Memory Fingerprinting: Identifying Individuals from a Collection of Recalled Words

Siddhant M. Kochrekar (smk371@scarletmail.rutgers.edu)
Department of Computer Science, 110 Frelinghuysen Rd
Piscataway, NJ 08854

John R. Gorgy (jrg234@scarletmail.rutgers.edu)
Department of Biological Sciences, 88 Lipman Dr
New Brunswick, NJ 08901

Introduction

The complexities of memory search has been a topic in the cognitive world for more than half a century. The underlying mechanisms that govern memory extraction are common amongst us. However, there are individual characteristics that contribute to differentiation during recall. Is there a way to leverage these unique differences and be able to identify someone by their specific implementation of the underlying cognitive tools? It's been proven that by studying Magnetoencephalography (MEG), an individual can be identified with a number of differential characteristics, including their real brain 'fingerprints' (Wu, Ramdas, & Wehbe, 2021). Other work has been done in exploiting these differences to suggest a correlation between the amount of words recalled and IQ (Healey, Crutchley, & Kahana, 2014). This literature found four factors: Primacy, Temporal, Semantic, and Recency, to be sources of variation that are significant in recall success and intelligence. Recent work has also proved that the properties within the corpus of words have a positive effect on word recall. There are six properties identified; concreteness, contextual diversity, word length, emotional valence, arousal, and animacy of words that contribute to the word-level and list-level recall performance (Aka, Phan, & Kahana, 2021)).

To better understand the individual's dynamics during memory recollection, we are focusing on not only the performance of an episodic memory recall task, but also unique attributes within the presented/recalled word lists and subjects. By encompassing prior literature, we are creating a holistic identification approach in attempts of creating a 'Memory Fingerprint' based off their usage of cognitive processes and discrete facts about the subject.

Methods

Participants

The data reported is from the Penn Electrophysiology Encoding and Retrieval Study (PEERS). The series of studies within attempts to illuminate the Electroencephalogram (EEG) correlates of memory processes with a particular focus on retrieval dynamics in free recall.

The dataset contains a total of 224 total subjects (186 young (<35 years of age) and 38 old(>60 years of age)) with each completing 20 sessions each. In each session, maximum of 16 sequential trials were conducted. In each trial, the subject is presented a list of 16 words (from a vocabulary of 1638 words) where they attempted to recall as many as they can. There exists subjects who have not completed all 20 sessions which were removed since there is no desired list of recalled words.

Data Processing

The PAR and session LOG files are converted to JSON object and eventually Pandas DataFrame for convenience of analysis and modeling. The study and recall phases form a single object, which depicts a trial as seen in Figure 1


```
{
    "actualWords_recallPhase":[
      "OFFICER",
      "SOLDIER",
      "VETERAN",
      "ARMY",
      "CASH",
      "MISSILE",
      "CAFE",
      "SNACK",
      "CHICKEN",
      "SUPPER",
      "STOMACH",
      "DIME",
      "HOME",
      "DORM"
    ],
  "actualWords_studyPhase":[
      "OFFICER",
      "CASH",
      "VETERAN",
      "MISSILE",
      "CAFE",
      "STOMACH",
      "SOLDIER",
      "HOME",
      "DORM",
      "ARMY",
      "LIST",
      "SNACK",
      "CHICKEN",
      "DIME",
      "SUPPER",
      "POSSUM"
  ],
    "listIndex": 2,
    "listLength_recallPhase": 14,
    "listLength_studyPhase": 16,
    "serialPositions_recallPhase": [16,6,2,9,1,3,4,11,12,14,5,13,7,8],
    "serialPositions_studyPhase": [0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15],
    "sessionIndex": 5,
    "subjectIndex": 63,
    "wordIndices_recallPhase": [973,1344,1561,39,252,900,216,1331,297,1411,1384,466,712,488],
    "wordIndices_studyPhase": [973,252,1561,900,216,1384,1344,712,488,39,830,1331,297,466,1411,1112]
}
```


Figure 1: Data Format used for a single trial in the dataset

After getting a workable dataframe, the next task tackled was the choice of word embedding. Our first intuition was to choose the Continuous Bag of Words model since that is one of the most common approaches to embedding. However, after some further research we used the Continuous Skip Gram model and knew that was the right method as it made semantic sense. In the CBOW model, the representations of context (or surrounding words) are combined to predict the word in the middle of a sentence. While in the Skip-gram model, the representation of the input word is used to predict the context as illustrated in Figure 2 (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). Here we do not have sentences, but a single trial is considered as a word cloud in multi dimensional space.

