Deep Science Publishing, 2025
https://doi.org/10.70593/978-93-7185-061-2_7

Chapter 7: Generative AI in Predictive Risk Modeling

7.1. Introduction to Predictive Risk Modeling

Predictive risk modeling can be defined as the practice of using statistics and machine learning techniques to assess and evaluate the likelihood of risk occurrence or non-occurrence in the future. Accurate prediction of risk is critical for most contemporary decision-making in both the private and public sectors.

Recent years have witnessed a boom in the development of generative artificial intelligence, which refers to a category of AI models used to generate new content akin to human creation. Full reservoir simulation is a crucial component of the oil and gas industry, yet it is beset with several challenges: expensive costs, lengthy periods, and complex operations. To meet evolving requirements, the adoption of predictive analytics in the oil and gas sector is on the rise. This approach seeks to leverage historical, current, and real-time data to generate risk estimates, striving to create a predictive risk model capable of outperforming humans in predictive efficiency. Generative AI currently ranks among the most potent weapons against fraud, indicative, for instance, of how generative pre-trained transformer technology offers advanced tools to the financial industry for analysis and crime prevention.

7.1.1. The Importance of Predictive Risk Modeling in Today's Landscape

Predictive risk modeling is a methodology that leverages data and algorithms to forecast potential risks and their probable impacts while refining decision-making processes. Its applications are diverse, spanning investment risk analysis, operational risk monitoring, human resource risk assessment, and predictive behaviors in distinct domains. Within these contexts, predictive risk models guide decisions aimed at mitigating risk or capitalizing on potential advantages that risk-taking may yield. The accuracy of risk prediction plays a pivotal role in minimizing risk-related losses generated during the occurrence of risky events. Contemporary society is increasingly reliant on advanced technologies to predict events before they happen and to prescribe optimal decisions. Predictive risk models, underpinned by artificial intelligence and deep learning, embody such technologies.

Artificial Intelligence (AI) methods are gaining greater attention in risk evaluation either as standalone predictive risk models or as tools assisting other models. They achieve a high level of comprehension by simulating complex properties inherent in nature, thought, or human intelligence [1-3]. The merits of AI-based risk prediction are rooted in algorithmic structures that emulate human neural interactions. In this context, Generative AI stands as a subclass of AI capable of generating synthetic content—ranging from text and images to videos, voice, code, and animations—that convincingly imitates human output.

7.2. Overview of Generative AI

Predictive risk modeling refers to the process of using past observations to build a model capable of predicting future outcomes. Such models are crucial in modern society due to the complexity of events and the surge in available data. Typically, conventional artificial intelligence (AI) methods have been employed for these tasks; however, the recent Artificial Intelligence (AI) Revolution, also known as the Fourth Industrial Revolution, has yielded generative AI, which is expected to improve risk prediction accuracy. Within the specific domain of predictive risk modeling, numerous generative models exist, among which Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) are currently the most popular.

A Variational Autoencoder is an autoencoder configured so that the latent variables encode a multivariate Gaussian distribution. While an autoencoder comprises an encoder that compresses the data into a lower-dimensional space and a decoder that reconstructs the original data—typically optimized using Reconstruction Loss—a VAE models these latent variables using this specific probability distribution. In contrast, Generative Adversarial Networks represent an alternative approach to generative modeling. Additionally, the Progression and Bills Erosion Models of Generative Duel Networks utilize a generative adversarial architecture.

7.2.1. Key Concepts and Techniques in Generative AI

Generative AI (GenAI), as the name suggests, is an artificially-intelligent system that produces various types of content, such as text, images, or videos in response to prompts. GenAI techniques, which are based on generative modelling, use large datasets to learn the underlying patterns of the analysed data. At its core, generative models need to be able to extract the essential features—referred to as latent variables—from the observed data and work towards reconstructing the data from those extracted features. The primary goal is to create new data points that share some similarity with the original dataset.

