aajed.com ISSN 3067-4190

AMERICAN ADVANCED JOURNAL FOR EMERGING DISCIPLINARIES (AAJED)

OPEN ACCESS. PEER-REVIEWED. GLOBALLY FOCUSED.

Applying Deep Learning to Mobile Home and Flood Insurance Risk Evaluation

Sneha Singireddy,
Software Development Engineer is Testing

Abstract

Flooding from storm surges and precipitation, high winds from storms and tornadoes, and wildfires are now responsible for about 90% of all catastrophe-related insurance losses. The increasing intensity and frequency of these perils means more and more billing, claims, and losses are incurred by property insurance because the buildings are not built to withstand natural disasters. The increasing number of wildfires in the USA and other parts of the world, with their long tail of severity, makes wildfire risk evaluation and property holders aware of this growing peril even more critical. The day of judgment is coming when insurance companies will need to prove that areas they have written hurricanes and water damage insurance are not a loss risk afterward.

During intense flooding events, inland and coastal flooding along with wind damage create billions in losses with the vast majority of claims coming from flood insurance. Residential property underwriting has to change moving forward. Businesses need to recognize the true cost of a policy before flood damage occurs. Insurance companies must take action for underwriting properties and businesses that are prone to higher frequency and severity of residential and business interruption losses due to flooding, high winds from hurricanes or tornadoes, and wildfires. In addition, the insurance market must adapt to the next stage of climate risk management tied to the financial ramifications of the growing frequency and severity of the perils across the globe. This is important since property insurance is the primary means of transferring the risk of a peril.

Keywords: Climate Risk Management, Property Insurance Losses, Flood Insurance Challenges, Wildfire Risk Evaluation, Storm Surge Damage, Hurricane Wind Damage, Tornado Catastrophe Impact, Residential Property Underwriting, Business Interruption Insurance, Long-Tail Severity Events, Natural Disaster Risk Mitigation, Insurance Market Adaptation, Financial Impact Of Climate Change, Peril Frequency And Severity, Flood Damage Assessment, Wildfire Insurance Claims, Catastrophe-Related Loss Trends, Coastal And Inland Flooding Risks, Climate Resilience In Underwriting, Global Perils Insurance Challenges.

1. Introduction

In the United States, mobile homes are a significant part of the housing inventory, representing approximately 7.4% of the total residential units. However, mobile homes represent a disproportionate share of total estimated losses from severe natural hazard events. Over the last 30 years, an upward trend in the number of mobile homes in high-risk flood zones has also been observed.

Flood insurance is often underwritten by corporations through the National Flood Insurance Program sponsored by the Federal Emergency Management Agency. Premium revenues, collected in stable periods, are expected to cover the total expected losses and associated expenses from necessary loss payments during disaster years. However, the program has accumulated substantial debt over its entire existence, making an alternative solution necessary. Achieving a more accurate estimation of actual policyholder risk, which would enable better premium setting, is essentially needed. In this context, one such strategy is improving risk assessment methods in collaboration with private insurance companies.

Natural catastrophe insurance loss estimation is conventionally carried out by catastrophe risk models that simulate the impact of seismic, windstorm, flood, and other hazards. These models need to estimate assessment issues at diverse levels of granularity, from a single building to large geographic regions. To date, statistical modeling of physical and economic aspects of the complete catastrophe scenario is used, where input data is obtained from various sources following a hierarchical structure. Input issues also include the adoption of geodetic and remote sensing techniques, spatial-temporal modeling, and location risk evaluation. The existing research utilizes gradient-boost decision trees, random forests, and extreme gradient-boosting techniques to address the issues and challenges of catastrophe risk modeling tasks.

1.1. Overview of the Study and Objectives

As natural disasters become more common and the likelihood of catastrophic loss increases, larger losses are expected in the insurance portfolios that cover these incidents. Predicting new losses is therefore extremely important for insurance companies to avoid insolvencies and to help them grow correctly. We propose the use of rooted deep learning algorithms, benefiting from multinomial regressions, to obtain a risk evaluation system capable of predicting losses for companies that cover events such as floods and mobile home fires. The results obtained are promising since, by utilizing risk evaluation models created through neural and deep learning, we can perform superior predictions when compared to the traditional methods that insurance companies employ nowadays. In this sense, we hope that rooted deep learning models can become an excellent tool to create new options for risk evaluation in the insurance world.

In recent years, neural networks (and machine learning models in general) have been endowed with great capabilities. Deep learning algorithms, which can also be seen as a particular case of generalized hierarchical models, allow companies to extract hidden information from their data. When these models are used properly, they can transform data entries to reveal deep and hidden relationships without incursions of gross specifications and without the informative loss chosen by the user when estimating model parameters. Nowadays, deep learning algorithms have exploded in many areas, including health, finance, vision systems, and natural language processing. However, in many areas of social science, they continue to wait until their potential is properly explored and revealed. This paper aims to penetrate a portion of the insurance area connected to deep learning. In particular, we will use two digital neural network models to obtain estimates for modeling insurance claim losses in the area of mobile home and flood insurance.

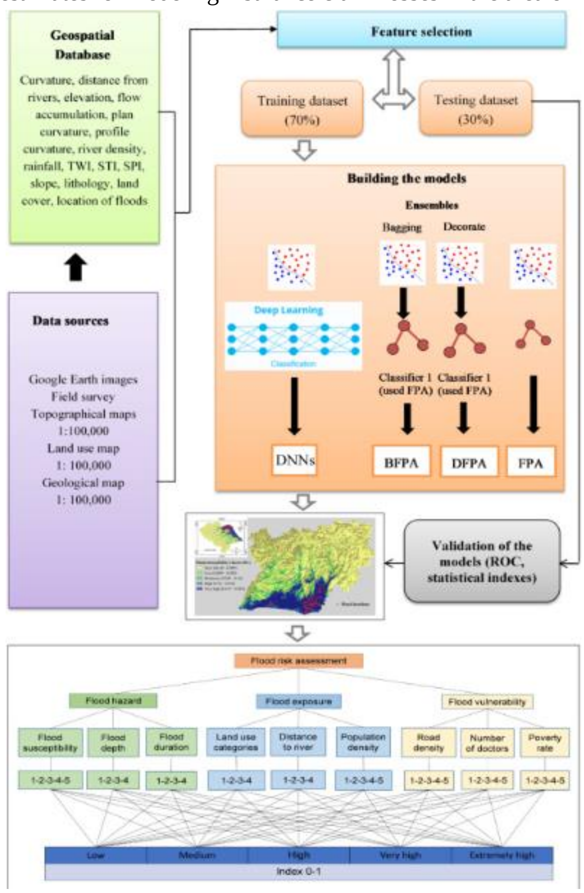

Fig 1 : Flood risk assessment using deep learning integrated

2. Background

1. Historical Context and Significance

Within the United States, neither private nor governmental insurance offers comprehensive coverage to mobile/manufactured homeowners. Different private insurers offer financial protection from claims made as a result of specific perils adjacent to mobile homes, but they typically only provide coverage for the home itself or the personal property within it and do not bundle other liability protection policies. This has left many mobile home residents in vulnerable situations, given the well-documented disparity in wealth between owners of mobile homes located in owner-occupied parks and owners of traditional homes. Furthermore, while mobile homes constitute less than 8% of all homes in the United States, they account for nearly a third of fire deaths every year, representing 15 times the risk of traditional homes. In many other regards, mobile homes present safety concerns, especially regarding tornadoes, hurricanes, and floods. Given the above statistics, understanding the insurance risk associated with mobile homes should be a priority. While studies have acknowledged mobile homes to be one of the housing types most at risk from flooding, issues arise from there not being a federal requirement for flood insurance coverage by all lenders for mobile home loans resulting in coverage gaps for unprivileged residents who own mobile homes, especially in states that have a long history of mobile home placements adjacent to rivers. Those living in mobile home communities are for the most part low to moderate-income households and more than 80% of mobile home residents in Mississippi are African American, suggesting that racial equity in insurance coverage is an issue.

