KONGSBERG

Whitepaper

Hybrid Modeling

Unlocking the business value of AI in heavy asset industries

kongsberg.com

Content

EXECUTIVE SUMMARY

Hybrid Machine Learning can provide the edge in moving up the ladder of digitalization and unlocking solutions that can transform your value chain. In this white paper we take a closer look at the limitations of the current methods and how hybrid machine learning can help overcome these.

Traditional methods for data analysis come short with the complexity of constantly evolving assets in heavy industry. Actionable insights require significant domain expertise, knowledge about the specifics of the facility of interest, as well as overcoming technical and administrative challenges to get access to and analyze the data.

For example, when it comes to Digital Twin deployments, we see that they often start with the concept of improved situational awareness and collaboration. However, the real game-changer comes when you can fully integrate forecasting and optimization and build robust solutions for data-driven insights at scale.

Using Hybrid Machine Learning we have seen that we can dramatically speed up the time to market for advanced applications such as real-time forecasting and optimization. These have in turn the potential to significantly reduce OPEX and CAPEX in the facilities where it is deployed.

If you aspire to become a more data-driven company, and want to learn more about how to utilize the true potential of emerging technologies to drive real business value, then keep on reading.

Do you want to know more about how to unlock the business value of AI in heavy asset industries?

INDUSTRIAL AI: FROM HYPE TO REAL-WORLD APPLICATIONS

How to utilize emerging technologies to drive business value.

AI "HYPE"

- Pros: Attention concerning possibilities offered
- Cons: Potential mismatch between expectation and reality

VALUE vs. HYPE

- AI solutions offer great value when applied on the right use-cases
- Important: Have a clear understanding of applicability!

Artificial Intelligence (AI) and Machine Learning (ML) are all over the media, and everyone wants to be involved in the technology race. Progress during the past few years has been tremendous, and you have probably heard claims such as "Data is the new oil" and "AI is the new electricity". We will not comment on those statements, but what we can all safely agree upon, is that there is definitively a lot of interest concerning these technologies. However, all this attention begs the important question: **can it really live up to the hype?**

Within certain areas, the technology has caught up to and even surpassed the hype. For example, within image recognition, the task of identifying objects and extracting information from images, AI is now going beyond human-level performance (i.e., machines are actually getting better than humans in identifying objects and images).

The pros about all the technology buzz include a lot of attention concerning the possibilities these technologies offer. However, many have started to become a bit fed up with the overpromise of AI, and how it can be used to just magically solve any issue. The approach of throwing ML at any problem where you have lots of data, expecting magic to happen and useful insights to emerge on the other side is not really ideal.

This is why it can lead to a mismatch between expectations and reality in some cases. This makes it even more important to go beyond the hype and show how AI/ML and data analytics can be applied correctly in order to solve business-relevant cases and provide real value.

Recent statistics claim that more than 80% of all ML projects fail, and never make it beyond a limited proof of concept. But, importantly, it does not have to be that way.

One of the main reasons behind this grim statistic is the lack of proper cross-functional collaboration.

Data and algorithms are not enough. First and foremost, **it is crucial to have a good understanding of the problem you are trying to solve**. You need someone with the domain knowledge, who understands how the data is captured, the possibilities and limitations within a system, and how the solution can be applied in practice. When applied correctly on the right use-cases, **AI solutions can offer great value to your business**. This is a crucial point. The Technology itself is only an enabling tool. It is how and when we should use it which is really the key factor.

The AI revolution: Why now?

Why do we suddenly have this hype around AI?
AI and ML have been around for a long time, but some key factors explain why these technologies have started to kick off during the last couple of years.

One crucial factor is of course the amount of available data. Big data is suddenly everywhere. From scarcity and difficulty to find data (and information), we now have a deluge of data. In recent years, the amount of available data has been growing at an exponential pace. This is in turn made possible due to the immense growth in the number of devices recording data, as well as the connectivity between all these devices through the internet of things. This is also becoming ever more relevant in the realm of industrial internet of things, with an ever-increasing amount of connected sensors and various other real-time data sources.

Of course, having access to a vast amount of data is one thing. Another is being able to process these data in order to extract useful information. Access to cheap and powerful computing resources has been vital to driving progress within AI and ML.

As an example of the tremendous development in computing power, one can get essentially the same amount of computing power from an ordinary gaming computer today as that of a state-of-the-art supercomputing facility 15 years ago.

The complementing factors of access to Big Data and massive computing power have made it possible to process data on a scale that was not feasible just a few years ago.

Digital transformation, digitalization, Industry 4.0, ...

These are all terms you have probably heard or read about before. However, behind all these buzz words, the main goal is the use of technology and data to increase productivity and efficiency. The connectivity and flow of information and data between devices and sensors allow for an abundance of available data. The key enabler is then being able to use these vast amounts of available data and extract useful information, making it possible to reduce costs, optimize capacity, and keep downtime to a minimum. This is where machine learning and data analytics comes into play.

