Artificial Intelligence/Machine Learning + Supply Chain Planning

Summary Report

Cambridge, Mass.

November 27–28, 2018

Moderated by:

Sergio Caballero PhD
Research Scientist
MIT Center for Transportation & Logistics

James B. Rice, Jr.
Deputy Director
MIT Center for Transportation & Logistics

Table of Contents

Executive Summary

MIT's Center for Transportation and Logistics (CTL) held a highly interactive one-and-a-half-day roundtable on the use of machine learning (ML) in supply chains. Representatives from 16 companies in a diversity of industries discussed their organizations' uses of ML in a variety of forecasting, optimization, and management applications in their supply chains. To ensure candor at the event, this report was prepared under the Chatham House Rule of not identifying the specific speakers or affiliations associated with their anecdotes, insights, or recommendations.

During the first half day of the roundtable, presenters from CTL introduced some of the fundamentals of ML. Short tutorials covered supervised learning, unsupervised learning, reinforcement learning, and neural networks. The presenters described several common data-driven algorithms for prediction, classification, and clustering. A later tour of CTL's Computational and Visual Education (CAVE) Lab showed how visualization can enable informed, data-driven decision making.

The second full day of the roundtable focused on specific supply chain applications of machine learning related to demand forecasting, revenue management, and transportation. Discussion of each application began with a kick-off case study presented by one of the industry participants. This was followed by in-depth discussions of the participants' experiences and issues with machine learning for that application. A beverage distributor described how it uses ML for demand forecasting to more accurately plan from the interplay of sales trends, holidays, weather, and promotions on sales volumes. An omnichannel apparel retailer presented its use of ML to optimize price markdowns on fashion items. An ocean freight data company and a 3PL presented half a dozen uses of ML in transportation to predict transportation asset activities, in-transit risks, spot market prices, and other applications. A presentation on autonomous vehicles illustrated the power of ML as well as the weaknesses of ML. This led to a recommendation of creating man-machine collaboration for vehicle operations.

In the final session, the participants discussed many cross-cutting issues relating to how organizations design, develop, and deploy machine learning systems. Key organizational issues included: how to find or create ML talent with the required knowledge of business, math, statistics, and computer science; where to place ML teams in the organization's structure; and how to solve change management issues in deploying data-driven automation. Key takeaways included:

- ML can improve forecasting of supply, demand, pricing, timing, etc. to proactively manage the future.
- ML can cluster and classify supply chain conditions, events, product, and customers, which can help manage complexity through differentiated responses and tailored best practices.
- ML requires data that needs to be gathered, aggregated, cleaned, and manipulated.
- ML requires more math, statistics, and computer science knowledge (and tools) than what most business data analysts and IT professionals have.
- Future supply chain leaders will need to understand enough about what is possible using ML both technologically and organizationally in order to improve business performance.

The Fundamentals of Machine Learning

After a welcome and introductions, Dr. Daniel Merchán began the roundtable by presenting the fundamentals of machine learning (ML). He first asked participants if they were using ML in their workplace, and about half of the roundtable participants raised their hands. Then he asked who was using ML at home, but only four participants raised their hands. Dr. Merchán, however, pointed out that everyone should have raised their hands to both questions. If they used Netflix, Spotify, Uber or Alexa—all of those have ML-enabled applications. Even if a participant only used email, email programs use ML algorithms to detect spam. Everyone in the room had been using ML, whether they knew it or not.

Next, Dr. Merchán defined artificial intelligence (AI) as "machines capable of performing cognitive functions associated with the human mind." He positioned machine learning as a subfield of AI, along with robotics, natural language processing, computer vision, and speech recognition. ML is the most dominant subfield of AI, using past data to build models capable of making predictions on future data.

Although AI dates back to the 1950s, ML's tremendous advances have been achieved only in the past few years due to the increased amounts of computing power and data that were not available before. Indeed, 90% of the world's data has been produced in the last two years. For example, four million videos are uploaded to YouTube every minute.

Machine Learning Methods

Machine learning uses data, probabilistic models, and algorithms. Because ML uses probabilistic models, the output should be assessed using statistical confidence levels. The machine learning process requires:

- problem identification
- cleaning the data
- implementing the model
- training and testing
- evaluation
- deployment
- updating

Machine learning methods can be classified into three major families. First, **supervised learning methods use labeled training data** to make predictions on future data such as predicting demand, classifying images, detecting fraud, or making medical diagnoses. Second, **unsupervised learning methods find previously unknown patterns in data** and can be used for customer segmentation and product recommendations. Third, **reinforcement learning methods use some notion of reward to guide training** and can be used for skill acquisition. Dr. Daniel Merchán, Dr. Sergio Caballero, and Connor Makowski gave the group tutorials on these different machine learning methods as well as some more advanced neural network methods.

Supervised Learning for Prediction and Classification

Supervised learning algorithms are used for classification and prediction in which the value of the outcome of interest is known in historical data or training data. As Dr. Caballero said, "You know how to label the existing input data and the type of behavior you want to predict, but you want the algorithm to calculate it for you on new data."

Dr. Caballero briefly described many types supervised learning techniques including linear regression, logistic regression, classification and regression trees (CART), and random forests, explaining the advantages and weaknesses of each. Regression techniques, for example, find the best-fit formula that explains or predicts an outcome, such as level of demand or productivity.

A classification tree, on the other hand, lets a company classify something, such as classifying bank customers as being acceptors or non-acceptors based on various variables such as income, education and credit card expenditure. Trees are good off-the-shelf classifiers and predictors, and they are useful for variable selection but they are sensitive to changes in the data. Slightly different sets of training data could affect the outcome a great deal. A big advantage of decision trees is that they make the logic of the decision easy to see and explain.