Equation 1 : Risk Prediction Output (RPO):

$$R_s = f_{\mathrm{DL}}(X)$$

R_s = Predicted insurance risk score

f_{DL} = Deep learning model (e.g., CNN, RNN)

X = Input features (e.g., location, structure type, flood zone data)

2.1. Historical Context and Significance

Applying machine learning methods to insurance risk evaluation is a niche but growing area of research and practice, with the most common techniques involving the use of models such as Bayesian additive regression trees, boosted regression trees, and elastic net solutions to generalized linear models. Insurance risk assessment remains one of the pillars of actuarial science, responding to a set of research questions that derive fundamentally from the selected distributional choices on loss behavior, across the data source, product coverage, and risk factor selection. In the United States, the home insurance and flood insurance industries face unique challenges regarding loss evaluation from historical data, especially due to limited data, data Heaping, data censoring, and potentially biased estimation following years of litigation and reinsurance issues. Importantly, rigorous and objective model testing validation would eventually be needed before adoption for practice, as is typical in actuarial work.

This paper shares new findings based on the application of state-of-the-art deep learning techniques to insurance modeling and forecasting, including risk premium curve creation, pricing and market segmentation, loss severity, and claim frequency, and new company establishment in the mobile home and flood insurance markets in Florida. The previously discussed motivation guiding this work stems from both the deep learning methods and the Florida insurance market itself. In terms of the specific data context, the state of Florida has been the highest volume state for home insurance claim filings and subsequent payouts in the last decade, with loss estimates due partially to infrastructure degradation from population dynamics, climate change, and the catastrophe cyclical nature of extreme weather events.

3. Deep Learning Overview

Deep learning is a subfield of machine learning that applies models built on neural networks to achieve its results. Long believed to be too large and complex to be useful, the deep learning models have recently witnessed a renaissance thanks to hardware advances, notably the advent of graphical processing unit computing, algorithmic advances like the introduction of dropout, and access to much larger training datasets. Such models have achieved superior performance on problems in computer vision, natural language processing, sound processing, and game playing, and show great promise in addressing a wider array of supervised, unsupervised, and reinforcement learning problems than previously addressed.

The neural network has long been viewed as a type of function approximator technically tuned to minimize the error of applied functions. Traditional methods of constructing neural network architectures involved expert ontology-based encapsulation of source problems to be solved, notably the input training vector, the output training vector, and rules for encoding the necessary features of these vectors. Deep learning is distinguished from other types of neural networks in that its architecture is predefined and more complex than in traditional networks, involving many more hidden units and layers, and it is conceptually easier to train. In the deep learning context, the feature extractor is a component of the neural network itself and is trained along with the other components of the network. Many of the technical challenges associated with training deep networks have been addressed, particularly around the initialization of the latent vectors; better training and inference algorithms; and better initialization, regularization, and adaptation phase techniques.

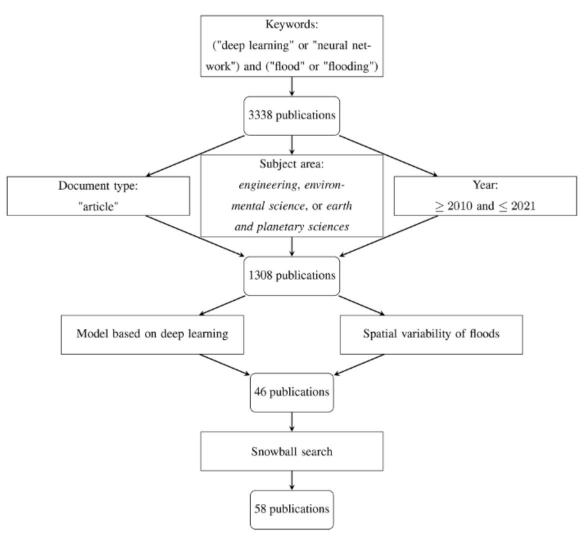

Fig 2 : Deep learning methods

3.1. Definition and Key Concepts

Several definitions, variations, and nuances of deep learning are found in the literature. In general, deep learning is a specific class of a broader machine learning method designed to optimize the performance of neural networks with many hidden layers using backpropagation and stochastic optimization methodologies. Deep learning can also refer to other algorithms designed to perform unsupervised or semi-supervised representation learning based on learned hierarchies that transform data using basic components, so-called neurons, in a similar way to neural networks.

It is critical to clarify the difference between deep learning and neural networks. Deep learning is often presented synonymously with deep neural networks; however, it is important to emphasize that deep neural networks represent one of the possible configurations of a deep learning method. The history of general neural networks starts with the initial contributions of Rosenblatt, who invented the Perceptron algorithm based on the biological model of the neuron, and later by others who created the multi-layer perceptron that can approximate arbitrary functions. While these neural network models were the first non-linear models to be introduced to machine learning, the difficulties of training them, in particular, getting stuck in local minima and overfitting or generalizing poorly, made these algorithms unattractive options compared to other machine learning methods.

3.2. Types of Deep Learning Models

Deep Learning is not a single algorithm, but rather a collection of models with different topologies that vary from each other in the enrichment architecture and the optimization architecture. Enrichment architecture allows different topologies, both directed and not, gated or not, deepened or shallow, and taught with different rules. There are convolutional neural networks for processing images or 3D data, recurrent neural networks suitable for processing time and discrete sequences of various sizes, and Hopfield, GMLP, and Boltzmann models, which are auto-associators, addressed, and energy models, respectively.

On the optimization side, the type of approximate input/output function pattern to be realized and the ways of connecting the network units are not indefinite. Each of the neurocommunicative models can be organized as a model, producing a probability distribution over its output patterns, given specific input. It is known that if the model metric is properly selected, the supervised training of the type highlighted refers to the problem of minimizing the maximum of the assumed metric input-output errors. Potty learning can be effectively used to explain away the uncertainty not contained in the input data. Different measures of network errors can be used for multi-task learning, where a group of input patterns realizes the distribution for input and output, and different groups for input and output are associated. All these various models and modifications are predicted by the general framework and tested in their capability to solve the specific pattern task. For example, convolutional networks are known to be suitable for processing images and 2D data arrays.

4. Insurance Risk Evaluation

1. Understanding Risk in Insurance

Insurance risk assessment is an essential basis of the insurance business. Insurance poor performance due to incorrect risk assessment is common. For example, in the year 2000, most of the insurance companies went bankrupt because of incorrect risk prediction of the dot-com bubble. Life and health insurance companies could go bankrupt if the numbers on death and hospitalization prediction were not appropriate. It could also be an inevitable situation for non-life insurance companies such as automobile or flood insurance companies. Therefore, proper methods are needed for accurate estimation of the intrigued risk in the insurance field.

"The earlier, the better" is a common saying to reduce the probability of accidents. The ideal way to manage risks is usually prediction instead of post-reaction. Recently, an increased emphasis has been placed on the introduction of prediction procedures for risk management. However, assessing the risk in the insurance business is a complicated process, consisting of several major interrelated factors. Also, the relationships between risk factors and risk functions are sometimes difficult to identify, as the available data may cover different periods, or the available data may be inconsistent with the underlying risk factors. Computer and expert systems allow actuaries as well as risk managers to create more complex models with a greater degree of certainty.

2. Traditional Risk Assessment Methods

In general, the most common method used for risk forecast in insurance business is derived from past experience of similar events occurrence. This technique, known as "component-based", examines historical data to develop a model to forecast loss amounts based on their relationship to several independent accessible variables. The advantages of this approach include its transparency and the fact that it is easy for non-specialists to understand.

4.1. Understanding Risk in Insurance

One of the most important questions that insurance companies face is whether they will pay out more in a certain period than what they take in as premiums. If the latter is lower, then insurance companies will make a profit and remain in good shape, but if not, they will experience a loss. Generally, when the amount going out is higher than the amount coming in, the insurance company will go bust. This is what makes understanding the risk underlying insurance models for such lines of business vital. On a theoretical level, the potential loss can be defined as the difference between the direct and indirect losses to the insurance company. The latter is involved in two stages: Firstly, direct losses occur when a catastrophe or adverse event happens, such as a flood or home fire, such that the customer claims to be compensated by the insurance company. Indirect losses occur due to excess economic disturbances, which lead to the company paying out a higher amount than expected. Though this definition sounds simple, and the detection of direct losses is fairly straightforward, quantifying indirect losses is more complicated since they are essentially unobservable.

