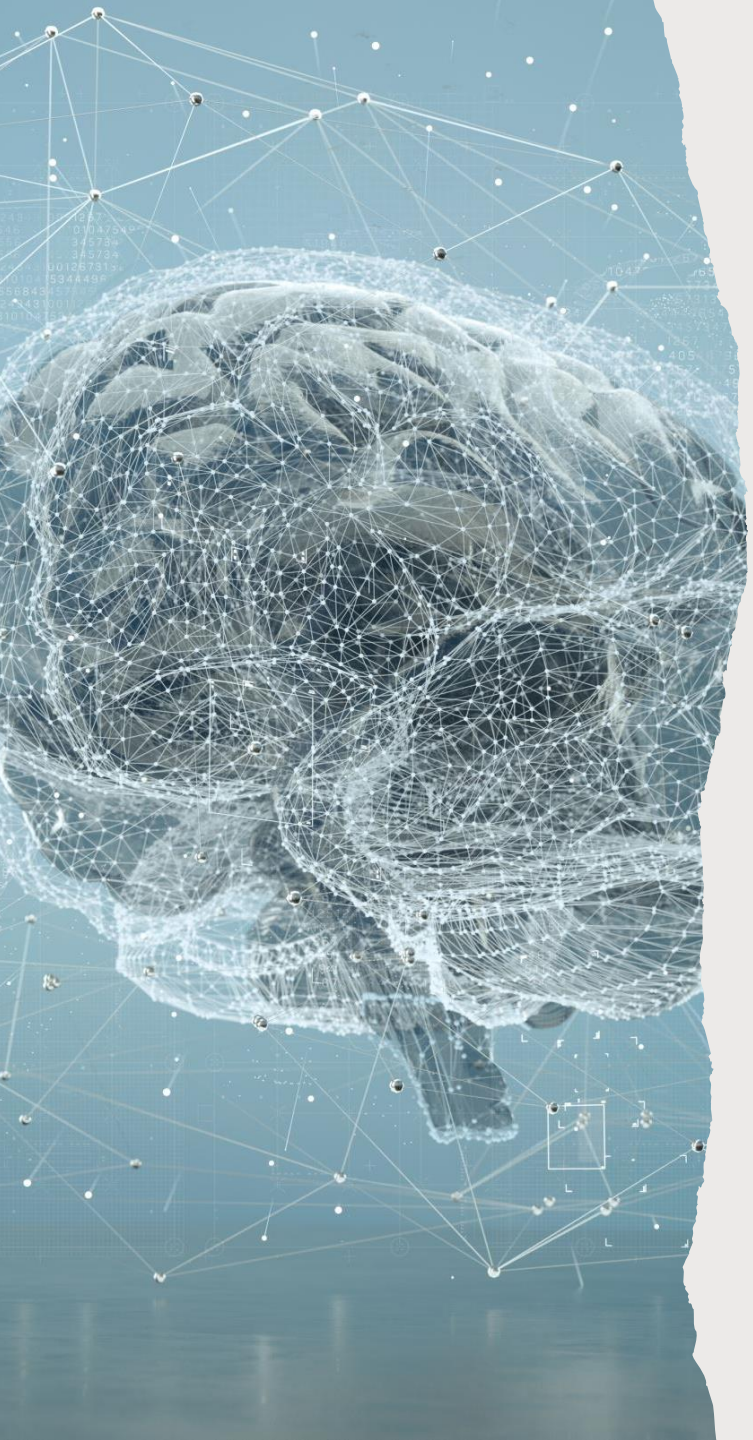


Artificial Intelligence in Regional Science

dr Maria Kubara
prof. Katarzyna Kopczewska

Faculty of Economic Sciences,

Chair of Data Science,

University of Warsaw

OpenAI
https://openai.com › index › chatgpt

Introducing ChatGPT

30 lis 2022 — **ChatGPT** is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.

Download · Introducing GPT-4o and more... · OpenAI announces new... · Research

Time Magazine
https://time.com › ... › Technology

How the AI Revolution Will Reshape the World

1 wrz 2023 — We are now facing a new wave of technology, centered around **AI** but including synthetic biology, quantum computing, and abundant new sources of ...

CMSWire.com
https://www.cmswire.com › gene...

ChatGPT's Traffic Surges by 115.9% Year-Over-Year

7 lis 2024 — **ChatGPT** has seen a 115.9% year-over-year increase in traffic, reaching 3.7 billion month visits globally, with a 17.2% month-over-month growth ...

Transform Business with AI

AI for Smarter Interactions — Discover scalable **AI** tools to simplify processes and unlock your **business** potential. Redefine workflows with DRUID's advanced conversational **AI** for smarter processes. Lower costs.

IBM
https://www.ibm.com › topics › a...

AI for Digital Transformation

AI transformation is a strategic initiative whereby a business adopts and integrates artificial intelligence (AI) into its operations, products and services ...

Growing popularity of AI – in daily life, in business and in general research

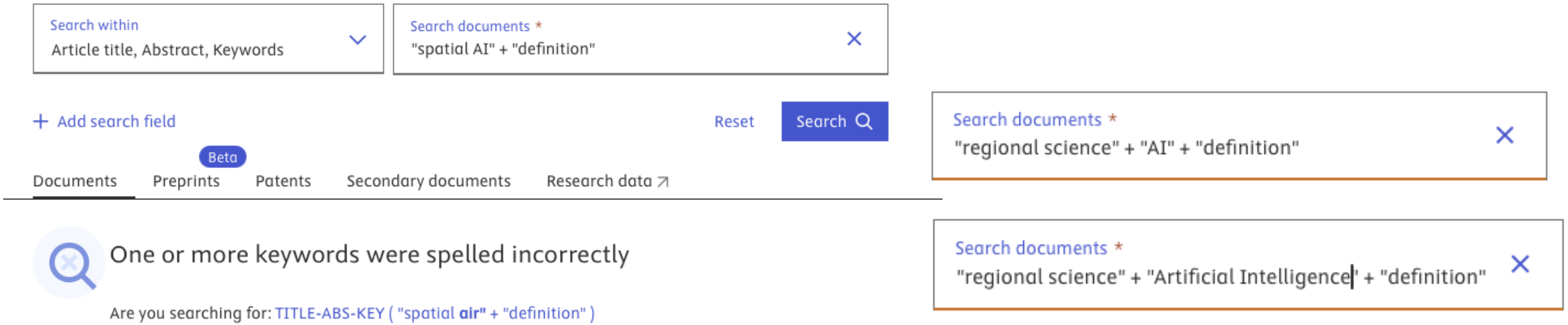
[HTML] Pharma's Bio-**AI revolution**
I Bentwich - Drug Discovery Today, 2023 - Elsevier
… Pharma is undergoing a 'Bio-**AI**' **revolution**. It is a **revolution** that has the potential to drastically cut the time and cost of developing drugs. Its prerequisite technologies are already here, …
☆ Zapisz 99 Cytuj Cytowane przez 13 Powiązane artykuły Wszystkie wersje 5

[KSIĄŻKA] The **AI revolution** in medicine: GPT-4 and beyond
P Lee, C Goldberg, I Kohane - 2023 - books.google.com
… This book represents the sort of effort that every sphere affected by **AI** will need to … if **AI** can be used to raise the bar for human health worldwide. This is a tremendously exciting time in **AI** …
☆ Zapisz 99 Cytuj Cytowane przez 142 Powiązane artykuły

The forthcoming Artificial Intelligence (**AI**) **revolution**: Its impact on society and firms
S Makridakis - Futures, 2017 - Elsevier
… the new ones being brought by the **AI revolution**. It must be emphasized that the stakes of correctly predicting the impact of the **AI revolution** are far reaching as intelligent machines may …
☆ Zapisz 99 Cytuj Cytowane przez 2117 Powiązane artykuły Wszystkie wersje 9

The **ai revolution**: opportunities and challenges for the finance sector
C Maple, L Szpruch, G Epiphaniou, K Staykova… - arXiv preprint arXiv …, 2023 - arxiv.org
… **AI** and the integration of **AI** with blockchain and Decentralised Finance (DeFi). It calls for further research into how **AI** … to lead the development of Explainable **AI**(XAI) and interpretable **AI**. …
☆ Zapisz 99 Cytuj Cytowane przez 30 Powiązane artykuły Wszystkie wersje 4 »

Reboot for the **AI revolution**
YN Harari - Nature, 2017 - nature.com
… The automation **revolution** is emerging from the confluence of two scientific tidal waves. Computer scientists are developing artificial intelligence (… The **AI revolution** won't be a single …
☆ Zapisz 99 Cytuj Cytowane przez 159 Powiązane artykuły Wszystkie wersje 8



No definition in the literature – what is AI in regional science context?

Goal: Define AI and ML in regional science and demonstrate AI's utility.

What is AI?

