Machine learning applications in central banking

Douglas Araujo, Giuseppe Bruno, Juri Marcucci, Rafael Schmidt, Bruno Tissot[1]

Executive summary

On 18–22 October 2021, the Irving Fisher Committee on Central Bank Statistics (IFC) and the Bank of Italy co-organised, with the support of the European Central Bank (ECB) and the South African Reserve Bank (SARB), a **workshop on "Data science in central banking" that focused on machine learning (ML) applications**. This event was an opportunity to take stock of how central banks are deploying ML across a variety of use cases. It also illustrated the importance of these new techniques in improving the efficiency and effectiveness of their related operations, including by increasing their availability to deal with larger and new sources of information in a more automatised way.

Indeed, the workshop underlined the diversity and maturity of ML approaches already developed and used by central banks. **This reflects their potential and usefulness for central banks in dealing with the increasingly complex environment in which they operate**.

To start with, the new techniques can **help gather more and better information**, which is key for central banks that rely heavily on data. ML can help respond to this demand by enhancing the data quality, eg dealing with outliers, addressing the problems posed by missing values, limited frequency and/or timeliness, and by providing richer contextual insights.

In addition, **a key issue for central banks is to make sense of the wealth of data available to derive useful insights on specific economic and financial situations.** This needs to happen in a reasonably fast and largely automated fashion, considering the constantly changing environment. Coping with the often exponential growth of data and associated complexity of the statistical analysis is a challenge for central bank statisticians. Fortunately, ML can greatly help central banks in this

1 Respectively, Economist, Monetary and Economic Department, Bank for International Settlements (BIS) (Douglas.Araujo@bis.org); Director, Economics and Statistics Directorate, Bank of Italy (Giuseppe.Bruno@bancaditalia.it); Economist, Economics and Statistics Directorate, Bank of Italy (Juri.Marcucci@bancaditalia.it); Head of IT, Monetary and Economics Department, BIS (Rafael.Schmidt@bis.org); and Head of Statistics and Research Support, BIS & Head of the Secretariat of the Irving Fischer Committee on Central Bank Statistics (IFC) (Bruno.Tissot@bis.org).

The views expressed here are those of the authors and do not necessarily reflect those of the Bank of Italy, the BIS, the ECB, the IFC, the SARB or any of those institutions represented at the workshop.

We thank Thomas Gottron, Robert Kirchner, Silke Stapel-Weber and Ulf von Kalckreuth for helpful comments and suggestions.

context by facilitating the modelling of economic and financial problems and supporting the related statistical exercises.

In turn, **the insights gained can effectively back the conduct of evidence-based central bank policies.** This is obviously the case regarding monetary stability, not least in terms of better understanding the drivers of monetary policy decisions that can be provided by ML. Similarly, applying ML in suptech can be instrumental in helping financial supervisors to perform their oversight tasks, including identifying and tackling micro-level fragilities and other emerging threats such as climate-related financial risks. Turning to the macroprudential perspective, central banks can benefit from the increased use of ML to interpret information from various, often unrelated, data sources to assess system-wide vulnerabilities and their evolution over time. Moreover, the new techniques can support other tasks that are also relevant from a financial stability perspective, including the functioning of the payment system, financial inclusion, consumer protection, anti-money laundering and the secure printing of money.

At a more practical level, the workshop provided useful benchmarking, feedback and training on ML models for the participants. **Several lessons and observations are worth noting for those in charge of deploying ML-based tools in their central banks.**

First, there is a wealth of alternative information sources that have barely been tapped by central banks and which can provide new, useful insights if explored with ML techniques. The ultimate goal is that policymakers have at their disposal better-quality, timelier and interpretable data when taking decisions, especially in uncertain times such as the Covid-19 pandemic. Second, complementarity is essential: ML methods can provide additional insights to traditional approaches but have to be blended with other types of exercises as well as with strong business expertise. Third, there are benefits to calibrating many ML tools, not just one, since combining different approaches can provide better results with usually limited additional effort. Particular emphasis needs to be placed on avoiding ML model overfitting, eg through cross-validation. Fourth, there is merit in following a pragmatic and gradual approach when implementing the new tools. A considerably varied set of ML methods can be considered, and it is important to carefully assess them before actual deployment, with due consideration of the available skill set and computing environment. Fifth, having more data is often better than increasing the sophistication of the ML model. Sixth, while ML can be instrumental in dealing with complexity, there is also a risk of developing black box solutions that would compound the challenges faced by users as their functionality is rarely intuitive. The focus should therefore be on the interpretability of the results obtained and on addressing well defined use cases. Lastly, ML exploratory work has only started, and substantial staff and IT investment as well as business adjustments will continue to be needed to make the most of the new techniques, computing equipment and data.

Addressing these issues will require **further modifications in central banks' current operational processes** – eg in developing software ("DevOps") and putting ML algorithms into production ("MLOps") – **and collaboration models** – with close cooperation between core IT experts, data scientists and business specialists. It also puts a premium on the IFC's mission to promote cooperation between central banks through the sharing of national use cases and to draw relevant lessons from the experiences observed outside the public community.

1. Introduction: increased central bank use of ML techniques

One of the IFC's raisons d'être is to foster cooperation between central banks on statistical issues based on showcasing projects and sharing national experiences. To this end, **the Committee has initiated recurrent workshops on "Data science in central banking"** aimed at a broad audience of practitioners and technicians, with the goal of reviewing the adoption of data analytics and business intelligence techniques and developments in the big data ecosystem. The first event, hosted by the Bank of Italy in October 2021 with the support of the ECB and the SARB, focused on the contribution of ML applications to central banking. This virtual event was attended by almost 500 participants, representing about 180 institutions from the public and private sectors.

ML can be defined as an algorithm – a method of designing a sequence of actions to solve a problem – that optimises automatically through experience (ie from data) and with limited or no human intervention (FSB (2017)). ML algorithms are a subset of Artificial Intelligence (AI) techniques and are typically divided into four main types: supervised, unsupervised, reinforcement and deep learning (Wibisono et al (2019)). They have been increasingly used in economic and financial academic and practitioner settings, and central banks are not far behind. One reason is that the **compilation of large and/or complex granular databases (Israël and Tissot (2021)) and the development of big data analytics (IFC (2019)) have spurred their ability to use ML tools** to support the conduct of their policies, especially in the areas of monetary and financial stability, including the associated statistical, analytical and communication tasks (Chakraborty and Joseph (2017), Doerr et al (2021) and Bruno and Marcucci (2021)).

Indeed, the IFC workshop highlighted the **diversity of ML approaches developed in central banking**. Authorities are exploring, and in some cases already deploying, ML techniques to support a wide range of use cases that encompass macroeconomic modelling, economic and inflation analysis, the support of monetary and financial stability policies (including microprudential tasks for those central banks in charge of financial supervision) and specific statistical work (eg detection of data anomalies). Moreover, the range of central banks involved in this exploratory work is broad and comprises most advanced economies as well as a growing number of emerging market economies.

The fact that diverse ML tools have been successfully applied across a wide spectrum of use cases underscores the great value of the analytical insights they can provide as well as the operational gains brought about by automatising and making more efficient various production processes. **One key benefit for central banks is, in particular, the ability to deal with the increasingly complex environment in which they operate**. As regards monetary policy for example, the communication of policy decisions has become increasingly important and multifaceted, especially after the Great Financial Crisis (GFC) of 2007–09 (Gros (2018), Cieslak and Schrimpf (2019) and Hansen and McMahon (2018)); it has in particular benefited from the use of natural language processing (NLP) techniques (Gentzkow et al (2019)) to facilitate dealing with textual information (Apel et al (2021), Ferreira (2021), Hansen et al (2018) and Ahrens and McMahon (2021)). Another example relates to financial supervisory tasks, for which new and complicated topics are constantly emerging, such as those related to climate-related financial risks (Hernández de Cos (2022)) or to the consequences of the Covid-19 pandemic (Casanova et al (2021)); it has in fact been argued that ML can increase central banks' efficiency in the supervisory area by helping them cover more ground with the same resources (Beerman et al (2021)).

To be successful, **ML projects require the availability of adequate staff and IT resources as well as good coordination with the business areas** (IFC (2020a)). Fortunately, the wide variety of the techniques already deployed by central banks suggests that they have been able to both rely on adequate human skills – including subject matter and IT experts, and data scientists – and address the associated complex IT requirements. They have benefited in particular from the fact that many ML computing frameworks are widely available in the public domain as "open source" – cf SCIKIT-LEARN (Pedregosa et al (2011)), PYTORCH (Paszke et al (2019)) and TENSORFLOW (Abadi et al (2015)). Moreover, new approaches such as transfer learning are further improving the accessibility to central banks of large, high-performance ML models (Box 1).

