

6-1-2019

## Models for Predicting Business Bankruptcies and Their Application to Banking and Financial Regulation

James Ming Chen

Follow this and additional works at: <https://elibrary.law.psu.edu/pslr>



Part of the [Law Commons](#)

---

### Recommended Citation

Chen, James Ming (2019) "Models for Predicting Business Bankruptcies and Their Application to Banking and Financial Regulation," *Penn State Law Review*: Vol. 123 : Iss. 3 , Article 5.

Available at: <https://elibrary.law.psu.edu/pslr/vol123/iss3/5>

This Article is brought to you for free and open access by Penn State Law eLibrary. It has been accepted for inclusion in Penn State Law Review by an authorized editor of Penn State Law eLibrary. For more information, please contact [ram6023@psu.edu](mailto:ram6023@psu.edu).

---

---

# Models for Predicting Business Bankruptcies and Their Application to Banking and Financial Regulation

James Ming Chen\*

## ABSTRACT

Models for predicting business bankruptcies have evolved rapidly as machine learning has displaced traditional statistical methodologies. Three distinct techniques for approaching the classification problem in bankruptcy prediction have emerged: single classification, hybrid classifiers, and classifier ensembles. Methodological heterogeneity through the introduction and integration of machine-learning algorithms (especially support vector machines, decision trees, and genetic algorithms) has improved the accuracy of bankruptcy prediction models. Improved natural language processing has enabled machine learning to combine textual analysis of corporate filings with evaluation of numerical data. Greater accuracy promotes external processes of banks by minimizing credit risk and facilitating regulatory compliance.

*Keywords:* bankruptcy prediction, classifier ensembles, hybrid classifiers, support vector machine, genetic algorithm, credit risk

---

\*Justin Smith Morrill Chair in Law, Michigan State University; candidate for the degree of Master of Science in Data Science, Northwestern University; visiting scholar, School of Economics and Business of the University of Zagreb (Ekonomski Fakultet, Sveučilište u Zagrebu). This paper was presented on March 22, 2019, at K&L Gates in Pittsburgh. Klaira Bielonko, Candice Bradley, Najda Dupanović, and Christopher C. French provided helpful comments. Special thanks to Heather Elaine Worland Chen.

### Table of Contents

I.	INTRODUCTION .....	736
II.	METHODS FOR PREDICTING BANKRUPTCIES .....	737
	A. Single classifiers. ....	738
	B. Hybrid classifiers. ....	741
	C. Classifier ensembles.....	742
III.	TEXTUAL ANALYSIS OF FINANCIAL DISCLOSURES .....	743
IV.	CONTRIBUTIONS TO BANK MANAGEMENT.....	747
	A. Credit evaluation.....	747
	B. Regulatory compliance. ....	749
V.	CONCLUSION .....	751

#### I. INTRODUCTION

Business failures radiate negative consequences far beyond the immediate confines of the fallen firm.<sup>1</sup> A business bankruptcy devastates not only shareholders and employees, but also customers, suppliers, and creditors.<sup>2</sup> “Credit risk assessment” is the “first task” that banks and other creditors should undertake when “considering a new transaction or business relationship with a counterparty.”<sup>3</sup> Credit risk assessment requires a lender or investor to estimate “the adverse probability” that a counterparty “may not fulfill its future obligations,” in order to decide whether to proceed with the transaction or relationship and, if so, under what conditions.<sup>4</sup>

The vast, diverse literature on this form of predictive analysis uses a profusion of synonyms, including “bankruptcy, insolvency, liquidation, financial distress, . . . dissolution,” and “corporate failure.”<sup>5</sup> Rarely do sources define “bankruptcy” specifically by reference to filing for protection under chapter 7 (liquidation) or chapter 11 (reorganization) of the Bankruptcy Code.<sup>6</sup> As recently as 2007, bankruptcy prediction models

1. See generally Ben S. Bernanke, *Bankruptcy, Liquidity, and Recession*, 71 AM. ECON. REV. 2, May 1981, at 155.

2. Wei-Wen Wu, *Beyond Business Failure Prediction*, 37 EXPERT SYSTEMS WITH APPLICATIONS 2371, 2371 (2010).

3. Jiří WITZANY, CREDIT RISK MANAGEMENT: PRICING, MEASUREMENT, AND MODELING 2 (2017).

4. *Id.*

5. Kingsley Opoku Appiah, Amon Chizema & Joseph Arthur, *Predicting Corporate Failure: A Systematic Literature Review of Methodological Issues*, 57 INT’L J.L. & MGMT. 461, 466 (2015).

6. Feng Mai, Shaonan Tian, Chihoon Lee & Ling Ma, *Deep Learning Models for Bankruptcy Prediction Using Textual Disclosures*, 274 EUR. J. OPERATIONAL RES. 743, 746 (2019).

relied most heavily on neural network methodologies<sup>7</sup> as the leading alternative to traditional statistical models.<sup>8</sup>

This paper reviews more recent methodologies for predicting business bankruptcy. It emphasizes “intelligent techniques,” which forgo traditional statistical assumptions about the data and rely instead on machine learning, artificial intelligence, deep learning, and operations research to extract insights that might otherwise elude conventional linear modeling.<sup>9</sup>

Part II identifies three distinct techniques for addressing bankruptcy prediction as a classification problem: single classification techniques, hybrid classification, and classifier ensembles. Part III identifies the potential application of machine learning to textual as well as financial data. Part IV discusses the contribution of improved bankruptcy prediction to two external processes in bank management: credit risk assessment and regulatory compliance. Part V concludes that legal applications will demand that these models’ predictive power match or exceed that of conventional statistical tools.

## II. METHODS FOR PREDICTING BANKRUPTCIES

“Bankruptcy prediction can be regarded as a typical classification problem” in machine learning.<sup>10</sup> Classification in machine learning,<sup>11</sup> or at least in *supervised* machine learning,<sup>12</sup> affects the accuracy of a model that seeks to distinguish between firms bound for bankruptcy and those that are not. Ideally, models can predict bankruptcy not just one, but several years

---

7. M. Adnan Aziz & Humayon A. Dar, *Predicting Corporate Bankruptcy: Where We Stand?*, 6 CORP. GOVERNANCE 18, 19–22 (2006); see also P. Ravi Kumar & V. Ravi, *Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques*, 180 EUR. J. OPERATIONAL RES. 1, 4–17 (2007); Muriel Perez, *Artificial Neural Networks and Bankruptcy Forecasting: A State of the Art*, 15 NEURAL COMPUTING & APPLICATIONS 154, 156–62 (2006).

8. See Sofie Balcaen & Hubert Ooghe, *35 Years of Studies on Business Failure: An Overview of the Classical Statistical Methodologies and their Related Problems*, 38 BRIT. ACCT. REV. 63, 65–71 (2006); see also Gordon V. Karels & Arun J. Prakash, *Multivariate Normality and Forecasting of Business Bankruptcy*, 14 J. BUS. FIN. & ACCT. 573, 579–89 (1987). See generally Michal Tkáč & Robert Verner, *Artificial Neural Networks in Business: Two Decades of Research*, 38 APPLIED SOFT COMPUTING 788, 789–94 (2016).

