ChatGPT

Session 1: Language in Cognitive Science

1. **Which of the following best captures the definition of human language according to linguistics?**
   A. A set of hand signals and gestures.
   B. A biological reflex with no structure.
   C. A structured system of communication with grammar and vocabulary [1].
   D. A fixed set of sentences.
   Answer: C. *Human language is defined as "a structured system of communication that consists of grammar and vocabulary" [1]. It is rule-governed and allows infinite utterances, unlike a fixed set of sentences or mere gestures.*

2. **What is the scientific study of language called?**
   A. Philology
   B. Linguistics [1]
   C. Anthropology
   D. Semiotics
   Answer: B. *The study of language as a scientific discipline is called linguistics [1]. Philology is historical study of texts, but linguistics covers all scientific aspects of language.*

3. **Who introduced the structuralist view of language as a formal system of signs governed by grammar?**
   A. Noam Chomsky
   B. Ferdinand de Saussure [2]
   C. Leonard Bloomfield
   D. Edward Sapir
   Answer: B. *Ferdinand de Saussure introduced the structuralist view, defining language as "a formal system of signs governed by grammar" [2]. Chomsky later introduced generative grammar (see next question).*

4. **Which linguist proposed that grammatical rules are an innate feature of the human mind?**
   A. B.F. Skinner
   B. Ferdinand de Saussure
   C. Noam Chomsky [3]
   D. Roman Jakobson
   Answer: C. *Noam Chomsky is known for the generative theory, considering grammatical rules to be innate to the human mind [3]. This contrasts with behaviorist views (e.g., Skinner).*

5. **Which of the following is a formal (artificial) language rather than a natural human language?**
   A. Spanish
   B. Java (programming language) [4]
   C. Mandarin
   D. Arabic

Answer: B. *Java is a programming (formal) language, not a natural human language. Formal languages are "artificially constructed systems for encoding information"* [4] *, whereas Spanish, Mandarin, and Arabic are natural languages used by human communities.*

Session 2: Language Analysis and Computational Linguistics

1. **In linguistics, what do paradigmatic relations involve?**
   A. Sequential order of words in a sentence.
   B. Substitution possibilities among words of the same class [5] .
   C. Phonetic articulation.
   D. Semantic roles.
   Answer: B. *Paradigmatic relations are about elements that can substitute each other in context (the "OR" choices)* [5] *. For example, "red", "blue", "green" can substitute in "the ___ car". This contrasts with syntagmatic (sequence/combination) relations.*

2. **Which of the following is NOT a standard level of linguistic analysis?**
   A. Phonology
   B. Syntax
   C. Morphology
   D. Sociolinguistics [6]
   Answer: D. *The main levels of linguistic analysis include phonetics, phonology, morphology, syntax, semantics, pragmatics, and discourse* [6] *. Sociolinguistics is a subfield of linguistics (social context), not a 'level' like the others.*

3. **According to IBM's explanation, how is NLP related to computational linguistics?**
   A. NLP is purely rule-based while CL is data-driven.
   B. NLP evolved from computational linguistics and emphasizes machine learning [7] .
   C. CL is a subset of NLP.
   D. They are unrelated fields.
   Answer: B. *NLP evolved from computational linguistics, and NLP tends to use machine learning on large datasets, whereas computational linguistics focused more on the formal linguistic aspects* [7] *.*

4. **In NLP, what does Natural Language Understanding (NLU) refer to?**
   A. Generating human-like text.
   B. Extracting knowledge from images.
   C. Comprehending and interpreting meaning of text [8] .
   D. Converting text to speech.
   Answer: C. *NLU is about machine comprehension of text, using syntactic and semantic analysis to determine meaning* [8] *. NLG (Natural Language Generation), by contrast, is about generating text.*

5. **What is the primary goal of Natural Language Generation (NLG)?**
   A. Parsing sentences into grammatical trees.
   B. Translating text between languages.
   C. Generating coherent text from data or abstract representations [9] .

D. Classifying texts by sentiment.
Answer: C. *NLG involves producing (writing) human-like text from structured data or representations* [9] *. It is the complement of NLU within NLP.*

Session 3: Shallow Parsing and NLP Tools

1. **What is tokenization in NLP?**
   A. Converting text into lower case.
   B. Removing stopwords.
   C. Dividing text into smaller units called tokens (words, characters, etc.) [10] .
   D. Tagging parts of speech.
   Answer: C. *Tokenization breaks raw text into tokens (words, subwords, characters)* [10] *. For example, splitting "Cats, dogs." into ["Cats", ",", "dogs", "."].*

2. **What does part-of-speech (POS) tagging accomplish?**
   A. Identifying named entities in text.
   B. Assigning each word its grammatical part of speech (noun, verb, etc.) [11] .
   C. Parsing the sentence tree.
   D. Breaking text into sentences.
   Answer: B. *POS tagging marks each word with its part of speech based on definition and context* [11] *(e.g., tagging "run" as a noun or verb).*