Because these losses cannot be measured or estimated directly, more traditional actuarial risk evaluation methods use predictive modeling to explore the relationships between a small number of claims and a set of insurance and policyholder-related factors, such as loss limit, deductible selected, construction type, and area. While this approach provides a simple means for estimating individual loss probability, it cannot perform well on larger less frequent events because of the small loss sample size combined with the very small probability of loss, which means that the predictions extrapolated from the model are likely to be subject to a great deal of error. However, this is not the only problem: Demographic risk assessments develop probability scores over time and use them to assign broader demographic segments into homogeneous risk categories, which leads to a higher loss. Currently, the use of catastrophe models has become the target when it comes to helping insurance model risks with a more probabilistic nature.

4.2. Traditional Risk Assessment Methods

The standard insurance risk management techniques focus on the pooling of similar risks into homogeneous groups, premium risk pricing, and claims monitoring by insurance actuaries and analysts. Such a model is inefficient when risks are heterogeneous for two reasons. First, the focus on hedging a unique property makes the loss-given claim event very large. Heterogeneous risks also imply a very low property diversification factor at the peril area or region level, when compared to the available natural hazard risk transfer protection provided by the insurance products. Low asset diversification reduces economic efficiency by requiring a very high capital cushion to absorb shocks in the event of claim payments above the pooled insurance policy risk premium.

There are two traditional methods for identifying risk and pricing risk premiums available for management decision use. The first is actuarial studies of loss distributions in the past. The ability to price based on a very short time scale is problematic, requiring analysts to substitute their belief about the prospects of risk for evidence of extensive past diversifying. Such an approach can lead to feedback between risk evaluation and business cycles, affecting the future diversification of returns from risk exposures considered. The second is to assign a 'best guess' to the probability of a specific risk scenario occurring in a specific place and to set a scale return factor based on the magnitude of the scenario. Extended loss estimating studies can show large mean premiums and large losses occurring rarely and can readily support such probabilistic risk scenario ranking.

5. Mobile Home Insurance Challenges

Mobile homes represent one of the most cost-effective, functional, and creative housing solutions, especially in the wake of the pandemic. Most often built at a significantly lower price point than traditional homes, mobile homes have become a more attractive asset during uncertain economic conditions. However, mobile homes are more exposed to the elements and are at increased risk during climate-related events, given their low weight and lack of foundations. Due to their cost and the region-specific nature of related perils, mobile home insurance is often subsidized or guaranteed by the federal government. However, lack of underwriting controls, inflation of material costs, and increased activity by catastrophe insurers pose a growing risk to the industry and create variability in the predicted loss estimates. The growth of more advanced analytic techniques, such as deep learning, poses an opportunity for insurers to replace traditional loss models with more advanced predictive models, allowing them to keep up with the growing uncertainty surrounding the industry.

Fulfilling an essential social function, mobile homes serve specific homeowners and are tied to unique risk structure components that differ from the characteristics of ordinary critical and insurable properties. Some of the most significant property risk factors that influence choices made by mobile home insurers but that are not as relevant for their property insurers include the physical characteristics of a property, location and physical environment exposure, and unique homeowner characteristics. During routine weather events in the summer, such as thunderstorms, hurricanes, and wildfires, loss ratios for mobile home insurers are often much higher than loss ratios for property insurers. Moreover, mobile homes also often come with negative social connotations, as their owners are often considered financially underprivileged and retirees without a steady income stream.

Equation 2 : Flood Hazard Impact Score (FHIS):

$$H_f = w_1 \cdot D + w_2 \cdot E + w_3 \cdot V$$

where:

- H_f = Flood hazard impact score
- D = Distance to nearest water body
- E = Elevation above sea level
- V = Historical flood event severity
- w_1, w_2, w_3 = Model-learned weights

5.1. Market Dynamics

In the United States and across the world, the insurance industry currently is in a period of rapid change. Rising flood exposure due to climate change and its effects is becoming clear, and companies are moving to raise prices if they can get away with it. Insurance penetration is low relative to real flood risk. Despite flood risk being presented as a public good, with government funders offering heavily subsidized policies, this is not persuasive for many families, who believe the government will assist if they are flooded and thus find flood insurance unnecessary. Although a recent natural catastrophe has been shown to lead to increases in penetration of insurance against that particular catastrophe, attempts to force higher penetration have been ineffective.

Within the mobile home and flood insurance space, many other mini-dynamics at play create difficulties for insurers. Although mobile home replacement crackdowns provide some level of cost control, mobile homes appreciate in a much more volatile way compared to conventional homes. Reflecting this added risk, mobile home insurance has, traditionally, come at a larger discount. Another unique challenge is that established mobile home insurers do not always interact with the flood insurance market. For mobile homes that are located in flood zones, that means a depth of expertise is not always there. For other markets, such as the auto market, the interactions are much more prevalent and make for a more holistic view of risk. Logically, interaction effects expand the sample size beyond any single type of subject and thus create an improved signal.

5.2. Risk Factors Specific to Mobile Homes

Mobile homes have a unique construction; primarily they are lightweight and elevated above the ground regardless of the underlying soil condition. These characteristics make them particularly vulnerable to damage from natural hazards, the like of which had occurred frequently in the past years in the Southern part of the United States where a large concentration of mobile homes is located, especially during hurricane seasons. Unfortunately, such disasters tend to occur in areas with a high concentration of mobile homes, where not only the property of the mobile home owner is exposed, but mobile homes are often found to have been set up in RV parks or rather this type of construction is used as temporary residences during a large project that has arrived for some time. The damage caused to mobile homes in the context of disasters mentioned above is not only about the damage to the mobile home owner but damage to a mobile home located in those areas can cascade, resulting in injuries to nearby permanent home residents. Such malfunctions are also consequential to nearby areas such as roads which can be blocked by the debris of the mobile homes after the hazard hits or the agency which would have to provide temporary spaces, placing survivors' lives at risk.

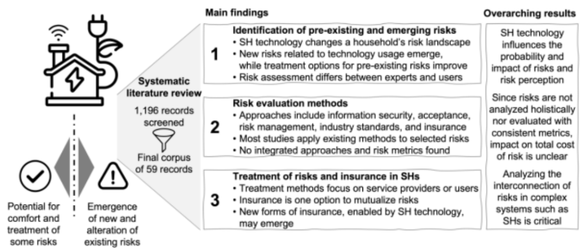

Fig 3 : Treatment of Risks in Smart Homes

Hurricanes, tornadoes, floods, and other natural hazards will provide information for such modeling for identifying damage susceptible locations regarding risk maps; models based on quantity distance or gravity models can be used for inclusion. After going down about their non-linear behavior, finding the exceedance probability of dredging through the Subject Quadratic programming models can be used for identifying risk-sensitive policies; a creative approach for extracting digital information in geographical mapping in real-time has been proposed.

6. Flood Insurance Insights

Flood insurance is an essential risk transfer mechanism, allowing homeowners to quell uncertainty regarding future loss liabilities resulting from flooding. Insurance works by displacing potential economic losses to the insurer, and in turn requires the insured to pay a premium that is determined using current and future risk. It is the calculation of risk that is informed by models that are crucial in setting premiums. Risk must be accurately estimated, as premiums set too low (relative to actual risk) open the insurance to undue exposure risk, while premiums set too high (relative to actual risk) discourage property owners from purchasing coverage, denying themselves a mechanism for loss transfer, while also increasing the costs to the government for disaster relief.

Several models have identified risk using varying methodologies: flood mapping, flood frequency analysis, and hydraulic modeling, among others. These models vary in both methodology and time-indexed risk; however, all metrics and models share the same limitations: reliance on subjective flood history data, limited spatial resolution, outdated data (in the case of mapping), or invasive and expensive hurricane simulations (in the case of hydraulic models). External factors also complicate regulatory estimation due to federal budget constraints, information asymmetries, protection incentives, and the high complexity of model outputs. The advent of deep-learning models offers an alternative risk assessment approach that both overlays and augments existing models. Providing better insurance risk assessment can reduce uncertainties and moral hazard effects in property insurance by improving pricing, product design, and risk management.

