Applications of NLP in Online Privacy *

Ankita Maity[1[2021701017]] and Pawan Sasanka Amamnamanchi[1[2020701028]]

International Institute of Information Technology, Hyderabad

Abstract. [1] established the average readability of privacy policies was high. In this paper, we explore if recent advances in Natural Language Processing can be used to make privacy policies more readable while maintaining most important intact. We report our findings and evaluate the readability through automatic and human evaluations.

Keywords: Online Privacy · Privacy Policies · Readability · Natural Language Processing

1 Introduction

A privacy policy is a document/statement/webpage used by organization to declare and announce the various privacy policies and measures at place at the company. A privacy policy discloses the ways the entity/party will gather information regarding the user in case of user interaction and how it will store/use it and what rights the user might have in case of accessing it. Privacy policies are usually filled with a lot of legalese and technical terms which might increase the complexity of the document, hence bringing the readability of this privacy policy down.

Readability of privacy policy is an important aspect of our increasingly digital lives. Whenever a user interacts with a website, his data might stored or used in a variety of ways and reading this privacy policy will inform him of his rights and inform him of what he can do with the data or if he has any access at all, this helps him take an informed decision around if he wants to continue interacting with the website.

[2], a survey conducted for 2700 people, found that 87% of people accept privacy policies without reading them and on average, social media privacy policies take over 47 minutes to read fully. Most users claimed that they didn't read privacy policies because they were too long and a majority also agreed that they would consent regardless of having read or not and some claimed they wouldn't understand what they were reading.

[1] at readability of the average privacy policy from the top 500 online websites is that similar to someone who has attended high school, while the average reading level of the American adult is at the 7th grade. This work aims to establish if recent advancements in natural language processing in summarization and language modeling can help enhance readability of privacy policies.

* Supported by IIIT Hyderabad.

We find that,

- Some models do enhance readability by a much greater margin compared to the source privacy policy.
- Current models aren't universally applicable to making privacy policies more readable and that human guidance is still needed.

2 Methodology

Fig. 1. Methodology of this paper. Collection of data first, then application of models and finally evaluation.

2.1 Data Collection

We first collect a wide variety of privacy policies primarily from OPP-115 [3], it is a corpus of 115 privacy policies with manual annotations for 23K fine-grained data practices. We also additionally collect data this time accommodating a variety of domains like health, finance, information technology, social media.

2.2 Models

Recent advances in Natural Language Processing have furthered the state of the art in a variety of problems like Language Modeling, Summarization, Neural Machine Translation, Simplification, Question Answering. We use some of these models to generate simplified/summarized privacy policies for further testing.

- GPT3 [4], is a recent language model. It is an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and performs at the state of the art in a variety of downstream tasks. We use the a prompt to generate the output, "Paraphrase the following privacy policy so that a 2nd grader could understand it" along with the privacy policy.
- BART [5] is a recent state of the art sequence to sequence language model. It achieved the best scores in a variety of tasks, mainly summarization.

2.3 Automatic Evaluation

For automatic evaluation, we use a variety of metrics like SMOG, Gunning Fog, Flesch-Kincaid Reading score and Flesch Reading Ease Score.

- Flesch Kincaid Reading Score (the lower the better) - Approximate reading grade level of a text.
- Flesch Reading Ease Score (the higher the better) – Scores along complexity of sentences and words, but assumes that long words are necessarily complex.
- Gunning Fog Score (the lower the better) – Upgrade of Flesch Reading Ease Score, calculates percentage of complex words as percentage of total while measuring.
- SMOG Score (the lower the better) – Useful for health readability and domains with high technicality. Adjusts score for complexity of words.

3 Discussion and Results

3.1 Automatic Evaluation

We evaluate the models using a variety of automatic evaluation metrics and come to the following conclusions.

- GPT-3 always results in improving readability of the privacy policy.
- BART improves on SMOG and FOG scores but consistently sacrifices plain readability of the document.
- BART outputs are very precise.
- GPT-3 outputs are more open ended but understandable.
- Both the approaches improve on readability of underlying document based on Flesch Reading Ease Scores.

