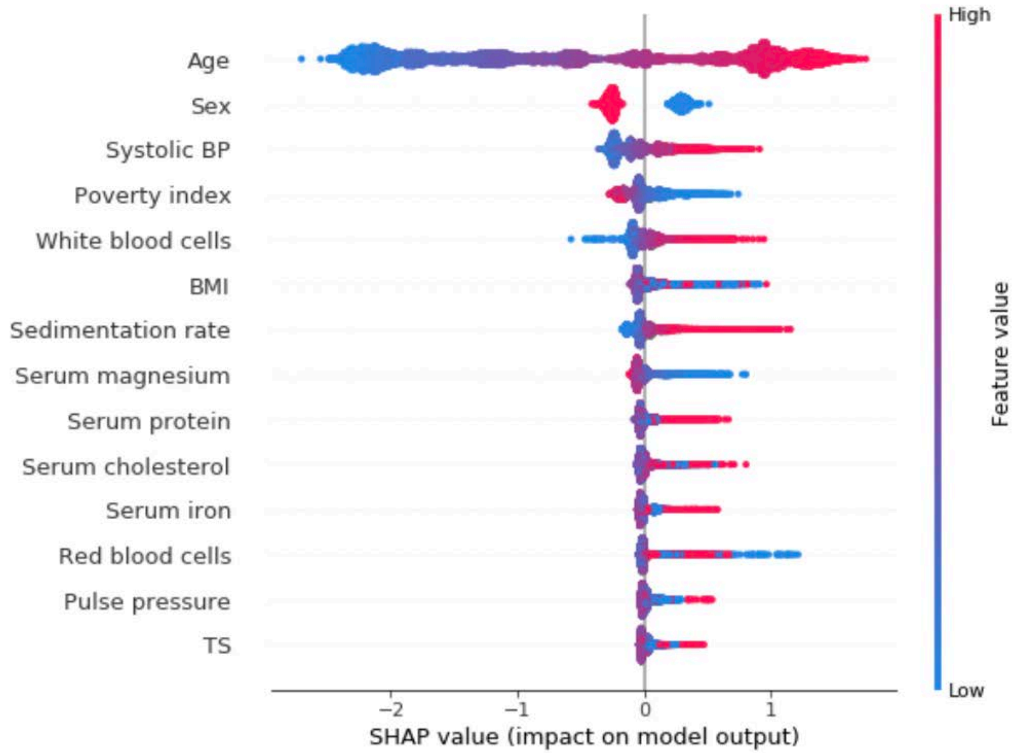
SHAP



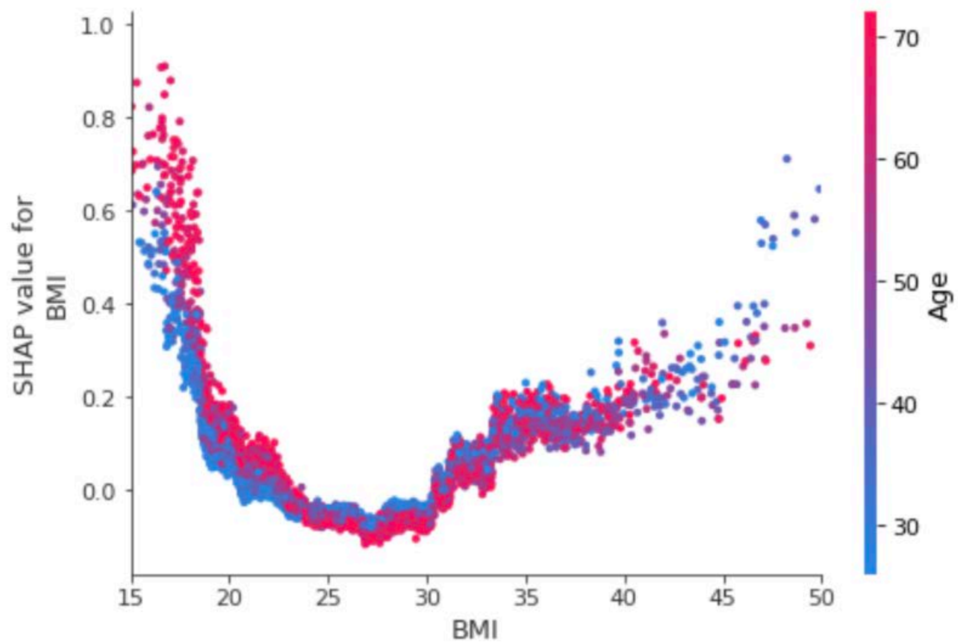
$$\hat{y}_i = shap_0 + shap(X_{1i}) + shap(X_{2i}) + \ldots + shap(X_{pi})$$

$$0.4 = 0.1 + (0.1) + (0.1) + (-0.3) + (0.4)$$

Prediction | Base rate | BMI | BP | Sex | Age



Feature Importance



Partial Dependence Plot

Today

- What is eXplainable AI (XAI) and why it matters?
- A case study of modelling ridesharing preferences using XAI and big data.

Ride-hailing market

- Uber and Lyft (and others) are common in cities.
- Jupiter Research (Dec 2021):
  - Consumer spending on ride-hailing will approach US $937 billion by 2026.
  - = 50 times the total annual revenue of Transport for London, New York City's MTA, and Beijing Metro in 2021.

