# MOVING YOUR MACHINE LEARNING EFFORTS INTO THE MAINSTREAM WITH THE CONVERGENCE OF BUSINESS INTELLIGENCE AND MACHINE LEARNING

Deploying and Obtaining the Benefits of Machine Learning Solutions is Easier with Recently Released Vendor Integration Capabilities.

The convergence of machine learning (ML) and business intelligence (BI) is upon us according to Gartner. BI tool vendors, particularly the three dominant cloud vendors, AWS, Azure, and Google, are enabling interoperability between their BI platforms and their ML platforms. This includes both rendering ML results such as scoring in their BI platforms as well as enabling basic ML capabilities within their BI platforms.

For example, within PowerBI there are now ML capabilities to develop and train models for classification and regression. Amazon AWS and Google have similar capabilities.
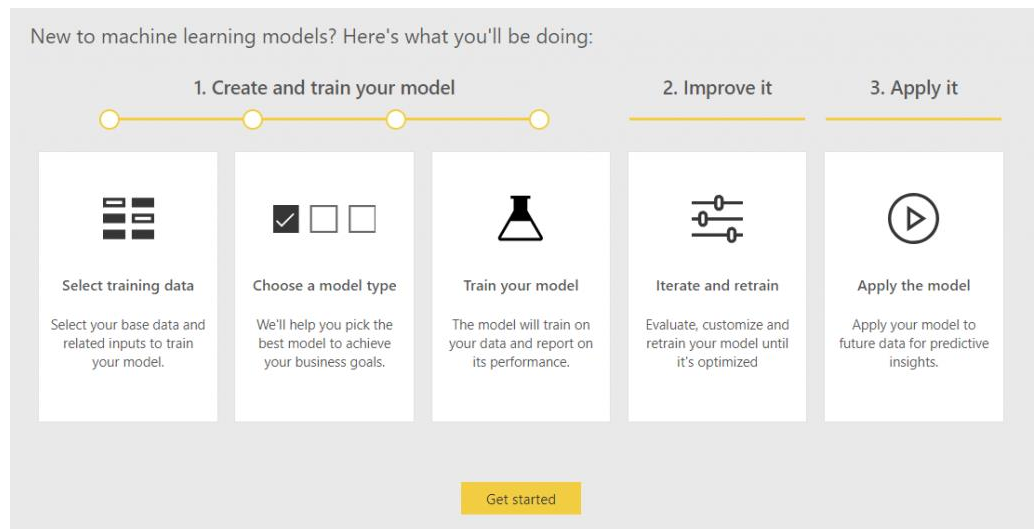
LACK OF BI AND ML CONVERGENCE IS AN IMPEDIMENT TO THE DEPLOYMENT OF ML CAPABILITIES

Historically BI and ML have been parallel and complementary capabilities, BI reported what happened, whereas ML predicted what would happen. However, in the past complementary BI and ML platforms did not interoperate, or if they did it required a significant software engineering development on the part of the customer to enable the interoperability.

BI systems have been in place in most organizations for over ten years and are now mature capabilities which are well understood and established within the organization. These BI systems excel at providing visualizations and reports to users to inform them of what happened historically. While BI systems provide a significant benefit to the organization, a related and arguably more important need is to predict what will happen, which is a capability that ML excels at.

Large digital companies such as Google, Amazon, Facebook, Uber, and others were the leading-edge innovators and adopters of leveraging ML as a competitive differentiator along with delivering significant top and bottom line financial results. And in fact, their enabling innovations for big data infrastructure have been released into open source. For example, Facebook developed and then released Presto into open source.

Other industries and firms took note of their success and looked to replicate the benefits of ML in their firm. A typical scenario is that the firm established an analytics functional area, several data scientists were hired, and a number of ML "proof of concept" projects were proposed, worked on, and then delivered. In many cases these "proof of concept" ML projects were successful in demonstrating the value of ML to the firm's use-cases, but the majority of "proof of concept" projects never become on-going production ML deliverables for the firm, due to many factors including organizational readiness (note: it is estimated that over 70% of data science projects fail to become on-going production deliverables).

One of the technical reasons for failure in moving the ML "proof of concept" into an operational production setting was due to interfacing with the organization's reporting and other production systems. Typically, the data science personnel created their ML models using python or R. While these languages and their respective environments, such as Anaconda or RStudio, work well for ML ideation and iterative development, establishing an interface from python or R to the organization's existing production BI and operational systems requires significant software engineering time, resources, and experience. Large digital organizations such as Google have a deep bench of big data engineering personnel who are able to take what the ML capabilities that data science personnel have developed and move it into a productionized state. However, for many firms, they lack the data and software engineering resources to enable this capability, while at the same time most data science personnel lack production oriented software engineering expertise.

For example, for a healthcare payer the data science team may create an ML model to predict the likelihood of fraud for new claims. During the "proof of concept" phase the scoring for fraud of new claims is presented using the reporting and visualization capabilities in python or R. However, python and R are not reporting and visualization tools that are used by the broader business. Instead, the organization as a whole typically uses a BI platform such as Tableau or PowerBI and will want the scoring for fraudulent claims to be


rendered in their BI platform. To enable the scoring and rendering of new claims in the BI platform requires integration capabilities developed by software engineering teams to interface the ML scoring engine which is in python or R with the BI platform which would be Tableau or PowerBI. This engineering effort could easily take months and multiple software engineering personnel, assuming such personnel were available. If real-time interface for scoring is required, then this will require the development of an API type interface. For many non-digital firms who do not have prior experience with this use case and integration requirement, the ML "proof of concept" will likely never reach a production state.