3. **Which statement correctly distinguishes lemmatization from stemming?**
   A. Both produce non-word stems.
   B. Stemming requires dictionaries, lemmatization does not.
   C. Lemmatization yields dictionary (valid) base forms; stemming simply trims suffixes [12] .
   D. They are identical processes.
   Answer: C. *Lemmatization reduces a word to its dictionary base form (lemma), ensuring it's a real word, while stemming crudely removes affixes without guaranteeing a valid word* [12] *.*

4. **What is Named-Entity Recognition (NER)?**
   A. A translation task.
   B. Identifying and classifying named entities (people, locations, organizations) in text [13] .
   C. Generating questions from text.
   D. Speech recognition.
   Answer: B. *NER finds and categorizes names of people, places, organizations, etc. in text* [13] *. For example, extracting "Alice" as a Person and "Paris" as a Location.*

5. **Which NLP task involves determining the correct meaning of an ambiguous word in context?**
   A. Tokenization
   B. Dependency parsing
   C. Word Sense Disambiguation (WSD) [14]
   D. Machine translation
   Answer: C. *WSD identifies which sense of a polysemous word is intended in context* [14] *. For instance, deciding that "bank" means "financial institution" vs "river bank" based on surrounding words.*

Session 4: Deep Parsing and NLP Tools

1. **What is Information Extraction (IE) in NLP?**
   A. Translating text to another language.
   B. Extracting structured information (e.g., relations, entities) from unstructured text [15].
   C. Summarizing documents.
   D. Speech-to-text conversion.
   Answer: B. *IE converts unstructured text into structured data, such as extracting "X works at Y" relations from text [15]. It is used to populate databases or knowledge bases.*

2. **What is the goal of automatic text summarization?**
   A. Translate text to another language.
   B. Condense text to a shorter form while preserving key information [16].
   C. Remove all stopwords from text.
   D. Convert speech to text.
   Answer: B. *Text summarization produces a concise summary of one or more documents, retaining main ideas [16]. For example, summarizing an article into a few bullet points.*

3. **Which type of summarization generates entirely new sentences not present in the source text?**
   A. Extractive summarization
   B. Text clustering
   C. Abstractive summarization [17]
   D. Keyword extraction
   Answer: C. *Abstractive summarization creates novel sentences to capture meaning, whereas extractive summarization selects existing sentences [17].*

4. **What is the Cocke–Younger–Kasami (CYK) algorithm used for?**
   A. Named Entity Recognition
   B. Parsing context-free grammars (CFGs) in Chomsky Normal Form [18]
   C. Clustering word embeddings
   D. Optimizing neural networks
   Answer: B. *CYK is a parsing algorithm for context-free grammars, requiring the grammar to be in Chomsky Normal Form [18]. It efficiently decides if a string belongs to a CFG.*

5. **In the context of the semantic web, what is an ontology?**
   A. A spoken language
   B. A database table
   C. A formal specification of concepts and relations in a domain [19]
   D. A compression algorithm
   Answer: C. *In semantic web and AI, an ontology defines the concepts and relationships of a domain to enable knowledge sharing [19]. For instance, an ontology might define "Person", "Organization", and their relationships.*

Session 5: Statistical NLP Methods

1. **What does the Markov property of a stochastic process imply?**
   A. Future states depend only on present state, not on past states [20].
   B. All events are independent.
   C. Past states determine future.
   D. All states are equally probable.
   Answer: A. *The Markov property states that given the present, the future is independent of the past [20]. In language modeling, this means the next word depends only on the current word (in a first-order Markov model).*

2. **In information theory, what does entropy measure?**
   A. Probability of an event.
   B. Average uncertainty or information content of a random variable [21].
   C. The length of a message.
   D. The frequency of the most likely outcome.
   Answer: B. *Entropy quantifies the average amount of uncertainty (or information) in possible outcomes of a random variable [21]. For language, high entropy means more unpredictability of words.*

3. **What is the main idea behind Support Vector Machines (SVM) in classification?**
   A. Use decision trees to separate classes.
   B. Find the hyperplane that maximizes the margin between classes [22].
   C. Cluster points and assign majority label.
   D. Use nearest neighbors rule.
   Answer: B. *SVMs are supervised models that find the optimal separating hyperplane maximizing the distance (margin) between classes [22]. This helps generalize well on unseen data.*

4. **What is collaborative filtering in recommender systems?**
   A. Filtering images for computer vision.
   B. Grouping users by similar behavior to recommend items [23].
   C. Analyzing text sentiment.
   D. Removing stopwords in text.
   Answer: B. *Collaborative filtering makes recommendations by finding users with similar preferences and recommending items they liked [23]. For example, if User A and B liked many of the same movies, recommend to A what B liked that A hasn't seen yet.*

5. **What technique is used in Latent Semantic Analysis (LSA) to reduce dimensionality of text data?**
   A. Decision Trees
   B. Clustering
   C. Singular Value Decomposition (SVD) [24]
   D. Tokenization
   Answer: C. *LSA applies Singular Value Decomposition to the term-document matrix to reduce dimensionality and uncover latent topics [24]. This helps capture the main themes in the data.*

Session 6: Machine Learning and Deep Learning for NLP

1. **How does deep learning differ from traditional machine learning in NLP?**
   A. Deep learning always uses small datasets.
   B. Deep learning automates feature extraction and often requires more data [25].
   C. Traditional ML is always neural network-based.
   D. There is no difference.
   Answer: B. *Deep learning is a subset of ML that "automates much of the feature extraction process" [25] and typically requires large datasets. Traditional ML often relies on manual feature engineering.*