6.1. Flood Risk Assessment Techniques

Located in districts responsible for emergency disaster management and housing development, mobile homes are more susceptible to damage from catastrophic flooding than any other type of structure. Owners of mobile homes are usually at a disadvantage since many place less value on their home and live on a low income, making it hard for them to recover from flood damage, especially as climate change impacts such as sea level rise and its acceleration lead to greater severity and higher frequency of disasters. With historical data on flood hazard events being scarce, it calls for advanced flood risk assessment to assess and mitigate risk as well as the impact of climate change on disaster resilience. This study aims to apply satellite imagery and deep learning to screen mobile homes in susceptible flood locations using social economic data. Land use inventory and remote sensing technologies have been widely used in flood risk assessment. A variety of remote imaging methods are employed to monitor flood-prone areas. Remote sensing-based flood analyses focus on single flood events and multi-event comparisons to calibrate disaster parameters and gain further insights. Specific remote sensing techniques include aerial photographs and satellites fitted with synthetic aperture radar-scale imaging for inundation timing and duration. Satellite imagery can also provide monitoring for river and mouth discharges and in-canopy flood processes. Airborne Light Detection and Ranging, bathymetric Lidar, moderate resolution imaging spectral, and Digital Elevation Models have proven effective in modeling inundation patterns, estimating flows, performing discharge monitoring, calculating revisit frequencies, performing flood time series mapping, and identifying flowing water patterns. Although satellite remote sensing has proven successful in flood area estimation from observed water levels, it needs to be supplemented and verified through hydrodynamic models as well as local scale data and information.

6.2. Impact of Climate Change on Flooding

The land surface temperature is predicted to gradually warm relative to the historical averages. This warming will enhance the rate of evaporation and is very likely to result in more rainfall for most areas, although in general bringing rainwater closer to the point of generation within the clouds and implying an increase in the number and intensity of heavy downpours. This aspect is certainly compatible with climate change predictions for many coastal and small island developing states. A warmer atmosphere can carry more water vapor, which can cause some locations and regions to suffer increased flooding due to more frequent and intense rain falling in some locations while adjacent areas continue to experience drought. Longer droughts and shorter, heavier downpours carried by more intense rain can result in soil compaction that may limit the ability of rains to replenish underground aquifers and increase flood magnitude and frequency as rainfall runs off and is not absorbed.

Likewise today, there will still need to be storm surges caused by winds associated with the landfalling hurricanes, with an increase in the proportion of the surge resulting from sea level rise. Global mean sea level rise from thermal expansion and the melting of land ice from Greenland and Antarctica will increase the height of storm surges associated with the most intense hurricanes and tropical storms if the storm path and intensity do not change, as well as increase the flooding threat from the weaker storms. Also, hurricane intensity is expected to increase because of sea surface warming. Global mean sea level has risen due to several factors over the last 25 years, with sea level rise recently becoming a significant part of the rip current hazard for many coastal areas due both to an increase in the mean level and an increase in the number of historic storms of hurricane intensity making landfall, as well as an increase in the associated storm tide.

7. Deep Learning Applications in Insurance

The application of Natural Language Processing algorithms built on neural language models has been exploited in the Insurance industry primarily for annotating, summarizing, classifying, and creating dialogues for insurance-related texts. These capabilities themselves can lead to many downstream applications that focus on either the customer experience or the internal process. The customer experience area includes query handling for customer service for life insurance and retirement plans, insurance claims sentiment detection, and insurance claim decision help. Other applications in this area include the automatic assessment of loyalty programs, proposal risk understanding determination, automatic business strategy recommendation for customer loyalty programs, and the classification of insurance domain query intents and dialogue scenarios. The internal insurance process includes the automatic classification of insurance plan documents, insurance market information concept extraction, insurance knowledge assessment, seal and signature detection in documents, and sentiment prediction from user descriptions. Other applications in this area include the full document automatic understanding of claim submission, digital insurance query classification to optimize operational response, detection of breaches from regulatory notifications, and automatic fire insurance report creation.

Based on image processing, specific claims processes such as automatic damage assessment for home insurance claims, automated claims routing based on received image analysis for car insurance, detection of counterfeit images in the claims process, automated decision support regarding claim rejections, scan of insurance tasks images to assess risks, detection of defects in images presented in mobile insurance claims, as well as the application of distance to centroid algorithms for image analytics in claim decisions. Closer to the proposed research, advanced neural network algorithms for multispectral optical image assessments have been applied to solar plant insurance modeling, wherein an image was processed to examine the risk for photovoltaic power plants located in different geographical areas.

7.1. Predictive Analytics

Predictive analytics can be defined as the use of data to identify the probabilities of future outcomes. Predictive analytics is determined through the efforts of organizations to provide products or services with an anticipation of profit while minimizing the risk of loss. One of the most commonly recognized forms of predictive analytics is risk assessment. Predictive analytics has been employed for several years to build risk assessment models, and in many cases, such models have become standard practice in the insurance industry. Underwriters at property/casualty insurance companies rely on scoring models to evaluate risk and make underwriting decisions. Insurance premium prices are based on an insurance company's assessment of the likelihood that an event resulting in a loss will ever happen and the financial consequences if it does.

A well-established property insurance practice is informed underwriting. Insurers use their proprietary pricing algorithms, which utilize predictive models that assign a certain level of risk to an individual property, to build pricing models at the county or ZIP code level, which do not include property-specific details. These models are then used to determine the price of the coverage offered to the customer, allowing insurers to remain competitive with other insurers while ensuring that the price paid by customers with the same level of risk is approximately the same. Thus, it is counterproductive to risk assess property locations where detailed risk attributes are known, which reduces geographic competition and promotes adverse selection. In the case of mobile homes and flood insurance, unique risk characteristics need to be developed so that adequate coverage can be undertaken by the underwriting company.

7.2. Fraud Detection

Fraud is a recognized aspect affecting the insurance industry negatively. Insurance fraud is mainly carried out in two ways: intentional deception on the part of an insured to obtain some advantage or benefit, and "neutral" fraud, which occurs when a person suffers a loss and declares a claim in a carrier, with product delivery or acceptance occurred before, but when this individual has hidden elements about the loss that have produced damages or have worsened the same. In this section, we approach explicitly the first situation, the most recognized insurance fraud commission. Although negative cases are less than the average of the collected variables for all claims, these events produce important consequences and costs when detected and processes are needed. Therefore, company costs increase due to fraud existence, and as a chain reaction, final customers must pay higher premiums.

Deep learning and AI are rapidly being adopted across many industries to improve safety, security, and welfare. Due in part to the use of these new technologies, this type of insurance fraud becomes easier and less burdensome for perpetrators in today's society. The proposed option to detect and identify fake insurance requests is the use of rules and analytics, models based on actuarial tables and association rules, or some hybrid of the two systems mentioned. Both analytic models and rule-based systems can be used exclusively or combined to integrate their advantages and disadvantages. Such types of models are indeed present in the industry as systems that allow the generation of risk functions and are oriented to fraud detection. However, these tools show limitations, due to their majority training with actuarial models, which detect frauds that have certain characteristics.

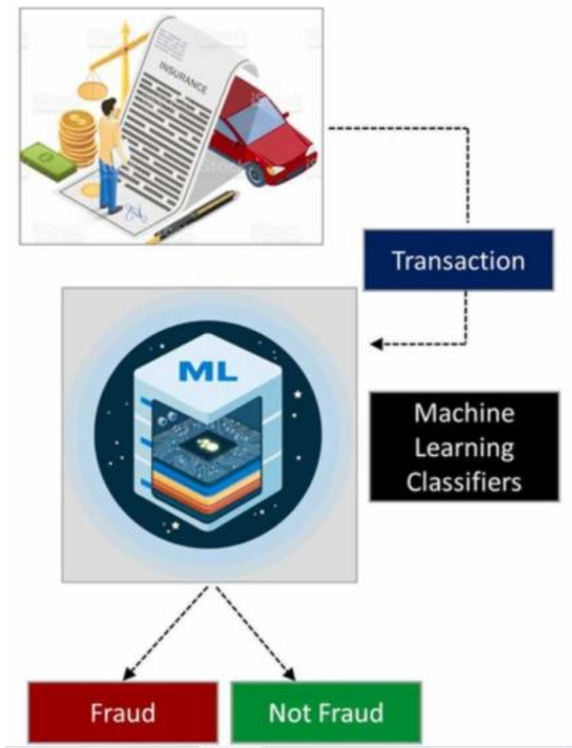

Fig 4 : Insurance fraud detection

7.3. Claims Processing Automation

Considering the different stages of claims processing, the automation of this major insurance function can materialize in multiple ways:

1) Intelligent First Notice of Loss: Despite the limitations described above, deep neural networks applied to natural language processing can be employed to analyze unstructured text data associated with the FNOL process and advance automation of the claims processing workflow at this early point within the claims process.

