

Question 1

Which of the following statements is true about fundamental frequency F_0 .

- A. It is the frequency at which vocal cords vibrate in voiced sounds.
- B. It is the musical pitch of a note that is perceived as the lowest partial present.

- 1. Only A
- 2. Only B
- 3. Both A and B
- 4. None of the above

Question 2

Which of the following are the examples of prosody features?

- A. Fundamental Frequency
- B. Formant Frequency
- C. Phoneme Rate
- D. All of the above

Question 3

Dataset is presented by IITKGP-SESC is an example of

- A. Natural Dataset
- B. Induced Dataset
- C. Simulated Dataset
- D. All of the above

Question 4

Speech-rate signifies:

- A. Loudness
- B. Pitch
- C. Voice quality
- D. Velocity of speech

Question 5

Angry speech has a higher fundamental frequency f_0

- A. True
- B. False

Question 6

In iterative feature normalization observations that are labeled as neutral are used to re-estimate the normalization parameters.

- A. True
- B. False

Question 7

Identical Feature Normalization can adversely affect the emotion discrimination of the feature.

- A. True
- B. False

Question 8

Which of the following statements are true.

(A) Positive voices are generally loud with considerable variability in loudness, have high and variable pitch, and are high in the first two formant frequencies

(B) Variations in pitch show differences between high arousal emotions (joy) and low arousal emotions (tenderness), when compared with neutral vocalizations.

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

Question 9

To account for variations brought on by the speaker's vocal qualities, speaker normalization is applied.

- A. True
- B. False

Question 10

Women's voice have a lower fundamental frequency than men's voice?

- A. True
- B. False

Question 11

What is the primary motivation for pursuing voice as a primary modality in affective computing, ~~according to the text?~~

- A. Facial expressions are unreliable. *NO, FE are not unreliable*
- B. Voice analysis can provide valuable information about emotions even without observing facial expressions or gestures. *True*
- C. Voice analysis is less intrusive than facial expression analysis. *False*
- D. Voice analysis is more accurate than other modalities.

Question 12

Which scenario illustrates the use of affective computing in indexing videos?

- A. Analyzing facial expressions of actors in movies.
- B. Identifying keywords in movie scripts.
- C. Searching for videos based on emotional content, such as happy events.
- D. Creating subtitles for movies in different languages.

Question 13

How is affective computing utilized in automatic translation systems?

- A) By translating between languages by ignoring emotional nuances.
- B) By analyzing the emotional and behavioral aspects of communication to enhance translation accuracy.
- C) By focusing on linguistic patterns to ensure accurate translation.
- D) All of the above

Question 14

What are the three major factors identified by Borden and others as challenges in understanding emotion through voice?

- A) The speaker's accent, pitch, and volume.
- B) The speaker's age, gender, and cultural background.
- C) The content of the speech, the paralinguistic information, and the speaker's attributes.
- D) The speaker's emotional state, tone of voice, and body language.

Question 15

Which of the following best describes the challenge of "how it is said" in understanding emotion through voice?

- A) Analyzing the speaker's accent and pronunciation.
- B) Understanding the speaker's tone and manner of speech.
- C) Recognizing the linguistic origin of the words spoken.
- D) Identifying the speaker's age and gender based on voice characteristics.

Question 16

Which category of voice-based data involves inducing emotions in participants through stimuli, such as videos, and then capturing their speech responses?

- A) Natural data
- B) Simulated data
- C) Scripted data
- D) Elicited emotion data


Question 17

Which dataset consists of recordings of 10 German sentences spoken in different emotions by actors?

- A) AIBO Database
- B) Berlin Database of Emotional Speech
- C) Ryerson Audio Visual Database of Emotional Speech and Song
- D) None of the above

Question 18

Which of the following is not a major challenge in voice-based affect analysis?

1. Limited availability of emotional speech synthesis
2.  Lesser accuracy than FER systems
3. Scalability of datasets
4. Cross dataset performance across different languages

Question 19

Which features are commonly used for automatic feature extraction in speech analysis?

