

Review

Potential Applications of Explainable Artificial Intelligence to Actuarial Problems

Catalina Lozano-Murcia ^{1,2,†,‡}, Francisco P. Romero ^{1,*,†,‡} , Jesus Serrano-Guerrero ^{1,†,‡} , Arturo Peralta ^{3,4,‡} and Jose A. Olivas ^{1,†,‡}

¹ Department Information Systems and Technologies, University of Castilla La Mancha, Paseo de la Universidad, 4, 13071 Ciudad Real, Spain; catalina.lozano@alu.uclm.es (C.L.-M.); jesus.serrano@uclm.es (J.S.-G.); joseangel.olivas@uclm.es (J.A.O.)

² Department of Mathematics, Escuela Colombiana de Ingeniería Julio Garavito, Bogotá 111166, Colombia

³ Escuela Superior de Ingeniería y Tecnología, Universidad Internacional de La Rioja, Avda. de la Paz 93-103, 26006 Logroño, Spain; arturo.peralta@unir.net

⁴ Ciencia y Tecnología, Escuela Superior de Ingeniería, Universidad Internacional de Valencia, Calle Pintor Sorolla, 21, 46002 Valencia, Spain

* Correspondence: franciscop.romero@uclm.es

† Current address: School of Computer Science, University of Castilla la Mancha, Paseo de la Universidad, 4, 13071 Ciudad Real, Spain.

‡ These authors contributed equally to this work.

Abstract: Explainable artificial intelligence (XAI) is a group of techniques and evaluations that allows users to understand artificial intelligence knowledge and increase the reliability of the results produced using artificial intelligence. XAI can assist actuaries in achieving better estimations and decisions. This study reviews the current literature to summarize XAI in common actuarial problems. We proposed a research process based on understanding the type of AI used in actuarial practice in the financial industry and insurance pricing and then researched XAI implementation. This study systematically reviews the literature on the need for implementation options and the current use of explanatory artificial intelligence (XAI) techniques for actuarial problems. The study begins with a contextual introduction outlining the use of artificial intelligence techniques and their potential limitations, followed by the definition of the search equations used in the research process, the analysis of the results, and the identification of the main potential fields for exploitation in actuarial problems, as well as pointers for potential future work in this area.

Keywords: machine learning; artificial intelligence; deep learning; explainable machine learning; accuracy; interpretability

MSC: 62R07



Citation: Lozano-Murcia, C.; Romero, F.P.; Serrano-Guerrero, J.; Peralta, A.; Olivas, J.A. Potential Applications of Explainable Artificial Intelligence to Actuarial Problems. *Mathematics* **2024**, *12*, 635. <https://doi.org/10.3390/math12050635>

Academic Editors: Eric Ulm and Budhi Surya

Received: 2 January 2024

Revised: 2 February 2024

Accepted: 5 February 2024

Published: 21 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Data science has been developing over the last century due to improvements in processing capacity (hardware) and the general interest from different industries in accelerating the exploitation of the potential of the tools developed to benefit their processes.

Notably, in the last 20 years, we have seen a growing volume of research being applied to industries susceptible to data, such as the financial industry, where the knowledge that can be extracted from the combination of machine learning techniques at different stages of the process, such as product generation, marketing, customer attraction and retention, asset management, technical liability management, and, in general, joint risk management. This represents unimaginable potential for the industry.

Actuarial science focuses on solving problems related to risk management and their economic quantification, especially complex financial problems. It generally seeks to understand, measure, and manage current risk sharing and potential future sharing scenarios,

considering their potential impact on the development of the analysis environment. Specific risks are commonly modeled, such as credit risk (default), market risk (price variation and reference rates), technical life risks (death or survival), morbidity risk (disability and health), non-life risks (automobile, home, casualty, etc.), among other more specific risks, and their potential economic impact.

This means that modeling will always be very challenging in terms of the efficiency and pertinence of the predictions, with limitations related to information availability, regulatory restrictions, and the need for transparency in a highly regulated sector due to the capture of public resources and the systemic risk industry for the economy, which emphasizes the need to make informed and transparent decisions for clients and investors in a simple and clear manner. Within this framework of potential, professions such as the actuarial profession have been focusing on taking advantage of these sophisticated models to develop their primary function, which, regardless of the sector in which they operate, is the informed, timely, and adequate management of risks.

However, the processes of information generation through the use of data science techniques have also led to general difficulties within the implementation processes in productive sectors due to some generalized limitations [1]:

- Models with results that are difficult to explain and interpret;
- Complex models like machine-learning techniques make it difficult to extract general patterns or rules and generate models that are difficult to handle and replicate;
- In many cases, models that are impossible to replicate with accurate results hinder the audit and review processes, which are essential in industries regulated by the collection of public funds, such as banking and insurance;
- The generation of biases not allowed by law, such as positive or negative discrimination in pricing or surcharges, which are not permitted in some jurisdictions;
- Communicating the processes, groupings, and rules generated using the models to a non-technically experienced public is difficult, preventing their implementation due to the lack of a guarantee that decision makers will understand.

All these limitations are targeted via the processes and methods developed to understand artificial intelligence (XAI) models [2]. These techniques, which are increasingly becoming part of the model governance process, allow us to understand the relationships established by the models and thus address the explainability processes for decision making, auditing algorithms and models at a lower cost, understanding and reducing complex models, and feature selection, among others. Given all these potential advantages, we believe that XAI [3] can be a valuable approach to managing the implementation of this model in business practice and, therefore, in the day-to-day life of actuaries beyond the academic approach with which machine-learning techniques have been used so far.

