Machine Learning and the Insurance Underwriter

Hassan Shuman
CTO | GenAI Pioneer | AWS & Azure Expert
Transforming Enterprises with GenAI, Clo...
21 articles + Follow

January 3, 2023

Open Immersive Reader

Last week, I discussed how ML can help an insurance actuary. This week, I wanted to explore the underwriters role and how ML can help them.

For clarity, I wanted to articulate my understanding of what an Insurance Underwriter does, at a very high level as my speciality is in the Cloud and AI/ML space does.

I will then focus on a specific example of where ML can help the underwriter and finish off this article with out-of-the-box Azure AND AWS architectures that can easily be set up to help Insurance companies with their analysis/predictions.

So what does an insurance underwriter do

An insurance underwriter is a professional who evaluates insurance applications and decides whether to offer

coverage to an individual or business. They are responsible for determining the level of risk associated with insuring a person or entity, and for setting the premiums for the insurance policy based on that risk. This involves reviewing the applicant's medical history, financial records, and other relevant information to assess their likelihood of making a claim. Insurance underwriters may also be responsible for evaluating the terms and conditions of insurance policies to ensure that they are fair and reasonable for both the applicant and the insurance company.

How can machine learning help/replace the insurance underwriter

Machine learning algorithms can be used to automate the process of insurance underwriting to some extent. By analyzing large amounts of data and identifying patterns and trends, machine learning models can help underwriters make more informed and accurate risk assessments. For example, a machine learning model might be trained on data about an applicant's medical history, driving record, and other factors to predict the likelihood that they will file an insurance claim.

There are a few potential benefits to using machine learning in insurance underwriting. One is that it can be faster and more efficient than manual underwriting, as the machine learning model can analyze a large number of variables and make a decision in a matter of seconds. Another benefit is that it can be more objective and unbiased than human underwriters, as it is not subject to the same cognitive biases and prejudices.

Four areas in which Machine Learning can help underwriters

1. Risk assessment: Machine learning algorithms can analyze a wide range of data about an applicant, including medical records, financial records, and

demographic information, to predict the likelihood that they will file an insurance claim.

2. Fraud detection: Machine learning can be used to identify patterns and anomalies in insurance claims data that may indicate fraudulent activity. This can help underwriters identify and reject fraudulent claims more quickly and accurately.
3. Premium pricing: Machine learning can be used to analyze data about an applicant's risk profile and the market conditions to determine appropriate premiums for insurance policies.
4. Policy recommendation: Machine learning can be used to recommend insurance policies to applicants based on their risk profile and other relevant factors. This can help underwriters present the most suitable policy options to applicants and improve the efficiency of the underwriting process.

I discussed risk assessments before, so here I will focus more on Policy Recommendations, Fraud Detection and Premium pricing only.

Policy Recommendations

There are several types of machine learning algorithms that can be used to recommend insurance policies. Some examples include:

1. Decision trees: This type of algorithm uses a flowchart-like structure to make predictions based on a series of decisions. Each decision is based on a specific feature or attribute of the data, and the algorithm follows the path that leads to the most likely outcome.
2. Random forests: This is an ensemble method that combines the predictions of multiple decision trees to make a final prediction. Random forests are often used in insurance underwriting because they are able to handle large and complex datasets and are relatively resistant to overfitting.
3. Naive Bayes: This is a simple but effective classification algorithm that uses Bayes' theorem to

make predictions based on the probability of an event occurring. It is often used in insurance underwriting to predict the likelihood that an applicant will file a claim.

4. K-nearest neighbors: This is a non-parametric algorithm that makes predictions based on the "nearest" data points in the training set. It is often used in insurance underwriting to identify patterns and trends in the data that can be used to predict an applicant's risk profile.
5. Logistic regression: This is a linear model that is often used in classification tasks, including insurance underwriting. It estimates the probability of an event occurring based on a set of predictor variables.

Here are four examples of how machine learning has been used to recommend insurance policies:

1. A health insurance company implemented a machine learning system to recommend policies to prospective customers based on their medical history, demographic information, and other relevant factors. The system analyzed data about past policyholders and identified patterns and trends that were associated with successful policy outcomes. By using this system, the company was able to recommend policies that were more likely to be suitable for the individual applicant, resulting in improved customer satisfaction and retention.
2. A provider of pet insurance implemented a machine learning system to recommend policies to pet owners based on the breed, age, and medical history of their pets. The system analyzed data about past policyholders and identified patterns and trends that were associated with successful policy outcomes. By using this system, the company was able to recommend policies that were more likely to be suitable for the individual pet, resulting in improved customer satisfaction and retention.
3. An insurance company that specializes in coverage for small businesses implemented a machine learning system to recommend policies to business owners based on their industry, size, and location. The system

analyzed data about past policyholders and identified patterns and trends that were associated with successful policy outcomes. By using this system, the company was able to recommend policies that were more likely to be suitable for the individual business, resulting in improved customer satisfaction and retention.

