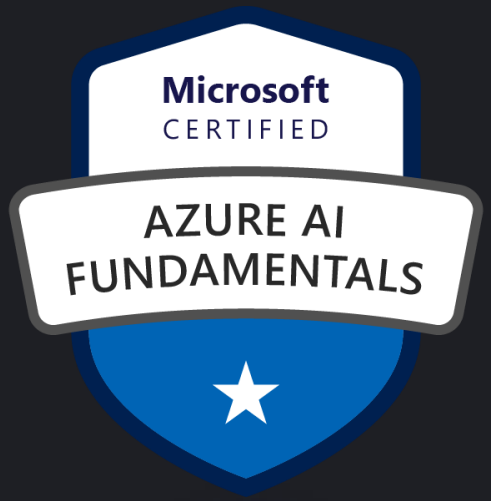
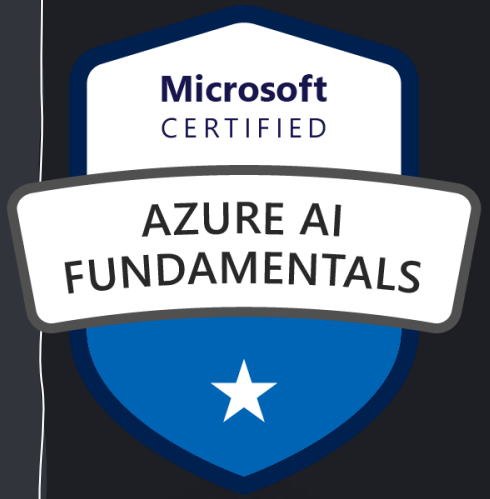
INTRODUCTION

Section 1:

About the exam & course setup



About the AI-900 Azure AI Fundamentals Certification



About AI-900 Exam

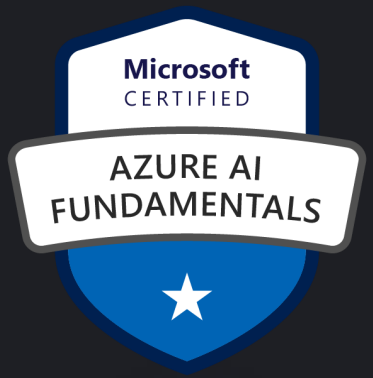
What is covered?

- ✓ Azure AI Fundamentals | AI-900 exam

https://learn.microsoft.com/en-us/credentials/certifications/azure-ai-fundamentals/

Why AI-900?

- ✓ Career Growth
- ✓ Future-proof skills
- ✓ High in-demand roles



Role

AI Engineer

About AI-900 Exam

Demos

- ✓ Not needed for exam
- ✓ Help with memorizing
- ✓ Practical foudation

Goal

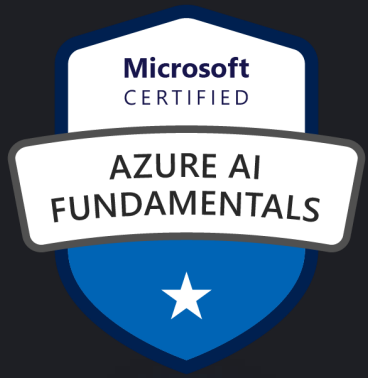
- ✓ Clear exam
- ✓ Knowledge working with Azure

Exam details

- ✓ Duration: 45min
- ✓ Questions: 40-60

Passing Score

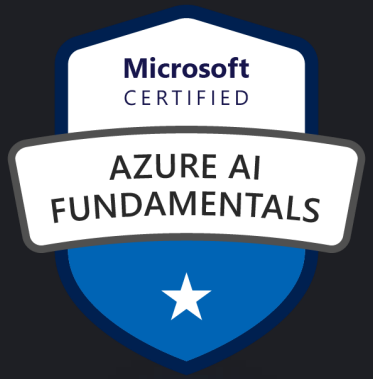
- ✓ 700 / 1000
- ✓ Goal: ~~850~~ 950 +



About AI-900 Exam

Exam Questions

- ✓ Fundamental AI & ML concepts
- ✓ AI & ML services
- ✓ No implementation



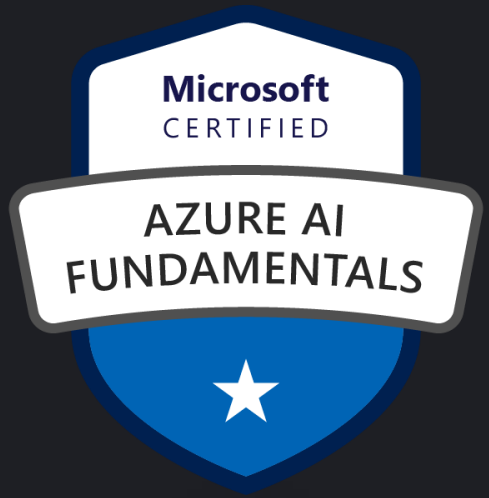
Example

What term describes the process where a system identifies a face in an image and highlights its position using a bounding box?

- ☑ Face Detection
- ❑ Pixel-Level Face Classification
- ❑ Face Categorization
- ❑ Semantic Segmentation

Working with the course

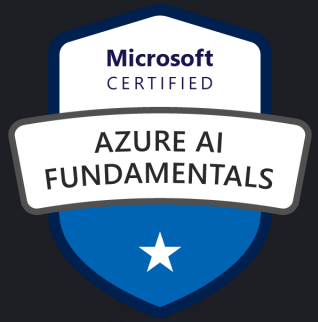
How to pass the exam



Recipe to clear the exam

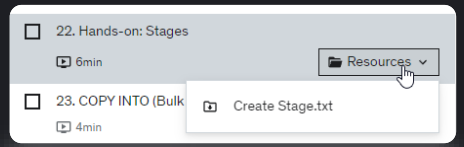
- **Lectures** — Step-by-step incl. Demos ~ 30-60 min / day
- **Quizzes** — Practice and test your knowledge
- **Slides** — Repeat and go through important points
- **Practice Test** — Evaluate knowledge & weaknesses; Eliminate weaknesses
- **Book Exam**

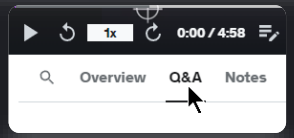
→ **Confident & Prepared**

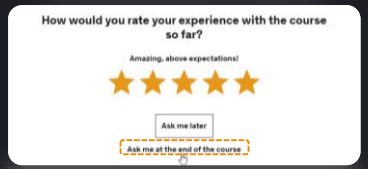


Working in Course & Passing Exam

- Resources
- Q&A Section
- Reviews
- Connect & Congratulate







About AI-900 Exam

Why AI-900?

- Career Growth
- Future-proof skills
- High in-demand roles

Role

- AI Engineer

Exam details

- Duration: 45min
- Questions: 40-60
- Passing Scrore: 700 / 1000
- Cost: $99 (country-specific)

Introduction to AI

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Introduction to AI

Artificial Intelligence (AI)

- AI is the imitation of human intelligence to perform tasks.
- Imitation is done by a machine or a software.
- Workloads: Machine Learning, Deep Learning, and Natural Processing.

Types of AI:

Traditional: Uses pre-programmed rules and algorithms.

Generative: Uses large datasets to learn patterns.

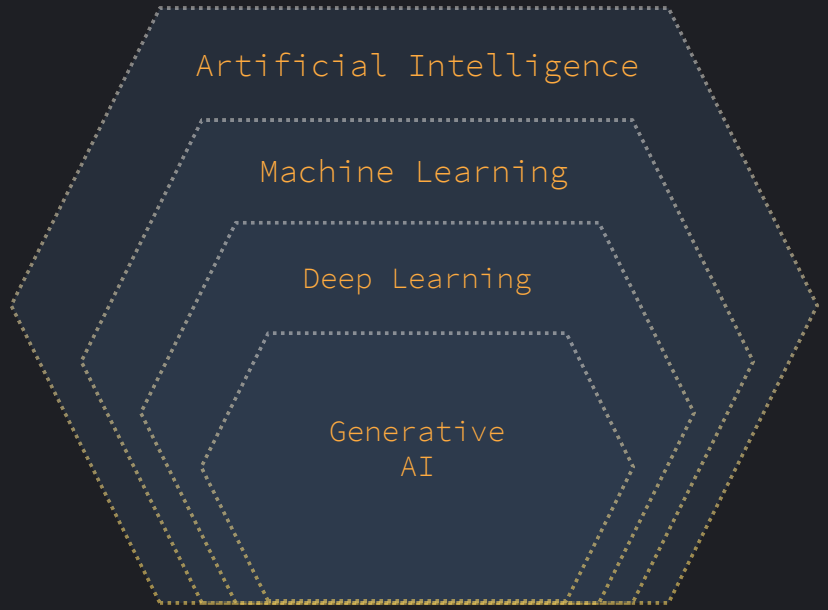
AI Use Cases in Business

- Healthcare: Diagnostics and treatment recommendations.
- Finance: Fraud detection and personalized advice.
- Retail: Inventory management and customer personalization.