A systematic literature review (SLR) has been carried out to cover the advances in the applications of XAI in actuarial science in recent years, the main sectors of application, and the advances in the processes of its incorporation in decision making. In conclusion, great practical potential is identified in the use of XAI techniques in different processes, including its current use, such as the strengthening of model governance, the simplification or selection of characteristics, facilitating the implementation of black-box models by allowing an understanding of the established relationships and the interpretation of the results and also offering tools for the development of audit processes to IA models, accompanied by challenges in its implementation, such as the lack of knowledge of the techniques that complement the key processes for estimation, computational capacity, regulatory issues, and generalization in the use of these techniques.

The rest of the study is structured as follows: Section 2 presents the background on AI and its challenges in applications in actuarial problems, as well as the definition of XAI, in the form of a related literature review; Section 3 describes the research method followed to perform the SLR, including the search equations and how we have classified the different papers selected; Section 4 presents the results of the SLR; Section 5 indicates the main identified contributions to specific actuarial problems; Section 6 explains the threats to the

validity of these methods and presents the discussion; and Section 7 describes the main conclusions and future work.

2. Related Research

AI can be helpful in risk modeling, mortality forecasting, claim calculation, and other actuary-related topics [4]. However, these models often encounter specific difficulties regarding the transparency and interpretation of the results; some involve the so-called black boxes, which do not allow a clear understanding of the relationships established and the replication and auditing processes. In AI, there are techniques categorized as explainable machine learning (XAI), which can help actuaries understand how to make decisions and forecasts and fully communicate results from their sources. This can be especially useful when making vital decisions based on the model's results, which is why this article has been developed as a contextual approach to the tools that can enhance efficient solutions to the risk management challenges that an actuary faces.

Before the systematic review in this document was undertaken, some works that served as a general framework to understand the methods [5] and their taxonomy [6] were identified. Another group of more specific research works refers to the applications of artificial intelligence (AI) for actuaries and their work sectors and, therefore, the need to use XAI techniques. In the banking domain, a systematic literature review is presented in this review [7], which identifies three key themes in AI applications in the banking sector: strategies, processes, and customers. It seeks to propose an AI banking services framework that bridges the gap between academic research and industry knowledge to formulate strategic decisions regarding the utilization and optimization of the value of artificial intelligence technologies in the banking sector. Moving into the specific area of an actuary, the work presented in [8] shows how actuarial science can adapt and evolve to incorporate AI techniques and methodologies in the coming years. Part 1 of this article includes benchmark AI techniques and deep learning, as well as potential applications, including examples of mortality modeling, claims reserves, non-life-insurance pricing, and telematics. Furthermore, another critical contribution to the literature is shown in [9]. This report presents a framework for actuaries by highlighting a literature review of how AI can be used in different lines of actuarial work and its impact on the profession.

On the other hand, although they do not entirely cover the focus of an actuary's work, some reviews develop reference approaches such as a SLR in a specific area of actuarial work. For example, the work presented in [10] analyzes the different applications of AI in the insurance industry, identifying and classifying them according to their levels of explainability and contributions at the different stages of the insurance cycle. In this sense, it identifies how XAI methods predominate in the claim, underwriting, and pricing stages, contributing to the simplification of models and extracting relationships or rules to understand the established relationships. Furthermore, the SLR presented in [11] is an overview of the current state of the art of XAI applications in the financial sector, classifying them by their explicability objective and thus extracting the main methods employed.

Although these publications generally cover XAI alternatives for specific problems in actuarial science, if we also consider publications from recent years, then there is a practical approach whereby the reliability of the sources is considered throughout an organization in terms of the classification of the results and their application to different lines of work within actuarial science; this approach is proposed in this article.

In systematic reviews of XAI applied to other industries, we can find diverse application approaches, such as XAI methods and evaluation metrics related to different application domains and tasks focused on AI/ML applications and deep learning, concluding that more attention is required to generate explanations for general users in sensitive domains such as finance and the judicial system [12,13]. For example, the work [14] provides an overview of the trends in XAI and answers the question of accuracy versus explainability, considering the extent of human involvement and explanation assessment. Finally, one promising XAI method, XAIR, is presented in [15], and the study presented in [16]

reviews the possibilities and potential of XAI applications by referring to publications of its applications in medicine between 2018 and 2022.

3. Planning of the SLR

Our SLR considered two steps: machine-learning applications and the explainability of machine-learning techniques. This section sets out the research questions guiding the search, the search strings used, the exclusion and inclusion criteria chosen to select the publications, and the process of selecting them, including an explanation of how the inclusion and exclusion criteria were applied.

3.1. Research Questions

The general questions posed to achieve our objectives were based on determining the type of machine-learning techniques used in insurance pricing problems, the need to explain the established relationships, and what type of explainability techniques have been implemented for pricing or actuarial problems.

The main question for the ML search was as follows:

- Q1.1. What are the machine-learning techniques used in insurance pricing?

Then, seeking to segment the analysis considering the primary nature of insurance risk, the questions we posed were as follows:

- Q1.2. Which technique would be most appropriate for modeling general insurance risks?
- Q1.3. Which techniques would be most suitable for modeling life insurance risks?

This question aims to identify techniques and compare the results in the main lines of business.

Once the application of difficult-to-trace techniques or the understanding of techniques has been identified, the second main search equation is proposed:

- Q2.1. What are the most common techniques for performing the explainability of machine-learning models in the financial sector?

This question refers broadly to applications in the financial sector. To specify the original problem raised, the complementary equations are defined:

- Q2.2. What are the most common techniques that allow the explicability of machine-learning models in actuarial problems?
- Q2.3. What are the most common techniques to explain machine-learning models in issues in the insurance sector?

This question seeks to broaden the identification of the XAI techniques that are potentially applicable to ML models in insurance pricing.

3.2. Search String

According to the two general research questions, the most appropriate keywords were “machine learning insurance” and “explainable artificial intelligence”.

