

To Prevent Algorithms From Heading Off Course, Regulators Consider Testing

September 08, 2022

As the various NAIC groups and state regulators continue to ascertain the seaworthiness of insurers' use of consumer data, algorithms, and machine learning, these lookouts have set their sights on unfair discrimination. The Accelerated Underwriting (A) Working Group's educational report states:

- Due to the fact accelerated underwriting relies on non-traditional, nonmedical data and predictive models or machine learning algorithms, it may lead to unexpected or unfairly discriminatory outcomes even though the input data may not be overtly discriminatory. It is critical to test the conclusions up front, on the back end, as well as, randomly, to ensure the machine learning algorithm does not produce unfairly discriminatory ratings or ones that are not actuarially sound. Testing can also be important in determining if a machine learning algorithm is accurate across demographic categories. Such scrutiny is especially important when behavioral data is utilized. Behavioral data may include gym membership, one's profession, marital status, family size, grocery shopping habits, wearable technology, and credit attributes. Although medical data has a scientific linkage with mortality, behavioral data may lead to questionable conclusions without reasonable explanation.

At the 2022 Summer National Meeting of the Big Data and Artificial Intelligence (H) Working Group (Big Data WG), Superintendent Elizabeth Dwyer confirmed that testing the results of an algorithm is especially important for algorithms that may change and evolve over time. Consumer representative Birny Birnbaum commented that testing consumer outcomes is an "essential component" to addressing bias and that a uniform approach is needed across insurers. Similarly, at the meeting of the Innovation, Cybersecurity, and Technology (H) Committee (H Committee):

- The Society of Actuaries noted that after implementation of an algorithm, insurers cannot just "set it and forget it" but must continue to evaluate the algorithm performance after deployment to improve the model's performance.

- Google noted that there must be oversight as an algorithm makes decisions and that responsibility needs to be baked in at every stage and suggested that (i) a model's outputs must be reviewed to evaluate whether the performance of the model compares to the model's ground truth and (ii) loss ratios must be tracked over time across communities to understand whether there are any systematic gaps across models or products.
- Professor Daniel Schwarcz commented that in all cases in which a problem with a model was discovered, for example with facial recognition models, it was discovered because the model was tested or audited on the back end.

Schwarcz indicated insurers might have to walk the plank for their reluctance to collect information about statutorily protected groups, which prevents them from determining when an algorithm produces biased results. Dwyer confirmed with the presenters that methods such as the Bayesian Improved Surname and Geocoding (BISG) method for inferring race could be used to attain the same result without the need to collect such consumer data. A similar approach is being considered by the Colorado Insurance Department.

At the Big Data WG summer meeting, Milliman consultants explained four different tacks that can be taken to test an algorithm:

1. **Control Variable Test** – Is the model/variable a proxy?

The protected class is added as a predictor in a model to account for the predictive effect of the protected class. The results of the model before and after the protected class is added are compared for differences.

2. **Interaction Test** – Is the predictive effect consistent across protected classes?

The protected class is added as an interaction item in a model to produce model indications for the evaluated variable for each protected class. The results of the model are compared across protected classes for consistency.

3. **Nonparametric Matching (Matched Pairs) Test** – Does the inclusion of the variable disproportionately impact otherwise similar risks?

Each policyholder of a protected class is matched with a policyholder with similar risk characteristics not of that protected class. The results of the model for each policyholder are compared for consistency.

4. **Double Lift Chart** – Does the variable improve predictions across protected classes?

The model predictions are compared when the protected class is included and when the protected class is excluded to assess which better predicts the response variable.

No overall compass heading was decided upon during the Summer National Meeting for finding the right means of testing algorithms and machine learning. Although it appears the regulators are still trying to get their bearings, insurers should be prepared to chart a course for testing their algorithms for unfair discrimination.

Authored By



Ann Young Black

Related Practices

[Life, Annuity, and Retirement Solutions](#)
[Financial Services Regulatory](#)

Related Industries

[Life, Annuity, and Retirement Solutions](#)
[Life, Annuity, and Retirement Solutions](#)
[Securities & Investment Companies](#)

©2024 Carlton Fields, P.A. Carlton Fields practices law in California through Carlton Fields, LLP. Carlton Fields publications should not be construed as legal advice on any specific facts or circumstances. The contents are intended for general information and educational purposes only, and should not be relied on as if it were advice about a particular fact situation. The distribution of this publication is not intended to create, and receipt of it does not constitute, an attorney-client relationship with Carlton Fields. This publication may not be quoted or referred to in any other publication or proceeding without the prior written consent of the firm, to be given or withheld at our discretion. To request reprint permission for any of our publications, please use our Contact Us form via the link below. The views set forth herein are the personal views of the author and do not necessarily reflect those of the firm. This site may contain hypertext links to information created and maintained by other entities. Carlton Fields does not control or guarantee the accuracy or completeness of this outside information, nor is the inclusion of a link to be intended as an endorsement of those outside sites.

