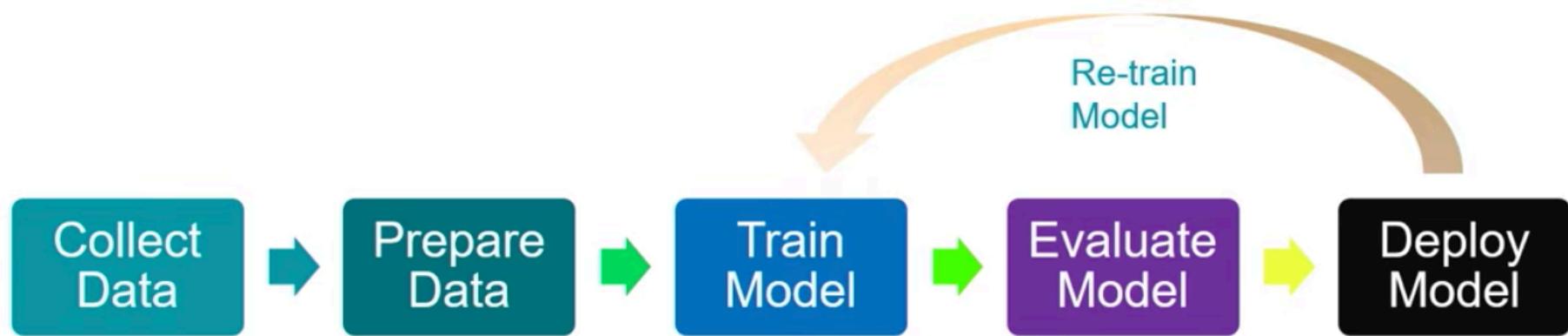
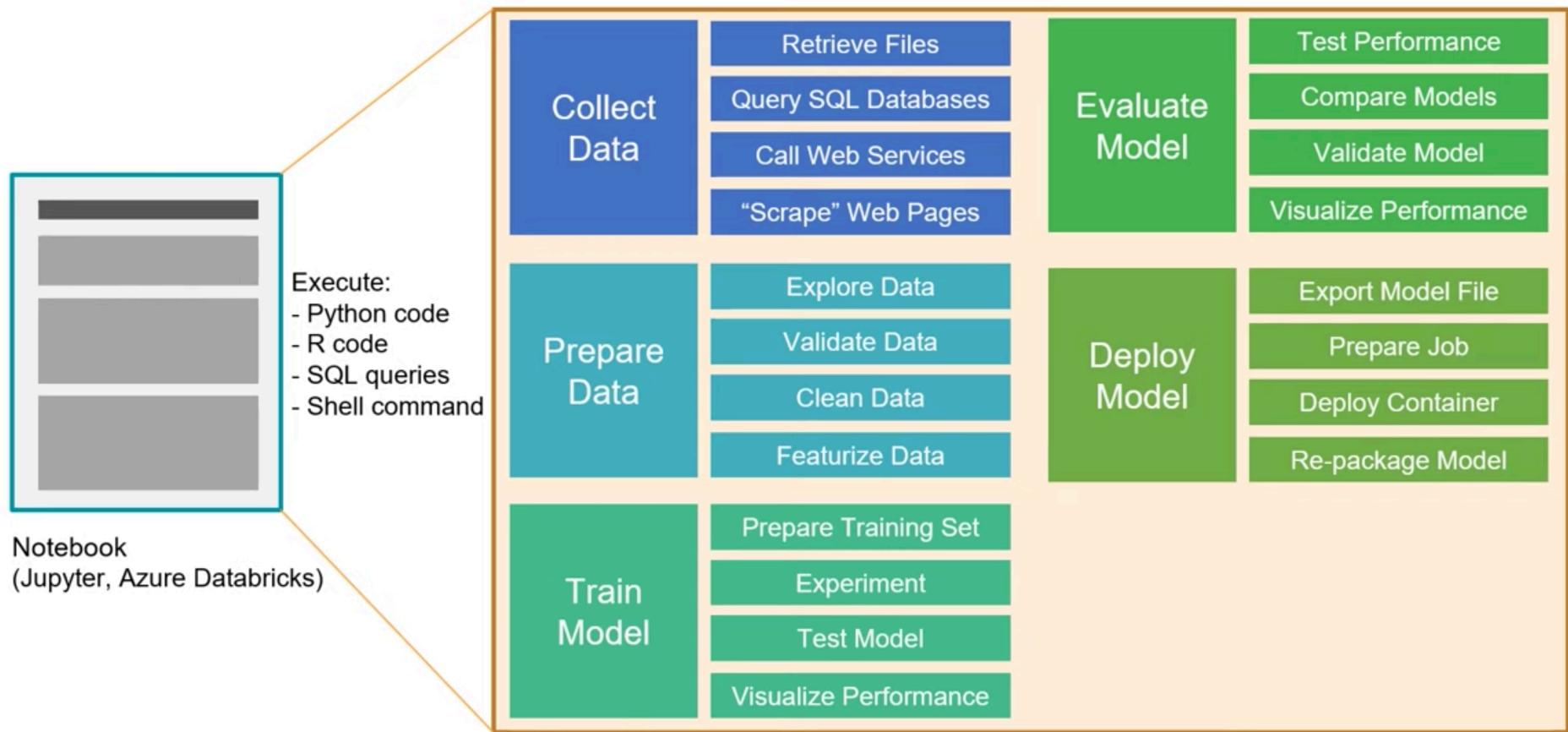
