Research Planning - Prediction of voting decisions of Proxy advisors toward election of director based on Proxy reports guided by Policy

Hongsang Yoo

1 Motivation

While there have been some researches to predict political election results [1] or select directors based on performance [2], there has been no attempt yet to predict the election of board of directors.

So far, less focus has been put on individual shareholders since the result of the voting is dependent on largest shareholders and individual shareholders often do not have enough knowledge that can lead them to vote properly. [3] Currently, companies which provide vote recommendation such as ISS and Glass Lewis hire a number of seasonal analysts in the time of annual shareholder meetings and train them based on their internal guidelines, which causes the company time and monetary loss. The proposed research can tackle these problems above.

I, myself worked as an analyst both in ISS and Glass Lewis and thought that this human labour of analysing agenda and vote recommendation could be automated. This has been the biggest motivation for planning the proposed research. Existing similar researches, where the authors tried to predict political election results, used sentiment extraction. However, there is a significant limitation in this sentiment-based approach: in most cases, directors are not familiar to the public, which means no sentiment polarity at all. Therefore, applying existing approach of sentiment analysis are not suitable for the goal of this research.

Since there has been no research in this specific topic, we will adopt and utilize findings and methods studied in previous literature review which are plausible to implement system. Firstly, we will build hierarchical representation of Proxy reports based on the methods studied in [6][7]. This method will be used again in the construction of ontology from Policy and matches between this two (Proxy reports and Policy) [8]. Lastly, we will adopt hierarchical deep learning model [9] for our final model. We will look at the terms mentioned above in the next section and discuss how to apply these methods in different phases in section 4.

2 Identification of the research question

The research question of the proposed research is Prediction of voting decisions of Proxy advisors toward election of directors based on Proxy reports guided by Policy. Let's look at important terms in this question.

2.1 Definition and description

2.1.1 Shareholder meeting

All listed companies should hold an annual shareholder meeting and list agendas such as 'Election of directors', which is the only agenda we are going to focus on among many, to earn the approval of shareholders.

2.1.2 Proxy reports

When companies list agendas, they upload relevant documents to reveal information to shareholders. For example, if the agenda is 'Election of directors', the company uploads the profile of directors and their activity in the board meetings during the past year, if they are already incumbent.

2.1.3 Shareholders and Proxy advisors

All shareholders can exercise one vote toward agenda per one share [4]. There are two types which exercise voting power: one is individual shareholders and the other is institutional shareholders, which take the authority to vote on behalf of individual shareholders. [4][5] In voting decisions, individual shareholders vote based on their interest [3]. On the other hand, institutional shareholders or consulting firms such as ISS and Glass Lewis which provide vote recommendation to them, depend on their internal policy [5]. We will refer to consulting firms that provide votes recommendation as "Proxy advisors" in this paper and we are interested in their decisions.

2.1.4 Policy

Proxy advisors use their internal policies to help analysis of proxy reports. Some examples of this policy are as follows [5] [1] :

- Vote against the nominee with attendance less than 75 % of the board meeting.
- The board must be majority independent.

2.2 Components of a question

There are three components of this research question: 1) Proxy reports 2) Policy 3) Labels. Therefore, we will need these three types of data, which will be covered in the next section. For labels, we will use voting decision of Proxy advisors (For/Against). The reason why the result of shareholder meetings is not used as labels is that the actual voting result tends to be largely dependent on largest shareholders no matter what the vote recommendation of consulting companies [3]. More specifically, we will consider this problem as text classification, where the input is the texts of Proxy reports and external Policy and the output is binary label, which is the voting decision (For/Against) of Proxy advisors.

3 Investigation method

3.1 Data gathering

While the purpose of the proposed research isn't limited in the shareholder meetings of a specific country, it is necessary to narrow down on one country due to the different language and system of the shareholder meeting. Here, the data of Korean companies will be used as I have experienced working towards Korean market as an analyst in ISS and using web API for gathering data is available in the official website for this country. There are mainly three types of data necessary to obtain: 1) Proxy reports 2) Policy 3) Labels.