Two of the most commonly used generative models are Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs). A VAE is a type of encoder-decoder architecture that consists of two major parts: a recognition model (encoder), which compresses the data into the lower-dimensional latent space, and a generative model (decoder), which reconstructs the original high-dimensional input from the latent representation. On the other hand, a GAN consists of two neural networks—a generator and a discriminator—that play an adversarial game. The generator synthesises data similar to the original dataset, while the discriminator evaluates the authenticity of the samples, distinguishing between real and fake.

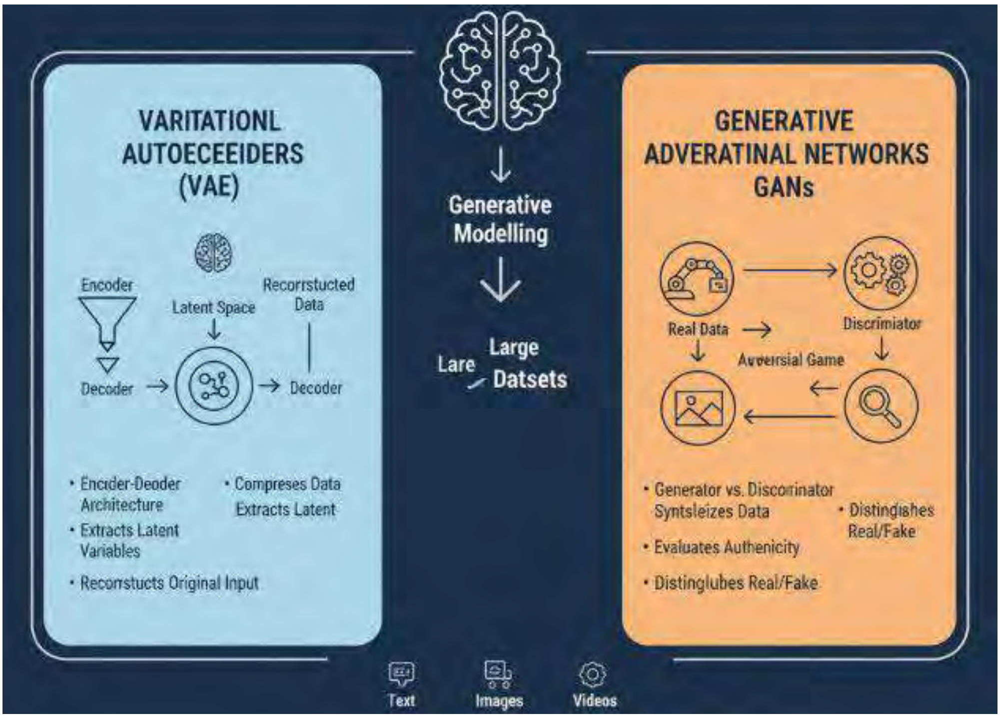

Fig 7 . 1 : Generative AI: VAEs and GANs

7.3. The Role of AI in Risk Assessment

Predictive risk modeling is an essential component of many decision processes. By analyzing past events, risk professionals are able to estimate the risk of future events. This exercise requires well-developed statistical models and plenty of historical data.

When these conditions are fulfilled, predictive models provide robust and valuable estimates about future losses, ratemaking, underwriting, and even reserving. Moreover, these models are used in marketing and customer retention strategies. The adoption of Artificial Intelligence (AI) and automation is expected to play a crucial role in everything that the future holds for predictive risk modeling.

Artificial Intelligence (AI) is extensively applied to risk assessment and classification in different domains in the present era. It can imitate human intelligence and behavior and display it while performing different tasks. AI applications include Expert Systems, Robotics, Natural Language Processing, Gesture Recognition, Face Recognition System, Vision System, Machine Learning, and more. Although AI is highly popular across organizations mainly for different kinds of risk assessment, it still has some limitations.

7.3.1. Significance of AI Innovations in Risk Evaluation

Significance of AI Innovations in Risk Evaluation Innovations in predictive algorithms and techniques have led to new methods to assess and manage risk more accurately. Today's predictive risk models help decision makers across industries predict when events may occur and and anticipate associated costs, negative impacts, or opportunities.

Predictive risk models combine business goals with probabilities in order to make predictions about the future. These models estimate: the likelihood of a negative event, such as defaulting on a loan, filing an insurance claim, or becoming seriously ill; the anticipated severity or impact of that event for the business, given the event actually occurs; and the optimum course of action.

