

COLLECTING MACHINE LEARNING QUIZ DATA USING WEB SCRAPING

questions number : 88

question example : What does "AI bias" mean when used in relation to artificial intelligence?

example answers :A.The tendency of AI systems to make decisions without data,B.The unintentional discrimination in AI systems due to biased training data,C.The ability of AI systems to make decisions without human intervention,D.The ethical considerations when designing AI algorithms.

website : [\[https://mcqprime.com/artificial-intelligence-mcq/\]](https://mcqprime.com/artificial-intelligence-mcq/)

quiz1 csv : [\[https://github.com/Minhtrna/AIL303m/blob/main/LAB-DATA/quiz1.csv\]](https://github.com/Minhtrna/AIL303m/blob/main/LAB-DATA/quiz1.csv)

quiz2 csv : [\[https://github.com/Minhtrna/AIL303m/blob/main/LAB-DATA/quiz2.csv\]](https://github.com/Minhtrna/AIL303m/blob/main/LAB-DATA/quiz2.csv)

```
In [2]: import pandas as pd
import requests
from bs4 import BeautifulSoup
```

```
In [3]: # Setup lấy data
quiz_items=0
def scap(url):
    global quiz_items
    headers = {'User-Agent': 'Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0'}
    response = requests.get(url, headers=headers)
    response.encoding = 'utf-8'
    soup = BeautifulSoup(response.text, 'html.parser')
    quiz_items = soup.find_all('ol', class_='wp-block-list')
    return quiz_items

print(scap('https://mcqprime.com/artificial-intelligence-mcq/')) # test
```

[<ol class="wp-block-list">
What does "AI bias" mean when used in relation to artificial intelligence?
a. The tendency of AI systems to make decisions without data
b. The unintentional discrimination in AI systems due to biased training data
c. The ability of AI systems to make decisions without human intervention
d. The ethical considerations when designing AI algorithms
, <ol class="wp-block-list" start="2">
Which AI technique is often used for clustering data into groups with similar characteristics?
a. Regression
b. Reinforcement Learning
c. Supervised Learning
d. Unsupervised Learning
, <ol class="wp-block-list" start="3">
What function does a Recurrent Neural Network (RNN) perform in AI?
a. To classify images
b. To analyze text sentiment
c. To process sequences of data with memory
d. To play board games
, <ol class="wp-block-list" start="4">
The main objective of artificial intelligence?
a. To replicate human intelligence exactly
b. To develop computer programs that can think for themselves
c. To mimic human behavior without understanding
d. To solve complex problems using algorithms
, <ol class="wp-block-list" start="5">
What does "NLU" mean when referring to natural language?
a. Natural Language Understanding
b. Neural Language Unit
c. New Linguistic Understanding
d. Neutral Language Understanding
, <ol class="wp-block-list" start="6">
What is the name of the kind of AI system that can complete particular tasks at or better than human levels of proficiency?
a. Strong AI
b. Weak AI
c. Narrow AI
d. General AI
, <ol class="wp-block-list" start="7">
What is the purpose of an AI chatbot?
a. To play video games
b. To automate routine customer service tasks
c. To generate random text
d. To translate languages
, <ol class="wp-block-list" start="8">
Which AI application involves teaching a computer to perform a task by example, rather than through explicit programming?
a. Supervised Learning
b. Unsupervised Learning
c. Reinforcement Learning
d. Machine Vision
, <ol class="wp-block-list" start="9">
What does "IoT" mean in terms of artificial intelligence?
a. Internet of Things
b. Intelligence of Technology
c. Input of Text
d. Internet of Training
, <ol class="wp-block-list" start="10">
Which of the following is not a common application of natural language processing (NLP)?
a. Sentiment analysis
b. Machine translation
c. Image recognition
d. Chatbots
, <ol class="wp-block-list" start="11">
Why is the Turing Test used in AI?
a. To evaluate the computational power of a computer
b. To measure a machine's ability to exhibit human-like intelligence
c. To assess the speed of algorithm execution
d. To determine if a computer can defeat a human in chess
, <ol class="wp-block-list" start="12">
Which programming language was designed for artificial intelligence and symbolic reasoning?
a. C++
b. Python
c. Lisp
d. Ruby
, <ol class="wp-block-list" start="13">
What does machine learning and AI mean when they refer to "data augmentation"?
a. Creating new data by generating random values
b. Increasing the size of the training dataset through transformations
c. Reducing the size

of the training dataset for efficiency

d. Scaling the data to fit a specific range

What is the name of the form of machine learning in which the algorithm learns to predict outcomes using labeled data?

