Office of the Superintendent of Financial Institutions Canada
Bureau du surintendant des institutions financières Canada

Data Science and the Future of Financial Supervision

Remarks by Superintendent, Jeremy Rudin to the 11th Symposium on Asian Banking and Finance

San Francisco, California
June 25, 2018

Speaking notes:
Please check against delivery

For additional information contact:

Kaitlin Sabourin
Communications and Consultations
kaitlin.sabourin@osfi-bsif.gc.ca
www.osfi-bsif.gc.ca

OSFI
BSIF

Canada

Remarks by Superintendent Jeremy Rudin[1]
Office of the Superintendent of Financial Institutions Canada (OSFI)
to the
11th Symposium on Asian Banking and Finance
San Francisco, California
June 25, 2018

Introduction

Good afternoon and thank you for the kind introduction. It is a pleasure to be involved in the 11th Symposium on Asian Banking and Finance. I would like to thank the Federal Reserve Bank of San Francisco and the Monetary Authority of Singapore for inviting me.

In this segment of the Symposium, the organizers have asked us to consider whether supervisors are adequately advancing their tools and capabilities to keep up with how the financial system is evolving, and with reason.

We have all seen how technological change is reshaping the financial services industry. To my mind, the most important of these developments is the remarkable explosion in our ability to collect and analyze data. The data sets available now are so large that they call for new approaches to analyzing them; approaches that harness the continuing growth of computational power by using machine learning and artificial intelligence more generally.

When people refer to this new field as “data science” it is not just marketing: the analytical techniques applied to massive data sets represent a break from traditional statistics and econometrics.[2]

You already know that the financial services industry has embraced data science, and is rushing further into that embrace. Financial institutions are using these techniques in insurance underwriting, credit adjudication, anti-money laundering and many other areas.

That is a very important topic for supervisors, but it is not my topic today. I am not going to address how data science will change what *financial institutions* do, but rather how data science will change what their *supervisors* do.[3]

---

[1] I am grateful to Solon Angel, Jamil Abou Saleh, Andrew Kriegler and a number of my OSFI colleagues for their comments and insights, none of whom are responsible for possible errors in the text.

[2] For a comparison of machine learning to applied econometrics, see Sendhil Mullainathan and Jann Spiess, “Machine Learning: An Applied Econometric Approach,” *Journal of Economic Perspectives,* Spring 2017.

[3] For a perspective from one of the early adopters (the Monetary Authority of Singapore) see: David Hardoon, “Data Science and Machine Learning in Practice,” *Keynote Speech at the 7th Annual Sim Kee Boon Institute Conference on Advances in Data Science and Implications for Business*, May 26, 2017, available at http://www.mas.gov.sg/News-and-Publications/Speeches-and-Monetary-Policy-Statements/Speeches/2017/Data-Science-and-Machine-Learning-in-Practice.aspx.

Data science and its impact on supervisors

In my view, data science will have a major impact on how we supervise financial institutions.

My message to fellow supervisors is this: data science is too important to leave to the data scientists alone. As supervisors, we need to understand what data science can, and cannot do to assist us. If we fail to understand how to use data science wisely, we will fail to gain the full benefits of these powerful techniques. Moreover, we could divert our attention away from some important risks; a mistake for which we would only have ourselves to blame.

To try to convince you that data science is too important to leave to the data scientists alone, I am going to consider how it can apply to the work of two different supervisors. The first supervisor works in conduct supervision -- specifically detecting and prosecuting illegal insider trading. I am going to refer to that supervisor as *she*. The second supervisor works in the prudential supervision of large and complex banks. I am going to refer to that supervisor as *me*.

Data science and conduct supervision: detecting insider trading

Let us allow our conduct supervision colleague to go first.

Being a conduct supervisor, she has a number of goals. These include deterring, if not preventing, illegal insider trading. Her job also requires her to detect past instances of illegal insider trading so that she can prosecute it.

She has a vast amount of data at her disposal; very high frequency data on market transactions across a wide range of financial instruments, going back a number of years. She has reason to believe that there are signs in her data set that point to cases of illegal insider trading, if only she knew where to look.

She turns to data science for help.[4] What can she do?

Let us assume that she has already identified and prosecuted a number of cases of insider trading throughout her career. She uses that information to label the known insider trades in her data set. She then programs a machine-learning algorithm to look across the entire data set to find the patterns of trades that best match these labelled cases of insider trading. The algorithm can then look for those patterns elsewhere in the data.

What she gets is a list of the most promising trades to investigate for potential illegal insider trading, as seen by the algorithm. What will she find as she investigates those cases?

---

[4] There are already a number of applications of data science in this area. See, for example: “SEC’s advanced analytics helps detect even the smallest illicit market activity,” Reuters, June 30, 2017, available at: https://www.reuters.com/article/bc-finreg-data-analytics/secs-advanced-data-analytics-helps-detect-even-the-smallest-illicit-market-activity-idUSKBN19L28C, and James Langton, “IIROC to strengthen market surveillance with new technology,” *Investment Executive*, July 13, 2017, available at https://www.investmentexecutive.com/news/from-the-regulators/iiroc-to-strengthen-market-surveillance-with-new-technology/.