2. **Which activation function is less prone to the vanishing gradient problem compared to sigmoid?**
   A. Tanh
   B. Softmax
   C. ReLU (Rectified Linear Unit) [26]
   D. Step function
   Answer: C. *The ReLU activation only saturates in one direction and avoids squashing small inputs toward zero, so it "suffers less from the vanishing gradient problem" [26] than sigmoid or tanh, making it preferred in deep networks.*

3. **What problem were LSTM networks specifically designed to address in RNNs?**
   A. Overfitting
   B. Lack of parallelism
   C. Vanishing gradient and long-range dependencies [27]
   D. Data augmentation
   Answer: C. *LSTMs (Long Short-Term Memory networks) were designed to solve the vanishing gradient problem in RNNs and to better capture long-term dependencies [27]. They use gating mechanisms to preserve gradients.*

4. **Compared to LSTMs, Gated Recurrent Units (GRUs) have what structural difference?**
   A. GRUs have three gates, LSTMs have two.
   B. GRUs have no gates at all.
   C. GRUs have two gates (reset and update), while LSTMs have three (input, forget, output) [28].
   D. LSTMs use convolutions internally, GRUs do not.
   Answer: C. *A key difference is that GRUs have two gates (reset and update) whereas LSTMs use three gates (input/update, forget, and output) [28]. This makes GRUs slightly simpler.*

5. **Which statement is true about GRUs and LSTMs?**
   A. GRUs are more complex than LSTMs.
   B. LSTMs train faster than GRUs.
   C. GRUs train faster due to simpler structure; LSTMs can capture longer dependencies [29] [30].
   D. They have identical performance in all cases.
   Answer: C. *GRUs are simpler (fewer gates) and tend to train faster [30]. LSTMs, with a memory cell and*

additional gate, can handle longer-range dependencies and often perform better when capturing long-distance context [29] .

Session 7: NLP Algorithms and Case Studies

1. **Why is text preprocessing important in NLP?**
   A. It translates text to another language.
   B. It cleans and normalizes raw text, improving data quality and model performance [31] .
   C. It encrypts the text for security.
   D. It has no effect on NLP tasks.
   Answer: B. *Preprocessing (e.g., tokenization, cleaning, normalizing) removes noise and inconsistencies in text, which improves the quality of input data and boosts model performance* [31] *. For example, lowercasing and removing punctuation helps models learn better.*

2. **Which of the following is a typical text preprocessing step?**
   A. Sorting words alphabetically.
   B. Removing all verbs.
   C. Tokenization (breaking text into words/tokens) [32] .
   D. Expanding abbreviations.
   Answer: C. *Tokenization—breaking text into tokens (words, sentences, etc.)—is a standard preprocessing step* [32] *. Other common steps include lowercasing, stopword removal, and lemmatization.*

3. **NLTK was originally designed for which purpose?**
   A. Industrial-scale text processing.
   B. Educational and research use by linguists and scholars [33] .
   C. Image analysis.
   D. Real-time speech recognition.
   Answer: B. *NLTK (Natural Language Toolkit) was built by researchers as an educational toolkit to develop NLP functions, particularly useful in academia and research* [33] *.*

4. **spaCy is known for what advantage over NLTK?**
   A. It is only for image data.
   B. It focuses on slow but accurate parsing.
   C. It provides faster, production-oriented NLP pipelines with accurate syntactic analysis [34] .
   D. It translates between any two languages.
   Answer: C. *spaCy offers a fast and efficient NLP pipeline, providing "the fastest and most accurate syntactic analysis" for practical use* [34] *. It is oriented towards production, unlike NLTK's educational focus.*

5. **Which of these libraries is optimized for large-scale, distributed NLP tasks?**
   A. NLTK
   B. spaCy
   C. Spark NLP [35]
   D. OpenCV
   Answer: C. *Spark NLP (from John Snow Labs) is built on Apache Spark to scale NLP tasks in a distributed environment. For example, it provides scalable ASR and text processing pipelines* [35] *. (OpenCV is for vision.)*

Session 8: Advanced RNN Architectures – GRU and LSTM

1. **What advantage do Long Short-Term Memory (LSTM) networks have over vanilla RNNs?**
   A. They require less data to train.
   B. They avoid the vanishing gradient problem, enabling learning of long-term dependencies [27].
   C. They use convolution instead of recurrence.
   D. They do not use activation functions.
   Answer: B. *LSTMs mitigate the vanishing gradient issue that plagues basic RNNs, allowing them to learn dependencies over longer time spans [27]. This is achieved through their gated architecture.*

2. **Compared to LSTMs, Gated Recurrent Units (GRUs) have how many gates?**
   A. Zero gates.
   B. Two gates (reset and update) [28].
   C. Three gates (same as LSTM).
   D. Five gates.
   Answer: B. *A GRU has two gates (a reset gate and an update gate), whereas an LSTM has three (input/update, forget, and output gates) [28]. This makes GRUs simpler to train in some cases.*