The search strings were formed on these bases by concatenating the main keywords and other relevant words related to the specific questions, such as machine-learning insurance AND “non-life”, “XAI techniques”, AND insurance. The quotation marks are included so that the search contains exactly those expressions.

The determination of the equations (see Table 1) is the product of different tests, where the need to specify expressions such as non-life with “—” and quotation marks was verified.

Additionally, searching XAI techniques using the acronym contributed as much as searching the complete expression. It is worth mentioning that the equations were not applied using the OR expression for a generalized search but were considered in independent searches to facilitate traceability in the result groups.

Table 1. Research string.

ID	Question	Equation
Q1	Machine learning	
Q1.1		Machine learning insurance pricing
Q1.2		Machine learning AND “non-life insurance” pricing
Q1.3		Machine learning AND “life insurance pricing”
Q2	XAI techniques	
Q2.1		“eXplainable Artificial Intelligence” AND financial models
Q2.2		“eXplainable Artificial Intelligence” AND actuarial models
Q2.3		“XAI techniques” AND insurance

In addition, the search strings were applied in the arXiv platform to show the contrast in results with the most recently published works; this meant that some characteristics of the equations were validated to be adapted to each website, identifying non-empty searches on each website for each equation in different iterations, which led to the definition of the specific questions, especially in the case of XAI, and to changes in the structure of some equations, such as the use of “-” in “non-life”. As additional considerations in the searches, the following were included in arXiv: the search was performed twice, searching first in the Quantitative Finance category and then in Computer Science, and the search equation had to appear in the title or abstract of review articles or research articles for them to initially be considered. On the other hand, different filters were applied to refine the searches on each platform, as detailed below.

3.3. Inclusion/Exclusion Criteria

The stated inclusion criteria were as follows:

- The publication was either a journal paper, conference paper, review article, or research article;
- The publication dealt with ML or XAI topics in finance, insurance, or actuarial context applications;
- It was published or updated between 2019 and 2023.

The exclusion criteria were as follows:

- The abstract content was related to the search equation applied to finance;
- Results whose content was unrelated to the application environment were excluded from the specific equations.

3.4. Study Classification

Once the search string and criteria were available, the selection process is shown in Figure 1: the abstract, a quick reading of the papers, and, finally, a detailed reading of the candidate papers. As shown in Figure 1, the search chain returned 510 documents, for which the abstract, title, and keywords were assessed by applying the inclusion and exclusion criteria. After the first step, 210 papers were left; papers were removed because they did not address the challenges of machine learning or the target XAI applications. After an overall reading, mainly focused on the conclusions, 110 published papers remained. Finally, we carried out a closer reading of every publication in a third iteration, in which we eliminated all the papers that did not specifically deal with every research string. Figures 1 and 2 show the principal results, including 44 publications.

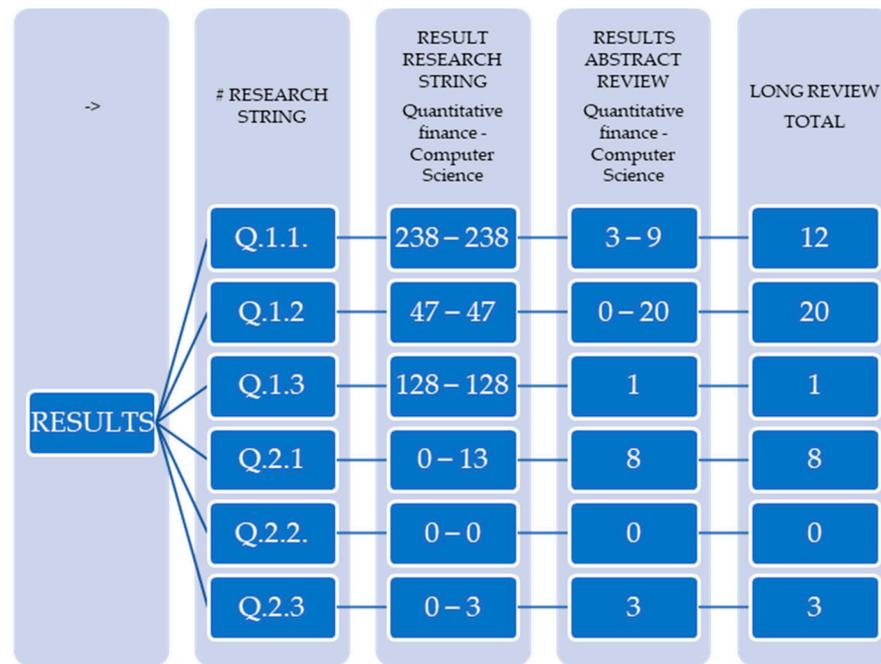


Figure 1. General selection process.

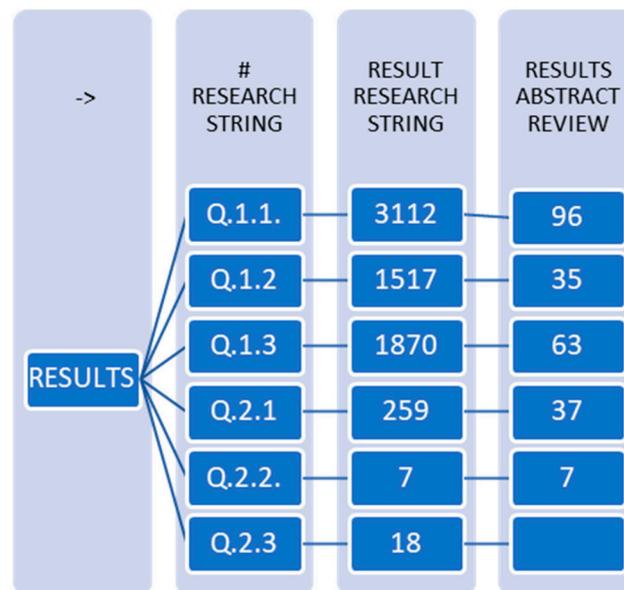


Figure 2. Selection process in arXiv.