If she only has a small number of proven insider trades identified to start with, she will probably find that the predictions of the algorithm are not fully reliable. The problem, ironically, is that there is too much data relative to the number of proven insider trades. The algorithm can look at a nearly endless list of characteristics: how profitable the trades were; how many people made similar trades; how many times they traded; how quickly they traded; what the same traders were doing in the associated derivatives; and on, and on. Any handful of trades is likely to share a number of common features in this extremely long list, if only by chance. A machine learning algorithm will find many, if not all of those commonalities. Some of these are reliable indicators of insider trading while others are just coincidences.

This problem, called "overfitting", is familiar from econometrics. It is very common in this sort of machine learning exercise, and it is more acute when there is only a relatively small number of identified examples of the event that we are trying to detect.[5]

Overfitting is a problem, but it is not a fatal problem in her case. She is not using data science to conclude that a particular trade was an insider trade. She is using data science to identify suspicious-looking trades for further investigation. Once she has a list of promising places to look, she can collect other evidence and decide which of the suspicious trades identified by the algorithm to prosecute. Indeed, she is obliged to gather other evidence; she cannot sanction someone for insider trading based on data science alone.

Data science and prudential supervision: possible contributions

Now it is my turn.

As I am a prudential supervisor, my job is to reduce the likelihood of failures of the large and complex banks in my country to some acceptably low level. How do I do this?

There are many aspects of prudential supervision. Prudential supervisors of banks typically review or examine banks' risk management practices, identifying weaknesses and requiring remedial actions to address those weaknesses. We also set capital and liquidity requirements that go beyond minimum regulatory requirements.

In support of those activities, we look for indicators that point to banks that are too risky, or specific activities in banks that are too risky. We can then take steps to bring the risk back down to the acceptably low level.

My colleagues and I already do considerable data analysis. By combining increased computing power with more granular data, we can increase the speed, accuracy, and level of detail of our existing data analysis. That is a step forward; perhaps it will be a big step forward.

[5] There are many treatments of overfitting and various ways to guard against it. A good intuitive explanation of overfitting in big data is: Vincent Spruyt, "The Curse of Dimensionality in Classification," available at http://www.visiondummy.com/2014/04/curse-dimensionality-affect-classification/.

But it is not the big prize I am looking for. I have to believe that there are more powerful, more insightful, more useful indicators to discover in the data. I am going to try to use data science to find them.

The data set that I get to use draws on information from a single country, so it pertains to a limited number of large and complex banks. However, it has a vast amount of detail about each of those banks, going back a number of years.

If I am like many prudential supervisors, I will only have a few failures of large and complex banks in my data set. Why? Because there are not that many large and complex banks, and they do not fail that often. Like my conduct supervision colleague, my results will also be vulnerable to overfitting.

However, I cannot use the same approach to mitigate the overfitting problem that she used. I am not looking to detect something that happened in the past; I am looking to predict the unrealized future. I have no additional evidence that can conclusively confirm or disprove the identifications made by the algorithm.

In my case, asking for conclusive evidence is asking too much. I could instead decide to devote particular scrutiny to the banks that the algorithm picks out. This is feasible, but I only have a few large and complex banks and I am already keeping a close eye on them, so it is not very helpful.

It would be more useful if I could use the predictions made by the algorithm to tell me where to look more closely. Here I will have two issues that my conduct colleague will not have.

The first issue is that it is unlikely that there is a clear or intuitive interpretation of the predictions made by the algorithm. The algorithm will pick out what it sees as the riskiest banks, using a combination of indicators that could be scattered across several bank activities, with various lead times to failure. That will not tell me where or what to look for. This is what I should expect; machine learning algorithms are typically not set up to explain why they predict what they predict.[6]

Second, as I am a prudential regulator, I am not in the prediction business, I am in the prevention business. The algorithm may indeed have found the best *predictors* of future failure. What I need to know is the most likely *causes* of future failure. Those may be two different things. There is no point in taking supervisory action to change something that is known to predict bank failures if that "something" does not also contribute to failure. Stated more simply: silencing the rooster will not keep the sun from rising.

My colleague does not have this problem. Once she has identified a suspicious trade, she is able to confirm if it can be prosecuted or not. Anything that points her to promising trades to investigate, hidden among the millions of trades in her data set, is useful to her. It does not matter to her if the machine has found the fire, or only the smoke.

---

[6] There are a number of efforts at mitigating this well-known deficiency. See, for example, Lars Hulstaert, "Interpreting machine learning models," available at https://towardsdatascience.com/interpretability-in-machine-learning-70c30694a05f.

The feedback issue: How our practices affect the data

We do share another issue, however. We both have to deal with the fact that the way we supervise will affect the data that we are using.

