SILVER CREEK
CAPITAL MANAGEMENT

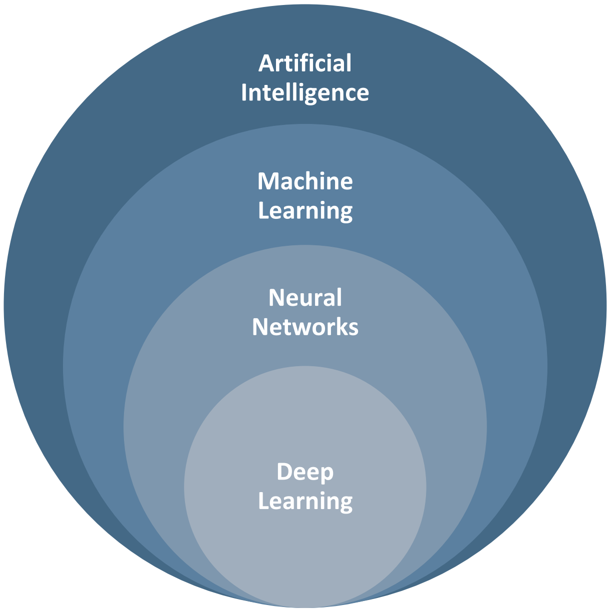
OVERVIEW
QUANTITATIVE TECHNOLOGIES
OCTOBER 2017

Introduction
Quantitative investing relies heavily on technology. As technology advances, new opportunities arise to use technology in the investment process. This article focuses on three ongoing technological changes that are reshaping what is possible in the investment realm: artificial intelligence, big data, and low latency.

ARTIFICIAL INTELLIGENCE
In 1950 Alan Turing proposed the Turing test, a test to determine whether a machine exhibited intelligence. Since that time, computer scientists have been attempting to develop machines with intelligence, commonly referred to as artificial intelligence. The definition of artificial intelligence is broad. It captures the ability for computers to behave like humans. Some of the dimensions this includes are perception, natural language processing, action, and learning.

While artificial intelligence is a large field of study, the aspect that is most interesting for the asset management community is the learning component. The sub-field of machine learning has been slow to evolve but in the last few years has begun to advance rapidly, especially after key advancements in a specialty of neural network algorithms known as deep learning. Machine learning is of particular interest for investors as it can learn from the past in ways that overcome human limitations.

Figure 1. The relationship between Artificial Intelligence, Machine Learning, Neural Networks, and Deep Learning.



Below we discuss machine learning in detail and highlight the recent implementation technique of deep learning.

I. MACHINE LEARNING
WHAT IS MACHINE LEARNING?
The concept of machine learning is simple: have a computer learn a task without explicitly programming it to conduct such a task. Traditionally, a computer worked by being told a specific set of instructions. With a programmer giving the computer instructions, the computer knew how to react to a specific data input. What distinguishes machine learning from other more traditional data analysis frameworks is its ability to work out new patterns and relationships from the data itself. By feeding a computer data, machine learning uses

predictive algorithms to independently discover new patterns and connections that allow the computer to improve its ability to make a decision or prediction.

Standard if-then programming is a quick and efficient method for analyzing straightforward problems. For complex problems, however, the number of lines of code necessary to tackle all possibilities and the degree of ex-ante understanding of relationships to understand them is enormous. For such problems it is best to have a computer program itself through its experience.

Machine learning falls within the realm of artificial intelligence and data analytics. It is a hybrid discipline that applies tools from various different fields. For instance, it uses prediction-making algorithms from statistics, takes advantage of methods used in mathematical optimization, and has overlap with data mining.

Machine learning has already crept into many aspects of our lives. For instance, it now influences Netflix's suggestions, Pandora radio's playlist, Google's translation, and whether an email is classified as spam or not. These, and many other systems, learn over time to improve their predictive accuracy and, as a result, their value to humans. While humans are excellent pattern detectors for data that comes in small sizes, we lack the ability to process large and complex data. Big data has allowed us to capture large and complex data while machine learning lets us learn from it.

