Machine Learning in Electronic Trading

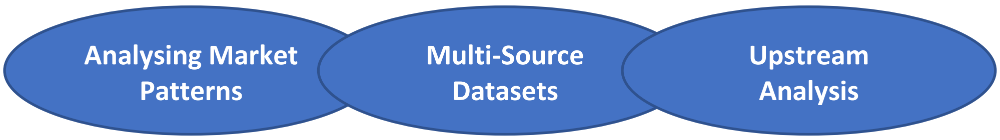
Professors Fethi Rabhi and Nikolay Mehandjiev[1]

Machine Learning (ML) is increasingly important in electronic trading. To analyse the state of its adoption in 2020, we conducted a number of semi-structured yet theory-driven interviews with industry experts from UK, Australia, USA and Germany. This report summarizes our findings organised around 3 questions[2]: (1) What is ML and how is it being applied in electronic trading? (2) What are the key management and technical aspects of ML in electronic trading? (3) What are the gaps and areas of future research and actions?

1 ML in Electronic Trading

Lack of clarity. This was not surprising since people interpret ML in so many different ways. Many respondents used *informal definitions* like "when there is no explicit algorithm", the use of big data to make decisions, extracting knowledge, etc. Interestingly, some defined ML by contrasting it with Artificial Intelligence (AI) and stressing that the two are different. Many associate ML with a specific technique such as neural networks or text analysis. Although widely published literature considers ML as "*a methodology in which programs fit a model or recognize patterns from data, without explicitly programmed and with limited or no human intervention*", many respondents consider ML just an *incremental improvement from statistical methods* in terms of complexity.

Academic research is scarce. This is despite recent fundamental changes in electronic trading such as an increase in the number of regulations and a rising number of trading venues resulting in complex decision-making across multiple jurisdictions and time zones. The largest body of academic literature focuses on *analysing financial market data patterns* (e.g., identifying stock price movements), as ML models can potentially address the limitations of traditional econometric models and able to process complex, imprecise, and large amounts of data. In addition, ML methods enable the use of different types of data (qualitative and quantitative) and are not subject to rigid assumptions such as those imposed on econometric models. As the direct use of low-level market data in any type of ML model is not recommended, there is also a large body of literature that is developing innovative ways of *gathering and processing data from multiple sources to produce good quality datasets*. For example, news sentiment datasets enable the exploration of relationships between news content (including social media) and financial markets. Finally, although decision-making has become very complex at trading desks, most academic research tends to *focus on the "up-stream" part* (like investment analysis) and less on *"down-stream"* activities such as trading order placement, intraday automated trading and operational efficiency problems during high frequency trading.



In conclusion, there is *a wide gap between academic finance and professional finance* when it comes to analysing big datasets with ML methods and putting these models into operation within a realistic trading environment. This motivated us to seek out new evidence from industry experts.