Generative AI represents a significant step forward for predictive risk modeling. Unlike other forms of AI, Generative AI analyzes a company's raw data, extracts underlying patterns, and creates a model that can generate new data that shares similar characteristics with the raw data. As a result, the generated risk predictions tend to be more accurate.

A number of additional associated advances are also improving predictive risk modeling: the expansion of Internet of Things (IoT) and wearable devices connected through 5G or other broadband services; the large personal and business databases resulting from these connected devices; and a wealth of data analytics services provided by the cloud. Together these components offer an unprecedented opportunity to expand the range and granularity of real-time risk assessment.

7.4. Types of Predictive Models

Predictive models play an essential role in transforming data into actionable insights for risk assessment and be broadly categorized in three classes: statistical models, machine learning models, and deep learning models. Statistical models, such as linear or logistic regression, are among the oldest forecasting methods using interpretable functions of selected a priori features. Machine learning models extend statistical models by considering nonlinear transformations of the features, automatically applying transformations based on optimization, and exploring different feature interactions. Deep learning models use layered neural networks to simultaneously perform feature-engineering and prediction.

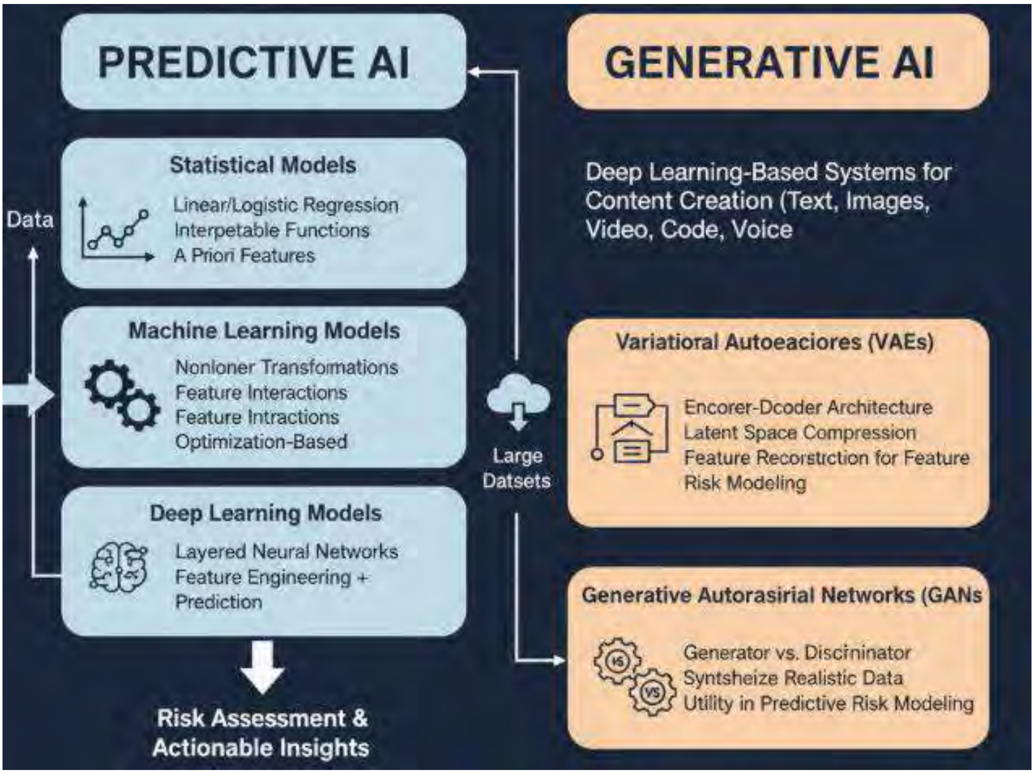

Fig 7 . 2 : Predictive vs. Generative AI

Generative AI, by contrast, is the overarching term for deep learning-based systems that are capable of generating content, including text, images, videos, code, and voice. Many generative AI models are trained using unsupervised or self-supervised learning, which does not require labeled training data, and employ large model architectures and massive volumes of training data [2,4,5]. Within deep-learning, Variational Autoencoders (VAEs) are generative architectures consisting of an encoder and a decoder neural network, which respectively compress original features into a lower-dimensional latent space and reconstruct features back into the input space. This capability renders VAEs valuable for feature extraction, dimensionality reduction, and information extraction in predictive risk modeling. Generative Adversarial Networks (GANs) are designed using two competing neural networks: the generator and discriminator. The generator creates synthetic data samples, whereas the discriminator attempts to discern real inputs from the counterfeit. The iterative training of these adversaries culminates in the generator's proficiency in producing highly realistic data samples, offering promising utility to predictive risk modeling.

7.4.1. Statistical Models

Statistical models are mathematical expressions that describe the relationship between dependent and independent variables. In the context of predictive risk modeling, logistic regression is a widely used statistical method for estimating the probability of an event occurring. Logistic regression enables nonlinear modeling by employing the logit transformation of the event probability. One of the primary objectives of risk modeling is to accurately predict a customer's payment behavior, which may be classified as good, bad, or doubtful. Logistic regression allows for calibration, determining a threshold below which a customer is considered a bad or doubtful risk.