3. **Which of the following is true about GRUs?**
   A. GRUs include a separate memory cell.
   B. GRUs are more complex than LSTMs.
   C. GRUs have fewer parameters and often train faster than LSTMs [30].
   D. GRUs cannot handle sequential data.
   Answer: C. *GRUs are simpler (fewer parameters) than LSTMs, so they often train faster and are more efficient [30]. They omit the separate memory cell of LSTMs.*

4. **In what scenario might an LSTM outperform a GRU?**
   A. When modeling extremely short sequences.
   B. When training time is very limited.
   C. When long-distance (long-term) dependencies in sequence data are crucial [29].
   D. When only linear relationships exist.
   Answer: C. *LSTMs, with their extra gate and memory cell, "in theory remember longer sequences" than GRUs and tend to perform better on tasks requiring long-range dependencies [29].*

5. **Which gate is present in an LSTM but not in a GRU?**
   A. Update gate
   B. Forget gate [28]
   C. Reset gate
   D. Input gate
   Answer: B. *The forget gate is used in LSTMs to determine what information to discard from the cell state; GRUs do not have a separate forget gate (they combine roles via their update gate) [28].*

Session 9: Sequence Modeling and Word Representation

1. **What are word embeddings?**
   A. One-hot vectors with all zeros.
   B. Numeric vector representations of words that capture semantic similarity [36] .
   C. Counts of word occurrences.
   D. Syntax trees.
   Answer: B. *Word embeddings map words to continuous vectors in a semantic space, so that similar words have similar vectors [36] . This enables models to understand word meaning and context.*

2. **Which of the following is an advantage of good word embeddings?**
   A. Similar words have dissimilar representations.
   B. They eliminate the need for models.
   C. Words with similar meanings have similar vectors [37] .
   D. They are all zero except one dimension.
   Answer: C. *A property of effective embeddings is that semantically or syntactically similar words obtain similar vector representations [37] . For example, "king" and "queen" should be close in the embedding space.*

3. **What is a drawback of one-hot encoding for words compared to embeddings?**
   A. It captures word order poorly.
   B. It has fixed length for any vocabulary.
   C. It produces very high-dimensional sparse vectors and fails to capture semantic similarity [38] .
   D. It requires neural networks to train.
   Answer: C. *One-hot vectors are high-dimensional (dimension = vocabulary size) and sparse, and they do not encode similarity (each word is orthogonal) [38] . Embeddings address this by dense, lower-dimensional representations.*

4. **What is the goal of sentiment classification in NLP?**
   A. Translating text to another language.
   B. Identifying the emotional polarity (positive, negative, neutral) of text [39] .
   C. Summarizing text.
   D. Extracting named entities.
   Answer: B. *Sentiment analysis (or classification) aims to detect the emotional tone of text, usually classifying it as positive, negative, or neutral [39] . It helps gauge opinions in reviews, tweets, etc.*

5. **Which task would involve assigning a text a label such as "positive" or "negative"?**
   A. Part-of-speech tagging
   B. Sentiment analysis [39]
   C. Topic modeling
   D. Syntax parsing
   Answer: B. *Sentiment analysis (opinion mining) is exactly the task of labeling text with sentiment polarity, such as positive or negative [39] .*

Session 10: Word2Vec, Seq2Seq, and Transformers

1. **In Word2Vec models, how do CBOW and Skip-gram differ?**
   A. CBOW predicts context words from a target; Skip-gram predicts target from context.
   B. Both do the same thing.
   C. CBOW predicts a word given its context; Skip-gram predicts surrounding context words given a target word [40].
   D. They are unrelated to word embeddings.
   Answer: C. *CBOW (Continuous Bag of Words) predicts the target word from its surrounding context, whereas Skip-gram does the opposite: it uses the target word to predict its context words [40].*

2. **What does GloVe stand for, and what is its main approach?**
   A. Global Vocabulary – a rule-based model.
   B. Global Vectors – it learns word vectors by factorizing global word-word co-occurrence statistics [41].
   C. Global Library – a translation model.
   D. Generalized Language – a language model.
   Answer: B. *GloVe stands for Global Vectors. It is an unsupervised model that trains on global word co-occurrence counts to produce word embeddings [41].*

3. **What is a sequence-to-sequence (Seq2Seq) model used for?**
   A. Image classification.
   B. Mapping an input sequence to an output sequence, such as in machine translation [42].
   C. Sorting numbers.
   D. Speech recognition.
   Answer: B. *A Seq2Seq model is a neural architecture for transforming one sequence into another (of possibly different length) [42]. It's widely used in tasks like machine translation and text summarization.*

4. **In a Seq2Seq model, what is the role of the decoder?**
   A. To tokenize the input.
   B. To encode the input into vectors.
   C. To take the encoded context and generate the output sequence token by token [43].
   D. To label each token's part of speech.
   Answer: C. *The decoder receives the encoded representation (context vector) from the encoder and generates the target sequence one token at a time [43]. In translation, for example, it produces the translated sentence.*