Chapter 1 - Clinical decision support systems: Benefits, potential challenges, and applications in pneumothorax segmentation

Sudha Subramaniam [a], K.B. Jayanthi [b]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/B978-0-323-99031-8.00009-0 ↗ Get rights and content ↗

Artificial intelligence (AI) is a branch of science that simulates intelligent behavior into computers to solve complex problems better than humans;

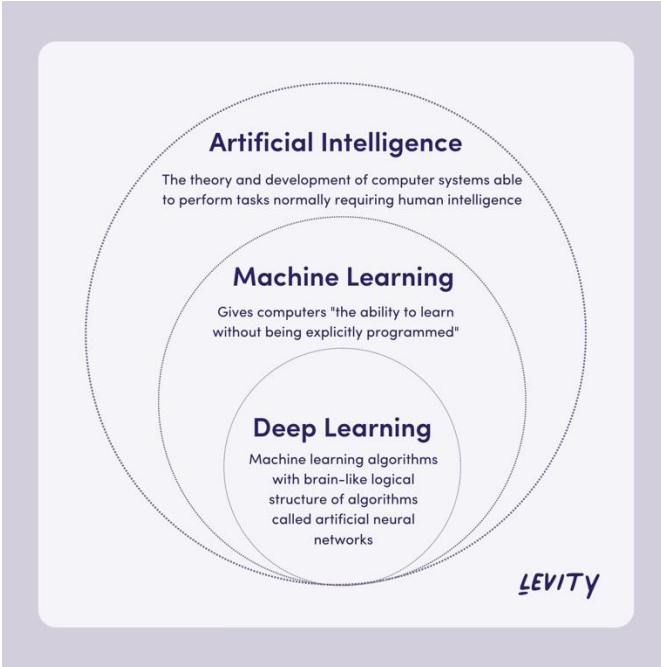
https://doi.org/10.1016/B978-0-323-99031-8.00009-0

Artificial intelligence in business is **the use of AI tools such as machine learning, natural language processing, and computer vision to optimize business functions, boost employee productivity, and drive business value.** 20 lut 2024

IBM
https://www.ibm.com › topics › artificial-intelligence-b...

What is Artificial Intelligence (AI) in Business? - IBM

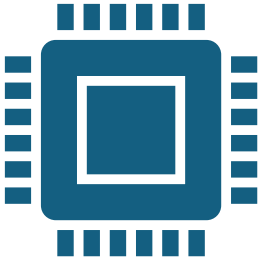
Informacje o fragmentach z odpowiedzią • Prześlij opinię



https://levity.ai/blog/difference-machine-learning-deep-learning

AI and ML are often used interchangeably but **differ significantly.**
Many definitions **lack clarity** or **fail to distinguish between tools and goals.**

What is AI? Clearing the Confusion



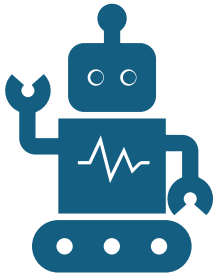
Machine Learning (ML):

Focus: Data analysis, pattern recognition, and prediction.

Methods: Algorithms like decision trees, SVMs, neural networks.

Goal: Explore and learn patterns from data.

Example: Predicting house prices or classifying images.



Artificial Intelligence (AI):

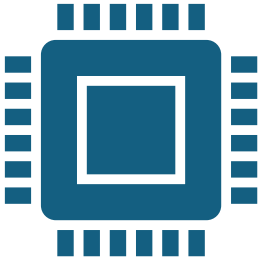
Focus: Enabling machines to make informed decisions or recommendations.

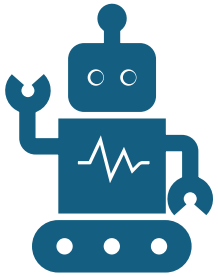
Broader scope: Combines ML, reasoning, optimization, and explainability.

Goal: Assist or automate decision-making processes.

Example: Recommending a business location based on multiple data layers.

What is AI? Clearing the Confusion





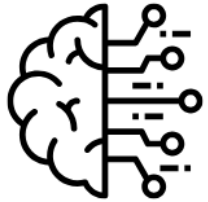
Machine Learning (ML): **Artificial Intelligence (AI):**

AI is for decision-making;
ML is for analysis and exploration.

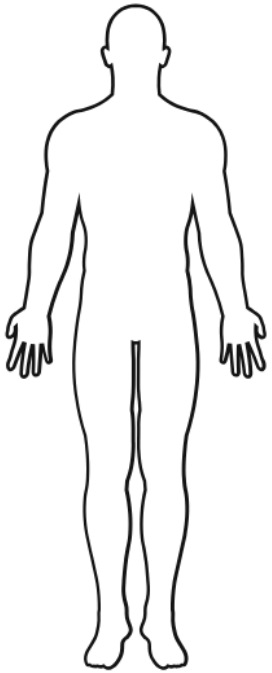
AI leverages ML insights but **goes further** by integrating rules, reasoning, and context to **support actionable decisions.**

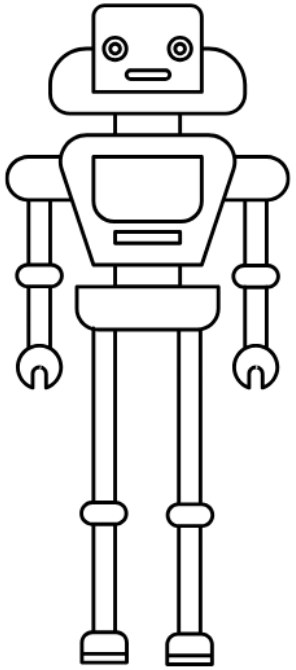
Brain

Machine Learning models



Human Intelligence





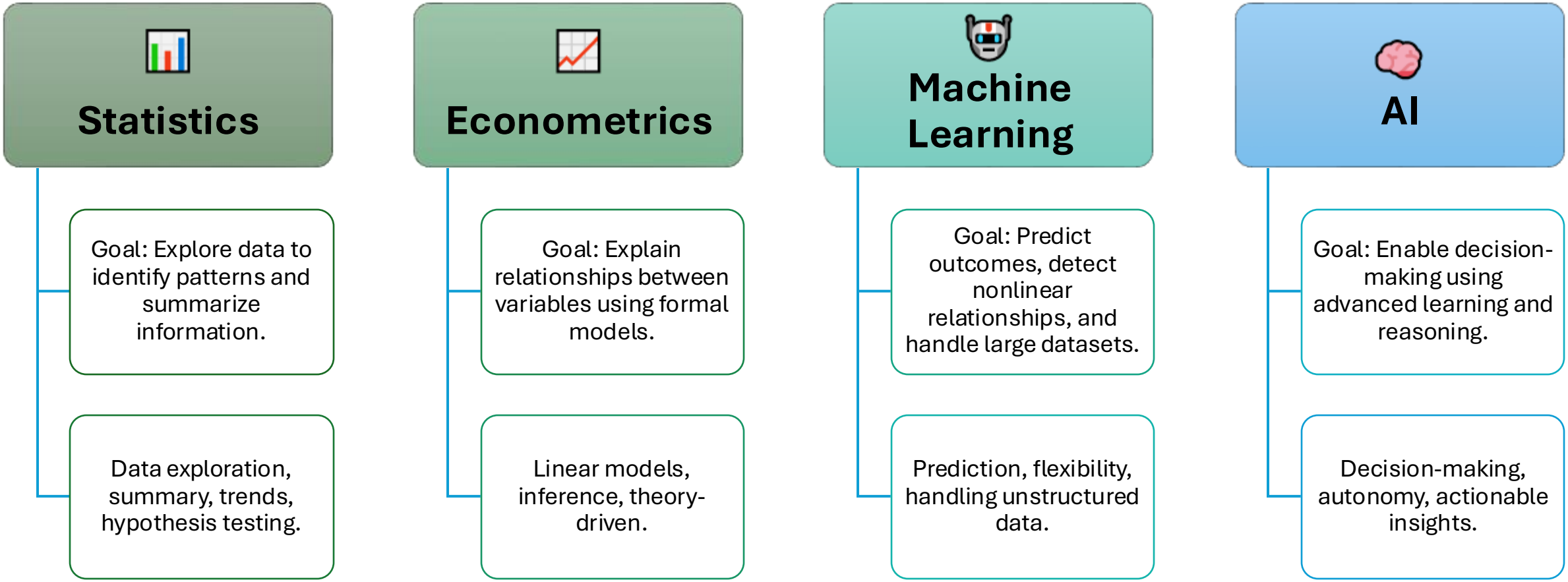
Artificial Intelligence

General goal of intelligence:

