Cognitive Services and machine learning

08/22/2019 • 5 minutes to read • 👤👤👤👤👤 +3

In this article

Cognitive Services provides machine learning capabilities to solve general problems such as analyzing text for emotional sentiment or analyzing images to recognize objects or faces. You don't need special machine learning or data science knowledge to use these services.

Cognitive Services is a group of services, each supporting different, generalized prediction capabilities. The services are divided into different categories to help you find the right service.

Service category	Purpose
Decision	Build apps that surface recommendations for informed and efficient decision-making.
Language	Allow your apps to process natural language with pre-built scripts, evaluate sentiment and learn how to recognize what users want.
Search	Add Bing Search APIs to your apps and harness the ability to comb billions of webpages, images, videos, and news with a single API call.

Service category	Purpose
Speech	Convert speech into text and text into natural-sounding speech. Translate from one language to another and enable speaker verification and recognition.
Vision	Recognize, identify, caption, index, and moderate your pictures, videos, and digital ink content.

Use Cognitive Services when you:

- Can use a generalized solution.
- Access solution from a programming REST API or SDK.

Use another machine-learning solution when you:

- Need to choose the algorithm and need to train on very specific data.

What is machine learning?

Machine learning is a concept where you bring together data and an algorithm to solve a specific need. Once the data and algorithm are trained, the output is a model that you can use again with different data. The trained model provides insights based on the new data.

The process of building a machine learning system requires some knowledge of machine learning or data science.

Machine learning is provided using Azure Machine Learning (AML) products and services.

What is a Cognitive Service?

A Cognitive Service provides part or all of the components in a machine learning solution: data, algorithm, and trained model. These services are meant to require general knowledge about your data without needing experience with machine learning or data science. These services provide both REST API(s) and language-based SDKs. As a result, you need to have programming language knowledge to use the services.

How are Cognitive Services and Azure Machine Learning (AML) similar?

Both have the end-goal of applying artificial intelligence (AI) to enhance business operations, though how each provides this in the respective offerings is different.

Generally, the audiences are different:

- Cognitive Services are for developers without machine-learning experience.
- Azure Machine Learning is tailored for data scientists.

How is a Cognitive Service different from machine learning?

A Cognitive Service provides a trained model for you. This brings data and an algorithm together, available from a REST API(s) or SDK. You can implement this service within minutes, depending on your scenario. A Cognitive Service provides answers to general problems such as key phrases in text or item identification in images.

Machine learning is a process that generally requires a longer period of time to implement successfully. This time is spent on data collection, cleaning, transformation, algorithm selection, model training, and deployment to get to the same level of functionality provided by a Cognitive Service. With machine learning, it is possible to provide answers to highly specialized and/or specific problems. Machine learning problems require familiarity with the specific subject matter and data of the problem under consideration, as well as expertise in data science.

What kind of data do you have?

Cognitive Services, as a group of services, can require none, some, or all custom data for the trained model.

No additional training data required

Services that provide a fully-trained model can be treated as a *opaque box*. You don't need to know how they work or what data was used to train them. You bring your data to a fully trained model to get a prediction.

Some or all training data required

Some services allow you to bring your own data, then train a model. This allows you to extend the model using the Service's data and algorithm with your own data. The output matches your needs. When you bring your own data, you may need to tag the data in a way specific to the service. For example, if you are training a model to identify flowers, you can provide a catalog of flower images along with the location of the flower in each image to train the model.

A service may *allow* you to provide data to enhance its own data. A service may *require* you to provide data.

Real-time or near real-time data required

A service may need real-time or near-real time data to build an effective model. These services process significant amounts of model data.

Service requirements for the data model

The following data categorizes each service by which kind of data it allows or requires.

Cognitive Service	No training data required	You provide some or all training data	Real-time or near real-time data collection
Anomaly Detector	x	x	x
Bing Search	x		
Computer Vision	x		
Content Moderator	x		x
Custom Vision		x	
Face	x	x	
Form Recognizer		x	
Immersive Reader	x		
Ink Recognizer	x	x	

Cognitive Service	No training data required	You provide some or all training data	Real-time or near real-time data collection
Language Understanding (LUIS)		x	
Personalizer	x*	x*	x
QnA Maker		x	
Speaker Recognizer		x	
Speech Text-to-speech (TTS)	x	x	
Speech Speech-to-text (STT)	x	x	
Speech Translation	x		
Text Analytics	x		
Translator	x		
Translator - custom translator		x	

*Personalizer only needs training data collected by the service (as it operates in real-time) to evaluate your policy and data. Personalizer does not need large historical datasets for up-front or batch training.

Where can you use Cognitive Services?

The services are used in any application that can make REST API(s) or SDK calls. Examples of applications include web sites, bots, virtual or mixed reality, desktop and mobile applications.

How is Azure Cognitive Search related to Cognitive Services?

Azure Cognitive Search is a separate cloud search service that optionally uses Cognitive Services to add image and natural language processing to indexing workloads. Cognitive Services is exposed in Azure Cognitive Search through built-in skills that wrap individual APIs. You can use a free resource for walkthroughs, but plan on creating and attaching a billable resource for larger volumes.

How can you use Cognitive Services?

Each service provides information about your data. You can combine services together to chain solutions such as converting speech (audio) to text, translating the text into many languages, then using the translated languages to get answers from a knowledge base. While Cognitive Services can be used to create intelligent solutions on their own, they can also be combined with traditional machine learning projects to supplement models or accelerate the development process.

Cognitive Services that provide exported models for other machine learning tools:

Cognitive Service	Model information
Custom Vision	Export for Tensorflow for Android, CoreML for iOS11, ONNX for Windows ML

Learn more

- Architecture Guide - What are the machine learning products at Microsoft?
- Machine learning - Introduction to deep learning vs. machine learning

Next steps

- Create your Cognitive Service account in the Azure portal or with Azure CLI.
- Learn how to authenticate to a Cognitive Service.
- Use diagnostic logging for issue identification and debugging.
- Deploy a Cognitive Service in a Docker container.
- Keep up to date with service updates.

Is this page helpful?

👍 Yes 👎 No