ECONOMIC NEWSLETTER
No. 21-03 | December 2021

THINKING AI-HEAD: Exploring Machine Learning Applications in Central Banks

By:
Rosemarie A. Amodia, Vidal Marvin C. Gabriel, and Cherrie R. Mapa[1]

[1] Ms. Amodia is Bank Officer V, Ms. Mapa is Acting Deputy Director, while Mr. Gabriel is a former Data Analytics Specialist at the Department of Economic Research.

Rapid improvements in computing power and the advent of big data paved way for remarkable progress in the field of artificial intelligence (AI). AI is no longer just a feature in futuristic sci-fi movies or novels. In today's data-driven world, AI is revolutionizing business processes and society's way of life. It is increasingly permeating various aspects of daily living because of the expanding reach of the internet and social media platforms, rising accessibility of smart devices, and increasing adoption of AI-powered applications across different organizations. AI brings the benefits of convenience, timely responses, operational efficiency, and cost-effectiveness, among others. Thus, many people view AI as the defining technology that will continue to shape the world and could very well be considered a must-have strategic capability.

One of the subfields of AI is ***machine learning (ML)*** or algorithms that are capable of learning patterns from data that enable them to make predictions, derive new information, and formulate decisions. Given the wide-ranging scope of possible applications, ML has been attracting the interest of numerous researchers, including those in central banking circles.

This newsletter provides an overview of ML and how central banks have been exploring the technology. It also cites some of the challenges in adopting ML-based solutions and briefly discusses some of the early initiatives at the Bangko Sentral ng Pilipinas (BSP).

Machine learning defined

While often used interchangeably, AI and ML are not quite the same. AI refers to computer systems that can perform tasks that normally require human intelligence. This encompasses a range of applications from simple rule-based systems to more complex and advanced models capable of learning. Notably, responses of rule-based AI could be limited by the set of conditions it was provided. Thus, comes the appeal of ML, an AI subfield focused on algorithms that deliver outputs based on patterns learned from data, including their own. However, since ML algorithms are dependent on what they were trained on, outcomes are not pre-defined and can result in some inaccuracies. In such cases, human intervention may be needed to aid the models in their learning process.

Meanwhile, a subset of ML known as deep learning requires no human intervention to learn from its inaccuracies. Deep learning utilizes multiple layers of artificial neural networks that attempt to mimic the human brain. It is often used in image detection, speech recognition, and language translation.

ML applications can be found in various fields. They are often used to provide solutions for identified needs. For example, in ecommerce, recommender systems built on ML algorithms work on big data[2] to help enhance customers' shopping experience. This allows personalized product recommendations and provides an avenue for businesses to further boost sales through cross-selling and up-selling products that would likely entice a customer to purchase. Shoppers of Lazada and Shopee may notice product suggestions that are pushed by these online shopping platforms under Lazada's Just for you and Shopee's Daily Discover features, respectively. Such product recommendations rely on technology that profiles each customer's past purchases and search history to predict product preferences. Online entertainment platforms, such as Netflix and Spotify, also employ similar recommender systems to understand and deliver the content preferred by its respective subscribers.

In finance, ML applications can be used to enhance credit scoring, client interface, and insurance risk management activities. ML algorithms help provide a data-driven assessment of a client's credit risk and subsequent loan pricing. Financial companies also leverage AI to enhance efficiency in interacting with clients via chatbots. Meanwhile, insurtech[3] relies on big data processing to assist insurance companies in underwriting and claims processing.[4]

The Philippine government has similarly utilized ML techniques in the conduct of its functions. For instance, the Remote Sensing and Data Science: DATOS Help Desk (DATOS Project) of the Department of Science and Technology – Advanced Science and Technology Institute (DOST-ASTI) developed an AI model that could detect flooded areas from satellite images with no human intervention, which is useful for disaster mapping.[5] Meanwhile, the Sugar Regulatory Administration (SRA), through its work with the DATOS Project, uses an AI model for its nationwide sugarcane mapping operations.[6]

---

[2] Big data generally refers to voluminous and variedly structured datasets produced at high frequency that can deliver valuable insights.
[3] Insurtech broadly refers to technological applications in the insurance industry.
[4] Financial Stability Board (2017).
[5] de la Cruz, et al.(2020).
[6] Olfindo Jr., et al. (2020).

