**CHƯƠNG 5: TEXT**

**📝 I. Kiến thức cơ bản về dữ liệu văn bản**

**1. Văn bản là gì trong xử lý ngôn ngữ tự nhiên?**

A. Hình ảnh chứa chữ  
B. Chuỗi ký hiệu mang ý nghĩa thông tin  
C. Mã nhị phân của ngôn ngữ máy  
D. Tín hiệu âm thanh dài

👉 **Đáp án:** B

**2. Dạng mã hóa văn bản phổ biến hiện nay là?**

A. UTF-8  
B. ASCII 5-bit  
C. ISO-9999  
D. JPEG

👉 **Đáp án:** A

**3. Trong xử lý văn bản, tokenization là gì?**

A. Nén văn bản  
B. Xác định ngôn ngữ  
C. Tách văn bản thành từ và câu  
D. Gán nhãn từ loại

👉 **Đáp án:** C

**📝 II. Mã hóa & biểu diễn văn bản**

**4. ASCII đại diện cho bao nhiêu ký tự chuẩn?**

A. 64  
B. 95  
C. 128  
D. 256

👉 **Đáp án:** C

**5. Unicode hỗ trợ bao nhiêu loại hệ thống chữ viết?**

A. Chỉ Latin  
B. Khoảng 50  
C. Khoảng 100  
D. Hơn 150

👉 **Đáp án:** D

**📝 III. Chỉ số hóa văn bản**

**6. Phương pháp Bag of Words không xem xét yếu tố nào?**

A. Từ vựng  
B. Tần suất từ  
C. Vị trí từ  
D. Có/không xuất hiện

👉 **Đáp án:** C

**7. TF-IDF dùng để:**

A. Nén văn bản  
B. Đo mức độ quan trọng của từ  
C. Dự đoán câu tiếp theo  
D. Dịch văn bản

👉 **Đáp án:** B

**8. Trong mô hình TF-IDF, nếu một từ xuất hiện nhiều trong toàn bộ văn bản thì:**

A. Có IDF cao  
B. Có TF thấp  
C. Có IDF thấp  
D. Không ảnh hưởng đến kết quả

👉 **Đáp án:** C

**📝 IV. Truy vấn và truy xuất thông tin**

**9. Truy vấn Boolean là:**

A. So sánh số lượng từ  
B. Truy vấn có AND, OR, NOT  
C. Truy vấn học sâu  
D. Truy vấn theo xác suất

👉 **Đáp án:** B

**10. Khoảng cách cosin được dùng để:**

A. Đo độ tương đồng giữa 2 văn bản  
B. Mã hóa văn bản  
C. Xác định ngôn ngữ  
D. Nén dữ liệu

👉 **Đáp án:** A

**📝 V. Đánh giá hệ thống phân loại**

**11. Công thức độ chính xác (Accuracy) là:**

A. TP / (TP + FP)  
B. TP / (TP + FN)  
C. (TP + TN) / (TP + TN + FP + FN)  
D. FN / (FN + TP)

👉 **Đáp án:** C

**12. F1-score là:**

A. Trung bình cộng của precision và recall  
B. Trung bình nhân của precision và recall  
C. Trung bình điều hòa của precision và recall  
D. Trung bình số học của precision và recall