HOW DOES IT WORK

Machine learning depends on data and processing power. The exponential growth in data and in computational power has led to improvements in machine learning. But how does it work? The scope of knowledge around machine learning is growing rapidly. But the core framework is digestible in a few short paragraphs.

Machine learning detects relationships between data points and uses these historical patterns to make predictions on new data.

There are four key methods by which computers learn:

SUPERVISED LEARNING	The learning methodology when the data are labelled and there are examples of the desired outcome. This type of machine learning is common in the application to investments. An algorithm will be developed based on the existing (often called training) data and will then be applied to new data where the outcome variable is to be predicted. A good example of supervised learning is using past weather patterns to predict the weather in the future.
UNSUPERVISED LEARNING	The objective of unsupervised learning is very different from that of supervised learning. It is used to categorize data. That is, the computer is given data that is unstructured and its purpose is to find a way to group and classify the data. Unsupervised learning is often used for visual recognition. Give a computer pictures along with text and with unsupervised learning it can try to match up which pictures should be associated with which text, with no training data required.
SEMI-SUPERVISED LEARNING	Semi-supervised learning is employed when the data structure has a component that is well defined and another component that is not. The first step is to categorize the unlabeled data. Such a process uses the unsupervised learning technique. The second step is to then apply supervised learning on the data to make predictions on the outcome variable. In practice, these two steps are intermingled so as to optimize the final outcome. A good example of semi-supervised learning is using facial recognition to identify an individual. Facial recognition

	uses past pictures of people to identify them in a new picture. To do so, different components of the facial structure are measured e.g. distance between nose and mouth. However, pictures do not classify what is a nose and what is a mouth. So, first the algorithm must classify these body parts. Second, it must try to predict to whom the facial characteristics belong.
REINFORCEMENT LEARNING	Reinforcement learning aims to maximize rewards through trial and error. The actor has to take actions in a particular environment. A good example of this type of learning is driving a vehicle. A driver (actor) steers and adjusts the car's speed (actions) while on a road (environment) in order to reach a destination without an accident (reward). The objective is for the actor to take actions in the environment that achieve the reward. Actions that increase the likelihood of receiving the reward are encouraged (reinforced).

Each type of learning has a number of algorithms used to carry out the process. Which specific algorithm works best depends on the data being analyzed and the particular purpose of the analysis.

WHY MACHINE LEARNING IS A BIG DEAL FOR INVESTING

Machine learning can be very valuable for investment managers. It excels at taking complex data and finding hidden patterns and relationships. The investment application is straightforward. Give a machine learning algorithm as much financial data as possible and let it sort out any links between observable behavior and possible future returns.

It is impossible for a human to keep up with even fundamental information such as regulatory filings, social media behavior, newspaper articles, corporate press releases, conference calls, and all the broader industry and market related activities. Yet, all of these activities have implications for firm value. For investors conducting statistical arbitrage where firm and competitor returns, limit order book information, options and futures data all can lead to trading signals, the limits of human pattern recognition are even more obvious. For a machine, all of these data points are accessible learning opportunities about how firms are likely to perform in the future.

While classical data analysis with a human prior can use these data sources to test for market predictions, its value is limited. Classical data analysis focuses on linear relationships and requires a human to determine which relationships to consider. Machine learning overcomes these limitations. The technique can consider non-linearities and need not be endowed with *a priori* knowledge about what relationships matter.

Machine learning still has limitations. While it is built to learn, it does not have intelligence in the human sense. It will find relationships but will not consider the sensibility of such relationships. For instance, perhaps in the past the U.S. stock market is found to increase when it snows on Mt. Rainier. While a fact about the past, there is no economic intuition for why the two should be linked. It is simply a spurious correlation. Hence, machine learning linkages should still meet human sensibilities to prevent spurious relationships from dragging down investment performance. Machine learning's ability to process data and learn paired with human sanity checks make possible large improvements in the investment process.

II. DEEP LEARNING

WHAT IS DEEP LEARNING?

Deep learning sounds intimidating. While the underlying math is technical, the general idea is not earth shattering. Deep learning tries to learn like a human does, through trial and error. Patterns that lead to the right outcome are supported while patterns that are unsuccessful are diminished.