The above developments have allowed central banks to take advantage of the wide range of ML techniques to address increasingly sophisticated use cases, as illustrated in Professor Michael McMahon's keynote speech. For instance, traditional NLP techniques can be combined with algorithms that recognise the temporal dimension of texts (cf Chang and Manning (2012)) to assess the information content of monetary policy statements. Another interesting case is the use of NLP techniques calibrated according to specific macroeconomic variables to analyse central bank communication.

Yet the continuous development of ML algorithms and related areas, such as big data analytics and cloud computing, is likely to require further substantial investment in skilled staff (eg data scientists). As emphasised in Bank of Italy Deputy Governor Piero Cipollone's introductory speech, the necessary skills transcend the technical to include also the ability to recognise and address the challenges and risks inherent in data science, such as the presence of bias in big data sets and the imperative to consider data integrity, confidentiality and privacy.

Further investment in IT equipment remains necessary. In fact, many central banks are in the process of implementing, or have already implemented, a cloud adoption strategy to, inter alia, better enable ML/high-performance computing use cases and to facilitate related collaboration with external researchers (as observed eg in the case of the Bank of Canada). A key reason is that cloud computing offers more agility for data analysis and experimentation; access to computing power is easier to scale up, and, importantly, the responsibility for maintaining an updated hardware and software environment in a so-called "plug-and-play model" (ie requiring little involvement by the end user in the necessary IT setting) can be shifted from internal staff to external service providers.

To sum up, the availability of large sets of data from many different and unstructured sources along with the development of new, innovative analytics and IT tools are **changing the financial landscape in which central banks operate**. This provides a key opportunity to leverage (automatised) ML algorithms to strengthen their economic and financial modelling toolkits, analytical capabilities and risk management tools, in turn helping them to pursue their mandates more effectively.

The following sections elaborate on the various types of ML applications that can be used effectively in view of the projects presented at the workshop. Section 2 describes ways to support central banks in making the best use of the data available as a key input for their operations. Sections 3 to 6 outline ML use cases in the specific areas of macroeconomic analyses, monetary policy, micro-financial supervision and (macro-)financial stability.

Box 1

Facilitating the use of machine learning by central banks with transfer learning

Douglas Araujo

Some ML models aim to achieve or even exceed human-level performance based on large and complex information, such as big data sets, text or images. Using them typically requires powerful and sophisticated IT systems due to their training on vast amounts of data. Such models can only achieve the desired performance if they are trained with a correspondingly large amount of data. For example, GPT-3,① a model famous for its ability to create human-like text, comprises 175 billion parameters that were trained on petabytes of text from two broad internet text corpora, numerous books and the whole of English language Wikipedia. Thus, the development of models with the highest performance in certain fields is usually done only by a few organisations with sufficient resources and specialised computer engineering capabilities.

Transfer learning is a technique that facilitates the use of these large models. The starting point is to download a pre-trained version of them and then fine-tune their specifications to the particular use cases at hand. In practice, organisations such as big tech firms (eg Google, Meta) or boutique ML firms (eg HuggingFace, DeepMind) will typically train reference models on selected big data sets. Then, they usually make the models publicly available by storing them in so-called "model hubs",② where the general public can search, compare and download desired models. Once a suitable pre-trained model is selected, the user can run it off the shelf on more specific data sets, or more commonly, fine-tune the model to the data for that particular use case. Importantly, given all the extensive pre-training work already done, the fine-tuning step can be effective even with a limited volume of data. Thus, the user benefits from high performance without the need to spend considerable time and resources creating or refining the model. This technique also facilitates the use of several large models for comparison or as an ensemble.

Transfer learning can help advance the use of ML by central banks. First, there is a broad range of models available in the main hubs, so that central banks can easily find adequate models to address a wide range of specific use cases they might have. This supports central banks in jumpstarting ML-powered projects in different areas without having to invest excessive resources or time. Second, central banks can explore and combine different models depending on the circumstances and analytical needs. Third, comparing these external "state-of-the-art" models with their own models and algorithms can help central banks enhance the performance and accuracy of such internally developed applications.

However, there are important challenges associated with transfer learning. One relates to quality assurance: model hubs are typically administered by reputable institutions, but in many instances the models themselves are developed and posted in the hubs by third parties. Hence, proper care during model selection is warranted. For instance, it is important to correctly test models trained by third-party entities to avoid biases or limitations that might be important in the particular use case of interest. Another issue to bear in mind is that, because transfer learning entails downloading the original models, the application would need to be (often manually) updated in case a new version of the original model is made available.

In summary, transfer learning can facilitate the use of high-performance ML models to support a variety of central bank applications, in turn supporting their use of big data analytics and ML tools. Moreover, transfer learning is operationally much simpler than developing models from scratch and managing them, reducing the associated burden in terms of resources (eg necessary skill set, IT equipment, budget). Finally, the development phase of ML models may require considerable energy. It has, for instance, been reported that the CO_2 emissions from training a single large model can reach multiples of the emissions of an average car during its whole lifetime.③ By enabling the use of pre-trained models instead of developing very similar models by multiple central bank users, transfer learning can help to limit the carbon footprint from ML usage.

① See T Brown, B Mann, N Ryder, M Subbiah, J D Kaplan, P Dhariwal, A Neelakantan, P Shyam, G Sastry, A Askell, S Agarwal, A Herbert-Voss, G Krueger, T Henighan, R Child, A Ramesh, D Ziegler, J Wu, C Winter, C Hesse, M Chen, E Sigler, M Litwin, S Gray, B Chess, J Clark, C Berner, S McCandlish, A Radford, I Sutskever, D Amodei, "Language models are few-shot learners", *NeurIPS Proceedings*, 2020. ② Examples of widely used model hubs are TensorflowHub (https://tfhub.dev), PyTorchHub (https://pytorch.org/hub) and HuggingFace Models (https://huggingface.co/models). ③ E Strubell, A Ganesh and A McCallum, "Energy and policy considerations for deep learning in NLP", *ACL*, 2019.

2. Gathering better and more information

Apart from being a key pillar of national statistical systems as producers of official statistics, central banks are heavy users of information to support the conduct of their policies, which are increasingly based on quantitative evidence. ML can help to respond to this appetite by enhancing the data quality, eg dealing with outliers and addressing the problems posed by missing values, limited frequency and/or timeliness, and by providing richer contextual insights.

Setting up adequate quality assurance frameworks

The growing availability of large and complex granular data sets ("financial big data"; IFC (2015)) obtained from statistical or supervisory reporting often at the transaction level and at very high frequencies (such as daily), has amplified **the need for better and faster data quality management frameworks** to ensure that the information collected can be reliably used for statistical production. In particular, many central banks have been leveraging on ML algorithms, sometimes in combination with traditional methods, to develop new data validation processes to better check the quality of the data at stake and correct them more effectively and/or efficiently.

One recent example is the ECB's new **anomaly detection project to support data quality checks in the production of statistics** on euro short-term interest rates. Their compilation is derived from a granular data set on individual transactions observed in money markets, ie the Money Market Statistical Reporting (MMSR), which involves 47 banks located in 10 countries and represents a total of around 50,000 daily transactions. There were important challenges related, in particular, to the presence of non-numerical variables, distribution skewness, the need for rapid data quality checks to support a daily production process, and the difficult interpretability of the results obtained for the users. These challenges were addressed with the use of various ML techniques to convert categorical variables into numerical ones, exploit the observed correlations, and detect anomalies through different models/algorithms, namely: standard regression analysis to compare observed data with model estimates; isolation forest and hierarchical clustering to isolate specific data points from the rest of the distribution; anomaly identification, based on the training on past data (supervised learning) using XGBoost (Chen and Guestrin (2016)); and the use of the Local Interpretable Model-Agnostic Explanations (LIME) algorithm

(Ribeiro et al (2016)) to facilitate the interpretation of the algorithm's results that could otherwise resemble a "black box" and provide users with a more interpretable model.

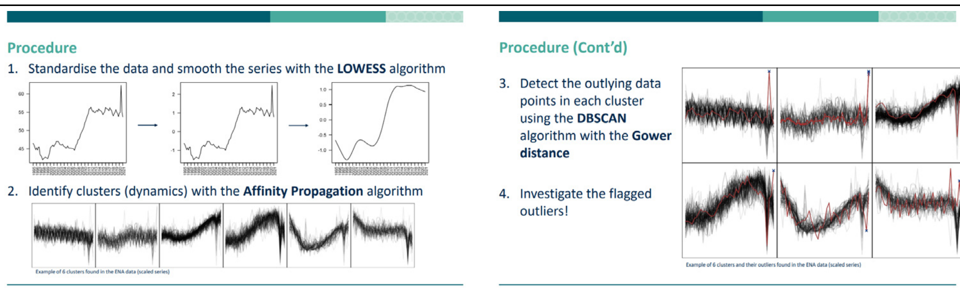
Another example is the BIS initiative to develop a **highly automated validation workflow relying on ML** tools and enhanced computer capacity. The data validation approach implemented is reported to be suitable for a large volume of indicators – about 3,000 daily time series of financial market data. It therefore clearly outperforms more traditional methods, such as graphical controls or threshold-based warnings. Moreover, the new solution appears better able to address the risk of errors that are "Type I" or "false positive" (ie mistaken rejection of the existence of an anomaly that in fact exists) as well as "Type II" or "false negative" (ie failure to reject the existence of an anomaly while the data are in fact correct). The workflow starts with the checking of potential data gaps, followed by a reduction in the dimensionality of the problem (by concentrating on a smaller set of series), and the use of a long short-term memory (LSTM) artificial neural network[2] that can process entire sequences of data (and not only single data points) to estimate prediction errors and detect possible anomalies.