2) Intelligent Categorization and Assignment: Shortcutting some of the non-value-adding steps that traditionally interrupt the claim process could be achieved by using deep neural networks applied to NLP tasks combined with image analysis models trained with deep learning. This technology could enable the categorization and prioritization of claims based on severity levels, the automatic assignment of those claims that require human intervention to a claims adjuster, and the identification of the non-complicated claims that otherwise require a significantly high cost for the insurer to process manually, for automatic reserving and payout processes by applying reserve estimators that inform necessary time and cost-slots.

3) Voice and Chatbot Integration: Neural networks used for Automatic Speech Recognition, Speech Synthesis, and Natural Language Understanding could bring to another level the use of AI in insurance customer service chatbots and voicebots for processing claims. The low costs associated with such interfaces and the reduced labor time they require could help insurers reduce cost-to-serve, improve identify and investigate fraud, and enhance customer relationships.

8. Data Sources for Risk Evaluation

Having described how to apply deep learning models to assist underwriters in making informed risk evaluation decisions, we now describe how to generate, assemble, and preprocess the data used for model training. Risk evaluation and pricing are usually derived from an understanding of the historical frequency and severity of actual claims for similar properties. This is the primary driver for the premium. Building models to assist in insurance risk evaluation is a supervised learning problem that estimates the conditional probability of flood damage, given the necessary input features. In this section, we discuss the sources for both input features and target labels used for these problems.

Data highlight unique geographic and climate patterns while utilizing deep learning for enhanced underwriting functions in the global insurance stack. Innovative deep learning technology uniquely offers scalable left-tail risk pricing solutions. The development of predictive models requires unique data sets of highly customized attributes. Well-chosen data sources help identify high-conflict insurance regions. Improved data analytics across industries continue growth trends for insurance technology. These trends facilitate game-changing product optimization for specialty and high-risk insurance lines. Exploiting these technology shifts enables exposure management for increased profitability and premium compounding from volatile areas. Enhanced pricing strategies are fostered from these decision-making frameworks using underlying data sources.

Both insured persons and entities stand to benefit from the stratified underwriting insights gained from our deep learning-driven models during the risk evaluation and pricing cycle. Key data providers have surfaced highly differentiated attributes in their curated data feeds. These data attributes benefit pricing optimization and risk reduction during the underwriting process. Most players utilize existing general base mapping and data sources. These include governmental mapping and data on hazard zones including flood zones, flood alerts, and flood observed forecasts. These hazard indices help set expectations for likelihood but not severity or claim costs. Existing general sources are limited, hence specialized sources augment accuracy and model performance.

8.1. Geospatial Data

Understanding risk is paramount to the insurance industry. Insurance pricing and risk evaluation require a lucid understanding of market needs and regional particulars. Actuaries evaluate significant loss events across the United States. Major factors for flooding risk evaluation are geospatial. Flood Insurance Rate Maps (FIRMs) were created to evaluate risk, create insurance baselines, and maintain updated elevation databases over the years. Although FIRMs are managed by a federal agency, all flood insurance in the USA is overseen by a national program. In response to catastrophic flood losses, a risk mapping and assessment initiative was developed. This initiative identifies areas at risk of flooding and also uses extensive amounts of risk and damage data to collect risk ratings. New risk rating systems incorporate available dam failure, levee analysis, landslide, stream flooding, and other relevant modeled geospatial data.

Geospatial data models are utilized extensively for disaster modeling and prediction systems. These approaches create flood maps and predictions: projected geographic information followed by risk probability assessments. Flood risk locations and projections are necessary both for urban planning and long-term resiliency planning by state and local partners. Statistical flood loss models consider financial impact as well as sample loss amount. Geospatial data models are utilized extensively for disaster modeling and prediction systems. The use of risk mapping has catapulted the field into digital geospatial technologies with expertise in hurricane, tropical storm, and general storm windspeed and wind value prediction through wind maps and models for risk evaluation.

8.2. Historical Insurance Claims Data

Insurance companies have accumulated many decades of claims experience that hold valuable insights into what has happened in the past and what may happen in the future. However, insurance data are notoriously sensitive—the companies guard it very closely. Our understanding of insurance risk is therefore limited, along with the techniques available to act on the intelligence. For example, in catastrophic modeling, surrogate data are often used because the companies can be persuaded to share aggregate results but not the underlying microdata.

Two main sources of direct insurance loss data are the Insurance Services Office and the National Oceanic and Atmospheric Administration storm and flood data archives. The former maintains a large database of unique insurance claims out of 30 U.S. states for private homeowners, commercial properties, and vehicles that have been affected by events declared presidential disaster areas. This is referred to as the fire tracking and storm loss databases. The claim records contain information about the insured properties, including location and the amount and type of damages, as well as the event type, date, and magnitude. The Enterprise Loss Data Store consists of anonymized data from participating U.S. insurers. It is allocated by event and zone and contains summary post-event information, such as covered and uncovered losses, claims closed without payment, and new claim counts. The information is available only for specific events and coverage types. Data collection is not exhaustive, with the data-sharing requirements potentially limiting participation.

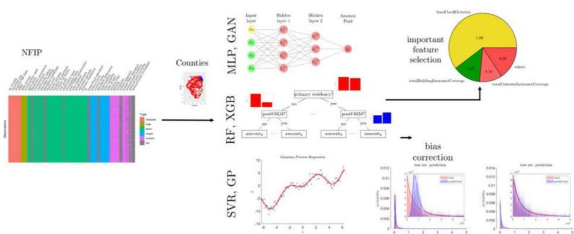

Fig 5 : historical flood insurance claims

8.3. Weather and Climate Data

Weather and climate data play a significant role in catastrophe modeling because the physical mechanisms leading to catastrophic losses, whether windstorm or earthquake, are based on natural laws. The modeling of these phenomena may vary in simplicity, from ad-hoc vulnerability functions to highly complex physical models solving the partial differential equations governing the event, but their starting point is always weather or climate data. Generally, a distinction has to be made between "actual" weather data, which contain the observed values for specific locations or stations, for some specific time, and climate data, which contain the expected or modeled average values, for defined time intervals. Eventually, an interpolation has to be performed to fill the spatial and temporal gaps that arise from the localized time series of the observed data and the coarse temporal and spatial resolution of the climate data.

The historical data usually consists of daily values for temperature, precipitation, dew point, cloud cover, and wind speed, but hourly data is also available for some stations. Several software packages are available to create average climate data for specific time intervals based on the historical weather data series. Climate data reports are also available for temperature, precipitation, wind speed, setup, and others. They provide conclusions about how risk-scoring variables will likely perform in the future.

9. Model Development

This chapter describes the implementation details of our models in this project. The original property data contains 382,261 records. We filtered that data set to eliminate claims during the hurricane from 2005 to the initial period of the records. Thereafter, we split this property dataset into two parts: the training set and validation temporal and spatial validation sets. The training set, which contains 362,478 records, is used to create the models; the temporal validation set includes 7,387 records of policies with a claim recorded from 2015 to 2022, which is used for testing the models based on our probability of prediction; the spatial validation set contains 12,396 records of policies with claims but has not been used to create the models; the features extracted from CNN and VGGNet are from only 17,697 policies. Then, that results in a combined spatial validation set.
1. Data Preprocessing
For applying the machine learning and deep learning methods, we need to do some preprocessing for data cleaning and to define a reasonable model input format. The categorical variables are transformed into binary dummy variables while we removed the Policy Key and the Claim Count variables, which are of no predictive value in the training sets. To cope with the Imbalanced Classification Problem, we add extra data pre-processing strategies to create a Balanced Training Data Set for building all ML models. The least occurred Class 0, with only 37K samples in the whole training set, is selected to crop records ten times higher than that of Class 1 through the program; hence, the final Balanced Training Data Set would have 75K samples for 2.5 Ve-Classification Ratio.