Cross-functional collaboration to build good products

However, to go beyond the hype and build solutions that provide real value to your business, there are several important ingredients to consider. The technical aspects concerning algorithms, computers and data represent one part of this equation, but to build good solutions, you need cross-functional collaborations.

Domain knowledge

On the one hand, you need someone with domain knowledge about the problem you are trying to solve. What are the possibilities and limitations within a system, and how can the solution be applied in practice?

Data Science

You also need the capabilities within Data Science, which involves everything related to analytics, statistics, signal processing, machine learning, artificial intelligence and deep learning. Essentially, methods and techniques to extract and utilize patterns and information in your data.

Software engineering

Software engineering skills are crucial for building good data -driven solutions. This involves setting up infrastructure for harvesting and processing data through proper pipelines and managing access to data, and in the end building functional and user-friendly software tools for the end-users.

AI IN HEAVY-ASSET INDUSTRIES: OPPORTUNITIES AND CHALLENGES

There is no lack of use cases, for example within the oil & gas industry. The capital cost of a new oil platform can be enormous, exemplified by the new oil platform Johan Sverdrup estimated to cost $ 10 billion.

Any anomalies slowing production will, therefore, incur large costs. A system that detects component fatigue, suggesting when parts should be replaced to avoid failures or give recommendations for optimizing production would provide substantial benefit. Increasing rig uptime and performance while at the same time improving safety could save these companies millions a year.

The challenges of Industrial Data Science

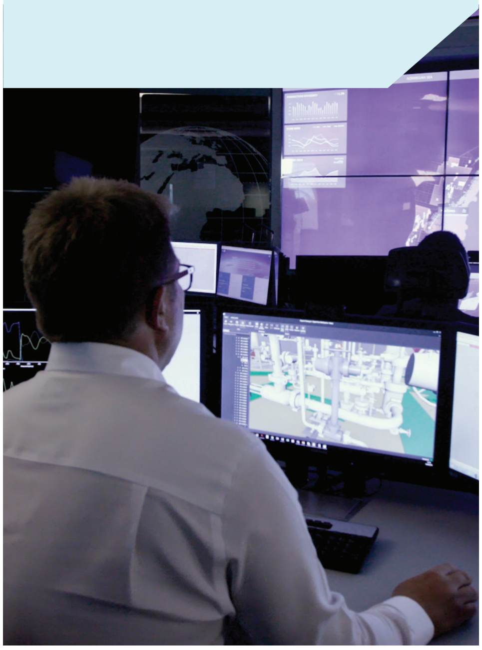
We want to be data-driven. However, traditional data science approaches don't cut it if you don't have the required data quality. It is commonly known that purely data-driven models have significant difficulties when used to analyze and predict time series in safety-critical applications. This could lead to poor decisions or at worst disastrous consequences.

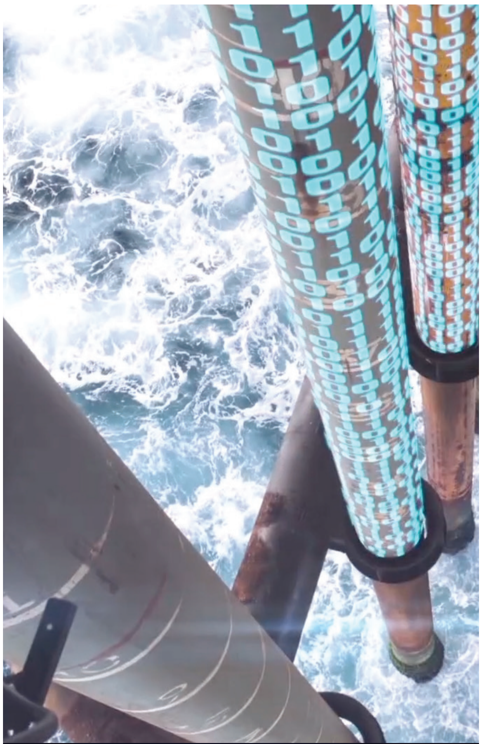
Traditional methods for data analysis come short with the complexity of constantly evolving assets in heavy industry. Actionable insights require significant domain expertise, knowledge about the specifics of the facility of interest as well as overcoming technical and administrative challenges to get access to and analyze the data.

One of the most common problems is that we only see a limited time span of data, in many cases not enough to learn all the insights you need. Adding to that, the process you are trying to model might be changing, so data might quickly get outdated. And maybe the worst of them all, the facility might not have explored enough of the operating space for the data-driven model to really learn the underlying physics. It is also difficult to design models that also incorporate energy and mass conservation, which makes the model predictions less reliable.

