risks

MDPI

Systematic Review

Explainable Artificial Intelligence (XAI) in Insurance

Emer Owens [1,*], Barry Sheehan [1], Martin Mullins [1], Martin Cunneen [1], Juliane Ressel [1,2] and German Castignani [3,4]

[1] Department of Accounting and Finance, University of Limerick, V94 PH93 Limerick, Ireland
[2] Research Center for the Insurance Market, Institute for Insurance Studies, TH Köln, 50968 Cologne, Germany
[3] Motion-S S.A., Avenue des Bains 4, Mondorf-les-Bains, L-5610 Luxembourg, Luxembourg
[4] Faculty of Science, Technology and Medicine (FSTM), University of Luxembourg, Esch-sur-Alzette, L-4365 Luxembourg, Luxembourg
* Correspondence: emer.owens@ul.ie

Abstract: Explainable Artificial Intelligence (XAI) models allow for a more transparent and understandable relationship between humans and machines. The insurance industry represents a fundamental opportunity to demonstrate the potential of XAI, with the industry's vast stores of sensitive data on policyholders and centrality in societal progress and innovation. This paper analyses current Artificial Intelligence (AI) applications in insurance industry practices and insurance research to assess their degree of explainability. Using search terms representative of (X)AI applications in insurance, 419 original research articles were screened from IEEE Xplore, ACM Digital Library, Scopus, Web of Science and Business Source Complete and EconLit. The resulting 103 articles (between the years 2000–2021) representing the current state-of-the-art of XAI in insurance literature are analysed and classified, highlighting the prevalence of XAI methods at the various stages of the insurance value chain. The study finds that XAI methods are particularly prevalent in claims management, underwriting and actuarial pricing practices. Simplification methods, called knowledge distillation and rule extraction, are identified as the primary XAI technique used within the insurance value chain. This is important as the combination of large models to create a smaller, more manageable model with distinct association rules aids in building XAI models which are regularly understandable. XAI is an important evolution of AI to ensure trust, transparency and moral values are embedded within the system's ecosystem. The assessment of these XAI foci in the context of the insurance industry proves a worthwhile exploration into the unique advantages of XAI, highlighting to industry professionals, regulators and XAI developers where particular focus should be directed in the further development of XAI. This is the first study to analyse XAI's current applications within the insurance industry, while simultaneously contributing to the interdisciplinary understanding of applied XAI. Advancing the literature on adequate XAI definitions, the authors propose an adapted definition of XAI informed by the systematic review of XAI literature in insurance.

Keywords: Explainable Artificial Intelligence; machine learning; insurance value chain; risk management; data governance

Citation: Owens, Emer, Barry Sheehan, Martin Mullins, Martin Cunneen, Juliane Ressel, and German Castignani. 2022. Explainable Artificial Intelligence (XAI) in Insurance. *Risks* 10: 230. https://doi.org/10.3390/risks10120230

Academic Editors: Gian Paolo Clemente, Nino Savelli and Diego Zappa

Received: 27 October 2022
Accepted: 24 November 2022
Published: 1 December 2022

1. Introduction

Artificial Intelligence (AI) revenues in insurance are expected to grow 23% to $3.4 billion between 2019–2024, yet the suitability of black-box AI models in insurance practices remains questionable (Bean 2021; Chen et al. 2019; GlobalData 2021). The growth of AI as an intelligent decision-making methodology that can perform complex computational tasks is revolutionising financial services, particularly within insurance practices. Data and its potential use are seen as a primary strategic asset and a source of competitive advantage in financial services firms, with AI models' leverage of such data providing numerous advantages (Kim and Gardner 2015). Such advantages of AI use in the insurance industry include enhanced fraud detection in claims management, granularity and personalisation when pricing insurance premiums, the creation of smart contracts, analysis of legal documents, virtual assistants (chatbots) and office operations (EIOPA 2021; Eling et al. 2021; McFall et al. 2020; Ngai et al. 2011; OECD, Organisation for Economic Co-operation and Development 2020; Riikkinen et al. 2018; Zarifis et al. 2019). AI encompasses the collation of multiple technologies in a single system which enables machines to interpret data and aid complex computational decision-making (Chi et al. 2020). Although AI models' advantages abound, recent literature highlights the AI models' opacity which is coined as black-box thinking (Adadi and Berrada 2018; Carabantes 2020). The Insurance Value Chain (IVC) makes extensive use of AI methods at every stage of the value creation process, with AI particularly impactful in claims management and underwriting and pricing departments (Eling et al. 2021). This research systematically reviews all peer-reviewed applications of (X)AI in insurance between 2000 and 2021 with a critical focus on explainability of the models. This is the first study to investigate XAI in an applied, insurance industry context.

The rationale for Explainable Artificial Intelligence (XAI) development is primarily driven by three main reasons: (i) demand for the production of more transparent models, (ii) necessity of techniques to allow for humans to interact with them, and (iii) trustworthy inferences from such transparent models (Došilović et al. 2018; Fox et al. 2017; Mullins et al. 2021). Decision-makers require an explanation of the AI system to aid in their understanding of their decision-making processes (Biran and Cotton 2017; Hoffman et al. 2018). Throughout this systematic review, AI is defined using recent recommendations by AI experts: "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real virtual environments such AI systems are designed to operate with varying levels of autonomy" (Krafft et al. 2020). As an extension of AI models, XAI involves enhancing current AI models by developing their transparency, interpretability and explicability, with such AI advancements ultimately aiming to make AI models more understandable to humans (Adadi and Berrada 2018; Floridi et al. 2018). By presenting an analysis of insurance's AI applications' degree of explainability, the reader gleans an insight into the progress made to-date in insurance practice and research to satisfy the want for transparency and explanations of AI-driven decisions. Practically, end-insurance consumers affected by AI-enhanced decisions will be less likely to trust in the decisions made by machines when they do not trust and understand the AI processes involved (Burrell 2016; Ribeiro et al. 2016).

Insurance's influence on socio-economic development cannot be understated, with the sound development of national insurance markets allowing for the promotion of financial stability, improved welfare and business innovation (Ferguson 2008; Ungur 2017). Insurance affordability is a key determinant of societal progress, with the modelling of insurance pricing practices playing a key role in this affordability (Daniels 2011), with actuarially fair pricing of insurance premiums allowing for a population to access insurance at rates which they can reasonably afford (Grant 2012). Transparency and explainability of AI models are core requirements to achieve impactful trustworthy AI in society (Felzmann et al. 2019; Maynard et al. 2022; Moradi and Samwald 2021). Trustworthiness is a core concept within the insurance industry, with enhanced XAI explanations directly affecting trust levels amongst insurance companies and their stakeholders.

This paper is structured as follows; Section 2 presents related works of this review while analysing current research on XAI's definition and related taxonomies, also outlining related work on (X)AI's impact on the IVC. Section 3 presents the methodological system to collect and analyse relevant literature on (X)AI use along the IVC. The search technique to arrive at relevant articles is especially emphasised to ensure the validity of eventual research results and allow for future research reproducibility. Section 4 outlines the review's findings of the systematically chosen sample of literature and their AI methods through the lens of defined XAI criteria. Section 5 presents a novel discussion of the review's results on the prevalence of XAI along the IVC, focusing on the extent to which AI applications along the IVC are explainable. Section 6 concludes the systematic review, reiterating points of interest regarding the future of XAI applications in insurance practices.

XAI Terminology

Kelley et al. (2018) define AI as "a computer system that can sense its environment, comprehend, learn, and take action from what it's learning", with XAI intuitively expanding on this description by allowing humans to be present at every stage of this AI decision-making lifespan. A common misnomer of AI models' explainability is that it is simply the improvement of trust in AI systems and their decision processes, through developing "causal structures in observational data" (Goodman and Flaxman 2017; Lipton 2018). Models' explainability enhances the interpretability, i.e., understanding how a model came to a certain decision (Lou et al. 2013), while also positively impacting fair and ethical decision making for high-computational tasks (Srihari 2020). Table 1 outlines the XAI variables and categories used within the systematic review to analyse the degree of explainability present in AI methods applied within the insurance industry. Additionally, the following categories of XAI methods are used to classify published applications of AI in insurance: (1) Feature Interaction and Importance, (2) Attention Mechanism, (3) Dimensionality Reduction, (4) Knowledge Distillation and Rule Extraction, and (5) Intrinsically Interpretable Models[1]. Additional categorisations and terminology determinations are summarised in Clinciu and Hastie (2019) and Arrieta et al. (2020).

Table 1. XAI Variables used during the literature analysis to assess the explainability of AI systems applied within insurance industry practices. Reference Appendix A for additional discussion on XAI variables used during analysis.

Intrinsic vs. Post hoc	Intrinsic Interpretability	Describes *how a model works* and is interpretable by itself. Interpretability is achieved through imposing constraints on the model.	Lipton (2018); Molnar (2019); Rudin (2018)
	Post hoc Interpretability	Analyses *what else the (original) model can tell us*, necessitating additional models to achieve explainability. The original model's explainability is analysed after training.	Du et al. (2019); Lipton (2018); Molnar (2019)
Local vs. Global	Local Interpretability	Reveals the impact of input features on the overall model's prediction.	Baehrens et al. (2010); Lundberg et al. (2020)
	Global Interpretability	Local explanations are combined to present the overall AI model's rules or features which determine their predictive outcome.	Kopitar et al. (2019); Lundberg et al. (2020)
Model-Specific vs. Model-Agnostic	Model-specific Interpretation	Interpretation is limited to specific model classes as each interpretation method is based on a specific model's internals.	Molnar (2019)
	Model-agnostic Interpretation	Applied to any AI model after the model's training. Analyses relationships between AI model's feature inputs and outputs.	Carvalho et al. (2019); Lipton (2018)

2. Fundamental Concepts & Background

2.1. Artificial Intelligence Applications in Insurance

AI use abounds across the entirety of the IVC with Eling et al. (2021) and EIOPA (2021) providing a thorough examination of the six main stages of the IVC and their goals. Tekaya et al. (2020) preface AI research in financial services by offering an overview of current use-cases and advantages of implementing Big Data and AI models in banking, credit risk management, fraud detection and the insurance industry. Several other articles highlight the importance and advantages of AI applications in the insurance industry, predicting major shifts in operations in the coming years (Paruchuri 2020; Riikkinen et al. 2018; Umamaheswari and Janakiraman 2014). Popular areas within insurance research where AI has been applied include fraud detection (Sithic and Balasubramanian 2013; Verma et al. 2017) and claims reserving (Baudry and Robert 2019; Blier-Wong et al. 2021; Lopez and Milhaud 2021; Wüthrich 2018). Grize et al. (2020) focus on ML applications in non-life insurance, highlighting AI's positive impact on risk assessment to improve the insurance companies' overall profitability in the long run.

Fang et al. (2016) used Big Data to develop a new profitability method for insurers using historical customer data, where they found that the Random Forest (RF) model outperformed other methods of forecasting (linear regression and SVM). Shapiro (2007) documents the extent to which fuzzy logic (FL) has been applied to insurance practices, which prompted Baser and Apaydin (2010)'s later research on claims reserving using hybrid fuzzy least squares regression and Khuong and Tuan (2016)'s creation of a neuro-fuzzy inference system for insurance forecasting. NallamReddy et al. (2014) present a robust review of clustering techniques used in insurance. Quan and Valdez (2018) use another understandable and transparent AI method, Decision Trees (DT), to investigate their use in insurance claims prediction. Interestingly, later research acknowledges the low predictive power of DTs and boosts their intrinsic interpretability to provide a more robust insurance pricing model (Henckaerts et al. 2021).

Sarkar (2020) argues that the insurance industry holds the potential for algorithmic capabilities to enhance each stage of the industry's value chain. Through highlighting AI's offerings at each stage of the IVC, the research prompted further studies from Walsh and Taylor (2020) and Eling et al. (2021) to determine precise AI opportunities available to the insurance industry. Walsh and Taylor (2020) highlight AI models' ability to mimic, or augment, human capabilities with NLP, Internet of Things (IoT) and computer vision. Eling et al. (2021) analyse AI's impact at each step on IVC and specifically highlights the potential for AI to enhance revenue streams, loss prediction and loss prevention measures for insurance practitioners.

Bias inherent to black-box AI systems threatens trust within the insurance industry, with this bias primarily driven by either humans' input or algorithmic bias (Koster 2020; Ntoutsi et al. 2020). There is potential for these models' impediments to compound and extenuate bias in their decision-making processes with unfair outcomes possible within the insurance industry (Confalonieri et al. 2021; Koster et al. 2021). This issue of bias is further aggravated when the lacking transparency in systems makes it difficult to dispute or appeal a biased decision by AI algorithms (von Eschenbach 2021). Bias in AI models could potentially lead to discriminatory behaviour of the AI system, caused by the model's tendency to use sensitive information resulting in unfair decisions (Barocas and Selbst 2016). There is strong research conducted on the determination of responsible AI, with Koster et al. (2021) providing a framework to create a responsible AI system, and Arrieta et al. (2020) outlining degrees of fairness to be implemented in an AI system to reduce discriminatory issues. Although a thorough examination of trust as it pertains to social sciences, leading into its importance in human-AI relationships, is beyond the scope of the current review, trust in AI systems is considered critical for the sustained use of AI technologies in insurance (Mayer et al. 1995; Siau and Wang 2018). Toreini et al. (2020) propose a Chain of Trust framework to further enhance users' trust in AI and ML technologies, while research on explanations in AI's use in medical diagnostic settings proves advantageous for clinician's trust and understanding in these technologies (Diprose et al. 2020; Tonekaboni et al. 2019). Jacovi et al. (2021) outline that the agreement between a human and AI system is contractual, therefore the interaction between a human and AI system must be explicit for trust to be present in the relationship between both parties (Hawley 2014; Tallant 2017). Trust derived from explanations in AI systems is enhanced by the provision of explanations and understandability, supporting the growth of XAI demand within the insurance industry.

2.2. Explainable Artificial Intelligence

XAI's recent history is firmly rooted in the field of AI, with contributions of explainability and transparency paving the way for XAI's growth. Lundberg and Lee (2017) cited explainability as the "interpretable approximation of the original complex [AI] model", while later Al-Shedivat et al. (2020) reference explainability as a "local approximation of a complex model (by another model)". What is clear from the increased research focus on AI in the late 2010's is that the notion of explainability did not drastically mature—research continues to ask the same questions pertaining to AI. Such issues include the fairness of an AI system, the transparency of decision pathways, and the explanation to be provided to the end user. A further important consideration is that XAI is merely the process of making AI understandable to humans, including its actions, recommendations and underlying decisions (Anjomshoae et al. 2019). Neither AI, nor XAI, are on the cusp of machine-led moral decisions or understanding (Ford 2018). Humans are still at the core of (X)AI, with bias and fairness central issues to contend with. This section outlines current research in XAI and its impact on the research field of AI.

The evaluation of the insurance industry's (X)AI applications' explainability contributes to the interdisciplinary literature on XAI. Through presenting the current discussion and taxonomies of XAI in the literature, the authors highlight the necessity of defined XAI criteria and categories in line with those used in this paper's analysis. Gade et al. (2019) outline the main challenges for XAI researchers which include (1) 'defining model explainability', (2) 'formulating explainability tasks for understanding model behaviour and developing solutions for these tasks', and (3) 'designing measures for evaluating the performance of models in explainability tasks'. Vilone and Longo (2020)'s later systematic study contributed a classification system for published XAI literature, aiming to establish boundaries in the field of XAI research. Four main clusters of research were found by Vilone and Longo (2020); (1) 'reviews focused on specific aspects of XAI', (2) 'the theories and notions related to the concept of explainability', (3) 'the methods aimed at explaining the inferential process of data-driven and knowledge-based modelling approaches', and (4) 'the ways to evaluate the methods for explainability'.

Extending on the above, the literature on XAI is attempting to determine a sound definition of XAI, which is commonly referred to as 'explainability' rather than 'interpretability'. Islam et al. (2020) note that explainability is more than interpretability in terms of importance and trust in the prediction. Interpretability is often the end goal with explanations acting as tools to reach interpretability (Honegger 2018). Additionally, the General Data Protection Regulations (GDPR) (EU 2016) which is discussed later in this paper covers only explainability (Došilović et al. 2018). These considerations encourage the authors to focus on the need for a domain-specific definition of XAI relevant to insurance practices. Instead of offering actionable definitions of XAI, other works classify the requirements that an explainable system should meet (Lipton 2018; Xie et al. 2020) or the methods of evaluations underhich an AI system can be deemed explainable (Doshi-Velez and Kim 2017; Hoffman et al. 2018; Lipton 2018; Rosenfeld 2021).

Reviews of XAI in medicine ignited the XAI research field, with many studies on the technology's effects on disease diagnosis, classification and treatment published in recent years. Payrovnaziri et al. (2020) involved the review of 49 articles published in the period 2009–2019 to group XAI methods used in the medical field. In this study, Payrovnaziri et al. (2020) grouped XAI methods into 5 different groups: (1) 'Knowledge Distillation and Rule Extraction', (2) 'Intrinsically Interpretable Models', (3) 'Data Dimensionality Reduction', (4) 'Attention Mechanism' and (5) 'Feature Interaction and Importance'. Antoniadi et al. (2021) outline challenges pertaining to AI's use for clinical decision support systems, emphasising lacking transparency as a key issue. Notwithstanding the obvious advantages of XAI methods to enhance understandability and aid medical practitioners' decisions which abound, their research finds a distinct lack of XAI applications in medicine.

Finance-related studies on XAI include Demajo et al. (2020); Hadji Misheva et al. (2021) and Biecek et al. (2021)'s research on credit scoring and risk management. Similarly, Bussmann et al. (2020) explore XAI in fintech risk management and peer-to-peer lending platforms, while Kute et al. (2021) also focus on risk management in finance applications through their review of DL and XAI technologies in the identification of suspicious money laundering practices. Gramegna and Giudici (2020) analyse XAI's potential to identify policyholders' reasons for buying or abandoning non-life insurance coverage. The grouping and assessment of like-minded policyholders allows for additional high-quality information on policyholders to be obtained, with transparent and accessible AI models used. Adadi and Berrada (2018) provide a foundational background to the main concepts and implications of an XAI system, citing data security and fair lending in financial services as key issues surrounding XAI use in financial services. Concerning banking and accounting practices, Burgt (2020) states that trust in AI systems in the banking industry is paramount and provides a discussion on the trade-off between explainability and predictability of AI systems. Gramespacher and Posth (2021) then utilise XAI to optimise the return target function of a loan portfolio, while Mehdiyev et al. (2021) add to the conversation by analysing tax auditing practices and public administration's appetite for XAI. Albeit the obvious advantages of developing transparent decision-making systems in public administration, this research cites the requirements of safe, reliable, and trustworthy AI systems as creating additional complexity in AI systems which take some time to implement widely. The interest in human-centred decision-making machines reaches beyond medical and finance domains. Putnam and Conati (2019) provide a survey that finds students seek additional explanations from their Intelligent Tutoring System to aid their education prospects. Natural Language Processing (NLP) is another research area with significant interest in XAI methods as revealed by Danilevsky et al. (2020) with sarcasm detection in dialogues later reviewed by Kumar et al. (2021). Anjomshoae et al. (2019) reviews inter-robot explainability and addresses the issue of explainability to non-users of ML robots through personalisation and context awareness.

The current systematic review builds upon previous research on XAI methods' classification and analysis of XAI literature during the systematic selection of literature. Although the above literature does provide a brief overview of the current understanding of XAI and related key concerns highlighted in the literature, this is the first paper to review XAI applications in the insurance industry.

2.3. The Importance of Explainability in Insurance Analytics

The personal data of EU citizens is described as a fundamental right by the EU Charter of Fundamental Rights and has been addressed since 1995 by the Data Protection Directive (Taylor 2017; Yeung et al. 2019). Citizen rights to privacy are operationalised through a number of data governance mechanisms ranging from consent platforms and data management systems which produce compliance measures to the control, use and lifespan of personal data. Accordingly, the data regulation environment is one of the most robust and sophisticated that is built on a strategy to both empower citizens to engage with the digital world and also to inform and guide commercial use of personal data. Data is protected by several regulatory instruments that provide a specific response to data use. These range from the data governance and the digital markets act to the GDPR (Andrew and Baker 2021; Goddard 2017). The range of different instruments speaks to the complexity of data use and data commercialisation scenarios. Insurance analytics often concerns the use of citizen and customer data to provide value to both the insureds and the insurance business model. Insurance analytics already uses personal data to optimise front- and back-end operations, risk modelling and risk pricing (Hollis and Strauss 2007; Keller et al. 2018; Ma et al. 2018; Mizgier et al. 2018; Naylor 2017). Furthermore, insurance analytics can provide important value in fraud management, claims management and better managing risk pooling by creating more accurate behavioural profiles of insureds (Barry and Charpentier 2020; Cevolini and Esposito 2020; Tanninen 2020). The commercial promise of insurance analytics also raises questions and concerns regarding the potential harms of undermining the core social solidarity of insurance by changing the pricing structure and limiting access to insurance products and services to those that meet stricter parameters of risk pricing. The importance of access to insurance is evident in compulsory products such as motor and, in some states, life insurance. Health insurance and insurance analytics are becoming a more controversial issue as increased reliance on private health care in parallel with increased use of insurance analytics are highlighting the tension between affordability and welfare. In short, insurance analytics offers scalable optimisation and high-value commercial solutions to IVCs and business models. Still, EU regulation is seeking to govern the use by the steering industry to more equitable, transparent and explainable (Kuo and Lupton 2020) uses of data analytics (EIOPA 2021; Mullins et al. 2021; van den Boom 2021).