In addition to facilitating the handling of a large number of series, the use of ML techniques can help to better deal with the fact that macroeconomic time series are often subject to sudden and unexpected shocks (eg the Covid-19 pandemic). **These changes imply that data quality monitoring procedures can be constantly challenged** as time passes. To address this issue, the approach developed jointly at the Bank of England and the ECB is based on a clustering procedure in order to automatically identify anomalies within an evolving database. This is done by analysing the correlations within the observations, representing 6,638 single time series from 31 countries in that particular case. The process involves the standardisation of the data, the smoothing of the series with a specific filter (the LOWESS algorithm), the identification of specific clusters using a dedicated ML algorithm (Affinity Propagation (AP); Frey and Dueck (2007)), and the detection of potential anomalies within each cluster through an algorithm grouping together similar observations – with the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) technique based on a specific metric that measures dissimilarities, the "Gower distance" (Graph 1). The method is data-driven, appears more robust to systemic shocks and allows for a high level of automation.

2 Neural networks (or "artificial neural networks") are ML algorithms that transmit a signal from one processing node to another (loosely analogous to the interactions between brain neurons) to identify non-linear relationships in the data. Such models are deemed capable of capturing and representing complex relationships (Richardson et al (2019)).

Procedure to detect anomalies developed at the Bank of England and the ECB

Graph 1



Source: A Maurin and N Benatti, "Time series outlier detection, a data-driven approach", *IFC Bulletin*, no 57, November 2022.

Outlier detection tools

One important focus point of the new quality approaches that are leveraging ML techniques is to **detect outliers in the vast and increasing amount of observations now collected in real time** by big data repositories, which are making traditional manual actions performed by humans (eg use of spreadsheets and simple graphical tools) increasingly inefficient if not impossible to perform. To address these issues, the Bank of Israel has developed dashboards using a specific package (R Shiny app, which uses the R programming language for statistical computing and graphics) that can check all the daily transactions in derivatives markets reported by financial institutions. Once the data are uploaded into a dashboard, the users can choose the variables to analyse, add filters, explore the data graphically and practice specific outlier detection algorithms – including detection graphical tools (eg Bootlier Plot) based on density histograms, isolation forest, etc.

The Deutsche Bundesbank has also adopted **unsupervised ML algorithms to detect outliers for a wide range of voluminous financial data sets** – eg on interest rates, money market statistics, sectoral securities holdings, investment fund holdings that differ markedly in terms of size (from 25,000 to 5 million rows), features (from 12 to 150) and number of outliers (from 0.04% to 5%). The approach relies on various ML algorithms to group information in specific clusters (eg tree-based methods like the isolation forest), assess dissimilarities (eg distance/density-based classification methods like the K-nearest neighbours (KNN) algorithm), compressing the information to be analysed (eg reduction in the size of input data that can be reconstructed afterwards with greater details through the use of self-supervised ML tools such as autoencoders),[3] and generate explanations so that humans can understand the decisions or predictions made (eg use of explainable AI/ML (XAI/XML) techniques like the Shapley value approach).

[3] An autoencoder is a type of neural network that learns the main features of the input information by constructing a lower-dimensional representation of it (similar to transforming an original photo into a lower-resolution version) and then reconstructing it (Rubio et al (2020)). The objective is to facilitate analytical and computing work that is easier to conduct when the dimensions are small.

Imputing missing values and interpolating data series

Reflecting the importance put by central banks on having enough data at their disposal to support their decision-making processes, **an important stream of work is to augment the information available, especially in the case of missing data points or when the data are not timely enough and/or not available with sufficient frequency**.

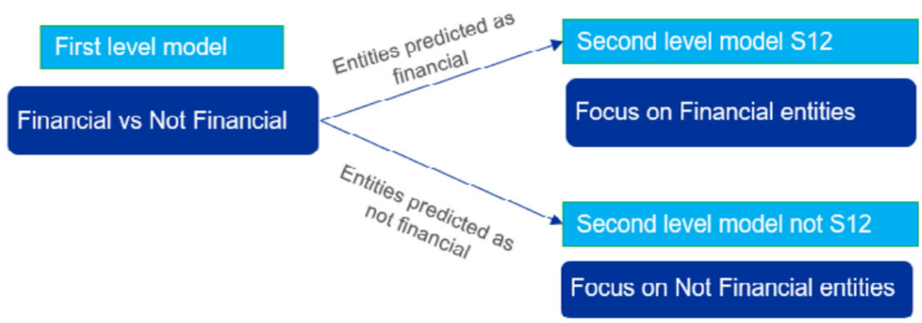
The problem posed by missing values may arise for various reasons, eg the information had not been reported or was collected with significant quality problems and had to be disregarded. This can create serious challenges, for instance in terms of the reduction of the sample of the data available or the introduction of potential biases, in turn possibly undermining the validity of the information and hence the relevance of the actions taken on its basis. While many statistical methods have been traditionally mobilised to address these issues, **ML approaches have become more popular ways to facilitate the imputation of missing data points** (cf the review of ML-based data augmentation and related management methods by Kumar et al (2017)). Frequently used techniques include random forest–based learning methods (Stekhoven and Bühlmann (2012), Rahman and Islam (2013), Tang and Ishwaran (2017) and Ramosaj and Pauly (2019)); the automatic discovery of regular data patterns through "generative deep learning" methods (cf Yoon et al (2018), Nazábal et al (2020), Hou et al (2022) and Qian et al (2022)); approaches that make use of observed differences between specific parts in the data, eg those using "discriminative deep learning" methods (Biessmann et al (2018)); and algorithms for forecasting time series such as the Deep Autoregressive model (DeepAR) (Salinas et al (2020)). These examples represent a small number of ML-based imputation methods available, and Jäger et al (2021) provide a useful review of their performance by looking at a large number of data sets under realistic conditions in terms of missing values.

A similar problem occurs with data sets that have a low frequency (eg annual or semiannual) and are typically available with too-long time lags. As was clearly obvious when the Covid-19 pandemic struck, the usefulness of such information is limited for those central banks willing to take decisions on a timelier and/or more frequent basis. Here also, **ML techniques can be particularly useful in mitigating these problems by helping to interpolate low-frequency series into higher-frequency ones or by speeding up their release using additional information**. For instance, the Central Bank of the Russian Federation (CBRF) has developed specific tools to facilitate the quarterly compilation of financial accounts and balance sheets in the System of National Accounts despite the fact that certain non-bank financial firms report their financial statements at an annual frequency only. Traditional interpolation methods were compared with various ML-based techniques, namely random forest, that basically relates to classification algorithms (Breiman (2001)), gradient boosting trees decision models that are used in regression and classification tasks (Friedman (2002)), and neural network-based tools to generate new data consistent with the information observed (eg the Wasserstein generative adversarial network or "GAN" approach (Arjovsky et al (2017)). The first two types of techniques were fine-tuned on annual values and their lags and used to estimate quarterly figures. The third involved a learning phase based on the data patterns observed for those companies producing quarterly statements and the simulation of the corresponding data for those reporting only annual figures. Yet one issue was the fact that these various approaches can lead to considerably different outcomes; moreover, their results are difficult to interpret for users, representing an important drawback compared with more traditional statistical techniques.

Provision of contextual information

ML also supports a richer augmentation of the data sets available, by incorporating complementary public information or records derived from administrative registers. For example, the newly established statistical reporting of the firms' Legal Entity Identifier (LEI) does not comprise information on the institutional sector of these entities (ie whether they are banks, money market funds, insurance firms, households, non-financial corporations etc), which can be important for supervisory monitoring purposes. The ECB has accordingly developed an ML-based way to augment the LEI database to also include estimates of the institutional sector of the reporting entities. The approach uses a random forest classifier technique to, first, separately identify financial companies from all other firms and, second, estimate specific subsectors in these two main groups (Graph 2). This two-level classification technique is initially estimated ("trained') on a specific data set for which the institutional sector is known, and then applied to the observed database for which the information is missing.

Entity sectoral classification model used by the ECB Graph 2



S12 represents financial sector subclassifications of the European System of National and Regional Accounts.

Source: F Benevolo, T Gottron, I Febbo, and N Pegoraro, "Supervised machine learning for estimating the institutional sectors of legal entities on a large scale", *IFC Bulletin*, no 57, November 2022.