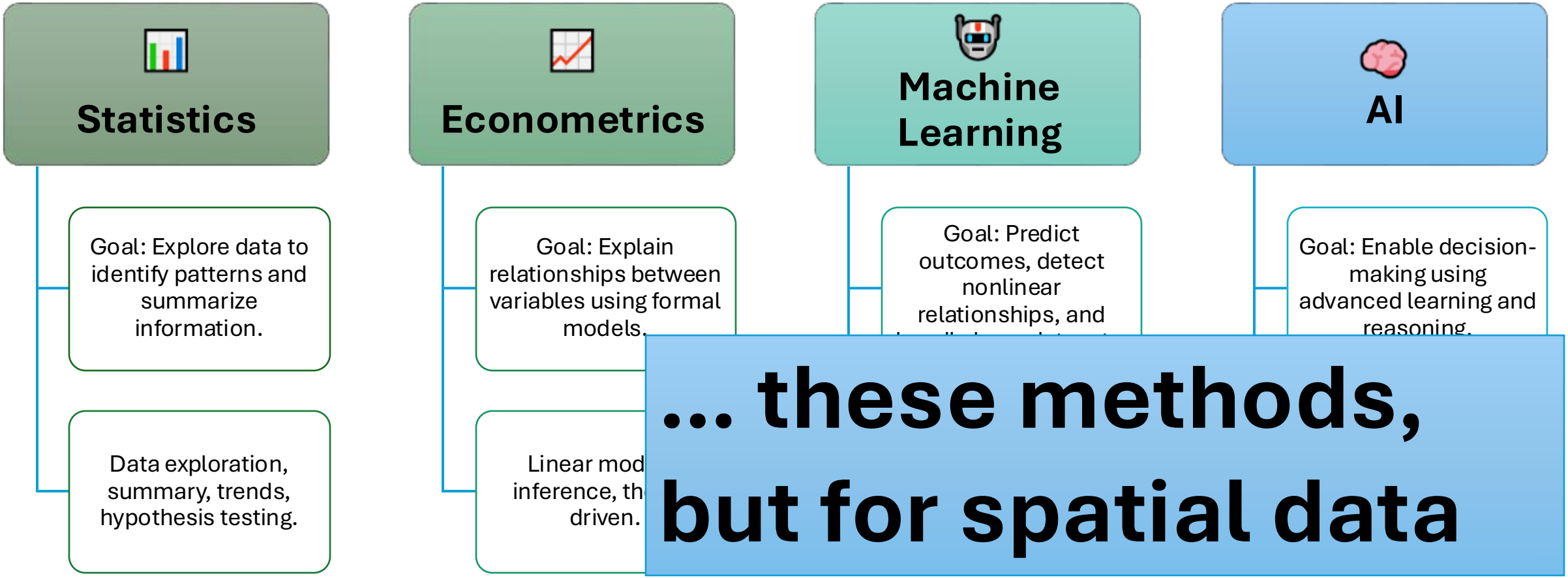
Learn from experience and draw insights

To make decisions and take actions

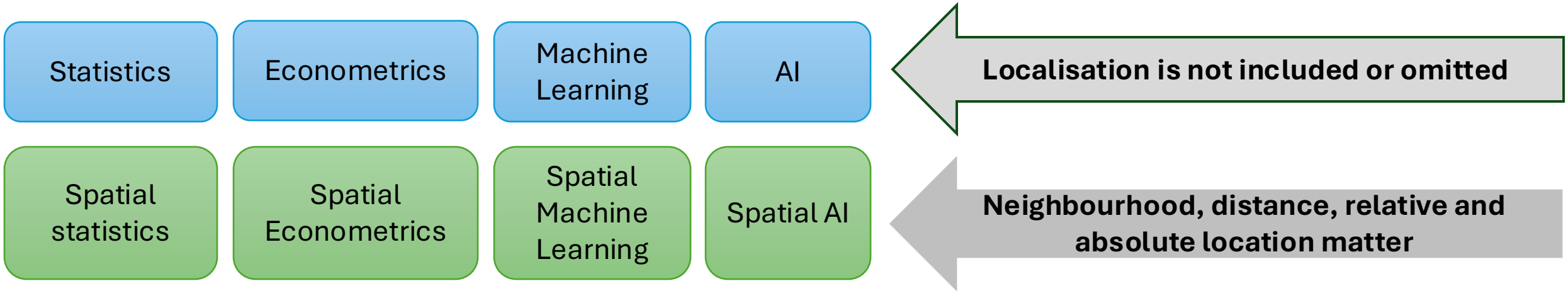
Towards AI: Where Are We Now?



Towards AI: Where Are We Now?



Add spatial component to the analysis



Why Spatial Data is Different:

- **Spatial heterogeneity**: Variations across locations; relationships aren't uniform.
- **Proximity-based relationships**: Observations are often interdependent (spatial autocorrelation).

Importance for Regional Science:

- Decisions are inherently spatial (e.g., business locations, infrastructure planning).
- AI must understand space to provide actionable insights.

Our proposition: FAIRS

FAIRS - a novel framework for AI in regional science that is:

- Fit-for-purpose,
- Actionable,
- Interpretable,
- Replicable
- Spatially-aware

FAIRS

Fit-for-purpose
Actionable
Interpretable
Replicable
Spatially-aware

New framework
for AI in regional science

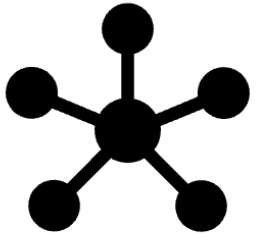
Use spatial AI for Regional Science

FAIRS - Framework for AI in Regional Science

Define problem to solve

Define sources of data that will leverage spatially-informed decisions

Define if decision is taken once or if it is a continous process based on new data

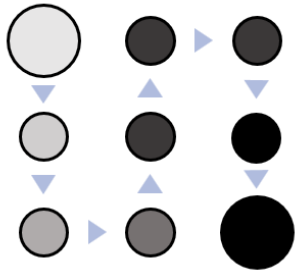


Spatial data processing

Integrating spatial data and handling spatial issues

Spatial Machine Learning

Train models that include information about location and neighbourhood



AI-supported decision

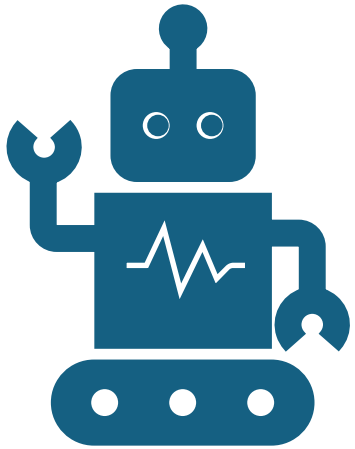
Make spatially-informed decisions that address place specificity

Special data preparation for spatial data

Spatial Machine Learning

Spatial AI for Regional Science

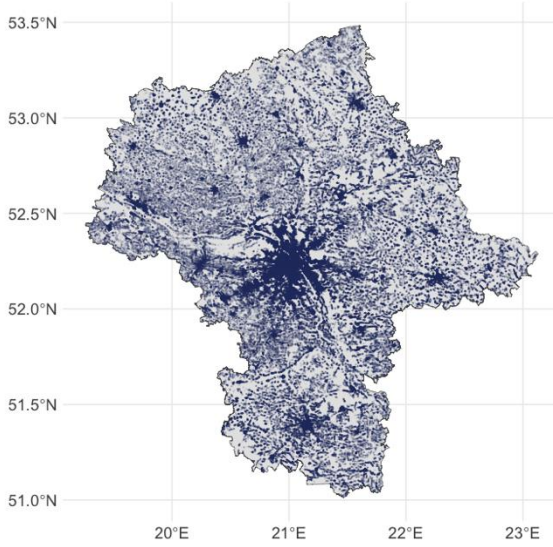
Empirical Study: Optimal Business Location



How to use spatial AI for Regional Science – a practical example

Empirical study

Actionable business insight:
For a specific location – what will be the best type of business?



Objective:

- Use AI to recommend optimal business locations in Poland.
- Focus: Identify which types of businesses are best suited to specific locations based on spatial and economic data on already existing patterns.

Data Overview:

- **Study area:** Whole population of firms from Mazowieckie, Poland (all companies operating there in 2012 – info from REGON database).
- **Spatial Data**: Proximity to competitors. Distances to different city types (core, midsize, regional, local).
- **Economic Data**: Business sectors (agriculture, service, construction, production). Local business aggregation by sector.
- **Other Variables**: Neighborhood indicators (e.g., population density, firm density). Demographic factors (e.g., local population characteristics).

scientific reports

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific reports › articles › article

Article | Open access | Published: 27 September 2024

Population density as the attractor of business to the place

Katarzyna Kopczewska ✉, Maria Kubara & Mateusz Kopyt

Scientific Reports **14**, Article number: 22234 (2024) | Cite this article

1208 Accesses | **8** Altmetric | Metrics

Study Design:

- Use ML classification models for the analysis.
- Incorporate spatial features into AI-driven recommendations and adjust the modelling procedure for spatial data.
- Steps:
  - **Analysis phase**: Classification of locations based on the already existing information on the types of businesses successfully operating there.
  - **Decision phase**: Use AI outputs for actionable location-based recommendations.