Unlike traditional regression, both dependent and independent variables in logistic regression can be either ratio or nominal. The dependent variable represents the event being modeled, with two possible values indicating the occurrence or non-occurrence of the event. Typically, the dependent variable is binary, yet it can encompass multiple values, although the binary case is most commonly applied in risk assessment. In addition to statistical models, machine learning and deep learning techniques significantly contribute to the construction of predictive models. In recent years, generative AI models—particularly Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN)—have attracted considerable attention for their potential to enhance the predictive risk modeling process.

7.4.2. Machine Learning Models

Generative AI principally applies to prediction models, which come in three types: statistical, machine learning, and deep-learning-based. Statistical predictive models, while simpler and quicker to train, are less adept at processing complex input features or relationships that require derivation or transformation from the raw data. Machine-learning models, by contrast, offer greater flexibility; they can derive the necessary features from data thanks to their intricate architecture and training process, enabling response to more complex scenarios. Deep-learning models, including generative models such as variational autoencoders and generative adversarial networks, represent the pinnacle of this hierarchy. Despite the ability of deep neural networks to function as classifiers or regressors, the term “deep learning” typically denotes models composed of multiple hidden layers that either generate new data from original training data (as in deep generative models) or produce lower-level features useful for subsequently generating data or acting as a classifier/regressor, as illustrated in recent research.

Machine-learning predictive models offer significant versatility, though at the cost of higher training time and the risk of under- or over-fitting due to a lack of embedded domain knowledge, unlike statistical models. Nevertheless, such knowledge can be incorporated during data-preprocessing or –transformation phases. The application of machine-learning models to risk prediction is growing rapidly, spurred not only by ongoing algorithmic and architectural improvements but also by an increase in available large datasets, whether of individual entities or key events and observations across organizations. The current phase of research aims to leverage generative AI capabilities to identify which designs are best suited for particular predictive-risk scenarios. Similar efforts are evident in studies applying AI techniques to other risk types, such as those employing statistical, machine-learning, or deep-learning models for country-risk prediction, as reviewed by Wang and Xiao.

7.4.3. Deep Learning Models

Deep learning models have risen to prominence due to their ability to detect subtle patterns in large and complicated datasets. Unlike machine learning methods that require manually designed features, deep learning architectures perform implicit feature engineering by discovering multiple levels of data representations. This multi-step automatic feature generation process considerably enhances the model’s generalization capabilities—its ability to predict accurately on new, unseen instances of a classification task.

Within the deep learning framework, neural networks consisting of three or more layers are termed deep neural networks (DNNs). When data exhibits significant spatial characteristics, convolutional neural networks (CNNs) are preferable, whereas recurrent neural networks (RNNs) arewell-suited to data with temporal dependencies. Predictive risk modelling frequently involves such data types. In a related discussion, generative AI constitutes a subset of artificial intelligence-based systems that utilize generative models like variational autoencoders (VAEs) and generative adversarial networks (GANs). Generative models can create new data instances such as text, images, or sounds that share the same underlying characteristics as the original data, thus offering valuable capabilities for risk modelling applications.