Introduction to AI

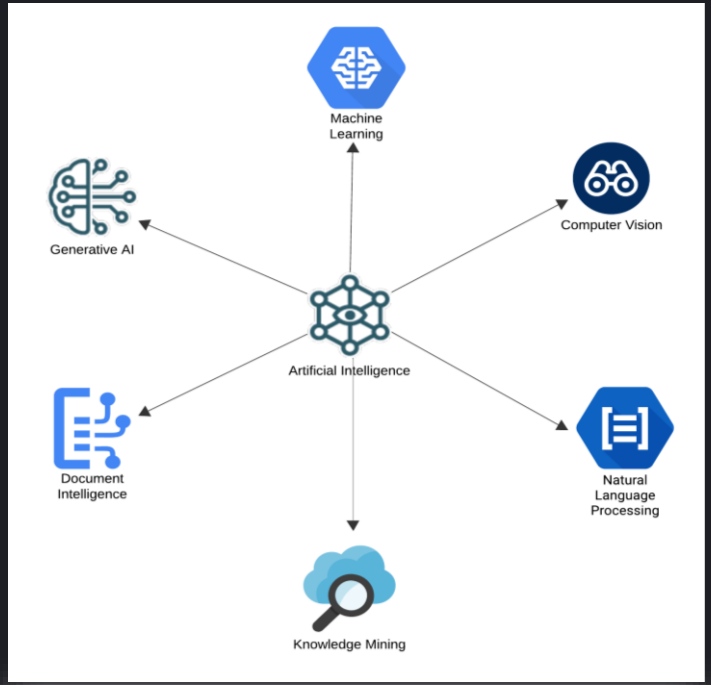
Artificial Intelligence vs Machine Learning vs Deep Learning vs Generative AI

- **Artificial Intelligence:**
  - Everything related to making machines smart.
- Machine Learning:
  - Teaching machines to learn from data.
- Deep Learning:
  - Designed to mimic the way brain work.
- Generative AI:
  - Not only learns from data but also creates new data.



Introduction to AI

Common AI Workloads in Azure



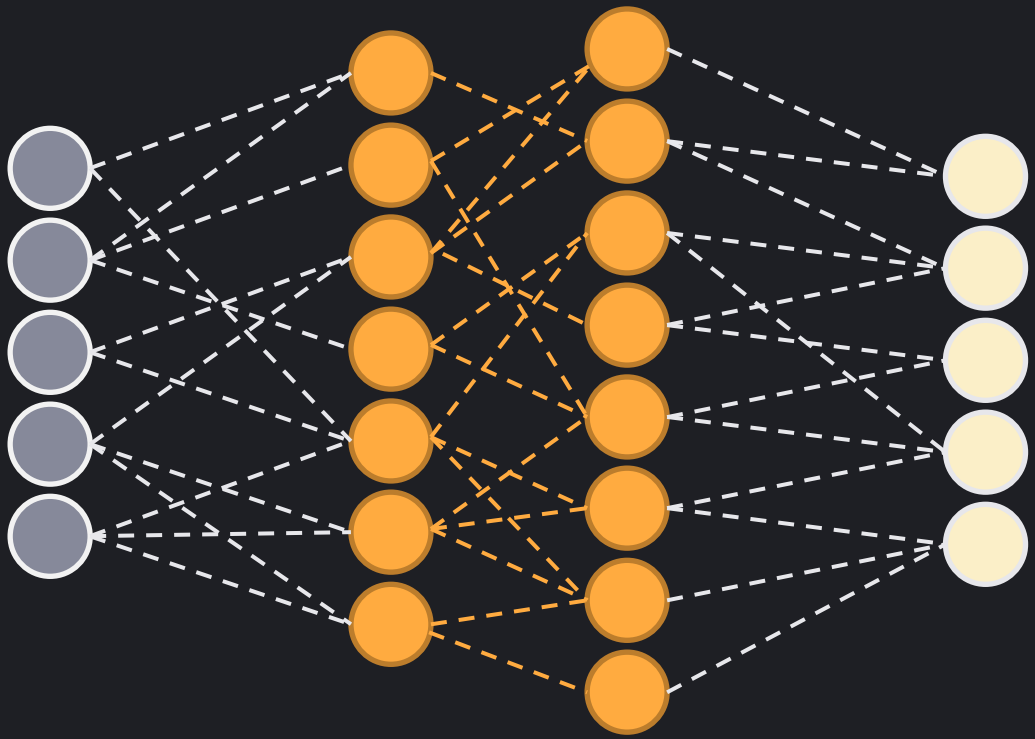
Deep Learning

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Deep Learning

- Advanced type of Machine Learning.
- Inspired by the human brain using artificial neural networks (ANNs).
- Learns from large amounts of data

