

FINAL QUESTION PAPER

1. 46. Which of the following is the primary computer vision task underlying the initial detection and tracking of a user's face in a virtual face filter application? (a) Object Classification (b) Object Segmentation (c) Facial Landmark Detection (d) Optical Flow Estimation
2. 20. When evaluating a recommendation system, which metric directly measures the difference between predicted and actual ratings on a continuous scale? (a) Precision@K (b) Recall@K (c) Root Mean Squared Error (RMSE) (d) Coverage
3. 2. Explain the core difference between a generative chatbot (like ChatGPT) and a retrieval-based chatbot in terms of how they formulate responses.
4. 35. The conversion of a Digital Voice Assistant's generated text response back into audible speech for the user is handled by which technology? (a) Natural Language Processing (NLP) (b) Speech Synthesis Markup Language (SSML) (c) Text-to-Speech (TTS) (d) Voice Activity Detection (VAD)
5. 34. In a typical Digital Voice Assistant architecture, the module that maintains the state of the conversation, tracks user context, and decides the next appropriate action or response is called: (a) Semantic Parser (b) Acoustic Model (c) Dialogue Management System (d) Emotion Recognition Engine
6. 60. If a virtual face filter application experiences repeated, brief failures in accurately detecting facial landmarks, what is the most likely immediate impact on the user experience? (a) The application will crash unexpectedly. (b) The filter will appear to jump, flicker, or momentarily disappear from the face. (c) The device's battery will drain rapidly. (d) The application will automatically switch to a different filter.
7. 5. Briefly describe the function of an intent classifier within a typical chatbot's Natural Language Understanding (NLU) module. What kind of machine learning model might be used for this task?
8. 17. In the context of recommendation systems, what does the "cold-start problem" primarily refer to? (a) The difficulty in recommending items that are unpopular. (b) The challenge of recommending items to new users or new items to any user due to lack of historical data. (c) The issue of slow system response times during peak usage. (d) The inability to handle dynamic changes in user preferences over time.
9. 45. Considering the future of AI interaction, how might Digital Voice Assistants conceptually integrate with technologies like Virtual Face Filters in a multimodal AI system? (a) By providing voice commands to activate or change filters based on spoken preferences. (b) By using face filter technology to enhance the assistant's microphone sensitivity. (c) By solely relying on facial expressions to understand user intent, replacing voice input entirely. (d) By converting facial movements directly into spoken words for TTS.
10. 26. Which of the following is NOT a typical component or phase in the lifecycle of a recommendation system? (a) Data collection and preprocessing. (b) Model training and evaluation. (c) Recommendation generation and serving. (d) Manual content creation by human experts for every recommendation.
11. 12. A chatbot is being developed for a smart home system. When a user says "Turn on the lights in the living room," the chatbot needs to identify the 'intent' (turn_on), the 'device' (lights), and the 'location' (living_room). What NLP task is being performed to extract 'device' and 'location' from the utterance?
12. 16. Which of the following is the primary goal of a recommendation algorithm? (a) To categorize items into predefined classes. (b) To predict user preferences and suggest relevant items. (c) To compress large datasets

for efficient storage. (d) To analyze the sentiment of user reviews.

13. 54. When evaluating the user experience of a real-time virtual face filter, which performance metric is most critical for ensuring smooth and interactive feedback without noticeable delays? (a) Model Accuracy (how well it detects features) (b) Frames Per Second (FPS) or Latency (c) Memory Footprint (RAM usage) (d) Model Size (disk space)

14. 51. To ensure a 3D virtual face filter (e.g., a hat or glasses) remains correctly positioned and oriented on a user's head as they move, which capability is essential for the underlying AI system? (a) Audio-to-text transcription (b) Gaze tracking and redirection (c) Head pose estimation and tracking (d) Voice recognition and command parsing

15. 48. A significant challenge in developing virtual face filters for real-time mobile applications, especially concerning resource consumption and user experience, is: (a) Storing large filter asset files on device. (b) Ensuring cross-platform compatibility across different operating systems. (c) Minimizing computational latency for frame-by-frame processing. (d) Managing user data for personalized filter recommendations.

