# Prediction of voting decisions of institutional shareholders toward nominees for director based on proxy reports guided by policy

Hongsang Yoo

## 1 Introduction

While there have been some researches to predict political election results [1][2] or select directors based on performance [13], there has been no attempt yet to predict the election of board of directors. While there are many factors affecting directors' election, the decision of institutional shareholders plays a significant role since their voting shares take a large portion of the total.

The aim of this paper is to review related literature to provide meaningful approach in prediction of voting decisions of institutional shareholders toward nominees for directors (hereafter called " directors "). More specifically, we will consider this problem as text classification, where the input is the texts of proxy reports and output is binary label, which is the voting decision (For/Against) of institutional shareholders. First, we are going to look at the limitation of applying existing approach in this problem. And we will study how to represent information of proxy reports hierarchically. Finally, in order to reinforce the deficiency of the inference, ontology-based model will be introduced.

## 2 Background

There are two types which exercise voting power: one is individual shareholders and the other is institutional shareholders, which take the authority to exercise of the vote on behalf of individual shareholders. In voting decisions, while institutional shareholders or consulting firms such as ISS and Glass Lewis which provide vote recommendation to them, which will be referred to as "analysts" in this paper, depend on their internal policy [12], individual shareholders vote based on their interest. Sentiment analysis may be considered in the case of latter but there is a limitation in this approach.

### 2.1 Limitation of existing approach

In the recent researches, sentiment analysis has been the most popular approach in the prediction of the vote result of political elections. In the research of KassraieAlireza et al., a large dataset of tweets was gathered and assigned to sentiment values, collecting relevant keywords from Google Trends and arranging various online poll results [1]. On the other hand, Parackal et al. used values-based method for predicting election results [2]. They carried out an experiment where they selected samples from consumer panels, from which extracting values such as well-being and power associated with the political party by using Laswell and Kaplan's societal value lexicon. Prediction was made regressing these extracted values onto voting decision.

Both researches above used sentiment extraction and showed that their prediction was consistent with the final election results quite accurately. However, there is a significant limitation in this sentiment-based approach: in most cases, directors are not familiar to the public, which means individual shareholders do not have sentiment polarity at all. Values-based method [2] may be an alternative in that this method does not directly rely on sentiment analysis towards a certain director. Still, there is a problem that, unlike political election, there is not clear ideology to be associated with the directors. Therefore, the existing approach of sentiment analysis may not be suitable for the goal of this research. To simplify the problem, we are going to consider only institutional shareholders and try to predict their decision.

## 3 Classification based on the proxy reports

As discussed in the background of the previous section, analysts depend on their internal policy when they analyse proxy reports. Here, proxy reports are the input document studied in this section and policy is covered in the next section regarding ontologies.

### 3.1 Limitation of general document representation

While word embedding is the most well-known for document representation, one of its limitations is that words with multiple senses may map into a single representation[1]. In addition, it does not capture the hierarchical sense of the texts. For example, in a pre-processed text "Director, name, position, 51, ...", it does not know whether '51' is the age of the director. While there are many state-of-the-art techniques to solve this problem, approaches to represent document hierarchically are studied in this paper.

### 3.2 Hierarchical representation of the document

With the reasons above, documents need to be represented in a hierarchical and structural way. First, taxonomy is one way for this. A taxonomy, which is in between the ontology spectrum [3], is a controlled vocabulary organized into a hierarchical or parent-child structure [4]. While ontologies were proposed to enable representations of knowledge [3], the main task of taxonomy is focused on constructing hierarchies.

However, unlike traditional taxonomies that treat knowledge as black and white, Wu et al. proposed an approach that uses probabilities to model ambiguous and uncertain information [6]. In each round of information extraction, their system accumulates knowledge and then uses this knowledge in the next round to help extract information missed previously. The author used a probabilistic approach for the detection of super-concept and sub-concept to make a hierarchical structure. Once large set of isA (sub-super concept) pairs were produced, the taxonomy is constructed by node merging operations based on similarity functions between nodes. With this hierarchical structure from taxonomy, if well structured, the documents can capture contents such as profiles of directors better. In addition, it is less likely to miss newly appeared information since they used previously acquired knowledge beyond syntactic patterns.

The effectiveness of the hierarchical text classification has not been assessed much yet. Steina et al. evaluated the models with measures appropriate for the hierarchical context [7]. They showed that using the right combination of word embeddings (i.e. GloVe and fastText) and machine learning algorithms leads to a good result. In the experiment of the paper, FastText achieved the best result. However, this result is limited to a certain dataset thus to find out the best result, an experiment with the right algorithms paired with word-embedding methods on the dataset is necessary.

### 3.3 Hierarchical text classification in deep learning

Now we need to merge this hierarchical representation into deep learning model. Kowsari et al. proposed a method of Hierarchical Deep Learning for Text classification, which uses stacks of deep learning architectures to provide understanding of the document hierarchy [8]. Their proposed basic architecture of DNN is extended to allow hierarchical classification. There are two DNN levels in this architecture. 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.

In this hierarchical deep learning architecture, the input can capture more structural information than flat word embedding. For example, if the output of the first model is labelled 'outside director', the DNN in the next hierarchical level is trained only with subordinate information within 'outside director' to predict the election of this 'outside director' rather than irrelevant information belonging to 'inside director'.

With hierarchical representation of the document merged into deep learning model, now the text can capture more structural meanings for classification. However, when analysts decide their votes based on proxy reports, they use their own policy. In the next section, we are going to look at how to build policy from external sources and apply it to text analysis.

