Machine learning in UK financial services

The Bank of England and Financial Conduct Authority conducted a second survey into the state of machine learning in UK financial services.

Published on 11 October 2022

Content

Executive summary

- **The number of UK financial services firms that use machine learning (ML) continues to increase.** Overall, 72% of firms that responded to the survey reported using or developing ML applications. These applications are becoming increasingly widespread across more business areas.
- **This trend looks set to continue and firms expect the overall median number of ML applications to increase by 3.5 times over the next three years.** The largest expected increase in absolute terms is in the insurance sector, followed by banking.
- **ML applications are now more advanced and increasingly embedded in day-to-day operations.** 79% of ML applications are in the latter stages of development, ie either deployed across a considerable share of business areas and/or critical to some business areas.
- **Financial services firms are thinking about ML strategically.** The majority of respondents that use ML (79%) have a strategy for the development, deployment, monitoring and use of the technology.
- **Firms use existing governance frameworks to address the use of ML.** 80% of respondents that use ML say their applications have data governance frameworks in place, with model risk management and operational risk frameworks also commonplace (67%).
- **Firms consider that ML presents a range of benefits.** Currently the most commonly identified benefits are enhanced data and analytics capabilities, increased operational efficiency, and improved detection of fraud and money laundering.
- **Respondents do not see ML, as currently used, as high risk.** The top risks identified for consumers relate to data bias and representativeness, while the top risks for firms are considered to be the lack of explainability and interpretability of ML applications.
- **The greatest constraint to ML adoption and deployment is legacy systems.** The difficulty integrating ML into business processes is the next highest ranked constraint.
- **Almost half of firms who responded to the survey said there are Prudential Regulation Authority and/or Financial Conduct Authority regulations that constrain ML deployment.** A quarter of firms (25%) said this is due to a lack of clarity within existing regulation.

1: Introduction

1.1: Context and objectives

Over the past few years the use of machine learning (ML) has continued to increase in the United Kingdom (UK) financial services sector. As with other technologies, ML can bring a range of benefits to consumers, firms, markets, and the wider economy. Many firms are already realising these benefits and deploying ML applications across various business lines, services and products. However, ML can also raise novel challenges (such as ethical issues) and amplify risks to consumers, the safety and soundness of firms, and even potentially financial stability. That is why it is important regulatory authorities monitor the state of ML deployment and ensure they understand the different use cases, maturity of applications, benefits, and risks.

In 2019, the Bank of England (Bank) and Financial Conduct Authority (FCA) conducted a **joint survey** to gain an understanding of the use of ML in the UK financial services sector. One of the key findings was the need for further dialogue between the public and private sector to ensure the safe and responsible adoption of ML. The Bank and FCA established the **Artificial Intelligence Public-Private Forum (AIPPF)** in 2020, which explored various barriers to adoption and challenges related to the use of artificial intelligence (AI)/ML, as well as ways to address such barriers and mitigate risks.

This survey builds on the 2019 survey, the **AIPPF final report**, and the wider domestic and international discussion about the use of ML in financial services (in which the Bank and FCA have been active participants). In publishing the findings, the Bank and FCA demonstrate their commitment to monitoring the state of ML deployment, improve their collective understanding, and support the safe and responsible adoption of ML technology in UK financial services.

This joint Bank-FCA report is the result of the analysis of the responses to the 2022 survey. This includes, in relation to the firms that responded to the survey:

- a quantitative overview of the use of ML;
- the ML implementation strategies of firms;
- the share of ML applications developed in-house or by third-party providers;
- approaches to the governance of ML;
- respondents' views on the benefits of ML;
- respondents' views on the risks of ML;
- perspectives on constraints to development and deployment of ML; and

- a snapshot of the use of different methods, data, safeguards performance metrics, validation techniques and perceived levels of complexity of ML.

The report closes with a selection of ML case studies, describing a sample of typical use cases, including:

- Insurance pricing and underwriting.
- Credit underwriting.
- Marketing.
- Fraud prevention and anti-money laundering (AML).

1.2: Methodology

In total, 168 firms received the questionnaire and 71 submitted responses (42% overall response rate). The Bank surveyed 48 dual-regulated firms, 17 firms applying for Prudential Regulation Authority (PRA) authorisation as a deposit-taker, and eight Bank-regulated financial market infrastructures (FMIs), and received 51 (70%) responses. The FCA surveyed 95 FCA-regulated firms and received 20 (21%) responses.

The Bank selected firms with the aim of surveying each type of FMI and PRA-regulated firm and covering a significant share of those firms. It also included several firms that are small in terms of their market share but were considered to be advanced in the use of ML and therefore of interest for horizon-scanning purposes. In addition, the sample included a number of FCA-regulated small-sized firms, who were undergoing the PRA authorisation process for deposit-taking permissions.

The FCA sent the survey to a representative list of firms from the following sectors: credit referencing agencies, crowdfunders, custody services, exchanges, fund management, alternatives, lifetime mortgage providers, multilateral trading facilities, non-bank lenders, principal trading firms, wealth manager and stock brokers, wholesale brokers, credit brokers, debt purchasers, debt administrators, consumer credit providers, motor finance providers, retail finance providers, payment services, and e-money issuers. It also included firms who responded to the 2019 survey.

Overall, the combined sample is skewed towards larger firms with no responses received from smaller fintech firms or start-ups. While firms may be more likely to respond to the survey if they are already using or developing ML, the sample can be seen to provide a broad representation of firms by types of activity, size, and areas of ML applications. However, the sample and survey findings should not be seen as representative for all types of firms or the entire UK financial services industry.

The results presented in this report are anonymised and aggregated with the respondents grouped into the sectors listed in Table A.

Table A: Sector classification used in the survey and report

Sector	Type of firms included
Banking	Building societies, international banks, retail banks, UK deposit-takers
Insurance	General insurers, health insurers, life insurers, personal and commercial lines insurers
Non-bank lending	Credit brokers, consumer credit lender, non-bank lenders
Investment and capital markets	Alternatives, asset managers, fund managers, wealth managers and stockbrokers, wholesale brokers
Financial market infrastructures (FMIs), payments and other	Credit reference agencies, e-money issuers, exchanges, financial market infrastructures, multilateral trading facilities

Sources: Bank of England and Financial Conduct Authority.

All charts in this report are based on data received from respondents from this survey. When designing the survey, the Bank and the FCA considered the Legislative and Regulatory Reform Act 2006 principle that regulatory activities should be carried out in a way which is transparent and proportionate.

Box A: Definitions of ML, application, algorithm and model

ML is a methodology whereby computer programmes build a model to fit a set of data that can be utilised to make predictions, recommendations, or decisions without being explicitly programmed to do so, instead learning from sample data or experience. There are many different approaches to the implementation of ML, which include techniques such as supervised learning, unsupervised learning, and reinforcement learning. For the purpose of this survey, this excludes simple linear regression – which we define as any regression techniques that does not employ subset selection methods, shrinkage methods or dimension reduction methods.

'ML application' refers to an entire system, including data collection, feature engineering, model engineering, and deployment. It also includes the underlying IT infrastructure (eg virtualisation, data storage, and integrated development environment). An ML application could include multiple models and algorithms. Respondents were asked to classify ML applications separately if they fulfil different business purposes or if their set up/components differ significantly.

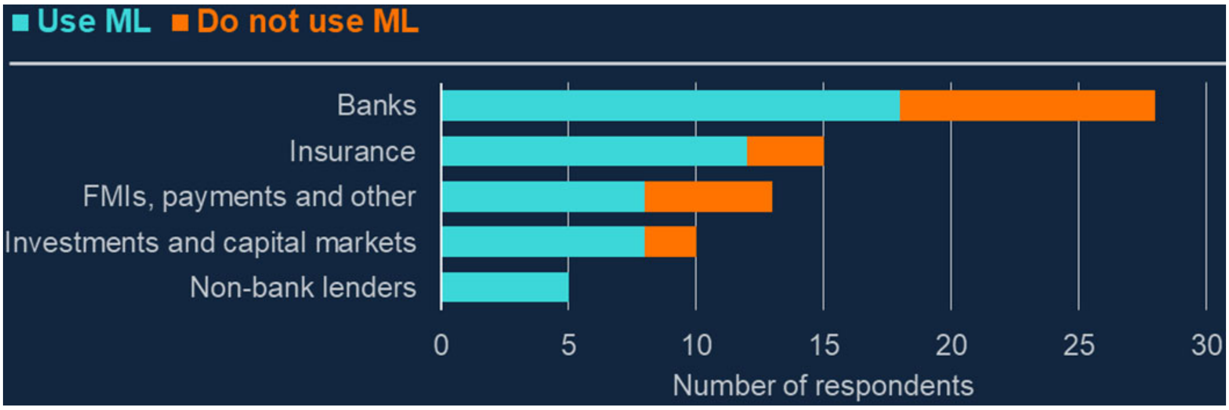
The term ‘algorithm’ means a set of mathematical instructions or rules that, especially if given to a computer, will help to calculate an answer to a problem. Whereas the term ‘model’ means a quantitative method, system, or approach that applies statistical, economic, financial or mathematical theories, techniques, and assumptions to process input data into output. The definition of a model includes input data that are quantitative and/or qualitative in nature or expert judgement-based, and output that are quantitative or qualitative. In ML, an algorithm is a procedure that is run on data to create a model.

2: Machine learning adoption and use

2.1: Financial services firms use an increasing number of ML applications

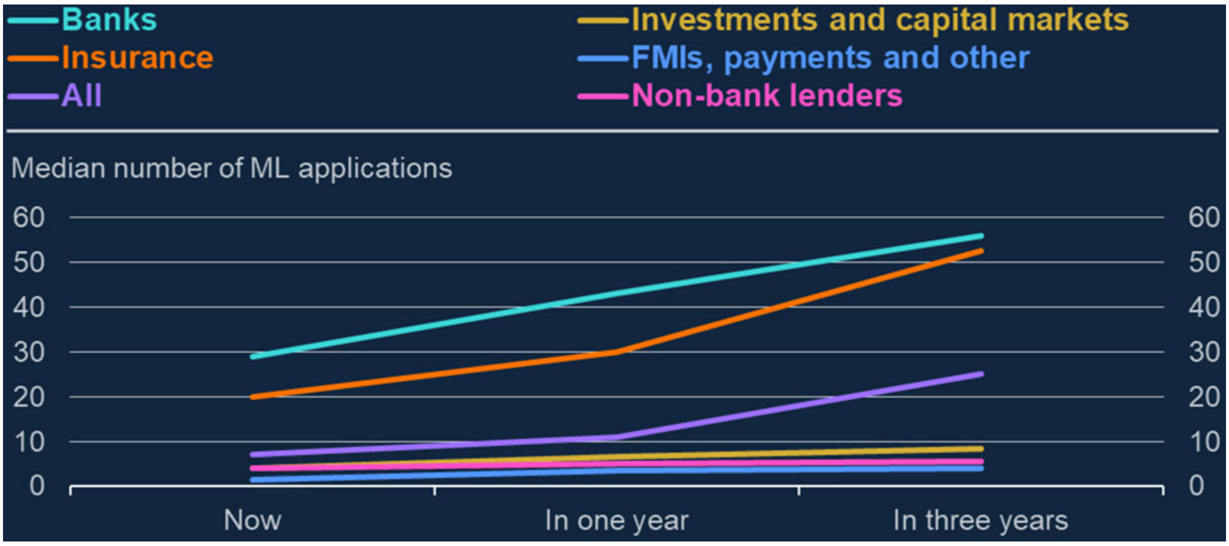
The number of ML applications used in UK financial services continues to increase. Overall, 72% of firms that responded to the survey reported using or developing ML applications. This compares to 67% of respondents to the 2019 survey, although it is worth noting the sample size and composition was different to the 2022 survey. Similar to 2019, respondents from the banking and insurance sectors have the highest number of ML applications.

