



[View, add and edit your notes in the app](#)

AzureAI

Generated on December 3, 2024

Summary

AI Notes

AI Slides

Text Notes

Screenshots

❖ 0

❖ 153

✉ 3

📷 230

What is the Microsoft Azure AI Fundamentals (AI-900)?

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



The Azure AI Fundamentals Certification is for those seeking an ML role such as **AI Engineer, or Data Scientist**

The certification will demonstrate a person can define and understand:

- Azure AI Services (e.g., Cognitive Services and Azure Applied AI Services)
- AI Concepts, Knowledge Mining, Responsible AI
- Basics of ML pipelines, Classical ML models, AutoML
- Generative AI workloads and Azure AI Studio

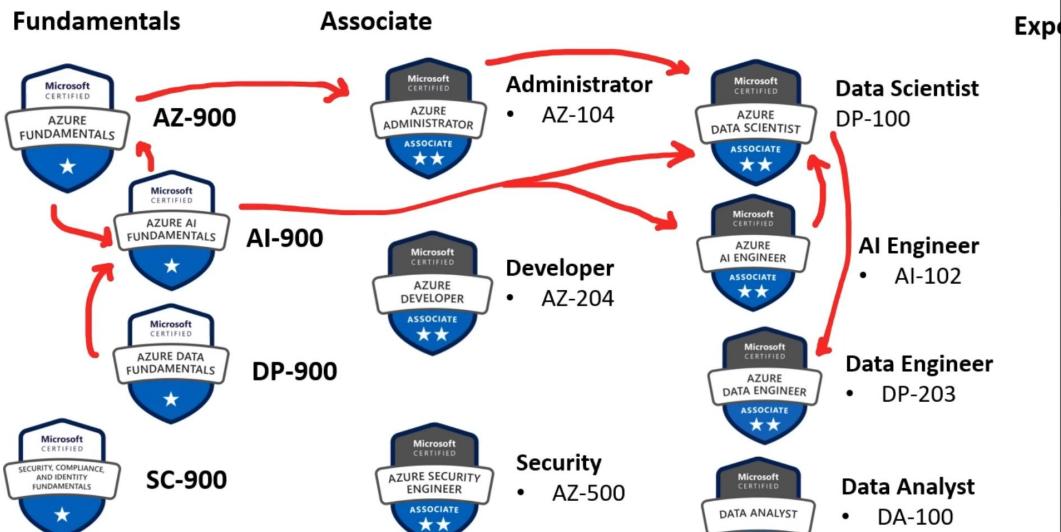
This certification is generally referred to by its course code **AI-900**

The **natural path** for the **Azure AI Engineer** or the **Azure Data Scientist**

This is an **easy course to pass**, and great for those **new to cloud or ML-related technology**

- ❖ The AI-900 certification is an accessible starting point for individuals new to cloud or machine learning technologies, and it serves as a foundational step before pursuing further certifications like the Azure AI engineer or Azure data scientist.

▶ 1:47



- Many individuals pursue the DP-900 certification before the AI-900 to build a solid data foundation, as this knowledge is beneficial for roles like data scientists and AI engineers.

▶ 2:32

How Long to Study to Pass?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Beginner

You are **New to ML and AI**
You are **New to Azure** or any cloud provider

Intermediate

You have passed **AZ-900 or DP-900**
You are familiar with Azure

Experienced

You have **1+ Year Azure experience** or another CSP e.g., AWS, GCP



8 hours (average study time)

- 50% lecture and labs
- 50% practice exams

Recommended to study **30 min to 1 hour** a day for 14 days.

- To effectively prepare for the exam, allocate about 8 hours of study time, dividing it equally between lectures and practice exams, while also engaging in hands-on labs and memorizing key information over a recommended period of 14 days.

▶ 4:30

What does it take to pass the exam?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



1. **Watch video lecture** and memorize key information
2. **Do hands-on labs** and **follow along** within your own account
3. **Do paid online practice exams** that simulate the real exam.

Practice Exam Console ➔

Practice Exam 1 50 questions 90 minutes	Take practice exam
Practice Exam 2 50 questions 90 minutes	Take practice exam
Practice Exam 3 50 questions 90 minutes	Take practice exam
Practice Exam 4 50 questions 90 minutes	Take practice exam
Practice Exam 5 50 questions 90 minutes	Take practice exam
Shuffled Practice Exam 50 questions 90 minutes	Take practice exam

Exam Pro

Signup and Redeem your **FREE** Practice Exam
No credit card required

<https://www.exampro.co/ai-900>

- 💡 It is advisable to take at least one practice exam or review a sample exam for Azure certifications, as they can be challenging and have five distinct domains of questions.

▷ 5:45

Exam Guide – Content Outline

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



The exam of questions in **5 domains**.



Each domain has its own weighting, this determines how many questions in a domain that will show up.

Skills measured

- 15-20% Describe Artificial Intelligence workloads and considerations
 - 6-9 questions
- 20-25% Describe fundamental principles of machine learning on Azure
 - 7-12 questions
- 15-20% Describe features of computer vision workloads on Azure
 - 6-9 questions
- 15-20% Describe features of Natural Language Processing (NLP) workloads on Azure
 - 6-9 questions
- 15-20% Describe features of generative AI workloads on Azure
 - 6-9 questions

- 💡 The exam can be taken in person at test centers like Certiport and Pearson VUE, or online from home, with a focus on describing features of Azure workloads and generative AI.

▷ 6:37



Passing Grade is *700/1000

You need to get “*around*” 70% to pass



Azure uses Scaled Scoring

- 💡 In-person assessments are generally less stressful and more controlled than online ones, with a passing grade of approximately 70% based on scaled scoring from 37 to 47 questions, allowing for about 10 to 13 incorrect answers.

▷ 7:16

Exam Guide – Response Types



There are ~37-47 Questions

You can afford to get *10-13 questions wrong



There is no penalty for wrong questions

- Some questions are worth **more than 1 point**
- Some questions cannot be skipped

Format of Questions

- Multiple Choice
- Multiple Answer
- Drag and Drop
- Hot Area

- 💡 The exam consists of various question formats, has a total duration of 60 minutes with no penalties for incorrect answers, and requires careful time management to allow for reading instructions and reviewing.

▷ 7:43

Exam Guide – Duration

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Duration of 60 mins

You get ~1mins per question

Exam Time is: 60mins

Seat Time is: 90mins

Seat time refer to the amount of time you should allocate for the exam.

It includes:

- Time to review instructions
- Read and accept NDA
- Complete the exam
- Provide feedback at the end.

- 💡 The exam has a total duration of 90 minutes, including time for instructions and feedback, and the certification obtained will remain valid indefinitely as long as the technology remains relevant.

▷ 8:02

Exam Guide – Valid Until

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Valid forever

This certification does not expire

Microsoft Fundamentals certifications such as the AZ-900 or MS-900 do not expire as long as the technology is still available or relevant.

- 💡 Microsoft fundamental certifications, such as the AI-900 and MS-900, do not expire as long as the technology remains available or relevant.

▷ 8:16

[Study guide for Exam AI-900](#)

Filter by title

- Credentials overview
- > Earn a certification
- > Earn an Applied Skills credential
- > Renew a certification
- > Manage Credentials in the Learn profile
- > Training and credentials news
- Offers
- > Program information
- > Credentials help
- > Study guides
 - Study guide for Exam AI-102
 - Study guide for Exam AI-900**
 - Study guide for Exam A2-104
 - Study guide for Exam A2-120
 - Study guide for Exam A2-140
 - Study guide for Exam A2-204
 - Study guide for Exam A2-309
 - Study guide for Exam A2-400
 - Study guide for Exam A2-500
 - Study guide for Exam A2-700
 - Study guide for Exam A2-800
 - Study guide for Exam A2-801
 - Study guide for Exam A2-900
 - Study guide for Exam DP-100
 - Study guide for Exam DP-200
 - Study guide for Exam DP-300
 - Study guide for Exam DP-400
 - Study guide for Exam DP-500
 - Study guide for Exam DP-600
- [Download PDF](#)

Describe fundamental principles of machine learning on Azure (20–25%)

Identify common machine learning techniques

- Identify regression machine learning scenarios
- Identify classification machine learning scenarios
- Identify clustering machine learning scenarios
- Identify features of deep learning techniques

Describe core machine learning concepts

- Identify features and labels in a dataset for machine learning
- Describe how training and validation datasets are used in machine learning

Describe Azure Machine Learning capabilities

- Describe capabilities of Automated machine learning
- Describe data and compute services for data science and machine learning
- Describe model management and deployment capabilities in Azure Machine Learning

Describe features of computer vision workloads on Azure (15–20%)

Identify common types of computer vision solution:

- Identify features of image classification solutions
- Identify features of object detection solutions
- Identify features of optical character recognition solutions
- Identify features of facial detection and facial analysis solutions

Additional resources

d0. Training

Module Fundamentals of Azure AI services - Training
This module is an introduction to how Azure AI services can be used to build applications.

Certification Microsoft Certified: Azure AI Fundamentals - Certifications
You can use Azure AI Fundamentals to validate foundational knowledge of machine learning and artificial intelligence concepts and related Microsoft Azure services.

□ Documentation

Exam AI-900: Microsoft Azure AI Fundamentals - Certifications
Exam AI-900: Microsoft Azure AI Fundamentals

Microsoft Certified: Azure AI Fundamentals - Certifications
You can use Azure AI Fundamentals to validate foundational knowledge of machine learning and artificial intelligence concepts and related Microsoft Azure services.

💡 The content covers essential machine learning principles, including regression, classification, clustering, deep learning features, and the identification of features and labels in datasets.

▶ 10:26

[Study guide for Exam AI-900](#)

Filter by title

- Credentials overview
- > Earn a certification
- > Earn an Applied Skills credential
- > Renew a certification
- > Manage Credentials in the Learn profile
- > Training and credentials news
- Offers
- > Program information
- > Credentials help
- > Study guides
 - Study guide for Exam AI-102
 - Study guide for Exam AI-900**
 - Study guide for Exam A2-104
 - Study guide for Exam A2-120
 - Study guide for Exam A2-140
 - Study guide for Exam A2-204
 - Study guide for Exam A2-309
 - Study guide for Exam A2-400
 - Study guide for Exam A2-500
 - Study guide for Exam A2-700
 - Study guide for Exam A2-800
 - Study guide for Exam A2-801
 - Study guide for Exam A2-900
 - Study guide for Exam DP-100
 - Study guide for Exam DP-200
 - Study guide for Exam DP-300
 - Study guide for Exam DP-400
 - Study guide for Exam DP-500
 - Study guide for Exam DP-600
- [Download PDF](#)

Describe fundamental principles of machine learning on Azure (20–25%)

Identify common machine learning techniques

- Identify regression machine learning scenarios
- Identify classification machine learning scenarios
- Identify clustering machine learning scenarios
- Identify features of deep learning techniques

Describe core machine learning concepts

- Identify features and labels in a dataset for machine learning
- Describe how training and validation datasets are used in machine learning

Describe Azure Machine Learning capabilities

- Describe capabilities of Automated machine learning
- Describe data and compute services for data science and machine learning
- Describe model management and deployment capabilities in Azure Machine Learning

Describe features of computer vision workloads on Azure (15–20%)

Identify common types of computer vision solution:

- Identify features of image classification solutions
- Identify features of object detection solutions
- Identify features of optical character recognition solutions
- Identify features of facial detection and facial analysis solutions

Additional resources

d0. Training

Module Fundamentals of Azure AI services - Training
This module is an introduction to how Azure AI services can be used to build applications.

Certification Microsoft Certified: Azure AI Fundamentals - Certifications
You can use Azure AI Fundamentals to validate foundational knowledge of machine learning and artificial intelligence concepts and related Microsoft Azure services.

□ Documentation

Exam AI-900: Microsoft Azure AI Fundamentals - Certifications
Exam AI-900: Microsoft Azure AI Fundamentals

Microsoft Certified: Azure AI Fundamentals - Certifications
You can use Azure AI Fundamentals to validate foundational knowledge of machine learning and artificial intelligence concepts and related Microsoft Azure services.

▶ 10:26



ML and AI Concepts

Layers of Machine Learning

- The Azure OpenAI service offers features such as natural language generation, code generation, and image generation, while also exploring the layers of machine learning and their interrelationships.

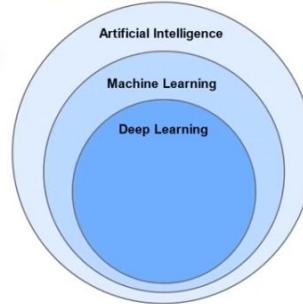
▷ 12:53

The Layers of Machine Learning

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

What is Artificial Intelligence (AI)?

Machines that perform jobs that **mimic human behavior**



- Artificial intelligence involves machines performing tasks that mimic human behavior, while machine learning enables them to improve at these tasks without explicit programming, and deep learning utilizes artificial neural networks for more complex processing.

▷ 13:01

The Layers of Machine Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Artificial Intelligence (AI)?

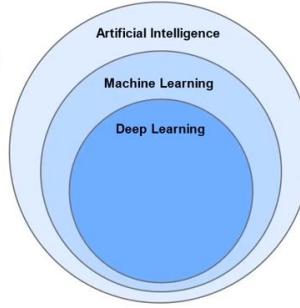
Machines that perform jobs that **mimic human behavior**

What is Machine Learning (ML)?

Machines that get better at a task **without explicit programming**

What is Deep Learning (DL)?

Machines that have an **artificial neural network** inspired by the human brain to solve complex problems.



What is a Data Scientist?

A person with **multi-disciplinary skills** in math, statistics, predictive modeling and machine learning **to make future predictions**

- 💡 A data scientist utilizes multi-disciplinary skills in math, statistics, and predictive modeling to solve complex problems and create AI systems that may incorporate machine learning or deep learning techniques.

▷ 13:48

The Layers of Machine Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Artificial Intelligence (AI)?

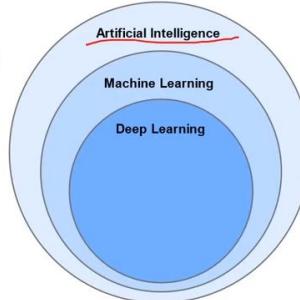
Machines that perform jobs that **mimic human behavior**

What is Machine Learning (ML)?

Machines that get better at a task **without explicit programming**

What is Deep Learning (DL)?

Machines that have an **artificial neural network** inspired by the human brain to solve complex problems.



What is a Data Scientist?

A person with **multi-disciplinary skills** in math, statistics, predictive modeling and machine learning **to make future predictions**

▷ 13:51

The Layers of Machine Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Artificial Intelligence (AI)?

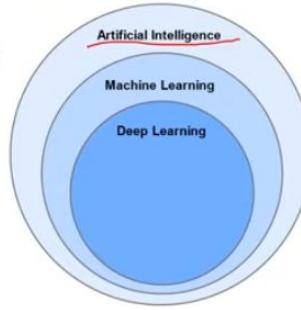
Machines that perform jobs that **mimic human behavior**

What is Machine Learning (ML)?

Machines that get better at a task **without explicit programming**

What is Deep Learning (DL)?

Machines that have an **artificial neural network** inspired by the human brain to solve complex problems.



▶ 13:52

The Layers of Machine Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Artificial Intelligence (AI)?

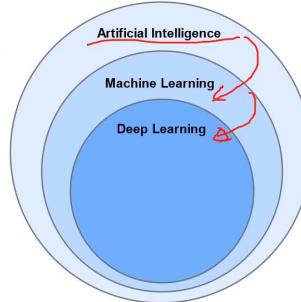
Machines that perform jobs that **mimic human behavior**

What is Machine Learning (ML)?

Machines that get better at a task **without explicit programming**

What is Deep Learning (DL)?

Machines that have an **artificial neural network** inspired by the human brain to solve complex problems.



What is a Data Scientist?

A person with **multi-disciplinary skills** in math, statistics, predictive modeling and machine learning **to make future predictions**

- 💡 AI is software that mimics human behaviors and capabilities, utilizing techniques such as machine learning, deep learning, or rule-based systems.

▶ 13:58



ML and AI Concepts

Key Elements of AI

- AI is software that mimics human behaviors and capabilities, incorporating elements such as machine learning, deep learning, or rule-based systems.

▶ 14:03

Key Elements of AI

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

AI is the software that **imitates human behaviors and capabilities**.

Key elements (*according to Microsoft/Azure*):

- Machine learning — the foundation of an AI system, learn and predict like a human
- Anomaly detection — detect outliers or things out of place like a human
- Computer vision — be able to see like a human
- Natural language processing — be able to process human languages and infer context
- Conversational AI — be able to hold a conversation with a human

- AI comprises several components, including machine learning, anomaly detection, computer vision, natural language processing (NLP), and conversational AI.

▶ 14:12

Key Elements of AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI is the software that **imitates human behaviors and capabilities**.

Key elements (*according to Microsoft/Azure*):

- Machine learning — the foundation of an AI system, learn and predict like a human
- Anomaly detection — detect outliers or things out of place like a human
- Computer vision — be able to see like a human
- Natural language processing — be able to process human languages and infer context
- Conversational AI — be able to hold a conversation with a human

▶ 14:38

Key Elements of AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI is the software that **imitates human behaviors and capabilities**.

Key elements (*according to Microsoft/Azure*):

- Machine learning — the foundation of an AI system, learn and predict like a human
- Anomaly detection — detect outliers or things out of place like a human
- Computer vision — be able to see like a human
- Natural language processing — be able to process human languages and infer context
- Conversational AI — be able to hold a conversation with a human

▶ 14:41

KEY ELEMENTS CAN BE ASKED IN EXAM AS QUESTION

▶ 14:56

Key Elements of AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI is the software that **imitates human behaviors and capabilities**.

Key elements (*according to Microsoft/Azure*):

- Machine learning — the foundation of an AI system, learn and predict like a human
- Anomaly detection — detect outliers or things out of place like a human
- Computer vision — be able to see like a human
- Natural language processing — be able to process human languages and infer context
- Conversational AI — be able to hold a conversation with a human

💡 A data set is a logical grouping of closely related units of data that share the same structure.

▷ 14:57

Key Elements of AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI is the software that **imitates human behaviors and capabilities**.

Key elements (*according to Microsoft/Azure*):

- Machine learning — the foundation of an AI system, learn and predict like a human
- Anomaly detection — detect outliers or things out of place like a human
- Computer vision — be able to see like a human
- Natural language processing — be able to process human languages and infer context
- Conversational AI — be able to hold a conversation with a human

💡 A data set is a logical grouping of closely related units of data that share the same structure.

▷ 14:57

Data Sets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a dataset?

A data set is a **logical grouping of units of data** that are closely related and/or share the same data structure.

There are **publicly available data** sets that are used in the **learning of statistics, data analytics, machine learning**

MNIST database

Images of **handwritten digits** used to test classification, clustering, and image processing algorithm.

Commonly used when learning how to build computer vision ML models to translate handwriting into digital text



0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

▶ 15:15

Data Sets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a dataset?

A data set is a **logical grouping of units of data** that are closely related and/or share the same data structure.

There are **publicly available data** sets that are used in the **learning of statistics, data analytics, machine learning**

MNIST database

Images of **handwritten digits** used to test classification, clustering, and image processing algorithm.

Commonly used when learning how to build computer vision ML models to translate handwriting into digital text

- 💡 The MNIST database consists of images of handwritten digits for testing and classifying image processing algorithms, while the COCO dataset contains various common images formatted in JSON.

▶ 15:33

Data Sets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a dataset?

A data set is a **logical grouping of units of data** that are closely related and/or share the same data structure.

There are **publicly available data** sets that are used in the **learning of statistics, data analytics, machine learning**

MNIST database

Images of **handwritten digits** used to test classification, clustering, and image processing algorithm.

Commonly used when learning how to build computer vision ML models to translate handwriting into digital text

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

- Common algorithms for image processing in computer vision include those used to convert handwritten text into digital format and analyze datasets like the COCO dataset, which identifies objects within images.

▶ 15:37

Data Sets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a dataset?

A data set is a **logical grouping of units of data** that are closely related and/or share the same data structure.

There are **publicly available data** sets that are used in the **learning of statistics, data analytics, machine learning**

MNIST database

Images of **handwritten digits** used to test classification, clustering, and image processing algorithm.

Commonly used when learning how to build computer vision ML models to translate handwriting into digital text

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

▶ 15:37

More Data Sets....

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Common Objects In Context (COCO) dataset

A dataset which contains many common images using a JSON file (coco format) that identify objects or segments within an image.

Dataset examples



This dataset features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 329K images (>200K labeled)
- 0.5 million object instances
- 79 object categories
- 90 stuff categories
- 4 captions per image
- 249,000 people with keypoints

- 💡 Coco datasets are important because Azure Machine Learning Studio's data labeling service can export to Coco formats, facilitating the use of open datasets in machine learning pipelines.

▷ 15:57

More Data Sets....

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Common Objects In Context (COCO) dataset

A dataset which contains many common images using a JSON file (coco format) that identify objects or segments within an image.

Dataset examples



This dataset features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 329K images (>200K labeled)
- 0.5 million object instances
- 79 object categories
- 90 stuff categories
- 4 captions per image
- 249,000 people with keypoints

▷ 15:57

More Data Sets....

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Common Objects In Context (COCO) dataset

A dataset which contains many common images using a JSON file (coco format) that identify objects or segments within an image.

Dataset examples



This dataset features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 329K images (>200K labeled)
- 0.5 million object instances
- 79 object categories
- 90 stuff categories
- 4 captions per image
- 249,000 people with keypoints

▶ 16:07

More Data Sets....

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Common Objects In Context (COCO) dataset

A dataset which contains many common images using a JSON file (coco format) that identify objects or segments within an image.

Dataset examples



This dataset features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 329K images (>200K labeled)
- 0.5 million object instances
- 79 object categories
- 90 stuff categories
- 4 captions per image
- 249,000 people with keypoints

▶ 16:10

More Data Sets....

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Common Objects In Context (COCO) dataset

A dataset which contains many common images using a JSON file (coco format) that identify objects or segments within an image.

Dataset examples



This dataset features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 329K images (>200K labeled)
- 0.5 million object instances
- 79 object categories
- 90 stuff categories
- 4 captions per image
- 249,000 people with keypoints

- ❖ Coco formats can be exported, and Azure machine learning pipelines provide access to common open datasets for use in projects.

▷ 16:27

More Data Sets....

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Common Objects In Context (COCO) dataset

A dataset which contains many common images using a JSON file (coco format) that identify objects or segments within an image.

Dataset examples



This dataset features:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 329K images (>200K labeled)
- 0.5 million object instances
- 79 object categories
- 90 stuff categories
- 4 captions per image
- 249,000 people with keypoints

- ❖ Data labeling involves identifying raw data such as images, text files, and videos, and adding meaningful labels to provide context.

▷ 16:36

What is Labeling?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Data Labeling?

the process of **identifying raw data** (images, text files, videos, etc.) **and adding one or more meaningful and informative labels to provide context** so that a machine learning model can learn

Task type [Info](#)

Task selection
Select the task that a human worker will perform to label objects in your dataset.

Image classification
Get workers to categorize images into specific classes. [Info](#)

Bounding box
Get workers to draw bounding boxes around specified objects in your images. [Info](#)





With **supervised machine learning**, labeling is a prerequisite to produce training data and **each piece of data will generally be labeled by a human**

With **unsupervised machine learning**, **labels will be produced by the machine**, and may not be human readable.

What is a ground truth?

A properly labeled dataset that you use as the objective standard to train and assess a given model is often called "ground truth." The accuracy of your trained model will depend on the accuracy of your ground truth

- Machine learning models rely on accurately labeled datasets, known as ground truth, for effective training and assessment, with the potential for machine-assisted labeling producing labels that may not be human-readable.

▷ 17:00

What is Labeling?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Data Labeling?

the process of **identifying raw data** (images, text files, videos, etc.) **and adding one or more meaningful and informative labels to provide context** so that a machine learning model can learn

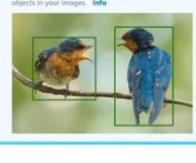
Task type [Info](#)

Task selection
Select the task that a human worker will perform to label objects in your dataset.

Image classification
Get workers to categorize images into specific classes. [Info](#)

Bounding box
Get workers to draw bounding boxes around specified objects in your images. [Info](#)





With **supervised machine learning**, labeling is a prerequisite to produce training data and **each piece of data will generally be labeled by a human**

With **unsupervised machine learning**, **labels will be produced by the machine**, and may not be human readable.

What is a ground truth?

A properly labeled dataset that you use as the objective standard to train and assess a given model is often called "ground truth." The accuracy of your trained model will depend on the accuracy of your ground truth

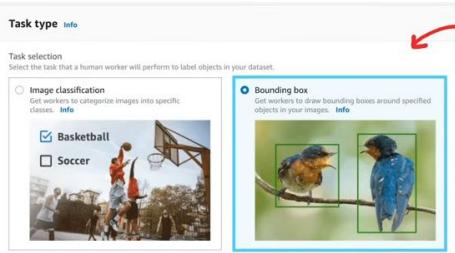
▷ 17:16

What is Labeling?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Data Labeling?

the process of **identifying raw data** (images, text files, videos, etc.) **and adding one or more meaningful and informative labels to provide context** so that a machine learning model can learn



The screenshot shows a user interface for selecting a labeling task. It includes a 'Task type' section with 'Image classification' and 'Bounding box' options, and a preview area showing images of basketball players and birds.

With **supervised machine learning**, labeling is a prerequisite to produce training data and **each piece of data will generally be labeled by a human**

With **unsupervised machine learning**, **labels will be produced by the machine**, and may not be human readable.

What is a ground truth?

A properly labeled dataset that you use as the objective standard to train and assess a given model is often called "ground truth." The accuracy of your trained model will depend on the accuracy of your ground truth

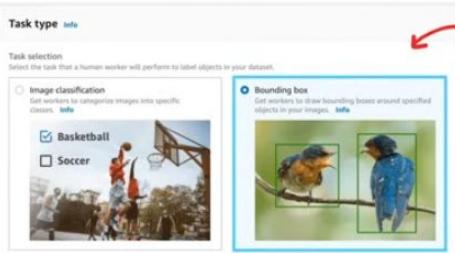
▶ 17:18

What is Labeling?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Data Labeling?

the process of **identifying raw data** (images, text files, videos, etc.) **and adding one or more meaningful and informative labels to provide context** so that a machine learning model can learn



The screenshot shows a user interface for selecting a labeling task. It includes a 'Task type' section with 'Image classification' and 'Bounding box' options, and a preview area showing images of basketball players and birds.

With **supervised machine learning**, labeling is a prerequisite to produce training data and **each piece of data will generally be labeled by a human**

With **unsupervised machine learning**, **labels will be produced by the machine**, and may not be human readable.

What is a ground truth?

A properly labeled dataset that you use as the objective standard to train and assess a given model is often called "ground truth." The accuracy of your trained model will depend on the accuracy of your ground truth

▶ 17:18

What is Labeling?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Data Labeling?

the process of **identifying raw data** (images, text files, videos, etc.) **and adding one or more meaningful and informative labels to provide context** so that a machine learning model can learn

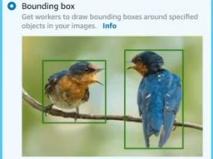
Task type [Info](#)

Task selection
Select the task that a human worker will perform to label objects in your dataset.

Image classification
Get workers to categorize images into specific classes. [Info](#)

Bounding box
Get workers to draw bounding boxes around specified objects in your images. [Info](#)





With **supervised machine learning**, labeling is a prerequisite to produce training data and **each piece of data will generally be labeled by a human**

With **unsupervised machine learning**, **labels will be produced by the machine**, and may not be human readable.

What is a ground truth?

A properly labeled dataset that you use as the objective standard to train and assess a given model is often called "ground truth." The accuracy of your trained model will depend on the accuracy of your ground truth

- ★ A properly labeled data set used as an objective standard for training and assessing a model is referred to as ground truth, which significantly impacts the accuracy of the trained model.

▷ 17:35

What is Labeling?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is Data Labeling?

the process of **identifying raw data** (images, text files, videos, etc.) **and adding one or more meaningful and informative labels to provide context** so that a machine learning model can learn

Task type [Info](#)

Task selection
Select the task that a human worker will perform to label objects in your dataset.

Image classification
Get workers to categorize images into specific classes. [Info](#)

Bounding box
Get workers to draw bounding boxes around specified objects in your images. [Info](#)





With **supervised machine learning**, labeling is a prerequisite to produce training data and **each piece of data will generally be labeled by a human**

With **unsupervised machine learning**, **labels will be produced by the machine**, and may not be human readable.

What is a ground truth?

A properly labeled dataset that you use as the objective standard to train and assess a given model is often called "ground truth." The accuracy of your trained model will depend on the accuracy of your ground truth

- ★ The accuracy of a trained model relies on the precision of its ground truth, and there are distinctions between supervised, unsupervised, and reinforcement learning.

▷ 17:41

Supervised vs Unsupervised vs Reinforcement

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Supervised Learning (SL)

Data that has been labeled for training
Task-driven – *make a prediction*

When the labels are known and you want a precise outcome.

When you need a specific value returned
eg. **Classification, Regression**

Unsupervised Learning (SL)

Data has not been labeled, the ML model needs to do its own labeling
Data-driven – *recognize a structure or pattern*

When the labels are not known and the outcome does not need to be precise.

When you're trying to make sense of data.
eg. **Clustering, Dimensionality Reduction, Association**

Reinforcement Learning (RI)

There is no data, there is an environment and an ML model generates data any many attempt to reach a goal
Decisions-driven – *Game AI, Learning Tasks, Robot Navigation*

- In machine learning, classification and regression are used for supervised learning with known labels to make precise predictions, while unsupervised learning involves data without labels where the model identifies patterns or structures independently.

▷ 17:54

Supervised vs Unsupervised vs Reinforcement

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Supervised Learning (SL)

Data that has been labeled for training
Task-driven – *make a prediction*

When the labels are known and you want a precise outcome.

When you need a specific value returned
eg. **Classification, Regression**

Unsupervised Learning (SL)

Data has not been labeled, the ML model needs to do its own labeling
Data-driven – *recognize a structure or pattern*

When the labels are not known and the outcome does not need to be precise.

When you're trying to make sense of data.
eg. **Clustering, Dimensionality Reduction, Association**

Reinforcement Learning (RI)

There is no data, there is an environment and an ML model generates data any many attempt to reach a goal

▷ 18:44

Supervised vs Unsupervised vs Reinforcement

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Supervised Learning (SL)

Data that has been labeled for training
Task-driven – *make a prediction*

When the labels are known and you want a precise outcome.

When you need a specific value returned
eg. **Classification, Regression**

Unsupervised Learning (UL)

Data has not been labeled, the ML model needs to do its own labeling
Data-driven – *recognize a structure or pattern*

When the labels are not known and the outcome does not need to be precise.

When you're trying to make sense of data.
eg. **Clustering, Dimensionality Reduction, Association**

Reinforcement Learning (RL)

There is no data, there is an environment and an ML model generates data any many attempt to reach a goal
Decisions-driven – *Game AI, Learning Tasks, Robot Navigation*

▷ 18:47

Supervised vs Unsupervised vs Reinforcement

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Supervised Learning (SL)

Data that has been labeled for training
Task-driven – *make a prediction*

When the labels are known and you want a precise outcome.

When you need a specific value returned
eg. **Classification, Regression**

Unsupervised Learning (UL)

Data has not been labeled, the ML model needs to do its own labeling
Data-driven – *recognize a structure or pattern*

When the labels are not known and the outcome does not need to be precise.

When you're trying to make sense of data.
eg. **Clustering, Dimensionality Reduction, Association**

Reinforcement Learning (RL)

There is no data, there is an environment and an ML model generates data any many attempt to reach a goal
Decisions-driven – *Game AI, Learning Tasks, Robot Navigation*

- Attempts to reach a goal in AI involve decision-driven processes, particularly in tasks like game AI and robot navigation, while classical machine learning relies heavily on statistics and math through supervised and unsupervised methods.

▷ 18:59

Supervised vs Unsupervised vs Reinforcement

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Supervised Learning (SL)

Data that has been labeled for training
Task-driven – **make a prediction**

When the labels are known and you want a precise outcome.

When you need a specific value returned
eg. **Classification, Regression**

Unsupervised Learning (UL)

Data has not been labeled, the ML model needs to do its own labeling
Data-driven – **recognize a structure or pattern**

When the labels are not known and the outcome does not need to be precise.

When you're trying to make sense of data.
eg. **Clustering, Dimensionality Reduction, Association**

Reinforcement Learning (RL)

There is no data, there is an environment and an ML model generates data any many attempt to reach a goal
Decisions-driven – **Game AI, Learning Tasks, Robot Navigation**

- 💡 A neural network is a computational model that mimics the brain, consisting of nodes or neurons that process input data through algorithms.

▶ 19:06

Supervised vs Unsupervised vs Reinforcement

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Supervised Learning (SL)

Data that has been labeled for training
Task-driven – **make a prediction**

When the labels are known and you want a precise outcome.

When you need a specific value returned
eg. **Classification, Regression**

Unsupervised Learning (UL)

Data has not been labeled, the ML model needs to do its own labeling
Data-driven – **recognize a structure or pattern**

When the labels are not known and the outcome does not need to be precise.

When you're trying to make sense of data.
eg. **Clustering, Dimensionality Reduction, Association**

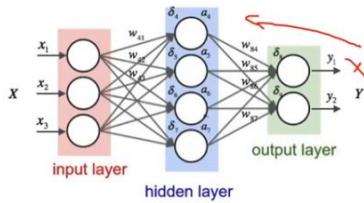
Reinforcement Learning (RL)

There is no data, there is an environment and an ML model generates data any many attempt to reach a goal
Decisions-driven – **Game AI, Learning Tasks, Robot Navigation**

▶ 19:07

Neural Networks and Deep Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



What are Neural Networks? (NN)

Often described as **mimicking the brain**, a **neuron/node represents an algorithm**. Data is inputted into a neuron and based on the output the data will be passed to one of many other connected neurons.

The connection between neurons is **weighted**.
The network is organized in **layers**.

There will be an **input layer**, 1 to many **hidden layers** and an **output layer**.

What is Deep Learning?

A neural network that has **3 or more hidden layers** is considered deep learning.

What is Feed Forward? (FNN)

Neural Networks where connections between nodes do not form a cycle
(always move forward)

What is Backpropagation (BP)?

Moves backwards through the neural network adjusting weights to improve outcome on next iteration. This is how a neural net learns.

Loss Function

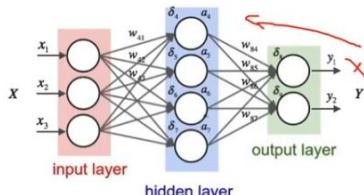
A function that compares the ground truth to the prediction to determine the error rate (how bad the network performed)

- A neural network consists of layers, including an input layer, one or more hidden layers, and an output layer, with connections between neurons that are weighted.

▷ 19:21

Neural Networks and Deep Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



What are Neural Networks? (NN)

Often described as **mimicking the brain**, a **neuron/node represents an algorithm**. Data is inputted into a neuron and based on the output the data will be passed to one of many other connected neurons.

The connection between neurons is **weighted**.
The network is organized in **layers**.

There will be an **input layer**, 1 to many **hidden layers** and an **output layer**.

What is Deep Learning?

A neural network that has **3 or more hidden layers** is considered deep learning.

What is Feed Forward? (FNN)

Neural Networks where connections between nodes do not form a cycle
(always move forward)

What is Backpropagation (BP)?

Moves backwards through the neural network adjusting weights to improve outcome on next iteration. This is how a neural net learns.

Loss Function

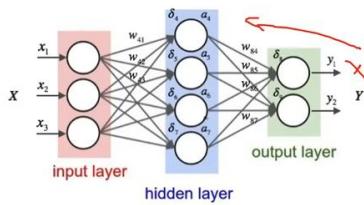
A function that compares the ground truth to the prediction to determine the error rate (how bad the network performed)

- The process involves calculating the error rate at the end of a network's performance, followed by backpropagation to adjust weights, and applying activation functions to hidden layer nodes that influence learning.

▷ 20:30

Neural Networks and Deep Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



What are Neural Networks? (NN)

Often described as **mimicking the brain**, a **neuron/node represents an algorithm**.

Data is inputted into a neuron and based on the output the data will be passed to one of many other connected neurons.
The connection between neurons is **weighted**.

The network is organized in layers.

There will be an input layer, 1 to many hidden layers and an output layer.

What is Deep Learning?

A neural network that has **3 or more hidden layers** is considered deep learning.

What is Feed Forward? (FNN)

Neural Networks where connections between nodes do not form a cycle
(always move forward)

What is Backpropagation (BP)?

Moves backwards through the neural network adjusting weights to improve outcome on next iteration. This is how a neural net learns.

Loss Function

A function that compares the ground truth to the prediction to determine the error rate (how bad the network performed)

Activation Functions

An algorithm applied to a hidden layer node that affects connected output e.g. ReLU

Dense

When the next layer increases the amount of nodes

Sparse

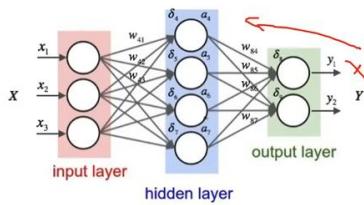
When the next layer decreases the amount of nodes

- Dimensionality reduction occurs when transitioning from a dense layer to a sparse layer in a neural network, as it reduces the number of nodes and thus the dimensions.

▶ 21:18

Neural Networks and Deep Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



What are Neural Networks? (NN)

Often described as **mimicking the brain**, a **neuron/node represents an algorithm**.

Data is inputted into a neuron and based on the output the data will be passed to one of many other connected neurons.
The connection between neurons is **weighted**.

The network is organized in layers.

There will be an input layer, 1 to many hidden layers and an output layer.

What is Deep Learning?

A neural network that has **3 or more hidden layers** is considered deep learning.

What is Feed Forward? (FNN)

Neural Networks where connections between nodes do not form a cycle
(always move forward)

What is Backpropagation (BP)?

Moves backwards through the neural network adjusting weights to improve outcome on next iteration. This is how a neural net learns.

Loss Function

A function that compares the ground truth to the prediction to determine the error rate (how bad the network performed)

Activation Functions

An algorithm applied to a hidden layer node that affects connected output e.g. ReLU

Dense

When the next layer increases the amount of nodes

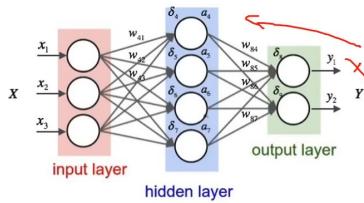
Sparse

When the next layer decreases the amount of nodes

▶ 21:20

Neural Networks and Deep Learning

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



What are Neural Networks? (NN)

Often described as **mimicking the brain**, a **neuron/node represents an algorithm**. Data is inputted into a neuron and based on the output the data will be passed to one of many other connected neurons.

The connection between neurons is **weighted**.
The network is organized in **layers**.

There will be an input layer, 1 to many hidden layers and an output layer.

What is Deep Learning?

A neural network that has **3 or more hidden layers** is considered deep learning.

What is Feed Forward? (FNN)

Neural Networks where connections between nodes do not form a cycle
(always move forward)

What is Backpropagation (BP)?

Moves backwards through the neural network adjusting weights to improve outcome on next iteration. This is how a neural net learns.

Loss Function

A function that compares the ground truth to the prediction to determine the error rate (how bad the network performed)

Activation Functions

An algorithm applied to a hidden layer node that affects connected output e.g. ReLU

Dense

When the next layer increases the amount of nodes

Sparse

When the next layer decreases the amount of nodes

- The transition from a dense layer to a sparse layer in a neural network is referred to as dimensionality reduction, as it reduces the number of dimensions by decreasing the amount of nodes.

▷ 21:22

What is a GPU?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a GPU?

A General Processing Unit (GPU) that is specially designed to quickly render high-resolution images and video **concurrently**.

GPUs can perform **parallel operations on multiple sets of data**, and so they are commonly used for **non-graphical tasks such as machine learning** and scientific computation.

CPU can have average 4 to 16 processor cores...

GPUs can **thousands of processor cores**

4 to 8 GPUs can provide as many as 40,000 cores

- GPUs, with their thousands of processor cores, are highly effective for tasks like machine learning and neural networks due to their ability to handle repetitive computations efficiently.

▷ 21:44

What is a GPU?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a GPU?

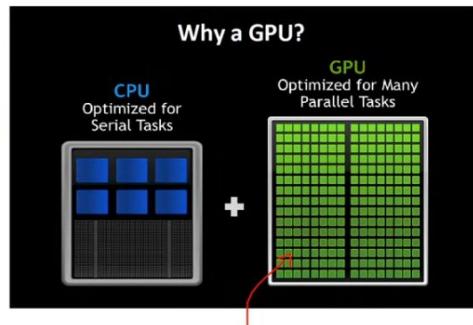
A General Processing Unit (GPU) that is specially designed to quickly render high-resolution images and video **concurrently**.

GPUs can perform **parallel operations on multiple sets of data**, and so they are commonly used for **non-graphical tasks such as machine learning** and scientific computation.

CPU can have average 4 to 16 processor cores...

GPUs can **thousands of processor cores**

4 to 8 GPUs can provide as many as 40,000 cores



GPUs are best suited for repetitive and highly-parallel computing tasks:

- Rendering graphics
- Cryptocurrency mining
- Deep Learning and ML

▶ 22:02

What is a GPU?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a GPU?

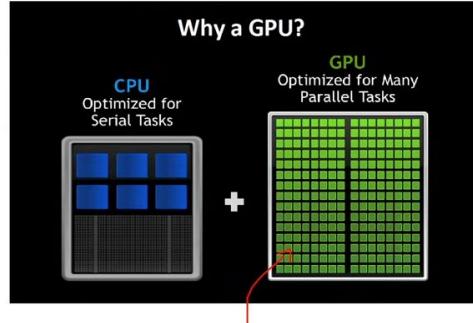
A General Processing Unit (GPU) that is specially designed to quickly render high-resolution images and video **concurrently**.

GPUs can perform **parallel operations on multiple sets of data**, and so they are commonly used for **non-graphical tasks such as machine learning** and scientific computation.

CPU can have average 4 to 16 processor cores...

GPUs can **thousands of processor cores**

4 to 8 GPUs can provide as many as 40,000 cores



GPUs are best suited for repetitive and highly-parallel computing tasks:

- Rendering graphics
- Cryptocurrency mining
- Deep Learning and ML

▶ 22:19

What is a GPU?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a GPU?

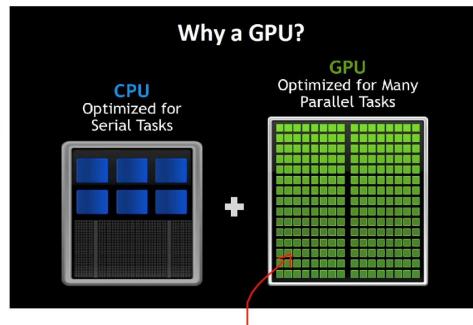
A General Processing Unit (GPU) that is specially designed to quickly render high-resolution images and video **concurrently**.

GPUs can perform **parallel operations on multiple sets of data**, and so they are commonly used for **non-graphical tasks such as machine learning** and scientific computation.

CPU can have average 4 to 16 processor cores...

GPUs can **thousands of processor cores**

4 to 8 GPUs can provide as many as 40,000 cores



GPUs are best suited for repetitive and highly-parallel computing tasks:

- Rendering graphics
- Cryptocurrency mining
- Deep Learning and ML

- 💡 GPUs are well-suited for repetitive and highly parallel computing tasks, making them ideal for applications like graphics rendering, cryptocurrency mining, deep learning, and machine learning.

▷ 22:19

What is a GPU?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a GPU?

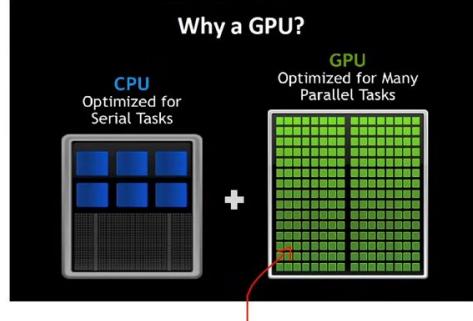
A General Processing Unit (GPU) that is specially designed to quickly render high-resolution images and video **concurrently**.

GPUs can perform **parallel operations on multiple sets of data**, and so they are commonly used for **non-graphical tasks such as machine learning** and scientific computation.

CPU can have average 4 to 16 processor cores...

GPUs can **thousands of processor cores**

4 to 8 GPUs can provide as many as 40,000 cores



GPUs are best suited for repetitive and highly-parallel computing tasks:

- Rendering graphics
- Cryptocurrency mining
- Deep Learning and ML

- 💡 Nvidia is a company that manufactures graphical processor units (GPUs) suited for repetitive and highly parallel computing tasks such as rendering graphics, cryptocurrency mining, deep learning, and machine learning.

▷ 22:20



ML and AI Concepts

Compute Unified Device Architecture

- GPUs are ideal for repetitive and highly parallel computing tasks like rendering graphics, cryptocurrency mining, deep learning, and machine learning, with Nvidia being a key manufacturer of these processors and CUDA serving as their unified computing platform.

▷ 22:24

What is CUDA?

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

What is NVIDIA?

NVIDIA is a company that manufactures **graphical processing units (GPUs)** for gaming and professional markets



What is CUDA?

Compute Unified Device Architecture (CUDA) is a **parallel computing platform** and **API** by NVIDIA that allows developers to use **CUDA-enabled GPUs** for general-purpose computing on GPUs (GPGPU)

All major deep learning frameworks are integrated with **NVIDIA Deep Learning SDK**

The NVIDIA Deep Learning SDK is a collection of NVIDIA libraries for deep learning.

