

# Foundation of AI and ML (4351601) - Winter 2023 Solution

Milav Dabgar

December 04, 2023

## Question 1 [a marks]

3 Define the following terms: (1) Artificial Intelligence (2) Expert System.

### Solution

Term	Definition
<b>Artificial Intelligence</b>	AI is a branch of computer science that creates machines capable of performing tasks that typically require human intelligence, such as learning, reasoning, and problem-solving.
<b>Expert System</b>	An expert system is a computer program that uses knowledge and inference rules to solve problems that normally require human expertise in a specific domain.

- **AI characteristics:** Learning, reasoning, perception
- **Expert system components:** Knowledge base, inference engine

### Mnemonic

"AI Learns, Expert Advises"

## Question 1 [b marks]

4 Compare Biological Neural Network and Artificial Neural Network.

### Solution

Aspect	Biological Neural Network	Artificial Neural Network
<b>Processing</b>	Parallel processing	Sequential/parallel processing
<b>Speed</b>	Slow (milliseconds)	Fast (nanoseconds)
<b>Learning</b>	Continuous learning	Batch/online learning
<b>Storage</b>	Distributed storage	Centralized storage

- **Biological:** Complex, fault-tolerant, self-repairing
- **Artificial:** Simple, precise, programmable

### Mnemonic

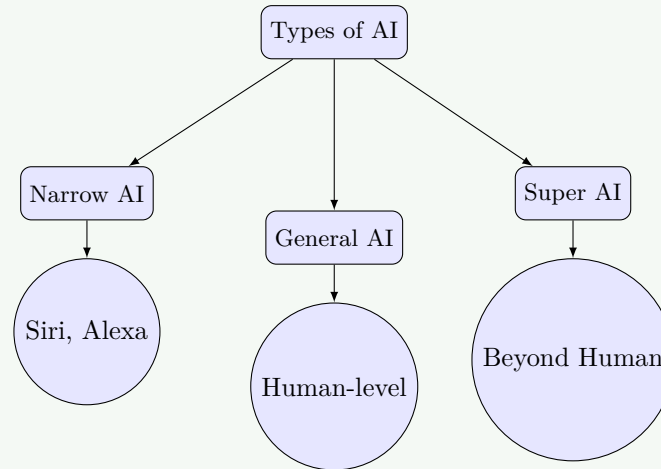
"Bio is Complex, AI is Simple"

## Question 1 [c marks]

7 Explain types of AI with its applications.

## Solution

Type of AI	Description	Applications
<b>Narrow AI</b>	AI designed for specific tasks	Voice assistants, recommendation systems
<b>General AI</b>	AI with human-level intelligence	Not yet achieved
<b>Super AI</b>	AI exceeding human intelligence	Theoretical concept



**Figure 1.** Types of Artificial Intelligence

- **Current focus:** Narrow AI dominates today's applications
- **Future goal:** Achieving General AI safely

#### Mnemonic

"Narrow Now, General Goal, Super Scary"

## Question 1 [c marks]

7 Explain AI ethics and limitations.

## Solution

Ethics Aspect	Description
<b>Privacy</b>	Protecting personal data and user information
<b>Bias</b>	Ensuring fairness across different groups
<b>Transparency</b>	Making AI decisions explainable
<b>Accountability</b>	Determining responsibility for AI actions

#### Limitations:

- **Data dependency:** Requires large, quality datasets
- **Computational power:** Needs significant processing resources
- **Lack of creativity:** Cannot truly create original concepts

#### Mnemonic

"Privacy, Bias, Transparency, Accountability"

## Question 2 [a marks]

3 Define the following terms: (1) Well posed Learning Problem (2) Machine Learning.

### Solution

Term	Definition
<b>Well posed Learning Problem</b>	A learning problem with clearly defined task (T), performance measure (P), and experience (E) where performance improves with experience.
<b>Machine Learning</b>	A subset of AI that enables computers to learn and improve automatically from experience without being explicitly programmed.

