DIGITAL FINANCE

This project has received funding from the Horizon Europe research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101119635

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

State Secretariat for Education, Research and Innovation SERI

Funded by the European Union

MARIE CURIE ACTIONS

DIGITAL

About IRP 16 / WP3

WP3 “Towards explainable and fair AI-generated decisions”

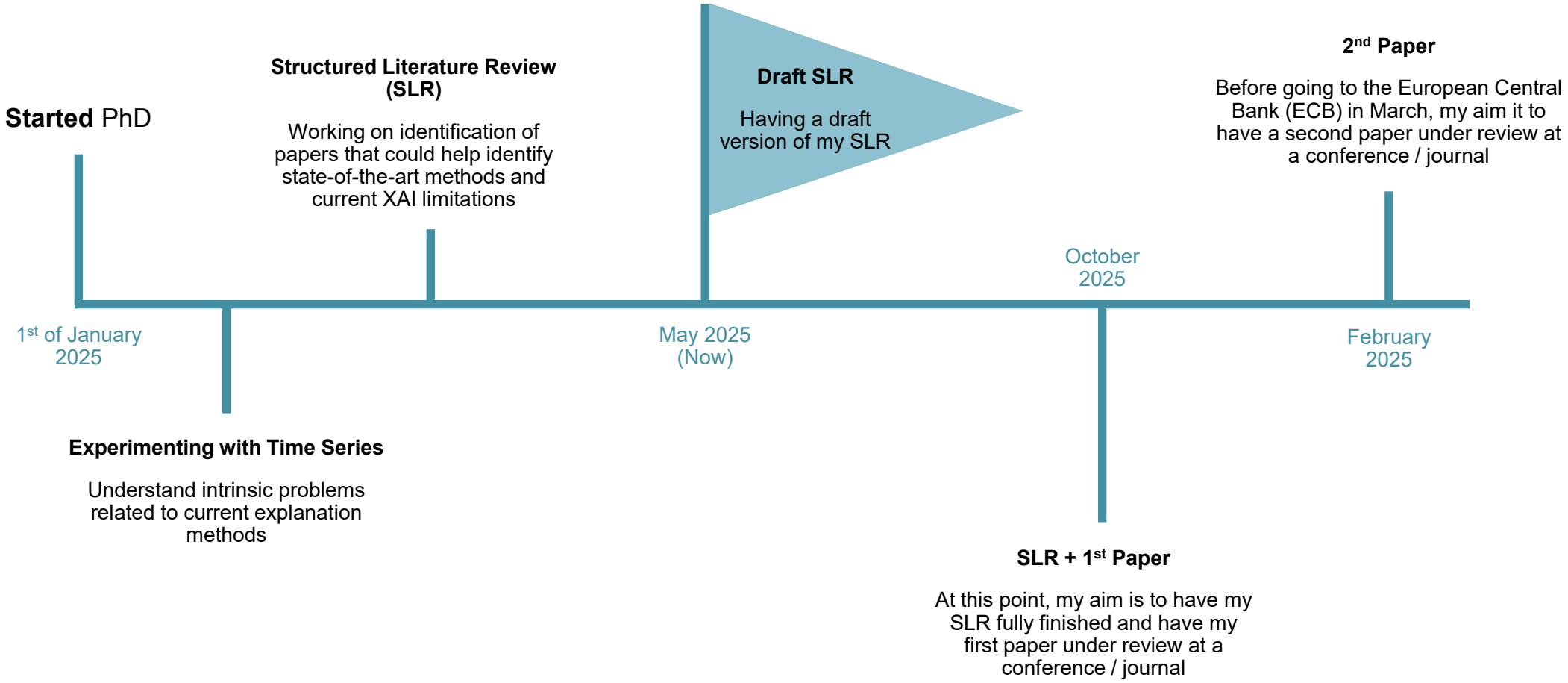
IRP16 “Investigating the utility of classical XAI methods in financial time series”

Planned Secondments

European Central Bank: Dr. Lukasz Kubicki, M27, 12 months, exposure to globally leading central bank research, training on EU principles

Fraunhofer (FRA): Prof. Dr. Ralf Korn, M39, 6 months

Timeline


Started PhD
1st of January 2025
Structured Literature Review (SLR)
Working on identification of papers that could help identify state-of-the-art methods and current XAI limitations
Experimenting with Time Series
Understand intrinsic problems related to current explanation methods
Draft SLR
Having a draft version of my SLR
May 2025 (Now)
October 2025
SLR + 1st Paper
At this point, my aim is to have my SLR fully finished and have my first paper under review at a conference / journal
2nd Paper
Before going to the European Central Bank (ECB) in March, my aim it to have a second paper under review at a conference / journal
February 2025

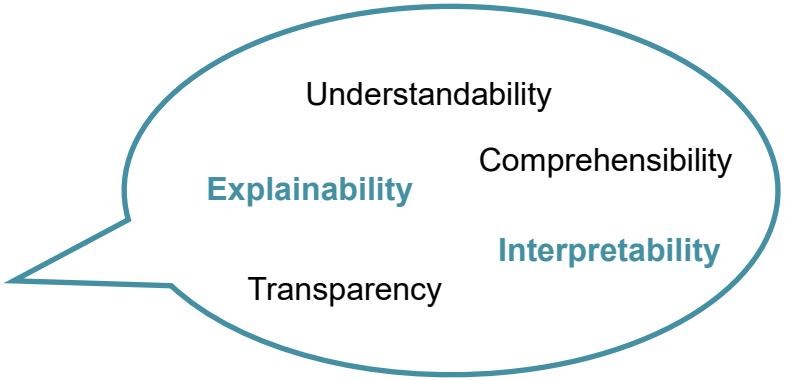
Definitions

XAI Explainable Artificial Intelligence

Papers Interchangeable misuse of terminology in the literature between explainability and interpretability

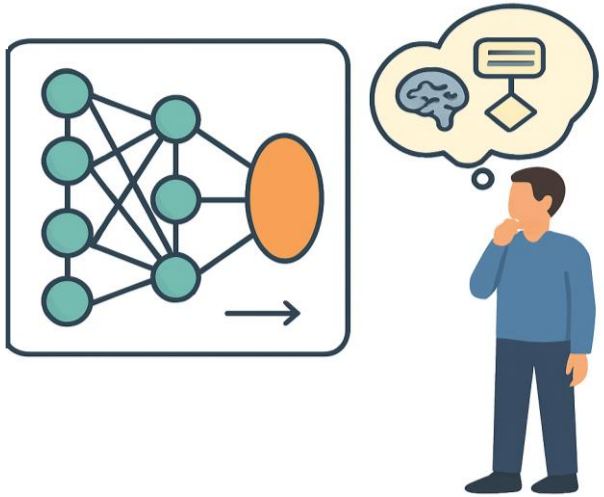
Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI

Alejandro Barredo Arrieta [a], Natalia Díaz-Rodríguez [b], Javier Del Ser [a c d] ✉, Adrien Bennetot [b e f], Siham Tabik [g], Alberto Barbado [h], Salvador Garcia [g], Sergio Gil-Lopez [a], Daniel Molina [g], Richard Benjamins [h], Raja Chatila [f], Francisco Herrera [g]