3.1.1 Proxy reports Data

Firstly, it is necessary to obtain Proxy reports uploaded by companies for the shareholder meeting. An official website for shareholder meetings by Korean government, DART[2], will be used as a source of this as all listed companies upload all the relevant documents here. As a method for gathering data, web-API that this website provides [3] will be used. Thus, there needs to implement a function to use API to gather the data. For the volume of data to be crawled by using API, the amount of data equivalent to one year is expected as there are about 10,000 annual shareholder meetings and this number will be sufficient to train networks.

3.1.2 Policy Data

Second type of data is Policy, which Proxy advisors use for their recommendation. Guidelines that can be found publicly on the web published by ISS or Glass Lewis [5] will be used for this data as this is an official policy from Proxy advisors that we seek for, which will be used to build ontology system.

3.1.3 Label Data

For labels, this website will be used [4]. The reason why we use separate website for gathering labels is that as discussed in identification section, we are going to need vote recommendation of Proxy advisors as labels instead of actual voting results. As a method for gathering data from this website, web-crawling will be used. This can be done via using web crawling software or implementing codes.

3.2 Procedure to implement the system

After securing data, the overall process to implement the system is as follows:

3.2.1 Proxy reports - Building hierarchical representation

As an alternative of general word embeddings, it is necessary to transform texts of Proxy reports into hierarchical representation. With this hierarchical structure, the documents can capture contents such as profiles of directors better.

3.2.2 Policy - Ontology construction

We will construct ontology from Policy that we obtained in 3.1.2

3.2.3 Classification using Hierarchical Deep learning guided by ontology

We will use hierarchical deep learning architecture to apply the result of 3.2.2 to the result of 3.2.1 to predict label. In this architecture, the

[1] https://www.issgovernance.com/file/policy/2017-2018-Australia-Voting-Guidelines.pdf

[2] https://englishdart.fss.or.kr

[3] https://opendart.fss.or.kr/intro/main.do

[4] http://www.cgs.or.kr/eng/main/main.jsp

input can capture more structural information than flat word embedding.

Models and methods to be used in each phase to implement the system as well as test methods will be covered in the next section.

4 Models and Test

4.1 Proxy reports - Building hierarchical representation

4.1.1 Algorithms description

1) Produce a set of isA (sub-super concept) pairs, where super-concept means the higher concept that sub-concept belongs to.

2) Perform node merging operations based on similarity functions between nodes.

4.1.2 Methods

We will use probabilistic approach and syntactic patterns.

1) Detect relevant Named Entity We will rank candidates by using Web-based co-occurrence statistics by computing PMI (Point-wise Mutual Information [6] and query search engine to determine appropriate synonyms to a given query and the algorithms chooses a synonym among a given list [7].

2) Sub-Super concept mapping The next step is to discover possible super concepts for each named entity using Hearst's taxonomic linguistic patterns seen in Table 1. [7][8]

Pattern	Example
Such NP as {NP,}* (and|or) NP	Such countries as Poland
NP {,} such as {NP,}* (and|or) NP	Cities such as Barcelona
NP {,} including {NP,}* (and|or) NP	Capital cities including London
NP {,} specially {NP,}* (and|or) NP	Science fiction films, specially Matrix
NP {,} (and|or) other NP	The Sagrada Familia and other churches

Table 1: Hearst patterns

4.2 Policy - Ontology construction

Repeat the same process as 4.1. to for Policy data to construct ontology from Policy.

4.3 Classification using Hierarchical Deep learning guided by ontology

4.3.1 Concepts Match between Proxy reports and Policy

Matches between the super concepts of a named entity of Proxy reports and the ontological classes constructed from Policy are performed [8], using the same step as 4.1.2 concept mapping. The difference is that this time rather than sub-super concept mapping, 'Super concept' matches between Proxy reports and Policy should be performed.

4.3.2 Merging into deep learning

We will use two levels in our Deep Neural Network studied in [9]. The first level DNN is the same as widely known. The second level consists of a DNN trained for the output in the first level. That is, the input of second level DNN is the output of the first level. For example, if the output of the first level network is labelled 'outside director', the DNN in the next level is trained only with subordinate information within 'outside director' to predict the election of this director rather than irrelevant information belonging to 'inside director'.