Deep learning is a type of artificial neural network, a concept that tries to simulate the layer-upon-layer of synapses in the human brain. And neural networks are one strategy for machine learning, an important component towards the quest of artificial intelligence. Deep learning is being deployed so computers can learn to do tasks on their own that humans find easy to do but are surprisingly hard to directly program a computer to do. These include visual interpretation, speech recognition, content recommendations, and language translation.

HOW DOES IT WORK

As deep learning is an extension of neural network analysis, let us first discuss neural networks.

Neural networks are inspired by biology, specifically the interconnection between neurons in the brain. In biology, any neuron can connect with any other neuron within range. In artificial neural networks, the data propagation follow discrete layers, directions, and connections. With a lot of data and a lot of computational power, the neural network can begin to learn how to interpret its input. The human brain has no programmer in the background instilling it with direct commands. Humans learn by doing. Human neural connections that produce good outcomes are strengthened while those that produce bad outcomes are allowed to die out. In the same way, artificial neural networks are fed data without particular rules. Links between the neurons that work well at predicting the outcome are strengthened and links between the neurons that produce failures are weakened.

Deep learning is also commonly referred to as "deep neural networks", which is a more accurate description. It takes the idea of neural networks and grows the number of layers the data jump through between the initial input and its final output. The layers take a big question, like whether a picture is a stop sign or not, and breaks the evaluation down to many small questions.

Continuing with the stop sign example, early layers of the neural network ask simple questions like "is there red in the picture"? Later layers of the network take the answer to the previous layers, in addition to new information, and ask more complex questions, such as "is the shape an octagon"? Eventually an answer of Yes or No will be determined regarding whether the picture contains a stop sign. Importantly, the algorithm learns which characteristics help determine whether the picture contains a stop sign. In mathematical terms, the deep learning process is good at taking high dimensionality, reducing it to low dimensionality, and then rescaling it back to high dimensionality. In the stop sign example, the algorithm takes a picture that is complex to understand, breaks it down to its core components, and finally rebuilds it with the goal of determining whether it contains a stop sign.

Numerous types of deep learning network structures exist. However one underlying concept has been important in the progress of deep learning: backpropagation. Typically one thinks of a network as flowing from a starting point outward. However, the information flow need not be one way. Backpropagation is having earlier layers learn from the latter ones based on the success of the network. As a result modern neural networks can iterate over one example to learn from its mistakes. Such backpropagation allows for the deep learning algorithm to more quickly learn from its errors and therefore consider more layers with less data.

WHY DEEP LEARNING IS A BIG DEAL FOR INVESTING

The appeal of deep learning is that it offers the freedom to detect relationships without direct guidance from a programmer. By simply feeding data into the software, potential predictable signals can be ascertained.

For instance, historically, risk factors needed to be selected ex-ante based on economic intuition. While it is important to understand the economic rationale behind economic models, such a manual selection process can miss out on valuable clues hidden in the data. With a deep learning algorithm, a machine need not be told what the possible risk factors are and instead the computer can infer what the factors are and determine their pricing in the market. The tool is not limited to risk factor analysis; it is equally relevant in helping categorize features that are important, such as the characteristics of value firms or the nuanced parameters that can enhance a momentum strategy.

BIG DATA

WHAT IS BIG DATA?

Big data is currently a trendy buzzword. There is good reason why everyone is talking about it. It is one of the key ingredients necessary for computers to provide new insights.

The term is not well defined in the popular vernacular. Big data to some simply means analytics. To others it means the use of data that cannot fit into standard data management systems like Excel. Neither of these definitions quite captures what delineates big data. A more complete definition is that big data refers to the use of data sets beyond the ability of standard software tools to collect, manage, process, and analyze in reasonable amounts of time.

