EBU4203 Introduction to AI – Week 1 Tutorial 2024

Q1: STATE and EXPLAIN the THREE ways that an experiment can fail when uncertainties are not considered appropriately.

LAZINESS: too much work to list the complete set of antecedents or consequents needed to ensure an exceptionless rule and too hard to use such rules.

It can be too cumbersome to list all necessary conditions and rules, making it challenging to account for every scenario.

THEORETICAL ignorance: Medical science has no complete theory for the domain.

Some fields like medicine, don't have complete theories for all conditions, meaning the knowledge base is inherently limited.

PRACTICAL ignorance: Even if we know all the rules, we might be uncertain about a particular patient because not all the necessary tests have been or can be run.

Even if the rules are known, not every test or condition can be observed or verified in real life, introducing gaps in the system's ability to make accurate decisions.

(Slides page 77)

Q2: STATE and EXPLAIN the THREE main types of learning in AI.

The type of learning depends on the feedback to learn from, there are 3 main types of learning, namely unsupervised learning, supervised learning and reinforcement learning.

1.Unsupervised learning: the agent learns patterns in the input even though no explicit feedback is supplied. (Example: detecting potentially useful clusters of input examples.)

Unsupervised machine learning is the training of models on raw and unlabelled training data. It is often used to identify patterns and trends in raw datasets, or to cluster similar data into a specific number of groups. It's also often an approach used in the early exploratory phase to better understand the datasets.

The vast majority of available data is unlabelled, raw data. By grouping data along similar features or analysing datasets for underlying patterns, unsupervised learning is a powerful tool used to gain insight from this data. In contrast, supervised machine learning can be resource intensive because of the need for labelled data.

Unsupervised machine learning is mainly used to:

- Cluster datasets on similarities between features or segment data
- Understand relationship between different data point such as automated music recommendations
- Perform initial data analysis

2.Supervised learning: the agent observes some example input–output pairs and learns a function that maps from input to output. (Example: the inputs are percepts and the outputs are provided by a teacher who says "Brake. or "Turn left.").

Supervised machine learning requires labelled input and output data during the training phase of the machine learning model lifecycle. This training data is often labelled by a data scientist in the preparation phase, before being used to train and test the model. Once the model has learned the relationship between the input and output data, it can be used to classify new and unseen datasets and predict outcomes.

The reason it is called supervised machine learning is because at least part of this approach requires human oversight. The vast majority of available data is unlabelled, raw data. Human interaction is generally required to accurately label data ready for supervised learning. Naturally, this can be a resource intensive process, as large arrays of accurately labelled training data is needed.

Supervised machine learning is often used for:

- Classifying different file types such as images, documents, or written words.
- Forecasting future trends and outcomes through learning patterns in training data.

3.Reinforcement learning: the agent learns from a series of reinforcements—rewards or punishments. (Example, the lack of a tip (reward) at the end of the journey gives the taxi agent an indication that it did something wrong.)

In reinforcement learning, developers devise a method of rewarding desired behaviors and punishing negative behaviors. This method assigns positive values to the desired actions to encourage the agent to use them, while negative values are assigned to undesired behaviors to discourage them. This programs the agent to seek long-term and maximum overall rewards to achieve an optimal solution.

Current uses include but are not limited to the following:

- Gaming.
- Resource management.
- Personalized recommendations.
- Robotics.

(Slides page 152-157)

Q3: Discuss the challenges and limitations of current AI technologies.

Data limitations: AI heavily relies on large amounts of high-quality data for training and learning. Limited or biased data can lead to inaccurate or biased AI systems.

Lack of interpretability: Many AI algorithms, such as deep neural networks, are considered black boxes, making it challenging to understand and interpret their decision-making processes.

Computing power and resource requirements: AI algorithms, particularly deep learning, often require significant computing power and resources, which can be a limitation for certain applications or organizations.

Q4: In class, 30% of students study History, 45% study Maths, and 15% study both History and Maths. If a student is randomly selected, what is the probability that he/she study History or maths?

Given that 45% study History, i.e., P(H) = 45/100 = 9/20

30% study Maths, i.e., P(M) = 30/100 = 6/20

15% study both History and Maths, i.e., P(H and M) = 15/100 = 3/20

So, P(H or M) = P(H) + P(M) - P(H and M)

= 9/20 + 6/20 - 3/20

= 12/20

= 3/5

(Slides page 94)

eq:$P(A \cup B) = P(A)+P(B)-P(A\cap B)$

Q5: Suppose that 10% of the registered patients of a chest clinic have been diagnosed with cancer. Also based on the clinic's data, 50% of its registered patients are smokers, and 80% of patients diagnosed with cancer are smokers. Suppose that patient X comes into the clinic for the first time. What is the probability that X will be diagnosed with cancer if we know that he is a smoker?