Chart 1: 72% of firms that responded already use or are developing ML



Respondents expect this trend to continue, with the overall median number of ML applications expected to increase by over 3.5 times over the next three years. This increase is in line with the trend reported in the 2019 survey. The largest expected increase is in the insurance sector, with the median number of applications per firm expected to increase by 163%.

Chart 2: Median number of ML applications expected to increase by over 3.5x

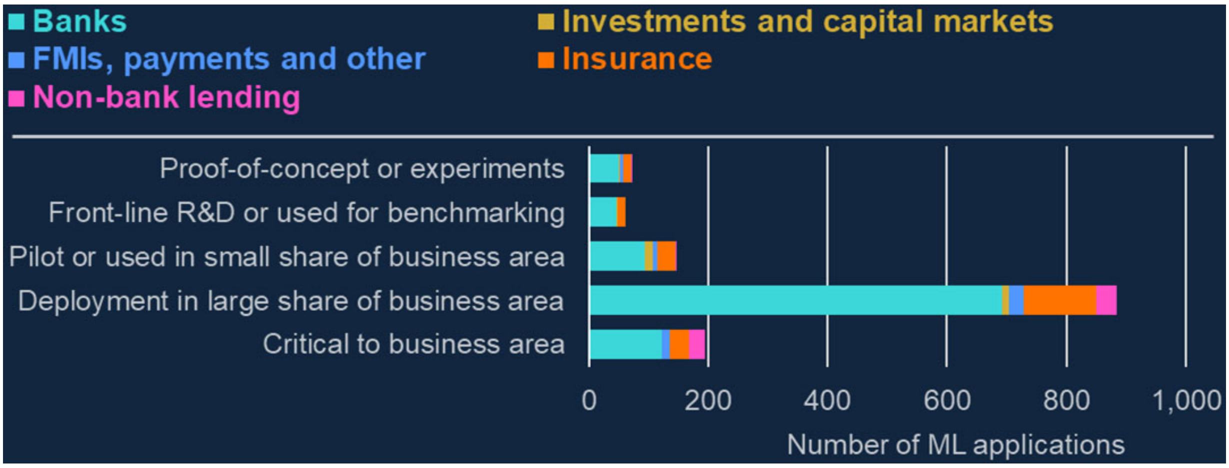


2.2: Deployment stage

ML applications pass through a number of development and deployment stages. The survey asked firms to report the number of applications they have at each of the five key stages: (i) proof-of-concept or experimental, (ii) front-line research and development and/or used for benchmarking existing models, (iii) pilot and/or used in small share of business area, (iv) deployed across considerable share of business area, and (v) critical to business area.

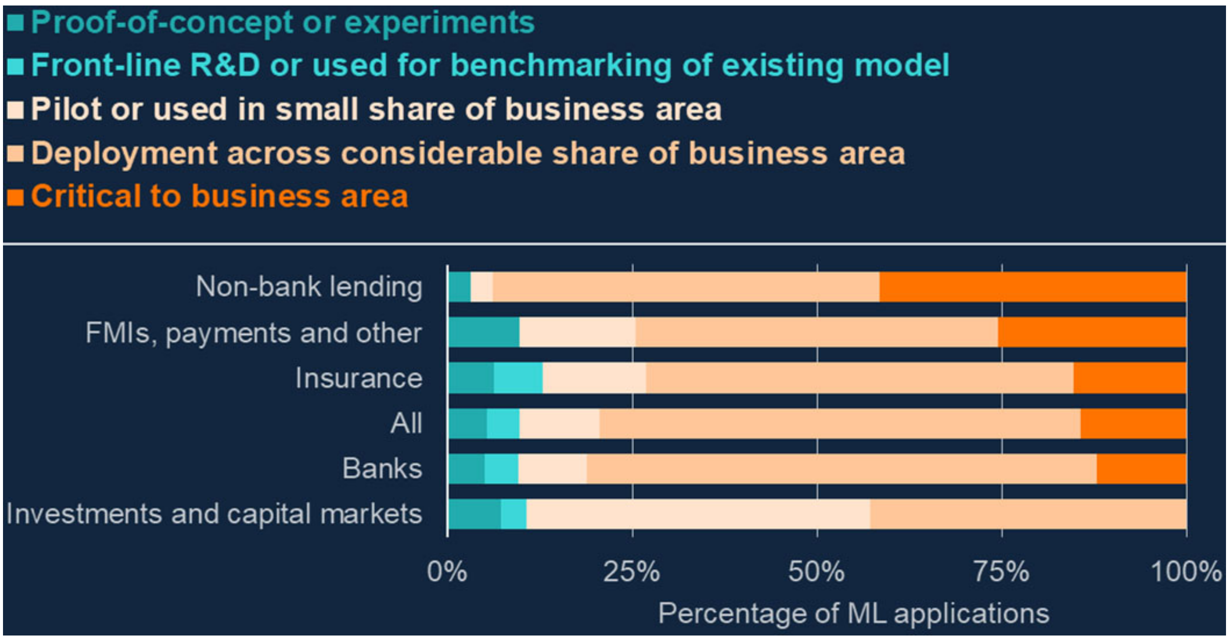
From the survey responses, 79% of ML applications are in deployment (Chart 3). In particular, 65% of applications are already deployed across a considerable share of business areas, with a further 14% of ML applications reported to be critical to the business area. Although the survey question was somewhat different in 2019, a significantly higher proportion of applications were in pre-deployment stages then, 44% in 2019 versus 10% in 2022. This suggests the survey respondents' ML applications are more advanced and increasingly embedded in day-to-day operations.

Chart 3: Overall, 80% of ML applications are in deployment or critical stages



Banks, insurance, and FMIs, payments and other firms broadly have a similar split between the different stages of deployment (Chart 4). Non-bank lenders have the highest percentage of ML applications (42%) that are critical to business areas with just 3% of applications in pre-deployment. At the other end of the scale, respondents from the investment and capital markets sector have the largest number of ML applications in the pilot or small share of business stage and no critical applications.

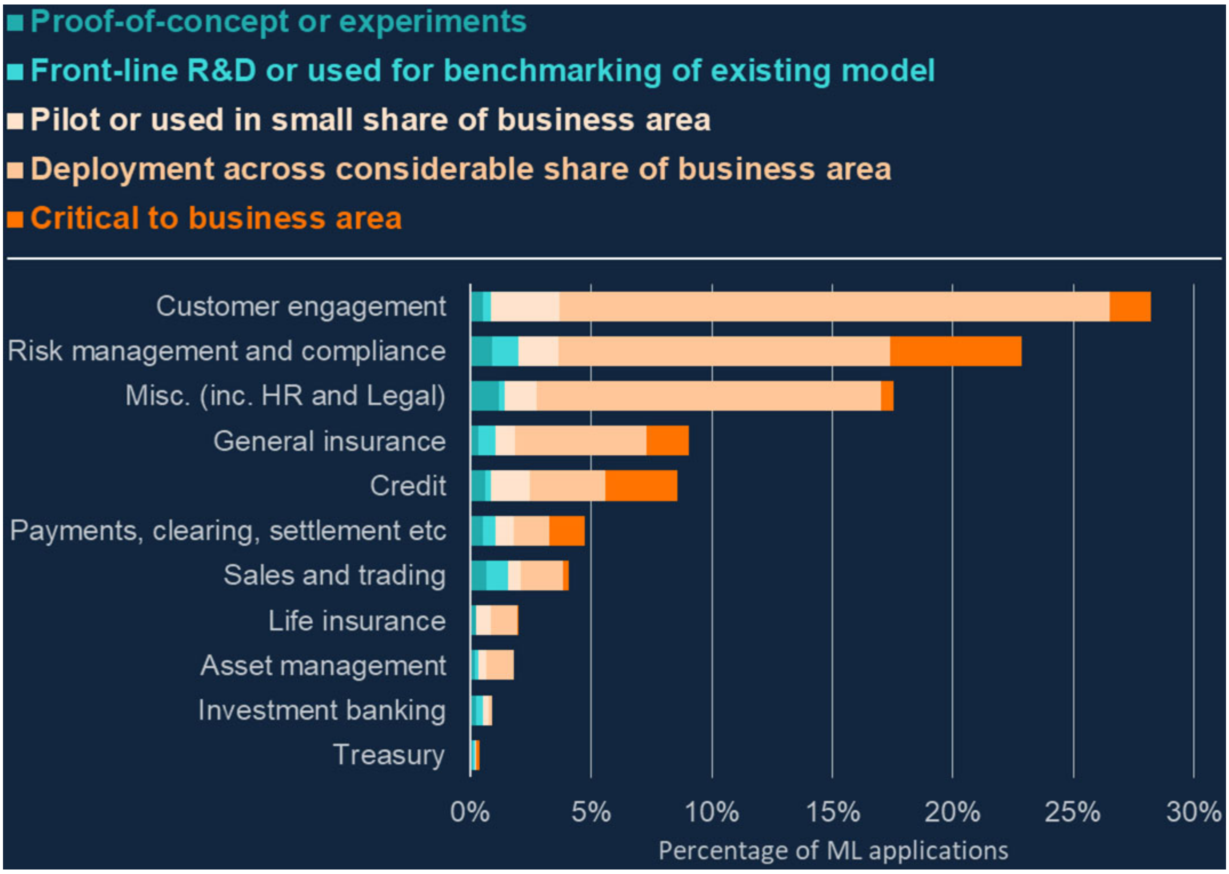
Chart 4: Non-bank lenders have the highest percentage of ML applications in deployment



2.3: Range of applications across sectors and business areas

In terms of the range of ML use cases (Chart 5), firms are developing or using ML across most business areas. As with the 2019 survey, 'customer engagement' and 'risk management' continue to be the areas with the most applications and account for 28% and 23% of all reported applications respectively. The 'miscellaneous' category, which included business areas like human resources and legal departments, had the third highest percentage of ML applications (18%). The business areas with the fewest ML applications are 'investment banking' (0.9%) and 'treasury' (0.4%), with the latter also being the business area with the fewest ML applications in the 2019 survey.

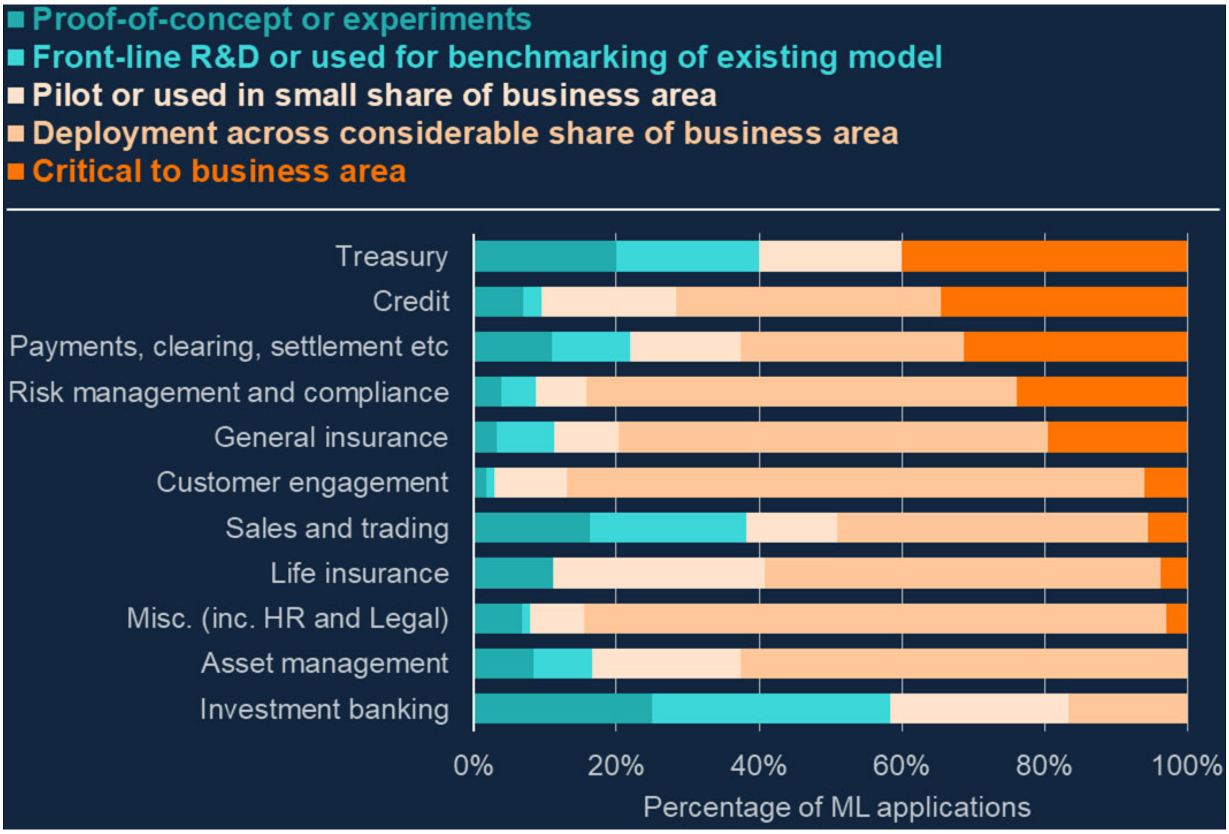
Chart 5: Over half of ML applications are in customer engagement or risk management



In terms of maturity of ML applications by business area (Chart 6), 'customer engagement' (97%) has the highest percentage of post-deployment applications. The areas with the highest proportion of ML applications at the pre-development stages are 'investment banking' and 'treasury' with 58% and 40% respectively. It is worth noting that treasury (40%) and credit

(34%) are the areas with most ML applications rated as ‘critical to business area’. The business areas with no critical ML applications are ‘investment banking’ and ‘asset management’.