Three attributes characterize big data:

VOLUME	The amount of data being stored. Big is relative. For instance, in 1992 Teradata systems were the first to store one terabyte of data. At the time this was clearly big data. Now a terabyte of data is large, but it's not obviously big data. Now, for example the New York Stock Exchange, one of over a dozen exchanges, captures a terabyte of information in a single day. Now some companies have data measured in petabytes (one petabyte equals 1000 terabytes, or one quadrillion bytes).
VARIETY	The type of data being stored. Standard data is typically structured in a format of rows and columns similar to Excel. Big data can have this organization. But it can also be unstructured. Unstructured data need not fit in to a pre-determined model, such as information on websites, satellite data, or mobile phone location data.
VELOCITY	The speed data can be generated and processed. Data from cars, phones, websites, exchanges, and other sources are continuous and voluminous. Each of us produce torrents of data throughout the day. Special techniques and technologies are needed to store and analyze such data in time horizons useful to decision makers.

The large pools of data provide more power for descriptive, predictive, and prescriptive analytics.

HOW IT WORKS

There are books, courses, and even college majors dedicated to understanding how big data works. Here we discuss a high-level overview of the inner workings of big data to familiarize the reader with the tools in this space.

As mentioned, large volumes of data is an integral part of big data. Kyder's law is the analogy of Moore's law in the storage space setting. Whereas Moore's law says the number of transistors on one inch of an integrated circuit will double every two years, Kyder's law says storage space will double per one inch of magnetic storage every thirteen months. New tools have been developed to deal with the ever-expanding and unstructured nature of data.

In the past, data had to be structured and contained in a central location. Then, when access was desired a command would linearly work with the data storage system. While the growth of data has been exponential these tools only grew linearly.

The key developments have been:

COLLECT	Tools that allow the vacuuming up of structured and unstructured data in real time or in batch processes.
STORE	Allow for storage and queries across multiple servers so database sizes are no longer restricted by server size.
PROCESS AND ANALYZE	Distributed parallel processing in which programs can be split into separate components and ran simultaneous, such as Apache Hadoop or Google's MapReduce, whereby the system Maps (splits) queries across parallel nodes and then Reduces (gathers and synthesizes) the results.

These developments dramatically improve processing speed and can make large datasets easy and fast to access and analyze.

WHY BIG DATA IS A BIG DEAL IN INVESTING

Big data is a big deal for investing. While big data allow for data-driven investment models, the majority of those data-driven approaches are on top of classical fundamentally-based and economically-driven investment models. The rise of big data can be used in three critical facets of the investment process.

First, big data expands the possible inputs to analyze. No longer do investors need to solely rely on data from Reuters, Yahoo Finance, or Securities and Exchange Commission filings. An investor can use unstructured and non-traditional resources. These include, for instance, examining satellite imagery of retail parking lots to determine customer flows, searching retail websites' product inventory and tracking in and outflows to determine the volume of sales, and performing textual analysis of quarterly earnings calls to glean additional insights into the views of the managers.

Second, big data can help uncover previously undetected connections. With extreme amounts of data, small patterns can be distinguished from noise. And recognizing patterns where none seem to exist can lift up the investment performance of a fund. These signals can come from numerous places. For instance, big data may allow for more nuanced pairs trading or it may detect lead-lag return patterns between companies that are interconnected in an unusual way.

Third, big data can enhance backtesting and stress testing scenarios. With more data on more dimensions of an investment, greater relationship matrices can be developed and more nuanced and varying episodic shocks can be tested to more fully understand how different positions are expected to perform.

While there is clear upside to big data there are concerns to keep in mind.

First, the data in big data are still historical, and there is no guarantee what is observed in the past will persist in the future. Second, with lots of data and variables, there will be lots of spurious correlations. Assuming a 5% statistical significance benchmark, if an investor looks at 100 investment strategies on average 5 of them will look promising but are in fact worthless. Third, it provides enough degrees of freedom to allow for subjective facts to be generated that support a pre-determined view. Fourth, it can encourage data-mining with no economic thesis.

Even with these caveats, the rise of big data when used correctly can be a valuable addition to the investment process. Many firms will integrate human decision making with data driven facts and analysis. The end result can be an improvement in investment performance.