7.5. Generative Models Explained

Generative AI represents a subset of Artificial Intelligence capable of generating new data, typically in text, images, or voice form, based on learned distributions. Variational Autoencoders (VAEs) serve as a specific technique designed to generate new data while simultaneously learning lower-dimensional features from the data set. VAEs compress input data into smaller encoded vectors before decompressing back to higher-dimensional spaces, thereby enabling both lower-dimensional feature extraction and robust data reconstruction [6-8].

Generative Adversarial Networks (GANs) employ an unsupervised machine learning approach capable of learning and replicating patterns in a given data set. GANs consist of two deep neural networks—the generator and the discriminator—which engage in a zero-sum game to create higher-quality models and generate new content. This adversarial process drives the system toward producing highly realistic data representations, enhancing performance in applications such as image, text, and video generation.

7.5.1. Variational Autoencoders

Predictive risk modeling refers to the quantitative techniques used to analyze future risks, with the aim of estimating their potential effects and deciding how to mitigate or manage them. Traditionally, the concept of risk analysis has been weakly grounded in theory and, consequently, has been treated in an informal and qualitative way. However, modern decision-making is marked by an increasing requirement to be both proactive and anticipatory, and consequently, risk prediction is becoming an urgent need. The capabilities of generative artificial intelligence (AI) promise to transform aspects of risk assessment, prediction, and evaluation in a variety of specialist domains. Risks emerge from uncertainty and an event that triggers that risk in a negative way. Therefore, prediction enables insight into potential risks, allowing the modeler to prepare mitigation or management arrangements. Predictive risk modeling forms the backbone of effective risk management and strategic planning, particularly within business when risk is often synonymous with uncertainty.

Predictive modeling applies data mining and probability to forecast results. Statistical procedures—such as logistic regression or K Means cluster analysis—and a carefully selected set of performance drivers enable a business to predict likely outcomes based on prior events. An example is credit scoring, used to assess the potential risk involved in lending money to a borrower. Through predictive modeling, an institution can estimate the likelihood of a customer failing to repay a credit card or loan balance before the next billing cycle, and adjust credit limits accordingly. Generative AI derives models that can synthesize new data instances that are similar to those in the training data. Variational autoencoders (VAEs) and generative adversarial networks (GANs)—two widely adopted methodologies in finance—represent the state of the art.

Fig 7 . 3 : Predictive Risk Modeling and Generative AI

7.5.2. Generative Adversarial Networks

Generative Adversarial Networks (GANs) are a deep learning method that produces synthetic data with similar characteristics to the training data. GANs consist of two components: a generator, which produces synthetic data, and a discriminator, which evaluates the lifelikeness of samples. These components operate competitively, with the generator aiming to create samples indistinguishable from real data and the discriminator striving to accurately differentiate between authentic and synthetic inputs. During training, the generator learns the data correlation and distribution of the training dataset, while the discriminator continuously evaluates its own performance and the generator's progress. This adversarial competition results in the generator eventually producing synthetic data indistinguishable from the real data. Capable of learning complex data distributions, GANs are ideal for predictive risk modeling as they can better forecast probable scenarios underlying the risk while providing insights into its severity.

Risk prediction can be challenging due to the lack of available data under specific hypothetical scenarios or because risk prediction involves estimating the likelihood of an abnormal event rather than modeling the normal trend or state of a phenomenon. GANs address these challenges by supplementing historical data with synthetic data tailored to different types of risk. For example, in weather forecasting, GANs can generate complementary weather prediction data. In financial applications, such as credit risk assessment, GANs generate different types of synthetic default forecasts. In image recognition, GANs create diverse synthetic images under specific conditions, aiding risk rating models of stormwater sewer systems. Currently, relatively little research has transferred GAN-generated data to supporting predictive models. Nonetheless, the adversarial architecture of GANs makes them highly efficient in generating data that can significantly enhance predictive risk assessment.

