

[Subscribe to email updates](#)[MACHINE LEARNING](#)

October 7, 2019

Why You Should Use Your Database for Machine Learning

[Sherry Tiao](#)

When it comes to taking a trained **data science model** and putting it into operational use at the company, **TDWI** found that **it took 14 percent of respondents over 9 months** for just that step alone.

But for more than 20 percent, that same step took only a few days to a few weeks.



How long does it take to put a defined model into operational use?



Figure 6. Length of time to put a model into production. Based on 141 respondents who stated they are doing this today.

There are several actions you can take to ensure that you'll be part of the fast 20 percent, not the slower 14 percent. One of them, of course, involves using your database for machine learning.

What Machine Learning Can Do

You've probably heard about some applications of machine learning in the news, like **computers creating art and music through machine learning**.

But what really excites people in the business world is machine learning's ability to use data to find patterns and trends.

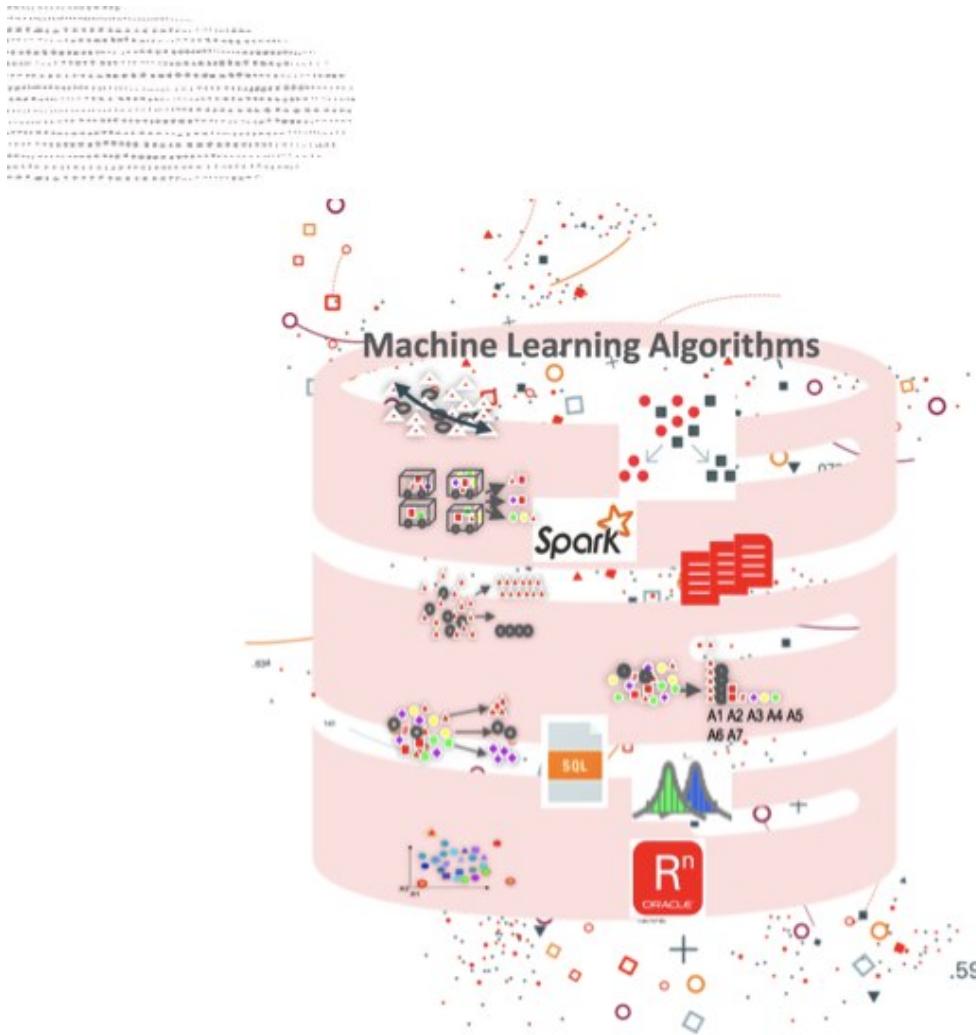
Machine learning is uniquely suited for this because it involves taking massive amounts of data and then using **computers with algorithms**. Computers are enabled to learn how to explore the data to find hidden information.