ARTICLES
https://doi.org/10.1038/s41893-020-00678-z
nature sustainability

Check for updates

Impacts of transportation network companies on urban mobility

Mi Diao[1], Hui Kong[2,3] and Jinhua Zhao[2]

**The role of transportation network companies (TNCs) in the urban transport system is under intense debate. In this study, we systematically assess three aspects of the net impacts of TNCs on urban mobility in the United States—road congestion, transit ridership and private vehicle ownership—and examine how these impacts have evolved over time. Based on a set of fixed-effect panel models estimated using metropolitan statistical area level data, we fi
congestion in terms of both intensity (by 0.9%) and duration (by 4.5%), an
change in vehicle ownership. Despite the ideal of providing a sustainable mo
our analysis suggests that TNCs have intensified urban transport challenge**

The hidden cost of ride-hailing: is Uber bad for the environment?

FEATURE | 18 March 2021

Research shows that Uber's carbon footprint depends largely on the characteristics of individual cities.

Why Uber is bad for cities

DAVID THORPE 26 NOVEMBER 2019

Los Angeles Times

BUSINESS

Column: Uber and Lyft increase traffic and pollution. Why do cities let it happen?

SUBSCRIBERS ARE READING

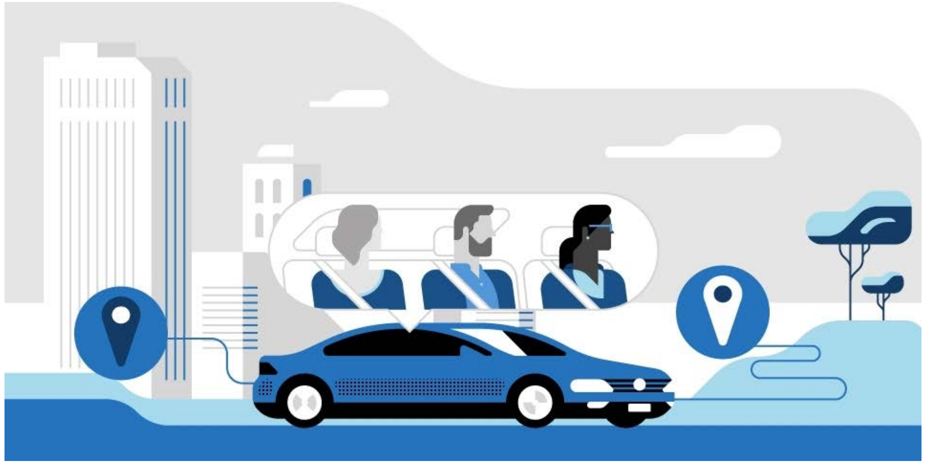
USC SPORTS
FOR SUBSCRIBERS
Inside USC baseball's long road back to prominence amid investigations and discontent

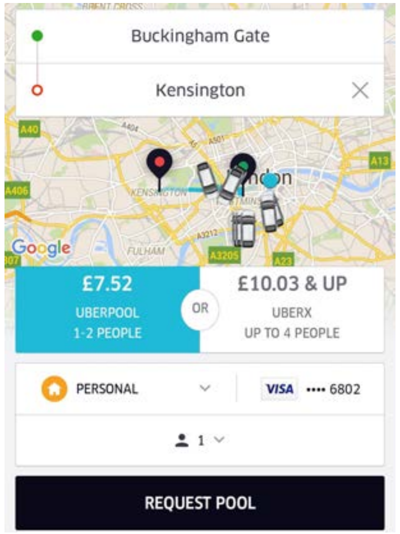
CALIFORNIA
FOR SUBSCRIBERS
He thought it was a date. Instead, he walked into a deadly MS-13 trap

TRAVEL
FOR SUBSCRIBERS

Ride-hailing services (solo vs. shared)

- Solo rides: Single-occupancy. e.g. UberX and Lyft
- Shared rides: shared carpool style. e.g. UberPool and Lyft Shared
- Shared ride benefits: shared rides can reduce traffic congestion, cut per-passenger carbon emissions, reduce parking infrastructure, and provide a more cost-effective way to travel (Shaheen and Cohen, 2019)
- Limited availability/popularity: Available only in selected cities, and occupies 15–25% of total trips in cities such as London, Hangzhou, Toronto, and Chicago.