3. Methodology

3.1. Literature Search Strategy

This literature search plan and related inclusion and exclusion criteria build upon the framework applied within Eling et al. (2021), with the aim of expanding upon their research to assess the prevalence of XAI methods in the IVC's AI applications. Eling et al. (2021)'s research assessed AI's impact on the IVC and the insurability of risks. The research presented in this paper expands on the abovementioned research to determine not only the impact on the IVC of AI systems being used, but also their degree of explainability. This framework is a suitable addition to the current study as a guide to literature inclusion criteria: inclusion of AI literature concerned with different stages along the IVC.

Analysis was conducted on a systematically selected body of literature from the following databases: EBSCOhost (Business Source Complete and EconLit), ACM Digital Library[2], Scopus, Web of Science and IEEE Xplore. These databases were chosen due to their wide breadth of content spanning both insurance and finance-related research, while also accounting for computer science journals to access research on AI applications. The above databases were chosen to feasibly and approximately align the current review with Eling et al. (2021)'s research, while considering database accessibility limitations.

Table 2 outlines the key search terms used interchangeably with AI in the abovementioned databases, alongside 'Insurance' OR 'Insurer' using Boolean terminology. This broad set of search terms ensures an all-encompassing article-base of the IVC's use of AI and are adapted from Eling et al. (2021)'s literature search method.

Table 2. Key Search Terms Interchangeable with (Explainable) Artificial Intelligence in the Literature Search Process.

Artificial Intelligence (AI)	Smart Devices	Analytics	Support Vector Machine (SVM)
Genetic Algorithm	Neural Network (NN)	Computational Intelligence	Machine Learning (ML)
Convolutional Neural Network (CNN)	Artificial Neural Network (ANN)	Explainable Artificial Intelligence (XAI)	Deep Learning
Data Mining	Big Data	Fuzzy Systems	Fuzzy Logic
Swarm Intelligence	Natural Language Processing (NLP)	Image Analysis	Machine Vision

Figure 1 outlines the systematic literature search process where an initial 419 articles were scanned for relevancy to this paper. Key relevancy criteria included the assessment of articles' contents concerning their place along the IVC. The IVC stages are extensively outlined in Table 3, as adapted from both Eling et al. (2021) and EIOPA (2021)'s research. The articles included in the systematic study of XAI in insurance are categorised according to the specific stage of the IVC which they refer to. This categorisation allows for further assessment of XAI use within the entire IVC process.

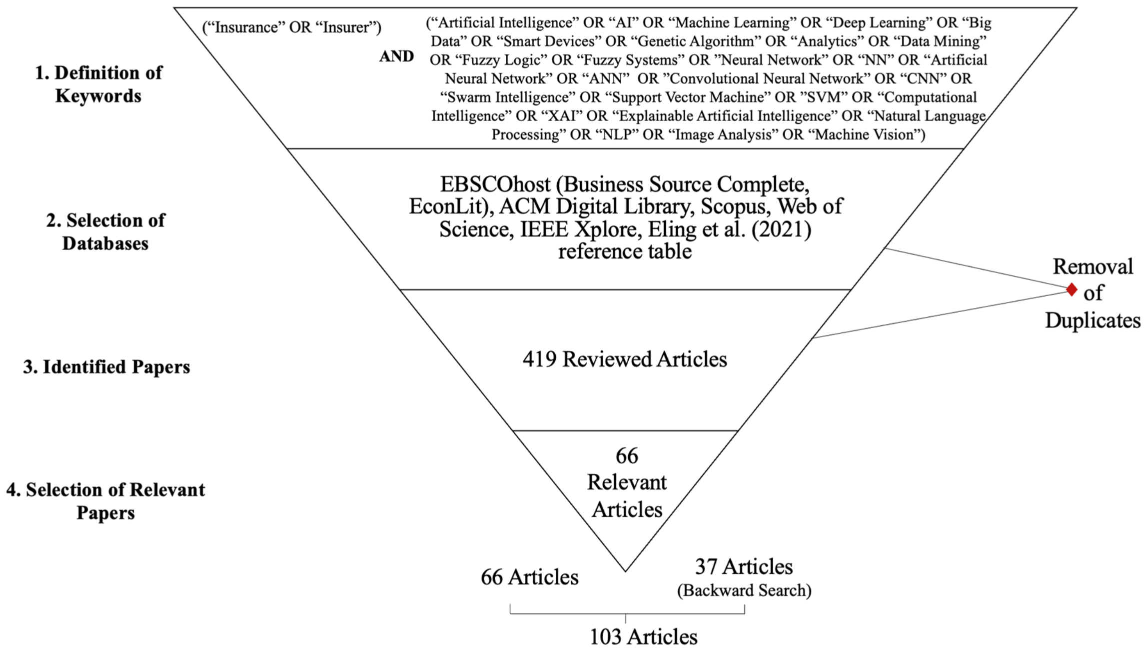

Figure 1. Literature Search Process. Backward Searching includes the assessment of the references in each of the 103 relevant articles for additional articles of relevance to the current review. Note: Eling et al. (2021).

In addition to the above, the articles' relevancy was filtered using the following criteria set:

- Time Period: Articles[3] published between 1 January 2000–31 December 2021 are included,
- Relevancy: The presence of keywords (Table 2) in the abstract is necessary for the article's inclusion. Additionally, the articles need to be relevant to the assessment of AI applications along the IVC directly (e.g., articles concerned with determining drivers' behaviour using telematics information, which may later inform insurance companies' pricing practices were excluded, as well as generalised surveys on AI uses in insurance[4]),
- Singularity: Duplicate articles found across the various databases are excluded,
- Accessibility: Only peer-reviewed articles that are accessible through the aforementioned databases and are accessible in full text are included (i.e., extended abstracts are not included),
- Language: Only articles published in English are included.

Articles published before 2000 are not included in the current review due to the increased understanding of AI from 2000 onwards (Liao et al. 2012), and the creation of the European GDPR in 2016 (implemented in the European Union in 2018) which is especially applicable to conversations on future XAI regulation.

Table 3. The Insurance Value Chain. The stages of the insurance industry's IVC is adapted from Grize et al. (2020); Eling et al. (2021) and EIOPA (2021).

Value Chain Stage	Main Tasks	Impact of Artificial Intelligence Applications
Marketing	Market and customer research Analysis of target groups Development of pricing strategies Design of advertisement and communication	- Improved prediction of customer lifetime value - Enhanced customer segmentation for personalised customer outreach and tailored communication strategies - Advanced insight about preferences in consumer purchasing behaviour for the identification of target product propositions and the generation of new ideas for product innovation - Churn models to enhance customer retention
Product Development	Configuration of products Verification of legal requirements	- The establishment of add-on services such as early detection of new diseases and their prevention enables the development of new revenue streams in addition to risk coverage - Entry into new markets and development of ecosystems with business partnerships in artificial intelligence-driven markets (e.g., autonomous driving, real-time health, and elderly care with nanobots, natural catastrophe management, smart home ecosystems) - Development of novel products utilising AI methods (e.g., usage-based, situational, and parametric insurance)
Sales and Distribution	Customer acquisition and consultation Sales conversations Product sale After-sales services	- Support of human sales agents by offering advanced sales insights (e.g., cross- and up-selling opportunities) through smart data-driven virtual sales assistants (chatbots) for improved customer consultation and tailored product recommendations - Proactive customer relationship management and improved after-sales services through increased client transparency - Chatbots for automated product consultation and sale of standardised insurance products - Customer Relationship Management (CRM) analytics used to inform nudging and cross-selling of related services ("next-best-action")
Underwriting and Pricing	Product pricing (actuarial methods) Application handling Risk assessment Assessment of final contract details	- Automated application handling, underwriting and risk assessment processes enable accurate insurance quotes within minutes - New data and insights allow the formation of small and homogenous risk pools, reduction in adverse selection and moral hazard in risk assessment - Micro-segmentation of insurance customers based on behavioural traits to provide personalised insurance pricing (e.g., dynamic online pricing)
Contract Administration and Customer Services	Change of contract data Customer Queries	- Development of chatbots for the automated answering of written and verbal customer queries using Natural Language Processing (NLP) - Offering advice about health and fitness goals or improved road safety to promote loss prevention - Proactive customer outreach and regular customer engagement
Claim Management	Claim settlement Investigation of fraud	- Automated claims management leads to decreasing claim settlement life cycles and increased payout accuracy - Improved fraud detection reduces fraud-related loss positions: anomaly detection, social network analytics and behavioural modelling - Loss reserving aided by AI estimating the value of losses
Asset and Risk Management	Asset allocation Asset liability management Risk control	- Automated investment research with more accurate and detailed market data enables portfolio management to make better-informed decisions due to new insights and more sophisticated analysis of data - Automated risk reporting - Development of robo-advisors for automated asset management - Automated trading systems improve asset allocation

The initial screening process included the assessment of 419 articles (following duplicate removal) based on their title, source, and abstract for the presence of the key search terms. In all, 66 articles were included for final review at this stage of the literature search. A backward search of the relevant articles (n = 66) was then conducted, which identified a further 37 articles. The backward search is a popular method of rigorous literature searching within systematic reviews in a range of disciplines including medicine (Mohamadloo et al. 2017), law (Siegel et al. 2021) and finance (Eckert and Hüsig 2021). The backward search entailed the assessment of the 66 relevant articles' bibliographies for additional articles of relevance to the current review. Based on this rigorous selection process, a total of 103 articles were identified as relevant for the current study (Reference Appendix A for the complete database of articles meeting the relevance threshold for inclusion in this systematic review). Figure 2 provides a breakdown of the publication year dispersion of each of these 103 articles. These results are comprised of ~75% journal articles (n = 77 and ~25% conference papers/proceedings (n = 26). The PRISMA flow diagram depicts the systematic review process (Figure 3). The PRISMA statement enhances transparency of systematic reviews, to ensure the research conducted during the course of a systematic review is robust and reliable (Page and Moher 2017). Each stage of the literature search for the systematic review is highlighted within the PRISMA diagram (Figure 3).

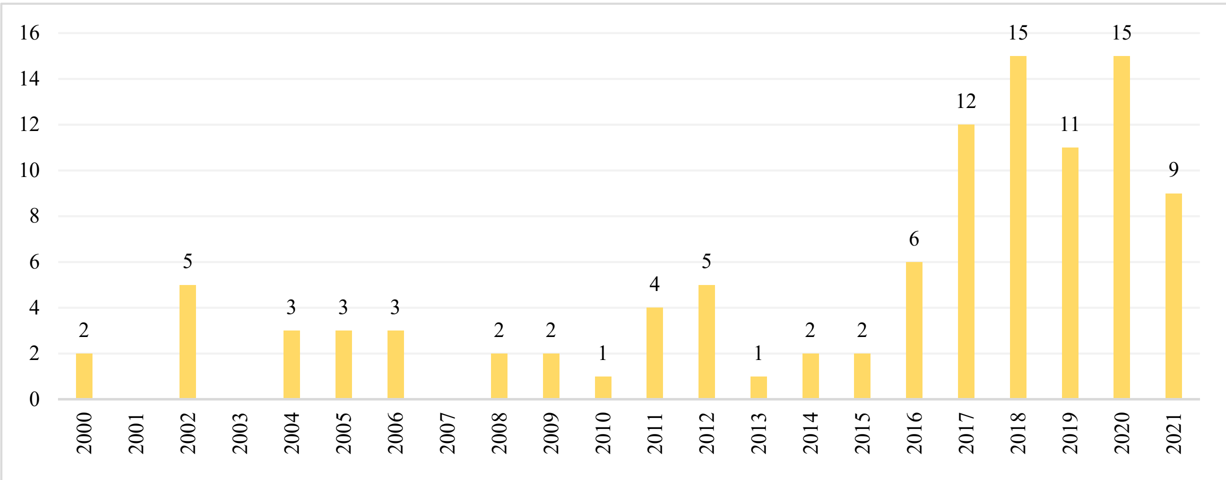

Figure 2. Insurance AI Articles Meeting Relevance Threshold (2000–2021) outlines the number of systematically reviewed articles by year according to the inclusion and exclusion criteria outlined in Section 3.1.

3.2. Literature Extraction Process

The evaluation of the full-text articles is sub sectioned into two distinct phases in line with both core contributions of this review. Initially, the articles' applied AI method was distinguished, alongside the prediction task(s) of this AI method. Secondly, the degree of explainability of the AI method employed is analysed. Here, the degree of explainability is evident in the XAI criteria applicable to each AI method employed in each article.

The criteria used in evaluating the AI methods' degree of explainability (Table 4) are adapted from Payrovnaziri et al. (2020)'s systematic review methodology and modified to suit this review on the insurance industry. The inclusion of the XAI variables and criteria is supported by previous research in XAI, with the criteria synthesised from Mueller et al. (2019); Du et al. (2019); Carvalho et al. (2019) and Payrovnaziri et al. (2020).

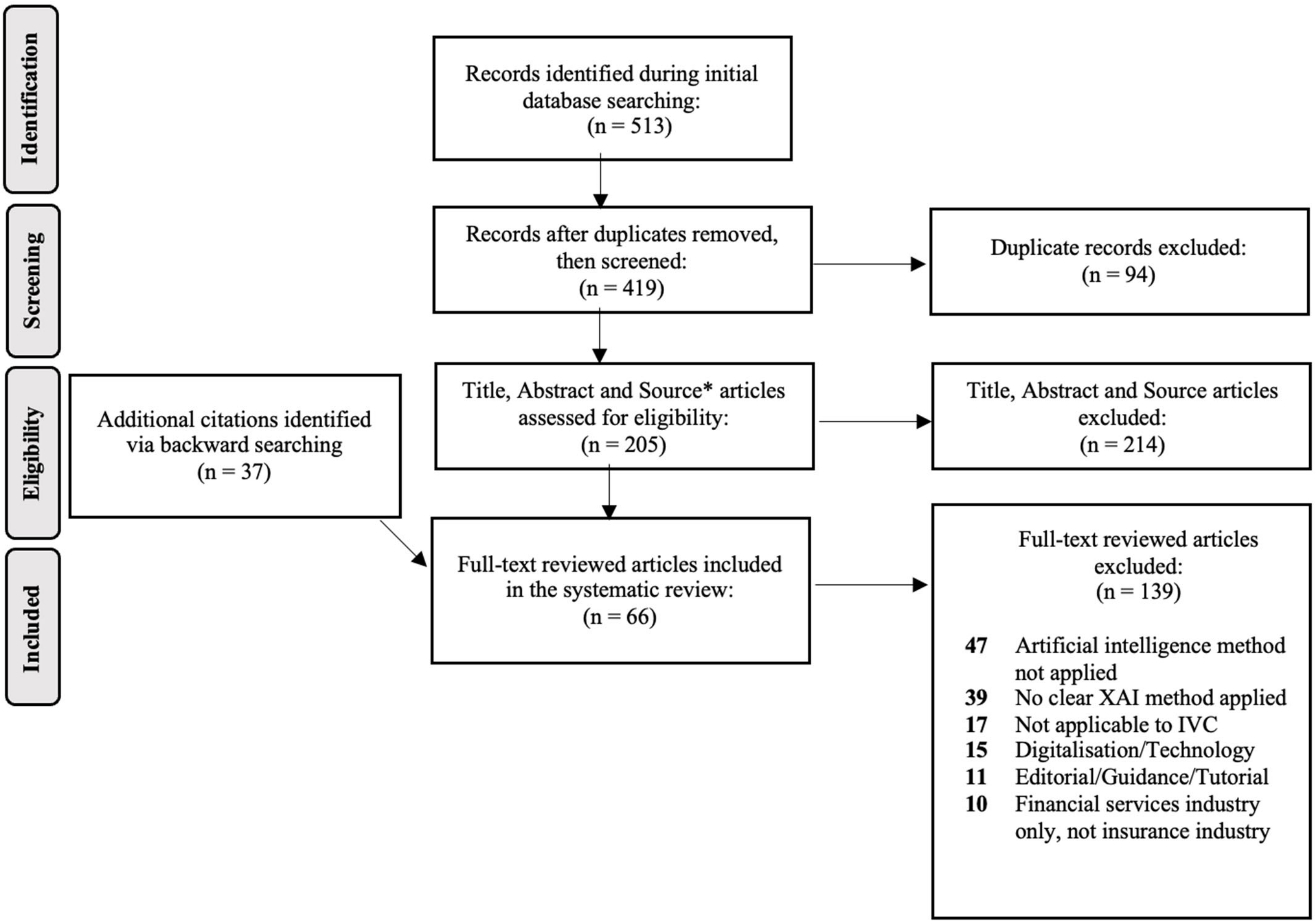

Figure 3. The PRISMA Flow Diagram is a recognised standard for systematic review literature search processes (Stovold et al. 2014). * 'Source' refers to the article inclusion criteria for this systematic review: journal articles and conference papers/proceedings are included.

Table 4. AI Methods and XAI Criteria used for the systematic analysis of the literature.

AI Method		XAI Criteria
Bayesian Network	Instance-based	Feature Interaction and Importance
Clustering	Regression	Attention Mechanism
Neural Network	Reinforcement Learning	(Data) Dimensionality Reduction
Decision Tree	Regularisation	Knowledge Distillation & Rule Extraction
Ensemble	Rule-based	Intrinsically Interpretable Models
Fuzzy Logic	Support Vector Machine	

3.3. Limitations of the Research

Limitations of the current review are outlined to ensure the validity and reliable reproducibility of results. In particular, the authors are unable to access 18 references which Eling et al. (2021) presented following their literature search process, while the industry reports reviewed within the same article are not included in the current systematic review. The lack of industry reports' analysis in this paper leads to an absence of articles concerning the Support Activities stage on the IVC. In Eling et al. (2021)'s research, all articles found pertaining to insurance companies' Support Activities were industry reports.

Industry reports were not included in this paper as access to articles with complete methodological processes outlined is pertinent to the current systematic review, a section which industry reports regularly omit in their publications. The inclusion of academic articles and conference articles ensures the methods of AI integration in each of the reviewed articles is outlined, in particular a coherent methodology discussion which can be assessed using the XAI criteria outlined in this paper.

The authors note the limitations of Payrovnaziri et al. (2020)'s research framework pertaining to XAI literature. In particular, the XAI categorisations presented feature some overlap across various XAI categories. For example, attention mechanism targets feature attribution, a category which is also covered under the feature interaction and importance categorisation. Nevertheless, this framework provides optimal categorisations for the scope of this work to assess the degree of explainability within AI applications in insurance, as defined boundaries of each XAI categorisation is provided.

4. Systematic Review Results

4.1. AI Methods and Prediction Tasks

The systematically chosen articles are first assessed based on the AI method employed and associated prediction task, with a focus on then distinguishing the degree of explainability evident in the literature. The stage of the IVC each article refers to is also clarified in the systematic research findings. Research on AI's use along the IVC over the twenty-one-year period of this review revealed AI is popular at every stage of the IVC, except for insurance companies' Support Activities. Such activities include general HR, IT and Public Relations departments in insurance companies. As mentioned above, a viable reason for the lack of articles concerned with this stage of the IVC is that Eling et al. (2021)'s study found articles on this subject through their review of industry reports, which the present systematic review did not include in the systematic review. The Underwriting and Pricing stage reveals significant research results (40%), with Claim Management (34%) also making extensive use of AI methods for fraud management and identification in particular.

Table 5 lists all the articles alongside the AI method employed and prediction task. A range of AI methods are used in the articles including; (1) Ensemble, (2) Neural Network (NN), (3) Clustering, (4) Regression (Linear and Logistic), (5) Fuzzy Logic, (6) Bayesian Network (BN), (7) Decision Tree, (8) Support Vector Machine (SVM). Other methods used include Instance- and Rule-based, Regularisation and Reinforcement Learning. The most popular AI method used is Ensemble (23%), with both NNs (20%) and Clustering (14%) also proving popular.