Machine Learning and Central Banks

Central banks have displayed an increasing interest in exploiting big data and the ML algorithms that are usually used to process and analyze said data structure. In a 2020 survey conducted by the Bank for International Settlements (BIS) Irving Fischer Committee (IFC), 80 percent of the respondent central banks noted that they use big data, up from the 30 percent response in 2015. The survey revealed that big data projects of central banks can be categorized under four main applications:

1. ***Natural language processing*** – the use of qualitative, text-based information to produce a quantitative summary, such as in an economic policy uncertainty index;

2. ***Nowcasting*** – the generation of real-time forecasts of key indicators that are usually produced at low frequencies, such as inflation and GDP, involving the use of both traditional economic data and non-traditional data sources, such as information scraped from the web;

3. ***Applications that extract economy-wide insights from granular financial datasets*** - the use of large datasets submitted for regulatory and statistical purposes to determine interconnectedness and spot emergence of systemic risk; and

4. ***Suptech and Regtech[7] applications*** – the use of firm-level information, such as those for micro-level risk assessment, to support micro-supervisory policies.[8]

   Table 1 below shows some ML-related initiatives pursued by a sample of central banks.

Machine Learning Applications at the BSP

Similar to other central banks, the BSP has explored applications of ML techniques in the areas of forecasting and banking supervision. This section provides an overview of two of these initiatives.[9]

Inflation Forecasting. The BSP's Department of Economic Research (DER) has developed the first regional inflation forecasting models for the Philippines employing ML approaches.[10] Using support vector regression (SVR), artificial neural network (ANN), and long-short term memory (LSTM) to forecast regional inflation,[11] the results

Table 1. Some ML-related initiatives pursued by a sample of central banks

CENTRAL BANK	DETAILS
Bank Indonesia (BI)	Using ML, the research developed a new index that measures stakeholders' expectations on BI's policy rate. Text mining techniques were employed to process the news articles ahead of further transformations and test five different ML classification models to identify likely actions of BI (rate hike, rate cut, rate unchanged, or no expectations). The resulting new index has a 76.6 percent correlation to the actual policy rate change and an F1-score[12] of 76.8 percent.
Bank of England	ML was applied to underpin early warning models that can predict financial crises. The Shapley value framework, a novel approach that helps in the interpretability of the ML models, was used to examine its predictors. The research found that ML models generally outperformed the benchmark method (logistic regression), with a consistent identification of similar predictors for financial crises.
Reserve Bank of New Zealand	The research explored various ML algorithms using around 600 domestic and international variables to produce nowcasting models for the gross domestic product (GDP) growth of New Zealand and determine if these models can outperform traditional and commonly used techniques (naïve autoregressive model, dynamic factor model), including their current official forecasts. The research discovered that some of the ML models were more accurate than their chosen benchmarks and could thus serve as complementary tools for understanding the current state of their economy.

Sources: IFC-Bank Indonesia International Workshop and Seminar on Big Data for Central Bank Policies/Building Pathways for Policymaking with Big Data, Bank of England Staff Working Paper, Reserve Bank of New Zealand Discussion Paper.

[7] Suptech and Regtech broadly refer to technological applications that aid in entity supervision and facilitate regulatory compliance, respectively.
[8] BIS IFC (2021).
[9] The BSP has been embarking on several ML and big data initiatives to improve and digitally transform its operations, as well as enrich its research and policy analysis capacity. In the interest of brevity, cited in this article are only two examples, which the authors had also previously worked on.
[10] Gabriel, Bautista and Mapa (2020).
[11] Support vector regression (SVR) is an algorithm that finds the best fit solution given an acceptable range of error. Artificial neural network (ANN) is a computational model that mimics the learning process of the human brain in converting inputs over layer(s) of transformation into relevant output(s). Long-short term memory (LSTM) is a special type of a recurring neural network (RNN) that uses historical learnings, retaining what could be relevant and discarding what are likely not, to process new information.
[12] F1-score is a classification accuracy measure that considers both the ratio of true positives against all predicted positives and the ratio of true positives against all actual positives

