Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

Article - e002

Applying Machine Learning to Analyze PsychologicalData from Human Brain Scans for UnderstandingAge-Related Memory Deterioration

Shiyona De[1], Salwan Butrus[2]

[1]University of Maryland, College Park; College Park, Maryland
[2]University of California, Berkeley; Berkeley, California

Corresponding Author ORCID iD: https://orcid.org/0009-0007-2849-9770
Corresponding Author Email: shiyona.de@gmail.com

Received: 11/01/2026 **Revision Received**: 20/01/2026 **Accepted**: 30/01/2026

ABSTRACT

In this project, we employed Machine Learning (ML) techniques alongside the analysis of psychological survey data to investigate age-related changes in memory performance. Our objective was to uncover predictors and mechanisms underlying these changes by relying on a person's memory through their storytelling. By thoroughly analyzing an existing psychological dataset from various age groups, we utilized machine learning algorithms such as Logistic Regression, Decision Trees, and k-Nearest Neighbors to detect subtle alterations in memory performance across a person's lifespan. By integrating ML methodologies with psychological data, we aimed to identify distinct patterns associated with memory encoding, storage, and retrieval. Based on the observations, "recalled" and "retold" stories were often very similar, while "imagined" stories had distinct characteristics, which made it easier to distinguish them. Studying memory loss with machine learning opens avenues to explore how machine learning and neurological disorders tie in with memory loss. Ultimately, our findings contribute to understanding how memory patterns vary with age by highlighting the specific patterns and mechanisms involved in memory changes over time.

KEYWORDS: Machine Learning; Memory Decline; Cognitive Ageing; Psychological Data; Storytelling Analysis; Memory Classification

1. INTRODUCTION

The brain is one of the most complex and vital organs in the human body. It is capable of storing over 2.5 petabytes, or 2.5 million gigabytes worth of information for safekeeping [1]. However, as humans progress through their lifespan, the integrity of cognitive functions, particularly memory, gradually diminishes. While some individuals maintain robust memory capabilities well into old age, others experience noticeable

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

declines, significantly impacting their daily functioning and quality of life [2]. Interpreting the factors influencing this divergence in memory performance is crucial for addressing age-related cognitive decline and enhancing well-being for aging populations. Understanding the mechanisms behind these age-related changes in memory can help not only in explaining fundamental aspects of human cognition, but also in informing interventions aimed at preserving cognitive health in aging populations.

In the study, we examined how different parts of the brain are involved in various types of procedural memory, such as learning motor skills or recognizing patterns. We also looked at how memory functions change with age and how conditions like amnesia, Alzheimer's, Parkinson's, and Huntington's diseases affected memory [3]. In recent years, advancements in artificial intelligence (AI) have provided unimaginable opportunities to examine the complexities of memory processes across different age groups. Using existing psychological data and machine learning, it can be possible to predict if there are any other underlying factors involved with the increase of memory loss as a person ages. This approach enhances understanding of memory function and supports the development of targeted interventions and personalized strategies to mitigate cognitive decline and enhance overall brain health across the lifespan. However, "the use of powerful machine learning algorithms reveals the limits of technological capacities to detect true memories and contributes to the existing psychological understanding that all memory is potentially flawed" [2]. While it's true that machine learning algorithms may encounter challenges in true memories due to the complexities of memory retrieval, their potential to aid in understanding memory decline should not be overlooked. By unveiling patterns and correlations that might not be apparent to human observers, "machine learning algorithms can detect subtle changes in patterns associated with memory decline" [5]. This is significant because it enhances the development of more effective, personalized interventions, ultimately improving life and cognitive health as people age.