9. Mai et al., *supra* note 6, at 744.

10. Efstathios Kirkos, *Assessing Methodologies for Intelligent Bankruptcy Prediction*, 43 ARTIFICIAL INTELLIGENCE REV. 83, 90 (2015).

11. CHRISTOPHER M. BISHOP, *PATTERN RECOGNITION AND MACHINE LEARNING* 1–4 (Michael Jordan et al. eds., 2006); see also ANDREW R. WEBB & KEITH D. COPSEY, *STATISTICAL PATTERN RECOGNITION* 291–92 (3d ed. 2011).

12. RICHARD O. DUDA, PETER E. HART & DAVID G. STORK, *PATTERN CLASSIFICATION* 84–85 (2d ed. 2001).

in advance.<sup>13</sup> Structuring data as time-varying panels should improve predictive performance relative to a static model using data from a single year.<sup>14</sup> Filtering out noise from financial datasets assumes pronounced importance since the absence of “generally agreed financial ratios for bankruptcy prediction and credit scoring” requires close examination of “collected variables” for “importance and explanatory power.”<sup>15</sup>

The data science literature has identified three distinct categories of classification techniques: single classifiers, hybrid classifiers, and classifier ensembles.<sup>16</sup> Methods for predicting financial distress form a progression from traditional statistical and semiparametric models, to individual machine-learning algorithms such as artificial neural networks, support vector machines, decision trees, and case-based reasoning, and ultimately to hybrid learning and multiple classifier systems.<sup>17</sup>

The three-way taxonomy of single, hybrid, and ensemble classifiers covers two additional categories: dynamic modeling and decision implementation.<sup>18</sup> Dynamic modeling seeks to update quantitative methods over time to avoid obsolescence.<sup>19</sup> Decision implementation integrates qualitative evaluations of financial distress from multiple human experts.<sup>20</sup> As Part III of this article will demonstrate, these tools work in tandem with natural language processing of financial disclosures.

#### A. *Single classifiers.*

Single classification techniques, which may involve statistical methods or supervised or unsupervised machine learning, rely on a single

---

13. Mária Režňáková & Michal Karas, *Bankruptcy Prediction Models: Can the Prediction Power of the Models be Improved by Using Dynamic Indicators?*, 12 *PROCEDIA ECON. & FIN.* 565, 566 (2014).

14. Tyler Shumway, *Forecasting Bankruptcy More Accurately: A Simple Hazard Model*, 74 *J. BUS.* 101, 123 (2001).

15. Deron Liang, Chih-Fong Tsai & Hsin-Ting Wu, *The Effect of Feature Selection on Financial Distress Prediction*, 73 *KNOWLEDGE-BASED SYSTEMS* 289, 289 (2015) [hereinafter Liang et al., *Effect of Feature Selection*]; see also Wei-Chao Lin, Yu-Hsin Lu & Chih-Fong Tsai, *Feature Selection in Single and Ensemble Learning-Based Bankruptcy Prediction Models*, 36 *EXPERT SYSTEMS* e12335, Aug. 16, 2018, at 1, 5.

16. Wei-Yang Lin, Ya-Han Hu & Chih-Fong Tsai, *Machine Learning in Financial Crisis Prediction: A Survey*, 42 *IEEE TRANSACTIONS ON SYSTEMS MAN & CYBERNETICS* 421, 422–26 (2012) [hereinafter Lin et al., *Financial Crisis Prediction*].

17. Ning Chen, Bernardete Ribeiro & An Chen, *Financial Credit Risk Assessment: A Recent Review*, 45 *ARTIFICIAL INTELLIGENCE REV.* 1, 4–10 (2016) [hereinafter Chen et al., *Financial Credit Risk*].

18. Jie Sun, Hui Li, Qing-Hua Huang & Kai-Yu He, *Predicting Financial Distress and Corporate Failure: A review from State-of-the-Art Definitions, Modeling, Sampling, and Featuring Approaches*, 57 *KNOWLEDGE-BASED SYSTEMS* 41, 43–48 (2014).

19. *Id.* at 48–49.

20. *Id.* at 49.

classifier.<sup>21</sup> Artificial neural networks, support vector machines, and decision trees are popular methods of supervised machine learning,<sup>22</sup> as is data envelopment analysis.<sup>23</sup>

Support vector machine models for bankruptcy and financial distress prediction measure the mathematical distance among variables that may be related to potential or even incipient business failure.<sup>24</sup> By overcoming noise and bias, the support vector machine method improves accuracy in financial predictions.<sup>25</sup> The method applies not only to bankruptcy prediction,<sup>26</sup> but also to broader questions of credit risk evaluation<sup>27</sup> and to failures of banks themselves.<sup>28</sup>

A rival machine learning algorithm, extreme learning machine,<sup>29</sup> and its nonrandom generalization, kernel extreme learning machine,<sup>30</sup> have informed bankruptcy prediction models.<sup>31</sup> These models have also

---

21. Lin et al., *Financial Crisis Prediction*, *supra* note 16, at 422.

22. *Id.* at 422–23.

23. See Anja Cielen, Ludo Peeters & Koen Vanhoof, *Bankruptcy Prediction Using a Data Envelopment Analysis*, 154 EUR. J. OPERATIONAL RES. 526, 527–32 (2004); Zhiyong Li, Jonathan Crook & Galina Andreeva, *Chinese Companies Distress Prediction: An Application of Data Envelopment Analysis*, 65 J. OPERATIONAL RES. SOC'Y 466 (2014); Maria Psilaki, Ioannis E. Tsolas & Dimitris Margaritis, *Evaluation of Credit Risk Based on Firm Performance*, 201 EUR. J. OPERATIONAL RES. 873, 874 (2010).

24. Ligang Zhou, Kin Keung Lai & Jerome Yen, *Bankruptcy Prediction Using SVM Models with a New Approach to Combine Features Selection and Parameter Optimisation*, 45 INT'L J. SYSTEMS SCI. 241, 245–46 (2014).

25. See Flavio Barboza, Herbert Kimura & Edward Altman, *Machine Learning Models and Bankruptcy Prediction*, 83 EXPERT SYSTEMS WITH APPLICATIONS 405, 407 (2017).

26. See Jae H. Min & Young-Chan Lee, *Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters*, 28 EXPERT SYSTEMS WITH APPLICATIONS 603, 604 (2005).

27. See Lean Yu, Wuyi Yue, Shouyang Wang & K.K. Lai, *Support Vector Machine Based Multiagent Ensemble Learning for Credit Risk Evaluation*, 37 EXPERT SYSTEMS WITH APPLICATIONS 1351, 1352 (2010).

28. See Shouwei Li, Mingliang Wang & Jianmin He, *Prediction of Banking Systemic Risk Based on Support Vector Machine*, MATHEMATICAL PROBLEMS IN ENGINEERING, 2013, Article ID 136030 at 1.