Data Description

agri	constr	prod	serv
253514	71579	63989	594637
25.77%	7.28%	6.50%	60.45%

Number of observations: 983719
Number of variables: 31

Target Variable

class4: Categorical variable representing the business sector:

Agriculture, Service, Construction, Production

Spatial Variables

- lon, lat: Longitude and latitude of the business locations, representing their spatial coordinates.

Distance Metrics:

- Distances to various city types at three thresholds (10 km, 25 km, 50 km):
- **Core Cities**: dist_core_10, dist_core_25, dist_core_50
- **Midsize Cities**: dist_midsize_10, dist_midsize_25, dist_midsize_50
- **Regional Cities**: dist_regional_10, dist_regional_25, dist_regional_50
- **Local Big Cities**: dist_localbig_10, dist_localbig_25, dist_localbig_50
- **Local Small Cities**: dist_localsmall_10, dist_localsmall_25, dist_localsmall_50

These variables capture **proximity to urban centers** of varying sizes and their influence on location attractiveness.

Core Density Checks:

- COREfirms: Indicates if the location is within a high-density core of firms.
- COREpopul: Indicates if the location is within a high-density population area.

Economic Variables

Local Aggregation of Businesses by Sector:

- locAggAgri.s: Number of businesses in agriculture.
- locAggProd.s: Number of businesses in production.
- locAggConstr.s: Number of businesses in construction.
- locAggServ.s: Number of businesses in services.
- locAggTOTAL.s: Total number of businesses in all sectors.

Local Business Diversity:

- locHH.s **(Herfindahl Index)**: A measure of business concentration in the area. Higher values indicate less diversity (more dominance of certain sectors).

Local Industry Specialization:

- locLQ.s **(Location Quotient)**: Indicates the degree of specialization in a specific industry relative to a broader reference region.

Demographic Variables

Local Population Density:

- locPdens.s: Population density in the immediate vicinity.

Numeric variables are standardised to 0-1 range

Considered models: XGBoost, Random Forest, LightGBM, LiblineaR

Handling large and and high-dimensional datasets

Multiclass classification: Native support for predicting multiple business categories.

XGBoost:

- High-performance, scalable predictions with robust regularization.
- Exceptional accuracy and efficiency for large, complex datasets.

Random Forest:

- Reliable ensemble predictions by combining multiple decision trees.
- Robustness to overfitting and ability to handle high-dimensional data.

LightGBM:

- Fast training with efficient memory usage for large-scale data.
- Optimized for speed and scalability in multiclass classification.

LiblineaR:

- Efficient linear classification for high-dimensional sparse data.
- Simplicity and speed for quick evaluation of linear relationships.

Linear (LiblineaR) and Nonlinear (XGBoost, Random Forest, LightGBM).

AI for Regional Science – empirical study

Take the raw spatial data

Spatial Cross-Validation (evaluation framework)

Model Comparison (performance and stability)

Final Model Fitting

Decision-Support Application (interactive tool)

Special data preparation for spatial data

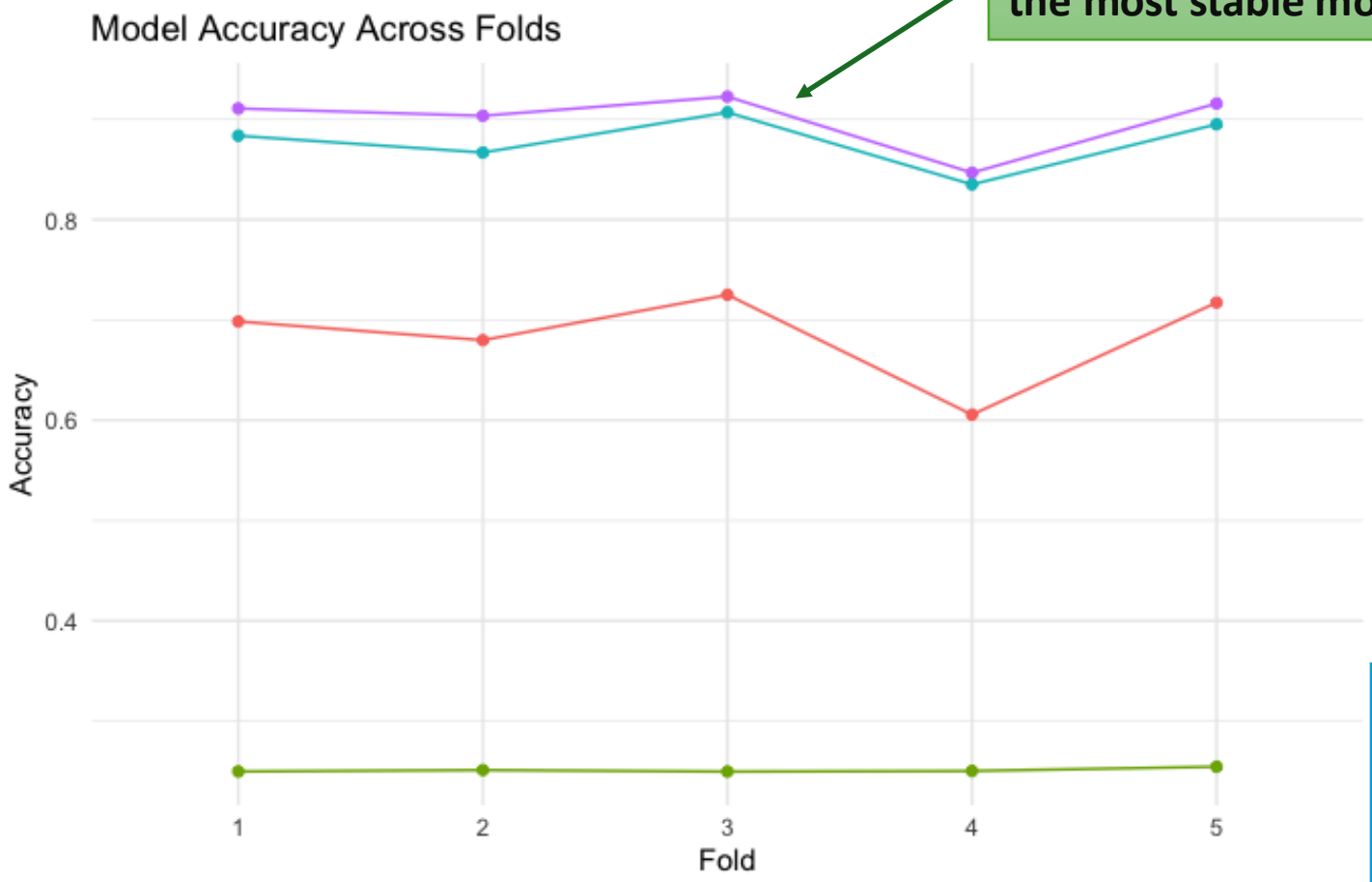
Spatial Machine Learning

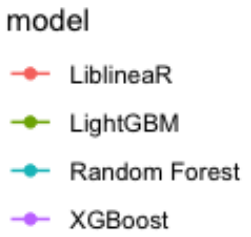
Spatial AI for Regional Science

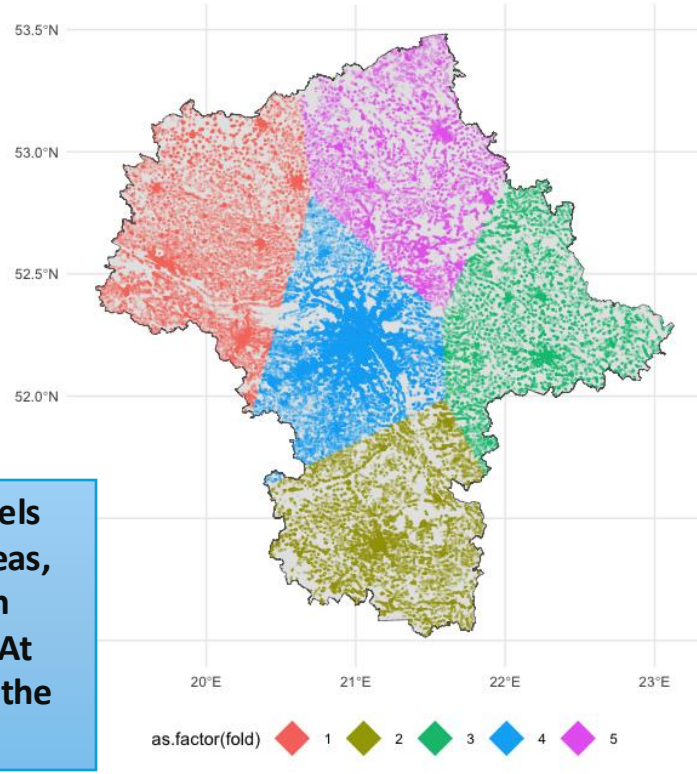
Model efficiency in spatial-cross validation