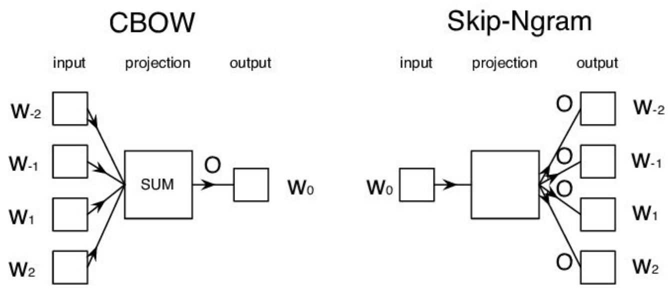

Figure 2: Comparsions of CBOW (Continous Bag of Words) and Skip-Ngram embedding(Ling et al., 2015)

Feature Engineering

Additional features on the basis of age of subject, length of word, recall rate, word position in study phase, living or non-living entity, average number of vowels in recalled words, recall repetition and number of external words were built to look out for additional lift in model metrics and also for the purpose of testing hypotheses.

Modeling

We treat this as a multi-class classification problem, where each individual subject is the target to be predicted given the input vector. Here input vector is the information that we have from a single trial of that subject. The word embeddings for all the 16 words of the study phase along with binary flags whether that word is recalled or not, form the input vector. Instead of using all 186 subjects, we tackled a 10 class classification problem. Randomly 10 subjects (having subject IDs [267, 87, 151, 90, 122, 242, 110, 92, 241, 194]) were picked and model metrics were reported using Random Forest Algorithm. Dimensionality reduction was performed using PCA for better stability and to tackle sparsity issue. The data was split in 70-20-10 format for train-val-test. 5 fold Cross Validation was performed to avoid overfitted models.

Results

Data Summary

There were 1638 unique words in the vocabulary and not all are shown to every subject uniformly. Some words are shown more frequently. In this research we majorly focus on the 186 young subjects constituting 17335 trials.

A peculiar observation was seen while analyzing most common words seen across trials. Most trials had words related to water, sea and ocean as seen in Figure 3. Further RCA needs to be done to check if this creates bias in studies.

Also there are words of varying lengths and eventually led us to testing the word difficulty/length hypothesis. Which stated that subjects tend to recall shorter words and it was rejected by looking at the results in Figure 4. We conclude that recall is not related (rather we see that longer words have slightly higher probability of recall) to the word length but more connected to the semantics of the word.

	word	occurrence in trials
0	MARINE	349
1	OTTER	348
2	BEAVER	342
3	DIVER	340
4	ANIMAL	339
5	CRAB	335
6	CREATURE	325
7	FISH	320
8	SWIMMER	318
9	SEAL	317

	word	occurrence in trials
0	GHETTO	44
1	CEMETERY	43
2	WEB	43
3	MUMMY	40
4	HUSBAND	39
5	SPOUSE	39
6	GYM	38
7	GHOST	30
8	PIMPLE	28
9	SUNSET	26

Figure 3: 10 most common words (left) and 10 least common words (right) in vocabulary