👉 **Đáp án:** C

**📝 VI. Mô hình ngôn ngữ & n-gram**

**13. Bigram là:**

A. Một từ đơn  
B. Hai từ liên tiếp  
C. Hai đoạn văn  
D. Một cụm danh từ

👉 **Đáp án:** B

**14. Mô hình ngôn ngữ n-gram được dùng để:**

A. Dự đoán câu trả lời  
B. Xác suất câu hoặc từ tiếp theo  
C. Xóa từ dừng  
D. Trích xuất đặc trưng âm thanh

👉 **Đáp án:** B

**📝 VII. Học máy trong xử lý văn bản**

**15. Học có giám sát dùng để:**

A. Phân nhóm văn bản  
B. Gán nhãn văn bản  
C. Tự học đặc trưng  
D. Tạo từ điển

👉 **Đáp án:** B

**16. Mô hình nào sau đây thường được dùng để phân loại văn bản?**

A. K-means  
B. SVM  
C. PCA  
D. CNN cho ảnh

👉 **Đáp án:** B

**🧹 I. Tiền xử lý văn bản (Text Preprocessing)**

**17. Mục đích của xử lý từ dừng (stop-word removal) là:**

A. Xóa từ quan trọng  
B. Giữ lại các từ khóa chính  
C. Loại bỏ từ không mang nhiều thông tin  
D. Dịch ngôn ngữ

👉 **Đáp án:** C

**18. Stemming là:**

A. Dịch nghĩa từ  
B. Tìm từ đồng nghĩa  
C. Cắt từ về gốc từ thô  
D. Xác định danh từ

👉 **Đáp án:** C

**19. Lemmatization khác stemming ở điểm nào?**

A. Không cần từ điển  
B. Cho ra gốc từ có nghĩa đúng  
C. Không cần xử lý ngữ pháp  
D. Không thay đổi từ gốc

👉 **Đáp án:** B

**📊 II. Biểu diễn văn bản**

**20. One-hot encoding có đặc điểm:**

A. Gọn nhẹ  
B. Mã hóa theo tần suất  
C. Vector rất thưa và dài  
D. Có ý nghĩa ngữ nghĩa

👉 **Đáp án:** C

**21. Word Embedding (như Word2Vec) giúp:**

A. Mã hóa ký tự  
B. Gán nhãn văn bản  
C. Biểu diễn từ dưới dạng vector gần nhau nếu nghĩa gần  
D. Xác định phần loại từ

👉 **Đáp án:** C

**🧠 III. Truy vấn và đánh giá**

**22. Precision là:**

A. Tỷ lệ dự đoán đúng trên tổng dự đoán đúng  
B. Tỷ lệ dự đoán đúng trên tổng dự đoán sai  
C. Tổng đúng chia tổng mẫu  
D. Xác suất sai

👉 **Đáp án:** A

**23. Recall là:**

A. Tỷ lệ tìm được tất cả mẫu đúng  
B. Tỷ lệ sai sót  
C. Tỷ lệ nhãn giả  
D. Tổng kết quả

👉 **Đáp án:** A

**🔁 IV. Mô hình xác suất & n-gram**

**24. Với mô hình bigram, xác suất câu “I love NLP” là:**

A. P(I)+P(love∣I)+P(NLP∣love)P(I) + P(love|I) + P(NLP|love)P(I)+P(love∣I)+P(NLP∣love)  
B. P(I)⋅P(love∣I)⋅P(NLP∣love)P(I) \cdot P(love|I) \cdot P(NLP|love)P(I)⋅P(love∣I)⋅P(NLP∣love)  
C. P(NLP)/P(I)P(NLP) / P(I)P(NLP)/P(I)  
D. Không xác định

👉 **Đáp án:** B

**25. Một nhược điểm lớn của n-gram là:**

A. Không xử lý được tiếng Anh  
B. Không áp dụng được cho máy học  
C. Bùng nổ dữ liệu với n lớn  
D. Chỉ dùng được với âm thanh

👉 **Đáp án:** C

**26. Mô hình n-gram hoạt động dựa trên giả định:**

A. Độc lập thời gian  
B. Dữ liệu độc lập hoàn toàn  
C. Mỗi từ chỉ phụ thuộc n-1 từ trước đó  
D. Mỗi từ phụ thuộc toàn bộ câu

👉 **Đáp án:** C

**✍️ V. Một số thuật ngữ và ứng dụng khác (tiếp)**

**27. Trong biểu diễn TF-IDF, nếu một từ xuất hiện trong hầu hết các tài liệu, thì:**

A. TF cao, IDF cao  
B. TF thấp, IDF cao  
C. TF cao, IDF thấp  
D. TF thấp, IDF thấp

👉 **Đáp án:** C

**28. Từ “the”, “is”, “of” thường được xem là:**

A. Stop words  
B. Root words  
C. Rare words  
D. Named entities

👉 **Đáp án:** A

**29. Lý do chính sử dụng stemming:**

A. Để rút gọn từ và giảm không gian từ vựng  
B. Để tạo từ điển  
C. Để phân biệt từ đồng nghĩa  
D. Để đánh trọng số cho từ

👉 **Đáp án:** A

**30. Một mô hình tìm kiếm văn bản hiệu quả cần đảm bảo:**

A. Lưu trữ nhanh nhất  
B. Truy xuất đầy đủ và chính xác  
C. Chỉ dựa vào tiêu đề văn bản  
D. Không cần xử lý tiền văn bản