The line of insurance business the research in each article refers to is also classified, with non-life insurance lines returning a high number of articles in the systematic review (55%). Motor insurance prediction problems are popular areas of research, including driving behaviour classification and automobile insurance fraud (44%). Articles concerning insurers' life-business shows health(care) insurance as a popular area of research (13%), with health insurance fraud prevention and the classification of health insureds the most prominent research areas.

Table 5. AI Methods and Prediction Tasks. Abbreviations in Table 5 are outlined in the Abbreviations section of this paper.

		AI Method	Prediction Task(s)	Life/Non-Life	Line of Insurance Business
Marketing					
1	Chang and Lai (2021)	Neural Network	ANNs used to predict the propensity of consumers to purchase an insurance policy	-	-
2	Desik et al. (2016)	Regression	Develop a predictive modelling solution to aid the identification of the best insurance product group for current insurance product group of customers	-	-
3	Fang et al. (2016)	Ensemble	Prediction of insurance customer profitability	Life	Health
4	Larivière and Van den Poel (2005)	Ensemble	Prediction of customer retention and profitability	-	-
5	Lin et al. (2017)	Ensemble	Classification to enhance the marketing of insurance products	Life	-
6	Morik et al. (2002)	Rule-based	Extraction of low-level knowledge data to answer high-level questions on customer acquisition, customer up- and cross-selling and customer retention within insurance companies	-	-
Product Development					
7	Alshamsi (2014)	Ensemble	Prediction of automobile insurance policies chosen by customers using Random Forest (RF)	Non-life	Motor
8	Karamizadeh and Zolfagharifar (2016)	Clustering	*K*-means used to identify clusters which contribute to the profit and loss of auto insurance companies	Non-life	Motor
9	Khodairy and Abosamra (2021)	Neural Network	Driving behaviour classification	Non-life	Motor
10	Shah and Guez (2009)	Neural Network	Calculation of life expectancy (mortality forecasting) based on the individual's health status	Life	Health
11	Sheehan et al. (2017)	Bayesian Network	BN risk estimation approach for the emergence of new risk structures, including autonomous vehicles	Non-life	Motor and ProductLiability
Sales and Distribution					
12	Desik and Behera (2012)	Decision Tree	Creation of business rules from customer-led data to improve insurer competitiveness	-	-
13	Gramegna and Giudici (2020)	Ensemble	XGBoost predictive classification algorithm provides Shapley values	Non-life	-
14	Jeong et al. (2018)	Rule-based	Association between policyholder switching after a claim and the associated change in premium	Non-life	Motor
15	Tillmanns et al. (2017)	Bayesian Network	Selection of promising prospective insurance customers from a vendor's address list	-	-
16	Wang (2020)	Ensemble	Prediction of auto-renewal using RF	Non-life	Motor
17	Yang et al. (2006)	Ensemble	Ensemble of DTs used to maximise the expected net profit of customers	-	-
18	Zahi and Achchab (2019)	Clustering	Grouping of health insured population	Life	Health
19	Zhang and Kong (2020)	Bayesian Network	Estimation of insurance product recommendation	-	-

Table 5. *Cont.*

		AI Method	Prediction Task(s)	Life/Non-Life	Line of Insurance Business
Underwriting and Pricing					
20	Aggour et al. (2006)	Fuzzy Logic	Encoded the underwriting guidelines to automate the underwriting procedures of long-term care and life insurance policies	Life	Long Term Care
21	Baecke and Bocca (2017)	Regression	Assess the enhanced accuracy of risk selection predictive models utilising driving behaviour variables in addition to traditional accident risk predictors	Non-life	Motor
22	Bian et al. (2018)	Ensemble	Ensemble learning-based approach to obtain information on a user's risk classification which informs the compensation payout	Non-life	Motor
23	Biddle et al. (2018)	Instance-based	Prediction of the applications of exclusions in life insurance policies when automated underwriting methods are employed	Life	-
24	Bonissone et al. (2002)	Fuzzy Logic	Automation of underwriting practices	-	-
25	Boodhun and Jayabalan (2018)	Neural Network	Predict the risk level of life insurance applicants	Life	-
26	Bove et al. (2021)	Rule-based	Predetermined feature values provided	Non-life	Motor
27	Carfora et al. (2019)	Clustering	Evaluation of UBI automobile insurance policies	Non-life	Motor
28	Cheng et al. (2011)	Support Vector Machine	Evaluation of loss risk and development of criteria for optimal insurance deductible decision making	Non-life	Construction
29	Christmann (2004)	Ensemble	Indirect estimation of the pure premium in motor vehicle insurance	Non-Life	Motor
30	David (2015)	Regression	Use of the GLM to establish policyholders' pure premium	Non-life	Motor
31	Denuit and Lang (2004)	Regression	GAMs used for rate-making	Non-life	Motor
32	Deprez et al. (2017)	Ensemble	Mortality modelling using boosting regression techniques	Life	-
33	Devriendt et al. (2021)	Regularisation	LASSO penalty development to aid regularisation techniques in ML	-	-
34	Gan (2013)	Clustering	Selection of representative policies for the assessment of variable annuity policy pricing	Life	-
35	Gan and Huang (2017)	Clustering	Valuation of variable annuity policies	Life	-
36	Gan and Valdez (2017)	Reinforcement Learning	Monte Carlo-based modelling for variable annuity portfolios	Life	-
37	Guelman (2012)	Ensemble	Gradient Boosting Trees used to predict insurance losses	Non-life	Motor
38	Gweon et al. (2020)	Ensemble	Bias-corrected bagging method used to improve predictive performance of regression trees	Non-life	-
39	Huang and Meng (2019)	Regression	Risk probability prediction based on telematics driving data	Non-life	Motor
40	Jain et al. (2019)	Ensemble	Risk assessment of potential policyholders using risk scores within numerous ensembles of AI methods	Life	-
41	Jiang et al. (2018)	Instance-based	A novel model for analysis of imbalanced datasets in end-to-end insurance processes	Life	-

Table 5. *Cont.*

		AI Method	Prediction Task(s)	Life/Non-Life	Line of Insurance Business
42	Joram et al. (2017)	Rule-based	Knowledge-based system to enhance life underwriting processes	Life	-
43	Kašćelan et al. (2016)	Clustering	Assessment and classification of premiums	Non-life	Motor
44	Kieu et al. (2018)	Clustering	Deal with inadequately labelled data trajectories with drivers' identifiers	Non-life	Motor
45	Kumar et al. (2010)	Support Vector Machine	Prediction of claims which need reworking due to errors	Life	Health
46	Kwak et al. (2020)	Ensemble	Driver identification using RF	Non-life	Motor
47	Lin (2009)	Neural Network	Price the correct premium rate for 'in-between' risks between predefined tariff rates	Non-life	Property & Casualty
48	Liu et al. (2014)	Ensemble	Adaboost to predict claim frequency of auto insurance	Non-life	Motor
49	Neumann et al. (2019)	Decision Tree	Prediction of insurance customers' decisions following an automobile accident	Non-life	Motor
50	Sakthivel and Rajitha (2017)	Neural Network	Prediction of an insurance portfolio's claim frequency for forthcoming years	Non-life	Motor
51	Samonte et al. (2018)	Neural Network	Automatic multi-class labelling of ICD-9 codes of patient notes	Life	Health
52	Sevim et al. (2016)	Neural Network	Determination of litigation risks for accounting professional liability insurance	Non-life	Professional Liability
53	Siami et al. (2020)	Instance-Based	Unsupervised pattern recognition framework for mobile telematics data to propose a solution to unlabelled telematics data	Non-life	Motor
54	Smith et al. (2000)	Neural Network	NNs used to classify policyholders as likely to renew or terminate, to aid the achievement of maximum potential profitability for the insurance company	Non-life	Motor
55	Wei and Dan (2019)	Support Vector Machine	Stock price prediction	Non-life	Agriculture
56	Wüthrich (2020)	Neural Network	Optimisation of NN insurance pricing models	Non-life	Motor
57	Yan and Bonissone (2006)	Neural Network	Classification to enhance NN functionality for automated insurance underwriting	-	-
58	Yan et al. (2020b)	Rule-based	Rating model for UBI automobile insurance rates	-	-
59	Yang et al. (2018)	Ensemble	Gradient Boosting Trees used to predict insurance premiums	Non-life	Motor
60	Yeo et al. (2002)	Clustering	Optimisation of insurance premium pricing	Non-life	Motor
Contract Administration and Customer Services					
61	Ravi et al. (2017)	Fuzzy Logic	Creation of association rules which analyse customer grievances and summarise them	-	-
62	Sadreddini et al. (2021)	Clustering	Prediction of airline customer clusters and appropriate Cancellation Protection Service insurance fee per customer group	Non-life	Airline
63	Sohail et al. (2021)	Bayesian Network	The optimal set of hyperparameters for the later used ML model is found using Bayesian optimisation methods	-	-
64	Vassiljeva et al. (2017)	Neural Network	Automobile insurance customers' risk estimate using ANN to inform contract development	Non-life	Motor
65	Vaziri and Beheshtinia (2016)	Fuzzy Logic	Value creation for insurance customers	Life	-

Table 5. *Cont.*

		AI Method	Prediction Task(s)	Life/Non-Life	Line of Insurance Business
Claim Management					
66	Baudry and Robert (2019)	Ensemble	Estimation of outstanding liabilities on a given policy using an ensemble of regression trees	-	-
67	Belhadji et al. (2000)	Regression	Calculate the probability of fraud in insurance files	Non-life	Motor
68	Benedek and László (2019)	Rule-based	Identification of fraud indicators	Non-Life	Motor
69	Bermúdez et al. (2008)	Bayesian Network	Bayesian skewed logit model used to fit an insurance database (binary data)	Non-life	Motor
70	Cao and Zhang (2019)	Instance-Based	SOFM NN used to extract characteristics of medical insurance fraud behaviour	Life	Health
71	Delong and Wüthrich (2020)	Neural Network	NNs testing of regression models	Non-life	Liability
72	Duval and Pigeon (2019)	Regression	Assessment of claim frequency		
73	Dhieb et al. (2019)	Ensemble	XGBoost used to detect automobile insurance fraudulent claims	Non-life	Motor
74	Frees and Valdez (2008)	Regression	Assessment of claim frequency	Non-life	Motor
75	Gabrielli (2021)	Neural Network	Estimation of claims reserves for individual reported claims	Non-life	
76	Ghani and Kumar (2011)	Support Vector Machine	Error detection in insurance claims	Life	Health
77	Ghorbani and Farzai (2018)	Clustering	Detection of fraud patterns	Non-life	Motor
78	Herland et al. (2018)	Ensemble	Medicare provider claims fraud	Life	Health
79	Johnson and Khoshgoftaar (2019)	Neural Network	Automation of fraud detection using ANN	Life	Health
80	Kose et al. (2015)	Clustering	Detection of fraudulent claims	Life	Health
81	Kowshalya and Nandhini (2018)	Rule-based	Fraudulent claim detection	Non-life	Motor
82	Kyu and Woraratpanya (2020)	Neural Network	CNN used to prevent claims leakage	Non-life	Motor
83	Lau and Tripathi (2011)	Rule-based	Association Rules' provision of actionable business insights for insurance claims data	Non-life	Liability
84	Lee et al. (2020)	Regression	GLM and GAM used in NLP to extract variables from text and use these variables in claims analysis	Non-life	Property &Casualty
85	Li et al. (2018)	Ensemble	Random Forest for automobile insurance fraud detection	Non-life	Motor
86	Liu and Chen (2012)	Clustering	Enhance the accuracy of claims fraud prediction	Non-life	Motor
87	Matloob et al. (2020)	Rule-based	Fraud detection	Life	Health
88	Pathak et al. (2005)	Fuzzy Logic	To distinguish whether fraudulent actions are involved in insurance claims settlement	-	-
89	Smyth and Jørgensen (2002)	Regression	GLM to model insurance costs' dispersion	Non-life	Motor

Table 5. *Cont.*

		AI Method	Prediction Task(s)	Life/Non-Life	Line of Insurance Business
90	Sun et al. (2018)	Instance-based	Determination of joint medical fraud through reducing the occurrence of false positives caused by non-fraudulent abnormal behaviour	Life	Health
91	Supraja and Saritha (2017)	Fuzzy Logic	Utilising fuzzy rule-based techniques to improve fraud detection	Non-life	Motor
92	Tao et al. (2012)	Fuzzy Logic	DFSVM used to solve the issue of misdiagnosed fraud detection due to the 'overlap' problem in insurance fraud samples	Non-life	Motor
93	Verma et al. (2017)	Clustering	*K*-means used to increase performance and reduce the complexity of the model	Life	Health
94	Viaene et al. (2002)	Regression	Fraud detection	Non-life	Motor
95	Viaene et al. (2004)	Ensemble	Adaboost used in insurance claim fraud detection	Non-life	Motor
96	Viaene et al. (2005)	Bayesian Network	NN for fraud detection	Non-life	Motor
97	Wang and Xu (2018)	Neural Network	NN used to detect automobile insurance fraud	Non-life	Motor
98	Xu et al. (2011)	Ensemble	Random rough subspace method	Non-life	Motor
99	Yan et al. (2020a)	Ensemble	Optimisation of BP Neural Network by combining it with an improved genetic algorithm	Non-life	Motor
Asset and Risk Management					
100	Cheng et al. (2020)	Neural Network	Optimal reinsurance and dividend strategies for insurance companies	-	-
101	Ibiwoye et al. (2012)	Neural Network	Insurer insolvency prediction	-	-
102	Jin et al. (2021)	Neural Network	Determine the optimal insurance, reinsurance, and investment strategies of an insurance company	-	-
103	Kiermayer and Weiß (2021)	Clustering	Grouping of insurance contracts	Life	Life

4.2. XAI Categories along the IVC

The following categories of XAI methods are highlighted within the article database; (1) Feature Interaction and Importance, (2) Attention Mechanism, (3) Dimensionality Reduction, (4) Knowledge Distillation and Rule Extraction, and (5) Intrinsically Interpretable Models. Figure 4 shows each stage on the IVC and the corresponding XAI method employed in the reviewed articles. The XAI methods' interpretability techniques are then categorised into (1) intrinsic or post hoc, (2) local or global and (3) model-specific or model-agnostic (Table 6). According to the reviewed articles, most of the research on AI applications in insurance is concerned with Knowledge Discovery and Distillation, which is also grouped with Rule Extraction (35%) XAI methods for the purpose of the current review.

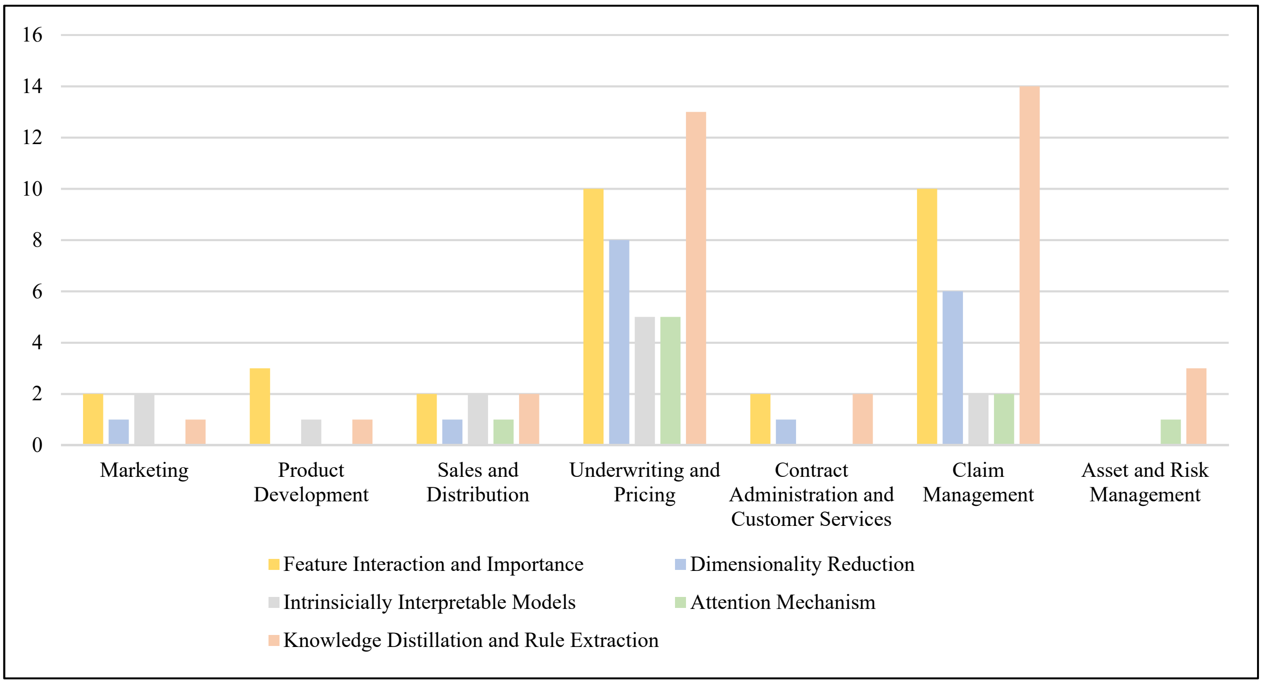

Figure 4. IVC Stage and Corresponding XAI Method Employed present the seven IVC stages assessed in the systematically chosen articles and the XAI method used in their methodology. Support Activities is not included in this paper as no articles returned in the systematic literature search presented prediction tasks in line with insurance companies' Support Activities.

Table 6. XAI Methods and their approach in the articles is outlined, with the additional XAI assessment of (i) intrinsic or post hoc, (ii) local or global, and (iii) model-specific or model-agnostic interpretability methods. Abbreviations in Table 6 are outlined in the Abbreviations section of this paper.