- **Well posed formula:**  $T + P + E = \text{Learning}$
- **ML advantage:** Automatic improvement from data

### Mnemonic

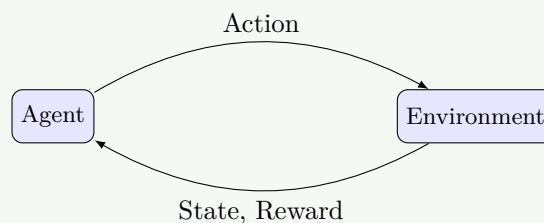
"Task, Performance, Experience"

## Question 2 [b marks]

4 Explain Reinforcement Learning along with terms used in it.

### Solution

Term	Description
<b>Agent</b>	The learner or decision maker
<b>Environment</b>	The world in which agent operates
<b>Action</b>	What agent can do in each state
<b>State</b>	Current situation of the agent
<b>Reward</b>	Feedback from environment



**Figure 2.** Reinforcement Learning Cycle

- **Learning process:** Trial and error approach
- **Goal:** Maximize cumulative reward

### Mnemonic

"Agent Acts, Environment States and Rewards"

## Question 2 [c marks]

7 Compare Supervised, Unsupervised and Reinforcement Learning.

## Solution

Aspect	Supervised	Unsupervised	Reinforcement
<b>Data</b>	Labeled data	Unlabeled data	Interactive data
<b>Goal</b>	Predict output	Find patterns	Maximize reward
<b>Feedback</b>	Immediate	None	Delayed
<b>Examples</b>	Classification	Clustering	Game playing

- **Supervised:** Teacher-guided learning
- **Unsupervised:** Self-discovery learning
- **Reinforcement:** Trial-and-error learning

## Mnemonic

"Supervised has Teacher, Unsupervised Discovers, Reinforcement Tries"

## Question 2 [a marks]

3 Write Key features of Reinforcement Learning.

## Solution

Feature	Description
<b>Trial and Error</b>	Learning through experimentation
<b>Delayed Reward</b>	Feedback comes after actions
<b>Sequential Decision</b>	Actions affect future states

- **No supervisor:** Agent learns independently
- **Exploration vs Exploitation:** Balance between trying new actions and using known good actions

## Mnemonic

"Try, Delay, Sequence"

## Question 2 [b marks]

4 Explain Types of Reinforcement learning.

## Solution

Type	Description
<b>Positive RL</b>	Adding positive stimulus to increase behavior
<b>Negative RL</b>	Removing negative stimulus to increase behavior

## Based on Learning:

- **Model-based:** Agent learns environment model
- **Model-free:** Agent learns directly from experience

## Mnemonic

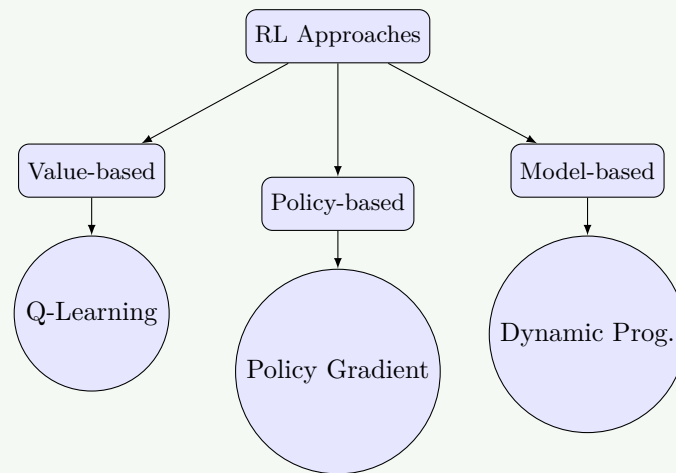
"Positive Adds, Negative Removes"

## Question 2 [c marks]

7 Explain approaches to implement Reinforcement Learning.

### Solution

Approach	Description	Example
<b>Value-based</b>	Learn value of states/actions	Q-Learning
<b>Policy-based</b>	Learn policy directly	Policy Gradient
<b>Model-based</b>	Learn environment model	Dynamic Programming



**Figure 3.** Reinforcement Learning Approaches

- **Value-based:** Estimates value functions
- **Policy-based:** Optimizes policy parameters
- **Model-based:** Uses environment model