16. 33. Which component of a Digital Voice Assistant is primarily responsible for understanding the user's intention and extracting relevant entities from the converted text? (a) Automatic Speech Recognition (ASR) (b) Natural Language Understanding (NLU) (c) Text-to-Speech (TTS) (d) Data Storage and Retrieval (DSR)

17. 6. The Transformer architecture, fundamental to models like ChatGPT, introduced the concept of "attention mechanisms." What is the primary benefit of the self-attention mechanism in processing input sequences? (a) It reduces the computational cost of processing long sequences by using fixed-size windows. (b) It allows the model to weigh the importance of different words in the input sequence relative to each other. (c) It enables the model to translate text directly from one language to another without intermediate steps. (d) It enforces a strict left-to-right processing order for all tokens in the sequence.

18. 55. Consider a scenario where a virtual face filter dynamically changes based on the user's emotional state (e.g., a happy filter for a smile). This advanced functionality would most directly involve the integration of which AI capability? (a) Natural Language Processing (NLP) (b) Reinforcement Learning (c) Facial Expression Recognition (d) Recommendation System Algorithms

19. 1. What is the primary purpose of a rule-based chatbot? (a) To generate human-like responses using large language models. (b) To follow predefined scripts and keywords for interactions. (c) To learn from user conversations without explicit programming. (d) To provide visual recommendations based on user sentiment.

20. 10. Explain the concept of "few-shot learning" or "in-context learning" as it applies to large language models. How does this capability reduce the need for extensive fine-tuning?

21. 53. Which programming library or framework is widely recognized and frequently utilized for its robust computer vision functionalities, including real-time image processing and feature detection, making it suitable for virtual face filter development? (a) Scikit-learn (b) TensorFlow.js (c) OpenCV (d) Pandas

22. 8. Outline a high-level architectural design for a chatbot system that integrates with an external knowledge base (e.g., a product catalog) to answer specific user queries. Mention the main components and their interactions.

23. 38. While both are conversational AI, a key distinction between a traditional chatbot and a Digital Voice Assistant often lies in: (a) The ability to access external databases. (b) Their capacity for deep learning. (c) Their primary interaction modality (text vs. voice). (d) The use of sentiment analysis.

24. 21. Consider a scenario where a digital assistant needs to suggest follow-up questions or actions based on a user's initial query. This functionality shares conceptual similarities with which aspect of recommendation systems? (a) Item-based collaborative filtering for suggesting similar items. (b) Content-based filtering using features of the initial query to match relevant follow-ups. (c) Matrix factorization for uncovering latent conversational topics. (d) All of the above, depending on the specific implementation strategy.

25. 36. A significant challenge for Natural Language Understanding (NLU) in Digital Voice Assistants involves: (a) Generating highly realistic synthesized voices. (b) Distinguishing between different user accents in Automatic Speech Recognition. (c) Resolving ambiguities, understanding context, and handling implicit requests. (d) Efficiently transmitting audio data over varying network conditions.

26. 30. In a large-scale e-commerce recommendation system, what would be the most significant challenge when scaling a traditional user-based collaborative filtering algorithm? (a) The lack of sufficiently powerful GPUs for model training. (b) The computational cost of calculating user-user similarities across millions of users. (c) The difficulty in obtaining explicit user ratings for all items. (d) The frequent need to update item descriptions.

27. 52. Virtual face filters are a prime application of Augmented Reality (AR) because they: (a) Completely replace the real-world view with a virtual environment. (b) Introduce virtual elements that are rendered and integrated into a live view of the real world. (c) Focus solely on collecting and analyzing user biometric data. (d) Create a fully immersive virtual experience using specialized head-mounted displays.

28. 3. Which component is primarily responsible for converting spoken user input into text for a digital voice assistant? (a) Natural Language Understanding (NLU) (b) Dialogue Manager (c) Speech Recognition (ASR) (d) Text-to-Speech (TTS)

29. 7. Compare and contrast the capabilities of a context-aware chatbot with a stateless chatbot. Provide a simple example where context awareness is crucial.

30. 32. The process within a Digital Voice Assistant that converts spoken words into machine-readable text is known as: (a) Natural Language Generation (NLG) (b) Automatic Speech Recognition (ASR) (c) Text-to-Speech (TTS) (d) Dialogue Management (DM)

31. 47. For generating highly realistic, novel facial expressions or styles beyond simple overlays (e.g., aging effects or cartoonization), which deep learning architecture is most frequently employed in advanced virtual face filters? (a) Recurrent Neural Network (RNN) (b) Support Vector Machine (SVM) (c) Generative Adversarial Network (GAN) (d) K-Means Clustering

32. 58. The fundamental difference in how a 2D virtual face filter (like a sticker overlay) and a 3D virtual face filter (like a virtual mask) interact with the user's face largely depends on: (a) The processing power of the device. (b) The amount of data transmitted to the cloud. (c) The dimensionality of the facial model and transformations used. (d) The color palette of the filter assets.

