A Note on the Use of Machine Learning in Central Banking[1]

Alexandra Kinywamaghana
Sascha Steffen

This version: July 13, 2021

Preliminary Version

Abstract

Artificial Intelligence (AI) and Machine Learning (ML) have gained significance in central banking decisions over the last years. After the 2008 global financial crisis, the mandate of central banks has broadened and new (big) data for central banks have emerged. In this note, we briefly discuss how AI and ML can support monetary policy decisions drawing from the recent literature and highlight several examples of how these methods are applied, *e.g.* at the European Central Bank and Bank of England. We document that AI and ML appear to be useful for central banks in four key areas, *i.e.* enhancing statistical information, macroeconomic analysis, and forecasting, monitoring of financial market indicators, and assessing financial risk. We also highlight some shortcomings associated with using AI/ ML. Finally, we introduce the "Artificial Intelligence and Monetary Policy Decisions" project as part of the Financial Big Data Cluster (safeFBDC), in which we explore AI and ML techniques to support monetary policy decisions.

Keywords: Artificial Intelligence, Big Data, Machine learning, Central Banks, forecasting, risk assessment, Financial Big Data Cluster (FBDC)

---

[1] Alexandra Kinywamaghana, Project Coordinator FIRE & FBDC, Frankfurt School of Finance & Management, Adickesallee 32-34, 60322 Frankfurt email: a.kinywamaghana@fs.de
Sascha Steffen, Professor of Finance, Frankfurt School of Finance & Management, Adickesallee 32-34, 60322 Frankfurt, email: s.steffen@fs.de
We are grateful for financial support from the BMWi and DLR.

Big data and monetary policy

Traditionally, the key objective of central banks (particularly, the European Central Bank) has been price stability and central banks have a set of conventional monetary policy tools available to achieve this objective. However, since the global financial crisis (GFC) of 2008/2009, the toolbox has changed, it also contains "unconventional" monetary policy tools, e.g. quantitative easing. Moreover, central banks have been tasked with new responsibilities to include the measurement of systemic risk, banking regulation and supervision, digital currencies, and climate change.[2] These responsibilities are in part a result of the collection and access to new sources of data — introducing central banks to **"Big Data"** (Chakraborty and Joseph, 2017).[3]

What is big data? While engineering and statistics might already have a clear definition as what they understand big data is, financial economists, do not yet have a broad-based definition. In a 2021 special issue on "Big Data in Finance" published in the *Review of Financial Studies*, Goldstein et al. (2021) try a tentative definition. In their view, big data has three properties: (1) **large size**, i.e. sample size; (2) **high dimension**, i.e. the data have many variables relative to sample size (see also Martin and Nagel, 2019); and (3) **complex structure**, i.e. these data are unstructured, including e.g. text, pictures or audio.

Where does this data come from? Central banks now have access to an unprecedented amount of data to facilitate monetary policy decisions, and they can also draw data from various sources including the internet. The biggest part of the data, however, is based on micro-transactions between firms (e-commerce, credit card transactions), public statistics, and financial market data.[4] Overall, about 80% of central banks used big data in 2020 as compared to 30% in 2015 (Doerr et al., 2021), and about 40% used big data to inform policy decisions. Compared to the private sector, however, only a few central banks have embraced big data (Tissot, 2018).

Artificial Intelligence (AI) and Machine Learning (ML) have gained importance in this process. AI and ML can further improve the data basis for monetary policy decisions of

[2] Central banks are increasingly acknowledging the risk climate change poses on price stability e.g., an increase in inflation due to weather-related events. On 25. January 2021, the ECB announced several important initiatives to address climate change, ECB, ECB sets up climate change centre, Jan 25, 2021, https://www.ecb.europa.eu/press/pr/date/2021/html/ecb.pr210125_1~3fc4ebb4c6.en.html .

[3] Several developments since the GFC e.g., the growth of FinTechs and Shadow banks have generated new data on the micro and macro level, that are used to facilitate monetary policy decisions.