Generally GPT-3 produces more consistent outputs and performs better when compared to BART, but BART on average captures more useful information owing to its summarization capabilities.

3.2 Human Evaluation

For our human evaluation, we show five privacy policies and their generated outputs to a set of human evaluators and ask them a variety of questions, at the end of it we also for any feedback if present. Questions asked include:

- Are you concerned about how your personal information is handled by different websites that you share your data with?
- Have you read a privacy policy before?
- Would you read a privacy policy, even if it was difficult to read, if you felt that a company was using your data for purposes you did not specify?
- How do you think is the readability of current privacy policies?

- Would you read a privacy policy that had less complicated words/was shorter in length (showing original and summarized policies from five websites)?

It was found that generally, the evaluators preferred a model output over the source policy always and within the models GPT3 is generally preferred over BART.

In some cases, people don't prefer either of the models, when asked to elaborate we found out that this could be because of language barrier or that the even the simplified text policy was difficult to read.

3.3 Issues of Language Models

Language Models aren't perfect and still face a variety of issues dealing with problems ranging from imperfect understanding of language, noisy or redundant generations or an inaccurate modeling of the task.

Summarization is a lossy compression of documents, when it comes to something as important as privacy policies, any loss of information is hurtful in the propagation of important information and any loss could also hide important points that were in the source privacy policy which couldn't be reflected in the generated output, thus leading to a incomplete and asymmetric knowledge. The loss of key terms is just as harmful.

Abstrative summarization/language models have biases and these biases morph information in a variety of ways. These biases are of two kinds, intrinsic bias and extrinsic bias. Intrinsic bias is when the model generates an output which captures information present in the source document in an inaccurate manner. Extrinsic bias is when the model generates information which is not present in the source document. Both of them are harmful in the context of privacy policies.

Human guidance is still needed for evaluating the outputs of these models and bringing them up to a level where they generate faithful and factual outputs and also add any legalese which is crucial for the privacy policy.

3.4 Reading comprehension of readable privacy policies

One aspect of evaluation we could not completely perform during human evaluation was reading comprehension of the outputs of the models. "Is a readable privacy policy more understandable" is a crucial question to ask and also helps make the case for using these models strong. For the scope of our project, we did the comprehension ourselves and read the source privacy policy and the output to establish if most of the important information was covered. While there were exceptions, in most cases salient information was captured by the output hence establishing a precedent for the usage of these models.

4 Conclusion & Future Work

We use advances in natural language processing to make privacy policies more readable. We find that models can generate good, human understandable privacy policies which can be comprehended more than the source privacy policies.

But human guidance is still desired to make these generated outputs perfect and faithful to the source privacy policy. Future work could involve generating outputs which also factor in preferences from the humans reading the privacy policy which would contribute to a better experience.