Let us use a prudential example to illustrate the point this time. Suppose that our data shows that a bank that engages in a particular practice is prone to later failure. Because we are in the prevention business, we could decide to prohibit that practice unless the bank takes some approved steps to offset the risk. Suppose further that this works to prevent failures arising from that risky practice. This new supervisory practice will feed back into the data.

As we collect data under our new regime, observing that particular risky practice in a bank will no longer predict the future failure of that bank. Our supervisory practice will cause the data to obscure the underlying risks.

There is a second, and perhaps more pernicious way that our supervisory practice can obscure the true risks. Suppose that we mistakenly prohibit a practice that is not truly risky. The data will never reveal that mistake, because we will not allow the practice to appear in the data and so prove its innocence.[7]

My colleague also faces this same issue. Potential insider traders may figure out the indicators she used to identify insider trades. The miscreants may then learn how to conduct insider trading without triggering those indicators. It will be important that she not be complacent if there is a drop in the number of cases that her algorithm flags. This drop could indicate that she is getting better at detecting insider trading, and that improved detection has a deterrent effect; or it could mean that the miscreants are getting better at avoiding detection.

Harnessing the power of unsupervised machine learning

So far, I have been referring to the prudential supervisor in my examples as *me*, but that is not entirely accurate. Currently, I cannot try the approach described above, because I do not have any failures of large and complex banks in my data set.

Is there anything data science can do for me? Indeed there is. When we are unable to label the cases of interest in our data set we can use an approach called, somewhat ominously, unsupervised machine learning. In unsupervised learning, the algorithm looks for anomalies in the data. The algorithm does not know that there is anything wrong with these cases; it just knows that they stand out in some way. That said, these cases might be good candidates for further investigation.

For example, if my conduct supervision colleague had no confirmed cases of illegal insider trading at hand, she could use unsupervised learning to identify anomalies for further investigation.

---

[7] The risk that a machine-learning algorithm will not be able to see and correct its mistakes is one of the themes of: Cathy O'Neil, *Weapons of Math Destruction*, 2016.

What am I going to do with my anomalous cases? I can certainly look into them further to see if they point to excessively risky activities. Unlike my colleague, I have no way to prove that the anomalies found are excessively risky; after all, there are no failures in my data set. I will instead use my judgement and experience as a prudential supervisor to try to determine which, if any, of the anomalies point to excessive risk-taking. As I do that, I have to be aware of two possible pitfalls.

First, having invested time and effort in searching for anomalies, I am going to tend to find reasons to believe that the anomalies found are indeed indicators of excessive risk. Confirmation bias, the tendency to think that you have found what you expect to find, is difficult to combat.

If we succumb to this bias now, we will not be able to correct it later. By discouraging banks from doing whatever it is that creates the anomaly, we will ensure that the anomaly will not reappear in the data. As we discussed earlier, it will not get a chance to prove its innocence.

Second, as I progressively stamp out the anomalies in the data, I will make my large and complex banks more similar to each other. That could be good, if I am stamping out excessively risky activities. That is the type of similarity that we seek. But it could be bad, if it reduces diversification across the banks, making all of them vulnerable to the same shocks.

There is a less ambitious, but perhaps more reliable way that I can use unsupervised learning on a data set without bank failures. I could train unsupervised learning on the regulatory submissions made by the banks, and then investigate whether the anomalies identified by the algorithm point to submissions that are incorrect or fraudulent.[8] This is something that I can independently verify, albeit with some effort, and it is certainly useful. Consistently incorrect or, worse, fraudulent regulatory reporting should definitely raise a red flag over the bank in question.[9]

Conclusion

In these few minutes, we have only scratched the surface of the topic. There are a number of other cases we could consider, and more techniques that are being made available through data science. I want to stop before I reach the limits of your patience and I need to stop before I reach the limits of my knowledge.

So let me leave you with this thought. Professions like prudential supervision, that rely heavily on judgement and experience, tend to resist the idea that algorithms can improve on what

[8] For an application of this approach in the insurance industry, see: Sutapat Thiprungsri and Miklos A. Vasarhelyi, “Cluster Analysis for Anomaly Detection in Accounting Data: An Audit Approach,” *International Journal of Digital Accounting Research*, 11, 2011. The idea of using mathematical anomalies to detect fraud predates the development of machine learning. See, for example, the history recounted in: Cindy Durtschi, William Hillison and Carl Pacini, “The Effective Use of Benford’s Law to Assist in Detecting Fraud in Accounting Data,” *Journal of Forensic Accounting*, 5, 2004.

[9] The Bank of England is experimenting with unsupervised learning in a few contexts. See: https://www.bankofengland.co.uk/research/fintech/proof-of-concept.

human expertise can accomplish alone.[10] That is a mistake and we cannot afford to make that mistake.

Data science will be important in financial supervision. It will be so important that it should not be left to the data scientists alone.

Thank you.

[10] This resistance long predates the development of data science. See for example, Daniel Kahneman, *Thinking, Fast and Slow*, 2011, notably chapters 21 and 22.