One of those libraries is the **CUDA Deep Neural Network library (cuDNN)**

cuDNN provides highly tuned implementations for standard routines such as:

- forward and backward convolution
- Pooling
- Normalization
- activation layers

- CUDA is a parallel computing platform and API by Nvidia that enables developers to utilize CUDA-enabled GPUs for general-purpose computing, with major deep learning frameworks integrated into Nvidia's deep learning SDK, which includes the CUDA Deep Neural Network library.

▷ 22:34

What is CUDA?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is NVIDIA?

NVIDIA is a company that manufactures **graphical processing units (GPUs)** for gaming and professional markets



What is CUDA?

Compute Unified Device Architecture (CUDA) is a **parallel computing platform** and **API** by NVIDIA that allows developers to use **CUDA-enabled GPUs** for general-purpose computing on GPUs (GPGPU)

All major deep learning frameworks are integrated with **NVIDIA Deep Learning SDK**

The NVIDIA Deep Learning SDK is a collection of NVIDIA libraries for deep learning.

One of those libraries is the **CUDA Deep Neural Network library (cuDNN)**

cuDNN provides highly tuned implementations for standard routines such as:

- forward and backward convolution
- Pooling
- Normalization
- activation layers

▶ 23:19

What is CUDA?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is NVIDIA?

NVIDIA is a company that manufactures **graphical processing units (GPUs)** for gaming and professional markets



What is CUDA?

Compute Unified Device Architecture (CUDA) is a **parallel computing platform** and **API** by NVIDIA that allows developers to use **CUDA-enabled GPUs** for general-purpose computing on GPUs (GPGPU)

All major deep learning frameworks are integrated with **NVIDIA Deep Learning SDK**

The NVIDIA Deep Learning SDK is a collection of NVIDIA libraries for deep learning.

One of those libraries is the **CUDA Deep Neural Network library (cuDNN)**

cuDNN provides highly tuned implementations for standard routines such as:

- forward and backward convolution
- Pooling
- Normalization
- activation layers

💡 Tuned implementations of standard routines enhance the performance of computer vision tasks, making GPUs significant despite their absence in certain certifications.

▶ 23:21

What is CUDA?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is NVIDIA?

NVIDIA is a company that manufactures **graphical processing units (GPUs)** for gaming and professional markets



What is CUDA?

Compute Unified Device Architecture (CUDA) is a **parallel computing platform** and **API** by NVIDIA that allows developers to use **CUDA-enabled GPUs** for general-purpose computing on GPUs (GPGPU)

All major deep learning frameworks are integrated with **NVIDIA Deep Learning SDK**

The NVIDIA Deep Learning SDK is a collection of NVIDIA libraries for deep learning.

One of those libraries is the **CUDA Deep Neural Network library (cuDNN)**

cuDNN provides highly tuned implementations for standard routines such as:

- forward and backward convolution

- Pooling

- Normalization

- activation layers

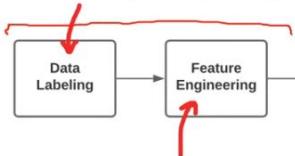
- 💡 Understanding pooling, normalization, and activation layers in computer vision is important for grasping the significance of GPUs, even though CUDA will not be covered in the AI 900 certification.

▷ 23:27

ML Pipeline

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

For supervised learning, you need to label your data so the ML model can learn by example during training

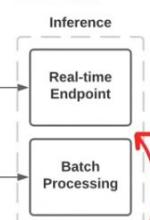


ML models only work with numerical data. So you need to translate it into a format that it can understand, extract out the important data that the ML needs to focus on.

Your model needs to learn how to become smarter. It will perform multiple iterations getting smarter with each iteration

An ML model can have different parameters, we can use ML to try out many different parameters to optimize the outcome

We need to make our ML model accessible, so we serve by hosting in a virtual machine or container.



Inference is the act of requesting to make a prediction

Very simplified ML pipeline

- 💡 Data labeling and feature engineering are essential pre-processing steps that prepare data for machine learning models by converting it into a numerical format that highlights important features.

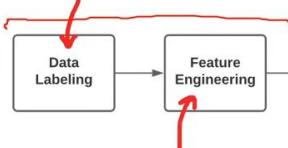
▷ 23:51

ML Pipeline

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

For supervised learning you need to label your data so the ML model can learn by example during training

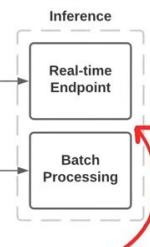
Your model needs to learn how to become smarter. It will perform multiple iterations getting smarter with each iteration



ML models only work with numerical data. So you need to translate it into a format that it can understand, extract out the important data that the ML needs to focus on.

An ML model can have different parameters, we can use ML to try out many different parameters to optimize the outcome

We need to make our ML model accessible, so we serve by hosting in a virtual machine or container.



Inference is the act of requesting to make a prediction

- Inference involves making predictions by sending a payload, with options for real-time or batch processing depending on whether single-item or multiple-item predictions are needed.

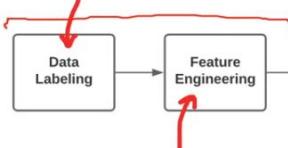
▶ 25:17

ML Pipeline

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

For supervised learning you need to label your data so the ML model can learn by example during training

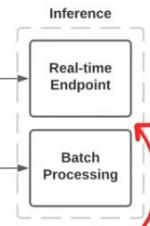
Your model needs to learn how to become smarter. It will perform multiple iterations getting smarter with each iteration



ML models only work with numerical data. So you need to translate it into a format that it can understand, extract out the important data that the ML needs to focus on.

An ML model can have different parameters, we can use ML to try out many different parameters to optimize the outcome

We need to make our ML model accessible, so we serve by hosting in a virtual machine or container.



Inference is the act of requesting to make a prediction

Very simplified ML pipeline

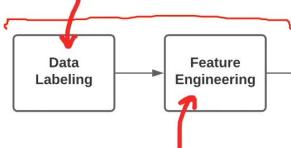
▶ 25:21

ML Pipeline

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

For supervised learning you need to label your data so the ML model can learn by example during training

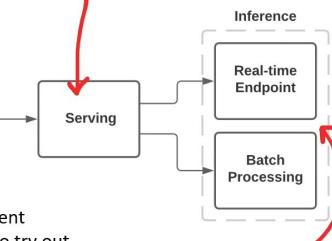
Your model needs to learn how to become smarter. It will perform multiple iterations getting smarter with each iteration



ML models only work with numerical data. So you need to translate it into a format that it can understand, extract out the important data that the ML needs to focus on.

An ML model can have different parameters, we can use ML to try out many different parameters to optimize the outcome

We need to make are ML model accessible, so we serve by hosting in a virtual machine or container.



Inference is the act of requesting to make a prediction

Very simplified ML pipeline

- Real-time endpoints allow for single-item predictions, while batch processing handles multiple data inputs at once, typically resulting in slower performance.

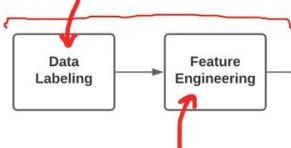
▶ 25:35

ML Pipeline

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

For supervised learning you need to label your data so the ML model can learn by example during training

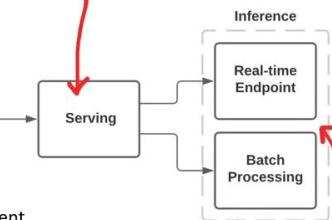
Your model needs to learn how to become smarter. It will perform multiple iterations getting smarter with each iteration



ML models only work with numerical data. So you need to translate it into a format that it can understand, extract out the important data that the ML needs to focus on.

An ML model can have different parameters, we can use ML to try out many different parameters to optimize the outcome

We need to make are ML model accessible, so we serve by hosting in a virtual machine or container.



Inference is the act of requesting to make a prediction

Very simplified ML pipeline

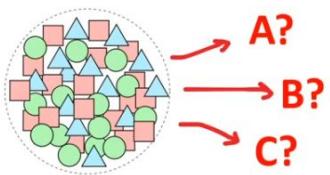
- Forecasting involves making informed predictions based on relevant data to analyze trends, rather than relying on guesswork.

▶ 25:38

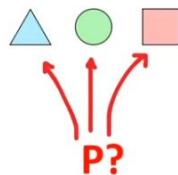
Forecasting vs Prediction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a **Forecasting**?



What is a **Prediction**?



- Makes a future prediction with **relevant data**
 - analysis of trends
 - Its not “guessing”

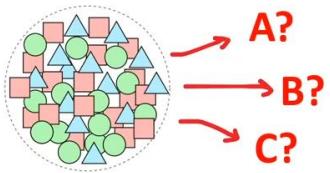
- Makes a future prediction **without relevant data**
 - uses statistics to predict future outcomes
 - Its more of “guessing”
 - Uses decision theory

▷ 25:54

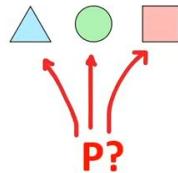
Forecasting vs Prediction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a **Forecasting**?



What is a **Prediction**?



- Makes a future prediction with **relevant data**
 - analysis of trends
 - Its not “guessing”

- Makes a future prediction **without relevant data**
 - uses statistics to predict future outcomes
 - Its more of “guessing”
 - Uses decision theory

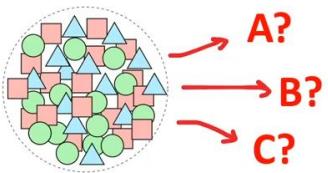
✳️ Performance or evaluation metrics are used to assess the effectiveness of various machine learning algorithms in making predictions based on limited data.

▷ 26:19

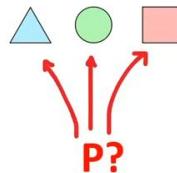
Forecasting vs Prediction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a **Forecasting**?



What is a **Prediction**?



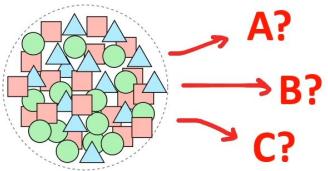
- Makes a future prediction with **relevant data**
 - analysis of trends
 - Its not “guessing”
- Makes a future prediction **without relevant data**
 - uses statistics to predict future outcomes
 - Its more of “guessing”
 - Uses decision theory

▷ 26:21

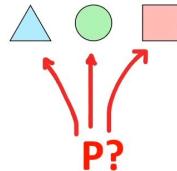
Forecasting vs Prediction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is a **Forecasting**?



What is a **Prediction**?



- Makes a future prediction with **relevant data**
 - analysis of trends
 - Its not “guessing”
- Makes a future prediction **without relevant data**
 - uses statistics to predict future outcomes
 - Its more of “guessing”
 - Uses decision theory

★ Performance or evaluation metrics are used to assess the effectiveness of different machine learning algorithms by analyzing their predictions.

▷ 26:22

What are Metrics?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Performance/Evaluation Metrics are used to evaluate different Machine Learning Algorithms

For different types of problems different metrics matter, (*this is not an exhaustive list*)

- Classification Metrics (accuracy, precision, recall, F1-score, ROC, AUC)
- Regression Metrics (MSE, RMSE MAE)
- Ranking Metrics (MRR, DCG, NDCG)
- Statistical Metrics (Correlation)
- Computer Vision Metrics (PSNR, SSIM, IoU)
- NLP Metrics (Perplexity, BLEU, METEOR, ROUGE)
- Deep Learning Related Metrics (Inception score, Frechet Inception distance)

There are two categories of evaluation metrics

- Internal Evaluation — metrics used to evaluate the internals of the ML model
 - Accuracy, F1 Score, Precision, Recall (The Famous Four) used in all kinds of models
- External Evaluation — metrics used to evaluate the final prediction of the ML model

💡 Different metrics are used to evaluate machine learning models, with classification metrics including accuracy, precision, recall, F1 score, and ROC AUC, while regression metrics include mean squared error (MSE).

▷ 26:35

What are Metrics?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Performance/Evaluation Metrics are used to evaluate different Machine Learning Algorithms

For different types of problems different metrics matter, (*this is not an exhaustive list*)

- Classification Metrics (accuracy, precision, recall, F1-score, ROC, AUC)
- Regression Metrics (MSE, RMSE MAE)
- Ranking Metrics (MRR, DCG, NDCG)
- Statistical Metrics (Correlation)
- Computer Vision Metrics (PSNR, SSIM, IoU)
- NLP Metrics (Perplexity, BLEU, METEOR, ROUGE)
- Deep Learning Related Metrics (Inception score, Frechet Inception distance)

There are two categories of evaluation metrics

- Internal Evaluation — metrics used to evaluate the internals of the ML model
 - Accuracy, F1 Score, Precision, Recall (The Famous Four) used in all kinds of models
- External Evaluation — metrics used to evaluate the final prediction of the ML model

▷ 27:48

What are Metrics?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Performance/Evaluation Metrics are used to evaluate different Machine Learning Algorithms

For different types of problems different metrics matter, (*this is not an exhaustive list*)

- Classification Metrics (accuracy, precision, recall, F1-score, ROC, AUC)
- Regression Metrics (MSE, RMSE MAE)
- Ranking Metrics (MRR, DCG, NDCG)
- Statistical Metrics (Correlation)
- Computer Vision Metrics (PSNR, SSIM, IoU)
- NLP Metrics (Perplexity, BLEU, METEOR, ROUGE)
- Deep Learning Related Metrics (Inception score, Frechet Inception distance)

There are two categories of evaluation metrics

- Internal Evaluation — metrics used to evaluate the internals of the ML model
 - Accuracy, F1 Score, Precision, Recall (The Famous Four) used in all kinds of models
- External Evaluation — metrics used to evaluate the final prediction of the ML model

- ❖ Various models and external evaluation metrics are used to assess the final predictions of machine learning models, with an emphasis on key metrics that will be revisited.

▷ 27:56

What are Metrics?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Performance/Evaluation Metrics are used to evaluate different Machine Learning Algorithms

For different types of problems different metrics matter, (*this is not an exhaustive list*)

- Classification Metrics (accuracy, precision, recall, F1-score, ROC, AUC)
- Regression Metrics (MSE, RMSE MAE)
- Ranking Metrics (MRR, DCG, NDCG)
- Statistical Metrics (Correlation)
- Computer Vision Metrics (PSNR, SSIM, IoU)
- NLP Metrics (Perplexity, BLEU, METEOR, ROUGE)
- Deep Learning Related Metrics (Inception score, Frechet Inception distance)

There are two categories of evaluation metrics

- Internal Evaluation — metrics used to evaluate the internals of the ML model
 - Accuracy, F1 Score, Precision, Recall (The Famous Four) used in all kinds of models
- External Evaluation — metrics used to evaluate the final prediction of the ML model

- ❖ The discussion covers the use of models and external evaluation metrics for assessing machine learning predictions, along with an introduction to Jupyter notebooks as tools for combining live code, narrative text, equations, and visualizations.

▷ 27:58

What is Jupyter Notebook?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Jupyter Notebook

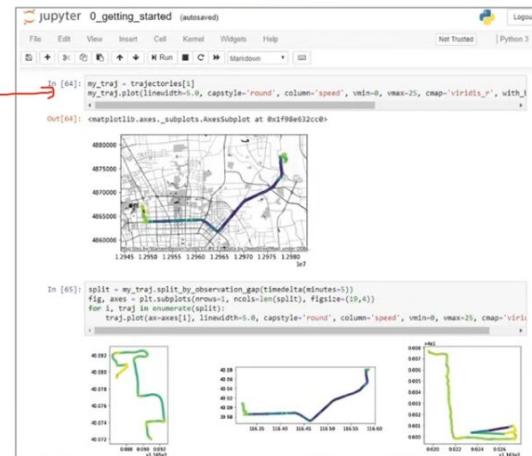
A Web-based application for authoring documents that combine:

- live-code
- narrative text
- equations
- visualizations

iPython's notebook feature *became* Jupyter Notebook

Jupyter Notebooks were overhauled and better integrated into an IDE called **JupyterLab**

You generally want to open Notebooks in Labs
The legacy web-based interface is known as
Jupyter classic notebook



In [64]: my_traj = trajectories[1]
my_traj.plot(linewidth=5.0, capstyle='round', column='speed', vmin=0, vmax=25, cmap='viridis_r', with_cbar=True)
Out[64]: <matplotlib.axes._subplots.AxesSubplot at 0x1f90e632cc0>

In [65]: split = my_traj.split_by_observation_gap(timedelta=minutes-5)
fig, axes = plt.subplots(1, len(split), figsize=(9,4))
for i, traj in enumerate(split):
 traj.plot(ax=axes[i], linewidth=5.0, capstyle='round', column='speed', vmin=0, vmax=25, cmap='viridis_r')
 axes[i].set_yticks([40.00, 40.05, 40.10, 40.15, 40.20, 40.25, 40.30])
 axes[i].set_xticks([114.35, 114.40, 114.45, 114.50, 114.55, 114.60])
 axes[i].set_xlabel('lat')
 axes[i].set_ylabel('lon')
 axes[i].set_title(f'{i+1}/{len(split)}')

- ★ Jupyter notebooks, originally derived from IPython, serve as an interactive environment for executing Python code and are commonly used in data science and machine learning applications.

▶ 28:08

What is Jupyter Notebook?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Jupyter Notebook

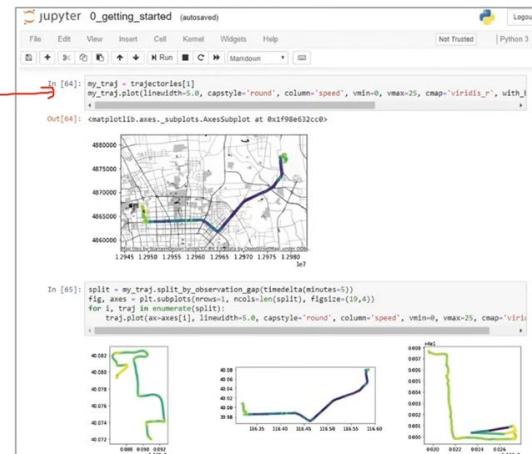
A Web-based application for authoring documents that combine:

- live-code
- narrative text
- equations
- visualizations

iPython's notebook feature *became* Jupyter Notebook

Jupyter Notebooks were overhauled and better integrated into an IDE called **JupyterLab**

You generally want to open Notebooks in Labs
The legacy web-based interface is known as
Jupyter classic notebook



In [64]: my_traj = trajectories[1]
my_traj.plot(linewidth=5.0, capstyle='round', column='speed', vmin=0, vmax=25, cmap='viridis_r', with_cbar=True)
Out[64]: <matplotlib.axes._subplots.AxesSubplot at 0x1f90e632cc0>

In [65]: split = my_traj.split_by_observation_gap(timedelta=minutes-5)
fig, axes = plt.subplots(1, len(split), figsize=(9,4))
for i, traj in enumerate(split):
 traj.plot(ax=axes[i], linewidth=5.0, capstyle='round', column='speed', vmin=0, vmax=25, cmap='viridis_r')
 axes[i].set_yticks([40.00, 40.05, 40.10, 40.15, 40.20, 40.25, 40.30])
 axes[i].set_x_ticks([114.35, 114.40, 114.45, 114.50, 114.55, 114.60])
 axes[i].set_xlabel('lat')
 axes[i].set_ylabel('lon')
 axes[i].set_title(f'{i+1}/{len(split)}')

- ★ Jupyter Labs is the next generation web-based user interface that incorporates all familiar features of classic Jupyter notebooks in a more flexible format.

▶ 28:46

What is Jupyter Notebook?

Cheat sheets, Practice Exams and Flash cards  www.exapro.co/ai-900



Jupyter Notebook

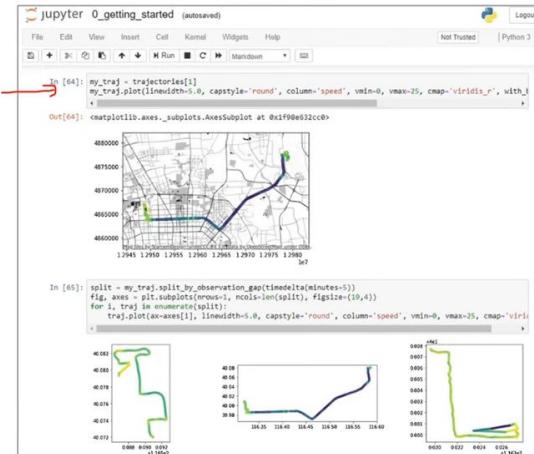
A Web-based application for authoring documents that combine:

- live-code
- narrative text
- equations
- visualizations

iPython's notebook feature *became* Jupyter Notebook

Jupyter Notebooks were overhauled and better integrated into an IDE called **JupyterLab**

You generally want to open Notebooks in Labs
The legacy web-based interface is known as
Jupyter classic notebook



► 28:52

What is Jupyter Notebook?

Cheat sheets, Practice Exams and Flash cards  www.exapro.co/ai-900



Jupyter Notebook

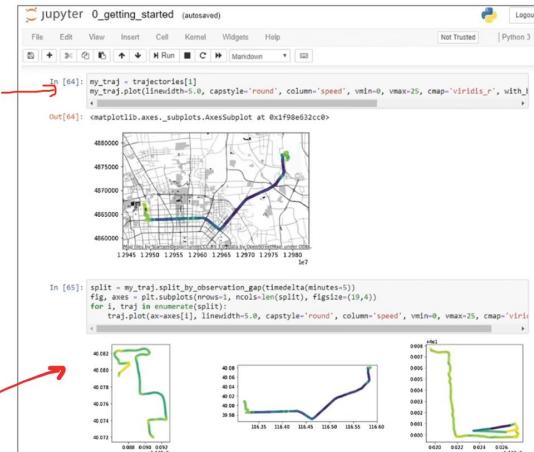
A Web-based application for authoring documents that combine:

- live-code
- narrative text
- equations
- visualizations

iPython's notebook feature *became* Jupyter Notebook

Jupyter Notebooks were overhauled and better integrated into an IDE called **JupyterLab**

You generally want to open Notebooks in Labs
The legacy web-based interface is known as
Jupyter classic notebook



★ Jupyter Labs is the next-generation web-based user interface that integrates familiar features of classic Jupyter notebooks with a more flexible and powerful design, including tools like notebooks, terminals, text editors, file browsers, and rich outputs.

► 28:53

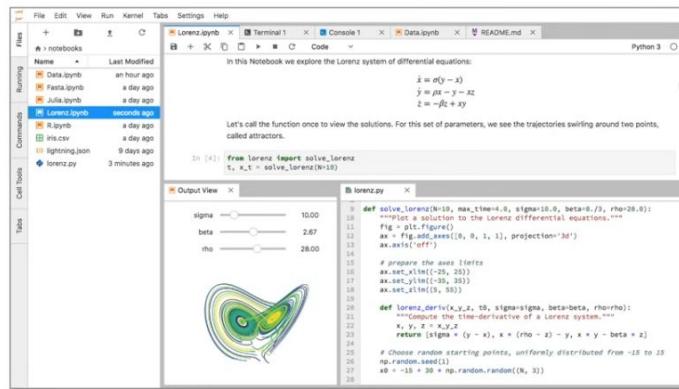
What is JupyterLab?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



JupyterLab is a a **next-generation web-based user interface**

All the familiar features of the classic Jupyter Notebook in a flexible and powerful user interface:



The screenshot shows the JupyterLab interface with several tabs open: 'Lorenz.ipynb' (active), 'Terminal 1', 'Console 1', 'Data.ipynb', and 'README.md'. The 'File' tab is selected. The 'Running' section lists notebooks like 'R.ipynb', 'Data.ipynb', 'Fasta.ipynb', 'Julia.ipynb', and 'Lorenz.ipynb'. The 'Code' tab shows a code cell with Python code for solving the Lorenz system. The 'Output View' tab shows a 3D plot of the Lorenz attractor.

```
In [4]: from Lorenz import solve_lorenz
t, x, t = solve_lorenz(N=10)

In [5]: %matplotlib inline
import Lorenz
Lorenz.plot_lorenz_3d(x)
```

```
def solve_lorenz(N=10, max_time=4.0, sigma=10.0, beta=8./3., rho=28.0):
    """Plot a solution to the Lorenz differential equations."""
    t = np.linspace(0, max_time, N+1)
    x = np.zeros((3, N+1))
    x[:, 0] = np.random.rand(3, 1) * 10
    for i in range(1, N+1):
        x[:, i] = rk4_step(lorenz_deriv, x[:, i-1], t[i]-t[i-1])
    return x

def lorenz_deriv(x_y_z, t, sigma=sigma, beta=beta, rho=rho):
    """Compute the time-derivative of a Lorenz system."""
    x, y, z = x_y_z
    x_dot = sigma * (y - x)
    y_dot = rho * x - y - z
    z_dot = x * (rho - z) - y
    return [x_dot, y_dot, z_dot]

# Choose random starting points, uniformly distributed from -10 to 10
np.random.seed(1)
x0 = -10 + 30 * np.random.rand(3, 1)
```

JupyterLab will eventually replace the classic Jupyter Notebook

▶ 29:02

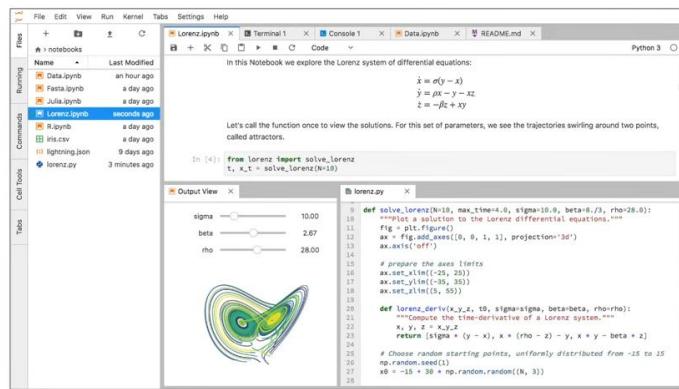
What is JupyterLab?

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



JupyterLab is a a **next-generation web-based user interface**

All the familiar features of the classic Jupyter Notebook in a flexible and powerful user interface:



The screenshot shows the JupyterLab interface with several tabs open: 'Lorenz.ipynb' (active), 'Terminal 1', 'Console 1', 'Data.ipynb', and 'README.md'. The 'File' tab is selected. The 'Running' section lists notebooks like 'R.ipynb', 'Data.ipynb', 'Fasta.ipynb', 'Julia.ipynb', and 'Lorenz.ipynb'. The 'Code' tab shows a code cell with Python code for solving the Lorenz system. The 'Output View' tab shows a 3D plot of the Lorenz attractor.

```
In [4]: from Lorenz import solve_lorenz
t, x, t = solve_lorenz(N=10)

In [5]: %matplotlib inline
import Lorenz
Lorenz.plot_lorenz_3d(x)
```

```
def solve_lorenz(N=10, max_time=4.0, sigma=10.0, beta=8./3., rho=28.0):
    """Plot a solution to the Lorenz differential equations."""
    t = np.linspace(0, max_time, N+1)
    x = np.zeros((3, N+1))
    x[:, 0] = np.random.rand(3, 1) * 10
    for i in range(1, N+1):
        x[:, i] = rk4_step(lorenz_deriv, x[:, i-1], t[i]-t[i-1])
    return x

def lorenz_deriv(x_y_z, t, sigma=sigma, beta=beta, rho=rho):
    """Compute the time-derivative of a Lorenz system."""
    x, y, z = x_y_z
    x_dot = sigma * (y - x)
    y_dot = rho * x - y - z
    z_dot = x * (rho - z) - y
    return [x_dot, y_dot, z_dot]

# Choose random starting points, uniformly distributed from -10 to 10
np.random.seed(1)
x0 = -10 + 30 * np.random.rand(3, 1)
```

JupyterLab will eventually replace the classic Jupyter Notebook

- JupyterLab is a web-based user interface that enhances the features of classic Jupyter notebooks with a more flexible and powerful design, incorporating elements like notebooks, terminals, text editors, file browsers, and rich outputs.

▶ 29:09

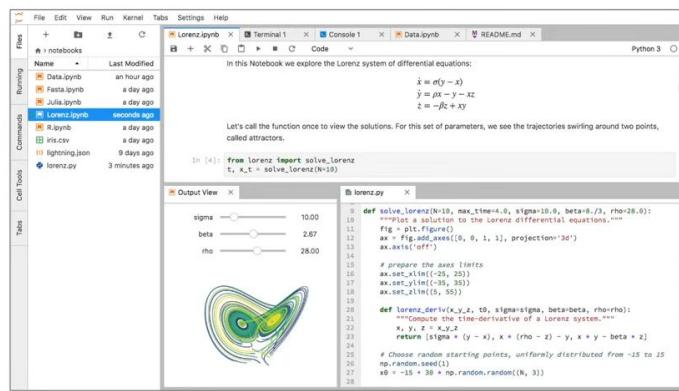
What is JupyterLab?

Cheat sheets, Practice Exams and Flash cards  www.exapro.co/ai-900



JupyterLab is a a **next-generation web-based user interface**

All the familiar features of the classic Jupyter Notebook in a flexible and powerful user interface:



The screenshot shows the JupyterLab interface. On the left is a file browser with notebooks like 'Data.ipynb' and 'Lorenz.ipynb'. In the center is a code editor with Python code for solving the Lorenz system. On the right is an output viewer displaying a 3D plot of the Lorenz attractor.

```
In [4]: from lorenz import solve_lorenz
t, x, t = solve_lorenz(N=10)

In [5]: lorenz.py
def solve_lorenz(N=10, max_time=4.0, sigma=10.0, beta=8./3., rho=28.0):
    """Plot a solution to the Lorenz differential equations.

    Parameters
    ----------
    N : int
        The number of points to plot
    max_time : float
        The time at which to stop integrating
    sigma : float
        The sigma parameter in the Lorenz system
    beta : float
        The beta parameter in the Lorenz system
    rho : float
        The rho parameter in the Lorenz system

    Returns
    -------
    x : array
        The x component of the solution
    y : array
        The y component of the solution
    z : array
        The z component of the solution
    """
    def lorenz_deriv(x_y_z, t0, sigma=sigma, beta=beta, rho=rho):
        """Compute the time-derivative of a Lorenz system.

        Parameters
        ----------
        x_y_z : array
            An array containing the current state (x, y, z)
        t0 : float
            The current time
        sigma : float
            The sigma parameter in the Lorenz system
        beta : float
            The beta parameter in the Lorenz system
        rho : float
            The rho parameter in the Lorenz system

        Returns
        -------
        x_dot : float
            The derivative of x
        y_dot : float
            The derivative of y
        z_dot : float
            The derivative of z
        """
        x, y, z = x_y_z
        x_dot = sigma * (y - x)
        y_dot = rho * x - y - z
        z_dot = x * (rho - z) - y
        return [x_dot, y_dot, z_dot]

    # Choose random starting points, uniformly distributed from -10 to 10
    np.random.seed(1)
    x0 = -15 + 30 * np.random.random((N, 3))

    # Integrate
    t = np.linspace(0, max_time, N)
    sol = integrate.odeint(lorenz_deriv, x0, t)

    # Plot
    fig = plt.figure()
    ax = fig.add_axes([0, 0, 1, 1], projection='3d')
    ax.set_xlim([-35, 35])
    ax.set_ylim([-35, 35])
    ax.set_zlim([0, 55])
    ax.set_xlabel('x')
    ax.set_ylabel('y')
    ax.set_zlabel('z')
    ax.view_init(30, 30)
    ax.plot(sol[:, 0], sol[:, 1], sol[:, 2], c='k')
```

JupyterLab will eventually replace the classic Jupyter Notebook

▷ 29:10

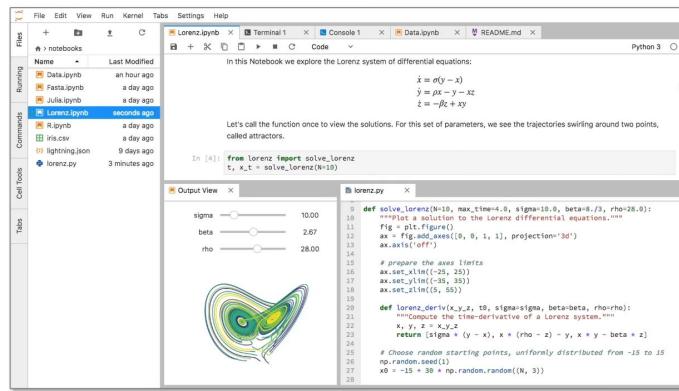
What is JupyterLab?

Cheat sheets, Practice Exams and Flash cards  www.exapro.co/ai-900



JupyterLab is a a **next-generation web-based user interface**

All the familiar features of the classic Jupyter Notebook in a flexible and powerful user interface:



The screenshot shows the JupyterLab interface. On the left is a file browser with notebooks like 'Data.ipynb' and 'Lorenz.ipynb'. In the center is a code editor with Python code for solving the Lorenz system. On the right is an output viewer displaying a 3D plot of the Lorenz attractor.

```
In [4]: from lorenz import solve_lorenz
t, x, t = solve_lorenz(N=10)

In [5]: lorenz.py
def solve_lorenz(N=10, max_time=4.0, sigma=10.0, beta=8./3., rho=28.0):
    """Plot a solution to the Lorenz differential equations.

    Parameters
    ----------
    N : int
        The number of points to plot
    max_time : float
        The time at which to stop integrating
    sigma : float
        The sigma parameter in the Lorenz system
    beta : float
        The beta parameter in the Lorenz system
    rho : float
        The rho parameter in the Lorenz system

    Returns
    -------
    x : array
        The x component of the solution
    y : array
        The y component of the solution
    z : array
        The z component of the solution
    """
    def lorenz_deriv(x_y_z, t0, sigma=sigma, beta=beta, rho=rho):
        """Compute the time-derivative of a Lorenz system.

        Parameters
        ----------
        x_y_z : array
            An array containing the current state (x, y, z)
        t0 : float
            The current time
        sigma : float
            The sigma parameter in the Lorenz system
        beta : float
            The beta parameter in the Lorenz system
        rho : float
            The rho parameter in the Lorenz system

        Returns
        -------
        x_dot : float
            The derivative of x
        y_dot : float
            The derivative of y
        z_dot : float
            The derivative of z
        """
        x, y, z = x_y_z
        x_dot = sigma * (y - x)
        y_dot = rho * x - y - z
        z_dot = x * (rho - z) - y
        return [x_dot, y_dot, z_dot]

    # Choose random starting points, uniformly distributed from -10 to 10
    np.random.seed(1)
    x0 = -15 + 30 * np.random.random((N, 3))

    # Integrate
    t = np.linspace(0, max_time, N)
    sol = integrate.odeint(lorenz_deriv, x0, t)

    # Plot
    fig = plt.figure()
    ax = fig.add_axes([0, 0, 1, 1], projection='3d')
    ax.set_xlim([-35, 35])
    ax.set_ylim([-35, 35])
    ax.set_zlim([0, 55])
    ax.set_xlabel('x')
    ax.set_ylabel('y')
    ax.set_zlabel('z')
    ax.view_init(30, 30)
    ax.plot(sol[:, 0], sol[:, 1], sol[:, 2], c='k')
```

JupyterLab will eventually replace the classic Jupyter Notebook

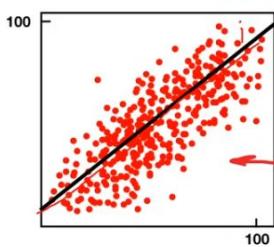
💡 Regression is the process of finding a function that relates to a labeled data set.

▷ 29:13

Regression

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Regression is a process of finding a function to correlate a labeled dataset into continuous variable/number.



Outcome: Predict this variable in the future

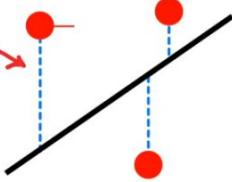


What will the temperature be next week?
eg. 20°C

Vectors (dots) are plotted on a graph in multiple dimensions eg (X,Y)
A regression line is drawn through the dataset.

The distance of the vector from the regression line called an Error
Different Regression algorithms use the error to predict future variables:

- Mean squared error (MSE)
- Root mean squared error (RMSE)
- Mean absolute error (MAE)



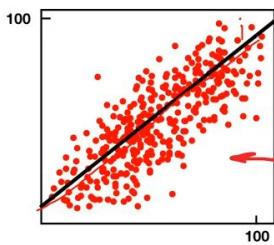
- ✿ The process of regression involves predicting a continuous variable using multiple dimensions represented by vectors plotted on a graph, ultimately resulting in a regression line.

▷ 29:29

Regression

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Regression is a process of finding a function to correlate a labeled dataset into continuous variable/number.



Outcome: Predict this variable in the future

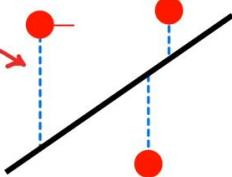


What will the temperature be next week?
eg. 20°C

Vectors (dots) are plotted on a graph in multiple dimensions eg (X,Y)
A regression line is drawn through the dataset.

The distance of the vector from the regression line called an Error
Different Regression algorithms use the error to predict future variables:

- Mean squared error (MSE)
- Root mean squared error (RMSE)
- Mean absolute error (MAE)



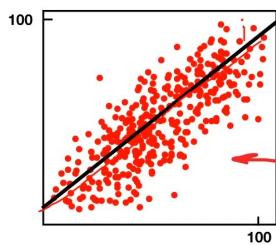
- ✿ The actual distance from a line will be used in a machine learning model to assess similarity with other lines through metrics such as mean squared error, root mean squared error, and mean absolute error.

▷ 30:40

Regression

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Regression is a process of finding a function to correlate a labeled dataset into continuous variable/number.



Outcome: Predict this **variable** in the **future**



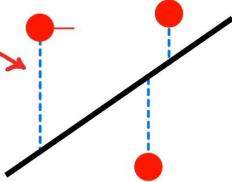
What will the temperature be next week?
eg. 20°C

Vectors (dots) are plotted on a graph in multiple dimensions eg (X,Y)
A regression line is drawn through the dataset.

The **distance** of the vector from the regression line called an **Error**

Different **Regression algorithms** use the error to predict future variables:

- Mean squared error (MSE)
- Root mean squared error (RMSE)
- Mean absolute error (MAE)



- 💡 The process involves comparing lines to determine similarity using metrics like mean squared error, root mean squared error, and mean absolute error, while classification aims to find a function that separates a labeled data set.

▷ 30:45

Classification

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

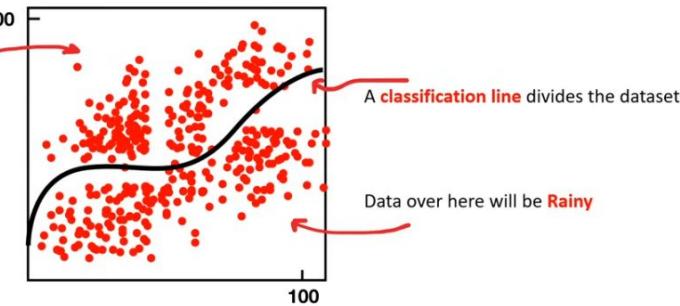
Classification is a process of finding a function to divide a labeled dataset into classes/categories

Outcome: Predict **category** to apply to the inputted data



Will it rain next Saturday?
eg. **Sunny, Rainy**

Data over here will be **Sunny**



A **classification line** divides the dataset

Data over here will be **Rainy**

Classification Algorithms

- Logistic Regression
- Decision Tree/Random Forest
- Neural Networks
- Naive Bayes
- K-Nearest Neighbors
- Support Vector Machines

- 💡 Classification algorithms categorize input data into classes, using methods like logistic regression, decision trees, random forests, and neural networks to determine which side of a dividing line the data falls on.

▷ 31:04

Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

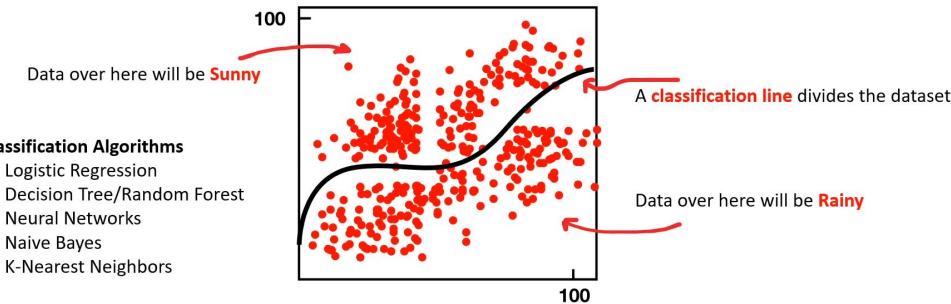
Classification is a process of finding a function to **divide a labeled dataset into classes/categories**

Outcome: Predict **category** to apply to the inputted data



Will it rain next Saturday?

eg. **Sunny, Rainy**



- ❖ Classification algorithms include logistic regression, decision trees, random forests, neural networks, naive Bayes, K-nearest neighbors (KNN), and support vector machines (SVMs), while clustering involves grouping unlabeled data in unsupervised learning.

▷ 31:39

Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

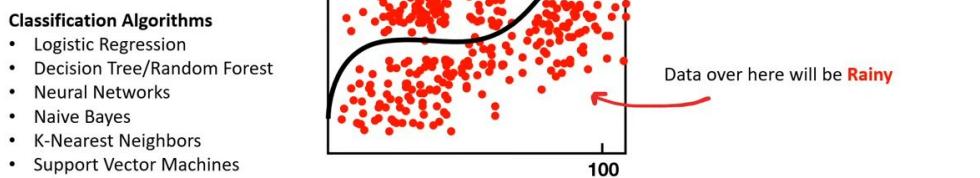
Classification is a process of finding a function to **divide a labeled dataset into classes/categories**

Outcome: Predict **category** to apply to the inputted data



Will it rain next Saturday?

eg. **Sunny, Rainy**



- ❖ Classification algorithms include logistic regression, decision trees, random forests, neural networks, naive Bayes, K-nearest neighbors (KNN), and support vector machines (SVMs), while clustering involves grouping unlabeled data based on similarities and differences.

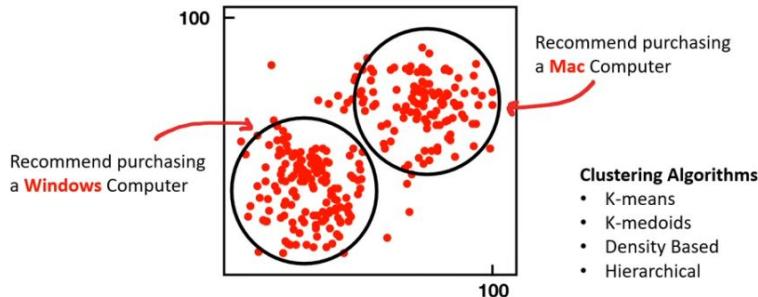
▷ 31:43

Clustering

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Clustering is a process **grouping unlabeled data based on similarities and differences.**

Outcome: Group data based on their similarities or differences



- 💡 Clustering algorithms, such as K-means, K-medoids, and hierarchical methods, are used to group unlabeled data based on similarities or differences.

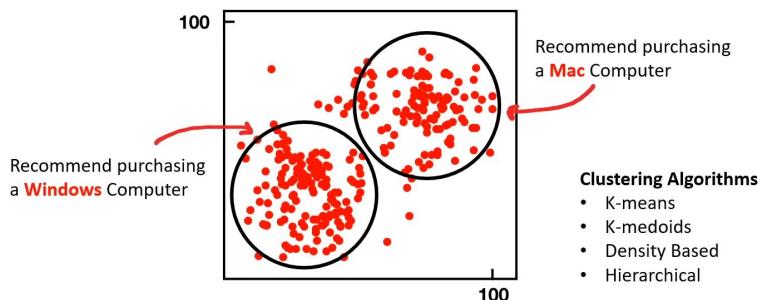
▷ 31:57

Clustering

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Clustering is a process **grouping unlabeled data based on similarities and differences.**

Outcome: Group data based on their similarities or differences



- 💡 The discussion involves clustering algorithms and the use of a confusion matrix to visualize model predictions against actual labels.

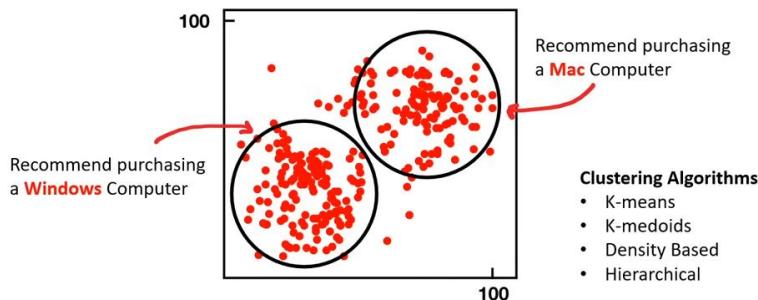
▷ 32:27

Clustering

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Clustering is a process **grouping unlabeled data based on similarities and differences.**

Outcome: Group data based on their similarities or differences



- 💡 Clustering algorithms like K-means, K-medoids, and hierarchical methods are used to infer labels from unlabeled data, while a confusion matrix visualizes model predictions against actual labels.

▷ 32:28

Classification Metrics – Confusion Matrix

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

A confusion matrix is table to visualize the **model predictions** (predicted) vs **ground truth labels** (actual)

Also known as an error matrix. They are useful in classification problems



How many people ate the banana?

		Predicted NO negative	Predicted YES positive	
		75 False Negatives (FN)	25 False Positives (FP)	Our ground truth had 100 labeled items Total False (tF)
Actual NO false	Actual YES true	50 True Negatives (TN)	20 True Positives (TP)	Our model made 70 predictions Total True (tT)
	100 were NO Total Negative (tN)	75 were YES Total Positive (tP)		

The size of matrix is dependent on the labels:

Apple, Banana, Orange $3 \times 2 = 6$ cells

We have total 170 items

Total (t)

- 💡 Classification problems involve comparing predicted outcomes to actual results to evaluate the accuracy of a model's predictions.