---

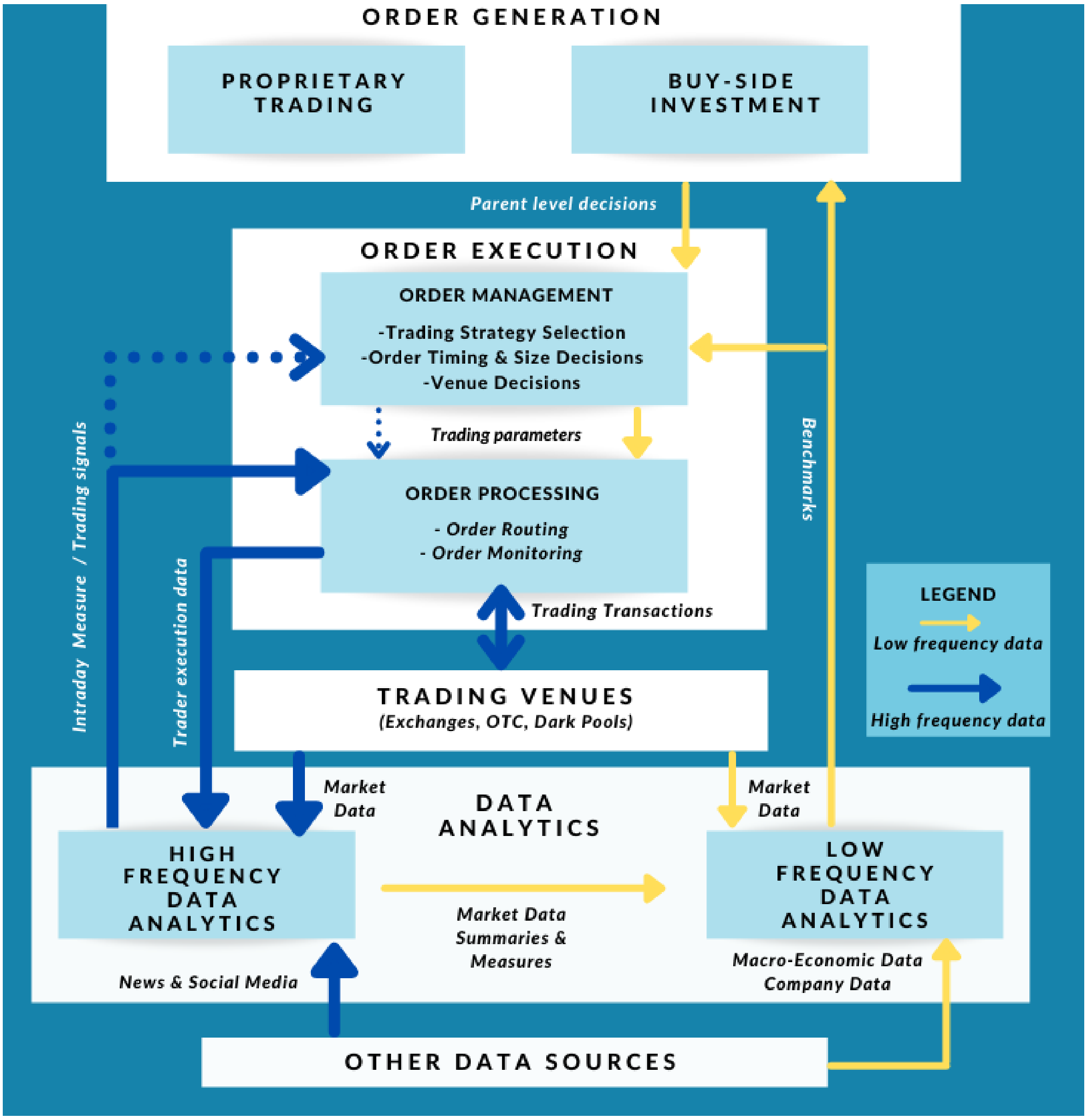
[1] Disclosures: This paper has been sponsored by the Plato Partnership. However, the views expressed in the paper are the views of the authors and do not necessarily reflect the views of the Plato Partnership. Fethi Rabhi is a Professor in Software Engineering at the University of New South Wales and Nikolay Mehandjiev is a Professor in Information Systems at the University of Manchester.

[2] A detailed report of these findings (complete with references) is also available from the authors

ML is used more in the 'low frequency' parts of electronic trading. It is hard to pinpoint where exactly ML techniques can be used in electronic trading systems because such systems have become highly complex and can be viewed from different angles depending on a participant's role (buy-side, sell-side etc.) as well as business model (agency trading, market making, internalisation etc.). For example, the information flow within a sell-side system to manage customer and proprietary order flow will comprise the components illustrated in Figure 1. At the upstream level, *order generation* comes as a result of a client's decision-making (e.g., buy-side investment, proprietary trading etc.) which will specify some values and constraints regarding order size, possible venues, timing etc. These will need to be preserved when the order is further refined downstream. *Order execution* consists of the activities typically conducted by trading desks that can be differentiated by:

(1) *the order management part* where trading decisions are made (e.g., trading strategy selection) and order parameters are further refined (e.g., order timing and size, venue selection) and
(2) *the order processing part* directly interacts with *trading venues* on which trades take place. The *data analytics* component analyses data from different sources to support both pre-trade and post-trade analytics activities (data flow can be high or low frequency).