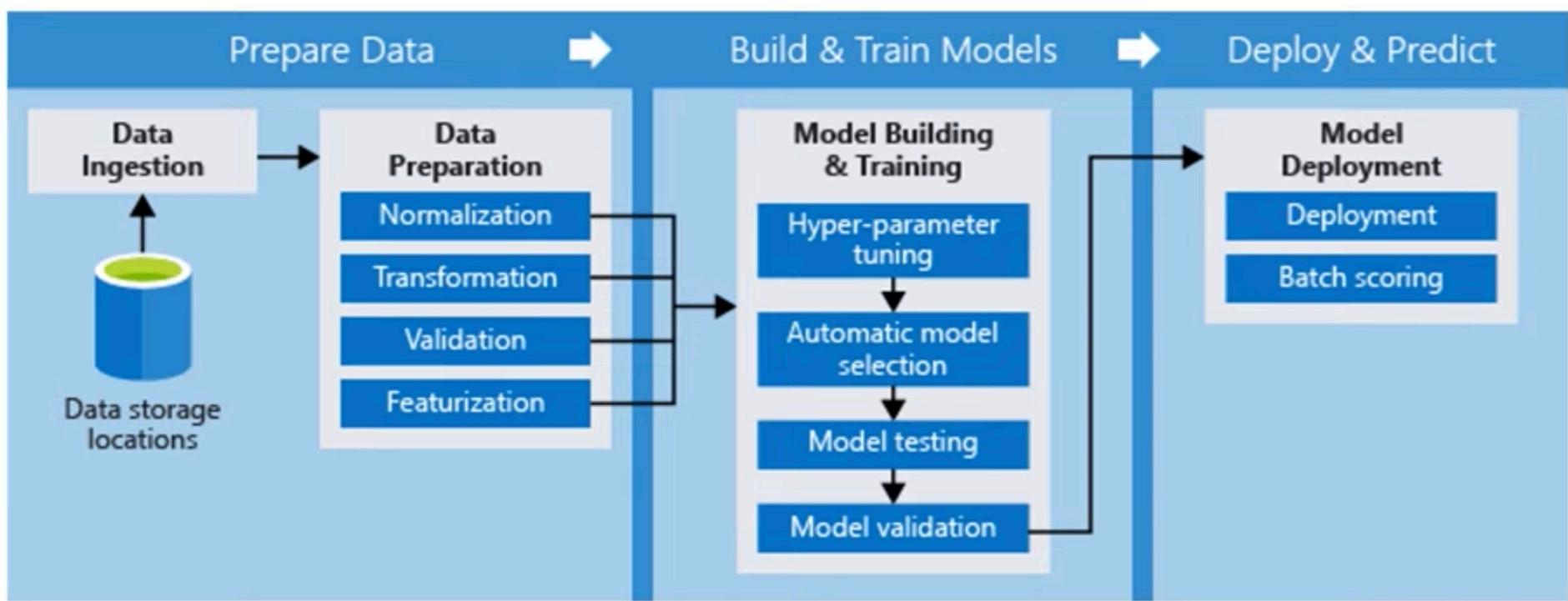


