Machine Learning Opportunities in the Energy Sector

Vic Batson, VP Sales & Marketing, Sparsix Corporation
Serge Terekhov, CSO, Neurok Techsoft, LLC
September 2009

This white paper introduces the basics and potential of machine learning (ML), a field of computational science that incorporates numerous technologies to create systems that can learn and evolve from the data in their environment. Machine learning was designed to enable the adaptive construction of descriptive and predictive models from real world data flows. This paper will briefly examine the following questions:

- What is machine learning? What are neural networks?
- What are machine learning methods useful for? Where are they best applied?
- What are some of the benefits to using machine learning?
- How are machine learning solutions developed and implemented?
- What are some specific examples of ML-based solutions applicable to the energy sector?
- How much experience does the Sparsix/Neurok team have in developing ML-based solutions?

The challenges of today's economic environment are forcing business in all industries to re-evaluate existing practices, methods and tools, looking for new ways to increase efficiency, reduce costs and create new value. Some of the most promising advances coming from these efforts are based on existing company resources – data already generated by existing processes.

From finance to engineering to operations, managers are asking how they can make better use of the data they already have:

- What does my historical data tell me about what is happening now?
- Can I predict future events from my existing data?
- Can I make real-time, operational decisions based on my data?
- Are there cheaper, faster ways to analyze the data I have?
- Can my systems learn from historical data and adaptively respond to changing conditions?

These questions, and others, pose daunting new challenges for existing data modeling and analysis techniques. As information technology continues to advance and more complex data is amassed, data modeling and analysis tools are confronted with a much more complex and demanding environment:

- Data sets are increasingly multi-dimensional, since target output variables depend on many input factors, making it infeasible to reduce the number of influential factors to an amount small enough to be analyzed by conventional statistics methods.
- Data sets can be extremely "noisy" and may contain a significant number of non-relevant samples. Simple, descriptive statistics cannot deal with such a case.
- Data is non-stationary. It is no longer possible to assume that observed data is coming from some fixed (though unknown) distribution. Fast developing phenomena also expose another complication. Since the time scale of data collection is typically comparable with the time scale of the modeled process, different samples in the data set may correspond to different system states making traditional analysis of a "uniform statistical data corpus" no longer possible.

- Existing data may not provide all of the relevant information necessary to reliably predict the output of a system. If outputs are not strictly dependent on the known inputs, then some relevant inputs are presumably unknown or unobserved and the data set can be assumed to contain missing values or blanks in the sample records.

This demanding environment requires new approaches and new tools to realize the full potential of a company's data assets. The most critical feature in developing new data modeling and analysis techniques is the ability to formulate a proper model, based on both data and domain-specific human knowledge, that is capable of learning or evolving as new data becomes available. This learning process must be capable of making more than just small corrections to the model – new data may signal a new system state and the model should be able to recognize and capture these cases.

What is machine learning? What are neural networks?

Machine learning originated from early research in cybernetics and robotics, where the goal was to understand how to create an artificial mechanism that was able to survive and evolve in an arbitrary environment such as the deep ocean or outer space. To survive, the machine would have to rely mainly on its own experience and sensory interactions and be able to solve certain critical tasks including:

- Sensory information storage and data compression
- Repeatable pattern detection and classification
- Regression analysis of noisy historical pattern sequences to find new expected patterns
- Decision making and control of its own components

Today, machine learning is mostly concentrated on the mathematical algorithms and software devoted to performing these tasks, and others, but in a much broader scope of applications than just unmanned exploration. Machine learning is not a single, particular technique or technology, but is rather a field of computational science that incorporates numerous technologies to create systems that can learn from the data in their environment and then make predictions and take actions when confronted with a new situation. Machine learning is strongly grounded in modern mathematics, drawing on expertise in function analysis, probability theory, sets theory, chaos and dynamic systems and calculus of variations among other areas.

Machine learning algorithms cannot be completely preprogrammed and fixed because application contexts can vary greatly. Instead, a rather broad family of algorithms is selected for a given situation and their variable parameters are tuned (learned) to fit a specific application's data.