Figure 1. An Example of Information Flow in an Electronic Trading System



Many interviewees referred to specific places where ML is useful such as determining order timing and size decisions. *ML was perceived as less useful for the 'high frequency' trading flows with simpler decision choices*, where conventional statistical techniques were more useful. *ML works well in areas as such trade strategy selection where conventional automation is challenging*. In general, ML can work hand in hand with any type of model as it allows software to "twiddle the knobs", i.e., to search the space of possible combinations of parameter values, to find good settings.

Most interviewees considered *structured market data as primary source for ML methods*, particularly fine-grained data (e.g. order-book) which originates from exchanges, but also derived data such as prices and market measures. ML forecasting uses not only historical data of the same type but also *other sources such as macroeconomic data, company data and even air quality*. Unstructured data was mentioned often, yet, surprisingly, *interviewees were quick to point out that its value has been quite disappointing for ML, especially* social network data. *An opinion was voiced that analysing unstructured data is more suitable for investing and long-term trading rather than for real-time e-Trading.*

ML helps companies increase the scale and scope of their operations. ML was perceived as helping companies to *increase the scale of their operations* by supporting efficient design of core trading algorithms, which can then scale outwards to other financial instruments, markets or exchanges. ML can also *identify patterns* more accurately and quicker than using conventional observations, bringing competitive advantage. The *costs are also reduced,* allowing decisions to be based on larger sets of finer-grained observations. ML was also reported to help address problems which are difficult to formalise and hence to automate, thus *extending the scope* of what an organisation can try.

ML helps companies learn. In addition, ML helps companies improve their operations by learning from past experiences. One respondent stated that one of the best features enabled by ML was the ability of an algo-wheel to gather real-time feedback on its performance, and this data can help determine the extent to which the new learning is incorporated in core algorithmic code. Thus, ML helps *avoid "muscle memory"* cognitive biases, where successes are remembered better than failures, leading to traders always choosing the same trading algorithms.

Assumptions constrain the applicability of ML. ML is based on two fundamental assumptions: (a) the observed data is not random, which contradicts the random-walk presumption about markets; and (b) the observed data will remain static for a sufficiently long period for the ML to work after training. The problem with the second assumption is that markets change in both dynamics and nature, as other traders keep discovering new ways to deliver excess returns.

The use of ML is constrained by data quality and availability. The data available often lacks the quality necessary to create effective ML. Training ML on a subset of the data brings set effect error and results in over-fitting, where *predictions are brittle when applied to out-of-sample data*.

Other inherent constraints on using ML for Electronic Trading. ML is perceived *too slow* to bring value downstream where the speed of execution is much higher than upstream. *The Goodhart's Law* was also seen as expressing itself, with ML delivering optimal results on isolated parameters but failing to achieve the overall aims of gaining the best profit for the overall trade. Often ML-driven algorithms would not trade "unprofitable" parts of the order, leaving them to be traded at a loss on the next day. Another problem is the *"black-box" nature of ML*, which impedes efforts to convince regulators that ML-driven trades comply with regulations. The final perceived constraint is the scarcity and the cost of ML experts and the infrastructure costs necessary to run ML with sufficient performance.

2 Key Management and Technical Aspects

Managing ML should be aligned with business strategy. Our respondents supported the need to align the introduction and control of ML with the business strategy, and to communicate clearly business goals to the ML activity in order to gain the best advantage from it. According to one respondent, ML was seen as "tool suitable for the job", and so different company goals would pose different requirements to the ML operations. Many respondents said that their companies are client-driven, so the scope of an ML activity is focused on “executing clients’ orders at minimal cost.” There were different opinions related to governance roles and policy and how to allocate roles to various committees, *e.g.* whether to have Model validation or Execution Review committees, a specialist ML team, etc.

(1) For trading companies and companies focused on pricing of instruments (e.g. derivatives market-making firms), *ML was seen as a supporting capability* improving their core activities. The proposed configuration there was for the ML sub-system to steer the core algorithmic trading part of the system in ways which do not break regulatory compliance.
(2) For companies specialising in adding value by developing innovative trading algorithms and providing pre-trading analytics services, *ML is perceived as central activity* and therefore subject to full regulatory compliance policies.