👉 **Đáp án:** B

**31. Mô hình ngôn ngữ được dùng để:**

A. Dự đoán ký tự tiếp theo  
B. Mã hóa văn bản thành ảnh  
C. Dịch văn bản sang hình ảnh  
D. Tăng tốc xử lý hình ảnh

👉 **Đáp án:** A

**32. Biểu diễn văn bản dưới dạng vector giúp:**

A. Dễ nén hơn  
B. Dễ truyền qua mạng  
C. Phân tích bằng mô hình học máy  
D. Dịch văn bản nhanh hơn

👉 **Đáp án:** C

**33. Cosine similarity trong truy vấn văn bản có giá trị từ:**

A. -∞ đến +∞  
B. 0 đến 100  
C. 0 đến 1  
D. -1 đến 1

👉 **Đáp án:** D

**34. Ưu điểm của biểu diễn bằng Word Embedding (so với one-hot):**

A. Đơn giản hơn  
B. Có thể học ý nghĩa từ và quan hệ giữa từ  
C. Dễ lưu hơn  
D. Chỉ dùng được cho tiếng Anh

👉 **Đáp án:** B

**35. Dạng biểu diễn Word2Vec nào sử dụng mục tiêu đoán từ xung quanh từ trung tâm?**

A. CBOW  
B. Skip-gram  
C. BERT  
D. FastText

👉 **Đáp án:** A

**36. Dạng biểu diễn nào đoán từ trung tâm từ các từ xung quanh?**

A. CBOW  
B. Skip-gram  
C. LSTM  
D. GPT

👉 **Đáp án:** B

**37. LSTM, GRU, BERT, GPT thường được dùng cho:**

A. Nén âm thanh  
B. Tăng độ phân giải ảnh  
C. Xử lý văn bản theo chuỗi  
D. Tìm đường đi ngắn nhất

👉 **Đáp án:** C

**CHƯƠNG 7: SPEECH**

**🔊 I. Khái niệm tiếng nói và đặc điểm âm thanh**

**1. Âm vị (phoneme) là gì?**

A. Ký tự trong bảng chữ cái  
B. Tín hiệu âm thanh dài  
C. Đơn vị âm giúp phân biệt từ  
D. Dạng sóng ngẫu nhiên  
👉 **Đáp án:** C

**2. Phân biệt âm hữu thanh và vô thanh dựa vào:**

A. Tần số lấy mẫu  
B. Sự rung của dây thanh âm  
C. Ngữ nghĩa câu nói  
D. Độ dài sóng âm  
👉 **Đáp án:** B

**3. Nguyên âm thường có đặc điểm:**

A. Vô thanh, ngắn và yếu  
B. Có sự cản trở luồng khí  
C. Hữu thanh, kéo dài và mạnh  
D. Không ảnh hưởng đến thanh điệu  
👉 **Đáp án:** C

**🔤 II. Bảng mã phiên âm và đặc điểm tiếng Việt**

**4. Mục đích của IPA là gì?**

A. Mã hóa văn bản  
B. Phiên âm âm thanh cho mọi ngôn ngữ  
C. Dịch tự động  
D. Lưu trữ tín hiệu giọng nói  
👉 **Đáp án:** B

**5. Tiếng Việt khác tiếng Anh vì:**

A. Có phụ âm nhiều hơn  
B. Không dùng nguyên âm  
C. Có thanh điệu ảnh hưởng đến nghĩa từ  
D. Chỉ có âm vô thanh  
👉 **Đáp án:** C

**📈 III. Tín hiệu tiếng nói và phân tích phổ**

**6. Tần số cơ bản F₀ là:**

A. Số lượng nguyên âm trong 1 giây  
B. Tần số rung dây thanh quản  
C. Biên độ trung bình của âm thanh  
D. Tần số cao nhất có thể nghe  
👉 **Đáp án:** B

**7. Công thức tính F₀ là:**

A. F₀ = số đỉnh / thời gian  
B. F₀ = năng lượng / độ dài tín hiệu  
C. F₀ = log(âm lượng)  
D. F₀ = khoảng cách hai từ  
👉 **Đáp án:** A

**8. Formant (F1, F2...) phản ánh:**

A. Cường độ âm thanh  
B. Tần số lặp lại  
C. Tần số cộng hưởng trong bộ phát âm  
D. Độ dài tín hiệu  
👉 **Đáp án:** C

**🔉 IV. Các phương pháp xác định F₀**

**9. Phương pháp nào sau đây dùng để tìm F₀?**

A. Autocorrelation  
B. TF-IDF  
C. Histogram  
D. N-gram  
👉 **Đáp án:** A

**10. AMDF là gì?**

A. Tính năng lượng trung bình  
B. Đo sự khác biệt tuyệt đối giữa tín hiệu và bản sao trễ  
C. Đo thời lượng âm thanh  
D. Là công cụ ghi âm  
👉 **Đáp án:** B

**🔊 V. Nhận dạng tiếng nói và mô hình hóa**

**11. ASR là viết tắt của:**

A. Audio Signal Rate  
B. Automatic Speech Recognition  
C. Auto Sound Recording  
D. Acoustic Standard Recognition  
👉 **Đáp án:** B