THE DANGERS OF BASING YOUR DATA-DRIVEN DECISIONS ON POOR DATA ARE NUMEROUS:

- Outdated decisions
- Biased decisions
- Missing out on potential improvements
- Danger of damaging equipment, environment or personnel.





On the other side, we have the tried and proven high fidelity simulators. Relying on first principle physics models ensures that the system can be modelled in the full domain of operations. However, these models are computationally very expensive, making frequent predictions and real-time optimization very difficult. There can also be a small mismatch between the plant and the model as it often does not model deterioration, or it relies on uncertain or changing hidden parameters.

All in all, neither are on their own enough to reach the predictive and prescriptive level of digitalization desired as part of a Digital Twin solution for your asset.

To understand how Hybrid Machine Learning can provide the edge in moving up the ladder of digitalization and unlocking solutions that can transform your value chain, we need to have a closer look at the limitations of the methods currently used to analyze and optimize the heavy asset industry, and how hybrid machine learning can help overcome these.

WHAT IS HYBRID MACHINE LEARNING?

Hybrid Machine Learning enhances and constrains data-driven models by incorporating knowledge of the physical world using high fidelity simulators.

Data-driven models often have some advantages over first principle models, as they can learn statistical dependencies of unmodeled dependencies and run inferences orders of magnitude faster than the simulators.

Physical models still have several advantages in terms of robustness and transparency facilitating the use in the real world.

The goal of Hybrid Machine Learning is to take advantage of the benefits of both data-driven models and physical simulators to allow new solutions for the heavy asset industry.

	PHYSICS-BASED MODELS	DATA-DRIVEN MODELS	HYBRID ML
Extrapolation Good predictions outside observed data	✓		✓
Proven accuracy	✓		✓
Explainable predictions “Gray box” predictions	✓		✓
Can generate **synthetic data** and **what-if scenarios**	✓		✓
Reliable for safety-critical operations	✓		✓
Does not require access to **high-quality data**	✓		✓
Fast inference Fast and real-time predictions		✓	✓
Quick implementation Quick and scalable development and deployment		✓	✓
Unmodeled dependencies Capability to learn things that you maybe do not notice		✓	✓
Good cost-to value			✓

Machine learning versus physics-based modeling: A gentle introduction

With enough information about the current situation, well-made physics-based models and simulator tools enable us to understand complex processes and predict future events.

Such models have already been applied across our modern society for vastly different processes, such as predicting the orbits of massive space rockets or the behavior of nano-sized objects which are at the heart of modern electronics.

The ability to make predictions is also one of the important applications of machine learning (ML). A common key question is how you choose between a physics-based model and a data-driven ML model. The answer depends on what problem you are trying to solve. In this setting, there are two main classes of problems:

1. We have no direct theoretical knowledge about the system, but we have a lot of experimental data on how it behaves.

If for instance, you have no direct knowledge about the behavior of a system, you cannot formulate any mathematical model to describe it and make accurate predictions.

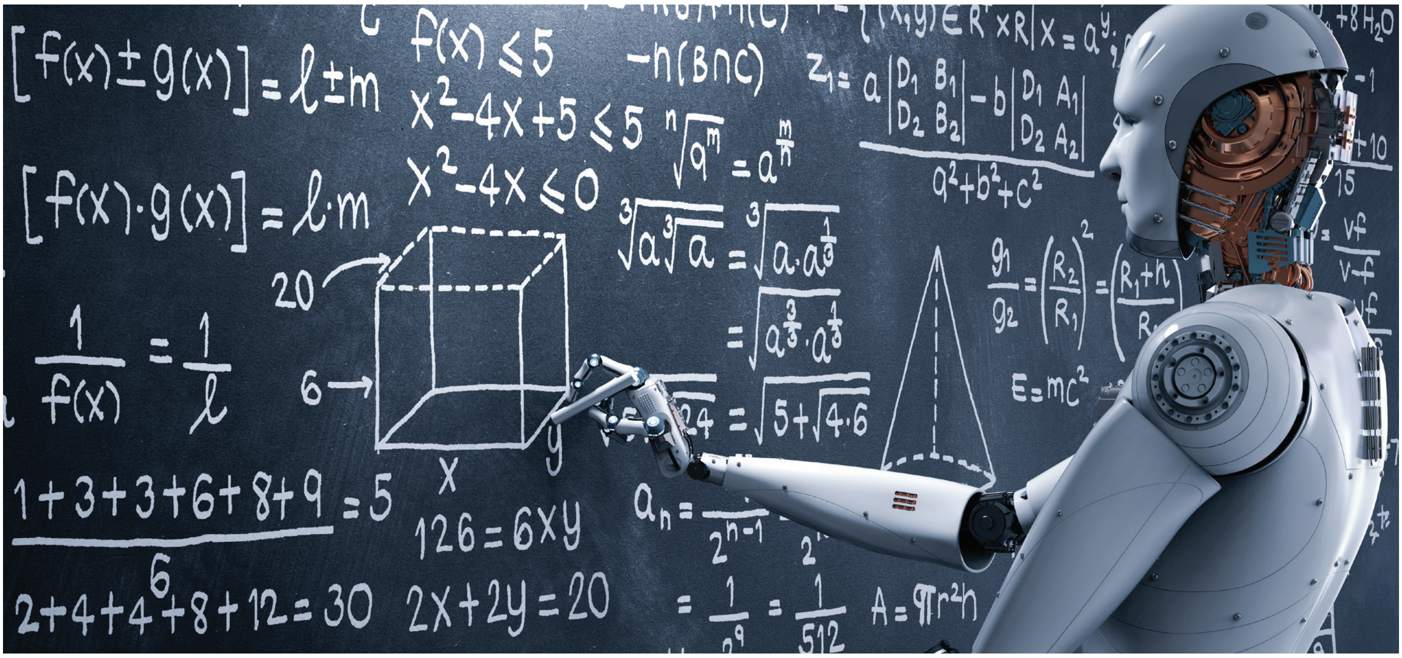
Luckily, all is not lost. If you have a lot of example outcomes, you could use an ML-based model. Given enough example outcomes (the training data), a ML model should be able to learn any underlying pattern between the information you have about the system (the input variables) and the outcome you would like to predict (the output variables).