The Data Science Process



The Notebook Paradigm



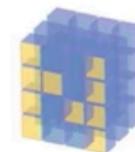


Libraries for Machine Learning

Core Framework and Tools



pandas
 $y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



NumPy



jupyter

Machine Learning & Deep Learning



Keras

PyTorch

Visualization



plotly

seaborn

matplotlib



Bokeh

Approaches to Machine Learning

Supervised Learning

Learns from data that contains both the inputs and expected outputs

Classification

Outputs are categorical

Regression

Outputs are numerical

Similarity Learning

Learns from examples using a similarity function

Feature Learning

Features are learned using labeled data

Anomaly Detection

Learns from data labeled as normal/abnormal

Unsupervised Learning

Learns from data that contains only inputs and finds hidden structures in data

Clustering

Assigns entities to clusters or groups

Feature Learning

Features are learned from unlabeled data

Anomaly Detection

Learns from unlabeled data assuming most entities are normal

Reinforcement Learning

Learns how an agent should take actions in an environment to maximize a reward function

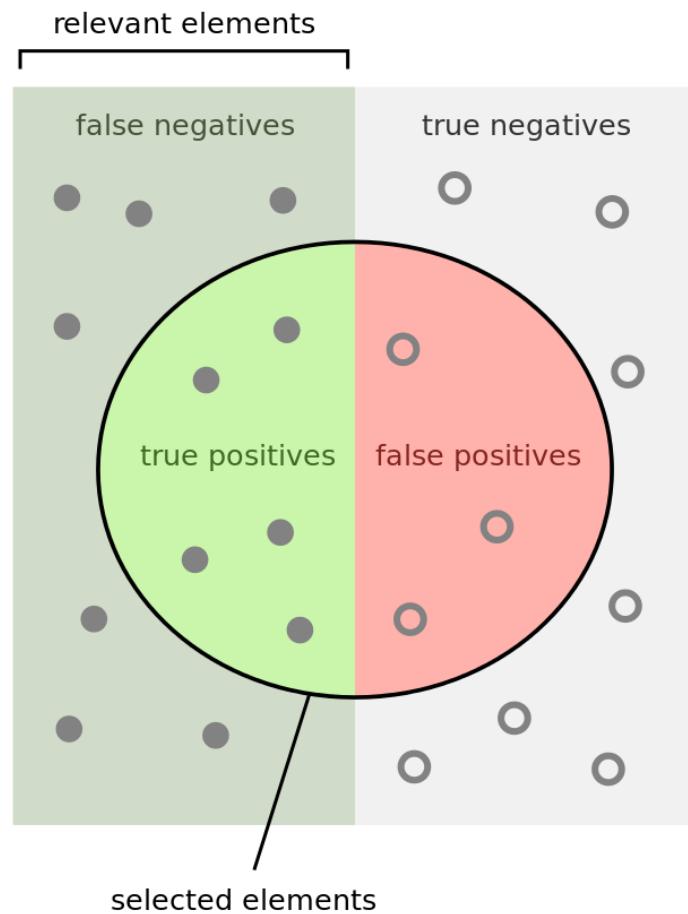
Markov Decision Process

Does not assume knowledge of an exact mathematical model

Specialized case	Approach
Similarity Learning	Supervised
Text Classification	Supervised (classification)
Feature Learning	Supervised (classification) Unsupervised (clustering)
Anomaly Detection	Supervised (classification) Unsupervised (clustering)
Forecasting	Supervised

Typical Model Deployment

- Get the model file (any format)
- Create a scoring script (.py)
- Optionally create a schema file describing the web service input (.json)
- Create a real-time scoring web service
- Call the web service from your applications
- Repeat the process each time you re-train the model



How many selected items are relevant?

