Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021
Adopted by the Market Information Systems (D) Task Force, June 16, 2022

"general AI," possess generalized autonomous problem-solving capacities that are comparable to the processes of the human brain, and they are able to adapt to novel situations or information (Macnish et al., 2019).

It is important to emphasize the ways in which AI modeling techniques contrast to the standard scientific model employed in classical or traditional statistics:

Classical Statistics: Method of hypothetical-deductive reasoning in which hypotheses are clearly and narrowly specified *prior* to data testing, often with a prior understanding of the underlying causal nature of the relationships between variables. Purpose: To further causal understanding.

AI: Often employs a type of "data mining" in which a machine pattern-seeking algorithm is released "into the wild" to identify possible correlations between variables that may be predictive of some independent variable. Hypotheses are not specified prior to data analysis, and the algorithm may very well identify correlations that would not have occurred to an analyst and whose causal relationship is constructed post-hoc (to the degree that AI users are concerned with causality at all). Purpose: Predict future outcomes or events.

The difference between these two approaches is not trivial, and significant disagreements about the advantages and disadvantages of AI remain. It is of note that AI did not emerge principally from university statistics departments, but rather from the field of computer science. Many statisticians remain skeptical of the techniques and have offered up a variety of caveats for their use. For example, recently the American Statistical Society (ASA) reacted regulators and insurers regarding the meaning of statistical relationships appearing in predictive models that lack intuitive or, in many cases, even plausible explanations. See Appendix A for further discussion of the ASA statement.

The discussion above is not intended to sway state insurance regulators one way or the other with respect to AI. The purpose is simply to proffer some caveats shared by many statisticians. A final caveat is the AI techniques were developed to analyze very large data sets consisting of millions of records and possibly thousands or tens of thousands of variables. It is said to have an advantage in that algorithms can perform a large volume of analyses across different constellations of variables in a way that would be highly impractical employing traditional (and manual) model building. For small data sets, such as the limited data currently available to market analysts, it is unclear whether the expense associated with developing AI techniques can be justified, nor whether AI is at all superior to traditional model building methods. This is not an unimportant point and is discussed in more depth elsewhere in this recommendation.

Current Status of Market Analysis

Quantitative market analysis relies on just a handful of data sources:

The Complaint Database System (CDS): The NAIC compiles complaints against insurers received by state insurance regulators. Thus, each state has access to a national

regulators.

entities that may merit additional scrutiny and to narrow focus on a much more limited subset of companies out of a larger pool of companies. It therefore primarily prioritizes limited regulatory resources.

State insurance regulators avail themselves of the formal analytical processes adopted by the NAIC. Quantitative or "baseline" analysis identifies entities with anomalous indicators that significantly depart for industry-wide values. A "level 1" analysis may be pursued, in which an analyst devotes additional scrutiny to such things as complaint trends, common reasons complaints are lodged against an insurer, similarities in RIRS actions, etc. If concern still remains (or additional concerns are identified) subsequent to level 1 analysis, a structured level 2 analysis may be performed. A level 2 analysis requires a much greater commitment of time and resources. For example, rather than just manually reviewing complaint data to identify patterns, an analyst may manually review actual complaint documentation to garner a more detailed understanding of the nature of complaints.

As a preliminary to the following discussion, AI/statistical analysis may have two primary functions within the context of the current market analysis structure:

1. More accurately identify companies that merit the additional expenditure of resources necessary to perform the more labor-intensive level 1 and level 2 analyses. Analysis processes that more

implemented, RIRS will also capture much more detailed data related to the specific misconduct that garnered a regulatory response. The RIRS proposal is currently under discussion with the Market Information Systems (D) Task Force, to which Working Group reports.

The remainder of availab ID)

Recommendation 2: In conjunction with recommendation 1 (assess data quality), state insurance regulators should adopt a much more rigorously statistical approach to identify the predictive power of market scoring systems, assess how each variable should be weighted in terms of its unique contribution to productiveness, and drop those that lack analytic utility. In addition, effort should be made to integrate data into a single overall analysis. For example, the MAPT does not incorporate MCAS data, which is typically subject to a separate analysis. The Working Group believes that a "piecemeal" approach is likely less effective than a more integrated approach.