5. **What is a key advantage of the Transformer architecture in NLP?**
   A. It only handles fixed-length inputs.
   B. It uses recurrent layers exclusively.
   C. It relies on attention mechanisms to allow parallel processing of sequences. (See contexts)
   D. It requires no training.
   Answer: C. *The Transformer model dispenses with recurrence and uses self-attention mechanisms to process all positions of the input in parallel, enabling efficient modeling of long-range dependencies.*

Session 11: BERT and NLP Deployment

1. **What does BERT stand for?**
   A. Bidirectional Encoder Representations from Transformers [44]
   B. Bidirectional English Recognition Tool
   C. Bayesian Encoder for Recurrent Texts
   D. Basic Encoder Representation Trainer
   Answer: A. *BERT is an acronym for Bidirectional Encoder Representations from Transformers [44]. It is a transformer-based language model introduced by Google.*

2. **What training tasks are used in BERT's pre-training?**
   A. Autoencoding and clustering.
   B. Machine translation.
   C. Masked language modeling and next-sentence prediction [45].
   D. Speech synthesis.
   Answer: C. *BERT is pretrained using two self-supervised tasks: predicting masked tokens (Masked LM) and predicting whether two sentences follow each other (Next Sentence Prediction) [45].*

3. **Which of these is NOT a component of BERT's architecture?**
   A. Transformer encoder stack.
   B. WordPiece tokenizer.
   C. Unsupervised text generation head (decoder).
   D. Optional task-specific output head.
   Answer: C. *BERT uses an encoder-only transformer architecture [46]. It does not include a traditional decoder for language generation during pre-training; instead, it has a classification/prediction "task head" used for masked token predictions.*

4. **Which NLP tasks can BERT be fine-tuned to perform?**
   A. Image captioning.
   B. Sentiment analysis and question answering [47].
   C. Audio synthesis.
   D. Database indexing.
   Answer: B. *BERT's learned representations can be fine-tuned for many text tasks, such as question answering and sentiment classification [47]. It provides contextual word embeddings that help in downstream tasks.*

5. **How would you typically deploy a trained BERT model as a service?**
   A. Compile it to C code.
   B. Use a web framework (e.g. Flask) to serve inference requests.
   C. Email it as an attachment.
   D. Deploy it on a GPU in the air.
   Answer: B. *In practice, trained NLP models like BERT are often served via lightweight web APIs (for example, using Python Flask) so that applications can send text and receive predictions. (While citations are not provided here, frameworks like Flask are common for model deployment.)*

Session 12: Speech Processing – Phonetics and Signal Processing

1. **What does articulatory phonetics study?**
   A. The acoustic properties of speech waves.
   B. How speech sounds are produced by the vocal tract [48].
   C. The grammar of language.
   D. Speech recognition algorithms.
   Answer: B. *Articulatory phonetics examines how speech sounds are formed by the vocal organs (tongue, lips, etc.)* [48]*. Acoustic phonetics (contrasted) deals with the sound waves.*

2. **What does acoustic phonetics focus on?**
   A. The meaning of speech.
   B. Physical properties of sound waves produced in speech [48].
   C. How sounds are heard by the ear.
   D. The syntax of spoken sentences.
   Answer: B. *Acoustic phonetics studies the physical properties of speech sounds (frequency, amplitude, duration) as vibrations in the air* [48].

3. **According to the Nyquist–Shannon sampling theorem, what sampling rate is needed to avoid aliasing of a signal of bandwidth B?**
   A. Greater than 0.5×B.
   B. At least 2×B [49].
   C. Exactly equal to B.
   D. Any rate is fine if filters are used.
   Answer: B. *The sampling theorem states that to capture all information from a band-limited signal, the sampling rate must be at least twice the maximum frequency (bandwidth) of the signal* [49].

4. **What is Automatic Speech Recognition (ASR)?**
   A. Generating audio from text.
   B. Transcribing spoken language into text [35].
   C. Classifying language identity.
   D. Emotion detection in speech.
   Answer: B. *ASR (Speech-to-Text) converts audio speech input into text output* [35]*. For example, dictation software transcribes spoken words into a text document.*

5. **What is Text-to-Speech (TTS) in speech processing?**
   A. A system to compress audio signals.
   B. Synthesizing spoken audio from written text.
   C. Translating speech to another language.
   D. Detecting the language of a text.
   Answer: B. *TTS converts written text into natural-sounding spoken audio. It is used in virtual assistants, screen readers, etc. (Definition from standard NLP/AI knowledge.)*

Session 13: NLP Applications

1. **Which application would involve transliteration from one script to another?**
   A. Spell checker
   B. Language identification
   C. Machine translation between scripts
   D. Transliteration (mapping text between scripts)
   Answer: D. *Transliteration converts text from one writing system into another (e.g., Devanagari to Latin). It is distinct from translation (different languages) and is specifically listed as an NLP application in the syllabus.*

2. **What is a typical method for spell checking in NLP?**
   A. Using OCR technology.
   B. Rule-based pattern matching or probabilistic edit distance algorithms.
   C. Neural machine translation.
   D. Spectrogram analysis.
   Answer: B. *Spell checkers often use dictionaries and algorithms like edit distance (e.g., Levenshtein distance) or probability models to detect and correct misspelled words. (Common NLP knowledge.)*