5 Code

```
import nltk
import stanza
import os
import openai


nlp = stanza.Pipeline(lang='en', processors='tokenize')


y = """""" # Privacy Policy

doc = nlp(y)

num_sentences = len(doc.sentences)
print(num_sentences)

# Sentence segmentation of privacy policy
sentences = [sent.tokens for sent in doc.sentences]
print(len(sentences))
sents = []
text = []
for tokens in sentences:
   text.append([token.text for token in tokens])
sentences = [" ".join(sent) for sent in text]
print(sentences)


#Chunking of privacy policy

max_chunk_length = 512 #for default summarization pipeline
chunks = []
current = 0
for s in sentences:
   if (len(chunks) == current + 1) and (len(chunks[current]) +
       len(s.split(' ')) <= max_chunk_length):
          chunks[current].extend(s.split(' ')) #put shorter
              sentence in current chunk
```

```
    elif (len(chunks) == current + 1) and (len(chunks[current]) +
        len(s.split(' ')) > max_chunk_length):
            current += 1 #new chunk
            chunks.append(s.split(' ')) #put longer sentence in new
                 chunk
    else:
        chunks.append(s.split(' ')) #append words of sentence to
            new chunk

l = len(chunks)
for i in range (l): #index to loop through the chunks
    chunks[i] = ' '.join(chunks[i])

for i in chunks:
    print(i)


import requests

API_URL = "https://api-inference.huggingface.co/models/facebook/
    bart-large-cnn"
headers = {"Authorization": ""} #ADD API CODE

def query(payload):
    response = requests.post(API_URL, headers=headers, json=
        payload)
    return response.json()

outputs = []
for chunk in chunks:
    output = query({
    "inputs": chunk})
    outputs.append(output[0]["summary_text"])


openai.api_key = # ADD OPEN AI API KEY


outputs_openai = []
for chunk in chunks:
    prompt = "My second grader asked me what this passage means:\n
        \"\"\"\n" + chunk + "\n\"\"\"\nI rephrased it for him, in
        plain language a second grader can understand:\n\"\"\"\n"
    response = openai.Completion.create(
      engine="davinci",
```

```
      prompt=prompt,
      temperature=0.5,
      max_tokens=110,
      top_p=1,
      best_of=2,
      frequency_penalty=0.2,
      presence_penalty=0,
      stop=["\"\"\""]
    )
    outputs_openai.append(response['choices'][0]["text"])



## DOCUMENT
print(textstat.flesch_kincaid_grade(y))
print(textstat.flesch_reading_ease(y))
print(textstat.smog_index(y))
print(textstat.gunning_fog(y))


## BART SUMMARY
summary = "".join(outputs)
print(textstat.flesch_kincaid_grade(summary))
print(textstat.flesch_reading_ease(summary))
print(textstat.smog_index(summary))
print(textstat.gunning_fog(summary))


outputs_openai = "".join(outputs_openai)

## OPEN AI OUTPUT
print(textstat.flesch_kincaid_grade(outputs_openai))
print(textstat.flesch_reading_ease(outputs_openai))
print(textstat.smog_index(outputs_openai))
print(textstat.gunning_fog(outputs_openai))




print(summary)

print(outputs_openai)
```

References

1. Gabriele Meiselwitz.: Readability Assessment of Policies and Procedures of Social Networking Sites. In: Proceedings of the Fifth International Conference on Online Communities and Social Computing, Springer International Publishing. https://doi.org/10.1007/978 − 3 − 642 − 39371 − 6_8
2. Many people accept privacy policies without reading them, study reveals, https://thepaypers.com/payments-general/many-people-accept-privacy-policies-without-reading-them-study-reveals–1245901. Last accessed 2nd Dec 2021.
3. Wilson, Shomir Schaub, Florian Dara, Aswarth Abhilash Liu, Frederick Cherivirala, Sushain Giovanni Leon, Pedro Schaarup Andersen, Mads Zimmeck, Sebastian , Sathyendra, Kanthashree Mysore , Russell, N. Cameron , Norton, Thomas B. , Hovy, Eduard , Reidenberg, Joel , Sadeh, Norman. : The Creation and Analysis of a Website Privacy Policy Corpus. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). https://doi.org/10.18653/v1/P16-1126
4. Brown, Tom, Mann, Benjamin , Ryder, Nick , Subbiah, Melanie , Kaplan, Jared D , Dhariwal, Prafulla , Neelakantan, Arvind , Shyam, Pranav , Sastry, Girish , Askell, Am Agarwal, Herbert-Voss, Ariel , Krueger, Gretchen , Henighan, Tom , Child, Rewon , Ramesh, Aditya , Ziegler, Daniel , Wu, Jeffrey , Winter, Clemens , Hesse, Chris , Chen, Mark , Sigler, Eric , Litwin, Mateusz , Gray, Scott , Chess, Benjamin , Clark, Jack , Berner, Christopher , McC ,lish, Sam , Radford, Alec , Sutskever, Ilya , Amodei, Dario. Language Models are fewshot learners. In: Advances in Neural Information Processing Systems. url = https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf
5. Lewis, Mike , Liu, Yinhan , Goyal, Naman , Ghazvininejad, Marjan , Mohamed, Abdelrahman , Levy, Omer , Stoyanov, Veselin , Zettlemoyer, Luke". BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.703