Assignment 18
Python & ML Essentials for GenAI Engineers

Part E — Engineering Essentials

16 questions

★	**Must-know** — asked very frequently in GenAI / ML interviews (8 questions). Prepare these first.
[CODE]	**Hands-on task** — write and run the code (4 tasks).

1. What is the difference between multithreading and multiprocessing in Python? What is the GIL, and how does it affect GenAI backends? ★ · Asked: TCS, Infosys
2. Concurrency vs parallelism: why use async programming in an AI application? If you must call three independent APIs, would you call them one by one? ★ · Asked: Infosys
3. What are generators and decorators in Python? How do generators relate to streaming LLM responses? ★ · Asked: TCS, Google
4. How do you process very large data (e.g. millions of rows or huge files) in Python without Out-Of-Memory errors? · Asked: TCS
5. Write a Python function that calls an LLM API with exponential backoff retries (with jitter) for transient errors. [CODE] · Asked: Wipro, EPAM
6. Implement an in-memory cache (with TTL) for LLM queries so identical requests don't call the API again. [CODE] · Asked: Wipro
7. Why is FastAPI preferred for GenAI backends, and how does Pydantic fit in? ★ · Asked: EPAM, Google
8. Quick-fire Python: list vs tuple, shallow vs deep copy, and what is a context manager? · Asked: Google
9. Pandas: given a DataFrame with columns 'Student' and 'Marks', mark students with marks below 30 as 'Fail' (others 'Pass') in a new column 'Result', and list the failed students. [CODE] · Asked: PwC
10. SQL: find the second-highest salary, and find duplicate records in a table. [CODE] · Asked: Google
11. Precision vs recall vs F1 score – when is recall more important than precision? ★ · Asked: Google
12. What is overfitting, and what is data leakage? How do you prevent each? ★ · Asked: Google
13. How do you handle class imbalance (e.g. 1% fraud vs 99% normal transactions)? ★ · Asked: NLP interview rounds, Google

14. How do you detect and handle outliers before feeding data into ML or embedding models? · Asked: PwC
15. Why are traditional ML concepts (classification vs regression, loss functions, metrics) still asked in GenAI interviews? · Asked: PwC
16. What is data drift vs model (concept) drift vs training–serving skew? Batch vs online inference? ★ · Asked: Google