Chart 6: Treasury and credit have the most ML applications that are critical to the business area



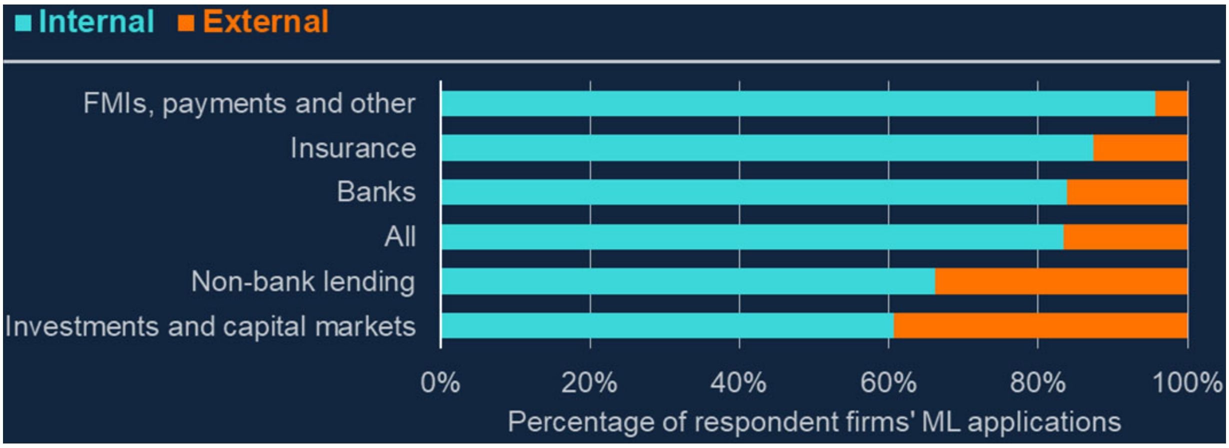
2.4: Internal versus external implementation and cloud computing

The survey asked firms about the number of ML applications that were implemented internally, compared to the number of applications that were implemented externally by third-party providers (ie applications where the majority of the development or deployment actions were implemented by a third party). However, the line between ‘internal’ and ‘external’ implementation is becoming increasingly blurred. This is because ML systems are becoming more complex and rely on a mix of internal and external components (data inputs, ML models, software packages, cloud computing storage, etc). For example, firms may develop models and code algorithms internally but use third-party data sets or base their algorithms on third-party cloud computing platforms.

While the **2020 Bank survey on the impact of Covid on ML in UK banking** suggested an increase in outsourcing and use of external vendors, this is not reflected in this survey. Instead, 83% of respondents develop and implement ML applications internally at their firm.

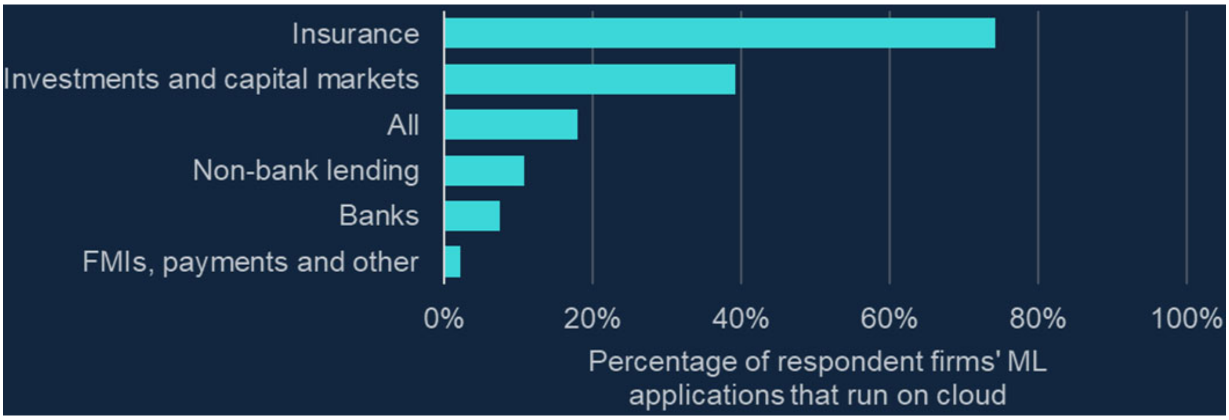
This is in line with the 76% figure reported in the 2019 survey. Investments and capital markets is the sector of respondents with the largest use of third-party ML applications (39%).

Chart 7: Investment and capital markets have the most externally implemented ML applications



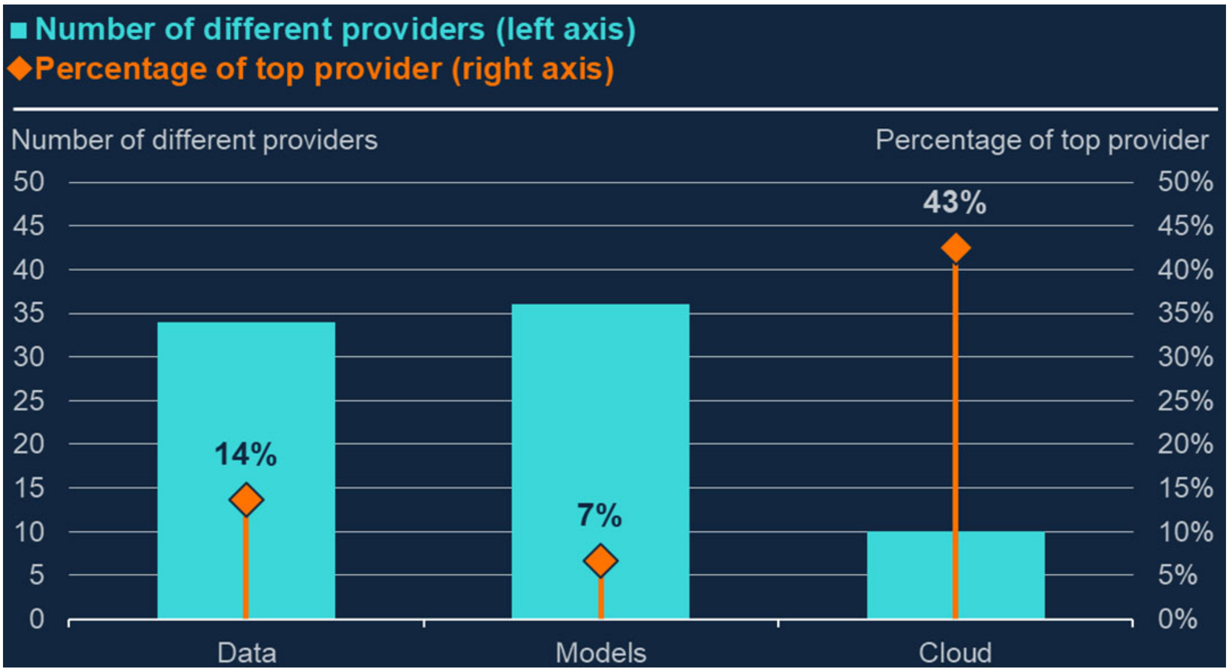
The survey also asked about the number of ML applications that are run on cloud computing platforms (Chart 8). Overall, 18% of ML applications are run on cloud computing platforms. While this is in line with the 22% reported in the 2019 survey, there are marked differences for particular market segments. In particular, insurance firms reported that nearly three quarters (74%) of their applications run on cloud computing platforms, compared to 39% of investment and capital markets firms, 11% of non-bank lenders, and 8% of banks. The percentage of insurance ML applications run on cloud computing platforms in this survey is also significantly higher than the 31% reported by insurers in 2019.

Chart 8: Almost three quarters of insurance firms' ML applications run on the cloud



The survey also asked firms which third-party providers they used for data, models, and cloud computing services. As Chart 9 shows, there was a wide range of data and model providers. However, there is a degree of concentration in cloud service providers with the top two providers accounting for 75% of firms that use cloud services (and the top provider accounting for 43% of firms that use cloud services).

Chart 9: Little diversity in cloud providers, with top provider named by 43% of respondents



3: Strategies and governance

3.1: Firms' ML strategies

The 2019 survey found that some firms were already thinking about ML strategically, mostly in the banking and insurance sectors. This trend has continued and expanded to other sectors. The majority of respondents in the 2022 survey (79%) had some form of strategy for the development, deployment, monitoring and use of ML (Figure 1). In terms of the sector breakdown by respondents, 100% of insurance firms had a strategy, as did 83% of non-bank lenders, 78% of banks, 70% of investment and capital market firms, and 60% of firms in the FMIs, payments and other category.

Survey respondents tend to use different elements of existing strategies, sometimes combining multiple approaches or elements into one, to support their use of ML. For example, 38% of respondents have a model risk management (MRM) strategy that incorporates ML and 25% of survey respondents include ML as part of their wider data, innovation or technology strategy. Elements of firms' wider governance frameworks are used as part of the overall ML strategy, as are ethical principles such as those related to fairness or bias in decision-making.

Respondents also explained how they operationalise the strategies. 29% of firms have specific teams that are responsible for the development and deployment of ML within the firm. These often consist of cross-functional teams that allow ML models to be deployed across multiple business cases within the firm. Sometimes these teams are also responsible for monitoring ML model risks and outputs.

Many of the firms that do not use ML report that it is not a priority given the size, scope or focus of their business.