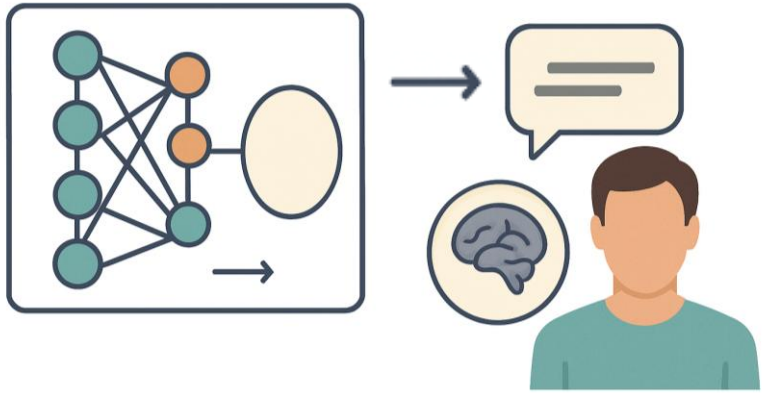


Definitions

Explainability	An active characteristic of a model, denoting any action or procedure taken by a model with the intent of clarifying or detailing its internal functions
Interpretability	Refers to a passive characteristic of a model referring to the level at which a given model makes sense for a human observer



Interpretability

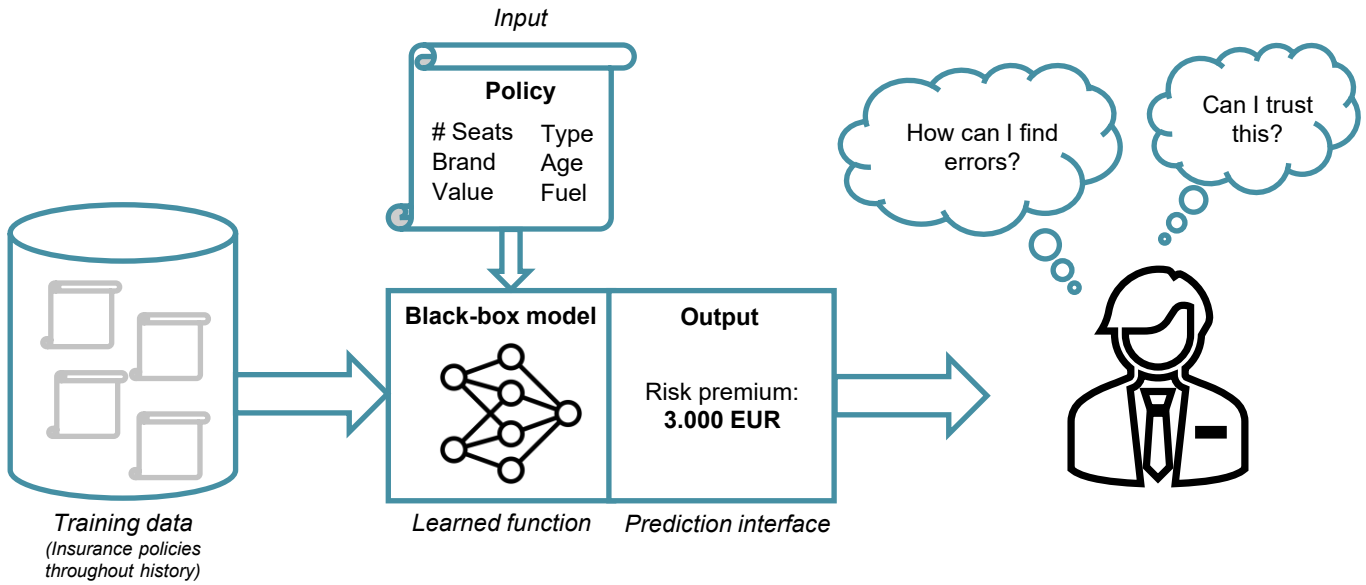


Explainability

What is XAI and why does it matter?

XAI Explainable Artificial Intelligence

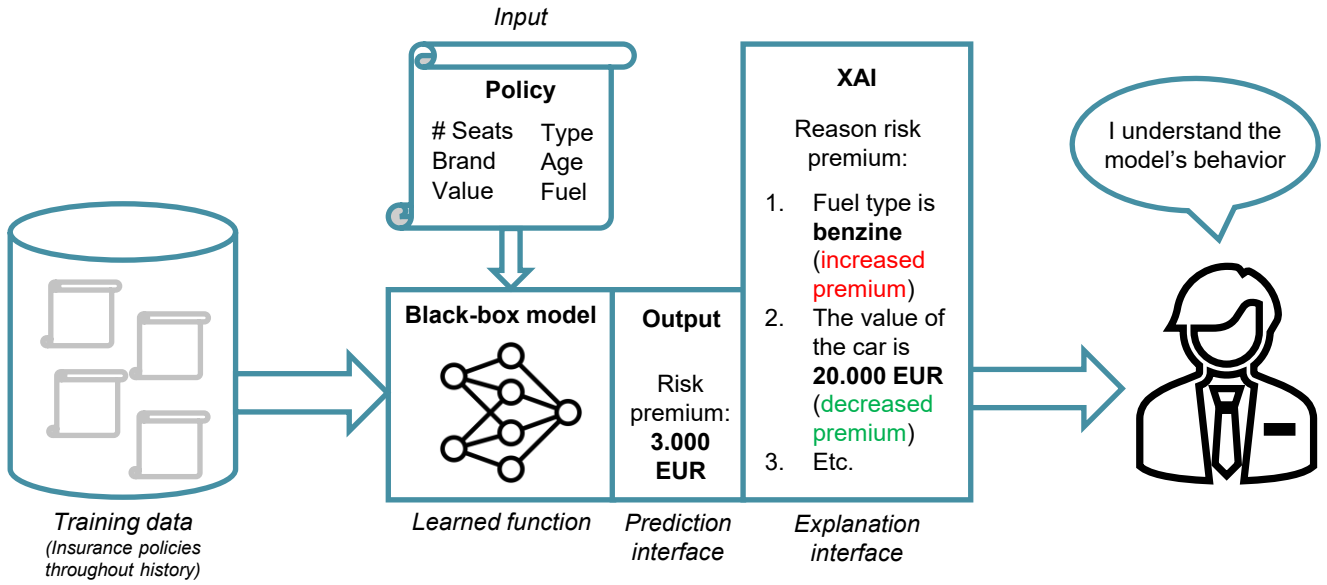
Why? XAI is critical because black-box models dominate, but understanding "why" matters more than ever



What is XAI and why does it matter?

XAI Explainable Artificial Intelligence

Why? XAI is critical because black-box models dominate, but understanding "why" matters more than ever



The rise of XAI

Growth The field of XAI is a fast-growing research field

Citations XAI has rapidly gained momentum with methods like SHAP and LIME. Thousands of citations highlight the growing impact on interpretability research

arXiv
https://arxiv.org › cs

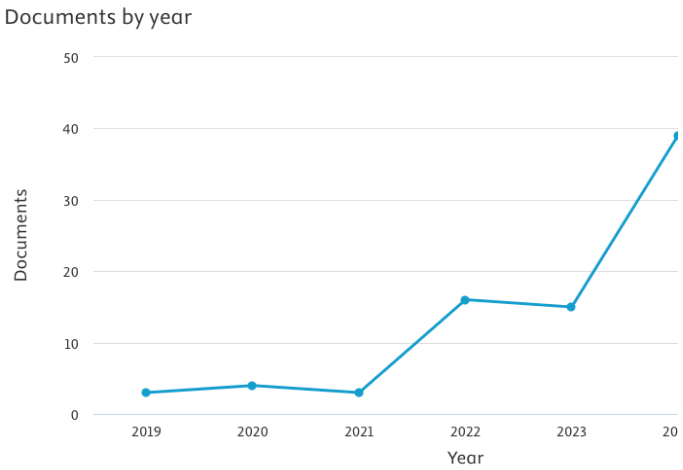
A Unified Approach to Interpreting Model Predictions

door S Lundberg · 2017 · Geciteerd door 34448 — We present a unified framework for interpreting predictions, **SHAP** (SHapley Additive exPlanations). SHAP assigns each feature an importance value f…

arXiv
https://arxiv.org › cs

"Why Should I Trust You?": Explaining the Predictions of ...

door MT Ribeiro · 2016 · Geciteerd door 23338 — In this work, we propose **LIME**, a novel explanation technique that explains the predictions of any classifier in an interpretable and faithful manner.