3. **Which task involves determining the language of a given text sample?**
   A. Topic modeling
   B. Language identification (language detection)
   C. Named entity recognition
   D. Summarization
   Answer: B. *Language identification is the task of automatically detecting which language a text is written in, based on its statistical and lexical features. It is a standard NLP task.*

4. **What is a common approach to gender prediction from text?**
   A. Counting letters.
   B. Machine learning on features like word usage and names.
   C. Checking IP address.
   D. Acoustic analysis.
   Answer: B. *Gender prediction models often use machine learning to find patterns in word usage, topics, or names that correlate with gender. For example, using lexical features from a person's writing to guess gender.*

5. **Which of the following is a challenge specific to processing Indian languages?**
   A. Handling Latin script.
   B. Dealing with multiple scripts and rich morphology.
   C. Lack of any script.
   D. Abundance of rhyme detection.
   Answer: B. *Indian languages often come in various scripts (Devanagari, Tamil, etc.) and tend to have complex morphology (agglutination, inflection), making tasks like tokenization and normalization challenging. (This is a known issue in multilingual NLP.)*

Session 14: Introduction to Computer Vision

1. **What are the "three Rs" of computer vision?**
   A. Recognition, Reconstruction, Rendering.
   B. Reading, Writing, Arithmetic.
   C. Repetitions, Reductions, Relations.
   D. Regression, Relaxation, Randomness.
   Answer: A. *The "three Rs" in vision are commonly given as Recognition (identifying objects), Reconstruction (understanding 3D shape from images), and Rendering (synthesizing images) – fundamental tasks in computer vision.*

2. **Which level of vision processing involves detecting edges and colors?**
   A. High-level vision
   B. Mid-level vision
   C. Low-level vision
   D. Semantic-level vision
   Answer: C. *Low-level vision concerns basic image processing such as edge and line detection, color computation, filtering, etc. Mid-level involves grouping and segmentation, and high-level deals with object recognition and reasoning.*

3. **What does image filtering using a convolution kernel achieve?**
   A. It converts images to text.
   B. It applies operations like blurring, sharpening, or edge detection [50] .
   C. It encrypts the image.
   D. It compresses the image to JPEG.
   Answer: B. *Applying a convolution kernel (mask) to an image can perform effects such as blurring, sharpening, or highlighting edges [50] . The kernel's values determine the filter's function.*

4. **Which technique is commonly used to equalize image brightness distribution?**
   A. Fourier transform
   B. Histogram equalization
   C. Gaussian Mixture Models
   D. Motion tracking
   Answer: B. *Histogram equalization redistributes the image's intensity histogram to enhance contrast, a basic image processing method.*

5. **What is a typical first step in processing a color image in computer vision?**
   A. Fourier analysis
   B. Converting to grayscale or separating into color channels
   C. Speech recognition
   D. Object detection
   Answer: B. *Often, color images are converted to grayscale or split into channels (RGB/YUV) as a preprocessing step for algorithms that operate on intensity values.*

Session 15: Image Processing Basics

1. **What is edge detection used for in images?**
   A. Increasing image resolution.
   B. Identifying sharp changes in image intensity [50] .
   C. Colorizing black-and-white photos.
   D. Compressing the image.
   Answer: B. *Edge detection algorithms (like Sobel or Canny) find pixels where intensity changes sharply, which often correspond to object boundaries or edges in an image.*

2. **Which filter is effective at removing "salt-and-pepper" noise while preserving edges?**
   A. Mean blur
   B. Median filter
   C. Sobel filter
   D. Histogram equalization
   Answer: B. *A median filter replaces each pixel with the median in its neighborhood, effectively removing outlier noise (salt-and-pepper) while keeping edges sharp. (Common CV knowledge.)*

3. **What does morphological dilation do to a binary image?**
   A. Erodes object boundaries.
   B. Expands (grows) object shapes by adding pixels to edges.
   C. Changes the color of objects.
   D. Inverts the image.
   Answer: B. *Dilation is a morphological operation that expands objects in a binary image, filling small holes and connecting nearby objects. (It is the inverse of erosion.)*

4. **What is the main purpose of Fourier transform in image processing?**
   A. Object recognition.
   B. Change of basis to frequency domain for filtering and analysis.
   C. Segmenting images into superpixels.
   D. Generating random images.
   Answer: B. *The Fourier transform converts spatial image data into frequency components. This is useful for analyzing periodic patterns and applying frequency-domain filters.*

5. **In an object tracking pipeline, what is typically done after detecting objects in each frame?**
   A. Immediately retrain the model.
   B. Compare detected objects across frames to maintain identities (tracking).
   C. Convert images to text.
   D. Discard previous frame data.
   Answer: B. *After detection, tracking algorithms link detections frame-to-frame (data association) so each object is consistently identified over time. Background subtraction or motion estimation may be used first, but final step is linking detections.*

Session 16: Feature Detection and Image Classification

1. **What is the main idea of corner detection (e.g., Harris detector)?**
   A. To find pixels with large gradient in all directions (corner points).
   B. To detect edges only.
   C. To classify objects.
   D. To reduce image noise.
   Answer: A. *Corner detectors identify points in an image where intensity changes significantly in orthogonal directions (high variation), which are often good features for matching or tracking.*