Company Example of ML Usage in Last-Mile Productivity

CTL researchers helped a beverage company understand and improve its last-mile productivity. To do this, the researchers developed a regression model that predicts productivity as a function of 18 route, service, and time factors such as: 1) the distance of the route, including the actual distance, the planned distance and deviations (absolute and relative) 2) duration of the route (actual, planned, and deviation) 3) stop sequence (actual, planned, deviation), and 4) vehicle capacity occupation. The analysis determined which variables were good predictors and enabled the researchers to predict whether a planned route would be of low, medium, high, or very high productivity.

Next, the researchers used a classification tree based on the variables to predict the productivity classification of other planned routes. However, they found that the classification tree was quite sensitive to the data, so they switched to a random forest model and also assessed the explanatory and predictive power of the variables. The most relevant factors identified were: route volume, vehicle capacity occupation by percentage, planned service time (hours), average drop size, number of customers, and planned route duration. The team used 110 trees to a depth of nine nodes, yielding a test accuracy of 63 percent.

Classifiers can also be trained to recognize physical objects. For example, the goal might be to classify whether an image shows a picture of a pallet or not. Supervised learning would be used on a database of images that had been labeled as showing or not showing a pallet. That data would then be used to train a model to have the highest possible pallet recognition accuracy. In the end, the trained machine can be presented with a new image, perhaps from a camera on an automated forklift, and the machine will predict whether the image shows a pallet or not.

Unsupervised Learning for Pattern Discovery

Unsupervised learning methods can identify new patterns and categories from data that were not known beforehand. Clustering methods such as k-means clustering, hierarchical clustering, k-mediods clustering, and Gaussian mixture models are common unsupervised learning methods. Each method uses different mathematical functions to aggregate similar data points together and split dissimilar ones into separate groups.

For example, additional work by CTL researchers on behalf of the previously mentioned beverage company looked at ways to tailor the company's last-mile strategies to different urban conditions. But rather than attempt to predefine these conditions for a foreign megacity, the researchers used unsupervised machine learning to discover them in the data. They used principle components analysis and k-means clustering to identify regions of a megacity with similar logistics profiles based on various dimensions of delivery patterns, population density, and road infrastructure properties. From this analysis, the researchers could further analyze critical areas of the city and propose various multi-tier distribution pilots.

Affinity analysis is another kind of unsupervised learning method. For example, Amazon and Netflix use this technique to derive product recommendation "rules" based on co-occurrences of events, such as people who buy one book also bought a certain other book. Amazon looks at conditional probabilities: what is the probability that if a customer buys A s/he will also buy B? If the probability is very high, Amazon can then recommend B to other buyers of A.

Reinforcement Learning for Self-Learning Machines

Reinforcement learning trains a machine through many iterations of decision making and provides reinforcement signals when the machine achieves a good outcome. Reinforcement learning can train a machine to successfully play a game or optimize a task without explicitly encoding the rules of play or strategies for winning. As with unsupervised learning, reinforcement learning can be used when humans don't even know the correct answer. In the case of reinforcement learning, the trainer only needs to be able to recognize a better answer from a worse one.

An example supply chain application of reinforcement learning could be the task of organizing inventory in a warehouse. The machine would try to minimize pick-and-pack labor for complex orders while avoiding congestion in any part of the warehouse. The machine might try various inventory placements and permutations of placements, which are then rewarded or penalized based on the amount of labor hours spent on fulfillment. Reinforcement learning often relies on computer simulations. Simulations are a very inexpensive and fast way to give the machine a lot of experience and a lot of time to try different strategies and tactics.

Machine Learning with Neural Networks

Neural networks are an important class of current-day machine learning algorithms that can be adapted to solve supervised, unsupervised, and reinforcement learning problems. Modeled very loosely on biological nerve tissue, a neural network for machine learning consists of one or more layers of nodes (neurons) connected to each other and to a set of inputs and outputs. The training process for a neural network adjusts the weights on the connections and other parameters to optimize the outputs that the network produces when given a set of inputs. As with other machine learning methods, neural networks come in many types, each suited for different applications such as speech recognition, natural language processing, and image recognition.

Deep learning algorithms use more complex multi-level architectures to model complex relationships. Deep learning requires less feature engineering to pre-structure the inputs and outputs for the learning model. However, deep learning requires both much more data and more computer resources to successfully train the model. Examples of deep learning include voice assistants such as Alexa and Siri, advanced game-playing AI such as AlphaGo, and autonomous vehicle control systems.

Supply chains could also deploy deep learning technologies such as voice recognition that could be used in pick-n-pack, with a voice telling the person what to pick rather than them having to read it. Another application could be voice interfaces for truck drivers, who could communicate with their trucks such as saying, "tell me where the terminal is."

Similarly, image recognition could be used for product identification (for picking or inventory management) or for detecting cargo damage, such as if a box has been partially crushed. Inside a truck, image recognition could tell whether the truck was empty or full or what it was carrying. Inventory cams could take pictures of the warehouse and estimate how many cases of product there were.

Advanced Visualization

Participants toured the MIT CTL Computational and Visual Education (CAVE) Lab to learn how advanced visualization can accelerate decision making. The lab has a large table-top and floor-to-ceiling touch screen monitors driven by a powerful array of computers. These facilities enable researchers and executives to model and visualize complex supply chain problems through interlinked and color-coded 3-D maps, graphs, diagrams, and animations.

For example, a large retailer and a large chemical company have each sponsored research using the CAVE to help understand and optimize their distribution networks. The visualization takes into account factors such as the numbers and locations of warehouses, proximity to customer populations, inbound and outbound transport costs, market share, sales volume, and profit. The visualization lets an executive, for example, open or close warehouses or change which products are distributed through which warehouses to see how it changes service metrics, market share, and total profit. Interactive systems such as CAVE enable informed, data-driven decision making and bridge the gap between inscrutable optimization algorithms and intuitive images of system performance.