[4] The major sources of data for monetary policy decisions are internet-based indicators (mainly web-based), commercial data sets, financial market indicators and administrative records. One example of big data obtained by central banks are credit registries, which have been used widely in academic research (e.g., Peydro et al., 2020).

central banks: for example, (1) by providing more complete, immediate and granular information to complement existing macroeconomic indicators; (2) using new sources of data such as Google searches, real estate, consumer prices on the internet, or social media[5]; and (3) introducing new data collection techniques, e.g., web-scraping, text-mining, or matching of different data sources.

An important question is then how to analyse big data most efficiently. Data might be high-dimensional or unstructured and financial economists need to extract information from this data (Goldstein et al., 2021). Moreover, financial big data might exhibit noise, heavy-tailed distributions, nonlinear patterns, or temporal dependencies, rendering traditional econometric methods insufficient to analyse them (Petropoulos et al., 2018). AL and ML emerged as new tools (together with other approaches) to address challenges arising from big data.

ML provides not only new tools but also solves a distinct problem compared to traditional economic applications and approaches. Importantly, while traditional approaches and applications revolved around parameter estimation, ML tools revolve around the problem of prediction, being able to uncover generalisable patterns in data (Mullanaithan and Spiess, 2017).

Overall, the use of AI/ML by central banks raises important questions about how it can support monetary policy decision making. This note summarises a recent discussion on how central banks can use AI/ML in their monetary decisions focusing on four questions: (1) How can AI/ML support improve monetary policy decision making? (2) In what area is AI/ML applied? (3) How is AI/ML applied today? (4) And, what are potential problems associated with using AI/ML?

How can AI / ML support/ improve monetary policy decision making?

As described above, central banks now have a broad range of responsibilities including price stability, the measurement of systemic risk, banking regulation and supervision, the future of digital currencies, as well as addressing climate change. And central banks apply AI/ML

[5] Particularly smartphones and cloud computing produce unprecedented amounts of data that finance sector and central banks can use to aid decisions. For example, market participants can track the spread of information in social networks through social media. Wibisono et al. (2018) provide a general overview on the use of AI/ML in central banking.

methods to address challenges, which, for financial economists, provides fruitful areas for future research. Below, we describe and highlight related research in these areas.

- **Banking supervision**: The ECB is implementing several Supervisory Technology (SupTech) projects, including advanced data analytics and textual analysis.[6] SupTech is any form of innovative technology used to support financial supervision, e.g., cloud computing or machine learning techniques.[7] For example, Hannes et al. (2018) investigate early warning systems applied to euro area banks. Goldsmith-Pinkham et al. (2016) use a computational linguistics approach to compare the approaches of supervisors and market monitors.
- **Economic forecasting and nowcasting**. Use big data and AI and ML methods to forecast business cycles (incl. components such as GDP, inflation, and other monetary aggregates), for example, through monitoring of consumer durable goods and understanding the labour market, *e.g.* using job adverts (Tissot, 2018). The suitability of ML methods for predictive analysis makes them especially attractive in the context of financial asset return predictability and risk premium measurement (*e.g.*, Gu et al., 2020).
- **Measuring (policy) uncertainty.** Using textual analysis and machine-reading of business news to construct an index to measure uncertainty in economic policy and the state of the economy (e.g., Bloom et al., 2016; Bybee et al., 2021). Saltzman and Yung (2018) use a machine learning approach to identify different types of uncertainty from Federal Reserve Beige Books.
- **Financial stability assessments.** Network analysis appears to be a useful tool to study the interconnectedness of financial (and non-financial institutions) to study potential sources of systemic risk improving on "small data" approaches, as in Cai et al. (2018).

[6] See ECB, "Digital transformation of supervision picks up pace," February 17, 2021, https://www.bankingsupervision.europa.eu/press/publications/newsletter/2021/html/ssm.nl210217_3.en.html.
[7] See ECB, "Exploring the potential of supervisory technology," August 12, 2020, https://www.bankingsupervision.europa.eu/press/publications/newsletter/2020/html/ssm.nl200812_3.en.html.