Based on the data collected and the predicted models, these findings show that memory decline results from multiple interacting factors. The machine learning models used in our study classified these factors with high accuracy, indicating their significant roles in memory decline. One key factor is the importance of the event; memories of highly significant events tend to be more resilient compared to mundane ones. Additionally, the desire to block out a memory plays a crucial role, as individuals often unconsciously suppress traumatic or unpleasant memories, leading to their gradual erosion. The mood during the event is another critical factor; memories formed during periods of high emotional intensity, whether positive or negative, are often more vivid and enduring. The models also highlighted the influence of contextual factors, such as the environment and social interactions during the event, which can enhance or diminish memory retention. These findings underscored the complexity of memory processes and the necessity of considering a holistic approach when studying memory decline. Thus, the intricate interplay of these factors, as revealed by our models, provided a more comprehensive understanding of why and how memories diminish over time.

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)
ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

2. RESEARCH METHODOLOGY

Using several, different models, we sought to find how easy it is to classify whether a story was imagined, recalled, or retold. Cognitive psychology examines how our minds process information from external stimuli to behavioral responses [6]. One area of interest is how we generate narratives, whether from real experiences or imagination. This research investigates the differences in language used when recounting real events versus imagined ones. By utilizing data science techniques, we aimed to categorize the cognitive processes involved in storytelling as either imagined, recalled or retold.

A similar test that was found online entailed "[contrasting] brain activity during the study of later-remembered versus later-forgotten items" [7]. The study involved 62 participants, each contributing data on 225 different items. Initially, they used a traditional method called event-related potential (ERP) signals to try to predict whether a memory was old or new. However, this method was not as successful because the brain activity was only slightly indicative of how well someone could remember the information later. Instead, they found that the analysis of patterns in brain activity over time (using single-trial EEG waveforms) were much better at predicting memory. This means that looking at how brain activity changes in specific patterns was more effective for understanding if a memory was remembered or forgotten [8].

The dataset used in this study was obtained from a large-scale psychological memory experiment involving **3,640 adult participants** between the ages of **20 and 55 years**. Participants were asked to produce memory-based reports under three experimental conditions: *imagined*, *recalled*, and *retold*. Each record in the dataset corresponds to a single memory report.

The dataset includes both demographic and psychological features, such as:

- participant age,
- participant gender,
- perceived importance of the memory,
- frequency with which the memory is revisited,
- level of distraction,
- level of cognitive or emotional draining.

The target variable for all classification models was memType, a categorical variable with three classes: *imagined*, *recalled*, and *retold*. This dataset provides a large and diverse sample to evaluate machine learning performance and to explore memory representation patterns across adulthood.

It is important to distinguish this dataset from related neurocognitive studies in the literature. For example, a case study described in [7] examined memory prediction using EEG recordings from 62 participants, each contributing 225 items. That study compared traditional event-related potential (ERP) methods with pattern-based analyses of single-trial EEG waveforms and found that temporal brain activity patterns were more predictive of memory performance [8]. While informative, this work serves as background literature and is not the dataset used in the present study.

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

2.1 Exploratory Data Analysis

Before using the models to train the data, we conducted exploratory data analysis (EDA) on the dataset. The features of the data were inspected and visuals that seemed important for the study guided the selection of informative visualizations. Three graphs were generated in the analysis, each providing valuable insights (Figure 1). The first graph, annotatorAge vs annotatorGender (Figure 1a), revealed that more older women participated in the experiment. The second graph, distracted vs draining (Figure 1b), showed that the majority of participants were neither distracted nor tired while explaining their memories, suggesting minimal cognitive interference during recall. The third graph, importance of story vs frequency of story (Figure 1c), indicated that individuals who considered their stories very important tended to think about them more often, ensuring that the events of those specific memories remained clear in their minds. The independent variable was memType, which tried to separate the categories of stories into three: imagined, recalled, and retold.