▷ 32:42

Classification Metrics – Confusion Matrix

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

A confusion matrix is table to visualize the **model predictions** (predicted) vs **ground truth labels** (actual)

Also known as an error matrix. They are useful in classification problems



How many people ate the banana?

	Predicted NO negative	Predicted YES positive	
Actual NO false	75 False Negatives (FN)	25 False Positives (FP)	Our ground truth had 100 labeled items Total False (tF)
Actual YES true	50 True Negatives (TN)	20 True Positives (TP)	Our model made 70 predictions Total True (tT)
100 were NO Total Negative (tN)		75 were YES Total Positive (tP)	

The size of matrix is dependent on the labels:

Apple, Banana, Orange 3x2 = 6 cells

We have total 170 items

Total (t)

- A binary classifier can be extended to handle multiple labels by adding additional cells for each label, allowing for the comparison of ground truth versus predictions.

▷ 33:41

Classification Metrics – Confusion Matrix

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

A confusion matrix is table to visualize the **model predictions** (predicted) vs **ground truth labels** (actual)

Also known as an error matrix. They are useful in classification problems



How many people ate the banana?

	Predicted NO negative	Predicted YES positive	
Actual NO false	75 False Negatives (FN)	25 False Positives (FP)	Our ground truth had 100 labeled items Total False (tF)
Actual YES true	50 True Negatives (TN)	20 True Positives (TP)	Our model made 70 predictions Total True (tT)
100 were NO Total Negative (tN)		75 were YES Total Positive (tP)	

The size of matrix is dependent on the labels:

Apple, Banana, Orange 3x2 = 6 cells

We have total 170 items

Total (t)

- An anomaly is defined as an abnormal occurrence that can be identified through a comparison of actual versus predicted values.

▷ 34:01

Classification Metrics – Confusion Matrix

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

A confusion matrix is table to visualize the **model predictions** (predicted) vs **ground truth labels (actual)**

Also known as an error matrix. They are useful in classification problems



How many people ate the banana?

	Predicted NO negative	Predicted YES positive	
Actual NO false	75 False Negatives (FN)	25 False Positives (FP)	Our ground truth had 100 labeled items Total False (tF)
Actual YES true	50 True Negatives (TN)	20 True Positives (TP)	Our model made 70 predictions Total True (tT)
100 were NO Total Negative (tN)	75 were YES Total Positive (tP)		

The size of matrix is dependent on the labels:

Apple, Banana, Orange 3x2 = 6 cells

We have total 170 items

Total (t)

- An anomaly is defined as an abnormal occurrence characterized by a deviation from the norm or standard.

▷ 34:04

Anomaly Detection AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is an anomaly?

An abnormal thing; a marked deviation from the norm or a standard

What is anomaly detection?

Anomaly Detection is the process of finding outliers within a dataset called an **anomaly**

Detecting when a piece of data or access patterns appear suspicious or malicious

Use cases for anomaly detection

- Data cleaning
- **Intrusion detection**
- **Fraud detection**
- Systems health monitoring
- Event detection in sensor networks
- Ecosystem disturbances
- Detection of critical and cascading flaws

Anomaly detection by hand is a very tedious process.

Using machine learning for anomaly detection is more efficient and accurate



Anomaly detector Detect anomalies in data to quickly identify and troubleshoot issues.

▷ 34:28

Azure has the service of anomaly detection to detect the anomaly

▷ 34:58

Anomaly Detection AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is an anomaly?

An abnormal thing; a marked deviation from the norm or a standard

What is anomaly detection?

Anomaly Detection is the process of finding outliers within a dataset called an **anomaly**

Detecting when a piece of data or access patterns appear suspicious or malicious

Use cases for anomaly detection

- Data cleaning
- **Intrusion detection**
- **Fraud detection**
- Systems health monitoring
- Event detection in sensor networks
- Ecosystem disturbances
- Detection of critical and cascading flaws

Anomaly detection by hand is a very tedious process.

Using machine learning for anomaly detection is more efficient and accurate



Anomaly detector Detect anomalies in data to quickly identify and troubleshoot issues.

- Machine learning, particularly through services like Azure's anomaly detector, enhances the efficiency and accuracy of detecting anomalies in data, while computer vision utilizes deep learning algorithms for advanced analysis of images and videos.

▶ 34:58

Anomaly Detection AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

What is an anomaly?

An abnormal thing; a marked deviation from the norm or a standard

What is anomaly detection?

Anomaly Detection is the process of finding outliers within a dataset called an **anomaly**

Detecting when a piece of data or access patterns appear suspicious or malicious

Use cases for anomaly detection

- Data cleaning
- **Intrusion detection**
- **Fraud detection**
- Systems health monitoring
- Event detection in sensor networks
- Ecosystem disturbances
- Detection of critical and cascading flaws

Anomaly detection by hand is a very tedious process.

Using machine learning for anomaly detection is more efficient and accurate



Anomaly detector Detect anomalies in data to quickly identify and troubleshoot issues.

- Machine learning, particularly through services like Azure's anomaly detector, enhances the efficiency and accuracy of detecting anomalies in data compared to manual methods, while deep learning algorithms are utilized in computer vision for high-level understanding of images and videos.

▶ 34:58

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Computer Vision is when we use Machine Learning Neural Networks to
gain high-level understanding from digital images or video

Computer Vision Deep Learning Algorithms:

- **Convolutional neural network (CNN)** — image and video recognition
 - Inspired after how human eyes actually process information and send it back to brain to be processed
- **Recurrent neural network (RNN)** — handwriting recognition or speech recognition

Types of Computer Vision

- **Image Classification** — look at an image or video and classify (place it in a category)
- **Object Detection** — identify objects within an image or video and apply labels and location boundaries
- **Semantic Segmentation** — identify segments or objects by drawing pixel mask (great for objects in movement)
- **Image Analysis** — analyze a and image or video to apply descriptive and context labels
 - eg. An employee sitting at a Desk in Tokyo
- **Optical Character Recognition** — Find text in images or videos and extract them into digital text for editing
- **Facial Detection** — detect faces in a photo or video, draw a location boundary, label their expression

💡 Deep learning algorithms include convolutional neural networks for image and video recognition and recurrent neural networks for handwriting and speech recognition, with common applications in image classification and object detection.

▷ 35:10

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Computer Vision is when we use Machine Learning Neural Networks to
gain high-level understanding from digital images or video

Computer Vision Deep Learning Algorithms:

- **Convolutional neural network (CNN)** — image and video recognition
 - Inspired after how human eyes actually process information and send it back to brain to be processed
- **Recurrent neural network (RNN)** — handwriting recognition or speech recognition

Types of Computer Vision

- **Image Classification** — look at an image or video and classify (place it in a category)
- **Object Detection** — identify objects within an image or video and apply labels and location boundaries
- **Semantic Segmentation** — identify segments or objects by drawing pixel mask (great for objects in movement)
- **Image Analysis** — analyze a and image or video to apply descriptive and context labels
 - eg. An employee sitting at a Desk in Tokyo
- **Optical Character Recognition** — Find text in images or videos and extract them into digital text for editing
- **Facial Detection** — detect faces in a photo or video, draw a location boundary, label their expression

💡 Image analysis technologies, such as optical character recognition and facial detection, can extract text from images and videos while identifying faces and their expressions.

▷ 36:01

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Computer Vision is when we use Machine Learning Neural Networks to
gain high-level understanding from digital images or video

Computer Vision Deep Learning Algorithms:

- **Convolutional neural network (CNN)** — image and video recognition
 - Inspired after how human eyes actually process information and send it back to brain to be processed
- **Recurrent neural network (RNN)** — handwriting recognition or speech recognition

Types of Computer Vision

- **Image Classification** — look at an image or video and classify (place it in a category)
- **Object Detection** — identify objects within an image or video and apply labels and location boundaries
- **Semantic Segmentation** — identify segments or objects by drawing pixel mask (great for objects in movement)
- **Image Analysis** — analyze a and image or video to apply descriptive and context labels
 - eg. An employee sitting at a Desk in Tokyo
- **Optical Character Recognition** — Find text in images or videos and extract them into digital text for editing
- **Facial Detection** — detect faces in a photo or video, draw a location boundary, label their expression

💡 Image analysis techniques such as optical character recognition and facial detection can extract text from images and identify faces along with their expressions, exemplified by Microsoft's Seeing AI app for iOS.

▶ 36:12

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Computer Vision is when we use Machine Learning Neural Networks to
gain high-level understanding from digital images or video

Computer Vision Deep Learning Algorithms:

- **Convolutional neural network (CNN)** — image and video recognition
 - Inspired after how human eyes actually process information and send it back to brain to be processed
- **Recurrent neural network (RNN)** — handwriting recognition or speech recognition

Types of Computer Vision

- **Image Classification** — look at an image or video and classify (place it in a category)
- **Object Detection** — identify objects within an image or video and apply labels and location boundaries
- **Semantic Segmentation** — identify segments or objects by drawing pixel mask (great for objects in movement)
- **Image Analysis** — analyze a and image or video to apply descriptive and context labels
 - eg. An employee sitting at a Desk in Tokyo
- **Optical Character Recognition** — Find text in images or videos and extract them into digital text for editing
- **Facial Detection** — detect faces in a photo or video, draw a location boundary, label their expression

💡 Image analysis techniques such as optical character recognition and facial detection can extract text from images and videos, identify faces, draw boundaries around them, label expressions, and are exemplified by Microsoft's Seeing AI app for iOS.

▶ 36:16

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Seeing AI is an AI app developed by Microsoft for iOS
Seeing AI uses the device camera to **identify people and objects**, and then the app audibly **describes those objects for people with visual impairment**.

Azure's Computer Vision Service Offering:

-  **Computer Vision** analyze images and video, and extract descriptions, tags, objects, and text
-  **Custom Vision** custom image classification and object detection models using your own images
-  **Face** Detect and identify people and emotions in images.
-  **Form Recogniser** translate scanned documents into key /value or tabular editable data

- 💡 There are various computer vision services available that can analyze images and videos, extract descriptions and tags, create custom image classification models, detect faces and emotions, and translate scanned documents.

▷ 36:29

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Seeing AI is an AI app developed by Microsoft for iOS
Seeing AI uses the device camera to **identify people and objects**, and then the app audibly **describes those objects for people with visual impairment**.

Azure's Computer Vision Service Offering:

-  **Computer Vision** analyze images and video, and extract descriptions, tags, objects, and text
-  **Custom Vision** custom image classification and object detection models using your own images
-  **Face** Detect and identify people and emotions in images.
-  **Form Recogniser** translate scanned documents into key /value or tabular editable data

- 💡 Custom Vision enables the creation of image classification and object detection models using personal images, while natural language processing (NLP) utilizes machine learning to comprehend the context of a text corpus.

▷ 37:00

Computer Vision

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Seeing AI is an AI app developed by Microsoft for iOS
Seeing AI uses the device camera to **identify people and objects**, and then the app audibly **describes those objects for people with visual impairment**.

Azure's Computer Vision Service Offering:

-  **Computer Vision** analyze images and video, and extract descriptions, tags, objects, and text
-  **Custom Vision** custom image classification and object detection models using your own images
-  **Face** Detect and identify people and emotions in images.
-  **Form Recogniser** translate scanned documents into key /value or tabular editable data

- ❖ Custom Vision enables the creation of image classification and object detection models using personal images, while Natural Language Processing (NLP) utilizes machine learning to comprehend the context of related text.

▷ 37:02

Natural Language Processing (NLP)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Natural Language Processing is Machine Learning that can **understand the context of a corpus (a body of related text)**.

NLP enables you to:

- Analyze and interpret text within documents, email messages
- Interpret or contextualise spoken token eg sentiment analysis
- Synthesize speech eg. a voice assistance talking to you
- Automatically translate spoken or written phrases and sentences between languages.
- Interpret spoken or written commands and determine appropriate actions.



Cortana is a **virtual assistant** developed by Microsoft which uses the Bing search engine to perform tasks such as setting reminders and answering questions for the user.

- ❖ Natural language processing (NLP) enables the analysis and interpretation of text and speech, facilitating tasks such as sentiment analysis, speech synthesis, translation, command interpretation, and virtual assistance.

▷ 37:16

Natural Language Processing (NLP)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Natural Language Processing is Machine Learning that can **understand the context of a corpus (a body of related text)**.

NLP enables you to:

- Analyze and interpret text within documents, email messages
- Interpret or contextualise spoken token eg sentiment analysis
- Synthesize speech eg. a voice assistance talking to you
- Automatically translate spoken or written phrases and sentences between languages.
- Interpret spoken or written commands and determine appropriate actions.



Cortana is a **virtual assistant** developed by Microsoft which uses the Bing search engine to perform tasks such as setting reminders and answering questions for the user.

Hi, I'm Cortana.

💡 Cortana is a voice assistant that utilizes the Bing search engine to set reminders and answer questions, while Azure's MLP offering includes text analytics features like sentiment analysis and key phrase extraction.

▷ 37:50

Natural Language Processing (NLP)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Natural Language Processing is Machine Learning that can **understand the context of a corpus (a body of related text)**.

NLP enables you to:

- Analyze and interpret text within documents, email messages
- Interpret or contextualise spoken token eg sentiment analysis
- Synthesize speech eg. a voice assistance talking to you
- Automatically translate spoken or written phrases and sentences between languages.
- Interpret spoken or written commands and determine appropriate actions.



Cortana is a **virtual assistant** developed by Microsoft which uses the Bing search engine to perform tasks such as setting reminders and answering questions for the user.

Hi, I'm Cortana.

💡 Cortana is a voice assistant that utilizes the Bing search engine for tasks like setting reminders and answering questions, while Azure's machine learning platform offers text analytics features such as sentiment analysis, key phrase extraction, and language detection.

▷ 37:53

Natural Language Processing (NLP)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure's NLP Service Offering:



Text Analytics

- sentiment analysis to find out what customers think
- Find topic-relevant phrases using key phrase extraction
- identify the language of the text with language detection
- Detect and categorize entities in your text with named entity recognition



Translator

- real-time text translation
- multi-language support



Speech

- transcribe audible speech into readable, searchable text



Language Understanding (LUIS)

- natural language processing service that enables you to understand human language in your own application, website, chatbot, IoT device, and more

💡 The text describes various language processing services, including key phrase extraction, language detection, named entity recognition, real-time translation, speech transcription, and natural language understanding.

▷ 38:02

Natural Language Processing (NLP)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure's NLP Service Offering:



Text Analytics

- sentiment analysis to find out what customers think
- Find topic-relevant phrases using key phrase extraction
- identify the language of the text with language detection
- Detect and categorize entities in your text with named entity recognition



Translator

- real-time text translation
- multi-language support



Speech

- transcribe audible speech into readable, searchable text



Language Understanding (LUIS)

- natural language processing service that enables you to understand human language in your own application, website, chatbot, IoT device, and more

▷ 38:40

Natural Language Processing (NLP)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure's NLP Service Offering:



Text Analytics

- sentiment analysis to find out what customers think
- Find topic-relevant phrases using key phrase extraction
- identify the language of the text with language detection
- Detect and categorize entities in your text with named entity recognition



Translator

- real-time text translation
- multi-language support



Speech

- transcribe audible speech into readable, searchable text



Language Understanding (LUIS)

- natural language processing service that enables you to understand human language in your own application, website, chatbot, IoT device, and more

💡 Conversational AI utilizes natural language processing to enable applications, websites, chatbots, and IoT devices to understand and engage in human conversations.

▶ 38:40

Conversational AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Conversational AI is technology that can **participate in conversations with humans**.

- Chatbots
- Voice Assistants
- Interactive Voice Recognition Systems (IVRS)

Use Cases

- **Online Customer Support** — replaces human agents for replying about customer FAQs, shipping

💡 The second version of interactive voice response systems includes speech recognition capabilities that can automate customer support, assist visually impaired users, and streamline HR processes.

▶ 38:54

Conversational AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Conversational AI is technology that can **participate in conversations with humans.**

- Chatbots
- Voice Assistants
- Interactive Voice Recognition Systems (IVRS)

Use Cases

- **Online Customer Support** — replaces human agents for replying about customer FAQs, shipping
- **Accessibility** — voice operated UI for those who are visually impaired
- **HR processes** — employee training, onboarding, updating employee information
- **Health Care** — accessible and affordable health care eg. claim processes
- **Internet of Things (IoT)** — Amazon Alexa, Apple Siri and Google Home
- **Computer Software** — autocomplete search on phone or desktop



QnA Maker Create a conversational question-and-answer bot from your existing content (Knowledge base).



Azure Bot Service Intelligent, serverless bot service that scales on demand. Used for creating, publishing, and managing bots

- 💡 A conversational question and answer bot can be created using existing content and deployed through Azure's intelligent serverless bot service, while adhering to responsible AI principles that emphasize ethical, transparent, and accountable uses of AI technology.

▷ 40:13

Conversational AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Conversational AI is technology that can **participate in conversations with humans.**

- Chatbots
- Voice Assistants
- Interactive Voice Recognition Systems (IVRS)

Use Cases

- **Online Customer Support** — replaces human agents for replying about customer FAQs, shipping
- **Accessibility** — voice operated UI for those who are visually impaired
- **HR processes** — employee training, onboarding, updating employee information
- **Health Care** — accessible and affordable health care eg. claim processes
- **Internet of Things (IoT)** — Amazon Alexa, Apple Siri and Google Home
- **Computer Software** — autocomplete search on phone or desktop



QnA Maker Create a conversational question-and-answer bot from your existing content (Knowledge base).



Azure Bot Service Intelligent, serverless bot service that scales on demand. Used for creating, publishing, and managing bots

- 💡 A knowledge base can be used to create and manage intelligent bots through a serverless service that scales on demand, while responsible AI emphasizes ethical, transparent, and accountable uses of AI technology.

▷ 40:14

Responsible AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Responsible AI focuses on **ethical, transparent and accountable** use of AI technologies

Microsoft puts into practice Responsible AI via its six **Microsoft AI principles**

1. **Fairness** — AI systems should treat all people fairly
2. **Reliability and Safety** — AI systems should perform reliably and safely
3. **Privacy and Security** — AI systems should be secure and respect privacy
4. **Inclusiveness** — AI systems should empower everyone and engage people
5. **Transparency** — AI systems should be understandable
6. **Accountability** — People should be accountable for AI systems

💡 The six Microsoft AI principles emphasize fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability in the development and deployment of AI systems.

▷ 40:25

Responsible AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Responsible AI focuses on **ethical, transparent and accountable** use of AI technologies

Microsoft puts into practice Responsible AI via its six **Microsoft AI principles**

1. **Fairness** — AI systems should treat all people fairly
2. **Reliability and Safety** — AI systems should perform reliably and safely
3. **Privacy and Security** — AI systems should be secure and respect privacy
4. **Inclusiveness** — AI systems should empower everyone and engage people
5. **Transparency** — AI systems should be understandable
6. **Accountability** — People should be accountable for AI systems

💡 AI systems should be secure, respect privacy, empower everyone, ensure transparency and accountability, and treat all people fairly.

▷ 41:02

Responsible AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Responsible AI focuses on **ethical, transparent and accountable** use of AI technologies

Microsoft puts into practice Responsible AI via its six **Microsoft AI principles**

1. **Fairness** — AI systems should treat all people fairly
2. **Reliability and Safety** — AI systems should perform reliably and safely
3. **Privacy and Security** — AI systems should be secure and respect privacy
4. **Inclusiveness** — AI systems should empower everyone and engage people
5. **Transparency** — AI systems should be understandable
6. **Accountability** — People should be accountable for AI systems

💡 AI systems should be designed to ensure fairness, transparency, inclusiveness, and accountability while avoiding the reinforcement of existing social biases.

▷ 41:07

Responsible AI – Fairness

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should treat all people fairly

AI systems can reinforce existing societal stereotypical
Bias can be introduced during the development of a pipeline

AI systems that are used to allocate or withhold:	In domains:
• opportunities	• Criminal Justice
• resources	• Employment and Hiring
• Information	• Finance and Credit

e.g. an ML model designed to select final applicants for a hiring pipeline
without incorporating any bias based on gender, ethnicity or may result in an unfair advantage

Azure ML can tell you how each feature can influence a model's prediction for bias

☰ **Fairlearn** Fairlearn is an open-source python project to help data scientist to improve fairness in their AI systems

💡 Stereotypical bias can be introduced in machine learning pipelines, affecting decisions related to opportunities and resources, but tools like Azure ML can help identify how features influence model predictions for bias.

▷ 41:21

Responsible AI – Fairness

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should treat all people fairly

AI systems can reinforce existing societal stereotypical
Bias can be introduced during the development of a pipeline

- | | |
|--|-------------------------|
| AI systems that are used to
allocate or withhold: | In domains: |
| • opportunities | • Criminal Justice |
| • resources | • Employment and Hiring |
| • Information | • Finance and Credit |

eg. an ML model designed to select final applicants for a hiring pipeline
without incorporating any bias based on gender, ethnicity or may result in an unfair advantage

Azure ML can tell you how each feature can influence a model's prediction for bias

 **Fairlearn** Fairlearn is an open-source python project to help data scientist to improve fairness in their AI systems

 Fairlearn is an open-source Python project designed to help data scientists enhance fairness in AI systems, although its features were still in preview at the time of discussion.

▷ 42:00

Responsible AI – Fairness

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should treat all people fairly

AI systems can reinforce existing societal stereotypical
Bias can be introduced during the development of a pipeline

- | | |
|--|-------------------------|
| AI systems that are used to
allocate or withhold: | In domains: |
| • opportunities | • Criminal Justice |
| • resources | • Employment and Hiring |
| • Information | • Finance and Credit |

eg. an ML model designed to select final applicants for a hiring pipeline
without incorporating any bias based on gender, ethnicity or may result in an unfair advantage

Azure ML can tell you how each feature can influence a model's prediction for bias

 **Fairlearn** Fairlearn is an open-source python project to help data scientist to improve fairness in their AI systems

 Fair Learn is an open-source Python project aimed at improving fairness in AI systems, although its fairness components are still in preview, and AI systems should be designed to perform reliably and safely.

▷ 42:04

Responsible AI – Reliability and safety

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should perform reliably and safely

AI software must be **rigorous tested** to ensure they work as expected before release to the end user

If there are scenarios where AI is making mistakes its important to release a report **quantified risks and harms** to end-users so they are informed of the short-comings of an AI solution

AI where concern for reliability and safety for humans is critically important:

- Autonomous Vehicle
- AI health diagnosis, AI suggesting prescriptions
- **Autonomous Weapon Systems**

- ★ It is crucial to provide a report on quantified risks and harms of AI solutions before their release to inform end users about potential shortcomings, particularly in high-stakes applications like autonomous vehicles and health diagnostics.

▷ 42:21

Responsible AI – Reliability and safety

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should perform reliably and safely

AI software must be **rigorous tested** to ensure they work as expected before release to the end user

If there are scenarios where AI is making mistakes its important to release a report **quantified risks and harms** to end-users so they are informed of the short-comings of an AI solution

AI where concern for reliability and safety for humans is critically important:

- Autonomous Vehicle
- AI health diagnosis, AI suggesting prescriptions
- **Autonomous Weapon Systems**

- ★ Autonomous systems, including vehicles, health diagnostics, and weaponry, raise significant concerns about the potential consequences of errors in their operation.

▷ 42:53

Responsible AI – Reliability and safety

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should perform reliably and safely

AI software must be **rigorous tested** to ensure they work as expected before release to the end user

If there are scenarios where AI is making mistakes its important to release a report **quantified risks and harms** to end-users so they are informed of the short-comings of an AI solution

AI where concern for reliability and safety for humans is critically important:

- Autonomous Vehicle
- AI health diagnosis, AI suggesting prescriptions
- **Autonomous Weapon Systems**

💡 Automated weapons pose significant risks due to the potential for mistakes, and AI systems should prioritize security and respect for privacy while requiring extensive data for training.

▷ 42:59

Responsible AI – Privacy and security

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should be secure and respect privacy

AI can require vast amounts of data to train Deep Learning ML models.

The nature of the ML model may require **Personally identifiable information (PII)**

It is important that we ensure protection of user data that it is not leaked or disclosed

In some cases ML Models can be run locally on a user's device so their PII remains on their device avoiding that vulnerability

AI Security Principles to detect malicious actors:

- Data Origin and Lineage
- Data Use Internal vs External
- Data Corruption Considerations
- Anomaly detection

▷ 43:22

Responsible AI – Privacy and security

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should be secure and respect privacy

AI can require vast amounts of data to train Deep Learning ML models.

The nature of the ML model may require **Personally identifiable information (PII)**

It is important that we ensure protection of user data that it is not leaked or disclosed

In some cases ML Models can be run locally on a user's device so their PII remains on their device avoiding that vulnerability

AI Security Principles to detect malicious actors:

- Data Origin and Lineage
- Data Use Internal vs External
- Data Corruption Considerations
- Anomaly detection

- ❖ AI systems should be designed to empower and engage all users, ensuring that solutions cater to the majority by considering the needs of minority users.

▶ 43:44

Responsible AI – Privacy and security

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should be secure and respect privacy

AI can require vast amounts of data to train Deep Learning ML models.

The nature of the ML model may require **Personally identifiable information (PII)**

It is important that we ensure protection of user data that it is not leaked or disclosed

In some cases ML Models can be run locally on a user's device so their PII remains on their device avoiding that vulnerability

AI Security Principles to detect malicious actors:

- Data Origin and Lineage
- Data Use Internal vs External
- Data Corruption Considerations
- Anomaly detection

- ❖ AI systems should be designed to empower and engage all users, ensuring that solutions cater to the majority by considering the needs of minority users.

▶ 43:44

Responsible AI – Inclusiveness

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should empower everyone and engage people

If we can design AI solutions for the **minority** of users
Then we can design AI solutions for the majority of users

Minority Groups

- physical ability
- gender
- sexual orientation
- ethnicity
- other factors

▶ 43:58

Responsible AI – Inclusiveness

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should empower everyone and engage people

If we can design AI solutions for the **minority** of users
Then we can design AI solutions for the majority of users

Minority Groups

- physical ability
- gender
- sexual orientation
- ethnicity
- other factors

💡 Designing technology for individuals with specialized needs can lead to solutions that benefit everyone, emphasizing the importance of transparency and understandability in AI systems.

▶ 44:19

Responsible AI – Inclusiveness

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should empower everyone and engage people

If we can design AI solutions for the **minority** of users
Then we can design AI solutions for the majority of users

Minority Groups

- physical ability
- gender
- sexual orientation
- ethnicity
- other factors

💡 Designing technology for minority groups, such as those who are deaf and blind, can lead to solutions that benefit everyone, emphasizing the importance of transparency and understandability in AI systems.

▷ 44:24

Responsible AI – Transparency

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should be understandable

Interpretability / Intelligibility is when end-users can understand the behaviour of the UI

Transparency of AI systems can result in

- Mitigating unfairness
- Help developers debug their AI systems
- Gaining more trust from our users

Those build AI systems should be:

- open about the why they are using AI
- open about the limitations of their AI systems

Adopting an open-source AI framework can provide transparency (at least from a technical perspective) on the internal workings of an AI systems

▷ 44:36

Responsible AI – Transparency

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should be understandable

Interpretability / Intelligibility is when end-users can understand the behaviour of the UI

Transparency of AI systems can result in

- Mitigating unfairness
- Help developers debug their AI systems
- Gaining more trust from our users

Those build AI systems should be:

- open about the why they are using AI
- open about the limitations of their AI systems

Adopting an open-source AI framework can provide transparency (at least from a technical perspective) on the internal workings of an AI systems

- 💡 Developers of AI systems should be transparent about their use and limitations, and adopting open-source frameworks can enhance trust and understanding of the technology.

▶ 44:53

Responsible AI – Transparency

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

AI systems should be understandable

Interpretability / Intelligibility is when end-users can understand the behaviour of the UI

Transparency of AI systems can result in

- Mitigating unfairness
- Help developers debug their AI systems
- Gaining more trust from our users

Those build AI systems should be:

- open about the why they are using AI
- open about the limitations of their AI systems

Adopting an open-source AI framework can provide transparency (at least from a technical perspective) on the internal workings of an AI systems

- 💡 Users who build AI systems should be transparent about their purposes and limitations, adopt open-source frameworks for technical clarity, and ensure accountability in the implementation of AI principles.

▶ 44:59

Responsible AI – Accountability

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

People should be accountable for AI systems

The structure put in place to consistently enacting AI principles and taking them into account

AI systems should work within:

- framework of governance
- organizational principles

ethical and legal standards
that are clearly defined

Principles guide Microsoft on how they **Develop, Sell and Advocate** when working with third-parties and this can push towards regulations towards AI Principles

▶ 45:13

Responsible AI – Accountability

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

People should be accountable for AI systems

The structure put in place to consistently enacting AI principles and taking them into account

AI systems should work within:

- framework of governance
- organizational principles

ethical and legal standards
that are clearly defined

Principles guide Microsoft on how they **Develop, Sell and Advocate** when working with third-parties and this can push towards regulations towards AI Principles

- ★ The emphasis is on the need for more well-developed ethical and legal standards in organizational principles, particularly as Microsoft advocates for its model to be adopted by others.

▶ 45:29

Responsible AI – Accountability

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

People should be accountable for AI systems

The structure put in place to consistently enacting AI principles and taking them into account

AI systems should work within:

- framework of governance
- organizational principles

ethical and legal standards
that are clearly defined

Principles guide Microsoft on how they **Develop, Sell and Advocate** when working with third-parties and this can push towards regulations towards AI Principles

- 💡 Legal standards guide the development, sale, and advocacy of Microsoft in collaboration with third parties, promoting a regulatory model that could benefit from further refinement.

▷ 45:32

Responsible AI – Accountability

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

People should be accountable for AI systems

The structure put in place to consistently enacting AI principles and taking them into account

AI systems should work within:

- framework of governance
- organizational principles

ethical and legal standards
that are clearly defined

Principles guide Microsoft on how they **Develop, Sell and Advocate** when working with third-parties and this can push towards regulations towards AI Principles

- 💡 A free web app with color-coded cards provides practical scenarios to help users understand and apply Microsoft AI principles.

▷ 45:37

Guidelines for Human-AI Interaction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Microsoft has a free web-app that goes through **practical scenarios** to teach Microsoft AI Principles



<https://aidemos.microsoft.com/guidelines-for-human-ai-interaction/demo>

- A free web app provides color-coded cards to help users understand and apply Microsoft AI principles through practical scenarios.

▷ 46:01

Guidelines for Human-AI Interaction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Microsoft has a free web-app that goes through **practical scenarios** to teach Microsoft AI Principles



<https://aidemos.microsoft.com/guidelines-for-human-ai-interaction/demo>

- A free web app provides a tool with 18 color-coded cards to help users understand and apply Microsoft AI principles in practical scenarios.

▷ 46:02



Azure AI Fundamentals

Responsible AI

Follow Along Guidelines for Human AI Interaction



Follow Along

The screenshot shows a digital interface for 'Follow Along Guidelines for Human AI Interaction'. A yellow card titled '8 WHEN WRONG Support efficient dismissal.' is displayed. Below it, a section titled 'EXAMPLE IN PRACTICE' shows a screenshot of Microsoft Forms with instructions on how to dismiss or ignore undesired AI system services. At the bottom, there are four small cards: blue, red, yellow, and green, each with a title and some text.

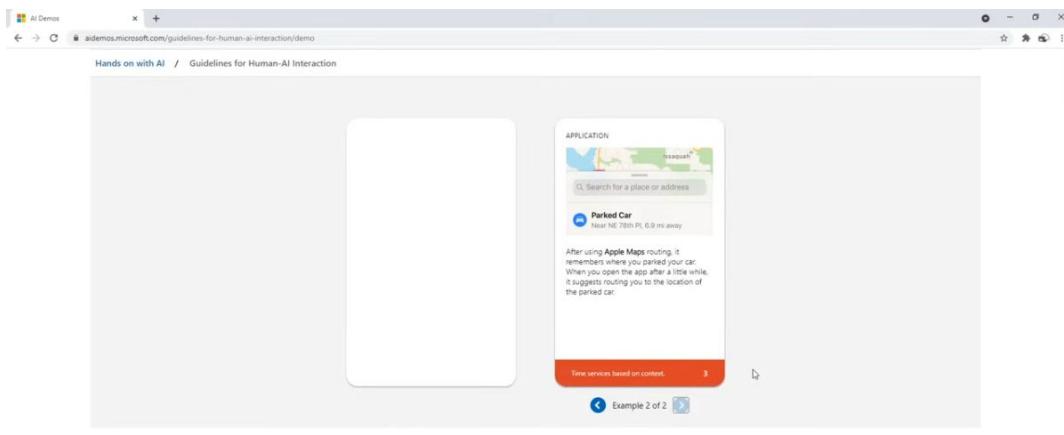
- A free web app provides a tool with color-coded cards for practical scenarios related to human-AI interaction, based on Microsoft AI principles.

▷ 46:07

The screenshot shows a digital interface for 'Follow Along Guidelines for Human AI Interaction'. A white card titled 'Does Focused inbox organize all the mail in my inbox?' is displayed. Below it, a section titled 'EXAMPLE IN PRACTICE' shows a screenshot of Outlook Web Mail with instructions on how to teach Outlook what matters most to you. At the bottom, there are four small cards: blue, red, yellow, and green, each with a title and some text.

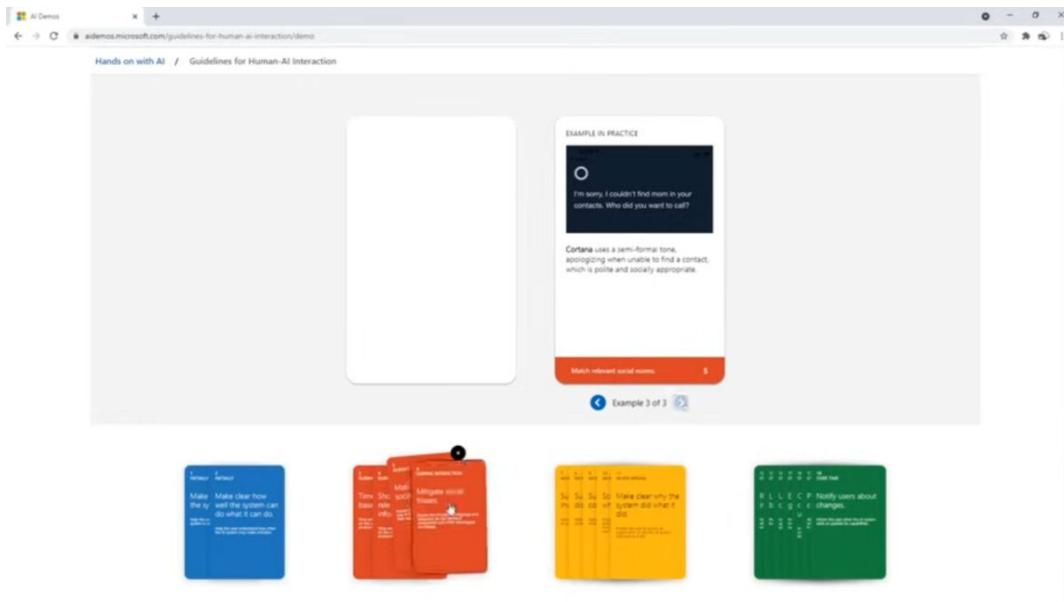
- The text discusses the use of language and features in technology products to enhance user experience, manage expectations, and communicate uncertainty while emphasizing continuous improvement through usage.

▷ 46:54



- Microsoft Outlook provides time-to-leave notifications with directions based on real-time traffic and location, while Apple Maps remembers where users parked their cars and suggests routes back to them, leading to speculation about a potential partnership between Microsoft and Apple.

▷ 47:38

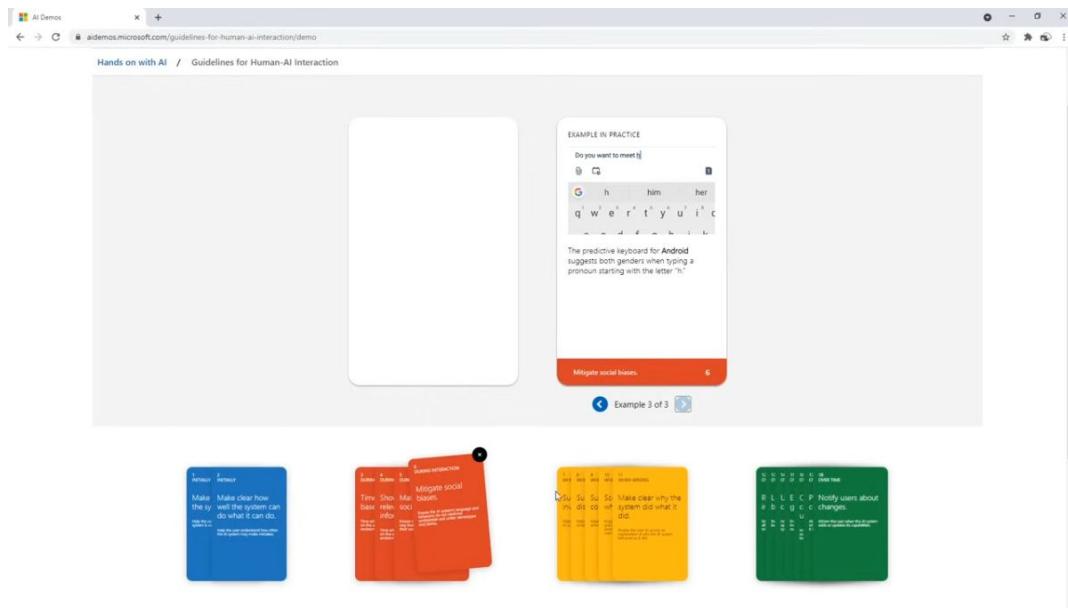


- Machine learning can enhance user experience by providing timely recommendations and relevant information based on the user's current task and environment.

▷ 48:10

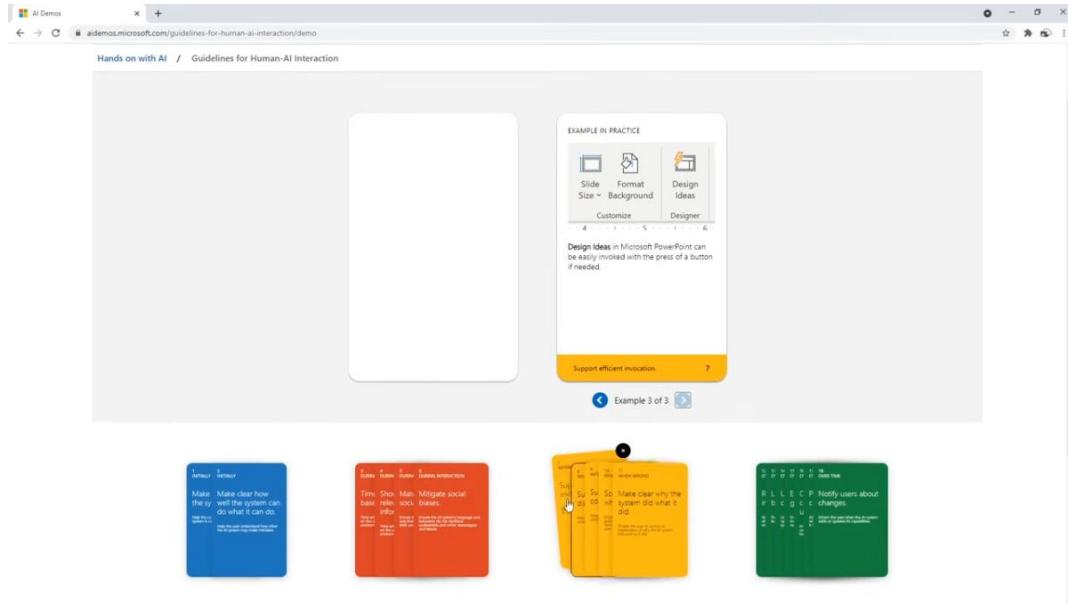
Need to go through the video unable to get the correct URL

▷ 49:20



- ❖ Using gender-neutral icons and diverse imagery can help mitigate bias in professional representations.

▷ 49:40



- ❖ There are features in various software applications that provide recommendations and design ideas, which can be easily accessed but may also become intrusive for some users.

▷ 50:14

A screenshot of a web browser window titled "AI Demos" showing a specific card from the "Guidelines for Human-AI Interaction". The card is titled "EXAMPLE IN PRACTICE" and shows a screenshot of a Siri interaction where the user says "Hey Siri never mind" and Siri responds with "OK". Below the screenshot, the text reads "Siri can be easily dismissed by saying 'never mind'". At the bottom of the card, there is a yellow button labeled "Support efficient dismissal" with the number "8" next to it. Below the card, there is a circular icon with a play button symbol and the text "Example 3 of 3".



- Several AI systems and services provide users with straightforward options to dismiss or ignore unwanted suggestions, ads, or prompts.

▷ 50:52

A screenshot of a web browser window titled "AI Demos" showing a specific card from the "Guidelines for Human-AI Interaction". The card is titled "EXAMPLE IN PRACTICE" and shows a screenshot of the Google Photos app interface. The text on the screen reads "On, the places you'll see" and "Turn on location history to see photos grouped by where you've been". Below the screenshot, the text reads "Google Photos allows users to turn location history on/off for all future photos." At the bottom of the card, there is a green button labeled "Provide global controls" with the number "17" next to it. Below the card, there is a circular icon with a play button symbol and the text "Example 3 of 3".



- The effectiveness of spellchecking in Word is criticized, and there is a suggestion for improved AI-based spellchecking, while also noting that search engines like Bing and features in Google Photos offer customizable settings.

▷ 56:15

The screenshot shows a browser window with the URL aidemos.microsoft.com/guidelines-for-human-ai-interaction/demo. The page title is "Hands on with AI / Guidelines for Human-AI Interaction". A central content area displays a "EXAMPLE IN PRACTICE" dialog box from "Outlook Web" showing a "Help" screen with a "What's new" section. Below this, a green button says "Notify users about changes." with the number "18" next to it. At the bottom, a blue circular icon indicates "Example 2 of 2".



- 💡 The "What's New" dialogue in Office provides an overview of the latest features and updates, including AI enhancements in Outlook web, while the help tab also includes a section on these updates.

▷ 56:59

Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Azure Cognitive Services is a **comprehensive family of AI services** and cognitive APIs to help you build intelligent apps



Create customizable, pretrained models built with "*breakthrough*" AI research



Deploy Cognitive Services anywhere from the cloud to the edge with containers



Get started quickly—no machine-learning expertise required



Developed with strict ethical standards, empowering responsible use with industry-leading tools and guidelines

▷ 57:47



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Decision

- **Anomaly Detector** — Identify potential problems early on.
- **Content Moderator** — Detect potentially offensive or unwanted content.
- **Personaliser** — Create rich, personalised experiences for every user.

Language

- **Language Understanding** — Build natural language understanding into apps, bots and IoT devices.
- **QnA Maker** — Create a conversational question and answer layer over your data.
- **Text Analytics** — Detect sentiment, key phrases and named entities.
- **Translator** — Detect and translate more than 90 supported languages.

Speech

- **Speech to Text** — Transcribe audible speech into readable, searchable text.
- **Text to Speech** — Convert text to lifelike speech for more natural interfaces.
- **Speech Translation** — Integrate real-time speech translation into your apps.
- **Speaker Recognition** — Identify and verify the people speaking based on audio.

Vision

- **Computer Vision** — Analyze content in images and video.
- **Custom Vision** — Customize image recognition to fit your business needs.
- **Face** — Detect and identify people and emotions in images.

▶ 58:07



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Azure Cognitive Services is a **comprehensive family of AI services** and cognitive APIs to help you build intelligent apps



Create customizable, pretrained models built with “*breakthrough*” AI research



Deploy Cognitive Services anywhere from the cloud to the edge with containers



Get started quickly—no machine-learning expertise required



Developed with strict ethical standards, empowering responsible use with industry-leading tools and guidelines

💡 The services offered include tools for anomaly detection and content moderation, emphasizing the importance of ethical standards and responsible AI use.

▶ 58:14



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Azure Cognitive Services is a **comprehensive family of AI services** and cognitive APIs to help you build intelligent apps



Create customizable, pretrained models built with “*breakthrough*” AI research



Deploy Cognitive Services anywhere from the cloud to the edge with containers



Get started quickly—no machine-learning expertise required



Developed with strict ethical standards, empowering responsible use with industry-leading tools and guidelines

- Containers enable quick development without requiring machine learning expertise, although some background knowledge is beneficial, and there are various services available for responsible AI use, including anomaly detection and content moderation.

▷ 58:17



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Decision

- **Anomaly Detector** — Identify potential problems early on.
- **Content Moderator** — Detect potentially offensive or unwanted content.
- **Personaliser** — Create rich, personalised experiences for every user.

Language

- **Language Understanding** — Build natural language understanding into apps, bots and IoT devices.
- **QnA Maker** — Create a conversational question and answer layer over your data.
- **Text Analytics** — Detect sentiment, key phrases and named entities.
- **Translator** — Detect and translate more than 90 supported languages.

Speech

- **Speech to Text** — Transcribe audible speech into readable, searchable text.
- **Text to Speech** — Convert text to lifelike speech for more natural interfaces.
- **Speech Translation** — Integrate real-time speech translation into your apps.
- **Speaker Recognition** — Identify and verify the people speaking based on audio.

Vision

- **Computer Vision** — Analyze content in images and video.
- **Custom Vision** — Customize image recognition to fit your business needs.
- **Face** — Detect and identify people and emotions in images.

