Predicting the Topics to Review in Preparation of Your Next Meeting

Seyed Ali Bahrainian and Fabio Crestani

Faculty of Informatics, University of Lugano (USI), Switzerland
{bahres, fabio.crestani}@usi.ch

Abstract. Memory augmentation is the process of providing human memory with information that facilitates and complements the recall of an event in a person's past. In this paper, we propose a novel time-series method for predicting the topics that one should review in preparation of one's next meeting. This can be seen as a way of augmenting human memory. Since the number of topics that are discussed in different meetings of a typical person might be very large, there is a need for detecting topics that are more likely to continue in subsequent meetings, in order to focus one's attention just on them. Our experimental results on real-world data demonstrate that our proposed method significantly outperforms the state-of-the-art Hidden Markov Model (HMM) baseline.

Keywords: Human Memory Augmentation; Topic Prediction; Workplace Meetings

1 Introduction

Human memory is a critically important cognitive ability that we constantly rely on. However, sometimes due to the volume and intensity of information that we are exposed to, or due to lack of adequate attention, or due to aging, this critical cognitive ability fails to recall important events in our past. Augmentation of human memory in a workplace environment can effectively serve as a solution in preventing failure to recall past events [3].

In this paper, we focus on tracking one's meetings with the purpose of memory augmentation. We use a real-world dataset of weekly meetings of seven groups of people. In this dataset, we recorded real workplace meetings of each group (where a group consisted of two individuals) over a span of an entire month. We compare our novel method against the HMM baseline for predicting the topics of a conversation that will be continued from previous meetings.

2 Background

Psychology of human memory has comprehensively studied how human memory recalls events or forgets them. One important work in this domain was the *forgetting curve* by Ebbinghaus. The forgetting curve (which is an exponentially decreasing curve) shows that a human forgets on average about 77% of the details of what one has learned after six days. This motivated our goal in augmenting human memory to assist one in recalling more details of one's past events. Additionally, our study is motivated by a memory augmentation tool that we have already developed and deployed in the context of a project[1] for aiding people's memory in their workplace meetings [1]. This system takes as input transcriptions of audio recordings of one's conversations and images taken automatically by one's wearable camera. Both media types are time synchronized. The tool then processes the data by modeling the topics of the transcribed conversations and connecting the topics with their corresponding images. We use Latent Dirichlet Allocation (LDA) [4] as our topic modeling approach.

Furthermore, the benefit of this research work is endorsed by relevant studies [8] which showed that for people, replaying their lives to them have significant effect in helping them better recall and remember past events.

3 Methodology

In this section, we briefly present our new method as well as the baseline method for predicting the topics to be reviewed in preparation of one's next meeting.

3.1 Our method

Our method is a hybrid that combines two effects that we refer to them as recency and establishment.

Recency. The recency effect, modifies the ranking of topics by assigning higher weights to topics of the most recent previous meetings.

Therefore, this effect assigns higher weight to a word which has occurred in the most recent meeting. As a result, a word vector based on the recency effect is produced.

Establishment According to the establishment effect the assigned probability scores are higher for the words which have persisted over time.

Therefore, the word vector constructed by averaging all topics in all n previous meetings based on the establishment effect will be a representation of average occurrence of each word, where the most established (i.e. persisting in occurrence) words are assigned higher weights.

Combining Recency and Establishment. We combine the two effects using a dynamic method by weighting each effect in an evolutionary process. This method integrates scores from the recency and the establishment effects using linear interpolation for a meeting at time slice t such that:

$$Score_t = w_{e,t} * Score_{establishment} + w_{r,t} * Score_{recency} \tag{1}$$

where $Score_{establishment}$ and $Score_{recency}$ are computed by the establishment and the recency effects, respectively. Furthermore, $w_{e,t}$ and $w_{r,t}$ are establishment weights and recency weights computed in a measurement process.

[1] http://recall-fet.eu/

The measurement process consists in using a topic linking module. In our previous work [2], we introduced a topic model for tracking the evolution of intermittent topics over time. We use this system to link together similar topics over time and based on that compute the weights of each effect. At the end in order to determine the probability of a topic continuing in a future meeting, we use an energy function.