29. See Guang-Bin Huang, Qin-Yu Zhu & Chee-Kheong Siew, *Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks*, 2 IEEE INT'L JOINT CONF. ON NEURAL NETWORKS 985, 985 (2004); Guang-Bin Huang, Qin-Yu Zhu & Chee-Kheong Siew, *Extreme Learning Machine: Theory and Applications*, 70 NEUROCOMPUTING 489, 490 (2006).

30. See Guang-Bin Huang, Hongming Zhou, Xiaojian Ding & Rui Zhang, *Extreme Learning Machine for Regression and Multiclass Classification*, 42 IEEE TRANSACTIONS ON SYSTEMS MAN & CYBERNETICS 513, 514 (2012).

31. See Qi Yu, Yoan Miche, Eric Severin & Amaury Lendasse, *Bankruptcy Prediction Using Extreme Learning Machine and Financial Expertise*, 128 NEUROCOMPUTING 296, 296–97 (2014) [hereinafter Yu et al., *Bankruptcy Prediction*]; Dong Zhao, Chunyu Huang, Yan Wei, Fanhua Yu, Mingjing Wang & Huiling Chen, *An Effective Computational Model for Bankruptcy Prediction Using Kernel Extreme Learning Machine Approach*, 49 COMPUTATIONAL ECON. 325, 326 (2017).

furnished broader descriptions of the corporate life cycle.<sup>32</sup> The *k*-nearest neighbor method epitomizes *unsupervised* machine learning methodology.<sup>33</sup>

Traditional analytical methods, such as decision trees<sup>34</sup> and artificial neural networks,<sup>35</sup> survive in contemporary literature on bankruptcy prediction.<sup>36</sup> An outgrowth of the classification and regression tree method (CART),<sup>37</sup> random forest techniques repeatedly generate classification results from portions of available data and then randomly select a subset of characteristics from each node of each decision tree for further testing.<sup>38</sup>

Inasmuch as cerebellar model neural networks “can be thought of as a learning mechanism imitating the cerebellum of a human being,”<sup>39</sup> one may reclassify the literature on artificial neural networks, one of the oldest and most popular methods of bankruptcy prediction,<sup>40</sup> as a branch of genetic and evolutionary algorithms.<sup>41</sup> Exotic variations on this biologically inspired theme include particle swarm optimization<sup>42</sup> and ant

---

32. See Sin-Jin Lin, Chingho Chang & Ming-Fu Hsu, *Multiple Extreme Learning Machines for a Two-class Imbalance Corporate Life Cycle Prediction*, 39 KNOWLEDGE-BASED SYSTEMS 214, 214–15 (2013).

33. See Hui-Ling Chen, Bo Yang, Gang Wang, Jie Liu, Xin Xu, Su-Jing Wang & Da-You Liu, *A Novel Bankruptcy Prediction Model Based on an Adaptive Fuzzy k-nearest Neighbor Method*, 24 KNOWLEDGE-BASED SYSTEMS 1348, 1349 (2011).

34. See Tae Kyung Sung, Namsik Chang & Gunhee Lee, *Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy Prediction*, 16 J. MGMT. INFO. SYSTEMS 63, 67–68 (1999).

35. See Guoqiang Zhang, Michael Y. Hu, B. Eddy Patuwo & Daniel C. Indro, *Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-validation Analysis*, 116 EUR. J. OPERATIONAL RES. 16, 16–17 (1999).

36. See, e.g., Sihem Khemakhem & Younes Boujelbene, *Predicting Credit Risk on the Basis of Financial and Non-financial Variables and Data Mining*, 17 REV. ACCT. & FIN. 316, 317 (2018); Felix J. López Iturriaga & Iván Pastor Sanz, *Bankruptcy Visualization and Prediction Using Neural Networks: A Study of U.S. Commercial Banks*, 42 EXPERT SYSTEMS WITH APPLICATIONS 2857, 2857 (2015); Alireza Mehrazin, Mohammad Taghipour, Omid Froutan, Bashir Ghabdian & Hamid Soleimani, *Radial Basis Function in Artificial Neural Network for Prediction of Bankruptcy*, 6 INT’L BUS. RES. 121, 122 (2013).

37. See Hui Li, Jie Sun & Jian Wu, *Predicting Business Failure Using Classification and Regression Tree: An Empirical Comparison with Popular Classical Statistical Methods and Top Classification Mining Methods*, 37 EXPERT SYSTEMS WITH APPLICATIONS 5895, 5895–96 (2010).

38. See Ching-Chiang Yeh, *Going-concern Prediction Using Hybrid Random Forests and Rough Set Approach*, 254 INFO. SCI. 98, 101 (2014).

39. Chang-Chih Chung, Tsung-Shih Chen, Lee-Hsuan Lin, Yu-Chen Lin & Chih-Min Lin, *Bankruptcy Prediction Using Cerebellar Model Neural Networks*, 18 INT’L J. FUZZY SYSTEMS 160, 161 (2016) (citing Chih-Min Lin et al., *Adaptive Filter Design Using Recurrent Cerebellar Model Articulation Controller*, 21 IEEE TRANSACTIONS ON NEURAL NETWORKS 1149 (2010)).

40. See Lin et al., *Financial Crisis Prediction*, *supra* note 16, at 421.

41. See Jae H. Min & Chulwoo Jeong, *A Binary Classification Method for Bankruptcy Prediction*, 36 EXPERT SYSTEMS WITH APPLICATIONS 5256, 5256 (2009).

42. See Mu-Yen Chen, *Bankruptcy Prediction in Firms with Statistical and Intelligent Techniques and a Comparison of Evolutionary Computation Approaches*, 62 COMPUTERS

colony optimization.<sup>43</sup> Some sources have concluded that genetic algorithms provide less accuracy than other predictive methods.<sup>44</sup> At a minimum, however, the broader class of evolutionary algorithms can supplement other statistical and machine-learning techniques.<sup>45</sup>

Relative to older techniques such as logistic regression, inductive learning algorithms, and genetic algorithms, newer methods based on “[s]upport vector machine, neural network with dropout, and autoencoder” offer improvements such as “control for overfitting, the improved probability of finding the global maxima, and the ability to handle large feature spaces.”<sup>46</sup> These are the leading virtues of hybrid classifiers<sup>47</sup> and classifier ensembles.<sup>48</sup> This article will discuss each of these strategies.

### B. Hybrid classifiers.

Methodological diversity strengthens hybrid classifier and classifier ensemble methods for predicting bankruptcy and financial distress. Hybrid learning methods deploy multiple “heterogeneous techniques,” but usually apply “only one . . . for the final prediction.”<sup>49</sup> Quite often a “hybrid paradigm” combining two methodologies will “use the first technique for feature selection and the second for classification.”<sup>50</sup>

---

& MATHEMATICS WITH APPLICATIONS 4514, 4516–17 (2011). See generally Shafiq Alam, *Research on Particle Swarm Optimization Based Clustering: A Systematic Review of Literature and Techniques*, 17 SWARM & EVOLUTIONARY COMPUTATION 1 (2014).