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
Marketing						
1	Chang and Lai (2021)	Feature Interaction and Importance	Dataset is pre-processed with three feature selection methods; (1) Neighbourhood Component Analysis (NCA), (2) Sequential Forward Selection (SFS) and, (3) Sequential Backward Selection (SBS)	Intrinsic	Global	Model-agnostic
2	Desik et al. (2016)	Dimensionality Reduction	Identification of relevant data clusters to inform model development for differing product groups	Post hoc	Local	Model-agnostic
3	Fang et al. (2016)	Intrinsically Interpretable Model	RF regression	Intrinsic	Global	Model-specific
4	Larivière and Van den Poel (2005)	Feature Interaction and Importance	Exploration of three major predictor categories as explanatory variables	Intrinsic	Local	Model-specific
5	Lin et al. (2017)	Intrinsically Interpretable Model	RF provides automatic feature selection which aids interpretability of the model	Intrinsic	Global	Model-specific
6	Morik et al. (2002)	Knowledge Distillation and Rule Extraction	Bridge the gap between databases and their users by implementing KDD methods	Intrinsic	Local	Model-specific
Product Development						
7	Alshamsi (2014)	Feature Interaction and Importance	Classification of data into different sets according to different policy options available	Intrinsic	Local	Model-specific
8	Karamizadeh and Zolfagharifar (2016)	Intrinsically Interpretable Model	Pattern recognition with clustering algorithms to find missing data to minimise insurance losses	Intrinsic	Global	Model-specific
9	Khodairy and Abosamra (2021)	Feature Interaction and Importance	Extraction of relevant features	Post hoc	Local	Model-agnostic
10	Shah and Guez (2009)	Feature Interaction and Importance	NN proposed as a better predictor of life expectancy than the Lee–Carter model due to the ability to adapt for each sex and each cause of life expectancy through a learning algorithm using historical data	Post hoc	Local	Model-agnostic
11	Sheehan et al. (2017)	Knowledge Distillation and Rule Extraction	Determination of causal and probabilistic dependencies through subjective assumptions (of the data)	Intrinsic	Local	Model-specific

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
Sales and Distribution						
12	Desik and Behera (2012)	Feature Interaction and Importance	CHAID used to create groups and gain an understanding of their impact on the dependent variable	Intrinsic	Local	Model-specific
13	Gramegna and Giudici (2020)	Intrinsically Interpretable Model	Similarity clustering of the returned Shapley values to analyse customers' insurance buying behaviour	Intrinsic	Global	Model-specific
14	Jeong et al. (2018)	Knowledge Distillation and Rule Extraction	Association rule learning to identify relationships among variables	Intrinsic	Global	Model-specific
15	Tillmanns et al. (2017)	Feature Interaction and Importance	PCA is used to reduce the dimensionality of the features and reduce the chance of overfitting	Post hoc	Local	Model-agnostic
16	Wang (2020)	Dimensionality Reduction	Removal of dataset features which have no bearing on the customers' likelihood to renew	Intrinsic	Local	Model-specific
17	Yang et al. (2006)	Knowledge Distillation and Rule Extraction	Development of postprocessing step to extract actionable knowledge from DTs to obtain actions which are associated with attribute-value changes	Intrinsic	Local	Model-specific
18	Zahi and Achchab (2019)	Intrinsically Interpretable Model	Clustering the insured population using k-means	Intrinsic	Global	Model-specific
19	Zhang and Kong (2020)	Attention Mechanism	Parameter optimisation for NB model	Post hoc	Local	Model-agnostic
Underwriting and Pricing						
20	Aggour et al. (2006)	Feature Interaction and Importance	Use of NLP and explanation of the interaction of different model features which alters the model	Intrinsic	Global	Model-specific
21	Baecke and Bocca (2017)	Feature Interaction and Importance	Stepwise feature selection	Intrinsic	Global	Model-specific
22	Bian et al. (2018)	Dimensionality Reduction	Found the 5 most relevant features to inform driving behaviour	Intrinsic	Local	Model-specific
23	Biddle et al. (2018)	Feature Interaction and Importance	Recursive Feature Elimination to provide feature rankings for feature subsets	Post hoc	Global	Model-agnostic

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
24	Bonissone et al. (2002)	Knowledge Distillation and Rule Extraction	Fuzzy rule-based decision systems used to encode risk classification of complex underwriting tasks	Intrinsic	Local	Model-specific
25	Boodhun and Jayabalan (2018)	Dimensionality Reduction	Correlation-Based Feature Selection and PCA	Intrinsic	Local	Model-specific
26	Bove et al. (2021)	Feature Interaction and Importance	SHAP is used to provide the contribution of each feature value to the prediction in comparison to the average prediction	Post hoc	Local	Model-agnostic
27	Carfora et al. (2019)	Intrinsically Interpretable Model	Identification of driver behaviour using ML algorithms	Intrinsic	Global	Model-specific
28	Cheng et al. (2011)	Knowledge Distillation and Rule Extraction	Development of loss prediction model using the ESIM	Intrinsic	Global	Model-specific
29	Christmann (2004)	Dimensionality Reduction	Exploitation of knowledge from certain characteristics of datasets to estimate conditional probabilities and conditional expectations given the knowledge of the variable representing the pure premium	Intrinsic	Local	Model-specific
30	David (2015)	Dimensionality Reduction	Use of policyholders' relevant characteristics to determine the pure premium	Intrinsic	Local	Model-specific
31	Denuit and Lang (2004)	Knowledge Distillation and Rule Extraction	Bayesian GAMs developed using MCAM inference	Intrinsic	Local	Model-specific
32	Deprez et al. (2017)	Attention Mechanism	Back-testing parametric mortality models	Post hoc	Global	Model-agnostic
33	Devriendt et al. (2021)	Knowledge Distillation and Rule Extraction	Development of SMuRF algorithm to allow for Sparse Multi-type Regularised Feature modelling	Intrinsic	Global	Model-specific
34	Gan (2013)	Knowledge Distillation and Rule Extraction	Gaussian Process Regression employed to value variable annuity policies	Intrinsic	Local	Model-specific
35	Gan and Huang (2017)	Knowledge Distillation and Rule Extraction	Kriging Regression method employed to value variable annuity policies	Intrinsic	Local	Model-specific

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
36	Gan and Valdez (2017)	Knowledge Distillation and Rule Extraction	Generalised Beta of the Second Kind (GB2) Regression method employed to value variable annuity policies	Intrinsic	Local	Model-specific
37	Guelman (2012)	Intrinsically Interpretable Model	Interpretable results given by the simple linear model through showcasing the relative influence of the input variables and their partial dependence plots	Intrinsic	Global	Model-specific
38	Gweon et al. (2020)	Knowledge Distillation and Rule Extraction	Bagging creates several regression trees which fits a bootstrap sample of the training data and makes a prediction through averaging the predicted outcomes from the bootstrapped trees	Post hoc	Local	Model-agnostic
39	(Huang and Meng 2019)	Dimensionality Reduction	Variables are binned to discretise continuous variables and construct tariff classes with significant predictive effects to improve interpretability of UBI predictive models	Post hoc	Intrinsic	Model-agnostic
40	Jain et al. (2019)	Feature Interaction and Importance	Using WEKA software, the dimensional feature set was reduced for use	Intrinsic	Global	Model-specific
41	Jiang et al. (2018)	Feature Interaction and Importance	Imbalanced data trend forecasting using learning descriptions and sequences and adjusting the CPLF	Post hoc	Local	Model-specific
42	Kašćelan et al. (2016)	Knowledge Distillation and Rule Extraction	Containment of the sets of rules with similar purpose and/or structure which defines the knowledge bases	Intrinsic	Global	Model-agnostic
43	Kieu et al. (2018)	Intrinsically Interpretable Model	Clustering provides homogeneity within classifications of risk and heterogeneity between risk classifications	Intrinsic	Global	Model-specific
44	Kumar et al. (2010)	Intrinsically Interpretable Model	Gradient Boosting DTs used to classify (unlabelled) trajectories	Post hoc	Local	Model-specific
45	Kwak et al. (2020)	Dimensionality Reduction	Frequency-based feature selection technique	Intrinsic	Global	Model-specific
46	Lin (2009)	Dimensionality Reduction	Reduction in feature values' noise (normalisation of sensing data)	Intrinsic	Local	Model-specific

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
47	Liu et al. (2014)	Attention Mechanism	Use of premium rate determination rules as network inputs in the BPNN to create the 'missing rates' of in-between risks	Post hoc	Local	Model-specific
48	Neumann et al. (2019)	Dimensionality Reduction	Reduction in claim frequency prediction problem to multi-class problem	Post hoc	Global	Model-specific
49	Sakthivel and Rajitha (2017)	Knowledge Distillation and Rule Extraction	Combination of simple linear weights and residual components to replicate non-linear effects to resemble a fully parametrised PPCI-like (Payments per Claim Incurred) model	Intrinsic	Local	Model-specific
50	Samonte et al. (2018)	Knowledge Distillation and Rule Extraction	Built a predictive model using previous Bayesian credibility inputs to predict the value of another field	Post hoc	Local	Model-specific
51	Carfora et al. (2019)	Attention Mechanism	NLP used for document classification of medical record notes, with RNNs employed to encode vectors in Bi-LTSM model	Intrinsic	Local	Model-specific
52	Sevim et al. (2016)	Attention Mechanism	Model is developed from the relationships between the variables gained from previous data and then tested	Post hoc	Local	Model-specific
53	Siami et al. (2020)	Feature Interaction and Importance	SOM to reduce data complexity	Intrinsic	Global	Model-specific
54	Smith et al. (2000)	Feature Interaction and Importance	Assessed the variables of relevance to the current task through rejecting variables with $x^5 < 3.92$	Post hoc	Local	Model-agnostic
55	Wei and Dan (2019)	Attention Mechanism	Parameter optimisation for SVM model	Intrinsic	Global	Model-specific
56	Wüthrich (2020)	Feature Interaction and Importance	Enhancement of neural network efficiency through feature selection	Intrinsic	Global	Model-specific
57	Yan and Bonissone (2006)	Knowledge Distillation and Rule Extraction	Comparison of four NN models for automated insurance underwriting	Post hoc	Local	Model-specific
58	Yan et al. (2020b)	Knowledge Distillation and Rule Extraction	Combination of the CNN and HVSVM models to create a model with higher discrimination accuracy than either model presents alone	Post hoc	Global	Model-specific

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
59	Yang et al. (2018)	Intrinsically Interpretable Model	TDBoost package provides interpretable results	Intrinsic	Local	Model-specific
60	Yeo et al. (2002)	Feature Interaction and Importance	Grouping of important clusters to input in NN model for insurance retention rates and price sensitivity prediction	Intrinsic	Local	Model-specific
Contract Administration and Customer Services						
61	Ravi et al. (2017)	Knowledge Distillation and Rule Extraction	Treatment of each variable as having a certain degree of membership with certain rules to categorise complaints	Intrinsic	Global	Model-specific
62	Sadreddini et al. (2021)	Feature Interaction and Importance	Cancellation Protection Service insurance fee is calculated based on the relevant weight of each cluster	Intrinsic	Global	Model-specific
63	Sohail et al. (2021)	Feature Interaction and Importance	SHAP is used in evaluating the feature importance in predicting the output level	Post hoc	Global	Model-agnostic
64	Vassiljeva et al. (2017)	Dimensionality Reduction	Only relevant parameters are considered in the ANN model	Intrinsic	Local	Model-specific
65	Vaziri and Beheshtinia (2016)	Knowledge Distillation and Rule Extraction	Development of integrated ML model to carry out the prediction task	Intrinsic	Local	Model-specific
Claim Management						
66	Baudry and Robert (2019)	Feature Interaction and Importance	Definition of policy subsets within the synthetic dataset	Post hoc	Local	Model-agnostic
67	Belhadji et al. (2000)	Feature Interaction and Importance	Regression used to isolate significant contributory variables in fraud	Intrinsic	Local	Model-specific
68	Benedek and László (2019)	Intrinsically Interpretable Model	Comparison of various intrinsic AI methods for fraud indicator identification	Intrinsic	Local	Model-specific
69	Bermúdez et al. (2008)	Knowledge Distillation and Rule Extraction	Use of a skewed logit model to more accurately classify fraudulent insurance claims	Post hoc	Global	Model-agnostic
70	Cao and Zhang (2019)	Dimensionality Reduction	PCA in the reduction in data's dimensionality	Post hoc	Local	Model-agnostic

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
71	Dhieb et al. (2019)	Dimensionality Reduction	Extraction of relevant features	Post hoc	Global	Model-specific
72	Delong and Wüthrich (2020)	Attention Mechanism	Describe the joint development process of individual claim payments and claims incurred	Intrinsic	Global	Model-agnostic
73	Duval and Pigeon (2019)	Knowledge Distillation and Rule Extraction	Combination of many regression trees together in order to optimise the objective function and then learn a prediction function	Intrinsic	Global	Model-agnostic
74	Frees and Valdez (2008)	Knowledge Distillation and Rule Extraction	Comparison of various fitted models which summarise all the covariates' effects on claim frequency	Intrinsic	Global	Model-specific
75	Gabrielli (2021)	Knowledge Distillation and Rule Extraction	NN proposed which is modelled through learning from one probability/regression function to the other via parameter sharing	Post hoc	Local	Model-specific
76	Ghani and Kumar (2011)	Knowledge Distillation and Rule Extraction	Development of an interactive prioritisation component to aid auditors in their fraud detection	Post hoc	Local	Model-specific
77	Ghorbani and Farzai (2018)	Knowledge Distillation and Rule Extraction	Definition of rules based on each cluster to determine future fraud propensity (using WEKA)	Intrinsic	Global	Model-specific
78	Herland et al. (2018)	Feature Interaction and Importance	Removed unnecessary data features	Intrinsic	Local	Model-specific
79	Johnson and Khoshgoftaar (2019)	Feature Interaction and Importance	Class imbalance within the dataset is rectified using one-hot encoding	Post hoc	Local	Model-specific
80	Kose et al. (2015)	Knowledge Distillation and Rule Extraction	Development of an electronic fraud & abuse detection model	Post hoc	Global	Model-agnostic
81	Kowshalya and Nandhini (2018)	Feature Interaction and Importance	Classifier construction using NB	Intrinsic	Local	Model-specific
82	Kyu and Woraratpanya (2020)	Feature Interaction and Importance	Fine-tuning of the dataset	Post hoc	Local	Model-specific

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
83	Lau and Tripathi (2011)	Knowledge Distillation and Rule Extraction	Development of Association Rules function for Workers' Compensation claim data analysis	Intrinsic	Global	Model-specific
84	Lee et al. (2020)	Knowledge Distillation and Rule Extraction	Transformation of words to vectors, where each vector represents some feature of the word	Intrinsic	Local	Model-specific
85	Li et al. (2018)	Dimensionality Reduction	PCA used to transform data at each node to another space when computing the best split at that node	Intrinsic	Global	Model-specific
86	Matloob et al. (2020)	Knowledge Distillation and Rule Extraction	Sequence generation to inform predictive model for fraudulent behaviour	Intrinsic	Local	Model-specific
87	Liu and Chen (2012)	Knowledge Distillation and Rule Extraction	Two evolutionary data mining (EvoDM) algorithms developed to improve insurance fraud prediction; (1) GAK-means (combination of *K*-means algorithm with genetic algorithm) and, (2) MPSO-*K*-means (combination of *K*-means algorithm with Momentum-type Particle Swarm Optimisation (MPSO))	Post-hoc	Local	Model-specific
88	Pathak et al. (2005)	Knowledge Distillation andRule Extraction	Mimic the expertise of the human insurance auditors in real life insurance claim settlement scenarios	Post-hoc	Local	Model-agnostic
89	Smyth and Jørgensen (2002)	Intrinsically Interpretable Model	Modelling of insurance costs' dispersion and mean	Intrinsic	Local	Model-specific
90	Sun et al. (2018)	Feature Interaction and Importance	Formulation of compact clusters of individual behaviour in a large dataset	Intrinsic	Local	Model-specific
91	Supraja and Saritha (2017)	Feature Interaction and Importance	*K*-means clustering used to prepare dataset prior to FL technique application	Intrinsic	Local	Model-specific
92	Tao et al. (2012)	Feature Interaction and Importance	Avoidance of curse of dimensionality problem through kernel function use for SVM's calculation	Post hoc	Global	Model-agnostic
93	Verma et al. (2017)	Knowledge Distillation and Rule Extraction	Association rule learning to identify frequent fraud occurring patterns for varying groups	Intrinsic	Local	Model-specific

Table 6. *Cont.*

		XAI Category	XAI Approach	Intrinsic/Post- hoc	Local/Global	Model-Specific/Agnostic
94	Viaene et al. (2002)	Dimensionality Reduction	Removal of fraud indicators with 10 or less instances to aid model convergence and stability during estimation	Intrinsic	Global	Model-specific
95	Viaene et al. (2004)	Attention Mechanism	Computation of the relative importance (weight) of individual components of suspicious claim occurrences	Intrinsic	Global	Model-specific
96	Viaene et al. (2005)	Feature Interaction and Importance	Determination of relevant inputs for the NN model	Post hoc	Local	Model-agnostic
97	Wang and Xu (2018)	Dimensionality Reduction	Extraction of text features hiding in the text descriptions of claims (Latent Dirichlet Allocation-based deep learning for text analytics)	Post hoc	Local	Model-agnostic
98	Xu et al. (2011)	Knowledge Distillation and Rule Extraction	Random rough subspace method incorporated into NN to detect insurance fraud	Intrinsic	Global	Model-specific
99	Yan et al. (2020a)	Dimensionality Reduction	PCA used to reduce dimensions of the multi-dimensional feature matrix, where the reduced data retains the main information of the original data	Intrinsic	Global	Model-specific
Asset and Risk Management						
100	Cheng et al. (2020)	Knowledge Distillation and Rule Extraction	Development of deep learning Markov chain approximation method (MCAM)	Intrinsic	Global	Model-specific
101	Ibiwoye et al. (2012)	Attention Mechanism	Tuning of the NN	Intrinsic	Local	Model-specific
102	Jin et al. (2021)	Knowledge Distillation and Rule Extraction	MCAM to estimate the initial guess of the NN	Intrinsic	Global	Model-specific
103	Kiermayer and Weiß (2021)	Knowledge Distillation and Rule Extraction	Approximation of representative portfolio groups to then nest in NN	Post hoc	Local	Model-specific

4.3. Feature Interaction and Importance

Analysing (X)AI models' input features' importance and interaction is a popular XAI method, with ~27% of reviewed articles utilising this method. The determination of features' importance contributed to the development of thorough XAI methods to complete many prediction tasks at each stage on the IVC. Smith et al. (2000) utilise Artificial Neural Networks (ANN) to gain an insight into customer policies which were likely to renew or terminate at the close of the policy period through analysing those factors which contribute to policy termination. This assessment of optimal premium pricing through data mining and ML methods instructs research on insurance customer retention and profitability. Additionally, addressing customer retention is Larivière and Van den Poel (2005)'s research which explored three predictor variables which encompass potential explanatory variables to inform insurance customer retention. Their RF model provides an importance measure between the explanatory and dependence variables for the prediction task.

Claim management and insurance fraud detection are areas which benefit from analysing the interaction and importance of feature inputs in AI applications through the isolation of important features which contribute to fraud (Belhadji et al. 2000). Similarly, Tao et al. (2012) avoid the curse of dimensionality through using the kernel function for SVMs in their XAI approach for insurance fraud identification, while Supraja and Saritha (2017) use this XAI method to ready their data for automobile fraud detection using fuzzy rule-based predictive techniques.

Feature interaction and importance is also useful in assessing risk across a wide range of insurance activities and informing underwriting and pricing of premiums. Biddle et al. (2018) add to literature on automated underwriting in life insurance applications using the XAI method of Feature Interaction and Importance. Recursive Feature Elimination is used to reduce the feature space through iteratively wrapping and training a classifier on several feature subsets and then providing feature rankings for each subset. Premium pricing of automobile insurance is researched by Yeo et al. (2002)'s, where cluster grouping of policyholders according to relative features aids in determining the price sensitivity of policyholder groups to premium prices.

4.4. Attention Mechanism

The Attention Mechanism within an AI model primarily attempts to find a set of positions in a sequence with the most relevant information on a prediction task (Payrovnaziri et al. 2020), which in turn enhances interpretability, according to Mascharka et al. (2018).

In line with the current review, Attention Mechanism is used to compute the weight of claim occurrences to inform fraud detection (Viaene et al. 2004) and inform insurer insolvency prediction (Ibiwoye et al. 2012). Lin and Chang (2009) apply Attention Mechanism in their determination of premium rates of 'in-between' risks through weight classification of different tariff classes. The method also aids in the determination of litigation risk of liability insurance within the accountancy profession, as Sevim et al. (2016) incorporate Attention Mechanism in their development of an ANN model, while Deprez et al. (2017) apply Attention Mechanism to mortality modelling through back-testing parametric mortality models. Samonte et al. (2018) use this XAI method for automatic document classification of medical record notes using NLP. The enhancement of the Hierarchical Attention Network model (EnHAN) assigns topics for each word in a given text and learns topical word embedding in a hierarchical manner. Topical word embedding models solve the multi-label, multi-class classification problem within medical records to inform cluster processes for billing and insurance claims.

Wei and Dan (2019) apply Attention Mechanism to parameter optimisation of SVM features, while Zhang and Kong (2020) also optimised parameters for input in NB model to inform insurance product recommendations. In terms of sequence generation, this XAI method was used by Matloob et al. (2020) to inform their predictive model for fraudulent behaviour in health insurance.

4.5. Dimensionality Reduction

Researchers typically use dimensionality reduction techniques in order to reduce the set of features inputted in the model principally to improve a model's efficiency (Motoda and Liu 2002). Kumar et al. (2010), for instance, use a frequency-based feature selection technique to reduce the dataset dimensions. This action aided in developing a model for error prevention in health insurance claims processing through reducing data storage requirements and improved model execution time. They found that using a lower frequency threshold and limiting the input feature improved the predictive accuracy. Finding similar results in terms of improved predictive accuracy, Li et al. (2018) use Principal Component Analysis (PCA) to increase the diversity of each of the 100 trees used in a RF model. This action improves the overall accuracy of the algorithm. In this instance, PCA transforms the data at each node to another space when computing the best split at that node which contributed to satisfactory feature selection in the development of the RF algorithm for fraud detection. PCA is also used in Underwriting and Pricing of life insurance through model development for risk assessment of life insurance customers (Boodhun and Jayabalan 2018).

For the popular prediction tasks related to automobile insurance, the reduction in dataset dimensionality is also useful. Liu et al. (2014) reduce their large claim frequency prediction to a multi-class prediction problem to aid the eventual implementation of Adaptive Boosting (AdaBoost) to automobile insurance data. The act of reducing the number of frequency classes contributes to AdaBoost presenting as superior to SVM, NN, DTs and GLM in terms of prediction ability and interpretability. Huang and Meng (2019) bin variables to approximate continuous variables in the dataset and construct tariff classes with high-level predictive power which enhances the model's accuracy and predictive power in the classification of usage-based insurance (UBI) products. An ANN model is optimised in Vassiljeva et al. (2017) to inform automobile contract development through assessing drivers' risk, while Bian et al. (2018) reduced their data dimensions to include only the five most relevant factors in determining drivers' behaviour.

