TAKING LIQUIDITY FORECASTING TO THE NEXT LEVEL WITH ARTIFICIAL INTELLIGENCE AND DATA ANALYTICS
Modern technologies such as AI and ML, along with the now widespread use of decision-making systems, pattern recognition and chatbots, could be used in financial departments for the task of analyzing and predicting liquidity levels with the goal of reducing the liquidity reserve buffer and using funds more efficiently.

As in any AI/ML application, data preparation and prior automation of data aggregation tasks are essential for successful implementation and efficiency gains.

In this paper, we explore applied examples of liquidity management with AI in banking and corporate treasuries.

The liquidity forecast considered in this paper can be improved step by step by increasing the level of detail and granularity of the input data and adding macroeconomic factors to the analysis.
INTRODUCTION

The extraordinary circumstances of 2020-2022 caused by the pandemic have prompted banks and corporate treasuries across the globe to re-evaluate their liquidity management practices. The Swiss and European treasurers have faced multiple challenges — increasingly demanding regulatory requirements, volatile markets, growing counterparty risk, and negative interest rates, to name a few.

At the same time, we are clearly seeing the trend of digitalization along with a widespread use of artificial intelligence (AI) and various kinds of intelligent automation methods. Many banks and corporations already actively use these in their daily practice. Some of the most prominent examples are, for example, AI-based loan decisions, chatbots, suspicious transaction tracking, etc. Thus, in the context of liquidity forecasting, the following question arises: can AI also be useful for the task of analyzing and predicting liquidity and related ratios?

When speaking of predicting future values of ratios or KPIs, we mean dealing with a time series, for example the development of liquidity coverage ratio over time. To process time series data, statistical/econometric methods are commonly used. Today, many discussions are taking place about the capabilities of AI and machine learning (ML) methods in forecasting. Can advanced AI algorithms do a better job than current statistical methods? To answer this question, we must consider the following:

1. **Extreme dynamics of series.** Unpredictable changes and leaps in financial series are a new reality (black swan\(^1\) effect).

2. **Statistical methods are based on rigid assumptions** about properties of time series and often do not reflect such characteristics of financial data as non-linearity, randomness, irregular periodicity. However, these methods are not outdated and can still be successfully used when chosen appropriately.

3. **AI/ML methods do not replace econometric theory,** they guide it and overcome some of its limitations. The most important property of AI/ML-based systems is the fact that **the particular model is fitted to the input data** and as new values of the series appear, it **automatically recognizes changes**, learns from them and **adjusts the model** if necessary. Econometrics may be good enough to succeed in the financial sector for now but succeeding in business in the future would require machine learning (ML).

Before diving into the specifics of AI and its application for modeling/forecasting liquidity metrics, it is worth remembering that legacy-dominated CFO and treasury offices pose a real challenge to the implementation of intelligent automation and that solid data foundations are required to leverage the AI’s full potential. More specifically, automation across the data journey in any treasury or finance application is pivotal. In this paper, we consider the fundamental aspects of using AI for liquidity forecasting, in particular:

- Approaches for modeling financial ratios and deriving liquidity and capital risks, including the main AI methods that can be used for this class of tasks
- Required steps towards intelligent automation in the treasury department
- Examples of application for banks and corporate treasurers
1. **AI MAIN CONCEPTS AND REQUIREMENTS**

Today, everyone is talking about artificial intelligence and machine learning, but there is still some misunderstanding about these concepts. What is the difference between AI and ML? We should consider **AI as an umbrella term** describing the science of how to create intelligent programs that can imitate human behaviour. **ML is a subset of AI methods** that can automatically learn and improve from experience without explicit programming. There are multiple ML techniques which can be used for building models for generating predictions in treasury as well as for building sustainable views of real time liquidity management:

1. **Supervised** – the algorithm is fed training data that has labeled tags characterizing the data. The algorithm then learns a general rule of classification (for discrete data: “Yes”/“No”, “Low”/“Medium”/“High”) or regression (for continuous data: price, temperature) to predict the labels/values for the remaining observations in the data set. These methods are often used to predict future values of financial indicators.