a. Supervised Learning

b. Unsupervised Learning

c. Reinforcement Learning

d. Semi-supervised Learning

Which AI strategy makes use of neural networks with numerous layers and aims to mimic how people think and learn?

a. Symbolic AI

b. Evolutionary Computing

c. Deep Learning

d. Genetic Algorithms

What does “AI ethics” mean when used in artificial intelligence?

a. The study of AI algorithms

b. The design of AI systems to be ethically neutral

c. The ethical considerations and guidelines for AI development and use

d. The use of AI for ethical decision-making

Which AI method is motivated by the laws of natural selection and the survival of the fittest?

a. Reinforcement Learning

b. Genetic Algorithms

c. Deep Learning

d. Fuzzy Logic

What function does an AI-based recommendation system provide?

a. To classify images

b. To optimize supply chain management

c. To suggest personalized content or products to users

d. To control industrial robots

What do you know by “named entity recognition” in natural language processing?

a. Identifying and classifying entities in text, such as names of people, places, and organizations

b. Analyzing sentiment in text

c. Translating text from one language to another

d. Summarizing lengthy documents

What is the name of machine learning where the algorithm looks for patterns or groupings in data even when the labels are not explicit?

a. Supervised Learning

b. Unsupervised Learning

c. Reinforcement Learning

d. Semi-supervised Learning

Which of the following cannot be considered a subset of AI?

a. Machine Learning

b. Natural Language Processing (NLP)

c. Expert Systems

d. Virtual Reality (VR)

What do you call a computer system that is capable of activities like speech recognition, visual perception, and decision-making, which generally need human intelligence?

a. Expert System

b. Neural Network

c. Turing Machine

d. Supercomputer

Which AI strategy maximizes a payoff in an uncertain environment when making decisions?

a. Genetic Algorithms

b. Reinforcement Learning

c. Expert Systems

d. Fuzzy Logic

What is the most popular programming language for developing AI and machine learning applications?

a. C++

b. Java

c. Python

d. Ruby

Which of the following is not a Python module or framework for natural language processing?

a. NLTK (Natural Language Toolkit)

b. TensorFlow

c. SpaCy

d. Gensim

What is the main advantage of using convolutional neural networks in image recognition work?
a. They require less training data
b. They can process text data effectively
c. They capture spatial relationships in images
d. They are more interpretable

Which AI application promotes vehicle delivery routes with the use of algorithms?
a. Natural Language Processing (NLP)
b. Robotics
c. Genetic Algorithms
d. Machine Vision

What does the term “overfitting” mean in AI?
a. The model performs well on the training data but poorly on new data
b. The model is too simple to capture complex patterns in the data
c. The model is biased towards a specific class of data
d. The model is unable to learn from data

What is the name of the AI system which can comprehend, interpret, and produce human language?
a. Natural Language Processing (NLP)
b. Machine Vision
c. Reinforcement Learning
d. Genetic Algorithms

Which of the following is an example of a task requiring supervised learning?
a. Image classification
b. Clustering customer data
c. Training a chatbot to have a conversation
d. Reinforcement learning in a game

What is the name of a particular type of AI system that can carry out a variety of tasks and displays intelligence similar to that of a human?
a. Strong AI
b. Weak AI
c. Narrow AI
d. General AI

What does “CNN” stand for in the field of artificial intelligence?
a. Convolutional Neural Network
b. Common Neural Network
c. Continuous Neural Network
d. Complex Neural Network

What is the name of an AI system’s ability to become more effective when additional data is made available to it?
a. Static Learning
b. Transfer Learning
c. Dynamic Learning
d. Incremental Learning

What application of artificial intelligence makes use of algorithms to examine and analyse visual data, frequently for tasks like object detection or facial recognition?
a. Natural Language Processing (NLP)
b. Robotics
c. Machine Vision
d. Genetic Algorithms