XGBoost – The best performing and the most stable model across folds

	model	mean_accuracy
1	LiblineaR	0.685
2	LightGBM	0.251
3	Random Forest	0.878
4	XGBoost	0.900







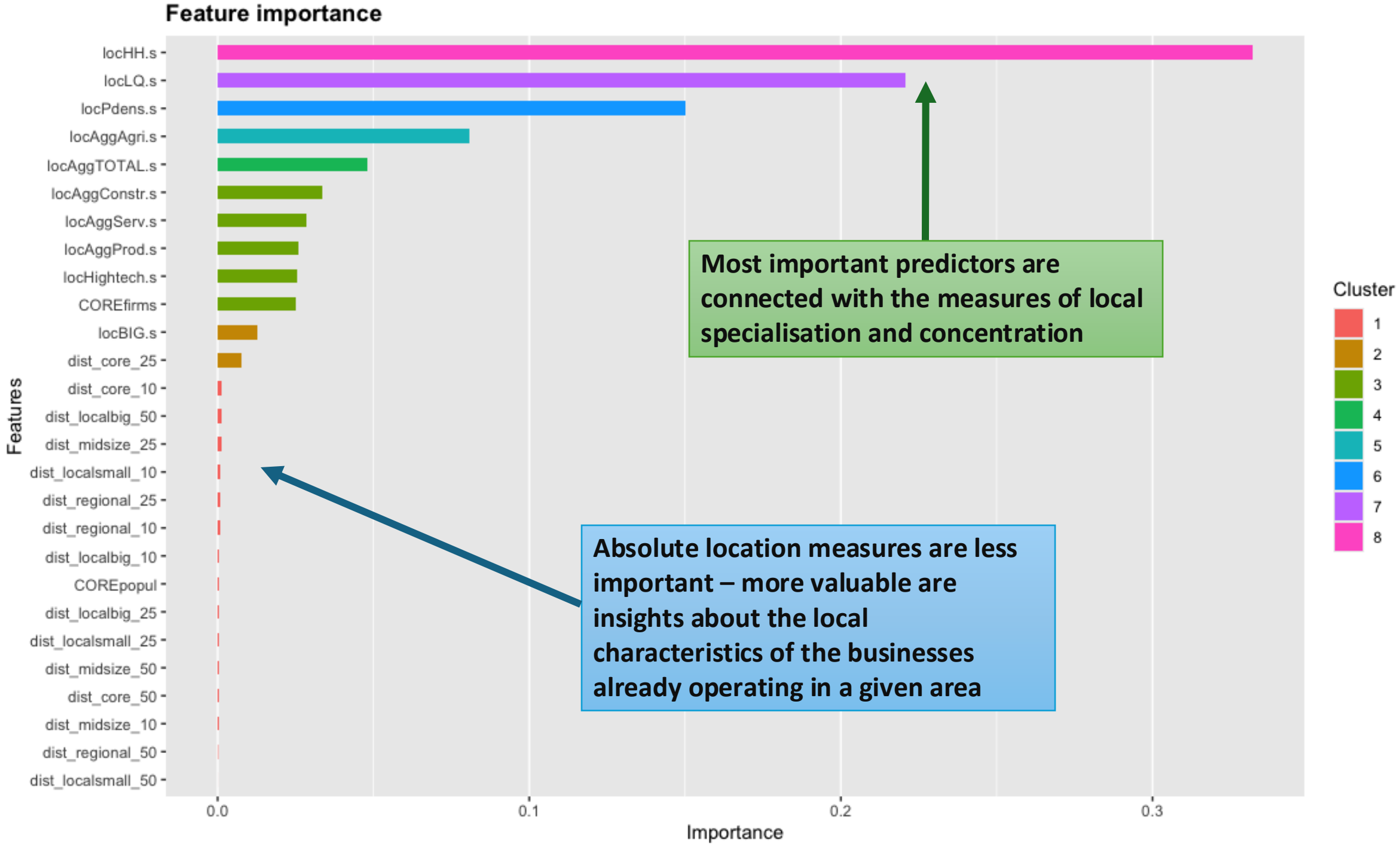
Spatial CV tests the models across diverse spatial areas, checking their stability in different environments. At the same time, it avoids the data leakage issue

Final evaluation of the XGBoost model

```
Overall Statistics

               Accuracy : 0.9296
                 95% CI : (0.9291, 0.9301)
    No Information Rate : 0.6045
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.8683

 Mcnemar's Test P-Value : < 2.2e-16

Statistics by Class:

                     Class: agri Class: constr Class: prod Class: serv
Sensitivity               0.9884       0.56938     0.61432      0.9818
Specificity               0.9904       0.99458     0.99644      0.8612
Pos Pred Value            0.9728       0.89182     0.92303      0.9153
Neg Pred Value            0.9960       0.96714     0.97378      0.9687
Prevalence                0.2577       0.07276     0.06505      0.6045
Detection Rate            0.2547       0.04143     0.03996      0.5935
Detection Prevalence      0.2619       0.04646     0.04329      0.6484
Balanced Accuracy         0.9894       0.78198     0.80538      0.9215
```

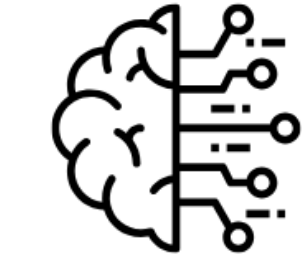
```
Confusion Matrix and Statistics

          Reference
Prediction   agri constr   prod   serv
    agri   250579    883    994   5139
    constr     17  40756   2006   2921
    prod        2    521  39310   2755
    serv     2916  29419  21679 583822
```

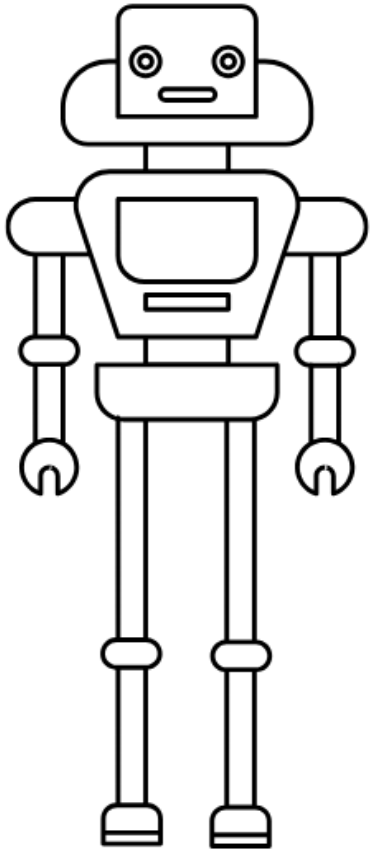
Quite good and stable prediction, general accuracy at the level of 0.9296. Still, possible room for improvement for better handling of categories “construction” and “production” – they are frequently misclassified as service.


Feature importance
locHH.s
locLQ.s
locPdens.s
locAggAgri.s
locAggTOTAL.s
locAggConstr.s
locAggServ.s
locAggProd.s
locHightech.s
COREfirms
locBIG.s
dist_core_25
dist_core_10
dist_localbig_50
dist_midsize_25
dist_localsmall_10
dist_regional_25
dist_regional_10
dist_localbig_10
COREpopul
dist_localbig_25
dist_localsmall_25
dist_midsize_50
dist_core_50
dist_midsize_10
dist_regional_50
dist_localsmall_50
Features
0.0
0.1
0.2
0.3
Importance
Most important predictors are connected with the measures of local specialisation and concentration
Absolute location measures are less important – more valuable are insights about the local characteristics of the businesses already operating in a given area
Cluster
1
2
3
4
5
6
7
8

Transform the ML analysis into AI-driven decision-making tool

Actionable business insight:
For a specific location – what will be the best type of business?





Artificial Intelligence for Regional Science

General goal of intelligence:

Learn from experience and draw insights

To make decisions and take actions

App for AI-informed decision making

Business Sector Prediction App

Input Features

Adjust the sliders and checkboxes to set the values for each feature.

