

Technical Disclosure Commons

Defensive Publications Series

March 30, 2018

Automatic evaluation of contractual terms

Alex Zheng

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Zheng, Alex, "Automatic evaluation of contractual terms", Technical Disclosure Commons, (March 30, 2018)
https://www.tdcommons.org/dpubs_series/1123



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Automatic evaluation of contractual terms

ABSTRACT

This disclosure describes use of machine learning and natural language processing techniques to assist users in reviewing terms and conditions (T&C), e.g., provided by online product and service providers. A database of T&C is created and indexed according to factors such as country, industry, application, etc. The database is utilized to identify T&C content that is common for a given industry application in a specific jurisdiction. When a user provides permission to analyze and provide assistance for a T&C document, such document is analyzed to determine a risk score. Further, any anomalous content in the T&C is flagged. The user is provided options to conduct in-depth review of the T&C, and to accept or decline various terms. When user permit, individual preferences for various terms are stored and matched to the T&C document.

KEYWORDS

- User preference
- Risk assessment
- Terms and conditions
- Online services
- Contracts

BACKGROUND

Providers of online services or products often require users to read and accept terms and conditions (“T&C”) prior to provision of the services or products. Typically, the T&C include, e.g., terms regarding user privacy, e.g., information regarding user data that may be obtained,

stored, and/or used by the provider, agreements between users and providers with respect to the product/service provided, disclaimers, etc.

Users often accept the T&C from a provider without conducting an in-depth review, e.g., since the T&C documents are often lengthy and time consuming to read. For example, some surveys indicate that only about 7% of users read T&C when shopping online. This percentage may be even lower for mobile device users, e.g., if the T&C are harder to read on smaller screens.

The terms and conditions for different providers may often be similar. However, some providers may include additional requirements that users may be unwilling to accept.

DESCRIPTION

This disclosure describes techniques to utilize natural language processing and machine learning techniques to assist users in reviewing terms and conditions (T&C) presented in various contexts. A database of T&C used by multiple providers is created and indexed, e.g., by country, industry, application, etc. The indexed T&C are analyzed to determine T&C content that is relatively common for a given industry application in a specific country.

When a user is presented with a set of T&C for review and/or acceptance, with user permission and express consent, the T&C is analyzed and compared with the database. Based on the analysis, unusual or uncommon content or clauses are flagged for user review. Analysis of the T&C takes into account contextual factors such as the industry, application, geography, etc., when permitted by the user.

The T&C are analyzed to determine whether the documents propose commonly accepted content or whether closer user scrutiny is necessary. Anomalous content is extracted from the T&C document for user review. Additionally, the T&C document is scored based on overall risk

assessment and the score is shared with the user. The score can be represented by a number, an alphabet or a level, e.g., high, medium, or low.

If the T&C document is determined as not requiring user scrutiny, the T&C analyzer can indicate to the user that the document is fine to accept, e.g., “low risk, all clear.” Users are provided an opportunity to review excerpts of the extracted anomalous T&C, e.g., to determine whether to scrutinize the entire T&C document. Based on NLP and machine learning techniques, only some portions of T&C documents that are determined as requiring user review are extracted and presented to the users. This enables users to review T&C documents efficiently, e.g., by eliminating the need to review the entire document.

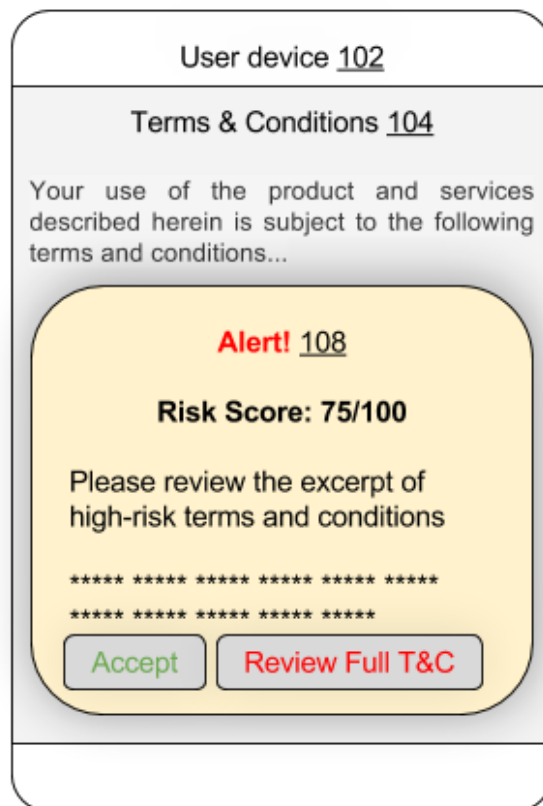


Fig. 1: Terms and conditions alert with risk assessment score

Fig. 1 illustrates a user device (102) that is displaying a terms and conditions document (104). In the example illustrated in Fig. 1, the user has provided permission for comparison of the document with the database, e.g., using indexed T&C document for the relevant country and industry application. A risk score is computed and anomalous terms and conditions are extracted. The risk score and anomalous content is displayed, e.g., as a user alert (108) on the user device. The user is provided with the options of accepting the T&C or reviewing the full T&C.

Users are enabled to specify individual preferences regarding acceptable and objectionable types of terms and conditions. With user permission, such individual preferences are used to filter in or out content within a given T&C. For example, a user that is willing to share a particular type of data can choose to universally apply acceptance of the type of data to all T&C that are presented to the user. With user permission and express content, machine learning techniques are used to determine user preferences over time based on the user's history of acceptance or rejection of terms and conditions.

The use of machine learning enables adaptation and indexing of novel T&C concepts as the corresponding industry practices, laws, and regulations of the particular jurisdiction evolve. NLP techniques are utilized to determine intent for particular terms in the T&C. The use of wordplay and new phrases, e.g., to mask T&C content, can be identified and flagged as unusual, anomalous or high-risk.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user's social network, social actions or activities, profession, a user's preferences, or a user's current location), and if the user

is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes use of machine learning and natural language processing techniques to assist users in reviewing terms and conditions (T&C), e.g., provided by online product and service providers. A database of T&C is created and indexed according to factors such as country, industry, application, etc. The database is utilized to identify T&C content that is common for a given industry application in a specific jurisdiction. When a user provides permission to analyze and provide assistance for a T&C document, such document is analyzed to determine a risk score. Further, any anomalous content in the T&C is flagged. The user is provided options to conduct in-depth review of the T&C, and to accept or decline various terms. When user permit, individual preferences for various terms are stored and matched to the T&C document.