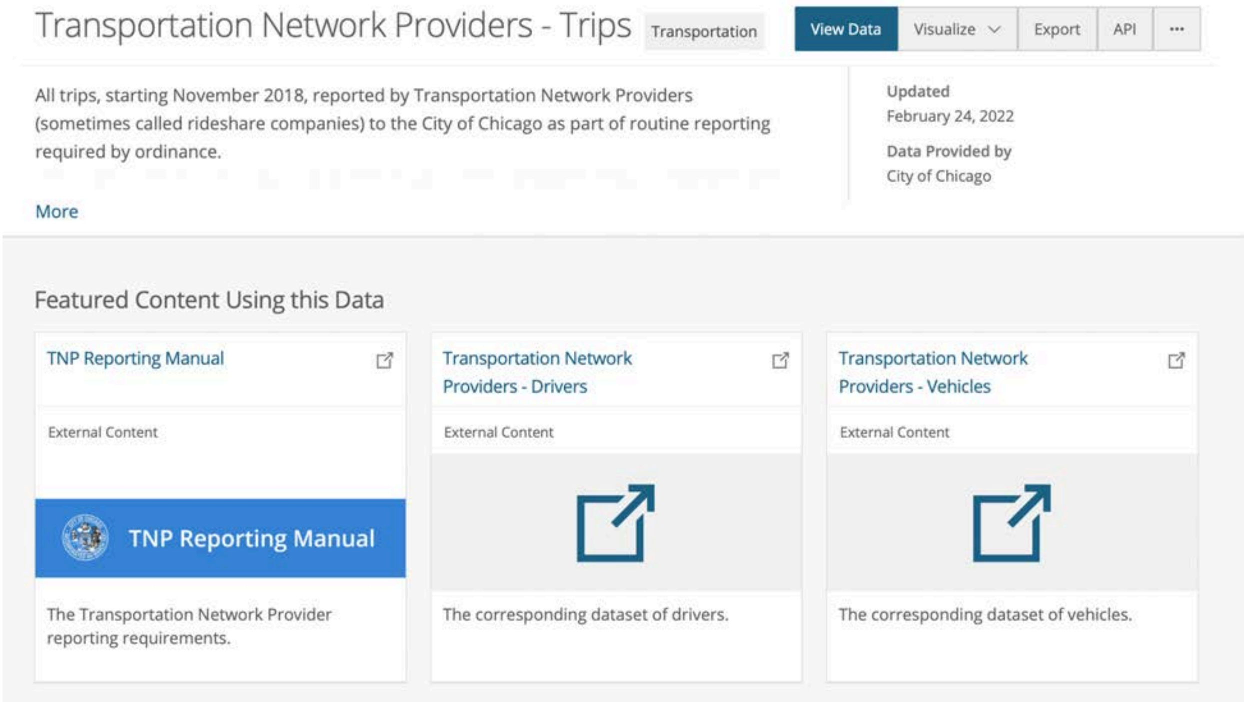
Understanding willingness to share

- To promote the proportion of shared trips, we need to understand the factors influencing people's willingness/reluctance to share.

?

Big trip data

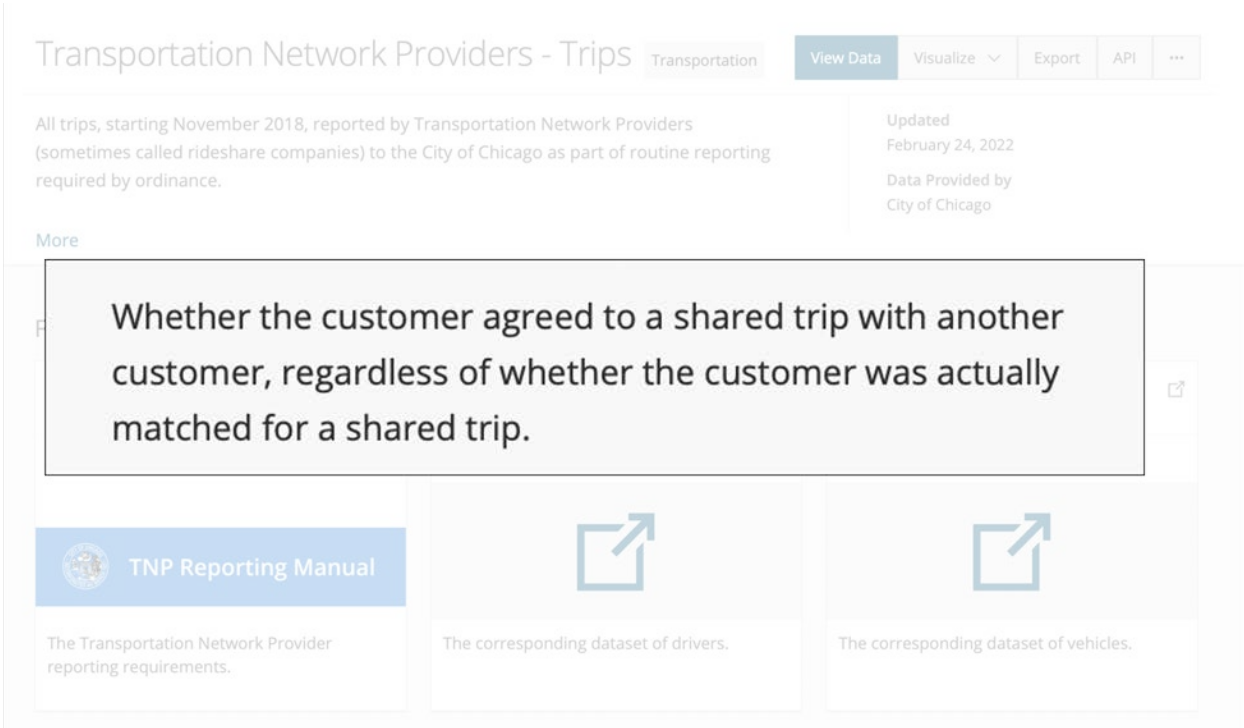
- City: Chicago, USA
- Trip records are time-stamped.
- O-D are geo-referenced at the census tract level.
- Label: shared trip? 1 : 0