2. **What does HOG (Histogram of Oriented Gradients) descriptor capture?**
   A. Color histograms.
   B. Distribution of gradient orientations (edge directions) in localized regions.
   C. Frequency domain features.
   D. Object bounding boxes.
   Answer: B. *HOG captures the distribution of gradient directions in patches of the image, which is effective for describing local shape and is used for pedestrian and object detection.*

3. **What is a common pipeline for image classification and recognition?**
   A. Edge detection → morphological transform → FFT.
   B. Feature extraction (SIFT/HOG) → coding/pooling (e.g., bag-of-features) → classifier (e.g., SVM).
   C. OCR → POS tagging → NER.
   D. Tokenization → Lemmatization → Parser.
   Answer: B. *A classic pipeline is to extract local features (SIFT, HOG), encode them (bag-of-features or dictionary model), then use a classifier (like SVM) for image category recognition.*

4. **What is transfer learning in the context of CNNs?**
   A. Transferring images from one camera to another.
   B. Using features or models learned on one task or dataset (like ImageNet) to help train a model on a new task.
   C. Converting a CNN to an RNN.
   D. Transferring weights over the network.
   Answer: B. *Transfer learning in vision means using a pretrained model (e.g., AlexNet trained on ImageNet) and fine-tuning or using its features to solve a different, but related, task, often with less data required.*

5. **Which of these is NOT a feature descriptor used in early vision?**
   A. SIFT
   B. SURF
   C. HOG
   D. R-CNN
   Answer: D. *SIFT, SURF, HOG are feature descriptors for images. R-CNN is a whole object detection framework, not a local feature descriptor.*

Session 17: Deep CNN Architectures – AlexNet, ResNet

1. **Which CNN won the 2012 ImageNet classification challenge and popularized deep learning?**
   A. LeNet
   B. AlexNet [51]
   C. VGGNet
   D. ResNet
   Answer: B. *AlexNet (2012) achieved a top-5 error of 15.3% on ImageNet, far surpassing others [51]. This demonstrated the power of deep CNNs and GPUs for vision.*

2. **What innovation did ResNet introduce?**
   A. Max-pooling layers.
   B. VGG-style networks.
   C. Residual (skip) connections to train much deeper networks.
   D. Using RGB instead of grayscale.
   Answer: C. *ResNet (2015) introduced residual skip connections, allowing the network to train much deeper architectures (dozens of layers) without vanishing gradients. (Common knowledge in DL.)*

3. **What is the ImageNet dataset?**
   A. A collection of medical images.
   B. A large labeled dataset of millions of natural images used for training/benchmarking image classifiers.
   C. A set of synthetic shapes.
   D. A video dataset.
   Answer: B. *ImageNet contains over a million labeled images in 1000+ categories. It powered the ImageNet Large Scale Visual Recognition Challenge, advancing CNN research.*

4. **What does "transfer learning" allow in CNN-based image tasks?**
   A. Training a model on GPUs only.
   B. Quickly adapting a model pretrained on one large dataset (e.g., ImageNet) to a new task with less data.
   C. Converting images to text.
   D. Communicating between different neural nets.
   Answer: B. *Transfer learning means reusing a CNN pretrained on a large dataset (like ImageNet) as a starting point for a new task, which speeds up training and often improves performance on smaller datasets.*

5. **Which activation function was used in AlexNet that helped its performance?**
   A. Sigmoid
   B. Tanh
   C. ReLU (Rectified Linear Unit) [52]
   D. Softmax
   Answer: C. *AlexNet used ReLU activations, which "trained better" than sigmoid or tanh for deep networks [52], as they reduce vanishing gradients and speed up convergence.*

Session 18: Object Detection – YOLO and R-CNN Variants

1. **What is the key idea behind the YOLO family of detectors?**
   A. Performing object proposals with selective search.
   B. Using a single neural network to predict bounding boxes and class probabilities in one pass over the image.
   C. Cascading weak classifiers.
   D. Matching templates exhaustively.
   Answer: B. *"You Only Look Once" (YOLO) frames detection as a single regression problem, predicting all bounding boxes and classes in one forward CNN pass, enabling real-time detection.*

2. **What is a distinguishing feature of the Faster R-CNN model?**
   A. It replaces the softmax with a decision tree.
   B. It uses a Region Proposal Network (RPN) to generate object proposals internally.
   C. It does not use any convolutional layers.
   D. It only works on black-and-white images.
   Answer: B. *Faster R-CNN introduced the RPN module that shares convolutional features to quickly propose regions of interest, making detection much faster than earlier R-CNN versions.*

3. **Mask R-CNN extends Faster R-CNN by adding what component?**
   A. An LSTM for sequence modeling.
   B. A parallel branch that outputs a segmentation mask for each detected object.
   C. A second RPN.
   D. Optical flow computation.
   Answer: B. *Mask R-CNN adds a mask-prediction branch to Faster R-CNN, enabling instance segmentation by generating a pixel-level mask for each bounding box.*

4. **In object detection, what is "non-maximum suppression" used for?**
   A. Merging overlapping bounding boxes to produce the final detection.
   B. Increasing image brightness.
   C. Speeding up training by skipping epochs.
   D. Splitting images into patches.
   Answer: A. *Non-maximum suppression eliminates redundant overlapping detections by keeping only the highest-confidence box for each object, a standard post-processing step in detection pipelines.*

