

# 時系列データを用いた機械学習モデルの構築

高田 梢

## プロセス産業におけるIoT問題点

プロセス産業の現場におけるIoT 推進は組立産業と比較してIoT への取組が遅れている。一方で、IoT推進により解決が期待される課題が多岐にわたって存在する。下記にいくつか列挙する。

- ・インラインでの測定が不可能なパラメータが存在する
- ・目視による品質検査が難しく、製造過程若しくは製品のサンプリングによる検査となる。
- ・設備故障による製造停止の損失が莫大である

## 手法・ツールの適用による解決

ソフトセンサー技術の活用、ビッグデータ解析、機械学習、深層学習の活用により、現場における課題解決を実現し、品質向上、作業標準化、安心安全の確保、設備故障による製造停止の削減（コスト削減）などが期待できる。今後、プロセス産業の現場において機械学習、深層学習を適用していくために、機械学習によるモデル構築の手法を獲得する。

## 実験

本修了制作では、プロセス産業の現場において機械学習、深層学習を適用していくことを想定し、機械学習によるモデル構築の手法を体得することを目的とする。尚、装置のセンサーデータの取得が難しいため、オープンデータの時系列データ（ビットコインの日足の始値、安値、高値、終値、出来高）を用いてモデルの構築を行う。

【実験内容】 当日の始値、安値、高値、終値、出来高より、翌日の終値の上げ下げを予測する

【入力情報】 2016年1月1日～2018年9月30日  
国内平均価格(OHLC)及び出来高 約1,000件

【適用予測モデル】 サポートベクターマシン、ランダムフォレスト、ロジスティック回帰

【出力情報】 0：翌日の終値が当日終値より下がる  
1：翌日の終値が当日終値より上がる

特徴変数（ロジスティック回帰）

目的変数

特徴変数（SVM, ランダムフォレスト）

日付	特徴変数（ロジスティック回帰）					予測	特徴変数（SVM, ランダムフォレスト）			
	始値	高値	安値	終値	出来高		出来高 上げ下げ	前日 上げ下げ	終値 > 前日高値	終値 > 前日安値
2018/9/25	742,318	748,647	719,684	725,328	42,056.07	1	1	0	0	1
2018/9/26	725,328	738,085	716,857	736,371	39,163.4	1	0	1	0	0

正規化

## 実験結果

テストデータを用いたビットコイン価格の翌日の上げ下げの予測精度は下記の通りとなった。

	テストデータ			
	20%	25%	30%	35%
サポートベクターマシン	<b>0.62</b>	0.61	0.59	0.61
ランダムフォレスト	0.58	0.61	0.59	0.61
ロジスティック回帰	0.59	0.58	0.56	0.56

- サポートベクターマシンのテストデータ20%の場合が最も精度が高かった。（62%）
- ロジスティック回帰は他の二つのモデルと比較して精度が低かった。
- アルゴリズム、特徴変数、テストデータ等のパラメータを調整することで、60%以上の精度を得ることができた。

## 現場での活用に向けて

下記の図は、データ解析のプロセスを示したものである。本実験では、「分析」「評価」に対して主に取り組んだ。実務においては、データ準備の「加工」の部分に大半の時間を費やすことが想定されるため、今後の勉強ではデータの前処理を含めた実験に取り組み、体得する。



## タイトル

所属

名前

メールアドレス(任意)

## 開発における問題点

ここでは、今回の修了制作で解決した開発における課題・問題点について述べる。例えば「XXシステム開発の際に、セキュリティ上の攻撃とそれに対する対策を系統的、網羅的に分析する必要がある」といったもの。「YYアプリに機能Aがなかった」等アプリケーション自体の問題点ではないことに注意する。

## 手法・ツールの適用による解決

ここでは、左で挙げた問題について、どのような手法・ツールの適用や提案によって解決したかを述べる。手法・ツールの名前を出すだけでなく、性質を明示し問題との対応がわかるようにする。例えば「達成目標を分析してシステムの構成要素を導出する系統的な方法を定めた要求分析手法KAOSを用いて・・・」といった感じ。

## ポスターの構成

上の概要のタイトルも必要であれば変更してください。人によっては「適用」ではなく「拡張」「提案」「連携」等となると思います。ある程度であればマスタの方の大きさを変えていただいても構いません。

概要より下の部分の構成(スペースの分割等)についてはお任せします。フタは残して下さい。

このフォーマットはA3になっていますが、実際にはポスターA0印刷、配布用A4印刷を行います。文字の大きさは最低13pt程度としてください。

## 注意点

Webにてアクセス制限なく公開するものであることに留意して下さい。

審査会での発表とは異なり、外部の方々・TopSEのツール・手法を知らない方々も対象となるため、

- ・モデリング方針
- ・ツールの設定

等の詳細よりも、

- ・その手法・ツールは何ができるか
- ・端的には、その図は何を表しており、その図を使って何をしたのか

といったことを概念的に説明することとなります。

## 提出

LMSから提出して下さい。

最終的には3月の修了式等のイベントにて、(上司の方々や外部の来賓にも)掲示します。

- ・こちらで印刷時の様子を見て多少レイアウト等調整を行う可能性があります。
- ・印刷はNIIで(事務局が)行います。