2. **Unsupervised** refers to the situation where the data has no labels, and no training data is provided. The algorithm then detects patterns in the data by identifying clusters of observations that depend on similar characteristics. As a rule, these methods are often used for clustering, detecting anomalies and trends.

3. **Reinforced** learning falls in between supervised and unsupervised. In this case, the algorithm is fed an unlabelled set of data, it then chooses an action for each data point, and receives feedback (“supervision”, perhaps from a human) that helps the algorithm learn and take a new action.

4. **Deep learning (DL)** algorithms try to emulate the human brain and rely on neural networks. They can be used mostly for complex problems when data is very diverse, unstructured, and inter-connected.

Overall, these methods promise faster, more holistic, and more connected insights, compared with traditional statistics techniques. They try to make sense of the data ingested and extract patterns from the data to improve financial forecasting, predict supply/demand and ultimately improve business performance.
As a subfield of AI, hundreds of ML algorithms have been developed over the past few decades, and each can be used to directly or indirectly generate new business knowledge. However, regardless of the business application chosen, any generic AI/ML model consists of five steps, which are independent of the choice of algorithm. Figure 2 underlines how fundamental the data provisioning and feature selection are to generating business insights.

First, any AI/ML application starts with data pre-processing and data automation; this step is critical as it ensures the data used for modeling purposes is representative, of high quality and governed. Extricating, changing, stacking and further cleaning of the data represents around 80 percent of the time for an AI undertaking. Second, decisions about feature selection and the choice of algorithm are made. This entails translating the business problem in datasets/samples and the definition of the variables of interest. Then, we proceed with training of the algorithm, which means fitting it with the data ingested and fine-tuning the parameters. Lastly, to make sure that the business insights and the AI/ML process are successful, performance is evaluated using chosen metrics. To apply an ML model, there are common requirements and a common denominator that need to be fulfilled:

1. Availability of the data (labeled, categorized in the case of supervised learning) as ML models are not programmed but “trained”

2. Quality of the data (correctness, completeness, timeliness of the data provisioned)

3. Representativeness of the testing/production data (in order for AI/ML models to generate meaningful outputs, the input data needs to be representative of the actual data)

4. Computing power depending on the AI/ML method chosen.

Indeed, in terms of a business-relevant application, investments in core tech will become critical to fully leverage the potential of AI and meet increasing demands for scalability, flexibility, and speed. Additionally, the choice of use cases needs to be business-relevant and add value from a business perspective, be it due to the scalability of the platform/tooling chosen, the improved data strategy or the actual forecasting benefits. This will be further described in a later section.

Figure 2: Five steps present in any generic ML model
### 3. AI/ML USE CASES IN TREASURY AT BANKS AND CORPORATES

In order to reduce variance and forecasting errors, AI/ML can be used (1) to better understand cash inflow and outflows with ML algorithms that can better mimic human behavior (sequential memory) and (2) to increase confidence levels due to more precise input values. Below are some use cases applicable in treasury:

1. **Variance analysis**
   - Understanding variance from expected mean e.g., deviations from “normal” amounts
   - Identifying and investigating data outliers

2. **Cash flow pattern recognition**
   - Understanding gaps in payments or cashflows, detecting anomalies, deductions behaviours, identifying group of flows with similar rules

3. **Correlation analysis**
   - Understanding correlations between indicators/ seasonality/ macroeconomics/ interest rates

4. **Automated benchmarking within competitor peer groups**

In the following section, we will go through the application of two treasury examples, one for banks and one for corporates, using the use cases described above.

### 3.1. Liquidity management example 1: Application for banks

The financial crisis of 2008 brought serious adjustments to bank liquidity requirements around the world. The Basel Directive III requires banks to provide an adequate and sufficient level of short-term liquidity, regulated by the Liquidity Coverage Ratio (LCR) as defined in equation 1.

