

Random Forest を用いた企業価値に影響を与える CSR 活動の予測

三塚協† 金盛克俊‡ 大和田勇人‡

†東京理科大学大学院 理工学研究科 経営工学専攻 ‡東京理科大学 理工学部 経営工学科

1 はじめに

CSR 活動への取り組みが各企業の企業価値に大きな影響を与えるとされている。これに伴い CSR に対する関心は近年日に日に高まっており、企業は様々な CSR 活動への取り組みを行っている。しかしながら CSR は十分に定義されないまま成長した概念であり、具体的な取り組みや活動が実際にどの程度企業の価値に影響を与えるかは明らかにされていない。そのことから、多くの企業は現在も CSR 活動への取り組み方を模索し続けている状況である[1]。

そこで本研究は、CSR 企業総覧に記載されている各企業の CSR 活動の項目に対して機械学習の手法の一つである Random Forest を用いることで、各企業における企業価値が高いか低いかの分類予測を行い、その精度を元になどのような企業の特徴が企業価値に対して影響を与えているかを検証する。

2 提案手法

本研究の提案手法を以下に示す。

- Step1 各企業の財務データから企業ごとの企業価値を求める。企業価値は、Ohlson Model[2]に基づく企業評価モデルを用いて計算式を用いて算出する。
- Step2 Step1 によって算出した企業価値から、企業価値の高い上位 40%を正事例、下位 40%を負事例と設定する。本手法では中央の 20%は、正事例と負事例の差異を明確にするために考慮しないものとする。
- Step3 CSR 企業総覧に記載されている CSR 活動を表している項目の内容を、Random Forest の特徴量として用いる。項目の特徴量は、例えば『1.CSR 兼任部署の有無』といった項目では「専任部署あり」「兼任部署あり」「その他」「無し」「無回答」等といった選択肢が存在し、その選択肢に応じてラベル付けを行った値を置く。
- Step4 作成した特徴量に対し Random Forest を用いて学習・分類を行い、その結果から分類予

測精度に高い影響・低い影響を与える CSR 活動項目を検証する。

2.1 Random Forest

図 1 に Random Forest の概要を示す。

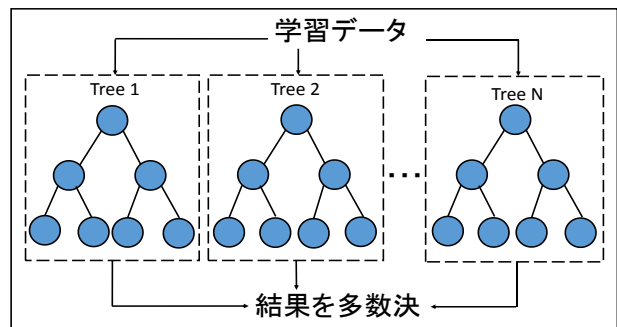


図 1. Random Forest 概要図

Random Forest は機械学習の手法の一つであり、学習データからランダムに変数をサンプリングした後に決定木を N 個生成し、それらの結果を統合することによって精度を向上させる手法である。Random Forest の利点は、各説明変数の重要度が得られることである。重要度は、各決定木のジニ係数の減少量を標準偏差で割った値から求められる。

本手法ではこの重要度を利用して、企業価値に大きな影響を与える項目を検証する。

3 実験

CSR 企業総覧 2014 年、2015 年度版に記載されている企業を対象として実験を行った。2014 年と 2015 年度の 264 社の企業の企業価値を算出し、提案手法の Step2 を用いて企業価値の高い方から上位 105 社、下位 105 社ずつをそれぞれ正事例、負事例とした合計 210 社を使用データとした。そして企業の CSR 活動を表す 37 項目を Random Forest の特徴量とした。表 1 にその一部を示す。

表 1. 使用する項目

No.	項目
1	CSR専任部署の有無
2	CSR担当役員の有無
3	マテリアリティ(重要性)の設定
	⋮
37	気候変動対応についての取り組み

Prediction of CSR activities that affects the corporate value using Random Forest

† Department of Industrial Administration, Graduate school of Science and Technology Tokyo University of Science.