In practice, these approaches cannot rely on the simple running of algorithmic techniques and require significant subject matter expertise. The ECB project, for instance, benefited from extensive business area knowledge to detect the presence of specific words in the entity names at stake – eg words similar to "bank" or "manufacturing" had to be selected as relevant by statistical experts and were therefore included in the classification process. Further, an LSTM neural network (cf above) was applied to deal with the names of similar entities that can be expressed in multiple languages. Finally, the models were selected with due consideration of users' preferences; for instance, a key element was their ability to reduce the risk of wrongly classifying a firm as a non-financial company, reflecting the business need to focus on the monitoring of financial entities as a priority.

3. Macroeconomic and financial analytical tasks

With central banks' decisions becoming increasingly based on factual evidence, **a key issue for them is to make sense of the wealth of existing data to derive useful insights on the economic and financial situation** so that proper policies can be conducted. Fortunately, ML techniques can greatly support this task, by: (i) making sense of the economic and financial data available; (ii) facilitating the modelling of the economy; and (iii) supporting forecasting exercises.

Making sense of the data available

Central banks' policy decision-making hinges on thorough, continuous analyses of a large set of variables to estimate the current state and outlook for the economy. Because of their ability to deal rapidly with vast and complex sets of observations, ML techniques can facilitate these analytical tasks.

One example is the ECB project to **explore alternative sources of data to extract useful insights in almost real time**. These new sources have become increasingly relevant with the digitalisation of economic activities, as was particularly evident with the use of online platforms for shopping, trading and entertainment during the Covid-19 pandemic. This project uses credit card data for developing supplementary indicators in partnership with the Fable Data firm,[4] which has specialised in the European alternative market to provide real-time banking and credit card data. And a further source of useful information relates to the development of fintech firms, as it provides further opportunities for exploring and analysing new types of data as a complement to the more "traditional" supervisory reporting exercises organised by financial supervisors and monetary authorities.

These various initiatives have underscored the importance of continuously innovating in order to make progress and in particular: (i) to maximise the use of the data available; (ii) to explore untapped, alternative sources of information; and (iii) to enhance cooperation with the related new private sector entities that are increasingly producing vast amounts of data. Yet a key drawback for central banks is that **large numbers of data points are not sufficient to guarantee the veracity of the indicators compiled**. Indeed, big data sets may present important composition bias, hampering their accuracy (Bender et al (2021), IFC (2017)). For instance, the information collected by one or a few firms may not represent the whole underlying economic and financial reality – not everybody is paying with a credit card, or at least not in all circumstances.

The exploration of untapped alternative information sources can not only help to improve the data available in a specific area but also shed light on phenomena for which reliable data are notoriously difficult to find. A good example relates to how inflation is perceived by households, which can be affected by various psychological factors, may differ markedly from headline inflation figures (cf in Europe with the launch of the euro in the early 2000s), and is difficult to gauge –

[4] See www.fabledata.com/.

typically requiring ad hoc surveys that are complex to set up and depend on the type of reporters questioned (eg the ECB's Survey of Professional Forecasters, the household inflation survey by the French national statistical agency INSEE). To address these issues, the Bank of France has harnessed non-traditional indicators from social networks such as Twitter to estimate inflation perceptions. All the relevant tweets were analysed with a dictionary-based filter – word2vec, a neural network-based NLP algorithm that associates words out of a large corpus of text. This allowed them to be classified into topics, so as to produce a price perception indicator predicated on the difference between the number of inflation- and deflation-related tweets.

Modelling

Turning to **macroeconomic modelling exercises, central banks' experience shows that they can benefit from the availability of unconventional data sources and new ML-based methods** such as deep learning. This lesson is in line with the existing literature, especially when dealing with cases when the data/expertise is limited. For example, Chauvet and Guimarães (2021) have trained a tool on US data and proposed a transfer learning strategy to identify business cycle phases in Brazil and the euro area. One interest of this approach is to make use of the knowledge gained from one region's economic experts and apply it to other geographical areas, for instance in the absence of a well recognised business cycle dating committee.

Another **important ML use case is to agnostically understand what the drivers of macroeconomic variables are** – that is, by following a pure data-driven approach instead of relying on ex ante assumptions. For instance, Kohlscheen (2021, 2022) has applied the random forest technique to analyse the drivers of inflation and in particular the role of financial factors that are typically disregarded in the toolkit of macroeconomic modellers.

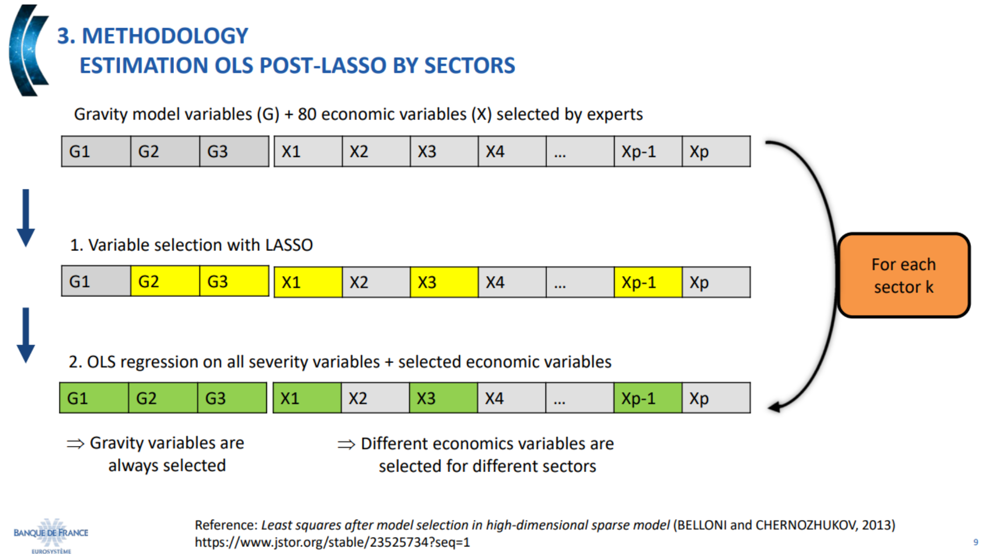
More generally, **ML-based models appear particularly well suited to uncovering explanatory factors from a multitude of candidate variables**. One recent example is the Bank of France project, BIZMAP, to support the internationalisation efforts of French small and medium-sized enterprises (SMEs) (Graph 3). The aim is to make sense of the wide range of publicly available information[5] to help identify attractive EU regions in terms of exports or direct investment. The intelligence behind the tool is programmed as follows: missing data are imputed using ML tools (eg Kalman filter or missForests), relevant variables are selected to explain exports and foreign direct investment, and a gravity trade model is estimated using the least absolute shrinkage and selection operator (Lasso) methodology – a regression analysis method to select more accurate explanatory variables (Tibshirani (1996)). In a similar way, the CBRF has used ML-based methodologies to estimate financial flows in the economy to cope with the fact that a large number of unknown parameters would need to be considered if one followed a more traditional, deductive approach. The project relied on the Variational Bayes (VB) methodology, an ML-based inference technique for making the necessary

5 Eighty-two publicly available variables from seven sources: Eurostat, the Organisation for Economic Co-operation and Development, the World Bank, the ECB, the European Investment Bank, the European Commission and the Centre for Research and Expertise on the World Economy.

approximations and that appears suitable for dealing with large data sets and complex models – both in terms of computational efficiency and estimation precision.

Use of ML to select covariates in the Bank of France's trade model to estimate exports and foreign direct investment

Graph 3



Source: C B L Kerhor, Y Houri, J-N Kien, E Kintzler and L Richardet, "Fostering European SME's internationalization using big data: the BIZMAP application", *IFC Bulletin*, no 57, November 2022.

Reflecting the above factors, **ML, and deep learning techniques in particular, are being increasingly utilised to support macroeconomic modelling exercises.** One example relates to the solving of convex optimisation problems and overcoming the "curse of dimensionality" (Bach (2017)) – that is, the problems faced when coping with an avalanche of data with increasing dimensions, including with respect to the computational efforts required for their processing and analysis. More generally, ML-based techniques are gradually and flexibly used for complex model estimations (Fernández-Villaverde et al (2020a,b), Maliar et al (2021) and Maliar and Maliar (2022)). A promising avenue relates to the area of dynamic stochastic general equilibrium (DSGE) models,[6] as argued by Fernández-Villaverde and Guerrón-Quintana (2020) and illustrated by the Bank of Canada project using deep learning methods to solve a neoclassical growth model. Lastly, neural networks are more and more popular among macroeconomic modelers, spurred by the availability of popular open source libraries, such as PyTorch and TensorFlow (cf above).

