POL-I-SEE

Samriddh Singh
IIIT
samriddh20466@iiitd.ac.in
Delhi, India

Varun Parashar
IIIT
varun20482@iiitd.ac.in
Delhi, India

Aaditya Gupta
IIIT
aaditya20552@iiitd.ac.in
Delhi, India

Harjeet Singh Yadav
IIIT
harjeet20561@iiitd.ac.in
Delhi, India

Diya Ahuja
IIIT
diya20431@iiitd.ac.in
Delhi, India

Ishita Sindhwani
IIIT
ishita20305@iiitd.ac.in
Delhi, India

February 21, 2023

Abstract

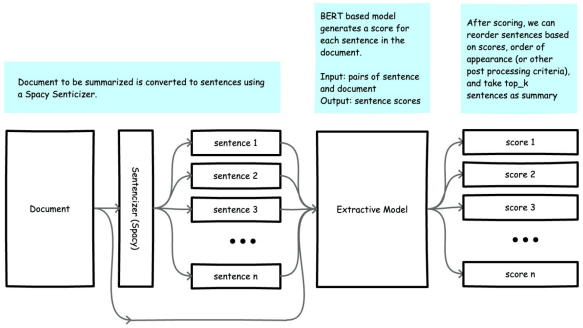
In this project, we plan to make an app that would simplify the mundane task of policy reading, which most of us skip, by presenting the policies of apps in a summarized manner. The summary of an app provides the user with the critical aspects of the privacy policy with which they can decide whether they want to give the app access to the information mentioned in the guidelines. Our motivation was to protect the users from digital predators who seek to access users' personal information without their knowledge. Privacy policies are a gateway through which people fall prey to digital predators like in a busy world like ours; no one has the time to read through tens of pages just before downloading an application. The challenge here is to make the people trust our application and make an algorithm that could distinguish between the junk and essential parts of a privacy policy. These problems can be solved by digital transparency by us towards the users and by legitimizing are apps through various app testing agencies. The algorithm for finding out the key points can be done using keywords to be searched in the complete policy. The algorithm will require optimizations done at every stage, but that is work that we can put in for a safer digital tomorrow.

1 Introduction

First, a privacy policy outlines how a company collects, uses, and shares our personal information. By skipping it, we may not know how your data is being used and who it is being shared with. This can leave us vulnerable to identity theft, data breaches, and other privacy violations. Second, a privacy policy may contain important information about our rights and personal data. For example, it may explain how we can request access to our data or request that it be deleted. By skipping the policy, we may not be aware of these rights and may miss out on the opportunity to exercise them. Third, many companies require us to agree to their privacy policy before using their services. By agreeing without reading, we may unknowingly be giving up substantial rights and consenting to practices that we are uncomfortable with. In short, skipping a privacy policy can be dangerous because it can leave us vulnerable to privacy violations, limit our rights, and expose us to risks we may not be aware of. Reading the policy carefully before agreeing to it or using the company's services is always a good idea.

2 Techniques and Algorithms

We will collect the app policies from the different online sources and play store. We will maintain a database containing the app policies and use positional inverted indexing to retrieve the required app policies from the database. For the summarization task we will use NLP transformers models which are the best for the task. BART and PEGASUS are widely used for retrieving the keywords from a big corpus of unstructured data. We are going to try other transformer models as well and will choose the one that performs well.



Neural Extractive Summarizing technique with transformers

3 Related Works

In the field of contract and policy summarization, AI and natural language processing are used to identify critical and dangerous points in legal documents. The goal is to help individuals and companies better understand the potential risks and consequences of the agreements they are entering into. One such project is LexCheck, an AI-powered contract review platform that identifies risky clauses in legal contracts, highlighting clauses that may be non-compliant with legal and regulatory requirements [1]. Similarly, Seal Software is an AI-powered contract analytics platform that identifies critical information from legal contracts and flags potentially risky clauses or language [2]. LawGeex is another AI-powered contract review platform that identifies critical clauses in legal contracts, including issues related to data privacy, compliance, and other potential risks [3]. Evisort is an AI-powered contract management platform that uses natural language processing to extract key information from legal contracts and highlight any potential issues [4]. Finally, Luminance is an AI-powered contract review platform that identifies potential areas of risk and highlights critical clauses in legal contracts [5]. All of these projects use AI and natural language processing to identify potentially risky clauses or language in contracts and policies. This can be especially important for legal contracts and policies, which can be complex and difficult to understand for those without legal training. By identifying critical and dangerous points in legal documents, these projects can help individuals and companies avoid potential legal issues and protect themselves from financial and reputational harm. As AI and natural language processing technologies continue to advance, it is likely that we will see further innovations in the field of contract and policy summarization, making it easier than ever for individuals and companies to understand the legal agreements they are entering into.

4 Novelty

In the current models, comparison of different apps at the time of download has not been implemented. We plan to maintain a database of standard policies required by different categories of applications. Our model will then browse through various applications of the same category and compare the policies and terms and conditions of the newly downloaded app to the ones in our database. It will help us to understand which permissions are required by the newly downloaded application to function properly and which are extra and unnecessary. This will further enable the user to decide whether to download the app or to look for the alternatives.

5 Evaluation

There are various ways in which we can evaluate our model. Some of these ways are: 1). Using user feedback from various sources including google play store ratings, survey forms and polls. 2). We can evaluate how well our application is performing by manually creating a database with some applications which contains summarized information about the policies and comparing it to those generated by our application. 3). We can also see if it gives correct reviews about whether an application is asking for more permissions that it needs to serve its purpose by comparing it to an application performing a similar job.

6 Potential Contributors

1).Varun Parashar (2020482) - Front-end Development, Database Management, Indexing and NLP Techniques.

2).Samriddh Singh (2020466) - Front-end Development, Database Management, Indexing and Model Training.

3).Aaditya Gupta (2020552) - Database Management, Indexing, NLP Techniques and Model Training.

4).Harjeet Singh Yadav (2020561) - Database management, Indexing, NLP techniques and Model Training.

5).Ishita Sindhwani (2020305) - Database Management, Indexing, NLP techniques and Model Training.

6).Diya Ahuja (2020431) - Database Management, Indexing and NLP techniques.

References

[1] J. Kosseff, "How artificial intelligence is changing contracts," *American Bar Association*, 2019.

[2] P. Rohner. (2020) Seal software uses ai to flag risky language in contracts. [Online]. Available: https://www.law.com/legaltechnews/2020/03/04/seal-software-uses-ai-to-flag-risky-language-in-contracts/

[3] Legal Tech News. (2020) Lawgeex expands ai-contract review platform. [Online]. Available: https://www.law.com/legaltechnews/2020/06/16/lawgeex-expands-ai-contract-review-platform/?slreturn=20230017092024

[4] M. Araujo. (2018) Evisort launches ai-powered contract management platform. [Online]. Available: https://www.legaltechnews.com/home/id=1202795735316/Evisort-Launches-AIPowered-Contract-Management-Platform?mcode=1395264250916curindex=1

[5] Ukonomics. (2018) Ai: The future of law? [Online]. Available: https://www.ukonomics.com/single-post/2018/06/06/AI-The-Future-of-Law