An example of this could be predicting the housing prices in a city. If you have enough examples of the selling prices of similar houses in the same area, you should be able to make a fair prediction of the price for a house that is put up for sale.

2. We have a good understanding of the system, and we are also able to describe it mathematically.

If a problem can be well described using a physics-based model, this approach will often be a good solution.

This does not mean that machine learning is useless for any problem that can be described using physics-based modeling. On the contrary, combining physics with machine learning in a hybrid modeling scheme is a very exciting prospect. So exciting, in fact, that it is being studied in-depth. In connection with our work, we have recently been deep-diving into this intersection between machine learning and physics-based modeling.



Hybrid analytics: Combining machine learning and physics-based modeling

Even if a system, at least in principle, can be described using a physics-based model, this does not mean that a machine learning approach would not work. The ability of ML models to learn from experience means they can also learn physics: Given enough examples of how a physical system behaves, the ML model can learn this behavior and make accurate predictions.

This ability of learning physics through experience rather than through mathematical equations is familiar to many of us, although we may not realize it: If for instance, you have ever played football, you probably tried to make the perfect shot. And to do that, you had to predict the path of the ball accurately. This is a somewhat complicated physics problem that includes several variables such as the force at which you kick the ball, the angle of your foot, the weight of the ball, the air resistance, the friction of the grass, and so on and so forth.

However, when a football player kicks the ball it is not a result of complicated physics calculations performed within a fraction of a second. Rather, he has learned the right movements from experience and obtained a gut feeling about making the perfect shot.