Forecasting

Given their growing contributions to economic and financial analysis and the modelling of agents' behaviour, it should not be surprising that **ML is increasingly called upon to support forecasting exercises covering the short-term – ie "nowcasting" exercises that try to predict the very recent past and the**

[6] See Tovar (2008) for a discussion of the main usage of DSGE models by central banks.

present – to the longer-term horizon – including risk scenarios. The focus has been primarily on enhancing the accuracy of "standard" central bank forecasting exercises that typically focus on real GDP and inflation as key variables influencing their policy decisions.

As regards **economic activity**, the Central Bank of Malaysia has shown the relevance of using ML techniques to extract sentiment indicators from newspaper text, which can in turn improve the forecasting accuracy of key macroeconomic indicators, ie GDP growth and its demand side components. The approach relied on building a corpus including over 720,000 business and financial news articles from 16 news portals. Interestingly, the positive results observed prior to the Covid-19 pandemic remained basically valid after this macroeconomic shock. However, the estimates also suggested that ML-based forecasts do not always outperform other models, as this can depend on the variable at stake. For instance, the computed news sentiment was deemed to improve the forecast of private investment compared with the benchmark autoregressive model, but not for the other components of economic activity.

Turning to **inflation**, the CBRF and the Higher School of Economics have analysed the contribution of various ML techniques to forecasts of consumer price inflation (CPI) using real-time versus adjusted data. To this end, a horse race was run among four popular ML algorithms: random forest, gradient boosting, the Bayesian neural network and regularised regression – ie a type of linear regression adapted to deal with a high number of variables to avoid overfitting, such as elastic net (Zou and Hastie (2005)). All of them were found to outperform an autoregressive model, providing further evidence of the usefulness of ML methods in forecasting. However, the selection of the best performing model was different depending on the forecasting horizon. For instance, gradient boosting and neural networks were found to perform better for one-month forecasts, while the elastic net had the top performance at a six-month horizon. Another important lesson was the need to assess the forecasting performance of these different models depending on the vintages of the data considered, for instance by using data available on a real-time basis or after successive statistical revisions.

Lastly, one benefit of ML techniques is to **allow the expansion of forecasting exercises to cover a wider range of potential variables of interest compared with more traditional approaches**. For instance, Bank Indonesia has been using news articles to enhance the forecasting of the situation in the labour market. The approach involved building a statistical index of employment vulnerability, computed from a corpus of around 27,000 monthly news texts covering a period of 23 years and based on NLP techniques. It facilitated the provision of forecasts on the weakening of the labour market and assessment of unemployment risks at a certain horizon and in specific sectors.

4. Monetary policy

As noted above, **ML techniques can be used to enhance the analysis and forecasts of economic output and inflation, two key variables of interest indirectly determining central banks' monetary policy reaction functions** (Taylor (1993)). In addition, these techniques also allow the integration of a much wider set of variables and contribute to a better understanding of the monetary policy decision process itself.

Assessing the influence of a wider range of factors

Monetary policy decisions can deviate from the "pure" influence of macroeconomic developments in terms of output and inflation because of **additional factors**. Yet the relationships involved are typically complex to analyse, not least because of non-linearity (eg the occurrence of an economic shock) and time-dependency issues (eg the different contributions of specific elements between expansionary and recessionary phases), hence representing an important potential use case for ML techniques.

For instance, Bank Indonesia has developed an ML-based approach to better take into consideration the **impact of foreign investors' behaviour (in terms of external capital flows into Indonesian government bonds) on exchange rate developments and, in turn, on monetary policy decisions**. The exercise involved analysing approximately 2,000 variables, derived from private data providers plus a supervisory data set of government bond transactions. Tree-based classification algorithms – the decision tree of Breiman et al (1984), random forest and XGBoost – were first used to select the most meaningful variables and prediction lags. Second, the more limited set of variables and lags obtained was kept to predict individual investors' daily investment amounts, again using a variety of ML techniques – logistic regression; support vector machine (SVM), a supervised learning algorithm used to predict discrete values (Boser et al (1992)); KNN; decision tree; random forest; XGBoost; and LSTM. Third, the LIME algorithm (cf section 2 above) was applied to explain the predictions of the best models for each investor and allow users to check the plausibility of the outcomes. The result showed that bond yields were important predictors of external investment flows, depending on the investor type (eg short-term versus long-term focus). Future work is planned to build similar ML models for analysing the stock and currency markets and to disseminate the results through a dashboard in order to facilitate feeding them into monetary policy operations.

Shedding light on the monetary policy decision process

In addition to facilitating the capture of a wider set of determinants, **ML tools appear to support a better understanding of the monetary policy decision process itself.** A study by the International Monetary Fund, also published as Edison and Carcel (2021), focused on the US monetary policy discussions and used a dedicated NLP technique – the Latent Dirichlet Allocation (LDA) of Blei et al (2003) – to analyse the topics discussed by the Federal Open Market Committee (FOMC) members over 2003–12. The meeting transcripts were divided into about 45,000 text entries, consisting of sentences or paragraphs said by the Governors. The algorithm was implemented with the goal of splitting the whole text data into eight topics: forecasting, economic modelling, statement language, risks, banking, voting decisions, economic activity and communication. This work showed the issues which were most discussed by the FOMC when taking policy decisions. For instance, it was found that discussions on economic modelling predominated during the GFC, but the main topic was on banking in the subsequent periods, and, later on, communication.

Relatedly, the joint work presented by the Swiss National Bank has been relying on a dedicated algorithm to **study the interlinkages between central bank independence and the evolution of inflation**, extending previous work by Baumann et al (2021). The related causal inference question – eg whether independence can help to reduce inflation – has been an issue of much interest but is difficult to answer using standard regression approaches. This provides an opportunity for using ML techniques, which are arguably better suited to dealing with complicated model specifications, non-linear relationships and a large number of potential explanatory variables compared with sample size. In this particular case, the application of the longitudinal targeted maximum likelihood estimation (LTMLE) of Tran et al (2019) showed that the role of central bank independence was complex and could vary depending on the historical path of inflation; it also highlighted the aptitude of ML for dealing with the causal inference problem.

5. Financial microsupervision

The field of financial supervision (principally of banking entities for those central banks tasked with their oversight, but also for other non-bank financial institutions when applicable) has seen a **marked rise in the use of ML, most notably to enable suptech**, that is, the use of new technologies and big data analytics to support supervision (Broeders and Prenio (2018), Beerman et al (2021)). These techniques can support supervisors' efficiency in: (i) covering traditional supervisory tasks (eg quality reporting, anomaly detection, sending of instructions); (ii) facilitating the assessment of micro-level fragilities; and (iii) identifying and tackling new emerging topics, such as climate-related financial risks, vulnerabilities from the Covid-19 pandemic, or the consequence of increased digitisation in finance (eg the development of fintechs).

Enhanced supervisory process

As regards the traditional micro financial supervisory tasks, the deployment of ML can strengthen the flow of information and communication between authorities and the entities they oversee.

On the one hand, the supervisory information flow starts with the reporting of individual records from monitored firms to the authorities. As for macro statistical exercises (cf Section 3), ML can make this reporting more efficient by strengthening the quality of the data in question. For instance, the Bank of Spain has developed, in collaboration with the Knowledge Engineering Institute, an ML-powered tool that imputes missing information and also detects outliers in non-financial firms' accounting statements. Out of 6.2 million statements, this tool facilitated the correction of those with insufficient data quality and those with missing data (0.5 million in both cases). Among the various methodologies tested to detect outliers – eg principal component analysis (PCA), the Mahalanobis distance and KNN – the one selected was a version of isolation forest ("missolation forest"). In parallel, missing values were imputed through regression analysis, so that information absent for one variable of a firm's report could be estimated from the values of its other variables. All in all, this project highlighted the importance of selecting the appropriate features of the model being contemplated, duly considering expert domain knowledge and factoring in the impact of computation costs during the training phase.

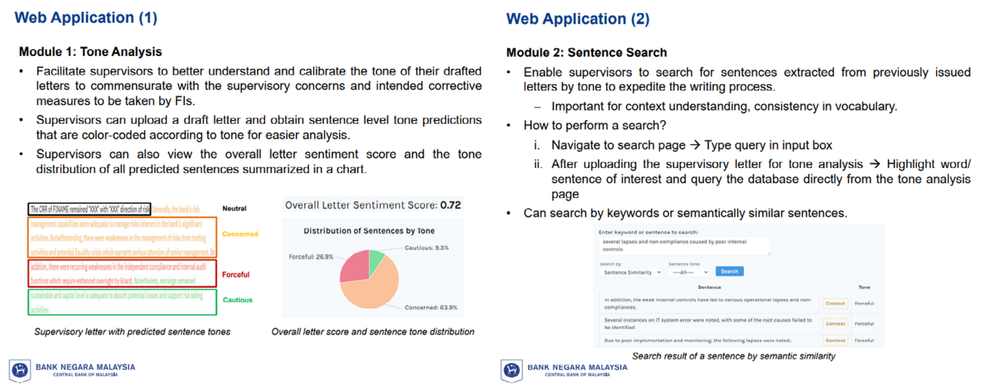
In a similar vein, the Bank of Canada has developed **a novel method based on ML to detect anomalies in data reported by financial institutions**. The objective was to enhance the efficiency and quality of the existing process that deals with millions of data points per month and that can be sensitive to support economic policy. The project relied on a two-step procedure, where financial institutions designated "similar" were clustered in a group and then analysed jointly using a supervised ML algorithm. By overcoming the traditional rule-based approach followed previously, the new method helped to detect anomalies not found before, while saving significant time. Moreover, the procedures were designed to be scalable, fully explainable and able to be run in either a cloud environment or a proprietary data lake.