Need for Transparency. Our respondents were univocal about the challenge for ML to be understood by non-technical regulatory agents. Ensuring transparency of ML models is impeded by their black-box nature. Some interviewees reported that their *processes for approving algorithms* were “heavily governed and formalised” and focused on testing the algorithm’s performance. Certain aspects of these processes were perceived as slow to “catch up” with the specific nature of ML models. For example, the usual size of the data sample used (one day was mentioned) is sufficient for conventional algorithms yet too small for ML approaches, with significant *distortions in the results* produced based on the specific characteristics of that specific day of training. The statistical complexity of ML models meant that they are *perceived as black-boxes even by the computer-science specialists* responsible for their implementation. *Recording the features and documenting the testing conditions* for the ML models were deemed important to ensure the success of ML when applied to electronic trading.

Scope-of-Testing should be documented. Respondents highlighted that “sell-side lacks incentives for sufficient validation”, with no standards forcing them to include “out-of-sample performance” in the documentation made available to buy-side users. This lack of performance and scope-of-testing data creates significant risks of over-fitting, and difficulties in comparing the performance of two alternative implementations of a ML technique, or two different ML techniques when selecting the best tool for the job at hand. Even if the outputs are deemed correct on observation, this lack of transparency impedes assuring risk managers that the tool will continue to produce correct outputs for different combinations of input parameters.

Explainability is particularly important for ML. The problem of trying to explain complex models and black-box operations is endemic with the wider adoption of artificial intelligence throughout our society. However, it is particularly important for ML in electronic trading because of the speed with which we need to manage risk and the monetary impact of ML-based trading decisions.

Supervised learning models are widely used in electronic trading. Interviewees mentioned supervised learning frequently as having numerous applications in electronic trading (regressions for market forecasting, classification of trader behaviour etc.). Neural networks, in contrast, were considered poorly performing for larger data sets, susceptible to overfitting, difficult to tune, etc. Other ML methods mentioned include non-linear extensions of classic regression techniques, tree based methods such as Random Forest, and ensemble methods where the best model is selected based on accuracy score or other performance measure. Unsupervised ML, where patterns in data can be identified without a training set were only occasionally mentioned, with clustering criticised for difficulties in balancing cluster populations.

Lack of clarity motivates the need for ML taxonomy. Participants answers indicate perceived overlap between different ML methods and mention many techniques that are not proper ML techniques but are used alongside ML such as pre-processing, dimensionality reduction, feature selection and visualization. The establishment of a taxonomy of ML techniques used for electronic trading can thus support its uptake.

Testing of ML models is a very important yet complex task. Our respondents confirmed that thorough validation is being applied across all algorithms and ML models used for electronic trading in their organisations. Testing was mentioned as taking place at four levels:

1. *Is the ML model selecting the right features?* Are these present in the data fed into the model? Huge datasets with high dimensionality need pre-processing such as dimensionality reduction.
2. *Is the model correct?* Would it work outside the data sample used? Is the data with sufficient volume to ensure statistically valid results? Have we removed start-of-trading test data?
3. *Does the model work in determining the correct trading strategy when back-tested?* There is a need to demonstrate the consistent operation of the algorithm across a number of data samples and across entire data sets, using a classic experimental design by sector and by market.
4. Is the model working in practice? Outcomes should be measured against "benchmarks" selected to match the goals of trading. Human traders also need to monitor for strange patterns (for example black fields or all orders submitted to the same algo).

Strong technical backing is needed to safeguard operations. In addition to the use of conventional IT security mechanisms (e.g. firewalls, duplication, backup system etc.), a number of precautions should be taken at different levels:

1. *At the model design level*, we need to automate important tasks such as *cross validation* (segmenting data into similar buckets when training and testing), *back testing* (using randomized testing and randomized agents whenever necessary) and other model validation techniques based on real output data (e.g. comparing predicted with actual outcomes).
2. *When implementing the software,* ML processes need to work within the constraints imposed by the trading core, which is often fully compliant and validated. Voting or averaging is also mentioned alongside duplication of code to avoid spurious decisions caused by a software bug.
3. *At deployment time,* the software must be gradually deployed on a partial use basis.
4. *During operations,* various safeguards around the system need to be implemented such as applying limit prices not far away from the arrival price to avoid disrupting the market.