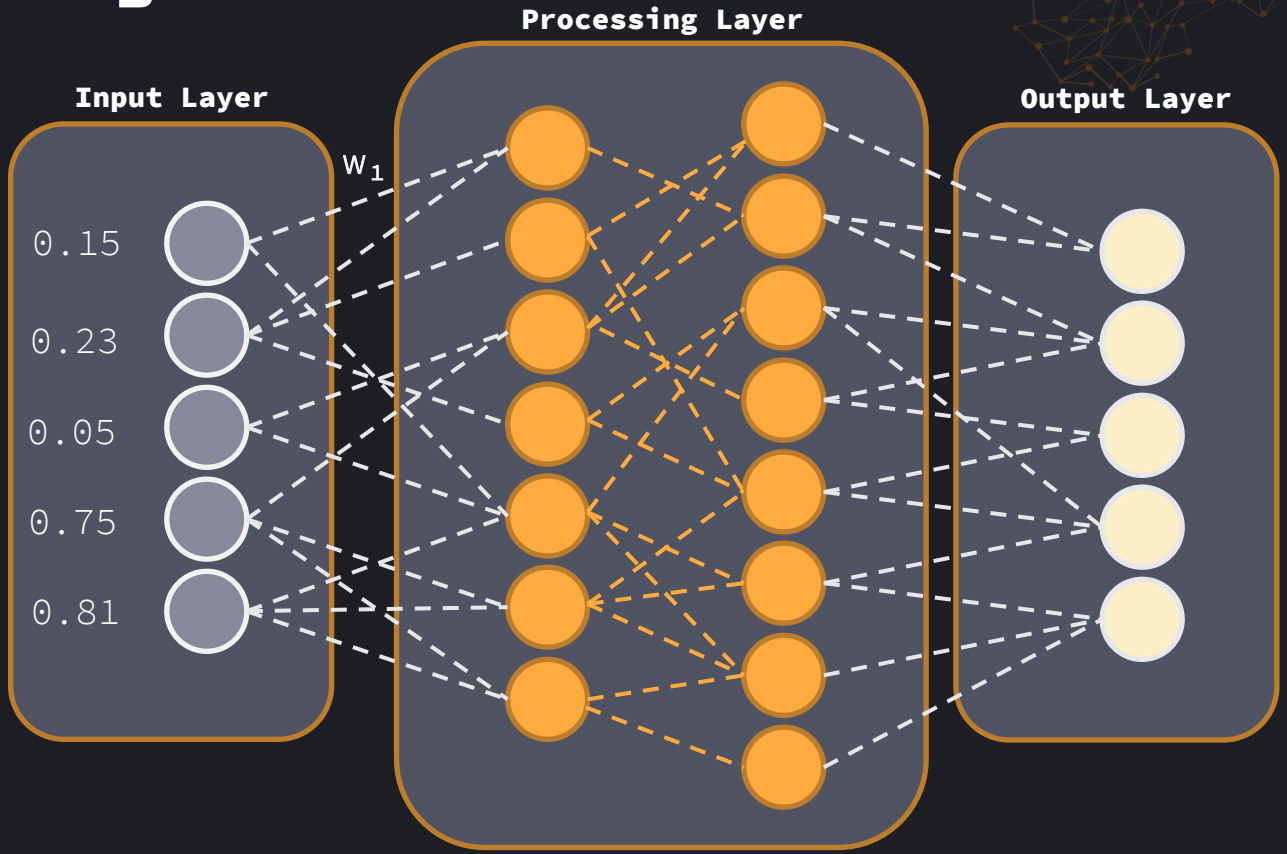
Deep Learning



Deep Learning

Data Representation

Numerical values



Deep Learning

Data Representation

Numerical values

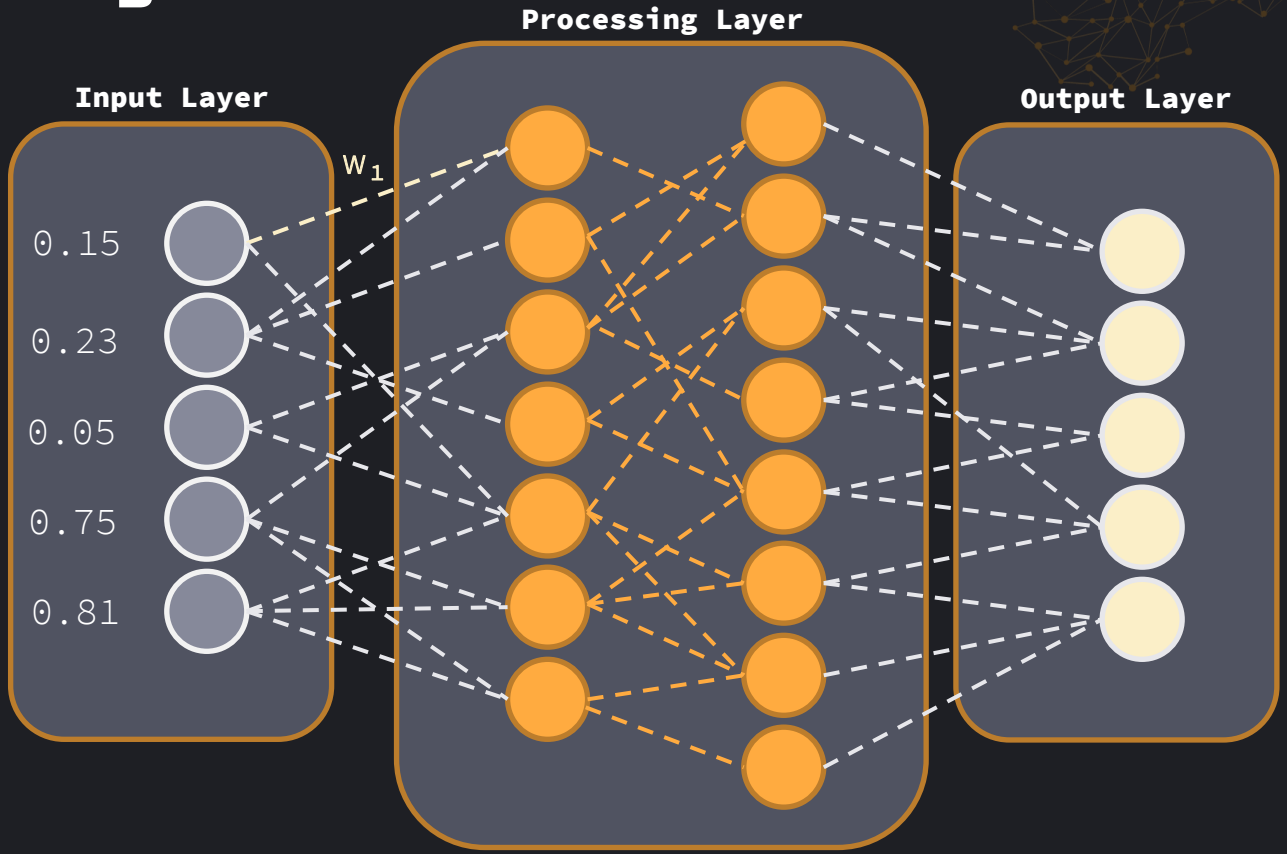
Hierarchical Learning

Patterns are learnt

Scalable

Leverage GPUs

Parallel processing



Deep Learning

Different types of architectures

- Feedforward Networks
- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Transformers

Deep Learning

Machine Learning and Deep Learning Comparison

Feature	(Other) Machine Learning	Deep Learning
Learning Approach	Uses algorithms to learn hidden patterns	Uses artificial Neural Networks to learn patterns
Dataset Size	Smaller datasets	Larger datasets
Use Cases	For simpler tasks with low complexity	For complex tasks e.g. NLP
Training Time	Less model training time	Higher model training time
Computing Requirements	CPU or less computing power.	High performance Computing with GPUs

Azure AI Services

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Azure AI Services

A collection of tools and APIs for integrating AI capabilities/workloads

⇒ Prerequisites: Azure Subscription and Machine Learning Workspace (for custom models)

Key features

- Over a dozen AI services
- Can be used separately or together.
- Can be tailored to specific needs.

Examples of AI Services

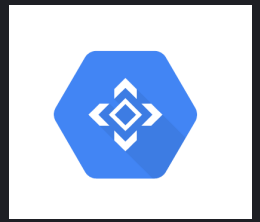
- Azure AI Content Safety
- Azure AI Metrics Advisor
- Azure OpenAI
- Azure AI Vision

Azure AI Services

Three Core Principles

Prebuilt and Ready to Use:

- Uses pre-trained ML models.
- High-Performance Computing.



Accessed Through APIs:

- REST APIs
- Client libraries

Availability and Security

- Enterprise-grade Security.
- Managed as Azure resources.

Azure AI Services

Authentication for Azure AI Services

Users and applications are authenticated through keys and endpoints.

Endpoint

Unique URL for the resource.

```
https://<your-resource-name>.cognitiveservices.azure.com/vision/
```

Resource Key

Protects resource privacy.

Should be updated regularly.

Azure AI Services

Resource Types in Azure AI Services

Multi-service Resources

- Access multiple AI services with one key and endpoint.
- Suitable for exploring multiple capabilities.
- All services are billed together

Single-service Resource

- Access a single AI service.
- Allows detailed cost tracking.

Azure AI Services

Accessing Azure AI Services

APIs

- Define communication between software components.
- Ensures easier integration and updates.

Studio Interfaces

- User-friendly tools like Vision Studio, Language Studio, etc.
- Test and evaluate with sample or custom content.

SDKs

- Access to prebuilt libraries.
- Compatibility with multiple programming languages.

Challenges and Risks with AI

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Challenges and Risks with AI

Problem: Bias Can Affect Results
Solution: Use diverse datasets and fairness audits.

Problem: Errors May Cause Harm
Solution: Conduct rigorous testing and validation.

Problem: Data Could Be Exposed
Solution: Encrypt data and adhere to privacy laws.

Challenges and Risks with AI

Problem: Solutions May Not Work for Everyone
Solution: Design inclusively and test across demographics.

Problem: Users Must Trust a Complex System
Solution: Provide transparency about system processes.

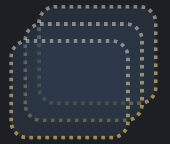
Problem: Who’s Liable for AI-Driven Decisions?
Solution: Establish legal and governance frameworks.

Introduction to Computer Vision

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Computer Vision

- Devices/applications that can understand and interact with visuals.
- Automates recognizing objects and patterns in images and videos.



What is an Image

0	0	0	0
0	255	255	0
0	255	255	0
0	0	0	0

Red

120	95	0	0
0	255	255	0
0	255	255	5
12	0	0	120

Green

255	0	0	0
0	255	255	0
0	255	255	0
120	0	0	190

Blue

0	170	0	110
190	255	255	0
0	255	255	0
180	0	134	0

Computer Vision

Filters on Images

- Common way to perform image processing tasks.

-1	-1	-1
-1	8	-1
-1	-1	-1

Edge Detection Kernel

Scans through each pixels

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	255	255	255	0	0
0	0	255	255	255	0	0
0	0	255	255	255	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Original Image

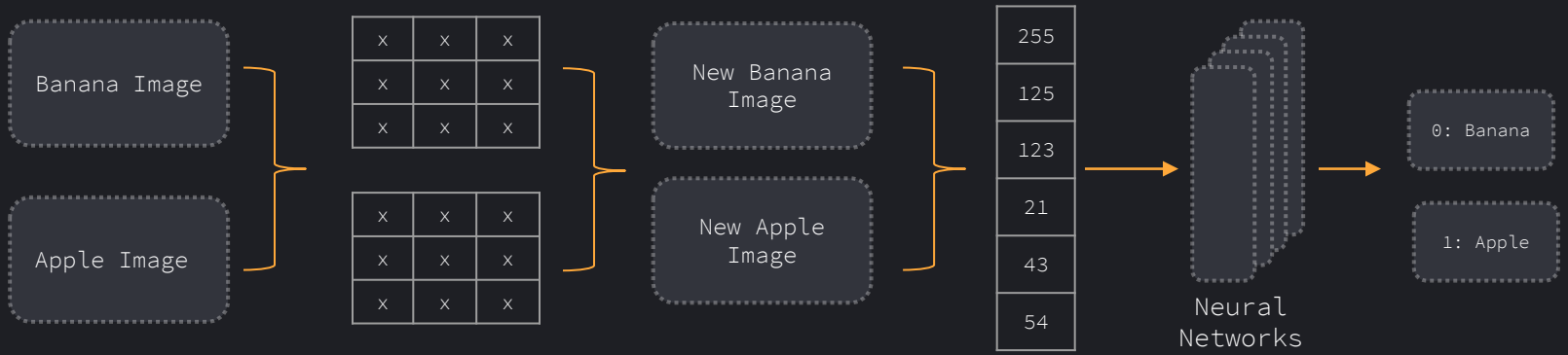
-255	-510	-765	-510	-255
-510	1275	765	1275	-510
-765	765	0	765	-765
-510	1275	765	1275	-510
-255	-510	-765	-510	-255

Filter Applied
New Image

Computer Vision

Convolutional Neural Networks (CNNs)

