

(株)数理システム ユーザコンファレンス2008

VMStudioで株価予測システムの構築
5年で+1200%の実績紹介

- ・ 2004年 1月 1日 構築開始(株価データ収集開始)
- ・ 2004年 2月23日 実践運用(テスト)開始
- ・ 2004年 5月 1日 本運用開始: 初期資金300万円
- ・ 2006年 6月30日 全日本株式投資選手権 選抜レース完全優勝
- ・ 2008年 7月29日 運用資産+1250%達成(過去の最高点)
- ・ 2008年11月14日現在 +1XXX %



石川 慎也

ごあいさつ

石川慎也
「データマイニングの宝箱」制作者
<http://www5.ocn.ne.jp/~shinya91>

所属: 東京電機大学 研究員
シンボリックデータ(量質混在データ)・データ解析の基礎研究

協産学共同システム研究所
データマイニングコンサルティング・他

解析手法開発
趣味のマイニング

小規模分析
データ解析
コンサルティング

産学連携ビジネス
ベンチャー支援

はじめに:
今回構築した株価予測モデルについて

- 売買スタイルは?
 - ➡ 数億円以下の資金規模。(個人向け)
- ハイリスクハイリターン? たまたま儲かった?
 - ➡ 少なくともハイリスクな戦略はとっていない。
過去54ヶ月のうち80%以上で市況平均より良い成績であり、半年以上利益の最高点を更新できなかったことはない。
- このシステムを使えば誰でも同じように利益が出せるか?
 - ➡ 私が売買するより、利益は出せるとは思います。
算出結果通りに売買できずに、多くを取り逃がしています。
- 誰でも利益の出せるマイニングモデルが作れるか?
 - ➡
- 株価予測ができたのか?
 - ➡

本日の内容

システム構築の過程

- 目標設定、データ収集
- モデル方針の検討
- マイニングモデル構築
- 検証・評価

実績紹介、VMStudioデモ

- 実績紹介
- VMStudioデモ
- 考え方

株価予測に挑戦!!

テーマ (ビジネス上の課題など)

株で儲けたい。

マイニングの方針決め

目的

← →

獲得可能なデータ

分析対象となる分野の専門知識が「ある程度」ないと、明確な目的を出すこともできない。
過去の方法を検討するなりで、抽象的な目的を、より明確な目的にしたい

- 超短期予測(動きの観察とスピード判断が主体)
- 短期 予測(テクニカル分析が主体)
- 中期予測(ファンダメンタルズ分析が主体)

買板・売板状況
ニュース速報
掲示板の書込頻度

オシレーター
トレンドフォロー

ファンダメンタルズ
回帰直線の当てはめ
割安株の発見
テキストマイニング

株価予測に挑戦!!

情報収集/既存手法検証

「世の中には、株で大儲けするノウハウが溢れているぞ!」
.....?????

2004年1月1日

- ①テーマ決め
- ②目的- 方向性決め
- ③情報収集 既存手法検証

★超短期(秒・分単位)予測を目的として考えてみる
→ 経験則に基づく。秒刻みのスピード判断が必要
そもそも、秒(分)刻みの過去データを集められない。
データがない。→ つくれぬ

★中長期(数ヶ月・数年単位)予測を目的として考えてみる
集められるデータ
「会社の経営状態データ」「アナリストのレポート(格付)」「チャート」

「現在の経営状況」と「未来の株価」を分析 → 因果関係はない。
「アナリスト予測」と「未来の株価」を分析検証 → 因果関係はない。
「チャートに回帰式を当てはめる(データ解析者)」 → あてはまらない。
因果関係を見つけたぞ! 「来年の会社の経営状態」と「来年の株価」だ

→ 未来の経営状態を予測できるのであれば株価を予測することは可能? これから起こり得ることにつながるデータが集められるか?????
仮説による過程が多い。→ つくっても信頼できない

株価予測に挑戦!!

2004年1月1日

- ①テーマ決め
- ②目的一方向性決め
- ③情報収集
既存手法検証

情報収集/既存手法検証

★短期(日・週単位)予測を目的として考えてみる
データ「チャート(日足)」を主体

Aさんの値上がりルール

でも、これは経験則?

データマイニングに限らず、視覚化作業はとも重要。「目的を決める」「解析の方向性を決める」「意思決定をする」等、視覚化なしに良い判断はできない。

株価予測に挑戦!!

2004年1月1日

- ①テーマ決め
- ②目的一方向性決め
- ③情報収集
既存手法検証

目的の明確化

「株」データマイニングで儲けたい

短期予測で上昇株を発見する

予測する形態をどうするか?

どのようにして、予測、知識へと結びつけるか?

違い (例: 上昇銘柄と下落銘柄の違いから)

関係 (例: 株価とテクニカル指標や条件の関係から)

予測 (例: 株価の予測)

知識 (例: 類似の動きをする銘柄グループに分類してから)

理解 (例: 多種多様の分析指標を、要約した理解しやすい指標にしてから)

分解 (例: 類似の動きをする銘柄グループに分類してから)

2004年1月15日

- ④目的の明確化
・解析タイプ

株価予測に挑戦!!