This allows businesses to do things like:

- Improve retail sales with a recommendation algorithm
- Anticipate equipment breakdown with a prediction algorithm
- Detect fraud through anomaly detection

Machine learning can do this in ways that traditional analytics just can't.

Machine Learning in Databases



You might not know that some databases come with machine learning inside them.

What this means is that you don't have to go out and acquire a data science platform, and you don't have to learn how to use Hadoop, and you don't have to learn how to utilize **data lakes** when you're just starting out.

Actually, you have everything you already need to get started if you're [using a database that comes with machine learning actually inside it.](#)

Now, this isn't how every database provider does things, and it's not how every machine learning platform does things. But it makes things easier for you in so many ways. Here are a few examples:

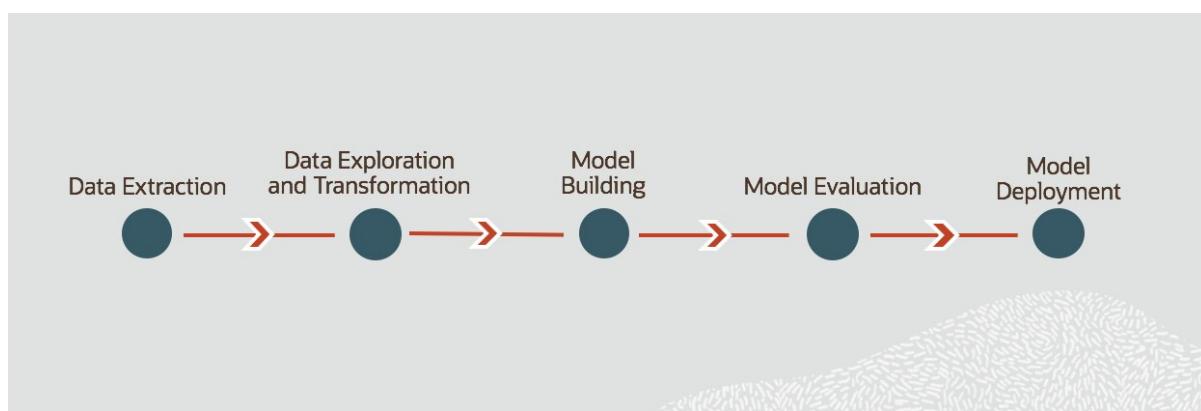
- Do more than ever with your existing data, while maintaining control of data by using your database as a single source of truth and eliminating movement of data out of your database.
- Experiment with machine learning products that come with the database and are optimized to start running. With the right products, you can even perform experimentation without needing data scientist skills, although they do make it easier.
- Use machine learning in the database to develop solutions for everything from fraud detection to predicting customer behavior to identifying selling opportunities.
- Make it easy for teams to use the machine learning models you've created, and operationalize your machine learning more easily than ever.

So, what is it about machine learning in the database that makes this possible, and what are the benefits?

Before I explain the benefits of using existing data in your database, I want to explain the data science process just in case you need a refresher. It will make some of the benefits of machine learning in the database clearer.

The Data Science Process for Machine Learning

Here's a simplified data science process for developing machine learning models. It starts with identifying the data needed for the model. If you're using a separate platform for machine learning, then that data has to be extracted from the source and loaded into that platform.



When you're working in the database, the assumption is that the data is already there so you don't need to extract it, which avoids an often time-

consuming step.

The middle three steps are the core work of understanding and preparing the data, building the model, then testing and evaluating the model. These steps may be repeated multiple times.

When a satisfactory result is achieved, then the model needs to be deployed so it can be used with new data and used by people and applications as needed. This can be one of the most challenging steps when working with an open-source machine-learning platform. What hardware is it going to run on? How will it access data? Does the code need to be converted into a different language? Does it need to be accessed from an API?

But when you build your machine learning in the database, you also run it in the database. There's no need to convert code. Just call the model from a SQL statement. Additionally, we provide a way to expose that model as a REST API if needed.

In this process, you can see that you can significantly simplify 40 percent of the activities needed to build and deploy the model. That's where you can really start to maximize the value of your existing data.