Reset Inputs

locPdens.s

0 — 1

locAggAgri.s

0 — 0.5 — 1

locAggProd.s

0 — 0.2 — 1

locAggConstr.s

0 — 1

- [ ] dist_localsmall_10
- [ ] dist_localsmall_25
- [ ] dist_localsmall_50
- [x] COREfirms
- [ ] COREpopul

Predict

Prediction Results

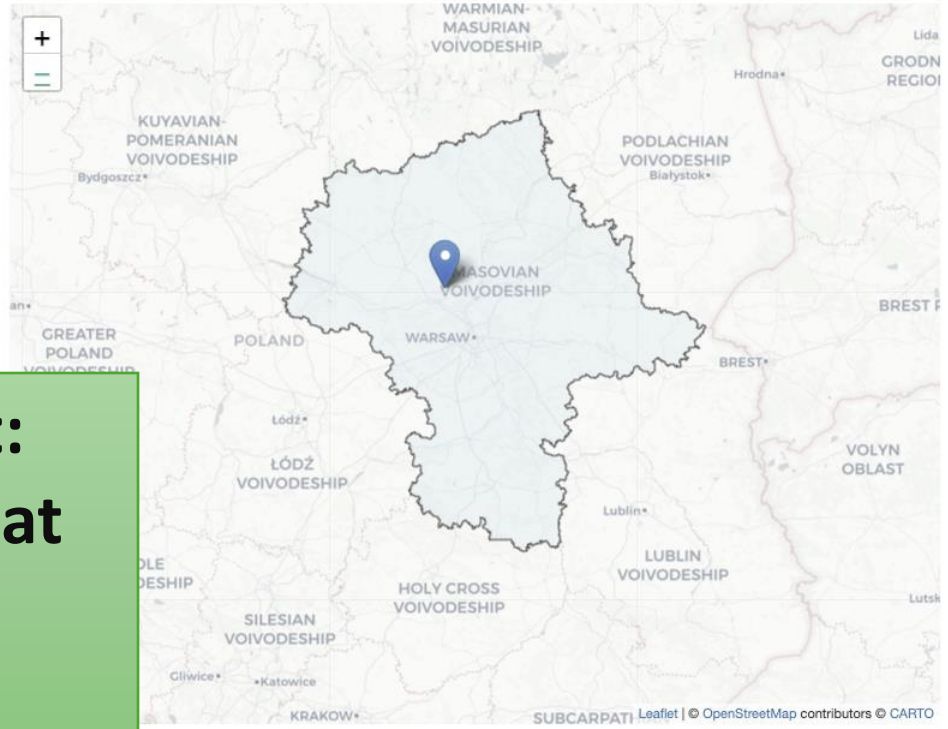
Class	Probability
agri	0.0000
constr	0.0001
prod	0.0004
serv	0.9995

Business Sector Prediction Map

Instructions

Click on the map to select a location. The app will find the nearest point with available data and display the predicted probabilities for each business sector at that location.

Reset Map



Actionable business insight: For a specific location – what will be the best type of business?

Prediction Results

Class	Probability
serv	56.38
agri	38.60
prod	3.59
constr	1.44

App for AI-informed decision making

Business Sector Prediction App

Input Features

Adjust the sliders and checkboxes to set the values for each feature.

Reset Inputs

locPdens.s

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 (1)

locAggAgri.s

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 (0.5)

locAggProd.s

0 0.1 0.2 0.3 0.4 0.5 (0.2)

locAggConstr.s

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 (1)

- [ ] dist_localsmall_10
- [ ] dist_localsmall_25
- [ ] dist_localsmall_50
- [x] COREfirms
- [ ] COREpopul

Predict

Choose parameters of the location

Prediction Results

Class	Probability
agri	0.0000
constr	0.0001
prod	0.0004
serv	0.9995

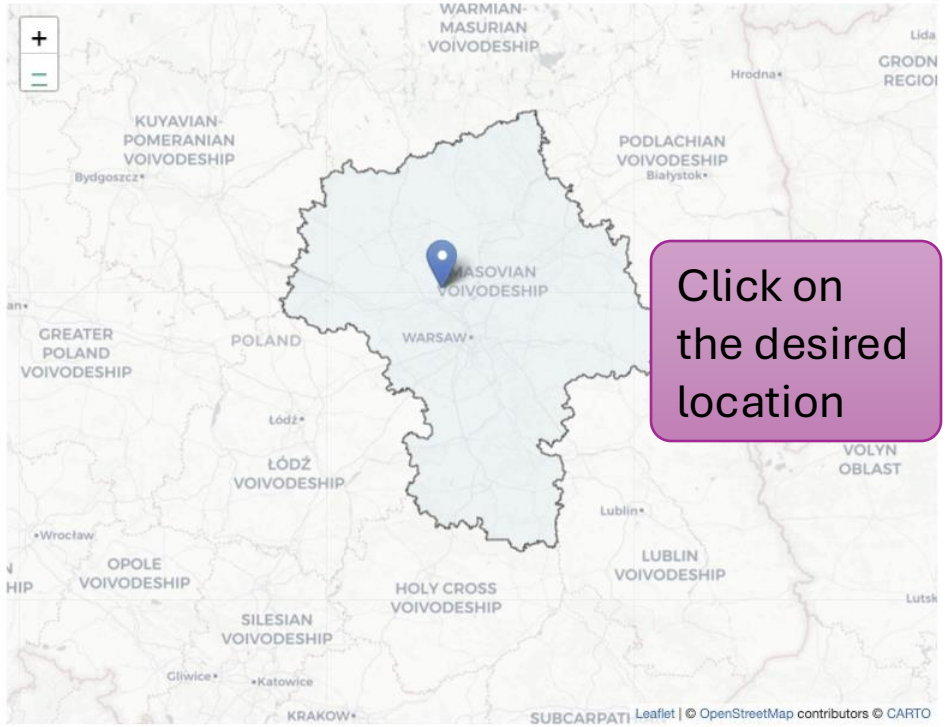
Predict the best business type based on the values or get them directly from the map

Business Sector Prediction Map

Instructions

Click on the map to select a location. The app will find the nearest point with available data and display the predicted probabilities for each business sector at that location.

Reset Map



Click on the desired location

Prediction Results

Class	Probability
serv	56.38
agri	38.60
prod	3.59
constr	1.44

Get the best business type

Business Sector Prediction Map

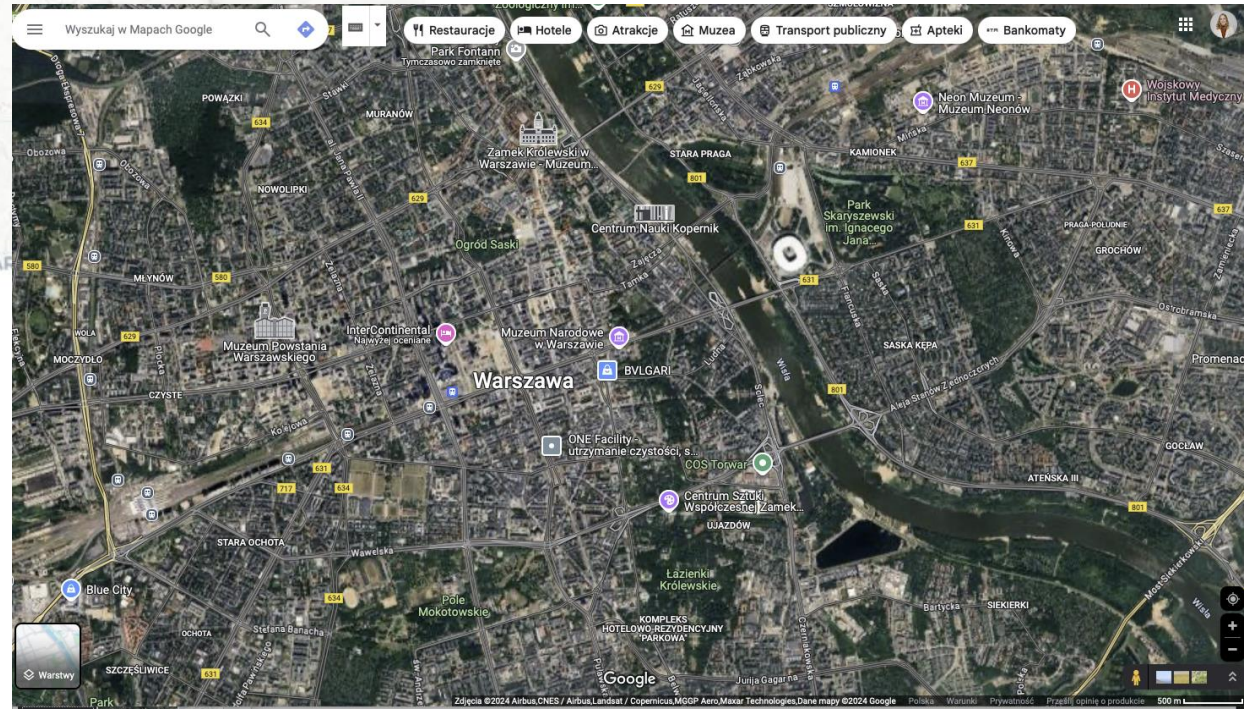
Instructions

Click on the map to select a location. The app will find the nearest point with available data and display the predicted probabilities for each business sector at that location.

