



What is Low Shot Learning?

Machine learning (ML) is a powerful set of mathematical algorithms that are used to predict outcomes and to classify objects. Low Shot Learning Occurs when the algorithms significantly reduce the size of the training data set.

Most machine learning setups require large training data sets, ranging from thousands to millions of examples. Here are some data points on the size of training sets used in other ML systems to extract data from documents. (See our recent [blog post](#)).

ML results for document classification and extracting unstructured data from documents

[Dropbox](#): 87% accuracy after training on more than 300,000 invoices

[Cloudscan](#): 84% accuracy after training on more than 326,000 invoices

[Jatana.ai](#): 85% accuracy after training on more than 14,000 documents

[GLYNT](#): 98% accuracy after training on 2 - 7 invoices

Each document in the training set is “marked” or “labeled” with the correct answer, and setting up the training data set for even 100’s of examples can become prohibitively costly.

In contrast, GLYNT’s Low Shot machine learning system reliably extracts unstructured data from documents using just 2 - 7 example documents.

Why is Low Shot Learning so important?

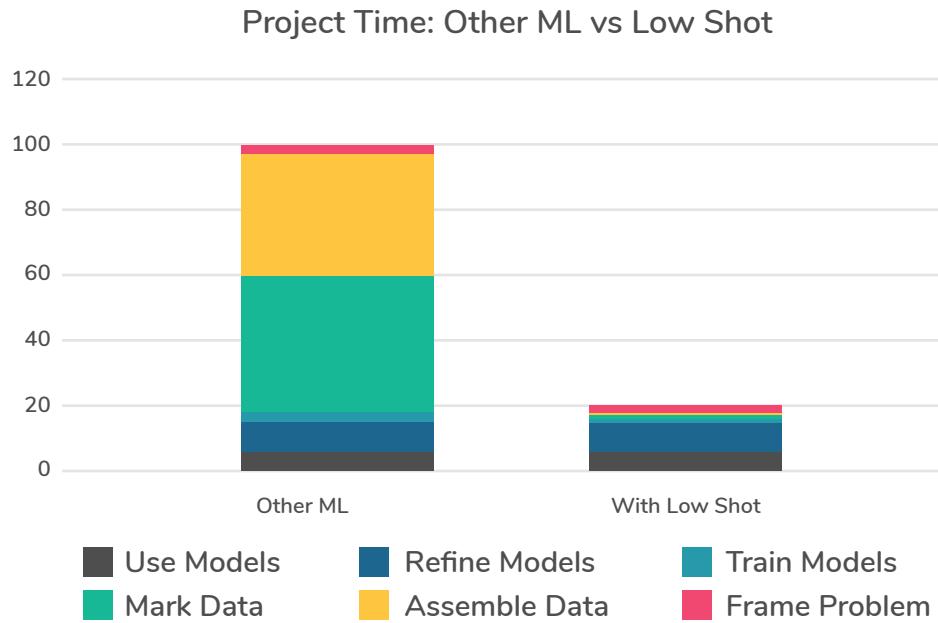
A typical machine learning project have six phases:

1. Frame the Problem
2. Assemble the Data and change it into a form suitable for machine learning
3. Mark the Correct Answer on Each Element of a Large Training Data Set
4. Train the Models
5. Refine the Models
6. Use the “production” Models on New Data

Steps 2 - 5 take up 90% of the time and effort.¹ Low Shot speeds through data marking and training because just a few documents are needed. So instead of months and months of time to first results, Low Shot enables answers in under an hour.

An important benefit of Low Shot Learning is risk reduction. When deploying machine learning , one never knows if their solution will work. Use of Low Shot Learning removes the risk of spending valuable time setting up large data sets only to get mediocre results.

¹Source: <https://aiworld.com/2017>, Keynote speech by David Kiron



Who else is doing Low Shot?

GLYNT has incorporated Low Shot Learning for a specific application, unstructured data extraction from complex documents, and for this application, only GLYNT has this feature built in. In other applications, Low Shot Learning is used by [Google's Deep Mind](#) for classification, [Salesforce](#) for its voice applications, and [Facebook](#) for image labels.

It has also been used by academics, including [MIT](#) researchers for drug discovery, [Stanford](#) researchers for image classification, [Oxford](#) researchers for image classification, and Ireland's [Center for Insight](#) for natural language processing. We would not be surprised if the results from academic papers are used internally by Google, Facebook and Amazon to improve product performance without the user even realizing the technology feat that is working behind the scenes.