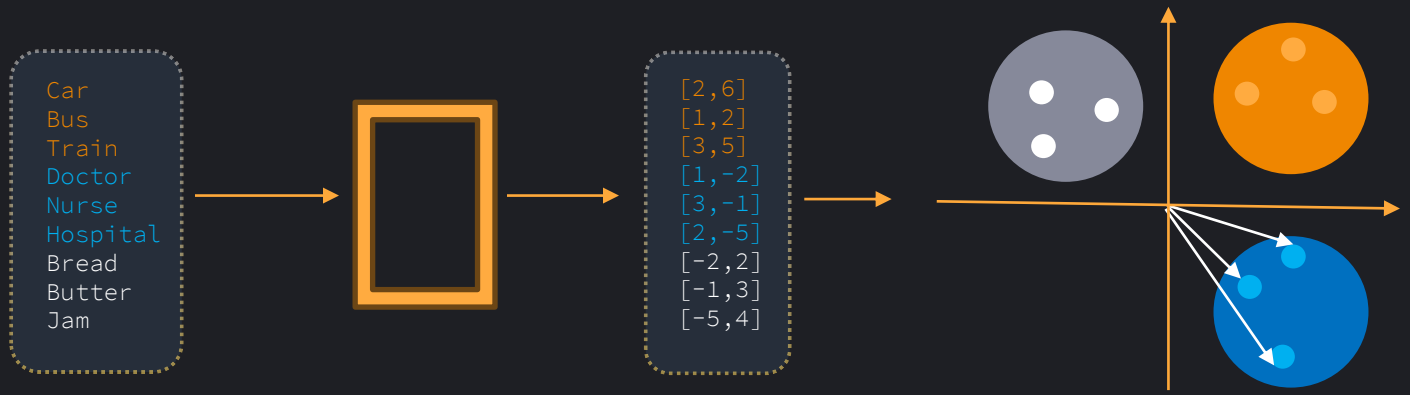
- Computer vision aims to automatically learn features by automating filter generation.
- Convolutional Neural Network (CNNs):
  - Specialized type of deep learning models.
  - Applies convolution layers, like filters, but weights of kernel learned automatically.



Computer Vision

NLP to Computer Vision

- Each word represented as vectors within Transformers.
- Model can understand meanings with numerical expressions.
- Multi-modal models combine language encoder with an image encoder.



Azure AI Vision

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Azure AI Vision

- Cloud-based image and video analysis service in Azure Cloud
- Prebuilt and customizable computer vision models
- Azure Vision Studio to test features easily
- Supports Python, Java and more
- Custom Vision for creating custom image classification and detection models

Features

Image Analysis

OCR

Face Detection

Object Detection

Object Segmentation

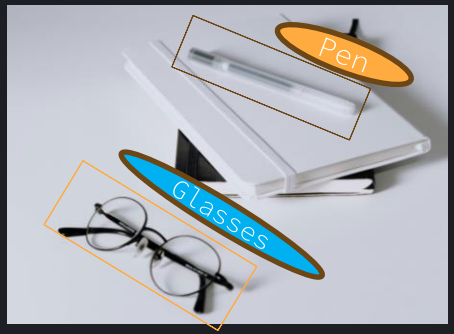
Azure AI Vision

Capabilities

Image Description

Elephant in a jungle

Object Detection



Segmentation

Optical Character Recognition (OCR)

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

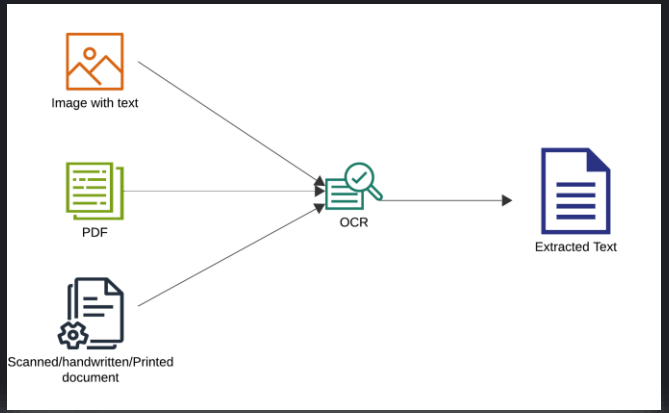
Optical Character Recognition (OCR)

AI capability to convert text in images to machine readable formats.

- Combines Computer Vision and NLP to interpret it.

Use Cases

- Digitizing Documents.
- Automating data entry.
- Scanning checks and medical records.



Azure AI Vision's OCR Engine

AI Vision's Read API

- Also known as OCR Engine, Powers OCR in Azure AI Vision.
- Extracts printed text, handwriting, layout elements, and data.
- Optimised for images with significant text or visual noise.

Access Options

- Vision Studio (no coding required)
- REST API
- SDKs: Python, C#, JavaScript

Optical Character Recognition(OCR)

Facial recognition and Azure AI Face

Understanding Facial Analysis

The use of algorithms to identify and analyse faces in images and videos

Subsets of Facial analysis

- **Face Detection**
  Identifies parts of an image that contains a human face.
- **Face Analysis**
  Extract facial features.
- **Face Recognition**
  Identify known individuals from their features.
  - Face Verification
  - Face Identification
  - Find Similar Faces
- **Face Liveness**
  Determines if face in video stream is real or fake

Azure AI Face

Pre-trained models for detecting, recognizing and analyzing faces

⇒ Wide range of facial analysis capabilities

⇒ Detecting face locations by use of a rectangular bounding box

Azure AI Face

Access services for facial analysis

- Azure AI Vision: Easy-to-use interface to access Face services
- Azure AI Video Indexer: Identify and detect faces in a video.
- SDK and REST API: Face services can be accessed liked that

Limited access feature : For responsible AI

⇒ Requires an intake form to perform certain facial analysis

Azure AI Face

Limited Facial Recognition

⇒ Retired: Emotion and gender

⇒ Limited: Age, smile, hair, makeup

Image requirements for accurate Detection

⇒ Avoid extreme lighting or angles

⇒ Minimize Occlusions

⇒ File size 6MB or less

⇒ Formats: JPEG, PNG, GIF, BMP.

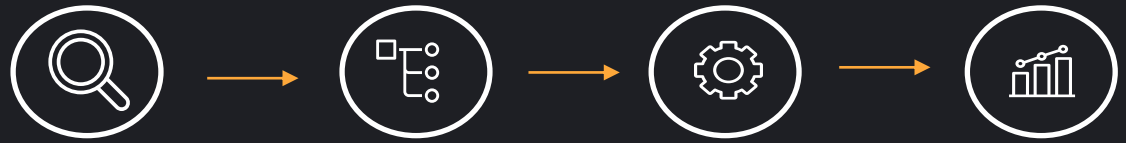
Introduction to Machine Learning (ML)

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Machine Learning (ML)

- A subset of Artificial Intelligence.
- Allows Computers to learn from data without being explicitly programmed.
- Uses the data to make predictions or decisions.

Machine Learning Process



Algorithm vs Model

- An algorithm is a procedure used to find patterns within data.
- A model is an output of an algorithm.

Machine Learning (ML)

Datasets in Machine Learning

A collection of data point used for training and testing machine learning models.

Components

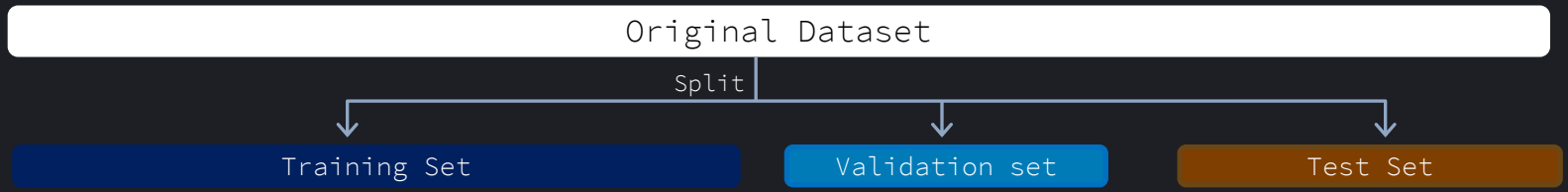
- Features: Input attributes used to make predictions.
- Values: Actual data points in the features.
- Labels: Target output.

Types of Datasets

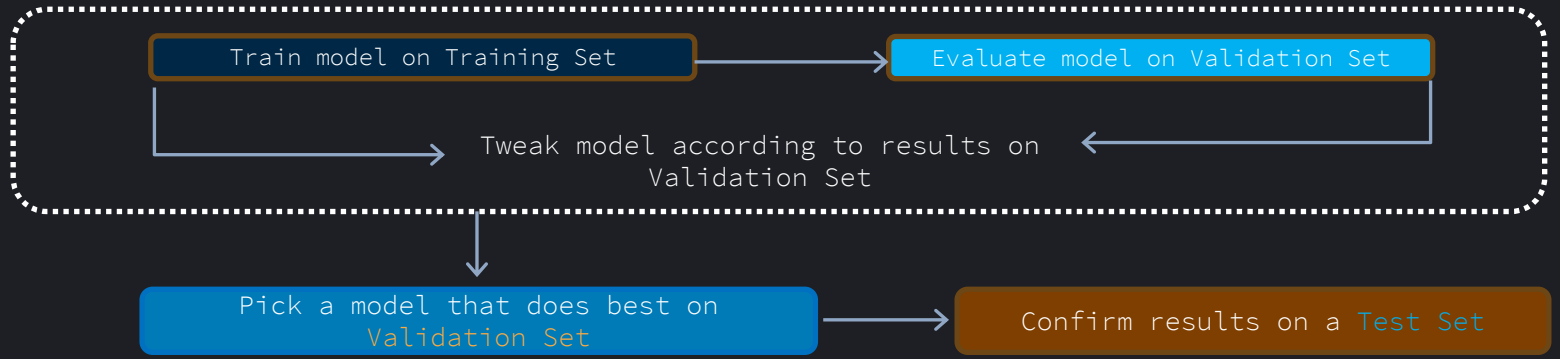
- Training set: Used to train the model.
- Validation: Used to assess the model’s predictive quality.
- Testing set: Evaluate the model’s performance.