- A. Syntax and semantics features
- B. Prosody features and spectral characteristics
- C. Phonetics and phonology features
- D. Pragmatics and discourse features

Question 20

Which library can be used to extract MFCCs from audio signals?

- A) TensorFlow
- B) Scikit-learn
- C) Librosa
- D) NumPy

Question 21

Which feature is commonly used to measure the energy in an input acoustic signal?

- A. Loudness
- B. Fundamental frequency
- C. Formant frequencies
- D. Speech rate

Question 22

Which attribute is commonly associated with positive voices in speech analysis?

- A) Low loudness and consistent pitch
- B) High loudness and variable pitch
- C) Low pitch and variable loudness
- D) High pitch and consistent loudness

Question 23

Why is speaker normalization applied in speech-based emotion recognition?

- A) To standardize the emotional content of speech across different speakers
- B) To enhance the differences between angry and neutral speech
- C) To accommodate variations in speech features introduced by differences in speakers' vocal qualities
- D) To eliminate differences in emotion expression caused by speaker variability

Question 24

What technique is commonly used for speech-based emotion recognition?

- A) Hidden Markov Models (HMMs) → ML
- ~~B) Spectrograms~~ → ML
- C) Support Vector Machines (SVMs)
- D) 1-D CNN → DL

Question 25

What is a major concern regarding the real-time applicability of speech-based affect recognition systems?

- A) Lack of robust algorithms for feature extraction
- B) Variability in emotional expression across different individuals
- ~~C) Computational complexity and potential exhaustion of resources~~
- D) Limited availability of labeled emotional speech datasets

1) What is one of the primary challenges in voice-based affect analysis across different languages?

- ☐ Limited availability of recording devices
- ☐ Difficulty in understanding speech patterns
- ☒ Cross-dataset performance variations
- ☐ Lack of computational resources

2) What does MFCC stand for in the context of audio signal processing?

- ☐ Multifunctional Frequency Cepstral Coefficients
- ☒ Mel-Frequency Cepstral Coefficients
- ☐ Modulated Frequency Cepstral Coefficients
- ☐ Multi-Feature Cepstral Coefficients

3) What is a spectrogram?

- ☐ A plot of amplitude against time
- ☒ A visual representation of frequency content over time
- ☐ A measure of signal power over time
- ☐ A representation of phase information in the frequency domain

4) In the context of operator safety in complex heavy machinery operation, how can voice analysis be utilized?

- ☐ To control the speed of the machinery
- ☐ To monitor the environmental conditions
- ☒ To assess the state of the driver/operator
- ☐ To detect physical defects in the machinery

5) Voice analysis is not as effective as facial expressions in determining emotional states in affective computing.

- ☐ True
- ☒ False

6) Considering meta-information about a user's background and culture is unnecessary for understanding emotions through voice in affective computing.

- ☐ True
- ☒ False

7) What is the concern regarding simulated or acted data in emotion analysis datasets?

- ☐ It provides the most accurate representation of real-world user behavior
- ☒ Not all examples captured may reflect genuine user behavior
- ☐ It is easier to collect compared to other types of data
- ☐ Simulated data is less prone to biases

8) Confirmation bias is the tendency to interpret information in a way that confirms one's preexisting beliefs.

- ☒ True
- ☐ False

9) What is the purpose of conducting statistical analysis on labels generated by multiple labelers for the same samples in a database?

- ☐ To simplify the labeling process for future samples
- ☒ To identify inconsistencies and discrepancies in the labels
- ☐ To increase the complexity of the dataset
- ☐ To exhaust the research grant funds

10) In the context of feature selection for emotion analysis, what is the suggested approach to achieve balance between different attributes of the signal?

- ☐ Using a single feature for both arousal and valence
- ☐ Selecting features randomly for analysis
- ☒ Applying fusion techniques to combine multiple features
- ☐ Ignoring feature selection and relying solely on raw data