Augmented reality (AR) and virtual reality (VR) are two additional advanced visualization technologies that can enhance supply chain design, operations, and consumer interfaces. AR-goggles could provide a heads-up display for the contents inside vehicles, pallets, and so forth. In picking in a warehouse, an item could be highlighted in green and the picker could say, "I picked the green one." Or, AR could display schematics for maintenance as needed when looking at a broken conveyor belt, for example, or it could highlight the bolt that needs to be replaced and which way to rotate it, with the end goal of speeding up repair. VR could be used to design new warehouses or train workers to drive forklifts. AR or VR could also help a customer view a product, such as what a couch would look like in their living room.

Machine Learning in the Supply Chain

Machine learning can be used for many categories of supply chain applications. ML can be used for prediction or forecasting of demand, supply, on-time deliveries, and risks. ML can help automate many routine elements of supply chain operations and help detect or predict exceptions to routine operations. ML can be used for planning and design such as of networks, inventories, schedules, and routes. Finally, ML is a key component of autonomous supply chain vehicles such as trucks, ocean freighters, delivery drones, and forklifts.

The second day of the roundtable focused on specific supply chain applications of machine learning. Discussions of each application began with a kick-off case study presented by one of the industry participants. Then participants shared their own experiences with that application of ML, and asked and answered each others' questions about the application.

Demand Forecasting

In the first session of day two of the roundtable, a beverage distribution company described how it uses ML to forecast demand. The company's analytics journey began three years ago to turbocharge a high-value post-merger integration project on consolidating cross-docks. Previously, analysis was primarily done in Excel spreadsheets. About a year and half ago, the company started to do more predictive analytics, hiring data scientists, and using open source ML software.

Context and Goals

Overall, the distributor handles 40,000 SKUs that include a mix of high-volume mass-market brands, low-volume must-carry products, seasonals, and volatile product lines. They equip their sales reps with a tool that helps them make product volume recommendations to each account based on which products would be relevant to that account. The company's goals in better forecasting were to help integrate acquisitions (growth), manage SKU complexity, and optimize business objectives (e.g. by reducing out of stocks, days on hand, or improving quality control and accuracy).

The company's beverage inventory levels are driven by several business factors in addition to demand. By contract, they must carry a certain number of days in stock, but the amount differs by supplier. Furthermore, some suppliers pre-build inventory prior to the summer, which risks the beverage going bad. Suppliers also continually introduce new products that they think will transform the business, but in reality most of those products die within three years while some others sell so well they outstrip production. In general, the company prefers overstocking to understocking. Out-of-stocks have high costs and although out-of-date product must be destroyed, the four-to-six month shelf life usually provides sufficient leeway to clear excess inventory.

Building a Model

When building its model, the company used historic sales data as a baseline, because overall volumes in its industry are fairly stable. The company then factored in a wide range of external and supplier events that could boost or mute sales, such as weather, sports events, holidays, suppliers' monthly business objectives, and promotions. In terms of models, the company began with ARIMA (AutoRegressive Integrated Moving Average) first and then used RNN (Recurrent Neural Nets) and ensemble techniques.

Specifically, weather data included not only temperatures but also rain data in each location. The company discovered that temperature categories (high, medium, low) weren't accurate enough, so it used numerical temperature values pulled from NOAA (National Oceanic and Atmospheric Administration). When forecasting future demand, the upcoming weather is only a forecast itself, but it is quite accurate for 10 days, less accurate for the next 20, and the company uses historical weather for the longer-horizon forecasts. The point is that even if the weather is only a forecast, it's sufficiently accurate for estimating near-term future demand for purposes of near-term operational planning.

External event data included sporting events such as the World Series, Super Bowl, Stanley Cup and so forth. Holidays included some that impacted particular beverage brands, such as St. Patrick's Day and Cinco de Mayo as well as holidays that had broader impact, such as July 4th and Thanksgiving. Supplier-related event data included monthly business objectives, which were set by suppliers. Finally, promotions were incorporated. Some promotions, like BOGOs (buy-one-get-one-free) and display promotions, were encoded as yes/no classifications in the data. Other promotions were encoded as continuous values, such as the percentage discount or the push promotion volume. Supplier event data was often presented in unstructured emails or semi-structured Excel files, so the company created parsers and a way of reading the files such that it would also know whether there was a sales incentive for the salespeople because sales incentives influenced brand choice a great deal. Overall, the model was a hybrid of both math and input by people. Human input was used if the forecast model missed something. If it did, S&OP (sales & operations planning) was used.

Getting Results

The company found that machine learning significantly improved forecast accuracy as measured by mean absolute percentage error (MAPE). Overall forecast error for using baseline historical sales was a lackluster 60%; the legacy model had 37% error; the first ARIMA machine learning model delivered 23% forecasting error; and the final ensemble model provided 10.5% MAPE. Overall, the company reduced the forecast error by 100,000 cases per week.

The company found that sales incentives, such as the potential for winning a trip to Cabo, were the #1 predictor of sales for that timeframe. Other sales promotions created huge spikes. A given product may sell 100 cases a week but a BOGO promotion could surge demand to 2000 cases a week. The company also had to factor in anomalous demand spikes from special events from previous years, such as when the local team won the World Series one year but not the next.

The impact of the weather depended on location. For example, 60 degrees Fahrenheit in a northern city is pleasant and has little effect on cold beverage consumption. But in Florida, 60 degrees Fahrenheit is considered cold and does has an effect. Furthermore, whether it rained only for one hour or rained all day made a difference. Unusual rain patterns such as the prolonged heavy rains in Los Angeles in the previous winter also had a big impact. Finally, extreme weather like hurricanes had an out-sized effect because unlike food, which is always needed every day, some beverages are more often consumed for special occasions such as picnics, parties, and sporting events. Sales lost to bad weather were not made up later.

The company also uses the ML tool to recommend product sales patterns to suppliers. A supplier might have a new product that they want places in 50% of locations, but which locations? The distributor negotiates the number of points of distribution based on its model and helps predict if there are correlations or cannibalizations between brands.