Solution:

We are given: P(cancer)=0.1, P(smoker)=0.5 and P(smoker|cancer)=0.8

From simple Bayes, it follows that P(cancer|smoker) = P(smoker|cancer)*P(cancer) / P(smoker) = (0.8x0.1)/0.5 = 0.16

(Slides page 106)

Eq: $\boldsymbol{P(A \mid B)\, P(B) = P(B \mid A)\, P(A)}$

Q6. Discuss the importance of understanding prior probabilities when interpreting predictions from AI models.

Prior probabilities represent the baseline likelihood of an event or condition before considering new evidence.

In AI predictions, especially in medical or criminal justice applications, understanding these priors is crucial as they influence the model's overall prediction confidence. Ignoring prior probabilities may lead to over-reliance on the model's output without considering the context.

For example, even if a model predicts high probability for a rare disease, the low prior probability of the disease should temper the interpretation. Bayes' theorem exemplifies this by combining prior probabilities with likelihoods to produce more balanced, context-aware predictions.

(This is more about the conditional probability (slide 120) and Bayes Rule)

Conditional Probability

- Definition:

P(A∩B)=P(A|B)×P(B)

where:

  - P(A|B): Probability of A given that we know B
  - P(A): Known as the *prior probability* of A
  - P(A|B): Referred to as the *posterior* or *conditional probability* of A given B.

The prior probability, *P(A)*, reflects the likelihood of event *A* occurring independently of other factors. When additional evidence *B* is available, conditional probability, *P(A|B)*, is used, allowing the interpretation of AI predictions to consider how prior information impacts the likelihood of outcomes.

For example, if *A* is "having a disease" and *B* is "testing positive for the disease," the conditional probability *P(A|B)* tells us the likelihood of actually having the disease after knowing the test result.

The relationship is governed by Bayes' Rule: *P(A|B) = P(B|A) . P(A) / P(B|A)*. Here, *P(A)* is the prior, and P(B|A) is how likely we are to see the evidence B if A were true.

Q7: The dataset in the table below has 10 observations belonging to two classes 'Yes' and 'No'. In the outcome, 6 observations belong to the class 'Yes', and 4 observations belong to class 'No'. Answer the following questions accordingly:

Grade	Colour	Outcome
Low	Red	Yes
Low	Red	No
Low	Yellow	Yes
Low	Yellow	Yes
Low	Red	Yes
High	Yellow	Yes
High	Red	No
High	Red	No
High	Red	Yes
High	Yellow	No

i) Compute the entropy of the outcome $\mathrm{H}(outcome)$.
Solution to i):

E(S) = - (6/10 * log $_2$ * 6/10 + 4/10 * log $_2$ * 4/10) ≈ 0.971

(Slides page 170)

Using the equation for calculating the entropy

$$\mathrm{H}(X) = -\sum_{i=1}^{n} P(X = i) \log_2 P(X = i)$$

ii) Explain what will happen to the Entropy for $\mathrm{H}(outcome)$ if all 10 observations belong to 1 class only, e.g., all outcomes are 'Yes'.

Solution to ii)

E(S) = - (1 $\log_2$ 1) = 0

(Slides page 170)

Still using the equation for calculating the entropy

$$\mathrm{H}(X) = -\sum_{i=1}^{n} P(X = i) \log_2 P(X = i)$$

iii) Compute the entropy for $\mathrm{H}(outcome|colour)$ and $\mathrm{H}(outcome|observations)$.

$$\mathrm{H}(outcome|colour) = -\frac{6}{10}\left(\frac{1}{2}\log_2\frac{1}{2}+\frac{1}{2}\log_2\frac{1}{2}\right) - \frac{4}{10}\left(\frac{3}{4}\log_2\frac{3}{4}+\frac{1}{4}\log_2\frac{1}{4}\right) = 0.92$$

$$\mathrm{H}(outcome|grade) = -\frac{1}{2}\left(\frac{4}{5}\log_2\frac{4}{5}+\frac{1}{5}\log_2\frac{1}{5}\right) - \frac{1}{2}\left(\frac{2}{5}\log_2\frac{2}{5}+\frac{3}{5}\log_2\frac{3}{5}\right) = 0.84$$

(Slides page 171-172)

Using the equation for calculating the conditional entropy

$$\mathrm{H}(Y|X) = -\sum_{x\in X}\sum_{y\in Y} p(x,y) \log_2 p(y|x)$$

iv) Based on your answers in iii), Compute the Information Gain for both conditions.

IG $(outcome, colour)$= $H(outcome)$- $H(outcome|colour)$=0.971-0.92 = 0.051

IG $(outcome, grade)$= $H(outcome)$- $H(outcome|grade)$=0.971-0.84 = 0.131

Eq for calculating the information gain: $I(X_n, Y) = \mathrm{H(Y)} - \mathrm{H}(Y|X_n)$

(Slides page 177)

v) Based on your answers in iv), state which attribute should be used as the root of a decision tree?

Grade

Choose the decision tree based on how much information we would gain from the decision. Since IG $(outcome, grade)$ > IG $(outcome, colour)$, we choose Grade.