▷ 59:32



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Decision

- **Anomaly Detector** — Identify potential problems early on.
- **Content Moderator** — Detect potentially offensive or unwanted content.
- **Personaliser** — Create rich, personalised experiences for every user.

Language

- **Language Understanding** — Build natural language understanding into apps, bots and IoT devices.
- **QnA Maker** — Create a conversational question and answer layer over your data.
- **Text Analytics** — Detect sentiment, key phrases and named entities.
- **Translator** — Detect and translate more than 90 supported languages.

Speech

- **Speech to Text** — Transcribe audible speech into readable, searchable text.
- **Text to Speech** — Convert text to lifelike speech for more natural interfaces.
- **Speech Translation** — Integrate real-time speech translation into your apps.
- **Speaker Recognition** — Identify and verify the people speaking based on audio.

Vision

- **Computer Vision** — Analyze content in images and video.
- **Custom Vision** — Customize image recognition to fit your business needs.
- **Face** — Detect and identify people and emotions in images.

💡 Azure Cognitive Services offers a range of AI services, including computer vision for analyzing images and videos, custom image recognition tailored to business needs, and the ability to detect and identify people and emotions in images.

▷ 59:36



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Decision

- **Anomaly Detector** — Identify potential problems early on.
- **Content Moderator** — Detect potentially offensive or unwanted content.
- **Personaliser** — Create rich, personalised experiences for every user.

Language

- **Language Understanding** — Build natural language understanding into apps, bots and IoT devices.
- **QnA Maker** — Create a conversational question and answer layer over your data.
- **Text Analytics** — Detect sentiment, key phrases and named entities.
- **Translator** — Detect and translate more than 90 supported languages.

Speech

- **Speech to Text** — Transcribe audible speech into readable, searchable text.
- **Text to Speech** — Convert text to lifelike speech for more natural interfaces.
- **Speech Translation** — Integrate real-time speech translation into your apps.
- **Speaker Recognition** — Identify and verify the people speaking based on audio.

Vision

- **Computer Vision** — Analyze content in images and video.
- **Custom Vision** — Customize image recognition to fit your business needs.
- **Face** — Detect and identify people and emotions in images.

💡 Azure Cognitive Services offers a range of AI capabilities, including custom image recognition and emotion detection in images, accessible through a single API key and endpoint.

▷ 59:39



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

Cognitive Services is an umbrella AI service that enables customers to
access multiple AI services with an API key and an API Endpoint

The screenshot shows the Azure Cognitive Services management interface. On the left, a list view displays a single resource named "myCognitiveServices734". On the right, a detailed view shows the "Show Keys" section with two keys (KEY 1 and KEY 2) and their corresponding endpoints. The endpoint for both keys is set to <https://mycognitiveservices734.cognitiveservices.azure.com/>. The location is set to "eastus". Red arrows highlight the resource name in the list and the endpoint field in the details view.

▶ 59:57



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

Cognitive Services is an umbrella AI service that enables customers to
access multiple AI services with an API key and an API Endpoint

The screenshot shows the Azure Cognitive Services management interface. On the left, a list view displays a single resource named "myCognitiveServices734". On the right, a detailed view shows the "Show Keys" section with two keys (KEY 1 and KEY 2) and their corresponding endpoints. The endpoint for both keys is set to <https://mycognitiveservices734.cognitiveservices.azure.com/>. The location is set to "eastus". Red arrows highlight the resource name in the list and the endpoint field in the details view.

- Azure Cognitive Services provides access to multiple AI services through a single API key and endpoint, which are essential for authentication.

▶ 59:59



Azure Cognitive Services

Cheat sheets, Practice Exams and Flash cards [👉 www.exampro.co/ai-900](http://www.exampro.co/ai-900)

Cognitive Services is an umbrella AI service that enables customers to access multiple AI services with an **API key and an API Endpoint**

The screenshot shows the Azure Cognitive Services management interface. On the left, a list view displays a single resource named 'myCognitiveServices734'. On the right, a detailed view shows the 'Show Keys' section with two keys (KEY 1 and KEY 2) and their corresponding endpoints. A red arrow points from the text 'access multiple AI services' to the 'Subscription' filter in the list view. Another red arrow points from the text 'with an API key and an API Endpoint' to the 'Endpoint' field in the details view.

Key	Value
KEY 1	[REDACTED]
KEY 2	[REDACTED]
Endpoint	https://mycognitiveservices734.cognitiveservices.azure.com/
Location	eastus

- AI services require an authentication key and API endpoint for access, enabling knowledge mining to efficiently analyze large volumes of information.

▶ 1:00:06

Knowledge Mining

Cheat sheets, Practice Exams and Flash cards [👉 www.exampro.co/ai-900](http://www.exampro.co/ai-900)

Knowledge mining is a **discipline** in AI that uses a **combination of intelligent services to quickly learn from vast amounts of information**.

It allows organizations to deeply understand and easily explore information, uncover hidden insights, and find relationships and patterns at scale.

Ingest

Enrich

Explore

- The process involves ingesting content from various structured and unstructured sources, enriching it with AI capabilities, and exploring the information to uncover insights and identify patterns.

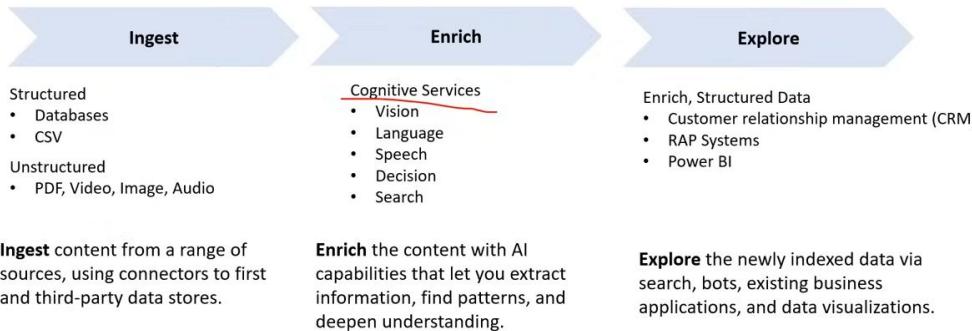
▶ 1:00:21

Knowledge Mining

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Knowledge mining is a **discipline** in AI that uses a **combination of intelligent services to quickly learn from vast amounts of information.**

It allows organizations to deeply understand and easily explore information, uncover hidden insights, and find relationships and patterns at scale.



💡 Azure's cognitive services provide solutions for efficiently conducting content research by streamlining the review of dense technical data.

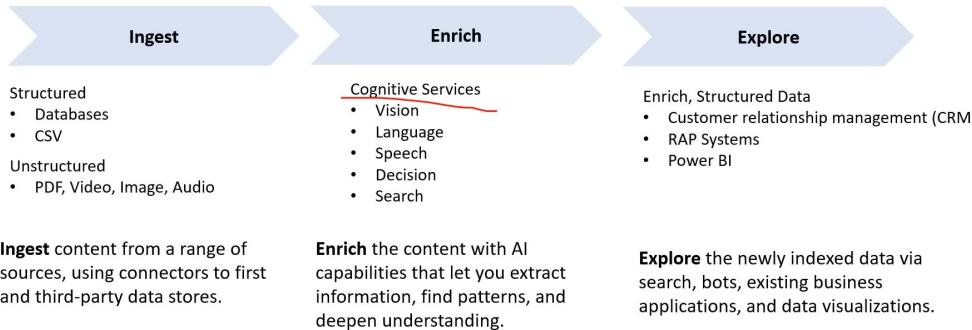
▶ 1:01:23

Knowledge Mining

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Knowledge mining is a **discipline** in AI that uses a **combination of intelligent services to quickly learn from vast amounts of information.**

It allows organizations to deeply understand and easily explore information, uncover hidden insights, and find relationships and patterns at scale.



💡 Knowledge mining leverages Azure's cognitive services to streamline the review and research of technical data, making it easier for employees to process dense information efficiently.

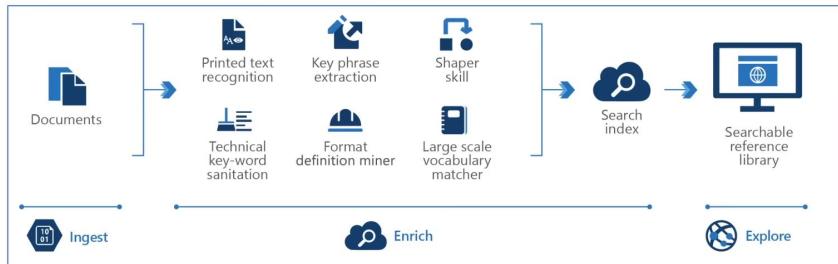
▶ 1:01:25

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Content research

When organizations task employees to review and research of technical data, it can be tedious to read page after page of dense text. Knowledge mining helps employees quickly review these dense materials.



- ✿ Azure's cognitive services enable efficient content research by utilizing knowledge mining to help employees quickly review dense technical documents through processes like printed text recognition.

▷ 1:01:31

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Content research

When organizations task employees to review and research of technical data, it can be tedious to read page after page of dense text. Knowledge mining helps employees quickly review these dense materials.



- ✿ Knowledge mining enhances the review of dense texts by utilizing techniques such as printed text recognition, key phrase extraction, and vocabulary matching to create a searchable reference library for improved audit risk compliance management.

▷ 1:01:36

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Content research

When organizations task employees to review and research of technical data, it can be tedious to read page after page of dense text. Knowledge mining helps employees quickly review these dense materials.



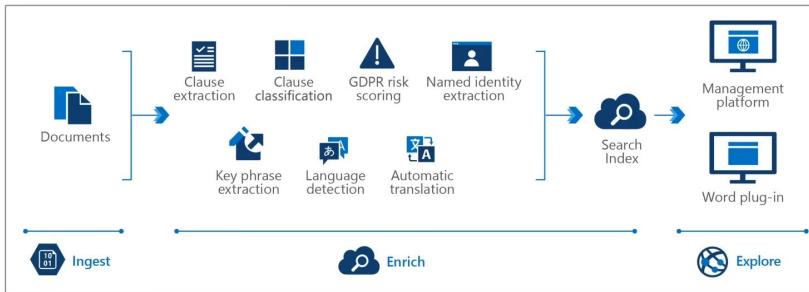
▶ 1:01:59

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Auditing, risk, and compliance management

Developers could use knowledge mining to help attorneys quickly identify entities of importance from discovery documents and flag important ideas across documents.



- 💡 A search service enables the efficient management of audit risk compliance by allowing developers to utilize knowledge mining for entity identification and keyphrase extraction in legal documents.

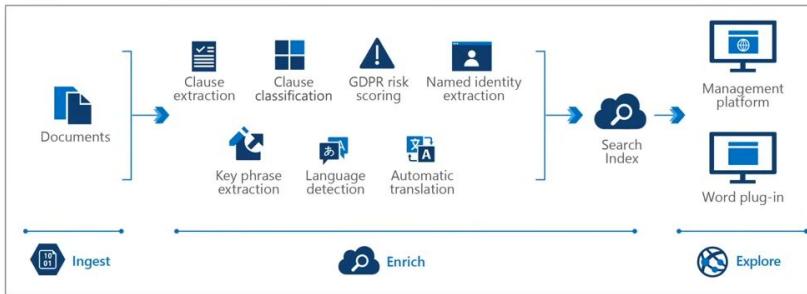
▶ 1:02:06

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Auditing, risk, and compliance management

Developers could use knowledge mining to help attorneys quickly identify entities of importance from discovery documents and flag important ideas across documents.



- ★ A search service enhances vocabulary matching and knowledge mining for developers, enabling efficient identification of important entities and ideas in documents while automating tasks like clause extraction, classification, GDPR risk assessment, named entity recognition, keyphrase extraction, and language detection.

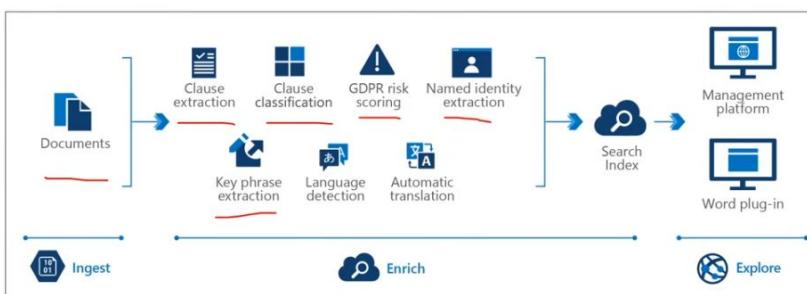
▷ 1:02:08

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Auditing, risk, and compliance management

Developers could use knowledge mining to help attorneys quickly identify entities of importance from discovery documents and flag important ideas across documents.



- ★ The process involves extracting and classifying clauses from documents, assessing GDPR risks, identifying key phrases, detecting languages, automating translations, and integrating the results into a searchable management platform for efficient business process management.

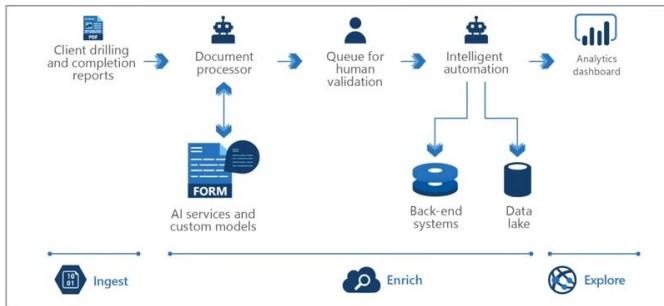
▷ 1:02:10

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Business process management

In industries where bidding competition is fierce, or when the diagnosis of a problem must be quick or in near real-time, companies can use knowledge mining to avoid costly mistakes.



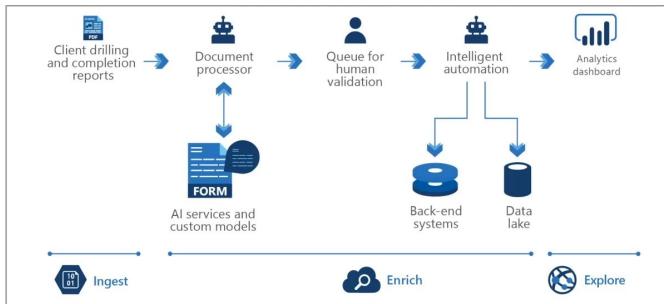
▶ 1:02:38

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Business process management

In industries where bidding competition is fierce, or when the diagnosis of a problem must be quick or in near real-time, companies can use knowledge mining to avoid costly mistakes.



- ★ Knowledge mining and AI services are utilized in business process management to enhance efficiency, facilitate quick problem diagnosis, and reduce costly errors through intelligent automation and human validation.

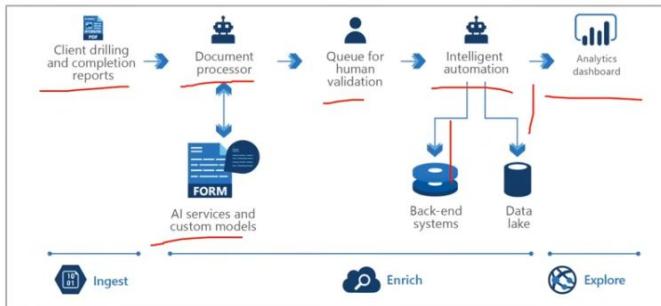
▶ 1:02:39

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Business process management

In industries where bidding competition is fierce, or when the diagnosis of a problem must be quick or in near real-time, companies can use knowledge mining to avoid costly mistakes.



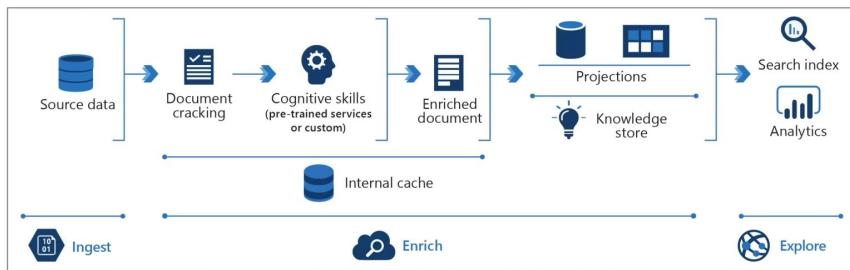
▶ 1:02:40

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Customer support and feedback analysis

For many companies, customer support is costly and inefficient. Knowledge mining can help customer support teams quickly find the right answer for a customer inquiry or assess customer sentiment at scale.



- ★ Intelligent automation can enhance customer support efficiency by utilizing knowledge mining to quickly provide answers and assess sentiment through data processing and analytics.

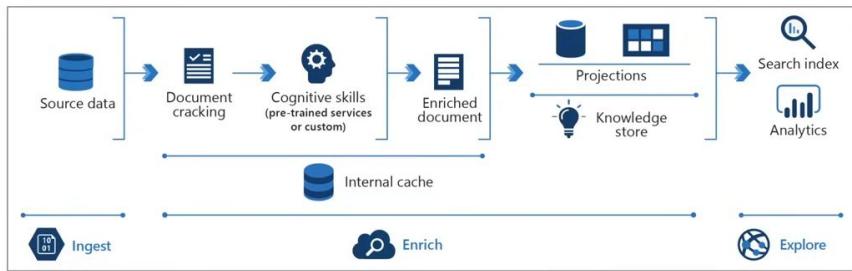
▶ 1:03:04

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Customer support and feedback analysis

For many companies, customer support is costly and inefficient. Knowledge mining can help customer support teams quickly find the right answer for a customer inquiry or assess customer sentiment at scale.



- Automation can streamline customer support by utilizing knowledge mining and cognitive skills to efficiently analyze inquiries and assess sentiment, ultimately improving analytics and reducing costs.

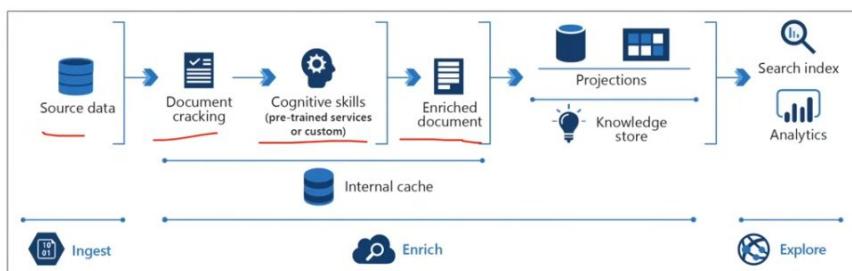
 1:03:06

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Customer support and feedback analysis

For many companies, customer support is costly and inefficient. Knowledge mining can help customer support teams quickly find the right answer for a customer inquiry or assess customer sentiment at scale.



- Cognitive services can enhance customer support by enabling quick access to answers, sentiment analysis, document enrichment, and analytics through structured data management.

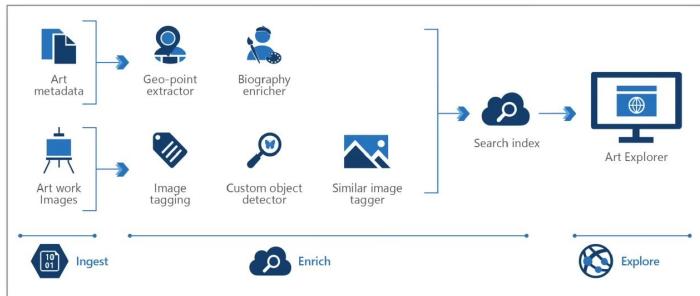
 1:03:06

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Digital asset management

Given the amount of unstructured data created daily, many companies are struggling to make use of or find information within their files. Knowledge mining through a search index makes it easy for end customers and employees to locate what they are looking for faster.



- 💡 Cognitive services enhance the ability to manage and analyze unstructured data by enabling efficient knowledge mining through search indexes, facilitating quicker information retrieval for users.

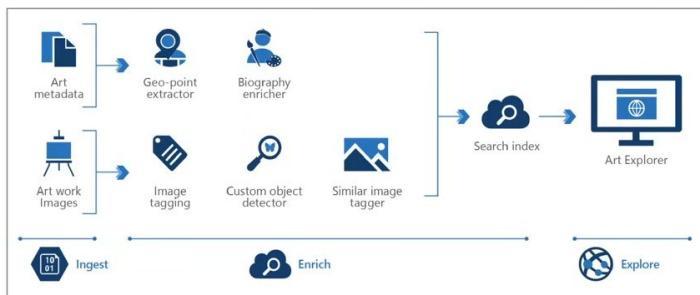
▶ 1:03:33

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Digital asset management

Given the amount of unstructured data created daily, many companies are struggling to make use of or find information within their files. Knowledge mining through a search index makes it easy for end customers and employees to locate what they are looking for faster.



- 💡 Cognitive services and knowledge mining through a search index can significantly enhance the ability of companies to manage and locate unstructured data efficiently.

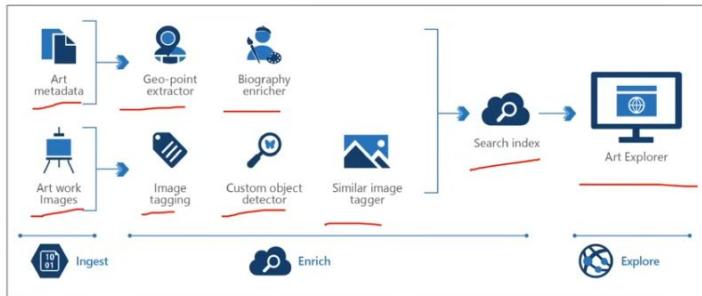
▶ 1:03:37

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards [👉 www.exampro.co/ai-900](http://www.exampro.co/ai-900)

Digital asset management

Given the amount of unstructured data created daily, many companies are struggling to make use of or find information within their files. Knowledge mining through a search index makes it easy for end customers and employees to locate what they are looking for faster.



- Many companies struggle to find information in their files, but knowledge mining through search indexes can help users locate what they need more quickly by utilizing metadata and various tagging techniques.

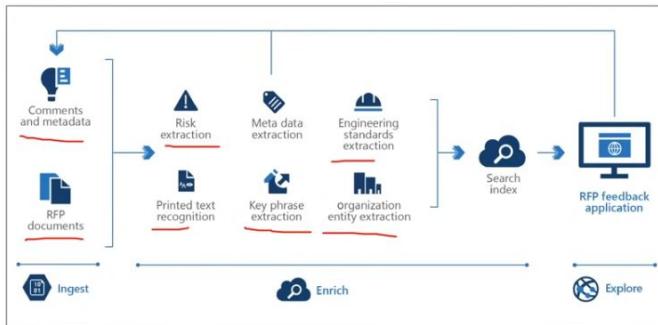
▷ 1:03:38

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards [👉 www.exampro.co/ai-900](http://www.exampro.co/ai-900)

Contract management

Many companies create products for multiple sectors, hence the business opportunities with different vendors and buyers increases exponentially. Knowledge mining can help organizations to scour thousands of pages of sources to create an accurate bid.



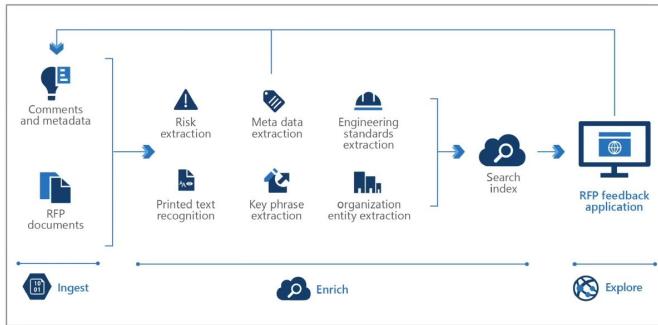
▷ 1:04:10

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Contract management

Many companies create products for multiple sectors, hence the business opportunities with different vendors and buyers increases exponentially. Knowledge mining can help organizations to scour thousands of pages of sources to create an accurate bid.



- Organizations can leverage knowledge mining to efficiently analyze extensive sources for accurate bidding in contract management across various sectors.

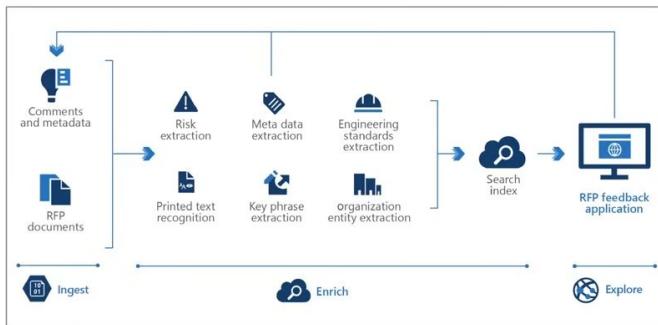
▶ 1:04:12

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Contract management

Many companies create products for multiple sectors, hence the business opportunities with different vendors and buyers increases exponentially. Knowledge mining can help organizations to scour thousands of pages of sources to create an accurate bid.



- Knowledge mining can enhance contract management by enabling organizations to analyze extensive sources for accurate bids through techniques like risk extraction, text recognition, keyphrase extraction, and organizational extraction.

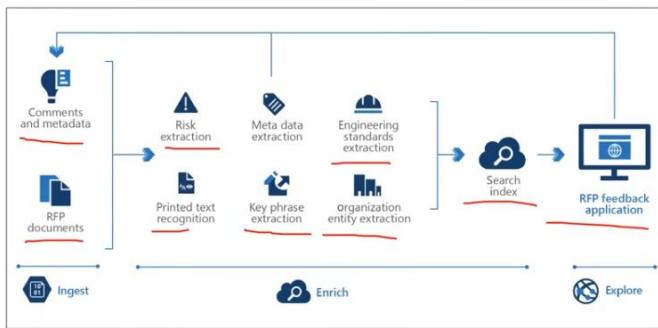
▶ 1:04:19

Knowledge Mining – Use Cases

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Contract management

Many companies create products for multiple sectors, hence the business opportunities with different vendors and buyers increases exponentially. Knowledge mining can help organizations to scour thousands of pages of sources to create an accurate bid.



💡 Azure Face Service offers an AI algorithm for detecting, recognizing, and analyzing human faces in images.

▷ 1:04:26

Face Service

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure Face service provides AI algorithms that **detect, recognize, and analyze human faces** in images

Azure Face can detect:

- faces in an image
- faces with specific attributes
- face landmarks
- similar faces
- the same face as a specific identity across a gallery of images

▷ 1:04:54



Azure Face service provides AI algorithms that **detect, recognize, and analyze human faces** in images

Azure Face can detect:

- faces in an image
- faces with specific attributes
- face landmarks
- similar faces
- the same face as a specific identity across a gallery of images

💡 Azure Face Service utilizes AI algorithms to detect, recognize, and analyze human faces in images by identifying specific attributes, landmarks, and matching identities across a gallery of images.

▶ 1:05:02



Face ID

unique identifier string for each detected face in an image

▶ 1:05:04



Azure Face service provides AI algorithms that **detect, recognize, and analyze human faces** in images

Azure Face can detect:

- faces in an image
- faces with specific attributes
- face landmarks
- similar faces
- the same face as a specific identity across a gallery of images

💡 The Azure Face Service uses AI algorithms to detect, recognize, and analyze human faces in images, providing unique identifiers for each detected face along with specific attributes and landmarks.

▶ 1:05:05



(1 faces detected)

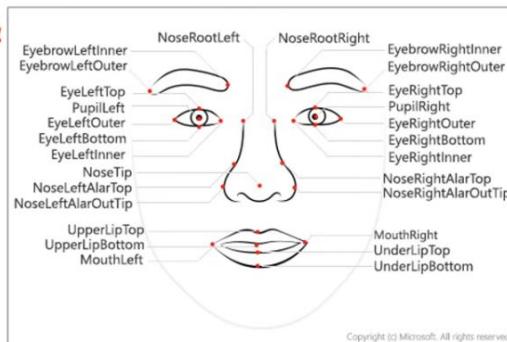


Face ID

unique identifier string for each detected face in an image

Face Landmarks

easy-to-find points on a face
27 predefined landmark points.



▶ 1:05:25



Face Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

(1 faces detected)

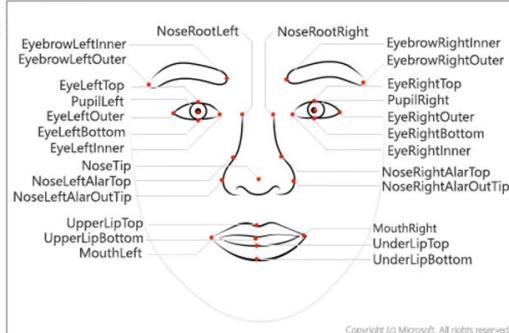


Face ID

unique identifier string for each detected face in an image

Face Landmarks

easy-to-find points on a face
27 predefined landmark points.



- The service provides unique identifiers for detected faces, identifies up to 27 predefined facial landmarks, and can determine whether individuals are wearing accessories like earrings.

▶ 1:05:29



Face Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

(1 faces detected)

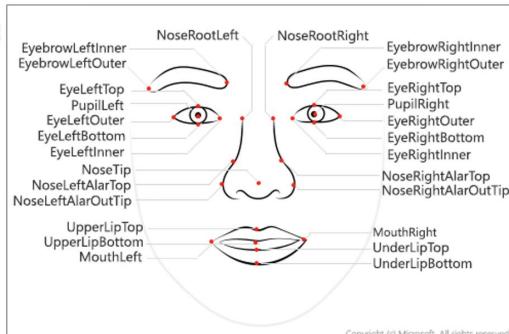


Face ID

unique identifier string for each detected face in an image

Face Landmarks

easy-to-find points on a face
27 predefined landmark points.



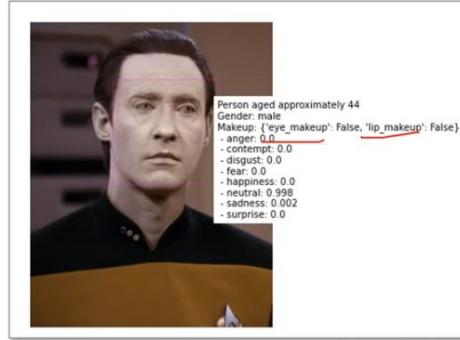
- The service can identify unique faces and their specific landmarks, assess attributes such as accessories and age, and evaluate image quality.

▶ 1:05:35



Face Attributes

- **Accessories.** (Wearing accessories)
- **Age**
- **Blur** (blurriness of the face in the image)
- **Emotion.**
- **Exposure**
- **Facial hair**
- **Gender**
- **Glasses**
- **Hair**
- **Head pose**
- **Makeup**
- **Mask.** (are they wearing a mask?)
- **Noise.** The visual noise detected in the face image
- **Occlusion.** (objects blocking parts of the face)
- **Smile**



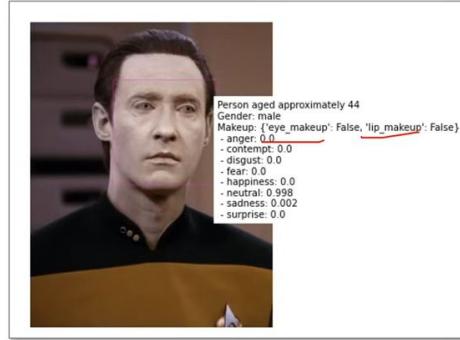
- ❖ Various factors such as lip rings, image clarity, emotional expression, exposure, contrast, facial hair, gender, glasses, hairstyle, head pose, makeup details (limited to eye and lip), mask usage, and visual artifacts contribute to the analysis of an image.

▷ 1:05:44



Face Attributes

- **Accessories.** (Wearing accessories)
- **Age**
- **Blur** (blurriness of the face in the image)
- **Emotion.**
- **Exposure**
- **Facial hair**
- **Gender**
- **Glasses**
- **Hair**
- **Head pose**
- **Makeup**
- **Mask.** (are they wearing a mask?)
- **Noise.** The visual noise detected in the face image
- **Occlusion.** (objects blocking parts of the face)
- **Smile**



- ❖ The system collects data on mask-wearing, noise levels, visual artifacts or occlusions, and includes a Boolean value indicating whether a person is smiling.

▷ 1:06:23

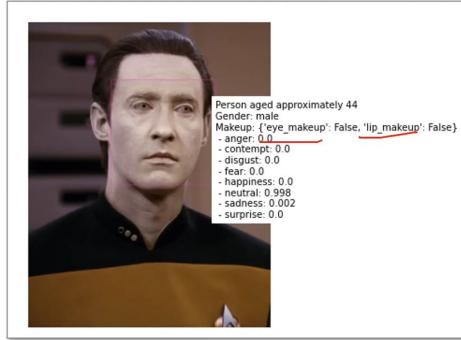


Face Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Face Attributes

- **Accessories.** (Wearing accessories)
- **Age**
- **Blur** (blurriness of the face in the image)
- **Emotion.**
- **Exposure**
- **Facial hair**
- **Gender**
- **Glasses**
- **Hair**
- **Head pose**
- **Makeup**
- **Mask.** (are they wearing a mask?)
- **Noise.** The visual noise detected in the face image
- **Occlusion.** (objects blocking parts of the face)
- **Smile**



- 💡 The service can translate 90 languages and dialects while also addressing issues like artifacts or occlusion that may affect facial recognition, including a Boolean value for detecting smiles.

▶ 1:06:30

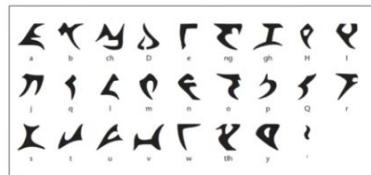
Speech and Translate Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Azure's Translate service is a **translation service**.

- It can translate 90 languages and dialects
 - It even supports **Klingon!**
- It uses **Neural Machine Translation (NMT)** replacing its legacy **Statistical Machine Translation (SMT)**
- **Custom Translator** allows you to extend the service for translation based on your business and domain use case



- 💡 A translation service has transitioned from statistical machine translation to neural machine translation for improved accuracy and offers customization based on specific business needs.

▶ 1:06:46

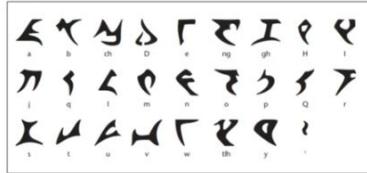
Speech and Translate Service

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Azure's Translate service is a **translation service**.

- It can translate 90 languages and dialects
 - It even supports **Klingon!** 
- It uses Neural Machine Translation (NMT) replacing its legacy Statistical Machine Translation (SMT)
- **Custom Translator** allows you to extend the service for translation based on your business and domain use case



Azure Speech service can **speech synthesis service** speech-to-text, text-to-speech, and speech-translation

Speech-to-Text

- Real-time Speech-to-text
- Batch Speech-to-Text
- Multi-device Conversation
- Conversation Transcription
- Create Custom Speech Models

Text-to-Speech

- using Speech Synthesis Markup Language (SSML)
- Create Custom Voices

Voice Assistance

- integrates with Bot Framework SDK

Speech Recognition

- Speaker verification & identification

💡 The technology encompasses speech-to-text, text-to-speech, custom voice creation, conversation transcription, and integration with voice assistants using various frameworks.

▶ 1:07:30

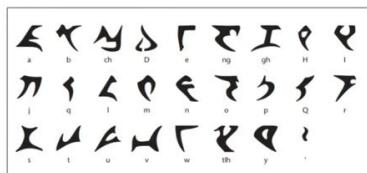
Speech and Translate Service

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Azure's Translate service is a **translation service**.

- It can translate 90 languages and dialects
 - It even supports **Klingon!** 
- It uses Neural Machine Translation (NMT) replacing its legacy Statistical Machine Translation (SMT)
- **Custom Translator** allows you to extend the service for translation based on your business and domain use case



Azure Speech service can **speech synthesis service** speech-to-text, text-to-speech, and speech-translation

Speech-to-Text

- Real-time Speech-to-text
- Batch Speech-to-Text
- Multi-device Conversation
- Conversation Transcription
- Create Custom Speech Models

Text-to-Speech

- using Speech Synthesis Markup Language (SSML)
- Create Custom Voices

Voice Assistance

- integrates with Bot Framework SDK

Speech Recognition

- Speaker verification & identification

💡 The content discusses text-to-speech technology, speech synthesis markup language, voice assistants, and natural language processing for text analytics.

▶ 1:08:01

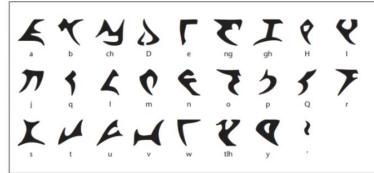
Speech and Translate Service

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Azure's Translate service is a **translation service**.

- It can translate 90 languages and dialects
 - It even supports **Klingon!** 
- It uses **Neural Machine Translation (NMT)** replacing its legacy **Statistical Machine Translation (SMT)**
- **Custom Translator** allows you to extend the service for translation based on your business and domain use case



Azure Speech service can **speech synthesis service** speech-to-text, text-to-speech, and speech-translation

Speech-to-Text

- Real-time Speech-to-text
- Batch Speech-to-Text
- Multi-device Conversation
- Conversation Transcription
- Create Custom Speech Models

Text-to-Speech

- using Speech Synthesis Markup Language (SSML)
- Create Custom Voices

Voice Assistance

- integrates with Bot Framework SDK

Speech Recognition

- Speaker verification & identification

💡 The text discusses the use of speech synthesis markup language for creating custom voices, integrating voice assistants with bot frameworks and speech recognition, as well as exploring natural language processing for text analytics.

▷ 1:08:01



Text Analytics

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Text Analytics API is a **Natural Language Processing (NLP)** service for **text mining and text analysis**

Text Analytics can perform:

- **sentiment analysis**
 - find out what people think of your brand or topic
 - feature provides sentiment labels (such as "negative", "neutral" and "positive")
- **opinion mining**
 - aspect-based sentiment analysis
 - granular information about the opinions related to aspects
- **key phrase extraction**
 - quickly identify the main concepts in text.
- **language detection**
 - detect the language an input text is written in
- **named entity recognition (NER)**
 - Identify and categorize entities in your text as people, places, organizations, quantities
 - Subset of NER is Personally Identifiable Information (PII)

▷ 1:08:26



Text Analytics

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Text Analytics API is a **Natural Language Processing (NLP)** service for **text mining and text analysis**

Text Analytics can perform:

- **sentiment analysis**
 - find out what people think of your brand or topic
 - feature provides sentiment labels (such as "negative", "neutral" and "positive")
- **opinion mining**
 - aspect-based sentiment analysis
 - granular information about the opinions related to aspects
- **key phrase extraction**
 - quickly identify the main concepts in text.
- **language detection**
 - detect the language an input text is written in
- **named entity recognition (NER)**
 - Identify and categorize entities in your text as people, places, organizations, quantities

💡 The text discusses various natural language processing techniques, including key phrase extraction, language detection, named entity recognition (NER), and personally identifiable information (PII) identification.

▷ 1:08:52



Text Analytics

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Text Analytics API is a **Natural Language Processing (NLP)** service for **text mining and text analysis**

Text Analytics can perform:

- **sentiment analysis**
 - find out what people think of your brand or topic
 - feature provides sentiment labels (such as "negative", "neutral" and "positive")
- **opinion mining**
 - aspect-based sentiment analysis
 - granular information about the opinions related to aspects
- **key phrase extraction**
 - quickly identify the main concepts in text.
- **language detection**
 - detect the language an input text is written in
- **named entity recognition (NER)**
 - Identify and categorize entities in your text as people, places, organizations, quantities
 - Subset of NER is Personally Identifiable Information (PII)

💡 Key phrase extraction is a technique used to identify the main concepts in larger amounts of text, while named entity recognition categorizes entities into types such as people, places, objects, and quantities.

▷ 1:08:58

NLP – Key Phrase Extraction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Key Phrase Extraction quickly identify the main concepts in text

- Key phrase extraction works best when you give it bigger amounts of text to work on
- This is opposite from sentiment analysis, which performs better on smaller amounts of text
- Document size must be 5,120 or fewer characters per document, and you can have up to 1,000 items (IDs) per collection

▶ 1:09:13

NLP – Key Phrase Extraction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Key Phrase Extraction quickly identify the main concepts in text

- Key phrase extraction works best when you give it bigger amounts of text to work on
- This is opposite from sentiment analysis, which performs better on smaller amounts of text
- Document size must be 5,120 or fewer characters per document, and you can have up to 1,000 items (IDs) per collection

When the Borg launch an attack on Earth, the Enterprise is sent to the neutral zone due to the Admiralty's mistrust of Picard's abilities as he had been assimilated in the past. The Enterprise however, disobeys and returns to help destroy the Borg ship. However a smaller ship escapes and travels back in time, causing the assimilation of Earth in the future. The Enterprise follows the ship back in time and have to undo the damage the ship did on the surface to an experimental warp drive unit that will lead Earth to its first contact with alien life. Meanwhile, on the Enterprise, survivors of the Borg ship begin to assimilate decks within the ship itself...

Key Phrases:
Borg ship
Enterprise
smaller ship escapes
time
assimilation of Earth
surface
experimental warp drive unit
Admiralty's mistrust of Picard's abilities
neutral zone
travels
contact
damage
attack
survivors
decks
alien life
future
past

▶ 1:09:25

NLP – Key Phrase Extraction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Key Phrase Extraction quickly identify the main concepts in text

- Key phrase extraction works best when you give it bigger amounts of text to work on
- This is opposite from sentiment analysis, which performs better on smaller amounts of text
- Document size must be 5,120 or fewer characters per document, and you can have up to 1,000 items (IDs) per collection

When the Borg launch an attack on Earth, the Enterprise is sent to the neutral zone due to the Admiralty's mistrust of Picard's abilities as he had been assimilated in the past. The Enterprise however, disobeys and returns to help destroy the Borg ship. However a smaller ship escapes and travels back in time, causing the assimilation of Earth in the future. The Enterprise follows the ship back in time and have to undo the damage the ship did on the surface to an experimental warp drive unit that will lead Earth to its first contact with alien life. Meanwhile, on the Enterprise, survivors of the Borg ship begin to assimilate decks within the ship itself....

Key Phrases:
Borg ship
Enterprise
smaller ship escapes
time
assimilation of Earth
surface
experimental warp drive unit
Admiralty's mistrust of Picard's abilities
neutral zone
travels
contact
damage
attack
survivors
decks
alien life
future
past

- Key phrases and named entities can be extracted from unstructured data, with a limit of 5,000 characters per document and up to 1,000 items per collection.

▶ 1:09:33

NLP – Key Phrase Extraction

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Key Phrase Extraction quickly identify the main concepts in text

- Key phrase extraction works best when you give it bigger amounts of text to work on
- This is opposite from sentiment analysis, which performs better on smaller amounts of text
- Document size must be 5,120 or fewer characters per document, and you can have up to 1,000 items (IDs) per collection

When the Borg launch an attack on Earth, the Enterprise is sent to the neutral zone due to the Admiralty's mistrust of Picard's abilities as he had been assimilated in the past. The Enterprise however, disobeys and returns to help destroy the Borg ship. However a smaller ship escapes and travels back in time, causing the assimilation of Earth in the future. The Enterprise follows the ship back in time and have to undo the damage the ship did on the surface to an experimental warp drive unit that will lead Earth to its first contact with alien life. Meanwhile, on the Enterprise, survivors of the Borg ship begin to assimilate decks within the ship itself....

Key Phrases:
Borg ship
Enterprise
smaller ship escapes
time
assimilation of Earth
surface
experimental warp drive unit
Admiralty's mistrust of Picard's abilities
neutral zone
travels
contact
damage
attack
survivors
decks
alien life
future
past

- The text discusses the extraction of key phrases and named entity recognition from unstructured data, highlighting their applications in identifying significant terms and semantic types.

▶ 1:09:35

Named Entity Recognition

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Named Entity Recognition detects **words and phrases mentioned in unstructured text** that can be **associated with one or more semantic types**.

Ribavirin [UMLS: C0035525] was also evaluated against SARS-CoV-2 infection , but the antiviral [UMLS: C0003451].

property of drugs [UMLS: C0013227] is still not well established against the SARS-CoV-2 [UMLS: C5203670] negation .

In addition, after oral administration, the drug was rapidly absorbed into the GI tract [UMLS: C0017189] .

The drug has oral bioavailability around 64 % with large volume of distribution.

Semantic types could be: *Location, Event, Location, Person, Diagnosis, Age*

▶ 1:10:09

Named Entity Recognition

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Named Entity Recognition detects **words and phrases mentioned in unstructured text** that can be **associated with one or more semantic types**.

Ribavirin [UMLS: C0035525] was also evaluated against SARS-CoV-2 infection , but the antiviral [UMLS: C0003451].

property of drugs [UMLS: C0013227] is still not well established against the SARS-CoV-2 [UMLS: C5203670] negation .

In addition, after oral administration, the drug was rapidly absorbed into the GI tract [UMLS: C0017189] .

The drug has oral bioavailability around 64 % with large volume of distribution.

Semantic types could be: *Location, Event, Location, Person, Diagnosis, Age*

- ★ The discussion involves applying semantic types to words or phrases, categorizing them into predefined sets such as diagnosis, medication class, location, and event, while also distinguishing between sentiment analysis and opinion mining.

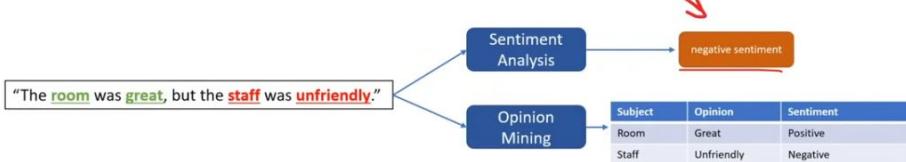
▶ 1:10:09

NLP – Sentiment Analysis

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Sentiment analysis will apply labels and confidence score to text at the **sentence and document level**.

