Practice Questions I
Topic: ML and NLP

Question 1 - A classifier is tested on 100 samples and makes the following predictions:

- True Positives (TP) = 40
- True Negatives (TN) = 30
- False Positives (FP) = 20
- False Negatives (FN) = 10

Find the accuracy using the formula

Solution: Accuracy = TP + TN / TP + TN + FP + FN

Accuracy = 40 + 30/ 40 + 30 + 20 + 10 = 70/100 = 0.7 = 70%

Answer - 70%

Question 2 - A K-means clustering algorithm groups 2000 data points into 4 clusters. If each cluster gets an equal number of points, how many points does each cluster contain?

Solution - Points per cluster = 2000/4 = 500

Answer - 500 points per cluster

Question 3 - Principal Component Analysis (PCA) reduces a dataset from 12 dimensions to 4 dimensions. If the original dataset has 6000 samples, how many feature values exist in the transformed dataset?

Solution - Total feature values = 6000 * 4 = 24000

Answer - 24000 feature values

Question 4 - A reinforcement learning agent receives rewards as follows:

- +10 for correct actions
- −5 for incorrect actions
- +3 for neutral actions

If the agent performs 6 correct actions, 4 incorrect actions, and 5 neutral actions, what is the total reward?

Solution - Total reward = (6 * 10) + (4 * -5) + (5 * 3)

= 60 - 20 + 15 = 55

Answer - 55

Question 5 - A dataset originally has 5000 samples, each with 15 features. After applying PCA, the dataset is reduced to 5 dimensions.

- The storage size of each feature is 8 bytes.
- What is the percentage reduction in total storage after PCA?

Solution -

Before PCA: Total size = 5000 * 15 * 8 = 600000 bytes

After PCA: New size = 5000 * 5 * 8 = 200000 bytes

Reduction Percentage: Reduction = (600000 -200000 / 600000) *100

= (400000 / 600000) * 100

= 66.67%

Answer - 66.67 % reduction

Question 6 - [contextual ambiguity]
A chatbot receives the message: *"Can you book a bank appointment for me?"*

- How would it figure out if "bank" means a financial institution or a riverbank?

Answer:

- The chatbot looks at surrounding words like *"appointment"*, which suggests a financial institution.
- It can use:
    - **Word Sense Disambiguation (WSD)** – Finding the best meaning based on context.
    - **Contextual Models (like BERT)** – Understanding the full sentence before deciding.
    - **Named Entity Recognition (NER)** – Recognizing "bank" as a company rather than a place.

Question 7 - [sentiment analysis]
A system classifies *"I can't believe how amazing this terrible service is!"* as positive.

- What mistake is it making?
- How can it be fixed?

Answer:

- The model sees *"amazing"* as positive but ignores the sarcasm and *"terrible"*.
- Fixes:
    - **Negation Handling** – Teach the model to recognize phrases like *"can't believe"*.
    - **Better Context Understanding** – Use models like BERT that analyze entire sentences.
    - **Post-processing Rules** – Adjust scores when positive and negative words appear together.

Question 8 - [metaphor understanding]
Why might an NLP model misunderstand *"He is drowning in work."*?
How can it be improved?

Answer:

- A simple model might think "drowning" means water instead of being overwhelmed.
- Fixes:
    - **Train on figurative language** – Teach the model to recognize metaphors.
    - **Use pre-trained models (GPT, BERT)** – These learn meanings from different contexts.
    - **Knowledge Graphs** – Show relationships between words (e.g., "drowning" and "work" = too much work).

Question 9 - [zero-shot learning]
A model summarizes legal documents without training on legal texts. How?
What are the risks?

Answer:

- It learned from a mix of texts (news, research papers), so it recognizes patterns in legal writing.
- **Risks:**
    - **Hallucination** – It might make things up.
    - **Lack of expertise** – It doesn't truly "understand" legal terms.
    - **Bias** – If trained on biased data, it may repeat unfair conclusions.

Question 10 - [bias]
A hiring model favors male names. Why? How can it be fixed?

Answer:

- **Why?**
    - Trained on biased data where men were hired more often.
    - Some word embeddings link "doctor" with "he" more than "she".

- **Fixes:**
    - **Debias word embeddings** – Remove gender bias from words.
    - **Fairness algorithms** – Adjust decisions to reduce bias.
    - **Audit the model** – Regularly check for unfair patterns.

Question 11 - [temporal understanding]
A chatbot receives: "Let's meet next Friday" on a Wednesday. Two days later, the user says "Is our meeting tomorrow?" and the bot confirms incorrectly.

- What's the core problem in date reference tracking?
- How can temporal reasoning be enhanced?