7.6. Future Trends in Predictive Risk Modeling

The evolution of predictive risk modeling is propelled by the relentless advancement of computing, analytics, and big data technologies, reshaping the predictive analytics environment. Demonstrated through practical, real-world examples, AI is rapidly moving from a mere trend to a critical component in operational business management—one that enables tasks and functions previously deemed impossible. While current AI capabilities are already impressive, the generative branch of AI—notably techniques such as Variational Autoencoders and Generative Adversarial Networks—suggests transformative applications for the near future.

Business intelligence experts concur that the integration of Internet of Things (IoT) devices with predictive analytics can revolutionize risk evaluation [7-9]. In the near future, IoT-enabled predictive analytics models will offer an end-to-end risk management solution by providing organizations with an ongoing assessment of risk as it truly stands at almost any given moment.

7.6.1. Integration with IoT

Predictive risk modeling plays a pivotal role, employing statistical methods and machine learning techniques to analyze algorithms and predict potential risks. This integration finds resonance in current applications that leverage the Internet of Things (IoT), where sensors and devices collect real-time data on elements like water level, moisture, and rainfall to feed into risk assessment models.

Generative AI techniques offer remarkable capabilities for model construction and operation, facilitating the development of models that capture feature representations and ensure robust data reconstruction. These models are foundational for advancements in IoT and real-time risk assessment, equipping risk processes with the analytical prowess of generative AI. The widespread application of such intelligent models is expected to enhance the precision of risk prediction technologies.

7.6.2. Real-Time Risk Assessment

Real-time risk assessment is an extremely critical sector where the capability of risk models to instantly assess when the dependent variable unravels into adversity determines the loss probability. Numerous sectors, including health care, insurance, and finance, are actively engaged in this domain, particularly following the COVID-19 pandemic generated by the SARS-CoV-2 virus. The demand for instantaneous risk evaluation has escalated to curtail losses caused by severe corporate adversities. Notably, in the health-care sector, the ability to detect viral variants promptly could result in saving hundreds of thousands of lives. With the advancement of Artificial Intelligence, particularly Generative AI, predictive risk modeling can significantly contribute to this subsector. Advanced generative models, such as Variational Autoencoders and Generative Adversarial Networks, exhibit a phenotype that facilitates feature mapping from latent space to the original space and vice versa, allowing generative modeling in both directions.

7.7. Conclusion

Predictive risk modeling involves the use of data-driven analytics to predict future risk with a certain degree of reliability. These models represent the relationship between a set of explanatory variables or features and the measure of risk, and are designed to reflect the social and economic impact of a risk event. Predictive usefulness depends on the accuracy of predicted risk values relative to actual outcomes. Today, accurate prediction of risk has become a crucial component in risk-sensitive decision-making; people need to know the probability of an attack, a pandemic, a health problem or even unemployment so that they are able to adopt a well prepared, safe and confident lifestyle.

Without these probabilities, they have to face uncertainties. AI plays a significant role in today’s risk evaluation process. Based on data, AI provides EU-wide assessments of the risk carried by an individual or a group of individuals21. Predictive risk models may appear at different scales and provide estimates of risk for individuals, companies, or even for a group of continents. Predictive risk modeling also plays a role with respect to managerial decision-making in the corporate and business sectors. When deciding on specific risk treatment options, it is important for managers to be able to anticipate the impact of their decisions on aspects such as profitability and business sustainability, especially when making decisions involving operational or capital expenditure. Predictive risk models may be designed with these specific business considerations in mind and provide forecasts of a company’s financial instability, profitability or future business performance. When the predictions involve financial stability there is considerable overlap with credit scoring. When the predictions relate to other aspects of business performance there is some overlap with forecasting.

