ORACLE®
Business Process
Services

LEVERAGING NEW DATA SETS IN ACTIVE ASSET
MANAGEMENT

SEARCHING FOR ALPHA-USING ALTERNATIVE DATA SETS & NEW GEN TECHNOLOGIES

Objective of this document is to examine the significance of leveraging alternative data in Active Asset Management, the challenges in consuming and using alternative data & the technology available for use to handle such data along with traditional data sets which would help asset managers to make swift investment decisions

RAMKUMAR SRINIVASAN
RONAK BHAGAT
SRIRAM GANESAN
PROMOD RADHAKRISHNAN

ORACLE®

DISCLAIMER

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Data is at the core of investing and finance. Active Asset managers base their analysis on historical data as well as current economic and business information for their investment decisions. Environmental, Social and Corporate Governance (ESG) are being increasingly considered for stock selection. Currently, investment decisions are based on information collected from traditional data sources like regulatory filing, press releases, Stock Exchange data, Data Service providers (Bloomberg, Reuters), Analysts reports etc. However, the issue with using traditional data is that they consist of the same datasets that other asset managers use; hence, may fail to give any information that provides a competitive edge. Also, these data sets are quite voluminous and asset managers use these data only to the extent humanly possible to them.

Today with the explosion of data that is generated by an increasingly interconnected world, combined with the emerging field of data science, has led to several asset management companies relying on alternative data sets in their quest for alpha. Today's portfolio managers have new tools in their arsenal like, Big Data, Image Recognition, Machine Learning, Natural language processing etc. to help them find this edge.

This blog examines the significance of leveraging alternative data in Active Asset Management, the challenges in consuming and using alternative data & the technology available for use to handle such data along with traditional data sets which would help asset managers to make swift investment decision.

2   ALTERNATIVE DATA: A POTENTIAL AGENT FOR CHANGE

Alternative data [1] is, put simply, any information that is non-market data. As such, any useable information or data that is not from a financial statement but can be used in obtaining insight within the investment process. The results that are generated from these data sets are often used by hedge fund managers and other institutional investment professionals to make a more informed investment decision. Alternative data will likely transform active investment management over the next five years [2]. Hedge fund management, long-only mutual funds, and even private equity managers will be impacted. Firms who don't update their investment processes to incorporate alternative data could face the strategic risk of being outmaneuvered by competitors who effectively incorporate big data investment into their securities valuation and trading signal process.

2.1   EXAMPLES OF ALTERNATIVE DATA IN INVESTMENT

On August 11, 2016, Brittany Weiss, a WBRZ (TV station in Baton Rouge, Louisiana) reporter tweeted about the fire going on at the Motiva oil refinery plant, the 28th largest US refinery, which was owned by Motiva Enterprises. Although Weiss' tweet was only one among the 350,000 tweets sent at that time and among the 500 million sent that day across the world, it allowed institutional investors to make money. This was because a large refinery going offline meant increased prices for oil-product futures in anticipation of the reduced supply. And that is exactly what happened when the media reported it hours later with a 4.8% gain in ultra-low sulfur diesel futures; a 4.4% rise in gasoline futures; and an increase of over 7% in the diesel crack spread. [4]

Some hedge funds are using satellite images to assess the real standing of the Chinese economy as the accuracy of official data has come under review. Image processing algorithms are used to evaluate snapshots in real time to understand the country's economy. Likewise, hedge funds are analyzing pictures of big-box retailers' car parks to get sales estimates. They are also looking at farm lands to determine corn yields and estimate retail sales with the use of consumer credit card transaction information. These forms of alternative data give them insights that traditional data does not. [4]

Tom Farms LLC and the digital business:- Tom Farms has gone all in on digitizing just about every part of the farming process it can -- from self-driving combines decked out with sensors that produce a stream of real-time data to mobile apps that help monitor and control irrigation systems. For Tom Farms, embracing data and analytics has helped grow the business from 700 acres back in the 1970s to 20,000 acres today, according to The New York Times. Success isn't just measured by an increase in acreage, according to Laney. Tom Farms uses technology to "reduce the need to diversify crops to hedge against weather and disease," he said [5]

2.2   ADVANTAGES OF USING ALTERNATIVE DATA SETS

- Non-Traditional data provides information on more than just financial performance, hence can provide better trend & performance predictions.
- Provides unbiased recommendations
- Large amounts of data generated continuously, hence more information for better analysis
- Additional Information aids in Investment decision and timing.
- Frequent inflow of data allows better monitoring of portfolio, prediction of trends and enhancing speed & accuracy of assets under management.