It is noted that the current state of data will likely prove limiting and that such efforts may not make much progress until additional data are

As noted above, AI techniques such as text analysis could potentially expand such exercises and improve the identification of concerning patterns at a deeper level, as well as assess ways to improve the efficiency of other qualitative tasks.

Recommendation 4: Assess ways AI can improve both the efficiency of *qualitative* analysis and facilitate pattern recognition across larger volumes of textual evidence, including most especially complaints, but perhaps other textual sources. For example, the "level 1" analysis formalized in NAIC market system may include a review of the "management discussion and analysis" of the financial annual statement.

V. Longer-Range Planning

As noted above, data mining and AI techniques were developed primarily as tools to analyze large volumes of data. For data past a certain magnitude, including especially those containing many hundreds or even thousands of variables, the traditional hypothetical-deductive cornerstone that is the cornerstone of traditional statistical inference may be ill-suited as well as cost-prohibitive in terms of time and resources.

are identified via AI and found useful, standard statistical models should also be employed to test whether different techniques yield superior predictive power. Additional discussion of caveats is presented in the appendix.

That said, there is much potential of AI in market analysis, *assuming that additional, more granular, data are available.* As noted, such techniques are most suited for large datasets whose very size would make a standard statistical approach impractical just given the sheer number of possible correlations available for testing.

Recommendation 5:

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND *P*-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative Science

March 7, 2016

The American Statistical Association (ASA) has released a "Statement on Statistical Significance and *P*-Values" with six principles underlying the proper use and interpretation of the *p*-value [http://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108#.Vt2XIOaE2MN]. The ASA releases this guidance on *p*-values to improve the conduct and interpretation of quantitative science and inform the growing emphasis on reproducibility of science research. The statement also notes that the increased quantification of scientific research and a proliferation of large, complex data sets has expanded the scope for statistics and the importance of appropriately chosen techniques, properly conducted analyses, and correct interpretation.

Good statistical practice is an essential component of good scientific practice, the statement observes, and such practice "emphasizes principles of good study design and conduct, a varietyof numerical and graphical summaries of data, understanding of the phenomenon under study, interpretation of results in context, complete reporting and proper logical and quantitative understanding of what data summaries mean."

"The *p*-value was never intended to be a substitute for scientific reasoning," said Ron

4. *Proper inference requires full reporting and transparency.*

5. *A p-value, or statistical significance, does not measure the size of an effect or theimportance of a result.*

6. *By itself, a p-value does not provide a good measure of evidence regarding a model orhypothesis.*

The statement has short paragraphs elaborating on each principle.

In light of misuses of and misconceptions concerning *p*-values, the statement notes that statisticians often supplement or even replace *p*-values with other approaches. These include methods "that emphasize estimation over testing such as confidence, credibility, or prediction intervals; Bayesian methods; alternative measures of evidence such as likelihood ratios or Bayes factors; and other approaches such as decision-theoretic modeling and false discovery rates."

"The contents of the ASA statement and the reasoning behind it are not new—statisticians and other scientists have been writing on the topic for decades," Utts said. "But this is the first time that the community of statisticians, as represented by the ASA Board of Directors, has issued a statement to address these issues."

"The issues involved in statistical inference are difficult because inference itself is challenging," Wasserstein said. He noted that more than a dozen discussion papers are being published in the ASA journal *The American Statistician* with the statement to provide more perspective on this broad and complex topic. "What we hope will follow is a broad discussion across the scientific community that leads to a more nuanced approach to interpreting, communicating, and using the results of statistical methods in research."

About the American Statistical Association

The ASA is the world's largest community of statisticians and the oldest continuously operating professional science society in the United States. Its members serve in industry, government and academia in more than 90 countries, advancing research and promoting sound statistical practice to inform public policy and improve human welfare. For additional information, please visit the ASA website at www.amstat.org.

For more information:

Ron
Wasserstein

Citations

Macnish, K., Ryan, M. & Stahl, B. (2019). Understanding ethics and human rights in smart information systems: A multi-case study approach. *The Orbit Journal*, *2*(2), 1 34.

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://science.sciencemag.org/content/349/6251/aac4716

Wasserstein, R.L., & Lazar, N.A. (2016).