#### Mnemonic

"Value, Policy, Model"

## Question 3 [a marks]

3 Describe the activation functions ReLU and sigmoid.

### Solution

Function	Formula	Range
<b>ReLU</b>	$f(x) = \max(0, x)$	$[0, \infty)$
<b>Sigmoid</b>	$f(x) = 1/(1 + e^{-x})$	$(0, 1)$

- **ReLU advantage:** No vanishing gradient problem
- **Sigmoid advantage:** Smooth gradient, probabilistic output

#### Mnemonic

"ReLU Rectifies, Sigmoid Squashes"

### Question 3 [b marks]

4 Explain Multi-layer feed forward ANN.

#### Solution

Component	Description
<b>Input Layer</b>	Receives input data
<b>Hidden Layers</b>	Process information (multiple layers)
<b>Output Layer</b>	Produces final result
<b>Connections</b>	Forward direction only
<ul style="list-style-type: none"> <li>• <b>Information flow:</b> Unidirectional from input to output</li> <li>• <b>No cycles:</b> No feedback connections</li> </ul>	

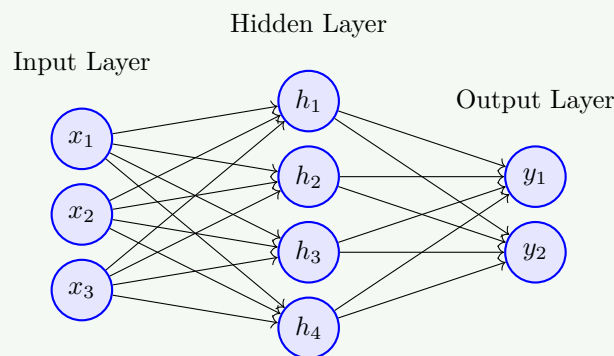
#### Mnemonic

"Input → Hidden → Output (Forward Only)"

### Question 3 [c marks]

7 Draw the structure of ANN and explain functionality of each of its components.

#### Solution



**Figure 4.** Artificial Neural Network Structure

Component	Functionality
<b>Neurons</b>	Processing units that receive inputs and produce outputs
<b>Weights</b>	Connection strengths between neurons
<b>Bias</b>	Additional parameter to shift activation function
<b>Activation Function</b>	Introduces non-linearity to the network
<ul style="list-style-type: none"> <li>• <b>Input layer:</b> Receives and distributes input data</li> <li>• <b>Hidden layers:</b> Extract features and patterns</li> <li>• <b>Output layer:</b> Produces final classification or prediction</li> <li>• <b>Connections:</b> Weighted links between neurons</li> </ul>	

#### Mnemonic

"Neurons with Weights, Bias, and Activation"

### Question 3 [a marks]

3 Write a short note on Backpropagation.

#### Solution

Aspect	Description
<b>Purpose</b>	Training algorithm for neural networks
<b>Method</b>	Gradient descent with chain rule
<b>Direction</b>	Backward error propagation
<ul style="list-style-type: none"> <li>• <b>Process:</b> Calculate error gradients backwards through network</li> <li>• <b>Update:</b> Adjust weights to minimize error</li> </ul>	

#### Mnemonic

"Back-ward Error Propagation"

### Question 3 [b marks]

4 Explain Single-layer feed forward network.

#### Solution

Feature	Description
<b>Structure</b>	Input layer directly connected to output layer
<b>Layers</b>	Only input and output layers
<b>Limitations</b>	Can only solve linearly separable problems
<b>Example</b>	Perceptron
<ul style="list-style-type: none"> <li>• <b>Capability:</b> Limited to linear decision boundaries</li> <li>• <b>Applications:</b> Simple classification tasks</li> </ul>	

#### Mnemonic

"Single Layer, Linear Limits"

### Question 3 [c marks]

7 Draw and explain the architecture of Recurrent neural network.

#### Solution



Figure 5. Recurrent Neural Network Architecture

Component	Function
<b>Hidden State</b>	Maintains memory of previous inputs
<b>Recurrent Connection</b>	Feedback from hidden state to itself
<b>Sequence Processing</b>	Handles sequential data
<ul style="list-style-type: none"> <li>• <b>Memory:</b> Retains information from previous time steps</li> <li>• <b>Applications:</b> Language modeling, speech recognition</li> <li>• <b>Advantage:</b> Can process variable-length sequences</li> </ul>	