Other stages on the IVC benefit from data dimensionality reduction, with Desik et al. (2016)'s identification of relevant data clusters to inform model development of marketing strategies within different insurance product groups proving successful. The Sales and Distribution stage of the IVC uses a similar reduction of dataset features which hold no bearing on insurance customers' likelihood of renewal (Kwak et al. 2020).

4.6. Knowledge Distillation and Rule Extraction

Knowledge Distillation and Rule Extraction components of AI models refers to the combination of large models to create a smaller, more manageable model (Hinton et al. 2015). For instance, both Cheng et al. (2020) and Jin et al. (2021) investigate optimal insurance strategies (insurance, reinsurance and investment) using the MCAM to develop adequate NN models for their respective prediction tasks. In another work concerning NNs and Knowledge Distillation XAI methods, Kiermayer and Weiß (2021) approximate representative portfolios of both term life insurance plans and Defined Contribution pension plans to aid in determining the insurer's solvency capital requirements. These representative portfolios are inputted in a NN model, which significantly outperforms k-means clustering for insurance portfolio grouping and the evaluation of insurers' investment surplus. The combination of models was also utilised by Xu et al. (2011) where a random rough subspace method is incorporated into a NN to aid optimised insurance fraud detection.

In terms of extracting actionable knowledge from models, Lee et al. (2020) propose a methodology for extracting variables from textual data (word similarities) to use such variables in claims analyses, thus improving actuarial modelling. Similarly, Wang and Xu (2018) apply LDA-based deep learning for the extraction of text features in claims data to detect automobile insurance fraud.

The development of association rules aids in building XAI models which are regularly understandable and useful for prediction tasks across the entirety of the IVC. Ravi et al. (2017) develop a model for analysing insurance customer complaints and categorising them for insurance customer service offices. Customer grievances are assigned an association rule which are categorised by treating grievance variables as holding a certain degree of membership with the different rules. Association rule learning is also implemented in fraud detection through the identification of frequent fraud occurrence patterns (Verma et al. 2017) and the computation of relative weights of variables related to suspicious claim activity using Adaboost AI methods (Viaene et al. 2004).

4.7. Intrinsically Interpretable Models

Aside from the interpretability techniques outlined above, other researchers have relied on the intrinsic predictive capabilities of models in their research. Through preserving the predictive capabilities of less complex AI models using boosting and optimisation techniques, the predictive power of Intrinsically Interpretable Models proves useful along the IVC.

Researchers implemented Intrinsically Interpretable Models for a range of prediction tasks including; (1) double GLMs to model insurance costs' dispersion and mean (Smyth and Jørgensen 2002), (2) prediction of insurance losses through boosting trees (Guelman 2012), (3) prediction of insurance customers' profitability (Fang et al. 2016), and (4) cluster identification and classification (Karamizadeh and Zolfagharifar 2016; Lin et al. 2017).

Carfora et al. (2019) identified clusters of driver behaviour to inform UBI pricing through unsupervised ML classification techniques and cluster analysis. *K*-means clustering is used to classify driver aggressiveness to inform a risk index of driving behaviour on different road types (primarily urban vs. highway). Benedek and László (2019) compare several interpretable AI techniques in their identification of insurance fraud indicators, which each facilitate the segmentation of such fraud indicators. DTs are highlighted as suitable AI methods for such indicator identification and classification.

5. Discussion

5.1. AI's Application on the Insurance Value Chain

The use of AI applications at each stage on the IVC is promising, with a variety of prediction tasks fulfilled by AI applications. In line with Eling et al. (2021)'s findings, AI is disrupting the insurance industry in a number of ways. The automation of underwriting tasks and the identification and prevention of fraudulent behaviour are key areas where AI is impacting the IVC. This is in line with a survey by the Coalition Against Insurance Fraud (2020) reporting 56% of insurance companies' surveyed AI as their primary mode of insurance fraud detection. An interesting note is the distinction between Eling et al. (2021)'s findings on AI's use in Support Activities and the presence of XAI methods in such activities. The literature search process for this review did not result in any articles concerning XAI use in insurance Support Activities (including HR, IT, Legal and General Management). The authors accept that this finding is likely attributed to restricted keyword searches which do not consider Support Activities, opening the possibility of further research on XAI's presence in insurance companies' Support Activities.

5.2. XAI Definition, Evaluation and Regulatory Compliance

Research on XAI (Section 2.1) highlight the disjointed understanding of XAI both across and within industries, thus providing motivation for the current review. There appears no consistent definition of XAI in the reviewed insurance literature, a finding which is in line with Payrovnaziri et al. (2020)'s findings of XAI's use and definition in medicine research. The main issue posed by this finding is that the evaluation of XAI methods is made increasingly difficult when there is no defined definition and scope of XAI. This review develops an XAI evaluation criteria, incorporating interpretability evaluation as either (i) intrinsic or post hoc, (ii) local or global and (iii) model-specific or model-agnostic. The results provide an extension to XAI survey research conducted by Adadi and Berrada (2018), Arrieta et al. (2020) and Das and Rad (2020) who each defined inter-related taxonomies of XAI. The development of an all-encompassing XAI definition for insurers and AI experts will allow for further adoption of XAI methods in the insurance industry.

Each definition of XAI discussed in Section 2.2 is derived from the early definition of explainability as the "assignment of causal responsibility" originally cited in Josephson and Josephson (1996). Although each paper providing additional insight into XAI definitions is useful, the lacking cohesion amongst these studies hampers the consolidation of each individual contribution into an interdisciplinarily accepted XAI definition. The authors acknowledge that an all-purpose XAI definition is difficult to determine, as both notions of explainability and interpretability (which are often used interchangeably and used in creating XAI definitions) are domain-specific notions (Freitas 2014; Rudin 2018). Lipton (2018) cites *interpretability* as an ill-defined concept as interpretability is not a fixed notion in and of itself. In efforts to define XAI specifically within the insurance industry, the authors accept all referenced definitions of XAI and findings of XAI use on the IVC to-date and propose the following XAI definition specific to the insurance industry:

> "XAI is the transfer of understanding to AI models' end-users by highlighting key decision- pathways in the model and allowing for human interpretability at various stages of the model's decision-process. XAI involves outlining the relationship between model inputs and prediction, meanwhile maintaining predictive accuracy of the model throughout"

In addition to benefitting XAI research, the authors note that a solid definition of XAI pertaining directly to the insurance industry (and financial services at large) will aid the development of adapted regulation, which is in line with recommendations from Palacio et al. (2021). The GDPR (EU 2016) established a regime of "algorithmic accountability" and (insureds) "right to explanation" from decision-making algorithms (Bayamlıoğlu 2021; Wulf and Seizov 2022). XAI promotes such transparent and interpretable traits, yet a comprehensive implementation of these methods necessitates regulatory compliance (Henckaerts et al. 2020). In the momentary absence of specific regulation of XAI models, the authors highlight the potential for XAI methods to be paired with existing governance measures in the insurance industry to mitigate concerns surrounding the use of novel AI methods until satisfactory regulation is developed. This recommendation is in line with governance guidelines from EIOPA (2021), for example the maintenance of human oversight in decision-making processes.

5.3. The Relationship between Explanation and Trust

The recent proliferation of XAI literature is partly driven by the need to maintain users' trust in AI to further develop AI adoption (Jacovi et al. 2021; Robinson 2020). Despite this rationale, prior XAI research has not considered the notion of trust in much detail. As a multidimensional and dynamic construct, a concise definition of trust has received considerable critical attention but remained elusive. The interplay between explainability and trust can be further substantiated by exploring what constitutes user trust in AI. So far, it has been established that explanations can positively affect users' trustworthiness assessment in several use cases, such as recommendation agents (Xiao and Benbasat 2007) or information security (Pieters 2011). In particular, explanations can foster cognitive-based trust that prevails early in the human-AI relationship. This initial trust development phase is often referred to as swift trust (Meyerson et al. 1996). This notion of interpersonal trust, following the common act of anthropomorphising machines, affects how humans interact with such machines (Hoffman 2017). Users are affected by the reliability of their 'partner' in the interpersonal relationship (the machine), however the lack of humane empathy and ability to apologise for mistakes during automated decision-making hinders the fostering of a truly anthropomorphised machine being involved in a real interpersonal relationship with a human (Beck et al. 2002). As interaction history is lacking, the extent to which a user can understand a given process or decision is paramount (Colaner 2022). However, the question remains whether there is a threshold after which this positive effect can be reversed. If users suffered from such *explanation overload*, more explanations would not be significantly associated with trust. This assessment is subjective and perceptual in nature and might well be influenced by a user's general propensity to trust AI models. This assumption accords with previous findings by McKnight et al. (2002) that the disposition to trust positively influences the trustworthiness assessment in e-commerce. Further work is thus required to examine how, precisely, the trust construct can be integrated into XAI research.

6. Conclusions

The primary contribution of this systematic review to widespread XAI understanding is an in-depth analysis of published literature on XAI in insurance practices. The growing commercialisation of AI applications leads to the potential of insurers to create high-value solutions in response to the industry's efficiency issues and respond appropriately to changes in the business landscape (Balasubramanian et al. 2018). The necessity to highlight transparent and understandable AI processes applied within the insurance industry prompts this investigation of XAI applications and their current use cases. This review of key literature provides a comprehensive analysis of XAI applications in insurance for both key insurance regulators and insurance practitioners which will allow for extensive application in future regulatory decision-making. Legally, the opacity of black-box AI systems hinders regulatory bodies from determining whether data is processed fairly (Carabantes 2020; Rieder and Simon 2017), with XAI enhancing the potential for AI systems' regulation under the GDPR in Europe (EU 2016).

This review assesses 103 articles (comprised of journal articles and conference papers/proceedings) which outline XAI applications at each stage of the IVC. The lack of explainability evaluation and consensus on XAI definitions hinders the potential progress of the XAI research field in insurance practices as there is no clear way to evaluate the degree of explainability in XAI. This review attempts to bridge this gap by defining XAI criteria and incorporating such criteria into a systematic review of XAI applications in insurance literature. Utilising this XAI criteria the degree of explainability in each XAI application is provided, assigning each AI method to a grouped XAI approach, and then evaluating the model's interpretability as either (i) intrinsic or post hoc, (ii) local or global, and (iii) model-specific or model-agnostic. Findings reiterate the authors' hypothesis that XAI methods are popular within insurance research, enabling the transparent use of AI methods in industry research. The transparency XAI methods afford insurance companies enhances the application of AI models in an industry striving for a basis of trust with multiple stakeholders.

Additionally, this paper analyses XAI definitions and proposes a revised definition of XAI. This proposed definition is informed by previous XAI definitions in XAI literature and systematic reviews of literature on AI applications on the IVC. The authors acknowledge this definition will not be widely applicable to a wide range of industries, therefore it's reiterated that the proposed XAI definition is applicable to financial services and the insurance industry. This definition will aid in adapting regulation in the insurance industry to suit an AI-rich insurance industry. Further clarification is necessary on the relationship between explanation and trust as both concepts pertain to XAI, with research recommendations centered on the extent to which explanations assist in the development of trust in AI models.

To achieve a substantial understanding of the entire potential of XAI research requires an interdisciplinary effort. The systematic review of XAI methods in different research areas is a stepping-stone to full understanding of the research field, with medicine reviews providing the bulk of knowledge on the topic at the time of writing. Considering the research gap regarding XAI applications along the IVC, this paper is one of the first attempts to provide an overview of XAI's current use within the insurance industry.

Author Contributions: Conceptualization, E.O., B.S., M.M. and M.C.; methodology, E.O. and B.S.; investigation, E.O. and B.S.; data curation, E.O.; writing—original draft preparation, E.O. and J.R.; writing—review and editing, B.S., M.M., M.C., J.R. and G.C.; visualization, E.O.; supervision, B.S., M.M. and M.C.; funding acquisition, G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Government of the Grand Duchy of Luxembourg, project LIAISON CVN-20210910RDI170010283594 RDI-REDDEX RDI REDIND.

Data Availability Statement: The authors can confirm that all relevant data are included in the article.

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

Abbreviations

AdaBoost	Adaptive Boosting
AI	Artificial Intelligence
ANN	Artificial Neural Network
BN	Bayesian Network
BPNN	Back Propagation Neural Network
CHAID	Chi-Squared Automatic Interaction Detection
CNN	Convolutional Neural Networks
CPLF	Cost-Sensitive Parallel Learning Framework
CRM	Customer Relationship Management
DFSVM	Dual Membership Fuzzy Support Vector Machine
DL	Deep Learning
ESIM	Evolutionary Support Vector Machine Inference Model
EvoDM	Evolutionary Data Mining
FL	Fuzzy Logic
GAM	Generalised Additive Model
GLM	Generalised Linear Model
HVSVM	Hull Vector Support Vector Machine
IoT	Internet of Things
IVC	Insurance Value Chain
KDD	
LASSO	Least Absolute Shrinkage and Selection Operator
MCAM	Markov Chain Approximation Method
ML	Machine Learning
NB	Naïve Bayes
NCA	Neighbourhood Component Analysis
NLP	Natural Language Processing
NN	Neural Network
PCA	Principal Component Analysis
RF	Random Forest
SBS	Sequential Backward Selection
SFS	Sequential Forward Selection
SHAP	Shapley Additive exPlanations
SOFM	Self-Organising Feature Map
SOM	Self-Organising Map
UBI	Usage-Based Insurance
WEKA	Waikato Environment for Knowledge Analysis
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting Algorithms

Appendix A. XAI Variables

Key XAI variables and criteria used both in the systematic review and throughout this paper are briefly outlined below as foundation for the paper's results and discussion. These XAI groupings are derived from Payrovnaziri et al. (2020), who synthesised the groupings from Du et al. (2019) and Carvalho et al. (2019)'s XAI reviews. This particular lens is suitable for the current study as key criteria for the determination of XAI's presence in AI method approaches.

Appendix A.1. Intrinsic vs. Post hoc *Interpretability*

The main differentiating aspect between an intrinsic and a post hoc interpretable explanation is whether interpretability is achieved through imposing constraints on the complexity of the model (intrinsic) or whether the models' explainability was analysed after training (post hoc) (Molnar 2019). Intrinsic methods primarily describe *how the model works*, which denotes a high degree of transparency in the model which is interpretable by itself (Lipton 2018; Rudin 2018). Lipton (2018) contrastingly summarises post hoc explainability as *what else can the model tell us*. Carvalho et al. (2019) clarifies that it is possible to apply post hoc methods to intrinsic models, as post hoc methods are usually derived from the main model. In summary, intrinsic models achieve their interpretability by incorporating it directly into their structures, while post hoc models require the creation of a second model to provide explanations for the existing model (Du et al. 2019).

Appendix A.2. Local vs. Global Interpretability

Local explanations primarily reveal the impact of input features on the overall model's prediction, while local explanations inspect model concepts to describe how the model works (Molnar 2019). Popular local explanation methods include: (1) the reporting of the decision path, (2) the assigning of credit to each input feature in the model and, (3) the application of several model-agnostic approaches which require the repeated execution of the model for each explanation (Baehrens et al. 2010; Lundberg et al. 2020; Štrumbelj and Kononenko 2014). A global explanation provides an overall view of the AI system, through listing the system's rules or features that eventually determine their predictive outcome (Lundberg et al. 2020). In terms of trustworthiness, Adadi and Berrada (2018) cite local explanations as more trustworthy than global ones as the latter connotes a sense of understanding of the mechanism by which the model works.

Appendix A.3. Model-Specific vs. Model-Agnostic Interpretation

Both model-specific and model-agnostic interpretation methods are derived from the above intrinsic vs. post hoc explainability criteria. As the name suggests, model-specific interpretation methods are limited to specific model classes as each method is based on a specific model's internals (Molnar 2019). Model-specific interpretability is by definition achieved from Intrinsically Interpretable Models (Adadi and Berrada 2018; Carvalho et al. 2019). Alternatively, model-agnostic methods can be applied to any model (black-box or otherwise) and are applied after the model has been trained (similar to post hoc interpretability). This method includes the analysis of relationships between the system's feature inputs and outputs, without sacrificing the model's predictive power (Carvalho et al. 2019; Lipton 2018). Table A1 below provides a summary of the above interpretability criteria and their generalised relationships.

Table A1. Association between XAI Interpretability Criteria where In-model and Post-model interpretability are defined using XAI variables.

In-Model	Intrinsic	Model-specific
Post-Model	Post hoc	Model-agnostic

Appendix B. Database of Reviewed Articles

Appendix B.1. Journal Articles Included in the Systematic Review

Reference	Title	Lead Author	Year	Source	Volume	Issue Number
Aggour et al. (2006)	Automating the underwriting of insurance applications	Aggour	2006	*AI Magazine*	27	3
Baecke and Bocca (2017)	The value of vehicle telematics data in insurance risk selection processes	Baecke	2017	*Decision Support Systems*	98	
Baudry and Robert (2019)	A machine learning approach for individual claims reserving in insurance	Baudry	2019	*Applied Stochastic Models in Business and Industry*	35	5
Belhadji et al. (2000)	A model for the detection of insurance fraud	Belhadji	2000	*The Geneva Papers on Risk and Insurance-Issues and Practice*	25	4
Benedek and László (2019)	Identifying Key Fraud Indicators in the Automobile Insurance Industry Using SQL Server Analysis Services	Benedek	2019	*Studia Universitatis Babes-Bolyai*	64	2
Bermúdez et al. (2008)	A Bayesian dichotomous model with asymmetric link for fraud in insurance	Bermúdez	2008	*Insurance: Mathematics and Economics*	42	2
Boodhun and Jayabalan (2018)	Risk prediction in life insurance industry using supervised learning algorithms	Boodhun	2018	*Complex & Intelligent Systems*	4	2
Carfora et al. (2019)	A "pay-how-you-drive" car insurance approach through cluster analysis	Carfora	2019	*Soft Computing*	23	9
Chang and Lai (2021)	A Neural Network-Based Approach in Predicting Consumers' Intentions of Purchasing Insurance Policies	Chang	2021	*Acta Informatica Pragensia*	10	2
Cheng et al. (2011)	Decision making for contractor insurance deductible using the evolutionary support vector machines inference model	Cheng	2011	*Expert Systems with Applications*	38	6
Cheng et al. (2020)	Optimal insurance strategies: A hybrid deep learning Markov chain approximation approach	Cheng	2020	*ASTIN Bulletin: The Journal of the IAA*	50	2
Christmann (2004)	An approach to model complex high–dimensional insurance data	Christmann	2004	*Allgemeines Statistisches Archiv*	88	4
David (2015)	Auto insurance premium calculation using generalized linear models	David	2015	*Procedia Economics and Finance*	20	
Delong and Wüthrich (2020)	Neural networks for the joint development of individual payments and claim incurred	Delong	2020	*Risks*	8	2
Denuit and Lang (2004)	Non-life rate-making with Bayesian GAMs	Denuit	2004	*Insurance: Mathematics and Economics*	35	3
Deprez et al. (2017)	Machine learning techniques for mortality modeling	Deprez	2017	*European Actuarial Journal*	7	2
Desik and Behera (2012)	Acquiring Insurance Customer: The CHAID Way	Desik	2012	*IUP Journal of Knowledge Management*	10	3
Desik et al. (2016)	Segmentation-Based Predictive Modeling Approach in Insurance Marketing Strategy	Desik	2016	*IUP Journal of Business Strategy*	13	2

Reference	Title	Lead Author	Year	Source	Volume	Issue Number
Devriendt et al. (2021)	Sparse regression with multi-type regularized feature modeling	Devriendt	2021	*Insurance: Mathematics and Economics*	96	
Duval and Pigeon (2019)	Individual loss reserving using a gradient boosting-based approach	Duval	2019	*Risks*	7	3
Fang et al. (2016)	Customer profitability forecasting using Big Data analytics: A case study of the insurance industry	Fang	2016	*Computers & Industrial Engineering*	101	
Frees and Valdez (2008)	Hierarchical insurance claims modeling	Frees	2008	*Journal of the American Statistical Association*	103	484
Gabrielli (2021)	An individual claims reserving model for reported claims	Gabrielli	2021	*European Actuarial Journal*	11	2
Gan (2013)	Application of data clustering and machine learning in variable annuity valuation	Gan	2013	*Journal of the American Statistical Association*	53	3
Gan and Valdez (2017)	Regression modeling for the valuation of large variable annuity portfolios	Gan	2018	*North American Actuarial Journal*	22	1
Ghorbani and Farzai (2018)	Fraud detection in automobile insurance using a data mining based approach	Ghorbani	2018	*International Journal of Mechatronics, Elektrical and Computer Technology (IJMEC)*	8	27
Gramegna and Giudici (2020)	Why to buy insurance? An Explainable Artificial Intelligence Approach	Gramegna	2020	*Risks*	8	4
Guelman (2012)	Gradient boosting trees for auto insurance loss cost modeling and prediction	Guelman	2012	*Expert Systems with Applications*	39	3
Gweon et al. (2020)	An effective bias-corrected bagging method for the valuation of large variable annuity portfolios	Gweon	2020	*ASTIN Bulletin: The Journal of the IAA*	50	3
Herland et al. (2018)	The detection of medicare fraud using machine learning methods with excluded provider labels	Herland	2018	*Journal of Big Data*	5	1
Huang and Meng (2019)	Automobile insurance classification ratemaking based on telematics driving data	Huang	2019	*Decision Support Systems*	127	
Ibiwoye et al. (2012)	Artificial neural network model for predicting insurance insolvency	Ibiwoye	2012	*International Journal of Management and Business Research*	2	1
Jain et al. (2019)	Assessing risk in life insurance using ensemble learning	Jain	2019	*Journal of Intelligent & Fuzzy Systems*	37	2
Jeong et al. (2018)	Association rules for understanding policyholder lapses	Jeong	2018	*Risks*	6	3
Jiang et al. (2018)	Cost-sensitive parallel learning framework for insurance intelligence operation	Jiang	2018	*IEEE Transactions on Industrial Electronics*	66	12
Jin et al. (2021)	A hybrid deep learning method for optimal insurance strategies: Algorithms and convergence analysis	Jin	2021	*Insurance: Mathematics and Economics*	96	