What is the main goal of reinforcement learning?
a. To classify data into categories
b. To optimize a system for maximum efficiency
c. To make predictions based on historical data
d. To learn optimal actions through trial and error

Which AI method uses a set of rules to describe knowledge and then uses those rules to solve problems?
a. Expert Systems
b. Neural Networks
c. Genetic Algorithms
d. Natural Language Processing (NLP)

What is the main drawback of applying deep learning models, like Convolutional Neural Networks (CNNs)?
a. They require a large amount of labeled data for training
b. They are not suitable for image recognition tasks
c. They cannot handle sequential data
d. They are computationally inefficient

What type of AI method is frequently used to simulate natural selection while maximizing solutions to challenging problems?
a. Supervised Learning
b. Reinforcement Learning
c. Genetic Algorithms
d. Unsupervised Learning

ning

, <ol class="wp-block-list" start="38">

What is the goal of a collaborative filtering-based AI recommendation system?
a. To generate content for a website
b. To provide personalized recommendations based on user behavior and preferences
c. To classify images
d. To optimize supply chain management

, <ol class="wp-block-list" start="39">

What is the process of training a model to carry out a task by giving it examples of appropriate behavior in AI?
a. Supervised Learning
b. Unsupervised Learning
c. Reinforcement Learning
d. Semi-supervised Learning

, <ol class="wp-block-list" start="40">

Which AI technology employs algorithms to produce text or speech that sounds human?
a. Machine Vision
b. Natural Language Processing (NLP)
c. Speech Recognition
d. Genetic Algorithms

, <ol class="wp-block-list" start="41">

What is the phrase used to describe a machine learning model's capacity to generalize from the training data to make precise predictions on additional data?
a. Overfitting
b. Bias
c. Variance
d. Generalization

, <ol class="wp-block-list" start="42">

What is the name of the AI technique where several models or experts are combined to create predictions or decisions?
a. Reinforcement Learning
b. Transfer Learning
c. Ensemble Learning
d. Deep Learning

, <ol class="wp-block-list" start="43">

Which of the following activation functions are frequently used in artificial neural networks?
a. Sigmoid
b. Regression
c. Fuzzy Logic
d. Principal Component Analysis (PCA)

, <ol class="wp-block-list" start="44">

What is the main benefit of applying reinforcement learning to AI for activities like playing games?
a. It requires a large amount of labeled data
b. It can handle complex sequences of actions
c. It is not suitable for real-time tasks
d. It relies on pre-defined rules

, <ol class="wp-block-list" start="45">

What is the name of the method used in AI that enables a model to still make predictions even when portion of the input data is missing?
a. Regularization
b. Imputation
c. Feature Engineering
d. Dimensionality Reduction

, <ol class="wp-block-list" start="46">

Which of the following approaches is frequently used to evaluate an AI classification model's effectiveness?
a. Mean Absolute Error (MAE)
b. R-squared (R^2)
c. Confusion Matrix
d. Root Mean Square Error (RMSE)

, <ol class="wp-block-list" start="47">

What does "backpropagation" mean in AI?
a. The process of training a neural network by adjusting its weights based on errors
b. The technique used for data augmentation
c. The process of encoding categorical data
d. The method for calculating feature importance

, <ol class="wp-block-list" start="48">

What is the name of the kind of AI model, such as a self-driving car, that can complete tasks without constant human intervention?
a. Strong AI
b. Weak AI
c. Narrow AI
d. General AI

, <ol class="wp-block-list" start="49">

What does "cross-validation" mean in AI?
a. The process of validating the results of a neural network
b. The process of training a model on multiple datasets
c. The process of splitting data into training and testing sets
d. The technique for evaluating model performance using multiple subsets of data

, <ol class="wp-block-list" start="50">

What is the process by which text data is transformed into a numerical representation for machine learning?
a. Text Analysis
b. Text Mining
c. Text Encoding
d. Text Embedding

, <ol class="wp-block-list" start="51">

Which of the following is a deep learning framework that is not frequently used?
a. TensorFlow
b. PyTorch
c. Keras
d. Scikit-Learn

, <ol class="wp-block-list" start="52">

What is the primary application of a generative adversarial network (GAN) in AI?
a. Image classification
b. Image generation
c. Natural language processing
d. Speech recognition