for the month-ahead and 12-month-ahead forecasting exercises on seven representative regions showed that two of the three ML methods (SVR and ANN) outperformed the benchmark ARIMA models, with the SVR having the least forecast errors. Given this, SVR was adopted to model inflation for the remaining 10 regions. Currently, all SVR regional inflation forecasting models are operational and regularly used by the DER to generate inflation nowcasts for all 17 regions in the country. A nationwide inflation nowcast is also produced using the weighted regional inflation nowcasts. These models complement the BSP's existing suite of models for macroeconomic forecasting.

Detecting Atypical Data. The BSP's Department of Supervisory Analytics (DSA) has explored the use of ML techniques to enhance its data validation process.[13] Applying a combination of ARIMA, linear regression, k-nearest neighbor (KNN), least absolute shrinkage and selection operator (LASSO) regression, decision tree, random forest, extra trees, gradient boosting method, and k-means clustering algorithms[14] to model account behavior, a framework was developed as a prototype to define the typicality of an account and consequently identify atypical data. In doing this, the framework considered an account's historical behavior, relationship with other accounts, and peer-based movements. Using a data validation tool designed to operationalize the prototype framework, the initial results of the study showed an estimated 98 percent time savings on exercising this validation layer compared with its manual process equivalent. Moreover, the study provided a proof of concept that ML techniques could support and strengthen the surveillance capacity of the BSP, offering a potential for further explorations for a more proactive, enriched, and efficient set of processes.

Some Limitations and Challenges

There are several limitations and challenges associated with ML processes, a few of which are discussed below.

1. The most often-cited limitation is the **black-box approach to ML.** While empirical studies provide evidence that ML techniques can outperform traditional regression analysis in forecasting, there is difficulty in interpreting the causal relationships in ML models. In contrast, traditional econometric models allow users to make inferences on causation and identify how an explanatory variable influences the dependent variable. This aspect is critical for economic analysis and policy formulation. Doerr et al. (2021) argue that an ML model that predicts an economic recession in the next period could have limited value in helping policymakers avert the recession if such a model cannot point to the specific indicators that could drive the downturn. The same could be argued for suptech and regtech applications. According to Doerr et al. (2021), an ML-enabled stress test that signals possible failures in the banking system may not provide adequate information that bank supervisors can act on pre-emptively if the sources of failures are not explicitly identified. It could be noted, however, that AI researchers have been actively looking for ways to address this concern.[15]

2. Similar to other traditional econometric models, ML algorithms may encounter **some challenges in accurately predicting tail risk events**. This is mainly because tail risk events are low-frequency events that may not always be captured in training datasets. While researchers have explored various methods that could address this (from data augmentation techniques to probabilistic type of models, among

[13] Amodia, et al. (2018).

[14] K-nearest neighbor (KNN) regression is an algorithm that considers the pattern or characteristics of similar data points in predicting the behavior of a target variable. Least absolute shrinkage and selection operator (LASSO) is a form of linear regression that penalizes some variables such that at an optimal calibration, features that could better explain the target variable are given more importance. Decision Tree is an algorithm that separates data into branches of subsets based on decision rules, while Random Forest is a collection of decision trees and Extra Tree is an algorithm that randomizes the tree-building process. Gradient boosting method is an algorithm that introduces predictors step-by-step while keeping residuals at a minimum. K-means clustering is an algorithm that identifies data clusters whose centroid is based on the average value of its members.

[15] As an example, Joseph (2020) proposed to address this gap with a generic statistical inference framework. Based on the Shapley-Taylor decomposition, the said framework allows for the estimation of the contribution and significance of explanatory variables to the predicted value of a dependent variable, on average.

others), it may not be sufficient to assume that the next tail risk event would be driven by the same factors that triggered the previous one, such as those of financial and economic crises. This suggests that while ML models can learn, they are limited by the dataset from which they learn. Hence, in their current state, there may still be a need to supplement and/or complement the results from some ML models with expert judgment.