On the other hand, **supervisory communication also includes the so-called drafting of supervisory letters.** This process is usually time-consuming and requires advanced analytical and communication skills. Moreover, maintaining consistency in the various messages prepared by the whole supervisory team and conveyed to a large number of firms can be challenging and burdensome. In view of these issues, the Central Bank of Malaysia has developed a suptech tool that supports communication with supervised entities, with the aim of enhancing both the efficiency of the process and the consistency of the messages conveyed. This tool had two main functionalities that complement each other: Tone Analysis and Sentence Search (Graph 4). Tone Analysis is based on a text classifier that can characterise any sentence from a supervisory letter as "neutral", "cautious", "concerned" or "forceful". The training data consisted of confidential supervisory letters from 2013 to 2016, yielding 5,000 individual sentences anonymised by a specific tool – ie applying a Named Entity Recognition (NER) algorithm. The process relied on the manual intervention of experienced supervisors and the use of the deep learning model DistilBERT[7] to classify new sentences. The second feature, Sentence Search, is a tool that searched text by keyword or similar semantics. The writing of new supervisory letters with the desired tone was supported by the SentenceBERT tool, which fosters semantic similarity between a reference document and the new draft being queried (by minimising the distance between their representative vectors). Lastly, the project illustrated well the complete workflow supporting the implementation of an ML-based solution, from the development of the model, its deployment in production, and its subsequent retraining for application to a next cycle – for instance, users were able to provide feedback, allowing administrators to refine the model for subsequent estimations.

[7] DistilBERT – a general-purpose language representation model with a compression technique to reduce the number of parameters to be estimated – was the best performing model among the alternatives explored by the authors (including logistic regression, XGBoost etc).

Overview of ML-powered suptech apps used by the Central Bank of Malaysia

Graph 4



Source: J Tan, C K Shum and M A Amri, "Supervisory letter writing app: expediting letter drafting and ensuring tone consistency", *IFC Bulletin*, no 57, November 2022.

Assessing micro-level fragilities

An important building block of the supervisory process is the assessment of micro-level fragilities to **identify the risks faced by a given financial institution and the potential actions warranted to mitigate these**. Given the large amount and complexity of granular data to be digested in this endeavour, ML has proved particularly well suited to addressing these tasks.

For instance, **ML techniques can help to enhance the quality of the firm-level data used in supervisory exercises**, as highlighted in a study by Hitotsubashi University Business School and the Bank of Japan. It compared predictions of the risk of a firm exiting the market in case of insolvency or of a voluntary exit made by humans (professional analysts of a Japanese credit bureau) with those made by ML algorithms (random forest). The results showed that the algorithms outperformed experts' predictions in general, although humans could perform better when assessing firms with less data available – possibly reflecting their greater ability to consider "soft information" compared with automatised algorithms. Hence, one key lesson was that ML techniques cannot fully replace expert judgment but should be used as a useful complement, depending on firm-level characteristics (eg the degree of information available) as well as on users' preferences for minimising the risk of errors of type I versus type II (cf above). Perhaps more importantly, the study also underscored the importance of systematically analysing the accuracy of the new tools in comparison with traditional methods, and in particular of analysing the causes of the respective errors observed.

Moreover, **ML techniques can be applied to enhance the models used for financial stability purposes by incorporating additional sources of information**. For instance, the CBRF has developed ML algorithms (eg logistic regression combined with random forest) that consider additional information on daily payments to improve the traditional default probability models that sit at the core of financial supervisory exercises and are typically based on firm-level accounting data. In that case too, it was found that the degree of accuracy of the respective techniques needed to be carefully analysed, with due consideration in particular of the differences observed across economic sectors.

Dealing with non-supervised entities

While micro supervisory tasks focus on the situation of the specific firms that have to be monitored by the authorities, **it is also important to consider other, less regulated, sectors, not least to prevent regulatory arbitrage** – that is, when non-regulated firms compete in the provision of services that are similar to the ones offered by regulated entities (cf discussion in Fleischer (2010)). On this front too, the use of ML techniques can provide useful insights to supervisors. It can also help supervisors to apply "proportionality" when considering new entrants in the financial system (BCBS and World Bank (2021)).

One telling example relates to the identification of entities involved in fintech, defined as technological innovation used to support or provide financial services (IFC (2020b)). The Bank of France and the Deutsche Bundesbank have created two complementary ML-based tools to identify and monitor these entities. The goal was to overcome the lack of sufficient information available about them due to their fast-paced development and churn. The projects, which are still in their early stages, required mostly public data and were designed to be replicable more broadly in other jurisdictions.

The tool developed by the **Bank of France** focused on classifying whether or not firms are potential fintechs, using publicly available data (covering 84 features) and the isolation forest outlier detection algorithm (Liu et al (2008)). Training and validation were conducted for 10,000 individual non-fintech firms, helping subject matter experts to identify around 350 firms as potential fintechs. The features found most relevant for supporting this identification exercise included newspaper articles about the firms, economic sectors, employees' job titles and the names of senior managers.

Turning to the tool of the **Deutsche Bundesbank**, it only needs an initial list of web addresses (belonging to already identified fintech and non-fintech firms)[8] as input for training and validation: the tool scrapes these websites and creates a graph database consisting of companies, named entities (persons, organisations and locations) and keywords as nodes. In setting up the graph, a large amount of information had to be processed – in this case 515,000 webpages with 1.1 million named entities. According to the location in the graph, a neural network algorithm will decide whether a new and hitherto unclassified company is a fintech or not. This approach appears particularly well suited for dealing with non-structured information.

6. Macro-financial stability policies

Independently of whether the central bank is in charge or not of micro-financial supervision, one of its key policy mandates relates to the macro dimension of financial stability (Crockett (2000)). The impact of the GFC has reinforced interest in developing

[8] The proof of concept uses a data set of 1,190 company web addresses, of which 390 are identified fintechs.

a **system-wide approach to monitoring financial risks, with a dual focus on the situation of different institutions together at a point in time and on the evolution of risks over time as the financial cycle evolves**. This duality calls for collecting and analysing huge amounts of data, covering a wide range of firms and over long periods. Hence, it should not be a surprise that the financial stability function of central banks can benefit from the increased use of ML – cf for instance Fouliard et al (2021), who document how it can enhance the ability to predict crises well before they take place. Two important elements which deserve to be highlighted from this perspective are: (i) the support of ML to match information from various, often unrelated corners that can help to identify system-wide vulnerabilities and their evolution over time; and (ii) the ability to support other policy tasks that are also relevant from a financial stability perspective.

Support of macroprudential exercises

Supporting the macro-financial function requires **collecting trustworthy statistics from various sectors of the economy, hence putting a premium on strong quality assurance processes** for dealing with databases that are not directly produced by the central bank alone. One example is the Bank of Portugal experience with the use of information from the Portuguese credit registry. This source is characterised by an extremely high level of granularity, resulting in a large number of complex observations (over 200 attributes) and calling for strong data quality controls to detect anomalies and identify subtle evolutions. To address the full range of potential anomalies, two automatic filters were created. The first was the Reporting Consistency test, to evaluate if all the financial instruments were reported in a consistent way until their maturity. A second test was the Concentration Check, to check the consistency of the reporting of categorical variables at the agent level. The detection of anomalies was then based on an isolation forest algorithm and was found to have facilitated the detection of reporting gaps, strange data patterns and structural breaks, in turn enhancing the quality of the information available to support macroprudential analyses.

Moreover, **the sheer scale of the detailed data sets of potential interest to policymakers is also a key factor supporting ML-based initiatives to develop a more structural framework**. The reason is that, even abstracting from data quality problems, analytical and computational limitations can prevent the use of these data for effective financial stability monitoring. To address this point, the ECB, Deloitte and Google have jointly developed a solution in the form of a dynamic multilayer network that helps supervisors to look at the available statistics in a comprehensive way and analyse them through various operations supported by data science tools – such as aggregation, filtering and bottom-up analyses from the individual data level. This solution is reported to have facilitated financial stability monitoring tasks in the face of systemic risk events, such as during the Covid-19-induced turmoil in financial markets in March 2020 (FSB (2020)).