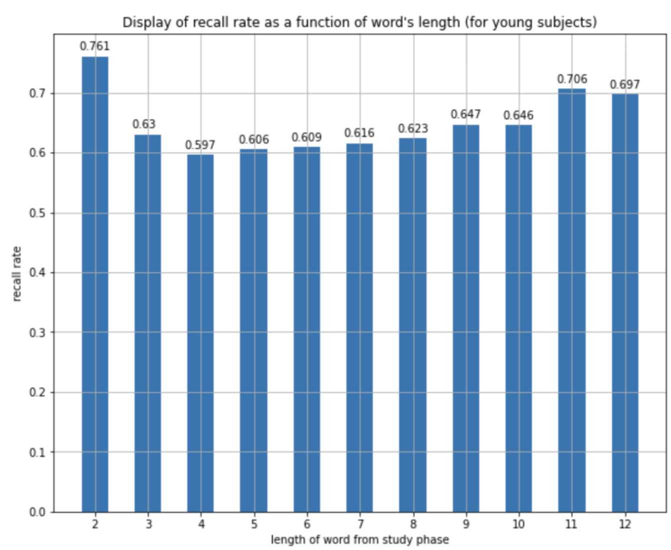

Figure 4: Recall rate as a function of Word Length

Hypothesis Testing

Primacy and recency effects can be seen in Figure 5, where recall rate is plotted for word positions in the study phase of experiment. The U-curve of total recall can be validated. Also it can be seen that subjects recall ending words more frequently than starting words and intermediate study space being least recalled. The primacy effect can be linked with rehearsal such that Items from the beginning of the list receive a larger number of rehearsals than do items from other list positions (Rundus, 1971; Tan & Ward, 2000).

In most cases, subjects tend to recall the last word of the study phase first. This is called as Recency Effect and can

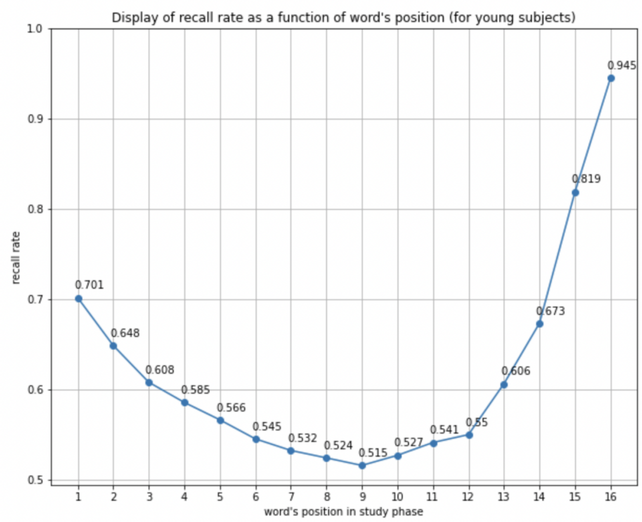

Figure 5: Recall curve for given word position

be seen in Figure 6. The first word in the recall phase comes from the ending portion of study phase.

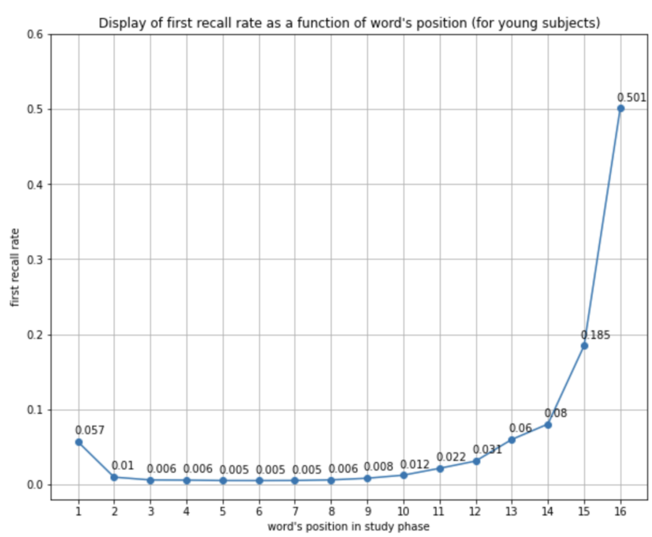

Figure 6: First Recall curve for given word position

For contiguity we see that a word is recalled with its neighboring words. This is also called as lag-recency effect in free recall. We can see in Figure 7 that nearby words are recalled more (peaks at +1 and -1) than words having a distance of more than 2 in the study phase. Also there is a forward asymmetry seen. Which means that the word after the current word in study phase is more likely to be recalled than the one before it.