LOW LATENCY

In parallel with the growth of the ability for computers to learn, a race to reduce latency has been under way. Latency is the time it takes between when a signal is produced to when an agent is able to react to it. For the majority of algorithmic based trading strategies, latency horizon is unimportant. If the arbitrage is available now, it will be available in the next second or minute as well. For a small subset of algorithmic trading strategies, specifically high-frequency trading, the need for speed is first-order.

To understand the advances in latency it is useful to review how information flows through the trading system. Among the most frequent events in financial markets is exchange message activity – limit orders, limit cancels, and market orders.

LATENCY IN MODERN MARKETS

What follows are the steps between new information and an investor's eventual trade. The new limit order the exchange received will be sent out to subscribers of the exchange's data feeds. The data feed will be received by the investor. The investor must take in the new piece of information, analyze it, and make a determination as to whether it wants to react. After processing the information, suppose the investor decides he wants to put in a market order to buy. He will then send a message to the exchange. The exchange will receive it and process it.

In each step there are have been significant reductions in latency in the last several years. The developments in latency have been along three dimensions.

EXCHANGE PROCESSING

The NASDAQ exchange was among the first electronic exchanges. Since then nearly all trading markets have gone electronic. This is the first step to reducing latency. Exchanges have gone much further since.

First, they have introduced colocation. Colocation is a service exchanges offer to let traders reduce their physical distance from the trading site. The term trading site is more accurate than trading floor as most trading floors have closed and instead trading takes place on matching engine servers. Colocation is space provided near the matching engine server. Firms can buy colocation space and place their own servers in the location. With a trading algorithm running on the co-located server the trader has eliminated latency due to information having to cross physical distance. In fact, the physical distance is such an important factor in latency that early on, co-located traders would prefer being on a rack closer to the matching engine than the co-located racks a few feet further away from the matching engine. The few-feet competitive edge has since been democratized as now the length of the cable connecting the server and matching engine is customized such that all co-located servers have the same amount of cable distance to travel regardless of their actual proximity to the matching engine.

Second, exchanges have invested heavily in upgrading their matching engines. The matching engine had been a human matching buyers and sellers. Now, it is a high performance server. As recently as the late 2000's, the time it took an exchange to process a new message was in terms of milliseconds. Now, however that time has been cut down dramatically so the time it takes between receiving a message and having it processed is in terms of microseconds. The reduction in latency has come about because of increased hardware processing power and more efficient software.

NETWORKS

The way in which information gets from its source to the trader has also evolved. On the exchange front, many exchanges now sell connectivity with different capacity abilities. This includes the sale of exchange data directly to the trader. Consolidated exchange data can be purchased through the Securities Information Processor with current lags of around 230 microseconds. With a purchase directly from the exchange the delay can be significantly reduced.

Beyond the exchange, network speed in general has grown ever faster. A main thoroughfare of information is between the futures exchanges in Chicago and the equities exchanges in New Jersey and New York. In the 1980s, a fiber optic cable was buried along existing rail line right-of-ways. This meant data had to be sent through a number of repeaters and that the route was indirect. It would take data about 14.5 milliseconds to make the round-trip. Since then, a number of different routes have cut down the time and now the use of microwaves to transfer information through the air has been adopted. As of 2013, the time it took to traverse the distance has fallen to 8.5 milliseconds. To put that in context, traveling at the speed of light between the two locations would take 7.86 milliseconds. While there is still room for improvement, there is not much.

TRADER PROCESSING

The third component that can be enhanced to reduce latency is the trader's systems. There are a variety of best practices to increase speed. Of course, the simpler the program the less processing is required to make a decision and respond. And the program should be running on the co-located server to minimize time going back and forth to the machine engine. Different programming languages also have different efficiencies, Java and C++ are more efficient than R or Python. Careful attention needs to be paid to infrastructure details including data structure, thread handling, CPU cache pollution, and parallelizing. Advances in GPU and CPU allow for more efficient processing with less heat generation. In early colocation, facilities server racks under the air vents were preferred. The direct air flow allowed the CPUs to cool more efficiently and therefore would run faster (now air vents are equally distributed over the racks). All of these tweaks have resulted in efficiency gains, reducing the time it takes for information to move within the server and the time it takes to carry out calculations.

