

**Sharda University**

## **5 Levels of Text Splitting**

Name: **Aamir Neyazi**

Enrollment No: **2023512205**

**Level 1: Character Splitting**

Text is divided into fixed-size character chunks. Simple and fast but may break meaning.

**Level 2: Recursive Character Splitting**

Text is split using multiple separators like paragraphs, sentences, and words to preserve structure.

**Level 3: Document Specific Splitting**

Chunking depends on document type such as PDF sections, code functions, or markdown headers.

**Level 4: Semantic Splitting**

Embeddings are used to group text with similar meaning into coherent chunks.

**Level 5: Agentic Splitting**

An AI agent decides how to split text based on task goals and retrieval performance.

**Workflow Summary:**

Raw Document → Splitting Strategy → Chunks → Embeddings → Retrieval → Language Model