3. Adopting ML as a solution to help address business problems **requires investments in infrastructure and a shift in organizational mindset**. Especially when using big data, institutions (including central banks) that wish to dabble with ML need to have sufficient computing power, and a sizable and secure data storage capability. These projects should also have trained staff who are proficient in the technical aspects of developing and operationalizing these algorithms, from data engineering, programming, the mathematics behind ML models, statistics, big data processing if handling such, data visualization, model deployment, as well as business domain knowledge. ML can be a strategic tool capable of enabling an organization in its digital transformation, but it can also be an expensive liability if not used according to its potential. Hence, management buy-in, concrete vision, and support for the initiatives are other important aspects to successfully integrate ML solutions as a strategic tool in an organization's roadmap. For instance, at the BSP, digital transformation is identified as one of the programs being developed and implemented to support the achievement of the BSP strategic objectives for 2020-2023.[16] The BSP takes on a holistic approach in paving its digital transformation journey, taking into consideration its resources, internal policies, processes, workforce, and stakeholders, as well as the evolving technology.

4. The use of big data in ML **could raise issues relating to data privacy and ethical data and modeling practice**. Data mined from the internet should be handled delicately. Large transactional databases (such as payment data or bank reports) must also be treated confidentially, as laws would dictate. In the Philippines, the Data Privacy Act of 2012 guarantees the privacy of individuals by regulating the processing of personal information. Data privacy issues also necessitate investments in cybersecurity tools and organizational guidelines that supplement existing laws to ensure proper handling of personal information and other confidential data. Meanwhile, prudence in model development should also be exercised such that models generate outputs that are reliable, fair and/or nondiscriminatory. Among other activities, this could include thorough data familiarization and preprocessing such that veracity is established and inputs to the model are balanced and/or representative data sets. Also, as Doerr et al. (2021) point out, the "garbage in garbage out" rule likewise applies to ML. It is, therefore, imperative that ML model developers ensure the reliability and representativeness of data fed into the algorithms, and that apt disclosure and interpretation of results are presented, especially in models that aid in decision or policymaking. In economic analysis, for instance, using big data to model business or consumer expectations would require careful consideration on the part of ML model developers in identifying its data sources and what indicators to consider/include in their models given the possible inherent bias in the available data. Particular to the financial sector, Prenio and Yong (2021) found five common principles among issuances on AI governance: reliability/soundness, accountability, transparency, fairness, and ethics. As countries cope to develop standards on AI governance, organizations could check existing laws and regulatory guidelines, along with the industry-specific dialogues on AI governance in developing their governance framework that could supplement these rules.

[16] BSP Organization Primer.

CONCLUSIONS

ML offers a myriad of opportunities in central banking, especially when combined with techniques from other disciplines such as econometrics and network science. Central banks have been actively pursuing initiatives that employ ML to aid in the discharge of their mandates. These projects have so far typically focused on nowcasting, text mining, as well as suptech and regtech applications to complement or enhance current practices in forecasting and surveillance activities.

In the BSP, these initiatives include, among others, nowcasting regional inflation that complemented the existing suite of macroeconomic models and a prototype framework for detecting atypical data that opened other explorations to enhance current data validation processes.

Looking ahead, continued advances in computing power and increases in digitized information could provide further impetus for ML. The COVID-19 pandemic might have the unintended but welcome impact of accelerating the digitalization of processes due to the required social distancing measures that resulted in increased demand for digital services. Adopting ML across various parts of an organization to improve processes or innovate products and delivery of services is one of the enabling tools toward digital transformation. In addition, the National AI Strategy Roadmap, launched in May 2021, serves as an enabling guide and resource for organizations looking to start or continue using AI in aiding their digital transformation journey to adapt to market and technological developments. Amid this, the BSP continues to explore ML applications that can be useful in the conduct of its key functions while carefully taking note of the challenges associated with adopting these algorithms.

REFERENCES

Amodia, R., Orencia, M., Soriano, C., and Sto. Domingo, J. (2019). Intelligent Algorithms for Detecting Potential Atypicality in Prudential Reports: A test case using Regulatory Submissions of Banks Supervised by the Bangko Sentral ng Pilipinas (Unpublished master's thesis). Asian Institute of Management, Makati, Philippines.