4. A provider of travel insurance implemented a machine learning system to recommend policies to travelers based on their destination, duration of travel, and other relevant factors. The system analyzed data about past policyholders and identified patterns and trends that were associated with successful policy outcomes. By using this system, the company was able to recommend policies that were more likely to be suitable for the individual traveler, resulting in improved customer satisfaction and retention.

Fraud Detection

Machine learning algorithms can be used to identify patterns and anomalies in insurance claims data that may indicate fraudulent activity. By analyzing large amounts of data, machine learning models can learn to identify patterns and trends that are indicative of fraudulent behavior. For example, a machine learning model might be trained on data about past fraudulent claims to identify common characteristics or patterns that are present in these claims.

Once the machine learning model has been trained, it can be used to analyze new claims data and flag any claims that exhibit these same characteristics or patterns. This can help underwriters identify potentially fraudulent claims more quickly and accurately, and allow them to take action to reject or investigate these claims.

Here are two examples of how machine learning has been used to help insurance underwriters detect and reject fraudulent applications:

1. In 2017, a major insurance company implemented a machine learning system to identify fraudulent auto

insurance claims. The system analyzed data about past claims, including information about the vehicles, drivers, and damages, to identify patterns and trends that were indicative of fraudulent activity. Using this system, the company was able to identify and reject fraudulent claims more quickly and accurately, resulting in significant cost savings.

2. In 2018, a leading provider of life insurance implemented a machine learning system to detect fraudulent applications. The system analyzed data about the applicants, including their medical records, financial history, and other relevant information, to identify patterns and anomalies that were indicative of fraudulent behavior. By using this system, the company was able to identify and reject a significant number of fraudulent applications, helping to protect against financial losses.

Premium pricing

Machine learning algorithms can be used to analyze data about an applicant's risk profile and the market conditions to determine appropriate premiums for insurance policies. By analyzing large amounts of data, machine learning models can learn to identify patterns and trends that are associated with different levels of risk.

For example, a machine learning model might be trained on data about an applicant's medical history, financial records, and other relevant factors to predict the likelihood that they will file an insurance claim. The model could then use this prediction to recommend a premium that reflects the level of risk associated with insuring the applicant.

By using machine learning to assist with premium pricing, underwriters can more accurately and efficiently set premiums that are fair and reasonable for both the applicant and the insurance company. This can help to improve the overall efficiency of the underwriting process and enhance customer satisfaction.

Here are two examples of how machine learning has been used to help insurance companies with premium pricing:

1. A major health insurance company implemented a machine learning system to assist with premium pricing. The system analyzed data about the medical histories and other risk factors of applicants to predict the likelihood that they would file a claim. The company used this prediction to set premiums that reflected the level of risk associated with insuring each applicant. By using this system, the company was able to more accurately set premiums and reduce the number of underpriced policies, resulting in improved financial performance.
2. A provider of home insurance implemented a machine learning system to assist with premium pricing. The system analyzed data about the characteristics of the homes being insured, as well as data about the locations and other risk factors, to predict the likelihood that the homes would be damaged. The company used this prediction to set premiums that reflected the level of risk associated with insuring each home. By using this system, the company was able to more accurately set premiums and reduce the number of underpriced policies, resulting in improved financial performance.

Possible Cloud Architectures for Azure and AWS

An Azure Landing Zone that can be used

1. Data collection: The first step in the process would be to collect the data that will be used to train the machine learning model. This might involve using Azure Data Factory to extract data from various sources, such as applicant databases, market data feeds, and past claims data.
2. Data preparation: Once the data has been collected, it would need to be cleaned, transformed, and standardized before it could be used to train the model. This might involve using Azure Data Lake Storage or Azure Blob Storage to store the data, and

Azure Databricks or Azure HDInsight to perform the data preparation tasks.

