

Forecasting Algorithm Research of Stock Price Trend Based on CART and Similar Stock

Jianzhong Liu¹, Qiwei Yin²

¹Tianjin Municipality State-Owned Assets Supervision and Administration Commission, Tianjin

²School of Computer and Control Engineering, Nankai University, Tianjin

Email: 1422624555@qq.com

Received: Jun. 6th, 2017; accepted: Jun. 25th, 2017; published: Jun. 28th, 2017

Abstract

Most of the current forecast stock prices are based on a single stock historical price data, and try to find out the law of its stock price changes, training a model which can be used for price forecasting. But the fluctuation of the stock price will be affected by many social real-time factors and investors' behavior, so the stock price forecast model based on historical data often fails. For this reason, in this paper, we study stocks with similar historical changes in 2100 stocks, and obtain training data and forecast data for forecasting model based on time series window sliding, then use the CART (Classification and Regression Trees) to determine the predicted data. And compared with the result of ARMA (Auto Regressive Moving Average Model), the validity of the method is verified.

Keywords

Stock Data, Slide Window, Similar Stock, CART, ARMA

基于CART和相似股的股票价格走势预测算法研究

刘建中¹, 殷其威²

¹天津市国资委, 天津

²南开大学计算机与控制工程学院, 天津

Email: 1422624555@qq.com

收稿日期: 2017年6月6日; 录用日期: 2017年6月25日; 发布日期: 2017年6月28日

摘要

目前预测股票价格大多都基于单支股票的历史价格数据, 并试图找出其股价变化规律, 训练出可预测价格的模型。但实际股票价格的波动会受众多社会实时因素和投资者行为的影响, 因此基于历史数据的股票价格预测模型往往失效。为此, 本文研究2100支股票中具有相似历史价格变化的股票, 并基于时间序列窗口滑动获得用于预测模型的训练的数据和预测数据, 使用决策树CART(Classification and Regression Trees)对预测结果数据做出判别。并与经典的时间序列分析模型ARMA(Auto regressive Moving Average Model)对股票价格走势预测的结果进行分析比较, 实验结果验证了本文提出的方法预测结果的有效性。

关键词

股票数据, 滑动窗口, 相似股票, CART, ARMA

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

现在常用的时间序列分析法主要是建立自回归模型(AR)、移动平均模型(MA)、自回归滑动平均模型(ARMA)和自回归移动平均模型(ARIMA), 其中 ARIMA 模型是较成熟的模型, 常被用来对股价(最高价、最低价、开盘价、收盘价)及综合指数进行预测[1] [2]。文献[3] [4]给出了时间序列一些常用方法。文献[5] [6]采用 SVM 对金融数据进行预测, 文献[7]使用神经网络预测, 这些预测方法都是基于单一股票信息的。投资人基于社会形势(如利息率、公司盈利、贸易收支、相关政策、流行趋势等), 利用一些理论或计算方法对股价进行评估, 因此股票价格预测是一个多信息综合的结果。不同股票在同一段时间内受到的社会因素的影响是一致的, 但这些社会因素很难被人总结出来, 因此很难对股价变化做出适当的分析。实验证明存在这样一种现象: 单独的看股票 A, 它的价格在任意浮动, 股票 B 也同样如此, 但实验发现有 7% 的股票的价格变化能准确预测一周后另外 7% 的股票的价格变化, 即存在一种可能, 用前一周股票 A 的股价走势能很好的预测下周股票 B 的走势。这种现象很难被人所总结出来, 且其规律不一定固定的公式。但是可以利用机器学习的方法实时的学习这种规律, 即可以从众多股票中学习到这种社会因素对股票价格的影响, 以达到预测的目的。本文研究的方法试图找出这样的 7% 股票 A, 用其一段时间的历史数据变化情况来预测另外 7% 的股票 B。

2. 股票数据研究

股票价格受社会因素和投资者行为影响。社会因素属于随机变量, 随着时间段变化, 而投资者行为具有稳定的偏好, 在一定程度上是可预见的。因此一段时间内的股票价格走势具有相似性的特点, 这种相似性对属于同一行业分类、概念分类的股票尤为明显。这并不意味着不同行业的股票之间没有相关性, 事实上不同行业间的股票价格变化有时也具有相关性, 这种相关性的产生很难用常识或经济知识来解释。