9.1. Data Preprocessing

In this work, we modeled a CNN with a Purpose Life Cycle (PLC). Notably, the PLC refers to the typical steps followed in experimentations with deep neural networks, including problem formulation, deep learning modeling, model training, testing, and deployment, to reveal the hyper-parameters that would provide acceptable accuracy levels in the risk classification problem, while also enhancing efficiencies in terms of model training and testing time. Due to the large volume of extreme weather-related data in existence, and the advent of technologies to store and process the data affordably, we performed rich feature selection. Moreover, due to the lack of predictive superiority of CNNs beyond complex feature representation by traditional machine or ensemble-based learning methods, and also due to factors such as the quantum of extracted features, and the size of the dataset, we limited our CNN to three neural layers in making the architectural design choice.
To convert the continuous values in the feature vectors, which also consisted of categorical variables, we used a library in Python. Also, we eliminated the missing values in the features, where applicable, as well as balancing the observations by either oversampling the minority class of losses or undersampling the majority class of no losses, to resolve the potential class imbalance in the target variable. Finally, before data preprocessing, as a common recommendation, we randomly split the dataset into a 75 percent training and 25 percent testing dataset. The role of the training dataset is to train the model on the relationship between input features and the target label. Conversely, the purpose of the testing dataset is only to test the model on its ability to predict the valid label correctly.

Equation 3 : Deep Risk Classification Loss (DRCL):

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(\hat{y}_{i,c})$$

$\mathcal{L}$ = Categorical cross-entropy loss

N = Number of samples

C = Number of risk categories (e.g., low, medium, high)

$y_{i,c}$ = True label for class c

$\hat{y}_{i,c}$ = Predicted probability for class c

9.2. Feature Selection

Obtaining satisfactory and generalized predictive models using deep learning necessitates large datasets. While typically deep learning models start with a large number of features, effective feature selection may help avoid key problems, such as an extended training period, difficulties in understanding and interpreting the model and acquiring models that lack generalization capabilities. Therefore, selecting relevant variables regularly leads to increased performance. In this work, we perform both methods, and the resulting input structures are examined and compared in terms of predictive performance and interpretability.
Upon reviewing the available features, we found that our final risk-scoring input structure included the following nine discrete variables: (1) Industry Volatility, (2) Financial Size Category, (3) Non-revenue premium, (4) Capitalization, (5) Combined ratio, (6) Rating and (8) Loss reserves to surplus ratio.
We decided not to use the FICO Score or CRC Variables due to great data sparsity, and we did not find them to be significantly informative in model training and validation results. Moreover, we have incorporated features beyond those that have been used to output predictive systems and scores for insurance risk evaluation.

Beyond what we describe, we wanted to have a relatively simple input structure that included a wide variety of economically relevant actuary factors, but which avoided some problems and limitations, such as the highly volatile nature of premiums and the sensitivity to changes in these markets, as well as the lack of data granularity. These are essential and could lead to the model overfitting and failing on unseen/uncertain data.

9.3. Model Training and Validation

With the optimal hyperparameters determined for the LSTM, the model is retrained on all available data to forecast holdout sample MAE. After defining X and y as N x n and N x 1 matrices respectively and performing the train-test split, all entries in the input matrix X are standardized to zero mean and unit variance. Then the training labels vector is reshaped into N_train x 1 x 1 and stored in a tensor y_train while the standardized input matrix X_train is stored in a tensor X_train of shape N_train x n x 1 for efficient model training. LSTM's architecture employs a mini-batch gradient descent method backpropagation through time via the Adam optimizer. Due to the stochastic behavior of mini-batch methods, the model has trained a total of 20 times, and the final training set MAE is selected based on which iteration yielded the smallest training set MAE. Overfit detection is performed using a validation set and entails terminating model training earlier during execution when the validation set MAE ceases to improve.

Standardization of test set data is also required but with the mean and standard deviation values identified from the training set. The training labels for the test set, stored as a tensor y_test, are reshaped. Test set MAE evaluation utilizes the standardized tensor using LSTM's predict() method. LSTM MAE values during training are stored and plotted against the iteration number to visualize model convergence. The same tensor shapes are used for LSTM input data because it requires 3D Tensors of dimensions N_samples x TimeSteps x Features.

10. Evaluation Metrics

The ability to reliably assess model performance through suitable statistical measures is critical to its utility. It is possible to apply regression-based loss functions as regression-like metrics, such as mean squared error. However, these types of losses may not correlate exactly well with the problem of practical interest, which is how well the model accurately captures the decisions we want to ultimately take using the model. We will begin with some of the simplest evaluation metrics used for classification settings to compare models.

1. Accuracy and Precision

For a binary classification problem with possible classes of either 0 or 1, we can define a True Positive (TP) case when the model predicted 1, class 1 was present in the ground truth and found by the model, and a True Negative (TN) case when the model predicted 0, class 0 was present in the ground truth and found by the model. This means there are also False Positive (FP) and False Negative (FN) cases by definition. Using the notation above, the accuracy of the model can be simply defined as:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Conversely, we can also define the Precision, also called Positive Predictive Value, which measures the fraction of retrieved instances that are relevant:

Precision = TP / (TP + FP)

2. Recall and F1 Score

The Recall, or Sensitivity or True Positive Rate, is the fraction of relevant instances that are retrieved by the model and can be defined as:

Recall = TP / (TP + FN)

The F1 score is the harmonic mean of the Precision and Recall and can be defined as:

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

10.1. Accuracy and Precision

The overall performance of our models was evaluated based on two derived metrics: accuracy and precision.

Accuracy measures the sensitivity of all predictions, combining true positives (TP) with false positives (FP) and false negatives (FN). Predictions based on building characteristics seldom yield predicted flood and inferred flood estimates equal to observed outcomes for a substantial part of the insurance portfolios with known claims, but yield very high accuracy scores. Before applying DL to our mobile home and flood risk estimates, we tuned the architecture of the selected model, minimizing the training error while avoiding overfitting.

Precision is an important metric to keep in mind in the insurance business, where excessive non-gainful predictions (false positives) reflect inconsistencies in model predictions. The ratio of mobile homes requesting aid by the federal government for reconstruction of damaged dwellings to the total number of mobile homes requesting such support is below 30%, well below predictions reflected on a model that fails to capture common characteristics of mobile homes requesting aid for damages.

10.2. Recall and F1 Score

The recall (also called sensitivity) is simply the ratio of correctly classified positive observations to the total actual positives. It indicates the model's ability to capture all positive samples. The Recall score takes values in the range [0, 1], and a Model Recall score of 1 means that the model has captured all positive samples. However, having a high Recall score comes with a trade-off of having a low Precision score. Recall is a metric used when a false negative rate is important. In practical applications, if we are going to scan all the pictures in a Sheet, we would want the model to find all the positive samples, even if it gives us some negatives as positive samples. We would want the Recall score to be high.

To understand why Recall might be preferred over Precision, let's consider an example of a Banking credit card company working on an application fraud detection model with two classes to be predicted: "Fraud" and "Not Fraud". The bank does not want any fraud to happen, even if there are some false positives for "Fraud", since this would cost them a lot of money. In such a scenario, the Recall score would be used to judge the efficacy of the model. The F1 score is simply the weighted average of Precision and Recall, but they are weighted in such a way that the best F1 score happens when both are at their highest score levels. When Precision and Recall are at their worst levels F1 has also taken a hit: F1 effectively balances the trade-off between Precision and Recall. In scenarios where one wants to keep high scores for both Precision and Recall, the F1 score can be used as an evaluation metric.

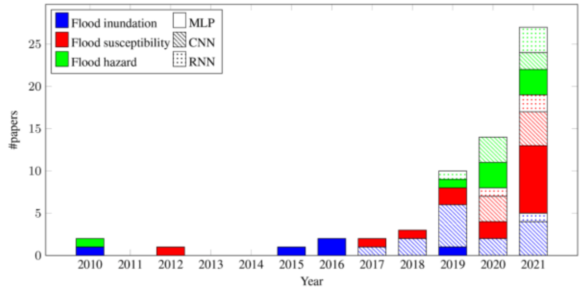

Fig 6 : Deep learning methods for flood mapping

10.3. ROC Curve Analysis

Receiver Operating Characteristic (ROC) curve analysis has emerged as a leading analysis approach for evaluating trained models, owing to its major advantages, including the opportunities to visualize and explore the performance of a classifier system as its discrimination threshold is varied, to show the relative tradeoffs between true positive rate (TPR) and false positive rate (FPR) at various threshold settings, to establish a relationship between the error rates (FPR and false negative rate (FNR)), and to provide a performance evaluation of a classifier.