4.4 Test methods

4.4.1 Baseline

TF-IDF with Logistic regression will be used as a baseline. TF-IDF, which is term frequency and inverse document frequency model, is often considered to be a good model for text classification. It fits many different kinds of supervised machine learning methods among which Logistic Regression will be used. The reason is that without implementation of word embeddings, TF-IDF can easily show baselines based on term-document frequency.

4.4.2 Evaluation metrics

F1 score toward negative labels (Vote against) will be used for evaluation metrics as it is more important to detect negative labels in that there are certain grounds in policy why it is negative.

- Precision: Precision is calculated as the number of correct positive results divided by the number of all positive results returned by the classifier.
- Recall: Recall is calculated as number of correct positive results divided by the number of all relevant samples
- F1 score: F1 score is calculated as harmonic mean of the precision and recall.

5 Contribution of the research

In this paper, we looked at planning and methods in order to implement the system to predict voting decisions of Proxy advisors toward election of directors. Based on prior literature reviews, procedure for implementing the system from data gathering to test phase will be pursued by using methods discussed in this paper. Although there has been some similar topics [1][2], there has been no research in this specific research question, that is the prediction of voting decisions of Proxy advisors toward the director elections.

Prediction of Proxy advisors' decision is significant since they provide vote recommendation to institutional shareholders and vote shares of institutional shareholders, which take the authority to vote on behalf of individual shareholders, take up a large portion of the total and thus influence the decision of the individual shareholders as well [4][5]. Therefore, while previous existing similar research such as [2] contributes to the benefit of company and the directors in the board, the proposed research contributes to decision of entire individual shareholders. This contribution is significant in that individual shareholders are often ignorant of important agenda in deciding vote for/against them [3] and they still hold large portion of the total number of voting.

In addition, Proxy advisors that recommend vote for/against agenda such as ISS and Glass Lewis often hire a great number of seasonal analysts in the time of annual shareholder meetings and train them based on their internal guidelines (Policy). The proposed research discussed in this paper, if properly implemented, can automate all the process in analysis of these companies thus saving money and time.

With reasons above, the proposed research contributes to the benefit of individual shareholders, who take up a great number of the total votes, and saving resources for the consulting companies which provides vote recommendation

References

[1] P. KassraieAlireza, M. Modirshanechi, H.K. Aghajan. Election Vote Share Prediction using a Sentiment-based Fusion of Twitter Data with Google Trends and Online Polls. In Proceedings of the 6th International Conference on Data Science, Technology and Applications, Madrid, 2017, pp. 363-370.

[2] I. Erel, L. Stern, C. Tan, M. Weisbach. Selecting Directors Using Machine Learning. IDEAS Working Paper Series from RePEc, 2018.

[3] E. Maug, K. Rydqvist. Do Shareholders Vote Strategically? Voting Behavior, Proposal Screening, and Majority Rules. Engineering Review of Finance, Jan 2009, Vol.13(1), pp.47

[4] H. Hansmann, M. Pargendler. The Evolution of Shareholder Voting Rights: Separation of Ownership and Consumption. Yale L.J., Vol.123 pp.948-3104, 2014

[5] Glass Lewis, 2019 Proxy Paper Guidelines: Shareholder Initiatives (Oct. 24, 2018)

[6] W. Wu, H. Li, H. Wang, K. Zhu. Probase: a probabilistic taxonomy for text understanding. Proceedings of the 2012 ACM SIGMOD International Conference on management of data, 20 May 2012, pp.481-492.

[7] M. Roche, Y. Kodratoff. Text and Web Mining Approaches in Order to Build Specialized Ontologies. JoDI: Journal of Digital Information, 2009, Vol.10(4).

[8] C. Vicient, D. Sánchez, A. Moreno. An automatic approach for ontology-based feature extraction from heterogeneous textual resources. Engineering Applications of Artificial Intelligence, March 2013, Vol.26(3), pp.1092-1106.

[9] K. Kowsari, D. Brown, M. Heidarysafa, M. Gerber, L. Barnes. HDLTex: Hierarchical Deep Learning for Text Classification. arXiv:1709.08267, Oct 2017.