Machine Learning (ML)

Model Training with Datasets

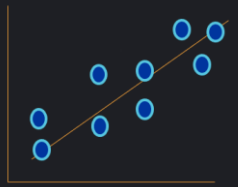


How it works



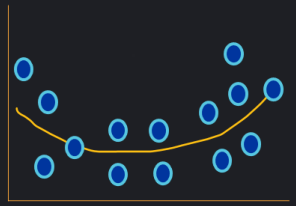
Machine Learning (ML)

Model Fit



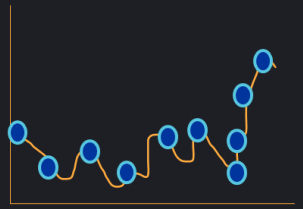
Underfitting Model

Performs poorly on training data



Balanced Model

Performs well on training and evaluation datasets



Overfitting Model

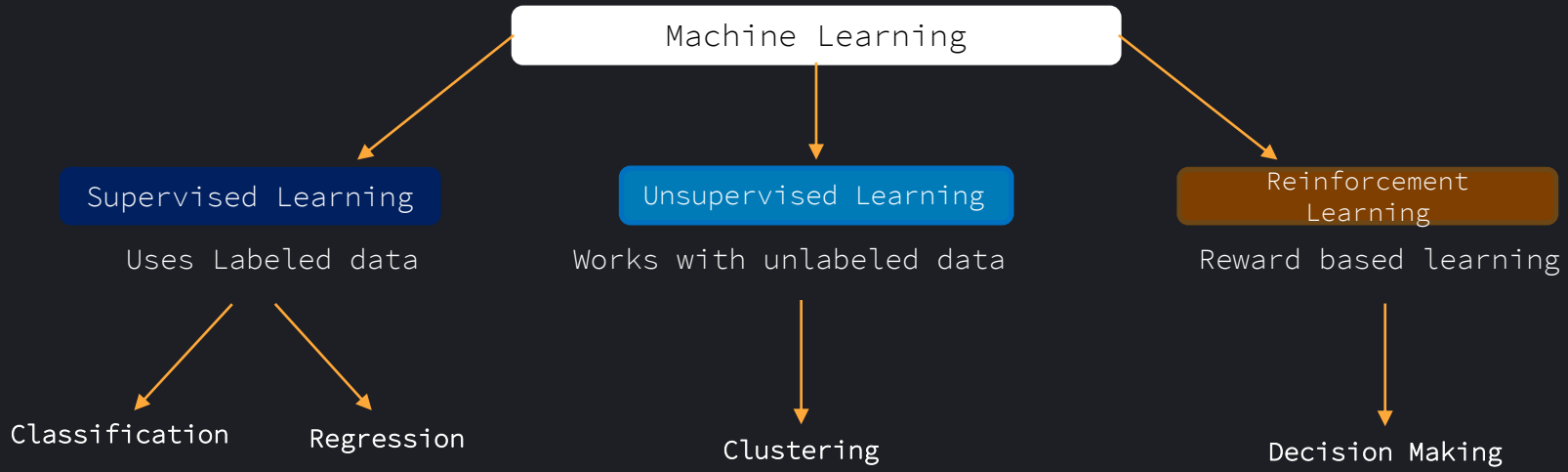
Performs well on training data but poorly on evaluation data

Machine Learning Types

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Machine Learning (ML)

Types of Machine Learning



Machine Learning (ML)

Machine Learning Types and their Use Cases

Supervised Learning

- Fraud Detection
- Customer Churn Prediction

Unsupervised Learning

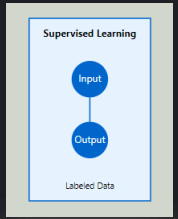
- Market Segmentation
- Anomaly Detection

Reinforcement Learning

- Game Playing
- Robotics

Machine Learning (ML)

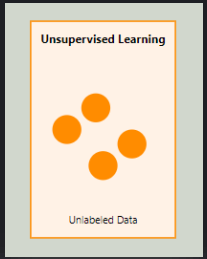
Understanding Machine Learning Models



Supervised Learning Models

- Classification Models : Categorize data into predefined classes
  ⇒ E.g.: Decision Tree Classification

- Regression Models : Predict continuous numeric values
  ⇒ E.g.: Linear Regression, Support Vector Machines



Unsupervised Learning Models

- Clustering Models : Group similar data points into clusters
  ⇒ E.g. : K-Means algorithm

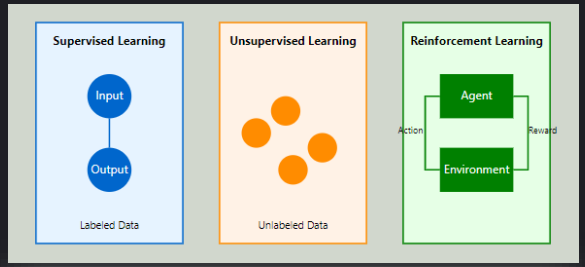
Machine Learning (ML)

Model Deciding Factors

- Data type (labeled vs. unlabeled)
- Problem type (classification vs. regression)
- Desired outcome (accuracy, interpretability)

Model Selection Guide

- Supervised Learning for predictive analytics
- Unsupervised Learning for exploratory data analysis
- Reinforcement Learning for dynamic decision-making tasks



Classifications

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Classification

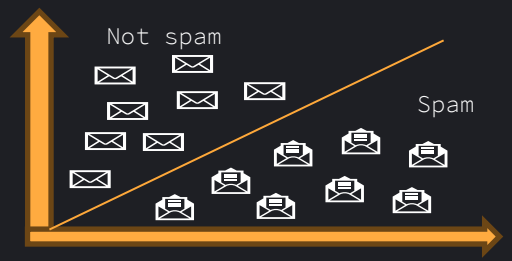
A supervised learning problem to categorize data into predefined classes/labels.

⇒ Algorithm used to solve this problem is a classifier

Types of Classification:

⇒ Binary Classification: Predicts two possible outcomes.

⇒ Multiclass Classification: Predicts one of multiple possible outcomes.



Binary Classification


Multiclass Classification

Evaluation Metrics for Classification Techniques

True Positive/Negative || False Positives/Negatives

- **True Positive (TP)**
  Correctly predicts positive class.
- **False Positive (FP) (Type I Error)**
  Incorrectly predicts positive class.
- **True Negative (TN)**
  Correctly predicts negative class.
- **False Negative (FP) (Type II Error)**
  Incorrectly predicts negative class.

Actual \ Predicted	**Positive** (Predicted Yes)	**Negative** (Predicted No)
Positive (Actual Yes)	True Positive (TP)	False Negative (FN)
Negative (Actual No)	False Positive (FP)	True Negative (TN)

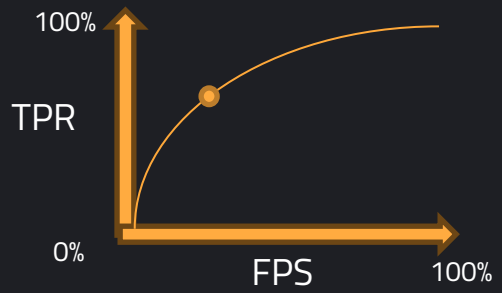
Evaluation Metrics for Classification Techniques

True Positive/Negative || False Positives/Negatives

- **True Positive (TP)**
  Correctly predicts positive class.
- **False Positive (FP) (Type I Error)**
  Incorrectly predicts positive class.
- **True Negative (TN)**
  Correctly predicts negative class.
- **False Negative (FP) (Type II Error)**
  Incorrectly predicts negative class.