Scopus analysis on the article distribution of articles describing financial forecasting, ML, and XAI

Why this SLR is needed

Reason No recent SLR focused on temporal data focusing on all fields next to financial data, focusing on limitations. Relevant for finding areas of improvement within the explainability domain for temporal data

Goal Need clarity on what makes an XAI method suitable for time-dependent contexts and address new advancements in literature over the last year and before

A Survey of Explainable Artificial Intelligence (XAI) in Financial Time Series Forecasting

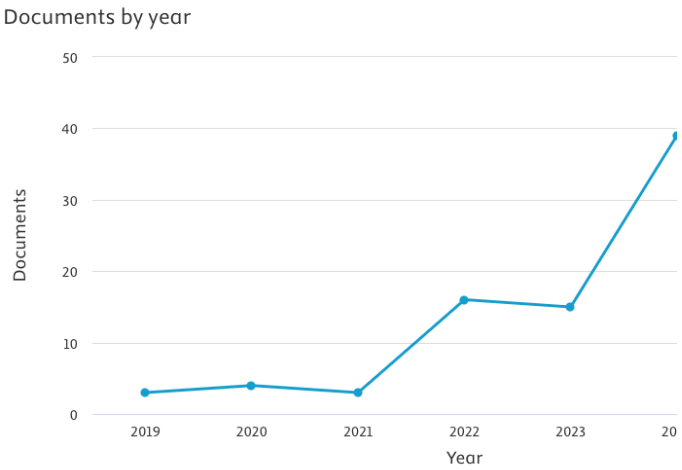
Pierre-Daniel Arsenault
pierre-daniel.arsenault@usherbrooke.ca
https://orcid.org/0009-0005-8911-696X

Shengrui Wang
shengrui.wang@usherbrooke.ca
https://orcid.org/0000-0001-6863-7022
Université de Sherbrooke 2500 Bd de l'UniversitéSherbrookeQuébec Canada J1K 2R1

Jean-Marc Patenaude
jeanmarc@laplaceinsights.com
Laplace Insights 230 King St. West suite 201SherbrookeQuébec Canada J1H 1P9

(2018; July 5, 2024)

"We limited our search to papers published **between 2018 and June 2023**."



Scopus analysis on the article distribution of articles describing financial forecasting, ML, and XAI

SLR overview

How? Papers reviewed are selected based on connections with the field of temporal data, ML, and XAI

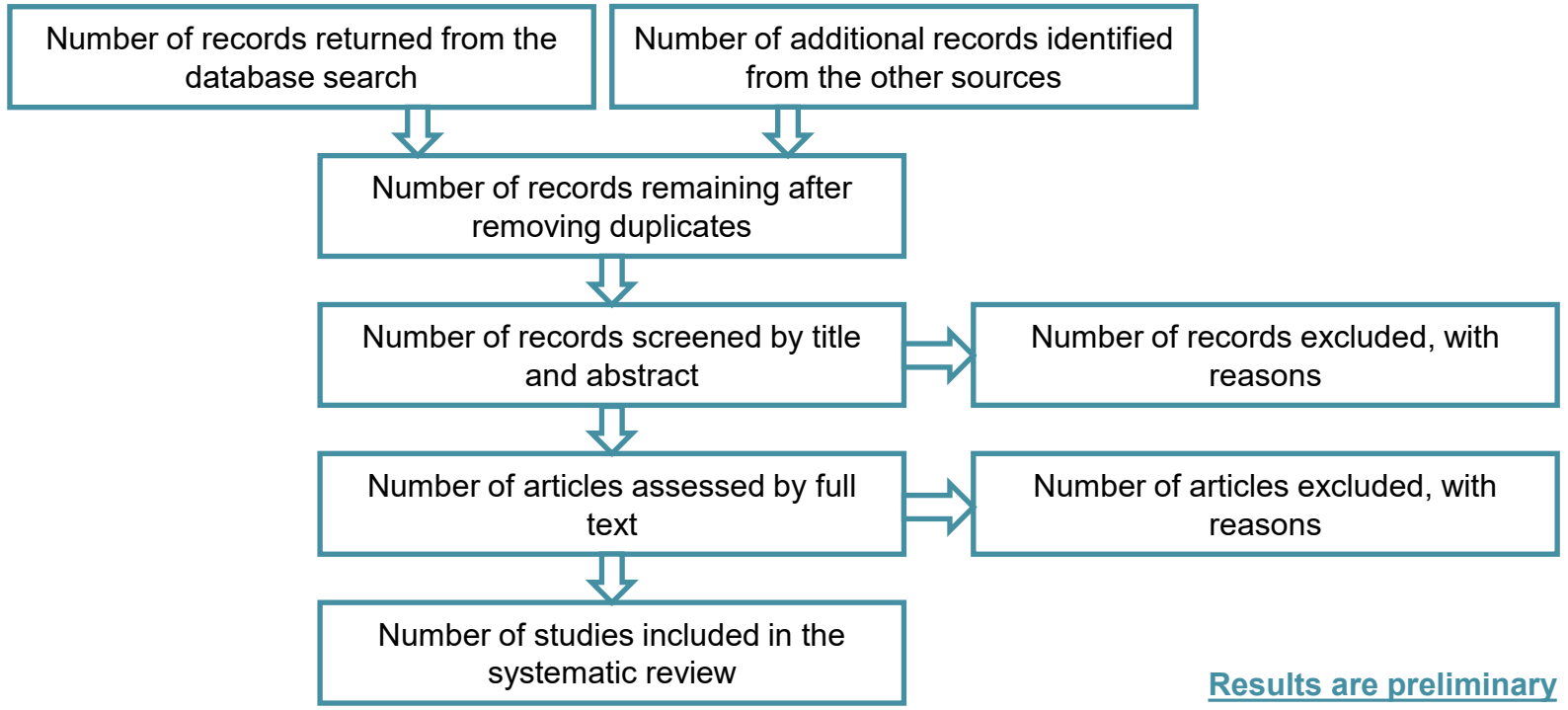
Method PRISMA

200+ Identification

200+ Screening

100+ Eligibility

43 Inclusion

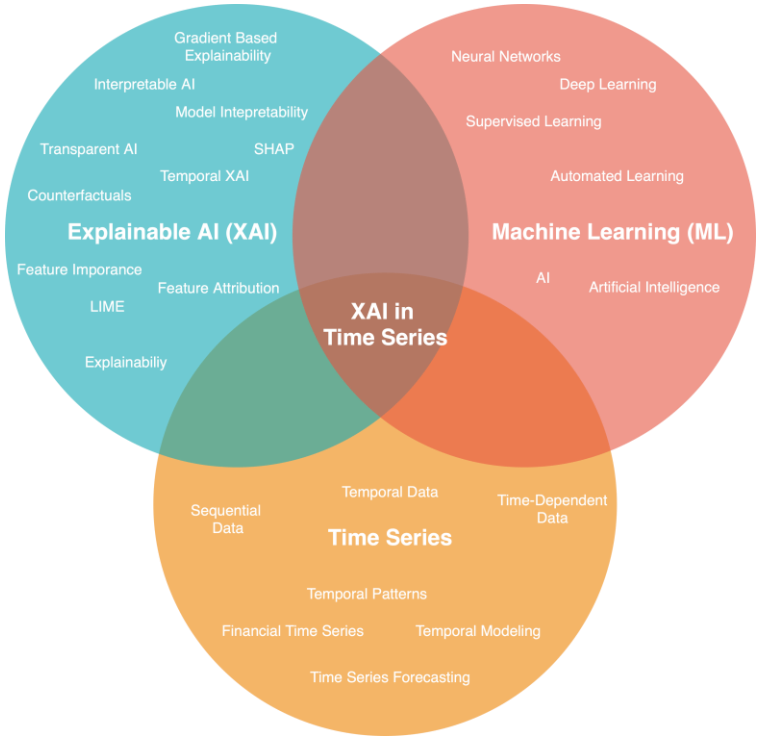


Results are preliminary

Assessing all fields (SLR)