Answer:

Core problems:

- Not maintaining reference time (original Wednesday)
- Not updating temporal references as time passes
- Failing to resolve relative time expressions ("next Friday", "tomorrow")

Improvements:

- Track conversation timestamp
- Maintain a temporal anchor for each conversation
- Update relative time references based on current date
- Use temporal reasoning system to resolve relative dates
- Create explicit timeline for each conversation

Question 12 - [semantic contradiction]
A fact-checking system analyzes: "The new policy will increase costs by reducing expenses."

- Why might basic sentiment analysis fail here?
- How can logical consistency checking be implemented?

Answer:

Basic sentiment analysis fails because:

- It focuses on individual words rather than logical relationships
- Doesn't detect mutual exclusivity ("increase" vs "reducing")
- Misses semantic contradiction

Solutions:

- Implement logical consistency checking
- Use natural language inference models
- Add contradiction detection layer
- Check for mutually exclusive term pairs

Question 13 -

A sentiment analyzer achieves these accuracy rates:

- Positive: 92%
- Neutral: 85%
- Negative: 78%
- Sarcastic: 45%

On a dataset of 10,000 customer reviews where 60% are positive, 20% neutral, 15% negative, and 5% sarcastic:

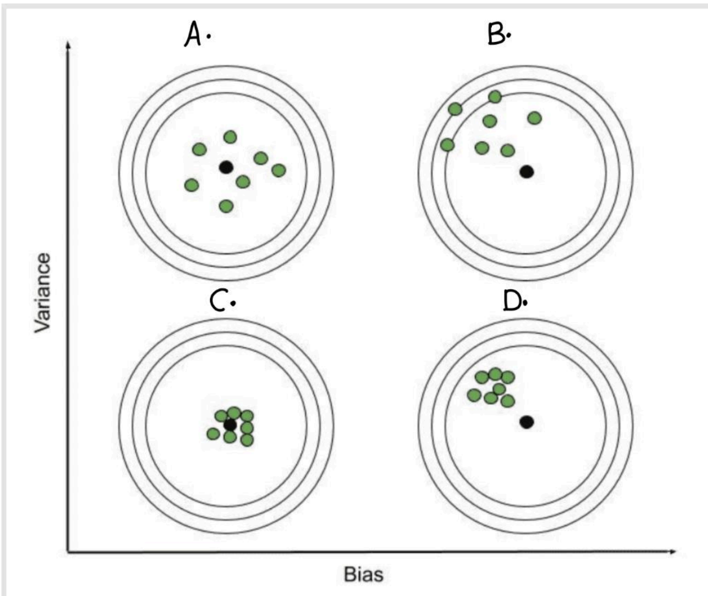
- Calculate the expected number of misclassified reviews

Answer:

Misclassified reviews calculation:

- Positive: 6000 × (1-0.92) = 480
- Neutral: 2000 × (1-0.85) = 300
- Negative: 1500 × (1-0.78) = 330
- Sarcastic: 500 × (1-0.45) = 275
- Total misclassified = 1,385 reviews

Question 14 - Given the bias variance trade off relation of 4 models. Classify models as underfitted, over fitted, Good Fit, and worst fit ?



Ans.

(D). High Bias, Low Variance: A model with high bias and low variance is said to be underfitting.

(A). High Variance, Low Bias: A model with high variance and low bias is said to be overfitting.

(B). High-Bias, High-Variance: A model has both high bias and high variance, which means that the model is not able to capture the underlying patterns in the data (high bias) and is also too sensitive to changes in the training data (high variance). As a result, the model will produce inconsistent and inaccurate predictions on average.

(C). Low Bias, Low Variance: A model that has low bias and low variance means that the model is able to capture the underlying patterns in the data (low bias) and is not too sensitive to changes in the training data (low variance). This is the ideal scenario for a machine learning model, as it is able to generalize well to new, unseen data and produce consistent and accurate predictions. But in practice, it's not possible.

Question 15. Consider a classification problem with three classes: A, B, and C. A neural network outputs the following probability distribution for a given input sample:

P(A) = 0.7, P(B) = 0.2, P(C) = 0.1

If the true class label for this sample is B, calculate the cross-entropy loss.

(a) 0.1

(b) 0.2

(c) 1.6

(d) 0.7

Ans.

$L = -\sum_{i=1}^{C} y_i \log(p_i)$

where:

- y_i is 1 if the true class is i, otherwise 0.
- p_i is the predicted probability for class i.

Step 1: Identify Given Values

- Predicted probabilities:

$$P(A) = 0.7, \quad P(B) = 0.2, \quad P(C) = 0.1$$

$$y(A) = 0, \quad y(B) = 1, \quad y(C) = 0$$

Using the formula:

$$L = -(0 \cdot \log(0.7) + 1 \cdot \log(0.2) + 0 \cdot \log(0.1))$$

Since only the term for class **B** contributes:

$$L = -\log(0.2)$$

Using logarithm (base **e**, natural log):

$$\log(0.2) \approx -1.609$$

Thus, the cross-entropy loss is:

$$L \approx 1.609$$