43. See David Martens, T. Van Gestel, M. De Backer, R. Haesen, J. Vanthienen & B. Baesens, *Credit Rating Prediction Using Ant Colony Optimization*, 61 J. OPERATIONAL RES. SOC'Y 561, 562 (2010); J. Uthayakumar, Noura Metawa, K. Shankar & S.K. Lakshmanaprabu, *Financial Crisis Prediction Model Using Ant Colony Optimization*, INT'L J. INFO. MGMT. (forthcoming 2019), <https://bit.ly/2yJcCDi>.

44. See, e.g., F. Mokhatab Rafiei, S.M. Manzari & S. Bostanian, *Financial Health Prediction Models Using Artificial Neural Networks, Genetic Algorithm and Multivariate Discriminant Analysis: Iranian Evidence*, 38 EXPERT SYSTEMS WITH APPLICATIONS 10216, 10216–17 (2011); Sun et al., *supra* note 18, at 44.

45. See generally Sacha Gobeyn, *Evolutionary Algorithms for Species Distribution Modelling: A Review in the Context of Machine Learning*, 392 ECOLOGICAL MODELLING 179 (2019).

46. Nanxi Wang, *Bankruptcy Prediction Using Machine Learning*, 7 J. MATHEMATICAL FIN. 908, 915 (2017).

47. See generally Chun-Ling Chuang, *Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction*, 236 INFO. SCI. 174 (2013).

48. See generally Zhi Xiao et al., *The prediction for listed companies' financial distress by using multiple prediction methods with rough set and Dempster-Shafer evidence theory*, 26 KNOWLEDGE-BASED SYSTEMS 196 (2012).

49. Chen et al., *Financial Credit Risk*, *supra* note 17, at 7.

50. Sun et al., *supra* note 18, at 45. See generally Philippe du Jardin, *A Two-Stage Classification Technique for Bankruptcy Prediction*, 254 EUR. J. OPERATIONAL RES. 236 (2016).



Some hybrid models enhance conventional logistic regression with decision trees.<sup>51</sup> Other hybrid models have bridged traditional statistics and machine learning by using “static and trend analysis models to construct and train a back-propagation neural network.”<sup>52</sup> Support vector machines have also become popular components in hybrid models.<sup>53</sup>

True to their biological inspiration, hybrids combining support vector machines with genetic algorithms have emerged. Genetic algorithms can improve the performance of support vector machines in two respects: feature subset selection and parameter optimization.<sup>54</sup> By optimizing support vector machine parameters’ ability to generalize, real-valued genetic algorithms can improve predictive accuracy relative to a support vector machine operating on its own.<sup>55</sup>

### C. Classifier ensembles.

Multiple classifier systems or classifier ensembles, the final broad category of bankruptcy prediction models, combine diverse, elementary classifiers—often with highly accurate predictive power in their own right, but “low intercorrelation so as to ensure [their] effective fusion”—in order to train and aggregate those methods.<sup>56</sup> Classifier ensembles outperform models relying on a single method.<sup>57</sup> By combining techniques as diverse

---

51. See Mu-Yen Chen, *Predicting Corporate Financial Distress Based on Integration of Decision Tree Classification and Logistic Regression*, 38 EXPERT SYSTEMS WITH APPLICATIONS 11261, 11262 (2011).

52. Shi-Ming Huang, Chih-Fong Tsai, David C. Yen & Yin-Lin Cheng, *A Hybrid Financial Analysis Model for Business Failure Prediction*, 35 EXPERT SYSTEMS WITH APPLICATIONS 1034, 1034 (2008).

53. See Fengyi Lin, Ching-Chiang Yeh & Meng-Yuan Lee, *The Use of Hybrid Manifold Learning and Support Vector Machines in the Prediction of Business Failure*, 24 KNOWLEDGE-BASED SYSTEMS 95, 95 (2011); Ching-Chiang Yeh, Der-Jang Chi & Ming-Fu Hsu, *A Hybrid Approach of DEA, Rough Set and Support Vector Machines for Business Failure Prediction*, 37 EXPERT SYSTEMS WITH APPLICATIONS 1535, 1536 (2010).

54. See Sung-Hwan Min, Jumin Lee & Ingoo Han, *Hybrid Genetic Algorithms and Support Vector Machines for Bankruptcy Prediction*, 31 EXPERT SYSTEMS WITH APPLICATIONS 652, 653–54 (2006).

55. See Chih-Hung Wu, Gwo-Hshiung Tzeng, Yeong-Jia Goo & Wen-Chang Fang, *A Real-Valued Genetic Algorithm to Optimize the Parameters of Support Vector Machine for Predicting Bankruptcy*, 32 EXPERT SYSTEMS WITH APPLICATIONS 1, 397–98 (2007).

56. Chen et al., *supra* note 17, at 8.

57. See *id.* at 9; Joaquín Abellán & Carlos J. Mantas, *Improving Experimental Studies About Ensembles of Classifiers for Bankruptcy*, 41 EXPERT SYSTEMS WITH APPLICATIONS 3825, 3825–26 (2014); Despina Deligianni & Sotiris Kotsiantis, *Forecasting Corporate Bankruptcy with an Ensemble of Classifiers*, in ARTIFICIAL INTELLIGENCE: THEORIES AND APPLICATIONS 65, 66 (Ilias Maglogiannis et al. eds., Springer LNAI Ser. No. 7297, 2012); Elena Fedorova, Evgenii Gilenko & Sergey Dovzhenko, *Bankruptcy Prediction for Russian Companies: Application of Combined Classifiers*, 40 EXPERT SYSTEMS WITH APPLICATIONS 7285, 7286 (2013); Myoung-Jong

as logistic regression, *k*-nearest neighbor, artificial neural networks, support vector machines, and decision tree algorithms, even “weak classifiers yield[ing] only moderate performance individually” have the “potential to increase the overall performance of an ensemble.”<sup>58</sup>

Classifier ensembles routinely improve the predictive accuracy of bankruptcy prediction models. Combining self-organizing maps with classifier ensembles, especially those including a multilayer-perceptron neural network, will yield “the best classifier” according to “the weighted voting approach.”<sup>59</sup> An even more stringent unanimous voting method requires all classifiers in an ensemble to agree.<sup>60</sup>

### III. TEXTUAL ANALYSIS OF FINANCIAL DISCLOSURES

Dramatic improvements in natural language processing<sup>61</sup> have made it possible to extend machine-learning models for bankruptcy prediction beyond traditional “market-based and accounting-based variables.”<sup>62</sup> Unlike “numeric data” such as stock market returns and volatility or profitability and liability ratios, which often appear “in a well-structured format,” textual evidence of firms’ health takes the “form of unstructured, qualitative data” such as regulatory data,<sup>63</sup> articles from such journalistic sources as the *Financial Times*,<sup>64</sup> the *New York Times*,<sup>65</sup> or the *Wall Street Journal*’s “Abreast of the Market” column,<sup>66</sup> or even tweets on Twitter.<sup>67</sup>

---

Kim & Dae-Ki Kang, *Ensemble with Neural Networks for Bankruptcy Prediction*, 37 EXPERT SYSTEMS WITH APPLICATIONS 3373, 3376 (2010).