The minimum value of the ratio is 100 percent. However, this does not mean that the bank should strive to maximize this value. The problem is that the higher this indicator, the more highly liquid assets the bank needs to keep on its accounts. Thus, the possibility of profitable investments of free cash is limited. Given the negative interest rates in the Euro and Swiss Franc zones, banks find themselves in a very difficult environment for generating profits. This means that the main task of the treasury is to determine and be able to predict the level of the ratio accurately enough, keeping it at a minimum, but sufficient level. The conservative approach to forecasting the LCR ratio means that the bank makes a certain necessary “reserve” of highly liquid assets in case of an unfavourable situation, thus, depriving itself of possibilities to earn extra income (for example, by investing excess liquidity in short-term transactions). A riskier approach means higher profit for the bank from investments, but, as the name suggests, it is also riskier. In this case, the bank must determine its risk appetite and formulate a buffer position accordingly.

There are many variants of using AI, which allow improving quality of forecasting and thereby enable the bank to use its liquid assets more efficiently and in parallel maintain their regulatory mandated LCR ratio. AI/ML methods can be used on multiple levels and together can significantly improve LCR prediction. For the purpose of this paper, we distinguish between three levels of forecast improvements using AI (see Figure 3):

### Equation 1: LCR calculation

\[
LCR = \frac{\text{High quality liquid asset amount (HQLA)}}{\text{Total net cash flow amount}}
\]
1. Evaluation and forecasting of the indicator based on seasonal trends and search for anomalies

2. Analysis of correlations with macroeconomic indicators and their integration into forecasting

3. Analysis of data at the level of transactional and client clusters – finding correlations, trends and anomalies at a more granular level.

Below we look at the different levels of implementation and what the analysis involves.

3.1.1. Searching for seasonal fluctuations and anomalies

The analysis of seasonal deviations is not a new topic. Speaking of retail banks, for example, it is fair to agree that people tend to spend a lot of money at Christmas, which directly affects the outflow of funds from bank accounts, and therefore a decrease in the denominator of LCR (Total Net CF Amount). Another example is the payroll in the last week of each month or bonus payments, and so forth. To find seasonal trends, classic statistical methods can be used, as well as modern methods of ML. The task of finding seasonal fluctuations is reduced to the task of pattern recognition. Patterns can have cyclic, seasonal and trend components. The choice of an appropriate model depends on the nature of the statistical probability distribution of the data (symmetry, skewness, kurtosis, etc.). One of the most common examples is the use of regression methods. A regression as a type of supervised ML technique is a function that describes the relationship between one or more independent variables and the target variable (see Figure 4). In other words, a regression allows to understand the time series behaviour and predict continuous outcomes. Different regression-based methods can be used, depending on the complexity of the data representation and the level of forecast accuracy required. Three classes of methods are distinguished in the literature:

1. Methods where the class of models considered is linear in the features (Regularized Linear Regression: Lasso, Ridge, and Elastic Nets)

2. Non-linear regression

3. Deep learning

Another approach to LCR analysis is looking for anomalies, e.g., extremely high liquidity outflows. Anomalies are outlier data points – that means there are very infrequent occurrences in the data sample observed, as shown in figure 5. In univariate cases, outliers or rare events can be easily spotted by the naked eye during visual analysis. However, in bidimensional cases, it is quite difficult to detangle the data as it is the combination between X and Y that identifies outliers. In those cases, AI/ML classification methods can be leveraged to compute the probability distribution of the data points and compare each new observation to an anomaly threshold. Classification techniques as a type of supervised ML in pattern recognition predict discrete outcomes – is this payment normal or abnormal?

**Recommended AI/ML methods for this use case:**
Regression methods such as Lasso/Ridge, nearest-neighbour classifiers, decision trees, neural networks.

3.1.2. Correlations with macroeconomic indicators

As discussed above, the biggest advantage of AI/ML methods is their capacity to learn from large amounts of data. This characteristic can be used to improve current metrics forecasting methods by increasing the data span. A particular trick to optimize forecast data is to analyze the correlation between the behaviour of certain customer groups (clusters) and changes in macroeconomic indicators (interest rates, FX rates, GDP, and other financial instruments). If we know that some variables are closely correlated, then it is possible to predict one variable from the other. For example, a change of the central bank interest rate (especially if we are talking about negative rates) is a serious prerequisite for both customers and banks to reconsider how much money should be kept (see Figure 6). Changes in rates in one direction or another tend to correlate closely with net cash inflows of liquidity. In other words, a model trained on historical data can predict cash inflows/outflows based on current interest rate values. There are several supervised and unsupervised methods related to correlation analysis.

