NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS

MEMORANDUM

To: Superintendent Elizabeth Kelleher Dwyer,
Chair of the Big Data and Artificial Intelligence (H) Working Group

From: Commissioner Kevin Gaffney,
Chair of Workstream One (Surveys) of the Big Data and Artificial Intelligence (H) Working Group

Cc: Nine-State Subject Matter Expert Group; Kris DeFrain (NAIC)

Date: December 8, 2022

Re: 2021 Private Passenger Auto Artificial Intelligence/Machine Learning Survey Analysis

---

The 2021 Private Passenger Auto Artificial Intelligence/Machine Learning Survey (PPA AI/ML Survey) [1] was conducted to inform the work of the Big Data and Artificial Intelligence (H) Working Group in support of its charge to:

> *Research the use of big data and artificial intelligence (AI) in the business of insurance, and evaluate existing regulatory frameworks for overseeing and monitoring their use. Present findings and recommend next steps, if any, to the Innovation and Technology (EX) Task Force, which may include model governance for the use of big data and AI for the insurance industry.*

The survey was conducted under the market examination authorities of nine (9) requesting states (Connecticut, Illinois, Iowa, Louisiana, Nevada, North Dakota, Pennsylvania, Rhode Island, and Wisconsin) and completed by insurers who write private passenger auto (PPA) insurance in one of the nine participating states and have at least $75 million in national PPA insurance premium for 2020. The following subject matter experts (SMEs) represented the nine states:

CT: George Bradner
IL: Erica Weyhenmeyer
IA: Andria Seip
LA: Nichole Torblaa
ND: Mike Andring and Chris Aufenthie
NV: Gennady Stolyarov

[1] The 2021 PPA AI/ML survey was conducted under the market conduct examination authority of nine states: Connecticut, Illinois, Iowa, Louisiana, Nevada, North Dakota, Pennsylvania, Rhode Island, and Wisconsin. Subject matter experts (SMEs) from these states opted to limit the survey request to only larger companies, defined as those PPA writers with more than $75 million in 2020 direct premium written. The SMEs also limited the scope to only "advanced" AI/ML models (which excludes models like generalized linear models [GLMs], which are used by 85% of companies in rate filings). A total of 193 responses were received.

PA: Michael McKenney
RI: Matt Gendron
WI: Timothy Cornelius

This memorandum contains the SMEs' summary of the survey analysis, key takeaways, and some recommendations for next steps. The SMEs also approved public distribution of the attached NAIC staff's survey analysis, which provides more detail about the survey results.

SURVEY ANALYSIS SUMMARY

"More Advanced" Artificial Intelligence/Machine Learning Model Use by Companies

The survey was intentionally limited to the reporting of "more advanced" types of AI models, so the data should be interpreted as applying to that subset of insurers' predictive models. Out of 193 companies completing the survey, 169 companies currently use, plan to use, or plan to explore using AI/machine learning (ML) as defined for this survey. This equates to approximately 88% of reporting companies.

Among insurer operations areas, companies reported varying levels of AI/ML use, from only 2% in the loss prevention area to 70% in claims operations. In order from maximum to minimum use, the percentage of companies using AI/ML were: claims, 70%; marketing, 50%; fraud detection, 49%; rating, 27%; underwriting, 18%; and loss prevention, 2%. The following shows the predominant uses, the levels of decision-making, and how often models are developed in-house or externally for each insurer operation.

MODELS BY INSURER OPERATIONS

Claims Models

Uses: In insurance claims operations, companies reported currently using AI/ML claims models mostly as an informational resource for adjusters. Other AI/ML claims model uses identified by at least 50 insurers are: 1) to determine claim settlement amounts; 2) to make claim assignment decisions; 3) to evaluate images of loss; and 4) for "other" claim-related functions.

Level of decision-making: Determination of settlement amount tends to include augmentation. Claim assignment decisions tend to be automated, or at least the models provide augmentation.

In-house or third-party: Models for claim approval, claim assignment decisions, adjusters' informational resource, and other claim-related functions tend to be developed in-house. Models used to determine settlement amounts and evaluate images of the loss tend to be developed by third parties.

Fraud Models

Uses (identified by at least 50 companies: In insurance fraud detection, companies reported currently using AI/ML models mostly to refer claims for further investigation.

Level of decision-making: For the referral of claims for further investigation, most of the levels of decisions are a mixture of augmentation and support. Other fraud-detection models are used for support.

In-house or third-party: Models to detect first-party and third-party liability tend to be developed by third parties. Other fraud detection is a mixture of in-house and third-party models.

Marketing Models

Uses (identified by at least 50 companies): Companies use marketing models for targeted online advertising.

Level of decision-making: Many of the marketing models are automated with no human intervention on execution. Marketing models are mostly automated when used for targeted online marketing, direct online sales, provision of offers to existing customers, and other marketing-related functions. When identifying recipients of mail or phone advertising, there is most often augmentation, where a model provides an answer and advises the human who is making the decision. When identifying potential customer groups, the number of models is evenly split between all three levels of decision-making. Demand modeling is evenly split between augmentation and support to the human.

In-house or third-party: Marketing models being used by insurance companies are equally developed in-house (with or without third-party assistance) and purchased from a third party. Two exceptions are that third-party models are used for targeted online advertising, and in-house models are used for the provision of offers to existing customers.

Rating, Underwriting, and Loss-Prevention Models

With a focus on "more advanced" AI/ML models, there are fewer rating, underwriting, and loss-prevention models reported. Therefore, the data for some detailed questions is less credible. This may, however, simply be a reflection of the limited extent of the deployment of such more advanced models to date. This corroborates the understanding of the SMEs that the majority of rating approaches that PPA insurers use today continue to involve more traditional ratemaking techniques and older-generation static predictive models. The more advanced AI/ML models currently constitute a minority of the models used by insurers in rating and underwriting.