All

This is the first part of the SLR, where all fields (e.g., robotics, healthcare, finance, etc.) were all considered within the domain. Used to give impression of XAI innovation within all fields

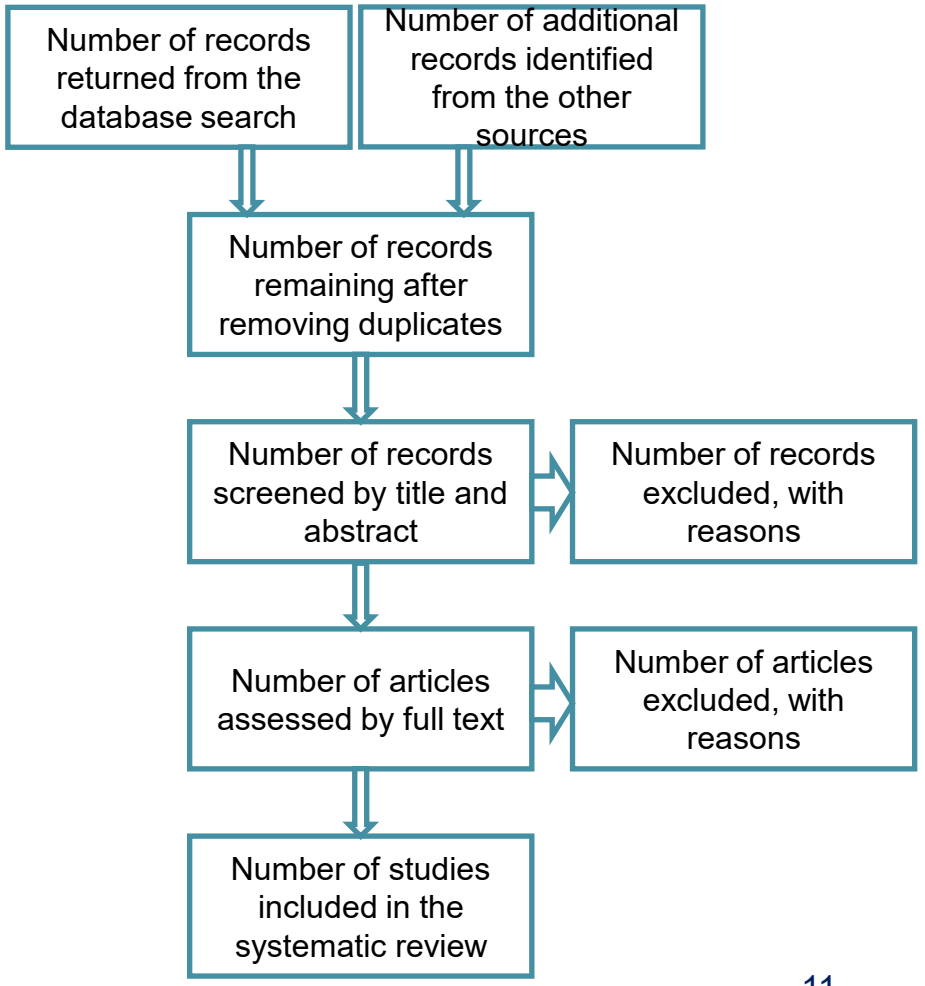


Identification
100+

Screening
100+

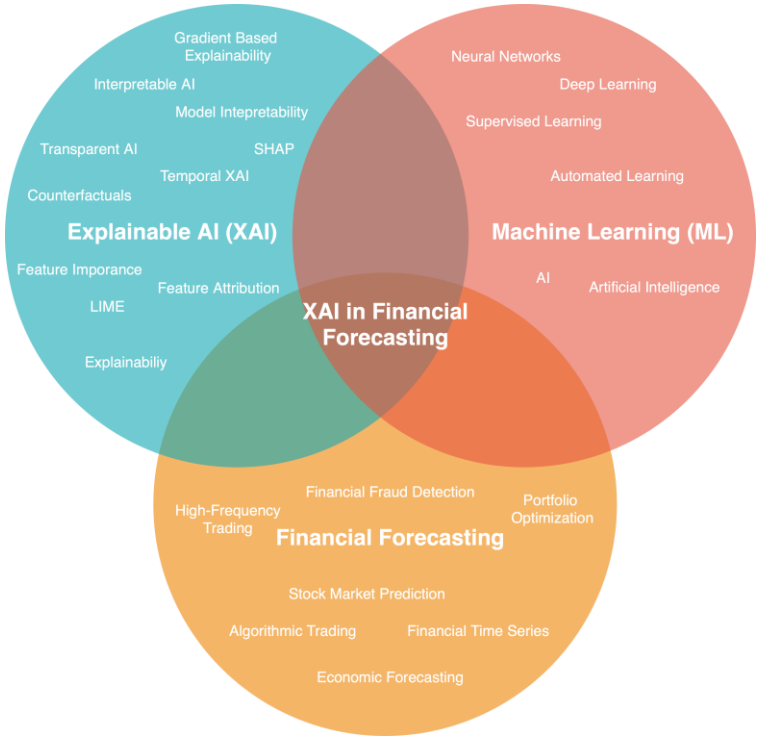
Eligibility
74

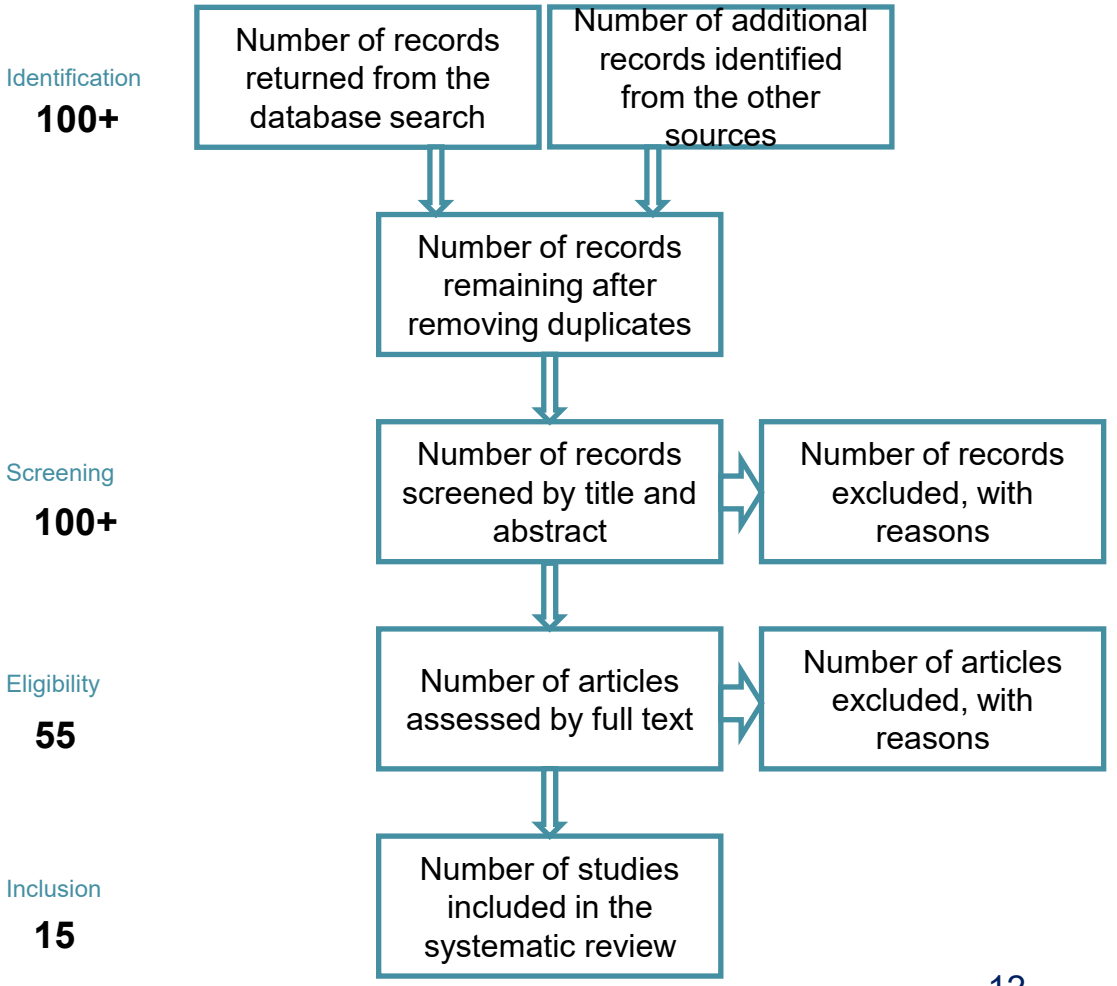
Inclusion
39



Assessing the financial field (SLR)

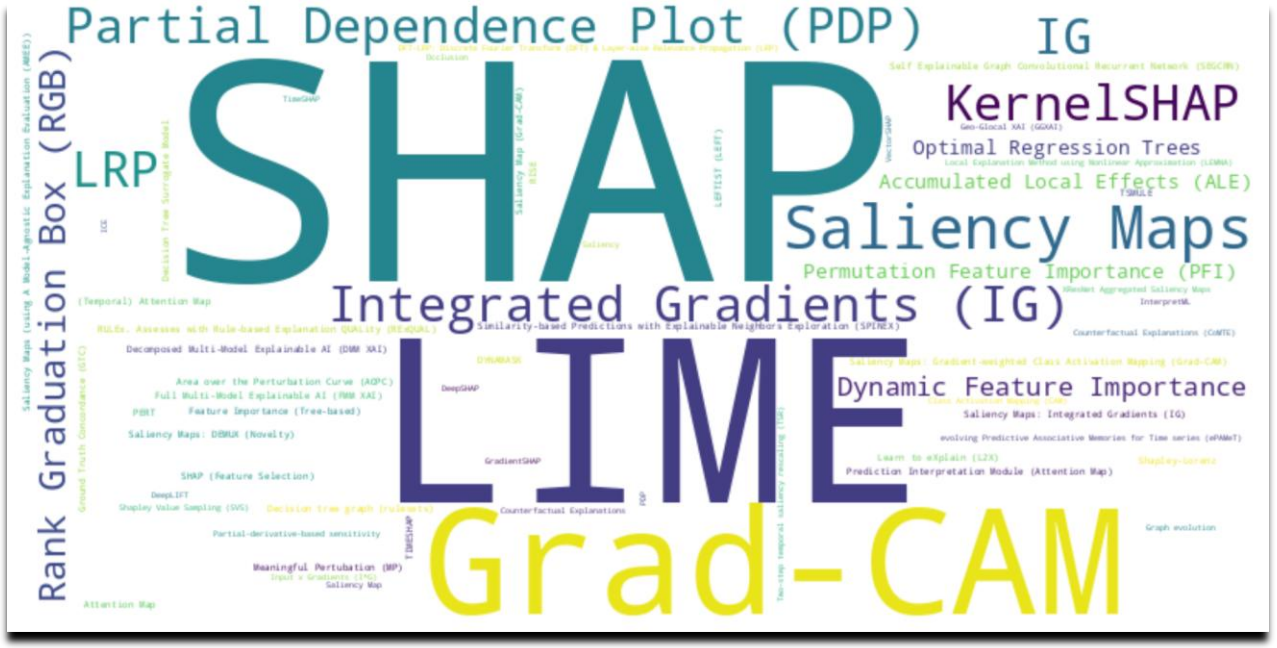
Finance This is the second part of the SLR, where only the financial field is considered. Used to give impression of XAI innovation within the financial domain





XAI for temporal data landscape

XAI	Various techniques are identified. However, despite growing diversity in methods due to decoupling of methods in literature, SHAP & LIME still dominate