**Mnemonic**

"Recurrent Remembers, Loops Back"

## Question 4 [a marks]

3 Define NLP and write down advantages of it.

**Solution**

Term	Definition
<b>NLP</b>	Natural Language Processing - enables computers to understand, interpret, and generate human language

**Advantages:**

- **Human-computer interaction:** Natural communication
- **Automation:** Automated text processing and analysis
- **Accessibility:** Voice interfaces for disabled users

**Mnemonic**

"Natural Language, Natural Interaction"

## Question 4 [b marks]

4 Compare NLU and NLG.

**Solution**

Aspect	NLU (Understanding)	NLG (Generation)
<b>Purpose</b>	Interpret human language	Generate human language
<b>Input</b>	Text/Speech	Structured data
<b>Output</b>	Structured data	Text/Speech
<b>Examples</b>	Sentiment analysis	Text summarization

- **NLU:** Converts unstructured text to structured data
- **NLG:** Converts structured data to natural text

**Mnemonic**

"NLU Understands, NLG Generates"



## Question 4 [c marks]

7 Explain word tokenization and frequency distribution of words with suitable example.

### Solution

Process	Description	Example
<b>Tokenization</b>	Breaking text into individual words/tokens	"Hello world" → ["Hello", "world"]
<b>Frequency Distribution</b>	Counting occurrence of each token	{"Hello": 1, "world": 1}

#### Example:

```

1 Text: "The cat sat on the mat"
2 Tokens: ["The", "cat", "sat", "on", "the", "mat"]
3 Frequency: {"The": 1, "cat": 1, "sat": 1, "on": 1, "the": 1, "mat": 1}
4
```

- **Case sensitivity:** "The" and "the" counted separately
- **Applications:** Text analysis, search engines
- **Preprocessing:** Essential step for NLP tasks

#### Mnemonic

"Tokenize then Count"

## Question 4 [a marks]

3 List disadvantages of NLP.

### Solution

Disadvantage	Description
<b>Ambiguity</b>	Multiple meanings of words/sentences
<b>Context dependency</b>	Meaning changes with context
<b>Language complexity</b>	Grammar rules and exceptions
<ul style="list-style-type: none"> <li>• <b>Cultural variations:</b> Different languages, dialects</li> <li>• <b>Computational cost:</b> Resource-intensive processing</li> </ul>	

#### Mnemonic

"Ambiguous, Contextual, Complex"

## Question 4 [b marks]

4 Explain types of ambiguities in NLP.

### Solution

Type	Description	Example
<b>Lexical</b>	Word has multiple meanings	"Bank" (financial/river)
<b>Syntactic</b>	Multiple parse trees possible	"I saw a man with a telescope"
<b>Semantic</b>	Multiple interpretations	"Flying planes can be dangerous"
<ul style="list-style-type: none"> <li>• <b>Resolution:</b> Context analysis, statistical models</li> </ul>		

- **Challenge:** Major hurdle in NLP systems

#### Mnemonic

"Lexical words, Syntactic structure, Semantic meaning"

## Question 4 [c marks]

7 Explain stemming words and parts of speech(POS) tagging with suitable example.

### Solution

Process	Description	Example
<b>Stemming</b>	Reducing words to root/stem form	"running" → "run", "flies" → "fli"
<b>POS Tagging</b>	Assigning grammatical categories	"The/DT cat/NN runs/VB fast/RB"

#### Stemming Example:

```
1 Original: ["running", "runs", "runner"]
2 Stemmed: ["run", "run", "runner"]
3
```

#### POS Tagging Example:

```
1 Sentence: "The quick brown fox jumps"
2 Tagged: "The/DT quick/JJ brown/JJ fox/NN jumps/VB"
3
```

- **Stemming purpose:** Reduce vocabulary size, group related words
- **POS purpose:** Understand grammatical structure
- **Applications:** Information retrieval, grammar checking