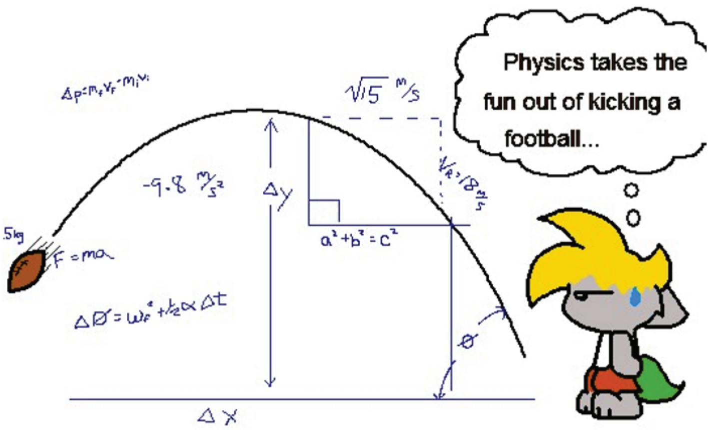


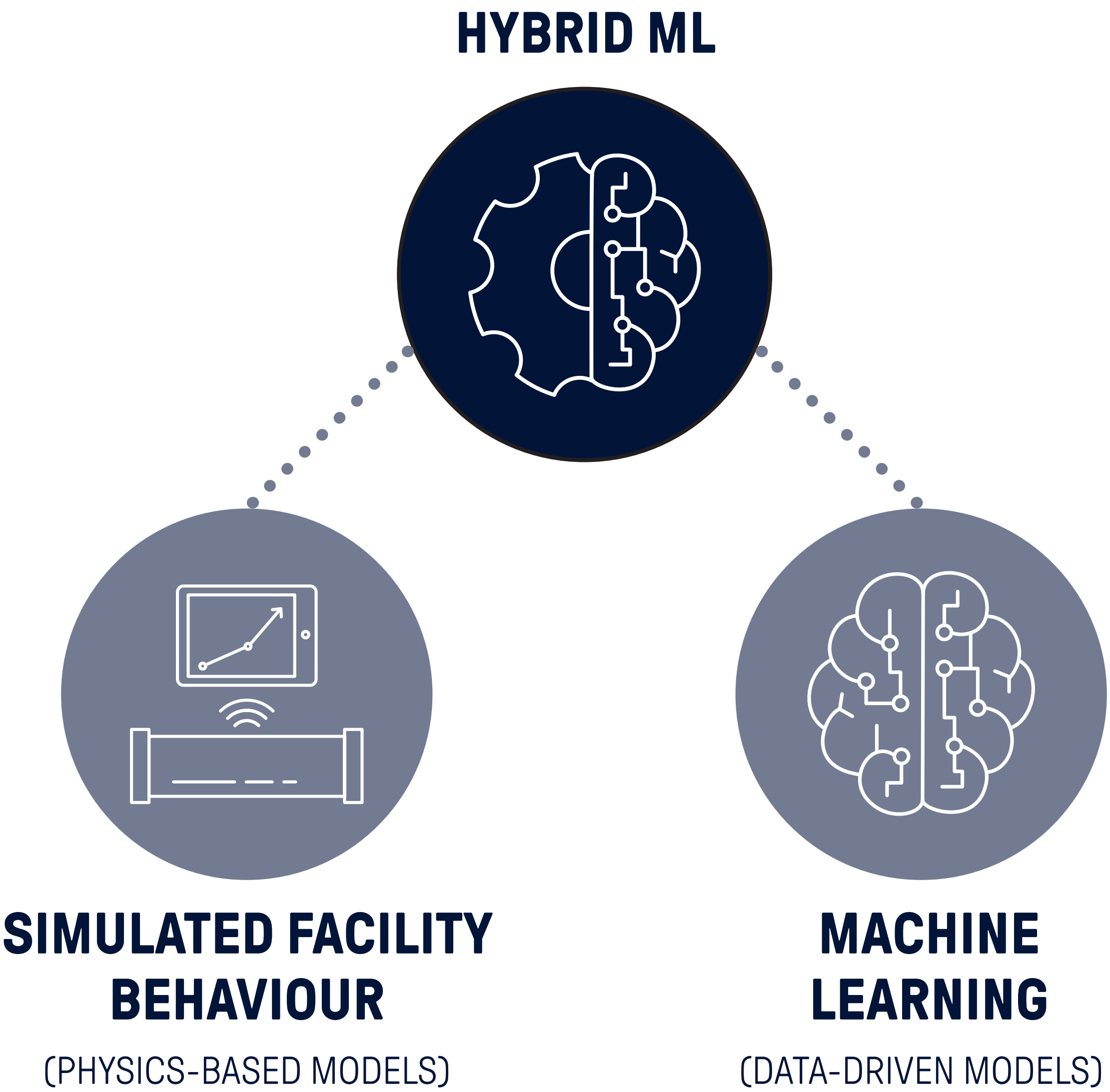
Illustration: Chaos Komori

The fact that ML models — or algorithms — learn from experience in principle **resembles the way humans learn.** A class of ML models called artificial neural networks are computing systems inspired by how the brain processes information and learns from experience.

This ability to learn from experience also inspired us to teach physics to ML models: Rather than using mathematical equations, **we train our models by showing them examples of the input variables and the correct solution.**

HYBRID ML: FROM CONCEPT TO DEPLOYABLE SOLUTION

We define Hybrid ML as: **"A method of doing data-driven modelling enhanced and constrained by first principle models using high fidelity simulators"**.
An important question is why we should want to implement a ML-based approach when we have a physics-based model that is able to describe the system in question. The basic idea is to leverage the **speed** and **flexibility** advantages of the data-driven methods but ensuring the **robustness** and **quality** of the high-fidelity simulators.

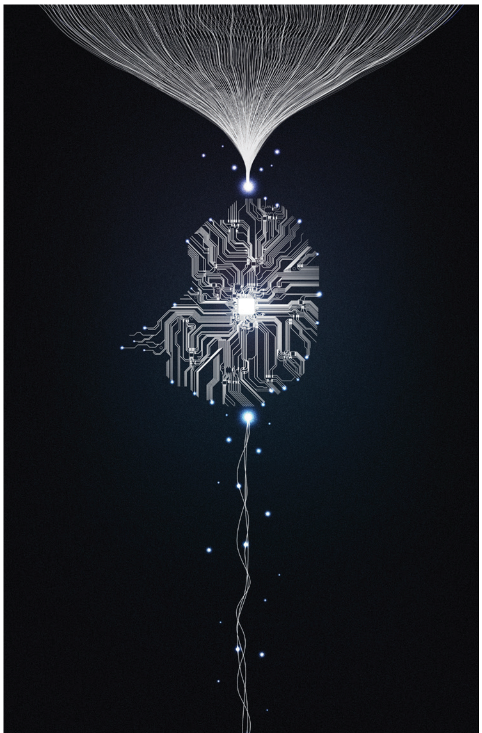


Simulated data enhances data-driven models

One way to improve data-driven models is by simulating data for a wide range of operating conditions.