CALL TO ACTION

The benefits of ML solutions are now attainable for more firms. Now that vendors are converging BI and ML capabilities, the time, cost, and expertise to move from a ML "proof of concept" to rendering results in a BI platform are straight-forward and are a configuration activity versus a software engineering development activity.
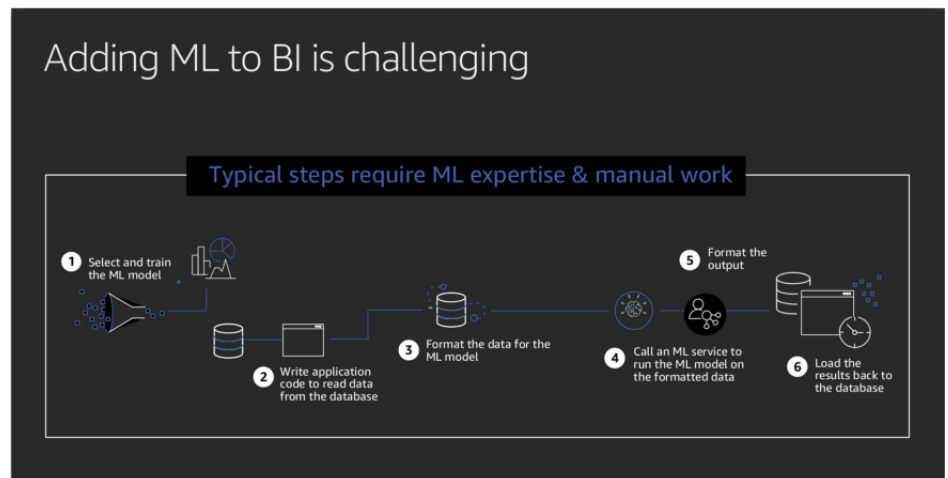
For example, AWS has integrated their ML AWS SageMaker platform with their BI QuickSight platform. Likewise, Microsoft has integrated Azure ML Studio with PowerBI.
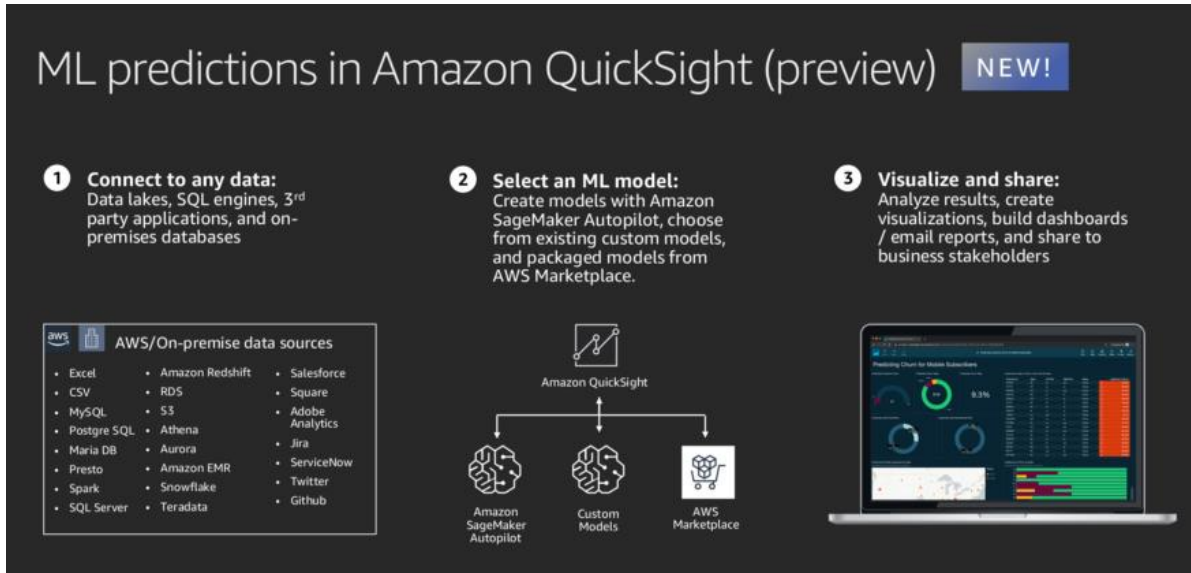
Enabling ML prediction models for scoring into the BI platform is now a configuration activity versus a software engineering development activity, and therefore does not require skilled data engineering personnel and can now be completed in a manner of days versus months.

Irrespective of whether or not you've already had a ML "proof of concept", now is the perfect time to jump in as it is easier than ever to operationalize ML given the convergence of BI and ML capabilities, and as a result, you can provide your firm with the competitive and financial benefits that ML can deliver.

Resources:

https://powerbiblogscdn.azureedge.net/wp-content/uploads/2019/03/AutoML-model-intro-1024x526.png
https://d1.awsstatic.com/events/reinvent/2019/Integrating_Amazon_SageMaker_and_Amazon_QuickSight_ANT341.pdf

This TIP was written by Trexin's Analytics Capability Lead, Michael Schillmoeller. Michael welcomes comments and discussion on this topic and can be reached at michael.schillmoeller@trexin.com.