4. Results and Findings

Numerous publications have been found regarding machine-learning applications for pricing, with the number growing in recent years. We did not find a correct distinction between life and non-life using the two search engines. However, in the selection process, more applications were identified for non-life—health than life; a significant bias was the challenge to health risk management in times of COVID-19 and the manner of classifying this type of insurance.

From the range of publications found (see Table 2), the final filter yielded 82 articles, concentrated on the first two questions and the first XAI question, with the distribution presented in Table 3.

Table 2. Research by question—initial selection of articles.

ID Question	ArXiv	Science Direct
Q1.1	476	3.112
Q1.2	94	1.517
Q1.3	256	1.870
Q2.1	13	259
Q2.2	0	7
Q2.3	3	18
Total	842	6.783

Table 3. Research by question—final selection of articles.

ID Question	Selected Articles
Q1.1	22
Q1.2	25
Q1.3	4
Q2.1	18
Q2.2	3
Q2.3	10
Total	82

Additionally, the distribution of the qualification of the reviews where these articles were published was reviewed (see Table 4), finding a majority concentration (42%) of ungraded publications, many of these being conferences or documents in qualification processes, followed by Q1 journals (35%) and Q2 (22%), with more than half of the publications being in highly rated journals. This result, together with the initial volume of publications, reflects the relevance of this topic within the industry and academia.

Table 4. Distribution of publication types by journal quality.

Journal Quality	Participation
Q1	35.1%
Q2	22.1%
Q3	0.0%
Q4	1.3%
Na	41.6%

Q1.1: Problems related to rate balancing, classification for group rate generation, and prediction are faced. The central axis includes the following models: GLM [17], adaptive [18], adaptive Gaussian models [18], logistic regression [19], auto-encoder LSTM [20], neural networks [21], boosting [22], SMuRF [19], TabNet DL [23], XGBoost [22], as well as the implementation of protocols such as the sum-product network (SPN) [24] or the development of processes for data pre-processing.

Q1.2: Regarding the pricing of non-life products, the problems focus on model simplification through feature selection, data cleaning, and the extraction of outliers, along with techniques to improve prediction capacity, such as RNN and SHAP [25,26], isotonic recalibration [27], tree-based ensemble [28], Hierarchical Risk-factors Adaptive Top-down (PHiRAT) [29], logistic regression, decision tree, random forest, XGBoost, feed-forward network [30], transaction models for in life IBNR, inconclusive [31], integration of graphic themes [32], and extreme event estimation [33]. Moreover, currently, more efficient prediction models have been developed with techniques such as extreme gradient boosting or XGBoost [34], Bayesian CART models [35], boosting [36], and deep neural networks [36–38], among others.

Q1.3: There is not a significant number of publications on machine-learning application models in the final selection, mainly because, despite initially being the category with

the most significant number of searches, they presented studies related to treating or diagnosing diseases and not specifically related to life insurance pricing. Regarding the articles reviewed, multi-modal models of the GLM type or Bayesian networks are applied to the prediction of longevity risks [39–42].

Q2.1: The applications refer to the need to guarantee transparency, equity, and accountability. On the one hand, it is essential to emphasize the relevance of transparent regulatory frameworks for artificial intelligence, differentiating between the explicability requirements of AI models themselves and the broader explicability obligations of AI systems under existing laws and regulations [43]. On the other hand, there are some challenges to address: for example, the definition of appropriate assessment methods for the banking sector, especially for fraud detection [34,35], and the interpretation of complex models like deep learning for various applications, including equity analysis and financial distress prediction [44–48]. Other challenges include reducing biases in research judgments [40] and emphasizing the role of cybersecurity in maintaining the integrity of AI systems [49].

Q2.2: Only three reference articles were obtained using Question 2.2, so its analysis is grouped with Question 2.3 as a related topic. In general, results are obtained with techniques [50–52] and frameworks of predictive processes with the best capacity explained using XAI techniques [53,54], auditing processes [55], classification problems [56], applications to the non-life automobile business [57], and a relevant spotlight on health risk assessment [58,59].

In Table 5, we wanted to show how 80% of the publications found in this question have been published in Q1 journals and have an average level of citations above 50, with an average H-index value for the primary author close to 15. The above reflects the boom of formal research developed on this topic for the sector. Although we do not have many specific publications, the related ones present a significant scientific endorsement. Finally, some establish the limitations related to the XAI technique selection and comparison process [60] or [61]. From the general search, three principal axes of XAI systems can be established, as shown in Figure 3.