Threshold: e.g. 0.8 / 80%

Actual \ Predicted	**Positive** (Predicted Yes)	**Negative** (Predicted No)
Positive (Actual Yes)	80	12
Negative (Actual No)	2	14



Evaluation Metrics for Classification Techniques

Shared Metrics for Binary and Multiclass

- **Accuracy**
  Percentage of correct predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precision**
  Ratio of true positives to total predicted positives

$$\text{Precision} = \frac{TP}{TP + FP}$$

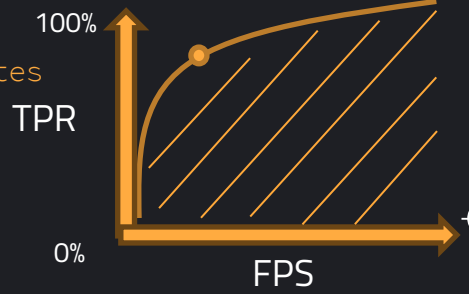
- **Recall (Sensitivity)**
  Ratio of true positives to actual positives.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1-Score**
  Harmonic mean of precision and recall

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Area Under Curve(AUC)**
  Measure trade-off between true positive and false positive rates



Regressions

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Regression

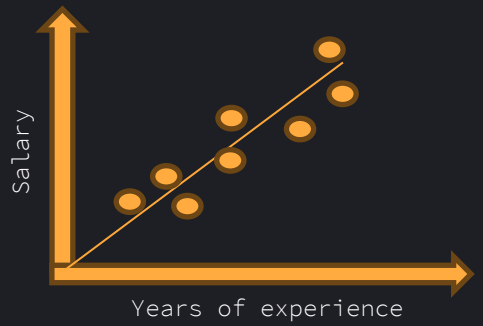
A problem with continuous numeric values as predictions based on input values.

Objective: Minimize difference between predicted and actual values

⇒ Achieved through iterative training

Use Cases:

Predicting house prices, stock market trends or sales forecasting



Regression

Evaluation Metrics

- **Mean Absolute Error(MAE)**
  Average absolute difference between predicted and actual values.
- **Mean Squared Error (MSE)**
  Amplifies larger errors by squaring them before averaging.
- **Root Mean Squared Error(RMSE)**
  Squares the errors values.
- **Coefficient of Determination (R-squared)**
  Measures how well the model explains the variance in data.

Clustering

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Clustering

A technique that groups data points into clusters based on similarities

⇒ Automatically selects the optimal number of clusters

Objective: Assign each observation to a cluster based solely on its feature

Clustering customers

Evaluation Metrics for Clustering Technique

- **Intra-Cluster Distance: Average Distance to Cluster Center**
  Measures compactness within cluster.
- **Inter-Cluster Distance: Average Distance to other Centres**
  Ratio of true positives to total predicted positives
- **Maximum Distance to Cluster Center**
  Detects outliers within clusters
- **Silhouette Score**
  Value between -1 and 1; high values indicate better-defined clusters

Azure Machine Learning

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Azure Machine Learning

Cloud Service for training, deploying, and managing Machine Learning models

Features

- ⇒ Centralized dataset storage and management
- ⇒ Azure Machine Learning Studio
- ⇒ Responsible AI tools
- ⇒ Model Deployment and Management

Azure Machine Learning

Automated Machine Learning(AutoML)

- Automates the process of building, training and deploying for ML models
- Finds the best model for your dataset
- Runs multiple training jobs with different hyperparameter combinations

How it works

- Uses algorithm and hyperparameter ranges you specify
- Chooses hyperparameter combinations that create the best model
- No code interface: training model without coding skills

Benefits

- Speeds up model development process
- High quality models

Large Language Models

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Generative AI - Models

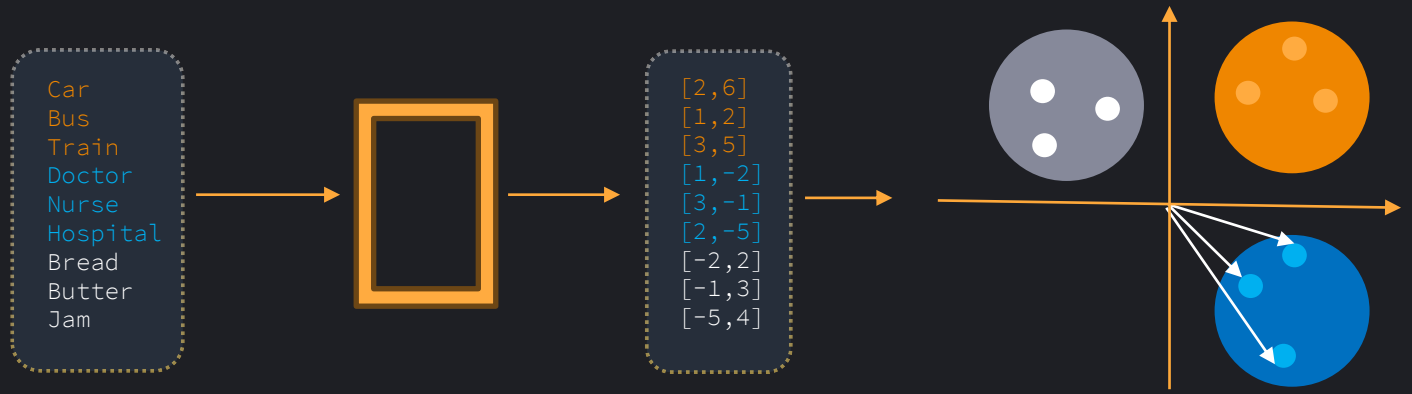
Large Language Models (LLMs)

- Understand and generate human-like texts.
- Tokens:
  - Basic units of text. “The quick brown fox jumps over the lazy dog” > “The”, “quick”, “brown”, “fox”, …
- Embeddings and Vectors:
  - Numerical representations of tokens.
  - Vectors help model understand context.
  - Vector of word “king” can represent semantic similarity of words “queen” and “monarch”.

Large Language Models

Vector Embeddings

- Each word represented as vectors (array of numbers).
- Model can understand meanings with numerical expressions.
- Vector Storing: AWS offers vector database solutions;
  - Amazon OpenSearch, pgvector extension in Amazon RDS for PostgreSQL and Amazon Kendra



Essentials of Prompt Engineering

Essentials of Prompt Engineering

Prompt:
Is the input or query provided to a language model to generate a response.

Can be

- Question
- Statement
- Set of Instructions

Prompt Engineering: is the process of designing and preparing prompts.

Essentials of Prompt Engineering

Elements of a Prompt

Instruction:

Summarize the given article in no more than 50 words.

Context:

The article discusses the impact of artificial intelligence on various industries.

Input Data:

"AI is revolutionizing healthcare, finance, and transportation by enhancing diagnostics, improving fraud detection, and enabling autonomous vehicles. These advancements present both opportunities and challenges for society."

Output Indicator:

Provide a concise summary.

Example Output:

AI is transforming healthcare, finance, and transportation, improving diagnostics, fraud detection, and enabling autonomous vehicles, presenting new opportunities and challenges.

Essentials of Prompt Engineering

Negative Prompting

- Used to guide the model away from producing certain types of content or exhibiting specific behaviors.
- Used to prevent the model from producing

Hate speech	Explicit content	Biased language

- Helps steer the output towards more appropriate content.

Prompt Engineering Techniques

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Prompt Engineering Techniques

- Prompt Techniques are ways to give instructions to an AI model so it knows how to respond or do a task.
- Most comment prompt techniques include

Zero-shot

Few-shot

Chain-of-Thought

Prompt Engineering Techniques

Zero-Shot Prompting

- Prompt does not include any examples or demonstrations.
- Effective on a larger and more capable model.

Prompt Engineering Techniques

Few-Shot Prompting

- Uses contextual examples
- Providing more examples can help the model better understand the task
- Too many examples might introduce noise or confusion

Prompt Engineering Techniques

Chain-of-Thought Prompting

- Divides complex reasoning tasks into smaller, intermediary steps.
- Can be employed using either.

  Zero-shot Few-shot

- To initiate the chain-of-thought reasoning process in a machine learning model, you can use the phrase "Think step by step".

Prompt Engineering Techniques

Using Zero-Shot

Prompt: *John has 3 apples. He buys 5 more apples and then eats 2. How many apples does he have now? Think step by step.*

Output:

John starts with 3 apples.

He buys 5 more apples: 3 + 5 = 8.

He eats 2 apples: 8 - 2 = 6.

John has 6 apples now.

Prompt Engineering Techniques

Using Few-Shot / One-Shot

Prompt:

Example #1:
Question: If there are 3 baskets and each basket contains 2 apples, how many apples are there in total?

Thought: There are 3 baskets and each basket contains 2 apples. To find the total number of apples, we multiply the number of baskets by the number of apples in each basket.
Calculation: 3 baskets * 2 apples/basket = 6 apples
Answer: 6 apples.

Now solve this in the same way:
Question: If there are 5 boxes and each box contains 4 books, how many books are there in total?

Answer:

Thought: There are 5 boxes and each box contains 4 books. To find the total number of books, we multiply the number of boxes by the number of books in each box.
Calculation: 5 boxes * 4 books/box = 20
Answer: 20 books

Azure AI Foundry Overview

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Azure AI Foundry

- Unified platform to simplify development & deployment of AI Solutions.
- Gathers pre-trained models, tools and needed infrastructure.

Integrated Ecosystem

Integrates with platforms like GitHub, Visual Studio and other Microsoft products.

Scalable

Scales effortlessly to meet your needs.