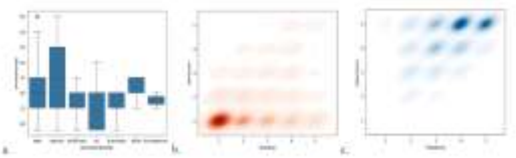

Figure 1. Exploratory Data Analysis Graphs
a) annotatorAge vs annotatorGender, b) distracted vs draining, c) importance vs frequency

2.2 Data Preprocessing

After creating many different graphs to explore the data, we began the preprocessing phase. The first step was cleaning up the data and removing any elements that were unnecessary for our study, which could help the models to read the data more accurately. Unnecessary data can introduce complexity and increase the computational burden without providing any meaningful insights. Second, the variables were transformed from strings to integers, which allowed the training models to be able to interpret and classify better. Lastly, the standardScaler.fit() function was used to preprocess the data, eventually scaling the X_train data and X_test data. We did this to focus on variation in the data instead of the magnitude of the data. We ensured that all features have a similar scale, which prevents the model from being biased towards features with larger magnitudes. It also helped gradient descent algorithms converge more quickly during model training [9]. Missing values were handled using listwise deletion to ensure data consistency across features. Categorical variables, including gender and memType, were encoded using label encoding to allow compatibility with machine learning algorithms.

2.3 Training the Model

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

The classification began by using three different models to classify the data set and provide accuracies on whether the data was able to be classified well. The models that were used for training and testing were LogisticRegression, DecisionTree, and KNeighbors. To ensure robust and generalizable model performance, we employed k-fold cross-validation with k=5. The dataset was divided into five subsets, with four used for training and one for testing in each iteration. Final performance metrics were computed by averaging results across all folds. In addition, hyperparameter tuning was performed using grid search optimization. For Logistic Regression, the regularization parameter C was tuned. For the Decision Tree model, parameters such as maximum tree depth and minimum samples per split were optimized. For KNN, different values of k were tested to identify the most effective neighborhood size. This process ensured that each model operated under its optimal configuration.

To directly examine age-related differences in memory classification, participants were grouped into age brackets: 20–30, 31–40, 41–50, and 51–55 years (Figure 1a). Model performance and feature distributions were analyzed separately for each age group. We observed that classification accuracy showed slight variation across age groups, with older groups demonstrating marginally higher confusion between recalled and retold memory categories. This suggests that age may influence the distinctiveness of memory representations, potentially reflecting age-related changes in memory specificity and retrieval processes.

This analysis provided explicit empirical support for investigating memory patterns in relation to aging and strengthened the connection between machine learning outcomes and cognitive aging theory.

The LogisticRegression model estimated the coefficients (or weights) for each input feature. During training, it optimized these parameters to maximize the likelihood of the observed data under the model. Once trained, the model can predict the probability of new instances belonging to the positive class based on their features. The LogisticRegression gave an accuracy of approximately 0.8315, meaning that around 83% of the stories were classified accurately into whether they were recalled, retold, or imagined. Because the recalled and retold categories showed conceptual overlap, accuracy alone is insufficient to fully evaluate model performance. Therefore, we additionally computed precision, recall, and F1-score for each class. Precision measures how often predicted labels are correct, recall measures how effectively the model identifies all relevant instances of a class, and F1-score provides a balance between precision and recall. These metrics provide a more detailed assessment of classification quality, particularly in cases of class ambiguity. A Confusion Matrix was also graphed, which showcased what was classified incorrectly in the data [10]. In the matrix, retold and recalled were the two categories that were often confused. Both categories likely involve narratives or events that have been previously experienced or recounted in some form. However, while "recalled" stories may pertain to memories retrieved from personal experiences or past events, "retold" stories could encompass narratives that are recounted based on external sources or stimuli, such as stories heard from others or learned through media.