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

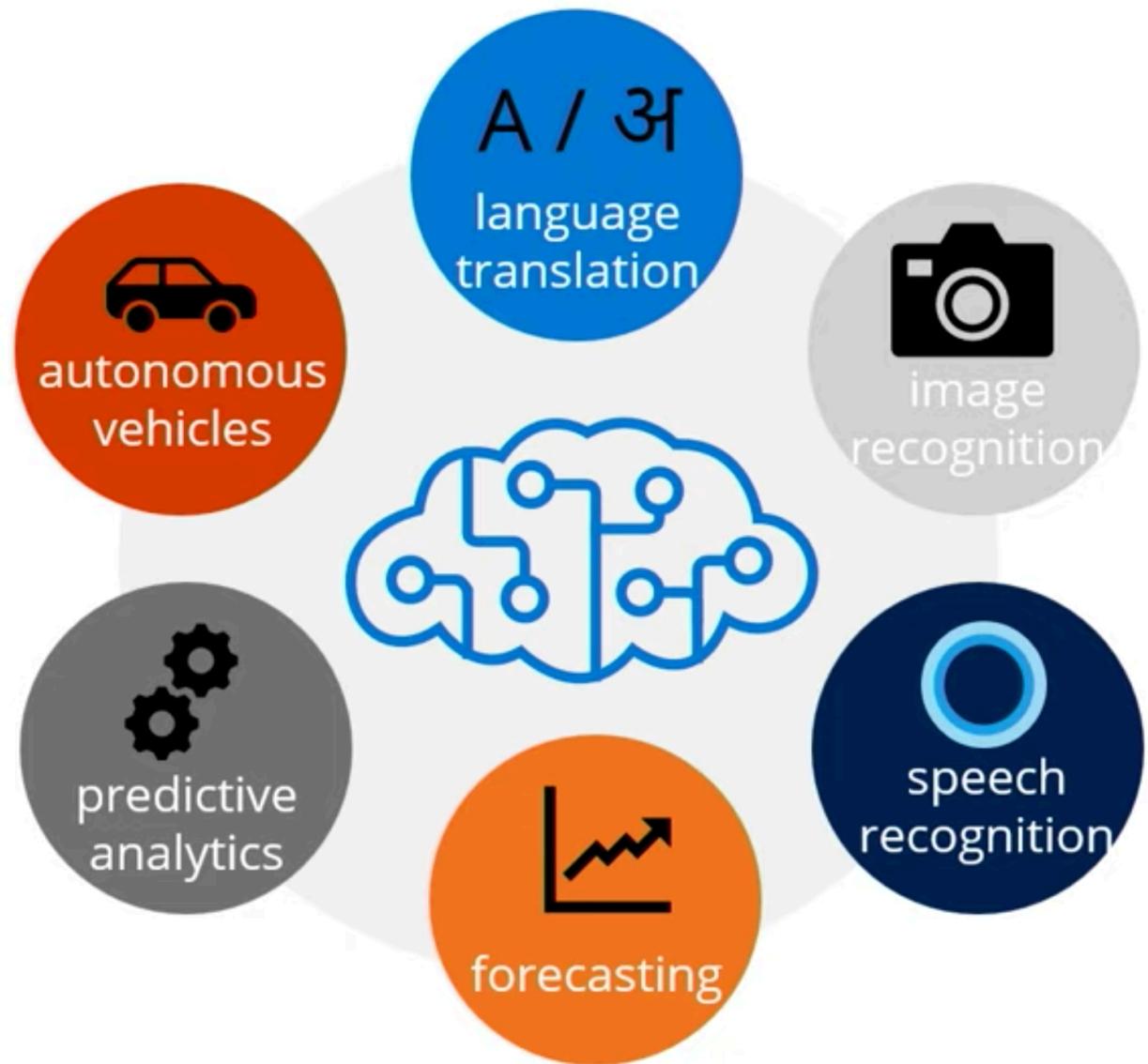
How many relevant items are selected?

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Azure Machine Learning service

- Azure Machine Learning service provides a comprehensive environment to implement model training processes, giving you a centralized place to work with all the artifacts involved in the process.

Where is Deep Learning applied?



AI Services

VISION	SPEECH	TEXT	CHATBOTS	PERSONALIZATION	FORECASTING
 AMAZON RECOGNITION	 AMAZON POLLY	 AMAZON TRANSLATE	 AMAZON LEX	 AMAZON PERSONALIZE	 AMAZON FORECAST
				FRAUD	DEVELOPMENT
	 AMAZON TRANSCRIBE + MEDICAL	 AMAZON COMPREHEND + MEDICAL		 AMAZON FRAUD DETECTOR	 AMAZON CODEGURU
		 AMAZON TTEXTACT		CONTACT CENTERS	SEARCH
			 CONTACT LENS		 AMAZON KENDRA

ML Services



ML Frameworks + Infrastructure

FRAMEWORKS	INTERFACES	INFRASTRUCTURE						
 TensorFlow  mxnet  PYTORCH	 GLUON  Keras  PyTorch  DeepGraphLibrary	 EC2 P3 & P3DN  EC2 G4 EC2 C5  FPGAs  DL CONTAINERS & AMIs	 ELASTIC CONTAINER SERVICE  ELASTIC KUBERNETES SERVICE  AWS IOT GREENGASS	 ELASTIC INFERENCE  INFERENTIA				

Deploying a trained model – step-by-step

- Get the model file (any format)
- Create a scoring script (.py)
- Optionally create a schema file describing the web service input (.json)
- Create a real-time scoring web service
- Call the web service from your applications
- Repeat the process each time you re-train the model