Access Different AI Models

Provides latest AI technology within Microsoft

Azure AI Foundry

Key Features

Model Catalog	⇒ Centralized AI Models repository. ⇒ Pre-Trained Models. ⇒ Third-Party Models. ⇒ Custom Models.
Customization & Fine-Tuning	⇒ Fine-Tuning: Modify pre-trained models for your use-case. ⇒ Grounding: Integrate your own dataset.
Development Environment	⇒ GitHub Copilot Studio. ⇒ Azure AI SDK. ⇒ Azure DevOps.
AI Management Center	⇒ Resource Management. ⇒ Monitoring. ⇒ Quotas and Governance.

Azure AI Foundry

- Provides tools and frameworks for Responsible AI.
  - Explainability
  - Fairness
  - Compliance

Use Cases

Generative AI Apps

- Automated report writings.
- Marketing.
- Image, video, sound creation.

Customer Support

- Provide accurate answers to customers
- Automate tasks.
- Personalized support.

Predictive Maintenance

- Real-time monitoring of equipment health.
- Predicting failures.

Microsoft Copilot Overview

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Microsoft Copilot

- AI-powered assistant integrated into Microsoft products.
- Context-aware assistance, automate tasks, generate insights.
- Combines LLMs with Microsoft Graph.
- Seamless integration with Microsoft products:
  - Word
  - Excel
  - PowerPoint
  - Teams
  - Outlook

Microsoft Copilot

Capabilities

- Drafting and Content Generation
  - In Word or Outlook, write initial drafts, generate contents.
- Summarization
  - Summarize long threads in Outlook or meeting transcripts in Teams.
- Insights & Data Analysis
  - Identify patterns in Excel sheets, generate tables and analysis.
- Presentation Enhancement
  - Turn Word documents into slides with PowerPoint, suggest visuals…
- Workflow Automation
  - Automate repetitive tasks, scheduling meetings

Microsoft Copilot

Copilot Studio

- Customize and create conversational AI with low-code environment.
- Designed for technically proficient users or developers.
- Define custom prompts, workflows and tasks to create customized AI Agent.

Low-Code Development

Fully Managed SaaS Solution

Integration with Microsoft Products

Azure AI Language

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Azure AI Language

- Combines multiple NLP capabilities to process and analyze text data.
- Organizations can build apps that can interact with human language.
- Fully managed service.
- Pay as you go.

Text Analytics

⇒ Sentiment analysis, extract key phrases, obtain people location names.

⇒ Detect Personally Identifiable Information (PII)

Azure AI Language

Conversational Language Understanding (CLU)

⇒ Build conversational AI applications, like chatbots.

⇒ Create and train custom models.

Question Answering

⇒ Create knowledge bases or FAQs.
⇒ Retrieve answers directly from structured or unstructured data.

Custom Text Classification

Custom Named Entity Recognition

Azure AI Language

- Integrates with other Azure services
  - Azure Bot Services
  - Azure Cognitive Search
  - Power Automate

- Customization and deployment
  - Prebuilt Models
  - Custom Models
  - Cloud-based deployment
  - Edge deployment

Conversational Language Understanding

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Conversational Language Understanding

Allows for building models that understand and act on natural language

Core Concepts

- Utterances: Represents user input or phrase that must be interpreted.
- Entities: Specific Items within an utterance.
- Intents: Desired outcome of the user input.

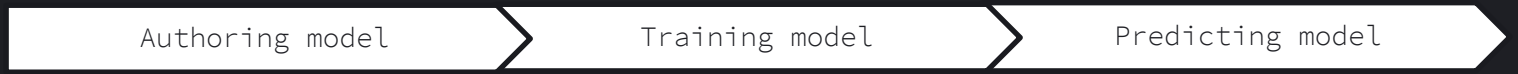
Common Use Cases

- Chatbots
- Voice Assistants
- Enterprise Bots

Conversational Language Understanding

Requirement:
Azure AI Language resource

Steps to building a CLU Model:



- Define Schema
  Identify intents & entities
- Label Data
  Label utterances with intents & entities
- Use Labeled data
  To train the model
- Evaluate Performance
  Test model with new data
- Deploy the Model
  To make real-time predictions

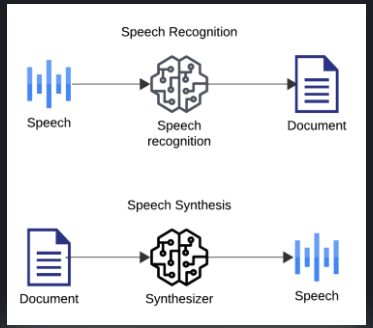
Azure AI Speech

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Speech Recognition and Synthesis

What is Azure AI Speech?

Advanced tools for transcription, synthesis, translation, and recognition



Speech-to-Text

Converts speech into text; real-time or batch.

- ⇒ Creating transcripts
- ⇒ Generating captions for videos

Text-to-Speech

Neural voices for audiobooks, chatbots, and accessibility.

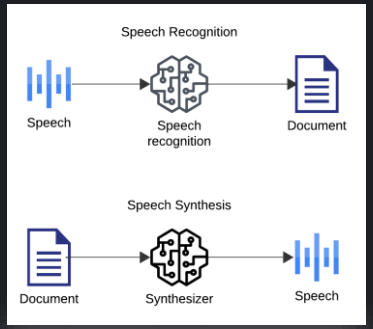
Speech Translation

Multilingual translation in real-time.

Speech Recognition and Synthesis

Speaker Recognition:
Identifies and authenticates speakers.

Pronunciation Assessment:
Improves fluency and accuracy.



Speech Recognition and Synthesis

Tools and Interface

Speech Studio

No-code platform for testing and exploring.

Speech SDK & REST APIs

Developer tools for integration.

Speech CLI

Command-line interface for managing and testing speech services

Speech Recognition and Synthesis

Deployment options

Cloud

Scalable and globally accessible

Edge containers

Scenarios requiring compliance, low latency, or offline capabilities.

Supports sovereign clouds

For regulated environments such as Azure Government

Speech Recognition and Synthesis

Responsible AI

Ethical AI principles.

Ensures transparency and privacy.

Azure AI Translator

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Azure AI Translator

Document translation feature using advanced AI models in Azure Cloud

⇒ Uses Neural Machine Translation model

Translation Operations:

Asynchronous Batch Document Analysis

Synchronous Document Translation.

Capabilities

- Custom Translation Models
- Supports live speech or text translation
- Handles translation across hundreds of languages

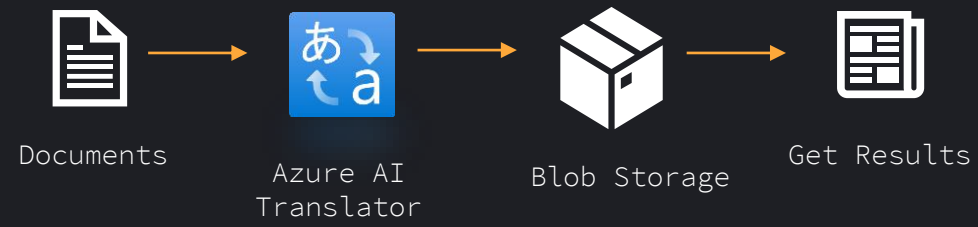
Azure AI Translator

Translation Operations sample architecture

Synchronous

For single-page documents

Asynchronous



For multi-page documents

Document Intelligence

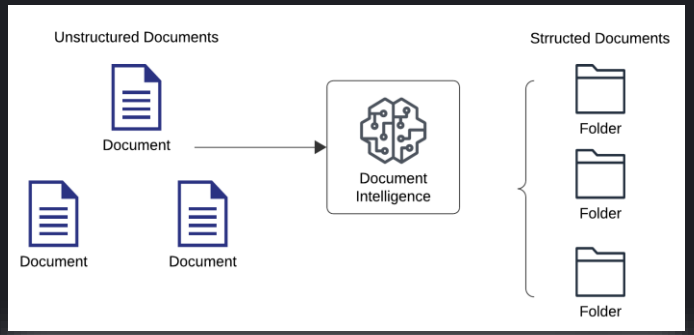
003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Document Intelligence

Uses AI to manage and process large amounts of data found in documents and forms.

Features

- Automates data extraction from scanned documents.
- Analyse documents in different formats.
- Reduces manual errors.
- Auto-categorization.



Azure AI Document Intelligence

Automates data processing by extracting and structuring text from documents

⇒ Uses OCR and AI

Capabilities

- Prebuilt Models: Ready-to-use templates for invoices, receipts, and more.
- Custom Models: Create models using your own labelled data.
- Streamlines documents: Returns structured data.

Real-World Applications

- Structuring healthcare records.
- Automating receipt processing.
- Simplifying contract reviews.

Knowledge Mining & AI Search

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Knowledge Mining and AI Search

Features

Data Extraction

Extracts valuable information from diverse sources

Semantic Search

Enhances search accuracy by understanding user intent

Data Organization

Creates structured repositories

Knowledge Mining

Capabilities

Ingest → Enrich → Explore

- Sources data
  From files, databases, or media formats

- Optical Character Recognition (OCR)
  Extracts text from documents
- Sentiment Analysis
  Assess content
- Entity Recognition
  Identifies key information
- Translation
  Makes it accessible to other languages

- Indexes data
  Advanced querying with Azure AI Search

Knowledge Mining and AI Search

Use Cases

Content Research

Simplifies information discovery

Auditing and Risk Management

Analyzes documents to improve compliance and mitigate risks

Customer Support

Quickly retrieves relevant customer information

Integrating Document Intelligence and Knowledge Mining

- Document Intelligence extracts data.
- Knowledge Mining organizes and makes it searchable.

