figure eight
an appen company

LAUNCHING AI:

HOW TO GET YOUR AI INITIATIVE OFF THE GROUND

Introduction

In 2017, research from MIT Sloan Management Review and Boston Consulting Group uncovered some surprising artificial intelligence-related trends. According to the publication's survey, 85 percent of surveyed executives believed, at the time, that "AI [would] allow their companies to obtain or sustain a competitive advantage." Despite that assertion, "Only one in 20 companies [had] extensively incorporated AI in offerings or processes," and "Less than 39% of all companies [had] an AI strategy in place." In the spring of 2018, Adobe released its annual "Digital Trends" report. Just 15 percent of enterprises were using AI at the time. Today, we are a year removed from Adobe's follow-up projection: 31 percent of enterprise executives said AI was on their agenda for the next 12 months.

As we move into 2019, two things are clear. While the desire to leverage AI is near-ubiquitous, the act of putting AI into practice lags behind those aspirations. Companies bridging the gap between their AI dreams and realities might be reminded of the growth of modern cloud computing. In 2008—the year Google released Google App Engine and two years after Amazon released its EC2 public cloud–research and analyst firm Gartner attempted to project IT cloud spending.

Then, Gartner predicted that Software-as-a-service (SaaS) would account for at "at least one-third" of business application spending over the following four years. The research firm also felt IT infrastructure would continue along a similar path: Companies would purchase 40 percent of their infrastructure as as-a-service solutions. Since 2008, the size of the public cloud computing market has ballooned nearly 2,700 percent.

Just as leaders viewed cloud computing as a competitive advantage a decade ago, today they understand the importance of AI in ushering in a new era of productivity and efficiency. And, just like with the advent of the as-a-service economy, there is no one-size-fits-all approach to finding the right machine learning (ML) model for your business purposes. Instead, we'll illustrate an AI decision tree line of business (LOB) owners can consult as they work to integrate AI into their existing or forthcoming business processes.

31 percent
of enterprise executives said AI was on their agenda for the next 12 months

Determine Your Business Case

An organization's approach to AI should be no different than its approach to any software purchase. When making a software purchase decision, organizational leadership does not simply dip into its pockets and buy the latest software platform without a plan in place. Leaders investing in AI should follow a similar tack.

Before selecting an algorithm or model to use, first determine your business use case.

Typically, AI is intended to replace or improve a current business process. Yet, before revving the AI engine it is pertinent to ask first if AI will even help solve the problem in question. If we use the example of a waste management company, we can better understand what these decisions might look like in practice.

In this example, consider a waste management company that is trying to make more efficient its sorting process. The company wants to integrate an AI-enabled camera system to help sort waste into different categories. First, the line of business owner must ask, "Is this sorting process a human-only process, can AI make it more efficient?" In this particular case, it becomes clear that an AI-enabled process can likely perform the sorting task much more efficiently than can humans alone.

Identify the Requisite Personnel

If AI can help make more efficient or replace the process in question, then it is important to understand what infrastructure is in place. Before even considering what technical tools are available, LOB owners must consider first if the right people are in place. It is unlikely companies experimenting with public clouds in 2008 did so without an expert on staff. It is just as crucial now to have a technical product manager (TPM) or someone similar on hand to lead any AI initiative. Organizations that successfully integrate AI will also, if necessary, recruit the services of an outside AI consultant to advise the TPM and business owner in their journey.

Build the Process Pipeline

With a clear objective for AI and the right leadership in place, it's time to begin laying out the AI roadmap. The process pipeline will help you understand how to turn your input into the output you desire. There are typically a number of moving parts that comprise a process pipeline.

In the case of the waste management company, there may be a camera sensor that deals with noisy data via an image classification model. The company will also likely need a recommendation model to tell the sorting system how to act on a given piece of waste and a text-to-speech model when the system requires voice interaction. The pipeline will also involve data acquisition and cleanup components. Break down the projects.

Develop the ML Model(s)

Once you outline a process pipeline, you can begin to build the ML model or models that will help you achieve your desired output. Organizations typically select between making their models in-house, paying a provider for an existing model, or outsourcing a model. Each of these models come with their own pros and cons, depending on the project and resources available to a company. Before choosing a purchase option, first determine if you need a pre-trained or totally customized model.

First determine
if you need a pre-trained
or totally customized model.

A pre-trained model typically serves a very specific purpose. These models do not allow for customization and have already been fine tuned with a data set over which you have no control. Models that are not pre-trained offer more flexibility but also require more leg work to build and adjust. The type of model you use will depend on the use case in question. If the process you're using AI to improve is not core to the business, then a pre-trained, already-built model is probably the way to go. The waste management company will likely rely on pre-trained models for cleaning up a CRM-related process, for instance.