Historical data records often lack the full span of data necessary to get robust and accurate models in unseen conditions.

However, artificially we can generate data in all the potential operating scenarios. The data-driven model will then work as a form of interpolator expanding our pre-generated simulations to the entire operational space. We can also estimate the accuracy in how well it emulates the simulated data. This also gives us the option to determine the uncertainty of the inference prediction. Further improvements can be made by fine-tuning the data-driven model to match the historical records available.

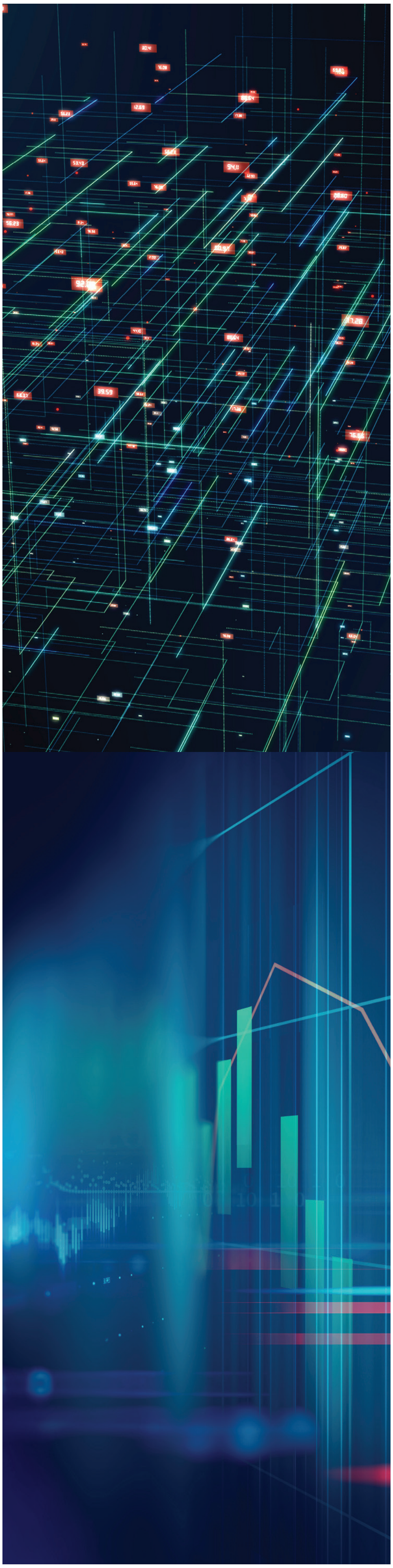


Data-driven models to enhance simulators

Hybrid ML can also be used to improve simulators to better match reality.

Sometimes we find dependencies in the model that are hard to determine from the datasheets. This could include deterioration parameters in equipment or boundary condition behavior outside the scope of the model. One such example is pipeline P-Q curves. These can change slowly over time as well as sudden step changes if well configuration changes on the other side of the pipeline.

Being able to estimate a good candidate for such models could be very useful for improving the match of the model in “what if” scenarios.