https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

Big trip data

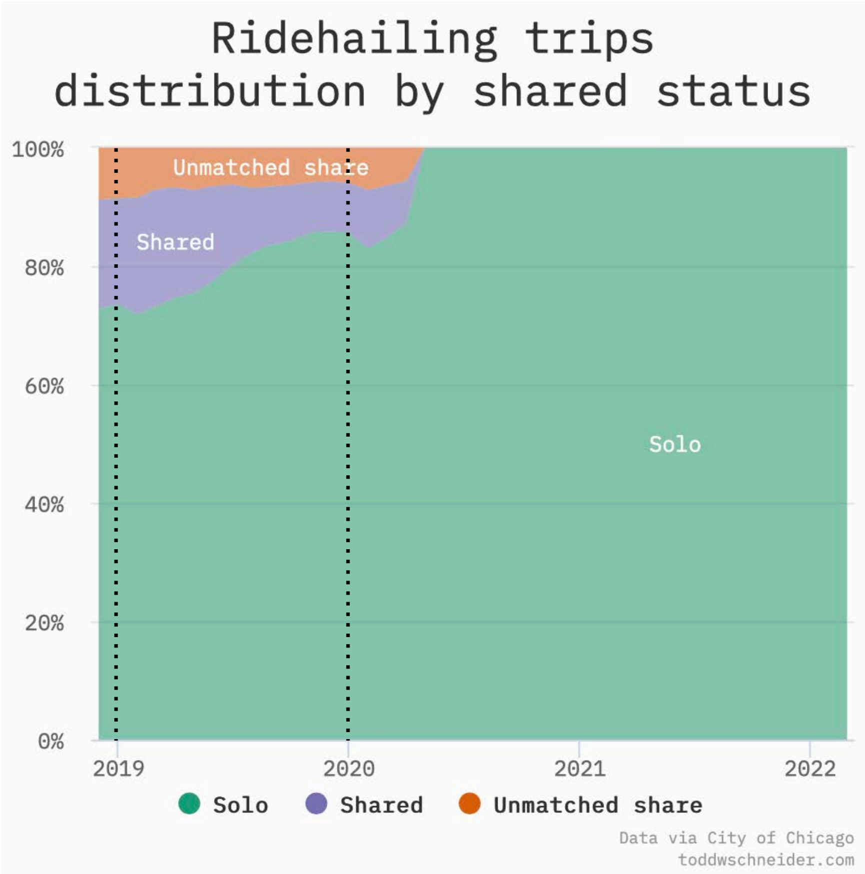
- City: Chicago, USA
- Trip records are time-stamped.
- O-D are geo-referenced at the census tract level.
- Label: shared trip? 1 : 0



https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

Big trip data

- City: Chicago, USA
- Trip records are time-stamped.
- O-D are geo-referenced at the census tract level.
- Label: shared trip? 1 : 0
- Date: whole year of 2019
  - Shared rides were suspended due to COVID.
- Total records: >10M



Modelling willingness to share

Upfront information

Trip base fare
Additional fees

Distances:
- Trip distance (straight)
- Distance to downtown

Built-environment:
- Walkability
- Distance to public transportation

Socio-economics
- Population density
- % no car
- Education (% College)
- Income (median household)
- Race/Ethnicity (% non white)
- Age (% 19-29)

Hour of the day
Week/weekend
Holiday
Weather:
- Rain
- Temperature
- Wind

No post-trip information
❌ Actual trip distance
❌ Actual trip duration
❌ Total cost (incl. tip, late fees, etc.)

XGBoost

- XGBoost is a gradient boosting method that uses a gradient descent optimisation algorithm to sequentially ensemble decision trees to minimise model error (Chen and Guestrin, 2016).

- XGBoost typically outperforms deep learning, random forests and other alternatives when handling tabular data (Grinsztajn et al., 2022; Shwartz-Ziv & Armon, 2022).

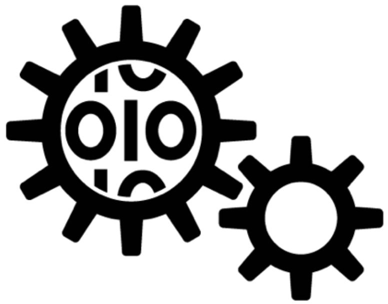
- XGBoost is highly scalable.

Workflow

Data

Machine Learning

Interpretation



Chicago

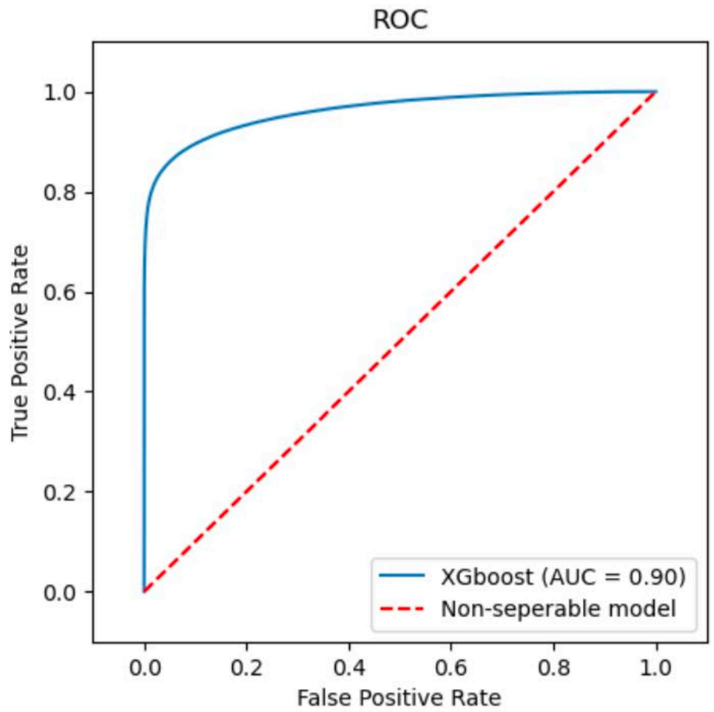
XGBoost

SHAP

Model accuracy

Table 2. Training and testing accuracy of the XGBoost model.

