# AI & ML: From Proof of Concept to Production

The current buzz around Artificial Intelligence (AI) & Machine Learning (ML) is at a peak in the data services, and business intelligence (BI) industry as every enterprise software vendor stakes a claim to having features and functionality leveraging it. Reading the marketing claims around AI & ML would lead one to believe that the development and use of the technologies is "push-button simple" and poised to enhance the decision-making power of every system that manages data.
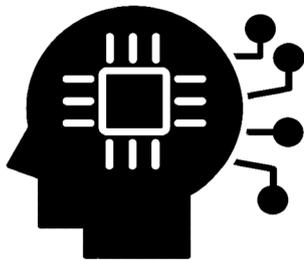
There's no doubt that machine learning algorithms have found a substantial place in many industries and for many valuable use cases. From credit card fraud detection to intelligent image recognition for quality control, ML models exist to carry out highly redundant tasks with high speed and accuracy. In each case, an objective is established (e.g., detect fraud, check quality, etc.), available data is evaluated, a ML model type is chosen (e.g., supervised learning, deep learning neural network, etc.), a model is developed and iterated over to achieve acceptable accuracy and performance characteristics, and finally the model is integrated into a production process to deliver value. While none of these steps are "push-button simple," the last step, integration into production, can be quite a bit easier when models share a common and robust production platform.

At ValueCentric, AI & ML is gaining a foothold as a valuable tool to improve productivity in managing data, improve the quality of the data we manage, and most recently, to provide insights that go beyond the traditional analytics approaches. Here's a brief look at our process of bringing AI & ML from proof of concept to production.

ESTABLISHING AN OBJECTIVE

WHAT PROBLEM ARE WE TRYING TO SOLVE?

It seems obvious, but probably the most difficult and most overlooked element of working with AI & ML is a clear understanding of the business question that a ML model is attempting to solve. The critical element in forming these business questions is, "what is the business value in solving this?" Examples from the credit card and retail sales industry, "Is this transaction fraudulent?", or in the healthcare value chain, "How likely is this patient to encounter obstacles on their therapy journey?", are easy to frame, and their business value is obvious. In any space with large volumes of data, the phrase "apply machine learning to it" is frequently uttered as a process to solve the most pressing problems. In practice, it is a challenge to derive the business question that can be solved, and that provides measurable value to the business.

"what is the business value in solving this?"

At ValueCentric, we are using AI & ML to solve an essential problem in data management: "How to efficiently and correctly assign and classify thousands of new variations of customer locations on inbound transactions each morning." By applying appropriate ML modeling to the inbound transactions to measure sequence similarity (LSH & TF-IDF), we were able to recognize the value in the form of more cost-effective resource utilization, higher quality of data matching, enabling massive volumes of transactional data to be ready for our customers' use in near real-time.

EVALUATION & CURATION OF AVAILABLE DATA
IS THE DATA UP TO THE TASK?

Recently, marketing emphasis on AI & ML suggest a certain mystique to the application of these techniques, but it's essential to understand that at its core, AI & ML is all math. The data elements used in a ML model become the model features. The answers to the business question that the model derives are the model labels. Inside the model, the features turn into numbers, and in turn, the model uses math to derive a very complex and deep set of relationships between the features and the resulting label or arrived answers. The definition of the features from the available data seeks to provide a broad enough view of the conditions, effectively arrive at the business outcome (the label). It is in this step of an AI & ML project where assumptions about the breadth and quality of the available data hurt the most. ValueCentric methodology for good ML modeling and development follows industry best practices by the most informed AI & ML organizations, includes a validation step that helps ensure that the available data is up to the task.

MODEL DEVELOPMENT, ITERATION, & VALIDATION
HOW TO SOLVE THE PROBLEM?

Depending on the objectives defined and the available data, data scientists and machine learning engineers have several ML model types at their disposal. Some problems can be solved with a single model type while most require integrating multiple model types for a complete and effective solution. The model development phase is iterative and based on a continuous improvement methodology where a vital step within each iteration is the validation of model performance and effectiveness. With each iteration, model features are refined, added, and deleted, and if necessary, new model types are introduced to drive toward ever-increasing model performance.

For ValueCentric's classification tools mentioned above, locality-sensitive hash forest (LSH Forest) and term frequency-inverse document frequency (TF-IDF) are two of the multiple model types employed. However, more complex problems may require alternate modeling approached. For instance, ValueCentric is piloting a Specialty Patient Journey Module to identify patients at risk of experiencing delays or impediments to drug therapy and prescribe a corrective approach, for which a long short-term memory neural network (LSTM NN) model is an integral part of the overall solution.

PRODUCT INTEGRATION
## HOW TO OPERATIONALIZE THE SOLUTION?

The key to effective production integration is to have a platform and engine in place that provides for new model training and model predictions driven by a standardized API. This allows for adding new models rapidly with light coding for the integration of the model results in user reporting or backend processing as required.

The implementation of AI & ML-driven solutions within ValueTrak required maintaining the key platform differentiators of the speed of implementation of new solutions, and near real-time access to data for our customers. To effectively do this, a separate, asynchronous, horizontally scalable AI & ML platform was built directly integrated with the ValueTrak processing and reporting engines. The new API driven platform allows new models to be brought online rapidly, and to directly access the required data assets to carry out real-time predictions as transactional data is entered in and processed by ValueTrak. This approach expedited the production use of multiple AI & ML classification tools and facilitate the production release of future AI & ML driven functionality with limited coding and infrastructure modifications.

IN SUMMARY

Much of the marketing around AI & ML remains highly academic, so beware of solutions that speak of theoretical applications of AI & ML without mention of applying real value to business like yours. There's no question that the scope and influence of AI & ML will increase exponentially within a short period of time. Like the introduction of any new technology over the years, momentum builds on the small successes of real, practical use cases as both business and technology see new pathways to value from the hard work of past projects. While the steps from proof of concept to production get easier with each new project, they are still the essential elements for achieving the ultimate business value. When selecting a vendor for your data management and business intelligence, it's fair to believe the hype around AI & ML, but smart to ask for the real examples of AI & ML in practice.
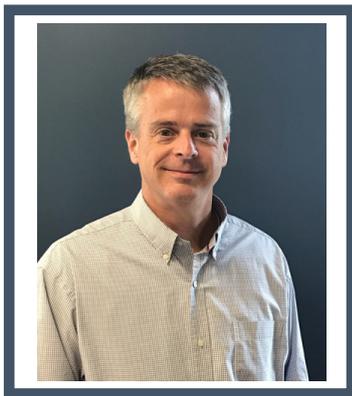
About the author

Steve Zagrobelny has been architecting, designing, and developing data management applications for over 30 years. Having designed and developed one of the first PC-based office automation programs for the US Navy starting in the late '80s through designing and building some of the first web-based SaaS applications in the late '90s and early 2000s, Steve has focused on deriving the highest value from the available data in real-time.

Steve joined Valuecentric in 2005 and has been the lead architect in designing and evolving the ValueTrak platform through to present day. More recently, he has applied his extensive experience in the development of SaaS applications to this new challenge of architecting and developing effective AI & ML solutions for the continued evolution of the ValueTrak platform.

valuecentric
an IQVIA company