ML Lab Research Proposal

Team Name: Project Mnemosyne
Lead Researcher: Andre Salehi
Date: 5/30/26

1. Problem Statement:
Human memory is selective rather than exhaustive. People rarely remember every detail of a visual scene, yet often retain information that later proves relevant. In contrast, many modern visual AI systems operating under memory constraints rely on recency-based retention, attention scores, or fixed memory allocation strategies. These approaches prioritize information that appears important in the present but may fail to preserve observations that become important later. This project investigates whether a visual AI system can learn a human-inspired memory strategy by predicting which observations are likely to become useful in the future and preferentially allocating memory to those observations. The central research question is:
Can an AI system learn what is worth remembering before it becomes important?
Answering this question may improve memory allocation in visual AI systems while also providing insight into how selective memory mechanisms can be modeled computationally.

2. Proposed Approach:
We propose a Future Utility Memory (FUM) module that assigns incoming visual observations a predicted future usefulness score. Rather than storing observations based solely on recency or attention weight, the system will estimate the probability that an observation will contribute to successful completion of a future task.
When memory capacity is reached, observations with lower predicted future utility will be discarded while higher-utility observations are retained.
The central hypothesis is:
Memory allocation based on predicted future utility will preserve more task-relevant information than existing retention strategies when operating under fixed memory budgets.
Performance will be compared against several baseline memory policies:

- FIFO (First-In, First-Out) retention
- Random retention
- Attention-based retention
- Recency-based retention

The goal is not to maximize memory size, but to maximize downstream task performance given a fixed memory budget.

3. Dataset:
Experiments will be conducted across multiple visual domains to evaluate whether future-utility memory generalizes across different environments.
Candidate datasets include:

- BDD100K driving sequences
- Something-Something V2 video understanding dataset
- Custom synthetic environments designed specifically for memory evaluation

Video frames will be converted into visual embeddings and paired with downstream tasks requiring information from earlier observations.
An observation will be considered useful if information contained within that observation contributes to successful completion of a future task while operating under constrained memory capacity.
Using multiple domains allows evaluation of whether future-utility memory captures a general memory principle rather than overfitting to a single application area.

4. Evaluation Metrics:
Performance will be evaluated using:

- Downstream task accuracy
- Future-useful observation retention rate
- Memory efficiency (task performance per stored observation)
- Precision and recall of retained observations
- Relative improvement over baseline memory strategies

Experiments will be conducted across multiple memory budget settings to evaluate robustness under increasing memory constraints.

5. Compute Requirements:
The project is expected to require:

- One modern consumer-grade GPU
- Standard PyTorch-based tooling
- Less than 12 GB GPU memory
- Approximately 1–3 days of training time depending on dataset and model configuration

No specialized hardware is expected to be required.

6. Expected Deliverables:
The project will produce:

- A working Future Utility Memory implementation
- Benchmark comparisons across multiple memory strategies
- Quantitative evaluation results
- Visual demonstrations of memory allocation decisions over time
- Open-source code repository
- Research paper describing methodology, experiments, and findings

7. Risk Assessment and Mitigation:
The proposed memory strategy may fail to outperform existing baselines.
If this occurs, the project will focus on identifying conditions under which different memory policies succeed or fail. Such results would still provide valuable empirical evidence regarding memory allocation in visual AI systems and establish a benchmark framework for future work.
Even a negative result would contribute useful information about the relationship between future prediction and memory retention, making the project scientifically valuable regardless of outcome.