**12. Mục tiêu của ASR là:**

A. Chuyển văn bản thành giọng nói  
B. Nhận diện người  
C. Chuyển giọng nói thành văn bản  
D. Đánh giá cao độ  
👉 **Đáp án:** C

**13. Mô hình phổ biến trong ASR là:**

A. CNN  
B. HMM-GMM  
C. Word2Vec  
D. SIFT  
👉 **Đáp án:** B

**🧠 VI. Mô hình xác suất và giải thuật**

**14. Mô hình kênh truyền sử dụng công thức:**

W^=arg⁡max⁡WP(O∣W)⋅P(W)\hat{W} = \arg\max\_{W} P(O|W) \cdot P(W)W^=argWmax​P(O∣W)⋅P(W)

Đây là áp dụng của:  
A. Luật Newton  
B. Mạng nơ-ron  
C. Định lý Bayes  
D. Thống kê mô tả  
👉 **Đáp án:** C

**15. WER là gì?**

A. Tỷ lệ từ đúng trên tổng từ dự đoán  
B. Tỷ lệ lỗi từ trong nhận dạng  
C. Số câu sai trong bài đọc  
D. Mức độ rung dây thanh âm  
👉 **Đáp án:** B

**16. Công thức WER là:**

A. (S + D + I) / N  
B. S / (S + D + I)  
C. (TP + TN) / N  
D. I / (S + D)  
👉 **Đáp án:** A  
(S = Thay thế, D = Xóa, I = Thêm, N = Số từ đúng)

**🛠 VII. Các công cụ và ứng dụng**

**17. Phần mềm nào sau đây là công cụ nhận dạng tiếng nói?**

A. Audacity  
B. Praat  
C. Kaldi  
D. VS Code  
👉 **Đáp án:** C

**18. MFCC là gì?**

A. Một dạng tần số lấy mẫu  
B. Đặc trưng âm học mô phỏng thính giác con người  
C. Ký hiệu trong IPA  
D. Loại âm vô thanh  
👉 **Đáp án:** B

Nếu bạn muốn mình:

* ✅ Tạo bản PDF tổng hợp các câu trắc nghiệm này kèm đáp án
* ✅ Sinh thêm bản **không có đáp án để luyện tập**
* ✅ Sinh **bộ đề trộn ngẫu nhiên** cho luyện thi
* **✅ VIII. Âm tiết, âm vị và cấu trúc âm học**
* **19. Âm vị là:**
* A. Đơn vị mang nghĩa trong từ  
  B. Đơn vị âm thanh nhỏ nhất có thể phân biệt từ  
  C. Tập hợp ký tự Latin  
  D. Tên khác của âm tiết  
  👉 **Đáp án:** B
* **20. Âm tiết là gì?**
* A. Một câu nói  
  B. Một cụm từ chứa ngữ điệu  
  C. Một tổ hợp phát âm tự nhiên trong lời nói  
  D. Một đơn vị viết tắt của chữ  
  👉 **Đáp án:** C
* **✅ IX. Phân loại âm trong tiếng nói**
* **21. Phụ âm được tạo ra bằng:**
* A. Rung dây thanh âm  
  B. Cản trở dòng khí  
  C. Mở rộng họng  
  D. Phát âm không tiếng  
  👉 **Đáp án:** B
* **22. Các ví dụ về âm hữu thanh là:**
* A. [s], [f], [k]  
  B. [z], [v], [d]  
  C. [h], [t], [p]  
  D. [ch], [sh], [th]  
  👉 **Đáp án:** B
* **✅ X. Biểu diễn tín hiệu tiếng nói**
* **23. Phổ tiếng nói được tạo ra bằng:**
* A. Fourier Transform  
  B. Convolution  
  C. Nén tín hiệu  
  D. Hàm sigmoid  
  👉 **Đáp án:** A
* **24. Spectrogram cho biết:**
* A. Cao độ và tốc độ nói  
  B. Biên độ, tần số và thời gian của tín hiệu  
  C. Tần số và sóng âm  
  D. Nhịp điệu và ngữ nghĩa  
  👉 **Đáp án:** B
* **25. Đặc điểm phổ của âm điều hòa là:**
* A. Có nhiều thành phần tần số rải rác ngẫu nhiên  
  B. Có các đỉnh tần số cách đều  
  C. Chỉ có 1 tần số  
  D. Không có tần số cụ thể  
  👉 **Đáp án:** B
* **26. Đặc trưng nổi bật để phân biệt nguyên âm là:**
* A. Năng lượng trung bình  
  B. Số từ mỗi giây  
  C. Các formant (F1, F2, F3)  
  D. Tần số lấy mẫu  
  👉 **Đáp án:** C
* **✅ XI. Đặc trưng âm học & MFCC**
* **27. MFCC viết tắt của:**
* A. Mel Frequency Cepstral Coefficients  
  B. Multi Fourier Cepstrum Conversion  
  C. Mel Filter Cepstral Coefficients  
  D. Mel Frame Coding Format  
  👉 **Đáp án:** A
* **28. MFCC mô phỏng đặc tính nghe của:**
* A. Loa điện tử  
  B. Tâm lý học  
  C. Hệ thống thính giác con người  
  D. Sóng não  
  👉 **Đáp án:** C
* **✅ XII. Ứng dụng và đánh giá hệ thống nhận dạng tiếng nói**
* **29. Hệ thống TTS là gì?**
* A. Nhận dạng tiếng nói  
  B. Dịch máy  
  C. Chuyển văn bản thành tiếng nói  
  D. Tìm kiếm văn bản  
  👉 **Đáp án:** C
* **30. Khi đánh giá hệ thống nhận dạng, chỉ số nào quan trọng?**
* A. BLEU  
  B. WER (Word Error Rate)  
  C. PSNR  
  D. F1-score  
  👉 **Đáp án:** B
* **31. WER tính dựa trên:**
* A. Khoảng cách Euclid  
  B. Số từ sai chia tổng từ đúng  
  C. Số từ thêm, xóa, thay chia tổng từ tham chiếu  
  D. Độ trùng khớp ký tự  
  👉 **Đáp án:** C
* **32. Để giảm WER, ta nên:**
* A. Tăng kích thước từ điển  
  B. Dùng mạng nơ-ron sâu hơn  
  C. Tăng số lượng lớp MFCC  
  D. Dùng từ điển ngữ nghĩa  
  👉 **Đáp án:** B