A recent example where high-frequency data might have been useful was the coordination of retail traders using social media during the GameStop episode in January 2021. The interaction of retail traders and sophisticated investors led to extreme market volatility and financial stability concerns.[8]

- **Central bank communication**. Textual analysis can be used to construct a measure of central bank communication and investigate what information is communicated and through what channels, and how markets react and to what signals they react. Central banks communicate through various channels, such as press conferences, meetings, but recently also social media as well as through forwarding guidance (Haldane and McMahon, 2018). Recent literature investigates different aspects of central bank communication using ML approaches. For example, Schmeling and Wagner (2019) show that central bank tone moves asset prices. Ehrmann and Wabitsch (2021) investigate central bank communication with non-experts using Twitter data. Ehrmann and Talmi (2020) investigate the implications of semantic similarity in central bank communication for market volatility.
- **Money laundering/financial crime**: AI can increase the effectiveness and efficiency of financial crime investigations and risk management in financial and non-financial institutions (Proudman, 2018). There has been an evolution from a rules-based approach to monitoring anti-money laundering to ML approaches that collect customer data with publicly available information from the internet to detect suspicious activities and flows of funds.
- **Credit risk analysis and scoring** helps banks assess credit scores and default risk and improve consistency by limiting human bias (Enria, 2019). Using natural language processing of reports and social media can provide firms with useful information on the credit quality of customers. Earlier, for example, applied these methods to consumer credit (e.g. Khandani et al., 2010).
- Banks may use AI-based models to enable better decision-making and to **provide products and services to a wider customer base**, including the underserved and under-banked population as a possible contribution to fight poverty (Blumenstock et al., 2015; Blumenstock, 2016).[9]

[8] The hedge fund Melvin Capital lost 53% in January 2021. J. Chung, "Melvin Capital Lost 53% in January, Hurt by GameStop and Other Bets," Wall Street Journal, January 31, 2021, https://www.wsj.com/articles/melvin capital-lost-53-in-january-hurt-by-gamestop-and-other-bets- 11612103117 .

[9] See ECB, "Bank to the future: supervisors take on fintech innovation," November 13, 2019, https://www.bankingsupervision.europa.eu/press/publications/newsletter/2019/html/ssm.nl191113_1.en.html.

- **Processing of complaints,** for example, Banca d'Italia applies AI techniques to process customer complaints and Banque de France is working on an algorithm to assess the compliance of banks' regular supervisory reports and other submissions automatically and rate information quality using natural language processing (Hakkarainen, 2020).

AI and ML also help central banks in their data capabilities:

- **Enhancing data collection**, it would be possible for regulators to extract data, at any level of granularity, directly from firms' systems in real-time (Proudman, 2020); sourcing data on consumer behaviour (spending and saving) through third parties.
- **Enhancing and refining analytical capacities**; AI can detect non-linear patterns in data, and handling datasets with many predictors.[10] According to Genberg and Karagedikli (2021), ML´s strength lies in its ability to predict—to uncover patterns in data that have not been pre-defined e.g., using model aggregation techniques from online machine learning to obtain the best out-of-sample predictor for systemic financial crises (Fouliard et al., 2020).
- **Reducing operational costs** by relieving supervisors of repetitive tasks, which may eliminate human error.
- **Shortening application processing time**, banks gain greater transparency on the decision-making process because they can now track their applications. Online application forms with mandatory fields mean that supervisors receive more complete applications, which reduces requests for additional documents after initial submission (Hakkarainen, 2020).

[10] See ECB, "Digital transformation of supervision picks up pace," February 17, 2021, https://www.bankingsupervision.europa.eu/press/publications/newsletter/2021/html/ssm.nl210217_3.en.html.

How is AI/ML applied today?

In this section, we explore (selected) **examples** of how market participants (e.g., ECB, banks) use AI/ML methods in practice.