58. Ning Chen & Bernardete Ribeiro, *A Consensus Approach for Combining Multiple Classifiers in Cost-Sensitive Bankruptcy Prediction*, in ADAPTIVE AND NATURAL COMPUTING ALGORITHMS 266, 275 (Marco Tomassini et al. eds., Springer LNCS Ser. No. 7824, 2013).

59. Chih-Fong Tsai, *Combining Cluster Analysis with Classifier Ensembles to Predict Financial Distress*, 16 INFO. FUSION 46, 55 (2014).

60. See Deron Liang, Chih-Fong Tsai, An-Jie Dai & William Eberle, *A Novel Classifier Ensemble Approach for Financial Distress Prediction*, 54 KNOWLEDGE & INFO. SYSTEMS 437, 438–39 (2018) [hereinafter Liang et al., *Novel Classifier Ensemble*].

61. See SHOLOM M. WEISS, NITIN INDURKHIA & TONG ZHANG, FUNDAMENTALS OF PREDICTIVE TEXT MINING 48 (David Gries & Fred B. Schneider eds., 2010). See generally DANIEL JURAFSKY & JAMES H. MARTIN, SPEECH AND LANGUAGE PROCESSING (2d ed. 2009); NITIN INDURKHIA & FRED J. DAMERAU, HANDBOOK OF NATURAL LANGUAGE PROCESSING (2d ed. 2010).

62. Mai et al., *supra* note 6, at 733.

63. *Id.*

64. See Merve Alanyali, Helen Susannah Moat & Tobias Preis, *Quantifying the Relationship Between Financial News and the Stock Market*, 3 SCI. REP. 3578, Dec. 20, 2013, at 1, 1.

65. See Diego Garcia, *Sentiment During Recessions*, 68 J. FIN. 1267, 1267 (2013).

66. See Paul C. Tetlock, *Giving Content to Investor Sentiment: The Role of Media*, 62 J. FIN. 1139, 1139 (2007).

67. See Johan Bollen, Huina Mao & Xiao-Jun Zeng, *Twitter Mood Predicts the Stock Market*, 2 J. COMPUTATIONAL SCI. 1, 1 (2011).

Machine learning and artificial intelligence overcome the traditionally formidable challenges of finding and quantifying textual data for the purpose of predicting financial distress.<sup>68</sup> Not long ago, natural language processing of textual data relied “on simple text summarization techniques such as word count, sentiment, and readability.”<sup>69</sup> Natural language processing is no longer limited to similar analytical methods treating textual sources as “bags-of-words.”<sup>70</sup> Advances in deep learning<sup>71</sup> and computational linguistics<sup>72</sup> have extended the predictive accuracy and reach of natural language processing.<sup>73</sup>

Just as hybrid models and classifier ensembles “capture more variations in the decision space and result in more stable and accurate predictions,” an expansion from purely numerical data sources to a richer informational environment that includes text should “improve[] prediction accuracy over . . . models using a single type of input.”<sup>74</sup> For instance, efforts to harness the predictive power of corporate governance criteria, such as ownership structure, board structure, and the composition of managerial leadership, would benefit from a synergistic evaluation of qualitative alongside quantitative data.<sup>75</sup>

---

68. See Mark Lang & Lorien Stice-Lawrence, *Textual Analysis and International Financial Reporting: Large Sample Evidence*, 60 J. ACCT. & ECON. 110, 110 (2015).

69. Mai et al., *supra* note 6, at 744. See generally Theresa Wilson, Janyce Wiebe & Paul Hoffmann, *Recognizing contextual polarity in phrase-level sentiment analysis*, in PROCEEDINGS OF THE HUMAN LANGUAGE TECHNOLOGY CONFERENCE AND CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING 347 (Ass’n for Computational Linguistics, Vancouver, B.C., 2005).

70. Youngjoong Ko, *A New Term-Weighting Scheme for Text Classification Using the Odds of Positive and Negative Class Probabilities*, 66 J. ASS’N FOR INFO. SCI. & TECH. 2253, 2259 (2015); Kilian Weinberger, Anirban Dasgupta, John Langford, Alex Smola & Josh Attenberg, *Feature Hashing for Large Scale Multitask Learning*, in PROCEEDINGS OF THE 26TH ANNUAL INTERNATIONAL CONFERENCE ON MACHINE LEARNING 1113, 1115 (Montreal, Que., 2009).

71. See Yann LeCun, Yoshua Bengio & Geoffrey Hinton, *Deep Learning*, 521 NATURE 436, 436 (2015).

72. See ROLAND HAUSSER, FOUNDATIONS OF COMPUTATIONAL LINGUISTICS 31 (2d ed. 2001).

73. See Richard Socher, Jeffrey Pennington, Eric Huang, Andrew Y. Ng & Christopher D. Manning, *Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions*, in PROCEEDINGS OF THE CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING 158 (2011).

74. Mai et al., *supra* note 6, at 744.

75. Cf. Deron Liang, Chia-Chi Lu, Chih-Fong Tsai & Guan-An Shih, *Financial Ratios and Corporate Governance Indicators in Bankruptcy Prediction: A Comprehensive Study*, 252 EUR. J. OPERATIONAL RES. 561, 564 (2016) [hereinafter Liang et al., *Financial Ratios*] (finding that a predictive model integrating quantitative financial ratios with qualitative corporate governance indicators was more effective than one using only financial ratios).

Of special interest to the law is the prospect of mining mandatory corporate filings for textual data that may reveal potential credit risks.<sup>76</sup> The managerial discussion and analysis (MD&A) section of 10-K forms filed by publicly listed corporations has attracted intense interest as a source of potential insights into these firms' financial vulnerabilities.<sup>77</sup>

The rules of the Securities and Exchange Commission (SEC) require corporations to provide "management's discussion and analysis of financial condition and results of operations" in five specific areas—liquidity, capital resources, results of operations, off-balance sheet arrangements, and tabular disclosure of contractual obligations—as well as "such other information that the registrant believes to be necessary to an understanding of its financial condition, changes in financial condition and results of operations."<sup>78</sup> The current MD&A requirement traces its origins to a 1980 SEC release,<sup>79</sup> and it has been clarified through occasional documents offering interpretation<sup>80</sup> and guidance.<sup>81</sup>

The SEC describes "MD&A [as] a narrative explanation of the financial statements and other statistical data that the registrant believes will enhance a readers' [*sic*] understanding of its financial condition, changes in financial condition and results of operation."<sup>82</sup> The SEC recommends "a narrative explanation of . . . financial statements, because

---

76. See Zahn Bozanic & Maya Thevenot, *Qualitative Disclosure and Changes in Sell-Side Financial Analysts' Information Environment*, 32 CONTEMP. ACCT. RES. 1595, 1598 (2015); John L. Campbell, Hsinchun Chen, Dan S. Dhaliwal, Hsin-min Lu & Logan B. Steele, *The Information Content of Mandatory Risk Factor Disclosures in Corporate Filings*, 19 REV. ACCT. STUD. 396, 398 (2014); Tim Loughran & Bill McDonald, *When Is a Liability Not a Liability? Textual Analysis, Dictionaries*, 66 J. FIN. 35, 35–36 (2011).