Among these machine learning algorithms, artificial neural networks (ANN) are the most common. ANNs are comprised of simple but non-linear basic functions: processing elements, or *neurons,* interlinked with adjustable connections, or *synapses*. Information is propagated through the system (algorithm) from the inputs of some neurons through intermediate units to the outputs of others. Thus, output signals are computed as a function of input signals. With the proper "learning" of the synaptic parameters, this functional dependence can be as complicated as required to represent any given system. Optimization theory provides solid, proven algorithms for optimal learning of these parameters. Well-trained ANNs can avoid, or ignore, local noisy data fluctuations and are able to predict outputs corresponding to new, unseen inputs.

Artificial neural networks may be the best known type of machine learning architecture, but there are many other types, each with their own advantages and disadvantages. Some other useful architectures include radial basis and kernel machines (support vector machines), self-organizing maps, Bayesian networks, probabilistic trees, evolutionary and genetic algorithms, fuzzy logic and neuro-fuzzy machines, and their hybrids. Selection of the proper architecture is dictated by the nature of specific applications and available data.

A more rigid introduction and technical review of machine learning can be found in textbooks [1 - 3].

What are ML-methods useful for? Where are they best applied?

Machine learning methods are most useful in applications where formal, complete algorithms for the solution either are not yet developed, or are too costly and complicated to be utilized in practice. “Costly and complicated” can apply not only to the time and cost to develop the algorithm, but also to the time and cost to compute the results.

For example, optimal control of a chemical reactor may require real-time execution of a non-linear optimization algorithm with the objective functions computed as the solution of a three-dimensional, non-linear system of partial derivative equations. The implementation of such a construct is by itself extremely difficult, but it is additionally complicated by various uncertainties of chemical parameters. Implementation of a traditional, computational approach would be both expensive and time consuming. Instead, the operational history of laboratory reactors can be recorded and used in conjunction with real-time data coming from an operational reactor to train a neural network model of control signals as a function of the process sensors. Of course, actual implementation of an industrial-scale neural controller is far more complicated than this simplified description suggests, but the idea is feasible and has been successfully implemented in real-world systems.

Another useful application for machine learning is forecasting an economic time series or an electricity demand load. In both cases, certain aspects of future behavior can be logically formulated, but a significant portion of variances cannot be captured by theoretical formulas. Machine learning methods assume that complicated dynamic factors frequently occur in a similar manner as they have in the past. A learned system can recognize and classify observed aspects of recent behavior, correlate them with multiple historical events and predict future events. In these applications, it is extremely important that the ML algorithms take into account the real data distributions as they actually exist, including all “heavy-tailed” distributions and “jumps”, and not just a few statistical indicators as is common in standard statistical applications. This robust knowledge of past and present behavior enables a truthful risk assessment as well as automatic accuracy estimation for all predictions.

Machine learning methods are also frequently used in combination with traditional computational methods providing new levels of adaptation and intelligent behavior. For example, one of the applications developed by Sparsix’s sister company, Neurok Techsoft, is an intelligent solver assistant. This automatic, neuro-genetic software adaptively optimizes the performance of an industrial solver of linear algebraic equations for petroleum reservoir simulations. On each iteration of the solver, the assistant compares both current and previous results against objectives and optimizes the parameters of the solver to speed convergence to a solution.

There are no simple rules for determining where to use or not use machine learning methods. Important prerequisites for their successful application are the availability and quality of data and clear formulation of the application requirements. As with any development project, it is important not to over-complicate the design. For example, if the main purpose of a forecast is to apply a simple threshold decision rule by comparing the model output with a predetermined level, there is no value in computing the detailed output magnitude using a complicated regression analysis method. It is sufficient to use a much simpler binary classification method where the model output *is* the final decision (yes-no) instead of some intermediate numerical value.

What are some of the benefits to using ML? Where is ML a better solution than a more traditional, computational approach?