Reset Map

Prediction Results

Class	Probability
serv	92.58
constr	5.98
prod	1.44
agri	0.00



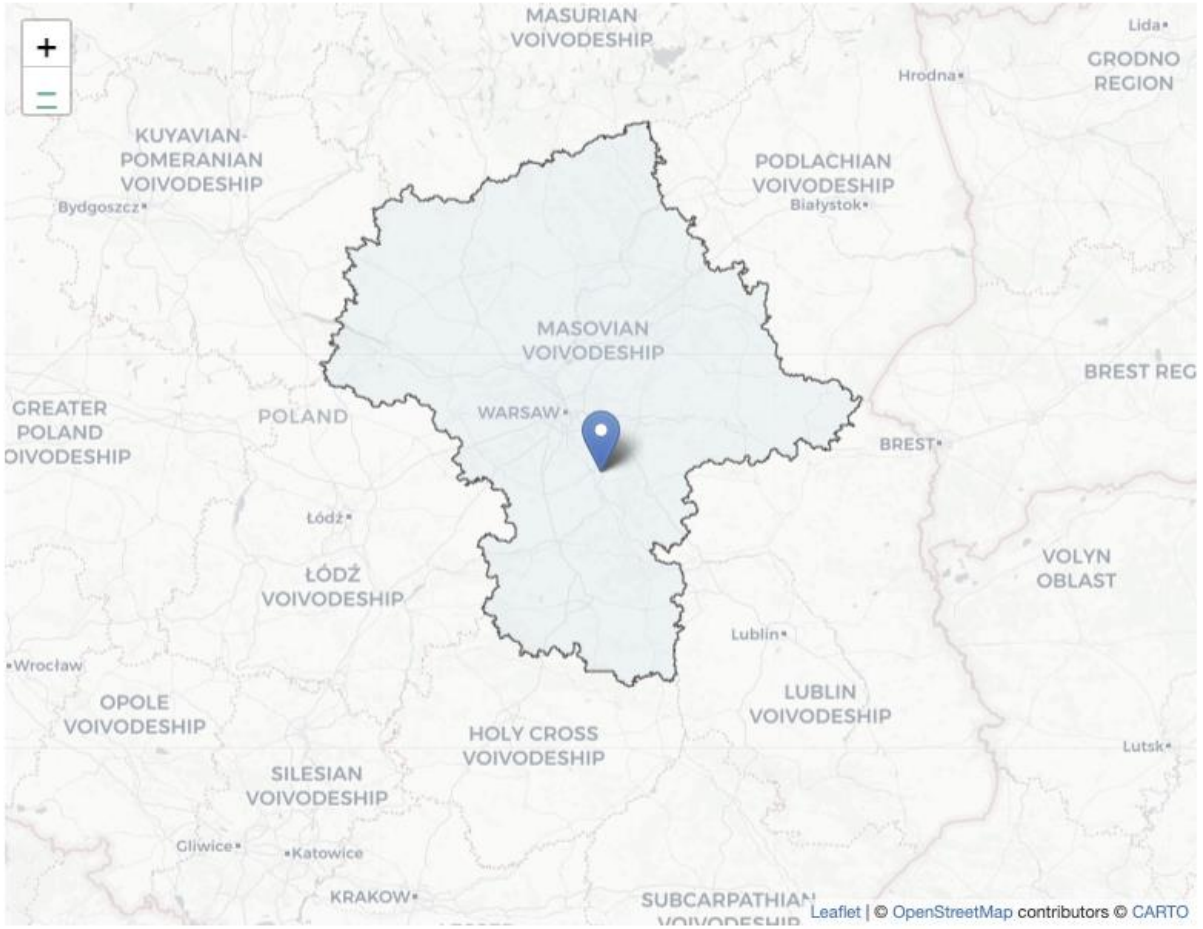
”Sanity check”

Business Sector Prediction Map

Instructions

Click on the map to select a location. The app will find the nearest point with available data and display the predicted probabilities for each business sector at that location.

Reset Map



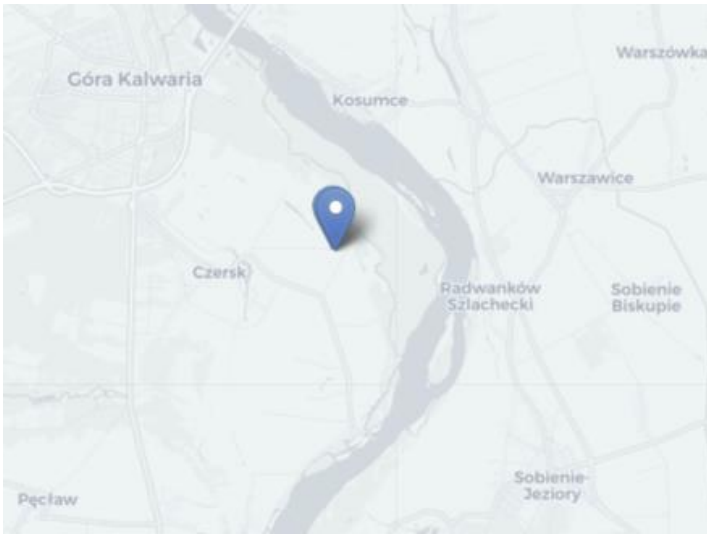
Prediction Results

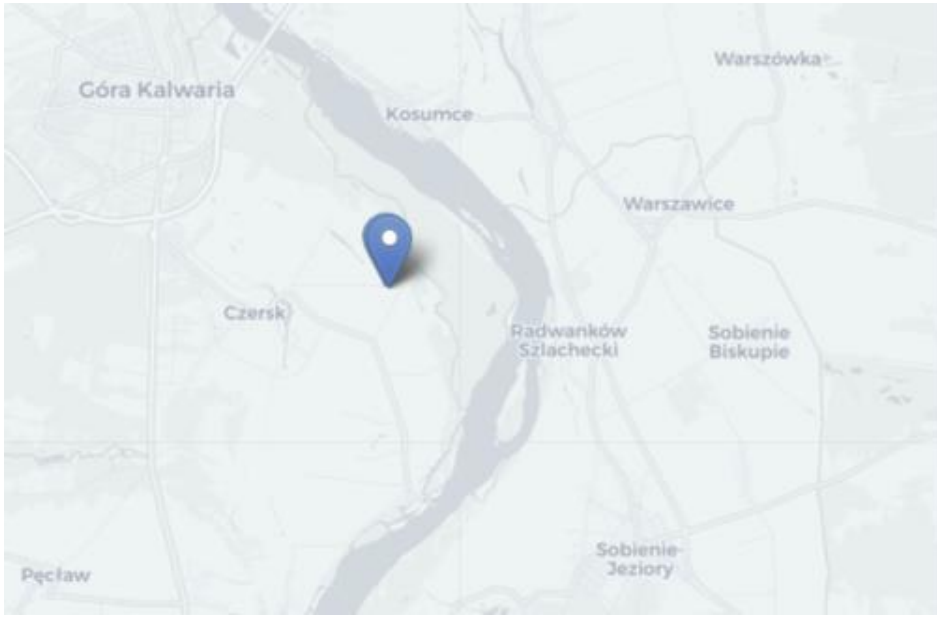
Class	Probability
agri	99.71
serv	0.24
constr	0.03
prod	0.02

”Sanity check”

