

Deep Learningを用いた 株価予測モデルの提案・検証

富士通株式会社 大山承剛 ohyama.shogo@jp.fujitsu.com

株価予測における問題点

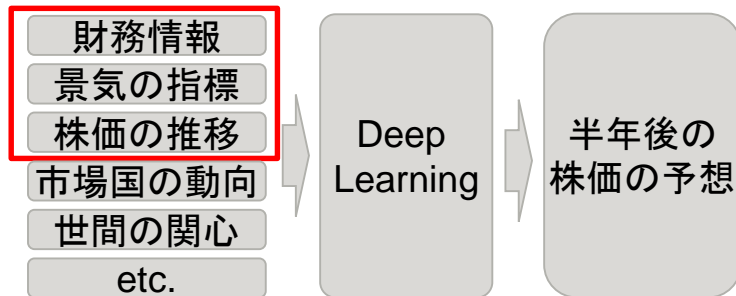
従来の株価予測のモデリングは高度な知識・ノウハウを要するが、株価予測には財務情報、ニュース、市場動向、景気指標、投資家の期待など様々な要因が関連し、モデリングそのものが非常に困難であった。

手法の適用による解決

Deep Learningは、高度な専門性を必要としたモデリングを行うことなくデータを分析することが出来るという特徴を持つ。この特徴を活かして、Deep Learningによる株価予測を提案・検証する。

モデルと実験

本実験での対象範囲



株価は様々な要因の基に変動するため、将来的には様々なデータを分析に利用する予定であるが、まずは財務情報、景気の指標、株価の推移をデータとして用いる。

実験: 半年後の景気(日経平均)が予測できれば、各企業の株価の変動を予測できるか

入力情報:

- 財務情報 : 過去5年の主要な財務科目の値
- 景気の指標: 報告日、1、2、3年前、半年後の日経平均の終値
- 株価の推移: 報告日、1、2、3年前の株価の終値

Deep Learningネットワーク

隠れ層: 2~7層, 荷重減衰(L2ノルム) 0~0.0004

出力

全企業の株価の変動を調査。株価の変動を指標に4分類
20%以上上昇、0-20%上昇、0-20%下降、20%以上下降

実験結果

完全正答を1		AIによる予想(%)			
上記+隣接0.5		over 20% up	0-20% up	0-20% down	over 20% down
up or down		72.0%			
実際 (%)	over 20% up	39.3	53.6	0.0	7.1
	0-20% up	17.9	64.3	3.6	14.3
	0-20% down	10.7	42.9	7.1	39.3
	over 20% down	7.1	25.0	14.3	53.6

- 上記表は2015年のデータを学習・評価した結果である。隠れ層の階層数6, 荷重減衰のλ値=0.0004で最も良い結果となり、4分類で42%、2分類で72%となった。
- 20%以上上がると予想して実際には下がった企業(全5社)を調査したところ、2社はおそらく仕手株、1社は敵対的買収への対抗、1社はゲーム会社でメジャータイトルの発表により、それぞれ2012~2013年ごろに株価が急騰していた。本モデルでは業績外の理由で株価が急騰した企業は正しく予測できないことが判明した。

- 日経平均の予想を±10%程度外した場合の影響を調べたが、正答率は上下した場合で約半数ずつであった。
- 2015年のデータを学習させて2016年のデータを評価した場合、ほぼ全ての企業を20%以上上がると予想した。これは景気の推移の差だと考えている。2015年は日経平均が下げ傾向で、2016年は年末に急騰した。この傾向の差が結果に大きく影響を及ぼしたと推測している。

今後の取り組み

- 仕手株などの特徴的な株について、例えば株価の時系列データをRNNに入力して判別させるモデルを検討し、予測できるようになるかを検証する。
- 2015年のデータだけでなく複数年のデータを学習することで2016年の結果が良くなるかを検証する。
- 日経平均株価の予測については本実験では実際の値を用いたが、この予測モデルを検討して組み込む。
- 今回使ったデータ以外の株価に影響を及ぼす情報を使うことで、実際に精度がよくなるかを検証する。