Bholat, D., Gharbawi, M., and Thew, O. (2020). The impact of COVID on machine learning and data science in UK banking. Bank of England Quarterly Bulletin 2020 Q4.

Bluwstein, K., Buckmann, M., Joseph, A., Kang, M., Kapadia, S., and Şimşek, Ö. (2020). Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach. Bank of England Staff Working Paper No. 848.

Chakraborty, C. and Joseph, A. (2017). Machine Learning at Central Banks. Bank of England Staff Working Paper no. 674.

Chen, S. (2019), Here's how Shopee is developing the next generation of data science and AI talent. Retrieved from https://www.techinasia.com/heres-shopee-developing-generation-data-science-ai-talent.

de la Cruz, R. M., Olfindo Jr., N. T., Felicen, M. M., Borlongan, N. J. B., Difuntorum, J. K. L., and Marciano Jr., J. J. S. (2020). Near-real time flood detection from multi-temporal Sentinel RADAR images using artificial intelligence. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B3-2020, 1663–1670. Retrieved from: https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1663-2020.

Department of Trade and Industry. (2021). DTI to launch National Artificial Intelligence Roadmap.

Doerr, S., Gambacorta, L., and J. Serena (2021). Big data and Machine Learning in Central Banking, BIS Working Paper no. 930.

Financial Stability Board (2017). Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications.

Gabriel, M., Bautista D., and C. Mapa. (2020). Forecasting regional inflation in the Philippines using Machine learning techniques: A New Approach. Bangko Sentral ng Pilipinas Working Paper no. 2020-10.

IBM Cloud Education. (2020). Machine Learning. Retrieved from https://www.ibm.com/cloud/learn/machine-learning.

Irving Fischer Committee on Statistics. (2021). Use of big data sources and applications at central banks, IFC Report no. 13. Bank for International Settlements.

Joseph, A. (2020). Parametric inference with universal function approximators. Bank of England Staff Working Paper No. 784.

Kavlakoglu, E. (2020). AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the difference? Retrieved from https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks

Olfindo Jr., N. T., de la Cruz, R. M., Borlongan, N. J. B., Marciano Jr., J. J. S., and Olalia, L. C. (2020). Sugarcane plantation mapping using dynamic time warping from multi-temporal Sentinel-1A RADAR images. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-3-2020, 519–524. Retrieved from https://doi.org/10.5194/isprs-annals-V-3-2020-519-2020.

Prenio, J. and Yong, J. (2021). Humans keeping AI in check – emerging regulatory expectations in the financial sector. FSI Insights on Policy Implementation No. 35, Bank for International Settlements. Retrieved from https://www.bis.org/fsi/publ/insights35.pdf.

Price, D. (2018). Jargon Alert: Machine Learning. Econ Focus (Federal Reserve Bank of Richmond). Third Quarter 2018, Volume 23, Number 3.

Richardson, A., van Florenstein Mulder, T. and Vehbi, T. (2019). Nowcasting GDP using machine learning algorithms: A real-time assessment. Reserve Bank of New Zealand Discussion Paper 2019-03.

Singapore Computer Society. (2020). Simplifying the Difference: Machine Learning vs Deep Learning. Retrieved from https://www.scs.org.sg/articles/machine-learning-vs-deep-learning.

Vichit-Vadakan, V. (2020). Lazada knows you! How the largest Southeast Asian e-commerce platform is using AI to enhance the online shopping experience. Retrieved from https://digital.hbs.edu/platform-digit/submission/lazada-knows-you-how-the-largest-southeast-asia-e-commence-platform-is-using-ai-to-enhance-the-online-shopping-experience/

Zulen, A. and Wibisono, O. (2018). Measuring stakeholders' expectations for the central bank's policy rate. IFC – Bank Indonesia International Workshop and Seminar on Big Data for Central Bank Policies/Building Pathways for Policymaking with Big Data. Irving Fisher Committee on Central Bank Statistics, Bank for International Settlements.