Only 52 companies reported current rating model uses, and the majority of those were for rating class determination or "other" uses; the levels of decision are a mix of all types; and almost all rating models were developed in-house.

Only 34 companies reported current underwriting model uses, and the majority of those were for "other" uses; the levels of decision are a mix of all types; and most underwriting models are developed in-house.

Only three companies reported current loss-prevention uses, and all of those were for "identification of high-risk consumers"; the levels of decision are support only; and most loss-prevention models are developed in-house.

DATA ELEMENTS BY INSURER OPERATIONS

The following are the data elements used by at least 50 companies in the different insurer operations. The data sources vary by data element, but for the most-cited data elements, the source tends to be internal.

- Claims
    - Data Elements: Vehicle-Specific Data, Loss Experience, and Medical
- Fraud Detection:
    - Data Elements: Loss Experience, Vehicle-Specific Data, and Medical
- Marketing:
    - Data Elements: Demographic

- Rating, Underwriting, and Loss Prevention: No data elements were used by at least 50 companies.

CUSTOMER DATA CORRECTION

Many companies discussed having a dispute process. The form of the dispute process ranged from calling the company or agent to dispute erroneous data to allowing policyholders to correct erroneous data themselves through an app.

Data Element Information Provided to Consumers

Insurers were asked to identify if they were providing additional information about data elements to consumers *other than what is required by law*. The answer, although the number of reporting companies is lower than expected, is almost unanimously "no" for each of the insurer operations, except for rating, which had about 32% of the responses reporting "yes." The second question is similar but asks whether consumers are told the purposes of data elements beyond what is required by law. For this question, the answer was almost unanimously "no," except for rating, which had about 26% of the responses reporting "yes."

Consumer Opportunity to Challenge or Correct Data

For the question on whether consumers have the opportunity to challenge or correct their specific data outside of processes for the federal Fair Credit Reporting Act (FCRA), many did not answer. Of those who answered this question, about 50% said "yes" for rating and underwriting; 40% said "yes" for claims and marketing; 15% said "yes" for fraud detection; and less than 10% said "yes" for loss prevention.

GOVERNANCE

The purpose of the model governance questions is to obtain a better understanding regarding a company's awareness of specific risk areas tied to selected categories in the NAIC Artificial Intelligence Principles. A sizable number of companies did not respond to these questions.

Insurers were asked if the following are *documented* in a governance program:

- Fairness and ethics considerations.
- Accountability for data algorithms' compliance with laws, as well as intended and unintended impacts.
- Appropriate resources and knowledge involved to ensure compliance with laws, including those related to unfair discrimination.
- Ensure transparency with appropriate disclosures, including notice to consumers specific to data being used and methods for appeal and recourse related to inaccurate data.
- AI systems are secure, safe, and robust, including decision traceability and security and privacy risk protections.

Insurers' answers were fairly consistent between each question. The answers for rating tended to be higher percentages of "yes" than for the other insurer operations. The transparency question received noticeably fewer affirmations than others. While the percentage of "yes" responses averaged 67% for most questions, the transparency question only received 56% "yes" responses.

THIRD-PARTY DATA SOURCES AND MODELS

Insurers identified third-party vendors they use to purchase models and/or data. There were 2,531 models listed in the survey (with some models being counted more than once because of separate uses for the same model); 1,073 (42%) are developed by a third party, and 1,458 (58%) are developed internally. After grouping the similarly named third parties, there are 76 unique third-party companies listed in the survey whose models are being used by insurers. Marketing has 39 different third parties listed, followed by claims with 28. For data purchases, there were 104 unique third parties listed as data sources in the survey.

CONCLUSION/NEXT STEPS

The insight gained from the survey will be used to supplement state insurance regulators' knowledge of the current regulatory framework around AI/ML, governance, consumers, and third parties and to evaluate whether any changes should be made to the regulatory frameworks.

Following are some potential next steps, including many activities already in progress. This list is not intended to be complete, but it may be helpful as a starting point for discussions and decision-making about what next steps to take at the NAIC:

- Evaluate the survey analysis and determine whether to further explore the following subjects:
    - Insurer AI/ML model usage and the level of decision-making (i.e., the amount of human involvement in decision-making).
    - Insurer data elements.
    - Insurers' governance frameworks and the documentation of such.
    - Consumer data recourse.
    - Third-party regulatory framework.
- Create a risk hierarchy to prioritize the need for more model governance and insurer oversight. The general concept is that more oversight of a model will be needed as the consumer risk or impact increases from the modeling or models.
- Evaluate consumer data recourse. Insurers report a wide variety of methods for consumers to evaluate and correct data used by insurers. Some methods are short and easy, such as using an app to correct data, and other methods are more time-consuming and require personal contact with the agent or company. In some cases, consumers may not know their data is being used, so consumer transparency is a priority. *(Privacy Protections (D) Working Group)*
- Evaluate the regulatory framework around the use of third-party models and third-party data. Evaluate the ability of insurers and state insurance regulators to obtain needed information from third parties and for regulators to oversee this work either through the insurers or third parties in some way. *(Workstream Two of the Big Data and Artificial Intelligence (H) Working Group)*
- Evaluate concerns about third-party concentration by insurer use. *(Workstream Two of the Big Data and Artificial Intelligence (H) Working Group)*
- Determine whether additional white papers on best practices would be useful on subjects in the AI/ML space.

Additional information was collected and analyzed in a confidential June 30, 2022, NAIC staff report, which is available to state insurance regulators by contacting Kris DeFrain, kdefrain@naic.org. This report is confidential because data was collected in a market conduct examination of the nine states and agreed confidentiality protections were applied.