**Weekly Reinforcement Table for Clare**

| **Weeks** | **Topic** | **3 Core Questions** | **Key Knowledge Points (for spaced repetition)** |
| --- | --- | --- | --- |
| 0 | Welcome | 1) What is Generative AI? 2) What are the course outcomes (LO1–LO5)? 3) What’s expected in the final deliverables? | • Course LO1–LO5 overview  • Why GenAI matters in real apps |
| 1 | Foundations of GenAI | 1) What defines an LLM? 2) How does the Transformer/self-attention work? 3) Why is pre-training + fine-tuning effective? | • Self-Attention & Transformer basics  • Perplexity (PPL) & eval notions |
| 2 | Understanding Foundation Models | 1) What makes a model “foundation”? 2) Domain-specific vs general-purpose models—trade-offs? 3) Scaling challenges? | • Multimodal FMs (e.g., CLIP, DALL·E)  • Data scale, bias & risks |
| 3 | Choosing Pre-trained Models | 1) Open-source vs proprietary—when to pick which? 2) Cost/latency vs accuracy trade-offs? 3) Why context length matters? | • Decision framework for LLM choice  • HF Hub & integration paths |
| 4 | Prompt Engineering | 1) Core principles for effective prompts? 2) Zero-shot, few-shot, CoT—when to use? 3) Prompt injection/jailbreak defenses? | • ReAct vs CoT differences  • Defensive prompt patterns |
| 5 | Building a Simple Chatbot | 1) What components form a bot pipeline? 2) How does memory improve dialogue? 3) Why use LangChain abstractions? | • Memory (short- vs long-term)  • Chains/Agents + Streamlit UI |
| 6 | Review Week | 1) Which prior module is your weakest? 2) How do model choice & prompting interplay? 3) What cross-module links matter most now? | • Cross-module consolidation  • Mid-course self-check prompts |
| 7 | Retrieval-Augmented Generation (RAG) | 1) Why does RAG improve factuality? 2) How do embedding-based retrieval & similarity work? 3) How to measure retrieval (precision/recall)? | • Hybrid retrieval (term + embedding)  • Indexes: IVFFlat vs HNSW |
| 8 | Agent & Agentic RAG | 1) What are an agent’s core components? 2) Why do agents need stronger models? 3) What is knowledge augmentation (vector DB)? | • Planning stages & tool use  • Action/extension tools & examples |
| 9 | Evaluating GenAI Apps | 1) What is “hallucination” & how to mitigate? 2) Bias/fairness—how to assess? 3) What metrics for GenAI evaluation? | • Evaluation patterns & “AI as judge”  • Human/auto mixed eval |
| 10 | Responsible AI Development | 1) Why Responsible AI? 2) Key risks (bias, privacy, misinformation)? 3) How do model cards/docs help governance? | • EU AI Act / OECD principles  • Prompt-level guardrails & RLAIF/CAI |

**Notes on Table Design:**

* Displayed as “tips/flashcards” beneath the chatbot dialogue.
* **Exam and project modules (11–14) have been removed**; only Modules 0–10 are included, with Week 6 reserved as a **consolidated review module**.
* **Citations in the table**: Course syllabus for general structure; Foundations / Model Selection / Prompting / RAG / Agents / Evaluation / Responsible AI sections reference the textbooks and Module discussions.
* **Weighting Spaced Repetition：**

1. **Current week:** **8%**; W-1: **12%**; W-2: **20%**; W-3: **25%**; W-4: **35%**
2. **Default weekly review (T+7):** Current week 8% → W-4 week 35%.
3. **Bi-weekly review (T+14) and Stage/monthly review (T+30) use adjusted distributions:**
   * **T+14:** W-6–5: 20% each, W-4–3: 15% each, W-2–1: 10% each, Current: 10%.
   * **T+30:** W-8–5: 15% each, W-4–3: 10% each, W-2–1: 7.5% each, Current: 5%.