- Labels include **negative, positive, mixed or neutral**
- Confidence scores ranging from 0 to 1



Opinion mining will provide more granular data with a **Subject** and **Opinion** tied to a Sentient

- Sentiment analysis can categorize sentiments as negative, positive, mixed, or neutral while providing a confidence score, but opinion mining offers more detailed insights by distinguishing between subjects and opinions.

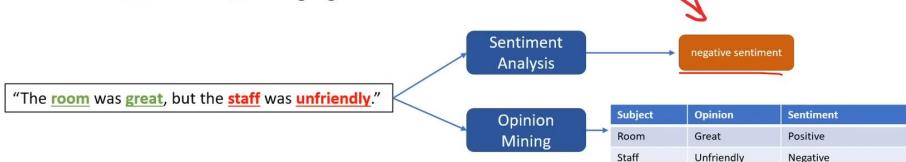
▶ 1:10:25

NLP – Sentiment Analysis

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Sentiment analysis will apply labels and confidence score to text at the **sentence and document level**.

- Labels include **negative, positive, mixed or neutral**
- Confidence scores ranging from 0 to 1



Opinion mining will provide more granular data with a **Subject** and **Opinion** tied to a Sentient

- Opinion mining provides granular data that allows for the identification of both positive and negative sentiments within a single subject.

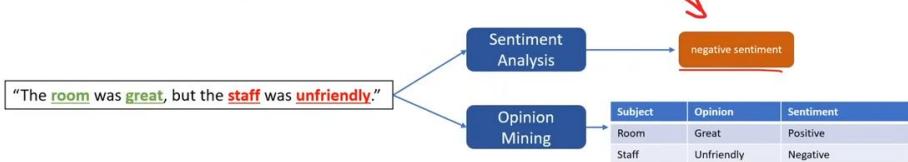
▶ 1:10:57

NLP – Sentiment Analysis

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Sentiment analysis will apply labels and confidence score to text at the **sentence and document level**.

- Labels include **negative, positive, mixed or neutral**
- Confidence scores ranging from 0 to 1



Opinion mining provides more granular data with a **Subject** and **Opinion** tied to a Sentiment

- Opinion mining provides granular data by analyzing both positive and negative sentiments about specific subjects.

▶ 1:10:58

[] Optical Character Recognition (OCR)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Optical character recognition (OCR) is the process of **extracting printed or handwritten text** into a digital and editable format

OCR can be applied to:

- photos of street signs
- Products**
- Documents
- Invoices
- Bills
- Financial Reports

▶ 1:11:16



Optical Character Recognition (OCR)

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Optical character recognition (OCR) is the process of extracting printed or handwritten text into a digital and editable format

OCR can be applied to:

- photos of street signs
- **Products** →
- Documents
- Invoices
- Bills
- Financial Reports
- Articles
- and more



Azure offers two APIs for extracting printed or handwritten text into a digital format: the OCR API and the Read API.

▶ 1:11:27



Optical Character Recognition (OCR)

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Optical character recognition (OCR) is the process of extracting printed or handwritten text into a digital and editable format

OCR can be applied to:

- photos of street signs
- **Products** →
- Documents
- Invoices
- Bills
- Financial Reports
- Articles
- and more



Azure offers two APIs for optical character recognition (OCR): the older OCR API and the newer Read API, which can extract text from various documents and images.

▶ 1:11:30



Optical Character Recognition (OCR)

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure has **two different APIs** that can perform OCR: **OCR API** and **Read API**

OCR API

- older recognition model
- supports only images
- executes synchronously
 - returning immediately with the detected text
 - Suited for less text
- Support more languages
- Easier to implement

Read API

- updated recognition model
- Supports images and PDFs
- Executes asynchronously
 - parallelizes tasks per line for faster results
 - Suited for lots of text
- Supports fewer languages
- A bit more difficult to implement



OCR is performed via the **Computer Vision SDK**

- ★ There are two recognition models: an older synchronous model that supports only images and is better for less text, and a newer asynchronous model that supports both images and PDFs, processes tasks in parallel for faster results, and is suited for larger amounts of text.

▷ 1:11:41



Optical Character Recognition (OCR)

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure has **two different APIs** that can perform OCR: **OCR API** and **Read API**

OCR API

- older recognition model
- supports only images
- executes synchronously
 - returning immediately with the detected text
 - Suited for less text
- Support more languages
- Easier to implement

Read API

- updated recognition model
- Supports images and PDFs
- Executes asynchronously
 - parallelizes tasks per line for faster results
 - Suited for lots of text
- Supports fewer languages
- A bit more difficult to implement



OCR is performed via the **Computer Vision SDK**

- ★ The service processes images and PDFs asynchronously, enabling faster task execution for large volumes of text while supporting fewer languages and being more complex to implement.

▷ 1:12:17



Optical Character Recognition (OCR)

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure has **two different APIs** that can perform OCR: **OCR API** and **Read API**

OCR API

- older recognition model
- supports only images
- executes synchronously
 - returning immediately with the detected text
 - Suited for less text
- Support more languages
- Easier to implement

Read API

- updated recognition model
- Supports images and PDFs
- Executes asynchronously
 - parallelizes tasks per line for faster results
 - Suited for lots of text
- Supports fewer languages
- A bit more difficult to implement



OCR is performed via the **Computer Vision SDK**

- ★ The service utilizes asynchronous processing to enhance task efficiency for text recognition, supports limited languages, and requires more complex implementation.

▷ 1:12:18



Form Recognizer Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure Form Recognizer is a **specialize OCR service** (translates printed text into digital and editable content) and **preserves that structure and relationship of form-like data**

Form Recognizer to **automate data entry** in your applications and **enrich your documents search capabilities**

Form Recognizer can identify:

- Key Value Pairs
- Selection Marks
- Table Structures

Form Recognizer outputs structures such as:

- Original file relationships
- Bounding boxes
- Confidence scores

Form Recognizer is composed of

- Custom document processing models
- Prebuilt models for invoices, receipts, IDs, business cards

- ★ Form Recognizer automates data entry and enhances document search by identifying key-value pairs, selection marks, and table structures while producing output that includes original file relationships, bounding boxes, and confidence scores.

▷ 1:12:38



Form Recognizer Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure Form Recognizer is a **specialize OCR service** (translates printed text into digital and editable content) and **preserves that structure and relationship of form-like data**

Form Recognizer to **automate data entry** in your applications and **enrich your documents search capabilities**

Form Recognizer can identify:

- Key Value Pairs
- Selection Marks
- Table Structures

Form Recognizer outputs structures such as:

- Original file relationships
- Bounding boxes
- Confidence scores

Form Recognizer is composed of

- Custom document processing models
- Prebuilt models for invoices, receipts, IDs, business cards
- The Layout Model

▶ 1:13:05

Form Recognizer Service – Layout

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Extract text, selection marks, and tables structures, along with their bounding box coordinates, from documents

Form Recognizer can extract text, selection marks, and table structure (the row and column numbers associated with the text) using high-definition optical character recognition (OCR) and an enhanced deep learning model from documents

The screenshot shows an invoice from Adventure Works Cycles. At the top, there is company information: "Adventure Works Cycles", "Adventure Works", "Marketing and Business division", "E-mail: sales@adventure-works.com", "Tel: +45 35 35 35 - Fax: +45 35 35 35". Below this is the invoice header: "00577710", "Invoice date: 01/01/2018", "Customer: N F 845 089 98", "Shipment date: 05/01/2018", "Ship to: AV DOLCE", "In: 15/12/18", "Out: 17/12/18", "Rate: ". The main body of the invoice contains a table with columns: "Description", "Line", "Quantity", "Unit price", "Amount". The table includes entries like "TAXI SP5", "Ship monitoring", "Entry", "Stay tax", "LightHouse", etc. At the bottom right of the invoice, there is a summary table with columns: "N.U.M.", "644889,00", "N.U.M.", "103184,00", "N.U.M.", "748072,00", "VAT", "0,00", and a total row "TOTAL VAT Inc." with a value of "2021,00".

💡 High-definition optical character enhancement models are ineffective for identifying row and column numbers related to text in marks and table structures.

▶ 1:13:14

Form Recognizer Service – Layout

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Extract text, selection marks, and tables structures, along with their bounding box coordinates, from documents

Form Recognizer can extract text, selection marks, and table structure (the row and column numbers associated with the text) using high-definition optical character recognition (OCR) and an enhanced deep learning model from documents

Adventure Works Cycles		Customer Information		Billing Information	
Marketing Department		Marketing and Business Division		Sales Tax	
Fax: 403 35 55 - Fax: 403 35 15				15/09/20	
00577110		Stopover		184450003	
Invoice no:		Ship to		All D/C/CCE	
N F 845 089 98		In		15/12/18	
		Out		17/12/18	
		Rate			
CUSTOMER: 029 VALEO		SEJ:	I NB	E/S:	1
		Quantity	Unit Price	Amount	
TAX / SPS		1	10000	10000,00	
Ship monitoring	TJ	18558	0,119	10156,03	
Customs	TJ	18558	0,175	3226,50	
Stay tax	TJ	18558	0,135	2484,70	
Lighthousier	TJ	18558	0,000	0,00	
Customs	TJ	18558	0,177	3226,50	
Stay tax	TJ	18558	0,135	2484,70	
Lighthousier	TJ	18558	0,033	529,01	
		N-UM		644889,00	
		U-UM			
		N-UM		103184,00	
		U-UM		478072,00	
		TOTAL VAT incl			
Invoice amount is the following: Seven hundred forty eight thousand seven two CFA					
Responsible	Service head		Director		

▶ 1:13:22

 A specific piece of information is conveyed.

▶ 1:13:30



Form Recognizer Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure Form Recognizer is a **specialize OCR service** (translates printed text into digital and editable content) and **preserves that structure and relationship of form-like data**

Form Recognizer to **automate data entry** in your applications and **enrich your documents search capabilities**

Form Recognizer can identify: Form Recognizer outputs structures such as:

- Key Value Pairs
- Selection Marks
- Bounding boxes
- Table Structures
- Confidence scores

Form Recognizer is composed of

- Custom document processing models
- Prebuilt models for invoices, receipts, IDs, business cards
- The Layout Model

The screenshot shows an invoice from 'Adventure Works Cycles' with a table of items. The table has columns for 'Description', 'Quantity', 'Unit Price', and 'Total'. The items listed include 'Tire - MTB', 'Grip - mountain', 'Candy', 'Water bottle', 'Lightweight', 'Water bottle', 'Lightweight', and 'Water bottle'. The total amount is 200000.00.

Description	Quantity	Unit Price	Total
Tire - MTB	10	18524	185240.00
Grip - mountain	10	81124	811240.00
Candy	10	12124	121240.00
Water bottle	10	18524	185240.00
Lightweight	10	18524	185240.00
Water bottle	10	18524	185240.00
Lightweight	10	18524	185240.00
Water bottle	10	18524	185240.00
Lightweight	10	18524	185240.00
			200000.00

Invoice address is in the following
Street Address: 1000 N. Capital Street, Seattle, WA 98101

Responsible: Service lead Director

▷ 1:14:13



Form Recognizer Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure Form Recognizer is a **specialize OCR service** (translates printed text into digital and editable content) and **preserves that structure and relationship of form-like data**

Form Recognizer to **automate data entry** in your applications and **enrich your documents search capabilities**

Form Recognizer can identify:

The screenshot shows an invoice from 'Adventure Works Cycles' with a table of items. The table has columns for 'Description', 'Quantity', 'Unit Price', and 'Total'. The items listed include 'Tire - MTB', 'Grip - mountain', 'Candy', 'Water bottle', 'Lightweight', 'Water bottle', 'Lightweight', and 'Water bottle'. The total amount is 200000.00.

Description	Quantity	Unit Price	Total
Tire - MTB	10	18524	185240.00
Grip - mountain	10	81124	811240.00
Candy	10	12124	121240.00
Water bottle	10	18524	185240.00
Lightweight	10	18524	185240.00
Water bottle	10	18524	185240.00
Lightweight	10	18524	185240.00
Water bottle	10	18524	185240.00
Lightweight	10	18524	185240.00
			200000.00

Invoice address is in the following
Street Address: 1000 N. Capital Street, Seattle, WA 98101

Responsible: Service lead Director

- 💡 Form Recognizer automates data entry by identifying key-value pairs, selection marks, and table structures in documents while providing structured outputs like bounding boxes and confidence scores through custom and pre-built models.

▷ 1:14:19



Form Recognizer Service

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

Azure Form Recognizer is a **specialized OCR service** (translates printed text into digital and editable content) and **preserves that structure and relationship of form-like data**

Form Recognizer to **automate data entry** in your applications and **enrich your documents search capabilities**

Form Recognizer can identify: Form Recognizer outputs structures such as:

- Key Value Pairs
- Selection Marks
- Bounding boxes
- Table Structures
- Confidence scores

Form Recognizer is composed of

- Custom document processing models
- Prebuilt models for invoices, receipts, IDs, business cards
- The Layout Model

Product ID	Description	Quantity	Unit Price	Extended Price
285104	Shimano 10-speed	10	100000,00	1000000,00
285105	Shimano 9-speed	10	85000,00	850000,00
285106	Shimano 7-speed	10	70000,00	700000,00
285107	Shimano 5-speed	10	50000,00	500000,00
285108	Shimano 3-speed	10	30000,00	300000,00
285109	Shimano 1-speed	10	20000,00	200000,00
285110	Shimano 1-speed	10	15000,00	150000,00
285111	Shimano 1-speed	10	10000,00	100000,00
285112	Shimano 1-speed	10	8000,00	80000,00
285113	Shimano 1-speed	10	6000,00	60000,00
285114	Shimano 1-speed	10	4000,00	40000,00
285115	Shimano 1-speed	10	3000,00	30000,00
285116	Shimano 1-speed	10	2000,00	20000,00
285117	Shimano 1-speed	10	1500,00	15000,00
285118	Shimano 1-speed	10	1000,00	10000,00
285119	Shimano 1-speed	10	800,00	8000,00
285120	Shimano 1-speed	10	600,00	6000,00
285121	Shimano 1-speed	10	400,00	4000,00
285122	Shimano 1-speed	10	300,00	3000,00
285123	Shimano 1-speed	10	200,00	2000,00
285124	Shimano 1-speed	10	150,00	1500,00
285125	Shimano 1-speed	10	100,00	1000,00
285126	Shimano 1-speed	10	80,00	800,00
285127	Shimano 1-speed	10	60,00	600,00
285128	Shimano 1-speed	10	40,00	400,00
285129	Shimano 1-speed	10	30,00	300,00
285130	Shimano 1-speed	10	20,00	200,00
285131	Shimano 1-speed	10	15,00	150,00
285132	Shimano 1-speed	10	10,00	100,00
285133	Shimano 1-speed	10	8,00	80,00
285134	Shimano 1-speed	10	6,00	60,00
285135	Shimano 1-speed	10	4,00	40,00
285136	Shimano 1-speed	10	3,00	30,00
285137	Shimano 1-speed	10	2,00	20,00
285138	Shimano 1-speed	10	1,50	15,00
285139	Shimano 1-speed	10	1,00	10,00
285140	Shimano 1-speed	10	0,80	8,00
285141	Shimano 1-speed	10	0,60	6,00
285142	Shimano 1-speed	10	0,40	4,00
285143	Shimano 1-speed	10	0,30	3,00
285144	Shimano 1-speed	10	0,20	2,00
285145	Shimano 1-speed	10	0,15	1,50
285146	Shimano 1-speed	10	0,10	1,00
285147	Shimano 1-speed	10	0,08	0,80
285148	Shimano 1-speed	10	0,06	0,60
285149	Shimano 1-speed	10	0,05	0,50
285150	Shimano 1-speed	10	0,04	0,40
285151	Shimano 1-speed	10	0,03	0,30
285152	Shimano 1-speed	10	0,02	0,20
285153	Shimano 1-speed	10	0,01	0,10
285154	Shimano 1-speed	10	0,005	0,05
285155	Shimano 1-speed	10	0,002	0,02
285156	Shimano 1-speed	10	0,001	0,01
285157	Shimano 1-speed	10	0,0005	0,005
285158	Shimano 1-speed	10	0,0002	0,002
285159	Shimano 1-speed	10	0,0001	0,001
285160	Shimano 1-speed	10	0,00005	0,0005
285161	Shimano 1-speed	10	0,00002	0,0002
285162	Shimano 1-speed	10	0,00001	0,0001
285163	Shimano 1-speed	10	0,000005	0,00005
285164	Shimano 1-speed	10	0,000002	0,00002
285165	Shimano 1-speed	10	0,000001	0,00001
285166	Shimano 1-speed	10	0,0000005	0,000005
285167	Shimano 1-speed	10	0,0000002	0,000002
285168	Shimano 1-speed	10	0,0000001	0,000001
285169	Shimano 1-speed	10	0,00000005	0,0000005
285170	Shimano 1-speed	10	0,00000002	0,0000002
285171	Shimano 1-speed	10	0,00000001	0,0000001
285172	Shimano 1-speed	10	0,000000005	0,00000005
285173	Shimano 1-speed	10	0,000000002	0,00000002
285174	Shimano 1-speed	10	0,000000001	0,00000001
285175	Shimano 1-speed	10	0,0000000005	0,000000005
285176	Shimano 1-speed	10	0,0000000002	0,000000002
285177	Shimano 1-speed	10	0,0000000001	0,000000001
285178	Shimano 1-speed	10	0,00000000005	0,0000000005
285179	Shimano 1-speed	10	0,00000000002	0,0000000002
285180	Shimano 1-speed	10	0,00000000001	0,0000000001
285181	Shimano 1-speed	10	0,000000000005	0,00000000005
285182	Shimano 1-speed	10	0,000000000002	0,00000000002
285183	Shimano 1-speed	10	0,000000000001	0,00000000001
285184	Shimano 1-speed	10	0,0000000000005	0,000000000005
285185	Shimano 1-speed	10	0,0000000000002	0,000000000002
285186	Shimano 1-speed	10	0,0000000000001	0,000000000001
285187	Shimano 1-speed	10	0,00000000000005	0,0000000000005
285188	Shimano 1-speed	10	0,00000000000002	0,0000000000002
285189	Shimano 1-speed	10	0,00000000000001	0,0000000000001
285190	Shimano 1-speed	10	0,000000000000005	0,00000000000005
285191	Shimano 1-speed	10	0,000000000000002	0,00000000000002
285192	Shimano 1-speed	10	0,000000000000001	0,00000000000001
285193	Shimano 1-speed	10	0,0000000000000005	0,000000000000005
285194	Shimano 1-speed	10	0,0000000000000002	0,000000000000002
285195	Shimano 1-speed	10	0,0000000000000001	0,000000000000001
285196	Shimano 1-speed	10	0,00000000000000005	0,0000000000000005
285197	Shimano 1-speed	10	0,00000000000000002	0,0000000000000002
285198	Shimano 1-speed	10	0,00000000000000001	0,0000000000000001
285199	Shimano 1-speed	10	0,000000000000000005	0,00000000000000005
285200	Shimano 1-speed	10	0,000000000000000002	0,00000000000000002
285201	Shimano 1-speed	10	0,000000000000000001	0,00000000000000001
285202	Shimano 1-speed	10	0,0000000000000000005	0,000000000000000005
285203	Shimano 1-speed	10	0,0000000000000000002	0,000000000000000002
285204	Shimano 1-speed	10	0,0000000000000000001	0,000000000000000001
285205	Shimano 1-speed	10	0,00000000000000000005	0,000000000000000005
285206	Shimano 1-speed	10	0,00000000000000000002	0,000000000000000002
285207	Shimano 1-speed	10	0,00000000000000000001	0,000000000000000001
285208	Shimano 1-speed	10	0,000000000000000000005	0,000000000000000005
285209	Shimano 1-speed	10	0,000000000000000000002	0,000000000000000002
285210	Shimano 1-speed	10	0,000000000000000000001	0,000000000000000001
285211	Shimano 1-speed	10	0,0000000000000000000005	0,000000000000000005
285212	Shimano 1-speed	10	0,0000000000000000000002	0,000000000000000002
285213	Shimano 1-speed	10	0,0000000000000000000001	0,000000000000000001
285214	Shimano 1-speed	10	0,00000000000000000000005	0,000000000000000005
285215	Shimano 1-speed	10	0,00000000000000000000002	0,000000000000000002
285216	Shimano 1-speed	10	0,00000000000000000000001	0,000000000000000001
285217	Shimano 1-speed	10	0,000000000000000000000005	0,000000000000000005
285218	Shimano 1-speed	10	0,000000000000000000000002	0,000000000000000002
285219	Shimano 1-speed	10	0,000000000000000000000001	0,000000000000000001
285220	Shimano 1-speed	10	0,0000000000000000000000005	0,000000000000000005
285221	Shimano 1-speed	10	0,0000000000000000000000002	0,000000000000000002
285222	Shimano 1-speed	10	0,0000000000000000000000001	0,000000000000000001
285223	Shimano 1-speed	10	0,00000000000000000000000005	0,000000000000000005
285224	Shimano 1-speed	10	0,00000000000000000000000002	0,000000000000000002
285225	Shimano 1-speed	10	0,00000000000000000000000001	0,000000000000000001
285226	Shimano 1-speed	10	0,000000000000000000000000005	0,000000000000000005
285227	Shimano 1-speed	10	0,000000000000000000000000002	0,000000000000000002
285228	Shimano 1-speed	10	0,000000000000000000000000001	0,000000000000000001
285229	Shimano 1-speed	10	0,0000000000000000000000000005	0,000000000000000005
285230	Shimano 1-speed	10	0,0000000000000000000000000002	0,000000000000000002
285231	Shimano 1-speed	10	0,0000000000000000000000000001	0,000000000000000001
285232	Shimano 1-speed	10	0,00000000000000000000000000005	0,000000000000000005
285233	Shimano 1-speed	10	0,00000000000000000000000000002	0,000000000000000002
285234	Shimano 1-speed	10	0,00000000000000000000000000001	0,000000000000000001
285235	Shimano 1-speed	10	0,000000000000000000000000000005	0,000000000000000005
285236	Shimano 1-speed	10	0,000000000000000000000000000002	0,000000000000000002
285237	Shimano 1-speed	10	0,000000000000000000000000000001	0,000000000000000001
285238	Shimano 1-speed	10	0,0000000000000000000000000000005	0,000000000000000005
285239	Shimano 1-speed	10	0,0000000000000000000000000000002	0,000000000000000002
285240	Shimano 1-speed	10	0,0000000000000000000000000000001	0,000000000000000001
285241	Shimano 1-speed	10	0,00000000000000000000000000000005	0,000000000000000005
285242	Shimano 1-speed	10	0,00000000000000000000000000000002	0,000000000000000002
285243	Shimano 1-speed	10	0,00000000000000000000000000000001	0,000000000000000001
285244	Shimano 1-speed	10	0,000000000000000000000000000000005	0,000000000000000005
285245	Shimano 1-speed	10	0,000000000000000000000000000000002	0,000000000000000002
285246	Shimano 1-speed	10	0,000000000000000000000000000000001	0,000000000000000001
285247	Shimano 1-speed	10	0,0000000000000000000000000000000005	0,000000000000000005
285248	Shimano 1-speed	10	0,0000000000000000000000000000000002	0,000000000000000002
285249	Shimano 1-speed	10	0,0000000000000000000000000000000001	0,000000000000000001
285250	Shimano 1-speed	10	0,00000000000000000000000000000000005	0,000000000000000005
285251	Shimano 1-speed	10	0,00000000000000000000000000000000002	0,000000000000000002
285252	Shimano 1-speed	10	0,00000000000000000000000000000000001	0,000000000000000001
285253	Shimano 1-speed	10	0,000000000000000000000000000000000005	0,000000000000000005
2852				

Form Recognizer Service – Custom Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Custom models allows you to **extract text, key/value pairs, selection marks, and table data** from forms

- These models are trained with your own data, so they're tailored to your forms
- you only need five sample input forms to start.
- A trained document processing model can output structured data that includes the relationships in the original form document.
- After you train the model, you can test and retrain it and eventually use it to reliably extract data from more forms according to your needs.

You have **2 learning options:**

Train without Labels

uses **unsupervised learning** to understand the layout and relationships between fields and entries in your forms

Train with Labels

uses **supervised learning** to extract values of interest, using the labeled forms you provide (trained data).

- The model allows for testing and retraining to reliably extract data from forms using unsupervised learning to understand layouts and relationships, as well as supervised learning for extracting specific values with labeled forms.

 1:15:28

Form Recognizer Service – Prebuilt Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Business Cards

English business cards

 1:15:44

Form Recognizer Service – Prebuilt Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Receipts

Sales receipts from Australia, Canada, Great Britain, India, and the United States



Fields Extracted

ReceiptType
MerchantName
MerchantPhoneNumber
MerchantAddress
TransactionDate
TransactionTime
Total
Subtotal
Tax
Tip
Items
Name
Quantity
Price
Total Price

- The system can extract various fields from receipts and English business cards, including merchant details, transaction information, and contact information.

▶ 1:16:05

Form Recognizer Service – Prebuilt Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Business Cards

English business cards



Fields Extracted

ContactNames
FirstName
LastName
CompanyNames
Departments
JobTitles
Emails
Websites
Addresses
MobilePhones
Faxes
WorkPhones
OtherPhones

▶ 1:16:26

Form Recognizer Service – Prebuilt Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Identity Documents (IDs)

world-wide passports and US driver licenses

Extracted Fields

CountryRegion
DateOfBirth
DateOfExpiration
DocumentNumber
FirstName
LastName
Nationality
Sex
MachineReadableZone
DocumentType
Address
Region



▶ 1:16:35



Form Recognizer Service – Prebuilt Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Invoices

extracts data from invoices in various formats
and returns structured data

Extracted Fields

		Extracted Line Item Data
CustomerName	ShippingAddress	Items
CustomerId	ShippingAddressRecipient	Amount
PurchaseOrder	SubTotal	Description
InvoiceId	TotalTax	Quantity
InvoiceDate	InvoiceTotal	UnitPrice
DueDate	AmountDue	ProductCode
VendorName	ServiceAddress	Unit
VendorAddress	ServiceAddressRecipient	Date
VendorAddressRecipient	RemittanceAddress	Tax
CustomerAddress	RemittanceAddressRecipient	
CustomerAddressRecipient	ServiceStartDate	
BillingAddress	ServiceEndDate	
BillingAddressRecipient	PreviousUnpaidBalance	

- 💡 The information includes various fields related to service addresses, remittance details, previous balances, line items with descriptions and pricing, as well as identification document specifics such as personal data and expiration details.

▶ 1:17:04

Form Recognizer Service – Prebuilt Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Identity Documents (IDs)

world-wide passports and US driver licenses

Extracted Fields

CountryRegion
DateOfBirth
DateOfExpiration
DocumentNumber
FirstName
LastName
Nationality
Sex
MachineReadableZone
DocumentType
Address
Region



- The text discusses various components of personal identification documents, including details such as birth date, expiration date, name, nationality, sex, and additional features.

▶ 1:17:28

{} Language Understanding Service (LUIS)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Language Understanding (LUIS) is a no-code ML service to build natural language into apps, bots, and IoT devices.

Quickly create enterprise-ready, custom models that continuously improve.

LUIS is accessed via its own isolate domain at luis.ai

LUIS utilizes Natural Language Processing (NLP) and **Natural Language Understanding (NLU)**

NLU is the ability to *transform* a linguistic statement to a representation that enables you to understand your users naturally

LUIS is intended to focus on **intention** and **extraction**:

- What the user wants
- What they are talking about

- Luis is a service that leverages natural language processing and understanding to create customizable, enterprise-ready models for applications, bots, and IoT devices.

▶ 1:17:46

{ } Language Understanding Service (LUIS)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Language Understanding (LUIS) is a no-code ML service to build natural language into apps, bots, and IoT devices.

Quickly create enterprise-ready, custom models that continuously improve.

LUIS is accessed via its own isolate domain at luis.ai

LUIS utilizes Natural Language Processing (NLP) and **Natural Language Understanding (NLU)**

NLU is the ability to *transform* a linguistic statement to a representation that enables you to understand your users naturally

LUIS is intended to focus on **intention** and **extraction**:

- What the user wants
- What they are talking about

- 💡 A schema is autogenerated in the Lewis AI web interface to help understand user intentions and extract relevant information without requiring manual coding.

▷ 1:18:23

{ } Language Understanding Service (LUIS)

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

A LUIS application is composed of a **schema** →
This schema is autogenerated for you when you use the LUIS.ai web interface

```
{
  "luis_schema_version": "7.0.0",
  "intents": [
    {
      "name": "None",
      "features": []
    }
  ],
  "entities": [],
  "hierarchicals": [],
  "composites": [],
  "closedLists": [],
  "prebuiltEntities": [],
  "utterances": [],
  "versionId": "0.1",
  "name": "example-app",
  "desc": "",
  "culture": "en-us",
  "tokenizerVersion": "1.0.0",
  "patternAnyEntities": [],
  "regex_entities": [],
  "phraselists": [
  ],
  "regex_features": [],
  "patterns": [],
  "settings": []
}
```

- 💡 The schema includes intents, entities, and utterances to define user requests and train a machine learning model for better response matching.

▷ 1:18:31



Language Understanding Service (LUIS)

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

A LUIS application is composed of a **schema**
This schema is autogenerated for you when you use the LUIS.ai web interface

The schema defines:

- **intentions** — what the user is asking for
 - a LUIS app always contains a **None** Intent
- **entities** — what parts of the intent is used to determine the answer
- **utterances** — Examples of user input that includes intent and entities to train the ML model to match predictions against real user input
 - An intent requires one or more example utterance for training
 - It is recommended to have 15-30 example utterances
 - To explicitly train to ignore an utterance use the None Intent

Intents **classify** user utterances
Entities **extract** data from utterance



```
{  
    "luis_schema_version": "7.0.0",  
    "intents": [  
        {  
            "name": "None",  
            "features": []  
        }  
    ],  
    "entities": [],  
    "hierarchicals": [],  
    "composites": [],  
    "closedLists": [],  
    "prebuiltEntities": [],  
    "utterances": [],  
    "versionId": "0.1",  
    "name": "example-app",  
    "desc": "",  
    "culture": "en-us",  
    "tokenizerVersion": "1.0.0",  
    "patternAnyEntities": [],  
    "regex_entities": [],  
    "phraselists": [  
    ],  
    "regex_features": [],  
    "patterns": [],  
    "settings": []  
}
```

💡 The discussion focuses on using classification to identify intents and entities in utterances for machine learning model training.

▷ 1:19:51



QnA Maker Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

QnA Maker is a **cloud-based Natural Language Processing (NLP) service** that allows you **to create a natural conversational layer** over your data.

QnAMaker is hosted on its own isolate domain at www.qnamaker.ai

It will find the most appropriate answer for any input from your **custom knowledge base (KB)** of information

Commonly used to build conversational client applications, which include:

- social media applications
- chat bots
- speech-enabled desktop applications

QnA Maker doesn't store customer data

All customer data is stored in the region the customer deploys the dependent service instances in

▷ 1:20:21



QnA Maker is a **cloud-based Natural Language Processing (NLP)** service that allows you **to create a natural conversational layer** over your data.

QnAMaker is hosted on its own isolate domain at www.qnamaker.ai

It will find the most appropriate answer for any input from your **custom knowledge base (KB)** of information

Commonly used to build conversational client applications, which include:

- social media applications
- chat bots
- speech-enabled desktop applications

QnA Maker doesn't store customer data

All customer data is stored in the region the customer deploys the dependent service instances in

💡 Q&A Maker is a tool that allows users to create custom knowledge bases for conversational applications, utilizing static information from documents like PDFs and URLs, while ensuring customer data remains stored in the user's chosen region.

▷ 1:20:37



When you have static information

Use QnA Maker when you have static information in your knowledge base of answers. This knowledge base is custom to your needs, which you've built with documents such as PDFs and URLs.

When you want to provide the same answer to a request, question, or command

when different users submit the same question, the same answer is returned.

When you want to filter static information based on meta-information

add metadata tags to provide additional filtering options relevant to your client application's users and the information. Common metadata information includes chit-chat, content type or format, content purpose, and content freshness.

When you want to manage a bot conversation that includes static information

your knowledge base takes a user's conversational text or command and answers it. If the answer is part of a pre-determined conversation flow, represented in your knowledge base with multi-turn context, the bot can easily provide this flow.

💡 A customized knowledge base allows for tailored responses to repeated questions by utilizing documents and metadata for filtering information relevant to users.

▷ 1:20:45



QnA Maker Service – Use Cases

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

When you have static information

Use QnA Maker when you have static information in your knowledge base of answers. This knowledge base is custom to your needs, which you've built with documents such as PDFs and URLs.

When you want to provide the same answer to a request, question, or command

when different users submit the same question, the same answer is returned.

When you want to filter static information based on meta-information

add metadata tags to provide additional filtering options relevant to your client application's users and the information. Common metadata information includes chit-chat, content type or format, content purpose, and content freshness.

When you want to manage a bot conversation that includes static information

your knowledge base takes a user's conversational text or command and answers it. If the answer is part of a pre-determined conversation flow, represented in your knowledge base with multi-turn context, the bot can easily provide this flow.

- ★ A knowledge base can manage bot conversations by using predetermined flows of static information to provide answers through a Q&A maker that imports content as question and answer pairs.

▶ 1:21:24



QnA Maker Service – Knowledgebase

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

QnA Maker imports your content into a knowledge base of question and answer pairs.

QnA Maker can build your knowledge base from an **existing document, manual or website (URL, DOCX, PDF)**

It will use ML to extract the question and answer pairs.

The content of the question and answer pair includes:

- All the alternate forms of the question
- Metadata tags used to filter answer choices during the search
- Follow-up prompts to continue the search refinement

QnA Maker stores answer text as **markdown**



- ★ A tool can automatically generate question and answer pairs from a document by extracting headings, text, and images, significantly saving time through machine learning techniques.

▶ 1:21:40



QnA Maker Service – Knowledgebase

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

QnA Maker imports your content into a knowledge base of question and answer pairs.

QnA Maker can build you knowledge base from an **existing document, manual or website (URL, DOCX, PDF)**

It will use ML to extract the question and answer pairs.

The content of the question and answer pair includes:

- All the alternate forms of the question
- Metadata tags used to filter answer choices during the search
- Follow-up prompts to continue the search refinement

QnA Maker stores answer text as **markdown**



- Machine learning is utilized to extract question and answer pairs, which can be refined and edited in a markdown format within a chat interface for improved interaction.

▶ 1:22:11



QnA Maker Service – Knowledgebase

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Once your Knowledge Base is imported you can **fine-tune the imported results** by editing the Question and Answer pairs

Question

Original source: <https://docs.microsoft.com/en-us/azure/cognitive-services/qnamaker/faqs>

I accidentally deleted a part of my QnA Maker, what should I do?

Can I undo deleted questions and answers?

Answer

All deletes are permanent, including question and answer pairs, files, URLs, custom questions and answers, knowledge bases, or Azure resources. Make sure you export your knowledge base from the **Settings** page before deleting any part of your knowledge base.

Metadata tags

Type : troubleshooting

Format : text-only

Nextstep : recover

▶ 1:22:20

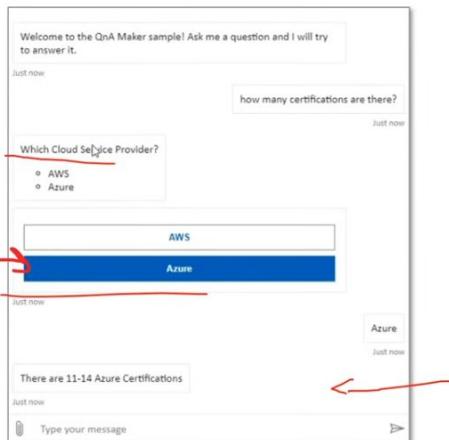
QnA Maker Service – Chat box

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

You converse with your bot through a Chat Box.
There are many opportunities to interact with your
bot in QNAMaker.ai, Azure Bot Service, Bot
Composer.

Via Channels you can even get embeddable
chatbox code

Multi-turn
conversation



- The Q&A maker AI, Azure bot service, and bot composer facilitate interactive conversations where users can ask questions and receive answers in a multi-turn format.

▶ 1:22:28

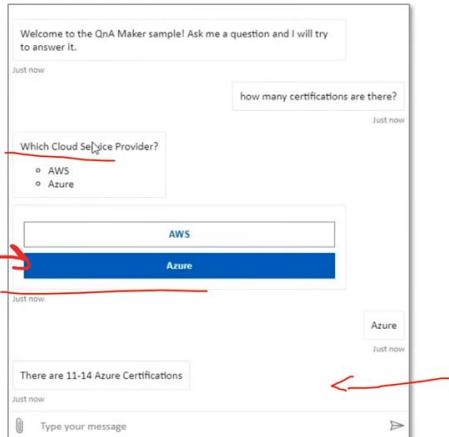
QnA Maker Service – Chat box

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

You converse with your bot through a Chat Box.
There are many opportunities to interact with your
bot in QNAMaker.ai, Azure Bot Service, Bot
Composer.

Via Channels you can even get embeddable
chatbox code

Multi-turn
conversation



▶ 1:22:49

QnA Maker Service – Chit Chat

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Chit-chat

- None
- Professional
- Friendly
- Witty
- Caring
- Enthusiastic

The chit-chat feature in QnA maker allows you to easily add a **pre-populated set of the top chit-chat**, into your knowledge base.

This dataset has about **100 scenarios** of chit-chat in the voice of multiple personas

 1:23:02

QnA Maker Service – Chit Chat

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Chit-chat

- None
- Professional
- Friendly
- Witty
- Caring
- Enthusiastic

The chit-chat feature in QnA maker allows you to easily add a **pre-populated set of the top chit-chat**, into your knowledge base.

This dataset has about **100 scenarios** of chit-chat in the voice of multiple personas

 A feature in Q&A Maker allows for the addition of prepopulated chitchat scenarios with varied persona responses, utilizing a layered ranking approach and storing data in Azure Search.

 1:23:12

QnA Maker Service – Layered Ranking

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

QnA Maker's system is a layered ranking approach.

The data is stored in Azure search, which also serves as the first ranking layer.

The top results from Azure search are then passed through QnA Maker's NLP re-ranking model to produce the final results and confidence score.

- ✿ The process involves using Azure Search for initial ranking, followed by a Q&A maker's NLP reranking model to refine results and manage multi-turn conversations through prompts and active learning.

▶ 1:23:28

QnA Maker Service – Layered Ranking

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

QnA Maker's system is a layered ranking approach.

The data is stored in Azure search, which also serves as the first ranking layer.

The top results from Azure search are then passed through QnA Maker's NLP re-ranking model to produce the final results and confidence score.

▶ 1:23:31

QnA Maker Service – Multi-turn conversation

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Multi-turn conversation is **follow-up prompts** and context to manage the multiple turns, known as *multi-turn*, for your bot from one question to another

When a question **can't be answered in a single turn**

QnA Maker provides multi-turn prompts and active learning to help you improve your basic question and answer pairs.

Multi-turn prompts give you the opportunity to connect question and answer pairs. This connection allows the client application to provide a top answer and provides more questions to refine the search for a final answer.

After the knowledge base receives questions from users at the published endpoint, QnA Maker applies **active learning** to these real-world questions to suggest changes to your knowledge base to improve the quality.

▶ 1:23:47

QnA Maker Service – Multi-turn conversation

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Multi-turn conversation is **follow-up prompts** and context to manage the multiple turns, known as *multi-turn*, for your bot from one question to another

When a question **can't be answered in a single turn**

QnA Maker provides multi-turn prompts and active learning to help you improve your basic question and answer pairs.

Multi-turn prompts give you the opportunity to connect question and answer pairs. This connection allows the client application to provide a top answer and provides more questions to refine the search for a final answer.

After the knowledge base receives questions from users at the published endpoint, QnA Maker applies **active learning** to these real-world questions to suggest changes to your knowledge base to improve the quality.

▶ 1:23:48

QnA Maker Service – Multi-turn conversation

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Multi-turn conversation is **follow-up prompts** and context to manage the multiple turns, known as *multi-turn*, for your bot from one question to another

When a question **can't be answered in a single turn**

QnA Maker provides multi-turn prompts and active learning to help you improve your basic question and answer pairs.

Multi-turn prompts give you the opportunity to connect question and answer pairs. This connection allows the client application to provide a top answer and provides more questions to refine the search for a final answer.

After the knowledge base receives questions from users at the published endpoint, QnA Maker applies **active learning** to these real-world questions to suggest changes to your knowledge base to improve the quality.

- ★ A system utilizes key and answer pairs to enhance search results through active learning, allowing for improved knowledge base quality by refining user questions.

▶ 1:24:12

Azure Bot Service

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900



Azure Bot Service Intelligent, serverless bot service that scales on demand.
Used for **creating, publishing, and managing bots**

You can **register and publish** a variety of bots from the Azure Portal

Azure Bot Service can integrate your bot with other Azure, Microsoft or Third Party services via **Channels**:

- Direct Line
- Alexa
- Office 365 email
- Facebook
- Kik
- LINE
- Microsoft Teams
- Skype
- Twilio
- ande more....

The screenshot shows a list of bot channels categorized into two columns:

Speech to text	Web App Bot
Cerebral Technologies Private Limited	Microsoft
Think AI Bot for Connectwise	Microsoft
Veracular.ai Intelligent Voice Assistant	Veracular.ai
devNXT- AI driven smart application development	Wipro Ltd
RxHealth	RxHealth
Zammo Ai SaaS	Zammo, Inc.
Audite Cloud	TALENTUM
Ma - Workplace Virtual Assistant	MHCM
Azure Health Bot	Microsoft

▶ 1:24:43



Azure Bot Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Azure Bot Service Intelligent, serverless bot service that scales on demand.
Used for **creating, publishing, and managing bots**

You can **register and publish** a variety of bots from the Azure Portal

Azure Bot Service can integrate your bot with other Azure, Microsoft or Third Party services via **Channels**:

- Direct Line
- Alexa
- Office 365 email
- Facebook
- Kik
- LINE
- Microsoft Teams
- Skype
- Twilio
- ande more....

Speech to text Cerebral Technologies Private Limited	Web App Bot Microsoft
Think AI Bot for Connectwise Think AI Consulting Corporation	Bot Channels Registration Microsoft
Veneracular.ai Intelligent Voice Assistant Veneracular.ai	Azure Bot Microsoft
devNXT- AI driven smart application development Wipro Ltd	
RxHealth RxHealth	
Zammo Ai SaaS Zammo, Inc.	
Audite Cloud TALENTUM	
Mia - Workplace Virtual Assistant MICROM	
Azure Health Bot Microsoft	

- ✿ The Azure Bot Service integrates with various third-party services and is commonly associated with the Bot Framework SDK and Bot Composer.

▷ 1:25:11

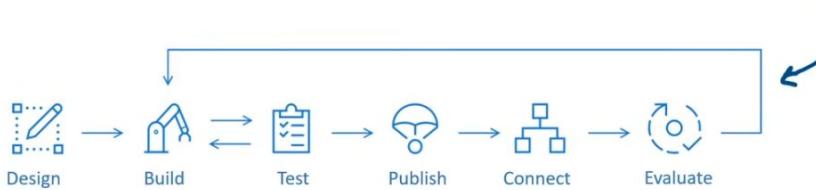


Bot Framework SDK

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

The Bot Framework SDK v4 is an **open-source SDK** that enable developers to **model and build sophisticated conversations**

The Bot Framework, along with the Azure Bot Service, provides an **end-to-end workflow**:



With this framework, developers can create bots that use speech, understand natural language, handle questions and answers, and more.

The Bot Framework includes a modular and extensible SDK for building bots, as well as tools, templates, and related AI services.

- ✿ The bot framework and Azure bot service offer a comprehensive workflow for designing, building, testing, publishing, connecting, and evaluating bots that utilize speech understanding and natural language processing.

▷ 1:25:31

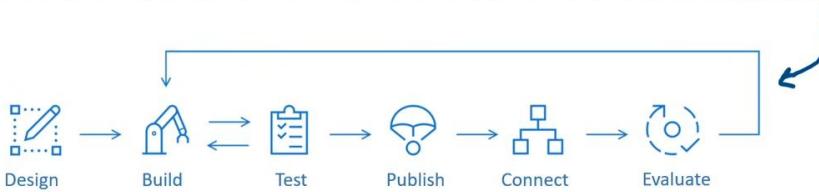


Bot Framework SDK

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

The Bot Framework SDK v4 is an **open-source SDK** that enable developers to **model and build sophisticated conversations**

The Bot Framework, along with the Azure Bot Service, provides and **end-to-end workflow**:



With this framework, developers can create bots that use speech, understand natural language, handle questions and answers, and more.

The Bot Framework includes a modular and extensible SDK for building bots, as well as tools, templates, and related AI services.

▶ 1:25:54



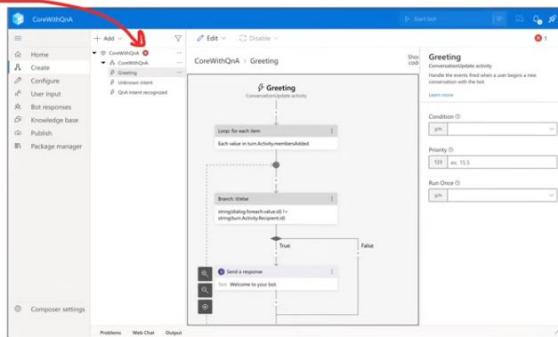
Bot Framework Composer

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Bot Framework Composer, built on the Bot Framework SDK, is an **open-source IDE for developers** to **author, test, provision and manage** conversational experiences.