Prediction Results

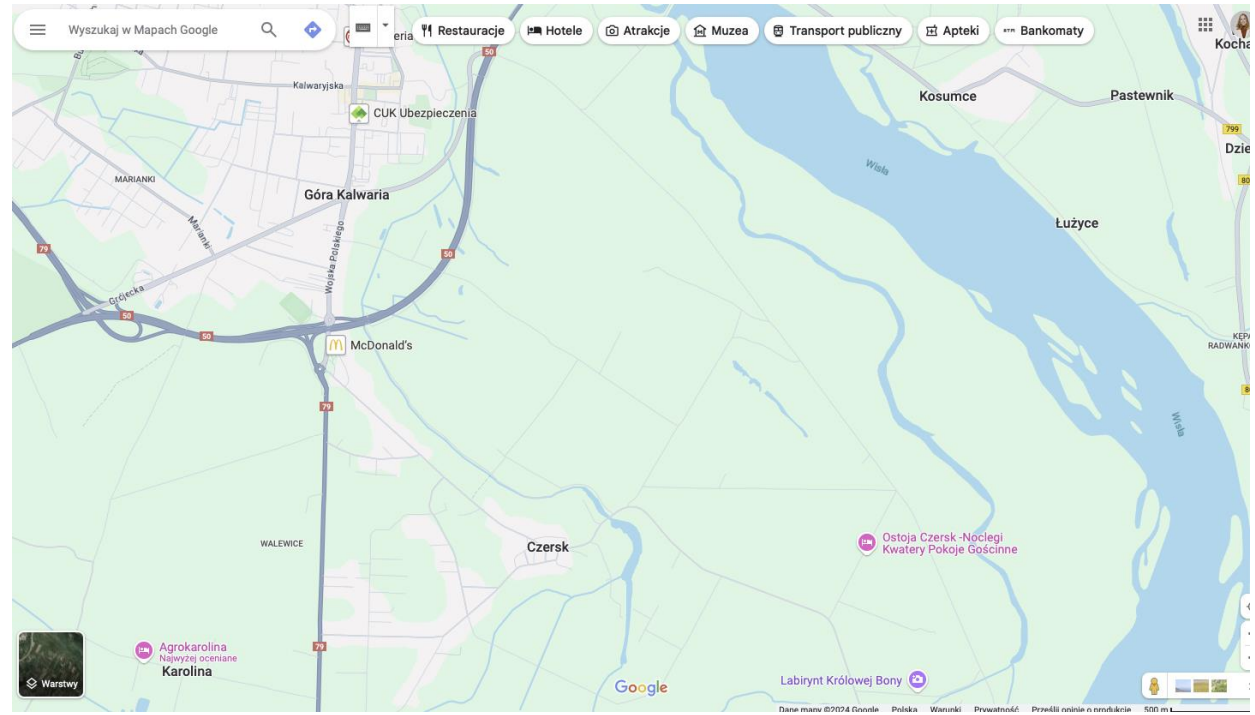
Class	Probability
agri	99.71
serv	0.24
constr	0.03
prod	0.02

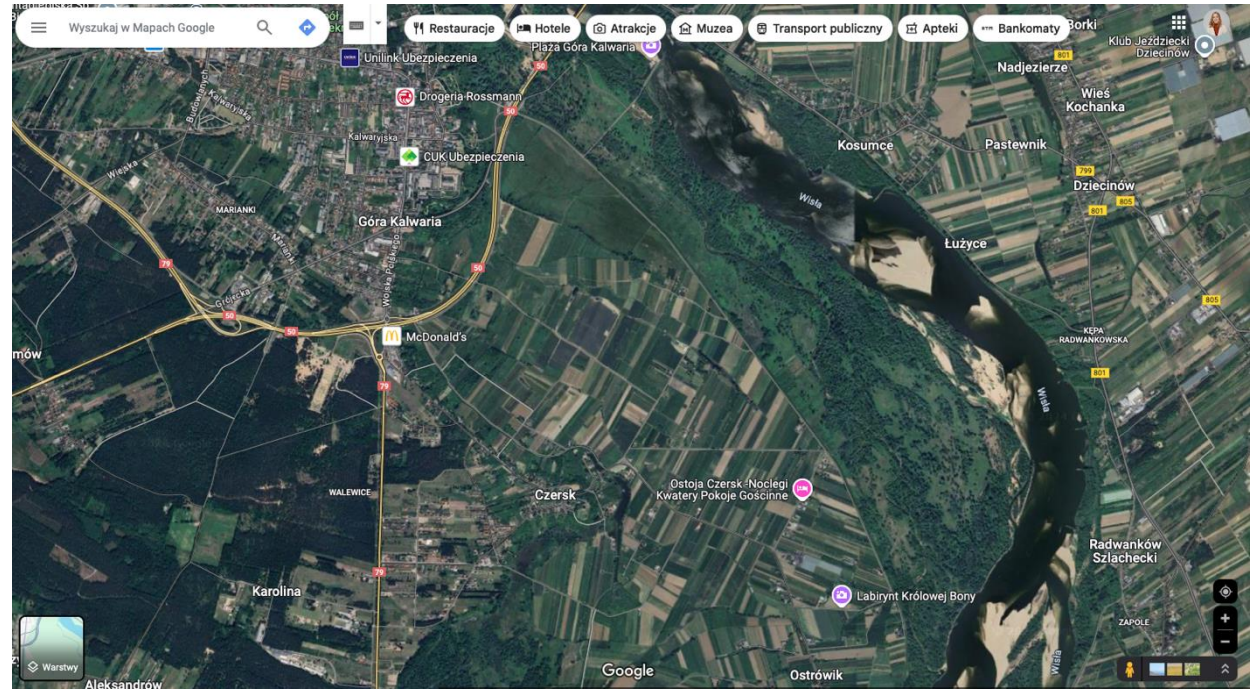




Prediction Results

Class	Probability
agri	99.71
serv	0.24
constr	0.03
prod	0.02





Key Takeaways

AI in Regional Science: A Decision-Making Tool

- Artificial Intelligence (AI) goes beyond pure Machine Learning (ML) by transforming data insights into actionable decisions.
- AI supports **decision-making** through integration of predictions, reasoning, and domain-specific considerations, bridging research and practical applications.

The Importance of Spatial Context

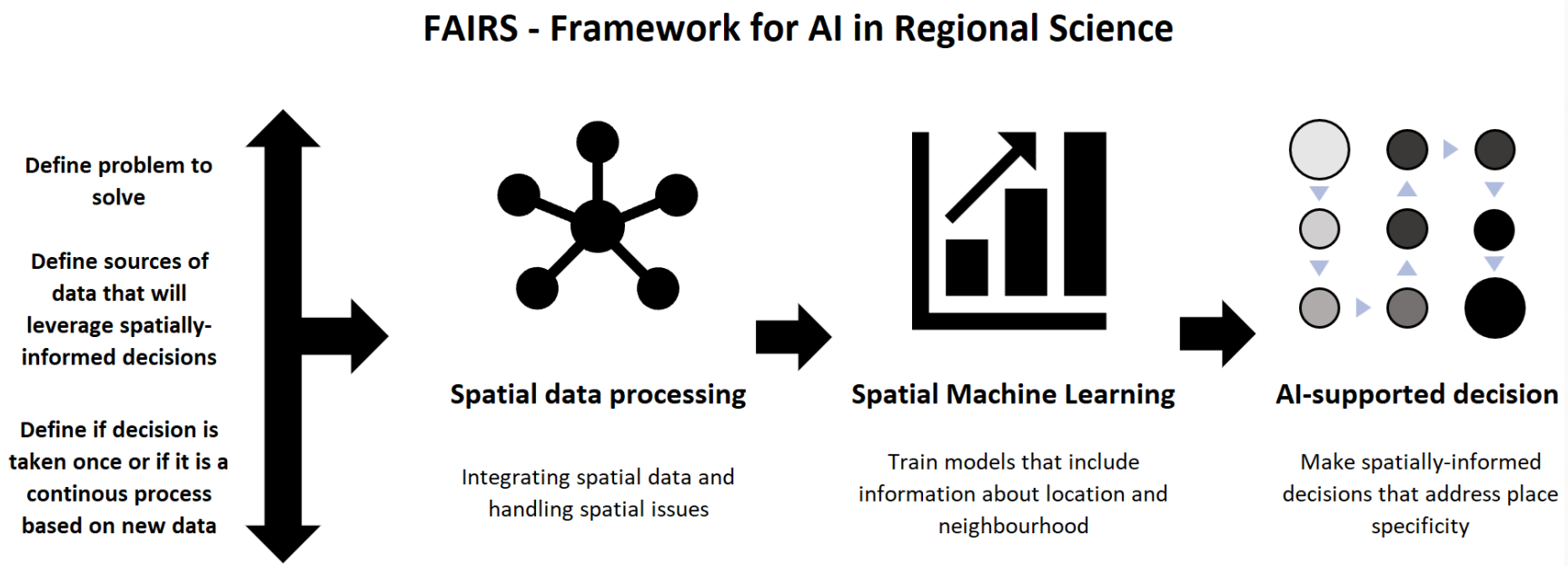
- Regional science requires methods that account for the unique characteristics of **spatial data**, such as heterogeneity and proximity-based relationships.
- **Spatial Machine Learning** must address these challenges, ensuring robust evaluation frameworks like spatial cross-validation and avoiding data leakage.
- AI in regional science must inherently incorporate spatial understanding, enriching insights and making location-sensitive decisions.

From Research to Real-World Applications

- Our example demonstrated how **spatial ML** can be applied to predict optimal business locations.
- The resulting AI-driven decision-making tool highlights how these methods can guide **business actors** in choosing locations or **policy makers** in supporting regional economic development.

Broader Implications

- The integration of spatially-informed AI opens new possibilities for answering regional science questions, combining methodological advancements and merging of ML with spatial econometrics insights with real-world utility.
- Such tools can be adapted beyond research to tackle practical challenges in urban planning, economic development, and resource allocation.



AI for regional science

Thank You!

Questions?

@SpatialWarsaw
@KubaraMaria maria.kubara@uw.edu.pl
@KathyKopczewska kkopczewska@wne.uw.edu.pl