- The ECB uses the "Truffle Analytics" implemented for the Supervisory Review and Evaluation Process (SREP) based on ML, to analyse SREP proceedings and decisions (Hakkarainen 2020). This helps the bank to spot similarities across different banks and identify trends.
- The ECB collects data daily from MTS Markets (a digital trading platform for governments) to produce eurozone daily yield curves[11] and uses Google data for nowcasting.
- The ECB is using data from Prisma on the prices of goods to look at volatility and resilience analysis of certain categories with the product basket.
- The ECB is working on a text mining project with Factiva. Text mining techniques make it possible for banks to analyse qualitative data. Some of the text data that central banks often analyse include news articles, financial contracts, social media, supervisory and market intelligence, and reports (Bholat, 2015). In addition, the Bank of Mexico (BoM), and the UK Financial Conduct Authority (FCA) combine web-scraping and text mining to audit promotional materials and financial advice documents from financial institutions (Di et al., 2019).
- The Bank of Canada is using ML to perform sentiment analysis on surveys and monetary policy reports to improve forecasting and detect anomalies.
- Crime detection and prevention: Banks in general use AI to detect and combat fraud. For example, HSBC is using AI to screen customer data against publicly available data to detect suspicious behaviours.[12]
- The Bank of Greece employs data mining algorithms on data collected from corporate and SME loans of the Greek banking system to reduce dimensionality and increase accuracy in predicting the future behaviour of corporate loans (Petropoulos et al., 2018).

[11] See Central Banking, February 26, 2021, "Digital policy-making in a data-driven world," https://www.centralbanking.com/central-banks/monetary-policy/7801436/digital-policy-making-in-a-data-driven-world.

[12] See M. Arnold, "HSBC brings in AI to help spot money laundering," April 8,2018, https://www.ft.com/content/b9d7daa6-3983-11e8-8b98-2f31af407cc8.

- AI is used in investments and trading activities e.g., ING's Katana tool uses predictive analytics, based on historical and real-time data to help traders to set bid and ask prices.

What are the potential problems?

AI/ML will transform all areas of society in positive ways, but innovations also introduce risks. In this section, we divide the risks into four categories (data, methodological, interpretability, and ethical).

Data challenge

- **Availability and quality of data:** a lack of access to data in a timely and usable manner and the presence of "Knightian uncertainty". AI assesses known knowns with a clearly defined question and uses historical data to infer conclusions, *e.g.*, fraud detection or insurance underwriting (Carney, 2018).
- **Embedded bias in data:** Bias in data and increased interconnections could lead to potentially pro-cyclical behaviour as AI/ML simply replicates the bias.
- **Data privacy and security issues:** very granular data might contain private information, which might be misused by third parties (Tissot, 2018). This calls for a need to have procedures in place to ensure data privacy. Central banks are built on trust, therefore, it´s important to have standards and a governance structure in place for the responsible use of data (Santor, 2021).
- **Reliability of insights:** e.g. data samples from social media may not be representative and/or reliable. Economists should conduct validation checks and quality controls to ensure the accuracy and reliability of information (Di Castri et al., 2019).

Methodological challenges

- **Data integration**: it is difficult to integrate AI collected data into a comprehensive information model.
- **Endogenous risk**: AI excels at measuring exogenous risk that is of little interest to macro authorities (Danielsson et al., 2020). Endogenous risks result from the interaction between agents in the market, e.g., regulators and financial institutions pursuing their objectives. These are interactions with unclear objectives and data,

established statistical techniques, and repetitive events that serve little purpose. Periods of stress may limit the behaviour of agents causing behavioural synchronisation and eventually major stress events or a crisis.

- **Reputational risk:** associated with ML as algorithms have stronger predictive power and weak explanatory. This may lead to public scrutiny if insights are used to justify policy decisions (Wibisono et al., 2019).
- **Biased algorithms:** Algorithms replicate biases that may be embedded in the data, *i.e.*, how data are collected can affect conclusions. Furthermore, ML predictions lead to biased estimates of causal effects (Athey, 2017).

Interpretability

- **Opacity and "Black box effect":** AI-based models lack transparency[13], it is not explicit how and why they reach a certain conclusion as users cannot see how nodes are analysing data; they only see the conclusion. Furthermore, "early adopters of AI" have found that complex computations are sometimes hard to understand.[14]
- **Private vs Public institutions:** a need for different skills to operate an effective oversight, risk, and control environment (Wibisono et al., 2019).
- **False sense of security:** A danger of supervisory staff relying on AI which can both undermine contingency planning and preventative regulatory measures and ultimately create a false sense of security (Danielsson et al., 2020).