Another hypothesis which was tested was whether number of words recalled reduces as humans get older. It can be seen in Figure 8 that older subjects recall lesser words, more precisely younger adults recall 9.88 words and older adults recall

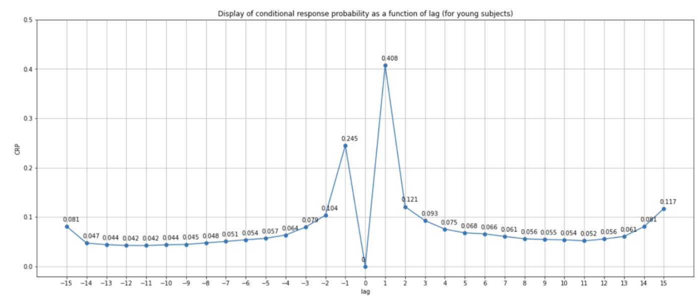

Figure 7: CRP as a function of word lag

6.97 words on an average out of 16 shown in study phase.

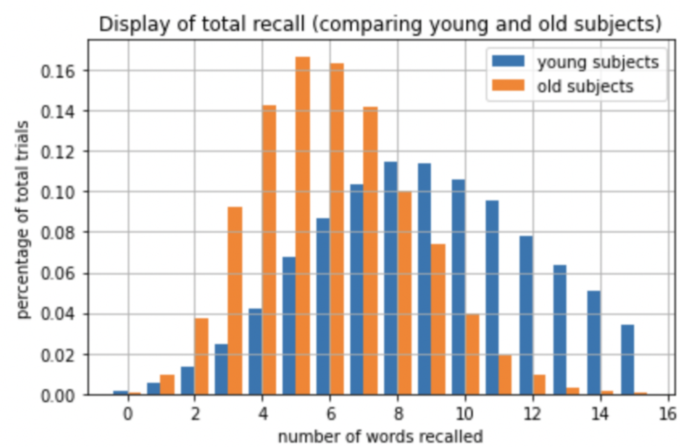

Figure 8: Recall success rate across age

Model Metrics

Baseline Random Forest model resulted in 22.91% accuracy (22 True Positives of the total 96 testing records).

PCA along with Random Forest model resulted in 30.21% accuracy (29 True Positives of the total 96 testing records).

Refer to Figure 9 and 10 for model evaluation metrics for the 10 subjects being tested using trials. Stabilized metrics using Cross Validation can be seen in Table 1.

Table 1: Effect of reducing dimensionality & 5-fold CV

Model	Accuracy
5 fold CV (Random Forest)	18.46%
5 fold CV (10 component PCA + Random Forest)	23.32%

General Discussion and Future Plans

PEERS Versions

The PEERS dataset has numerous versions of trials/subjects since it it spanned between 2010-2020. To further investigate our hypothesis and model, we are planning to explore various