, <ol class="wp-block-list" start="53">

What is the name of the process that turns raw data into a format that machine learning models can use?
a. Data Labeling
b. Data Engineering
c. Data Imputation
d. Data Visualization

, <ol class="wp-block-list" start="54">

Which AI algorithms are used to understand and interpret human movements and gestures?
a. Speech Recognition
b. Machine Vision
c. Natural Language Processing (NLP)
d. Gesture Recognition

, <ol class="wp-block-list" start="55">

What is the primary goal of reinforcement learning in AI?
a. To optimize the accuracy of predictions
b. To find patterns and relationships in data
c. To learn optimal actions based on rewards and punishments
d. To classify data into predefined categories

, <ol class="wp-block-list" start="56">

What is the name of the AI technique where a model is trained on one task, then used for another but similar task?
a. Reinforcement Learning
b. Transfer Learning
c. Ensemble Learning
d. Deep Learning

, <ol class="wp-block-list" start="57">

Which of the following is an example of a generative model?
a. Support Vector Machine (SVM)
b. Recurrent Neural Network (RNN)
c. K-Means Clustering
d. Principal Component Analysis (PCA)

, <ol class="wp-block-list" start="58">

What is the main goal of natural language processing?
a. To analyze and process human emotions
b. To generate random text
c. To understand, interpret, and generate human language
d. To recognize and understand visual information

, <ol class="wp-block-list" start="59">

Which AI approach uses rules and knowledge representation to arrive at logically sound decisions?
a. Supervised Learning
b. Expert Systems
c. Deep Learning
d. Reinforcement Learning

, <ol class="wp-block-list" start="60">

What is the name of the capability of an AI system to observe and comprehend the spatial structure of the environment?
a. Natural Language Processing (NLP)
b. Computer Vision
c. Speech Recognition
d. Sentiment Analysis

, <ol class="wp-block-list" start="61">

What does "bias-variance trade-off" mean in the context of AI?
a. The need for more data to reduce model bias
b. The balance between model complexity and generalization
c. The choice between supervised and unsupervised learning
d. The trade-off between accuracy and precision

, <ol class="wp-block-list" start="62">

Which of the following approaches is utilized in AI to reduce dimensionality?
a. K-Means Clustering
b. Principal Component Analysis (PCA)
c. Recurrent Neural Networks (RNNs)
d. Support Vector Machines (SVMs)

, <ol class="wp-block-list" start="63">
What is the name of the metric used in AI to evaluate a classification model's performance?
a. Root Mean Square Error (RMSE)
b. F1 Score
c. Gradient Descent
d. Perceptron Loss
, <ol class="wp-block-list" start="64">
What is the name for an AI model that has been trained to accomplish a single task and cannot simply be applied to other tasks?
a. Strong AI
b. Weak AI
c. Narrow AI
d. General AI
, <ol class="wp-block-list" start="65">
What kind of AI method uses labeled data to divide input into predetermined categories?
a. Reinforcement Learning
b. Unsupervised Learning
c. Supervised Learning
d. Semi-supervised Learning
, <ol class="wp-block-list" start="66">
What is the procedure for reducing the dimensionality of data while maintaining its essential characteristics?
a. Principal Component Analysis (PCA)
b. Regression
c. K-Means Clustering
d. Gradient Boosting
, <ol class="wp-block-list" start="67">
What does "chatbot" mean in terms of AI?
a. Chatting Robot
b. Chatter Robot
c. Computer Help Assistance Technology for Browsing Online Text
d. Communication Helper Automated Technology Based on Text
, <ol class="wp-block-list" start="68">
Which AI application uses algorithms to decode and interpret human emotions from speech or text?
a. Sentiment Analysis
b. Image Recognition
c. Autonomous Vehicles
d. Fraud Detection
, <ol class="wp-block-list" start="69">
RNN stand for __.
a. Recursive Neural Network
b. Randomized Neural Network
c. Recurrent Neural Network
d. Robotic Neural Network
, <ol class="wp-block-list" start="70">
What is the main advantage of using unsupervised learning in AI?
a. It can make predictions based on labeled data
b. It can perform image recognition tasks
c. It can discover hidden patterns and relationships in data
d. It requires less computational power
, <ol class="wp-block-list" start="71">
Which AI method is frequently applied to maximizing judgments in uncertain or inadequate information situations?
a. Reinforcement Learning
b. Supervised Learning
c. Genetic Algorithms
d. Fuzzy Logic
, <ol class="wp-block-list" start="72">
How you can describe an AI system's capacity to identify and comprehend human speech?
a. Natural Language Processing (NLP)
b. Speech Recognition
c. Audio Analysis
d. Language Understanding System
, <ol class="wp-block-list" start="73">
Which of the following open-source machine learning libraries is well-known for offering tools for modeling and data analysis?
a. TensorFlow
b. Keras
c. PyTorch
d. Pandas
, <ol class="wp-block-list" start="74">
What is the process of optimizing a deep learning model that has already been trained for a particular task using a smaller dataset?
a. Reinforcement Learning
b. Transfer Learning
c. Unsupervised Learning
d. Semi-supervised Learning
, <ol class="wp-block-list" start="75">
Which AI method is employed to solve optimization problems and is modeled after the behavior of ants, bees, and other social insects?
a. Reinforcement Learning
b. Swarm Intelligence
c. Genetic Algorithms
d. Expert Systems
, <ol class="wp-block-list" start="76">