3.2 Hidden Markov Model Baseline

HMMs [7] have been extensively used for modeling multivariate time series and predicting next states. In [6] a number of papers that describe domains were HMMs hold the state-of-the-art performance are enlisted. Thus, in this paper we use HMM as a baseline for our benchmark. The HMM we implement is Gaussian. For determining the number of HMM output states we use the Bayesian Information Criterion to find the optimal number of output states given the data of each set of four meetings. Finally, after training the HMM model with the topics of the first n meetings we measure the likelihood of each of the topics that we want to predict its continuation under the trained model. The result is a likelihood score per topic. We normalize the likelihood scores by dividing each of them by the maximum likelihood score. Finally, we compute the optimal threshold using n-fold cross validation.

4 Experimental Setup

4.1 Dataset Description

Our dataset consists of recordings of workplace meetings of 7 groups of people. Each group consisted of two members. For each group, the audio of 4 consecutive meetings over four weeks were recorded. Our dataset is real-world and captured in the wild, meaning that the involved participants were asked to simply have their usual meetings with no regulations imposed.

LDA topics were extracted from the transcriptions of meetings. Since the number of topics discussed in two different meetings might vary, it is important to estimate the number of topics per each meeting. For this purpose, similar to the method proposed in [5], we went through a model selection process.

The extracted LDA topics from the first 3 meetings of each group were then manually labeled based on whether or not they had continued in the 4_{th} meeting. The resulting number of labeled topics to be predicted out of all 7 groups is 205. Our aim is to correctly predict the assigned labels.

4.2 Evaluation

In our first experiment, we compute precision and recall values of our proposed method and compare it with the HMM baseline. Table 1 shows precision values at different levels of recall and for all decision thresholds. The values are obtained from interpolated precision-recall curves. The table shows that the our method outperforms the HMM baseline in terms of precision at all levels of recall.

Furthermore, we computed Mean Average Precision (MAP) of our method which was 69.73% against that of HMM which was 55.69%. Additionally, we computed the F_1 measures using 7-fold cross validation for each of the seven groups. We can observe through this experiment that our method significantly outperforms the HMM baseline with an F_1 measure of 76.87% versus 58.43%. We confirmed the significance in performance difference using a paired t-test.

Table 1. Precision of our method versus HMM at different Recall levels

Recall (%)	10	20	30	40	50	60	70	80	90	100
Prec. Our method (%)	76.19	73.58	73.58	73.58	73.58	73.58	72.05	69.35	66.77	65.28
Prec. HMM (%)	67.59	66.45	66.45	64.28	64.28	64.28	64.28	64.28	64.28	64.28

5 Conclusions

In this paper, we introduced the problem of predicting topics to be reviewed in preparation of one's next meeting to augment one's memory. For this purpose, we proposed a novel method and compared it against an HMM baseline.

The developed method could be implemented as a part of a proactive memory augmentation system that aids people in their every day lives.

References

1. S. A. Bahrainian and F. Crestani. Cued retrieval of personal memories of social interactions. In *Proc. of the First Workshop on Lifelogging Tools and Applications*, LTA '16, pages 3–12, 2016.
2. S. A. Bahrainian, I. Mele, and F. Crestani. Modeling discrete dynamic topics. In *Proc. of the Symposium on Applied Computing*, SAC '17, pages 858–865, 2017.
3. A. Bexheti, E. Niforatos, S. A. Bahrainian, M. Langheinrich, and F. Crestani. Measuring the effect of cued recall on work meetings. In *Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 1020–1026, 2016.
4. D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, pages 993–1022, 2003.
5. T. L. Griffiths and M. Steyvers. Finding scientific topics. *Proc. of the National academy of Sciences*, pages 5228–5235, 2004.
6. M. Pietrzykowskiand and W. Sałabun. Applications of hidden markov model: state-of-the-art. In *International Journal of Computer Technology and Applications*, pages 1384–1391, 2014.
7. L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. In *Proc. of the IEEE*, pages 257–286. 1989.
8. A. Sellen, A. Fogg, M. Aitken, S. Hodges, C. Rother, and K. R. Wood. Do lifelogging technologies support memory for the past? In *Proc. of the Conference on Human Factors in Computing Systems*, pages 81–90, 2007.