#### Mnemonic

"Stem to Root, Tag by Grammar"

## Question 5 [a marks]

3 Define the term word embedding and list various word embedding techniques.

### Solution

Term	Definition
<b>Word Embedding</b>	Dense vector representations of words that capture semantic relationships

#### Techniques:

- **TF-IDF:** Term Frequency-Inverse Document Frequency
- **Bag of Words (BoW):** Simple word occurrence counting
- **Word2Vec:** Neural network-based embeddings

#### Mnemonic

"TF-IDF counts, BoW bags, Word2Vec vectorizes"

## Question 5 [b marks]

4 Explain about Challenges with TF-IDF and BoW.

### Solution

Method	Challenges
TF-IDF	Sparse vectors, no semantic similarity, high dimensionality
BoW	Order ignored, context lost, sparse representation

#### Common Issues:

- **Sparsity:** Most vector elements are zero
- **No semantics:** Similar words have different vectors
- **High dimensions:** Memory and computation intensive

#### Mnemonic

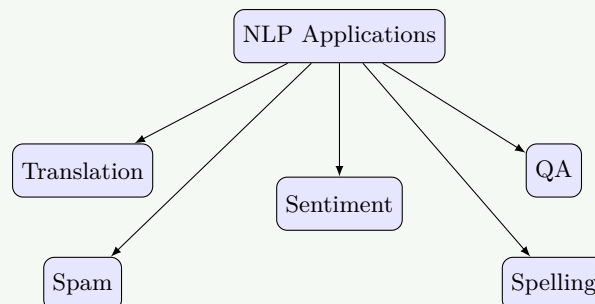
"Sparse, No Semantics, High Dimensions"

## Question 5 [c marks]

7 Explain applications of NLP with suitable examples.

### Solution

Application	Description	Example
Machine Translation	Translate between languages	Google Translate
Sentiment Analysis	Determine emotional tone	Product review analysis
Question Answering	Answer questions from text	Chatbots, virtual assistants
Spam Detection	Identify unwanted emails	Email filters
Spelling Correction	Fix spelling errors	Auto-correct in text editors



**Figure 6.** NLP Applications

- **Real-world impact:** Improves human-computer interaction
- **Business value:** Automates text processing tasks
- **Growing field:** New applications emerging constantly

#### Mnemonic

"Translate, Sentiment, Question, Spam, Spell"

## Question 5 [a marks]

3 Describe the Glove(Global Vector for word representation).

### Solution

Aspect	Description
<b>Purpose</b>	Create word vectors using global corpus statistics
<b>Method</b>	Combines global matrix factorization and local context
<b>Advantage</b>	Captures both global and local statistical information
<ul style="list-style-type: none"> <li>• <b>Global statistics:</b> Uses word co-occurrence information</li> <li>• <b>Pre-trained:</b> Available trained vectors for common use</li> </ul>	

### Mnemonic

"Global Vectors, Local Context"

## Question 5 [b marks]

4 Explain the Inverse Document Frequency (IDF).

### Solution

Component	Formula	Purpose
<b>IDF</b>	$\log(N/df)$	Measure word importance across documents
<b>N</b>	Total documents	Corpus size
<b>df</b>	Document frequency	Documents containing the term
<ul style="list-style-type: none"> <li>• <b>High IDF:</b> Rare words (more informative)</li> <li>• <b>Low IDF:</b> Common words (less informative)</li> <li>• <b>Application:</b> Part of TF-IDF weighting scheme</li> </ul>		

### Mnemonic

"Inverse Document, Rare is Important"

## Question 5 [c marks]

7 Explain calculation of TF(Term Frequency) for a document with suitable example.

### Solution

#### Methods:

- **Raw TF:** Simple count  $f(t,d)$
- **Normalized TF:**  $f(t,d)/\max(f(w,d))$
- **Log TF:**  $1 + \log(f(t,d))$

**Example:** Document "The cat sat on the mat." has "the" appearing 3 times.

#### Steps:

1. Count occurrences.
2. Apply formula.
3. Use in TF-IDF.

**Mnemonic**

"Count, Normalize, Log"