Revenue Management

A department store retailer presented the company's use of machine learning for revenue management that optimizes markdowns of items. The retailer sells through both online and offline channels, including having omnichannel abilities to serve one store's customer from another store's inventory. The retailer also operates a network of off-price outlet stores to handle surplus merchandise.

Revenue Forecasting

One of the retailer's early ML projects was to improve the accuracy of revenue forecasts. The company used an ensemble model with generalized additive models with auto-regressors. It proved that these models could improve forecast accuracy. The new models were also much more scalable and could generate fine-grained forecasts at the product or group level. "Once we have the forecast clear, then we can make better decisions on how to allocate product," the retailer said.

Revenue Management: Price Markdown Optimization

The retailer sells many fast-fashion products. It offers products both online and offline and then monitors their sales performance. If an item under-performs, the retailer can mark down the price. The question is, when should the retailer mark down the price and how much should it mark down the price? Overall, the company seeks to maximize revenue, which creates competing optimization goals of selling the items at the highest possible unit price while using markdowns to drive unit volume.

The company's legacy markdown process used a set of business rules maintained by each department. The rules included the list of eligible products, inventory levels of the products, and age of the products. Based on that, members of the merchandise team determined the markdown. Some products are ineligible for markdowns. Others could be marked down a total of three times. After that, any remaining inventory was sent to the retailer's own off-price outlet store. This old process was highly manual and was not data-driven.

The retailer wanted to have a data-driven optimization model to see how the price markdowns affected sales and revenue. The first step in optimizing the markdowns was to estimate the price elasticity—the percentage change in sales sparked by a percentage change in price. However, the products being marked down were typically trendy, seasonal items that never accumulate the kind of long sales history needed to accurately estimate elasticity. Thus, the retailer used unsupervised ML (K-means) to cluster similar types of products on product attributes (e.g., sleeve length) and then learned from the sales of clustered products to come up with a price function. Even with that, using product attributes to identify products of similar price elasticity was a challenge, so the retailer next tried hierarchical classification. It created a forecast tree of those attributes with "a women's top" as the uppermost class and then traveled down each node (blouse, sweater, t-shirt, etc.) to hit optimal price accuracy.

Other promotions, such as the retailer's spring sale and Black Friday sale, complicated the modeling and optimization. These short-term sales involved both a price drop for the sale and then a price increase back to the pre-sale price. This contrasted with the markdown logic in which the price never reverted to its old value. The Black Friday price drop was determined using the last several weeks of the price, so that the customer saw a significant drop. Product sales always spiked during these short-term promotions but then dropped off when the sale was over.

Revenue Management: Price Markdown Behavior

The data revealed a characteristic product sales life cycle with diminishing sales at the end, so the retailer incorporated that pattern and price range into its model. Not only did the retailer have to classify each product, but price had various factors as well, including the lifecycle pattern, age of the product, and seasonality of the trend. The retailer used a general additive model and polynomial regression to forecast demand. Analysis found that the number of markdowns played a more important role than the percentage of the markdown. The retailer could just mark the price down once or twice and then send it to the outlet store to free up space. The process has not been completely automated yet, because the massive Excel spreadsheet for business rules by department still existed.

The company found that markdowns had different effects on in-store revenues versus online revenues. The reason is that customers visiting a brick and mortar store may simply miss the marked-down item or fail to notice it based on where the product is located and the customer's route. In contrast, the online site has a prominent "Sale" tab including "New Markdowns" sub-items in each department that aggregate the discounted products in one place for bargain hunters.

The retailer did uncover a counterintuitive result. Many business people think that a 50% price reduction will create more sales volume than a 30% reduction, but the inventory levels affect sales, too. Based on their own experience, the retailer concluded that if there is little inventory left, the larger price reduction does not sell more units, it only sub-optimizes overall revenue. Similarly, they concluded that a price markdown may not boost demand as much as expected if the customer's size is missing. Thus, for nonreplenishable products, the company had to consider the total purchase order quantity or remaining inventory as the upper limit of possible sales.

Transportation Planning

Two participants, one in ocean freight data and the other a 3PL heavily involved in trucking shared their experiences with a number of ML examples related to transportation planning. The first company focused on machine learning to extract or forecast valuable data about ocean freight transportation systems. The second company focused on prediction of potential problems in the execution of a customer's supply chain so that the company can prevent problems before they occur.

Predicting Ships' Destinations

Tracking and monitoring the movement of important freight can use radio transponders, but these typically use cellular radio networks that are only available on land. Tracking freight on the ocean often uses vessel tracking as a proxy for tracking the cargo carried on the vessel. Using maritime data and vessel shipping data, one company was able to see when a vessel arrived at a port, and when it departed, so that the company could match the container that was on board and continue to track it. However, the movement on the ocean between ports was still hidden, meaning that the company couldn't tell if there were delays on the water. So, the company continued to look for missing information to analyze and understand the performance on the water, such as by analyzing how congestion and different contextual pieces impacted vessel speed. The challenge on the analytics side was to identify when a vessel started its journey. Was it when it arrived at a port or left the port? Moreover, the customer only cared about their own products on the vessels, not others. If the vessel was being operated by a 3PL, the company had little to no visibility into arrivals and departures.

The project required multiple steps to identify different types of data, and then learn from that data. For example, the company used AIS (Automatic Identification System) data that all ships are supposed to transmit to surrounding vessels as a safety measure. With AIS, each ship broadcasts its location, speed, heading, and other data every minute —500 million daily messages worldwide. The company used over 10 years of data as its training data for ML. To use data efficiently for the purpose of making predictions, however, it is better to use less data. Thus, the company down-samples it to use only data that represents a state change, such as a speedup, slowdown, or change in heading. That leaves 2 million data points per day. Looking only at commercial traffic reduces it down to 500,000 data points per day.