**CHƯƠNG 6: SOUND**

**✅ I. Kiến thức cơ bản về âm thanh**

**1. Âm thanh là gì?**

A. Sóng ánh sáng truyền trong không khí  
B. Rung động truyền dưới dạng sóng áp suất có thể nghe được  
C. Tín hiệu điện áp xoay chiều  
D. Mã nhị phân số học  
👉 **Đáp án:** B

**2. Sóng âm có thể truyền qua môi trường nào?**

A. Chỉ không khí  
B. Chỉ chất lỏng  
C. Bất kỳ chất khí, lỏng hoặc rắn nào  
D. Chỉ chân không  
👉 **Đáp án:** C

**✅ II. Biểu diễn âm thanh**

**3. Biên độ âm thanh trong miền thời gian đo bằng gì?**

A. Watt  
B. Decibel (dB)  
C. Hertz (Hz)  
D. Bit  
👉 **Đáp án:** B

**4. Công thức SPL (Sound Pressure Level) là:**

A. 10log⁡10(P1/P0)10 \log\_{10}(P\_1 / P\_0)10log10​(P1​/P0​)  
B. 20log⁡10(P1/P0)20 \log\_{10}(P\_1 / P\_0)20log10​(P1​/P0​)  
C. log⁡(P1+P0)\log(P\_1 + P\_0)log(P1​+P0​)  
D. 2P1+P02P\_1 + P\_02P1​+P0​  
👉 **Đáp án:** B

**✅ III. Đặc trưng âm thanh trong miền thời gian**

**5. Công thức tính năng lượng trung bình là:**

A. 1N∑x(n)\frac{1}{N} \sum x(n)N1​∑x(n)  
B. ∑∣x(n)∣\sum |x(n)|∑∣x(n)∣  
C. 1N∑x(n)2\frac{1}{N} \sum x(n)^2N1​∑x(n)2  
D. x(n)\sqrt{x(n)}x(n)​  
👉 **Đáp án:** C

**6. Tốc độ vượt qua 0 (Zero Crossing Rate) dùng để:**

A. Tính tần số trung bình  
B. Đo sự thay đổi dấu của tín hiệu  
C. Tính biên độ tối đa  
D. Tính độ dài tín hiệu  
👉 **Đáp án:** B

**✅ IV. Biểu diễn trong miền tần số**

**7. Biến đổi Fourier dùng để:**

A. Phân đoạn tín hiệu  
B. Nén âm thanh  
C. Phân rã tín hiệu thành các thành phần tần số  
D. Tách giọng nói khỏi tiếng ồn  
👉 **Đáp án:** C