Computational cost

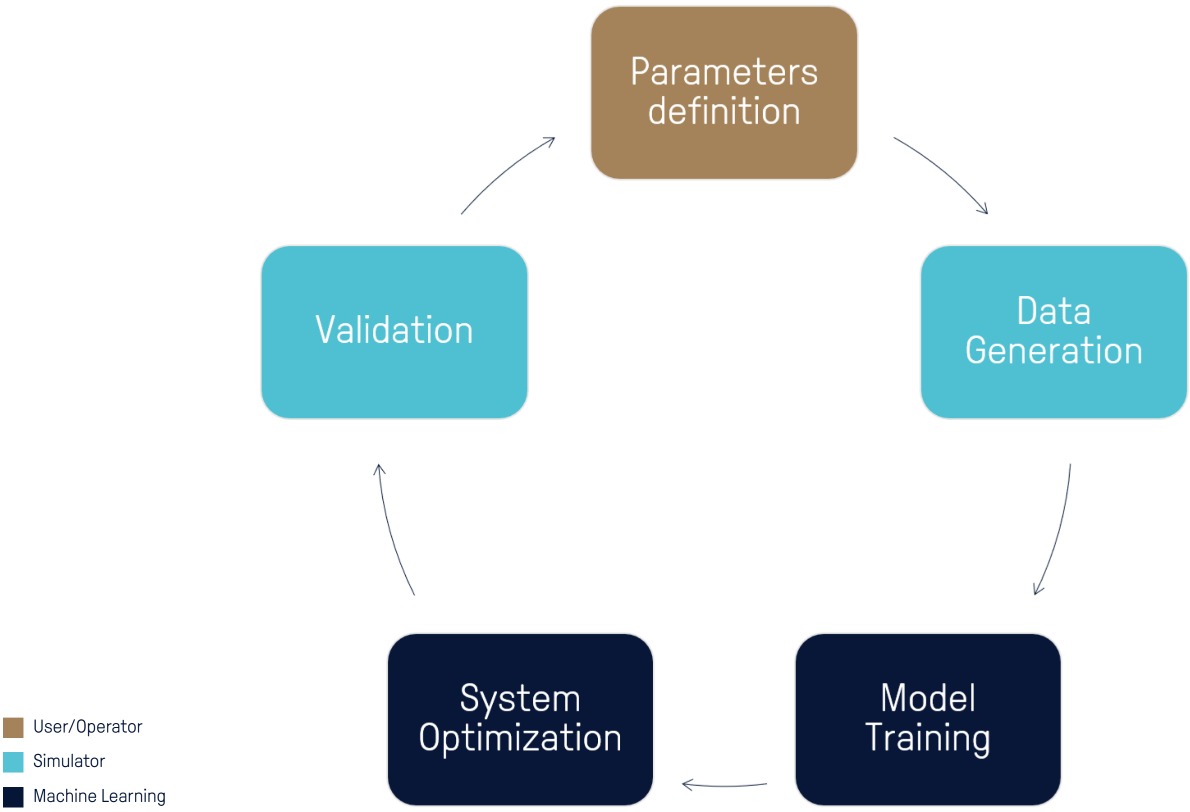
Another key aspect is the computational cost of the model.

We might be able to describe the system in detail using a physics-based model, but solving this model could be complicated and time-consuming.

Thus, a physics-based approach might break down if we aim for a model that can make real-time predictions on live data.

In this case, a simpler ML-based model could be an option. The computational complexity of an ML model is mainly seen in the training phase. Once the model has finished training, making predictions on new data is straightforward. This is one example where the hybrid approach of combining machine learning and physics-based modeling becomes highly interesting.

A typical Hybrid ML pipeline will consist of 5 steps



1. **Objective parameters definition:** We first choose the key objective parameters, an example of this could be the energy consumption of the facility. We also need to find the key constraints of the system as well the free variables and boundary conditions. Complex parts of the process will need to be modelled using high fidelity simulators.

2. **Data generation:** We use the high-fidelity simulator to generate training data in the full operating range of the equipment. Collecting this kind of data would not be possible in the real world, as you would not want to move your facility close to the limits of operations.

3. **Model training:** We can now use Machine Learning to train data-driven models with this data, before fine-tuning it to historical data to improve the model-plant mismatch.

4. **System Optimization:** The fast approximations to the full model can now be used inside an optimization loop to suggest new operating conditions given changes in the field in real-time.

5. **Validation:** Potential improved operating conditions can then be simulated and validated using the original simulator to ensure the robustness and safety of the proposed solution. Then, We feed the solution back to the operator.

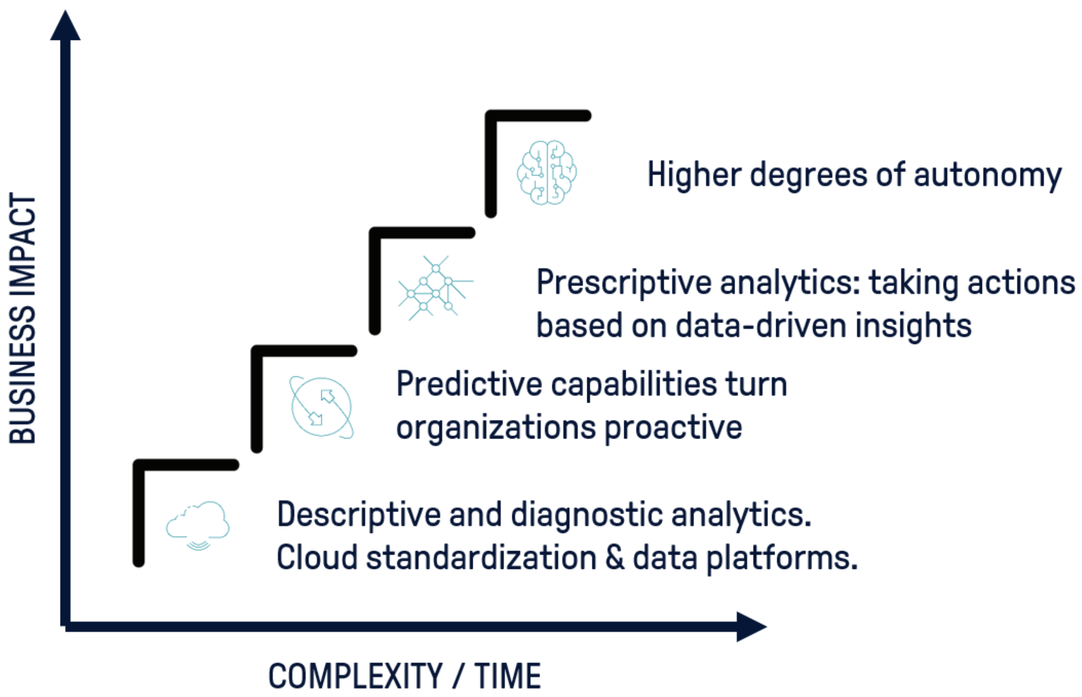
With Hybrid ML, you can **combine the best of both worlds**. The speed and versatility of the data-driven models, but with the explainability and reliability of the simulation models.