Other information comes from manual entries keyed in by the captain, such as the vessel's destination. Unfortunately, different captains use different codes for a port, such as SNG or SNP for Singapore. Or, the captain might key in "Panama Canal" but the ultimate destination is the port in Jacksonville, Mississippi. The company used ML to predict which port the captain meant by correlating it with the vessel's current location, speed, and direction. Another piece of information that helped predict the right port was the nature of the ship's cargo and the subset of worldwide ports that handled the type of cargo—some ports only handle particular bulk products. Based on the port of origin, the company could predict the type of cargo on board and then limit the possible destination ports to those handling that type of cargo. A key facet of ML using big data with today's fast computers is that it can leverage a great many noisy or weak sources of data to develop very accurate predictions.

Predicting Asset Availability

Both presenters during the transportation planning session spoke of efforts to predict asset availability. Shippers and brokers seek transportation assets that are available and willing to carry their cargo from a given origin to a given destination. In times of tight capacity, shippers struggle to lock in available assets before someone else does. And at all times, they seek the lowest-cost asset. The two participants had used ML to predict carriers' asset availability—who might have a ship or truck available at some future date, location, and destination? Both participants developed ML systems to forecast where assets were going and whether they might be available for hire at those forecast locations.

The ocean freight data company used ML techniques such as recurrent neural networks (RNN) for predicting vessels' port sequences. By knowing what a vessel has done previously, that information could be re-coded to convert the raw data into abstract data that would go into a predictor trained by a supervised learning method of what the company believed would be the next event/location. Then, the company could take that abstract sequence and convert it back to events/locations that would be applicable to the vessel. The company also used public information on who has contracted with which vessel, to predict whether a vessel was on a long-term charter or whether it might become available. Other publicly-available data, such as weather data, impacted vessel speed and thus could be used to enrich AIS data.

Other hidden factors also impacted availability. Bulk vessels can carry many types of cargo, but a vessel that carried dirty cargo such as crude oil could not be used for more refined cargos such as jet fuel without being cleaned. A vessel would remain dirty and thus not available for some kinds of cargo unless it visited a maintenance facility. The ocean freight company achieved a 95% accuracy on ships' clean/dirty status, which in turn improved the prediction of affected asset availability.

Similarly, a 3PL is modeling truck movements all over the country for its brokerage operations. For example, it may see that a certain truck moved from LA to Portland and then see it again some days later going from Chicago to LA. Thus, the broker knows that the truck must have moved from Portland to Chicago, but it was not carrying any freight from the broker. The hidden Portland to Chicago move is called "dark freight" because the truck is being used by a competitor. If the data shows that this truck always moves in an LA → Portland → Chicago → LA triangle, the company can contact the carrier once the truck leaves LA and ask whether they would like to take a load from Portland to Chicago. In this way, the broker can find additional assets to carry its freight, and the truck may have better capacity utilization.

The 3PL uses deep learning models to identify potential assets that a carrier has not yet announced as available because the carrier hasn't started to look for a new load yet. The 3PL used a neural net to model the carrier profile and forecast truck locations up to 10 days in the future. The 3PL can then proactively contact the carrier about future availability four to five days in advance, knowing both the likely upcoming destination of the truck (e.g., Portland) and the truck's preferred subsequent destination (e.g., Chicago). Also, calling ahead means the shipment is less likely to be canceled or rolled.

The situation is win/win: "It's likely that we have a shipment going there so we can offer them opportunity to get that capacity and add value to that carrier to find the best, most profitable opportunity triangles to reduce deadhead miles, and it keeps our customers happy," the 3PL said. Previously, the company had done this through manual methods, but ML brings the opportunity to handle greater scale given that there are tens of thousands of shipments daily and different rule sets are needed for each carrier. For example, some carriers do triangles, others do inbound, others go up and down the coast.

Predicting Spoilage and Problems in Transit

Several examples showed how ML could be used to predict problems in transportation. For example, a beverage company was writing off $20 million a year due to freezing that occurred during ocean voyages. The company wanted some way to forecast which shipments should be checked or reordered if there might be a problem. Analyzing a combination of historical data on vessel trajectories, weather, and quality control outcomes helped them determine that shipments were only in danger if they spent more than six hours in freezing conditions. With that rule, the company could monitor future shipments—vessel tracks and weather data—to predict which shipments had been spoiled and take corrective action.

Similarly, a large food manufacturer faced problems with freezing mayonnaise. To prevent this in the past, the manufacturer pre-emptively shipped the goods during winter months in temperature-controlled trucks. When a risk analytics company studied the situation, it found that the manufacturer was over-using the more costly trucks by a factor of two. A more accurate model helped them predict which loads would really need temperature control. The results were both a lower shipping bill and no loads lost to freezing.

Another client of the risk analytics firm is a car maker with only three hours' worth of inventory for certain parts. Any disruption can cost the manufacturer hundreds of thousands of dollars per hour of disruption. The analytics firm developed a model to predict whether a port would have to close due to weather. Forecasts of disruption enable the car maker to make alternate arrangements as necessary to keep parts coming.

Predicting Arrival Times

The ocean freight data company could also predict ETAs (estimated time of arrival) by using the current speed, heading, and predicted destination to draw a great circle path and calculate point A to Point B distances and travel times. The company used a year's worth of data of the whole world to generate a spoke map – the map would show where vessels had moved in order to find the most efficient route on the water. The lengths of the legs could be used to calculate the distance. Likewise, the company could forecast the line-up of inbound vessels at the port to predict if a given vessel might be delayed due to congestion.

This company used these kinds of predictive models of travel time to help a large client improve coordination among assets that needed to arrive together. The client company moved freight from China to the US, which took about 22 days but each day had a cost of $24 million on the business. Thus, improving velocity by one day could have an immediate savings of $24 million. Even more crucial, however, was that some of the ocean vessels needed to arrive at the same time. If 9 out of 10 ships arrived at port but the tenth one was late, there was a daily cost of holding the inventory on the other nine ships. Thus, coordination based on accurate predictions had huge value as well.