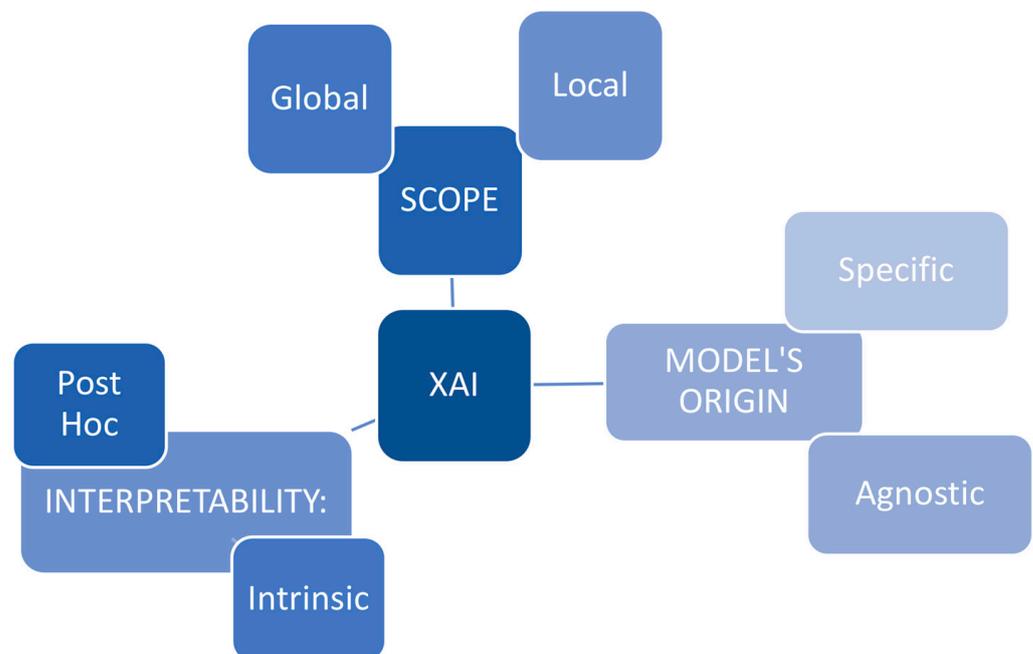


Figure 3. XAI axes.

Table 5. Types of models by axes.

Reference	Paper		Journal		
	Principal Author H-Index	Citations	H-Index	SJR	Quantile
[52]	6	4119	136	4.15	Q1
[60]	22	145	226	2.46	Q1
[51]	8	60	114	2.1	Q1
[61]	37	0	210	2.28	Q1
[59]	38	268	136	4.15	Q1
[53]	2	0	171	1.76	Q1
[55]	1	23	60	1.15	Q1
[62]	18	69	136	4.15	Q1

Within the three axes presented, a group of techniques is identified (see Table 6) within the review, which, according to their nature, are applied to different problems or stages in the actuarial modeling process.

Table 6. Types of models by axis.

Techniques	Interpretability		Destiny		Origin	
	Intrinsic	Post Hoc	Global	Local	Specific	Agnostic
Counterfactual explanations		X		X		X
Decision trees	X		X		X	
Feature importance		X	X	X		X
Lime		X		X		X
Partial dependence plot		X	X	X		X
Rule extraction		X	X			X
Sensitive analysis		X	X	X		X
Shape		X		X		X
Shapely explanations		X		X		X
Surrogate models		X	X	X		X

5. Contributions of XAI in the Actuarial Context

Four key areas of machine-learning applications were found with greater intensity in non-life and health risks. They are presented in the order of the intensity of use of these techniques and the limitations identified in the business use of this type of technique:

1. Marketing, product design, and commercialization: They are commonly used since there are no specific regulations regarding the use of AI in these processes. However, explaining them to management and decision makers’ understanding of the intuition of the results may limit the business use of this type of technique. XAI applications would complement the processes of understanding relationships, strengthening process development [2], and translating results into a common language [44];
2. Risk management (ALM, credit, liquidity, cybersecurity, etc.): Comprehensive risk management involves the development of complex models, which make it possible to identify trade-offs or objectives through the interaction of many variables, which makes AI techniques very useful for risk identification, prediction, management, and decision making. However, these techniques will have hardware limitations related to the processing capacity of large volumes of information and transformation through algorithms; many of these processes require the assurance of internal and external auditing processes, as well as supervisory approval; decision makers must understand what happens within the models, the risks around the implementation and interpretation of results, including the incorporation of prescriptive models. XAI techniques would facilitate audit processes as an alternative to model replication [63] for the management of specific risks, such as cybersecurity or fraud, and interpreting established relationships;

3. Prediction for pricing technical risks considering internal and external surcharges: This is one of the aspects where combining AI techniques would be most helpful, from data processing and cleaning, including risk and expense classification and prediction models. However, this application has extensive regulations, including requirements to not discriminate based on gender or health [21], such as the law of oncological oblivion, compliance with transparency in risk modeling, and the auditability and traceability of these processes, even without the consideration of the limitations of processing capacity and speed, which is a necessity within the efficiency of model governance. XAI techniques allow model debugging—which aims to reduce the number of variables irrelevant to the algorithms [19] and auditing processes—and the traceability, understanding, and general strengthening of model governance;
4. Prediction for reserving (pure prospective technical risk): This raises the same potentialities and difficulties as the pricing processes; however, the regulations are much stricter for this process, in addition to there being a multiplicity of reports and requirements or measurement conditions that work simultaneously (Solvency II, IFRS 17, local GAP) and complexify the financial reporting processes for the companies. This is compounded by the need to analyze variations and justify the variation in the results for the best estimate liabilities (BELs) that must accompany the different accounting and financial reports.

In addition to the already-mentioned advantages of implementing XAI techniques, their application to reserve processes would facilitate our understanding of the origins of the variation in the results from models of different types, but with the same purpose, which is required to contrast the different regulatory reports. It would also facilitate the identification of key variables in BEL processes and the generation of impacts due to relevant stresses. Finally, it gives rise to the generation of regulations that include the use of AI techniques in this type of process and guidelines or standards for regulators for the development of norms [43] and supervisory standards, allowing the use of more robust and comprehensive, and potentially better, predictive models.

6. Threats to Validity and Limitations

Although these types of publications are very useful for understanding the context and level of work related to a specific topic, it is essential to highlight some possible limitations of the present publication.

On the one hand, there could be a potential publication bias. As noted, most of the publications found do not stand out as belonging to qualified journals; some even belong to sponsored publications with commercial objectives. This may lead to an overestimation of the actual effects and affect the validity of the review.

On the other hand, if the studies included do not adequately represent the target population, this may limit the generalizability of the results. The search was not exhaustive; it only included publications from the last few years and two article search engines. It was not performed in an actuarial niche with a more academic approach, but it was considered the best approach for addressing this topic.