Importantly, in addition to the speedup and ability to **explore massive "what if" scenarios**, you can also verify model output according to the physical constraints of the system to ensure you are within safe operational bounds and **maintain asset integrity**.

This is one of the reasons we are using data-driven analytics like neural networks combined with physical models, to **verify and build trust in model output for safety-critical applications**.

DATA-DRIVEN INSIGHTS AT SCALE

For Digital Twin deployments we see that while they often start with the concept of improved situational awareness and collaboration, the real game-changer comes when you can fully integrate forecasting and optimization into the solution.



Kognitwin Energy, a dynamic digital twin for Energy operations, goes beyond the visuals by using proven Hybrid Machine Learning technologies, and leveraging the best of 30+ years of experience from high fidelity simulators with recent advances in Machine Learning.

Using Hybrid Machine Learning we have seen that we can dramatically speed up the time to market for advanced applications such as real-time forecasting and optimization. These have in turn the potential to significantly reduce OPEX and CAPEX in the facilities where it is deployed.

Kognitwin Energy is making this possible through:

- Enabling easy and secure data access
- Contextualizing the data to provide greater insights
- Helping you manage your data sources to automatically detect data quality issues
- Allowing you to connect physical simulators and ML models enabling real-time insights

Always be one step ahead

Through Kognitwin Energy, you can bridge the gap between your assets and your data-driven insights. Beyond being a virtual replica of your industrial facility, our dynamic digital twin delivers a rich framework for advanced digitalization and analytics, including a range of solutions that can be customized to attend to your needs.

The combination of physics-based models and data science approaches and cloud scalability lends operators to streamline and scale testing of hypothetical scenarios.

This allows for improved prediction of impact, options comparison, and increased quality in decisions making. Leading to overall improved performance and productivity, whilst upholding safety levels, and lowering energy consumption in energy facility operations.

Unify data, knowledge and people. Enable the best decision, every time.

Click here to watch the Kognitwin explained video!

AUTHORS

Vegard Flovik, PhD
Analytics Lead - Kongsberg Digital
vegard.flovik@kdi.kongsberg.com
LinkedIn: /vegard-flovik | Medium: @vflovik

Vegard combines a solid practical background from the energy industry with strong analytical skills, that enables him to approach complex problems through mathematical models and computational algorithms. Vegard holds a Master and PhD in Physics and has significant experience with AI/Machine Learning and data analysis. Based in Norway, he currently manages a team of scientist and engineers.

Eivind Roson Eide
Senior Engineer - Kongsberg Digital
eivind.roson.eide@kdi.kongsberg.com
LinkedIn: /eivind-roson-eide | Medium: @eivind.roson.eide

Eivind is a Senior Engineer at Kongsberg Digital. Based in Oslo, Norway, he has solid experience in Applied Machine Learning, Hybrid ML and software development for industrial applications. Today he is leading a team developing mathematical modeling powered SaaS products for the Energy Industry. Eivind holds a Master in Information Engineering from the University of Cambridge.

ABOUT KONGSBERG DIGITAL

Kongsberg Digital is a provider of next-generation software and digital solutions to customers within maritime, oil and gas, and renewables and utilities. The company consists of more than 500 software experts with leading competence within the internet of things, smart data, artificial intelligence, maritime simulation, automation and autonomous operations.

Find out more: www.kongsbergdigital.com

KONGSBERG

Kongsberg Digital AS
Drengsrudbekken 12, Asker - 1383 - Norway
www.kongsbergdigital.com