Knowledge Mining

Workflow

- Collects and processes
- Adds metadata and insights
- Creates searchable indexes

Knowledge Mining in Azure

Azure AI Search

- Private, enterprise search solution in the Azure Cloud
- Build searchable indexes for internal or public use.
- Leverages AI capabilities for data extraction and enrichment

Core Features

- Processes structured, semi-structures and unstructured data.
- Platform as a Service(PaaS): Managed by Microsoft.

Knowledge Mining in Azure

Deployment options

Cloud

Scalable and globally accessible

Edge containers

On-premise deployment in containers requiring compliance and low latency

Responsible AI Practices

003-1040559 1250 003-77156.8 1760 0009-14563.7 73273

Responsible AI

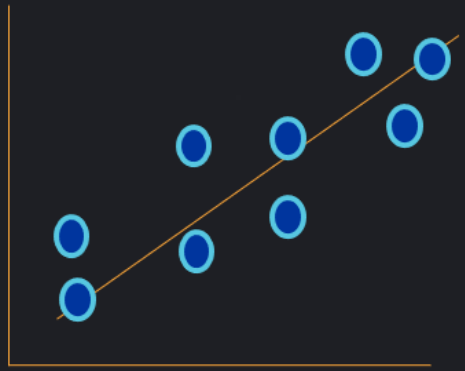
- **What is Responsible AI:**
  The practice of designing, developing, and deploying AI systems that adhere to ethical standards and principles

- **Types of AI:**
  Traditional AI
  Generative AI

- **Generative AI Business value:**
  - Creativity
  - Productivity
  - Connectivity

Responsible AI

Challenges of Responsible AI in Traditional AI and Generative AI



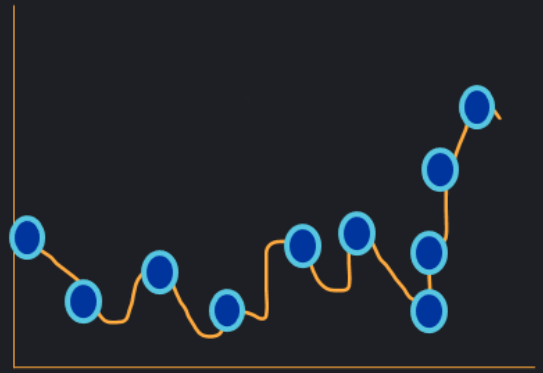
Underfitted Model example

- **Bias:**
  Model overlooks important features in the dataset

- **High Bias:**
  High bias indicates underfitting

Responsible AI

Challenges of Responsible AI in Traditional AI and Generative AI

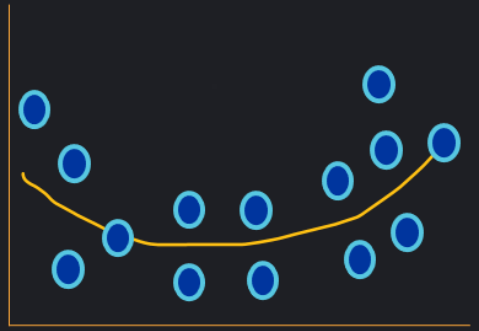


Overfitted Model example

- **Variance:**
  Model becomes sensitive to noise or fluctuations in the training data
- **High Variance:**
  High variance leads to overfitting

Responsible AI

Challenges of Responsible AI in Traditional AI and Generative AI



Balanced Model example

- **Bias-Variance Trade-Off:**
  Balance between underfitting (high bias) and overfitting (high variance).
- **Balanced Model:**
  Ensures optimal model performance on both training and new data.

Responsible AI

Strategies to Address Bias and Variance

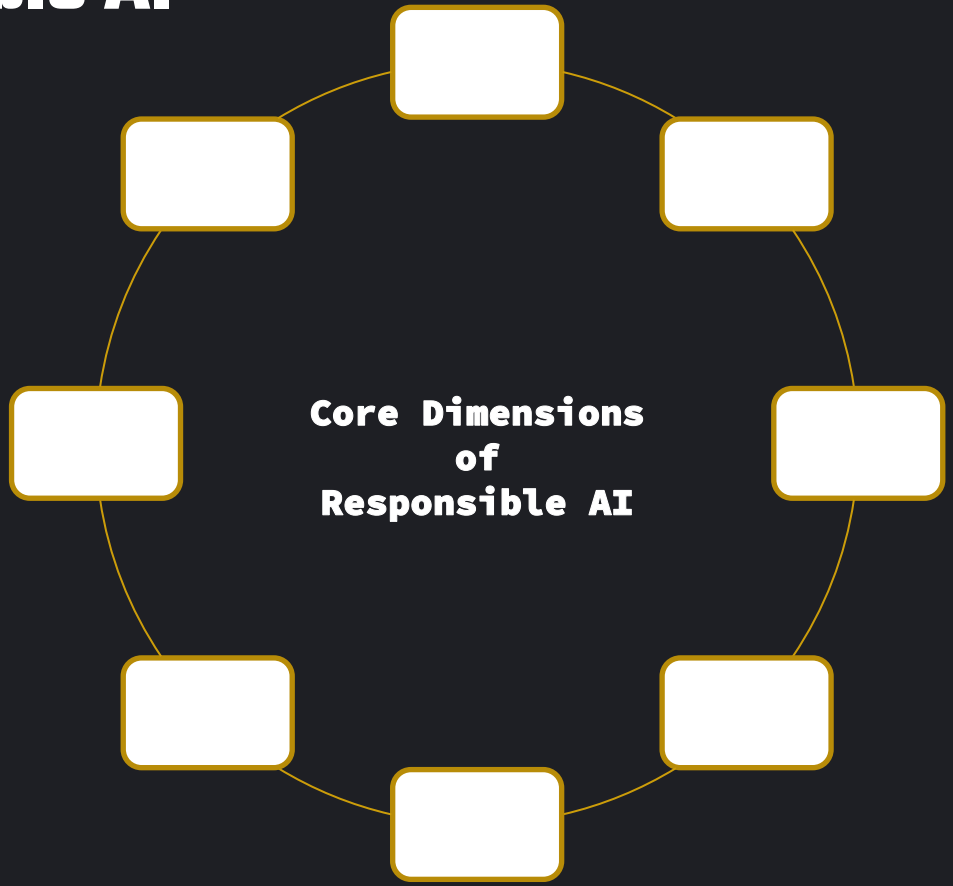
- **Cross-Validation**
  - Use multiple training and validation sets
- **Increasing Data**
  - Adding more data helps the model learn better
- **Regularization**
  - Penalize extreme model parameters
- **Simpler Models**
  - Simpler models can help avoid overfitting
- **Dimension Reduction (PCA)**
  - Reduces the number of input variables
- **Early Stopping**
  - Prevents the model from memorizing the training data

Responsible AI

Unique Challenges in Generative AI

- **Toxicity**
  - Can produce harmful or offensive content
- **Hallucinations**
  - Might generate inaccurate information
- **Intellectual Property**
  - Risks of replicating proprietary content without authorization
- **Plagiarism and Cheating**
  - Might be used to plagiarize or cheat
- **Disruption of the Nature of Work**
  - Significantly change job roles and work processes

Responsible AI



Responsible AI – Model Trade-Offs

- **Interpretability**
  How easily a human can understand the decisions made by an AI model

- **Performance**
  The accuracy and effectiveness of the model

VS

- **Safety**
  Ensure that AI systems avoid harmful outcomes and protect sensitive information

- **Transparency**
  Reveal how the AI system works to build trust and accountability

Responsible AI – Model Trade-Offs

- **Controllability**
  The ability to influence the model's predictions and behavior

- **Complexity**
  How intricate the model is

VS

- **Bias**
  Error due to overly simplistic models that miss relevant relations (underfitting)

- **Variance**
  Error due to models that are too sensitive to training data (overfitting)

Responsible AI

Principles of Human-Centered Design for Explainable AI

✔ Design for amplified decision-making

- Design systems to assist users in critical decisions.
- Key Aspects: Clarity, simplicity, usability, reflexivity, accountability.

✔ Design for unbiased decision-making

- Ensure decision-making processes are free from unfair biases.
- Steps: Identify biases, Transparent process and Training.

Responsible AI

Principles of Human-Centered Design for Explainable AI

✔ Design for human and AI learning

- Develop tools that enhance both human and AI learning.
- Key Aspects: Cognitive Apprenticeship, Personalization and User-Centered

✔ Reinforcement Learning from Human Feedback (RLHF)

- Use human feedback to refine and improve AI models.
- Tools: Amazon SageMaker Ground Truth.
- Amazon SageMaker Ground Truth : Incorporates human feedback to enhance model accuracy and relevance.