Question 1

Combining multiple modalities always ensures more accurate inference and expression?

A. True
B. False

Question 2

Which of the following is not a righteous feature fusion technique?

A. Early fusion
B. Late fusion
C. Decision Fusion
D. Selective Fusion

Question 3

Which of the following statements are true?

(A) The assumption behind early data fusion is the conditional independence between multiple modalities.

(B) Errors are uncorrelated during late fusion because errors from different models are handled independently.

1. A only
2. B only
3. A but not B
4. Both A and B

Question 4

Feature selection tries to eliminate features which are

A. Descriptive
B. Redundant
C. Relevant
D. Interactive

Question 5

The process of identifying and selecting the most helpful features in a dataset is known as feature selection. In the machine learning pipeline, it is an essential stage. We should be concerned about feature selection since unnecessary features could lead to

(A) Overfitting or Poor Generalization (B) Poor training speed (C) poor model explainability.

1. A only
2. C only
3. Both A and C
4. A, B and C

Question 6

We begin CFS by calculating all pairwise correlations. The feature with a higher mean absolute correlation with other features would be deleted if a pair of features' correlation exceeds a predetermined threshold.

1. True
2. False

Question 7

As opposed to treating each modality separately, affective stimuli should be classified simultaneously for all available modalities.

True

False

Question 8

Slow fusion eliminates the synchronization requirement while taking advantage of the correlations between the modalities.

True

False

Question 9

Real-time prediction from multimodal data stream requires larger window segments

True

False

Question 10

How much information each modality conveys about the expressed affect during multimodal interactions is determined by

Modality Selection

Feature Selection

Redundancy

All of the above

Question 11

What is the primary objective of combining multiple streams of data in an emotion recognition system?

A) Decreasing computational load

B) Enhancing accuracy with complementary information

C) Simplifying the analysis process

) Reducing redundancy in data

Question 12

What is crucial w.r.t. different modalities in multimodal systems?

A) Increasing the number of sensors

B) Understanding the temporal synchronization of data

C) Implementing complex fusion methods

D) Reducing the size of affective corpora

Question 13

What is the primary goal of projecting the data onto a higher-dimensional space in multimodal affect recognition?

A) To increase computational complexity

B) To reduce the amount of data

C) To enhance feature representation

D) To standardize feature representation across modalities

Question 14

Which challenges arise when performing early fusion with features of different natures or dimensions?

1. Lack of synchronization
2. Inconsistent feature extraction techniques
3. Dimensionality reduction
4. Feature normalization
A. 1 and 2
B. 3 and 4
C. 1,2 and 3
D. 1,2,3,4

Question 15

Identify the correct statements.

1. Soft fusion combines decisions using measures of confidence
2. Hard fusion uses logical operations such as AND , OR
3. Soft fusion combines feature vectors directly, while hard fusion uses statistical analysis.
A. 2 only
B. 1 and 2
C. 1 and 3
D. 1, 2 and 3

Question 17

What is the primary focus of the SEMAINE dataset?

A) Collecting textual data for dialogue systems

B) Creating virtual agents for gaming applications

C) Generating emotional responses from users through dialogue interactions

D) Analyzing physiological signals for health monitoring

Question 18

What trade-off is involved in choosing the duration of data segments for affect prediction?

A) Trade-off between capturing physical activity and conversation

B) Trade-off between real-time analysis and computation resources

C) Trade-off between hardware costs and system performance

D) Trade-off between user comfort and data accuracy

Question 19

What is the significance of considering longer windows for data segmentation in affect prediction?

A) Longer windows provide more temporal data, allowing for better prediction accuracy

B) Longer windows reduce computational resources required for analysis

C) Longer windows capture only the onset of expressions, leading to more accurate predictions

D) Longer windows result in fewer frames, making predictions less reliable

Question 20

In which scenario would the cost of an incorrect prediction be higher in a multimodal affect recognition system?

A) Suggesting movies to a user based on their emotional state

B) Monitoring students' engagement level during a class

C) Providing real-time feedback during therapy sessions

D) Analyzing sentiment in social media posts

Question 21

How many different personalities does the virtual avatar in the SEMAINE dataset have?

A) Two

B) Three

C) Four

D) Five

Question 22

What dimensions were annotated by raters in the SEMAINE dataset for each interaction?

A) Arousal, expectation, power, and valence

B) Happiness, sadness, anger, and fear

C) Positive affect, negative affect, neutrality, and intensity

D) Joy, surprise, disgust, and contempt

Question 23

What is one of the main challenges faced by participants in the EmotiW challenge?

A) Limited computational resources

B) Lack of diversity in the dataset

C) Intra-class variability due to varied environments

D) Overfitting of machine learning models

Question 24

Which of the following is a common dataset for multimodal affect recognition specifically for sensors data?

A. EmotiW
B. ASCERTAIN
C. AVEC
D. RAVDESS

Question 25

Adding and removing modalities is easier in early fusion.

A. True
B. False

---

1) Which statement best describes the expectation of a multimodal emotion recognition system?   *1 point*

- ○ Each modality is anticipated to offer redundant information.
- ◉ Each modality is expected to provide unique information.
- ○ Combining modalities is unnecessary for accurate emotion recognition.
- ○ Modality combination diminishes the effectiveness of emotion recognition systems.

2) If any two modalities convey identical emotional information, combining them may   *1 point* increase computational load without improving system performance.

- ◉ True
- ○ False

3) What is the primary challenge in early feature fusion of different modalities? *1 point*

- ◉ Synchronization issues due to differences in feature dimensions.
- ◯ Difficulties in selecting the appropriate machine learning algorithm.
- ◯ leads to an increase in computational complexity
- ◯ Lack of available datasets for feature combination experimentation.

4) Which of the following is NOT a technique used for combining outputs in late fusion? *1 point*

- ◯ Averaging
- ◉ Early fusion
- ◯ Voting
- ◯ All of the above

5) Soft late fusion involves incorporating a measure of confidence which is linked to the decisions being made. *1 point*

- ◉ True
- ◯ False

6) Having overlapping information can lead to the curse of dimensionality, particularly in scenarios like feature fusion where overlap is common. *1 point*

- ◉ True
- ◯ False

7) What assumption is made in decision-level fusion? *1 point*

- ◯ Conditional dependence between modalities
- ◉ Conditional independence between modalities
- ◯ Complete independence between modalities
- ◯ None of the above

8) Slow fusion provides us with the correlation between modalities while relaxing the requirement of synchronization. *1 point*

- ◉ True
- ◯ False

9) When performing decision-level fusion (late fusion), what potential loss of information could occur? **1 point**

- ○ Loss of computational efficiency
- ○ Loss of model interpretability
- ○ Loss of correlation between individual features
- ◉ Loss of feature diversity

No, the answer is incorrect.
Score: 0
Accepted Answers:
Loss of correlation between individual features

10) Support Vector Machine (SVM) is an example of which fusion technique? **1 point**

- ○ Early fusion
- ○ Late fusion
- ○ Feature fusion
- ◉ None of the above