Figure 1: Percentage of respondents with overall ML strategy and strategy elements

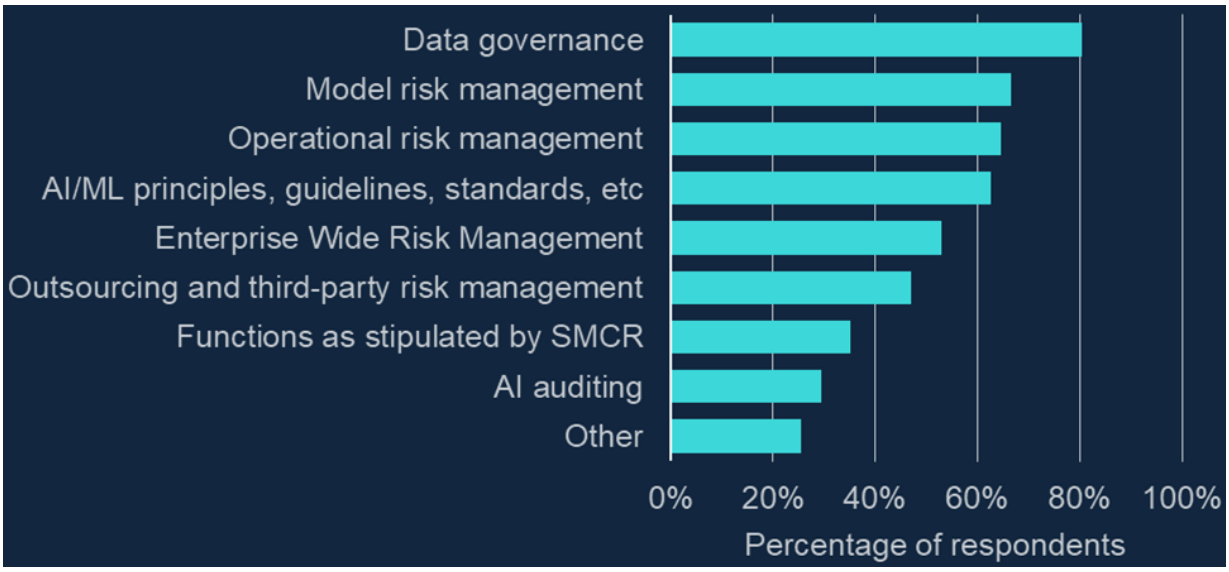
	Overall strategy	Strategy elements			
		Data	MRM	Governance	Ethics
Insurance	100%	31%	38%	15%	23%
Non-bank lending	83%	40%	20%	20%	0%
Banks	78%	11%	44%	33%	39%
Investments and capital markets	70%	40%	30%	10%	0%
FMIs, payments and other	60%	20%	40%	30%	30%
All	**79%**	**25%**	**38%**	**23%**	**23%**

3.2: Firm ML governance and accountability

Good governance is essential for the safe adoption and use of ML in financial services. Governance underpins effective risk management across the ML lifecycle by putting in place the set of rules, controls, policies, and processes for a firm's use of ML. Furthermore, governance ensures accountability for ML applications and is vital for ensuring that ML is used in a safe and responsible manner.

The survey asked firms which governance elements and frameworks they employ (Chart 10). As with firms' strategies for ML, many respondents use existing governance frameworks (such as MRM and operational risk management) to address the use of ML. As noted below, 80% of respondents say that they have data governance frameworks in place. Just over two thirds of respondents (67%) have AI/ML specific principles, guidelines, and standards in place as part of their approach to ML governance. This may reflect the fact that ML can pose novel challenges to financial services firms and, therefore, may need specific governance principles to address them (as noted in the **AIPPF final report** and **governance meeting minutes**).

Chart 10: 80% of firms have data governance frameworks in place for their ML applications



3.3: Lessons learnt from ML deployment

The survey asked respondents about the main lessons learned from the deployment and use of ML applications within their firms, including how to integrate these lessons into the broader approach to ML development. The most common lesson respondents highlight is the need for good and effective governance frameworks (such as MRM). In particular, respondents mentioned clear lines of ownership of the ML application within the business to allow for effective risk management.

Respondents also emphasise the importance of continuous reviews and assessments of ML applications to ensure the outputs are accurate and fair. This includes assessing potential bias in decisions that may affect consumers. Respondents also said it is key to know when best to retire and replace ML models. Respondents noted that understanding ML models and their outputs are also essential elements of ML explainability and transparency. Some respondents highlighted the need for clear metrics to assess what constitutes a successful outcome for a ML application and suggest setting a framework and baseline metrics to measure the impact of ML applications.

Survey respondents noted that well-trained teams are crucial for the effective implementation and monitoring of ML applications. Respondents reported that cross-functional, multi-disciplinary teams tend to be the most effective at developing and deploying ML applications that are safe and responsible. Survey respondents also mentioned the importance of comprehensive staff training and culture to ensure the safe and responsible adoption of ML.

4: Benefits, risks and constraints

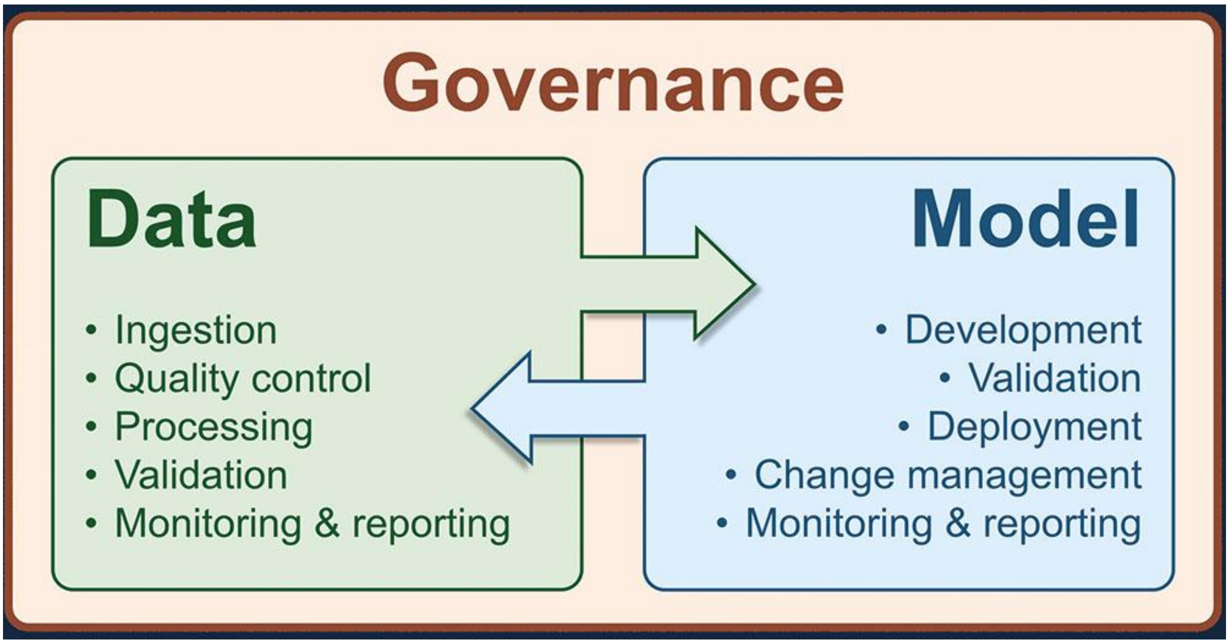
4.1: Benefits, risks and trade-offs

The use of ML can bring benefits to consumers, firms, and the wider financial system. For consumers, ML can create more personalised products and services as well as better customer engagement. Firms can benefit from improved data analytics and increased operational efficiencies. These benefits can in turn aggregate to the financial system and the economy as a whole.

However, the use of ML in financial services can also amplify existing challenges and risks. Many of these challenges and risks can be traced to three underlying drivers and stages of the ML lifecycle: (i) data, (ii) models, and (iii) governance (Figure 2).

For example, historical biases in data sets and poor-quality data can feed into the modelling and cause subsequent model risk. ML techniques can also increase model complexity and lack of explainability, as well as other forms of model risk. Although effective governance and accountability are key mitigants of those risks, weak and ineffective ML governance can exacerbate the issues and be a challenge in itself.

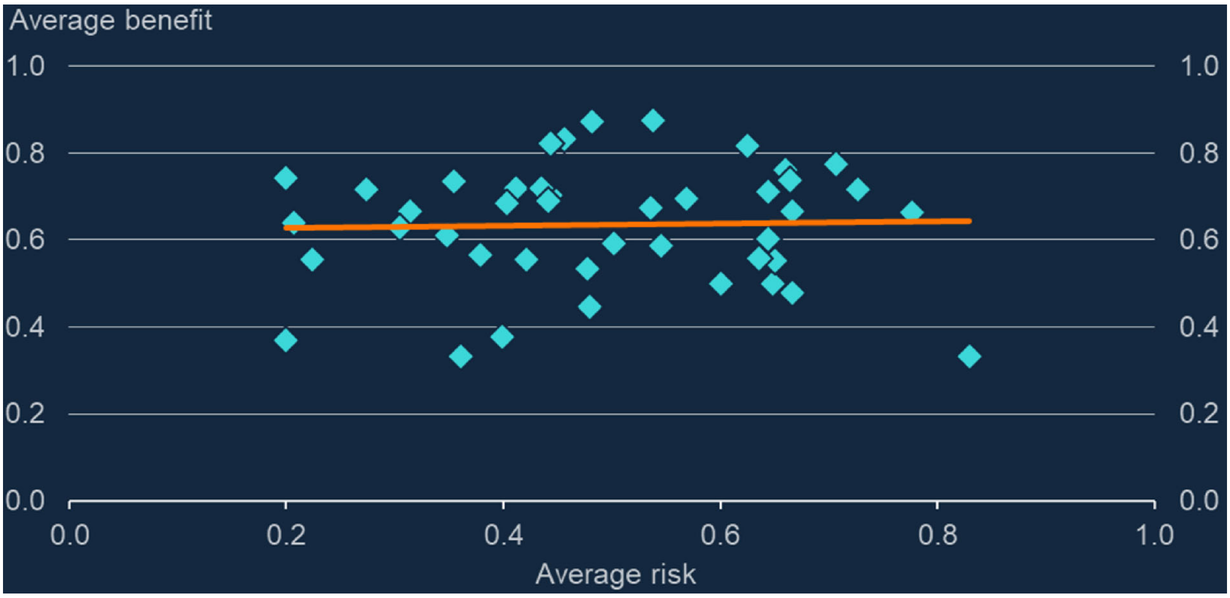
Figure 2: Stages of ML lifecycle



Source: Artificial Intelligence Public-Private Forum.

While there are many trade-offs in considering the adoption and use of ML, one could expect to find a correlation between the perceived benefits and the perceived risks of using ML, namely, that greater risk would be associated with greater benefit. However, the survey responses show no significant association between average perceived risk and average perceived benefit (Chart 11).[1]

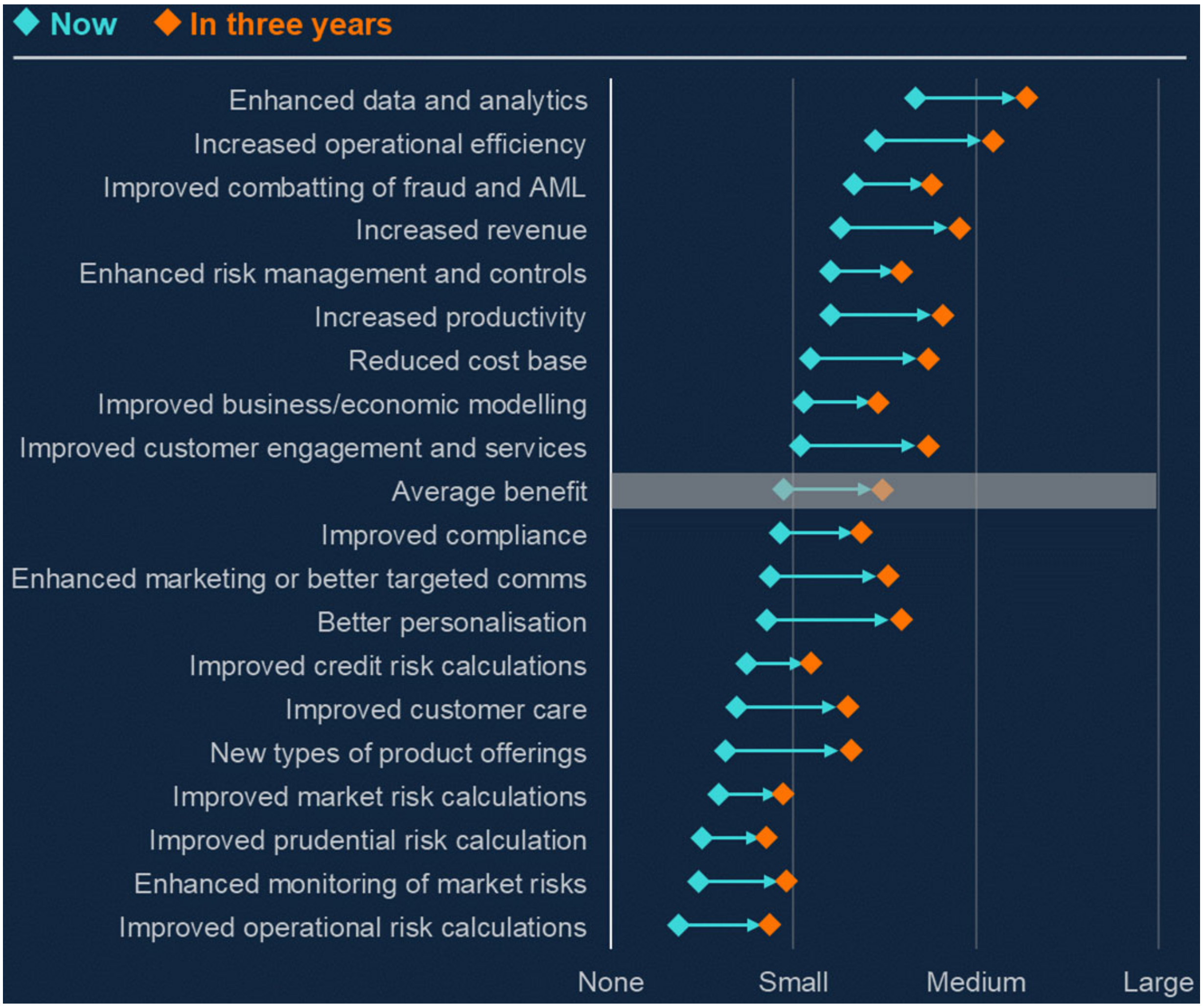
Chart 11: No significant association between perceived risk and perceived benefit



