

Questions & Answers on Chapter 1

1. List examples of real-world applications of NLP.

- Core apps:
 - ✉ Email spam filters
 - 🗣 Voice assistants (Alexa/Siri)
 - 🔍 Search engines (Google/Bing)
 - 🌐 Machine translation (Google Translate)
- Other apps:
 - 📱 Social media sentiment analysis
 - 🛒 E-commerce product tagging
 - 🏥 Healthcare report generation
 - 📄 Plagiarism detection (Turnitin)

2. Explain the following NLP tasks:

- Language modeling 📖 : Predict the next word (e.g., phone keyboard suggestions).
- Text classification 📁 : Sort text into categories (e.g., spam vs. not spam).
- Information extraction 🔍 : Pull key details (e.g., dates from emails).
- Information retrieval 📖 : Find docs matching a query (e.g., Google Search).
- Conversational agent 💬 : Build chatbots (e.g., Siri).
- Text summarization ✂ : Shorten long articles.
- Question answering ❓ : Answer questions (e.g., Watson AI).
- Machine translation 🌐 : Convert text between languages.
- Topic modeling 📁 : Group docs by themes (e.g., news categories).

3. What are the building blocks of language and their applications?

- 🗣 Phonemes: Sounds (used in speech-to-text).
- 📖 Morphemes: Smallest meaningful units (used in tokenization).
- 📐 Syntax: Sentence structure rules (used in parsing).
- 🌐 Context: Meaning from world knowledge (used in summarization).

4. Why is NLP challenging?

- 😲 Ambiguity: Words have multiple meanings.
- 🧠 Common knowledge: Machines lack human-like reasoning.
- 🎨 Creativity: Idioms, slang, poetry confuse models.
- 🌐 Diverse languages: Rules vary across languages.

5. How are NLP, ML, and DL related?

ML => Branch in AI that deal with algorithm that can learn to preform tasks
DL=> Branch in ML that based on ANN architecture
NLP=> use ML and DL to process language

- 🤖 AI → 🧠 ML (learns from data) → 🧠🔗 DL (neural networks).
- NLP uses ML/DL to process language.

6. What is heuristics-based NLP?

- 📄 Rule-based systems (e.g., regex for emails).
- Example: Count "positive" words 😊 to guess sentiment.

7. Explain Naïve Bayes, SVM, HMM, and CRF approaches.

- 📊 **Naive Bayes**: Uses word probabilities (fast but assumes independence).
 - 🪓 **SVM**: Draws boundaries between classes (robust but slow).
 - 🔄 **HMM**: Models hidden states (e.g., part-of-speech tagging).
 - 🏷️ **CRF**: Tags sequences (e.g., entity extraction).
-

8. What is the difference between RNN and LSTM?

- 🔄 **RNN**: Processes sequences but forgets long contexts.
 - 🧠💾 **LSTM**: Remembers important info, forgets irrelevant details.
-

9. How can CNN be used for text processing?

- 🖼️ Treats text as a matrix (words = rows, embeddings = columns).
 - Uses filters to detect word patterns (e.g., "very good" in reviews).
-

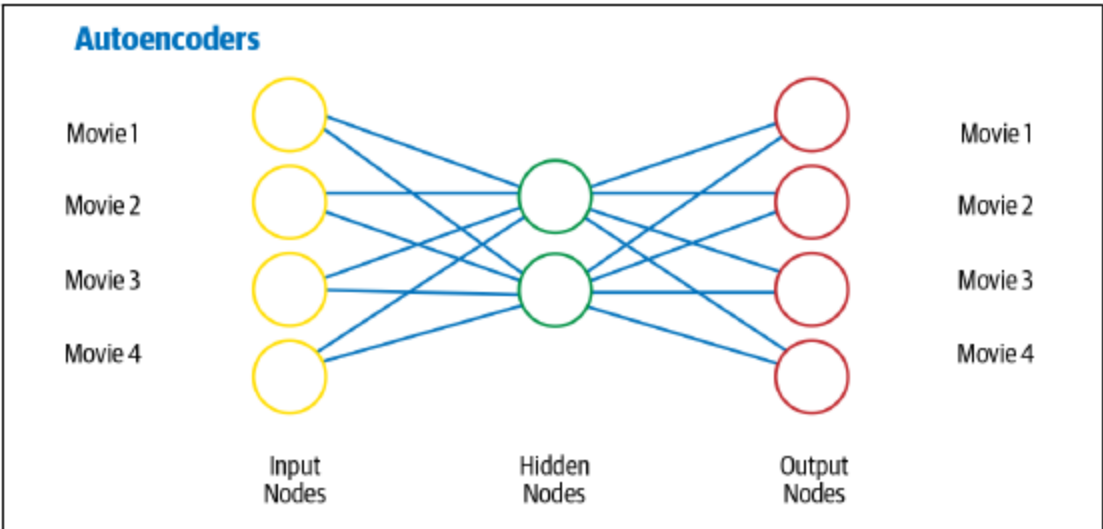
10. What is transfer learning?

- 🔄➡️🎯 Pre-train on big data (e.g., BERT), then fine-tune for specific tasks.

Transfer learning is a technique in AI where the knowledge gained while solving one problem is applied to a different but related problem

11. What is the architecture of an autoencoder?

- 🧑🏠♂️ **Input** → **Encoder** (compresses) → **Latent vector** → **Decoder** (reconstructs).



12. Why is deep learning not suitable for all NLP tasks?

- 📊 Needs huge data (overfits small datasets).
 - 🌐 Fails with domain shifts (e.g., medical vs. social media text).
 - 🤖 Black box (hard to explain decisions).
 - 💰 Expensive (GPUs, time, \$\$\$).
 - 🧑🏠♂️ Lacks common sense.
-

13. Explain the flow of conversational agents.

1. 🗣️➡️📄 **Speech-to-text**: Convert voice to words.
2. 🧐 **NLU**:
 - 😬 Sentiment analysis
 - 🏷️ Entity extraction
3. 🗣️ **Dialog management**:
 - Classify intent (question vs. command).

4. 🎵 **Response generation:**
- Use templates or retrieve info (e.g., "Playing song X").

إِنَّ اللَّهَ وَمَلَائِكَتَهُ يُصَلُّونَ عَلَى النَّبِيِّ يَا أَيُّهَا الَّذِينَ آمَنُوا صَلُّوا عَلَيْهِ وَسَلِّمُوا تَسْلِيمًا (56)