SAFEGUARDS FOR SLOW TRADERS

Being fast has clear advantages. There are two key aspects in which it matters. First, if two investors are looking for a piece of information, the one who receives it first will be able to react first. Being able to react first also depends on the ability of the investor to process and respond first. If it is a very short term mispricing, only the first investor will be able to capture the arbitrage. If it is a longer term arbitrage, an investor will still gain from being first as it will means they receive the best price.

For firms engaging in high-frequency trading, implementing low latency technologies is essential to remain competitive. For the rest of investors, there is concern that high-frequency traders will detect their activities and capture their alpha. The solution to such a concern is provided by most brokers and is referred to as a smart router. To achieve its objective of minimizing trading costs, a smart router does three activities differently.

First, it synchronizes its order submissions. A trader in Ohio will have different latencies to NASDAQ, BATs, and NYSE. If an order submission arrives at each exchange at the same time, other traders will have no time to respond on the other exchanges to the new order. Therefore, instead of sending the three order submissions from Ohio at the same time, the smart router will send the order submissions at slightly varying times such that they arrive at each exchange at precisely the same time.

Second, it randomizes orders. There is lots of activity on financial markets. An investor that leaves traces of their future intent in the data is asking to be front run. It is the equivalent of a quarterback pointing to where he is going to throw the football. A smart router will therefore randomize its submission patterns, especially for large trades, so that they are not reoccurring for the same size or happening at consistent intervals.

Third, it takes into account maker/taker fees and rebates, among other factors. While the Securities and Exchange Commission limits price increments to one penny, exchanges vary in the amount they charge and rebate for taking and providing liquidity. For large investors these differential prices can be important at the margin. Lower cost venues should be given preference when feasible. In addition, other factors such as desired time of implementation and optimal order submission size should be included in the routing decision.

CONCLUSION

Technology and quantitative investing are intrinsically linked. Quantitative investing regularly adopts the latest and most sophisticated technologies. The topics discussed here - artificial intelligence, big data, and low latency – will only grow in importance. As technology continues to evolve, investment managers will find creative ways to embrace it.

IMPORTANT DISCLOSURES:

This document is for informational purposes only and does not constitute an offer or a solicitation of an offer to buy an interest in any fund (each, a "Fund") managed by Silver Creek Advisory Partners LLC and/or any affiliated management company thereof, including without limitation Silver Creek Capital Management LLC (collectively, "Silver Creek"). Offers are made only pursuant to the Confidential Offering Memorandum and the Subscription Documents of the Fund, which should be read in their entirety. Funds are intended only for sophisticated investors who are able to assume the risks inherent in investment vehicles of this type and who meet the Funds' eligibility requirements. **No assurance can be given that any of the Funds will achieve their investment objective or any particular level of returns. An investor may lose money by investing in any of the Funds. Past results of Funds are not necessarily indicative of future performance, and performance may be volatile.**

Silver Creek does not necessarily have access to information from third-party managers to ensure the accuracy of the information presented, and any information received from third-party managers may be inaccurate or incomplete. Certain information presented is of a high-level, summary, condensed and aggregated nature, and is inherently limited, incomplete, and required the application of simplifications, generalizations and assumptions to produce. Silver Creek expressly disclaims any representation or warranty as to the accuracy, completeness, availability or timeliness of the information presented. The information provided does not necessarily reflect the most up to date or current information available.

By accepting receipt of this document, you hereby agree and acknowledge that the information contained herein (the "Confidential Information") is strictly confidential and may not be reproduced or distributed in any manner. You agree to not disclose any Confidential Information to third parties, except as provided below. You may only disclose Confidential Information upon a good faith determination that such disclosure is required by judicial or other governmental order or as otherwise required by law, provided that you have given reasonable notice to and shall consult with counsel of Silver Creek prior to such disclosure and you shall comply with any applicable protective order or equivalent. You may disclose Confidential Information to your employees or legal and financial advisors on a need-to-know basis.