Insight	Gradient- and attention-based methods are emerging for temporal data, integrated gradients and attention models show promise but are underused



How do current papers integrate XAI?

Insight Papers in the financial time series domain are mostly application-based, meaning these papers just apply existing XAI algorithms (like SHAP, LIME, etc.) to their produced algorithm and make their models “explainable”

But These authors don’t assess whether their XAI algorithms are showing the reality within the black-box as they just assume the XAI algorithm is correct

1

Proposal of time series forecasting algorithm that outperforms a certain benchmark

2

To ensure “explainability”, XAI algorithms are added in the paper as an additional application

3

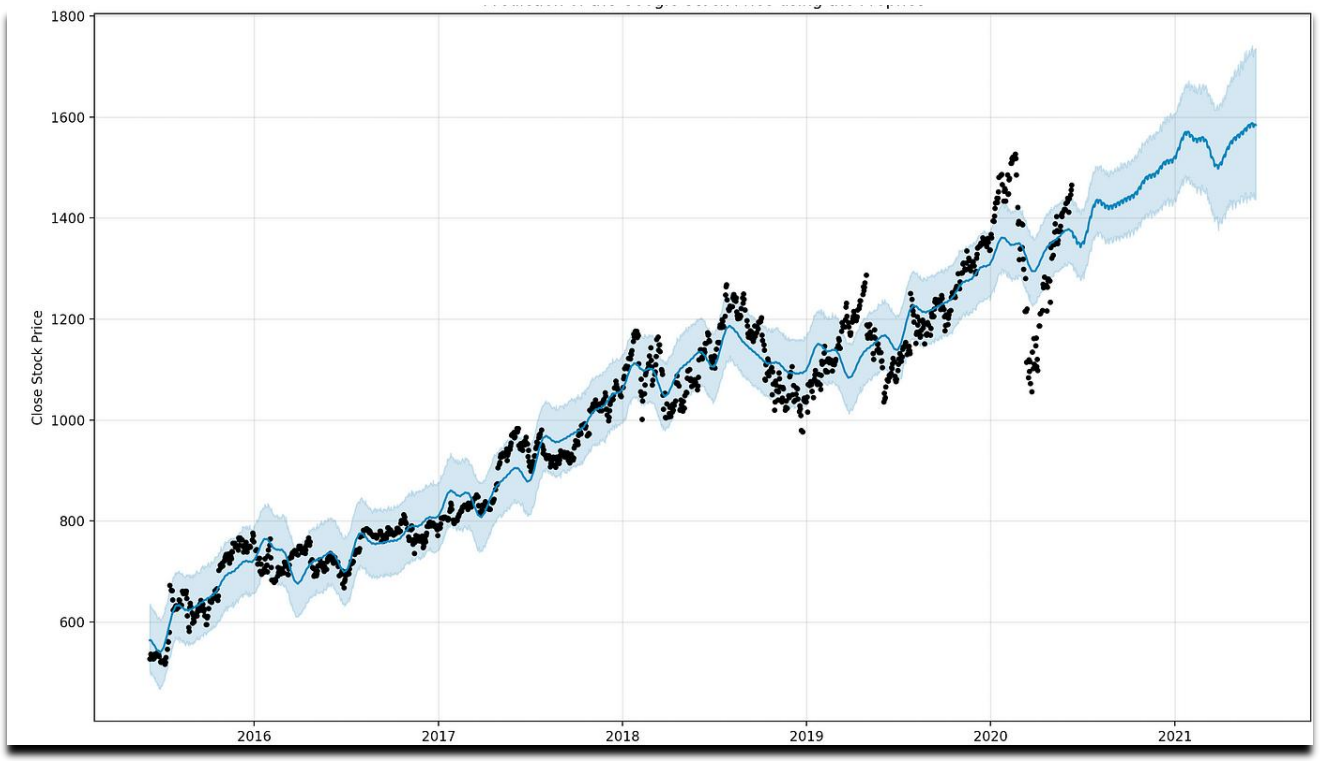
Based on this combination, various insights are retrieved and stated in the results and conclusion sections