2.1. 相似股票度量方法

股票价格走势是随时间序列变化的, 两个序列是否相似, 主要看它们的变化趋势是否一致。目前时

间序列相似度量, 最常用的是基于样本间“距离”(Distance)的计算, 其中大多是欧式距离(ED)及其变体形式。他们都具有“闵可夫斯基距离”(Minkowski Distance)的定义形式。

股票价格时间序列的欧式距离计算, 是以股票价格 x_i 和 y_i 的差值作为基础, 对于股票价格数据而言, 即使是经过归一化处理的数据, 走势相似的股票之间的距离也可能很大。如图 1 所示, 可以看出 X 和 Y 的走势是完全相同的, 但使用欧式距离计算出来的结果会很大。

对股票趋势进行预测的研究, 综上所述显然用样本间“距离”的计算方式是不合适的。本文基于皮尔森积矩相关系数(Pearson product-moment correlation coefficient)计算股票之间的相似度[8] [9]。以下以股票价格为例, 给出计算的具体描述。首先计算股票价格的变化率作为基础计算数据, 设某只股票的价格序列 $W = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n\}$, 序列 W 的变化率计算表达式 $C = \left\{ p_i \mid p_i = \frac{\omega_{i+1} - \omega_i}{\omega_i} \right\}$, 其中 $\omega_i \in W$, 最终的变化率序列 $W_p = \left\{ \frac{\omega_2 - \omega_1}{\omega_1}, \frac{\omega_3 - \omega_2}{\omega_2}, \frac{\omega_4 - \omega_3}{\omega_3}, \dots, \frac{\omega_n - \omega_{n-1}}{\omega_{n-1}} \right\}$ 。

设任意股票 X 价格变化率时间序列 $X = \{x_1, x_2, x_3, \dots, x_n\}$ 和股票 Y 价格变化率时间序列 $Y = \{y_1, y_2, y_3, \dots, y_n\}$, 序列 X 的标准差 $\sigma_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)^2}$ 其中 μ_x 为时间序列 X 的均值, 同理可计算出 Y 的标准差 σ_Y 。 $\text{cov}(X, Y) = \frac{\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}{n-1}$ 表示时间序列 X 与 Y 的协方差, 则股票之间相似度的计算公式如下:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

时间序列 X 和 Y 的距离 $D(X, Y) = 1 - \rho_{X,Y}$, 其中 $\rho_{X,Y} \in (-1, 1)$, $\rho_{X,Y} > 0$ 表示正相关, $\rho_{X,Y} < 0$ 表示负相关。

取股票代码为 600520 的 2014-03-27 到 2016-11-07 的价格数据, 利用以上方法计算出与其相关度 $\rho > 0.62$ 的股票一共 13 个, 绘制出它们的价格变化如图 2 所示,

上图中横坐标表示时间, 纵坐标表示股票价格, 从图可以看出不同种类的股票即使价格相差很大, 股票的波动情况却极为相似。实验证明使用基于统计的相关系数作为股票时间序列相似度量是有效的。同时透过上图价格走势相似的多支股票, 本文认为各种社会因素对它们的影响具有一致性。

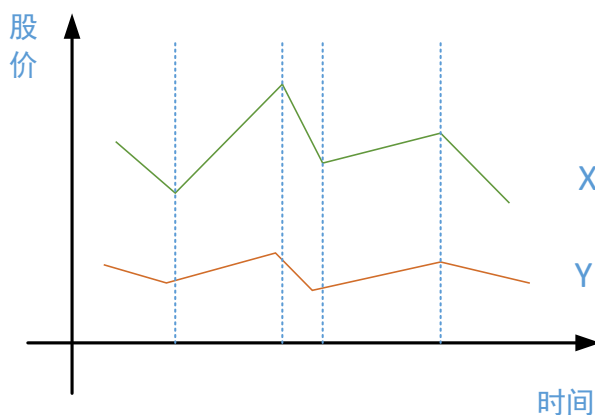


Figure 1. Distance calculation example
图 1. 距离计算示例

2.2. 基于时间序列滑动窗口获取数据

具有历史相似性的股票，在同一时间段的波动并不严格的相同，甚至在变化上有时会产生一种错位的现象，一些股票的价格变化会优先于另外一些股票。以下以股票闭市价格为例介绍如何用时间序列滑动得到错位现象的描述数据。