4.2: Benefits now and in three years

This survey suggests that benefits are the most likely in relation to enhanced data and analytics capabilities, increased operational efficiency, and improved combatting of fraud and money laundering (Chart 12). All benefits are expected to increase over the next three years, with the greatest increase expected to be in better personalisation and improved customer engagement.

Chart 12: All categories of benefits are expected to increase over the next three years



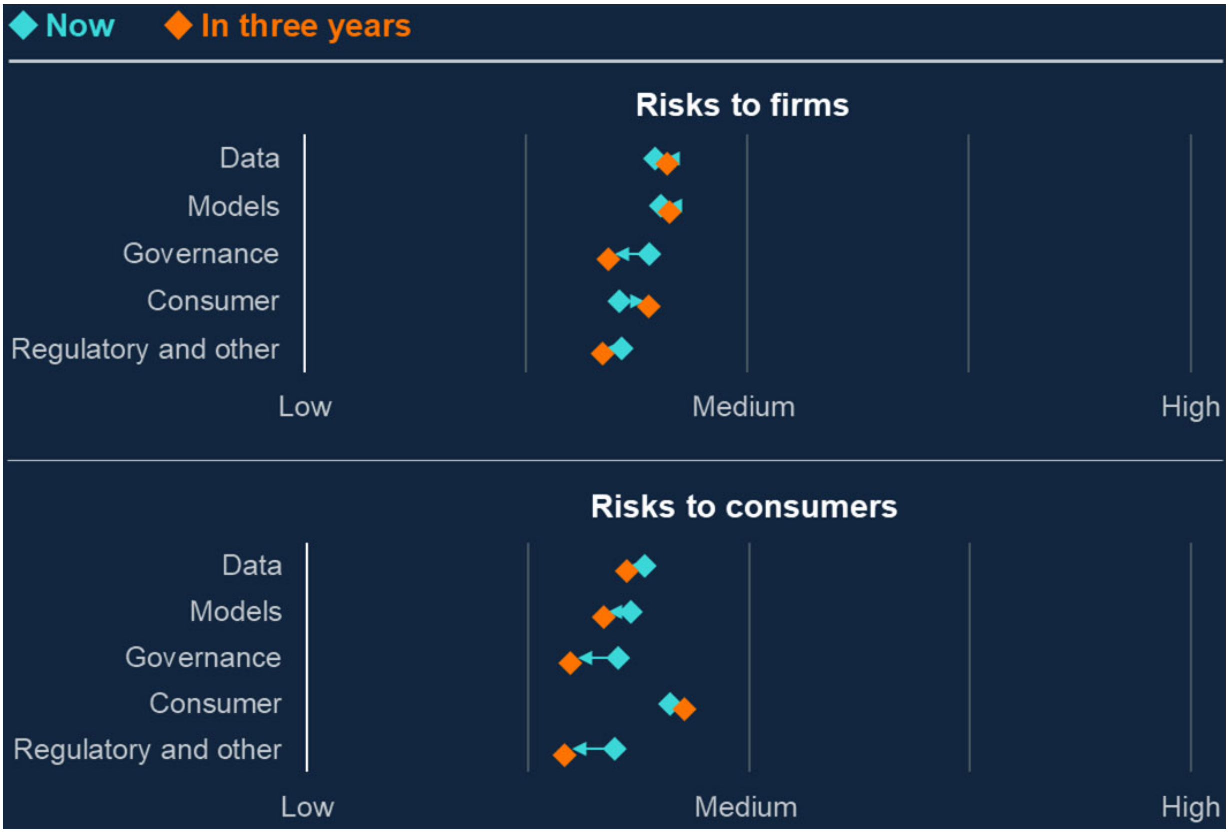
4.3: Risks and mitigants

Overview

As mentioned above, the primary drivers of ML risk in financial services relate to interconnected risks at the data level, which feed into the model level, and then raise broader challenges at the level of the firm and its overall governance of ML systems. The survey asked respondents to rate the level of various risks related to those three drivers, as well as any specific risks to consumers, regulation, and other risks such as cybersecurity and outsourcing risk.

Overall, respondents consider the current levels of risk to be low to medium across all risk categories and expect this to stay at similar levels over the next three years. While the highest perceived risks are for consumers, respondents consider risks related to ML models, rather than data or governance, as the highest risks for firms (Chart 13).

Chart 13: Overall levels are low to medium across all risk categories



The survey also asked firms to identify specific risks within each of the categories. Respondents said the lack of explainability in both the workings and outcomes of ML applications are a key risk, which can lead to both inaccurate consumer outcomes and subsequent reputational and legal risk to firms.

A further risk many respondents highlighted was the potential risks associated with data quality, structure, and bias issues.

More than half of firms (52%) noted that the use of ML causes concern about potential ethical and bias issues arising. The respondents noted that these issues can have a negative impact on consumers' ability to use products and services. This can, in turn, damage the firm's reputation and lead to operational costs, service breakdowns, and other negative consequences. In addition, respondents highlighted outsourcing risk, which can include issues around third-party compliance with firms' governance and data sharing frameworks.

Overall, respondents consider the top ML-related risks to be: biases in data, algorithms and outcomes (52%); data quality and structure issues (43%); lack of explainability within the model itself and the outcome (36%), which could lead to inaccurate predictions (34%)

resulting in poor decisions and reputational damage (11%); inadequate controls or governance (25%); and outsourcing or third-party risks (16%).

While 14% of respondents report that ML applications may amplify existing risks and challenges, the technology can also introduce novel ones. Half of the survey respondents say the use of ML within their firms could lead to novel risks, whereas 23% could not see any novel risks arising and 27% gave no comment.

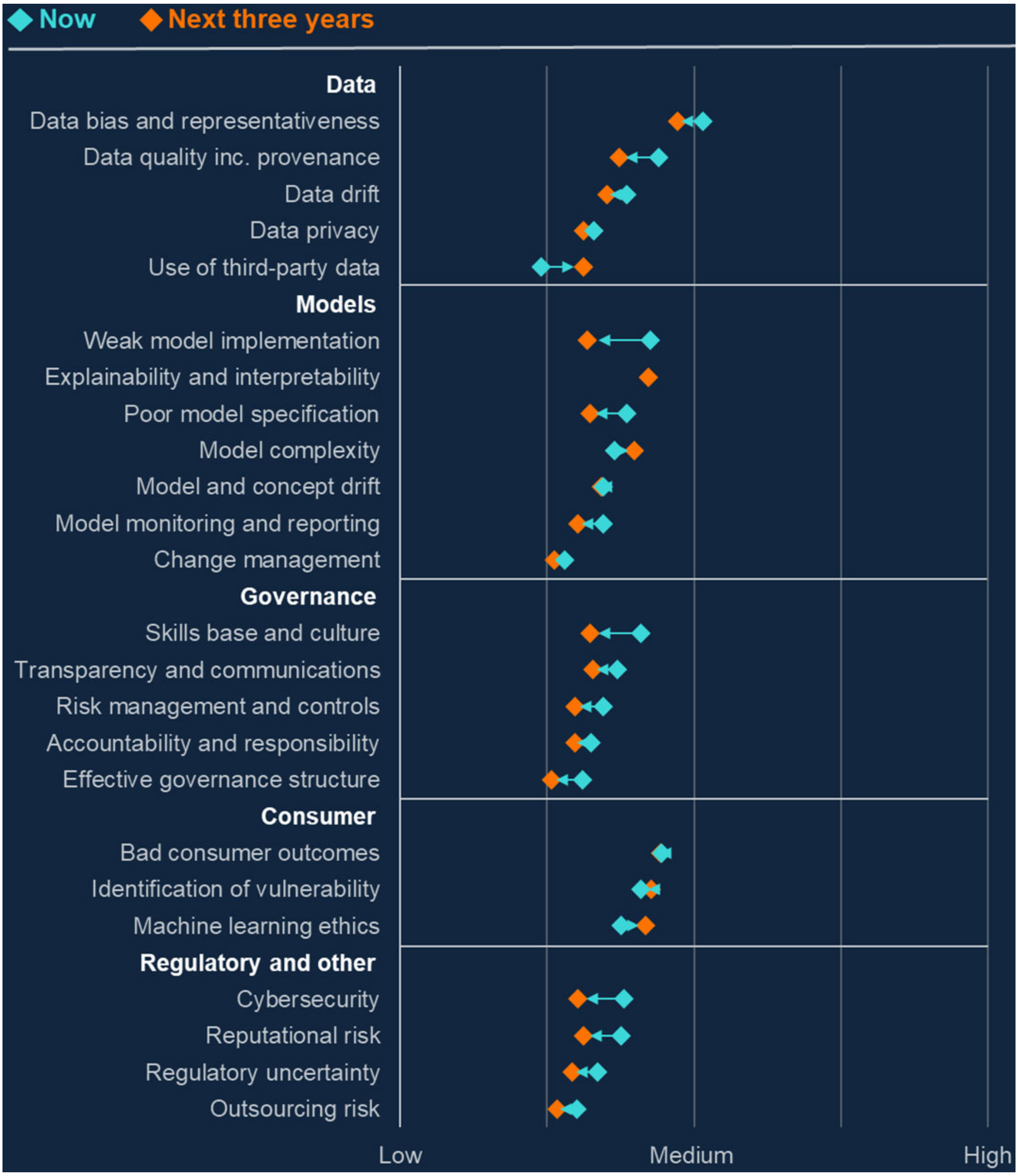
The survey also asked firms to identify what they considered to be novel risks associated with the use of ML. 13% of respondents highlighted the risk related to unethical and biased outcomes, which could be particularly harmful for vulnerable consumers. Another area highlighted by 9% of firms was the potential increase in reputational damage by deploying risky ML.

In terms of mitigating ML risks, respondents said effective governance frameworks (such as MRM and data-quality validation) are key and this should include effective assessment and reviews of ML models from development stages through to deployment. Survey respondents also said clear lines of accountability are key to mitigate risks with some respondents' highlighting the use of 'human in the loop' processes to ensure there is a human accountable for any autonomous decisions. Some respondents suggested consumer protection and prudential regulation could be an effective mitigant to some of the ML risks.

Risks to consumers

Respondents expect most risks to consumers to decrease over the next three years, with the exception of risks related to third-party data, ML ethics, and model complexity. The top three identified risks are data bias and representativeness, bad consumer outcomes, and identification of vulnerability. The risks that are expected to decrease the most are 'weak model implementation' and 'skills base and culture'.