**8. Phổ (spectrum) biểu diễn:**

A. Độ dài tín hiệu  
B. Biên độ theo thời gian  
C. Phân bố năng lượng theo tần số  
D. Tốc độ lấy mẫu  
👉 **Đáp án:** C

**✅ V. Ảnh phổ (Spectrogram)**

**9. Spectrogram thể hiện:**

A. Quan hệ giữa tần số, thời gian và biên độ  
B. Tần số trung bình  
C. Độ dài âm thanh  
D. Các tệp âm thanh  
👉 **Đáp án:** A

**10. Trong spectrogram, cường độ mạnh được biểu diễn bằng:**

A. Màu nhạt  
B. Màu xám đậm hoặc đỏ  
C. Màu xanh  
D. Dạng sóng nhọn  
👉 **Đáp án:** B

**✅ VI. Đặc trưng trong miền tần số**

**11. Tín hiệu điều hòa có đặc điểm:**

A. Không có tần số xác định  
B. Các thành phần phổ là bội số nguyên của tần số thấp nhất  
C. Rất nhiễu  
D. Ngẫu nhiên  
👉 **Đáp án:** B

**12. Băng thông là:**

A. Số bit để mã hóa âm thanh  
B. Số lượng mẫu trong một giây  
C. Dải tần số mà tín hiệu âm thanh chiếm  
D. Dung lượng tệp âm thanh  
👉 **Đáp án:** C

**13. Âm sắc (Timbre) giúp:**

A. Xác định cao độ  
B. Phân biệt hai âm thanh cùng cao độ và cường độ  
C. Xác định âm lượng  
D. Phân tích ảnh  
👉 **Đáp án:** B

**✅ VII. Số hóa âm thanh**

**14. Tần số lấy mẫu phải tuân theo định lý nào?**

A. Newton  
B. Fourier  
C. Shannon  
D. Gauss  
👉 **Đáp án:** C

**15. Định lý Shannon yêu cầu:**

A. Fs>2fmaxF\_s > 2f\_{\text{max}}Fs​>2fmax​  
B. Fs=fmaxF\_s = f\_{\text{max}}Fs​=fmax​  
C. Fs<2fminF\_s < 2f\_{\text{min}}Fs​<2fmin​  
D. Fs=fminF\_s = f\_{\text{min}}Fs​=fmin​  
👉 **Đáp án:** A

**16. Với 16 bit, mẫu âm thanh có thể biểu diễn bao nhiêu mức biên độ?**

A. 256  
B. 1024  
C. 32768  
D. 65536  
👉 **Đáp án:** D (từ -32768 đến +32767)

**✅ VIII. Lưu trữ và tìm kiếm âm thanh**

**17. MIDI là gì?**

A. Dạng âm thanh nén  
B. Tệp mẫu nhạc dạng cấu trúc  
C. Dạng nén âm thanh MP3  
D. Bộ mã hóa tín hiệu  
👉 **Đáp án:** B

**18. Hệ thống tìm kiếm âm thanh theo mẫu cần:**

A. Tín hiệu gốc  
B. Đặc trưng như MFCC, độ cao, băng thông  
C. Bản ghi nén  
D. Mã hóa nốt nhạc  
👉 **Đáp án:** B

**19. MPEG-4 Structured Audio dùng để:**

A. Mã hóa ảnh  
B. Truyền video  
C. Mã hóa âm thanh cấu trúc (dựa trên thuật toán)  
D. Định dạng tệp PDF  
👉 **Đáp án:** C

**✅ IX. Công cụ xử lý âm thanh**

**20. Phần mềm phổ biến để xử lý và thu âm là:**

A. Excel  
B. Audacity  
C. Chrome  
D. Matlab  
👉 **Đáp án:** B

**21. Praat thường dùng để:**

A. Dự đoán tiếng nói  
B. Tách nguồn tín hiệu  
C. Phân tích, hiển thị phổ, trích đặc trưng tiếng nói  
D. Mã hóa video  
👉 **Đáp án:** C

**✅ X. Dạng tín hiệu & kỹ thuật số hóa âm thanh**

**22. Tín hiệu tương tự là:**

A. Dữ liệu dạng nhị phân  
B. Dữ liệu số rời rạc  
C. Dữ liệu biến đổi liên tục theo thời gian  
D. Tín hiệu bị mã hóa  
👉 **Đáp án:** C

**23. Lượng tử hóa là:**

A. Tách tín hiệu theo thời gian  
B. Chuyển tín hiệu về miền tần số  
C. Gán giá trị rời rạc cho biên độ tín hiệu  
D. Đảo ngược tín hiệu  
👉 **Đáp án:** C

