Machine Learning and Credit Flow trading

How to do more with less in a capital constrained environment

Aldous Birchall
aldous_birchall@hotmail.com

Software to the rescue…

- Credit trading businesses have had a tough time of late: everyone is being forced to "do more with less".
- The winners will work smarter. The losers have cut costs which in many cases is synonymous with shutting down.
- Working smarter until recently meant shuffling org charts, risk limits, and product mix.
- However, the recent explosion in machine learning technology represents an alternative way of 'working smarter' in terms of managing and executing flow trading.

What is machine learning?

- Two main categories of machine learning ('ML'):
- Supervised learning algorithms automatically learn from existing data to make predictions given new information. They can categorise things or predict numerical values.
- Unsupervised learning algorithms automatically find hidden patterns in data, essentially looking for 'clusters' across multiple dimensions. This is a descriptive rather than predictive method.

What can machine learning do?

- Subject to constraints, machine learning algos will find optimal, non-trivial insights in complex data sets.
- They are especially good at generating robust empirical models that can do things like image recognition, asset price prediction, or autonomously drive cars.
- Machine learning algos are incredibly flexible. You can throw all sorts of data into them and they can combine the interesting bits whilst ignoring noise. This used to be incredibly difficult using standard statistical methods.

Why all the sudden interest?

- E-commerce, social media, and internet search have generated huge datasets that can be monetised.
- The resulting surge in interest in what is now being called data science has led to the proliferation of off-the-shelf machine learning software libraries*.
- What was until recently a specialist field of computer science is now available to anyone with reasonable numeracy, computer literacy, and some relatively straightforward training.
- The latest machine learning algorithms are extraordinarily powerful and can be run on desktop machines. They represent a quantum leap beyond what was considered best practice less than a decade ago.

*Matlab statistics and machine learning toolbox, Python Scikit learn, R machine learning packages.

Investment banks have gone from tech leaders to laggards

- The prevalent mode of data analysis remains excel based with some inflexible bespoke analytics (generally in-house data tools and Bloomberg functions).
- The analyst mindset revolves around easy exposition of ‘trade ideas’ which is not compatible with complex analytics.
- Whilst banks have been dealing with post-crisis realities, data analytics in other sectors have made a generational leap.
- Retail finance, insurance and ‘fintech’ all have business models based to varying degrees on machine learning methods.
- Many hedge funds have deployed machine learning methods with great success, especially in equities. Transaction costs have made this a more difficult proposition in credit.

Analytically, some areas of credit trading are stuck in the 1990's

- Credit instruments have always had the problem of sparse data compounded by instrument stratification.
- The result has been a research paradigm heavily focused on business fundamentals and documentation.
- Quant activity in credit has largely been limited to correlation products which for obvious reasons is now viewed with suspicion.
- Credit trading businesses have been preoccupied with punitive capital constraints.
- The good news is the intersection of improved data quality, *some* product convergence and the sheer power of machine learning algorithms has opened the door to a large opportunity set for ML methods in credit flow trading.

What can traders do with machine learning?

- Produce robust empirical models of single instrument and portfolio behaviour that takes into account many inputs *and* the complex interactions between them. This can give us:
  - Optimal hedging strategies for single trades and portfolios.
  - The best risk adjusted relative value plays.
  - Predictions on where any market variable *should be* given a particular scenario.
- Classify instruments, clients or even market regimes. If you can think of a way of labelling something, you can train an algo to do if for you.
- Search for hidden patterns across markets, clients and traders.
- Predict the 'market impact' of a trade as function of instrument, trade size and client.
- Search 'strategy space' for effective trading strategies.

…can’t we do that already?

- Under the Excel paradigm, you can solve simple optimisation problems, but it can be very unreliable (e.g. solver) and limited in scope. Multivariate regression is possible in Excel, but it’s inflexible and limited to small data sets.
- Brute force parameter optimisation is possible in code, but it quickly runs into a combinatorial explosion limiting users to low dimensionality.
- *The above methods represent the lower bound (often lower than that) of the most trivial machine learning methods.*
- Equity trading has large quant teams that build bespoke applications for predictive analytics and data mining, but this is viable only because of the business scale and data richness they enjoy.
- Machine learning provides flow credit businesses with the opportunity to bypass bespoke IT investment and dedicated quant headcount to tackle these problems rapidly and with more flexibility.

Do more with ML: desk managers

- Identify clear, weighted market drivers of desk PnL.
- Construct optimal multi-asset hedges to better manage desk risk as an optimised portfolio rather than using simple risk metrics / beta approach.
- Client intelligence:
  - Behavioural patterns.
  - Market impact.
- Team intelligence:
  - Behavioural patterns.
  - PnL drivers.