2.3   DISADVANTAGES IN USING ALTERNATIVE DATA SETS

- Require high level of commitment and investments
- Data collection process and retention needs to be carefully managed as data collected may contain personal data.
- Contextualizing the data
- Upcoming business area and hence sparsely regulated, providing little protection

3   HOW TO TURN ALTERNATIVE DATA INTO ALPHA

Asset managers are increasingly tapping into alternative data to generate alpha. They are surrounded by a seemingly endless stream of data. As technology advances, so does the range of sources from which we can gather information and alternative data, picked from uncommon avenues (3). New alternative data streams offer a clear business value to those who can utilize them, providing insights that fall outside of what the norm can offer, but they are often unstructured with significant challenges to integration with existing models. Few firms that have successfully leveraged alternative data for alpha have teams of data scientists analyzing and parsing the data. To make full use of their potential and gain a competitive edge, many organizations are finally realizing that it's critical to be able to aggregate and analyze alternative data before using it to make business and investment decisions.

4   CHALLENGES FACED BY ASSET MANAGERS IN USING ALTERNATIVE DATA

Alternative data promises to provide insights into behavior, performance and interactions. At the same time, the nature of the data sets poses challenges as not all data have the same quality. Few key issues pertaining to leveraging alternative data in Investment strategies are:

- Data Availability, Quality and Consistency
- Objectivity – Data Contextualization
- Technology Complexity
- Prudent Implementation of Big Data Strategies
- Poor systems integration
- Ethics
- Periodic review of Models and Strategies

4.1   Data Availability, Quality and Consistency

A key challenge faced by asset managers is securing access to reliable, potent and unique big data. To generate differentiated insights, asset managers are competing aggressively to secure access to unique big data sources. Some asset managers are also investing in capabilities with the 'secret sauce' such that they can source unique data and build databases for proprietary use. Given the diverse data sources, asset managers must test and validate the sources. Asset managers should also perform due diligence on data vendors and their sourcing methodologies and ensure that there are no legal or privacy issues. Notably, cleaning, validating and transforming raw data could take up to 80% of the effort in bringing the data into the investment research process. Identifying the right data sources is ranked as the biggest challenge.[6]

4.2   Objectivity – Data Contextualization

One of the major challenges in using the alternative data set being they are current information but cannot be directly converted into a forecast. For instance, a reduced foot fall in a mall can be because of weather related distress but may not reflect the strategic approach of the company that it had put in place to manage the adversity [11]. Hence data analysis without context using alternative data sets will create incorrect investment signals.

4.3   Technology Complexity

The state of art technology is the key in executing an effective digital strategy. One of the hindrances of asset managers looking to leverage alternative data is legacy IT. The trouble with overcoming this is that modern technologies are increasingly complex. While technology may become less complex in some cases, thanks to the development of smart and cloud solutions – meaning physical storage problems of the past may decline – on the other hand complexity of technology will grow because of higher expectations in automation of very complex processes. Processes that are currently manual will be automated in the near future, including the capture of complex data, administration, and integration with blockchains. [14]

4.4   Prudent Implementation of Big data strategies

The significant investments required in technology and people are another challenge for asset managers. The problem is that of many –there are one too many ideas -and the asset managers have the strenuous task of choosing where to spend their time, money or effort. As a result, asset managers are doing one of the two things at this stage –either evaluating data sources or commissioning projects and proof-of -concept to evaluate ways of integrating big data with their existing investment research process. Some asset managers, who do not want to get concerned with plausible non-exclusivity issues, are investing and developing in-house proprietary data sources. According to a senior quantitative leader and industry practitioner, a robust framework to be laid on how asset managers can develop their big data and big data analytics strategy over 2 phases. [6]

- Projects for addressing specific business challenges utilizing service providers in the AI domain implemented on existing infrastructure for which there is a prudent ROI. These are predominantly incubators in understanding both the advantages and pitfalls for a futuristic specialized solution.
- Futuristic solutions leveraging new-age technologies and capabilities such as machine learning and artificial intelligence.