Transportation Document Recognition

Another transportation challenge is classifying all the documents associated with transportation, such as which ones are bills-of-lading (BOL) and whether or not the BOLs have any mark-ups that might indicate damage to the freight. A transportation broker built a basic image recognition algorithm to do this so that when a customer requests the BOL, the company can find and deliver it quickly without manual labor. The automated system is still in the testing phase but is expected to help accelerate payments. With $2 billion per day tied up in the company's receivables, that acceleration would have a very big impact on the company.

Predicting Spot Quote Prices

Transportation brokers face the challenges of finding carriers to match shippers. A shipper might submit an RFQ (request for quote) with a long list of loads, but then the broker must ask various carriers to respond and then wait for the bids. Carriers don't always respond or respond quickly because they know they might only win two percent of bids, so they often perceive this as a waste of time. However, the shipper would prefer a timely response from the broker. A broker used ML to estimate real-time supply and demand and predict near-term spot market prices with enough statistical confidence that it could provide binding quotes to shippers even before it found a matching carrier.

The faster that the broker can respond, the more valuable it is for the customer, and it is much more scalable than trying to do it with people. "If 200 carriers are searching in our app, in that lane, posting trucks, posting offers, we can react in a short time period and in a few minutes we can change the rate so that if one of our sales reps quotes a lane it impacts other sales reps and they see it immediately. It can be coordinated much faster than people could because it's all in one single data stream." The broker can provide 5,000 quotes in 30 minutes, for example. The only downside is that occasionally, the broker cannot find a carrier for the price the broker has already quoted and must absorb the difference as a loss.

Looking Forward: Autonomous Vehicles

Dr. Lex Fridman described how successful autonomous vehicles (AV) first appeared in the 2005 DARPA Grand Challenge. That off-road race in the desert required a fully autonomous vehicle to successfully navigate many miles with absolutely no human intervention. No car had finished the previous 2004 race, but in 2005, Stanford's team won and a total of four cars finished. At that point, universities thought all the fundamental research was done and nothing new was needed. But DARPA came up with the 2007 Urban Grand Challenge, to navigate in a complex urban environment, obey all traffic laws, and avoid hitting other cars on the road. Again, when several cars finished, academics felt all the research was done, "The rest is just stamp collecting." The belief was that industry would polish the few remaining rough edges and soon be selling AV.

What Seemed Solved Wasn't

Belief in this victory arose from three notions: that the driving task was easy, that humans were bad at driving, and that humans plus automation didn't mix well, i.e., that humans over-trusted machines or misinterpreted the state of the machine. Dr. Fridman challenged each of those three notions. First, driving wasn't an easy task. Actual roads are fraught with pedestrians, cars, buses, and their irrational behavior. Studying pedestrians, Dr. Fridman and his colleagues learned that pedestrians cannily calculate whether they have enough time to cross the street ahead of a car. Sometimes the safest driver response to a pedestrian is to speed up so that the pedestrian doesn't attempt to cross. But robots are programmed to be extra safe. Would a car maker ever program a car to duplicate this human practice of speeding up if it observed a pedestrian?

Second, research has found no conclusive data that autonomous vehicles are safer than human drivers. Looking at fatalities per million miles driven, autonomous vehicles have a worse record with one fatality in 11 million miles. That rate is 10 times less safe than human drivers. Admittedly, one event cannot provide an accurate estimate of AV safety, but it certainly prevents a sure conclusion that AVs are safe.

Third, research shows that humans are giving Tesla's primitive "autopilot" feature control much more than expected, turning it on 34.4% of the time on highways. The belief that humans will ignore the road when on autopilot can be tempered by creating a collaboration between driver and car. For example, the autopilot could announce to the driver, "Elevated risk detected, I am stopping for a pedestrian" so that the human knows to pay special attention and understands what's going on. In this shared autonomy model, the AV could also ask the driver a question, such as, "A pedestrian is blocking our road, should I honk?"

Also missing from the supposed AV solution were the unintended consequences of mixing AVs with human-driven cars on the road. Prof. Yossi Sheffi mentioned talking to Eric Schmidt when Schmidt was head of Google, and saying how AVs "drive like grandmas"—very slowly and cautiously—which would create huge congestion problems in cities. A participant agreed that aggressive human drivers would learn to take advantage of cautious AVs.

Who's Minding the Road?

The performance of a collaborative human-AV system depends on the computer's ability to monitor the driver. Is the driver looking at the road? Even if the driver is looking at the road, that doesn't mean s/he is paying attention. Is the driver looking happy, frustrated, or distracted? These visual clues can be used not only for other drivers but the AV's own shared driver, so that that AV can alert the driver if the driver is looking drowsy, for example. However, monitoring the driver can have difficulties in accurately detecting or estimating the driver's condition under challenging lighting situations such as partial sun, shadows, and sunglasses.

A beverage distributor has cameras in each of its 8,500 trucks that can record what happens just prior and after a hard brake, but currently the cameras have no real-time processing ability to detect and alert a drowsy driver. Dr. Fridman noted that Cadillac's "super cruise" mode forces the driver to keep looking at the roadway. Driver-monitoring technology is now relatively cheap, but it still requires an investment in technology to make these systems work in a company's trucks.

Getting Started: People, Methods, Tools, and Data

The final session of the roundtable turned from discussing specific supply chain applications to talking about more general issues of how companies develop and deploy ML. This comes down to understanding:

- the important differences between data science compared to traditional data and technology roles in organizations, such as business analysts and enterprise IT,
- where to place ML in the broader context of the organization's structure, and
- how to bridge the gap between the people who create ML systems and the people who will be using those systems.