77. See Mark Cecchini, Haldun Aytug, Gary J. Koehler & Praveen Pathak, *Making Words Work: Using Financial Text as a Predictor of Financial Events*, 50 DECISION SUPPORT SYSTEMS 164, 164–65 (2010); Cathy J. Cole & Christopher L. Jones, *Management Discussion and Analysis: A Review and Implications for Future Research*, 24 J. ACCT. LITERATURE 135, 135–74 (2005); Loughran & McDonald, *supra* note 76, at 35; William J. Mayew, Mani Sethuraman & Mohan Venkatachalam, *MD&A Disclosure and the Firm's Ability to Continue as a Going Concern*, 90 ACCT. REV. 1621, 1621–51 (2015).

78. 17 C.F.R. § 229.303(a) (2018).

79. Amendment to Annual Report Form, Securities Act Release No. 6231, Exchange Act Release No. 17114, 45 Fed. Reg. 63630 (1980).

80. Management's Discussion and Analysis of Financial Condition and Results of Operations, Securities Act Release No. 6835, Exchange Act Release No. 26831, Investment Company Act Release No. 16961, 54 Fed. Reg. 22427 (1989).

81. Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations, Securities Act Release No. 8350, Exchange Act Release No. 48960, 68 Fed. Reg. 75056 (2003); Commission Guidance on Presentation of Liquidity and Capital Resources Disclosures in Management's Discussion and Analysis, Securities Act Release No. 9144, Exchange Act Release No. 62934, 75 Fed. Reg. 59894 (2010).

82. U.S. SEC. & EXCH. COMM'N, DIV. OF CORP. FIN., FINANCIAL REPORTING MANUAL § 9110.1, at 296 (2017).

a numerical presentation,” even with “brief accompanying footnotes,” “may be insufficient for an investor to judge the quality of earnings and the likelihood that past performance is indicative of future performance.”<sup>83</sup>

MD&A specifically invites “the investor . . . to look at the company *through the eyes of management* by providing both a short and long-term analysis of the business of the company.”<sup>84</sup> Management enjoys wide berth “to identify and address those key variables and other qualitative and quantitative factors which are peculiar to and necessary for an understanding and evaluation of the . . . company.”<sup>85</sup> MD&A “is intended to provide management with flexibility to describe . . . financial matters,” with the specific goal of “enabl[ing] investors to see the company through the eyes of management.”<sup>86</sup>

The MD&A requirement “grants firms considerable flexibility . . . to experiment with formats [for] conveying information.”<sup>87</sup> Management has cause “to hide bearish information or to use vague language” in its disclosures.<sup>88</sup> Despite this flexibility, the overarching goal of MD&A remains intact. This portion of the 10-K report “serve[s] as a qualitative disclosure for investors to make more accurate projections of future financial and operating results.”<sup>89</sup>

Many words found to be important in bankruptcy prediction are absent from standard dictionaries of positive and negative sentiment.<sup>90</sup> At a minimum, this outcome “suggests that bankruptcy prediction from text is more nuanced than sentiment analysis.”<sup>91</sup> The specific words—from capital structure (repurchase, dividend, tranches) to strategy (international, corus, and llamas) and shareholder satisfaction (wages, compensation, costs, suppliers)—invite deeper evaluation of emotionally dense information. “Future research can” transcend MD&A disclosures and “investigate the value of other channels such as news reports and user-generated content.”<sup>92</sup>

---

83. Concept Release on Management’s Discussion and Analysis of Financial Condition and Operations, Securities Act Release No. 6711, Exchange Act Release No. 24356, 75 Fed. Reg. 13715, 13717 (1987).

84. *See id.*

85. Disclosure and Analysis of Mutual Fund Performance Information, Securities Act Release No. 6850, Investment Company Act Release No. 17294, 55 Fed. Reg. 1460, 1462 (1990); *accord* SEC v. Conaway, 698 F. Supp. 2d 771, 818 (E.D. Mich. 2010) (quoting Securities Act Release No. 6349, 23 SEC Docket 962, 964 (Sept. 28, 1981)).

86. FINANCIAL REPORTING MANUAL, *supra* note 82, § 9110, at 296.

87. Mai et al., *supra* note 6, at 746.

88. *Id.*

89. *Id.* at 748.

90. *See* Loughran & McDonald, *supra* note 76, at 36; *see also* Wilson et al., *supra* note 69, at 347.

91. Mai et al., *supra* note 6, at 755.

92. *See* Mai et al., *supra* note 6, at 756.

## IV. CONTRIBUTIONS TO BANK MANAGEMENT

A. *Credit evaluation.*

The competitive application of bankruptcy prediction to the external processes of a bank—or, for that matter, those of any lender or investor supplying funds to a business—is obvious and powerful. “Effective prediction models” can help banks “make appropriate lending decisions.”<sup>93</sup> Proper prediction either reinforces or undermines legally required corporate disclosures of risk.<sup>94</sup>

The reasons for business failure, especially in a small, privately held firm, may not be transparent to a lender.<sup>95</sup> In business lending, bankruptcy prediction serves as the functional analog of credit scoring systems for individual consumers.<sup>96</sup> Indeed, at least one source has described its evaluation of lending risk as “corporate credit ratings.”<sup>97</sup> Even consumer credit evaluation, long the subject of traditional analytical techniques yielding a single “credit score,” faces the promise of improved predictive accuracy through the application of machine-learning algorithms.<sup>98</sup> Quite simply, bankruptcy prediction helps banks “better control their risk exposure” and reduce their “future number of bad debts.”<sup>99</sup>

The technological triumph of bankruptcy prediction models arises from the promise that machines can evaluate business borrowers as accurately and as ubiquitously as credit reports have come to dominate consumer lending. The legal pitfall lies in the absence of regulatory safeguards resembling those of the Fair Credit Reporting Act.<sup>100</sup> No comparable statutory scheme aspires to ensure fair and accurate credit reporting for business borrowers and to protect such borrowers’ privacy

---

93. Liang et al., *Financial Ratios*, *supra* note 75, at 561.

94. *See* Campbell et al., *supra* note 76, at 401.

95. *See* Peter L. Jennings & Graham Beaver, *The managerial dimension of small business failure*, 4 J. STRATEGIC CHANGE 185, 194 (1995).

96. *See* Liang et al., *Effect of Feature Selection*, *supra* note 15, at 289; *See also* Tsai, *supra* note 59, at 46.