**✅ XI. Windowing trong biến đổi Fourier**

**24. Mục đích của windowing là:**

A. Cắt tín hiệu thành các đoạn nhỏ để biến đổi  
B. Nén dữ liệu  
C. Xoay phổ tần số  
D. Giảm độ chính xác  
👉 **Đáp án:** A

**25. Khi áp dụng biến đổi Fourier, tín hiệu được giả định là:**

A. Không tuần hoàn  
B. Tuần hoàn trong cửa sổ thời gian  
C. Bị chập nhiễu  
D. Không thể phân tích  
👉 **Đáp án:** B

**✅ XII. So sánh âm nhạc và tiếng nói**

**26. Âm nhạc khác tiếng nói ở điểm nào?**

A. Tần số thấp hơn  
B. Không có nhịp điệu  
C. Có phổ điều hòa và băng thông rộng hơn  
D. Không thể số hóa  
👉 **Đáp án:** C

**27. Tín hiệu âm nhạc thường có:**

A. Điều hòa yếu và ngắn hạn  
B. Thành phần tần số rải rác  
C. Băng thông cao, phổ đều và nhiều harmonics  
D. Dạng sóng rối loạn  
👉 **Đáp án:** C

**✅ XIII. Hiệu ứng âm thanh & phần mềm thực hành**

**28. Hiệu ứng trong Audacity bao gồm:**

A. Xoay ảnh  
B. Giảm nhiễu, thay đổi cao độ, hồi âm  
C. Tăng độ phân giải  
D. Lưu dưới dạng video  
👉 **Đáp án:** B

**29. Trong thực hành với Praat, ta có thể:**

A. Chèn văn bản vào ảnh  
B. Trích xuất F₀, vẽ spectrogram, gán nhãn âm thanh  
C. Dịch file PDF  
D. Phân loại video  
👉 **Đáp án:** B

**30. Khi nhìn spectrogram, các đường ngang cách đều là dấu hiệu của:**

A. Nhiễu ngẫu nhiên  
B. Âm thanh không điều hòa  
C. Âm thanh điều hòa (harmonic)  
D. Tín hiệu bị lỗi  
👉 **Đáp án:** C

**CHƯƠNG 8**

**📘 I. Khái niệm & Tổng quan**

1. Phân tích cảm xúc đa phương thức sử dụng những loại dữ liệu nào?
   * A. Hình ảnh
   * B. Văn bản
   * C. Âm thanh
   * D. Cả 3 loại trên ✅
2. Mục tiêu chính của khai thác thông tin đa phương thức là:
   * A. Giảm kích thước dữ liệu
   * B. Tăng độ chính xác nhờ tổng hợp thông tin từ nhiều nguồn ✅
   * C. Tăng tốc độ truyền dữ liệu
   * D. Hạn chế sử dụng mạng nơ-ron
3. Khó khăn trong xử lý đa phương thức **KHÔNG BAO GỒM**:
   * A. Khác biệt định dạng
   * B. Đồng bộ thời gian
   * C. Thiếu thuật toán học máy
   * D. Mức độ tương quan khác nhau ✅

**📘 II. Mức kết hợp thông tin**

1. Kết hợp ở mức đặc trưng còn được gọi là:
   * A. Kết hợp muộn
   * B. Kết hợp lai
   * C. Kết hợp sớm ✅
   * D. Kết hợp tuyến tính
2. Kết hợp mức quyết định hoạt động chủ yếu ở:
   * A. Mức đặc trưng thô
   * B. Mức vector từ nhúng
   * C. Mức ngữ nghĩa ✅
   * D. Mức âm học
3. Ưu điểm **chính** của kết hợp mức quyết định là:
   * A. Khó mở rộng đầu vào
   * B. Dễ áp dụng các bộ phân tích khác nhau ✅
   * C. Độ chính xác thấp
   * D. Cần huấn luyện lại hệ thống

**📘 III. Phương pháp kết hợp**

1. Phương pháp kết hợp dựa trên luật thường bao gồm:
   * A. Truy vấn logic
   * B. Kết hợp trọng số tuyến tính ✅
   * C. Biểu đồ quyết định
   * D. Ma trận cộng hưởng
2. Phương pháp bầu chọn đa số là:
   * A. Kết hợp có trọng số
   * B. Chọn theo xác suất
   * C. Chọn theo quyết định được nhiều bộ phân tích đồng thuận nhất ✅
   * D. Dựa trên các đặc trưng siêu văn bản
3. Phương pháp phân lớp dùng để:
   * A. Ghép video và âm thanh
   * B. Dự đoán chuyển động
   * C. Phân loại các quan sát ĐPT thành các lớp ✅
   * D. Nén đa phương tiện
4. Phương pháp dựa trên ước lượng áp dụng nhiều nhất trong:
   * A. Nhận dạng chữ viết tay
   * B. Ước lượng chuyển động đối tượng ✅
   * C. Tóm tắt video
   * D. Mã hóa văn bản