There are a number of benefits frequently associated with machine learning applications:

- ML-based solutions can sometimes solve previously intractable problems by bypassing complicated and time-consuming calculations and providing a solution based on recognition of patterns and relationships within the data. One notable example of this is solutions for credit card fraud detection which have transformed this sector by providing performance that is simply unobtainable using traditional computation approaches.
- ML-based solutions are "lightweight" compared to mathematical simulators and extended database searches. Neural computations are extremely fast and can be massively parallel, which is critical for real-time applications.
- ML models are compact and can be implemented using embedded technology such as DSP processors, special-purpose computers and even custom chips. This makes ML-based solutions ideal for industrial applications.
- From a user's point of view, ML models are frequently more informative because they can quantify both the model output and also the whole distribution of results and uncertainties. This quantification is critical in risk management applications.
- ML algorithms and models can dramatically expand the range of a solution's capability and applicability. For example, in industrial monitoring solutions, standard linear PID-controllers are only reliable as long as a system's state can be approximated as a linear function. If the system goes "non-linear" PID-based controllers may fail to respond. However, a neural network-enabled controller typically has a much wider zone of controlled parameters and can quickly and accurately respond to both linear and non-linear changes in state.
- ML algorithms are good at improving on traditional technologies through more sophisticated data utilization. One example of this is *virtual sensors*. Consider two measuring devices, one of them much more expensive and comprehensive than other, but fragile and restricted to a laboratory environment. A neural algorithm can learn how to transform the output of the cheaper device to match the output of the expensive device but allow for the cheaper device to be deployed in industrial conditions.
- ML implementations can facilitate low-cost support and updates. Typically, updates for bug-fixes or functionality enhancements require only "plug-and-play" sets of model parameters that can be provided through XML files or similar methods making the update process quick and easy.

How are ML solutions developed and implemented? How much effort is required? What are the critical success factors in implementing a custom ML solution?

Small and medium-scale machine learning models and feasibility experiments can be performed by consumer-oriented software, both commercial and public domain. These applications are usually limited to separate, specific technologies and simplified, or "pure", model formulations.

Enterprise-level applications typically require customized solutions developed by ML specialists in cooperation with domain-specific specialists. The basic machine learning and optimization algorithms are organized into specialized software libraries that have been developed in-house and refined over the years to simplify the process of developing ML-based solutions. For the final solution, these libraries are integrated into domain-specific application environments. For example, in the case of the linear solver assistant, the neural net module was integrated with the linear solver into the customer's industrial petroleum reservoir simulation application.

Data collection and model testing are the most expensive, time-consuming and critical phases in the implementation of machine learning technology. The model's learning algorithms are subjected to statistical tests on independent data to quantify error limits and the stability of classifications and forecasts. For example, model testing for the pipeline defects recognition system lasted several months and even continued after operational deployment in industrial conditions.

In the case of mission-critical applications, several ML algorithms are usually developed and tested in parallel. In addition, "committees" of different algorithms can be used simultaneously to reduce the variance of predictions and further mitigate risks.

One critical success factor relates to the peculiarities of formulating specific machine learning tasks for a particular application. Many engineers and numeric analysts are not familiar with evaluating a problem or task from a machine learning perspective, for instance isolating classification or non-linear control tasks. Successful formulation of the specific task an ML-based system must perform is critical to the success of the project, and is frequently produced in an iterative process of several feasibility studies. Working with a team experienced in designing and implementing ML technologies can dramatically shorten this process.

The most critical factor in ensuring the success of a machine learning project is the quality and completeness of relevant data. Partial machine learning methods can be used to assist in the process of data preparation. For example:

- Self-organizing maps can visually represent a multidimensional corpus of data
- Missing-values imputation methods can test for and work to correct outliers and blanks
- Probabilistic trees can be used to estimate the relative importance of certain types of data and significantly reduce the amount of data necessary to ensure a robust and accurate model

What are some specific examples of ML-based solutions that are applicable to the energy sector?

There are numerous applications in the energy sector that could potentially benefit from machine learning technology. Some illustrative examples include:

- Demand load forecasting. These forecasts may improve short-term operations by avoiding bottlenecks in power distribution. Results of forecasts can be embedded into dispatch systems as additional constrains and as part of optimized objective functions (such as introducing penalty terms to the deviations of a trial solution from a forecasted value at a specific node). These forecasts may also be valuable when forecasting day-ahead spinning reserve requirements or proposing good initial values for the solution of locational pricing.

- Demand load monitoring. Machine learning systems can quickly discover unusual patterns in electricity consumption (over and above common and expected deviations such as time of day, seasonal, weather, etc.). These patterns may indicate potential malfunctions in grid components as well as loss factors not included in EMS models.