Reference	Title	Lead Author	Year	Source	Volume	Issue Number
Johnson and Khoshgoftaar (2019)	Medicare fraud detection using neural networks	Johnson	2019	*Journal of Big Data*	6	1
Joram et al. (2017)	A knowledge-based system for life insurance underwriting	Joram	2017	*International Journal of Information Technology and Computer Science*	3	
Karamizadeh and Zolfagharifar (2016)	Using the clustering algorithms and rule-based of data mining to identify affecting factors in the profit and loss of third party insurance, insurance company auto	Karamizadeh	2016	*Indian Journal of science and Technology*	9	7
Kašćelan et al. (2016)	A nonparametric data mining approach for risk prediction in car insurance: a case study from the Montenegrin market	Kašćelan	2016	*Economic research-Ekonomska istraživanja*	29	1
Khodairy and Abosamra (2021)	Driving Behavior Classification Based on Oversampled Signals of Smartphone Embedded Sensors Using an Optimized Stacked-LSTM Neural Networks	Khodairy	2021	*IEEE Access*	9	
Kiermayer and Weiß (2021)	Grouping of contracts in insurance using neural networks	Kiermayer	2021	*Scandinavian Actuarial Journal*	2021	4
Kose et al. (2015)	An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance	Kose	2015	*Applied Soft Computing*	36	
Kwak et al. (2020)	Driver Identification Based on Wavelet Transform Using Driving Patterns	Kwak	2020	*IEEE Transactions on Industrial Informatics*	17	4
Larivière and Van den Poel (2005)	Predicting customer retention and profitability by using random forests and regression forests techniques	Lariviere	2005	*Expert systems with applications*	29	2
Lee et al. (2020)	Actuarial applications of word embedding models	Lee	2020	*ASTIN Bulletin: The Journal of the IAA*	50	1
Li et al. (2018)	A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification	Li	2018	*Applied Soft Computing*	70	
Lin (2009)	Using neural networks as a support tool in the decision making for insurance industry	Lin	2009	*Expert Systems with Applications*	36	3
Lin et al. (2017)	An ensemble random forest algorithm for insurance big data analysis	Lin	2017	*IEEE Access*	5	
Liu et al. (2014)	Using multi-class AdaBoost tree for prediction frequency of auto insurance	Liu	2014	*Journal of Applied Finance and Banking*	4	5
Matloob et al. (2020)	Sequence Mining and Prediction-Based Healthcare Fraud Detection Methodology	Matloob	2020	*IEEE Access*	8	
Neumann et al. (2019)	Machine Learning-Based Predictions of Customers' Decisions in Car Insurance	Neumann	2019	*Applied Artificial Intelligence*	33	9

Reference	Title	Lead Author	Year	Source	Volume	Issue Number
Pathak et al. (2005)	A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims	Pathak	2005	*Managerial Auditing Journal*	20	6
Ravi et al. (2017)	Fuzzy formal concept analysis based opinion mining for CRM in financial services	Ravi	2017	*Applied Soft Computing*	60	
Sadreddini et al. (2021)	Cancel-for-Any-Reason Insurance Recommendation Using Customer Transaction-Based Clustering	Sadreddini	2021	*IEEE Access*	9	
Sakthivel and Rajitha (2017)	Artificial intelligence for estimation of future claim frequency in non-life insurance	Sakthivel	2017	*Global Journal of Pure and Applied Mathematics*	13	6
Sevim et al. (2016)	Risk Assessment for Accounting Professional Liability Insurance	Sevim	2016	*Sosyoekonomi*	24	29
Shah and Guez (2009)	Mortality forecasting using neural networks and an application to cause-specific data for insurance purposes	Shah	2009	*Journal of Forecasting*	28	6
Sheehan et al. (2017)	Semi-autonomous vehicle motor insurance: A Bayesian Network risk transfer approach	Sheehan	2017	*Transportation Research Part C: Emerging Technologies*	82	
Siami et al. (2020)	A mobile telematics pattern recognition framework for driving behavior extraction	Siami	2020	*IEEE Transactions on Intelligent Transportation Systems*	22	3
Smith et al. (2000)	An analysis of customer retention and insurance claim patterns using data mining: A case study	Smith	2000	*Journal of the Operational Research Society*	51	5
Smyth and Jørgensen (2002)	Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling	Smyth	2002	*ASTIN Bulletin: The Journal of the IAA*	32	1
Sun et al. (2018)	Abnormal group-based joint medical fraud detection	Sun	2018	*IEEE Access*	7	
Tillmanns et al. (2017)	How to separate the wheat from the chaff: Improved variable selection for new customer acquisition	Tillmanns	2017	*Journal of Marketing*	81	2
Vaziri and Beheshtinia (2016)	A holistic fuzzy approach to create competitive advantage via quality management in services industry (case study: life-insurance services)	Vaziri	2016	*Management decision*	54	8
Viaene et al. (2002)	Auto claim fraud detection using Bayesian learning neural networks	Viaene	2002	*Expert Systems with Applications*	29	3
Viaene et al. (2004)	A case study of applying boosting Naive Bayes to claim fraud diagnosis	Viaene	2004	*Journal of Risk and Insurance*	69	3
Viaene et al. (2005)	A case study of applying boosting Naive Bayes to claim fraud diagnosis	Viaene	2005	*IEEE Transactions on Knowledge and Data Engineering*	16	5

Reference	Title	Lead Author	Year	Source	Volume	Issue Number
Wang (2020)	Research on the Features of Car Insurance Data Based on Machine Learning	Wang	2020	*Procedia Computer Science*	166	
Wang and Xu (2018)	Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud	Wang	2018	*Decision Support Systems*	105	
Wei and Dan (2019)	Market fluctuation and agricultural insurance forecasting model based on machine learning algorithm of parameter optimization	Wei	2019	*Journal of Intelligent & Fuzzy Systems*	37	5
Wüthrich (2020)	Bias regularization in neural network models for general insurance pricing	Wüthrich	2020	*European Actuarial Journal*	10	1
Yan et al. (2020a)	Research on the UBI Car Insurance Rate Determination Model Based on the CNN-HVSVM Algorithm	Yan	2020	*IEEE Access*	8	
Yan et al. (2020b)	Improved adaptive genetic algorithm for the vehicle Insurance Fraud Identification Model based on a BP Neural Network	Yan	2020	*Theoretical Computer Science*	817	
Yang et al. (2006)	Extracting actionable knowledge from decision trees	Yang	2006	*IEEE Transactions on Knowledge and data Engineering*	19	1
Yang et al. (2018)	Insurance premium prediction via gradient tree-boosted Tweedie compound Poisson models	Yang	2018	*Journal of Business & Economic Statistics*	36	3
Yeo et al. (2002)	A mathematical programming approach to optimise insurance premium pricing within a data mining framework	Yeo	2002	*Journal of the Operational research Society*	53	11

Appendix B.2. Conference Papers Included in the Systematic Review

Reference	Title	Lead Author	Year	Source
Alshamsi (2014)	Predicting car insurance policies using random forest	Alshamsi	2014	2014 10th International Conference on Innovations in Information Technology (IIT)
Bian et al. (2018)	Good drivers pay less: A study of usage-based vehicle insurance models	Bian	2018	Transportation research part A: policy and practice
Biddle et al. (2018)	Transportation research part A: policy and practice	Biddle	2018	Australasian Database Conference
Bonissone et al. (2002)	Evolutionary optimization of fuzzy decision systems for automated insurance underwriting	Bonissone	2002	2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems
Bove et al. (2021)	Contextualising local explanations for non-expert users: an XAI pricing interface for insurance	Bove	2021	IUI Workshops
Cao and Zhang (2019)	Using PCA to improve the detection of medical insurance fraud in SOFM neural networks	Cao	2019	Proceedings of the 2019 3rd International Conference on Management Engineering, Software Engineering and Service Sciences
Dhieb et al. (2019)	Extreme gradient boosting machine learning algorithm for safe auto insurance operations	Dhieb	2019	2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES)
Gan and Huang (2017)	A data mining framework for valuing large portfolios of variable annuities	Gan	2017	Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
Ghani and Kumar (2011)	Interactive learning for efficiently detecting errors in insurance claims	Ghani	2011	Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining
Kieu et al. (2018)	Distinguishing trajectories from different drivers using incompletely labeled trajectories	Kieu	2018	Proceedings of the 27th ACM international conference on information and knowledge management
Kowshalya and Nandhini (2018)	Predicting fraudulent claims in automobile insurance	Kowshalya	2018	2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)
Kumar et al. (2010)	Data mining to predict and prevent errors in health insurance claims processing	Kumar	2010	Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining
Kyu and Woraratpanya (2020)	Car Damage Detection and Classification	Kyu	2020	Proceedings of the 11th International Conference on Advances in Information Technology
Lau and Tripathi (2011)	Mine your business—A novel application of association rules for insurance claims analytics	Lau	2011	CAS E-Forum. Arlington: Casualty Actuarial Society
Liu and Chen (2012)	Application of evolutionary data mining algorithms to insurance fraud prediction	Liu	2012	Proceedings of 2012 4th International Conference on Machine Learning and Computing IPCSIT
Morik et al. (2002)	End-user access to multiple sources-Incorporating knowledge discovery into knowledge management	Morik	2002	International Conference on Practical Aspects of Knowledge Management

Reference	Title	Lead Author	Year	Source
Samonte et al. (2018)	ICD-9 tagging of clinical notes using topical word embedding	Samonte	2018	Proceedings of the 2018 International Conference on Internet and e-Business
Sohail et al. (2021)	Feature importance analysis for customer management of insurance products	Sohail	2021	2021 International Joint Conference on Neural Networks (ICJNN)
Supraja and Saritha (2017)	Robust fuzzy rule based technique to detect frauds in vehicle insurance	Supraja	2017	2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)
Tao et al. (2012)	Insurance fraud identification research based on fuzzy support vector machine with dual membership	Tao	2012	2012 International Conference on Information Management, Innovation Management and Industrial Engineering
Vassiljeva et al. (2017)	Computational intelligence approach for estimation of vehicle insurance risk level	Vassiljeva	2017	2017 International Joint Conference on Neural Networks (IJCNN)
Verma et al. (2017)	Fraud detection and frequent pattern matching in insurance claims using data mining techniques	Verma	2017	2017 Tenth International Conference on Contemporary Computing (IC3)
Xu et al. (2011)	Random rough subspace based neural network ensemble for insurance fraud detection	Xu	2011	2011 Fourth International Joint Conference on Computational Sciences and Optimization
Yan and Bonissone (2006)	Designing a Neural Network Decision System for Automated Insurance Underwriting	Yan	2006	Insurance Studies
Zahi and Achchab (2019)	Clustering of the population benefiting from health insurance using *k*-means	Zahi	2019	Proceedings of the 4th International Conference on Smart City Applications
Zhang and Kong (2020)	Dynamic estimation model of insurance product recommendation based on Naive Bayesian model	Zhang	2020	Proceedings of the 2020 International Conference on Cyberspace Innovation of Advanced Technologies

Notes

1 The five XAI categories used were introduced to XAI literature by Payrovnaziri et al. (2020), adapted from research conducted by Du et al. (2019) and Carvalho et al. (2019).

2 Searched 'The ACM Guide to Computing Literature'.

3 'Articles' throughout this review refers to both academic articles and conference papers.

4 Several such surveys and reviews are discussed in Section 2.2.

5 Searched 'The ACM Guide to Computing Literature'.

References

Adadi, Amina, and Mohammed Berrada. 2018. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* 6: 52138–60. [CrossRef]

Aggour, Kareem S., Piero P. Bonissone, William E. Cheetham, and Richard P. Messmer. 2006. Automating the underwriting of insurance applications. *AI Magazine* 27: 36–36.

Alshamsi, Asma S. 2014. Predicting car insurance policies using random forest. Paper presented at the 2014 10th International Conference on Innovations in Information Technology (IIT), Al Ain, United Arab Emirates, November 9–11.

Al-Shedivat, Maruan, Avinava Dubey, and Eric P. Xing. 2020. Contextual Explanation Networks. *Journal of Machine Learning Research* 21: 194:1–94:44.

Andrew, Jane, and Max Baker. 2021. The general data protection regulation in the age of surveillance capitalism. *Journal of Business Ethics* 168: 565–78. [CrossRef]

Anjomshoae, Sule, Amro Najjar, Davide Calvaresi, and Kary Främling. 2019. Explainable agents and robots: Results from a systematic literature review. Paper presented at the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17.

Antoniadi, Anna Markella, Yuhan Du, Yasmine Guendouz, Lan Wei, Claudia Mazo, Brett A Becker, and Catherine Mooney. 2021. Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review. *Applied Sciences* 11: 5088. [CrossRef]

Arrieta, Alejandro Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García, Sergio Gil-López, Daniel Molina, and Richard Benjamins. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58: 82–115. [CrossRef]

Baecke, Philippe, and Lorenzo Bocca. 2017. The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems* 98: 69–79. [CrossRef]

Baehrens, David, Timon Schroeter, Stefan Harmeling, Motoaki Kawanabe, Katja Hansen, and Klaus-Robert Müller. 2010. How to explain individual classification decisions. *The Journal of Machine Learning Research* 11: 1803–31.

Balasubramanian, Ramnath, Ari Libarikian, and Doug McElhaney. 2018. *Insurance 2030—The Impact of AI on the Future of Insurance*. New York: McKinsey & Company.

Barocas, Solon, and Andrew D Selbst. 2016. Big data's disparate impact. *California Law Review* 104: 671. [CrossRef]

Barry, Laurence, and Arthur Charpentier. 2020. Personalization as a promise: Can Big Data change the practice of insurance? *Big Data & Society* 7: 2053951720935143.

Baser, Furkan, and Aysen Apaydin. 2010. Calculating insurance claim reserves with hybrid fuzzy least squares regression analysis. *Gazi University Journal of Science* 23: 163–70.

Baudry, Maximilien, and Christian Y. Robert. 2019. A machine learning approach for individual claims reserving in insurance. *Applied Stochastic Models in Business and Industry* 35: 1127–55. [CrossRef]

Bayamlıoğlu, Emre. 2021. The right to contest automated decisions under the General Data Protection Regulation: Beyond the so-called "right to explanation". *Regulation & Governance* 16: 1058–78.

Bean, Randy. 2021. Transforming the Insurance Industry with Big Data, Machine Learning and AI. *Forbes*. July 6. Available online: https://www.forbes.com/sites/randybean/2021/07/06/transforming-the-insurance-industry-with-big-data-machine-learning-and-ai/?sh=4004a662f8a6 (accessed on 11 August 2021).

Beck, Hall P., Mary T. Dzindolet, and Linda G. Pierce. 2002. Operators' automation usage decisions and the sources of misuse and disuse. In *Advances in Human Performance and Cognitive Engineering Research*. Bingley: Emerald Group Publishing Limited.

Belhadji, El Bachir, George Dionne, and Faouzi Tarkhani. 2000. A model for the detection of insurance fraud. *The Geneva Papers on Risk and Insurance-Issues and Practice* 25: 517–38. [CrossRef]

Benedek, Botond, and Ede László. 2019. Identifying Key Fraud Indicators in the Automobile Insurance Industry Using SQL Server Analysis Services. *Studia Universitatis Babes-Bolyai* 64: 53–71. [CrossRef]

Bermúdez, Lluís, José María Pérez, Mercedes Ayuso, Esther Gómez, and Francisco. J. Vázquez. 2008. A Bayesian dichotomous model with asymmetric link for fraud in insurance. *Insurance: Mathematics and Economics* 42: 779–86. [CrossRef]

Bian, Yiyang, Chen Yang, J. Leon Zhao, and Liang Liang. 2018. Good drivers pay less: A study of usage-based vehicle insurance models. *Transportation Research Part A: Policy and Practice* 107: 20–34. [CrossRef]

Biddle, Rhys, Shaowu Liu, and Guandong Xu. 2018. Automated Underwriting in Life Insurance: Predictions and Optimisation (Industry Track). Paper presented at Australasian Database Conference, Gold Coast, QLD, Australia, May 24–27.

Biecek, Przemysław, Marcin Chlebus, Janusz Gajda, Alicja Gosiewska, Anna Kozak, Dominik Ogonowski, Jakub Sztachelski, and Piotr Wojewnik. 2021. Enabling Machine Learning Algorithms for Credit Scoring—Explainable Artificial Intelligence (XAI) methods for clear understanding complex predictive models. *arXiv* arXiv:2104.06735.

Biran, Or, and Courtenay Cotton. 2017. Explanation and justification in machine learning: A survey. Paper presented at the IJCAI-17 Workshop on Explainable AI (XAI), Melbourne, VIC, Australia, August 19–21.

Blier-Wong, Christopher, Hélène Cossette, Luc Lamontagne, and Etienne Marceau. 2021. Machine Learning in P&C Insurance: A Review for Pricing and Reserving. *Risks* 9: 4.

Bonissone, Piero. P., Raj Subbu, and Kareem S. Aggour. 2002. Evolutionary optimization of fuzzy decision systems for automated insurance underwriting. Paper presented at the 2002 IEEE World Congress on Computational Intelligence, 2002 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE'02. Proceedings (Cat. No. 02CH37291), Honolulu, HI, USA, May 12–17.

Boodhun, Noorhannah, and Manoj Jayabalan. 2018. Risk prediction in life insurance industry using supervised learning algorithms. *Complex & Intelligent Systems* 4: 145–54.

Bove, Clara, Jonathan Aigrain, Marie-Jeanne Lesot, Charles Tijus, and Marcin Detyniecki. 2021. Contextualising local explanations for non-expert users: An XAI pricing interface for insurance. Paper presented at the IUI Workshops, College Station, TX, USA, April 13–17.

Burgt, Joost van der. 2020. Explainable AI in banking. *Journal of Digital Banking* 4: 344–50.

Burrell, Jenna. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* 3: 2053951715622512.

Bussmann, Niklas, Paolo Giudici, Dimitri Marinelli, and Jochen Papenbrock. 2020. Explainable ai in fintech risk management. *Frontiers in Artificial Intelligence* 3: 26. [CrossRef] [PubMed]

Cao, Hongfei, and Runtong Zhang. 2019. Using PCA to improve the detection of medical insurance fraud in SOFM neural networks. Paper presented at the 2019 3rd International Conference on Management Engineering, Software Engineering and Service Sciences, Wuhan, China, January 12–14.

Carabantes, Manuel. 2020. Black-box artificial intelligence: An epistemological and critical analysis. *AI & Society* 35: 309–17.

Carfora, Maria Francesca, Fabio Martinelli, Francesco Mercaldo, Vittoria Nardone, Albina Orlando, Antonella Santone, and Gigliola Vaglini. 2019. A "pay-how-you-drive" car insurance approach through cluster analysis. *Soft Computing* 23: 2863–75. [CrossRef]

Carvalho, Diogo V., Eduardo M. Pereira, and Jaime S. Cardoso. 2019. Machine learning interpretability: A survey on methods and metrics. *Electronics* 8: 832. [CrossRef]

Cevolini, Alberto, and Elena Esposito. 2020. From pool to profile: Social consequences of algorithmic prediction in insurance. *Big Data & Society* 7: 2053951720939228.

