

Analyzing Text Based on Dementia Blog Posts



Aron Bishop, Department of Computer Science
Ashanti Boone, Department of Computer Science
Soo-Yeon Ji, Ph. D, Department of Computer Science

OFFICE OF RESEARCH AND SPONSORED
PROGRAMS

Research and Innovation 2023

Abstract

Our goal is to investigate the viability of using Term Frequency-Inverse Document Frequency (TF-IDF) analysis in combination with artificial intelligence (AI) for analyzing dementia. The project's major aim is to analyze user post information from public discussion boards associated with dementia to identify emotional responses or caregivers' needs. The project contains four stages; data collection, preprocessing, feature extraction, and model generation. While our research is still in the preliminary stage, we have been able to analyze the TF-IDF of two comments.

Introduction

Dementia, a debilitating and progressive neurological disorder, encompasses various conditions such as Alzheimer's and Parkinson's. With an increasingly aging global population, dementia is expected to afflict hundreds of millions of individuals. By integrating AI and TF-IDF for early dementia detection, one could potentially expedite the diagnostic process, provide more accurate assessments, enable timely interventions, and ultimately improve patient care and quality of life.

Words	Frequency	Frequency
pay	0.004	0.013
I	0.000	0.013
new	0.000	0.006
Home	0.000	0.006
take	0.000	0.006
access	0.000	0.006
might	0.000	0.006
account	0.000	0.006
Protection	0.000	0.006
advice!	0.000	0.006
year	0.000	0.006
advised	0.000	0.006

Methods

Obtain text from Dementia Forum:



<https://forum.alzheimers.org.uk/>

```
2023-04-17 08:24:27 [scrapy.core.scrapy] DEBUG: Scraped from <200 https://forum.alzheimers.org.uk/threads/chc-continuing-healthcare-support-thread.86059/page-2>
{'forum_name': 'Legal and financial issues',
 'post_data': 'Hi 2jays,I think you could argue that your Mum should remain in '
              'familiar surroundings with staff who know how to interpret her '
              'needs. Everything should be geared in your Mums best '
              'interests.That's the argument I used to keep my Dad in his '
              'nursing home once he got NHS Continuing Health Care, and the '
              'mental health team and the nursing home backed me up.Of course '
              'the nursing home had to be able to meet his needs.I'm very '
              'sorry to hear your Mum has got to this stage '
              'though. xxxxxxxxxxxx',
 'post_date': '2015-11-04T18:25:06+0000',
 'post_url': 'https://forum.alzheimers.org.uk/threads/chc-continuing-healthcare-support-thread.86059/page-2',
 'post_User': '12271',
 'thread_name': 'CHC (Continuing Healthcare) support thread'}
```

Preprocess obtained text:

- Collect data
- Remove stop words (the, a, with)
- Compute Term Frequency (TF)
- Compute Inverse Document Frequency (IDF)
- Predictive model generation utilizing machine learning techniques

Results

 Hesketh New member	Jan 20, 2023 Hi All - new member here after some advice! My Dad has just moved to a Nursing Home and we need to arrange deputyship to access his bank account to pay the fees. Dad's social worker advised that some solicitors might pay the Court of Protection application fees and hold their costs until the Deputyships are approved, and then submit their invoice. Being as the process can take up to a year I find this implausible, but I wondered if any members had experience of this.	< #1
 thistlejak Registered User	Jan 20, 2023 Can't help you with Solicitors waiting for their money but have you considered completing the forms yourself? I did forms for both MIL & FIL at the same time - It is doable. We didn't consider using a solicitor after looking at the forms because we realised that we would have to supply them with all the information for them to just fill in boxes. The only form that can cause some difficulties is COP3 - the statement of capacity - we used an independent social worker as no one else would complete it even if we paid them. We did have to pay the money out with the application but reimbursed ourselves when we got Deputyship - It cost us about £600 each including the COP3 report. Have you passed the window to get your Dad to grant a POA?	< #2

This sheet shows the TF-IDF analysis of Comment 1 and 2. The words listed come from the two comments. The top number shows the frequency of the word in the first comment. The bottom number shows the frequency of the word in the second comment.

Conclusion

Our research is still in the preliminary stage. We have only been able to analyze two comments as shown in the results portion. Our goal is to broaden our word corpus to include more comments from the Dementia Forum to generate more accurate information. The resulting scores will reflect the term's relative importance in the comment compared to the entire post data. This approach will enhance our understanding of the most relevant topics and concerns within the dementia community.

Discussion

We still have to complete the model generation phase of our research. We have been able to analyze two comments from the dementia forum, but we hope to expand our corpus by preprocessing more comments from the website. We will categorize comments by forums and compute the TF-IDF for each forum. “I have dementia” and others will all have a separate term frequency. Then we will combine the words into a single corpus and measure the overall TF-IDF. This provides more information for the model generation phase of research and will assist with pattern recognition.

Acknowledgements

This material is based upon work supported by the Army Research Office (Grant No. W911NF-18-1-0460) and partially supported by an NSF grant (Grant No. 2219634)