3. Model training: Next, the machine learning model would be trained on the prepared data. This might involve using Azure Machine Learning Designer to select and configure the model, and Azure Machine Learning to train and evaluate the model. The model would be trained to predict the likelihood that an applicant would file an insurance claim based on their risk profile and other relevant factors.
4. Model deployment: Once the model has been trained and tested, it would need to be deployed so that it can be used to predict premiums for new applicants. This might involve using Azure Container Instances or Azure Kubernetes Service to host the model, and Azure API Management to manage and secure the API that provides access to the model.
5. Premium prediction: Underwriters would use the deployed machine learning model to predict premiums for new applicants. They would input the relevant data about the applicant into the model using an API call, and the model would return a prediction of the likelihood that the applicant would file an insurance claim. The underwriter could then use this prediction to set a premium that reflects the level of risk associated with insuring the applicant.
6. Monitoring and maintenance: The machine learning model would need to be regularly monitored and maintained to ensure that it is performing accurately and effectively. This might involve periodically retraining the model on new data to ensure that it is up to date, as well as monitoring the model's performance and making any necessary adjustments. This might involve using Azure Monitor or Azure Log Analytics to monitor the model's performance, and Azure Machine Learning to retrain the model as needed.

AWS Example Architecture

1. Data collection: The first step in the process would be to collect the data that will be used to train the machine learning model. This might involve using AWS Glue or AWS Data Pipeline to extract data from various sources, such as applicant databases, market data feeds, and past claims data.
2. Data preparation: Once the data has been collected, it would need to be cleaned, transformed, and standardized before it could be used to train the model. This might involve using Amazon S3 to store the data, and Amazon EMR or AWS Lake Formation to perform the data preparation tasks.
3. Model training: Next, the machine learning model would be trained on the prepared data. This might involve using Amazon SageMaker to select and configure the model, and to train and evaluate the model. The model would be trained to predict the likelihood that an applicant would file an insurance claim based on their risk profile and other relevant factors.
4. Model deployment: Once the model has been trained and tested, it would need to be deployed so that it can be used to predict premiums for new applicants. This might involve using Amazon Elastic Container Service or Amazon Elastic Kubernetes Service to host the model, and Amazon API Gateway to manage and secure the API that provides access to the model.
5. Premium prediction: Underwriters would use the deployed machine learning model to predict premiums for new applicants. They would input the relevant data about the applicant into the model using an API call, and the model would return a prediction of the likelihood that the applicant would file an insurance claim. The underwriter could then use this prediction to set a premium that reflects the level of risk associated with insuring the applicant.
6. Monitoring and maintenance: The machine learning model would need to be regularly monitored and maintained to ensure that it is performing accurately and effectively. This might involve periodically

retraining the model on new data to ensure that it is up to date, as well as monitoring the model's performance and making any necessary adjustments. This might involve using Amazon CloudWatch or AWS X-Ray to monitor the model's performance, and Amazon SageMaker to retrain the model as needed.

Report this

Published by

Hassan Shuman
CTO | GenAI Pioneer | AWS & Azure Expert Transforming Enterprises with...
Published • 1y

21 articles

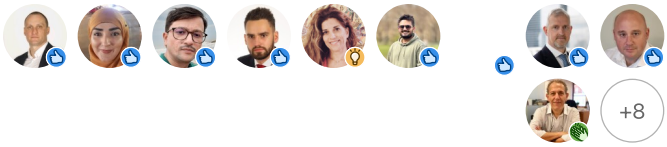
+ Follow

Last week, I discussed how ML can help an insurance actuary. This week, I wanted to explore the underwriters role and how ML can help them.

I will cover two high level architectures, one on AWS an another on Azure. Reach out if you want help in detailing the solution for your requirements on either cloud where I can even share my Terraform code to create the landing zones to help you start. #aws #cloud #azure #machinelearning #insurance #underwriting #sagemaker #azuredatabricks #azuremachinelearning #cloudinfrastructure #terraform #ml

Like Comment Share 20 • 2 comments

Reactions



2 Comments

Most relevant

Add a comment...

Asim Khan • 3rd+
Chief Risk Officer at Zurich Insurance Company
1y (edited) •••

Very interesting read Hassan Shuman. Hope you are doing well

Like · 1 | Reply

Yasmine Bark • 3rd+
NHS England | Pharmacy | Public Health | Healthcare Strategy | Digital Health Innovation
1y •••

Best of luck Hassan Shuman 🙏

Like · 1 | Reply

Hassan Shuman
CTO | GenAI Pioneer | AWS & Azure Expert Transforming Enterprises with GenAI, Cloud Migration, and Innovation | CIO/CTO Advisor | ex-IBM, Accenture,

+ Follow

More from Hassan Shuman

A Year On: Securing IoT Healthcare Through the Case of Pacemakers

Hassan Shuman on LinkedIn

OpenAI ChatGPT vs Amazon Bedrock: A Comparison of Generative...

Hassan Shuman on LinkedIn

The Unlikely Sewing Circle: How Tailors and Data Scientists Craft Perfection

Hassan Shuman on LinkedIn

See all 21 articles