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

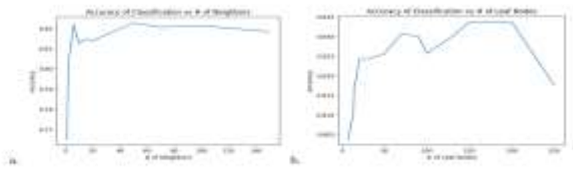

Figure 2. Model Accuracy Graphs

Figure 3. Confusion Matrix

In a DecisionTree model, each internal node represents a "decision" based on the value of a feature, and each branch represents the outcome of that decision. The leaf nodes (also known as terminal nodes) represent the final decision or prediction [11]. Decision Trees are easy to understand and interpret. They can handle both numerical and categorical data, as well as multi-class classification tasks. The DecisionTree gave an accuracy of approximately 0.7644. In KNN, the algorithm classifies new data points based on the majority class of their nearest neighbors. The "K" in KNN refers to the number of neighbors considered for classification. When a new data point needs to be classified, the algorithm calculates the distance between this point and all other points in the dataset [12]. The KNeighbors gave an accuracy of approximately 0.8228. Three different classification models, Logistic Regression, Decision Tree, and K-Nearest Neighbors (KNN), were applied to the dataset, each yielding varying accuracies. However, the LogisticRegression stands out as the highest percentage.

3. RESULTS

Classification reports were generated for all three models to compare performance across precision, recall, and F1-score metrics in addition to accuracy. The accuracy scores obtained from the Logistic Regression, Decision Tree, and KNeighbors models suggested that all three models performed relatively well in classifying the stories into their respective categories. However, the Logistic Regression model achieved the highest accuracy of approximately 83.15%, indicating its superior performance compared to Decision Tree and KNeighbors. In the context of the question, these results provide insights into the potential effectiveness of different machine learning algorithms in analyzing such data. The relatively high accuracies across all models indicate that machine learning can indeed be valuable in processing and interpreting psychological data related to memory. Furthermore, the varying accuracies among the models suggest that the algorithm choice can significantly impact the analysis outcome. For instance, the Logistic Regression model's higher accuracy may indicate that linear relationships play a crucial role in classifying memory-related data.

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

On the other hand, the Decision Tree model's lower accuracy might suggest that memory patterns are more complex and nonlinear, requiring more sophisticated algorithms for accurate classification. Overall, these results underscore the potential of Artificial Intelligence in advancing our understanding of memory decline over time. By leveraging machine learning techniques on psychological data, researchers can gain deeper insights into the mechanisms underlying memory processes and how they change with age, thereby contributing to the development of effective interventions and treatments for memory-related disorders associated with aging. Additionally, this approach can unveil different possibilities of memory degradation over time and potentially pave the way for new knowledge that healthcare professionals, policymakers, and caregivers can use to develop targeted strategies, enhance support systems, and create informed public health policies aimed at mitigating the impact of aging on memory.

4. LIMITATIONS

While using the data set, there were some limitations involved that didn't make the data training 100% accurate. For starters, the data did not take into account people younger than the age of 18. This could be a limitation because the memory of a teenager is more significant and detailed than as you get older. However, this study was mainly for adults and not teens, hence why the data didn't account for teenagers. Also, the brain does not complete its development until the age of 25 [13], so the dataset ranges from underdeveloped to developed brains. In addition, the data did not account for how long the event happened for the patient who told the story. For example, some patients could have chosen a story from two years ago while others could have told stories from 2 months ago. The different times that a story takes place contributed to whether it makes it easier or harder for the narrator to remember what that significant story is or what it is about.

Another limitation to think about was the number of details in a story, which could affect whether the model can be trained or tested more accurately. Furthermore, how significant the story was to that person could determine how strong their memory of that event could be. For example, if it was just a story about a regular work day, it might have been less meaningful than a person remembering their 40th birthday. Finally, the preprocessing and EDA of the data before classifying it could be another limitation. If the pre-processing was not the most accurate, it may have reduced classification accuracy. To determine the best pre-processing method, a comparative analysis would need to be conducted by applying various techniques such as normalization, feature scaling, or dimensionality reduction to the dataset and evaluating their impact on downstream tasks. This comparative analysis would have allowed me to systematically test different combinations of pre-processing steps and assess their effects on model performance metrics, aiding in finding the most effective approach for the specific dataset and task at hand.

5. CONCLUSION

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

In conclusion, this study demonstrates the effectiveness of applying machine learning techniques to psychological memory data to investigate age-related changes in memory performance. By analyzing memory reports categorized as imagined, recalled, and retold across different age groups, we showed that machine learning models can successfully identify meaningful patterns in memory representation and classification. Among the models tested, Logistic Regression achieved the highest performance, highlighting its suitability for capturing relationships within psychological memory features.