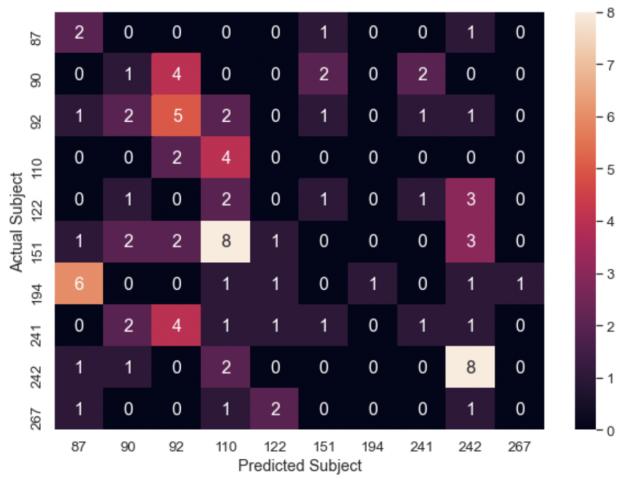

Figure 9: Confusion matrix for baseline RF

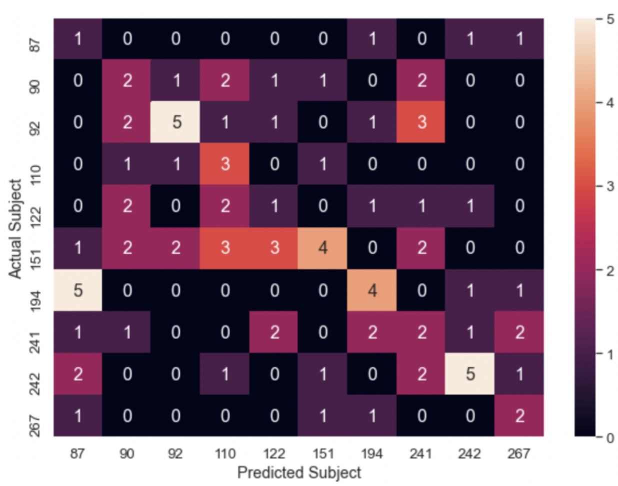

Figure 10: Confusion matrix for PCA + RF

versions of PEERS in order to get a larger training/test sets. As our selected features should be agnostic of specific subjects, we expect to find similar outcomes between different age groups and presented words.

Source of External words

Out of the total 17335 trials, 5074 (30%) trials had external words being recalled. We are trying to figure out where do these words come from. Word embeddings and semantic analysis would help in this task. Connecting a task with prior tasks for external word also seems to be one possibility.

Latent Semantic Analysis (LSA)

Along increasing the size of our dataset, we plan on implementing a Latent Semantic Analysis (LSA) on the presented word matrix. Latent Semantic Analysis (LSA) is a theory and method for extracting and representing the contextual-usage meaning of words by statistical computations applied to a large corpus of text (Landauer, Foltz, & Laham, 1998). The idea is that the combination of all the word contexts where a word is present or not provides a set of constraints that largely determines the similarity of meaning of words and sets of words to each other.

Log Files

With any modeling process, transparency of what is happening in the back end can be crucial to drawing certain conclusions. As we finalize our models, we are going to implement a LOG file to keep track of all the events that occur while running our model. One reasons for doing so is to have a single source of truth if any issues that might occur making trouble shooting much easier. Another purpose, is the metadata can be useful providing information about the assets, capacity, and configuration of our storage systems.

Modeling

Considering distance based algorithms (K nearest neighbors) forms one important task going forward. Clustering (K means/ hierarchical) data points in space to find the origin of subject's identity takes us to the root level of fingerprinting.

We need to incorporate order of recall in the input vector, so far we have flags which inform whether the word is recalled or not.

Siamese network/Artificial Neural Network

Finding similarities between features is the basis of many mathematical models. There a numerous methods of comparing two element vectors, depending on the final goal of the comparison (Euclidean distance, Pearson correlation coefficient, and others). But if the comparisons are not as straight forward and has to be applied to more complex data sets with features varying in dimensions/types that might need compression before processing, these measures could be ill-suited. In these cases, a siamese neural network may be the best choice: it consists of two identical artificial neural networks each capable of learning the hidden representation of an input vector (Chicco, 2021). Given the range of dimensions and complex features of our model, utilizing a siamese neural network will be beneficial. By having two parallel ANNs performing feed-forward perceptrons and back-propagation, it provides semantic similarity between the projected representation of the two input vectors. This can be extremely useful when comparing the vector spaces of two words and outputting the semantic similarity between both.

References

Aka, A., Phan, T. D., & Kahana, M. J. (2021). Predicting recall of words and lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(5), 765.

Chicco, D. (2021). Siamese neural networks: An overview. *Artificial Neural Networks*, 73–94.

Healey, M. K., Crutchley, P., & Kahana, M. J. (2014). Individual differences in memory search and their relation to intelligence. *Journal of Experimental Psychology: General*, *143*(4), 1553.

Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General*, *143*(2), 575.

Healey, M. K., & Kahana, M. J. (2016). A four-component model of age-related memory change. *Psychological Review*, *123*(1), 23.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, *25*(2-3), 259–284.

Ling, W., Dyer, C., Black, A., & Trancoso, I. (2015, 05). Two/too simple adaptations of word2vec for syntax problems.. doi: 10.3115/v1/N15-1142

Long, N. M., Danoff, M. S., & Kahana, M. J. (2015). Recall dynamics reveal the retrieval of emotional context. *Psychonomic bulletin & review*, *22*(5), 1328–1333.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 26). Curran Associates, Inc.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological review*, *115*(4), 893.

Wu, S., Ramdas, A., & Wehbe, L. (2021). Brainprints: identifying individuals from magnetoencephalograms. *bioRxiv*, 2020–06.