取任意股票(假设为 k)时间段长为 ε 的价格数据，滑动时间片长度 θ ，其中 ε 和 θ 的取值根据前期计算出的均值确认，一般取 4~10 天(使用 2016 年 9 月的 2100 支股票的历史价格信息，可计算出这种错位现象的时间均值为 4.21 天，方差为 2.6，取不同时间段值，结果可能不同)。然后与其具有历史相似度的股票(假设为 i)做滑动的相似度计算，相似度计算同样使用上文的皮尔森积矩相关系数。如图 3 所示，股票 k 和股票 i 具有相同的时间起始点，长度为 ε 的股票 k 的时间序列，在股票 i 的时间序列上向前滑动长度为 θ 。

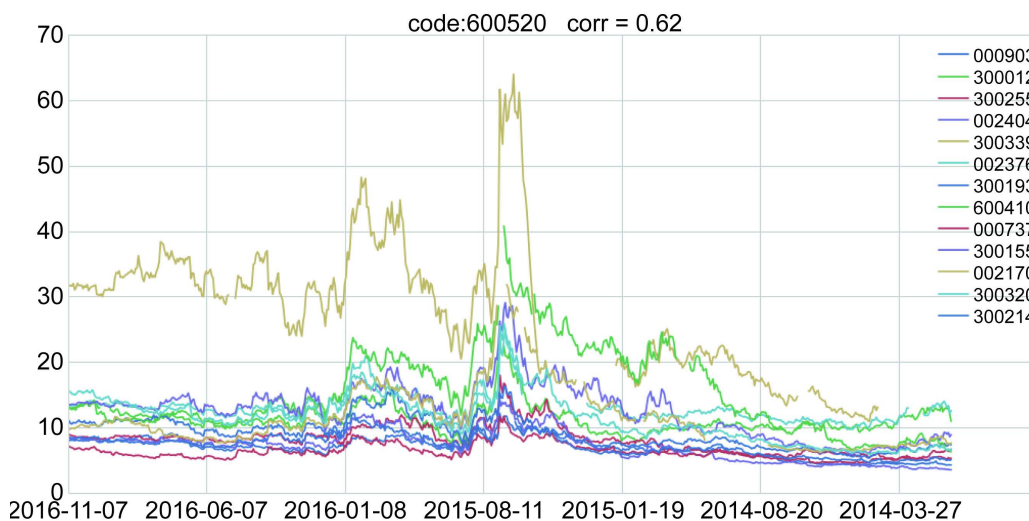


Figure 2. Historical price similar to stock price volatility
图 2. 历史价格相似股票价格波动图

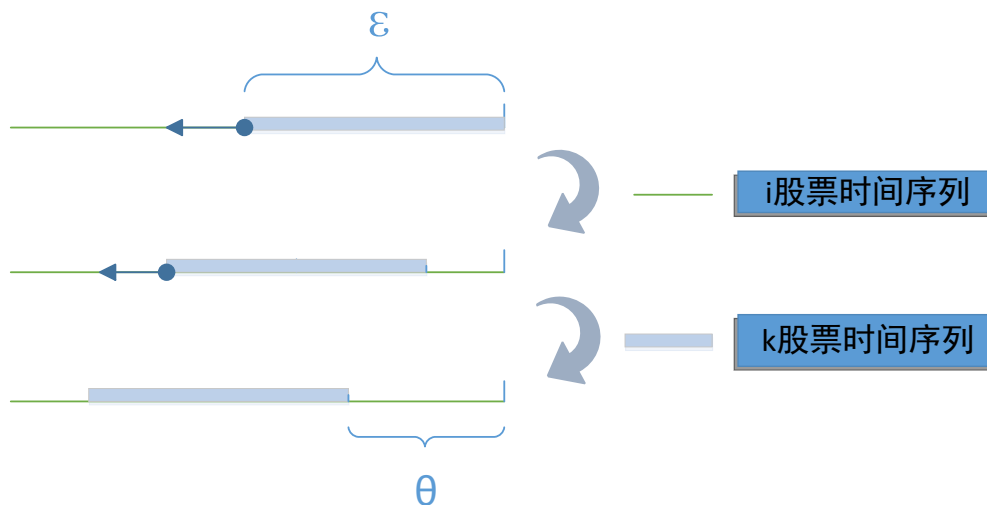


Figure 3. Slide calculation example diagram
图 3. 滑动计算示例图

首先计算出股票 k 具有历史相似度(通过相似度阈值 ρ 确认)的股票集合 G_k 。对于任意股票 $i \in G_k$ ，可以计算出时间片长度为 ε 的滑动相似度集合 $M_{k,n,i} (0 < n \leq \theta, n \in N)$ ，其中 n 表示滑动长度。若 $\max(M_{k,n,i}) > \mu$ ，其中 μ 为相似度阈值(一般取 0.5~0.7 之间)，则表示集合 $M_{k,n,i}$ 对应的股票 i 的数据可以用于股票 k 的预测分析。得到最终的描述数据集合：

$$S_k = \left\{ \max(M_{k,n,i}) \mid \max(M_{k,n,i}) > \mu, 0 < \mu < 1, i \in G_k \right\}$$

通过上面的数据计算可以得到如表 1 的数据字段。

同理，使用同样的方法可以得到价格日 K 线、3 日 K 线、周 K 线等的描述数据。