## 4 Ontology-guided text classification

Analysts use their internal policies to help analysis of proxy reports. Some examples of this policy[2] are as follows:

[1]https://en.wikipedia.org/wiki/Word_embedding

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

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

We need to 1) construct ontology from these policies written in texts and 2) apply this policy to text analysis.

### 4.1 What is Ontology?

Compared to taxonomy discussed in the previous section, ontologies, rather than defining how its structure should be, cover a spectrum of useful artifacts [3]. In AI, ontology means 1) representation vocabulary, often specialized to some domain and 2) a body of knowledge describing some domain, using the representation vocabulary. [5]. Then how can we construct ontology from the text? This is based on conceptualization. Every knowledge base has to be committed to a conceptualization and can be created by extracting the relevant instances from information to form ontologies [4]. We are going to look at the methods of ontology construction in this section.

### 4.2 Ontology construction from text

Lee et al. proposed a method for automated ontology construction from unstructured text [9]. Episode, which means a partially ordered collection of events occurring together, is extracted based on its occurring frequency. And then the terms are mapped to the result of the concept clustering of episodes. Finally, the fuzzy inference mechanism is adopted to obtain new instances. Their results showed that their proposed approach successfully constructed the Chinese text domain ontology. However, some algorithms used in this paper contained Chinese language specific rules. It is notable that they used fuzzy inference. It may be useful when we must construct ontology from the target document only. Still, some important terms that appear less frequently might not to be extracted. In order to tackle this problem, the statistical approach of web mining combined with text analysis could be an alternative.

Roche et al. applied a text-mining method to extract candidate terms and web-mining approach to validate the extracted candidates [10]. Their proposed system extracts the candidate terms from a corpus to associate with concepts. Although statistical measure is used at this stage, they are often inadequate for a small corpus. To solve this problem, they applied web-mining approach. This method ranks the candidates using the algorithm PMI-IR (Pointwise Mutual Information and Information Retrieval) and queries search engine to determine appropriate synonyms to a given query and PMI-IR chooses a synonym among a given list. Their results showed that their experiment improves the results when associated with statistical measure.

While the general approach of both researches above for extracting terms for conceptualization is similar, the statistical web-mining methods can complement the limitation that the previous research [9] had. However, as seen in the examples of the policy of this section, the ontological structure needs to handle axioms[3]. Due to the limitation of time and space, the detailed research for this was not included in this paper. We looked at how to construct ontology from texts so far and these will be used when building a policy from external text resources such as ISS's policy guidelines[2]. In order to apply this ontology in the analysis of proxy reports, methods of extracting features from reports based on the ontology will be introduced in the next section.

### 4.3 Ontology-based feature extraction

Vicient et al. proposed methods that extract relevant features, given an input document and an ontology stating which features should be extracted [11]. First, to detect the most relevant named entities, they used Web-based cooccurrence statistics by computing PMI (Pointwise Mutual Information in a similar fashion discussed in the previous study using web mining [10]. Once Named Entities were found, the next step is to discover possible super concepts for each named entity using Hearst's taxonomic linguistic patterns seen in Table 1, which have proved their effectiveness to retrieve hyponym/hypernym relationships. Hearst's patterns can be usefully used in the previous stage of this paper, which is ontology construction, too. Here it is notable that these super concepts are abstractions of the named entity and do not depend on any ontology. Finally, matches between the super concepts of a named entity and the ontological classes are performed.

Discovering first possible higher concepts of named entities is suitable for our work in that we already studied the way of hierarchical representation of the document [6][8] and it is easy to perform this task in the hierarchical

[3]http://www.orm.net/pdf/OntologicalModeling14.pdf

| 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

structure. However, again the constructed ontology may contain axioms as well as entities and methods of handling this should be dealt with in the further work.

## 5 Conclusions

There has been no research in the prediction of the director elections although there has been some similar topics [1][2][13]. This paper aimed to provide related literature reviews that are foundational for prediction of institutional shareholder's decision toward the directors' election based on proxy reports. As an alternative of general word embeddings, hierarchical document representation was studied. To provide expertise in an analysis of reports, the construction of ontology and ontology-aided feature extraction from documents were explained. Prediction of institutional shareholders' decision is significant since their vote shares take up a large portion of the total and thus influence the decision of the individual shareholders as well. As further work, it is a priority to study implementation on how to merge policy built in ontology (especially axioms) into document (e.g., Question Answering system). Other important line of work is to obtain proper internal policy guidelines[2].

## 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] M. Parackal, D. Mather, D. Holdsworth. Value-based prediction of election results using natural language processing: A case of the New Zealand General Election. International Journal of Market Research, March 2018, Vol.60(2), pp.156-168.

[3] M. Gruninger, O. Bodenreider, F. Olken, L. Obrst, P. Yim. Ontology Summit 2007–Ontology, taxonomy, folksonomy: Understanding the distinctions. Applied ontology 3(3), pp. 191-200.

[4] W. Wong, W. Liu, M. Bennamoun. Ontology Learning from Text: A Look Back and into the Future. ACM Computing Surveys (CSUR), August 2012, Vol.44(4), pp.1-36.

[5] B. Chandrasekaran, J.R. Josephson, V.R. Benjamins. What are ontologies, and why do we need them? IEEE Intelligent Systems and their Applications, January 1999, Vol.14(1), pp.20-26.

[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] R.A. Stein, P.A. Jaques, J.F. Valiati. An Analysis of Hierarchical Text Classification Using Word Embeddings. Information Sciences, January 2019, Vol.471, pp.216-232.

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

[9] C. Lee, Y.Kao, Y. Kuo, M. Wang. Automated ontology construction for unstructured text documents. Data & Knowledge Engineering, 2007, Vol.60(3), pp.547-566.

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

[11] 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.

[12] 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

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