Regarding general limitations that we have identified, there are several related to the regulatory requirements in the financial and insurance sector, where even models with good prediction or classification capacity cannot be used due to the nature of the variables and the potential relationships between discriminatory variables, among other factors. Another limitation may be related to the computational requirements, both in terms of processing capacity and time, and, finally, the implementation of these methods in the industry beyond scientific practice that, given the previous limitations, generates resistance against commercial use.

7. Conclusions and Future Challenges

Publications related to machine learning and XAI techniques have grown significantly in recent years, becoming a trending topic in research in the financial sector. XAI has proven

to be a valuable tool in this sector by providing clear explanations of how decisions are made and offering enhanced transparency and reliability, which are essential in increasing the confidence of users and regulators regarding the use of AI in estimation and modeling processes in the financial sector.

XAI techniques are also presented as tools for improving risk management, helping institutions better understand and mitigate the risks associated with implementing machine learning models. This understanding is fundamental in several fields: fraud detection, investment portfolio management, insurance pricing, and reserving. In general, the applications are mainly focused on overcoming problems around prediction capacity, data preparation, and classification, where they are complemented, either in the analysis or in the model review processes, using XAI techniques.

XAI is playing a growing and crucial role in the financial sector by improving artificial intelligence systems' transparency, reliability, and efficiency. However, it is critical to address ethical, privacy, and regulatory challenges while continuing to advance its development and applications in the financial industry.

There is an uncertainty that has not yet been addressed regarding the challenge posed by using and comparing XAI techniques. However, they are considered a relevant part of model governance and review processes. They are also identified as an alternative to consider from the point of view of audit and regulation processes due to the growing need to accept and regulate the use of AI and ML in the financial sector.

This study presents interesting results regarding the potential applications of XAI in actuarial problems and demonstrates that, in the future, it would be pertinent to continue researching and developing more advanced and accurate XAI models for the financial sector. Then, the development of techniques that facilitate a more transparent and detailed explanation of decisions made using machine learning models is implied, along with protocols for selecting and comparing these techniques.

Another relevant challenge pertains to creating robust ethical frameworks and implementing safeguards to protect sensitive customer data, which is directly linked with adapting regulations and compliance standards in the financial sector. Such an endeavor necessitates integrating these techniques into monitoring and auditing processes while ensuring that XAI systems comply with regulatory requirements and align with best practices in security.

Author Contributions: Conceptualization F.P.R. and C.L.-M.; methodology: C.L.-M.; validation: F.P.R. and J.S.-G.; writing: C.L.-M. and F.P.R.; visualization: A.P.; supervision: J.A.O.; project administration: J.A.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially supported by FEDER and the State Research Agency (AEI) 590 of the Spanish Ministry of Economy and Competition under grant SAFER: PID2019-104735RB-C42 591 (AEI/FEDER, UE).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bahalul, N.I.P.M. Haque Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technol. Forecast. Soc. Change* **2023**, *182*, 122120.
2. Langer, M.; Oster, D.; Speith, T.; Hermanns, H.; Kästner, L.; Schmidt, E.; Sesing, A.; Baum, K. What do we want from Explainable Artificial Intelligence (XAI)?—A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artif. Intell.* **2021**, *296*, 103473. [[CrossRef](#)]
3. Scholbeck, C.A.; Molnar, C.; Heumann, C.; Bischl, B. Sampling, Intervention, Prediction, Aggregation: A Generalized Framework for Model Agnostic Interpretations. In Proceedings of the International Workshops of ECML PKDD 2019, Würzburg, Germany, 16–20 September 2019.
4. Richman, R. AI in actuarial science—A review of recent advances part 1. *Ann. Actuar. Sci.* **2021**, *15*, 207–229. [[CrossRef](#)]
5. Minh, D.; Wang, X.; Li, F.; Nguyen, T. Explainable artificial intelligence: A comprehensive review. *Artif. Intell. Rev.* **2021**, *55*, 3503–3568. [[CrossRef](#)]
6. Speith, T. A Review of Taxonomies of Explainable Artificial Intelligence (XAI) Methods. In Proceedings of the FAccT '22: Actas de la Conferencia ACM de 2022 Sobre Equidad, Responsabilidad y Transparencia, Seoul, Republic of Korea, 21–24 June 2022.