33. 9. Discuss one potential ethical concern related to the deployment of large language model chatbots like ChatGPT in critical applications (e.g., medical advice, legal consultation).

34. 50. From an ethical standpoint, a major concern associated with the widespread use of highly realistic virtual face filters, especially when considering advancements like deepfakes, is: (a) The excessive battery consumption on mobile devices. (b) The potential for privacy breaches through facial data collection without explicit consent. (c) The increased bandwidth usage for filter downloads. (d) The difficulty in updating filter libraries for new trends.

35. 59. For highly realistic 3D virtual face filters that deform or blend with the user's face (e.g., changing facial structure), what additional input or estimation is often crucial beyond 2D facial landmarks? (a) User's geographical location data. (b) User's current heart rate. (c) Depth information or 3D mesh reconstruction of the face. (d)

Ambient temperature readings.

36. 15. A travel agency wants to build a chatbot to help users book flights. Besides natural language understanding, what other AI component, often associated with recommendation systems, could significantly enhance the user experience by suggesting destinations or flight times? Explain how.

37. 37. How do Large Language Models (LLMs), like those powering ChatGPT, enhance the capabilities of modern Digital Voice Assistants? (a) By exclusively improving the speed of Automatic Speech Recognition. (b) By enabling more natural, coherent, and context-aware conversational responses. (c) By directly controlling smart home devices without user commands. (d) By reducing the processing power required for Text-to-Speech synthesis.

38. 44. Which of the following represents a potential security vulnerability specific to Digital Voice Assistants? (a) Inability to update firmware remotely. (b) Susceptibility to "dolphin attacks" or inaudible commands. (c) Limited battery life of the device. (d) Lack of physical buttons for interaction.

39. 39. A critical ethical concern associated with the widespread adoption of Digital Voice Assistants is: (a) The limited variety of voice tones available for TTS. (b) The potential for continuous recording and privacy breaches. (c) The complexity of integrating with legacy smart devices. (d) The high cost of developing new wake words.

40. 25. A student is developing a recommendation system using a k-Nearest Neighbors (k-NN) approach. For calculating similarity between users based on their ratings, which of the following would be an appropriate similarity metric? (a) Manhattan Distance (b) Euclidean Distance (c) Cosine Similarity (d) All of the above, depending on data characteristics and desired behavior.

41. 11. Consider a simple chatbot designed to respond to "hello" with "Hi there!" and "how are you" with "I'm doing well, thank you." Write a basic Python pseudocode snippet that demonstrates how to implement this using conditional statements (if/elif/else) for pattern matching.

42. 23. Which type of recommendation system would be most susceptible to the "filter bubble" phenomenon? (a) Only hybrid recommendation systems. (b) Systems that heavily rely on past user preferences without introducing novelty. (c) Systems that solely recommend based on global popularity. (d) Systems that use random sampling for recommendations.

43. 40. How might a recommendation algorithm be integrated into a Digital Voice Assistant's functionality? (a) To suggest relevant products, music, or information based on user preferences and past interactions. (b) To optimize the audio compression algorithms used in ASR. (c) To prioritize which wake word to listen for in a multi-assistant environment. (d) To translate spoken language into different written scripts.

44. 13. How do digital voice assistants handle ambiguity in user commands (e.g., "Play some music" without specifying genre or artist)? Briefly describe one strategy.

45. 42. In a scenario where multiple people are speaking in a room, which AI technique would be crucial for a Digital Voice Assistant to correctly attribute spoken words to individual speakers? (a) Voice Activity Detection (VAD) (b) Speaker Diarization (c) Pitch Extraction (d) Emotion Recognition

46. 14. Name two distinct metrics commonly used to evaluate the performance of a chatbot in a customer service scenario, and briefly explain what each metric measures.