	Training	Testing
True Positive	86.8%	86.0%
True Negative	95.8%	95.0%
False Positive	13.2%	14.0%
False Negative	4.2%	5.1%
F1	0.91	0.90
Overall Accuracy	91.3%	90.5%



Feature group feature importance

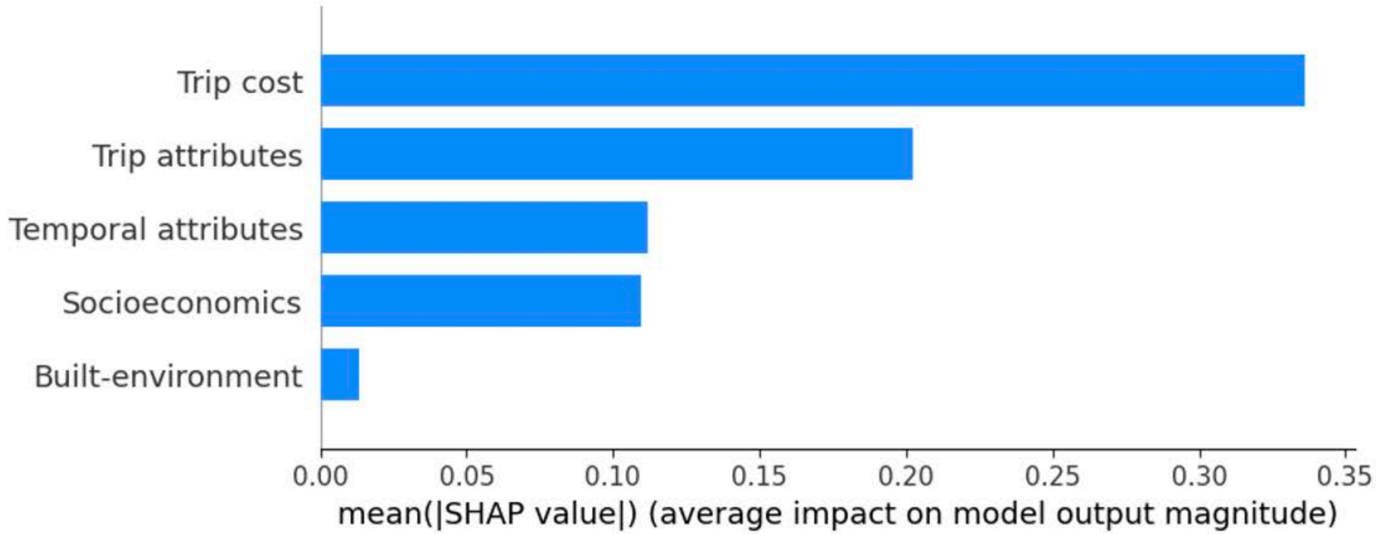
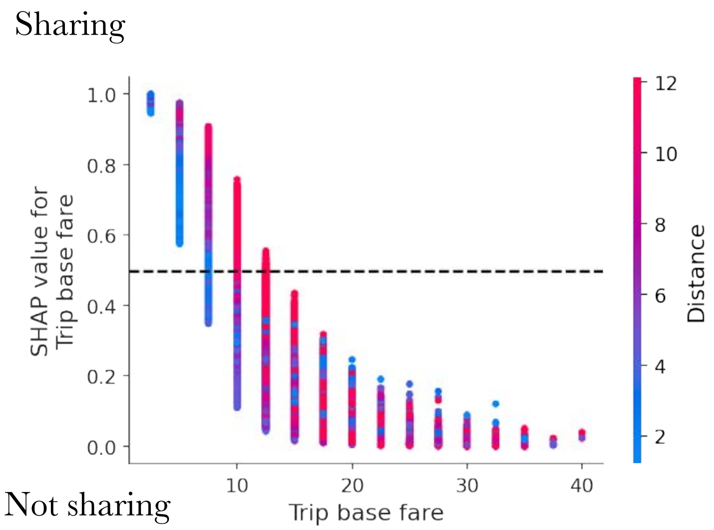
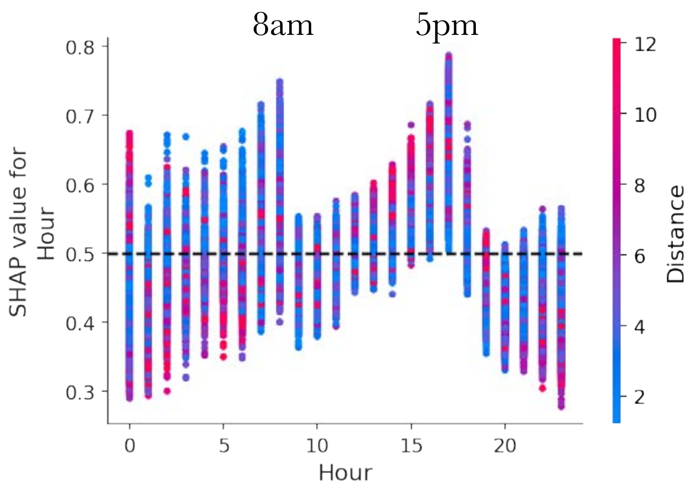