7. Omar, F.; Irfan, B.; Seung, M.L. Hwan Utilization of artificial intelligence in the banking sector: A systematic literature review. *J. Financ. Serv. Mark.* **2022**, *28*, 835–852.
8. Richman, R. AI in actuarial science—A review of recent advances—Part 2. *Ann. Actuar. Sci.* **2019**, *15*, 230–258. [[CrossRef](#)]
9. Yeo, N.; Lai, R.; Ooi, M.J.; Liew, J.Y. Literature Review: Artificial Intelligence. 12 2019. Available online: <https://www.soa.org/globalassets/assets/files/resources/research-report/2019/ai-actuarial-work.pdf> (accessed on 15 November 2023).
10. Owens, E.; Sheehan, B.; Mullins, M.; Cunneen, M.; Ressel, J.; Castignani, G. Explainable Artificial Intelligence (XAI) in Insurance. *Risks* **2022**, *10*, 230. [[CrossRef](#)]
11. Weber, P.; Carl, V.; Hinz, O. Applications of Explainable Artificial Intelligence in Finance—A systematic review of Finance, Information Systems, and Computer Science literature. *Manag. Rev. Q.* **2023**, *2023*, 1–41. [[CrossRef](#)]
12. Roussel, C.; Böhm, K. Geospatial XAI: A Review. *Int. J. Geo-Inf.* **2023**, *12*, 355. [[CrossRef](#)]
13. Le, T.-T.; Prihatno, A.T.; Oktian, Y.E.; Kang, H.; Kim, H. Exploring Local Explanation of Practical Industrial AI Applications: A Systematic Literature Review. *Appl. Sci.* **2023**, *13*, 5809. [[CrossRef](#)]
14. Nor, A.K.M.; Pedapati, S.R.; Muhammad, M.; Leiva, V. Overview of Explainable Artificial Intelligence for Prognostic and Health Management of Industrial Assets Based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses. *Sensors* **2021**, *21*, 8020. [[CrossRef](#)] [[PubMed](#)]
15. Clement, T.; Kemmerzell, N.; Abdelaal, M.; Amberg, M. XAIR: A Systematic Metareview of Explainable AI (XAI) Aligned to the Software Development Process. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 78–108. [[CrossRef](#)]
16. Ali, S.; Akhlaq, F.; Ali, S.I.; Kastrati, Z.; Muhammad, S.M.M. Daudpota The enlightening role of explainable artificial intelligence in medical & healthcare domains: A systematic literature review. *Comput. Biol. Med.* **2023**, *166*, 107555.
17. Li, H.-J.; Luo, X.-G.; Zhang, Z.-L.; Jiang, W.; Huang, S.-W. Driving risk prevention in usage-based insurance services based on interpretable machine learning and telematics data. *Decis. Support Syst.* **2023**, *172*, 113985. [[CrossRef](#)]
18. Yuqing, Z.; Neil, W. Adaptive Pricing in Insurance: Generalized Linear Models and Gaussian Process Regression Approaches. *arXiv* **2019**, arXiv:1907.05381.
19. Devriendt, S.; Antonio, K.; Reynkens, T.; Verbelen, R. Sparse regression with Multi-type Regularized Feature modeling. *Insur. Math. Econ.* **2021**, *96*, 248–261. [[CrossRef](#)]
20. Mayaki, M.Z.A.; Riveill, M. Multiple Inputs Neural Networks for Medicare fraud Detection. *arXiv* **2022**, arXiv:2203.05842.
21. Lindholm, M.; Richman, R.; Tsanakas, A.; Wüthrich, M. A Discussion of Discrimination and Fairness in Insurance Pricing. *arXiv* **2022**, arXiv:2209.00858. [[CrossRef](#)]
22. Henckaerts, R.; Côté, M.-P.; Antonio, K.; Verbelen, R. Boosting insights in insurance tariff plans with tree-based machine learning methods. *arXiv* **2019**, arXiv:1904.10890. [[CrossRef](#)]
23. McDonnell, K.; Murphy, F.; Sheehan, B.; Masello, L.; Castignani, G. Deep learning in insurance: Accuracy and model interpretability using TabNet. *Expert Syst. Appl.* **2023**, *217*, 119543. [[CrossRef](#)]
24. Varley, M.; Belle, V. Fairness in Machine Learning with Tractable Models. *Knowl.-Based Syst.* **2021**, *215*, 106715. [[CrossRef](#)]
25. Matthews, S.; Hartman, B. mSHAP: SHAP Values for Two-Part Models. *Risks* **2022**, *10*, 3. [[CrossRef](#)]
26. Yevhen, H.; Julia, H. Detection of Interacting Variables for Generalized Linear Models via Neural Networks. *Actuar. J.* **2023**, *30*, 1–30. [[CrossRef](#)]
27. Wüthrich, M.V.; Ziegel, J. Isotonic Recalibration under a Low Signal-to-Noise Ratio. Available online: <https://www.tandfonline.com/action/showCitFormats?doi=10.1080/03461238.2023.2246743> (accessed on 15 November 2023).
28. Terefe, E.M. Tree-Based Machine Learning Methods for Vehicle Insurance Claims Size Prediction. *arXiv* **2023**, arXiv:2302.10612.
29. Bravo, C.; Katrien, A. On clustering levels of a hierarchical categorical risk factor. *arXiv* **2023**, arXiv:2304.09046.
30. Baran, S.; Rola, P. Prediction of motor insurance claims occurrence as an imbalanced machine learning problem. *arXiv* **2022**, arXiv:2204.06109.
31. Buchardt, K.; Furrer, C.; Sandqvist, O.L. Transaction time models in multi-state life insurance. *Scand. Actuar. J.* **2023**, *2023*, 974–999. [[CrossRef](#)]
32. Blier-Wong, C.; Baillargeon, J.-T.; Cossette, H.; Lamontagne, L.; Marceau, E. Rethinking Representations in P&C Actuarial Science with Deep Neural Networks. *arXiv* **2021**, arXiv:2102.05784.
33. Bai, Y.; Lam, H.; Zhang, X. A Distributionally Robust Optimization Framework for Extreme Event Estimation. *arXiv* **2023**, arXiv:2301.01360.
34. Verschuren, R.M. Customer Price Sensitivities in Competitive Automobile Insurance Markets. *Expert Syst. Appl.* **2022**, *202*, 117133. [[CrossRef](#)]
35. Zhang, Y.; Ji, L.; Aivaliotis, G.; Taylor, C. Bayesian CART models for insurance claims frequency. *Insur. Math. Econ.* **2023**, *114*, 108–131. [[CrossRef](#)]
36. Kuo, K. DeepTriangle: A Deep Learning Approach to Loss Reserving. *Risks* **2019**, *7*, 97. [[CrossRef](#)]
37. Frey, R.; Köck, V. Deep Neural Network Algorithms for Parabolic PIDEs and Applications in Insurance Mathematics. *Computation* **2022**, *10*, 272–277. [[CrossRef](#)]
38. Jin, Z.; Yang, H.; Yin, G. A hybrid deep learning method for optimal insurance strategies: Algorithms and convergence analysis. *Insur. Math. Econ.* **2021**, *96*, 262–275. [[CrossRef](#)]
39. Souto, L.; Cirillo, P. Joint and survivor annuity valuation with a bivariate reinforced urn process. *Insur. Math. Econ.* **2021**, *99*, 174–189. [[CrossRef](#)]