Given these challenges, asset managers should initially commence an analyst-driven project model that leverages existing infrastructure and technology, external consultants and niche service providers. In a survey it was found that about 70% of the firms plan to create cross-functional teams with a mix of internal and external people consisting of quants, data scientists and fundamental analysts, while just under 60% of firms plan to exclusively explore off-the-shelf solutions and collaborate with external consultants and niche service providers.

4.5   POOR SYSTEMS INTEGRATION

One of the reasons why asset managers source ever greater volumes of data are to produce more valuable actionable insight – that is to use analytics tools to extract information from the data that will help deliver competitive advantage. Unfortunately, many asset managers are struggling to integrate systems, business processes and data-delivery formats in a way that facilitates the creation of this actionable insight. Moreover, their user interfaces are often poorly designed which lack intuitiveness, undermining the ability of analysts to work with the information they are seeing. [14]

4.6   ETHICS

Ensuring fairness while using alternative data to make decisions is one of the key challenges. Ethics is currently one of the most overlooked elements of big data science. Ethical practices enhance the ability for individuals to be able to make decisions about how, where, why and for how long the data is used. Ethical principles should include informed consent, data ownership, accountability, transparency, data protection and data access. Informed consent goes beyond merely making people aware of the terms of service or that data is being gathered about them, but a clear articulation of how data might be used, whether third parties will get access to that data, and how people can opt out or limit how much of their data is gathered.

Privacy self-management (notice and consent) gives users the option to control their data by allowing or revoking access to their data by opting in or out. However, this is not always the reality, because personal data are always collected, used, and analyzed and at times shared or abused. Consent is always not read, and if read not always completely understood, and if not accepted by the user he might have no access to a certain service, or choose an inappropriate alternative, or sometimes the lack of an alternatives, this makes agreeing to the terms obligatory and no more of an option. [13]

4.7   PERIODICAL REVIEW OF MODELS & STRATEGIES

Another challenge for asset managers is ensuring the big data -driven models are periodically reviewed and refined, for they could lose their edge as the external environment changes. Additionally, creating new big data strategies are necessary as competitors mimic successful strategies. Regulatory changes and shifts in macro and market factors such as volatility can also contribute to the failure of signals and strategies. Asset managers should periodically validate their models and strategies as the underlying regulatory and market environment changes. Asset managers may also end up using too many factors and over-fitting models and hence teams-internal and external-must be trained in basic statistics to understand big data analytics and insights better [6].

All the above issues must be confronted and managed for fulfilling the data needed to support asset management. While there are definite opportunities arising from access to users' preferences and behavior, there are equivalent challenges in sharing of data across jurisdictions

5   INDUSTRY VIEW - ROADBLOCKS IN USING ALTERNATIVE DATA

Note: Based on 50 responses.
Source: Greenwich Associates 2017 Alternative Data Usage Study

The following are the broad inferences from the above diagram

- The wide variance in responses between Asset Manager and Hedge fund indicates the varied level of progression in adapting the Alternative data. Hedge funds being early starters are already leaping ahead of traditional asset managers and hence there will be a qualitative inference difference on the questions and hence the responses

- Near consensus in Data availability, Data Evaluation challenges and Human Capital shows the issues exist not only for advanced practitioners but also for just entrants.

- Difficulty in understanding / Working with data sets is a surmountable challenge given that it is only felt widely by Asset Managers and not by Hedge Fund Managers.

- Wider the data sets and increased complexity in algorithms increases the cost of implementation can be inferred from the responses of Hedge Fund Managers.

- Lack of Explainability of complex models can be inferred by Hedge fund managers responses to "Management not convinced of data values".