What does “feature engineering” mean in the context of AI?
a. The process of designing user interfaces for AI applications
b. The process of selecting and transforming input data to improve model performance
c. The process of encoding rules in expert systems
d. The process of fine-tuning hyperparameters in machine learning models

Which method is frequently used to decrease overfitting in machine learning models?
a. Increasing the model complexity
b. Decreasing the amount of training data
c. Regularization techniques like L1 and L2
d. Removing all features except one

What is the name of the capability of an AI system to comprehend and analyze visual data from the real world?
a. Natural Language Processing (NLP)
b. Machine Vision
c. Reinforcement Learning
d. Genetic Algorithms

What is the name for a mathematical model in AI that use a system of equations to imitate the behavior of a complicated system?
a. Artificial Neural Network (ANN)
b. Genetic Algorithm (Ga.
c. Agent-Based Model (ABM)
d. Turing Machine

What is the name for a certain class of AI model that has been taught to forecast a continuous numerical value?
a. Classification Model
b. Regression Model
c. Clustering Model
d. Ensemble Model

What is the phrase used in AI to describe a model’s capacity to make decisions based on ambiguous or insufficient data?
a. Fuzzy Logic
b. Genetic Algorithms
c. Reinforcement Learning
d. Deep Learning

What is the name of the neural network architecture that has input, hidden, and output layers among its interconnected layers?
a. Recurrent Neural Network (RNN)
b. Convolutional Neural Network (CNN)
c. Feedforward Neural Network
d. Radial Basis Function Network (RBFN)

Which AI method is frequently used to divide data into groups or categories based on their similarities?
a. Reinforcement Learning
b. Clustering
c. Regression
d. Dimensionality Reduction

What is the name of the method used in AI that enables a model to learn from its errors and enhance its performance over time?
a. Reinforcement Learning
b. Supervised Learning
c. Unsupervised Learning
d. Semi-supervised Learning

What is the main objective of an AI system that does machine translation using natural language processing (NLP)?
a. To summarize text documents
b. To convert speech to text
c. To translate text from one language to another
d. To generate human-like text

What is the name of a particular class of AI model that may produce new data points that are comparable to current data points?
a. Discriminative Model
b. Generative Model
c. Ensemble Model
d. Reinforcement Model

Which of the following unsupervised learning techniques are often used for dimensionality reduction?
a. Decision Trees
b. Naive Bayes
c. Singular Value Decomposition (SVd.
d. Random Forest


```

</ol>, <ol class="wp-block-list" start="89">
<li><strong>What does the term “hyperparameter” refer in AI?</strong><br/>a. Paramet
ers learned by the model during training<br/>b. Parameters that define the structure
of the model<br/>c. Parameters used to make predictions<br/>d. Parameters related to
the loss function</li>
</ol>, <ol class="wp-block-list" start="90">
<li><strong>Which of the following is a common algorithm used for recommendation sys
tems in AI?</strong><br/>a. Naive Bayes<br/>b. Linear Regression<br/>c. Matrix Facto
rization<br/>d. Decision Trees</li>
</ol>]