Data Challenges

Computing power per dollar doubles every two years, and the cost of storing 1 gigabyte of data has dropped from $300 to $0.80 between 1995 and 2005. Together, large amounts of data and decreased computing costs are driving ML. However, all that data presents some major challenges. The specific data needed by a company may not be readily available or may be inconsistent across data sources. Furthermore, processing large amounts of data is very resource intensive, not only in terms of computing power but also in labor. Data scientists spend 50% of their time just trying to gather data, then clean it, and process it. They also have to identify salient features of the data—feature engineering takes much time and energy in order to prepare the data to gain value from it.

Given the big effort that data cleansing takes, one freight company wondered whether the data science team should influence the way data is captured and apps are built. A furniture retailer said it observed a Data Engineering role emerging for people who were good at working with fast-flowing high volumes of data, databasing it, and making it available. They were focused on the newest tools. This was better than giving the task to the IT department who might simply call a vendor to handle their data. There was nothing wrong with using a vendor, he added, just that it may not be what the data scientists want. A 3PL invested in data scientists and used an open source software application as a data lake to assemble data for training. A normal enterprise IT group may not be aware of those tools.

Data Analysts vs. Data Scientists

Several broader issues are the new roles and skills required by ML that draw a dividing line between traditional business data analysts and data scientists. A 3PL noted that data analysts know the business but don't understand the deeper math or computer science the way the data science team does and needs to. He wondered whether companies were training their analysts to take on a more computer science role. A transportation software company has paired analysts with data scientists to collect data, do descriptive analysis, and understand the business and technical issues together. On the other hand, a beverage distributor argued that everyone on a data science or ML team needs knowledge of all three domains: business, math, and computer science.

A grocer noted that a second difference between data analytics teams and data science teams is that they speak different languages and have different working styles. One is descriptive and the other prescriptive. One works on one-month projects and the other works on a tool for a year, so in their opinion it's better to have them separated. Both add value but at different stages. One company admitted to failing in trying to shift people from analyst roles to data science roles because of this specific issue. The project-focused analysts can address the problem but the work had to be done manually because the team wasn't experienced enough to create a supportable product for the company. A 3PL added that there have to be dedicated teams for the two styles of work because the workflow for projects and products is different.

A related divide is the pay-back horizon of data analyst versus data science work. At a large 3PL, their data science team debates how much time to spend on deep research versus decision support for the business. The participant's team is in R&D, so they spend 60%-70% of their time on deep, long-term projects that have an end goal of automating a decision process. Short term, they do data analysis to derive better business rules rather than relying on a person's own experience. The analytics team can use the same set of metrics as the person uses, but also gives the person a better threshold to improve the businesses' decision process. The data analysis part is used in both long-term and short-term projects.

The Quest for Talent

Two companies spoke of urgent needs to hire data scientists now, particularly people with computer science, statistics, and math skills. However, hiring university graduates can require waiting more than six months for them to graduate. The group had many suggestions for finding talent in unusual places.

Technology-focused online communities were good venues to find people. For example, Hacker News at Ycombinator has a lot of technology-minded members and has a monthly "who's hiring" thread. A participant recommended asking the company's tech team where they spend time online such as the various Google Groups or Reddit's subreddits that focus on discussing algorithms. "We find these hidden groups," he said. Other fruitful communities were those which contribute to open source products, repositories. HackerNews' "Data Tao" group and data science competitions such as Kaggle were also cited. A retailer added that it uses LinkedIn. A shoe company said it found people who had taken online courses such as Coursera or edX related to data science or ML MicroMasters programs.

Two participants complained that data science majors often did not have enough programming skills or fundamentals. For example, they may have had only one class on programming. They don't know how statistical confidence intervals or regression models work, because the Data Science program skips those fundamentals. They do know the tools such as how to train a model in Karat but don't know how to call data from an API or do the front-end coding needed for data manipulation. One company hired people who had masters or PhDs in computer science instead, to get the level of programming proficiency needed.

Universities could do more to deliver good new talent by aligning their curricula and student tool stack with industry practices. Universities still teach Minitab, but industry does not use it anymore. Universities' classroom projects work with small data, not big data found in the real world. Companies need talent that can work with big data using many databases. Students also need to understand that data cleaning and data manipulation will be a big part of the job. Creating an ML model comes after that.

A transportation software company asked how companies can convince would-be job candidates that working on transportation problems is sexy and fun, compared to working in finance. A 3PL replied that new hires like to work with the latest tools such as Python, R, and Tensorflow, so that could be attractive. New hires don't want to use tools from five years ago. It also implies letting new hires use Linux or Mac computers rather than be forced to use Windows-based machines.

Companies have other ways to get the talent they need. Acquiring another company can bring a pool of talent. Training is another option for larger companies. Amazon has created its own ML university to push understanding of this key new technology to all levels of the organization, even the executive level. Microsoft does this, too.

Tools: Open Source vs. Commercial

Many companies (and their data scientists) were reveling in all the free, open source tools—e.g., Python, R, CARET, Tensorflow—that have burst onto the data science and machine learning scene. Various companies praised Google for making TensorFlow available for free. Some companies just use the TensorFlow tool while other companies also give back to the open source community by making improvements to the tool and posting those.

Several comments highlighted a tension between the use of open source software—which are often programs written by programmers for programmers—and commercial software that might be designed for ease-of-use by less-technical users. One retailer mentioned that as an enterprise, they don't like working with open source software because there's no one to call if there's a problem. They don't want to be reading source code at midnight if something goes wrong. Instead, they want on-call professional support from someone who can be responsible for fixing it. On the other hand, a 3PL mentioned that his data scientists wanted to use open source Mapbox as a visualization tool instead of a commercial tool like Tableau because they didn't like the graphical user interface of the commercial tool.

A grocer said that his company moved its data to Microsoft's Azure platform, which bridges the gap between the two worlds. He didn't want his salespeople to be having to enter data into Shiny (an open source framework for making apps in R) or using open source. The Azure environment is easily scalable and has containers for tools so that if data scientists want to build a tool in Spark or R it can be easily integrated because it is all under one umbrella.