97. *See* Haoming Zhong, Chunyan Miao, Zhiqi Shen & Yuhong Feng, *Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings*, 128 NEUROCOMPUTING 285, 285–86 (2014).

98. *See* Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun-Hwa Chen & Soushan Wu, *Credit rating analysis with support vector machines and neural networks: a market comparative study*, 37 DECISION SUPPORT SYSTEMS 543, 546 (2004); *see also* Amir K. Khandani, Adlar J. Kim & Andrew W. Lo, *Consumer credit-risk models via machine-learning algorithms*, 34 J. BANKING & FIN. 2767, 2783 (2010).

99. Adrian Gepp & Kuldeep Kumar, *Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques*, 54 PROCEDIA COMPUTER SCI. 396, 396 (2015).

100. *See* Fair Credit Reporting Act of 1970, Pub. L. No. 91-508, 84 Stat. 1127 (1970) (codified as amended at 12 U.S.C. §§ 1830–1831, 15 U.S.C. §§ 1681–1681x).

---

---

interests alongside the broader goal of promoting efficiency in the banking system.<sup>101</sup>

Despite the presence of factors that clutter financial data and confound conventional statistical analysis, such as “high correlated variables, outliers, and missing values,” machine-learning models for predicting bankruptcy have achieved “substantial classification accuracy” relative to “traditional mechanisms such as linear discriminant analysis, logistic regression, and artificial neural networks.”<sup>102</sup> Even without resorting to “time series or survival analysis,” machine-learning models promise to “improve predictive accuracy” at temporal distances greater than “the year before bankruptcy.”<sup>103</sup>

Not all errors in bankruptcy prediction carry the same weight, however. Machine learning treats bankruptcy prediction as “a binary classification problem.”<sup>104</sup> Those predictions can err in either direction.<sup>105</sup> As between false positives and false negatives during the detection of financial distress,<sup>106</sup> Type II false negatives consisting of decisions to lend to business borrowers who ultimately fail pose a far deeper threat to the financial well-being of banks.<sup>107</sup> Sensitive to the differential impact of Type I and Type II errors, one model incorporating classifier ensembles has evaluated its results through analysis of variance, a standard statistical test.<sup>108</sup>

From the borrower’s perspective, by contrast, a Type I false positive delivered by a prediction model is the primary source of injury. In the language of the Fair Credit Reporting Act, denial of credit constitutes “adverse action.”<sup>109</sup> Denial of credit to a financially sound business could provide a rebuffed borrower with some basis for alleging systematic bias in a machine-based method for evaluating financial distress. In the consumer setting, laws addressing discrimination include the Equal Credit

---

101. See *Safeco Ins. Co. of Am. v. Burr*, 551 U.S. 47, 52 (2007); see also *TRW Inc. v. Andrews*, 534 U.S. 19, 23 (2001).

102. Barboza et al., *supra* note 25, at 415.

103. *Id.*

104. Yu et al., *Bankruptcy Prediction*, *supra* note 31, at 296.

105. See Tom Fawcett, *An introduction to ROC analysis*, 27 *PATTERN RECOGNITION LETTERS* 861, 870 (2006).

106. For background on false positives and negatives in hypothesis testing, see Amitav Banerjee et al., *Hypothesis Testing, Type I and Type II Errors*, 18 *INDUS. PSYCHIATRY J.* 127 (2009); Joseph F. Mudge et al., *Setting an Optimal  $\alpha$  That Minimizes Errors in Null Hypothesis Significance Tests*, 7 *PLoS ONE* e32734 (2012).

107. See Liang et al., *Novel Classifier Ensemble*, *supra* note 60, at 438; see also Randall S. Sexton, Ram S. Sriram & Harlan Etheridge, *Improving Decision Effectiveness of Artificial Neural Networks: A Modified Genetic Algorithm Approach*, 34 *DECISION SCI.* 421, 430–31 (2003).

108. See Tsai, *supra* note 59, at 57–58.

109. See 15 U.S.C. § 1681m(a) (2012 & Supp. 2017); *Safeco Ins. Co. of Am. v. Burr*, 551 U.S. 47, 52 (2007).

Opportunity Act,<sup>110</sup> the Home Mortgage Disclosure Act,<sup>111</sup> and the Community Reinvestment Act.<sup>112</sup> Absent intentional discrimination in feature selection, however, it is hard to imagine how machine-based business lending decisions could support a disparate impact approach to liability for denial of credit.<sup>113</sup>

*B. Regulatory compliance.*

More accurate models for predicting bankruptcy and other sources of credit risk may help banks improve their performance of a distinct external function: regulatory compliance. Regulators at the national and international levels have identified credit risk management as a bedrock obligation of banks and a primary supervisory objective to be attained by their official overseers.<sup>114</sup>

By the same token, regulators may ask banks to make credit more freely available during times of economic distress, precisely when the tightening of credit would rationally counsel banks to exercise greater vigilance toward businesses seeking to secure or renew credit.<sup>115</sup> Bankruptcy prediction models can help banks balance their own solvency against regulatory demands during downturns and recessions. Under difficult economic conditions, governmental concerns may become more political or macroeconomic in nature, and policy uncertainty may exacerbate banks' peril in every respect.<sup>116</sup>

Amid such uncertainty, machine learning and artificial intelligence provide clarifying value. Bankruptcy prediction models informed by these techniques can evaluate country-level data, at least with respect to the

---

110. Equal Credit Opportunity Act of 1974, Pub. L. No. 93-495, § 701, 88 Stat. 1521, 1521–22 (1974) (codified as amended at 15 U.S.C. § 1691).

111. Home Mortgage Disclosure Act of 1975, Pub. L. No. 94-200, 89 Stat. 1125 (1975) (codified as amended at 12 U.S.C. §§ 2801–2811).

112. Community Reinvestment Act of 1977, Pub. L. No. 95-128, 91 Stat. 1147 (1977) (codified as amended at 12 U.S.C. §§ 2901–2908).

113. See *Tex. Dep't of Hous. & Cmty. Affairs v. Inclusive Cmty. Project*, 135 S. Ct. 2507, 2536 (2015).

114. See U.S. OFFICE OF THE COMPTROLLER OF THE CURRENCY, *RATING CREDIT RISK: COMPTROLLER'S HANDBOOK 1–4* (June 2017), <http://bit.ly/2Qeuu0M>; BASEL COMM. ON BANKING SUPERVISION, *PRINCIPLES FOR THE MANAGEMENT OF CREDIT RISK 1–2* (Sept. 2000), <https://www.bis.org/publ/bcbs75.pdf>; see also BASEL COMM. ON BANKING SUPERVISION, *BASEL III: A GLOBAL REGULATORY FRAMEWORK FOR MORE RESILIENT BANKS AND BANKING SYSTEMS 29–30* (June 2011), <https://www.bis.org/publ/bcbs189.pdf>.