There is no investigation conducted into the suitability of the XAI algorithm for the specific task

Why further research into XAI is needed

Finding XAI methods are growing but the problem of explainability is not yet solved

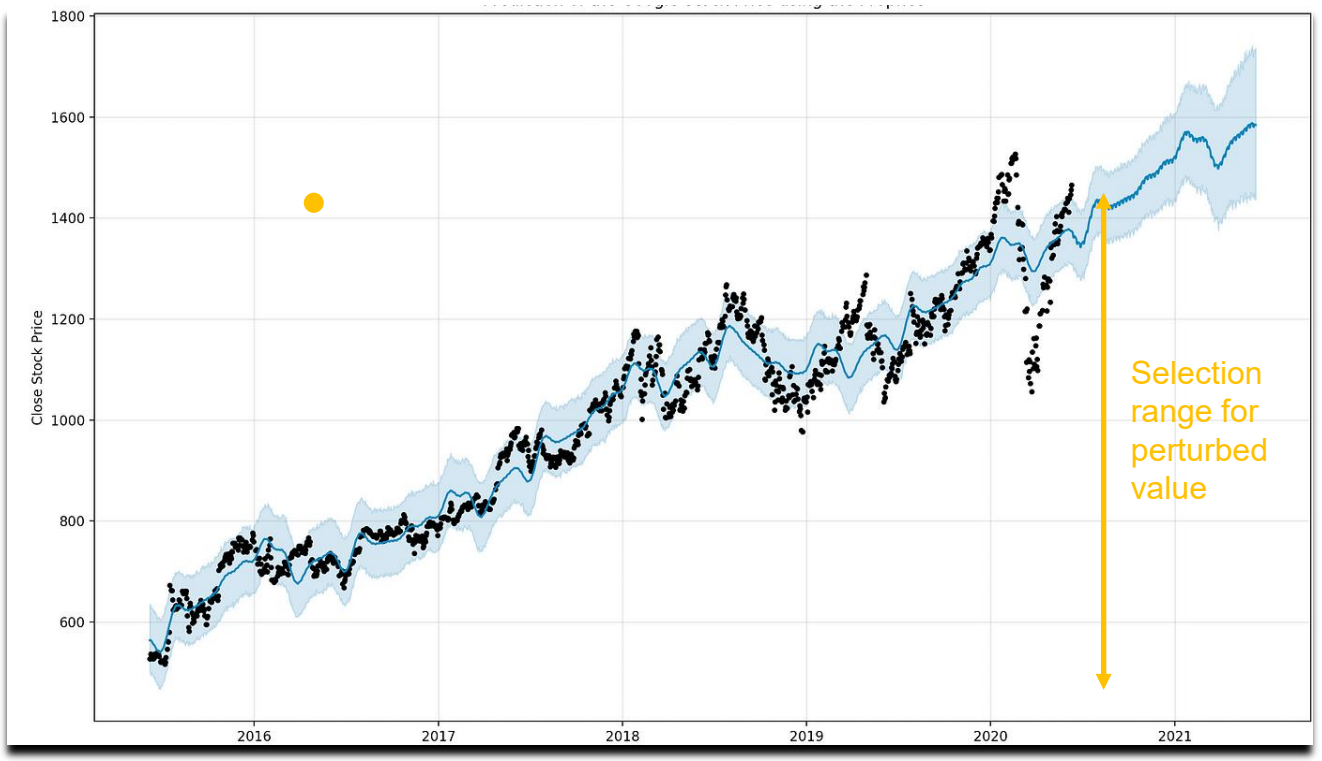
Status Most studies use SHAP & LIME, which are perturbation-based methods. Growing amount of work that uncovers limitations of state-of-art (perturbation-based) XAI methods when applied to temporal data



Why further research into XAI is needed

Finding XAI methods are growing but the problem of explainability is not yet solved

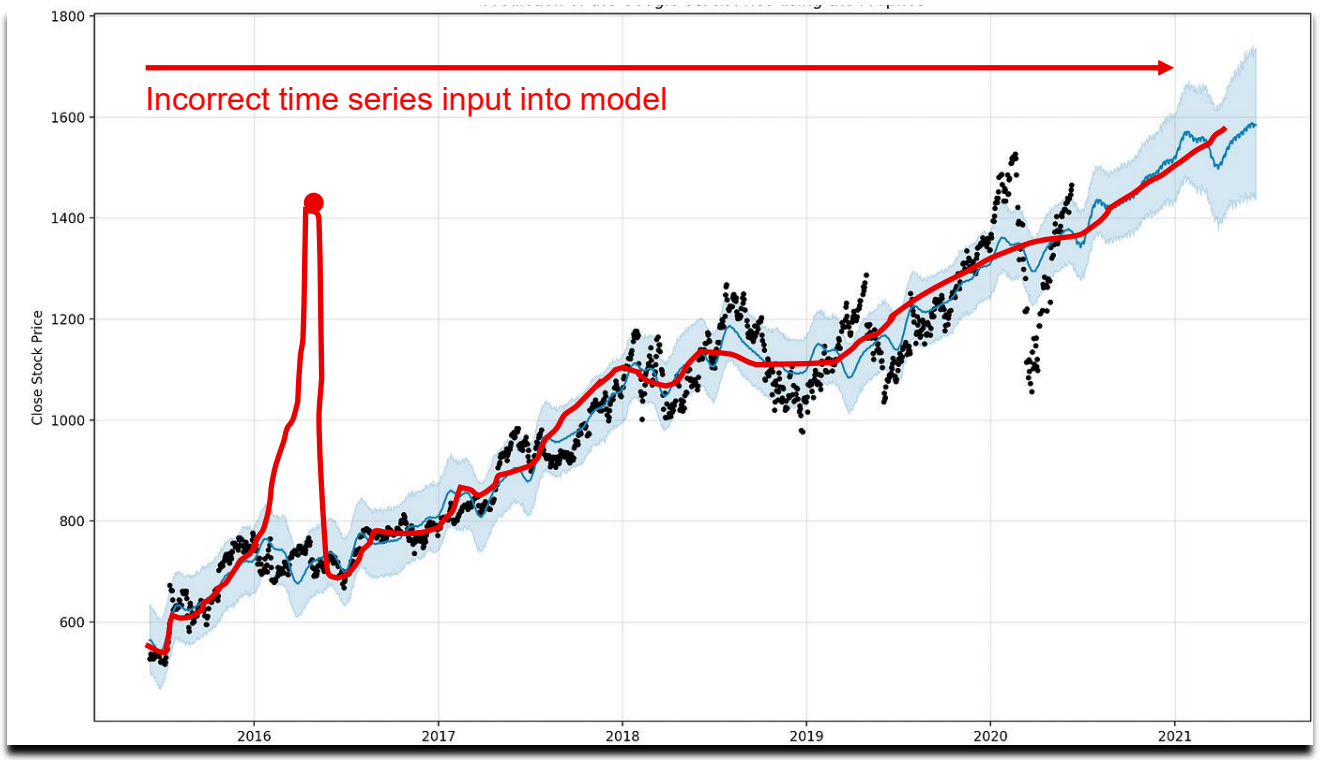
Status Most studies use SHAP & LIME, which are perturbation-based methods. Growing amount of work that uncovers limitations of state-of-art (perturbation-based) XAI methods when applied to temporal data