Ethical

- **Proxy discrimination.** ML analysis might involve practices that disproportionately harm members of a specific class when it relies on proxies for class membership, *e.g.* zip codes (Prince and Schwarcz, 2020).

[13] See e.g. ECB, "Bank to the future: supervisors take on fintech innovation," November 13, 2019, https://www.bankingsupervision.europa.eu/press/publications/newsletter/2019/html/ssm.nl191113_1.en.html.
[14] See e.g. ECB, "Bringing artificial intelligence to banking supervision, " November 13, 2019, https://www.bankingsupervision.europa.eu/press/publications/newsletter/2019/html/ssm.nl191113_4.en.html.

How does the use of AI/ML in monetary policy decisions relate to the Financial Big Data Cluster (safeFBDC) Project?

As described above, big data and AI/ML play an important role in supporting central banks´ monetary policy decision-making processes, notably in enhancing statistical information and refining analytical capacities, macroeconomic analysis, and forecasting, monitoring of financial market indicators, and assessing financial risk.

Between 2021- 2023, the "Artificial Intelligence and Monetary Policy Decisions" project (being part of the safe Financial Big Data Cluster (safeFBDC)) will investigate the importance and use of AI and ML-based in some of these areas described above in different, well-defined, case studies.[15] The project aims to improve monetary policy decision making in the Eurozone in two dimensions: (1) improve the data basis for monetary policy decisions; (2) use AI / ML methods to generate new information that is valuable for central banks and monetary policy to investigate questions related to macro-analysis and forecasting, supervision of financial indicators and the assessment of financial stability risks. More broadly, these case studies will also help market participants (such as financial and non-financial institutions, regulators, academics) and other "observers" to better understand the reasons and implications of monetary policy decisions.

[15] We provide more information and disseminate our research, papers and presentations through this website.

References

Athey, Susan, 2017. "Beyond Prediction: Using Big Data for Policy Problems." *Science*, 355 (6324), pp. 483-85.

Bholat, David, 2015. "Big Data and Central Banks." *Big Data & Society*, 1, pp. 1-6.

Baybee, Leland, Bryan Kelly, Asaf Manela, and Dacheng Xiu, 2021. The structure of economic news. Working Paper.

Baker, Scott, Nicholas Bloom, and Steve Davis, 2016. Measure economic policy uncertainty, *Quarterly Journal of Economics* 131(4). 1593-1636.

Blumenstock, Joshua E., Gabriel Cadamuro, and Robert On, 2015. Predicting Poverty and Wealth from Mobile Phone Metadata, *Science* 350(6264): 1073–76.

Blumenstock, Joshua E., 2016. Fighting Poverty Poverty with Data, *Science* 353(6301): 753–54.

Cai, Jian, Frederik Eidam, Anthony Saunders, and Sascha Steffen, 2018. Syndication, Interconnectedness and Systemic Risk. *Journal of Financial Stability* 34, 105-120.

Carney, Mark, 2018. "AI and the Global Economy." Speech, Machine Learning and the Market for Intelligence Conference, Rotman School of Management, University of Toronto, Toronto, October 23, 2018.

Chakraborty, Chiranjit, and Andreas Joseph, 2017. Machine Learning at Central Banks. Bank of England Working Paper.

Danielsson, Jon, Robert Macrae, and Andreas Uthemann, 2020. Artificial Intelligence and Systemic Risk. Working Paper.

Di Castri, Simone, Stefan Hohl, and Arend Kulenkampff, 2019. *The Suptech Generations.* Publication no. 19. October 2019.

Doerr, Sebastian, Leonardo Gambacorta, and Jose Maria Serena, 2021. "Big Data and Machine Learning in Central Banking." BIS Working Papers.

Ehrmann, Michael, and Jonathan Talmi, 2020. Starting from a blank page? Semantic similarity in central bank communication and market volatility, *Journal of Monetary Economics* 111, 48-62.

Ehrmann, Michael, and Alena Wabitsch, 2021. Central Bank Communication with Non-Experts – A Road to Nowhere? Working Paper.