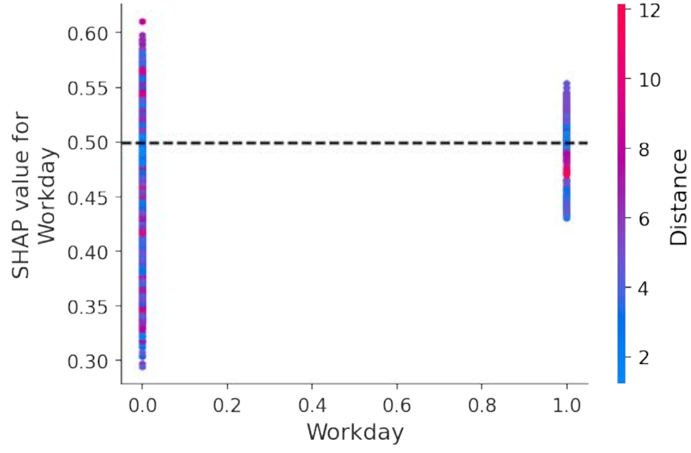
Figure 3 SHAP-based global feature importance ranking for five groups of features.

SHAP values have been converted to probability.

SHAP-based partial dependence plots





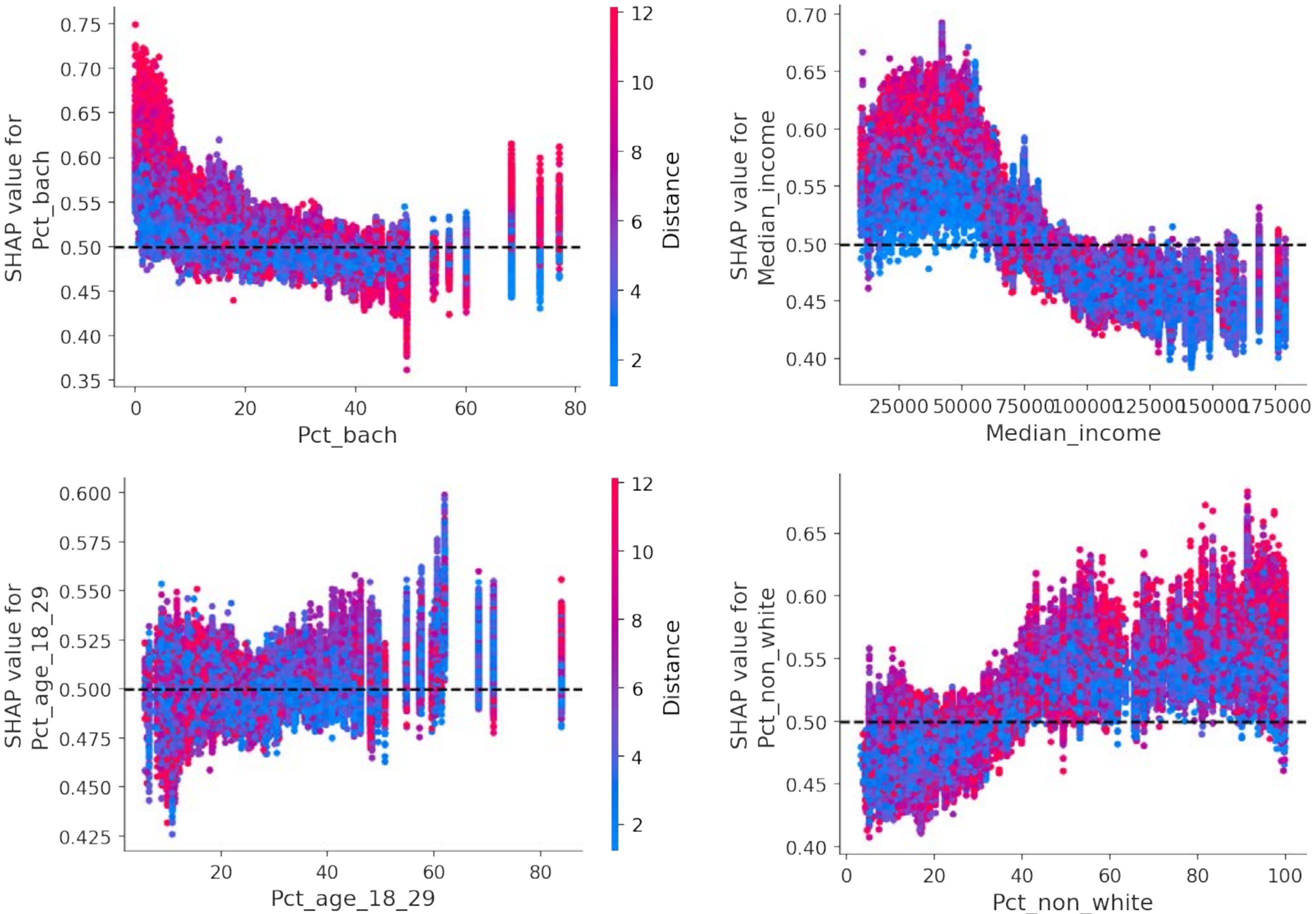


- People favour shared rides for longer distance trips while holding everything else constant.