Chang, Wen Teng, and Kee Huong Lai. 2021. A Neural Network-Based Approach in Predicting Consumers' Intentions of Purchasing Insurance Policies. *Acta Informatica Pragensia* 10: 138–54. [CrossRef]

Chen, Irene Y., Peter Szolovits, and Marzyeh Ghassemi. 2019. Can AI help reduce disparities in general medical and mental health care? *AMA Journal of Ethics* 21: 167–79.

Cheng, Min-Yuan, Hsien-Sheng Peng, Yu-Wei Wu, and Yi-Hung Liao. 2011. Decision making for contractor insurance deductible using the evolutionary support vector machines inference model. *Expert Systems with Applications* 38: 6547–55. [CrossRef]

Cheng, Xiang, Zhuo Jin, and Hailiang Yang. 2020. Optimal insurance strategies: A hybrid deep learning Markov chain approximation approach. *ASTIN Bulletin: The Journal of the IAA* 50: 449–77. [CrossRef]

Chi, Oscar Hengxuan, Gregory Denton, and Dogan Gursoy. 2020. Artificially intelligent device use in service delivery: A systematic review, synthesis, and research agenda. *Journal of Hospitality Marketing & Management* 29: 757–86.

Christmann, Andreas. 2004. An approach to model complex high–dimensional insurance data. *Allgemeines Statistisches Archiv* 88: 375–96. [CrossRef]

Clinciu, Miruna-Adriana, and Helen Hastie. 2019. A survey of explainable AI terminology. Paper presented at the 1st Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence (NL4XAI 2019), Tokyo, Japan, October 29–November 1.

Coalition Against Insurance Fraud. 2020. *Artificial Intelligence & Insurance Fraud*. Washington, DC: Coalition Against Insurance Fraud. Available online: https://insurancefraud.org/wp-content/uploads/Artificial-Intelligence-and-Insurance-Fraud-2020.pdf (accessed on 2 May 2021).

Colaner, Nathan. 2022. Is explainable artificial intelligence intrinsically valuable? *AI & Society* 37: 231–38.

Confalonieri, Roberto, Ludovik Coba, Benedikt Wagner, and Tarek R. Besold. 2021. A historical perspective of explainable Artificial Intelligence. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 11: e1391. [CrossRef]

Daniels, Norman. 2011. The ethics of health reform: Why we should care about who is missing coverage. *Connecticut Law Review* 44: 1057.

Danilevsky, Marina, Kun Qian, Ranit Aharonov, Yannis Katsis, Ban Kawas, and Prithviraj Sen. 2020. A survey of the state of explainable AI for natural language processing. *arXiv* arXiv:2010.00711.

Das, Arun, and Paul Rad. 2020. Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv* arXiv:2006.11371.

David, Mihaela. 2015. Auto insurance premium calculation using generalized linear models. *Procedia Economics and Finance* 20: 147–56. [CrossRef]

Delong, Łukasz, and Mario V. Wüthrich. 2020. Neural networks for the joint development of individual payments and claim incurred. *Risks* 8: 33. [CrossRef]

Demajo, Lara Marie, Vince Vella, and Alexiei Dingli. 2020. Explainable AI for interpretable credit scoring. *arXiv* arXiv:2012.03749.

Denuit, Michel, and Stefan Lang. 2004. Non-life rate-making with Bayesian GAMs. *Insurance: Mathematics and Economics* 35: 627–47. [CrossRef]

Deprez, Philippe, Pavel V. Shevchenko, and Mario V. Wüthrich. 2017. Machine learning techniques for mortality modeling. *European Actuarial Journal* 7: 337–52. [CrossRef]

Desik, P. H. Anantha, Samarendra Behera, Prashanth Soma, and Nirmala Sundari. 2016. Segmentation-Based Predictive Modeling Approach in Insurance Marketing Strategy. *IUP Journal of Business Strategy* 13: 35–45.

Desik, P. H. Anantha, and Samarendra Behera. 2012. Acquiring Insurance Customer: The CHAID Way. *IUP Journal of Knowledge Management* 10: 7–13.

Devriendt, Sander, Katrien Antonio, Tom Reynkens, and Roel Verbelen. 2021. Sparse regression with multi-type regularized feature modeling. *Insurance: Mathematics and Economics* 96: 248–61. [CrossRef]

Dhieb, Najmeddine, Hakim Ghazzai, Hichem Besbes, and Yehia Massoud. 2019. Extreme gradient boosting machine learning algorithm for safe auto insurance operations. Paper presented at the 2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Cairo, Egypt, September 4–6.

Diprose, William K., Nicholas Buist, Ning Hua, Quentin Thurier, George Shand, and Reece Robinson. 2020. Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. *Journal of the American Medical Informatics Association* 27: 592–600. [CrossRef]

Doshi-Velez, Finale, and Been Kim. 2017. Towards a rigorous science of interpretable machine learning. *arXiv* arXiv:1702.08608.

Došilović, Filip Karlo, Mario Brčić, and Nikica Hlupić. 2018. Explainable artificial intelligence: A survey. Paper presented at the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, May 21–25.

Du, Mengnan, Ninghao Liu, and Xia Hu. 2019. Techniques for interpretable machine learning. *Communications of the ACM* 63: 68–77. [CrossRef]

Duval, Francis, and Mathieu Pigeon. 2019. Individual loss reserving using a gradient boosting-based approach. *Risks* 7: 79.

Eckert, Theresa, and Stefan Hüsig. 2021. Innovation portfolio management: A systematic review and research agenda in regards to digital service innovations. *Management Review Quarterly* 72: 187–230. [CrossRef]

EIOPA. 2021. *Artificial Intelligence Governance Principles: Towards Ethical and Trustworthy Artificial Intelligence in the European Insurance Sector*. Luxembourg: European Insurance and Occupational Pensions Authority (EIOPA).

Eling, Martin, Davide Nuessle, and Julian Staubli. 2021. The impact of artificial intelligence along the insurance value chain and on the insurability of risks. *The Geneva Papers on Risk and Insurance-Issues and Practice* 47: 205–41. [CrossRef]

EU. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). Available online: https://eur-lex.europa.eu/eli/reg/2016/679/oj (accessed on 27 June 2020).

Fang, Kuangnan, Yefei Jiang, and Malin Song. 2016. Customer profitability forecasting using Big Data analytics: A case study of the insurance industry. *Computers & Industrial Engineering* 101: 554–64.

Felzmann, Heike, Eduard Fosch Villaronga, Christoph Lutz, and Aurelia Tamò-Larrieux. 2019. Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society* 6: 2053951719860542.

Ferguson, Niall. 2008. *The Ascent of Money: A Financial History of the World*. London: Penguin.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, and Francesca Rossi. 2018. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines* 28: 689–707. [CrossRef]

Ford, Martin. 2018. *Architects of Intelligence: The Truth about AI from the People Building it*. Birmingham: Packt Publishing Ltd.

Fox, Maria, Derek Long, and Daniele Magazzeni. 2017. Explainable planning. *arXiv* arXiv:1709.10256.

Frees, Edward W., and Emiliano A. Valdez. 2008. Hierarchical insurance claims modeling. *Journal of the American Statistical Association* 103: 1457–69. [CrossRef]

Freitas, Alex A. 2014. Comprehensible classification models: A position paper. *ACM SIGKDD Explorations Newsletter* 15: 1–10. [CrossRef]

Gabrielli, Andrea. 2021. An individual claims reserving model for reported claims. *European Actuarial Journal* 11: 541–77. [CrossRef]

Gade, Krishna, Sahin Cem Geyik, Krishnaram Kenthapadi, Varun Mithal, and Ankur Taly. 2019. Explainable AI in industry. Paper presented at the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, August 4–8.

Gan, Guojun, and Emiliano A. Valdez. 2017. Valuation of large variable annuity portfolios: Monte Carlo simulation and synthetic datasets. *Dependence Modeling* 5: 354–74. [CrossRef]

Gan, Guojun, and Jimmy Xiangji Huang. 2017. A data mining framework for valuing large portfolios of variable annuities. Paper presented at the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13–17.

Gan, Guojun. 2013. Application of data clustering and machine learning in variable annuity valuation. *Insurance: Mathematics and Economics* 53: 795–801.

Ghani, Rayid, and Mohit Kumar. 2011. Interactive learning for efficiently detecting errors in insurance claims. Paper presented at the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 21–24.

Ghorbani, Ali, and Sara Farzai. 2018. Fraud detection in automobile insurance using a data mining based approach. *International Journal of Mechatronics, Elektrical and Computer Technology (IJMEC)* 8: 3764–71.

GlobalData. 2021. *Artificial Intelligence (AI) in Insurance—Thematic Research*. London: GlobalData.

Goddard, Michelle. 2017. The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research* 59: 703–5. [CrossRef]

Goodman, Bryce, and Seth Flaxman. 2017. European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine* 38: 50–57. [CrossRef]

Gramegna, Alex, and Paolo Giudici. 2020. Why to Buy Insurance? An Explainable Artificial Intelligence Approach. *Risks* 8: 137. [CrossRef]

Gramespacher, Thomas, and Jan-Alexander Posth. 2021. Employing explainable AI to optimize the return target function of a loan portfolio. *Frontiers in Artificial Intelligence* 4: 693022. [CrossRef]

Grant, Eric. 2012. *The Social and Economic Value of Insurance*. Geneva: The Geneva Association (The International Association for the Study of Insurance Economics). Available online: https://www.genevaassociation.org/sites/default/files/research-topics-document-type/pdf_public/ga2012-the_social_and_economic_value_of_insurance.pdf (accessed on 3 July 2020).

Grize, Yves-Laurent, Wolfram Fischer, and Christian Lützelschwab. 2020. Machine learning applications in nonlife insurance. *Applied Stochastic Models in Business and Industry* 36: 523–37. [CrossRef]

Guelman, Leo. 2012. Gradient boosting trees for auto insurance loss cost modeling and prediction. *Expert Systems with Applications* 39: 3659–67. [CrossRef]

Gweon, Hyukjun, Shu Li, and Rogemar Mamon. 2020. An effective bias-corrected bagging method for the valuation of large variable annuity portfolios. *ASTIN Bulletin: The Journal of the IAA* 50: 853–71. [CrossRef]

Hadji Misheva, Branka, Ali Hirsa, Joerg Osterrieder, Onkar Kulkarni, and Stephen Fung Lin. 2021. Explainable AI in Credit Risk Management. *Credit Risk Management*, March 1.

Hawley, Katherine. 2014. Trust, distrust and commitment. *Noûs* 48: 1–20. [CrossRef]

Henckaerts, Roel, Katrien Antonio, and Marie-Pier Côté. 2020. Model-Agnostic Interpretable and Data-driven suRRogates suited for highly regulated industries. *Stat* 1050: 14.

Henckaerts, Roel, Marie-Pier Côté, Katrien Antonio, and Roel Verbelen. 2021. Boosting insights in insurance tariff plans with tree-based machine learning methods. *North American Actuarial Journal* 25: 255–85. [CrossRef]

Herland, Matthew, Taghi M. Khoshgoftaar, and Richard A. Bauder. 2018. Big data fraud detection using multiple medicare data sources. *Journal of Big Data* 5: 1–21. [CrossRef]

Hinton, Geoffrey, Oriol Vinyals, and Jeff Dean. 2015. Distilling the knowledge in a neural network. *arXiv* arXiv:1503.02531.

Hoffman, Robert R. 2017. A taxonomy of emergent trusting in the human–machine relationship. In *Cognitive Systems Engineering: The Future for a Changing World*. Boca Raton: CRC Press, pp. 137–64.

Hoffman, Robert R., Shane T. Mueller, Gary Klein, and Jordan Litman. 2018. Metrics for explainable AI: Challenges and prospects. *arXiv* arXiv:1812.04608.

Hollis, Aidan, and Jason Strauss. 2007. *Privacy, Driving Data and Automobile Insurance: An Economic Analysis*. Munich: University Library of Munich.

Honegger, Milo. 2018. Shedding light on black box machine learning algorithms: Development of an axiomatic framework to assess the quality of methods that explain individual predictions. *arXiv* arXiv:1808.05054.

Huang, Yifan, and Shengwang Meng. 2019. Automobile insurance classification ratemaking based on telematics driving data. *Decision Support Systems* 127: 113156. [CrossRef]

Ibiwoye, Ade, Olawale Olaniyi Ajibola, and Ashim Babatunde Sogunro. 2012. Artificial neural network model for predicting insurance insolvency. *International Journal of Management and Business Research* 2: 59–68.

Islam, Sheikh Rabiul, William Eberle, and Sheikh K. Ghafoor. 2020. Towards quantification of explainability in explainable artificial intelligence methods. Paper presented at the Thirty-Third International Flairs Conference, North Miami Beach, FL, USA, May 17–20.

Jacovi, Alon, Ana Marasović, Tim Miller, and Yoav Goldberg. 2021. Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. Paper presented at the 2021 ACM Conference on Fairness, Accountability, and Transparency, Toronto, ON, Canada, March 3–10.

Jain, Rachna, Jafar A. Alzubi, Nikita Jain, and Pawan Joshi. 2019. Assessing risk in life insurance using ensemble learning. *Journal of Intelligent & Fuzzy Systems* 37: 2969–80.

Jeong, Himchan, Guojun Gan, and Emiliano A. Valdez. 2018. Association rules for understanding policyholder lapses. *Risks* 6: 69. [CrossRef]

Jiang, Xinxin, Shirui Pan, Guodong Long, Fei Xiong, Jing Jiang, and Chengqi Zhang. 2018. Cost-sensitive parallel learning framework for insurance intelligence operation. *IEEE Transactions on Industrial Electronics* 66: 9713–23. [CrossRef]

Jin, Zhuo, Hailiang Yang, and George Yin. 2021. A hybrid deep learning method for optimal insurance strategies: Algorithms and convergence analysis. *Insurance: Mathematics and Economics* 96: 262–75. [CrossRef]

Johnson, Justin M., and Taghi M. Khoshgoftaar. 2019. Medicare fraud detection using neural networks. *Journal of Big Data* 6: 1–35. [CrossRef]

Joram, Mutai K., Bii K. Harrison, and Kiplang'at N. Joseph. 2017. A knowledge-based system for life insurance underwriting. *International Journal of Information Technology and Computer Science* 3: 40–49. [CrossRef]

Josephson, John R., and Susan G. Josephson. 1996. *Abductive Inference: Computation, Philosophy, Technology*. Cambridge: Cambridge University Press.

Karamizadeh, Faramarz, and Seyed Ahad Zolfagharifar. 2016. Using the clustering algorithms and rule-based of data mining to identify affecting factors in the profit and loss of third party insurance, insurance company auto. *Indian Journal of Science and Technology* 9: 1–9. [CrossRef]

Kašćelan, Vladimir, Ljiljana Kašćelan, and Milijana Novović Burić. 2016. A nonparametric data mining approach for risk prediction in car insurance: A case study from the Montenegrin market. *Economic Research-Ekonomska Istraživanja* 29: 545–58. [CrossRef]

Keller, Benno, Martin Eling, Hato Schmeiser, Markus Christen, and Michele Loi. 2018. *Big Data and Insurance: Implications for Innovation, Competition and Privacy*. Geneva: Geneva Association-International Association for the Study of Insurance.

Kelley, Kevin H., Lisa M. Fontanetta, Mark Heintzman, and Nikki Pereira. 2018. Artificial intelligence: Implications for social inflation and insurance. *Risk Management and Insurance Review* 21: 373–87. [CrossRef]

Khodairy, Moayed A., and Gibrael Abosamra. 2021. Driving Behavior Classification Based on Oversampled Signals of Smartphone Embedded Sensors Using an Optimized Stacked-LSTM Neural Networks. *IEEE Access* 9: 4957–72. [CrossRef]

Khuong, Mai Ngoc, and Tran Manh Tuan. 2016. A new neuro-fuzzy inference system for insurance forecasting. Paper presented at the International Conference on Advances in Information and Communication Technology, Bikaner, India, August 12–13.

Kiermayer, Mark, and Christian Weiß. 2021. Grouping of contracts in insurance using neural networks. *Scandinavian Actuarial Journal* 2021: 295–322. [CrossRef]

Kieu, Tung, Bin Yang, Chenjuan Guo, and Christian S. Jensen. 2018. Distinguishing trajectories from different drivers using incompletely labeled trajectories. Paper presented at the 27th ACM International Conference on Information and Knowledge Management, Torino, Italy, October 22–26.

Kim, Hyong, and Errol Gardner. 2015. The Science of Winning in Financial Services-Competing on Analytics: Opportunities to Unlock the Power of Data. *Journal of Financial Perspectives* 3: 1–34.

Kopitar, Leon, Leona Cilar, Primoz Kocbek, and Gregor Stiglic. 2019. Local vs. global interpretability of machine learning models in type 2 diabetes mellitus screening. In *Artificial Intelligence in Medicine: Knowledge Representation and Transparent and Explainable Systems*. Berlin: Springer, pp. 108–19.

Kose, Ilker, Mehmet Gokturk, and Kemal Kilic. 2015. An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. *Applied Soft Computing* 36: 283–99. [CrossRef]

Koster, Harold. 2020. Towards better implementation of the European Union's anti-money laundering and countering the financing of terrorism framework. *Journal of Money Laundering Control* 23: 379–86. [CrossRef]

Koster, Olivier, Ruud Kosman, and Joost Visser. 2021. A Checklist for Explainable AI in the Insurance Domain. Paper presented at the International Conference on the Quality of Information and Communications Technology, Algarve, Portugal, September 8–11.

Kowshalya, G., and M. Nandhini. 2018. Predicting fraudulent claims in automobile insurance. Paper presented at the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, April 20–21.

Krafft, Peaks, Meg Young, Michael Katell, Karen Huang, and Ghislain Bugingo. 2020. Defining AI in policy versus practice. Paper presented at the AAAI/ACM Conference on AI, Ethics, and Society, New York, NY, USA, February 7–8.

Kumar, Akshi, Shubham Dikshit, and Victor Hugo C. Albuquerque. 2021. Explainable Artificial Intelligence for Sarcasm Detection in Dialogues. *Wireless Communications and Mobile Computing* 2021: 2939334. [CrossRef]

Kumar, Mohit, Rayid Ghani, and Zhu-Song Mei. 2010. Data mining to predict and prevent errors in health insurance claims processing. Paper presented at the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 24–28.

Kuo, Kevin, and Daniel Lupton. 2020. Towards Explainability of Machine Learning Models in Insurance Pricing. *arXiv* arXiv:2003.10674.

Kute, Dattatray V., Biswajeet Pradhan, Nagesh Shukla, and Abdullah Alamri. 2021. Deep learning and explainable artificial intelligence techniques applied for detecting money laundering—A critical review. *IEEE Access* 9: 82300–17. [CrossRef]

Kwak, Byung Il, Mee Lan Han, and Huy Kang Kim. 2020. Driver Identification Based on Wavelet Transform Using Driving Patterns. *IEEE Transactions on Industrial Informatics* 17: 2400–10. [CrossRef]

Kyu, Phyu Mar, and Kuntpong Woraratpanya. 2020. Car Damage Detection and Classification. Paper presented at the 11th International Conference on Advances in Information Technology, Bangkok, Thailand, July 1–3.

Larivière, Bart, and Dirk Van den Poel. 2005. Predicting customer retention and profitability by using random forests and regression forests techniques. *Expert Systems with Applications* 29: 472–84. [CrossRef]

Lau, Lucas, and Arun Tripathi. 2011. Mine your business—A novel application of association rules for insurance claims analytics. In *CAS E-Forum*. Arlington: Casualty Actuarial Society.

Lee, Gee Y., Scott Manski, and Tapabrata Maiti. 2020. Actuarial applications of word embedding models. *ASTIN Bulletin: The Journal of the IAA* 50: 1–24. [CrossRef]

Li, Yaqi, Chun Yan, Wei Liu, and Maozhen Li. 2018. A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification. *Applied Soft Computing* 70: 1000–9. [CrossRef]

Liao, Shu-Hsien, Pei-Hui Chu, and Pei-Yuan Hsiao. 2012. Data mining techniques and applications–A decade review from 2000 to 2011. *Expert Systems with Applications* 39: 11303–11. [CrossRef]

Lin, Chaohsin. 2009. Using neural networks as a support tool in the decision making for insurance industry. *Expert Systems with Applications* 36: 6914–17. [CrossRef]

Lin, Justin, and Ha-Joon Chang. 2009. Should Industrial Policy in developing countries conform to comparative advantage or defy it? A debate between Justin Lin and Ha-Joon Chang. *Development Policy Review* 27: 483–502. [CrossRef]

Lin, Weiwei, Ziming Wu, Longxin Lin, Angzhan Wen, and Jin Li. 2017. An ensemble random forest algorithm for insurance big data analysis. *IEEE Access* 5: 16568–75. [CrossRef]

Lipton, Zachary C. 2018. The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* 16: 31–57. [CrossRef]

Liu, Jenn-Long, and Chien-Liang Chen. 2012. Application of evolutionary data mining algorithms to insurance fraud prediction. Paper presented at the 4th International Conference on Machine Learning and Computing IPCSIT, Hong Kong, China, March 10–11.