在股票交易中由成交量形成的“成交量柱”，简称“量柱”，能反应股民的交易动向和意图，因为交易量是真实货币堆砌起来的。因此股票“量柱”的变化和价格有非常强的关系。股票“量柱”的描述数据，可用上文中的数据集合 S_k (即滑动天数取 S_k 中的 n 天)结合交易量数据计算得到。数据描述见表 2。

利用以上数据获取方法，按以下步骤做简单的预测实验：

- 1) 假设预测日期 Date 为 2016-09-27，取 Date 之前的数据用于数据计算，Date 之后的数据用于结果验证。
- 2) 取窗口长度 $\varepsilon = 10$ ，滑动最大长度 $\theta = 10$ ，历史天数 300 天，历史相似度阈值 $\rho = 0.6$ ，滑动窗口相似度阈值 $\mu = 0.65$ 。
- 3) 选取随机 500 支持预测的股票，利用滑动窗口计算出表 1，表 2 的描述数据。
- 4) 获取待预测股票日期 Date 后的真实价格变化数据。
- 5) 利用描述数据集合 S_k 中的滑动天数数据，获取预测数据。
- 6) 比较真实数据与描述数据的走势情况，可以得到预测的准确率。

Table 1. Price sliding window data description table

表 1. 价格滑动窗口数据描述表

数据字段	描述
n	滑动天数
S_k	滑动相似度值
t_i	股票 i 的时间窗口为 ε 的价格标准差
t_k	股票 k 的时间窗口为 ε 的价格标准差
M_{ki}	股票 i 和股票 k 的历史相似度
m_k	股票 k 价格均值
m_i	股票 i 价格均值

Table 2. Transaction volume sliding window data description table

表 2. 交易量滑动窗口数据描述表

数据字段	描述
V_k	滑动相似度值
vt_i	股票 i 的时间窗口为 ε 的成交量标准差
vt_k	股票 k 的时间窗口为 ε 的成交量标准差
vm_k	股票 k 成交量均值
vm_i	股票 i 成交量均值

下面分别选取描述数据中的历史相似度 M_{ki} 和“柱量”的相似度 V_k ，分析其与预测准确率的变化情况。

图 4 中横坐标表是历史相似度，纵坐标表示走势预测的准确率，从图 4 中可以看到历史价格相似度与预测的准确率有一定的正相关性。历史价格相似度越大，预测的准确率也就越大。