**Recommended AI/ML methods for this use case:** Principal component analysis to find condense information in high-dimensional datasets, random forest, decision tree, high-dimensional factor analysis.
3.1.3. Defining customer and transaction clusters

Complicating the task, we can also analyze customer behaviour based on clusters mentioned above (see Figure 7). This problem then is no longer about forecasting values, but about clustering. The idea is to classify existing customers based on them having a certain characteristic (it can be information from account data, age, gender, income level, credit rating, retail or wholesale segment). Clustering is the most common type of unsupervised ML algorithms. Having clusters ready to work, the AI algorithm identifies the behaviour of specific groups of customers rather than the LCR parameter. Thus, knowing, for example, that wholesale customers have certain behavioural patterns (cyclicality), we can tune our prediction in more detail, using the methods described in section 3.1.1.

We can further trace patterns using transactional data, for example, interest payments on loans or payroll. We can also drill down by type of payment (credit card, wire transfer, etc.). Thus, with the help of AI it is possible to predict the actions on a particular bank account.

**Recommended AI/ML methods for this use case:**
3.2. Liquidity management example 2: Application for corporates

For corporate treasuries, as for banks, liquidity management issues are of primary importance. In this matter, firms operate with such terms as “liquidity position” or “free liquidity”, which with slight variations depending on the method of calculation, reflect the level of short-term liquidity of the firm. The formula for Free Liquidity in a simplified form is as follows:

Equation 2: Free liquidity position

\[
\text{Free liquidity} = \text{Freely available Cash} + \text{Unused credit lines}
\]

This key performance indicator (KPI) allows the financial manager to answer such questions as:

- Can the firm make the necessary payments (salaries, dividends)?
- Can the firm make short-term investments of excess cash?
- How can the firm strengthen its balance sheet?

To make these decisions, the estimated cashflow position is crucial. Continuous improvement of the forecast quality is a necessary and important task for the finance department of an enterprise.

As in the LCR banking example, similar basic steps can be distinguished when applying AI methods.

3.2.1. Analyzing cyclicality of cashflow components

The time series of a short-term liquidity position values can be analyzed for cyclicality and anomalies using ML methods. The idea of these methods is described in the section on LCR and can be fully applied to the financial status as a task of cycle and anomaly detection.

The traditional approach to forecasting is based on the analysis of historical data (see Figure 8). It is possible to improve the quality of forecasting by analyzing the components of the cash flow indicator: accounts receivable (AR), cash or unused credit lines, or to go even deeper, the analysis of transactional data. This will help improve the quality of the forecast (see Figure 9). For example, to analyze and forecast target revenue for the coming year and its distribution by months. In combination with an analysis of unused credit lines, potential points for increasing or decreasing the credit limit can be recognized.

Recommended AI/ML methods for this use case:

Regularized regression such as Lasso/Ridge, random forest that partitions the feature space, boosting methods.

3.2.2. Defining clusters in payments flows

Another approach for refining the forecast is the analysis of inflows. Depending on the month, the inflows ratio changes and has a certain seasonality. For example, January is traditionally a low-margin period, and we can expect that inflow payments will be made later than in other months. This also holds true in certain industries, e.g., hotel during off-peak season, etc.

Therefore, it is important to analyze not only the correlations, but also the cyclic nature of these relationships. Here we can talk about the task of finding clusters and modeling their seasonality (see Figure 10). By clusters we understand the time intervals for which the range of variation of values “invoice amount – days until payment” is within certain ranges. Although clusters have cyclic features, we should remember that recently identified clusters have a greater influence (mathematically speaking — more weight) on the forecast than the clusters identified last year. This is consistent with current market realities: in a pandemic, we cannot fully rely on patterns identified before the pandemic. The increase of bill payments period in the COVID crisis could not be determined on the pre-crisis cluster analysis, however it had a stronger influence on the forecast than the pre-crisis information (volatility clustering effect).