A downside of using the ROC curve for evaluation is that, in order to construct it, it requires the running of the classifier on all validation/testing data multiple times, once for each value of the threshold used to determine the predicted class. However, this is generally not as much of a problem, as it is commonly sufficient to use only a resolution of 100 or 200 quantiles of the class probability distributions in determining the ROC curve.

The area under the ROC curve (AUC) is also receiving increasing attention in the area of evaluation analysis. The value of the AUC lies in a single number for comparison with other classifiers. Its handicap, however, is that it can neither be used as a stand-alone performance measure nor provide much information by itself, other than a means of rank-ordering different classifiers on testing data. Thus, the ROC curve and AUC should be treated not as substitutes, but as complementary tools in the evaluating tool kit.

11. Conclusion

Many of the recent advances in deep learning have been made using special kinds of data, both in terms of the input manifolds and the output information. In this work, we have shown that deep neural networks can exploit simple kinds of geographic information, such as housing characteristics, region and state codes, and census data, to produce high-quality mobile home and flood insurance risk predictions. We also show that the models, trained using historic losses, are effective at predicting losses in years not used during training. Thus, not only can these models produce exotic risk maps, given an appropriate loss estimate, but we can also rank buildings according to their likelihood of experiencing a loss. Furthermore, unlike many commercial providers of insurance risk maps, the maps provided by these models can be cost-effectively updated annually. Using this procedure would allow a renter to estimate their required renter's insurance coverage limit and a homeowner to estimate their required flood insurance coverage limit.

We illustrate these capabilities using a model for mobile home replacement value, as well as a model for flood policy flood loss estimates. We then present 1,395 risk maps at 1.0° × 1.25° and 1.5° × 2.0° spatial resolutions for the mobile home replacement value model and the flood policy loss estimate model, respectively. The flood loss estimate model has been deployed to provide low-cost nationwide annual updates of the flood loss estimates. Future work should be focused on incorporating improved input demographic and geographic data into the models to improve their quality. Other areas of future research might focus on estimating models for buildings not covered by flood insurance and also for different building types other than mobile homes. Special consideration should be given to the validation of the models in geography and on space-time scales not used during training.

11.1. Final Thoughts and Future Directions

While we see great potential in our deep-learning flood and mobile home insurance risk work, and while we believe our results point toward a novel framework for studying insurance applications, we acknowledge that much of our analysis relies on supervised learning. That is, we do not explore generalizing to unseen data. This is especially important when considering commercial insurance applications, as a private insurer. However, we hope to address future work in this area with a few additional sparks. First and foremost, while we rely on proprietary data, our results may not generalize to unseen locations. This is particularly important since our model on mobile home damage predictions works best in areas around our historical loss data. This begs the question of whether we could easily retrain a neural network model to perform as well in another area as we do here in a new area – there have been discussions on the possibilities of transfer learning for spatially distinct areas. Therefore, while the approach to the prediction is quite general, we may be limited initially by the availability of insurance loss data.

Next, while we only consider using opening ceremony damage to a structure as the supervised labels for the neural network predictions, and framing cost as the damage type to show several building failures, perhaps a better supervised learning setup would be to use all damages to train the model and consider all types of damages. This would alleviate the that there is no guarantee that these opening ceremony damages are being caused by frame-type structures alone. While we see that a majority of the listed damages are data points with types typically associated with frame build types, there is a mixed number of other types as well, as we cannot be sure that these damages could also have non-frame type structures listed as part of the assets in a different area. Perhaps a better setup would be a segmentation model that could differentiate and visualize what damage types may be associated solely with each structure type.

11. References

[1] Sondinti, K., & Reddy, L. (2024). Financial Optimization in the Automotive Industry: Leveraging Cloud-Driven Big Data and AI for Cost Reduction and Revenue Growth. Financial Optimization in the Automotive Industry: Leveraging Cloud-Driven Big Data and AI for Cost Reduction and Revenue Growth (December 17, 2024).

[2] Venkata Krishna Azith Teja Ganti ,Kiran Kumar Maguluri ,Dr. P.R. Sudha Rani (2024). Neural Network Applications in Understanding Neurodegenerative Disease Progression. Frontiers in HealthInformatics, 13 (8) 471-485

[3] Sambasiva Rao Suura. (2024). Integrating Generative AI into Non-Invasive Genetic Testing: Enhancing Early Detection and Risk Assessment. Utilitas Mathematica, 121, 510–522. Retrieved from https://utilitasmathematica.com/index.php/Index/article/view/2046

[4] Venkata Narasareddy Annapareddy. (2024). Harnessing AI Neural Networks and Generative AI for Optimized Solar Energy Production and Residential Battery Storage Management. Utilitas Mathematica, 121, 501–509. Retrieved from https://utilitasmathematica.com/index.php/Index/article/view/2045

[5] Kannan, S. Revolutionizing Agricultural Efficiency: Leveraging AI Neural Networks and Generative AI for Precision Farming and Sustainable Resource Management.

[6] Harish Kumar Sriram. (2024). Leveraging AI and Machine Learning for Enhancing Secure Payment Processing: A Study on Generative AI Applications in Real-Time Fraud Detection and Prevention. Utilitas Mathematica, 121, 535–546. Retrieved from https://utilitasmathematica.com/index.php/Index/article/view/2048

[7] Karthik Chava. (2024). Harnessing Generative AI for Transformative Innovations in Healthcare Logistics: A Neural Network Framework for Intelligent Sample Management. Utilitas Mathematica, 121, 547–558. Retrieved from https://utilitasmathematica.com/index.php/Index/article/view/2049

[8] Komaragiri, V. B. The Role of Generative AI in Proactive Community Engagement: Developing Scalable Models for Enhancing Social Responsibility through Technological Innovations.

[9] Siramgari, D. (2024). Metadata Mastery : Charting the Future of Technical and Data Catalogs in the Era of AI and Cloud. Zenodo. https://doi.org/10.5281/ZENODO.14533320

[10] Daruvuri, R., Patibandla, K., & Mannem, P. (2024). Leveraging unsupervised learning for workload balancing and resource utilization in cloud architectures. International Research Journal of Modernization in Engineering Technology and Science, 6(10), 1776-1784.

[11] Ganesan, P. (2024). AI-Powered Sales Forecasting: Transforming Accuracy and Efficiency in Predictive Analytics. J Artif Intell Mach Learn & Data Sci 2024, 2(1), 1213-1216.

[12] Siramgari, D. R., & Sikha, V. K. (2024). From Raw Data to Actionable Insights: Leveraging LLMs for Automation. Zenodo. https://doi.org/10.5281/ZENODO.14128827

[13] Patibandla, K., Daruvuri, R., & Mannem, P. (2024). Streamlining workload management in AI-driven cloud architectures: A comparative algorithmic approach. International Research Journal of Engineering and Technology, 11(11), 113-121.

[14] Ganesan, P. (2024). Cloud-Based Disaster Recovery: Reducing Risk and Improving Continuity. Journal of Artificial Intelligence & Cloud Computing. SRC/JAICC-E162. DOI: doi. org/10.47363/JAICC/2024 (3) E162 J Arti Inte & Cloud Comp, 3(1), 2-4.

[15] Chaitran Chakilam. (2024). Revolutionizing Genetic Therapy Delivery: A Comprehensive Study on AI Neural Networks for Predictive Patient Support Systems in Rare Disease Management. Utilitas Mathematica, 121, 569–579. Retrieved from https://utilitasmathematica.com/index.php/Index/article/view/2051

[16] Murali Malempati. (2024). Generative AI-Driven Innovation in Digital Identity Verification: Leveraging Neural Networks for Next-Generation Financial Security. Utilitas Mathematica, 121, 580–592. Retrieved from https://utilitasmathematica.com/index.php/Index/article/view/2052

[17] Challa, S. R., Challa, K., Lakkarasu, P., Sriram, H. K., & Adusupalli, B. (2024). Strategic Financial Growth: Strengthening Investment Management, Secure Transactions, and Risk Protection in the Digital Era. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 97-108.

[18] Nuka, S. T. (2024). Exploring AI and Generative AI in Healthcare Reimbursement Policies: Challenges, Ethical Considerations, and Future Innovations. International Journal of Medical Toxicology and Legal Medicine, 27(5), 574-584.