Our findings indicate that memory decline is influenced by multiple interacting factors rather than a single cause, including emotional significance, contextual relevance, and cognitive engagement with an event. The observed overlap between recalled and retold memories further suggests that aging may reduce the distinctiveness of memory retrieval processes, emphasizing the complexity of memory representation over time.

Overall, this research bridges psychological theory and artificial intelligence by demonstrating how computational methods can enhance the analysis of cognitive aging. It provides a scalable framework for studying memory deterioration and supports the use of machine learning as a valuable tool for advancing research in cognitive science, neuroscience, and mental health. Future work can extend this approach by incorporating longitudinal datasets and neurobiological measures to further refine our understanding of memory changes across the human lifespan.

ACKNOWLEDGMENTS

The authors declare that no financial or institutional support was received for this research.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

References

1. National Institute of Neurological Disorders and Stroke. 2023. "Brain Basics: Know Your Brain | National Institute of Neurological Disorders and Stroke." Www.ninds.nih.gov. March 17, 2023. Brain Basics: Know Your Brain.
2. Murman, Daniel. 2015. "The Impact of Age on Cognition." Seminars in Hearing 36 (03): 111–21. https://doi.org/10.1055/s-0035-1555115.
3. Gabrieli, J. D. E. 1998. "COGNITIVE NEUROSCIENCE of HUMAN MEMORY." Annual Review of Psychology 49 (1): 87–115. https://doi.org/10.1146/annurev.psych.49.1.87.
4. Valerio, Kate E, Sarah Prieto, Alexander N Hasselbach, Jena N Moody, Scott M Hayes, and Jasmeet P Hayes. 2024. "Machine Learning Identifies Novel Markers Predicting Functional Decline in Older Adults." Oup.com. 2024.

ISSN: 3139-3527 | January-March-Issue, Vol. 1, No. 1 (2026) | DOI: 10.66261/cxc14x51

https://academic.oup.com/braincomms/article/3/3/fcab140/6310007.

5. Jordan, Candina. 2022. "What Is Cognitive Psychology?" WebMD. May 31, 2022. https://www.webmd.com/mental-health/what-is-cognitive-psychology.
6. Chakravarty, Sucheta, Yvonne Y Chen, and Jeremy B Caplan. 2020. "Predicting Memory from Study-Related Brain Activity." Journal of Neurophysiology 124 (6): 2060–75. https://doi.org/10.1152/jn.00193.2020.
7. Murphy, Emily R D, and Jesse Rissman. 2020. "Evidence of Memory from Brain Data." Journal of Law and the Biosciences 7 (1). https://doi.org/10.1093/jlb/lsaa078.
8. Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. The Elements of Statistical Learning, Second Edition: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer. https://hastie.su.domains/Papers/ESLII.pdf.
9. Kundu, Rohit. 2022. "Confusion Matrix: How to Use It & Interpret Results [Examples]." Www.v7labs.com. September 13, 2022. https://www.v7labs.com/blog/confusion-matrix-guide.
10. GeeksForGeeks. 2017. "Decision Tree - GeeksforGeeks." GeeksforGeeks. October 16, 2017. https://www.geeksforgeeks.org/decision-tree/.
11. Srivastava, Tavish. 2019. "Introduction to KNN, K-Nearest Neighbors : Simplified." Analytics Vidhya. March 7, 2019. Introduction to KNN, K-Nearest Neighbors
12. Journey To College. "When Does the Brain Reach Maturity? It's Later than You Think – Journey to College – Realize Your Potential." *Journeytocollege.mo.gov,* Journey to College – Realize Your Potential

Code:

https://colab.research.google.com/drive/1gzFk0pfQs7TZsdksHCpbS1bcGdHh1HA1?usp=sharing

Special thanks to Mr. Salwan Butrus from the University of California, Berkeley for his mentorship throughout this project.