2 volatility clustering effect
Recommended AI/ML methods for this use case: K-means, hierarchical clustering, unsupervised methods for finding clusters, supervised learning regression methods (GARCH).

3.2.3. Correlation to exogenous factors
External (exogenous) factors directly or indirectly influence the behaviour of financial series (indices). For example, volatility of market indices, the level of GDP, the size of interest rates. It is impossible to explain the variability of a forecast only by the changes of mathematical and statistical characteristics of a series. By involving external factors and finding correlations with clusters we can significantly improve the quality and reliability of a forecast.

As in the example of banks, we can use the data of the peer group to compare the behaviour of the indicator we are looking for with other firms. Benchmark values exist for some indicators. AI algorithms find correlations between parameters that are difficult for humans to find or pay attention to. The result of this analysis is a more accurate prediction.

Recommended AI/ML methods for this use case: Principal component analyses, random forest, decision tree, high-dimensional factor analysis.
4. BUSINESS IMPLEMENTATION – A PRACTICAL EXAMPLE

To summarize the above concepts and ideas, we present a conceptual approach of using AI methods in liquidity management tasks. It is based on the process described below (Figure 11) and summarizes the common data lifecycle, which regardless of the model complexity, exists for all data analytics applications. The key idea is that the business insights can only be as good as the sourced internal and external data that underpins them. Thus, the key issue is trust. Regardless of how powerful, accurate, or statistically reliable the results are, the data science capability needs to be trusted by the consumers when basing their decision-making on analytics insights.

Most ML Ops business implementations consist of five generic steps, which the following graph (Figure 12) illustrates.

1. **Data sourcing and ingestion** – determine what data is needed and the necessary onboarding steps.

2. **Data curation and pre-processing** – clean data based on requirements – in most platforms, data validation and pre-processing are done first.

3. **Data analytics** – normalize data if needed and apply ML models. Once the data has been prepared and stored, it is sent to the ML Pipeline, where all stages of searching, validating, and testing a model that fits the data are automatically performed.

4. **Evaluation results** – once the model is fitted, we can evaluate its results in terms of accuracy and validity.

5. **Business insights** – visualize results, generate reports, and evaluate business value and impact.

A cloud solution can be a convenient and efficient way to process, store, analyze and visualize data, as it provides a comprehensive automated approach to working with big data.

---

**Figure 11: Business process implementation of AI/ML in liquidity management**

---
Figure 12: Generic cloud-based ML Ops journey

Data sourcing and pre-processing
- Check Data Availability
- Clean and normalize Data
- Store Data in Data Lake

Data analytics in ML Pipeline
- Train model
- Package model
- Package model
- Deploy model
- Monitor model
- Retrain model

Results evaluation and visualization
- Evaluate results
- Visualize output
- Reportings
5. BARRIERS TO A WIDER ADOPTION OF AI/ML

Although the AI/ML methods described in this paper are very powerful tools, there are several limitations which restrict their wider application. Some of the challenges are described below:

1. **Unpredictable events.** The methods described in this paper are intended to improve the quality of modeling and predicting the financial performance of a business. However, sometimes events occur that cannot be predicted by machines. For example, unplanned outflows (reorganization or legal costs) or unpredictable inflows (unpredictable sales, capital inflows). Specific events known to a certain group of insiders cannot be known to a machine and such phenomena occur from time to time in any business. If these cases are known in advance, they can be added manually to adjust the quality of the forecast.

2. The problems of **machine algorithm bias** should also be mentioned here. Being trained on a limited set of data, the algorithm can “ignore” what it sees as “unimportant” events.

3. The legal aspects of using **cloud data processing and storage** systems may also be very difficult for large international companies, due to the differences in jurisdictional laws on the use and storage of data. Despite all the intelligence of AI, the implementation process requires the joint participation of specialists from different fields.

4. The current AI/ML tasks are generally about applying the existing best-in-class libraries to a specific space and use case. This raises the question of whether these methods are always applied correctly or whether their **“black box” results** are used naively to solve any business question. We hope to have provided a conceptual overview of the most common methods and their range of application and where these algorithms work, excel or stumble.