**📘 IV. Công nghệ & kỹ thuật**

1. Kỹ thuật nào sau đây thường áp dụng cho chiến lược kết hợp muộn?
   * A. DBN
   * B. SVM
   * C. Trọng số tuyến tính ✅
   * D. Multimodal Transformer
2. Attention và Fusion Network là những kỹ thuật:
   * A. Dựa trên luật
   * B. Dựa trên mạng học sâu ✅
   * C. Phân cụm
   * D. Ước lượng tuyến tính
3. Các độ đo để đánh giá tương quan mức đặc trưng **KHÔNG BAO GỒM**:
   * A. Mutual Information
   * B. Correlation Coefficient
   * C. Cosine Distance
   * D. Agreement Coefficient ✅
4. Thông tin ngữ cảnh có thể là:
   * A. Thời gian
   * B. Vị trí
   * C. Tâm trạng người dùng
   * D. Cả 3 đáp án trên ✅

**📘 V. Đồng bộ dữ liệu**

1. Đồng bộ mức đặc trưng là:
   * A. Kết hợp sau khi ra quyết định
   * B. Đồng bộ dữ liệu tại các khoảng thời gian trùng nhau ✅
   * C. Chỉ cần trên ảnh
   * D. Tự động thực hiện
2. Trong đồng bộ mức quyết định, yếu tố quan trọng nhất là:
   * A. Số đặc trưng
   * B. Thời gian xử lý
   * C. Xác định điểm kết hợp trên trục thời gian ✅
   * D. Số lượng ảnh

**📘 VI. Đồng bộ dữ liệu đa phương thức**

1. Lý do chính cần đồng bộ trước khi kết hợp các phương thức là:

* A. Để loại bỏ dữ liệu thừa
* B. Vì các phương thức có định dạng, tần số khác nhau ✅
* C. Để tiết kiệm bộ nhớ
* D. Để dễ kết hợp metadata

1. Khi kết hợp ở mức đặc trưng, cần đồng bộ theo:

* A. Vector ngữ nghĩa
* B. Tốc độ xử lý
* C. Thời gian ✅
* D. Cấu trúc ngữ pháp

1. Điểm khác biệt của đồng bộ mức quyết định là:

* A. Diễn ra sau huấn luyện
* B. Diễn ra trên tập đặc trưng
* C. Xác định các điểm trên trục thời gian để kết hợp các quyết định ✅
* D. Chỉ cần dùng cho video

**📘 VII. Các vấn đề khác**

1. Độ đo nào dùng để đánh giá tương quan giữa các phương thức ở mức đặc trưng?

* A. Causal Link Analysis
* B. Correlation Coefficient ✅
* C. Entropy
* D. Precision@k

1. Tại mức quyết định, độ đo nào dùng để đánh giá mức độ đồng thuận giữa các phương thức?

* A. Mutual Information
* B. TF-IDF
* C. Agreement Coefficient ✅
* D. Bag-of-Words

1. Ví dụ về ứng dụng dùng phân tích nguyên nhân-kết quả giữa các phương thức là:

* A. Dịch văn bản
* B. Trả lời câu hỏi
* C. Phát hiện sự kiện trong môi trường giám sát ĐPT ✅
* D. Tìm kiếm ảnh

1. Cách đơn giản để kết hợp thông tin ngữ cảnh là:

* A. Fusion Attention
* B. Mô hình học sâu
* C. Quy tắc if–then–else ✅
* D. Vector space model

**📘 VIII. Ứng dụng & ví dụ**

1. Trong bài toán “Early Fusion in satellite image retrieval”, việc kết hợp thông tin sớm giúp:

* A. Phân loại tiếng nói
* B. Hiểu rõ hơn cảnh vật từ nhiều cảm biến ✅
* C. Phát hiện chuyển động
* D. Tăng kích thước dữ liệu

1. Trong bài toán “Late Fusion for social media content retrieval”, loại dữ liệu được kết hợp là:

* A. Âm thanh và nhạc
* B. Video và email
* C. Mạng xã hội và ảnh vệ tinh ✅
* D. Tốc độ mạng và video

1. “Double fusion” trong phát hiện sự kiện từ video là sự kết hợp của:

* A. Kết hợp audio và text
* B. Kết hợp hai giai đoạn: early + late fusion ✅
* C. Kết hợp hai video song song
* D. Gộp hai mô hình học máy