47. 27. When a virtual face filter application suggests new filters to a user based on their past usage and current preferences (e.g., if they frequently use animal filters, suggesting new animal-themed ones), which recommendation approach is most likely being applied? (a) User-based collaborative filtering (b) Item-based

collaborative filtering (c) Content-based filtering (d) Popularity-based filtering

48. 22. What is the "sparsity problem" in recommendation systems, particularly for collaborative filtering? (a) The difficulty in storing a large volume of user-item interaction data. (b) The challenge when the user-item interaction matrix has very few recorded ratings. (c) The lack of computational resources to process complex recommendation algorithms. (d) The presence of too many irrelevant features in item descriptions.

49. 18. A content-based recommendation system would primarily rely on which of the following to suggest a new movie to a user? (a) Movies watched by users similar to the current user. (b) Features of movies the user has previously liked (e.g., genre, actors, director). (c) The average rating of all movies in the database. (d) Real-time trends of movie popularity across all users.

50. 28. The concept of "implicit feedback" in recommendation systems refers to: (a) User ratings provided on a numerical scale (e.g., 1-5 stars). (b) User actions that indicate preference without explicit rating, such as clicks, views, or purchases. (c) Feedback provided by system administrators for quality control. (d) User feedback obtained through direct surveys or interviews.

51. 31. Which of the following best describes the primary function of a Digital Voice Assistant? (a) To generate visual content based on text prompts. (b) To process spoken language commands and provide relevant responses or actions. (c) To optimize network traffic for voice over IP calls. (d) To analyze emotional states from facial expressions.

52. 43. The "wake word" detection mechanism in Digital Voice Assistants presents a significant technical challenge primarily due to: (a) The need to accurately identify a very short, specific audio pattern while minimizing false positives in noisy environments. (b) The difficulty in synthesizing a pleasant-sounding confirmation chime. (c) The computational overhead of displaying a visual confirmation. (d) The process of uploading user preferences to a cloud server.

53. 56. Virtual face filters must maintain consistent performance under varying environmental conditions. Which technical challenge is most significant when a user moves from brightly lit to dimly lit environments? (a) Degradation of network connectivity. (b) Increased complexity in audio processing. (c) Reduced accuracy of facial landmark detection and feature tracking. (d) Higher power consumption of the device.

54. 41. Which type of deep learning architecture is widely used in Automatic Speech Recognition (ASR) systems, particularly for capturing sequential dependencies in audio signals? (a) Convolutional Neural Networks (CNNs) for image feature extraction. (b) Generative Adversarial Networks (GANs) for synthetic data generation. (c) Recurrent Neural Networks (RNNs) like LSTMs or GRUs, and more recently Transformers. (d) Support Vector Machines (SVMs) for traditional classification tasks.

55. 4. A company wants to implement a chatbot for customer support. Describe two key challenges they might face when using a purely rule-based approach for handling complex customer queries, and suggest how a more advanced AI approach could mitigate one of these challenges.

56. 49. What type of data is most crucial for training machine learning models that accurately identify and localize facial key points (landmarks) across diverse individuals and conditions for virtual face filters? (a) Large datasets of text descriptions of facial features. (b) Audio recordings of speech patterns corresponding to facial movements. (c) Images/videos annotated with precise coordinates of facial key points. (d) Demographic information and user preferences for filter styles.

57. 24. If you are designing a recommendation system where user privacy is paramount, which approach would you prioritize to minimize direct sharing of individual user data? (a) User-based collaborative filtering with explicit ratings. (b) Item-based collaborative filtering with explicit ratings. (c) Content-based filtering using aggregated item features and user profiles. (d) Deep learning models requiring extensive personal data.

58. 19. Which of the following techniques is most commonly associated with collaborative filtering that leverages latent factors? (a) Apriori Algorithm for association rule mining. (b) Term Frequency-Inverse Document Frequency (TF-IDF). (c) Singular Value Decomposition (SVD) or Matrix Factorization. (d) K-Means Clustering for item segmentation.

59. 29. What is the primary benefit of using a hybrid recommendation system compared to a pure content-based or pure collaborative filtering system? (a) It is always faster to compute recommendations. (b) It completely eliminates the cold-start problem. (c) It can combine the strengths of different approaches to mitigate individual weaknesses (e.g., cold start, sparsity, filter bubble). (d) It requires significantly less data for training.

60. 57. How do robust virtual face filter systems typically handle the diversity of human face shapes, sizes, and proportions to apply filters effectively across different users? (a) By using a single, fixed filter overlay that stretches automatically. (b) By requiring users to manually adjust the filter's size and position. (c) By employing statistical shape models (e.g., Active Appearance Models) to adapt filter placement. (d) By collecting and storing a pre-trained model for every possible face type.