40. Blake, D.; Cairns, A.J. Longevity risk and capital markets: The 2019-20 update. *Insur. Math. Econ.* **2021**, *99*, 395–439. [[CrossRef](#)]
41. Bravo, J.M.; Ayuso, M.; Holzmann, R.; Palmer, E. Addressing the life expectancy gap in pension policy. *Insur. Math. Econ.* **2021**, *99*, 200–221. [[CrossRef](#)]
42. Albrecher, H.; Bladt, M.; Bladt, M.; Yslas, J. Mortality modeling and regression with matrix distributions. *Insur. Math. Econ.* **2022**, *107*, 68–87. [[CrossRef](#)]
43. Ouren, K.; Martin, B.; Joost, B.; Stefan, L. Exploring Explainable AI in the Financial Sector: Perspectives of Banks and Supervisory Authorities. In *Artificial Intelligence and Machine Learning, Proceedings of the 33rd Benelux Conference on Artificial Intelligence, BNAIC/Benelearn 2021, Esch-sur-Alzette, Luxembourg, 10–12 November 2021*; Springer: Cham, Switzerland, 2022; pp. 105–119.
44. Júnior, J.S.; Mendes, J.; Sousa, F.; Premebida, C. Survey on Deep Fuzzy Systems in regression applications: A view on interpretability. *Int. J. Fuzzy Syst.* **2023**, *25*, 2568–2589. [[CrossRef](#)]
45. Petersone, S.; Tan, A.; Allmendinger, R.; Roy, S.; Hales, J. A Data-Driven Framework for Identifying Investment Opportunities in Private Equity. *arXiv* **2022**, arXiv:2204.01852.
46. Zhang, Z.; Wu, C.; Qu, S.; Chen, X. An explainable artificial intelligence approach for financial distress prediction. *Inf. Process. Manag.* **2022**, *59*, 102988. [[CrossRef](#)]
47. Moscato, V.; Picariello, A.; Sperli, G. A benchmark of machine learning approaches for credit score prediction. *Expert Syst. Appl.* **2020**, *165*, 113986. [[CrossRef](#)]
48. Futagami, K.; Fukazawa, Y.; Kapoor, N.; Kito, T. Pairwise acquisition prediction with SHAP value interpretation. *J. Financ. Data Sci.* **2021**, *7*, 22–44. [[CrossRef](#)]
49. Zhang, Z.; Al Hamadi, H.; Damiani, E.; Yeun, C.Y.; Taher, F. Explainable Artificial Intelligence Applications in Cyber Security: State-of-the-Art in Research. *IEEE Access* **2022**, *10*, 93104–93139. [[CrossRef](#)]
50. Koster, O.; Kosman, R.; Visser, J. A Checklist for Explainable AI in the Insurance Domain. *arXiv* **2021**, arXiv:2107.14039.
51. Panigutti, C.; Perotti, A.; Panisson, A.; Bajardi, P.; Pedreschi, D. FairLens: Auditing Black-Box Clinical Decision Support Systems. *Inf. Process. Manag.* **2021**, *58*, 102657. [[CrossRef](#)]
52. Arrieta, A.B.; Diaz-Rodriguez, N.; Del Ser, J.; Bannetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [[CrossRef](#)]
53. Amini, M.; Bagheri, A.; Delen, D. Discovering injury severity risk factors in automobile crashes: A hybrid explainable AI framework for decision support. *Reliab. Eng. Syst. Saf.* **2022**, *226*, 108720. [[CrossRef](#)]
54. Yang, Z.; Zhang, A.; Sudjianto, A. GAMI-Net: An explainable neural network based on generalized additive models with structured interactions. *Pattern Recognit.* **2021**, *120*, 108192. [[CrossRef](#)]
55. Zhang, C.; Cho, S.; Vasarhelyi, M. Explainable Artificial Intelligence (XAI) in auditing. *Int. J. Account. Inf. Syst.* **2022**, *46*, 100572. [[CrossRef](#)]
56. Xie, S.; Lawniczak, A.; Gan, C. Optimal number of clusters in explainable data analysis of agent-based simulation experiments. *J. Comput. Sci.* **2021**, *62*, 101685. [[CrossRef](#)]
57. Masello, L.; Castignani, G.; Sheehan, B.; Guillen, M.; Murphy, F. Using contextual data to predict risky driving events: A novel methodology from explainable artificial intelligence. *Accid. Anal. Prev.* **2023**, *184*, 106997. [[CrossRef](#)] [[PubMed](#)]
58. Chen, L.; Tsao, Y.; Sheu, J.-T. Using Deep Learning and Explainable Artificial Intelligence in Patients' Choices of Hospital Levels. *arXiv* **2020**, arXiv:2006.13427.
59. Yang, G.; Ye, Q.; Xia, J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Inf. Fusion* **2022**, *77*, 29–52. [[CrossRef](#)] [[PubMed](#)]
60. Langer, M.; Landersb, R. The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Comput. Hum. Behav.* **2021**, *123*, 106878. [[CrossRef](#)]
61. Ding, W.; Abdel-Basset, M.; Hawash, H.; Ali, A.M. Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Inf. Sci.* **2022**, *615*, 238–292. [[CrossRef](#)]
62. Ali, S.; Abuhmed, T.; El-Sappagh, S.; Muhammad, K.; Alonso-Moral, J.M.; Confalonieri, R.; Guidotti, R.; Del Ser, J.; Diaz-Rodriguez, N.; Herrera, F. Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Inf. Fusion* **2023**, *99*, 101805. [[CrossRef](#)]
63. Gerlings, J.; Constantiou, I. Machine Learning in Transaction Monitoring: The Prospect of XAI. *arXiv* **2022**, arXiv:2210.07648.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.