The smallest obstacle to putting alternative data to use is Human expertise. As a group, the biggest roadblocks are the cost and technical implementation. [7]

6 TECHNOLOGY CHOICES FOR IMPLEMENTATION – GENERATING ALPHA

6.1 MACHINE LEARNING

Machine learning applications are those that can continuously learn from data by applying algorithms written on statistical models and make useful predictions without explicit programming. Machine Learning algorithms can be based on Supervised model (Trained by Humans) or Unsupervised model (No Human Training and uses iterative approach called Deep Learning). [8]

Machine learning algorithms analyze reams of data in a flash and generate insights, they can determine relationships between a vast number of data inputs in a way that adapts to changing data patterns

6.1.1 FEW USECASES OF MACHINE LEARNING

- **Document classification:** the system will analyze documents and, based on their content, automatically assign them to relevant categories or group them into clusters based on their similarity.

- [9] **Predicting probabilities for Trade Execution:** Many prestigious trading firms use proprietary models to analyze market liquidity. These are probabilistic model where even a trade with a relatively low probability, at a high enough volume or speed, can turn huge profits for the firms.

- [9] **Predicting probabilities for Post Trade Confirmation:** BNPP has created a tool that can predict those trades that will not match automatically and requires manual intervention.

- **Client Risk Profile Assessment:** Assess Client's Risk tolerance and investment suitability.

- **Investment Management (Robo-Advisor):** Portfolio Construction, Optimizing Portfolio Rebalance, Risk Modelling, Price Forecasting are some use cases which can be done effectively through AI and machine learning.

Machine learning is important for asset management and that most firms will be utilizing either machine learning tools or data within five years. It should be noted that Human involvement will still be critical for risk management and framework selection, but investment strategy innovation process will get automated.

6.2 NATURAL LANGUAGE PROCESSING (nlp)

Natural Language Processing (NLP) are applications that can "analyze" and "understand" human language as it is spoken and written [8]. The essence of NLP is extracting information to map fluid text into structured data..

6.2.1 APPLICATIONS OF NATURAL LANGUAGE PROCESSING (NLP)

- **User Experience:** A big focus of efforts in NLP has been in Chatbots, Bots that manage the chat communication, to understand the questions posed by the customer in natural language, provide the right answers automatically for standard questions and forward to Human officers when the clarification requires human intervention.

- **Machine translation:** There are lot of analyst reports generated internally by the Buy side organizations and much of them are available in the public domain. Machines are taught to understand the investment sentiments in the reports (Example: Opportunity, Threat, Upgrade, increase in market share etc.) and generate investment signals. However, the challenge with machine translation technologies is not in simply translating words but preserving the essential meaning of sentences.

- **Social media usage and analysis:** Using Tweets, Blogs and Facebook posts (to name a few), to understand customer sentiments to a product / company, demography analysis of product users are a few examples to generate investment signals.

7   CONCLUSION

Financial institutions are going through a wave of changes such as changing demographics, customer expectations, tighter regulations, disruptive digital technologies, margin pressures and rising competition.

In the constant search for alpha, signals generated by alternative data such as transaction data, satellite imagery, weather pattern data, package delivery data and social media are becoming increasingly important. However, a growing number of investment managers are recognizing the importance of context in driving the analysis while using alternative data sets they use in their investment decision. They are interweaving alternative data points with other information they have about the company, giving them a more rounded, informed perspective on overall future performance.

It should be remembered that the data landscape is constantly changing, and asset managers must evolve to keep up with it. However, every model has a shelf life. Rare, unique and proprietary data eventually diffuses and becomes commonplace, easily available, edgeless data. Newly discovered datasets are rare and valuable because they contain alpha and firms should be on the constant look out to bring in these datasets. [10]

Big data solutions provide the ability to manage large data sets. Big data applications now extend beyond investment management to the areas of sales and marketing, compliance and risk. However, given the substantial investment requirements, asset managers are increasingly tapping into external consultants and niche service providers to execute big data pilot projects, rather than invest in full-scale infrastructure and additional staff.