Composer is downloadable app available for Windows, OSX and Linux

- You can use either C# or Node to build your bot
- Deploy your bots to:
 - Azure Web App
 - Azure Functions
- Templates to build:
 - QnA Maker Bot
 - Enterprise or Personal Assistant Bot
 - Language Bot
 - Calendar or People Bot
- Test and debug via the Bot Framework Emulator
- Built in Package manager



▶ 1:26:20



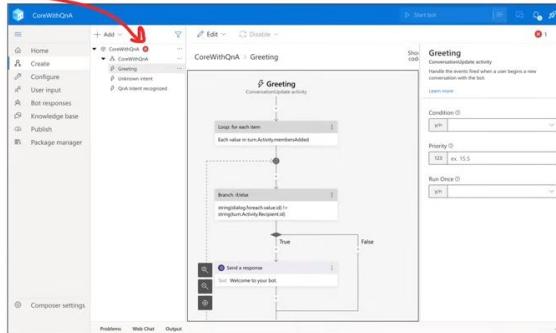
Bot Framework Composer

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Bot Framework Composer, built on the Bot Framework SDK, is an **open-source IDE for developers** to **author, test, provision and manage** conversational experiences.

Composer is downable app available for Windows, OSX and Linux



- You can use either C# or Node to build your bot
- Deploy your bots to:
 - Azure Web App
 - Azure Functions
- Templates to build:
 - QnA Maker Bot
 - Enterprise or Personal Assistant Bot
 - Language Bot
 - Calendar or People Bot
- Test and debug via the Bot Framework Emulator
- Built in Package manager

💡 Azure offers templates for building various types of bots, including Q&A maker and personal assistant bots, along with features like testing via the bot framework emulator and a built-in package manager.

▶ 1:26:40



Azure Machine Learning Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure Machine Learning Studio (classic)

An older service that manages AI/ML workloads. Does not have a pipeline and other limitations.
Workloads are not easily transferable from classic to the new service.



Azure Machine Learning Service

A service that simplifies running AI/ML related workloads allowing you to build flexible Automated ML Pipelines. Use Python or R, Run DL workloads such as Tensorflow

Jupyter Notebooks

- build and document your machine learning models as you build them, share and collaborate

Azure Machine Learning SDK for Python

- An SDK designed specifically to interact with Azure Machine Learning Services

MLOps

- end to end automation of ML model pipelines eg. CI/CD, training, inference

Azure Machine Learning Designer

- drag and drop interface to visually build, test, and deploy machine learning models

Data Labeling Service

- ensemble a team of humans to label your training data

Responsible Machine Learning

- model fairness through disparity metrics and mitigate unfairness

💡 The Azure Machine Learning service simplifies AI and machine learning workloads by enabling the creation of automated ML pipelines, supporting Python and deep learning frameworks like TensorFlow, and allowing for the use of Jupyter notebooks.

▶ 1:26:56



Azure Machine Learning Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure Machine Learning Studio (classic)

An older service that manages AI/ML workloads. Does not have a pipeline and other limitations.
Workloads are not easily transferable from classic to the new service.



Azure Machine Learning Service

A service that simplifies running AI/ML related workloads allowing you to build flexible
Automated ML Pipelines. Use Python or R, Run DL workloads such as Tensorflow

Jupyter Notebooks

- build and document your machine learning models as you build them, share and collaborate

Azure Machine Learning SDK for Python

- An SDK designed specifically to interact with Azure Machine Learning Services

MLOps

- end to end automation of ML model pipelines eg. CI/CD, training, inference

Azure Machine Learning Designer

- drag and drop interface to visually build, test, and deploy machine learning models

Data Labeling Service

- ensemble a team of humans to label your training data

Responsible Machine Learning

- model fairness through disparity metrics and mitigate unfairness

- A drag-and-drop interface will be developed to create machine learning model pipelines and a data labeling service, emphasizing responsible AI practices and fairness metrics.

▶ 1:28:02

Azure Machine Learning Studio – Overview

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Author
Notebooks
Automated ML
Designer
Assets
Datasets
Experiments
Pipelines
Models
Endpoints
Manage
Compute
Environments (preview)
Datastores
Data Labeling
Linked Services

Author

Notebooks — Jupyter Notebooks, an IDE to write python code to build ML models

AutoML — Completely automated process to build and train an ML model

Designer — Visual drag and drop designer to construct end to end ML pipelines

Assets

Datasets — data that you upload which will be used for training

Experiments — when you run a training job they are detailed here

Pipelines — ML workflows you have built, or you have used in the Designer

Models — a model registry containing trained models that can be deployed

Endpoints — when you deploy a model its hosted on an accessible endpoint eg. REST API

- The tools available for building machine learning models include Jupyter notebooks and IDEs for coding, an automated process with AutoML, a visual drag-and-drop designer for constructing ML pipelines, and the ability to upload datasets.

▶ 1:28:21

Azure Machine Learning Studio – Overview

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Author
Notebooks
Automated ML
Designer
Assets
Datasets
Experiments
Pipelines
Models
Endpoints
Manage
Compute
Environments (preview)
Datastores
Data Labeling
Linked Services

Author

Notebooks — Jupyter Notebooks, an IDE to write python code to build ML models

AutoML — Completely automated process to build and train an ML model

Designer — Visual drag and drop designer to construct end to end ML pipelines

Assets

Datasets — data that you upload which will be used for training

Experiments — when you run a training job they are detailed here

Pipelines — ML workflows you have built, or you have used in the Designer

Models — a model registry containing trained models that can be deployed

Endpoints — when you deploy a model its hosted on an accessible endpoint eg. REST API

Manage

Compute — the underlying computing instances used to for notebooks, training, inference

Environments — a reproducible Python environment for machine learning experiments

Datastores — a data repository where your dataset resides

Data Labeling — have humans with ML-assisted labeling to label your data for supervised learning

Linked Services — external services you can connect to the workspace eg. Azure Synapse Analytics

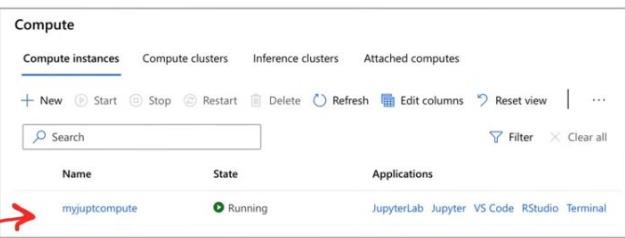
- A reproducible Python environment for machine learning experiments includes data storage, human-assisted labeling for supervised learning, and connections to external services like Azure Synapse Analytics, along with various types of compute options available in Azure Machine Learning Studio.

▶ 1:29:35

Azure Machine Learning Studio – Compute

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure Machine Learning Studio has **4 kinds of compute**:



Name	State	Applications
myjuptcompute	Running	JupyterLab

1. **Compute Instances** — Development workstations that data scientists can use to work with data and models.
2. **Compute Clusters** — Scalable clusters of virtual machines for on-demand processing of experiment code.
3. **Inference Clusters** — Deployment targets for predictive services that use your trained models.
4. **Attached Compute** — Links to existing Azure compute resources, such as Virtual Machines or Azure Databricks clusters.

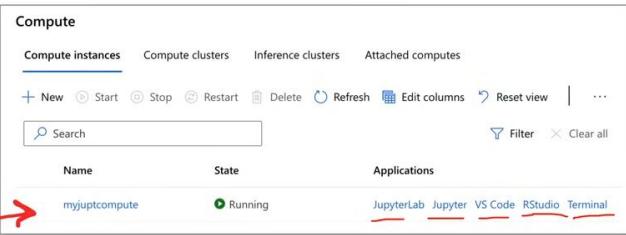
- Data scientists can utilize scalable compute clusters, including Azure VMs and Databricks, for on-demand processing and experimentation with trained models through various platforms like Jupyter Labs, VS Code, R Studio, and terminal.

▶ 1:29:53

Azure Machine Learning Studio – Compute

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure Machine Learning Studio has **4 kinds of compute:**



Name	State	Applications
myjuptcompute	Running	JupyterLab Jupyter VS Code RStudio Terminal

- 1. Compute Instances** — Development workstations that data scientists can use to work with data and models.
- 2. Compute Clusters** — Scalable clusters of virtual machines for on-demand processing of experiment code.
- 3. Inference Clusters** — Deployment targets for predictive services that use your trained models.
- 4. Attached Compute** — Links to existing Azure compute resources, such as Virtual Machines or Azure Databricks clusters.

▶ 1:30:20

Azure Machine Learning Studio – Data Labeling

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Create Data labeling jobs to prepare your Ground Truth
for supervised learning

Human-in-the-loop labeling

You have a team of humans that will apply labeling
These are humans you grant access to labeling

Machine-learning-assisted data labeling

You will use ML to perform labeling

You can export the label data for Machine Learning experimentation at any time
Users often export multiple times and train different models, rather than wait for all the
images to be labeled.
Image labels can be exported in:

- COCO format
- Azure Machine Learning dataset
 - dataset format makes it easy to use for training in Azure Machine Learning

- 💡 A team of humans will label data, while machine learning assists in daily labeling, allowing for the export of labeled data in COCO format for experimentation and model training.

▶ 1:30:59

Azure Machine Learning Studio – Data Labeling

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Create Data labeling jobs to prepare your Ground Truth for supervised learning

Human-in-the-loop labeling

You have a team of humans that will apply labeling. These are humans you grant access to labeling.

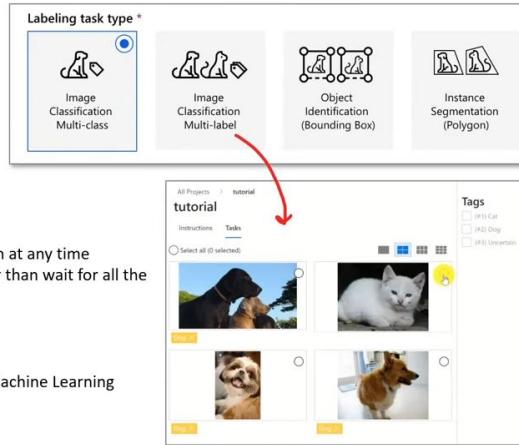
Machine-learning-assisted data labeling

You will use ML to perform labeling

You can export the label data for Machine Learning experimentation at any time. Users often export multiple times and train different models, rather than wait for all the images to be labeled.

Image labels can be exported in:

- COCO format
- Azure Machine Learning dataset
 - dataset format makes it easy to use for training in Azure Machine Learning



- 💡 A specific data set format is recommended for training in Azure Machine Learning, allowing users to easily perform labeling tasks through a user interface while securely connecting to Azure storage services without exposing authentication credentials.

▷ 1:31:40

Azure Machine Learning Studio – Data Stores

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Datastores **securely connect to your storage service** on Azure without **putting your authentication credentials** and the integrity of your original data source **at risk**.

Datastore type *
Azure Blob Storage
Azure Blob Storage
Azure file share
Azure Data Lake Storage Gen1
Azure Data Lake Storage Gen2
Azure SQL database
Azure PostgreSQL database
Azure MySQL database

Azure Blob Storage	data is stored as objects, distributed across many machines
Azure File Share	a mountable file share via SMB and NFS protocols
Azure Data Lake Storage (Gen 2)	Azure Blob storage designed for vasts amount of data for Big Data analytics
Azure SQL database	Full-managed MS SQL relational database
Azure Postgres database	open-source relational database
Azure MySQL Database	Open-source relational database

- 💡 The available data sources include Azure Blob Storage, Azure File Share, Azure Data Lake Storage Gen 2, Azure SQL Database, PostgreSQL, and MySQL.

▷ 1:31:57

Azure Machine Learning Studio – Data Stores

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Datastores securely connect to your storage service on Azure without putting your authentication credentials and the integrity of your original data source at risk.

Datastore type *
Azure Blob Storage
Azure Blob Storage
Azure file share
Azure Data Lake Storage Gen1
Azure Data Lake Storage Gen2
Azure SQL database
Azure PostgreSQL database
Azure MySQL database

Azure Blob Storage
data is stored as objects, distributed across many machines

Azure File Share
a mountable file share via SMB and NFS protocols

Azure Data Lake Storage (Gen 2)
Azure Blob storage designed for vasts amount of data for Big Data analytics

Azure SQL database
Full-managed MS SQL relational database

Azure Postgres database
open-source relational database

Azure MySQL Database
Open-source relational database

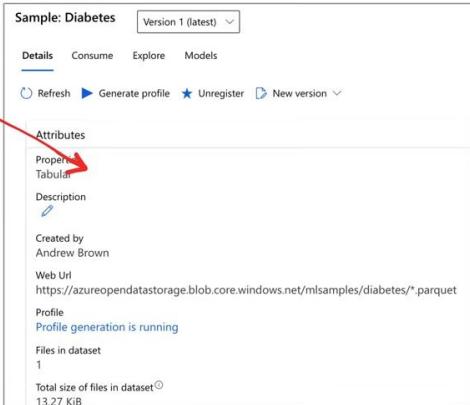
- Managed MSSQL and MySQL are open-source relational databases, with the latter being the most popular, while Azure ML facilitates easy registration of datasets for machine learning workloads by providing associated metadata.

▶ 1:32:32

Azure Machine Learning Studio – Datasets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure ML Datasets makes it easy to register your datasets for use with your ML workloads



The screenshot shows the 'Sample: Diabetes' dataset page. At the top, a note says 'There will be various metadata associated to your dataset'. On the left, a sidebar lists options to upload new datasets from local files, datastores, web files, or Open Datasets. A callout box highlights 'You can upload new datasets and they will be versioned'. On the right, the dataset details are shown: Version 1 (latest). The 'Attributes' section includes 'Properties' and 'Tabular'. Below that is 'Description' with a link. The 'Created by' field shows 'Andrew Brown'. The 'Web Url' is 'https://azureopendatastorage.blob.core.windows.net/mlsamples/diabetes/*_parquet'. The 'Profile' section indicates 'Profile generation is running'. It also shows 'Files in dataset 1' and 'Total size of files in dataset 13.27 KB'. A red arrow points from the 'versioned' text in the sidebar to the 'Version 1 (latest)' dropdown. Another red arrow points from the 'From local files' option in the sidebar to the 'Current version 1' and 'Latest version 1' labels in the main panel.

- You can register datasets for machine learning workloads, manage multiple versions, and utilize sample code from the Azure ML SDK in Jupyter notebooks to generate data profiles.

▶ 1:32:53

Azure Machine Learning Studio – Datasets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure provides a same code snippet with the **Azure Machine Learning SDK for Python** to start programmatically using datasets in your Jupyter Notebooks

```
Sample usage   
# azureml-core of version 1.0.72 or higher is required  
# azureml-dataprep[pandas] of version 1.1.34 or higher is required  
from azureml.core import Workspace, Dataset  
  
subscription_id = '7f3352cf-6c7d-456a-8ecb-83ef2128907b'  
resource_group = 'MyStudio'  
workspace_name = 'MyStudio'  
  
workspace = Workspace(subscription_id, resource_group, workspace_name)  
  
dataset = Dataset.get_by_name(workspace, name='Sample: Diabetes')  
dataset.to_pandas_dataframe()
```

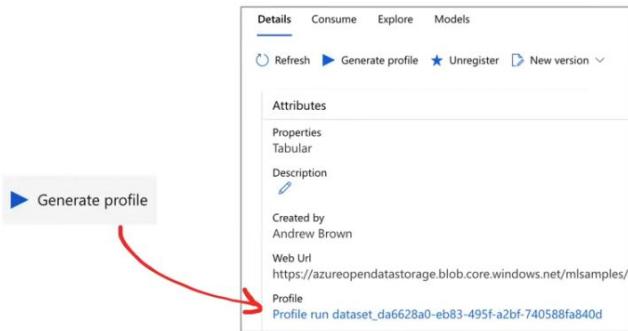
▶ 1:32:56

Azure Machine Learning Studio – Datasets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Generate Profile

You can create a data profile that has summary statistics, distribution of the data and more. You will need to launch a compute instance to generate a profile.



The screenshot shows the 'Details' tab of a dataset in Azure Machine Learning Studio. At the top, there are tabs for 'Details', 'Consume', 'Explore', and 'Models'. Below the tabs, there are buttons for 'Refresh', 'Generate profile' (which is highlighted with a red arrow), 'Unregister', and 'New version'. Under the 'Attributes' section, there are fields for 'Properties' (set to 'Tabular'), 'Description' (with a blue edit icon), and 'Created by' (set to 'Andrew Brown'). Below that is a 'Web Url' field containing a link to a blob storage location. At the bottom, there is a 'Profile' section with a link to a specific profile run.

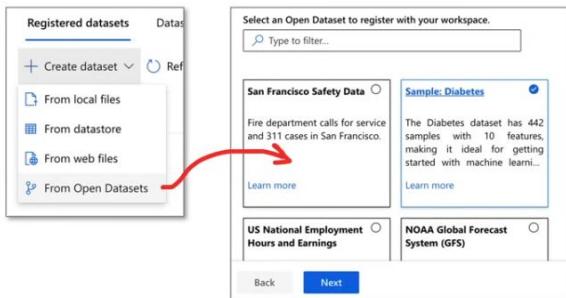
- 💡 You can generate profiles and summary statistics from data sets using a compute instance, with the results stored in Blob storage, while also accessing curated open data sets for learning machine learning model development.

▶ 1:33:04

Azure Machine Learning Studio – Open Datasets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Open DataSets are **publicly hosted datasets** that are commonly **used for learning how to build ML models**



Azure has a curated list of open-datasets that you can quickly add to your data store.

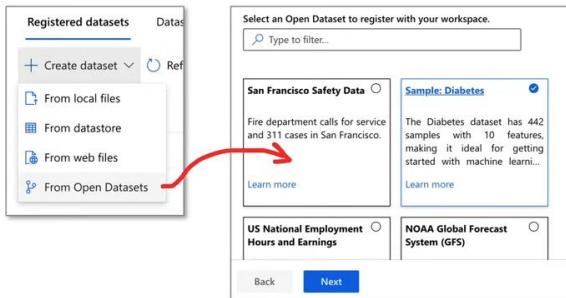
Great for learning how to use AutoML or Azure Machine Learning Designer

▶ 1:33:26

Azure Machine Learning Studio – Open Datasets

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Open DataSets are **publicly hosted datasets** that are commonly **used for learning how to build ML models**



Azure has a curated list of open-datasets that you can quickly add to your data store.

Great for learning how to use AutoML or Azure Machine Learning Designer

- 💡 A curated list of open datasets is available for learning how to build machine learning models, particularly useful for beginners using automated ML tools.

▶ 1:33:35

Azure Machine Learning Studio – Experiments

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Experiments is a **logical grouping Azure Runs**

Runs are the **act of running an ML task** on a virtual machine or container



Experiment	Latest run	Last submitted ↓	Created	Created by	Run types
dataset_profile	1	Jun 5, 2021 10:30 AM	Jun 5, 2021 10:30 AM	Andrew Brown	Script
MyDiabetes	2	Jun 4, 2021 4:39 PM	Jun 4, 2021 4:36 PM	Andrew Brown	Automated ML
MyExperiment	1	Jun 4, 2021 4:19 PM	Jun 4, 2021 4:19 PM	Andrew Brown	Pipeline

Experiments do not include *Inference*

The contents of a run will vary based on its **Run Type**

- Azure ML pipelines facilitate the execution of complete machine learning tasks, excluding inference after model deployment.

▶ 1:34:14

Azure Machine Learning Studio – Pipelines

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure ML Pipelines is an **executable workflow** of a complete machine learning task

Subtasks are encapsulated as a series of steps within the pipeline

Independent steps allow multiple data scientists to work on the same pipeline at the same time without over-taxing compute resources

Separate steps also make it easy to use different compute types/sizes for each step.

When you rerun a pipeline, the run jumps to the steps that need to be rerun, such as an updated training script.

Steps that do not need to be rerun are skipped

After a pipeline has been published, you can configure a REST endpoint, which allows you to rerun the pipeline from any platform or stack

You can build pipelines two ways:

- Using the Azure Machine Learning Designer
- Programmatically using the Azure Machine Learning Python SDK**

- Subas are independent, encapsulated steps within a pipeline that enable multiple data scientists to collaborate simultaneously while optimizing compute resources and allowing for selective reruns of updated steps.

▶ 1:34:28

Azure Machine Learning Studio – Pipelines

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure ML Pipelines is an **executable workflow** of a complete machine learning task

Subtasks are encapsulated as a series of steps within the pipeline

Independent steps allow multiple data scientists to work on the same pipeline at the same time without over-taxing compute resources

Separate steps also make it easy to use different compute types/sizes for each step.

When you rerun a pipeline, the run jumps to the steps that need to be rerun, such as an updated training script.

Steps that do not need to be rerun are skipped

After a pipeline has been published, you can configure a REST endpoint, which allows you to rerun the pipeline from any platform or stack

You can build pipelines two ways:

- Using the Azure Machine Learning Designer
- **Programmatically using the Azure Machine Learning Python SDK**

```
ws = Workspace.from_config()
blob_store = Datastore(ws, "workspaceblobstore")
compute_target = ws.compute_targets["STANDARD_NC6"]
experiment = Experiment(ws, 'MyExperiment')

input_data = Dataset.File.from_files(
    DataPath(blob_store, '20newsgroups/20news.pkl'))
prepped_data_path = OutputFileDatasetConfig(name="output_path")

dataprep_step = PythonScriptStep(
    name="prep_data",
    script_name="dataprep.py",
    source_directory="prep_src",
    compute_target=compute_target,
    arguments=["--prepped_data_path", prepped_data_path],
    inputs=[input_data.as_named_input('raw_data').as_mount()])

prepped_data = dataprep_step.run().outputs['output_path']

train_step = PythonScriptStep(
    name="train",
    script_name="train.py",
    compute_target=compute_target,
    arguments=["--prepped_data", prepped_data],
    source_directory="train_src"
)
steps = [dataprep_step, train_step]
pipeline = Pipeline(workspace=ws, steps=steps)
pipeline_run = experiment.submit(pipeline)
pipeline_run.wait_for_completion()
```

⊕ Pipelines can be built using Azure ML Designer or programmatically with the Azure Machine Learning Python SDK by creating and assembling steps.

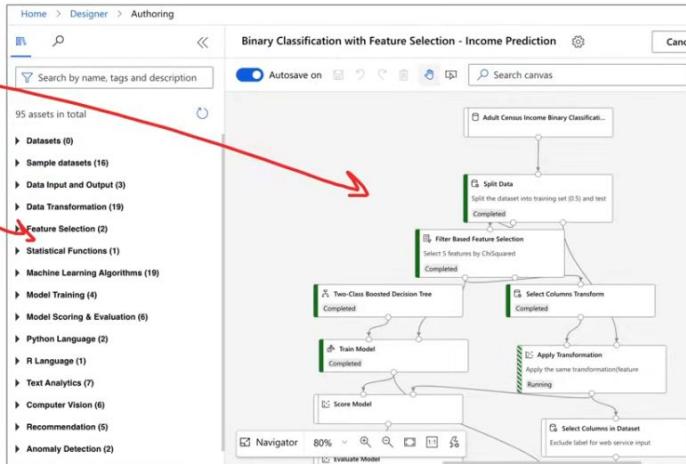
▷ 1:35:15

Azure Machine Learning Studio – ML Designer

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

The Azure Machine Learning Designer lets you **quickly build Azure ML Pipelines** without having to write code.

You can **drag out various templated steps** called assets to quickly prototype your pipeline



⊕ The service offers a visual interface for building machine learning pipelines with pre-built assets, allowing users to create both real-time and batch inference pipelines while requiring a solid understanding of end-to-end ML processes.

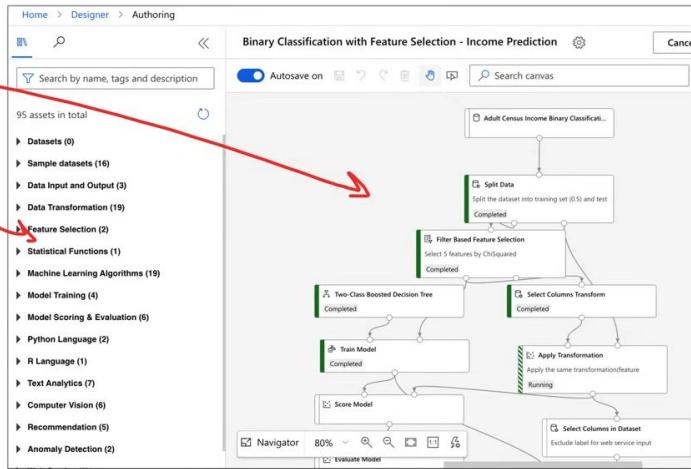
▷ 1:35:34

Azure Machine Learning Studio – ML Designer

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

The Azure Machine Learning Designer lets you **quickly build Azure ML Pipelines** without having to write code.

You can **drag out various templated steps** called assets to quickly prototype your pipeline

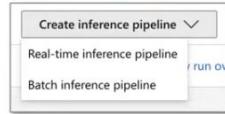


▶ 1:35:46

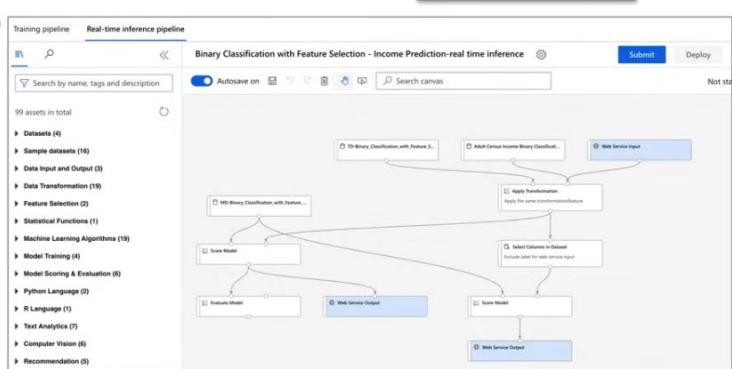
Azure Machine Learning Studio – Machine Learning Designer

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Once a pipeline is trained you can create an **inference pipeline**



You can **toggle** between your training in inference pipeline



💡 The model registry enables the creation, management, and tracking of registered models as incremental versions under the same name, allowing for metadata tagging and improved search capabilities.

▶ 1:35:57

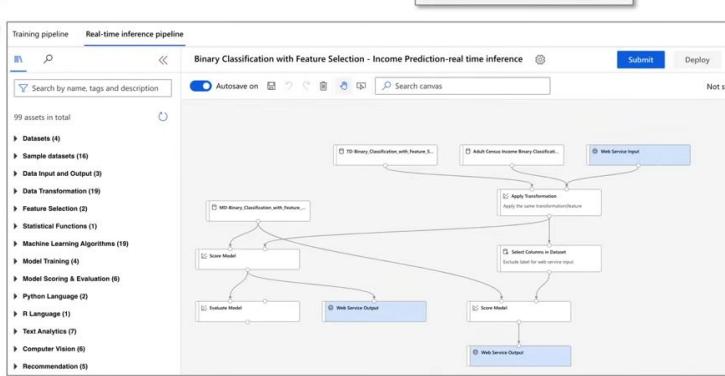
Azure Machine Learning Studio – Machine Learning Designer

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Once a pipeline is trained you can create an **inference pipeline**



You can **toggle** between your training in inference pipeline



▶ 1:35:59

Azure Machine Learning Studio – Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Model Registry allows you to **create, manage and track your registered models** as incremental versions under the same name

Each time you register a model with the same name as an existing one, the registry assures that **it's a new version.**

Additionally, you can provide metadata tags and use the tags when you search for models.

- Machine learning models can be easily shared, deployed, and downloaded by registering them with metadata tags, and deploying them as a web service involves several steps including model registration, script preparation, configuration setup, local deployment testing, and selecting a compute target.

▶ 1:36:20

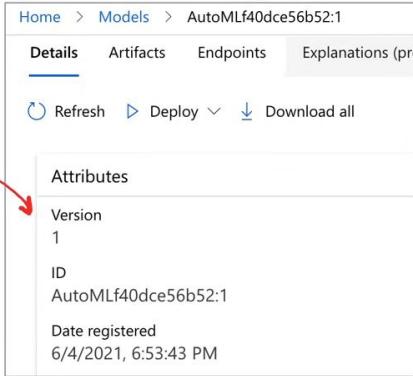
Azure Machine Learning Studio – Models

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Model Registry allows you to **create, manage and track your registered models** as incremental versions under the same name

Each time you register a model with the same name as an existing one, the registry assures that **it's a new version.**

Additionally, you can provide metadata tags and use the tags when you search for models.



The screenshot shows a detailed view of a registered model in the Azure Machine Learning Studio. The URL in the address bar is "Home > Models > AutoMLf40dce56b52:1". The page has tabs for "Details", "Artifacts", "Endpoints", and "Explanations (pre)". Below the tabs, there are buttons for "Refresh", "Deploy", and "Download all". The main content area is titled "Attributes" and lists the following information:

Version	1
ID	AutoMLf40dce56b52:1
Date registered	6/4/2021, 6:53:43 PM

▶ 1:36:27

Azure Machine Learning Studio – Endpoints

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure ML Endpoints allow you to **deploy machine learning models as a web service**

The workflow for deploying a model:

- Register the model
- Prepare an entry script
- Prepare an inference configuration
- Deploy the model locally to ensure everything works
- Choose a compute target
- Re-deploy the model to the cloud
- Test the resulting web service

Realtime endpoints

An endpoint that provides remote access to invoke the ML model service running on either:

- Azure Kubernetes Service (AKS)
- Azure Container Instance (ACI)

Pipeline endpoints

An endpoint that provides remote access to invoke an ML pipeline.

You can parametrize the pipeline endpoint for managed repeatability in batch scoring and retraining scenarios.

💡 The process involves preparing an inference configuration, deploying a model locally for testing, choosing a compute target, redeploying the model to the cloud, and then testing it through either real-time or pipeline endpoints.

▶ 1:36:44

Azure Machine Learning Studio – Endpoints

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure ML Endpoints allow you to deploy machine learning models as a web service

The workflow for deploying a model:

- Register the model
- Prepare an entry script
- Prepare an inference configuration
- Deploy the model locally to ensure everything works
- Choose a compute target
- Re-deploy the model to the cloud
- Test the resulting web service

Realtime endpoints

An endpoint that provides remote access to invoke the ML model service running on either:

- Azure Kubernetes Service (AKS)
- Azure Container Instance (ACI)

Pipeline endpoints

An endpoint that provide remote access to invoke an ML pipeline.

You can parametrize the pipeline endpoint for managed repeatability in batch scoring and retraining scenarios.

- Models can be deployed to an endpoint on either Azure Kubernetes Service (AKS) or Azure Container Instances (ACI), with the deployment primarily reflected under AKS.

▶ 1:37:16

Azure Machine Learning Studio – Endpoints

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

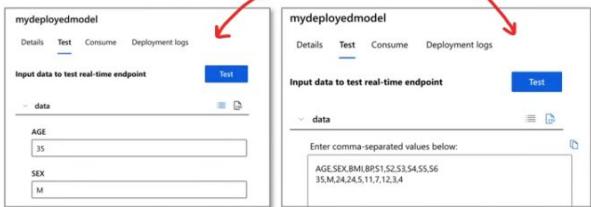
When you **deploy a model** to an endpoint it will either be deployed to:
• Azure Kubernetes Service (AKS)
• Azure Container Instance (ACI)

Compute type * ⓘ

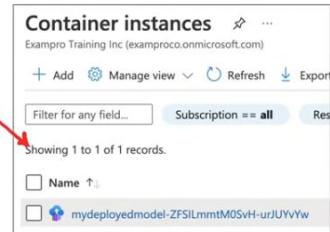
The computing resource will not show in Azure Machine Learning Studio

You need to check AKS or ACI

When you have deployed a real-time endpoint you can test the endpoint by sending a **single request** or a **batch request**.



Two screenshots of the 'mydeployedmodel' endpoint page. Both show a 'Test' button. The first screenshot shows input fields for AGE (35) and SEX (M). The second screenshot shows input data: 'AGE,SEX,BMI,BPS1,S2,S3,S4,S5,S6' and '35,M,24,24,5,11,7,12,3,4'.



Container instances

Name
mydeployedmodel-ZFSILmmmtM0SvH-urJUYvYw

- Azure provides a built-in Jupyter-like notebook editor for testing real-time endpoints with single or batch requests, which are displayed under AKS or ACI in the Azure portal rather than being consolidated in Azure Machine Learning Studio.

▶ 1:37:26

Azure Machine Learning Studio – Endpoints

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

- When you **deploy a model** to an endpoint it will either be deployed to:
- Azure Kubernetes Service (AKS)
 - Azure Container Instance (ACI)

Compute type * ⓘ

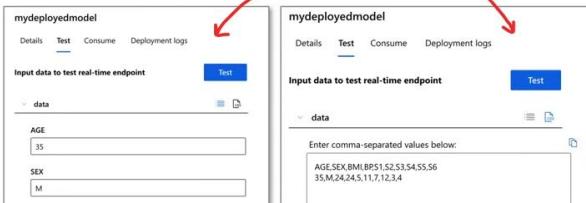
Azure Kubernetes Service

Azure Container Instance

The computing resource will not show in Azure Machine Learning Studio

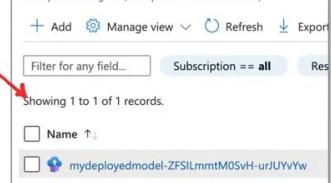
You need to check AKS or ACI

When you have deployed a real-time endpoint you can test the endpoint by sending a **single request** or a **batch request**.



Two screenshots of the Azure Machine Learning Studio interface for a deployed model named "mydeployedmodel". Both screenshots show a "Test" button. The first screenshot shows a "Single request" input field with "AGE: 35" and "SEX: M". The second screenshot shows a "Batch request" input field with comma-separated values: "AGE,SEX,BMIBPS1,52,53,54,55,56,35,M,24,24,5,11,7,12,3,4".

Container instances



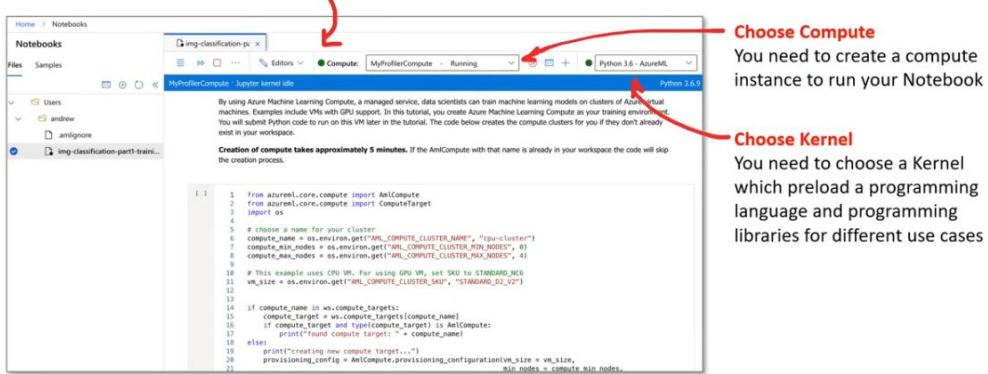
A screenshot of the "Container instances" view in the Azure Machine Learning Studio. It shows a single record: "mydeployedmodel-ZFSILmmtMOSvH-urJUYvYw".

▶ 1:37:46

Azure Machine Learning Studio – Notebooks

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Azure has a built in **Jupyter-like Notebook editor** so you can build and train your ML models



A screenshot of the Azure Machine Learning Studio Jupyter Notebook editor. The top navigation bar shows "Notebooks" and "MyProfileCompute - Jupyter kernel idle". The main area shows a code cell with Python code for creating an Azure Compute instance. The code includes imports for "azml", "Compute", and "ComputeTarget". It sets environment variables for cluster name and min nodes, and creates a new compute target if none exists. The code cell is labeled "Python 3.6 - AzureML". The status bar at the bottom indicates the cell is "Running". A red arrow points from the "Compute" dropdown to the text "Choose Compute". Another red arrow points from the "Kernel" dropdown to the text "Choose Kernel".

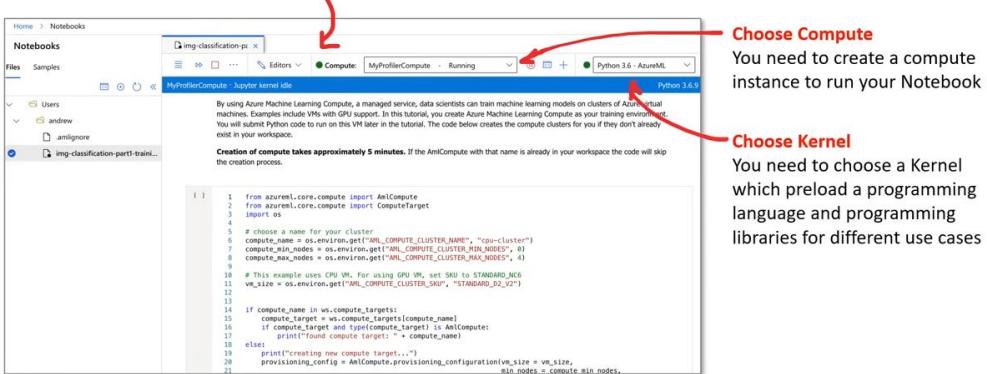
```
1 # choose a name for your cluster
2 compute_name = os.environ.get("AML_COMPUTE_CLUSTER_NAME", "cpu-cluster")
3 compute_min_nodes = os.environ.get("AML_COMPUTE_CLUSTER_MIN_NODES", 0)
4 compute_max_nodes = os.environ.get("AML_COMPUTE_CLUSTER_MAX_NODES", 4)
5
6 # This example uses CPU VM. For using GPU VM, set SKU to STANDARD_NC1
7 vm_size = os.environ.get("AML_COMPUTE_CLUSTER_SKU", "STANDARD_NC1")
8
9 if compute_name in ws.compute_targets:
10     compute_target = ws.compute_targets[compute_name]
11     if compute_target.type == "ComputeTarget":
12         print("found compute target: " + compute_name)
13     else:
14         print("creating new compute target...")
15         provisioning_config = ComputeTargetException(ws)
16         provisioning_config.vm_size = vm_size,
17         provisioning_config.min_nodes = compute_min_nodes,
```

▶ 1:38:04

Azure Machine Learning Studio – Notebooks

Cheat sheets, Practice Exams and Flash cards  www.exampopro.co/ai-900

Azure has a built in **Jupyter-like Notebook editor** so you can build and train your ML models



The screenshot shows the Azure Machine Learning Studio interface for Notebooks. At the top, there's a navigation bar with 'Home > Notebooks'. Below it, a sidebar shows a file structure under 'Users' and 'andrew'. The main area displays a Jupyter-like notebook titled 'MyProfilerCompute - Jupyter kernel idle'. A red arrow points from the 'Compute' dropdown menu at the top right to a callout box labeled 'Choose Compute'. Another red arrow points from the 'Python 3.6 - AzureML' dropdown menu to a callout box labeled 'Choose Kernel'.

Choose Compute
You need to create a compute instance to run your Notebook

Choose Kernel
You need to choose a Kernel which preload a programming language and programming libraries for different use cases

```
1 from azureml.core.compute import AmlCompute
2 from azureml.core.compute import ComputeTargetException
3
4 # choose a name for your cluster
5 compute_name = os.environ.get("AML_COMPUTE_CLUSTER_NAME", "cpu-cluster")
6 compute_min_nodes = os.environ.get("AML_COMPUTE_CLUSTER_MIN_NODES", 0)
7 compute_max_nodes = os.environ.get("AML_COMPUTE_CLUSTER_MAX_NODES", 4)
8
9 # This example uses a CPU VM. For using GPU VM, set SKU to STANDARD_NC6
10 vm_size = os.environ.get("AML_COMPUTE_CLUSTER_SKU", "STANDARD_D2_V2")
11
12
13
14 if compute_name in ws.compute_targets:
15     compute_target = ws.compute_targets[compute_name]
16     if compute_target and type(compute_target) is ComputeTargetException:
17         print("Compute target {} already exists".format(compute_name))
18     else:
19         print("Creating new compute target...")
20         provisioning_config = AmlCompute.provisioning_configuration(vm_size = vm_size,
21                         min_nodes = compute_min_nodes,
```

- You can select a compute instance and kernel to run a notebook, which can be opened in familiar IDEs like VS Code or Jupyter Lab.

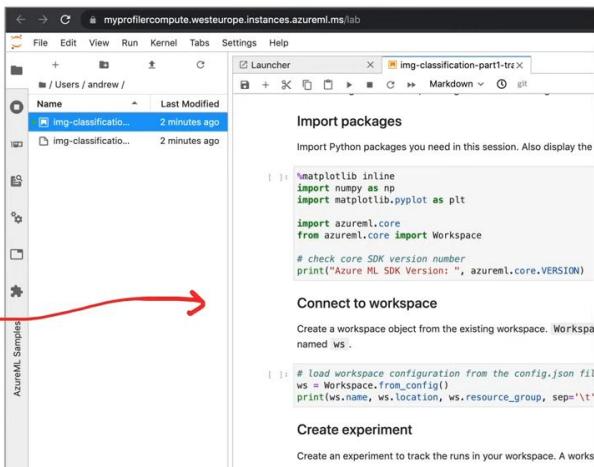
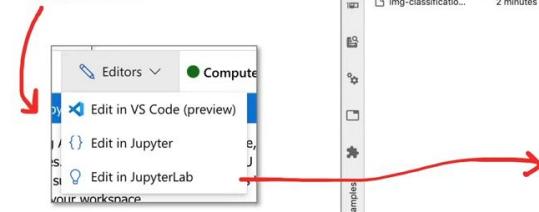
▶ 1:38:10

Azure Machine Learning Studio – Notebooks

Cheat sheets, Practice Exams and Flash cards  www.exampopro.co/ai-900

You can open the Notebook in a more familiar IDE:

- VSCode
- Jupyter Notebook (classic)
- Jupyter Labs**



The screenshot shows the Azure Machine Learning Studio interface for Notebooks. It features a left sidebar with a file tree and a right panel for editing code. A red arrow points from the 'Edit in Jupyter' and 'Edit in JupyterLab' buttons in the toolbar to the main code editor area.

```
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt

import azureml.core
from azureml.core import Workspace

# check core SDK version number
print("Azure ML SDK Version: ", azureml.core.VERSION)

# load workspace configuration from the config.json file
ws = Workspace.from_config()
print(ws.name, ws.location, ws.resource_group, sep='\t')

Create experiment
```

- The speaker prefers using Jupyter Lab over the VS Code interface for Jupyter notebooks, despite acknowledging the variety of options available in Azure automated machine learning.

▶ 1:38:35

AutoML

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Automated machine learning (AutoML) **automates the process of creating an ML model.**

With Azure AutoML you

- supply a dataset
- **Choose a Task Type** (Classification, Regression or Time Series Forecasting)
- Then AutoML will train and tune your model

Classification

When you need to make a prediction based on several classes:

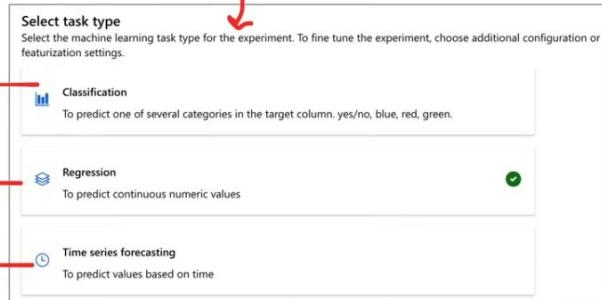
- binary classification: Yes or No
- multi-class classification: Red, Green, Blue

Regression

When you need to predict a continuous number value

Time Series Forecasting

When you need to predict the value based on time



💡 There are various task types in AutoML, including classification for predicting classes, regression for continuous values, and time series forecasting for predicting values based on time.

▶ 1:38:51

AutoML

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Automated machine learning (AutoML) **automates the process of creating an ML model.**

With Azure AutoML you

- supply a dataset
- **Choose a Task Type** (Classification, Regression or Time Series Forecasting)
- Then AutoML will train and tune your model

Classification

When you need to make a prediction based on several classes:

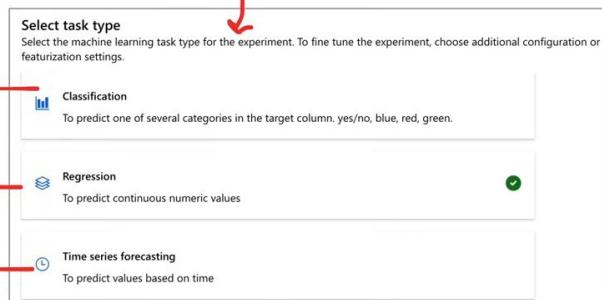
- binary classification: Yes or No
- multi-class classification: Red, Green, Blue

Regression

When you need to predict a continuous number value

Time Series Forecasting

When you need to predict the value based on time



▶ 1:39:11

AutoML – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Classification is a type of **supervised learning** in which **models learn using training data**, and apply those learnings to new data.



The screenshot shows a configuration panel for a machine learning model. At the top left is a blue icon with three horizontal bars. To its right is the word "Classification". Below this is a description: "To predict one of several categories in the target column. yes/no, blue, red, green." On the far right is a green checkmark icon. A red curved arrow points from the bottom of the "Classification" section down towards the "Enable deep learning" checkbox. Below the panel, a note says: "If you enable Deep Learning than you will likely want a GPU compute".

The goal of classification models is to **predict which categories new data will fall into** based on learnings from its training data:

- **binary classification:** a record is labeled out of two possible labels eg: true or false
- **multiclass classification:** a record is labeled out of range of labels: happy, sad, mad, rad

▶ 1:39:28

AutoML – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Classification is a type of **supervised learning** in which **models learn using training data**, and apply those learnings to new data.