ANSWER KEY

1. (c)

2. (c)

3. Generative chatbots create novel responses by generating new text based on patterns learned from vast training data, without relying on pre-existing responses. Retrieval-based chatbots select the most appropriate pre-written response from a predefined database of responses based on the user's query.

4. (c)

5. (c)

6. (b)

7. An intent classifier determines the underlying goal or purpose of the user's utterance (e.g., "book_flight", "check_balance"). A machine learning model such as a Support Vector Machine (SVM), Logistic Regression, or a deep learning model like a Feedforward Neural Network or a Text Classifier based on transformer embeddings (e.g., BERT) might be used for this task.

8. (b)

9. (a)

10. (d)

11. Named Entity Recognition (NER).

12. (b)

13. (b)

14. (c)

15. (c)

16. (b)

17. (b)

18. (c)

19. (b)

20. Few-shot learning (or in-context learning) refers to the capability of a large language model to learn a new task or concept and generate appropriate responses using only a small number of examples provided within the input prompt itself, without requiring any updates to the model's internal weights through traditional fine-tuning. This capability significantly reduces the need for extensive fine-tuning because the model can leverage its vast pre-trained knowledge and generalize from these few examples or instructions, rapidly adapting to new tasks based solely on the prompt's context.

21. (c)

22. High-level architectural design:

23. (c)

24. (d)

25. (c)

26. (b)

27. (b)

28. (c)

29. A context-aware chatbot maintains a memory of previous turns in a conversation, allowing it to understand and respond to user queries that refer back to earlier information. A stateless chatbot treats each user input as an independent event, without recalling any prior interaction history. Example: If a user asks "What is the capital of France?" and then immediately asks "What language do they speak there?", a context-aware chatbot would know "there" refers to France, whereas a stateless chatbot would not and might ask for clarification or provide an irrelevant answer.

30. (b)

31. (c)

32. (c)

33. One potential ethical concern is the generation of biased or misleading information. Large language models are trained on vast datasets that may contain societal biases, leading them to produce responses that are discriminatory, reinforce stereotypes, or provide inaccurate "facts." In critical applications like medical or legal advice, this could have severe real-world consequences, from misdiagnosis to incorrect legal counsel.

34. (b)

35. (c)

36. A recommendation algorithm could significantly enhance the user experience by suggesting destinations, flight times, or even travel packages based on explicit user preferences, implicit patterns from past bookings, popular routes, or current deals. For example, if a user frequently searches for flights to beach destinations in winter, the chatbot could proactively suggest tropical getaways. If a user inputs a departure city, the algorithm could recommend popular or cost-effective destinations based on historical data. This proactive suggestion capability personalizes the booking process, making it more efficient and tailored to the user's potential interests.

37. (b)

38. (b)

39. (b)

40. (d)

41. N/A

42. (b)

43. (a)

44. Digital voice assistants handle ambiguity in user commands by employing strategies such as: 1) Asking clarifying questions: If the command is too vague (e.g., "Play some music"), the assistant might ask, "What genre of music would you like to hear?" or "Who is your favorite artist?". 2) Using default settings or user preferences: If

no specific input is given, the assistant might default to a user's most listened-to genre or a popular playlist.

45. (b)

46. 1) Resolution Rate: This metric measures the percentage of user queries or issues that the chatbot successfully resolves completely without requiring human agent intervention. It indicates the chatbot's effectiveness and autonomy in handling user requests. 2) Customer Satisfaction (CSAT) Score: Typically collected via a post-interaction survey (e.g., "How satisfied were you with this chat?"), this metric measures the user's perceived quality of the interaction and the helpfulness of the chatbot.

47. (c)

48. (b)

49. (b)

50. (b)

51. (b)

52. (a)

53. (c)

54. (c)

55. Two key challenges for a purely rule-based approach: 1) Difficulty in handling variations in user phrasing or synonyms, requiring extensive manual rule creation for every possible input. 2) Limited ability to maintain conversational context across multiple turns, leading to disjointed interactions. A more advanced AI approach, such as Natural Language Understanding (NLU) with deep learning models, can mitigate challenge 1 by intelligently interpreting user intent and extracting entities even with unseen or varied phrasing, reducing the need for explicit rule programming.

56. (c)

57. (c)

58. (c)

59. (c)

60. (c)