The above approaches can be instrumental in **facilitating the analysis of interconnections observed at a given point in time across the various segments of the financial system** and that are a key source of attention for macroprudential authorities. One example relates to the relationships between the banking and housing sectors, as analysed by the Australian service provider Quant Property Solutions. In particular, an important feature mechanism is that lenders willing to foreclose on mortgaged real estate can trigger important developments in housing prices, with possibly severe financial stability implications because of imperfect available information. For instance, foreclosed houses are typically sold below market values, presumably reflecting a specific bias among those market participants willing to sell their collateral, not least because of banks' balance sheet considerations. ML techniques were used to support the market price discovery mechanism, by helping to disentangle the contributions of the multiple factors at play (eg the situation of the bank selling a property, geo-specific real estate features, the economic outlook).

Turning to the time dimension of systemic risk, ML methods can support dealing with large and complicated data sets that change over time. For instance, the CBRF has adopted this type of approach to facilitate work on (changing over time) micro-level databases on banking loans, with several benefits observed in terms of data quality assurance, scalability and automation of the operations, and higher interpretability of the results. Moreover, the solution also allowed for matching the database in question with other sources, namely the Federal Tax and the State Statistics services.

Additional financial stability dimensions

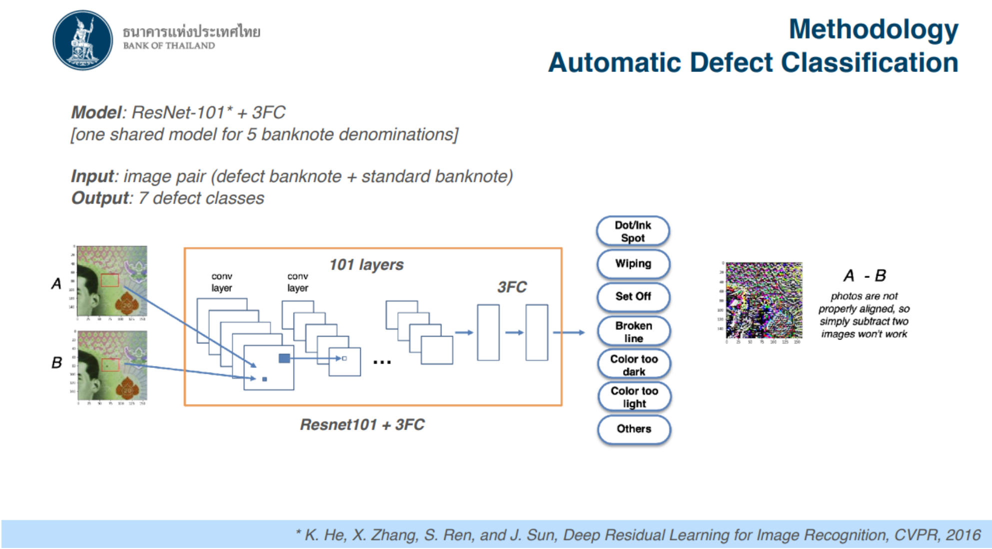
There are additional tasks performed by central banks that are dedicated to the service of society and that, by protecting prosperity and providing financial security and confidence, also play a role in supporting financial stability more generally. **Cases in point relate to financial inclusion, consumer protection and anti-money laundering**, three areas that are reported to have benefited significantly from the development of big data analytics in recent years.

Another important domain is **the safeguarding of the payment system**, whose monitoring sits at the core of a central bank's mandate to both ensure a smooth functioning of payments and prevent its misuse. ML techniques can be instrumental in coping with the large amount of individual transactions involved, as shown by the joint experience reported by the Central Bank of Ecuador in developing neural networks for outlier detection – in that case autoencoders (cf Section 2), with the goal of identifying abnormal transactions that might require closer scrutiny by the payment system's oversight team. As argued by Rubio et al (2020), this application has been able to identify a wide range of payment transaction anomalies. Their findings confirm the experience of previous similar projects, for instance at the Netherlands Bank (Triepels et al (2017)).

A final area where ML-based anomaly detection and classification techniques can support central banks' operations relates to their core mandate of printing money. Occasionally, some banknotes are produced with defects, which can happen at different steps of the production process. While the problematic notes are typically detected by cash machines, analysing the defects to identify their causes can be a laborious, time-consuming and repetitive task. To cope with these challenges, one can usefully deploy ML techniques in line with the example of the Bank of Thailand. This institution has implemented a convolutional neural network-based tool (ResNet-101), which is a type of artificial neural network commonly applied to analyse images (Graph 5). The experience so far is that the number of misprinted notes has fallen by more than half in Thailand.

Bank of Thailand’s computer vision model for detecting banknote defects

Graph 5



3FC refers to the three fully connected neural layers.

Source: J Kerdsri and P Treeratpituk, “Using deep learning technique to automate banknote defect classification”, *IFC Bulletin*, no 57, November 2022.

References

Abadi, M, A Agarwal, P Barham, E Brevdo, Z Chen, C Citro, G S Corrado, A Davis, J Dean, M Devin, S Ghemafwat, I Goodfellow, A Harp, G Irving, M Isard, R Jozefowicz, Y Jia, L Kaiser, M Kudlur, J Levenberg, D Mané, M Schuster, R Monga, S Moore, D Murray, C Olah, J Shlens, B Steiner, I Sutskever, K Talwar, P Tucker, V Vanhoucke, V Vasudevan, F Viégas, O Vinyals, P Warden, M Wattenberg, M Wicke, Y Yu and X Zheng (2015): "TensorFlow: Large-scale machine learning on heterogeneous systems", static.googleusercontent.com/media/research.google.com/en//pubs/archive/45166.pdf.

Ahrens, M and M McMahon (2021): "Extracting economic signals from central bank speeches", *Proceedings of the Third Workshop on Economics and Natural Language Processing*, pp 93–114.

Apel, M, M Grimaldi and I Hull (2021): "How much information do Monetary Policy Committees disclose? Evidence from the FOMC's minutes and transcripts", *Journal of Money, Credit and Banking*, doi.org/10.1111/jmcb.12885.

Arjovsky, M, S Chintala and L Bottou (2017): "Wasserstein generative adversarial network", *Proceedings of the 34th International Conference on Machine* Learning, PMLR no 70, pp 214–23.

Bach, F (2017): "Breaking the curse of dimensionality with convex neural networks", *Journal of Machine Learning Research*, no 18, pp 1–53.

Basel Committee on Banking Supervision (BCBS) and World Bank (2021): *Proportionality in bank regulation and supervision – a joint global survey*.

Baumann, P, E Rossi and A Volkmann (2021): "What drives inflation and how? Evidence from additive models selected by cAIC", *Swiss National Bank Working Papers*, vol 12.

Beerman, K, J Prenio and R Zamil (2021): "Suptech tools for prudential supervision and their use during the pandemic", *FSI Insights on policy implementation*, no 37.

Bender, E, T Gebru, A McMillan-Major and M Mitchell (2021): "On the dangers of stochastic parrots: can language models be too big?", *Proceedings of the 2021 Association for Computing Machinery Conference on Fairness, Accountability, and Transparency*.

Biessmann, F, D Salinas, S Schelter, P Schmidt, and D Lange (2018): "Deep learning for missing value imputation in tables with non-numerical data", *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*.

Blei, D, A Ng and M Jordan (2003): "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, no 3, pp 993–1022.

Boser, B, I Guyon and V Vapnik (1992): "A training algorithm for optimal margin classifiers", *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pp 144–52.

Breiman, L (2001): "Random forests", *Machine Learning*, no 45, pp 5–32.

Breiman, L, J Friedman, R Olshen and C Stone (1984): "Classification and regression trees", *Wadsworth Advanced Books and Software*.

Broeders, D and J Prenio (2018): "Innovative technology in financial supervision (suptech): the experience of early users", *FSI Insights on policy implementation*, no 9.

Bruno, G and J Marcucci (2021): "Data science and machine learning for a data-driven central bank", in P Nymand-Andersen (ed), *Data science in economics and finance for decision makers*, Chapter 10.

Casanova, C, B Hardy and M Onen (2021): "Covid-19 policy measures to support bank lending", *BIS Quarterly Review*, September, pp 45–59.

Chakraborty, C and A Joseph (2017): "Machine learning at central banks", *Bank of England Working Paper*, no 674, www.bankofengland.co.uk/working-paper/2017/machine-learning-at-central-banks.

Chang, A X and C Manning (2012): "SUTime: A library for recognizing and normalizing time expressions", *8th International Conference on Language Resources and Evaluation*.

Chauvet, M and R S Guimarães (2021): "Transfer learning for business cycle identification", *Banco Central do Brasil Working Paper*, no 545.

Chen, T and C Guestrin (2016): "XGBoost: a scalable tree boosting system", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp 785–94.

Cieslak, A and A Schrimpf (2019): "Non-monetary news in central bank communication", *Journal of International Economics*, no 118, pp 293–315.

Crockett, A (2000): "Marrying the micro- and macro-prudential dimensions of financial stability", remarks before the Eleventh International Conference of Banking Supervisors, Basel, 20–21 September.