The screenshot shows a configuration panel for a machine learning model. At the top left is a blue icon with three horizontal bars. To its right is the word "Classification". Below this is a description: "To predict one of several categories in the target column. yes/no, blue, red, green." On the far right is a green checkmark icon. A red curved arrow points from the bottom of the "Classification" section down towards the "Enable deep learning" checkbox. Below the panel, a note says: "If you enable Deep Learning than you will likely want a GPU compute".

The goal of classification models is to **predict which categories new data will fall into** based on learnings from its training data:

- **binary classification:** a record is labeled out of two possible labels eg: true or false
- **multiclass classification:** a record is labeled out of range of labels: happy, sad, mad, rad

💡 Deep learning requires GPUs for optimal performance, and regression is a type of supervised learning that applies learnings from training data to new data.

▶ 1:40:01

AutoML – Regression

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Regression is a type of **supervised learning** in which **models learn using training data**, and apply those learnings to new data.



Regression

To predict continuous numeric values



The goal of regression is to predict a variable in the future

- 💡 Deep learning primarily utilizes GPUs, and while regression is a type of supervised learning aimed at predicting future variables, time series forecasting functions similarly by predicting outcomes like revenue or demand through multivariate regression analysis.

▷ 1:40:20

AutoML – Time Series Forecasting

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Forecast revenue, inventory, sales, or customer demand

An automated time-series experiment is treated as a **multivariate regression problem**

Past time-series values are "pivoted" to become additional dimensions for the regressor together with other predictor

unlike classical time series methods, has an advantage of naturally incorporating multiple contextual variables and their relationship to one another during training

⌚ Time series forecasting
To predict values based on time

The time series forecasting method requires some additional information.

Time column * 

Time series identifier(s) 

Frequency * 
 Autodetect

Forecast horizon * 
 Autodetect

Enable deep learning 

▷ 1:40:31

AutoML – Time Series Forecasting

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Forecast revenue, inventory, sales, or customer demand

An automated time-series experiment is treated as a **multivariate regression problem**

Past time-series values are "pivoted" to become additional dimensions for the regressor together with other predictor

unlike classical time series methods, has an advantage of naturally incorporating multiple contextual variables and their relationship to one another during training

Time series forecasting
To predict values based on time

The time series forecasting method requires some additional information.

Time column * 

Select a time column...

Time series identifier(s) 

Select column(s)...

Frequency * 

Autodetect

Forecast horizon * 

Autodetect

Enable deep learning 

- Multivariate regression leverages past time series values as additional dimensions alongside other predictors, allowing for the integration of multiple contextual variables and their interrelationships during training, with applications in holiday detection, feature engineering, and advanced forecasting techniques like Auto ARIMA and deep learning models.

▶ 1:40:49

AutoML – Time Series Forecasting

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Advanced forecasting configuration includes:

- holiday detection and featurization
- time-series and DNN learners (Auto-ARIMA, Prophet, ForecastTCN)
- many models support through grouping
- rolling-origin cross validation
- configurable lags
- rolling window aggregate features

- AutoML includes features like Auto ARIMA for profit forecasting, model support through grouping and rolling origin cross-validation, and data guardrails that ensure high-quality input data during automatic featurization.

▶ 1:40:59

AutoML – Time Series Forecasting

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Advanced forecasting configuration includes:

- holiday detection and featurization
- time-series and DNN learners (Auto-ARIMA, Prophet, ForecastTCN)
- many models support through grouping
- rolling-origin cross validation
- configurable lags
- rolling window aggregate features

▶ 1:41:13

AutoML – Data Guard Rails

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Data guardrails are run by Azure AutoML when **automatic featurization** is enabled.

A **sequence of checks** to **ensure high quality input data** is being used to train model.



Type	Status	Description	✓
Validation split handling	Done	The input data has been split for validation to improve model performance. Learn more about validation data.	✓

+ View additional details

Type	Status	Description	✓
Missing feature values imputation	Passed	No feature missing values were detected in the training data. Learn more about missing value imputation.	✓

Type	Status	Description	✓
High cardinality feature detection	Passed	Your inputs were analyzed, and no high cardinality features were detected. Learn more about high cardinality feature detection.	✓

▶ 1:41:27

AutoML – Data Guard Rails

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Data guardrails are run by Azure AutoML when **automatic featurization** is enabled.

A **sequence of checks** to **ensure high quality input data** is being used to train model.



Type	Status	Description	✓
Validation split handling	Done	The input data has been split for validation to improve model performance. Learn more about validation data.	✓
+ View additional details			
Type	Status	Description	✓
Missing feature values imputation	Passed	No feature missing values were detected in the training data. Learn more about missing value imputation.	✓
+ View additional details			
Type	Status	Description	✓
High cardinality feature detection	Passed	Your inputs were analyzed, and no high cardinality features were detected. Learn more about high cardinality feature detection.	✓
+ View additional details			

- Automatic featurization includes checks for high-quality input data, such as validation split handling, missing value imputation, and detection of high cardinality features.

▶ 1:41:37

AutoML – Automatic Featurization

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

During model training with AutoML, one of the following **scaling or normalization techniques** will be applied to each model



StandardScaleWrapper — Standardize features by removing the mean and scaling to unit variance

MinMaxScalar — Transforms features by scaling each feature by that column's minimum and maximum

MaxAbsScaler — Scale each feature by its maximum absolute value

RobustScalar — Scales features by their quantile range

Principal component analysis (PCA) — Linear dimensionality reduction using Singular Value Decomposition of the data to project it to a lower dimensional space

TruncatedSVDWrapper — This transformer performs linear dimensionality reduction by means of truncated singular value decomposition (SVD). Contrary to PCA, this estimator does not center the data before computing the singular value decomposition, which means it can work with `scipy.sparse` matrices efficiently

SparseNormalizer — Each sample (that is, each row of the data matrix) with at least one non-zero component is rescaled independently of other samples so that its norm ($\|1$ or $\|2$) equals one

- Various scaling techniques, including min-max scaling, max absolute value scaling, robust scaling based on quantile range, and PCA for dimensionality reduction using singular value decomposition, are employed to manage complex data with numerous features or labels.

▶ 1:42:22

AutoML – Automatic Featurization

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

During model training with AutoML, one of the following **scaling or normalization techniques** will be applied to each model



StandardScaleWrapper — Standardize features by removing the mean and scaling to unit variance

MinMaxScalar — Transforms features by scaling each feature by that column's minimum and maximum

MaxAbsScaler — Scale each feature by its maximum absolute value

RobustScalar — Scales features by their quantile range

Principal component analysis (PCA) — Linear dimensionality reduction using Singular Value Decomposition of the data to project it to a lower dimensional space

TruncatedSVDWrapper — This transformer performs linear dimensionality reduction by means of truncated singular value decomposition (SVD). Contrary to PCA, this estimator does not center the data before computing the singular value decomposition, which means it can work with `scipy.sparse` matrices efficiently

SparseNormalizer — Each sample (that is, each row of the data matrix) with at least one non-zero component is rescaled independently of other samples so that its norm (ℓ_1 or ℓ_2) equals one

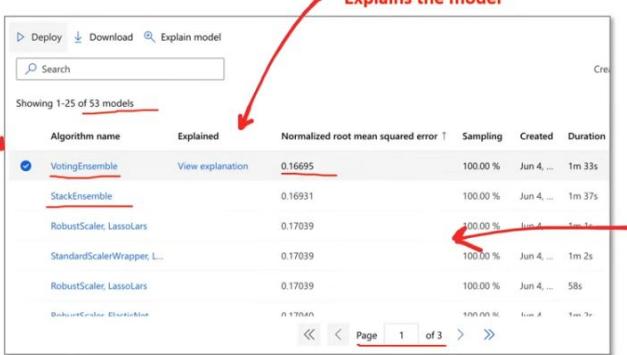
- Truncated singular value decomposition can efficiently handle sparse matrices without centering the data, allowing for independent rescaling of each sample based on its norm.

▶ 1:43:19

AutoML – Model Selection

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Model selection is the **task of selecting a statistical model** from a set of **candidate models**.
Azure AutoML will use **many different ML Algorithms** and will recommend the best **performing candidate**



The top candidate

Explains the model

The results of running for different ML algorithms

Algorithm name	Explained	Normalized root mean squared error ↑	Sampling	Created	Duration
VotingEnsemble	View explanation	0.16695	100.00 %	Jun 4, ...	1m 33s
StackEnsemble		0.16931	100.00 %	Jun 4, ...	1m 37s
RobustScaler, LassoLars		0.17039	100.00 %	Jun 4, ...	1m 1s
StandardScalerWrapper, L...		0.17039	100.00 %	Jun 4, ...	1m 2s
RobustScaler, LassoLars		0.17039	100.00 %	Jun 4, ...	58s
DecisionTree Classifier		0.17039	100.00 %	Jun 4, ...	1m 7s

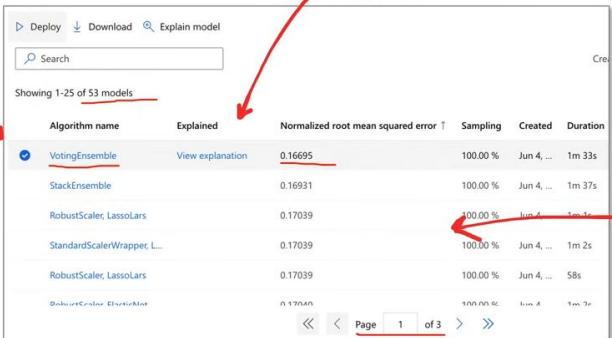
- The top-performing candidate among 53 models is the voting ensemble algorithm, which combines two weaker machine learning models to create a stronger one, achieving the highest primary metric value.

▶ 1:44:08

AutoML – Model Selection

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Model selection is the **task of selecting a statistical model** from a set of **candidate models**
Azure AutoML will use **many different ML Algorithms** and will recommend the best **performing candidate**



Algorithm name	Explained	Normalized root mean squared error ↑	Sampling	Created	Duration
VotingEnsemble	View explanation 0.16695	100.00 %	Jun 4, ...	1m 33s	
StackEnsemble	0.16931	100.00 %	Jun 4, ...	1m 37s	
RobustScaler, LassoLars	0.17039	100.00 %	Jun 4, ...	1m 4s	
StandardScalerWrapper, L...	0.17039	100.00 %	Jun 4, ...	1m 2s	
RobustScaler, LassoLars	0.17039	100.00 %	Jun 4, ...	58s	
OrdinalEncoder, ElastiCNet	0.17040	100.00 %	Jun 4, ...	2m 7s	

- The highest value of the primary metric indicates the preferred model to use, while explainability allows for understanding model choices, though less experienced users may simply select the top candidate.

▶ 1:44:49

AutoML – Explanation

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

ML Explainability (MLX) is the process of **explaining and interpreting** ML and deep learning models.
MLX can help machine learning developers to better understand and interpret the model's behavior

After your top candidate model is selected by Azure AutoML
you can get an explanation of the internals on various factors:

- Model Performance
- Dataset explorer

- Machine learning explainability involves interpreting models to understand their behavior, including factors affecting performance and feature importance.

▶ 1:45:06

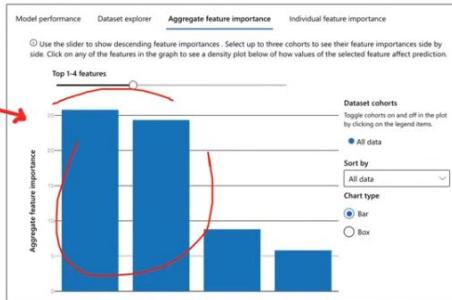
AutoML – Explanation

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

ML Explainability (MLX) is the process of **explaining and interpreting** ML and deep learning models.
MLX can help machine learning developers to better understand and interpret the model's behavior

After your top candidate model is selected by Azure AutoML
you can get an explanation of the internals on various factors:

- Model Performance
- Dataset explorer
- **Aggregate feature importance** 
- Individual feature importance



💡 The primary metric is a key parameter for optimizing model training,
with specific metrics available for classification, regression, and time
series tasks.

▷ 1:45:36

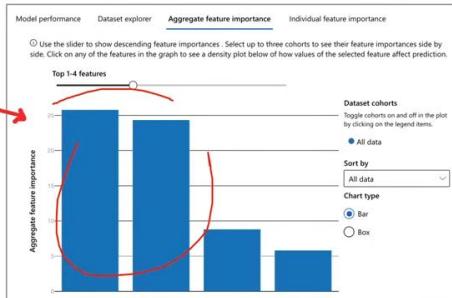
AutoML – Explanation

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

ML Explainability (MLX) is the process of **explaining and interpreting** ML and deep learning models.
MLX can help machine learning developers to better understand and interpret the model's behavior

After your top candidate model is selected by Azure AutoML
you can get an explanation of the internals on various factors:

- Model Performance
- Dataset explorer
- **Aggregate feature importance** 
- Individual feature importance



▷ 1:45:43

AutoML – Primary Metrics

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

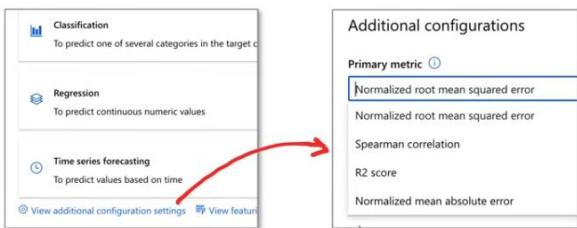
The primary metric parameter determines the metric to be used during model training for optimization.

Classification

- accuracy
- AUC_weighted
- average_precision_score_weighted
- norm_macro_recall
- precision_score_weighted

Regression and Time Series Forecasting

- spearman_correlation
- normalized_root_mean_squared_error
- r2_score
- normalized_mean_absolute_error



The screenshot shows the AutoML interface. On the left, there are three main categories: Classification, Regression, and Time series forecasting. The 'Time series forecasting' category is highlighted with a red arrow pointing towards the 'Additional configurations' panel on the right. The 'Additional configurations' panel lists several primary metrics: Normalized root mean squared error, Spearman correlation, R2 score, and Normalized mean absolute error. The 'Normalized root mean squared error' option is selected.

▶ 1:46:01

AutoML – Primary Metrics

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

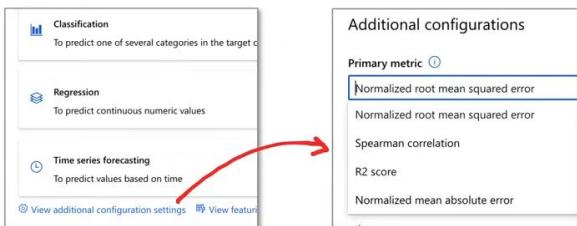
The primary metric parameter determines the metric to be used during model training for optimization.

Classification

- accuracy
- AUC_weighted
- average_precision_score_weighted
- norm_macro_recall
- precision_score_weighted

Regression and Time Series Forecasting

- spearman_correlation
- normalized_root_mean_squared_error
- r2_score
- normalized_mean_absolute_error



The screenshot shows the AutoML interface. On the left, there are three main categories: Classification, Regression, and Time series forecasting. The 'Time series forecasting' category is highlighted with a red arrow pointing towards the 'Additional configurations' panel on the right. The 'Additional configurations' panel lists several primary metrics: Normalized root mean squared error, Spearman correlation, R2 score, and Normalized mean absolute error. The 'Normalized root mean squared error' option is selected.

💡 The primary metric in model training can be auto-detected or manually overridden based on the task type, with specific configurations available for classification, regression, and time series.

▶ 1:46:07

AutoML – Primary Metrics – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

- 💡 Larger, well-balanced data sets are ideal for various classification tasks, with specific metrics like accuracy and average precision score being suitable for different applications.

▶ 1:46:21

AutoML – Primary Metrics – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Classification Scenarios

- Suited for larger datasets that are well-balanced
 - **accuracy** — Image classification, Sentiment analysis, Churn prediction
 - **average_precision_score_weighted** — Sentiment analysis
 - **norm_macro_recall** — Churn prediction
 - **precision_score_weighted**
- Suited for small datasets that are imbalanced
 - **AUC_weighted** — Fraud detection, Image classification, Anomaly detection/spam detection

- 💡 Precision score weighted is useful for evaluating models on imbalanced datasets, while various metrics like Spearman correlation and R2 score are applicable in different regression scenarios.

▶ 1:47:04

AutoML – Primary Metrics – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Regressions Scenarios

- Works well when value to predict encompasses a large range eg. 10K to 200K
 - **spearman_correlation**
 - **r2_score** — Airline delay, Salary estimation, Bug resolution time
- Works well when value to predict encompasses a smaller range eg. 10-20K
 - **normalized_root_mean_squared_error** — Price prediction (house/product/tip), Review score prediction
 - **normalized_mean_absolute_error**

💡 The discussion focuses on evaluating prediction accuracy using metrics like normalized root mean square error and normalized mean absolute error, particularly in relation to forecasting.

▷ 1:47:15

AutoML – Primary Metrics – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Regressions Scenarios

- Works well when value to predict encompasses a large range eg. 10K to 200K
 - **spearman_correlation**
 - **r2_score** — Airline delay, Salary estimation, Bug resolution time
- Works well when value to predict encompasses a smaller range eg. 10-20K
 - **normalized_root_mean_squared_error** — Price prediction (house/product/tip), Review score prediction
 - **normalized_mean_absolute_error**

▷ 1:47:20

AutoML – Primary Metrics – Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Regressions Scenarios

- Works well when value to predict encompasses a large range eg. 10K to 200K
 - **spearman_correlation**
 - **r2_score** — Airline delay, Salary estimation, Bug resolution time
- Works well when value to predict encompasses a smaller range eg. 10-20K
 - **normalized_root_mean_squared_error** — Price prediction (house/product/tip), Review score prediction
 - **normalized_mean_absolute_error**

▶ 1:47:30

AutoML – Primary Metrics – Time Series

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Time Series Scenarios

- Works well when value to predict encompasses a large range eg. 10K to 200K
 - **spearman_correlation**
 - **r2_score** — Price prediction (forecasting), Inventory optimization, Demand forecasting
- Works well when value to predict encompasses a smaller range eg. 10-20K
 - **normalized_root_mean_squared_error** — Price prediction (forecasting), Inventory optimization, Demand forecasting
 - **normalized_mean_absolute_error**

▶ 1:47:40

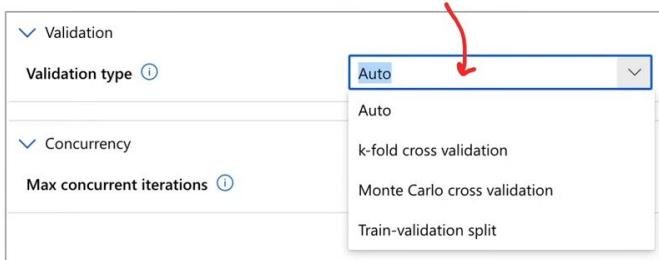
AutoML – Validation Type

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Model Validation is when we **compare the results of our training dataset to our test dataset**.

Model Validation occurs *after* we train the model

With AutoML you can change the validation type



- There are various options for model evaluation, including Auto k-fold cross-validation, Monte Carlo cross-validation, and train-validation split.

▷ 1:48:11

Custom Vision

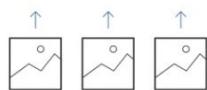
Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Custom Vision is a **fully-managed no-code** service to quickly build your own **Classification and Object Detection ML models**.

This service is hosted on its own isolate domain at www.customvision.ai

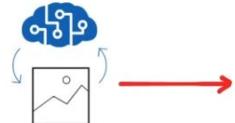
Upload Images

Bring your own labeled images, or use Custom Vision to quickly add tags to any unlabeled images.



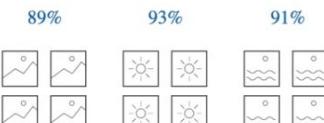
Train

Use your labeled images to teach Custom Vision the concepts you care about.



Evaluate

Use simple REST API calls to quickly tag images with your new custom computer vision model.



- To utilize custom Vision, users must upload labeled images to train a model and can then use a REST API to tag unlabeled images based on the concepts learned.

▷ 1:48:35



Custom Vision

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Custom Vision is a **fully-managed no-code** service to quickly build your own **Classification and Object Detection ML models**.

This service is hosted on its own isolate domain at www.customvision.ai

Upload Images

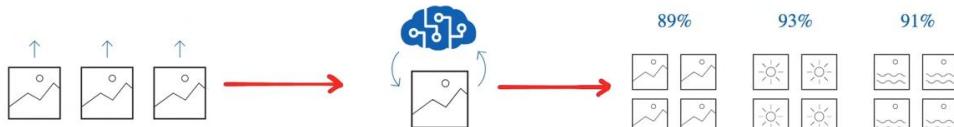
Bring your own labeled images, or use Custom Vision to quickly add tags to any unlabeled images.

Train

Use your labeled images to teach Custom Vision the concepts you care about.

Evaluate

Use simple REST API calls to quickly tag images with your new custom computer vision model.



▶ 1:48:50



Custom Vision – Project Types

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Within Custom Vision you setup projects and you need to select a **Project Type**

Project Types ⓘ

- Classification
- Object Detection

Classification Types ⓘ

- Multilabel (Multiple tags per image)
- Multiclass (Single tag per image)

Classification

• Multi-label

- When we want to apply many tags to an image
 - Image contains both a Cat and a Dog

• Multi-class

- when we only have one possible tag to apply to an image:
 - It is either a Apple, Banana, Orange

Object Detection

- When we want to detect various objects in an image

You will need to also choose a **Domain**

A Domain is a Microsoft Managed dataset that is used for training the ML model

There are different domains that suited for different use cases

▶ 1:49:16



Custom Vision – Project Types

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Within Custom Vision you setup projects and you need to select a **Project Type**

Project Types ⓘ

- Classification
- Object Detection

Classification Types ⓘ

- Multilabel (Multiple tags per image)
- Multiclass (Single tag per image)

Classification

• Multi-label

- When we want to apply many tags to an image
 - Image contains both a Cat and a Dog

• Multi-class

- when we only have one possible tag to apply to an image:
 - It is either a Apple, Banana, Orange

Object Detection

- When we to detect various objects in an image

You will need to also choose a **Domain**

A Domain is a Microsoft Managed dataset that is used for training the ML model

There are different domains that suited for different use cases

▶ 1:49:39



Custom Vision – Image Classification Domains

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Classification Domains

Project Types ⓘ

- Classification
- Object Detection

Classification Types ⓘ

- Multilabel (Multiple tags per image)
- Multiclass (Single tag per image)

Domains:

- General [A2]
- General [A1]
- General
- Food
- Landmarks
- Retail
- General (compact) [S1]
- General (compact)
- Food (compact)
- Landmarks (compact)
- Retail (compact)

General Optimized for a broad range of image classification tasks. If none of the other specific domains are appropriate, or if you're unsure of which domain to choose, select one of the General domains.

General [A1] Optimized for better accuracy with comparable inference time as General domain. Recommended for larger datasets or more difficult user scenarios. This domain requires more training time.

General [A2] Optimized for better accuracy with faster inference time than General[A1] and General domains. Recommended for most datasets. This domain requires less training time than General and General [A1] domains.

Food Optimized for photographs of dishes as you would see them on a restaurant menu. If you want to classify photographs of individual fruits or vegetables, use the Food domain.

Landmarks Optimized for recognizable landmarks, both natural and artificial. This domain works best when the landmark is clearly visible in the photograph. This domain works even if the landmark is slightly obstructed by people in front of it.

Retail Optimized for images that are found in a shopping catalog or shopping website. If you want high-precision classifying between dresses, pants, and shirts, use this domain.

Compact domains Optimized for the constraints of real-time classification on edge devices.

★ General is suitable for a wide range of image classification tasks, while A1 offers improved accuracy for larger datasets at the cost of longer training time, and A2 provides better accuracy with faster inference times than both A1 and General.

▶ 1:49:48



Custom Vision – Image Classification Domains

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Classification Domains

Project Types
 Classification
 Object Detection

Classification Types
 Multilabel (Multiple tags per image)
 Multiclass (Single tag per image)

Domains:
 General [A2]
 General [A1]
 General
 Food
 Landmarks
 Retail
 General (compact) [S1]
 General (compact)
 Food (compact)
 Landmarks (compact)
 Retail (compact)

A red curly brace on the right side groups the 'General' classification type and the 'General' domain.

General Optimized for a broad range of image classification tasks. If none of the other specific domains are appropriate, or if you're unsure of which domain to choose, select one of the General domains.

General [A1] Optimized for better accuracy with comparable inference time as General domain. Recommended for larger datasets or more difficult user scenarios. This domain requires more training time.

General [A2] Optimized for better accuracy with faster inference time than General[A1] and General domains. Recommended for most datasets. This domain requires less training time than General and General [A1] domains.

Food Optimized for photographs of dishes as you would see them on a restaurant menu. If you want to classify photographs of individual fruits or vegetables, use the Food domain.

Landmarks Optimized for recognizable landmarks, both natural and artificial. This domain works best when the landmark is clearly visible in the photograph. This domain works even if the landmark is slightly obstructed by people in front of it.

Retail Optimized for images that are found in a shopping catalog or shopping website. If you want high-precision classifying between dresses, pants, and shirts, use this domain.

Compact domains Optimized for the constraints of real-time classification on edge devices.

- Optimize image classification for retail items like clothing and use compact domains for real-time edge classification, while employing a broad object detection domain when other specific options are unsuitable.

▶ 1:51:06



Custom Vision – Object Detection Domains

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Object Detection Domains

Project Types
 Classification
 Object Detection

Domains:
 General [A1]
 General
 Logo
 Products on Shelves
 General (compact) [S1]
 General (compact)

General

Optimized for a broad range of object detection tasks. If none of the other domains are appropriate, or you are unsure of which domain to choose, select the General domain.

General [A1]

Optimized for better accuracy with comparable inference time as General domain. Recommended for more accurate region location needs, larger datasets, or more difficult user scenarios. This domain requires more training time, and results are not deterministic: expect a +/-1% mean Average Precision (mAP) difference with the same training data provided.

Logo

Optimized for finding brand logos in images.

Products on shelves

Optimized for detecting and classifying products on shelves.

- Choosing a domain optimized for specific tasks can improve accuracy and inference time, but more complex scenarios may require additional training with results that are not deterministic.

▶ 1:51:18

Custom Vision – Object Detection Domains

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Object Detection Domains

Project Types (i)

Classification

Object Detection

Domains:

General [A1]

General

Logo

Products on Shelves

General (compact) [S1]

General (compact)

General

Optimized for a broad range of object detection tasks. If none of the other domains are appropriate, or you are unsure of which domain to choose, select the General domain.

General [A1]

Optimized for better accuracy with comparable inference time as General domain.

Recommended for more accurate region location needs, larger datasets, or more difficult user scenarios. This domain requires more training time, and results are not deterministic: expect a +/- 1% mean Average Precision (mAP) difference with the same training data provided.

Logo

Optimized for finding brand logos in images.

Products on shelves

Optimized for detecting and classifying products on shelves.

- 💡 Deterministic models can achieve a mean average precision difference of plus or minus 1% when optimized for detecting and classifying brands, logos, and products on shelves using the same training data.

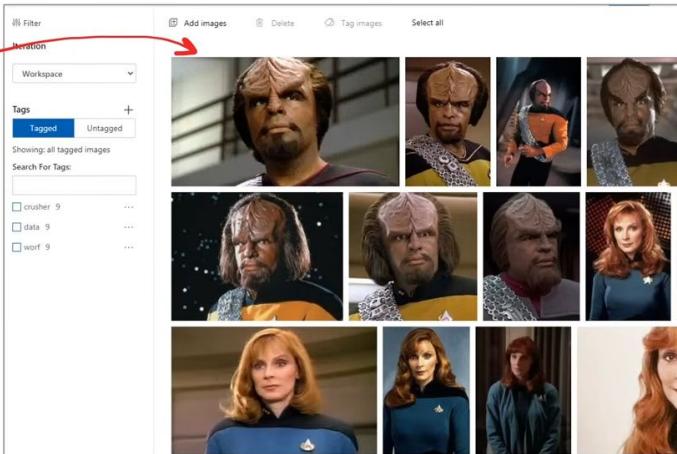
▷ 1:51:49



Custom Vision – Image Classification

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

For Image Classification you upload multiple images and you apply a single or multiple labels to the entire image.



- 💡 Image classification involves uploading multiple images and applying single or multiple labels, while object detection requires tagging specific objects within an image using bounding boxes generated by machine learning.

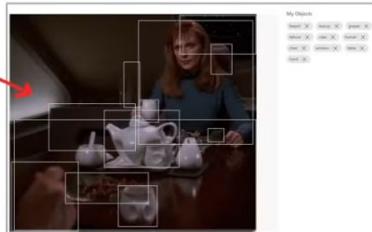
▷ 1:52:07



Custom Vision – Object Detection

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

For Object Detection you apply tags to objects in an image for data labeling



When you **hover your cursor over the image**

Custom Vision uses ML to show bounding boxes of possible objects but yet have no label applied

- To train a model, at least 50 images must be tagged for each object, and there are two training options: Quick Training for faster but less accurate results, and Advanced Training for improved accuracy at the cost of increased compute time.

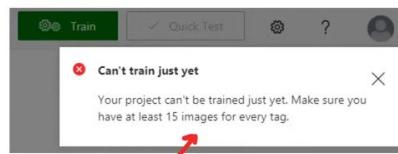
▷ 1:52:19



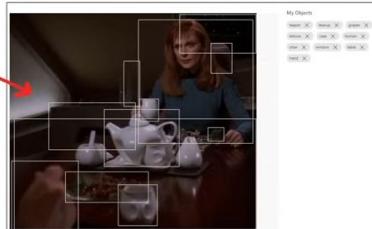
Custom Vision – Object Detection

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

For Object Detection you apply tags to objects in an image for data labeling



You have to **have at least 15 images** for every tag to train



When you **hover your cursor over the image**

Custom Vision uses ML to show bounding boxes of possible objects but yet have no label applied

- To train a model, at least 50 images must be tagged for each object, and there are two training options: Quick Training for faster but less accurate results, and Advanced Training for improved accuracy at the cost of increased compute time.

▷ 1:52:34

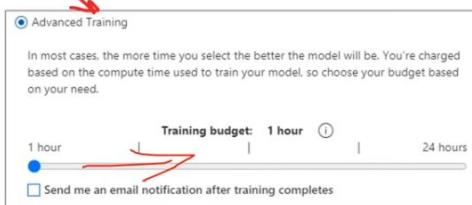


Custom Vision – Training

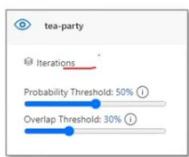
Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

When you are ready to train your model you have two options:

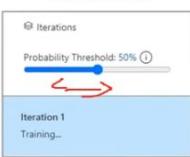
- **Quick Training** – trains quickly but can be less accurate
- **Advanced Training** – increase the compute time to improve your results



Object Detection



Classification



With each iteration of training our ML model will improve the evaluation metrics (**precision** and **recall**) The **probability threshold value** determines when to stop training when our evaluation metrics meet our desired threshold

- Advanced training increases compute time to enhance model performance, with adjustments made to a probability threshold value that determines when training stops based on desired evaluation metrics like precision and recall.

▷ 1:52:43



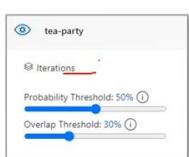
Custom Vision – Training

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

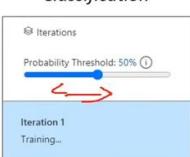
When you are ready to train your model you have two options:

- **Quick Training** – trains quickly but can be less accurate
- **Advanced Training** – increase the compute time to improve your results

Object Detection



Classification



With each iteration of training our ML model will improve the evaluation metrics (**precision** and **recall**) The **probability threshold value** determines when to stop training when our evaluation metrics meet our desired threshold

- The probability threshold value influences when to stop training based on evaluation metrics like precision and recall, which measure the accuracy of relevant item selection and sensitivity, respectively.

▷ 1:53:11



Custom Vision – Training

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

Once the **Classification** training job is complete we will get a report of the evaluation metrics **outcome**

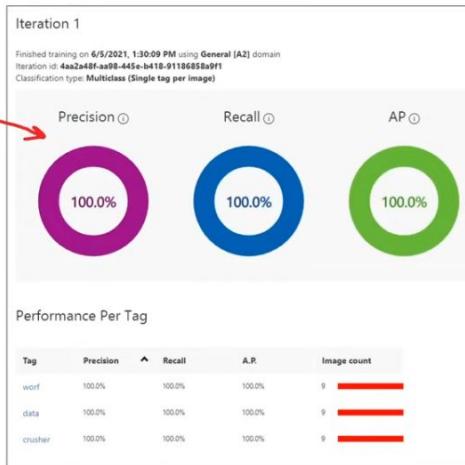
Precision

- being exact and accurate
- select items that are relevant

Recall (Sensitivity or True Positive Rate)

- How many relevant items returned

Average Precision (AP)



▶ 1:53:21



Custom Vision – Training

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

Once the **Classification** training job is complete we will get a report of the evaluation metrics **outcome**

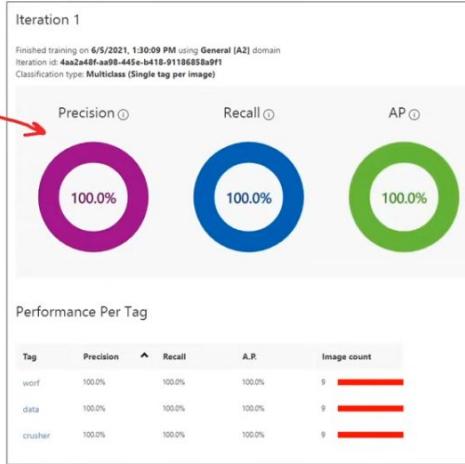
Precision

- being exact and accurate
- select items that are relevant

Recall (Sensitivity or True Positive Rate)

- How many relevant items returned

Average Precision (AP)



💡 Precision and recall are important evaluation metrics in object detection, with precision indicating the accuracy of relevant item selection and recall measuring the true positive rate.

▶ 1:53:22



Custom Vision – Training

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Once the **Classification** training job is complete we will get a report of the evaluation metrics **outcome**

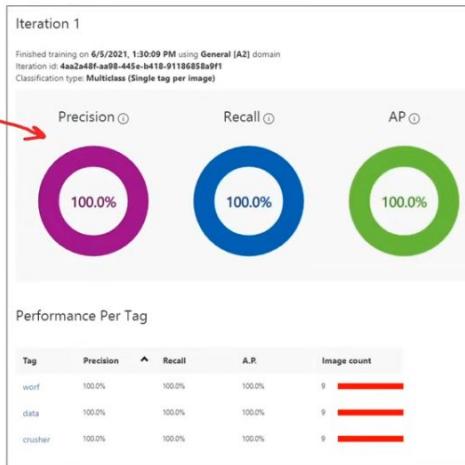
Precision

- being exact and accurate
- select items that are relevant

Recall (Sensitivity or True Positive Rate)

- How many relevant items returned

Average Precision (AP)



▶ 1:53:22



Custom Vision – Training

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Once the **Classification** training job is complete we will get a report of the evaluation metrics **outcome**

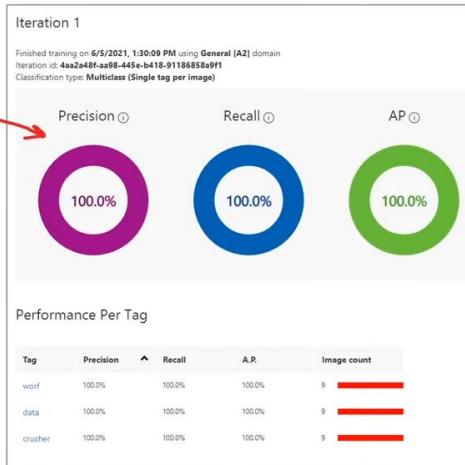
Precision

- being exact and accurate
- select items that are relevant

Recall (Sensitivity or True Positive Rate)

- How many relevant items returned

Average Precision (AP)



▶ 1:53:31

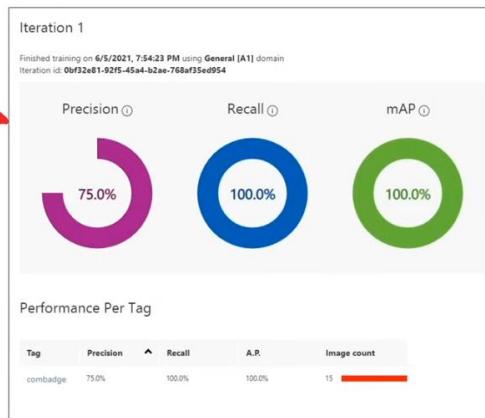


Custom Vision – Training

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Once the **Object Detection** training job is complete we will get a report of the evaluation metrics **outcome**

- Precision
- Recall
- Mean Average Precision (mAP)



▶ 1:53:44



Custom Vision – Quick Test

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Before deploying our final trained model that can be invoked via an API Endpoint
It is good practice to test our ML model using the **Quick Test** feature

Quick Test

Image URL: Enter Image URL →
or

File formats accepted: jpg, png, bmp
File size should not exceed: 4mb

Using model trained in:

Iteration: Iteration 1

Predictions:

Tag	Probability
worf	98.7%
data	1%
crusher	0.1%

★ After deploying a pipeline, users can test its functionality by uploading an image for prediction and then publish it to receive a prediction URL, while also utilizing a Smart labeler feature for training data.

▶ 1:53:56



Custom Vision – Publish

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

To deploy our ML model to be accessible using our API Key and Endpoint we need to trigger the **Publish** action.

Publish Model

We only support publishing to a prediction resource in the same region as the training resource the project resides in.

Please check if you have a prediction resource and if the prediction resource is in the same region as the training resource.

Model name: star-trek-crew

Prediction resource: myCognitiveServices734

Publish **Cancel**

How to use the Prediction API

If you have an image URL:

```
https://mycognitiveservices734.cognitiveservices.azure.com/customvision/v3.0/Pred  
Set Prediction-Key Header to : 8094c25bb05f4fd68092a5d9e59744a2  
Set Content-Type Header to : application/json  
Set Body to : {"Url": "https://example.com/image.png"};
```

If you have an image file:

```
https://mycognitiveservices734.cognitiveservices.azure.com/customvision/v3.0/Pred  
Set Prediction-Key Header to : 8094c25bb05f4fd68092a5d9e59744a2  
Set Content-Type Header to : application/octet-stream  
Set Body to : <image file>
```

Got it!

Once published we can get the **Prediction URL**

▶ 1:54:04

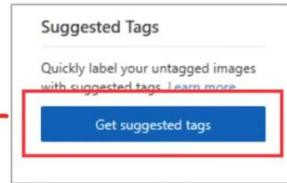
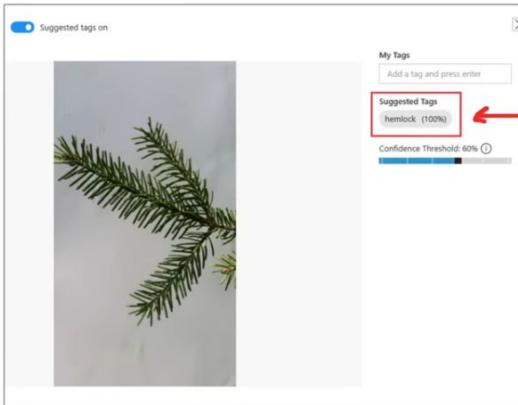


Custom Vision – Smart Labeler

Cheat sheets, Practice Exams and Flash cards www.exapro.co/ai-900

When you have a very large dataset you can use Smart Labeler to predict labels.

Smart Labeler only works if you have trained the label



Smart labeler is when you want to increase your training set, and want to ML-assisted labeling to speed up this process.

▶ 1:54:18



Custom Vision – Smart Labeler

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

When you have a very large dataset you can use Smart Labeler to predict labels.

Smart Labeler only works if you have trained the label

Suggested Tags

hemlock (100%)

Confidence Threshold: 60%

Get suggested tags

Suggested Tags

Quickly label your untagged images with suggested tags. [Learn more](#)

Smart labeler is when you want to increase your training set, and want to ML-assisted labeling to speed up this process.

▶ 1:54:20



Custom Vision – Smart Labeler

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

When you have a very large dataset you can use Smart Labeler to predict labels.

Smart Labeler only works if you have trained the label

Suggested Tags

hemlock (100%)

Confidence Threshold: 60%

Get suggested tags

Suggested Tags

Quickly label your untagged images with suggested tags. [Learn more](#)

Smart labeler is when you want to increase your training set, and want to ML-assisted labeling to speed up this process.

- Machine learning-assisted labeling can enhance the efficiency of building training datasets by making predictions based on initial data, particularly beneficial for large datasets.

▶ 1:54:27



AI vs Generative AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Artificial Intelligence (AI)

- Generative AI, including technologies like ChatGPT, is gaining recognition beyond tech circles and relies on advanced mathematical techniques from statistics, data science, and machine learning.

▶ 1:54:42



AI vs Generative AI

Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Artificial Intelligence (AI)



AI refers to the development of computer systems that can **perform tasks typically requiring human intelligence**. These include **problem-solving, decision-making, understanding natural language, recognizing speech and images**, and more.



The primary goal of traditional AI is to create systems that can **interpret, analyze, and respond to human actions** or environmental changes efficiently and accurately. It aims to replicate or simulate human intelligence in machines.



AI applications are vast and include areas like **expert systems, natural language processing, speech recognition, and robotics**.



AI is used in various industries for tasks such as **customer service chatbots, recommendation systems in e-commerce, autonomous vehicles, and medical diagnosis**.

▶ 1:55:26



AI vs Generative AI

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Artificial Intelligence (AI)



AI refers to the development of computer systems that can **perform tasks typically requiring human intelligence**. These include **problem-solving, decision-making, understanding natural language, recognizing speech and images**, and more.



The primary goal of traditional AI is to create systems that can **interpret, analyze, and respond to human actions** or environmental changes efficiently and accurately. It aims to replicate or simulate human intelligence in machines.



AI applications are vast and include areas like **expert systems, natural language processing, speech recognition, and robotics**.



AI is used in various industries for tasks such as **customer service chatbots, recommendation systems in e-commerce, autonomous vehicles, and medical diagnosis**.

- AI encompasses a wide range of applications across various industries, while generative AI specifically focuses on creating new and realistic content or data.

▶ 1:55:45



AI vs Generative AI

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Generative AI



Generative AI is a subset of AI that focuses on **creating new content or data** that is novel and realistic. It does not just interpret or analyze data but **generates new data itself**. It includes **generating text, images, music, speech, and other forms of media**.



It often involves advanced machine learning techniques, particularly deep learning models like **Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Transformer models (like GPT)**.



Generative AI is used in a range of applications including creating realistic **images and videos**, generating **human-like text**, composing **music**, creating virtual environments, and even drug discovery.



Examples: Tools like **GPT (Generative Pre-trained Transformer)** for text generation, **DALL-E** for image creation, and various deep learning models that compose music.

- Generative AI utilizes advanced machine learning techniques to autonomously create various forms of media, including text, images, music, and virtual environments.

▶ 1:55:56



AI vs Generative AI

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Generative AI



Generative AI is a subset of AI that focuses on **creating new content or data** that is novel and realistic. It does not just interpret or analyze data but **generates new data itself**. It includes **generating text, images, music, speech, and other forms of media**.



It often involves advanced machine learning techniques, particularly deep learning models like **Generative Adversarial Networks (GANs)**, **Variational Autoencoders (VAEs)**, and **Transformer models (like GPT)**.



Generative AI is used in a range of applications including creating realistic **images** and **videos**, generating **human-like text**, composing **music**, creating virtual environments, and even drug discovery.



Examples: Tools like **GPT (Generative Pre-trained Transformer)** for text generation, **DALL-E** for image creation, and various deep learning models that compose music.

- Generative AI differs from regular AI in its ability to create content, handle diverse data types, and apply to a wide range of applications such as image and text generation, music composition, virtual environment creation, and drug discovery.

▷ 1:56:31



AI vs Generative AI

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Feature	Artificial Intelligence (AI)	Generative AI
Functionality	Regular AI focuses on understanding and decision-making	Generative AI is about creating new, original outputs.
Data Handling	AI typically analyzes and makes decisions based on existing data	Generative AI uses existing data to generate new, unseen outputs.
Applications	Its applications span across various sectors, including data analysis, automation, natural language processing, and healthcare.	Its applications are more creative and innovative, focusing on content creation, synthetic data generation, deepfakes, and design.

- Regular AI focuses on analyzing existing data for decision-making, while generative AI creates new outputs and is used in more creative applications.

▷ 1:56:42



Feature	Artificial Intelligence (AI)	Generative AI
Functionality	Regular AI focuses on understanding and decision-making	Generative AI is about creating new, original outputs.
Data Handling	AI typically analyzes and makes decisions based on existing data	Generative AI uses existing data to generate new, unseen outputs.
Applications	Its applications span across various sectors, including data analysis, automation, natural language processing, and healthcare.	Its applications are more creative and innovative, focusing on content creation, synthetic data generation, deepfakes, and design.

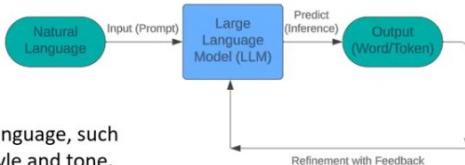
▶ 1:57:06



What is a Large Language Model (LLM)?

A Large Language Model (LLM) such as GPT (Generative Pre-trained Transformer) works in a way that's similar to a complex, **automatic system that recognizes patterns and makes predictions**.

Training on Large Datasets: Initially, the model is trained on massive amounts of text data. This data can include **books, articles, websites, and other written material**.



During this training phase, the model learns patterns in language, such as grammar, word usage, sentence structure, and even style and tone.