Putting ML into the Organizational Structure

Different companies at the roundtable put data science projects in different parts of the organization. For example, a department store had a centralized department focused on data science. The 100 data analysts in the group were organized into multiple areas: digital web analytics which do A/B testing; operations which includes supply chain; marketing campaign analytics; customer analytics; and corporate analytics. Supply chain management included network optimization and price optimization while another team analyzes product returns. Another team would analyze when a product should go to the discount outlet, based on sales percentage, age of the product, inventory on hand, and whether it has been marked down once or twice already.

At a large 3PL, data science teams work closely with operations people and are embedded with the business unit. At a beverage distributor, currently, the company has the function centralized and the debate focuses around the type of project and scale. The company's pipeline of ML projects is currently 80% focused to one product category, so adding a team to each business unit does not seem necessary. The relative advantage of organizing by business unit is that the data scientists will gain more business knowledge of that unit, which is more valuable than the transfer of best practices in coding that occurs with a centralized team. The distributor plans to continue with a center of excellence model for a few years until it has the scale to put teams in every business unit.

Two other participants structured data science within other functional departments. At a carrier, the data science team reported through IT rather than the business and said that structure was partially successful. At a furniture retailer, the data science team was inside of the engineering organization, which worked well because the engineering organization was focused on making money. He had seen how enterprise IT could stall advanced technology projects, but by embedding data science in the profit-focused side, IT did not block it.

Change Management: Putting ML into Practice

A shoe maker asked about best practices for getting business people on board, accepting new apps, and not fearing that "machines are taking over." Participants described five best practices. The first best practice in change management was to not change too much, such as changing people's processes. For example, if the people are accustomed to using the ERP system for purchasing, then let them continue to do so. The new ML system might simply inform them of the recommended items to purchase while letting them make the purchase in the familiar way.

The second was to create a trusted source of data. A freight data company said it was difficult working with companies where the "source of truth" depended on whom one asked and the company couldn't agree on the base data. Thus one challenge to gaining acceptance was to organize the data to make it easy to analyze, readily accessible and trusted by the many stakeholders of the business. Getting all the data into one place took time, however.

The third was to work with the target users. When beverage distributor rolled out its demand forecasting solution, the distributor had to deal with the egos of the salespeople who had traditionally used their experience to make these forecasts. So the company launched the model along with a survey asking salespeople if they used it or not, and if not, why not. The company then collected data on the forecasts the salespeople were using and fed those into the model to see if it would improve the model.

The fourth was to visualize the value to help convince the target users. The beverage distributor also had to assure salespeople that the demand forecasting system was not meant to take away their jobs. Instead, it was meant to help them sell more and to win more sales incentives. Visualization helped users understand the model. If they only saw code, it would only be a meaningless number, but visualization helped users understand the tool. This was particularly true with the exception management features of the demand forecast that helped salespeople focus their labor where it could make the most difference.

Fifth, the beverage distributor used training and seminars, and showed business units how the new apps will add value. Business units have to understand the ML model and how it was trained. The company estimated that 10-15% of an ML project spend goes to training that is tailored to the employee's level. For example, front line employees only have to know how to use the tool, middle managers need to understand how the tool will impact their departments, and C-level people need to understand the strategic implications. The company uses webinars, live conferences, and emails to help explain the new tools.

Supply chain projects can often have an added dimension of change management to the extent that the project affects the company's external supply chain partners, such as the organization's customers and/or suppliers. A 3PL considers model-building for such projects as the easy part, taking a couple of weeks or months to enact the idea. Taking the idea to market is the challenging part. The company has 120,000 customers and each have their own preferences and expectations. The company used ML to automate detection and classification of BOLs for customer service but has not rolled the system out yet. Automating customer service is trickier because a customer is involved, so that is a tougher change management issue.

Finally, sometimes a machine learning solution is not worth the effort of deploying it. For example, Netflix ran a competition with a $1 million award to the team that could improve its movie recommendation engine. The winning team wrote an algorithm that was 10% better, but Netflix decided that the engineering effort to gain that extra performance was not worth the incremental gain.

Creating Future Supply Chains with Machine Learning

A freight company asked about the best way to propose new projects: to hire a principal data scientist or to have a business person supported by junior data science resources? A 3PL recommended an outside hire because the company should want to disrupt itself and avoid being stuck doing things the way they have always been done. A furniture retailer recommended starting small when hiring data science teams, because it was possible to hire too large a team of smart people who cost more than the value they may immediately provide. Going after the low-hanging fruit, such as simply building something that interprets an Excel spreadsheet just a little better, could be enough to demonstrate the value of ML and justify investment in more complicated models later.

The roundtable showed that machines have much to learn about supply chains, and supply chains have much to learn from machines. The growing availability of data and the ability to train models on millions or billions of data points transcends the experience of even the most grizzled supply chain veteran. Understanding what ML can do is a prerequisite to deciding what ML should do.

Overall, the roundtable members showed how ML can be used for prediction of supply, demand, pricing, timing, etc. to proactively manage the future. And ML can cluster and classify supply chain conditions, events, and categories of products or customers. This can help manage complexity through better differentiated responses or tailored best practices. Furthermore, ML can forecast likely exceptions or potential problems such as frozen cargo, disrupted ports, or distracted truck drivers. ML has an endless list of possible supply chain applications.

However, the machines by themselves do not know which problems to solve, which data to use, or how to integrate an ML solution into the business. That duty of managing ML will fall on the shoulders of future supply chain leaders. Those leaders will need to understand enough about what is possible both technologically and organizationally to create what is profitable. Thus, the future of supply chains depends on understanding both the machines and the people well enough to help them work together.

MIT Center for Transportation & Logistics

Report recording prepared by:

Andrea Meyer and Dana Meyer
WorkingKnowledge®
3058 3rd St.
Boulder, CO 80304
workingknowledge.com

Edited by the moderators.