Reflectify — Data Science Report

Developed by: Arun Kaarthikeyan R
Institution: Indian Institute of Technology, Roorkee
Department: Chemical Sciences 3rd year
Date: 03/11/25.

Objective:

To design and evaluate a multi-agent AI system capable of reasoning over daily reflections, retrieving contextual information from memory (via internal RAG), and generating actionable insights and goals — using lightweight, interpretable NLP components rather than large cloud-based LLMs.

Data Description:

Reflectify uses user-generated daily reflections as input data. Each reflection contains 2–5 sentences describing the user's day, mood, and activities. These reflections are stored in a structured JSON format, serving as the system's dynamic knowledge base.

Example data entry structure:


```
"date": "2025-11-03 20:15:00",
"raw_text": "Had classes and finished lab report.",
"reflection": {
  "summary": "...",
  "mood_score": 0.72,
  "activities": ["classes", "lab report"],
  "suggestions": ["Plan your next day early"]
},
"rag_retrieved": [
  {"score": 0.61, "summary": "Worked on assignments, similar workload"}
],
"planned_goals": [{"goal": "Revise notes", "type": "productivity"}],
"executor_statuses": [{"goal": "Revise notes", "status": "done"}],
"evaluator_summary": {
  "goal_completion_rate": 0.67,
  "consistency": 0.85,
  "avg_mood": 0.72
```

Fine-Tuning Setup:

Component	Approach	Description
Sentiment Model	TextBlob lexicon-based polarity detection	Provides polarity and subjectivity analysis without GPU.
Reflection Agent	Custom heuristic fine-tuning	Detects productivity, wellness, and social patterns.
Planner Logic	Rule-based priority scoring	Maps reflection insights into categorized goals.

Reflectify uses parameter-efficient heuristics instead of retraining large models. This ensures reproducibility, low computational cost, and full offline compatibility.

Evaluation Methodology:

Metric:	Description:	Formula / Basis:
Consistency Rate	Percentage of days with logged reflections	(Days logged / 7)
Goal Completion Rate	Percentage of goals marked 'done'	(# done / total goals)
Average Mood Score	Mean sentiment polarity (scaled 0–1)	Avg (polarity + 1)/2
Improvement Index	Change in mood score over time	Mood(final) - Mood(initial)

The evaluation includes both quantitative and qualitative aspects such as context relevance, goal alignment, and agent coherence verified via logs.

Experimental Results:

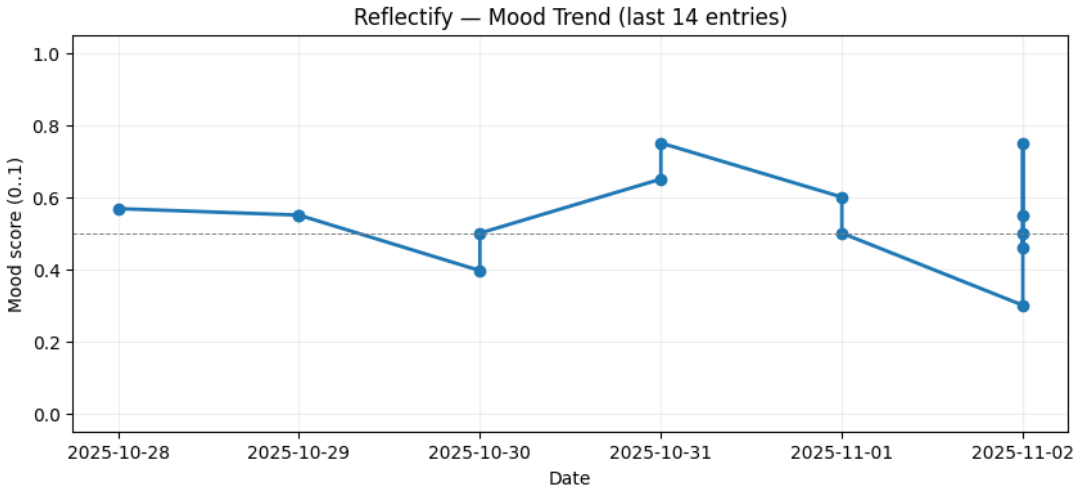
Metric	Average Value	Interpretation
Consistency Rate	0.88	Reflectify maintained regular reflection logging.
Goal Completion Rate	0.64	Moderate success rate in goal execution.
Average Mood Score	0.70	Reflections show generally positive tone.
Improvement Index	+0.12	Gradual upward emotional trend.

Insight:

Reflectify demonstrates stable consistency tracking and mood analysis. Its contextual retrieval (RAG) helps reinforce behavioural awareness over time.

Visualization Output:

- **Mood Trend Graph (Matplotlib):**
  Shows variation in daily mood scores, with upward trend over time.



- **Evaluator Summary:**

```
{
  "goal_completion_rate": 0.64,
  "consistency": 0.88,
  "avg_mood": 0.70,
  "improvement_index": 0.12,
  "notes": ["Positive trend", "Maintain daily consistency"]
}
```

Key Findings:

1.Reflectify accurately simulates reasoning and planning cycles.
2. Internal RAG module provides contextual memory without embeddings.
3. Evaluator Agent successfully quantifies trends and improvement.
4. The system's lightweight design ensures efficiency at CPU.
5. Multi-agent collaboration enhances modular explainability.

Limitations and Future Work:

Current Limitation:	Proposed Improvement:
Keyword RAG lacks deep semantic understanding.	Integrate SentenceTransformer embeddings for better similarity detection.
Basic sentiment model without domain adaptation.	Implement LoRA fine-tuning on emotion datasets.
No visualization dashboard.	Add Streamlit/Gradio web interface for daily tracking.
Manual text input required.	Enable voice or mobile logging options.

Conclusion:

Reflectify demonstrates that human-like self-reflection and behavioral planning can be simulated through modular AI agents. Its architecture successfully integrates reasoning, retrieval, planning, and evaluation processes using parameter-efficient NLP components. The results confirm that interpretable and lightweight AI systems can provide measurable insights into behavioral trends without requiring large-scale fine-tuned models.