Doerr, S, L Gambacorta and J M Serena (2021): "Big data and machine learning in central banking", *BIS Working Papers*, no 930.

Edison, H and H Carcel (2021): "Text data analysis using Latent Dirichlet Allocation: an application to FOMC transcripts", *Applied Economics Letters,* vol 28, no 1, pp 38–42.

Fernández-Villaverde, J and P A Guerrón-Quintana (2020): "Estimating DSGE models: recent advances and future challenges", *NBER Working Paper*, no 27715, August.

Fernández-Villaverde, J, S Hurtado and G Nuno (2020a): "Financial frictions and the wealth distribution", *Banco de España Working Paper,* no 2013.

Fernández-Villaverde, J, G Nuno , G Sorg-Langhans and M Vogler (2020b): "Solving high-dimensional dynamic programming problems using deep learning", mimeo.

Ferreira, L N (2021): "Forecasting with VAR-teXt and DFM-teXt models: exploring the predictive power of central bank communication", *Banco Central do Brasil Working Paper*, no 559.

Financial Stability Board (FSB) (2017): *Artificial intelligence and machine learning in financial services: market developments and financial stability implications*, www.fsb.org/wp-content/uploads/P011117.pdf.

——— (2020): *Holistic Review of the March Market Turmoil and COVID-19 pandemic: Financial stability impact and policy responses*, November.

Fleischer, V (2010): "Regulatory arbitrage", *Texas Law Review,* vol 89, no 7.

Fouliard, J, M Howell and H Rey (2021): "Answering the Queen: machine learning and financial crises", *National Bureau of Economic Research Working Paper,* no 28302.

Frey, B J and D Dueck (2007): "Clustering by passing messages between data points", *Science*, vol 315, no 5814, pp 972–76.

Friedman, J H (2002): "Stochastic gradient boosting", *Computational Statistics & Data Analysis*, vol 38, no 4, pp 367–78.

Gentzkow, M, B Kelly and M Taddy (2019): "Text as data", *Journal of Economic Literature*, vol 57, no 3, pp 535–74.

Gros, D (2018): "When communication becomes the policy", *Monetary dialogue*, European Parliament, September.

Hansen, S and M McMahon (2018): "How central bank communication generates market news", *Handbook of Macroeconomics*, Chapter 15.

Hansen, N, M McMahon and A Prat (2018): "Transparency and deliberation within the FOMC: A computational linguistics approach", *Quarterly Journal of Economics*, vol 133, no 2, pp 801–70.

Hernández de Cos, P (2022): "Old risks, news challenges, same objective: the work programme of the Basel Committee in 2022", speech, www.bis.org/speeches/sp220225.htm.

Hou, J, H Jiang, C Wan, L Yi, S Gao, Y Ding and S Xue (2022): "Deep learning and data augmentation based data imputation for structural health monitoring system in multi-sensor damaged state", *Measurement*, vol 196, 111206, doi.org/10.1016/j.measurement.2022.111206.

Irving Fisher Committee (IFC) (2015): "Central banks' use of and interest in 'big data", *IFC Report*, no 3, October.

——— (2017): "Big data", *IFC Bulletin*, no 44, March.

——— (2018): "IFC report on central banks and trade repositories derivatives data", *IFC Report*, no 7, October.

——— (2019): "The use of big data analytics and artificial intelligence in central banking", *IFC Bulletin*, no 50.

——— (2020a): "Computing platforms for big data analytics and artificial intelligence", *IFC Report*, no 11.

——— (2020b): "Towards monitoring financial innovation in central bank statistics", *IFC Report*, no 12.

Israël, J-M and B Tissot (2021): "Incorporating micro data into macro policy decision-making", *IFC Bulletin*, no 53.

Jäger, S, Allhorn A and F Biessmann (2021): "A benchmark for data imputation methods", *Front. Big Data*, 4:693674.

Kohlscheen, E (2021): "What does machine learning say about the drivers of inflation?", *BIS Working Papers*, no 980.

Kohlscheen, E (2022): "Quantifying the role of interest rates, the dollar and Covid in oil prices", *BIS Working Papers,* no 1040.

Kumar, A, M Boehm and J Yang (2017): "Data management in machine learning: Challenges, techniques, and systems", *Proceedings of the 2017 ACM International Conference on Management of Data*, pp 1717–22.

Liu, F, K Ting and Z Zhou (2008): "Isolation forest", *Eighth IEEE International Conference on Data Mining*.

Maliar, L, S Maliar and P Winant (2021): "Deep learning for solving dynamic economic models", *Journal of Monetary Economics*, vol 122, pp 76–101.

Maliar, L and S Maliar (2022): "Deep learning classification: modelling discrete labor choice", *Journal of Economic Dynamics and Control*, vol 135.

Nazábal, A, P Olmos, Z Ghahramani and I Valera (2020): "Handling incomplete heterogeneous data using VAEs", *Pattern Recognition*, vol 107:107501.

Paszke, A, S Gross, F Massa, A Lerer, J Bradbury, G Chanan, T Killeen, Z Lin, N Gimelshein, L Antiga, A Desmaison, A Kopf, E Yang, Z DeVito, M Raison, A Tejani, S Chilamkurthy, B Steiner, L Fang, J Bai and S Chintala (2019): "PyTorch: an imperative style, high-performance deep learning library", *NeurIPS proceedings*, pp 8024–35, papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

Pedregosa, F, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, P Prettenhofer, R Weiss, V Dubourg, J Vanderplas, A Passos, D Cournapeau, M Brucher, M Perrot and E Duchesnay (2011): "Scikit-learn: machine learning in Python", *Journal of Machine Learning Research*, no 12, pp 2825–30.

Qian, Y, L Tian, B Zhai, S Zhang and R Wu (2022): "Informer-WGAN: high missing rate time series imputation based on adversarial training and a self-attention mechanism", *Algorithms*, vol 15, no 252, www.doi.org/10.3390/a15070252.

Rahman, M G and M Islam (2013): "Missing value imputation using decision trees and decision forests by splitting and merging records: Two novel techniques", *Knowledge-Based Systems*, vol 53, pp 51–65.

Ramosaj, B and M Pauly (2019): "Predicting missing values: A comparative study on non-parametric approaches for imputation", *Comput Stat*, vol 34, no 4, pp 1741–64.

Ribeiro, M, S Singh and C Guestrin (2016): "'Why should I trust you?' Explaining the iredictions of any classifier", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp. 1135–44.

Richardson, A, T van Florenstein Mulder and T Vehbi (2019): "Nowcasting New Zealand GDP using machine learning algorithms", *IFC Bulletin*, no 50, May.

Rubio, J, P Barucca, G Gage, J Arroyo and R Morales-Resendiz (2020): "Classifying payment patterns with artificial neural networks: An autoencoder approach", *Latin American Journal of Central Banking*, vol 1, no1.

Salinas, D, V Flunkert, J Gasthaus and T Januschowski (2020): "DeepAR: Probabilistic forecasting with autoregressive recurrent networks", *International Journal of Forecasting*, 36, pp 1181–91.

Stekhoven, D J and P Bühlmann (2012): "MissForest – non-parametric missing value imputation for mixed-type data", *Bioinformatics*, no 28, pp 112–8. doi:10.1093/bioinformatics/btr597.

Tang, F and H Ishwaran (2017): "Random Forest missing data algorithms", *Stat Analysis Data Mining*, vol 10, no 6, pp 363–77.

Taylor, J (1993): "Discretion versus policy ules in practice", *Carnegie-Rochester Conference Series on Public Policy*, vol 39.

Tibshirani, R (1996): "Regression shrinkage and selection via the Lasso", *Journal of the Royal Statistical Society*, Series B (Methodological), vol 58, no 1, pp 267–88.

Tovar, C (2008): "DSGE models and central banks", *BIS Working Papers*, no 258, September.

Tran, L, C Yiannoutsos, K Wools-Kaloustian, A Siika, M Van Der Laan, M Petersen (2019): "Double robust efficient estimators of longitudinal treatment effects: Comparative performance in simulations and a case study", *The International Journal of Biostatistics*, vol 15, no 2.

Triepels, R, H Daniels and R Heijmans (2017): "Anomaly detection in real-time gross settlement systems", *ICEIS*, vol 1 .

Wibisono, O, H Ari, A Widjanarti, A Zulen and B Tissot (2019): "Using big data analytics and artificial intelligence: a central banking perspective", in "Data Analytics", *Capco Institute Journal of Financial Transformation*, 50th edition, pp 70–83.

Yoon, J, J Jordon and M van der Schaar (2018): "GAIN: missing data imputation using generative adversarial nets," *Proceedings of the 35th International Conference on Machine Learning*, pp 5675–84.

Zou, H and T Hastie (2005): "Regularisation and variable selection via the elastic net", *Journal of the Royal Statistical Society*, Series B (Methodological), vol 67, no 2, pp 301–20.