‡ Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science

Random Forest の学習においては本実験では K 分割交差検定で分析を行った。対象となる 210 社のデータを 10 分割し、1 つをテストデータ、残る 9 個をトレーニングデータとした検定を行い、予測精度を出力した。

4 実験結果及び考察

表 2 に本実験の実験結果を示す。表中の TP, TN, FN, FP はそれぞれ True-Positive, True-Negative, False-Negative, False-Negative を表している。

表 3 予測精度

	2015年	2014年
TP	87	82
TN	88	78
FN	18	24
FP	17	27
Accuracy	0.833	0.762
Precision	0.837	0.752
Recall	0.829	0.781
F値	0.833	0.766

表 2 より、Accuracy, Precision, Recall, F 値の値が 2015 年と 2014 年でいずれも高い値を示していることがわかる。よってこの分類モデルはそれなりに信頼できるということが言え、このモデルから得られる各項目の重要度は信頼性が高いものであるということが言えるであろう。

各項目の重要度の上位 5 項目と下位 5 項目の結果を図 2, 図 3 に示す。

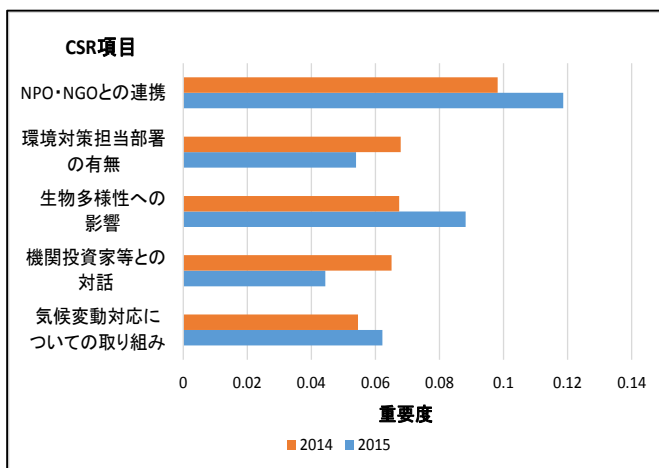


図 2. CSR 項目重要度(上位 5 項目)

図 2 「NPO・NGO との連携の有無」は、2014, 2015 年共に最も重要度が高いことから、予測精度結果を示す上で最も重要な項目であると判明した。

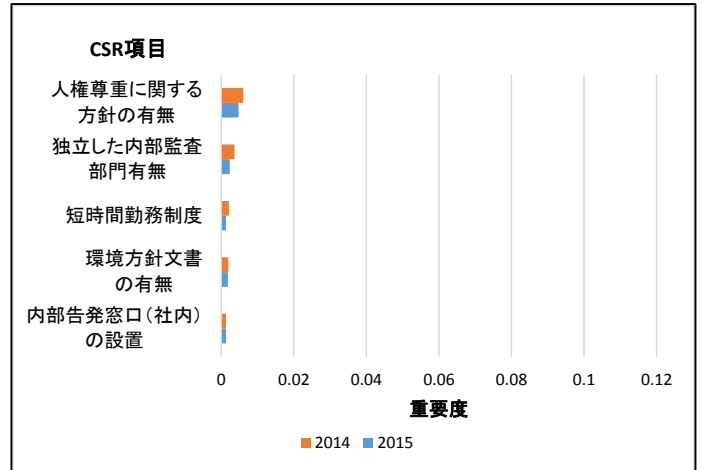


図 3. CSR 項目重要度(下位 5 項目)

一方で図 3 より、「内部告発窓口(社内)の設置の有無」は 2014 年, 2015 年共に最も重要度が低いことから、予測精度結果に殆ど寄与しないことが判明した。

5 まとめ及び今後の展望

本研究は CSR 企業総覧に記載されている各企業の特徴の項目に対して Random Forest を用いて企業価値の高低の分類予測を行うことで、どのような企業の CSR 活動の項目が企業価値に影響を与えるかを示した。

本手法より高い分類精度結果を得られ、分類結果に寄与する重要な CSR 項目を割り出すことが出来た。

しかし CSR 企業総覧の中には、本研究では考慮しなかったが、活動内容がテキスト文書で記述されている項目もいくつか存在する。よって今後は新たにテキストマイニング技術を用いて、テキスト中のどのようなキーワードが企業価値に強く影響を与えるか等について詳しく検証していく予定である。

参考文献

- [1] 殿崎正芳「CSR と財務成果の関係についての一考察」イノベーション・マネジメント No.11, pp.145-16 (2013).
- [2] Ohlson, J.A. "The theory of value and earnings, and an introduction to the Ball-Brown analysis" Contemporary Accounting Research, Vol. 8, No.2, pp.1-19 (1991).
- [3] 大菌陽子「企業業績が CSR の取組みに与える影響—日本企業の実証分析—」『成蹊大学一般研究報告』第 45 巻第 4 分冊, pp.1-14 (2012).