5. Often the problem is a **lack of competence** and using predominantly internal specialists. Typically, the minimum team composition for a pilot project in an AI field is seven people. These are the project manager, architect, subject matter expert, data science analyst, business analyst, technical writer, and infrastructure engineer. If a company is launching several pilots at the same time and implementing successful projects, it will be difficult to find the necessary number of qualified specialists within the company. It is necessary to compare the company’s costs for the development of a “self-written” solution (including the cost of equipment and developers’ time) and the costs of buying a ready-made solution - a leader in its field. In the first case, there is a possibility to spend a lot of time and effort and not get a quality result, as this is not the company’s main business and there are not enough resources to solve the problem.
6. CONCLUDING REMARKS

Summarizing the above, we can say that AI helps banks and corporates understand the behaviour of their customers at a deeper level and consequently have a more accurate and reasonable approach for determining and predicting the level of current liquidity. The main economic benefit results from maximizing the amount of cash that can be invested in more profitable assets by determining the minimum and sufficient level of liquidity (in the case of banks – optimal LCR minimization). More precise planning as well as an extended liquidity forecasting horizon allows to find ‘weaknesses’ in advance and to adjust liquidity management techniques, thus reducing volatility. Undoubtedly, this preventive approach has a positive effect on regulatory norms and reduces the risk of non-compliance. In summary, the use of AI can significantly improve the risk management situation, reducing the burden on risk management resources.

At the same time, fragmented data and silos will remain a challenge for most treasury departments. Making data available at the right time, in the right format should be the key focus in any data analytics journey.

Machine learning algorithms are nowadays readily available. There are a multitude of convenient packages in R/Python and tools (ML Ops, Automated ML4, Google Tensorflow, Theano, Torch, Scikit-learn, Jupiter Notebook) for automating the selection of model parameters and AL/ML methods. Most cloud services (Amazon ML, Google Prediction API und Cloud AutoML, Azure ML Services, SAP Leonardo, etc.) also offer data analytics capabilities for any kind of application.

However, significant economic effect can only be achieved by a complete transformation of the business on the principle of “AI as a mindset”. To reach this goal, it is necessary to form a pool of ideas, formulate clear objectives and scenarios, calculate risks, and prioritize use cases. With an agile approach, treasuries can track the development of prototypes for new methods using AI, thus consistently mastering new insights into forecasting and roles allocation. The resources freed up as a result can be directed to new and more important tasks.
1. A black swan is an unpredictable event that is beyond what is normally expected of a situation and has potentially severe consequences.

2. Volatility clustering effect described by Mandelbrot (1963): “large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes.”


3. Materials of Department of Mathematics, Florida State University https://people.sc.fsu.edu/~jpeterson/

4. Data Management: a foundation for effective data science, Capco Journal 50th edition, 2019

5. The CFO of the future, Capco Journal 50th edition, 2019

   Automated machine learning (AutoML) is the process of automating the tasks of applying machine learning to real-world problems.


ABOUT CAPCO

Capco, a Wipro company, is a global technology and management consultancy specializing in driving digital transformation in the financial services industry. With a growing client portfolio comprising of over 100 global organizations, Capco operates at the intersection of business and technology by combining innovative thinking with unrivalled industry knowledge to deliver end-to-end data-driven solutions and fast-track digital initiatives for banking and payments, capital markets, wealth and asset management, insurance, and the energy sector. Capco’s cutting-edge ingenuity is brought to life through its Innovation Labs and award-winning Be Yourself At Work culture and diverse talent.

To learn more, visit www.capco.com or follow us on Twitter, Facebook, YouTube, LinkedIn Instagram, and Xing.

WORLDWIDE OFFICES

APAC
Bangalore
Bangkok
Gurgaon
Hong Kong
Kuala Lumpur
Mumbai
Pune
Singapore

EUROPE
Berlin
Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
London
Munich
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA
Charlotte
Chicago
Dallas
Hartford
Houston
New York
Orlando
Toronto
Tysons Corner
Washington, DC

SOUTH AMERICA
São Paulo