Why further research into XAI is needed

Finding XAI methods are growing but the problem of explainability is not yet solved

Status Most studies use SHAP & LIME, which are perturbation-based methods. Growing amount of work that uncovers limitations of state-of-art (perturbation-based) XAI methods when applied to temporal data



Towards tailored XAI for finance

Identify Time dependencies are broken when data is randomly perturbed. Generating synthetic samples that distort sequence structure creates inconsistent explanations

Improve A future XAI solution should address or avoid the weaknesses within current XAI methods

Identified limitations

Explanations often fail to capture dependencies between features or long-term patterns in the data

Different configurations of XAI methods yield non-comparable or contradictory results

Explanation reliability: small model changes can cause large shifts in outputs

A temporal XAI method should include

Temporal awareness, maintain sequence information

Consistency in explanations, stable and repeatable explanations

Human-centered explanations, explain with clarity for stakeholders

Towards tailored XAI for finance

Identify Time dependencies are broken when data is randomly perturbed. Generating synthetic samples that distort sequence structure creates inconsistent explanations

Improve A future XAI solution should address or avoid the weaknesses within current XAI methods

Identified limitations

Explanations often fail to capture dependencies between features or long-term patterns in the data

Different configurations of XAI methods yield non-comparable or contradictory results

Explanation reliability: small model changes can cause large shifts in outputs

A temporal XAI method should include

Temporal awareness, maintain sequence information

Consistency in explanations, stable and repeatable explanations

Human-centered explanations, explain with clarity for stakeholders

Towards tailored XAI for finance

Identify Time dependencies are broken when data is randomly perturbed. Generating synthetic samples that distort sequence structure creates inconsistent explanations

Improve A future XAI solution should address or avoid the weaknesses within current XAI methods

Identified limitations

- Explanations often fail to capture dependencies between features or long-term patterns in the data
- Different configurations of XAI methods yield non-comparable or contradictory results
- Explanation reliability: small model changes can cause large shifts in outputs

A temporal XAI method should include

- Temporal awareness, maintain sequence information
- Consistency in explanations, stable and repeatable explanations
- Human-centered explanations, explain with clarity for stakeholders

Towards tailored XAI for finance

Identify Time dependencies are broken when data is randomly perturbed. Generating synthetic samples that distort sequence structure creates inconsistent explanations

Improve A future XAI solution should address or avoid the weaknesses within current XAI methods

Identified limitations

- Explanations often fail to capture dependencies between features or long-term patterns in the data
- Different configurations of XAI methods yield non-comparable or contradictory results
- Explanation reliability: small model changes can cause large shifts in outputs

A temporal XAI method should include

- Temporal awareness, maintain sequence information
- Consistency in explanations, stable and repeatable explanations
- Human-centered explanations, explain with clarity for stakeholders

Towards tailored XAI for finance

Identify Time dependencies are broken when data is randomly perturbed. Generating synthetic samples that distort sequence structure creates inconsistent explanations

Improve A future XAI solution should address or avoid the weaknesses within current XAI methods

Identified limitations

- Explanations often fail to capture dependencies between features or long-term patterns in the data
- Different configurations of XAI methods yield non-comparable or contradictory results
- Explanation reliability: small model changes can cause large shifts in outputs

A temporal XAI method should include

- Temporal awareness, maintain sequence information
- Consistency in explanations, stable and repeatable explanations
- Human-centered explanations, explain with clarity for stakeholders

Towards tailored XAI for finance

Identify Time dependencies are broken when data is randomly perturbed. Generating synthetic samples that distort sequence structure creates inconsistent explanations

Improve A future XAI solution should address or avoid the weaknesses within current XAI methods

Identified limitations

- Explanations often fail to capture dependencies between features or long-term patterns in the data
- Different configurations of XAI methods yield non-comparable or contradictory results
- Explanation reliability: small model changes can cause large shifts in outputs

A temporal XAI method should include

- Temporal awareness, maintain sequence information
- Consistency in explanations, stable and repeatable explanations
- Human-centered explanations, explain with clarity for stakeholders

DIGITAL

Funded by the European Union

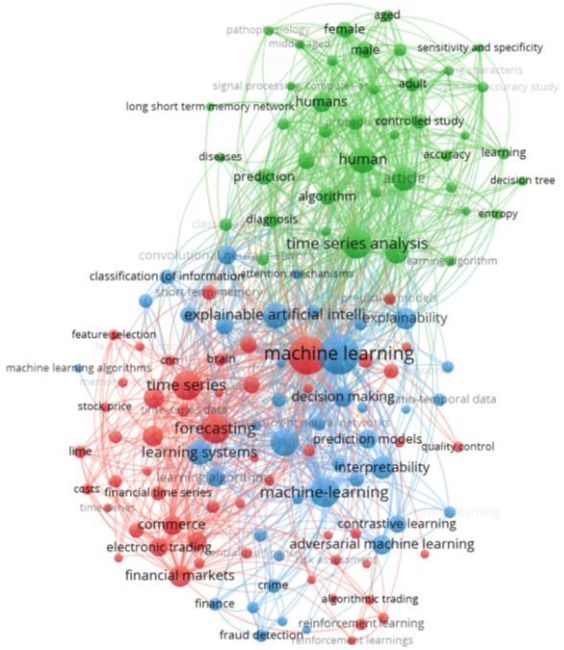
Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Horizon Europe: Marie Skłodowska-Curie Actions. Neither the European Union nor the granting authority can be held responsible for them.

This project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101119635

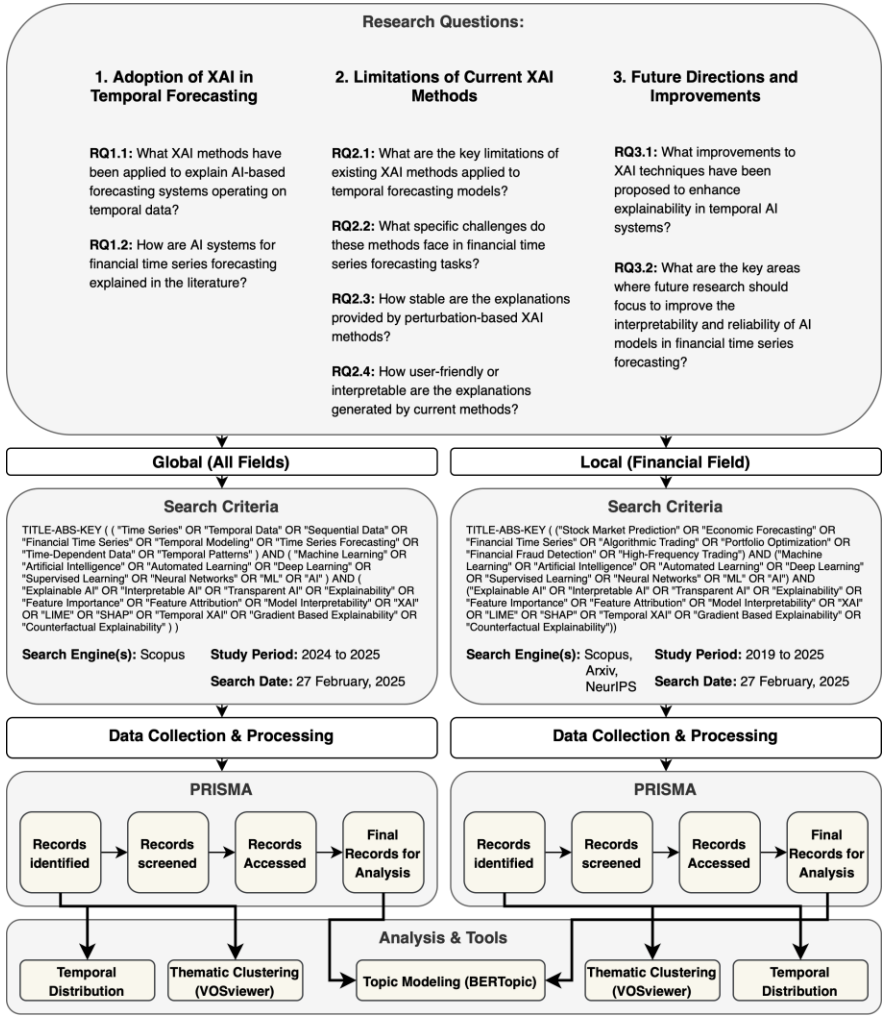
Appendix: SLR approach

SLR The total overview of the SLR is depicted on the right

RQs The research questions associated with the SLR are depicted and divided into three topics