[19] Burugulla, J. K. R. (2024). The Future of Digital Financial Security: Integrating AI, Cloud, and Big Data for Fraud Prevention and Real Time Transaction Monitoring in Payment Systems. MSW Management Journal, 34(2), 711-730.

[20] Intelligent Supply Chain Optimization: AI Driven Data Synchronization and Decision Making for Modern Logistics. (2024). MSW Management Journal, 34(2), 804-817.

[21] Pamisetty, V. (2024). AI Powered Decision Support Systems in Government Financial Management: Transforming Policy Implementation and Fiscal Responsibility. Journal of Computational Analysis & Applications, 33(8).

[22] Revolutionizing Automotive Manufacturing with AI-Driven Data Engineering: Enhancing Production Efficiency through Advanced Data Analytics and Cloud Integration . (2024). MSW Management Journal, 34(2), 900-923.

[23] Leveraging Deep Learning, Neural Networks, and Data Engineering for Intelligent Mortgage Loan Validation: A Data-Driven Approach to Automating Borrower Income, Employment, and Asset Verification. (2024). MSW Management Journal, 34(2), 924-945.

[24] Lahari Pandiri, Subrahmanyasarma Chitta. (2024). Machine Learning-Powered Actuarial Science: Revolutionizing Underwriting and Policy Pricing for Enhanced Predictive Analytics in Life and Health Insurance . South Eastern European Journal of Public Health, 3396–3417. https://doi.org/10.70135/seejph.vi.5903

[25] Mahesh Recharla, (2024). The Role of Agentic AI in Next-Generation Drug Discovery and Automated Pharmacovigilance for Rare and Neurological Diseases. Frontiers in Health Informatics, Vol. 13(8), 4999-5014

[26] Botlagunta Preethish Nandan. (2024). Revolutionizing Semiconductor Chip Design through Generative AI and Reinforcement Learning: A Novel Approach to Mask Patterning and Resolution Enhancement. International Journal of Medical Toxicology and Legal Medicine, 27(5), 759–772. https://doi.org/10.47059/ijmtlm/V27I5/096

[27] Balaji Adusupalli,. (2024). Agentic AI-Driven Identity and Access Management Framework for Secure Insurance Ecosystems . Journal of Computational Analysis and Applications (JoCAAA), 33(08), 2794–2814. Retrieved from https://eudoxuspress.com/index.php/pub/article/view/2225

[28] Paleti, S., Pamisetty, V., Challa, K., Burugulla, J. K. R., & Dodda, A. (2024). Innovative Intelligence Solutions for Secure Financial Management: Optimizing Regulatory Compliance, Transaction Security, and Digital Payment Frameworks Through Advanced Computational Models. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 125-136.

[29] Pallav Kumar Kaulwar,. (2024). Agentic Tax Intelligence: Designing Autonomous AI Advisors for Real-Time Tax Consulting and Compliance. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 2757–2775. Retrieved from https://eudoxuspress.com/index.php/pub/article/view/2224

[30] AI-Powered Revenue Management and Monetization: A Data Engineering Framework for Scalable Billing Systems in the Digital Economy . (2024). MSW Management Journal, 34(2), 776-787.

[31] Abhishek Dodda. (2023). Digital Trust and Transparency in Fintech: How AI and Blockchain Have Reshaped Consumer Confidence and Institutional Compliance. Educational Administration: Theory and Practice, 29(4), 4921–4934. https://doi.org/10.53555/kuey.v29i4.9806

[32] Singireddy, J. (2024). Deep Learning Architectures for Automated Fraud Detection in Payroll and Financial Management Services: Towards Safer Small Business Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 75-85.

[33] Sneha Singireddy. (2024). Leveraging Artificial Intelligence and Agentic AI Models for Personalized Risk Assessment and Policy Customization in the Modern Insurance Industry: A Case Study on Customer-Centric Service Innovations . Journal of Computational Analysis and Applications (JoCAAA), 33(08), 2532–2545. Retrieved from https://eudoxuspress.com/index.php/pub/article/view/2163

[34] Siramgari, D. (2024). Building the Future: Unveiling the AI Agent Stack. Zenodo. https://doi.org/10.5281/ZENODO.14533422

[35] Mannem, P., Daruvuri, R., & Patibandla, K. (2024). Leveraging supervised learning in cloud architectures for automated repetitive tasks. International Journal of Innovative Research in Science, Engineering and Technology, 13(11), 18127-18136.

[36] Ganesan, P. (2023). Revolutionizing Robotics with AI. Machine Learning, and Deep Learning: A Deep Dive into Current Trends and Challenges. J Artif Intell Mach Learn & Data Sci, 1(4), 1124-1128.

[37] The Future of Banking and Lending: Assessing the Impact of Digital Banking on Consumer Financial Behavior and Economic Inclusion. (2024). MSW Management Journal, 34(2), 731-748.

[38] Satyaveda Somepalli. (2024). Leveraging Technology and Customer Data to Conserve Resources in the Utility Industry: A Focus on Water and Gas Services. Journal of Scientific and Engineering Research. https://doi.org/10.5281/ZENODO.13884891

[39] Patibandla, K., & Daruvuri, R. (2023). Reinforcement deep learning approach for multi-user task offloading in edge-cloud joint computing systems. International Journal of Research in Electronics and Computer Engineering, 11(3), 47-58.

[40] Singireddy, S., Adusupalli, B., Pamisetty, A., Mashetty, S., & Kaulwar, P. K. (2024). Redefining Financial Risk Strategies: The Integration of Smart Automation, Secure Access Systems, and Predictive Intelligence in Insurance, Lending, and Asset Management. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 109-124.

[41] Intelligent Technologies for Modern Financial Ecosystems: Transforming Housing Finance, Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions. (2024). MSW Management Journal, 34(2), 953-971.

[42] Abhishek Dodda. (2023). NextGen Payment Ecosystems: A Study on the Role of Generative AI in Automating Payment Processing and Enhancing Consumer Trust. International Journal of Finance (IJFIN) - ABDC Journal Quality List, 36(6), 430-463. https://ijfin.com/index.php/ijfn/article/view/IJFIN_36_06_017

[43] Hara Krishna Reddy Koppolu, Venkata Bhardwaj Komaragiri, Venkata Narasareddy Annapareddy, Sai Teja Nuka, & Anil Lokesh Gadi. (2023). Enhancing Digital Connectivity, Smart Transportation, and Sustainable Energy Solutions Through Advanced Computational Models and Secure Network Architectures. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 1905–1920. https://doi.org/10.53555/jrtdd.v6i10s(2).3535

[44] Kaulwar, P. K. (2023). Tax Optimization and Compliance in Global Business Operations: Analyzing the Challenges and Opportunities of International Taxation Policies and Transfer Pricing. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 150-181.

[45] Paleti, S. Agentic AI in Financial Decision-Making: Enhancing Customer Risk Profiling, Predictive Loan Approvals, and Automated Treasury Management in Modern Banking.

[46] Adusupalli, B., & Insurity-Lead, A. C. E. The Role of Internal Audit in Enhancing Corporate Governance: A Comparative Analysis of Risk Management and Compliance Strategies.

[47] Moore, C., & Routhu, K. (2023). Leveraging Machine Learning Techniques for Predictive Analysis in Merger and Acquisition (M&A). Available at SSRN 5103189.

[48] Chinta, P. C. R., Katnapally, N., Ja, K., Bodepudi, V., Babu, S., & Boppana, M. S. (2022). Exploring the role of neural networks in big data-driven ERP systems for proactive cybersecurity management. Kurdish Studies.

[49] Routhu, K., Bodepudi, V., Jha, K. M., & Chinta, P. C. R. (2020). A Deep Learning Architectures for Enhancing Cyber Security Protocols in Big Data Integrated ERP Systems. Available at SSRN 5102662.

[50] Maka, S. R. (2023). Understanding the Fundamentals of Digital Transformation in Financial Services: Drivers and Strategic Insights. Available at SSRN 5116707.

[51] Sadaram, G., Sakuru, M., Karaka, L. M., Reddy, M. S., Bodepudi, V., Boppana, S. B., & Maka, S. R. (2022). Internet of Things (IoT) Cybersecurity Enhancement through Artificial Intelligence: A Study on Intrusion Detection Systems. Universal Library of Engineering Technology, (2022).