Low Shot Learning (aka Few Shot Learning), One Shot Learning and Zero Shot Learning are active research areas on the frontiers of machine learning. As a point of reference, [a child can recognize](#) 10,000 - 30,000 objects by age six. Machine learning has not yet caught up, and Low Shot Learning is part of effort.

How do users of GLYNT benefit from Low Shot Learning?

GLYNT users benefit from Low Shot Learning without lifting a finger. GLYNT incorporates Low Shot Learning within its machine learning algorithms, and is delivered to users in three key product features:

- Super-fast setup. Use our pre-set field lists or use your own. The field list determines which data is extracted from the documents.
- Super-fast training. Push the button and get results in minutes.
- Iterate and Adjust Quickly. Suppose you set up your documents and extracted data. Then a month later a new field appears on the document. Just take a few minutes to add it to the setup and re-train the model.

In short, Low Shot Learning delivers powerful machine learning with very little investment of your time. Also, you don't have to be a software engineer to use our system. GLYNT is set up so that junior staff can handle the tasks.

Can GLYNT's Low Shot Learning be used for other applications?

Let's break this question apart into the components that are required for Low Shot Learning:

Required Component	GLYNT?
Elastic AI Workbench	GLYNT's self-contained system, orchestrating data, documents, complex models and flows, is ready for additional machine learning applications
Application-Specific Modeling	For some applications, GLYNT's Low Shot Learning Models are directly applicable or could be easily adapted. For other applications, a new machine learning library would need to be built
Low Shot Learning Model Library	One abstract representation of Low Shot Learning is a two-layer schema: A foundational model that mathematically captures the key aspects of the data in the problem space, and a second set of machine learning models that predicts and classifies. The tight coordination two layers gives enables Low Shot Learning. Some applications will require the entire two-layer model to be rebuilt for the problem at hand. Others can re-use the GLYNT setup. In either case, GLYNT's Workbench can speed the process, as it is the right engineering infrastructure in a self-contained, easy to deploy system.

Looking ahead 5 years, what will the role of low shot be in the world of AI and machine learning?

GLYNT is first with Low Shot Learning for the extraction of unstructured data from complex documents, and has a platform that is extensible to other applications of Low Shot Learning. With such large benefits to Low Shot Learning, it is not surprising that it is an area of active research. We see three big changes in the next five years.

First, we're watching how [Low Shot develops for the field of Natural Language Processing \(NLP\)](#).

NLP is a high value application of artificial intelligence for voice – for example powering Amazon Alexa and Google Assistant – and is widely used in healthcare to extract summaries and context from medical records. NLP is also used to determine the sentiment of writing on the web and in social media. Low Shot Learning will speed the field of NLP, but is likely to require re-thinking many NLP libraries.

Second, Low Shot Learning also increases the AI talent wars. The form of Low Show Learning in GLYNT is roughly a two-layer model. The first layer consists of a mathematically based model of the domain. The second layer learns a particular skill, prediction, classification, and so on, based on the first model. This two-layer schema increases the demand for skilled software developers who can invent, specify and build the first layer, the foundational mathematical representation. One example of this first layer is from [Google](#), where a research team built a 3D image model from 2D images. This serves as the first layer and a second layer of models enabled Zero Shot Learning.

To advance Low Shot Learning, every team will need at least one person who can build such a foundational model. The perfect hire will have domain experience, graduate-level math and computer science skills, fearlessness and persistence. It's a pioneering effort. Demand for this talent will increase, even beyond today's levels, because with just a few of these hires products can be built that create enormous and easily measured value: time and money. Low Shot Learning leverages the mathy talent to remove significant costs and bottlenecks from elsewhere in the AI project flow.. But it needs those pioneers!

Third, Low Shot Learning works on small databases. Until now these have been excluded from the AI conversation. Now the data can be efficiently extracted, such as through GLYNT, and used in other Low Shot algorithms. Smaller data sets are now more valuable.

Current AI thinking is that AI requires very large data sets for training, lots of computing resources, and lots of cash to spend: AI projects take a long time and cost a lot of money. The diffusion of Low Shot Learning changes that. Smaller companies can compete on an even footing because so little data is needed. We think the diffusion of Low Shot Learning will lead to the faster adoption of AI, better encapsulation of AI into useful software products and services and a business case for AI with small data sets.

If software eats the world and AI eats software, then Low Shot Learning is an accelerant. GLYNT fuels that acceleration by liberating data, making it ready for use by other AI.

ABOUT US



GLYNT is a machine learning system that produces a stream of clean, labeled data from any document. Get started in minutes. GLYNT was developed by the team at WattzOn, which uses GLYNT in its products for the energy and credit markets. See us at GLYNT.AI and WattzOn.com