5. **Which method involves passing multiple region proposals through a CNN individually?**
   A. YOLO
   B. R-CNN (original)
   C. Sliding window with SVM
   D. Bag-of-features
   Answer: B. *Original R-CNN (Regional CNN) method feeds each of many region proposals (from selective search) through a CNN to classify them. This is computationally expensive, which Faster R-CNN later improved on.*

---

[1] [2] [3] [4] Language - Wikipedia
https://en.wikipedia.org/wiki/Language

[5] grammar - Paradigmatic vs syntagmatic relationship - Linguistics Stack Exchange
https://linguistics.stackexchange.com/questions/31803/paradigmatic-vs-syntagmatic-relationship

[6] Levels of Linguistic Analysis to Know for Intro to Linguistics
https://library.fiveable.me/lists/levels-of-linguistic-analysis

[7] [8] [9] NLP vs. NLU vs. NLG: What's the Difference? | IBM
https://www.ibm.com/think/topics/nlp-vs-nlu-vs-nlg

[10] Tokenization in NLP - GeeksforGeeks
https://www.geeksforgeeks.org/nlp/nlp-how-tokenizing-text-sentence-words-works/

[11] Part-of-speech tagging - Wikipedia
https://en.wikipedia.org/wiki/Part-of-speech_tagging

[12] What Are Stemming and Lemmatization? | IBM
https://www.ibm.com/think/topics/stemming-lemmatization

[13] What Is Named Entity Recognition? | IBM
https://www.ibm.com/think/topics/named-entity-recognition

[14] Word-sense disambiguation - Wikipedia
https://en.wikipedia.org/wiki/Word-sense_disambiguation

[15] What is Information Extraction? | IBM
https://www.ibm.com/think/topics/information-extraction

[16] [17] What is text summarization? | IBM
https://www.ibm.com/think/topics/text-summarization

[18] CYK Algorithm for Context Free Grammar - GeeksforGeeks
https://www.geeksforgeeks.org/theory-of-computation/cyk-algorithm-for-context-free-grammar/

[19] Semantic Web - Wikipedia
https://en.wikipedia.org/wiki/Semantic_Web

[20] Markov property - Wikipedia
https://en.wikipedia.org/wiki/Markov_property

[21] Entropy (information theory) - Wikipedia
https://en.wikipedia.org/wiki/Entropy_(information_theory)

[22] What Is Support Vector Machine? | IBM
https://www.ibm.com/think/topics/support-vector-machine

[23] What is collaborative filtering? | IBM
https://www.ibm.com/think/topics/collaborative-filtering

[24] What is latent semantic analysis? | IBM
https://www.ibm.com/think/topics/latent-semantic-analysis

[25] AI vs. Machine Learning vs. Deep Learning vs. Neural Networks | IBM
https://www.ibm.com/think/topics/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks

26 27 Vanishing gradient problem - Wikipedia
https://en.wikipedia.org/wiki/Vanishing_gradient_problem

28 29 30 neural network - When to use GRU over LSTM? - Data Science Stack Exchange
https://datascience.stackexchange.com/questions/14581/when-to-use-gru-over-lstm

31 32 Text Preprocessing in NLP - GeeksforGeeks
https://www.geeksforgeeks.org/nlp/text-preprocessing-for-nlp-tasks/

33 34 Natural Language Processing: NLTK vs spaCy - ActiveState
https://www.activestate.com/blog/natural-language-processing-nltk-vs-spacy/

35 Converting Speech to Text with Spark NLP and Python - John Snow Labs
https://www.johnsnowlabs.com/converting-speech-to-text-with-spark-nlp-and-python/

36 37 38 Word Embeddings in NLP - GeeksforGeeks
https://www.geeksforgeeks.org/nlp/word-embeddings-in-nlp/

39 What is Sentiment Analysis? | Definition from TechTarget
https://www.techtarget.com/searchbusinessanalytics/definition/opinion-mining-sentiment-mining

40 Comparison Between CBOW and Skip-Gram Models - GeeksforGeeks
https://www.geeksforgeeks.org/nlp/word-embeddings-in-nlp-comparison-between-cbow-and-skip-gram-models/

41 GloVe - Wikipedia
https://en.wikipedia.org/wiki/GloVe

42 43 seq2seq Model - GeeksforGeeks
https://www.geeksforgeeks.org/machine-learning/seq2seq-model-in-machine-learning/

44 45 46 47 BERT (language model) - Wikipedia
https://en.wikipedia.org/wiki/BERT_(language_model)

48 Phonetics | Definition, Types, Examples, & Facts | Britannica
https://www.britannica.com/science/phonetics

49 Nyquist–Shannon sampling theorem - Wikipedia
https://en.wikipedia.org/wiki/Nyquist%E2%80%93Shannon_sampling_theorem

50 Kernel (image processing) - Wikipedia
https://en.wikipedia.org/wiki/Kernel_(image_processing)

51 52 AlexNet - Wikipedia
https://en.wikipedia.org/wiki/AlexNet