```

```

In [4]: questions = []
        optionA = []
        optionB = []
        optionC = []
        optionD = []

```

```

In [5]: # Tách câu hỏi và đáp án
def createdata(quiz_items):
    global questions
    global optionA
    global optionB
    global optionC
    global optionD
    for item in quiz_items:
        # Loại bỏ khoảng trắng
        text = item.text.strip()
        #text = text.replace('Xem đáp án', '') test web vietnam
        text = text.replace(' ', '')

        # tách câu hỏi và đáp án ( đáp án bắt đầu từ a. b. c. d. ==> tách tất cả nh
        parts = text.split("a.", 1)
        if len(parts) == 2:
            # câu hỏi
            question = parts[0].strip()
            question = ' '.join(question.split()) # xóa khoảng trắng và cách dòng
            answers = parts[1]

            # đáp án
            options = answers.split("b.")
            if len(options) == 2:
                a = "A." + options[0].strip()
                rest = options[1]

                parts = rest.split("c.")
                if len(parts) == 2:
                    b = "B." + parts[0].strip()
                    rest = parts[1]

                    parts = rest.split("d.")
                    if len(parts) == 2:
                        c = "C." + parts[0].strip()
                        d = "D." + parts[1].strip()

                    questions.append(question)
                    optionA.append(a)

```

```
        optionB.append(b)
        optionC.append(c)
        optionD.append(d)
    return questions, optionA, optionB, optionC, optionD
```

```
In [6]: def createDf(quiz_items):
        global questions
        global optionA
        global optionB
        global optionC
        global optionD
        createdata(quiz_items)
        df = pd.DataFrame({
            'Question': questions,
            'Option A': optionA,
            'Option B': optionB,
            'Option C': optionC,
            'Option D': optionD
        })
        return df
```

```
In [7]: # Link 1 : https://mcqprime.com/machine-learning-mcq/
        # Link 2 : https://mcqprime.com/artificial-intelligence-mcq/
        quiz_items=0
        questions = []
        optionA = []
        optionB = []
        optionC = []
        optionD = []
        quiz_items=scap('https://mcqprime.com/artificial-intelligence-mcq/')
        df=createDf(quiz_items)
        df.head(10)
```

Out[7]:

	Question	Option A	Option B	Option C	Option D
0	What does "AI bias" mean when used in relation...	A.The tendency of AI systems to make decisions...	B.The unintentional discrimination in AI syste...	C.The ability of AI systems to make decisions ...	D.The ethical considerations when designing AI...
1	Which AI technique is often used for clusterin...	A.Regression	B.Reinforcement Learning	C.Supervised Learning	D.Unsupervised Learning
2	What function does a Recurrent Neural Network ...	A.To classify images	B.To analyze text sentiment	C.To process sequences of data with memory	D.To play board games
3	The main objective of artificial intelligence?	A.To replicate human intelligence exactly	B.To develop computer programs that can think ...	C.To mimic human behavior without understanding	D.To solve complex problems using algorithms
4	What does "NLU" mean when referring to natural...	A.Natural Language Understanding	B.Neural Language Unit	C.New Linguistic Understanding	D.Neutral Language Understanding
5	What is the name of the kind of AI system that...	A.Strong AI	B.Weak AI	C.Narrow AI	D.General AI
6	What is the purpose of an AI chatbot?	A.To play video games	B.To automate routine customer service tasks	C.To generate random text	D.To translate languages
7	Which AI application involves teaching a compu...	A.Supervised Learning	B.Unsupervised Learning	C.Reinforcement Learning	D.Machine Vision
8	What does "IoT" mean in terms of artificial in...	A.Internet of Things	B.Intelligence of Technology	C.Input of Text	D.Internet of Training
9	Which of the following is not a common applica...	A.Sentiment analysis	B.Machine translation	C.Image recognition	D.Chatbots

```
In [8]: # save thành csv (test)
df.to_csv('lab1a.csv', index=True)
```

```
# save thành json  
df.to_json('lab1a.json', orient='records')
```