- People tend to favour shared rides during morning and afternoon commute hours.
- Solo rides are preferred in the evening.

- Willingness has a higher variance on a non-working day, with a preference to take more solo rides.

SHAP-based partial dependence plots



- Shared rides are preferred in **low-income, low-education, non-white, young, high vehicle ownership** census tracts.

Trip-level explanation analysis

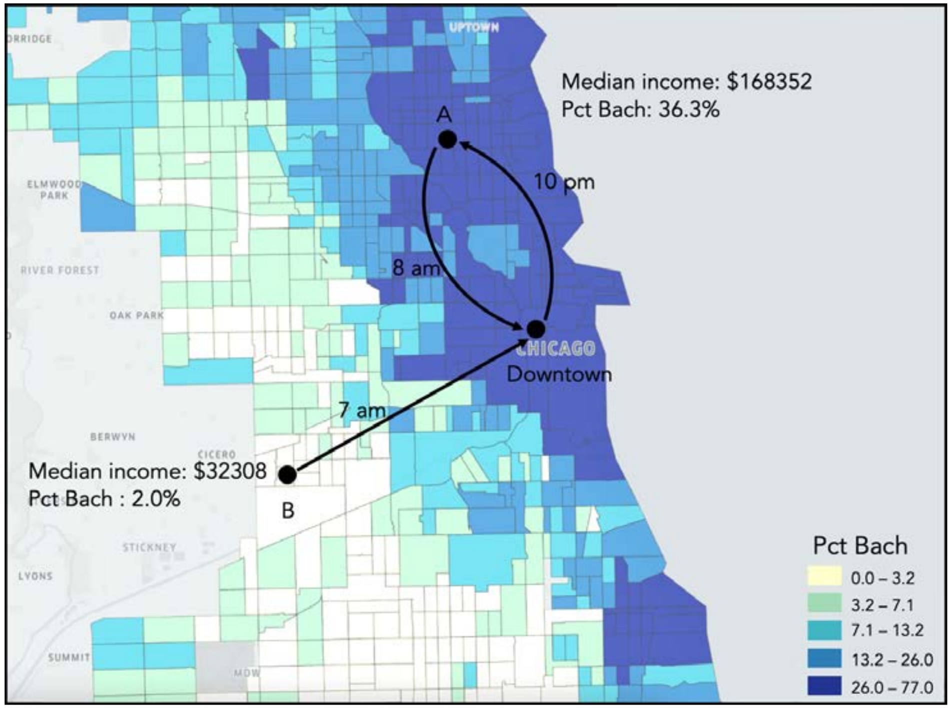


Table 3. Feature values and their SHAP contributes to three example trips in Figure 8.

Neighborhood	Trip	Shared	Time	SHAP (Hour)	Median Income	SHAP (Median Income)	Pct Bach	SHAP (Pct Bach)
A	A to Downtown	Yes	8am	+22.4%	168,352	- 0.2%	36.3%	+ 0.4%
	Downtown to A	No	10pm	-15.5%		- 2.6%		- 1.7%
B	B to Downtown	Yes	7am	+16.2%	32,308	+ 13.0%	2.0%	+17.4%

Conclusions

- Decision to rideshare is largely driven by economic considerations.
- User tend to prefer ridesharing during am/pm rush hours.
- Socio-economic disparities.

- How will COVID impact the ridesharing market?
- How does ridesharing compete with public transport?

Some thoughts on XAI

- We should adopt a simpler and more interpretable model if it has similar performance.

- If AI model is used, some degree of XAI should be presented to provide credibility of the model.

- XAI can be a good alternative to statistical approaches especially when data is large.

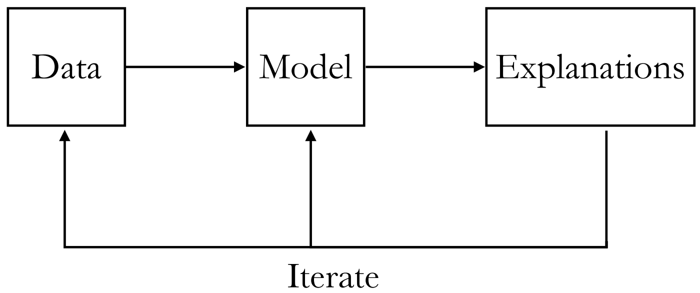
- Explanations are skewed if either data or model is biased.

Some further thoughts on XAI

- How do we evaluate explanation accuracy/faithfulness?
  - Ground truth validations.
  - Simulating simple data generating processes to validate model and explanations.