Chart 14: Most ML risks to consumers expected to decrease over the next three years

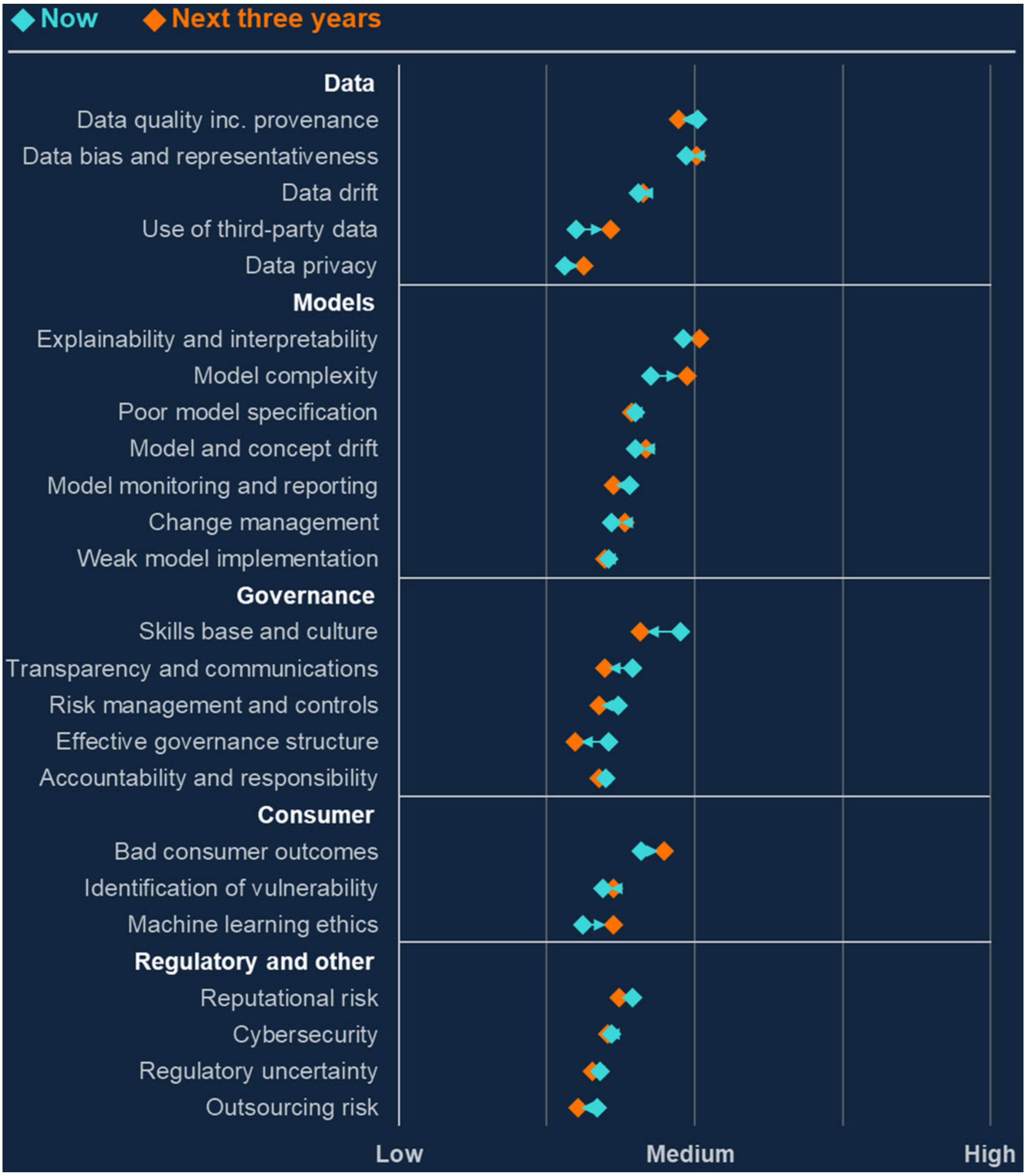


Risks to firms

While the overall level of perceived risks to firms is similar to the risks to consumers, there are differences in the ranking and rating of specific risks. The highest three perceived risks to firms are explainability and interpretability, data bias and representativeness, and model complexity.

Survey respondents expect the majority risks to firms to remain constant over the next three years. As with the perceived risks to consumers, the same three categories are expected to increase: use of third-party data, ML ethics, and model complexity. The risks expected to decrease the most are skills base and culture and effective governance structure. These results are largely aligned with the findings of the AIPPF.

Chart 15: Highest perceived ML risk to firms in next three years is explainability and interpretability

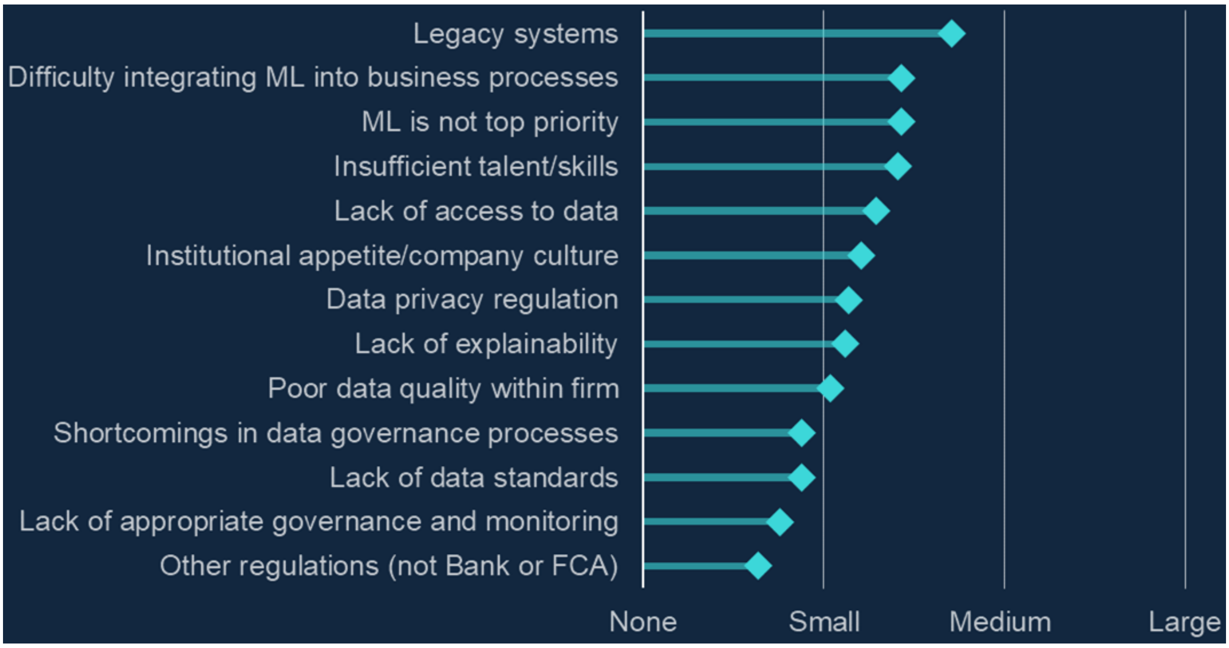


4.4: Constraints to deployment

There are various factors that may constrain deployment of ML systems among respondent firms. The greatest perceived constraint is legacy systems and associated technology infrastructure, similar to the 2019 survey. ML applications and the processes around them often need up-to-date hardware, database infrastructure, and operating systems to run effectively and efficiently. That is why developing, deploying and/or integrating ML with legacy systems can become challenging, which is the second highest reported challenge.

While the third highest reported challenge is that ML is not a priority at this time, the fourth biggest challenge is a lack of sufficient skills. As the **AIPPF final report** notes, this lack of skills can become a significant challenge not only at the development and implementation phases of ML models but also in ongoing monitoring and risk management. There may also be insufficient skills in understanding and managing third-party models or data.

Chart 16: Legacy systems remain the highest perceived constraint to ML deployment



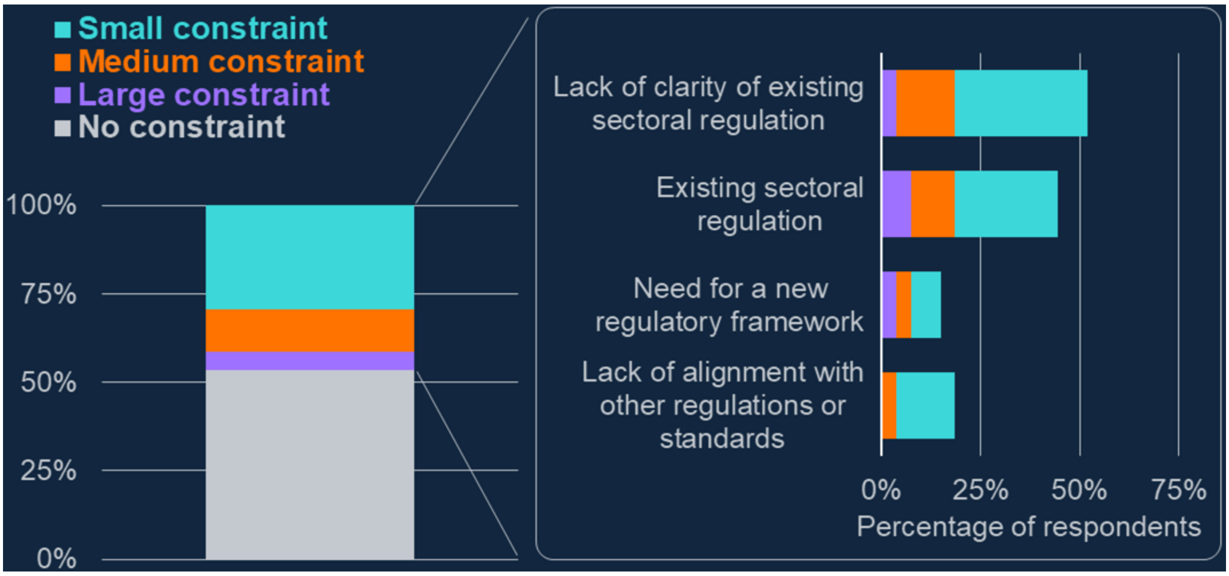
4.5: Regulation

Financial services regulation may act as a barrier to development or a constraint on deployment, especially if the perceived compliance burden and associated costs outweigh the perceived benefits. Almost half of firms who responded to the survey said that there are regulations (for which the PRA and/or FCA are the competent authorities) that constrain ML deployment, although 30% of respondents said that these are a small constraint with only 5%

noting that sectoral regulation is a large constraint (Chart 17). A quarter of respondents said it was due to a lack of clarity with existing regulation and just under a quarter of respondents (22%) thought that existing regulation is itself as a constraint.

The Bank, PRA, and FCA are exploring how the current regulatory framework applies to AI/ML, whether additional clarification of the existing regulatory framework may be helpful, and how policy can best support further safe and responsible AI/ML adoption via the joint **Bank-PRA-FCA Discussion Paper** (DP).

Chart 17: 47% of respondents said there are regulations that constrain ML deployment



5: Case studies

5.1: Purpose and background

The survey asked firms more detailed questions about two case studies: (i) the most advanced ML application in terms of deployment and (ii) the ML application deployed (even on a small scale) within a critical business area. These questions aimed to explore certain case studies in more depth and provide insights into the specific business contexts within which ML models were used.

