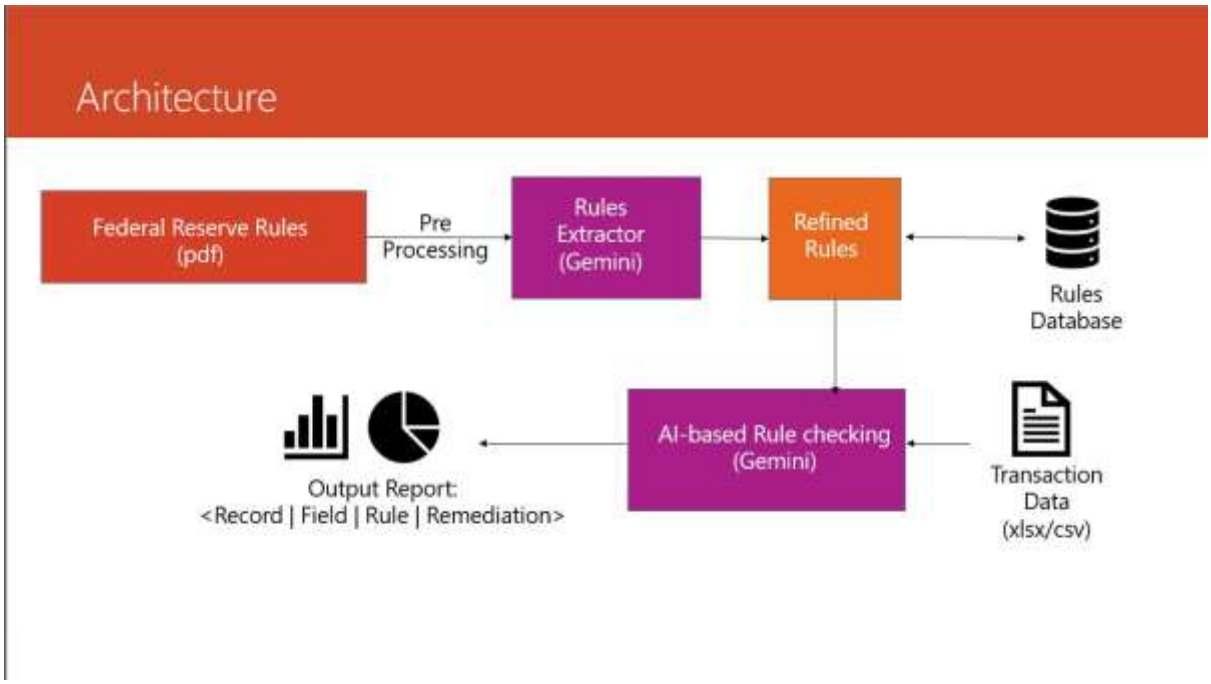
Architectural Diagram :-



Explaination:-

1. Federal Reserve Rules (Input Regulations)

- Regulatory authorities provide rules that define the required fields for various reporting schedules (H1, H2, H3, and B2).
- These rules are detailed in textual format and typically provided as PDF documents.
- The rules serve as a reference to ensure that all transactional data adheres to the specified compliance guidelines.

---

2. Rules Extractor (LLM-Powered Rule Extraction)

- The first LLM model (Gemini) in the system processes the input rules PDF.
- It extracts relevant business and regulatory rules and converts them into a structured, predefined JSON format.
- This structured format ensures that rules are machine-readable and can be efficiently applied to data validation.

---

3. Refined Rules (Processed Regulatory Rules)

- The extracted rules from the Rules Extractor serve as input for the next stage of processing.
- If a PDF is not provided, users can manually retrieve rules from the database.
- These refined rules act as a direct input to the AI-based Rule Checking model.

---

4. Rules Database (Storage and Retrieval System)

- The structured rules are stored in a JSON-based database.
- Users can retrieve rules on demand, eliminating the need to upload the PDF every time.
- This database enables quick rule retrieval for different financial schedules without reprocessing the original PDF.

---

5. Transaction Data (Input for Rule Validation)

- This is the client-provided dataset that needs to be validated.
- It contains transactional records with fields corresponding to the extracted rules.
- The same field structure is maintained to ensure seamless rule application.

---

6. AI-Based Rule Checking (Automated Validation Engine)

- The primary LLM-based model checks the given transactional data against the extracted rules.
- Users can either:
    - Manually select a schedule (H1, H2, H3, or B2) for validation.
    - Allow the system to automatically select the closest matching schedule based on field similarity.
- Every record is evaluated field-by-field, ensuring strict rule adherence.
- Any deviation or mismatch between the data and the rules is flagged as an anomaly.

---

7. Output Data (Anomaly Detection & Reporting)

- A CSV file is generated containing the results of the validation process.
- The output file includes:
    - Record number (for tracking the problematic entry).
    - Field value that does not comply with the rules.
    - Description of the rule that has been violated.
    - Potential remediation steps to correct the anomaly.
- This structured output enables easy review and correction of non-compliant data.