Liu, Qing, David Pitt, and Xueyuan Wu. 2014. On the prediction of claim duration for income protection insurance policyholders. *Annals of Actuarial Science* 8: 42–62. [CrossRef]

Lopez, Olivier, and Xavier Milhaud. 2021. Individual reserving and nonparametric estimation of claim amounts subject to large reporting delays. *Scandinavian Actuarial Journal* 2021: 34–53. [CrossRef]

Lou, Yin, Rich Caruana, Johannes Gehrke, and Giles Hooker. 2013. Accurate intelligible models with pairwise interactions. Paper presented at the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, August 11–14.

Lundberg, Scott M, and Su-In Lee. 2017. A unified approach to interpreting model predictions. Paper presented at the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, December 4–9.

Lundberg, Scott M, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* 2: 56–67. [CrossRef]

Ma, Yu-Luen, Xiaoyu Zhu, Xianbiao Hu, and Yi-Chang Chiu. 2018. The use of context-sensitive insurance telematics data in auto insurance rate making. *Transportation Research Part A: Policy and Practice* 113: 243–58. [CrossRef]

Mascharka, David, Philip Tran, Ryan Soklaski, and Arjun Majumdar. 2018. Transparency by design: Closing the gap between performance and interpretability in visual reasoning. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, June 18–23.

Matloob, Irum, Shoab Ahmed Khan, and Habib Ur Rahman. 2020. Sequence Mining and Prediction-Based Healthcare Fraud Detection Methodology. *IEEE Access* 8: 143256–73. [CrossRef]

Mayer, Roger C., James H. Davis, and F. David Schoorman. 1995. An integrative model of organizational trust. *Academy of Management Review* 20: 709–34. [CrossRef]

Maynard, Trevor, Luca Baldassarre, Yves-Alexandre de Montjoye, Liz McFall, and María Óskarsdóttir. 2022. AI: Coming of age? *Annals of Actuarial Science* 16: 1–5. [CrossRef]

McFall, Liz, Gert Meyers, and Ine Van Hoyweghen. 2020. The personalisation of insurance: Data, behaviour and innovation. *Big Data & Society* 7: 2053951720973707.

McKnight, D. Harrison, Vivek Choudhury, and Charles Kacmar. 2002. Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research* 13: 334–59. [CrossRef]

Mehdiyev, Nijat, Constantin Houy, Oliver Gutermuth, Lea Mayer, and Peter Fettke. 2021. *Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes–On the Potential of XAI in Tax Audit Processes*. Cham: Springer.

Meyerson, Debra, Karl E. Weick, and Roderick M. Kramer. 1996. Swift trust and temporary groups. *Trust in Organizations: Frontiers of Theory and Research* 166: 195.

Mizgier, Kamil J., Otto Kocsis, and Stephan M. Wagner. 2018. Zurich Insurance uses data analytics to leverage the BI insurance proposition. *Interfaces* 48: 94–107. [CrossRef]

Mohamadloo, Azam, Ali Ramezankhani, Saeed Zarein-Dolab, Jamshid Salamzadeh, and Fatemeh Mohamadloo. 2017. A systematic review of main factors leading to irrational prescription of medicine. *Iranian Journal of Psychiatry and Behavioral Sciences* 11: e10242. [CrossRef]

Molnar, Christoph. 2019. *Interpretable Machine Learning*. Morrisville: Lulu Press.

Moradi, Milad, and Matthias Samwald. 2021. Post-hoc explanation of black-box classifiers using confident itemsets. *Expert Systems with Applications* 165: 113941. [CrossRef]

Morik, Katharina, Christian Hüppej, and Klaus Unterstein. 2002. End-user access to multiple sources-Incorporating knowledge discovery into knowledge management. Paper presented at the International Conference on Practical Aspects of Knowledge Management, Vienna, Austria, December 2–3.

Motoda, Hiroshi, and Huan Liu. 2002. Feature selection, extraction and construction. *Communication of IICM (Institute of Information and Computing Machinery, Taiwan)* 5: 2.

Mueller, Shane T., Robert R. Hoffman, William Clancey, Abigail Emrey, and Gary Klein. 2019. Explanation in Human-AI Systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for Explainable AI. *arXiv* arXiv:1902.01876.

Mullins, Martin, Christopher P. Holland, and Martin Cunneen. 2021. Creating ethics guidelines for artificial intelligence and big data analytics customers: The case of the consumer European insurance market. *Patterns* 2: 100362. [CrossRef]

NallamReddy, Sundari, Samarandra Behera, Sanjeev Karadagi, and A Desik. 2014. Application of multiple random centroid (MRC) based k-means clustering algorithm in insurance—A review article. *Operations Research and Applications: An International Journal* 1: 15–21.

Naylor, Michael. 2017. *Insurance Transformed: Technological Disruption*. Berlin: Springer.

Neumann, Łukasz, Robert M. Nowak, Rafał Okuniewski, and Paweł Wawrzyński. 2019. Machine Learning-Based Predictions of Customers' Decisions in Car Insurance. *Applied Artificial Intelligence* 33: 817–28. [CrossRef]

Ngai, Eric W. T., Yong Hu, Yiu Hing Wong, Yijun Chen, and Xin Sun. 2011. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems* 50: 559–69. [CrossRef]

Ntoutsi, Eirini, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, and Emmanouil Krasanakis. 2020. Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10: e1356. [CrossRef]

OECD, Organisation for Economic Co-operation and Development. 2020. *The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector*. Paris: OECD. Available online: https://www.oecd.org/finance/Impact-Big-Data-AI-in-the-Insurance-Sector.pdf (accessed on 1 September 2021).

Page, Matthew J., and David Moher. 2017. Evaluations of the uptake and impact of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement and extensions: A scoping review. *Systematic Reviews* 6: 1–14. [CrossRef] [PubMed]

Palacio, Sebastian, Adriano Lucieri, Mohsin Munir, Jörn Hees, Sheraz Ahmed, and Andreas Dengel. 2021. XAI Handbook: Towards a Unified Framework for Explainable AI. *arXiv* arXiv:2105.06677.

Paruchuri, Harish. 2020. The Impact of Machine Learning on the Future of Insurance Industry. *American Journal of Trade and Policy* 7: 85–90. [CrossRef]

Pathak, Jagdish, Navneet Vidyarthi, and Scott L. Summers. 2005. A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims. *Managerial Auditing Journal* 20: 632–44. [CrossRef]

Payrovnaziri, Seyedeh Neelufar, Zhaoyi Chen, Pablo Rengifo-Moreno, Tim Miller, Jiang Bian, Jonathan H. Chen, Xiuwen Liu, and Zhe He. 2020. Explainable artificial intelligence models using real-world electronic health record data: A systematic scoping review. *Journal of the American Medical Informatics Association* 27: 1173–85. [CrossRef]

Pieters, Wolter. 2011. Explanation and trust: What to tell the user in security and AI? *Ethics and Information Technology* 13: 53–64. [CrossRef]

Putnam, Vanessa, and Cristina Conati. 2019. Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS). Paper presented at the IUI Workshops, Los Angeles, CA, USA, March 16–20.

Quan, Zhiyu, and Emiliano A. Valdez. 2018. Predictive analytics of insurance claims using multivariate decision trees. *Dependence Modeling* 6: 377–407. [CrossRef]

Ravi, Kumar, Vadlamani Ravi, and P. Sree Rama Krishna Prasad. 2017. Fuzzy formal concept analysis based opinion mining for CRM in financial services. *Applied Soft Computing* 60: 786–807. [CrossRef]

Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. 2016. "Why should i trust you?" Explaining the predictions of any classifier. Paper presented at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13–17.

Rieder, Gernot, and Judith Simon. 2017. Big data: A new empiricism and its epistemic and socio-political consequences. In *Berechenbarkeit der Welt?* Berlin: Springer, pp. 85–105.

Riikkinen, Mikko, Hannu Saarijärvi, Peter Sarlin, and Ilkka Lähteenmäki. 2018. Using artificial intelligence to create value in insurance. *International Journal of Bank Marketing* 36: 1145–68. [CrossRef]

Robinson, Stephen Cory. 2020. Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society* 63: 101421. [CrossRef]

Rosenfeld, Avi. 2021. Better metrics for evaluating explainable artificial intelligence. Paper presented at the 20th International Conference on Autonomous Agents and Multiagent Systems, London, UK, May 3–7.

Rudin, Cynthia. 2018. Please stop explaining black box models for high stakes decisions. *Stat* 1050: 26.

Sadreddini, Zhaleh, Ilknur Donmez, and Halim Yanikomeroglu. 2021. Cancel-for-Any-Reason Insurance Recommendation Using Customer Transaction-Based Clustering. *IEEE Access* 9: 39363–74. [CrossRef]

Sakthivel, K. M., and C. S. Rajitha. 2017. Artificial intelligence for estimation of future claim frequency in non-life insurance. *Global Journal of Pure and Applied Mathematics* 13: 1701–10.

Samonte, Mary Jane C., Bobby D. Gerardo, Arnel C. Fajardo, and Ruji P. Medina. 2018. ICD-9 tagging of clinical notes using topical word embedding. Paper presented at the 2018 International Conference on Internet and e-Business, Singapore, April 25–27.

Sarkar, Abhineet. 2020. Disrupting the Insurance Value Chain. In *The AI Book: The Artificial Intelligence Handbook for Investors, Entrepreneurs and FinTech Visionaries*. New York: Wiley, pp. 89–91.

Sevim, Şerafettin, Birol Yildiz, and Nilüfer Dalkiliç. 2016. Risk Assessment for Accounting Professional Liability Insurance. *Sosyoekonomi* 24: 93–112. [CrossRef]

Shah, Paras, and Allon Guez. 2009. Mortality forecasting using neural networks and an application to cause-specific data for insurance purposes. *Journal of Forecasting* 28: 535–48. [CrossRef]

Shapiro, Arnold F. 2007. An overview of insurance uses of fuzzy logic. In *Computational Intelligence in Economics and Finance*. Berlin: Springer, pp. 25–61.

Sheehan, Barry, Finbarr Murphy, Cian Ryan, Martin Mullins, and Hai Yue Liu. 2017. Semi-autonomous vehicle motor insurance: A Bayesian Network risk transfer approach. *Transportation Research Part C: Emerging Technologies* 82: 124–37. [CrossRef]

Siami, Mohammad, Mohsen Naderpour, and Jie Lu. 2020. A mobile telematics pattern recognition framework for driving behavior extraction. *IEEE Transactions on Intelligent Transportation Systems* 22: 1459–72. [CrossRef]

Siau, Keng, and Weiyu Wang. 2018. Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal* 31: 47–53.

Siegel, Magdalena, Constanze Assenmacher, Nathalie Meuwly, and Martina Zemp. 2021. The legal vulnerability model for same-sex parent families: A mixed methods systematic review and theoretical integration. *Frontiers in Psychology* 12: 683. [CrossRef] [PubMed]

Sithic, H. Lookman, and T. Balasubramanian. 2013. Survey of insurance fraud detection using data mining techniques. *arXiv* arXiv:1309.0806.

Smith, Kate A., Robert J. Willis, and Malcolm Brooks. 2000. An analysis of customer retention and insurance claim patterns using data mining: A case study. *Journal of the Operational Research Society* 51: 532–41. [CrossRef]

Smyth, Gordon K., and Bent Jørgensen. 2002. Fitting Tweedie's compound Poisson model to insurance claims data: Dispersion modelling. *ASTIN Bulletin: The Journal of the IAA* 32: 143–57. [CrossRef]

Sohail, Misbah, Pedro Peres, and Yuhua Li. 2021. Feature importance analysis for customer management of insurance products. Paper presented at the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, July 18–22.

Srihari, Sargur. 2020. Explainable Artificial Intelligence: An Overview. *Journal of the Washington Academy of Sciences*.

Stovold, Elizabeth, Deirdre Beecher, Ruth Foxlee, and Anna Noel-Storr. 2014. Study flow diagrams in Cochrane systematic review updates: An adapted PRISMA flow diagram. *Systematic Reviews* 3: 1–5. [CrossRef]

Štrumbelj, Erik, and Igor Kononenko. 2014. Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems* 41: 647–65. [CrossRef]

Sun, Chenfei, Zhongmin Yan, Qingzhong Li, Yongqing Zheng, Xudong Lu, and Lizhen Cui. 2018. Abnormal group-based joint medical fraud detection. *IEEE Access* 7: 13589–96. [CrossRef]

Supraja, K., and S. Jessica Saritha. 2017. Robust fuzzy rule based technique to detect frauds in vehicle insurance. Paper presented at the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, August 1–2.

Tallant, Jonathan. 2017. Commitment in cases of trust and distrust. *Thought* 6: 261–267. [CrossRef]

Tanninen, Maiju. 2020. Contested technology: Social scientific perspectives of behaviour-based insurance. *Big Data & Society* 7: 2053951720942536.

Tao, Han, Liu Zhixin, and Song Xiaodong. 2012. Insurance fraud identification research based on fuzzy support vector machine with dual membership. Paper presented at the 2012 International Conference on Information Management, Innovation Management and Industrial Engineering, Sanya, China, October 20–21.

Taylor, Linnet. 2017. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4: 2053951717736335.

Tekaya, Balkiss, Sirine El Feki, Tasnim Tekaya, and Hela Masri. 2020. Recent applications of big data in finance. Paper presented at the 2nd International Conference on Digital Tools & Uses Congress, Virtual Event, October 15–17.

Tillmanns, Sebastian, Frenkel Ter Hofstede, Manfred Krafft, and Oliver Goetz. 2017. How to separate the wheat from the chaff: Improved variable selection for new customer acquisition. *Journal of Marketing* 81: 99–113. [CrossRef]

Tonekaboni, Sana, Shalmali Joshi, Melissa D McCradden, and Anna Goldenberg. 2019. What clinicians want: Contextualizing explainable machine learning for clinical end use. Paper presented at the Machine Learning for Healthcare Conference, Ann Arbor, MI, USA, August 9–10.

Toreini, Ehsan, Mhairi Aitken, Kovila Coopamootoo, Karen Elliott, Carlos Gonzalez Zelaya, and Aad Van Moorsel. 2020. The relationship between trust in AI and trustworthy machine learning technologies. Paper presented at the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, January 27–30.

Umamaheswari, K., and S. Janakiraman. 2014. Role of data mining in insurance industry. *Int J Adv Comput Technol* 3: 961–66.

Ungur, Cristina. 2017. Socio-economic valences of insurance. *Revista Economia Contemporană* 2: 112–18.

van den Boom, Freyja. 2021. Regulating Telematics Insurance. In *Insurance Distribution Directive*. Berlin: Springer, pp. 293–325.

Vassiljeva, Kristina, Aleksei Tepljakov, Eduard Petlenkov, and Eduard Netšajev. 2017. Computational intelligence approach for estimation of vehicle insurance risk level. Paper presented at the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, May 14–19.

Vaziri, Jalil, and Mohammad Ali Beheshtinia. 2016. A holistic fuzzy approach to create competitive advantage via quality management in services industry (case study: Life-insurance services). *Management Decision* 54: 2035–62. [CrossRef]

Verma, Aayushi, Anu Taneja, and Anuja Arora. 2017. Fraud detection and frequent pattern matching in insurance claims using data mining techniques. Paper presented at the 2017 Tenth International Conference on Contemporary Computing (IC3), Noida, India, August 10–12.

Viaene, Stijn, Guido Dedene, and Richard A. Derrig. 2005. Auto claim fraud detection using Bayesian learning neural networks. *Expert Systems with Applications* 29: 653–66. [CrossRef]

Viaene, Stijn, Richard A. Derrig, and Guido Dedene. 2004. A case study of applying boosting Naive Bayes to claim fraud diagnosis. *IEEE Transactions on Knowledge and Data Engineering* 16: 612–20. [CrossRef]

Viaene, Stijn, Richard A. Derrig, Bart Baesens, and Guido Dedene. 2002. A comparison of state-of-the-art classification techniques for expert automobile insurance claim fraud detection. *Journal of Risk and Insurance* 69: 373–421. [CrossRef]

Vilone, Giulia, and Luca Longo. 2020. Explainable artificial intelligence: A systematic review. *arXiv* arXiv:2006.00093.

von Eschenbach, Warren J. 2021. Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology* 34: 1607–22.

Walsh, Nigel, and Mike Taylor. 2020. Cutting to the Chase: Mapping AI to the Real-World Insurance Value Chain. In *The AI Book: The Artificial Intelligence Handbook for Investors, Entrepreneurs and FinTech Visionaries*. New York: Wiley, pp. 92–97.

Wang, Hui Dong. 2020. Research on the Features of Car Insurance Data Based on Machine Learning. *Procedia Computer Science* 166: 582–87. [CrossRef]

Wang, Yibo, and Wei Xu. 2018. Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems* 105: 87–95. [CrossRef]

Wei, Cheng, and Li Dan. 2019. Market fluctuation and agricultural insurance forecasting model based on machine learning algorithm of parameter optimization. *Journal of Intelligent & Fuzzy Systems* 37: 6217–28.

Wulf, Alexander J., and Ognyan Seizov. 2022. "Please understand we cannot provide further information": Evaluating content and transparency of GDPR-mandated AI disclosures. *AI & Society*, 1–22. [CrossRef]

Wüthrich, Mario V. 2018. Machine learning in individual claims reserving. *Scandinavian Actuarial Journal* 2018: 465–80. [CrossRef]

Wüthrich, Mario V. 2020. Bias regularization in neural network models for general insurance pricing. *European Actuarial Journal* 10: 179–202. [CrossRef]

Xiao, Bo, and Izak Benbasat. 2007. E-commerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly* 31: 137–209. [CrossRef]

Xie, Ning, Gabrielle Ras, Marcel van Gerven, and Derek Doran. 2020. Explainable deep learning: A field guide for the uninitiated. *arXiv* arXiv:2004.14545.

Xu, Wei, Shengnan Wang, Dailing Zhang, and Bo Yang. 2011. Random rough subspace based neural network ensemble for insurance fraud detection. Paper presented at the 2011 Fourth International Joint Conference on Computational Sciences and Optimization, Kunming, China, April 15–19.

Yan, Chun, Meixuan Li, Wei Liu, and Man Qi. 2020a. Improved adaptive genetic algorithm for the vehicle Insurance Fraud Identification Model based on a BP Neural Network. *Theoretical Computer Science* 817: 12–23. [CrossRef]

Yan, Chun, Xindong Wang, Xinhong Liu, Wei Liu, and Jiahui Liu. 2020b. Research on the UBI Car Insurance Rate Determination Model Based on the CNN-HVSVM Algorithm. *IEEE Access* 8: 160762–73. [CrossRef]

Yan, Weizhong, and Piero P. Bonissone. 2006. Designing a Neural Network Decision System for Automated Insurance Underwriting. Paper presented at the 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, July 16–21.

Yang, Qiang, Jie Yin, Charles Ling, and Rong Pan. 2006. Extracting actionable knowledge from decision trees. *IEEE Transactions on Knowledge and Data Engineering* 19: 43–56. [CrossRef]

Yang, Yi, Wei Qian, and Hui Zou. 2018. Insurance premium prediction via gradient tree-boosted Tweedie compound Poisson models. *Journal of Business & Economic Statistics* 36: 456–70.

Yeo, Ai Cheo, Kate A. Smith, Robert J. Willis, and Malcolm Brooks. 2002. A mathematical programming approach to optimise insurance premium pricing within a data mining framework. *Journal of the Operational Research Society* 53: 1197–203. [CrossRef]

Yeung, Karen, Andrew Howes, and Ganna Pogrebna. 2019. AI governance by human rights-centred design, deliberation and oversight: An end to ethics washing. In *The Oxford Handbook of AI Ethics*. Oxford: Oxford University Press.

Zahi, Sara, and Boujemâa Achchab. 2019. Clustering of the population benefiting from health insurance using K-means. Paper presented at the 4th International Conference on Smart City Applications, Casablanca, Morocco, October 2–4.

Zarifis, Alex, Christopher P. Holland, and Alistair Milne. 2019. Evaluating the impact of AI on insurance: The four emerging AI-and data-driven business models. *Emerald Open Research* 1: 15. [CrossRef]

Zhang, Bo, and Dehua Kong. 2020. Dynamic estimation model of insurance product recommendation based on Naive Bayesian model. Paper presented at the 2020 International Conference on Cyberspace Innovation of Advanced Technologies, Guangzhou, China, December 4–6.