Enria, Andrea, 2019. "A Binary Future? How Digitalisation Might Change Banking." Speech, Banking Seminar Organised by De Nederlandsche Bank, Amsterdam, March 11, 2019.

Fouliard, Jeremy, Michael Howell, and Hélène Rey, 2020. Answering the Queen: Machine Learning and Financial Crises. Working paper.

Genberg, Hans, and Özer Karagedikli, 2021. Machine Learning and Central Banks: Ready for Prime Time? Working paper.

Goldsmith-Pinkham, Paul, Beverly Hirtle, and David Lucca, 2016. Parsing the content of bank supervision. Federal Reserve New York Staff Report No. 70.

Goldstein, Itay, Chester Spatt, and Mao Ye, 2021. Big Data in Finance, *Review of Financial Studies*, forthcoming.

Gu, S., B. Kelly, and D. Xiu. 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33:2223–73.

Hakkarainen, Pentti, 2020. "Digitalising Banking Supervision: An Ongoing Journey, Not a Final Destination." Speech, The Supervision Innovators Conference, November 30, 2020.

Khandani, Amir, Adlar Kim, and Andrew Lo, 2010. Consumer credit-risk models via machine-learning algorithms, *Journal of Banking & Finance* 34(11), 2767-2787.

Lang, Jan H., Tuomas Peltonen, and Peter Sarling, 2018. A framework for early-warning modelling with an application to banks, ECB Working Paper.

Martin, Ian and Stefan Nagel, 2019, Market Efficiency in the Age of Big Data, NBER Working Paper.

Mullainathan, Sendhil, and Jann Spiess, 2017, Machine Learning: An Applied Econometric Approach, *Journal of Economic Perspectives* 31(2), 87-106.

Petropoulos, Anastasios, Vasilis Siakoulis, Evaggelos Stavroulakis, and Aristotelis Klamargias, 2018. A Robust Machine Learning Approach for Credit Risk Analysis of Large Loan-level Datasets Using Deep Learning and Extreme Gradient Boosting. Proceedings of IFC – Bank Indonesia International Workshop and Seminar on "Big Data for Central Bank Policies / Building Pathways for Policy-Making with Big Data", Bali, 2018.

Peydro, Jose-Louis, Andrea Polo, and Enrico Setti, 2020. Monetary Policy at Work: Security and Credit Application Registers Evidence, *Journal of Financial Economics*, forthcoming.

Prince, Anya E.R., and Daniel Schwarcz, 2020. "Proxy Discrimination in the Age of Artificial Intelligence and Big Data." *Iowa Law Review*, 105(3), 1257-318.

Proudman, James, 2018. "Cyborg Supervision – the Application of Advanced Analytics in Prudential Supervision." Speech, Workshop on Research on Bank Supervision, London, November 09, 2018.

Proudman, James, 2020. "Supervisor-centred Automation – the Role of Human-centred Automation in Judgement-centred Prudential Supervision." Speech, Conference on the 'Impact of AI and Machine Learning on the UK Economy', March 26, 2020.

Robinson, Paul, 2018. "Big Data: New Insights for Economic Policy –The Bank of England Experience." Presentation, IFC – Bank Indonesia International Workshop and Seminar on "Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data", Bali, 2018.

Salzman, Bennett, and Julieta Yung, 2018. A machine learning approach to identifying different types of uncertainty. *Economics Letters* 171, 58-62.

Santor, Eric, 2021. "Digital Policy-making in a Data-driven World." Speech, Improved Central Bank Data Management in a Digital Age, January 21, 2021.

Schmeling, Maik, and Christian Wagner, 2019. Does Central Bank Tone Move Asset Prices? Working Paper.

Tissot, Bruno, 2018a. Big Data for Central Banks. Proceedings of IFC – Bank Indonesia International Workshop and Seminar on “Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data”, Bali. 2018.

Wibisono, Okiriza, Hidayah Dhini Ari, Anggraini Widjanarti, and Alvin Andhika Zulen, 2019. The Use of Big Data Analytics and Artificial Intelligence in Central Banking - An Overview. Proceedings of The Use of Big Data Analytics and Artificial Intelligence in Central Banking, Bali. May 2019.