Automated alert systems are considered critical. Performance needs to be monitored in real-time and kill switches must be activated to prevent disasters. The system should stop trading if unusual conditions or outputs are encountered as at the start of the COVID pandemic, or on Brexit day. The use of a "dead-man's handle", where the system will realise if human is not monitoring it and will take appropriate action (i.e. alert another human or stop trading) is also regarded as important.

3 Gaps and Future Research Areas

The study highlighted the following recent trends in industrial innovations:

1. *Interlocking functionalities.* Important changes in the regulatory frameworks (particularly the MiFID regulations in Europe) have resulted in a proliferation of software components performing different functions in the trading cycle. Therefore, designing a new ML model requires a good understanding of its impact on other parts of the system. For example, it is hard to assess whether such models will violate or not compliance requirements with relevant regulations.
2. *Explainable models.* The increased use of sophisticated ML techniques fuelled by (a) the rise in the availability, variety and volume of data throughout the trading cycle, and (b) by the diversification in the trading instruments, has resulted in formidable challenges to ensure ML models are "understandable" for regulators, accurate and aligned with business goals.
3. *Real-time data processing workflows.* Current models mostly rely on static data, yet real-time data becomes increasingly important, so we need cost-effective solutions able to acquire, process and analyse real-time data. In addition, the need to integrate systems from different vendors has motivated research in workflow design and efficiency.
4. *Risk quantification and management.* the automated trading risks are poorly understood and rarely quantified. We need to research risk management culture in financial institutions and regulators.

In addition to directly addressing the challenges above, the *research community* should also tackle the following more general issues to ensure efficient and effective application of ML in electronic trading:

Information representation for ML powered analytics: Feeding hundreds of low-level indicators into ML algorithms is impractical and empirical results have confirmed a severe drop in the performance of trading systems in such cases. We need research into better ways of representing and abstracting data, especially textual data. New types of information that can represent the richness and diversity of existing institutional environments, traders' behaviour including emotions, and market regulations need to be defined. This raises many challenges such as ensuring data quality, managing provenance information (needed for explainability) as well as organizing metadata when combining data from multiple sources.

Engineering sophisticated ML-based systems: We need to include constraints, incentives and biases of decision makers (fund managers, traders), and their impact on market valuations. In addition, new models for quantifying risk are urgently required. For these reasons, interest in the use of deep learning methods is growing for their ability to detect interactions in the data that are not captured by existing financial or economic theories. In addition, the use of ML methods as a part of a bigger system in which other models co-exist has many practical advantages. We expect to see shift of focus from individual ML models to engineering complex ML systems for electronic trading.

Automatic trading strategy elaboration and execution: the use of ML techniques in the elaboration of a trading strategy and its parameters such as order size and venue choice, taking into account a wide range of goals, time horizons and benchmarks, is still in its infancy. Trying to capture individual effects for a large number of financial instruments increases the number of predictors and the level of noise, thus reducing the performance of the trading system. We need deeper empirical studies involving the hyper-optimization of all parameters. It is also expected that optimization of trading processes will be increasingly automated, guided by applying ML techniques to learn from experiences captured in vast amounts of execution data.

For *organisations*, we recommend the following actions arising out of the identified challenges of applying ML to electronic trading:

1. Recognise the need to have an **ML governance and implementation framework**, and work towards developing industry standards and guidelines in this area with peer organisations.
2. Invest in **explainable solutions** and ensure that the **risk management boards have inter-professional membership** with stakeholders from the whole organisation including IT support.
3. Support community-wide **multidisciplinary educational and research initiatives** involving different fields such as computer science, finance and econometrics that can facilitate the development of expertise and education resources in this area.

An overview of these directions is presented in Figure 2 below.

Figure 2. Suggested Directions for Future Work