115. See Patrick M. Parkinson, *Interagency Statement on Meeting the Needs of Creditworthy Small Business Borrowers*, BD. OF GOVERNORS OF THE FED. RESERVE SYS. 1 (Feb. 12, 2010), <https://www.federalreserve.gov/boarddocs/srletters/2010/sr1002.pdf>.

116. See Scott R. Baker, Nicholas Bloom & Steven J. Davis, *Measuring economic policy uncertainty*, 131 Q.J. ECON. 1593, 1626–34 (2016).



impact of energy efficiency policies on companies in the energy sector.<sup>117</sup> Advances in data science can reveal how effective the Troubled Asset Relief Program was in relieving pressure on American banks during the financial crisis of 2008 to 2009.<sup>118</sup>

The precise relationship between credit risk, liquidity risk,<sup>119</sup> operational risk,<sup>120</sup> and other threats to bank solvency remains obscure. One particular paradox of credit risk regulation gives pause, even as it lights a path for future research. By increasing correlations among different aspects of banks' risk-taking behavior, the reduction of credit risk within individual banks can have the perverse effect of raising the entire sector's *systemic* risk.<sup>121</sup> Prudential bank regulation aspires to patrol against systemic and cyclical risks arising from the unwitting coordination of banks' collective behavior.<sup>122</sup> The best methods for evaluating the threat that bank failures pose to the broader economy may arise from the very toolkit that is emerging for the prediction of credit risk through business bankruptcy: analytical models combining multiple methods drawn from statistics and machine learning.<sup>123</sup>

---

117. See Michalis Doumpos, Kostas Andriosopoulos, Emiliós Galariotis, Georgia Makridou & Constantin Zopounidis, *Corporate failure prediction in the European energy sector: A multicriteria approach and the effect of country characteristics*, 262 EUR. J. OPERATIONAL RES. 347, 351 (2017).

118. See Raffaella Calabrese, Marta Degl-Innocenti & Silvia Angela Osmetti, *The effectiveness of TARP-CPP on the US banking industry: A new copula-based approach*, 269 EUR. J. OPERATIONAL RES. 1029, 1036 (2017).

119. See Björn Imbierowicz & Christian Rauch, *The relationship between liquidity risk and credit risk in banks*, 40 J. BANKING & FIN. 242, 254–55 (2014).

120. See J. David Cummins, Christopher M. Lewis & Ran Wei, *The market value impact of operational loss events for US banks and insurers*, 30 J. BANKING & FIN. 2605, 2631 (2006).

121. See Rob Nijskens & Wolf Wagner, *Credit risk transfer activities and systemic risk: How banks became less risky individually but posed greater risks to the financial system at the same time*, 35 J. BANKING & FIN. 1391, 1398 (2011).

122. See Gabriel Jiménez & Jesús Saurina, *Credit cycles, credit risk, and prudential regulation*, 2 INT'L J. CENT. BANKING 65, 94–95 (2006).

123. See Yuliya Demyanyk & Iftekhar Hasan, *Financial crises and bank failures: A review of prediction methods*, 38 OMEGA 315, 319 (2010); Aykut Ekinci & Halil I. Erdal, *Forecasting bank failure: Base learners, ensembles, and hybrid ensembles*, 49 COMPUTATIONAL ECON. 677, 682; Hong H. Le & Jean-Laurent Viviani, *Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios*, 44 RES. INT'L BUSINESS & FIN. 16, 24 (2018); Shukai Li, Whye L. Tung, & Wee K. Ng, *A novelty detection machine and its application to bank failure prediction*, 130 NEUROCOMPUTING 63, 71 (2014).

## V. CONCLUSION

Bankruptcy prediction as a special case of data-driven decision-making<sup>124</sup> must meet a high standard of performance. It no longer suffices for a machine to persuade a human observer that its reactions are indistinguishable from that of a human.<sup>125</sup> Instead, in terms of accuracy, precision, sensitivity, and specificity—the quantitative criteria defining the receiver operating characteristic curve and the confusion matrix<sup>126</sup>—machine-based bankruptcy prediction must represent a meaningful improvement over human judgment standing alone, or perhaps even business judgment informed by conventional mathematical and statistical tools.

Standards for evaluating market risk models in banking provide a useful benchmark. *Traditional* back testing “considers” whether a “risk measure procedure is correct.”<sup>127</sup> This form of validation is based on the number of exceedances of a target level of risk, whether measured according to value-at-risk, expected shortfall, or some other mathematical representation of market risk in a regulated financial institution’s trading book.<sup>128</sup>

By contrast, *comparative* back testing asks whether a predictive model outperforms its rivals.<sup>129</sup> Evaluation of “a risk measure in practice” hinges upon “a panorama of the mathematical properties” of *all* measures and the assessment of each measure’s “advantages and disadvantages.”<sup>130</sup> The line between traditional and comparative back testing parallels the line dividing internal from external risk measures. “Internal risk measures are applied in the interest of an institution’s shareholders or managers, whereas external risk measures are used by regulatory agencies to maintain safety and soundness of the financial system.”<sup>131</sup>

Like any other external measure of risk worthy of application throughout the financial industry, a model for predicting financial distress

---

124. See FOSTER PROVOST & TOM FAWCETT, *DATA SCIENCE FOR BUSINESS: WHAT YOU NEED TO KNOW ABOUT DATA MINING AND DATA-ANALYTIC THINKING* 4–7 (2013); Foster Provost & Tom Fawcett, *Data science and its relationship to big data and data-driven decision making*, 1 *BIG DATA* 51, 52 (2013).

125. See Alan M. Turing, *Computing machinery and intelligence*, 59 *MIND* 433, 455–57 (1950).

126. See Fawcett, *supra* note 105, at 871.

127. See Natalia Nolde & Johanna F. Ziegel, *Elicitability and backtesting: Perspectives for banking regulation*, 11 *ANNALS APPLIED STAT.* 1833, 1840 (2017).

128. See Nick Costanzino & Michael Curran, *A simple traffic light approach to backtesting expected shortfall*, *RISKS*, Jan. 9, 2018, at 1, 2.

129. See Nolde & Ziegel, *supra* note 127, at 1833, 1848–49.

130. See Susanne Emmer, Marie Kratz & Dirk Tasche, *What is the best risk measure in practice?*, 18 *J. RISK* 31, 31–33 (2015).

131. See Steven Kou, Xianhua Peng & Chris Heyde, *External risk measures and Basel accords*, 38 *MATHEMATICS OPERATIONS RES.* 393, 394 (2013).

“must be unambiguous, stable, and capable of being implemented consistently.”<sup>132</sup> Robustness “with respect to underlying models and data” ensures that “different judges will reach similar conclusions when they implement” a predictive model.<sup>133</sup> By this lofty standard, novel methods for evaluating financial risk must outperform all previous analytical baselines.<sup>134</sup>

---

132. *See id.* at 401.

133. *See id.*

134. *See* James M. Chen, *On exactitude in financial regulation: Value-at-risk, expected shortfall, and expectiles*, 6(2) RISKS 61, 1, 9 (June 1, 2018); Nolde & Ziegel, *supra* note 127, at 1870.