Understanding Context: The model's design allows it to consider a wide context. This means it doesn't just focus on single words, but understands them in **relation to the words and sentences** that come **before and after**. This context understanding is important for generating coherent and relevant text.

- 💡 The model is trained on extensive text data to learn language patterns, grammar, and context for generating coherent responses.

▶ 1:57:30

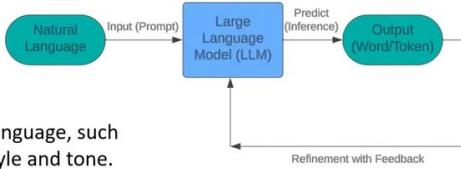


What is a Large Language Model (LLM)?

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

A Large Language Model (LLM) such as GPT (Generative Pre-trained Transformer) works in a way that's similar to a complex, **automatic system that recognizes patterns and makes predictions.**

Training on Large Datasets: Initially, the model is trained on massive amounts of text data. This data can include **books, articles, websites, and other written material.**



During this training phase, the model learns patterns in language, such as grammar, word usage, sentence structure, and even style and tone.

Understanding Context: The model's design allows it to consider a wide context. This means it doesn't just focus on single words, but understands them in **relation to the words and sentences** that come **before and after**. This context understanding is important for generating coherent and relevant text.

▶ 1:57:58

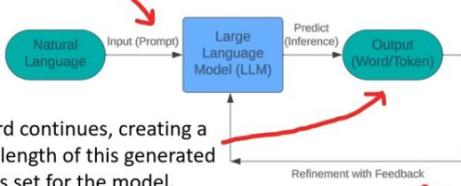


What is a Large Language Model (LLM)?

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Predicting the Next Word: When you give the model a **prompt** (a starting piece of text), it uses what it has learned to predict the next most likely word.

It then adds this word to the prompt and repeats the process, continually **predicting the next word** based on the extended sequence.



Generating Text: This process of predicting the next word continues, creating a **chain of words** that forms a coherent piece of text. The length of this generated text can vary based on specific instructions or limitations set for the model.

Refinement with Feedback: The model can be further **refined** and **improved** over time with **feedback**. This means it gets better at understanding and generating text as it is exposed to more data and usage.

- ★ A model generates coherent text by continually predicting the next word based on an extended sequence, improving over time through feedback and exposure to more data.

▶ 1:58:12

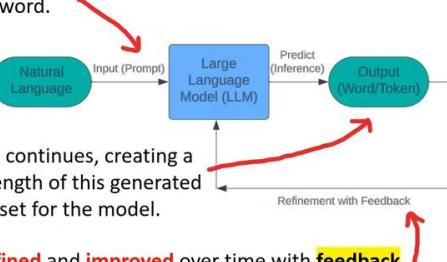


What is a Large Language Model (LLM)?

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Predicting the Next Word: When you give the model a **prompt** (a starting piece of text), it uses what it has learned to predict the next most likely word.

It then adds this word to the prompt and repeats the process, continually **predicting the next word** based on the extended sequence.



Generating Text: This process of predicting the next word continues, creating a **chain of words** that forms a coherent piece of text. The length of this generated text can vary based on specific instructions or limitations set for the model.

Refinement with Feedback: The model can be further **refined** and **improved** over time with **feedback**. This means it gets better at understanding and generating text as it is exposed to more data and usage.

- ★ A large language model improves its ability to understand and generate text by learning from extensive data and user feedback, utilizing a Transformer architecture.

▶ 1:58:48



Transformer models

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

A transformer model is a type of machine learning model that's especially good at **understanding and generating language**.

It's built using a structure called the transformer architecture, which is really effective for tasks involving **natural language processing (NLP)**, like **translating languages or writing text**.

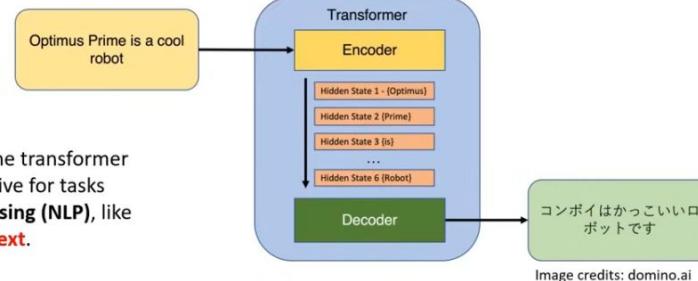


Image credits: domino.ai

Transformer model architecture consists of **two components**, or **blocks**:

- 1. Encoder:** This part **reads and understands the input text**. It's like a smart system that goes through everything it's been taught (which is a lot of text) and picks up on the meanings of words and how they're used in different contexts.
- 2. Decoder:** Based on what the encoder has learned, this part **generates new pieces of text**. It's like a skilled writer that can make up sentences that flow well and make sense.

- ★ The Transformer architecture consists of an encoder that comprehends input text and a decoder that generates new text based on the encoder's understanding.

▶ 1:59:10



Transformer models

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

A transformer model is a type of machine learning model that's especially good at **understanding and generating language**.

It's built using a structure called the transformer architecture, which is really effective for tasks involving **natural language processing (NLP)**, like **translating languages or writing text**.

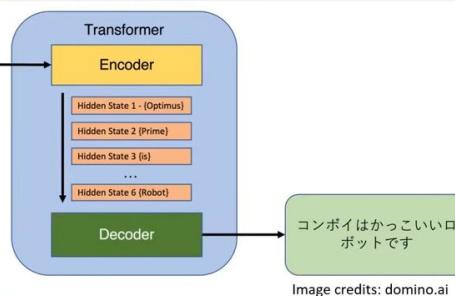


Image credits: domino.ai

Transformer model architecture consists of **two components**, or *blocks*:

- 1. Encoder:** This part **reads and understands the input text**. It's like a smart system that goes through everything it's been taught (which is a lot of text) and picks up on the meanings of words and how they're used in different contexts.
- 2. Decoder:** Based on what the encoder has learned, this part **generates new pieces of text**. It's like a skilled writer that can make up sentences that flow well and make sense.

▶ 1:59:38



Transformer models

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

There are different types of transformer models with specific jobs. For example:



BERT is good at understanding the language. It's like a librarian who knows where every book is and what's inside them. **Google** uses it to help its search engine understand what you're looking for.



GPT is good at creating text. It's like a skilled author who can write stories, articles, or conversations based on what it has learned.

- ★ A transformer model utilizes tokenization to process and understand text, enabling it to generate coherent written content.

▶ 1:59:52



Transformer models

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

There are different types of transformer models with specific jobs. For example:



BERT is good at understanding the language. It's like a librarian who knows where every book is and what's inside them. **Google** uses it to help its search engine understand what you're looking for.



GPT is good at creating text. It's like a skilled author who can write stories, articles, or conversations based on what it has learned.

▶ 2:00:03



Tokenization

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Tokenization in a transformer model is like turning a sentence into a puzzle. For example, you have the sentence: "I heard a dog bark loudly at a cat." To help a computer understand it, we chop up the sentence into pieces called **'tokens'**. Each piece can be a word or even a part of a word. So, for our sentence, we give each word a number, like this:

- "I" might be 1
- "heard" might be 2
- "a" might be 3
- "dog" might be 4
- "bark" might be 5
- "loudly" might be 6
- "at" might be 7
- "a" is already tokenized as 3
- "cat" might be 8



Now, our sentence becomes a series of numbers: [1, 2, 3, 4, 5, 6, 7, 3, 8]. This is like giving each word a **special code**.

The computer uses these codes to **learn about the words and how they fit together**.

If a word repeats, like "a", we use its code again instead of making a new one.

- As the computer reads more text, it keeps turning new words into new tokens with new numbers.
- If it learns the word "meow," it might call it 9, and "skateboard" could be 10.

💡 Tokens are created by breaking down sentences into numbered pieces, allowing a computer to represent words and parts of words as unique codes.

▶ 2:00:27



Tokenization

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Tokenization in a transformer model is like turning a sentence into a puzzle. For example, you have the sentence: "I heard a dog bark loudly at a cat." To help a computer understand it, we chop up the sentence into pieces called **'tokens'**. Each piece can be a word or even a part of a word.

So, for our sentence, we give each word a number, like this:

- "I" might be 1
- "heard" might be 2
- "a" might be 3
- "dog" might be 4
- "bark" might be 5
- "loudly" might be 6
- "at" might be 7
- "a" is already tokenized as 3
- "cat" might be 8



Now, our sentence becomes a series of numbers: [1, 2, 3, 4, 5, 6, 7, 3, 8]. This is like giving each word a **special code**.

The computer uses these codes to **learn about the words and how they fit together**.

If a word repeats, like "a", we use its code again instead of making a new one.

- As the computer reads more text, it keeps turning new words into new tokens with new numbers.
- If it learns the word "meow," it might call it 9, and "skateboard" could be 10.

- 💡 A computer generates a unique list of tokens for words based on their frequency and usage in text, similar to creating a dictionary.

▷ 2:01:11



Embeddings

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

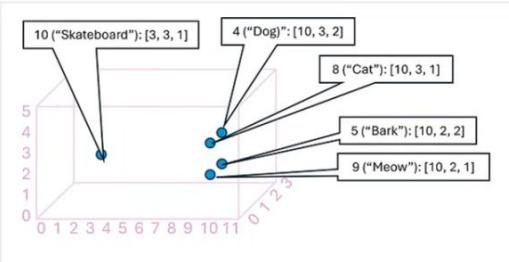
To help a computer understand language, we turn words into tokens and then give each token a special **numeric code**, called an **embedding**. These embeddings are like a secret code that captures the meaning of the word.

As a simple example, suppose the embeddings for our tokens consist of **vectors** with three elements, for example:

- 4 ("dog"): [10,3,2]
- 5 ("bark"): [10,2,2]
- 8 ("cat"): [10,3,1]
- 9 ("meow"): [10,2,1]
- 10 ("skateboard"): [3,3,1]

Words that have **similar meanings** or are used in similar ways get **codes that look alike**.

So, "dog" and "bark" might have similar codes because they are **related**.



Embedding: Turning words (tokens) into vectors (lists of numbers)

This way, the computer can figure out which words are **similar to each other** just by looking at their codes.

It's like giving each word a home on a map, and words that are neighbors on this map have related meanings.

- 💡 Tokens are assigned numeric codes called embeddings, which are vector representations that capture the meanings of words.

▷ 2:01:34



Embeddings

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

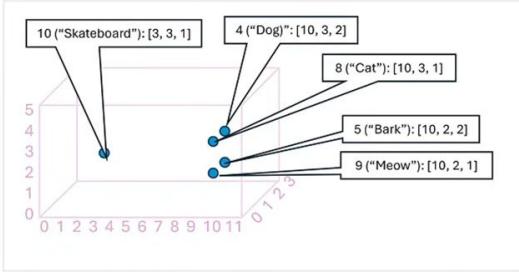
To help a computer understand language, we turn words into tokens and then give each token a special **numeric code**, called an **embedding**. These embeddings are like a secret code that captures the meaning of the word.

As a simple example, suppose the embeddings for our tokens consist of **vectors** with three elements, for example:

- 4 ("dog"): [10,3,2]
- 5 ("bark"): [10,2,2]
- 8 ("cat"): [10,3,1]
- 9 ("meow"): [10,2,1]
- 10 ("skateboard"): [3,3,1]

Words that have **similar meanings** or are used in similar ways get **codes that look alike**.

So, "dog" and "bark" might have similar codes because they are **related**.



Embedding: Turning words (tokens) into vectors (lists of numbers)

This way, the computer can figure out which words are **similar to each other** just by looking at their codes.

It's like giving each word a home on a map, and words that are neighbors on this map have related meanings.

- Words can be represented as codes in a multi-dimensional space, where similar words are located near each other based on their meanings.

▶ 2:02:25



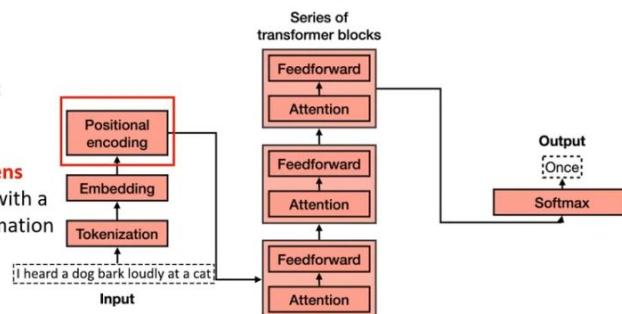
Positional encoding

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Positional encoding is a technique used to ensure that a language model, such as GPT (Generative Pre-trained Transformer) doesn't lose the **order of words** when processing natural language. This is important because the order in which words appear can change the meaning of a sentence.

Let's take the sentence "**I heard a dog bark loudly at a cat**" from our previous example:

Without positional encoding, if we simply tokenize this sentence and convert the **tokens** into **embedding vectors**, we might end up with a set of vectors that **lose the sequence** information



Positional encoding adds a **positional vector** to each word to keep track of the positions of the words.

▶ 2:02:58



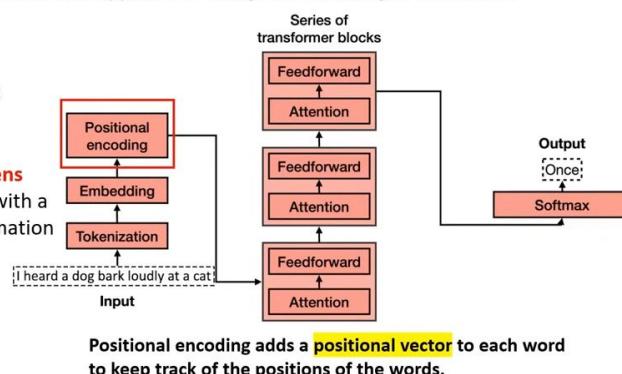
Positional encoding

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Positional encoding is a technique used to ensure that a language model, such as GPT (Generative Pre-trained Transformer) doesn't lose the **order of words** when processing natural language. This is important because the order in which words appear can change the meaning of a sentence.

Let's take the sentence "**I heard a dog bark loudly at a cat**" from our previous example:

Without positional encoding, if we simply tokenize this sentence and convert the **tokens** into **embedding vectors**, we might end up with a set of vectors that **lose the sequence** information



- Positional encoding enhances word embeddings by adding unique positional vectors to maintain the sequence information of words in a sentence.

▶ 2:03:20



Positional encoding

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

However, by adding **positional encoding vectors to each word's embedding**, we ensure that each **position** in the sentence is uniquely identified:

- The embedding for "I" would be modified by adding a positional vector corresponding to position 1, labeled "**I (1)**".
- The embedding for "heard" would be altered by a vector for position 2, labeled "**heard (2)**".
- The embedding for "a" would be updated with a vector for position 3, labeled "**a (3)**", and reused with the same positional vector for its second occurrence.

- Each word token in a sentence is modified by adding a unique positional vector that corresponds to its position in the sequence.

▶ 2:03:30



Positional encoding

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

However, by adding **positional encoding vectors to each word's embedding**, we ensure that each **position** in the sentence is uniquely identified:

- The embedding for "I" would be modified by adding a positional vector corresponding to position 1, labeled "I (1)".
- The embedding for "heard" would be altered by a vector for position 2, labeled "heard (2)".
- The embedding for "a" would be updated with a vector for position 3, labeled "a (3)", and reused with the same positional vector for its second occurrence.
- This process continues for each word/token in the sentence, with "dog (4)", "bark (5)", "loudly (6)", "at (7)", and "cat (8)" all receiving their unique positional encodings.

As a result, the sentence "**I heard a dog bark loudly at a cat**" is represented not just by a sequence of vectors for its words, but by a sequence of vectors that are influenced by the **position** of each word in the sentence.

This means that even if another sentence had the same words in a different order, its overall representation would be different because the positional encodings would differ, reflecting the different sequence of words.

- Positional encodings influence the representation of words in a sentence by accounting for their order, resulting in different vector sequences for sentences with the same words arranged differently.

2:04:15



Attention

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Attention in AI, especially in transformer models, is a way the model figures out how **important each word (or token) is to the meaning of a sentence**, particularly in **relation** to the other words around it. Let's reuse the sentence "**I heard a dog bark loudly at a cat**" to explain this better:

- 1 **Self-Attention:** Imagine each word in the sentence shining a flashlight on the other words. The brightness of the light shows how much one word should pay attention to the others when understanding the sentence. For "bark", the light might shine brightest on "dog" because they're closely related.

- Self-attention in a transformer model allows each word to focus on other words in a sentence, with the intensity of attention reflecting their relevance, which helps create contextually appropriate numerical representations for the words.

2:04:44



Attention

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Attention in AI, especially in transformer models, is a way the model figures out how **important each word (or token) is to the meaning of a sentence**, particularly in **relation** to the other words around it. Let's reuse the sentence "I heard a dog bark loudly at a cat" to explain this better:

Self-Attention: Imagine each word in the sentence shining a flashlight on the other words.

- 1 The brightness of the light shows how much one word should pay attention to the others when understanding the sentence.

For "bark", the light might shine brightest on "dog" because they're closely related.

Encoder's Role: In the encoder part of a transformer model, attention helps decide **how to represent each word as a number (or vector)**. It's not just the word itself, but also its context that matters.

For example, "bark" in "the bark of a tree" would have a different representation than "bark" in "I heard a dog bark", because the surrounding words are different.

💡 Attention in a transformer model helps determine word representations based on their context, influencing how the decoder generates new text by identifying important preceding words.

▶ 2:05:00



Attention

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Attention in AI, especially in transformer models, is a way the model figures out how **important each word (or token) is to the meaning of a sentence**, particularly in **relation** to the other words around it. Let's reuse the sentence "I heard a dog bark loudly at a cat" to explain this better:

Self-Attention: Imagine each word in the sentence shining a flashlight on the other words.

- 1 The brightness of the light shows how much one word should pay attention to the others when understanding the sentence.

For "bark", the light might shine brightest on "dog" because they're closely related.

Encoder's Role: In the encoder part of a transformer model, attention helps decide **how to represent each word as a number (or vector)**. It's not just the word itself, but also its context that matters.

For example, "bark" in "the bark of a tree" would have a different representation than "bark" in "I heard a dog bark", because the surrounding words are different.

▶ 2:05:17



Attention

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Decoder's Role: When generating new text, like completing a sentence, the decoder uses attention to figure out which words it already has are **most important** for **deciding what comes next**. If our sentence is "I heard a dog," the model uses **attention** to know that "heard" and "dog" are key to adding the next word, which might be "bark."

Multi-Head Attention: It's like having multiple flashlights, each **highlighting different aspects of the words**. Maybe one flashlight looks at the **meaning** of the word, another looks at its **role** in the sentence (like subject or object), and so on. This helps the model get a richer understanding of the text.

Building the Output: The decoder builds the sentence one word at a time, using **attention** at each step. It looks at the sentence so far, decides what's important, and then **predicts** the next word. It's an ongoing process, with each **new word influencing the next**.

Attention in transformer models is like a guide that helps the AI understand and create language by focusing on the most relevant parts of the text, considering both individual word meanings and their relationships within the sentence.

The model utilizes multi-head attention to analyze key words in a sentence, enhancing its understanding and enabling it to generate output word by word.

▶ 2:05:26



Attention

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Decoder's Role: When generating new text, like completing a sentence, the decoder uses attention to figure out which words it already has are **most important** for **deciding what comes next**. If our sentence is "I heard a dog," the model uses **attention** to know that "heard" and "dog" are key to adding the next word, which might be "bark."

Multi-Head Attention: It's like having multiple flashlights, each **highlighting different aspects of the words**. Maybe one flashlight looks at the **meaning** of the word, another looks at its **role** in the sentence (like subject or object), and so on. This helps the model get a richer understanding of the text.

Building the Output: The decoder builds the sentence one word at a time, using **attention** at each step. It looks at the sentence so far, decides what's important, and then **predicts** the next word. It's an ongoing process, with each **new word influencing the next**.

Attention in transformer models is like a guide that helps the AI understand and create language by focusing on the most relevant parts of the text, considering both individual word meanings and their relationships within the sentence.

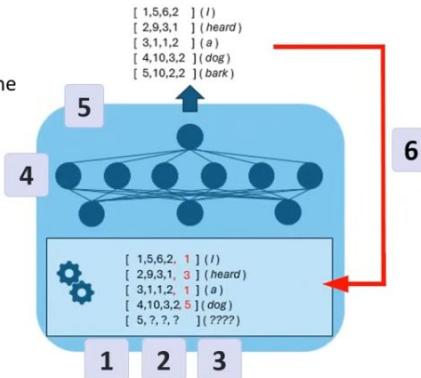
▶ 2:06:10



Attention – Process

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

- 1 **Token Embeddings:** Each word in the sentence is represented as a **vector** of numbers (its **embedding**).
- 2 **Predicting the Next Token:** The goal is to figure out what the next **word (token)** should be, also represented as a vector.
- 3 **Assigning Weights:** The attention layer looks at the sentence so far and decides how much **influence (weight)** each word should have on the next one.
- 4 **Calculating Attention Scores:** Using these weights, a new vector for the next token is calculated, which includes an attention score. **Multi-head attention** does this several times, focusing on different aspects of the words.
- 5 **Choosing the Most Likely Word:** A neural network takes these vectors with **attention scores** and picks the **word** from the vocabulary that most likely comes next.
- 6 **Adding to the Sequence:** The chosen word is **added to the existing sequence**, and the process **repeats for each new word**.



💡 The process involves predicting the next word by calculating attention scores for each word in a sentence, using multi-head attention to focus on different aspects, and employing a neural network to select the most likely subsequent word from the vocabulary.

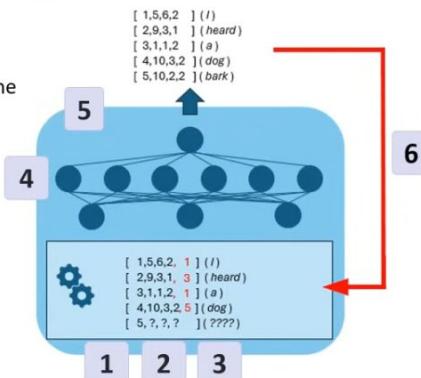
▷ 2:06:24



Attention – Process

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

- 1 **Token Embeddings:** Each word in the sentence is represented as a **vector** of numbers (its **embedding**).
- 2 **Predicting the Next Token:** The goal is to figure out what the next **word (token)** should be, also represented as a vector.
- 3 **Assigning Weights:** The attention layer looks at the sentence so far and decides how much **influence (weight)** each word should have on the next one.
- 4 **Calculating Attention Scores:** Using these weights, a new vector for the next token is calculated, which includes an attention score. **Multi-head attention** does this several times, focusing on different aspects of the words.
- 5 **Choosing the Most Likely Word:** A neural network takes these vectors with **attention scores** and picks the **word** from the vocabulary that most likely comes next.
- 6 **Adding to the Sequence:** The chosen word is **added to the existing sequence**, and the process **repeats for each new word**.



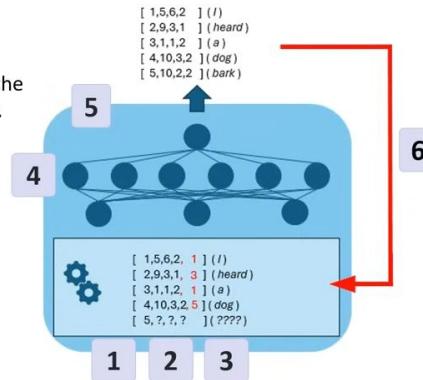
▷ 2:06:55



Attention – Process

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

- 1** **Token Embeddings:** Each word in the sentence is represented as a **vector** of numbers (its **embedding**).
- 2** **Predicting the Next Token:** The goal is to figure out what the next **word (token)** should be, also represented as a vector.
- 3** **Assigning Weights:** The attention layer looks at the sentence so far and decides how much **influence (weight)** each word should have on the next one.
- 4** **Calculating Attention Scores:** Using these weights, a new vector for the next token is calculated, which includes an attention score. **Multi-head attention** does this several times, focusing on different aspects of the words.
- 5** **Choosing the Most Likely Word:** A neural network takes these vectors with **attention scores** and picks the **word** from the vocabulary that most likely comes next.
- 6** **Adding to the Sequence:** The chosen word is **added to the existing sequence**, and the process **repeats for each new word**.



- ✿ A neural network generates text by repeatedly selecting the most likely next word based on attention scores and existing sequences.

▷ 2:06:59



GPT-4

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



A transformer model like **GPT-4** works by taking a text **input (prompt)** and producing a well-structured **output (completion)**. During training, it learns from a vast array of text data, understanding how words are typically arranged in sentences.

The model knows the correct sequence of words but **hides (masks)** future words to learn how to **predict** them. When it tries to predict a word, it compares its guess to the actual word, gradually adjusting to reduce errors.

In practice, the model uses its training to assign **importance (weights)** to each word in a sequence, helping it guess the next word accurately. The result is that GPT-4 can create **sentences** that sound like they were written by a human.

However, this doesn't mean the model "knows" things or is "intelligent" in the human sense. It's simply very good at using its **large vocabulary and training to generate realistic text based on word relationships**.

- ✿ The model learns to predict the next word in a sequence by comparing its guesses to actual words and adjusting based on errors, allowing it to generate human-like sentences without possessing true understanding or intelligence.

▷ 2:07:17



GPT-4

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



A transformer model like GPT-4 works by taking a text **input (prompt)** and producing a well-structured **output (completion)**. During training, it learns from a vast array of text data, understanding how words are typically arranged in sentences.

The model knows the correct sequence of words but **hides (masks)** future words to learn how to **predict** them. When it tries to predict a word, it compares its guess to the actual word, gradually adjusting to reduce errors.

In practice, the model uses its training to assign **importance (weights)** to each word in a sequence, helping it guess the next word accurately. The result is that GPT-4 can create **sentences** that sound like they were written by a human.

However, this doesn't mean the model "knows" things or is "intelligent" in the human sense. It's simply very good at using its **large vocabulary and training to generate realistic text based on word relationships**.

- ◆ The Azure OpenAI Service utilizes advanced language models to generate realistic text based on word relationships, demonstrating a high level of vocabulary and training without possessing true human intelligence.

▷ 2:07:55



Azure OpenAI Service

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Azure OpenAI Service is a cloud-based platform designed to **deploy and manage advanced language models from OpenAI**. This service combines OpenAI's latest language model developments with the robust security and scalability of Azure's cloud infrastructure.

Azure OpenAI offers several **types of models** for different purposes:



- **GPT-4 Models:** These are the newest in the line of GPT models and can create **text and programming code** when given a prompt written in natural language.
- **GPT-3.5 Models:** Similar to GPT-4, these models also create text and code from natural language prompts. The GPT-3.5-turbo version is specially designed for **conversations**, making it a great choice for **chat applications and other interactive AI tasks**.

- ◆ A cloud-based platform integrates OpenAI's advanced language models with Azure's security and scalability, offering various model types for generating text and code from natural language prompts.

▷ 2:08:09



Azure OpenAI Service

Cheat sheets, Practice Exams and Flash cards

www.exampro.co/ai-900

Azure OpenAI Service is a cloud-based platform designed to **deploy and manage advanced language models from OpenAI**. This service combines OpenAI's latest language model developments with the robust security and scalability of Azure's cloud infrastructure.

Azure OpenAI offers several **types of models** for different purposes:



- **GPT-4 Models:** These are the newest in the line of GPT models and can create **text and programming code** when given a prompt written in natural language.
- **GPT-3.5 Models:** Similar to GPT-4, these models also create text and code from natural language prompts. The GPT-3.5-turbo version is specially designed for **conversations**, making it a great choice for **chat applications and other interactive AI tasks**.
- **Embedding Models:** These models turn written **text into number sequences**, which is helpful for analyzing and comparing different pieces of text to find out how **similar** they are.
- **DALL-E Models:** These models can make **images from descriptions** given in words. The DALL-E models are still being tested and aren't shown in the Azure OpenAI Studio, so you don't have to set them up for use manually.



- Azure OpenAI provides models that generate images from text descriptions and allows users to analyze text similarity through prompts, completions, tokens, and other resources.

▶ 2:09:07



Azure OpenAI Service

Cheat sheets, Practice Exams and Flash cards

www.exampro.co/ai-900

Key concepts in using Azure OpenAI include **prompts and completions, tokens, resources, deployments, prompt engineering, and various models**:

Prompts & Completions: Users interact with the **API** by providing a text command in English, known as a **prompt**, and the model generates a text response, or completion.

- E.g., a prompt to count to five in a loop results in the model returning appropriate code.

Tokens: Azure OpenAI breaks down text into **tokens**, which are words or character chunks, to process requests. The number of tokens affects response latency and throughput.

- For images, token cost varies with image size and detail setting, with low-detail images costing fewer tokens and high-detail images costing more.

Resources: Azure OpenAI operates like other **Azure products** where users create a resource within their Azure Subscription.

Deployments: To use the service, users must deploy a **model via Deployment APIs**, choosing the specific model for their needs.

Prompt Engineering: Crafting **prompts** is crucial as they guide the model's output.

This requires skill, as prompt construction is nuanced and impacts the model's response.

Models: Various models offer different capabilities and pricing. **DALL-E** creates images from text, while **Whisper** transcribes and translates speech to text. Each has unique features suitable for different tasks.

- Different AI models have unique capabilities and pricing, with some generating images from text while others transcribe and translate speech, all of which can be utilized by developers in Azure OpenAI Studio.

▶ 2:10:22



Azure OpenAI Studio

Cheat sheets, Practice Exams and Flash cards [👉 www.exampro.co/ai-900](https://www.exampro.co/ai-900)

Developers can work with these models in **Azure OpenAI Studio**, a web-based environment where AI professionals can **deploy, test, and manage LLMs** that support generative AI app development on Azure.

Access is currently **limited** due to the high demand, upcoming product improvements, and Microsoft's commitment to responsible AI.

Presently, collaborations are being prioritized for those who already have a **partnership with Microsoft**, are engaged in lower-risk use cases, and are dedicated to including necessary safeguards.

The screenshot shows the Azure OpenAI Studio portal. On the left, there is a sidebar with navigation links: Azure OpenAI, Playgrounds, Chat, Completions, DALLE (Preview), Management, Deployments, Models, Data files, Quotas, and Content filters (Preview). The main area is titled "Welcome to Azure OpenAI service" with the sub-instruction "Explore the generative AI models, craft unique prompts for your use cases, and fine-tune select models.". A red arrow points to the "Create new deployment" button. Below this, there are two sections: "Get started" with "Chat playground" and "Completions playground", each with a "Try it now" button.

- Access to large language models for generative AI app development on Azure is currently limited due to high demand and a focus on responsible AI, prioritizing collaborations with existing partners engaged in lower-risk use cases.

▶ 2:10:41



Azure OpenAI Studio

Cheat sheets, Practice Exams and Flash cards [👉 www.exampro.co/ai-900](https://www.exampro.co/ai-900)

Developers can work with these models in **Azure OpenAI Studio**, a web-based environment where AI professionals can **deploy, test, and manage LLMs** that support generative AI app development on Azure.

Access is currently **limited** due to the high demand, upcoming product improvements, and Microsoft's commitment to responsible AI.

Presently, collaborations are being prioritized for those who already have a **partnership with Microsoft**, are engaged in lower-risk use cases, and are dedicated to including necessary safeguards.

The screenshot shows the Azure OpenAI Studio portal. On the left, there is a sidebar with navigation links: Azure OpenAI, Playgrounds, Chat, Completions, DALLE (Preview), Management, Deployments, Models, Data files, Quotas, and Content filters (Preview). The main area is titled "Welcome to Azure OpenAI service" with the sub-instruction "Explore the generative AI models, craft unique prompts for your use cases, and fine-tune select models.". A red arrow points to the "Create new deployment" button. Below this, there are two sections: "Get started" with "Chat playground" and "Completions playground", each with a "Try it now" button.

- Collaborations for AI product improvements are currently limited to existing partners engaged in lower risk use cases with necessary safeguards.

▶ 2:11:00



Azure OpenAI Studio

In Azure AI Studio, you can deploy large language models, provide few-shot examples, and test them in Azure OpenAI Studio's Chat playground.

The image shows **Azure OpenAI's Chat playground interface**, where users can test and configure an AI chatbot.

In the middle, there's a **chat area** to type user messages and see the assistant's replies.

On the left, there's a menu for **navigation** and a section to set up the **assistant**, including a reminder to save changes.

On the right, adjustable **parameters control** the AI's response behavior, like **length, randomness, and repetition**. Users input queries, adjust settings, and observe how the AI responds to fine-tune its performance.

- Azure OpenAI Studio allows users to adjust parameters that control the AI's response behavior and fine-tune its performance, with pricing information available for models like GPT-3.5 Turbo.

▶ 2:11:37



Azure OpenAI Service pricing

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Pricing for Language models

Models	Context	Prompt (Per 1,000 tokens)	Completion (Per 1,000 tokens)
GPT-3.5-Turbo	4K	\$0.0015	\$0.002
GPT-3.5-Turbo	16K	\$0.003	\$0.004
GPT-3.5-Turbo-1106	16K	N/A	N/A
GPT-4-Turbo	128K	N/A	N/A
GPT-4-Turbo-Vision	128K	N/A	N/A
GPT-4	8K	\$0.03	\$0.06
GPT-4	32K	\$0.06	\$0.12

- Azure OpenAI Service offers GPT-3.5 Turbo models with different token limits and associated costs for prompts and completions, with prices varying based on the model's capabilities.

▶ 2:12:02



Azure OpenAI Service pricing

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Base models		Image models		
Models	Usage per 1,000 tokens	Models	Quality	Resolution
Babbage-002	\$0.0004	Dall-E-3	Standard	1024 * 1024
Davinci-002	\$0.002		Standard	1024 * 1792, 1792 * 1024
Fine-tuning models				
Models	Training per compute hour	Hosting per hour	Input Usage per 1,000 tokens	Output Usage per 1,000 tokens
Babbage-002	\$34	\$1.70	\$0.0004	\$0.0004
Davinci-002	\$68	\$3	\$0.002	\$0.002
GPT-3.5-Turbo	\$102	\$7	\$0.0015	\$0.002
Embedding models				
Models	Per 1,000 tokens			
Ada	\$0.0001			
Speech models				
Models	Per hour			
Whisper	\$0.36			

- Various models, including image, embedding, and speech models, have different pricing structures based on usage, with higher quality models generally costing more.

▶ 2:13:03

What are copilots?

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Copilots are a new type of computing tool that integrates with applications to help users with common tasks using generative AI models. They are designed using a standard architecture, allowing developers to create custom copilots tailored to specific business needs and applications.

- Copilots might appear as a chat feature beside your document or file, and they utilize the content within the product to generate specific results.

Creating a copilot involves several steps:

1. Training a large language model with a vast amount of data.
2. Utilizing services like Azure OpenAI Service, which provide pre-trained models that developers can either use as-is or fine-tune with their own data for more specific tasks.
3. Deploying the model to make it available for use within applications.
4. Building copilots that prompt the models to generate usable content.
5. Business users can use copilots to boost their productivity and creativity with AI-generated content.

Copilots have the potential to revolutionize the way we work. These copilots use generative AI to help with first drafts, information synthesis, strategic planning, and much more.

- Co-pilots utilizing generative AI can significantly enhance productivity and creativity for business users by assisting with tasks such as information synthesis and strategic planning.

▶ 2:14:16

Copilot Examples

Cheat sheets, Practice Exams and Flash cards [👉 www.exapro.co/ai-900](http://www.exapro.co/ai-900)



Microsoft Copilot is integrated into various applications to assist users in creating **documents, spreadsheets, presentations, and more**, by generating content, summarizing information, and aiding in strategic planning.

- It is used across Microsoft's suite of products and services to enhance user **experience and efficiency**.

A screenshot of the Microsoft AI interface. On the left, there's a preview of a cover letter for a job application with a "masters degree in political science". On the right, there are options for "Professional", "Casual", and "Entertaining" styles, along with "Length" settings for "Short", "Medium", and "Long". A "Generate draft" button is visible at the bottom.



The Microsoft Bing search engine provides a copilot to help when **browsing or searching** the Internet by generating **natural language answers** to questions based on context rather than just search results of indexed pages.

A screenshot of the Microsoft Bing AI Search interface. It shows a search query for "Create a cover letter for a job application for a candidate with a masters degree in political science". Below the search bar, there are buttons for "More Creative", "More Balanced" (which is selected), and "More Precise". A generated cover letter is displayed, starting with "Dear Hiring Manager," followed by a paragraph about the user's interest in the position and their qualifications.

- Microsoft's search engine features an integrated co-pilot that enhances the browsing experience by generating natural language answers and is also designed to assist users in productivity tasks across various applications.

▷ 2:14:59

Copilot Examples

Cheat sheets, Practice Exams and Flash cards [👉 www.exapro.co/ai-900](http://www.exapro.co/ai-900)



Microsoft 365 Copilot is designed to be a partner in your workflow, integrated with productivity and communication tools like **PowerPoint** and **Outlook**.

- It's there to help you craft effective **documents**, design **spreadsheets**, put together **presentations**, manage emails, and streamline other tasks.

A screenshot of the Microsoft 365 Copilot interface. It shows a "Create with Copilot" tool where a user has described the desire to create a "bulletin board showing postcards, sticky notes, and photos, black and white, artistic photography style". Below the description, there are several thumbnail previews of different image compositions. At the bottom, there are "Insert" and "Cancel" buttons.

GitHub Copilot

GitHub Copilot is a tool that helps software developers, offering real-time assistance as they **write code**. It offers more than suggesting code snippets; it can help in **documenting the code** for better understanding and maintenance.

A screenshot of the GitHub Copilot interface integrated into Visual Studio Code. The code editor shows a file named "test.js" with the following content:

```
C:\> Users > berryvor > JS testjs > findimagesWithoutAlt
1 // find all images without alternate text
2 // and give them a red border
3 function findImagesWithoutAlt() {
4     var images = document.getElementsByTagName("img");
5     for (var i = 0; i < images.length; i++) {
6         if (images[i].alt == "") {
7             images[i].style.border = "2px solid red";
8         }
9     }
10 }
```

- GitHub Copilot is a tool that assists software developers by providing real-time coding support, helping with documentation and testing to enhance efficiency and reduce errors.

▷ 2:15:27



Prompt engineering

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900



Prompt engineering is a process that **improves the interaction between humans and generative AI**. It involves **refining the prompts** or instructions given to an AI application to **generate higher quality responses**. This process is valuable for both the developers who create AI-driven applications and the end-users who interact with them.

For example, developers may build a generative AI application for teachers to create multiple-choice questions related to text students read. During the development of the application, developers can **add other rules** for what the program should do with the prompts it receives.

System messages

Prompt engineering techniques include defining a system message. The message sets the **context for the model** by describing **expectations** and **constraints**.

For example, "You're a helpful assistant that responds in a **cheerful, friendly manner**".

These system messages determine constraints and styles for the model's responses.

Writing good prompts

To maximize the utility of AI responses, it is essential to be **precise and explicit** in your prompts.

A well-structured prompt, such as "**Create a list of 10 things to do in Edinburgh during August**," directs the AI to produce a targeted and relevant output, achieving better results.

- Effective AI prompts should be precise and structured to enhance the relevance of responses, with zero shot learning enabling task performance without prior examples and one shot learning allowing for adaptation from a single instance.

▷ 2:16:52

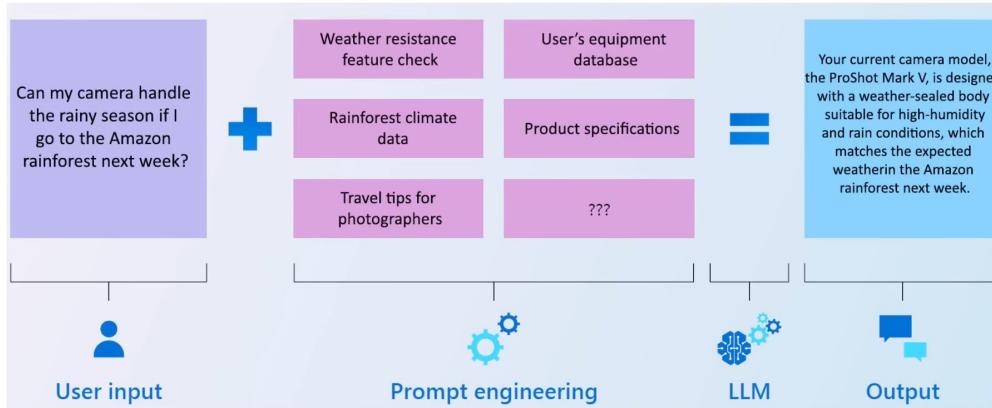


Prompt engineering

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Zero-shot learning refers to an AI model's ability to correctly perform a task **without any prior examples or training** on that specific task.

One-shot learning involves the AI model learning from a **single example or instance** to perform a task.



- Prompt engineering involves integrating user queries with relevant data, such as equipment specifications and environmental conditions, to provide tailored responses.

▷ 2:17:19



Prompt engineering workflow

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

1. Task Understanding: Know what you want the AI to do.

2. Crafting Prompts: Write instructions for the AI.

3. Prompt Alignment: Make sure instructions match what the AI can do.

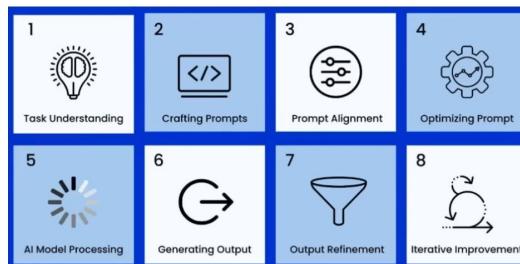
4. Optimizing Prompt: Improve the instructions for better AI responses.

5. AI Model Processing: The AI thinks about the instructions.

6. Generating Output: The AI gives an answer or result.

7. Output Refinement: Fix or tweak the AI's answer.

8. Iterative Improvement: Keep improving the instructions and answers.



💡 Prompt engineering involves iterative improvement of instructions and outputs to enhance the performance of large language models.

▷ 2:18:44

Grounding

Cheat sheets, Practice Exams and Flash cards www.exampro.co/ai-900

Grounding in prompt engineering is a technique used in large language models (LLMs) where you provide **specific, relevant context within a prompt**. This helps the AI to produce a more accurate and related response.

For example, if you want an LLM to summarize an email, you would include the actual email text in the prompt along with a command to summarize it. This approach allows you to leverage the LLM for tasks it wasn't explicitly trained on, without the need for retraining the model.

Prompt engineering vs Grounding

Prompt engineering broadly refers to the art of crafting effective prompts to produce the desired output from an AI model. **Grounding** specifically involves enriching prompts with relevant context to improve the model's understanding and responses.

Grounding ensures the AI has **enough information to process the prompt** correctly, whereas **prompt engineering** can also include techniques like format, style, and the strategic use of examples or questions to guide the AI.

💡 Effective prompt engineering enhances AI model understanding and response accuracy by incorporating relevant context, strategic formatting, and guided examples or questions.

▷ 2:19:44

Grounding options

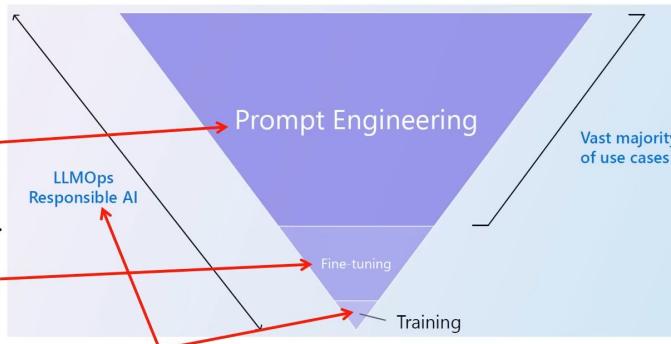
Cheat sheets, Practice Exams and Flash cards  www.exampro.co/ai-900

Grounding Options: These are techniques to ensure LLM outputs are accurate and adhere to responsible AI principles.

Prompt Engineering: Placed at the top, indicating its broad applicability, this involves designing prompts to direct the AI toward generating the desired output.

Fine-Tuning: A step below in complexity, where LLMs are trained on specific data to improve their task performance.

Training: The most resource-intensive process, at the triangle's base, suggesting its use in more extensive customization needs.



LLMOps and Responsible AI: These foundational aspects emphasize the importance of operational efficiency and ethical standards across all stages of LLM application development.

- The text discusses the significance of operational efficiency and ethical standards in large language model operations and responsible AI, while also introducing a demo on using Co-Pilot with GPT-4 on Microsoft Bing.

▶ 2:20:29

Name	Type	Resource group	Location	Subscription
b094978d080944c79480e1bd41acf9fd	Container registry	MyStudio	East US	Azure subscription 1
b1nangpipeline-2eUvQkx0t0J0OPI9Qz5jQ	Container instances	MyStudio	East US	Azure subscription 1
diabeteprod-2eUvQkx0t0J0OPI9Qz5jQ	Container instances	MyStudio	East US	Azure subscription 1
kkas	Storage Sync Service	kkas	Canada East	Azure subscription 1
MyM3Workplace	Machine learning	MyStudio	East US	Azure subscription 1
nymworkplace1828757807	Key vault	MyStudio	East US	Azure subscription 1
nymworkplace2545665410	Application insights	MyStudio	East US	Azure subscription 1
nymworkplace9413440532	Storage account	MyStudio	East US	Azure subscription 1
NetworkWatcher_centralus	Network Watcher	NetworkWatcherRG	Central US	Azure subscription 1
NetworkWatcher_eastus	Network Watcher	NetworkWatcherRG	East US	Azure subscription 1
picardult	Recovery Services vault	picard	Central US	Azure subscription 1

- Check all resources and delete anything that is running.

▶ 4:23:39