If the business case in question is core to the business, then building a model in-house or selecting a customizable model likely offers longer-term success.

Paying for a vendor-produced model. There is a growing market for ML model providers. AWS offers an extensive market with models for a variety of purposes. Google, Microsoft, and IBM Watson are just a few of the big-name enterprises that offer machine learning-as-a-service solutions for sale.

+ **Pros:** The models provided by vendors are generally cheaper than building a model in-house. The APIs that these vendors offer typically work with a number of use cases. Finally, purchasing a model typically allows an organization to get AI up and running faster.

+ **Cons:** Despite the powerful APIs these models contain, they're very generic and offer limited, specific functionality. If your use case does not align well with what one of these models offers, then it will be of little help.

Building a model in-house. Working on a process that's core to your business? Then building a model your organization owns and can tweak when necessary is probably the best scenario.

+ **Pros:** Building a proprietary model offers an organization control over the IP. This route will also yield a model that fits very specifically the use case in question. Finally, if more models are required in the future, building an initial test case in house can offer a strong foundation from which to work when developing subsequent models.

+ **Cons:** Building a model in-house is likely the most expensive—at least at the outset—option of the three. It requires an appropriate team in place, as well as the raw resources necessary to build from scratch, train, test, and maintain in production.

Outsourcing model creation. This is the option that perhaps best suits an organization that needs a more tailored model than what providers offer but doesn't have the resources to build one from scratch.

+ **Pros:** Outsourcing a model allows for a more customized model without the initial overhead required of building a model from scratch. Outsourcing is a more flexible option than using pre-built algorithms and allows for more agile long-term planning.

+ **Cons:** There is a higher risk associated with leveraging a third party. Paying a third party to create a model can end up being more expensive than anticipated if said party doesn't produce a model adequate for your purposes.

Generate High-Quality Data

No matter which option your organization chooses, you'll have to

feed high-quality data into the ML model in order to obtain the results you want.

Data annotation, people resources, technology, contributor cost, proof of concepts, and more considerations all feed into the acquisition of high-quality data. There are a number of different approaches organizations can take to ensure the data they send through their model(s) of choice is of the appropriate quality.

Annotation. To create high-quality data for the waste management model, organizations will need to have human annotators create data labels. In this case, the annotators will perform tasks, such as drawing bounding boxes around and labeling pieces of waste in thousands of images. The annotation process is especially crucial for ensuring data quality and accuracy. To improve quality and accuracy, businesses should ensure the tools they use to annotate data apply human intelligence to the process adequately.

Data augmentation. Using the contents of a Google image search for "garbage" might seem appropriate for a waste management model. But, chances are, the images don't look exactly like what the cameras will see as they attempt to automate the sorting practice. In cases like this, it can be useful to augment data, such as by adjusting the photos so that they look more like the actual waste passing through the system. Data augmentation helps ensure organizations are feeding their model data that is as similar as possible to the data they'll use in the real world.

Transfer learning. If the company does not have the resources to build a model in-house, they might leverage an open-source image identification model. This model won't be trained on images of waste. However, rather than build a model from scratch, the company can strip down the existing model to a more essential component. In doing so, it can then feed the images of waste into the model, while the model transfers what it has already learned about identifying images to the waste-sorting problem.

Conclusion

No matter how cutting edge a model happens to be, it's useless without enough high-quality training data. Leveraging processes that help create quality data is essential to building an effective AI initiative.

Figure Eight combines the best of human and machine intelligence to provide high-quality annotated training data that powers the world's most innovative machine learning (ML) and business solutions. With over a decade of experience and over 10 billion judgements, Figure Eight's enterprise-ready data annotation platform enables ML and data-driven business solutions to scale across a diverse set of industries including retail, automotive, finance, manufacturing, agriculture, life sciences, robotics, and more. The Figure Eight platform transforms audio, video, text and images into high-quality annotated data to support a variety of use cases ranging from computer vision and search relevance to data categorization and natural language processing (NLP).

Learn more at figure-eight.com

Authored through the combined efforts of the Figure Eight Machine Learning Team:

Monchu Chen, Director of HCI, Machine Learning

Phoebe Liu, Machine Learning Scientist

Rahul Parundekar, Sr. Machine Learning Engineer

Joan Xiao, Lead Machine Learning Scientist

Humayun Irshad, Lead Machine Learning Scientist, Computer Vision

Qazaleh Mirsharif, Machine Learning Scientist, Computer Vision

Kiran Vajapey, Sr. HCI Developer