Do more with ML: index trading

- Deploy optimised, adaptive market making algos.
- Produce best in class index relative value analysis:
  - vs. other credit instruments.
  - vs. macro futures.
  - 'Auto' suggest best replacement asset(s) against a client trade.
- Quasi index arb: construct weighted baskets of liquid single names and / or macro instruments to capture basis without the hassle of full replication.
- As above for block trade hedging where basis is favourable or liquidity in the traded instrument is limited.

Do more with ML: single names

- Provide traders with real time instrument rich/cheap indication and relative ranking based on potentially hundreds of market data inputs including fundamental credit metrics, economic data and ‘red flag’ classification.
- ‘Auto’ suggest alternative market neutral replacement trades after a client trade.
- Construct optimal trade and book hedges using weighted baskets of liquid instruments.
- Predict market impact of trades on specific instruments by client, trade size, trade direction and market backdrop.

Systematic portfolio optimisation

- The buy side frequently uses machine learning to construct market timing strategies for both index and single names. Relatively wide bid-offer in credit markets means that these strategies can actually work better if tweaked and applied to a flow desk:

- Use for client order pre-positioning or market neutral inventory management:
  - The trader is presented with optimal replacement trades after each client trade based on the strategy.
  - **Every client trade has the potential to become one side of a 'smart' revenue generating position whilst keeping the transaction market neutral.**

- In a limited number of cases, deploy ML driven market making algos in the most liquid products (indices being the obvious case).

The future - next few years

- All trading units will need a ‘head of data science’ to:
    - Identify, prioritise, design and deploy machine learning algos in areas that stand to benefit.
    - Understand informational limits on what is and isn’t worth doing.
    - Manage the data library to ensure quality and temporal integrity.
    - Liaise with IT and middle office to achieve these objectives.
- As regulators push for ever more transparency in OTC markets, credit markets *will* become increasingly data rich.
- To leave this technology unused in *any* data rich environment is just bad business.

The future - next 5-10 years

- Instrument liquidity will be *the* organising principle of flow desks.
- Liquid instruments will be managed as a central portfolio with ML algos decomposing risk into the most liquid factors and optimising idiosyncratic exposures on an RV basis.
- Market makers with individual books can thus spend more time further up the value chain in the more illiquid and idiosyncratic instruments.
- ‘Sales traders’ can take over liquid instrument execution with risk managed centrally.
- As the market evolves with liquid credit moving ‘on exchange’ or some other technology, a clear illiquid OTC vs. liquid product delineation will emerge.

The future – industry wide

- Individual flow desks will be defined entirely by which parts of the liquidity spectrum they occupy.
- Large, fully committed players will run full service credit trading operations.
- Smaller players will most likely run centrally managed portfolios of only liquid instruments with teams of sales traders there to facilitate primary activity and act as an institutional conduit for credit risk.

Risks

- Machine learning methods are far more robust than ‘last gen’ approaches, but like all algo strategies, they have no sense of informational context or meaningful appreciation of how other agents in the market may be behaving. They are best used to augment human decision making.

- Markets are a special and difficult data problem; essentially a strategic, self referential, iterated game. This means historical data *information content* varies over time in often quite a complex way.

- This is a non-trivial issue, and therein lies the art of non-naïve use of algos in managing risk. As enticing as it is, simple ‘engineering’ approach to algo design and risk management is almost guaranteed to end in failure.

Safeguards

- Established machine learning best practices exist to avoid over-fitting to noisy and random data.
- Data integrity is key. Raw CDS data needs to be adjusted for rolls and defaults. Bond data needs to be scrubbed.
- Domain specific knowledge (i.e. common sense to experienced traders) is key when designing and using ML algos. It is not about throwing a bunch of data into a black box.
- The bottom line is you need experienced traders trained in data science. Risk management experience *should* lead the process with ML algos there to enhance and optimise trader's decision making.

Limitations

- Focus has to be on what can be done with most liquid products including macro futures.
- Providing flow traders with ‘menus’ of good decisions mitigates some liquidity problems but there will always be some product areas that are too illiquid / idiosyncratic to benefit from some strategies beyond good generic hedging.
- Model risk: all trading approaches are subject to periodic regime shifts, algo driven or discretionary; even if it’s just because they work well and become crowded. It would be prudent to impose some independent risk framework on model driven positioning even though individual traders would be making informed choices rather than running systematic strategies per se.

Conclusion

- Machine learning methods can now be implemented by numerate, computer literate traders with relatively little training.
- In much the same way Excel created an analytical explosion in the 1990’s within finance, intelligently deployed machine learning algos promise to do the same now.
- In a capital constrained environment, machine learning offers the opportunity to help more intelligently manage risk and allocate capital, as well as understand market and client dynamics.
- Machine learning will allow credit flow desks to do more with scarce resources and over time will re-shape the business.