2004年1月1日

- ①テーマ決め
- ②目的一方向性決め
- ③情報収集
既存手法検証

目的の明確化 (VMStudioによる予備解析)

2004年1月15日

- ④目的の明確化
・解析タイプ

株価予測に挑戦!!

2004年1月1日

- ①テーマ決め
- ②目的一方向性決め
- ③情報収集
既存手法検証

目的の明確化 (具体的に目標を設定する)

「株」データマイニングで儲けたい

短期予測で上昇株を発見する

明確な目的が決定しました!(テーマが決まって1ヶ月)

2004年1月15日

- ④目的の明確化
・解析タイプ
・詳細設定

株価予測に挑戦!!

2004年1月1日

- ①テーマ決め
- ②目的一方向性決め
- ③情報収集
既存手法検証

マイニングモデル作成

VMStudioのモデル

2004年1月15日

- ④目的の明確化
・解析タイプ
・詳細設定
- ⑤モデル作成

①学習データ入力部 ②検証データ入力部 ③予測データ入力部
④学習データクリーニング部 ⑤第一段階予測モデル
⑥第一段階の5つの予測モデルを取りまとめるモデル
⑦取りまとめた結果、評価の高い上位数%を選定し、さらに少ないサンプルで再解析。
⑧適合率の高い銘柄を抽出する

株価予測に挑戦!!

2004年1月1日

- ①テーマ決め
- ②目的一方向性決め
- ③情報収集
既存手法検証

マイニングモデル作成

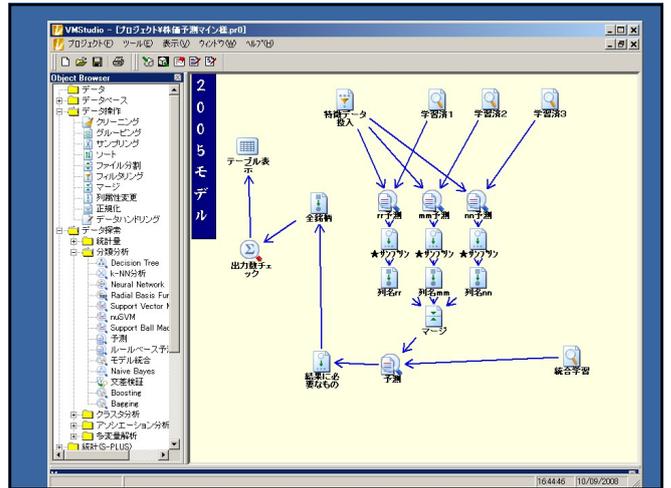
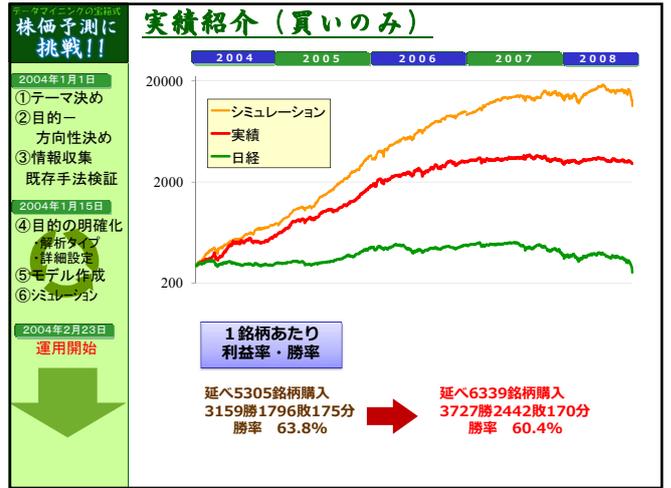
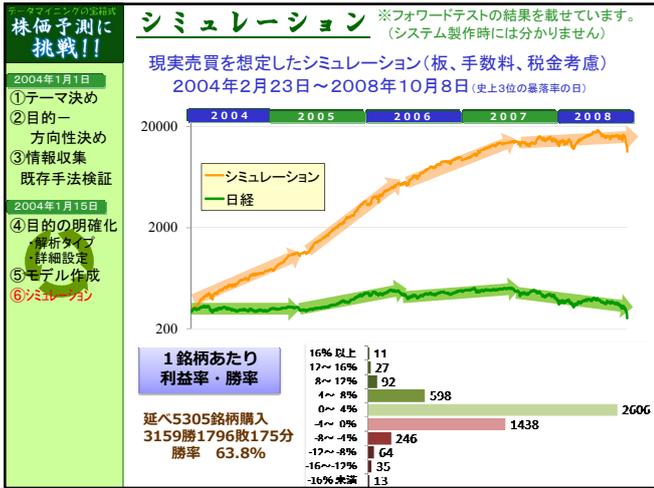
モデルの精度

予備調査

作成したマイニングモデル

2004年1月15日

- ④目的の明確化
・解析タイプ
・詳細設定
- ⑤モデル作成



まとめ

株価予測のシステム作りについて

- 個別銘柄の日次データ(四本値+出来高)は、入手が容易で、比較的整理されており、大量(2000銘柄×日数)であり一貫性がある。
- 結果の汎用性が高い
- わずかな精度向上を実現できれば、複利効果により、大きな影響がでる

データマイニングを楽しんでください。
投資は自己責任で