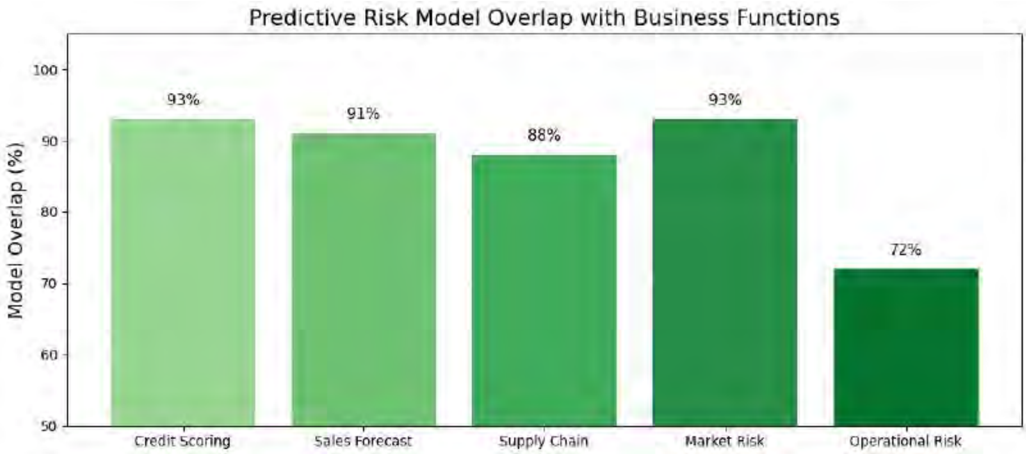

Fig 7 . 4 : Predictive Risk Model Overlap with Business Functions

7.7.1. Summary and Future Directions of Predictive Risk Modeling

Predictive risk modeling applies statistics to forecast risks in diverse scenarios. Its use is increasingly essential as more decision-making areas experience greater complexity and volatility. The United States Federal Emergency Management Agency's (FEMA) recent budget request specifically highlights a need for enhanced predictive risk modeling in natural disaster areas in response to climate change. Techniques such as IoT integration and alert systems for natural disasters aim—primarily—to anticipate risks, allowing more time to respond accordingly.

Generative models represent a subset of unsupervised machine learning algorithms designed to identify a schema underlying data in order to create new examples that could have originated from the original dataset. Variational Autoencoders (VAEs) reconstruct new data using a set of latent variables, while Generative Adversarial Networks (GANs) use two neural networks contending in a zero-sum game framework where the generator creates new samples, and the discriminator evaluates whether the samples are real (coming from the actual data) or fake (generated).

References

[1] Singh KD. (2023). Security and Privacy in Fog/Cloud-based IoT Systems for AI and Robotics. EAI Endorsed Transactions on AI and Robotics, 2(1).

[2] Munnangi, A. S. M., Reddy Koppolu, H. K., Nayeem, S. M., Polineni, I., & Reddy Munnangi, S. (2025). Experimental Measurements, Molecular Dynamics Simulations, and Machine Learning Predictions for the γ-Butyrolactone–N, N-Dimethylacetamide Binary System. ChemistrySelect, 10(28), e02391.

[3] AI-Based Financial Advisory Systems: Revolutionizing Personalized Investment Strategies. (2021). International Journal of Engineering and Computer Science, 10(12).

[4] Fang Z, Li Q. (2025). A Survey on Privacy and Security Issues in IoT-based Environments: Technologies, Protection Measures and Future Directions. Computers & Security, 148.

[5] Sneha Singireddy. (2024). The Integration of AI and Machine Learning in Transforming Underwriting and Risk Assessment Across Personal and Commercial Insurance Lines . Journal of Computational Analysis and Applications (JoCAAA), 33(08), 3966–3991. Retrieved from

[6] Alwarafy A, Al-Thelaya KA, Abdallah M, Schneider J, Hamdi M. (2020). A Survey on Security and Privacy Issues in Edge Computing-Assisted Internet of Things. arXiv preprint.

[7] Researcher. (2023). Decision Support Systems for Government Auditing: The Role of AI in Ensuring Transparency and Compliance. Zenodo. https://doi.org/10.5281/ZENODO.15489803

[8] Radanliev P, De Roure D, Maple C, Nurse JRC, Nicolescu R, Ani U. (2024). AI Security and Cyber Risk in IoT Systems. Frontiers in Big Data, 7.

[9] Koppolu, H. K. R., & Sheelam, G. K. (2024). Machine Learning-Driven Optimization in 6G Telecommunications: The Role of Intelligent Wireless and Semiconductor Innovation. Global Research Development (GRD) ISSN: 2455-5703, 9(12).