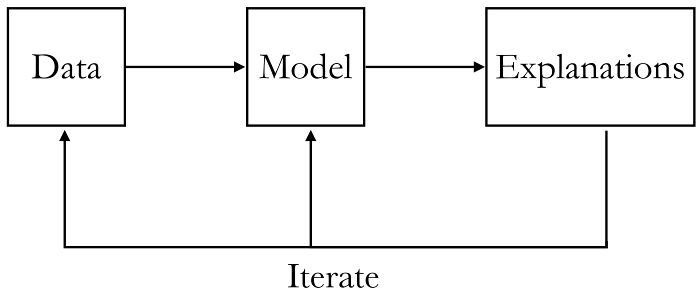
Some further thoughts on XAI

- How do we evaluate explanation accuracy/faithfulness?
    - Ground truth validations.
    - Simulating simple data generating processes to validate model and explanations.
- How can XAI inform us with the development of (Geo)AI?
    - XAI provides a great way to diagnose and improve model.



Some further thoughts on XAI

- How do we evaluate explanation accuracy/faithfulness?
    - Ground truth validations.
    - Simulating simple data generating processes to validate model and explanations.
- How can XAI inform us with the development of (Geo)AI?
    - XAI provides a great way to diagnose and improve model.
- Can XAI give us spatial explanations?



Some thoughts on XAI

ELSEVIER

Computers, Environment and Urban Systems

Volume 96, September 2022, 101845

Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost

Ziqi Li

Show more

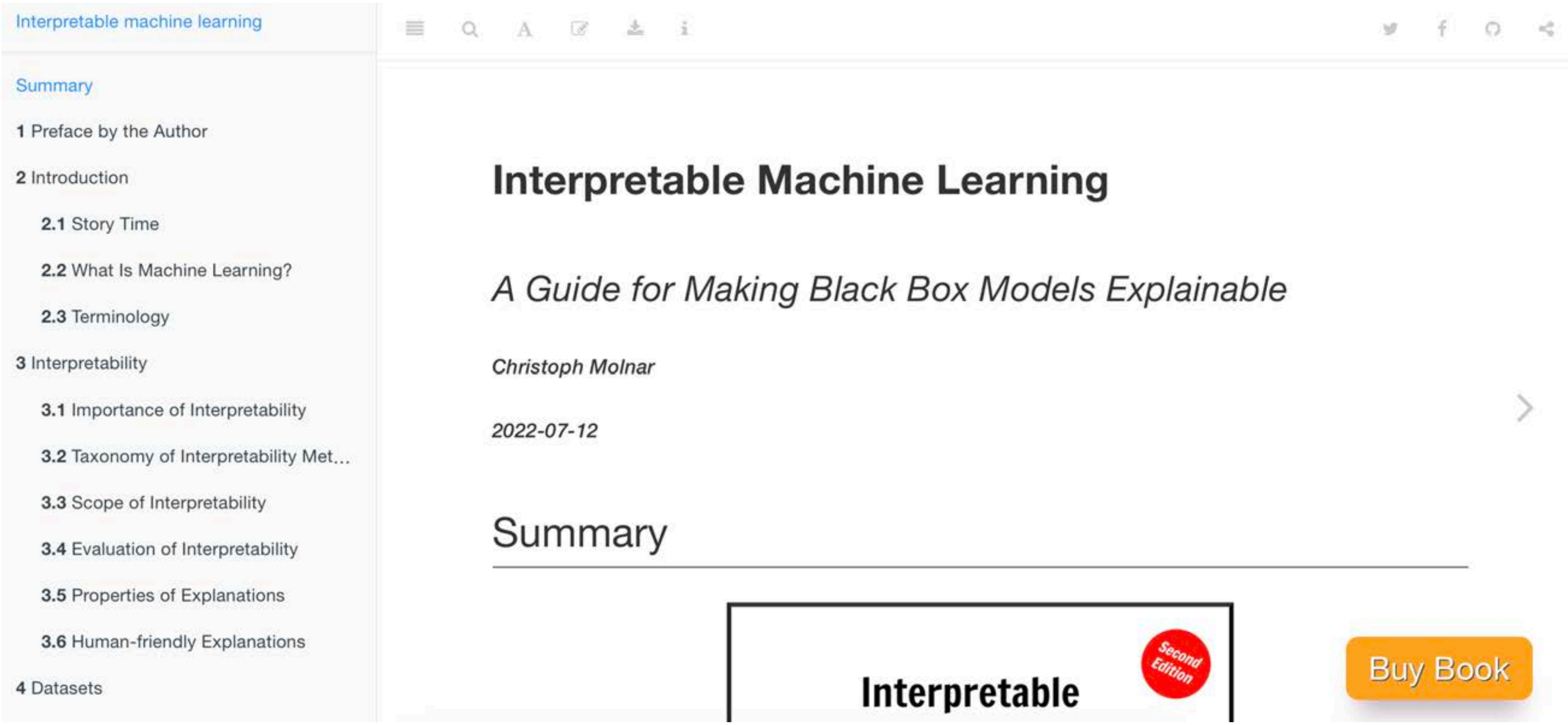
Share Cite

https://doi.org/10.1016/j.compenvurbsys.2022.101845 Get rights and content

 Open access

- Benchmark with spatial models
- Validate explanations
- Discuss the use of XAI with spatial data

XAI Book



https://christophm.github.io/interpretable-ml-book/

Thank you!

- Email: Ziqi.Li@glasgow.ac.uk
- Twitter: @geo_ziqi