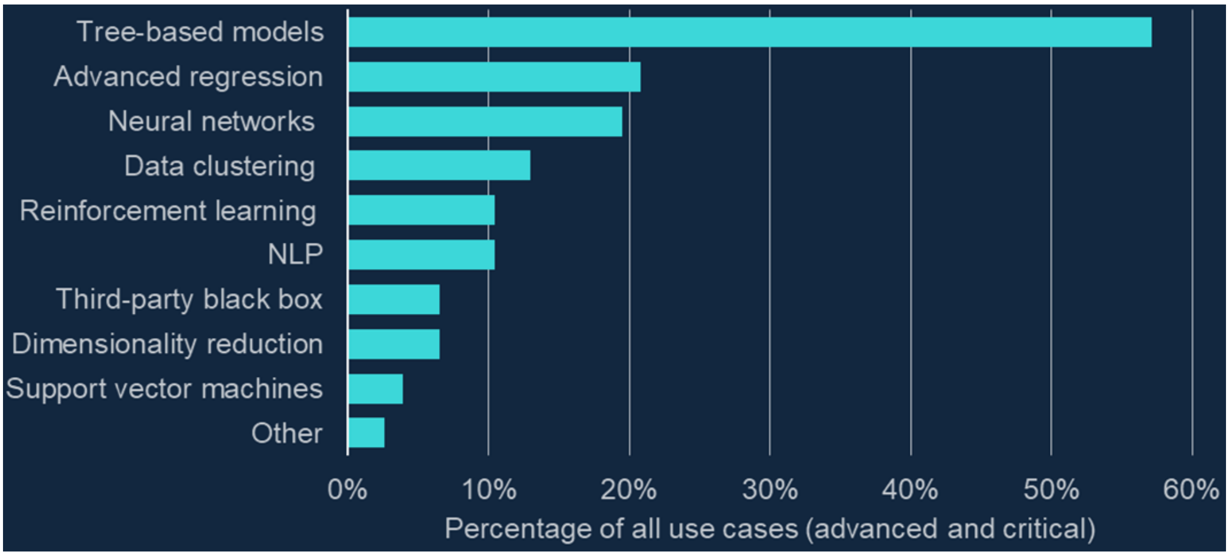
In total, 45 firms responded to the first case study (most advanced ML application) and 34 responded to the second question (ML use in critical business area). There was some, limited overlap between these use cases. Ten respondents that use ML elsewhere in their firms said they do not use the technology in any critical business areas.

The case studies presented here were selected based on the number of responses received, so reflect the most commonly reported use cases within the survey sample.

Types of ML techniques and data

According to firms' responses for both case studies, the ML techniques most often used are on the more complex end of the current spectrum. The most commonly used ML methods are tree-based models, regression (excluding simple linear regression), and neural networks (Chart 18). Models in the 'other' category included Bayesian approaches, alternating least squares, grid search techniques, and image recognition techniques. Some firms were unsure about the underlying model because it was created and developed by a third-party provider.

Chart 18: Tree-based models are the most popular ML techniques



Firms often use tree-based ensemble methods, such as 'random forests' and 'XGBoost', which consist of a multitude of – often large – decision trees whose individual predictions are averaged. Respondents said these methods have been relatively successful for prediction in traditional financial data analysis contexts. Neural networks tend to be used to make forecasts based on historical information and find complex relations between non-linear variables. Respondents also used some ML regression techniques, such as logistic and penalised regression, for forecasting and time series modelling. Most respondents' ML applications used, on average, a combination of two methods. In one use case, a firm used nine separate ML techniques in a single application. (See Box B for descriptions of different ML methods.)

Almost all the use cases where ML was deployed in both advanced and critical areas relied on structured data. Less than 10% of respondents used novel data sources in critical areas (such as unstructured and semi-structured data), but these were more common in some of the advanced deployment case studies.

Model validation

Model validation is key to ensuring ML models work as intended. These validation techniques might be used in the pre-deployment phase (where the application is being trained and tested) or post-deployment (where the application is live in the market) to continuously assess if the model and ML application is performing within accepted thresholds.

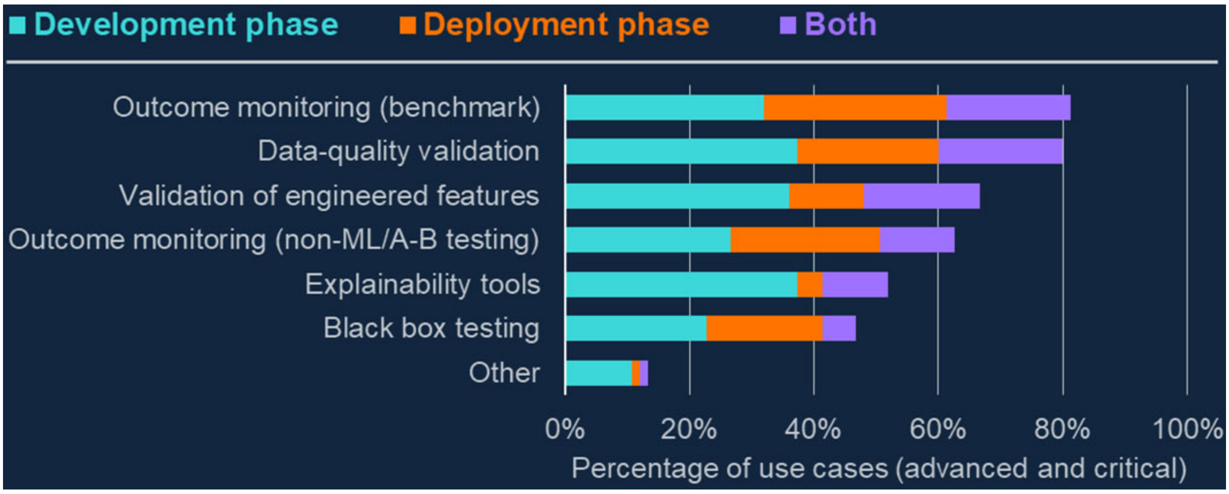
Table B: Firms use a variety of validation techniques to assess ML

Validation method	Description
Outcome monitoring against a benchmark	Decisions or actions associated with the ML system are monitored using one or multiple metrics. Performance is assessed against a certain benchmark value of those metrics.
Outcome monitoring against non-ML model or A–B testing	Decisions or actions associated with the ML system are monitored using one or multiple metrics. Performance is assessed by comparing it to the performance of a separate, non-ML model. The same approach is used in A–B testing (also known as split testing).
'Black box' testing	Input-output testing without reference to the internal structure of the ML application. The developer 'experiments' with the model, feeding it different data inputs to better understand how the model makes its predictions.
Explainability tools	Tools aimed at explaining the inner workings of the ML model (going beyond input-output testing).
Validation of engineered features	Engineering features used in the ML application are scrutinised, including potential impacts on model performance.
Data-quality validation	One or more techniques are used to ensure potential issues with data (such as class imbalances, missing or erroneous data) are understood and considered in the model development and deployment process. Examples of these include data certification, source-to-source verification or data issues tracking.

Sources: Bank of England and Financial Conduct Authority.

Chart 19 summarises which ML model validation techniques and frameworks are most frequently used by respondents (as described in Table B). The most common method is outcome-focused monitoring and testing against benchmarks, which enables firms to scrutinise how ML models perform against historical benchmarks in terms of profitability, customer satisfaction or pricing, for example. Data-quality validation, including detecting errors, biases, and risks in the data, is the next most frequently used method. Overall, these methods were used by 81% of the respondents. In over half of the cases (63%), outcomes were benchmarked against a non-ML model. Black box testing techniques were used in less than half of the cases.

Chart 19: Outcome monitoring and testing against benchmarks is the most common ML validation technique

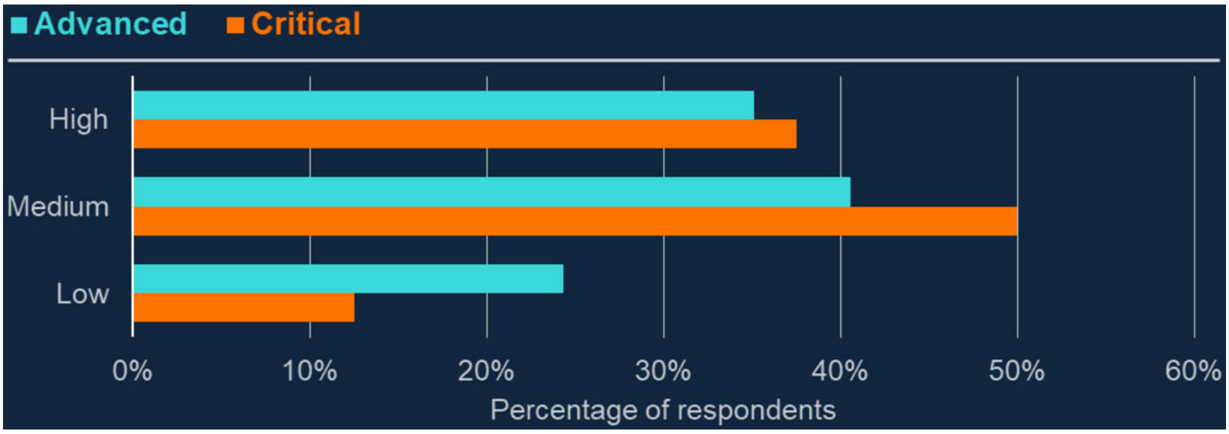


Complexity

The survey asked firms a range of questions about the complexity of their ML applications. Firms often mentioned that it was difficult to clearly define what model 'complexity' means but attempted to make an assessment based on the number of components, data sources, and algorithms in the ML application. Given this, it is clear that there are significant differences in the way firms rate the complexity of ML models and applications.

Respondents gave their best estimates of the complexity of the models and related processes. These were then grouped into three categories ('Low', 'Medium', or 'High'). For example, some firms emphasised the low complexity of the model, considering this to be important given that the area for deployment was critical or there was a critical need for the model to be explainable. Others highlighted the large volumes of data processed, varying data sources (including third parties), the interaction of different applications, and the number of trainable parameters, as factors that would result in the application to be considered medium or high complexity.

Chart 20: Respondents rated more than a third of ML use cases as 'High' complexity



The majority of respondents design and develop ML applications in-house. However, they sometimes rely on third-party providers for the underlying platforms and infrastructure, such as cloud computing (see above in Section 2.4). In one case where ML models were provided by a third party, the firm stated that it was difficult to classify the complexity of the models as they did not know the types of underlying models nor did they have oversight of the development of the ML applications. As summarised in Chart 20, two thirds of ML applications are considered to be of medium to high complexity.

Safeguards

Firms use a range of mechanisms and controls to manage the risks associated with ML applications. This includes the novel challenges related to ML, such as the additional complexity of ML techniques, issues with explainability, and the continuous lifecycle of some ML applications.

42% of respondents said that they use some form of monitoring but did not specify the safeguards in place for the applications. The most common controls among respondents are 'alert systems', 'human-in-the-loop', and 'back-up systems'. 'Alert systems' flag unusual or unexpected actions to employees, 'human-in-the-loop' are systems where decisions made by the ML application are only executed after review or approval from a human, and 'back-up systems' (sometimes also known as 'shadow models') perform the same or similar function as the ML application and can be used instead of it if needed.

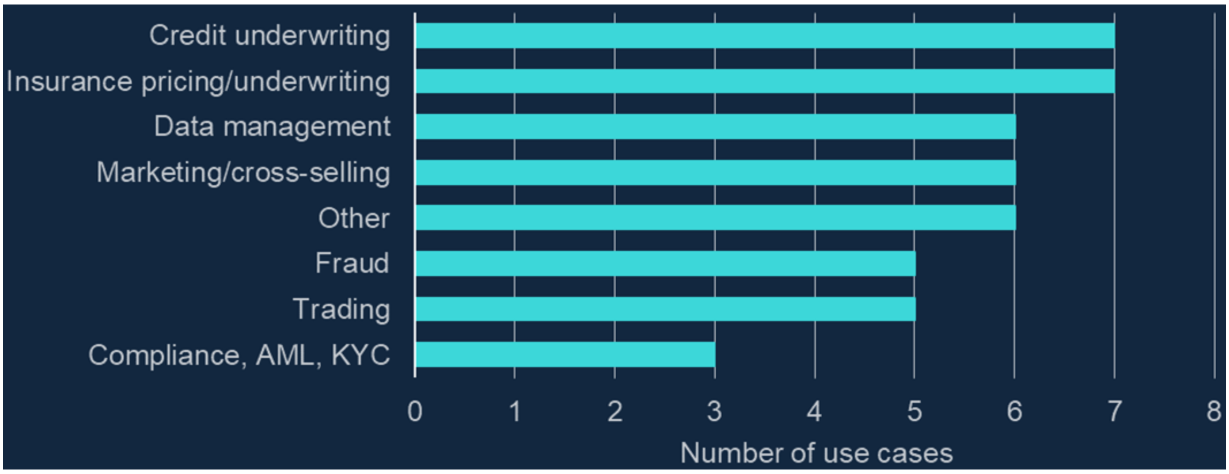
5.2: Cross-firm themes

Most advanced use cases

Overall, the most common areas where ML use is most advanced were (Chart 21) credit underwriting and insurance pricing/underwriting. Among insurers, the most advanced deployment of ML is typically within their core business to support either pricing or underwriting. Other sectors showed greater variation in the most advanced use case.