Okay, so now that we have that out of the way, let's talk about the benefits of existing data to you.

Three Benefits of Machine Learning in the Database

Database Machine Learning Benefit #1: You Get Simplicity

You know your data. For data scientists or anyone else, working with **data in the database versus data in the data lake** is like being a kid in a candy shop. The data is clean, it's managed, and you can often just jump ahead and apply analytical techniques.

If you're using a database with machine learning that your company is familiar with, then it means that you already have people who work with it and know it. Instead of hiring five people who are each an expert in one of five software platforms that you think you need in your machine learning workflow, just hire one or two who are well-versed in your existing ecosystem. You can try using in-house talent too. And you can ensure that

this way, everyone is on the same page because they're on the same platform.

And don't forget that your database provider has likely optimized its database to work best with its machine learning offering.

We talk a lot about making machine learning easier, but it's still hard—so believe me, you want anything that can simplify your process.

Database Machine Learning Benefit #2: You Get Time

At Oracle, we approach machine learning differently from other companies. About 20 years ago, we saw that moving the ever-growing volumes of data to where you could run your algorithms was going to get more and more difficult.

It can often take hours or days just to move the data to another platform where the algorithms reside. And doing that introduces all this complexity, like potential data loss. But this is actually still standard for most other companies.

Here at Oracle, we said it makes much more sense to move the algorithms over to the database where they can use the power of the database to run very quickly and where you can easily use the database to get access to data in other databases, a native feature of Oracle Database.

So, pairing machine learning with the database just makes a lot of sense. It's so much faster. You save time, and you save a lot of effort.

Database Machine Learning Benefit #3: You Get Results

The fact is that although it may take a lot of time and effort to build a machine learning model, train it, gain results, and analyze the results, machine learning doesn't usually matter to businesses until the model is *deployed into production*; until ordinary people at your company can make use of it.

For example, you can build a beautiful lead scoring model for your marketing team and express happiness about the superiority of the leads it categorizes. But until that lead scoring model is integrated with your marketing team's systems, until they are actually using it as part of their workflow, that machine learning project isn't valuable.

Now, this step of operationalization and deployment into production is what we discussed at the very beginning of this article, with that survey by TDWI. Remember, it took 20 percent of respondents under a month. But it took 14 percent over 9 months.

Operationalization can be simple or it can be very complex, because deployment and integration into applications like business intelligence (BI) dashboards, call centers, ATMs, websites, and mobile devices can be a very big challenge for IT.

Errors can be introduced at multiple stages. It's a very complicated process and it can be a very time-consuming and expensive model deployment phase.

What you want is an operationalized model with business benefits that you can show to your executive team, and to be able to tell them that you have actually improved the business, through machine learning, with results that anyone can point to as proof of something to be proud of.

If your model has been in the database all along, you don't have to do as much complicated deployment work. That makes the process just that much easier, with results that are gained that much faster.

Conclusion

So what does machine learning with existing data in the database mean for your business? It means that with the right [database and machine learning provider](#), you can minimize the number of steps you need to take for more efficient, faster, easier-to-operationalize machine learning.

Here's what you get:

- More simplicity for you and your employees, since you're starting with tools and data you're familiar with
- More time, with algorithms in the database that ensure minimized data movement and more speed, which saves time and costs
- More results faster, with models in the database that are easier to deploy and operationalize

All of this equals faster and easier time to impact, which is what everyone wants when it comes to a project like this.

To learn how you can benefit from Oracle Big Data, visit [Oracle.com/Big-Data](#), and don't forget to [subscribe to the Oracle Big Data blog](#) and get the latest posts sent to your inbox.

Written by Sherry Tiao and Wes Prichard

Join the discussion

Comments (1)

Recent Content

MACHINE LEARNING

Announcing Python Virtual Environment Support for Spark Applications

More Enterprises are building data lakes in the cloud to unlock the cloud's benefits of scale and agility. When it comes to Big Data, Spark..

MACHINE LEARNING

Disruptive Effects of Cloud Native Machine Learning Systems and Tools

In 1970, E. F. Codd proposed the relational database. Fast forward today, and completely different cloud native architectures in application...

Oracle Content Marketing Login