An increasingly considered option being Cloud Implementation of Big Data and Analytics because of scalable infrastructure, breadth of technology stack and high computing abilities. In cloud environment the idea to spin up the cluster in minutes based on the volume and analysis complexities. These cloud services are entitled to:

- Integrate Cloud Infrastructure & Platform services for rapid development & deployment, at the same time enable inter-departmental data sharing. This will help to adopt Big Data related technologies without needing to build in-house expertise to create and manage Hadoop environment

- Define/deploy entire stacks without getting into integration complexities between various services.

- Take the guess work out of right-sizing your Big Data environment by automating the scaling of Big Data environment and extract the maximum value out of the infrastructure to deliver advance analytics, machine learning with superior performance

Oracle Big Data Cloud Service – Compute Edition is integrated with Oracle Public Cloud at both infrastructure and platform layers, providing a seamless experience for application development and deployment. In addition, Oracle Analytics Cloud extends the self-services BI to Smart insights, Advance Analytics and automation of analytical models.

It's high time for every fund house to ideate, experiment with modern technology that can help them achieve tangible business goals and differentiate themselves in an ultra-competitive marketplace. A great future to imagine, but one that will belong only to those who are acting now: acquiring datasets now, implementing latest technology solutions and hiring new talent. Embracing newer technologies such as AI, Machine Learning, Cognitive computing, Robotic process automation (RPA), Image Recognition, NLP etc., will help banks unlock the value of data to drive more informed decision making which is very important to business growth.

8 GLOSSARY OF REFERENCES

1. https://www.quinlanandassociates.com/wp-content/uploads/2017/09/Quinlan-Associates-Alternative-Alpha.pdf
2. https://www2.deloitte.com/us/en/pages/financial-services/articles/alternative-data-for-investment-management-decisions.html
3. https://near.co/blog/how-to-turn-alternative-data-into-alpha.php
4. https://www.linkedin.com/pulse/alternative-data-new-source-alpha-hedge-funds-akash-singhal
5. https://searchcio.techtarget.com/opinion/Ten-analytics-success-stories-in-a-nutshell
6. https://www.crisil.com/content/dam/crisil/our-analysis/reports/gr-a/whitepapers/big-data-in-asset-management-may2017.pdf
7. https://www.arcadiadata.com/blog/keys-to-a-successful-alternative-data-strategy
8. https://www.finantix.com/wp-content/uploads/2018/01/Applying-Artificial-Intelligence-in-Wealth-Management.pdf
9. https://www.globalcustodian.com/bnp-paribas-securities-services-adopts-ai-for-trade-processing
10. https://www.tcs.com/content/dam/tcs/pdf/Industries/Banking%20and%20Financial%20Services/analytics-artificial-intelligence-machine-learning-0817-1.pdf
11. http://resource.datamanagementreview.com/deriving-alpha-white-paper-download?utm_campaign=White%20Paper%3A%20Visible%20Alpha%20Nowcast_Nov%202018&utm_source=hs_email&utm_medium=email&utm_content=67886858&_hsenc=p2ANqtz-8Bpj7lixLja3Inbyh_hQvrbYoTcvGV_pFA2tN9_6RZfrgPqjoStvNypvHav8rG0-4LVQZKaVq9Wz9yWj0z_dyeRA3p6Q&_hsmi=67886858
12. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431
13. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5654190/
14. https://blog.statpro.com/blog/five-reasons-why-poor-data-control-is-holding-back-asset-managers
15. https://www.consultancy.uk/news/17145/five-challenges-for-asset-managers-to-remain-competitive

ORACLE CORPORATION

Worldwide Headquarters
500 Oracle Parkway, Redwood Shores, CA 94065 USA

Worldwide Inquiries
TELE   + 1.650.506.7000   + 1.800.ORACLE1
FAX   + 1.650.506.7200
oracle.com

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com. Outside North America, find your local office at oracle.com/contact.

blogs.oracle.com/oracle   facebook.com/oracle   twitter.com/oracle

Integrated Cloud Applications & Platform Services

White Paper Title
January 2017
Author: [OPTIONAL]
Contributing Authors: [OPTIONAL]

Oracle is committed to developing practices and products that help protect the environment