Other examples of ML models in the most advanced stage of deployment were: expected loss accounting models; claims handing; monitoring for insider trading or market manipulation; to direct queries within customer interfaces; 'know your customer' (KYC) checks; trading strategy and execution; payments authorisation; and staff wellness programmes.

Chart 21: Areas where respondents reported their most advanced ML applications



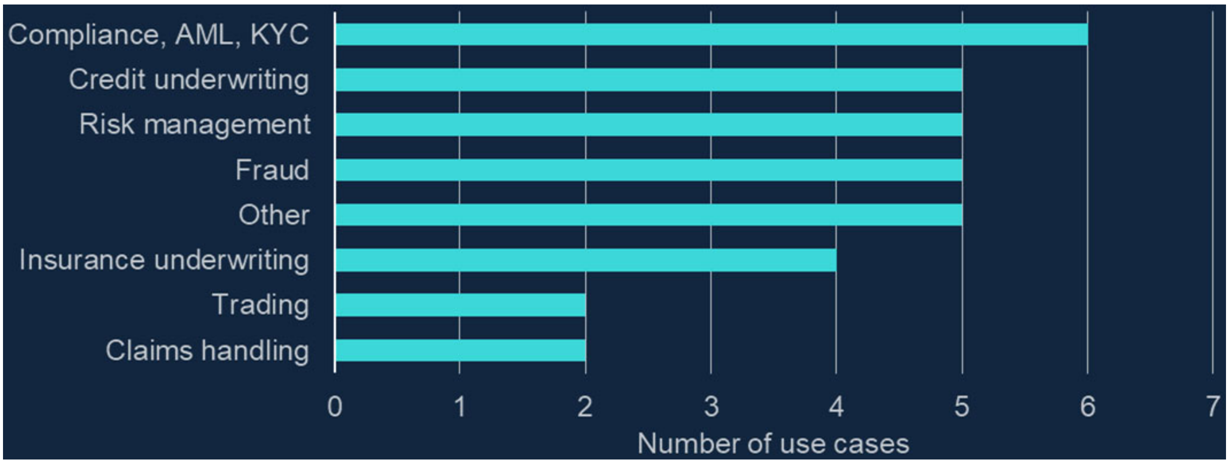
Most critical use cases

The uses of ML within critical areas varied considerably, including within similar types of firm (Chart 22). Those most commonly cited were as an input to compliance, AML and KYC. The critical use cases for credit underwriting were similar to the advanced ones, with ML used to support lending decisions, sometimes as part of scorecards and sometimes as a direct input to automated underwriting (in retail credit).

Within risk management, respondents described how ML was used within models predicting expected cash flows, redemptions, delinquencies, inappropriate account use (eg a personal account being used for business purposes) or excess losses.

Other activities in critical areas making use of ML were: insurance underwriting (particularly within life insurance); claims handling; monitoring for insider trading or market manipulation; informing trading strategies; directing queries within customer interfaces; managing fund transfers; IT security; and managing a network of third-party experts. Relative to the most advanced cases of ML deployment, respondents were much less likely to cite marketing or cross-selling, with only one respondent describing the use of ML in this area as critical (the specific example being in the optimisation of mortgage rates to retain customers).

Chart 22: Areas where respondents reported their critical ML applications



5.3: Prominent use cases

Insurance pricing and underwriting

Respondents gave examples of how ML applications are used in motor risk pricing and life insurance underwriting, as well as in making use of novel data sets to improve pricing or risk management around other products. For example, one insurer uses an ML-based application to pre-approve consumers for life insurance using data already available from bank accounts and credit rating.

In motor insurance, firms use ML models to predict the frequency and severity of expected claims for all new business and renewal quotes. Some motor technical risk models are being migrated from traditional generalised linear models to ML models, which use techniques like gradient boosting. These models seek to predict the cost of a written policy and can be used in subsequent processes ranging from performance monitoring to pricing. Some respondents use telematics risk modelling based on Deep Neural Networks to estimate driver behaviour and, thereby, predict the magnitude of the claim and determine the premiums charged to the consumer. In other firms, the assessment of motor claim liability to provide a recommendation to the claim handler is based on ML models. This involves automated early liability decisioning and can improve subsequent claim journeys and outcomes for the consumer.

Credit underwriting

ML use cases around credit underwriting typically supplemented existing credit scorecards, either when a customer applied for a loan or as part of a pre-approval process. ML applications are used to process unstructured data or large data volumes, significantly reducing the need for manual processing.

However, one firm described their most advanced use of ML as determining the key input in deciding whether or not to extend credit. In this case, the ML algorithm and model uses features at the point of new credit card applications to predict likelihood of customer delinquency within the first year of the loan.

ML-based credit decisioning in relation to personal loans was not widespread among the respondents but starting to be introduced by some firms, including for products like loan underwriting in auto finance. One model in the early stages of operation predicts the probability of the applicant entering significant arrears in the first 12 months of the loan being opened. Real-time authorisations are also increasingly specific to the personal account of the consumer drawing on historical authorisation data.

Marketing

ML applications can be used for various aspects of marketing and customer engagement. Instead of presenting the same default combination of products and services to all potential customers, ML applications use the inputs provided by the potential customer as part of the online quote journey to predict a bundle of tailored product or service options. This can then be further manually adapted to the consumers' needs. Other ML applications can be used to prevent customers' policies from lapsing. Some firms use product recommendation ML applications to identify new product sales opportunities. One respondent noted that this is the most advanced ML application owing to its consumption of broad financial behaviour features from global data assets.

Fraud prevention and anti-money laundering

As this survey and the 2019 survey found, ML is a becoming a common technique for fraud detection and AML in payments transactions, especially for card authorisations that flow through the payments network. As transactions are processed, the ML application assigns a risk score, allowing card issuers to stop potential fraud losses before the transaction goes through.

Similarly, automated fraud screening of customer transactions on the retail non-plastics channels (internet banking, telephony, branch and Open Banking) is taking place. As fraudulent transaction risks faced by investments funds are growing, ML-based transaction monitoring and detection capabilities are developed. This often involves a combination of ML and rule-based engine monitor activity and highlights anomalies in bond claims and online portal journeys.

Box B: ML methods

Advanced regression includes penalised regression techniques such as ridge regression and least absolute shrinkage and selection operator (LASSO). Penalised regression involves the application of regularisation techniques to a standard regression models (such as logistic regression). This penalises the regression weights, leading to models that use less input variables, or generalise better to new data. These models tend to be less complex and more interpretable.

Tree-based models use decision trees, which produce an output using a cascade of binary (yes/no) decisions, they are capable of handling both regression and classification tasks. The most common algorithms in this class are typically ensemble methods, in which an output is determined by a collection of decision trees that vote on or average their individual outputs.

Neural networks are loosely based on neuronal structures in the brain and consist of layers of nodes each having weighted connections to nodes in other layers. The neural network learns by adjusting these weights in response to training data. Neural networks can be used for both regression and classification and have wide-ranging applications including computer vision and natural language processing.

Natural language processing (NLP) itself involves the application of algorithms – often neural networks – to identify and extract natural language rules such that unstructured language data are converted into a form that computers can understand.

Dimensionality reduction techniques reduce the number of variables under consideration by constructing combinations of those variables to form a smaller set of principle variables. Approaches can be divided into feature selection and feature extraction.

6: Conclusion based on survey findings

Overall, the adoption of ML in UK financial services has increased since 2019 and is more mature and widespread today. 72% of the firms that responded to this (2022) survey reported using or developing ML applications. This trend looks set to continue and survey respondents expect the overall median number of ML applications to increase by over 3.5 times over the next three years. That is why the Bank and FCA will continue to monitor ML developments and conduct similar surveys going forward.

Firms now use the technology across a greater range and variety of business areas. While AML and fraud detection applications are widespread, ML is also used: in banks, for enhanced credit risk analytics; by other lenders, to automate loan underwriting; in insurers, to analyse driving behaviour and claims risk; and, among investment and capital markets firms, to extract unstructured alternative investment data from multiple sources.

Moreover, the ML applications that are used in these business areas are more advanced and increasingly embedded in day-to-day operations. 79% of ML applications among are respondents are deployed across a considerable share of business areas and/or are critical to the business area. In contrast, in 2019, 44% of applications were still in the pre-deployment phase (ie proof-of-concepts) and only 32% were deployed across a considerable share or all of a business area.

The increase in ML adoption is mirrored by an increase in the number of firms that are taking a strategic approach to the technology. The majority of respondents in the 2022 survey (79%) had some form of strategy for the development, deployment, monitoring and use of ML. Similarly, many respondents have developed governance frameworks to manage their use of ML and 80% of respondents to this (2022) survey say this fits into their data governance framework. This is promising as the Bank and the FCA consider good governance essential for the safe and responsible adoption of ML in financial services.

Safe adoption of ML also means that consumers, firms and the wider financial system can benefit from the technology. Currently the greatest identified benefits are from enhanced data and analytics capabilities, increased operational efficiency, and improved combatting of fraud and money laundering. While small overall, all benefits are expected to increase over the next three years as the technology becomes even more widely used, which may improve the personalisation of financial products and services as well as overall customer engagement.

However, there are trade-offs with the use of ML and the technology can pose risks. Respondents consider the current levels of risk to consumers and firms as low to medium. The top risks to consumers related to data bias and representativeness, while the top risks to firms are the lack of explainability and interpretability of ML applications.

Firms are also contending the practical constraints to ML adoption and deployment. The greatest constraints are associated with legacy systems and the associated technology infrastructure at firms, which was the case in 2019. These are followed closely by the lack of sufficient skills. For the Bank and FCA, it is helpful to understand that almost half of the firms who responded to the survey said there are regulations (for which the PRA and/or FCA are the competent authorities) that constrain ML deployment, with 11% of those saying that these are a large constraint. Of those who thought that regulation was a constraint, over half said it is a lack of clarity with existing regulation.

The Bank, PRA and FCA have published a Discussion Paper (DP) on how the current UK regulatory framework applies to AI/ML. The DP also explores how policy can best support further safe and responsible AI/ML adoption and whether additional clarification of the existing regulatory framework may be helpful. Readers of this survey are encouraged to engage with the DP and submit a response.

7: Acknowledgements

The authors of this report are Kathleen Blake, Mohammed Gharbawi, Oliver Thew, Seema Visavadia (Bank), Leo Gosland and Henrike Mueller (FCA).

We are grateful to Tom Mutton, Louise Eggett, Jessica Rusu, Ian Phoenix and Ravi Bhalla for their helpful comments and support.

We would like to thank Marvin Tewarrie, Zac Lovell, Itua Etiobhio (Bank), Haydn Robinson, Vanessa Do, Nishad Hoque and Joyce Wong (FCA) for assistance with the data collection and analysis. We would also like to thank Rahul Vyas for their support in getting the report ready for publication.

We are also grateful to colleagues from across the Bank, FCA and PRA for their input, including the supervisors of surveyed firms for their support throughout the process. Finally, we would like to thank all of the firms that participated in the survey for their input.

1. Individual firm averages across all benefit and risk categories rated on a scale of 0 to 1, with 1 being highest risk or benefit and 0 being no risk or benefit.