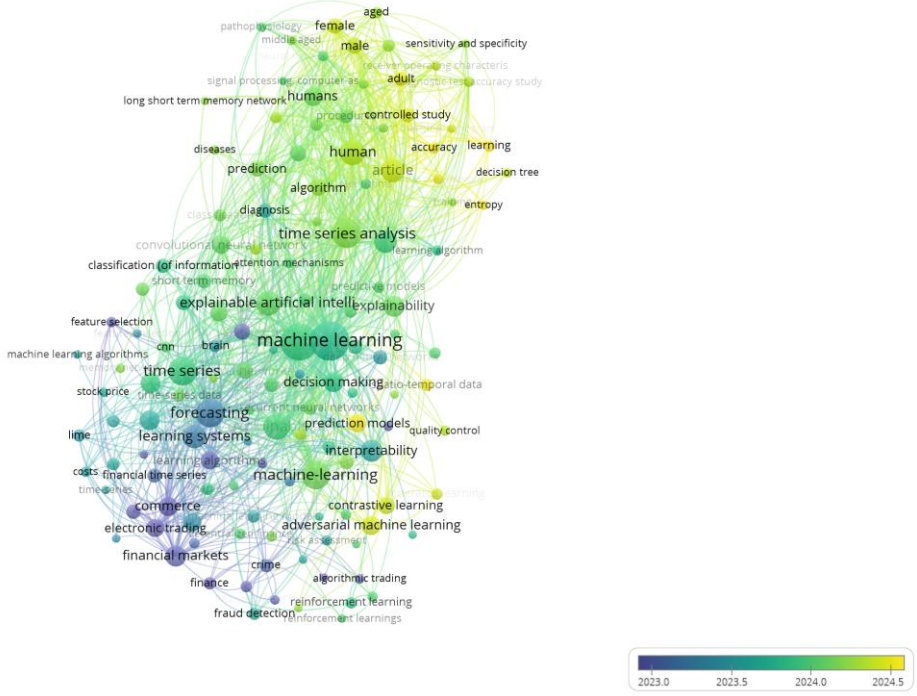
Network Visualization (VOSviewer)



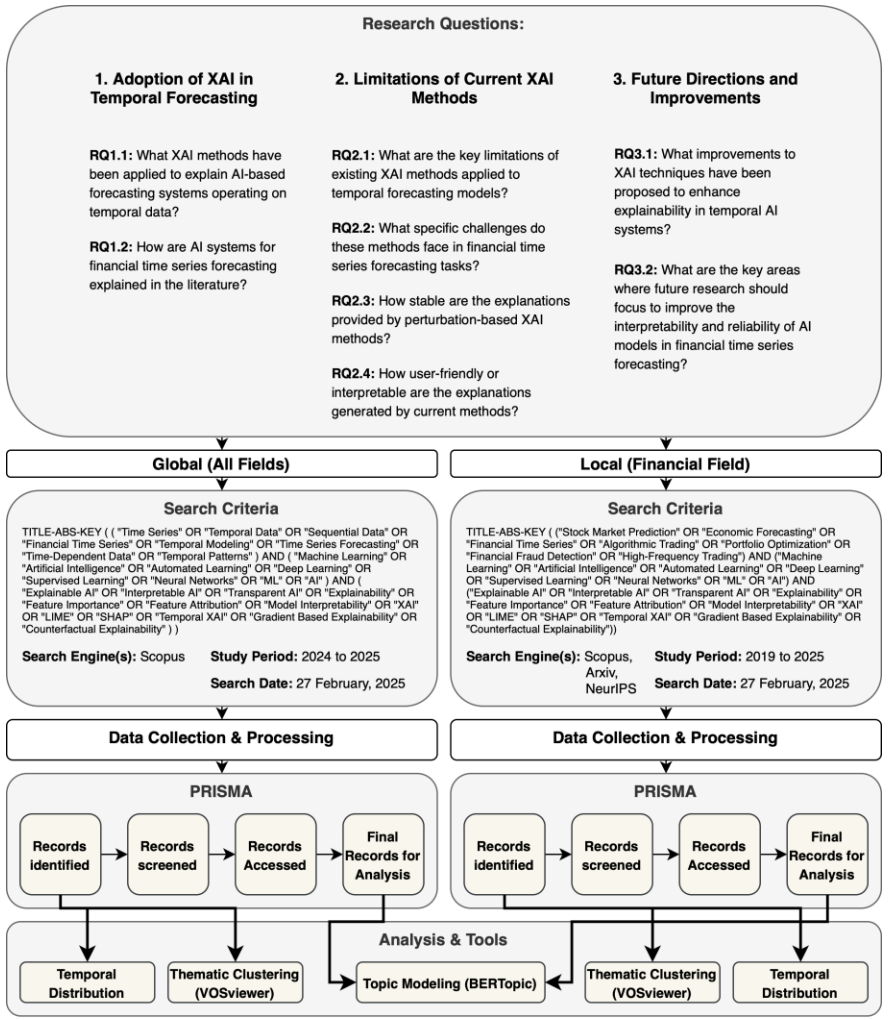
Appendix: SLR approach

SLR The total overview of the SLR is depicted on the right

RQs The research questions associated with the SLR are depicted and divided into three topics



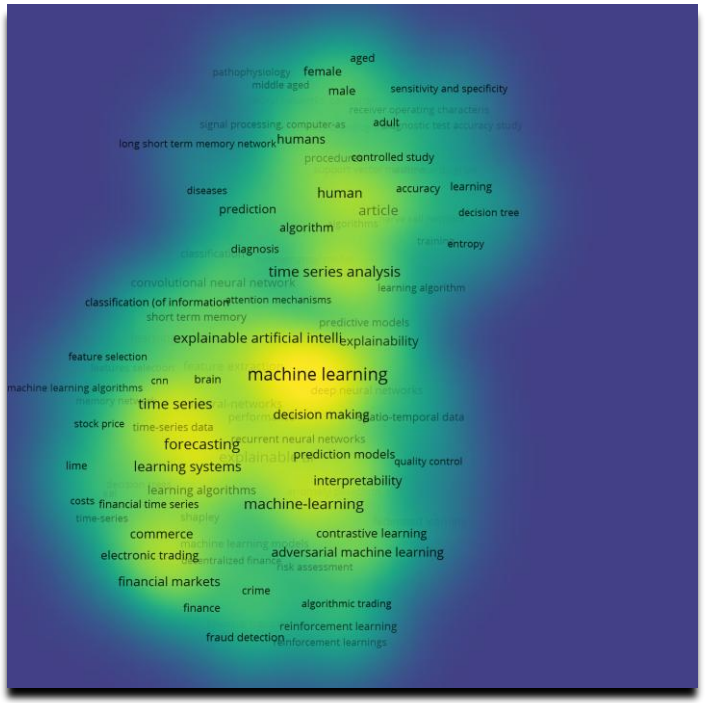
Overlay Visualization (VOSviewer)



Appendix: SLR approach

SLR The total overview of the SLR is depicted on the right

RQs The research questions associated with the SLR are depicted and divided into three topics



Density Visualization (VOSviewer)