- Data health monitoring. When deviations between measured values and forecasted values are found in only a small fraction of observed data points, these deviations may be classified as a sensor fault. In these cases the data readings should be considered as "missing" and can be imputed, or modeled, from relationships in the majority of valid data points.

- Neural network-based predictive controllers for nonlinear control systems. This topic has been widely discussed in academic papers over the past few years, but little if anything has been published about actual implementations. Working in parallel with standard linear controllers, these controllers could correct for large critical errors when the standard controllers allow a system to move outside of its operational boundaries because of nonlinearities. To achieve this, the neural network would need to be “taught” using high-deviation data instead of “normal” data as contemplated in the majority of academic papers.

- Forecasts of financial instruments such as financial transmission rights and day-ahead pricing. The primary purpose of these forecasts would be to provide currently used optimization solvers with good initial approximations of the solutions thus leading to significant improvement in computation speeds. These forecasts could also be used to detect and isolate unusual and unstable pricing situations which may indicate internal problems with market structures.

How much experience does the Sparsix/Neurok team have in developing ML-based solutions? What are some specific examples of solutions developed by the team?

For nearly two decades, the Sparsix/Neurok team has been developing and deploying ML-based solutions across numerous different industries. Some of those applications include:

- Industrial chemical process control. This solution predicts the correct chemical process formulation based on past history and current process and environmental properties. The system was implemented for the beauty care division of one of the world’s largest consumer products companies.

- Non-destructive recognition, evaluation and quantification of defects in industrial oil pipelines. Neural networks and probabilistic trees were used to discover anomalies in magnetic flux leakage and ultrasonic sensor signals. By carefully “teaching” the system on test samples of defects with known sizes, the virtual sensing technology is now able to predict defect size and depth for industrial measurements in pipelines, allowing the system to quickly characterize defects as critical or non-critical.

- Adaptive performance control of iterative linear solver. A neural network engine with local memory optimized by a genetic algorithm is able to track the performance of the iterative linear solver in real time, during computations, and recommend optimal sets of solver parameters for the next iteration providing for fast, accurate convergence.

- Prediction of financial indices from textual news streams. In this application news streams from 20 high-quality sources are digitized and used as the input to a neural network predictor of a financial time series. The technology complements traditional forecasts based solely on time series history providing for a more robust overall forecast.

- Recognition of high explosives for airport security checkpoints. A microwave signal is used to detect the presence of explosive material, but the signal suffers from numerous variable influences such as body conditions, mass and shape of the explosives, reflections from the surrounding environment, etc. A neural network was used to find certain non-linear combinations between different frequency readings which are most stable under all of these varying perturbations. This is a very difficult and highly non-linear task, but with sufficient training a neural network provides an excellent solution.

Conclusion

Professionals in the energy sector will face many increasingly daunting challenges in the future: balancing the ever-increasing demand for energy while maintaining and expanding a reliable grid infrastructure, integrating new forms of renewable energy and ensuring reliable delivery, and increasing transparency and liquidity in rapidly expanding energy trading markets, to name just a few. Addressing these challenges will require advances across many disciplines of science and engineering, but a common factor in future successes will be better decision-making based on data generated by both existing systems and new technologies being deployed.

Machine learning offers great potential to empower new approaches to these challenges. Like all technologies, ML is no “silver bullet” that can solve any problem, but it is a proven methodology that can, and will, provide a significant competitive advantage to those who choose to utilize it. Academics have spent years studying the potential applications of ML technologies to challenges in the energy sector, but so far few if any of these ideas have been developed and deployed in a real-world solution. Now is the time to take this next step.

Developing an operational, ML-based solution takes a team experienced in the intricacies of both machine learning and the application domain. The Sparsix/Neurok team has been developing and deploying real-world ML solutions for almost two decades by partnering with experts in specific application domains to create world-class teams. Sparsix focuses its resources on the challenges in the energy sector, is actively partnering with leading energy sector participants, and welcomes contact from those ready to explore the benefits of machine learning and to learn how machine learning can help an organization achieve its goals.

----

[1] C. M. Bishop. Pattern Recognition and Machine Learning. Springer, 2006

[2] D.J.C MacKay. Information Theory, Inference, and Learning Algorithms. Cambridge, 2003

[3] B. Scholkopf, A.J. Smola. Learning with Kernels. MIT, 2002