图 5 中横坐标表是“柱量”相似度，纵坐标表示走势预测的准确率。从图 5 中可以看出，“柱量”相似度越大，一定范围内的预测的准确率也有变大的趋势。

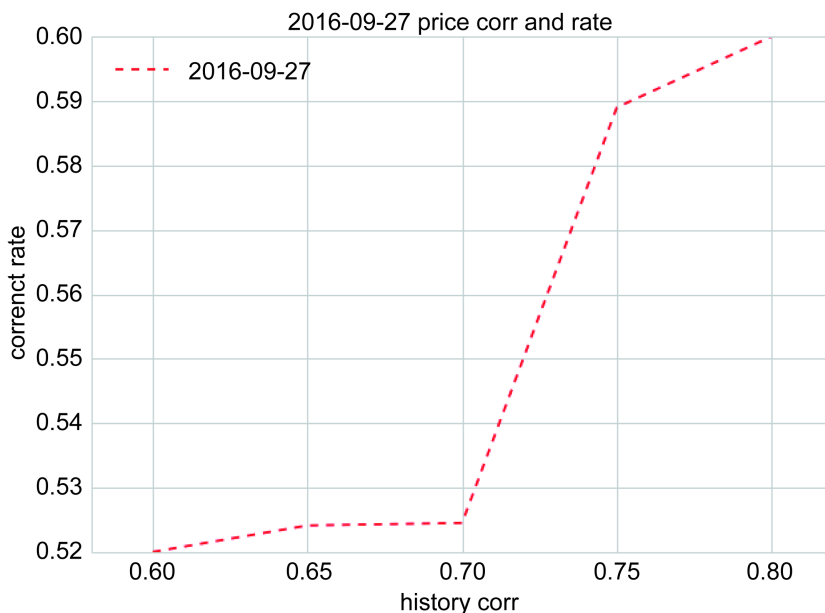


Figure 4. Historical price similarity and forecast accuracy
图 4. 历史价格相似度与预测准确率

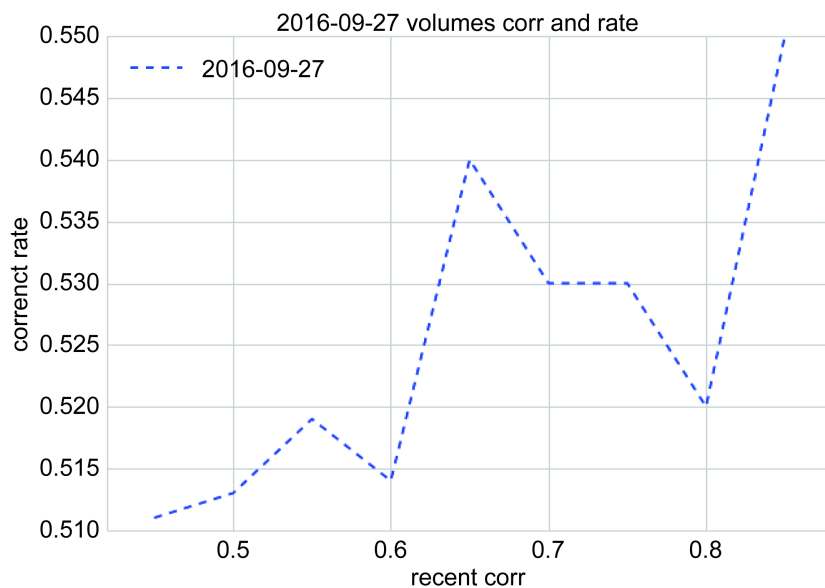


Figure 5. The similarity of columns and the prediction accuracy
图 5. “柱量”相似度与预测准确率

实验利用本文提出的滑动窗口计算方法得到的预测数据的准确性与“柱量”相似度、历史相似度等有一定的相关性，为进一步验证这一结论的有效性，下面用 CART 模型来试图学习出这种相关性。

3. 基于 CART 的预测模型

通过上文的预测实验可知描述数据与预测的准确率之间是存在一定相关性的，为验证这种相关性，本文使用决策树(Decision Trees)算法中的 CART (Classification and Regression Trees)算法，使用上文的描述数据拟合出一个预测模型。CART 与 C4.5 非常相似，区别在于 CART 支持数值变量且不产生显式的决策规则集合[10]。

模型选用决策树是因为其易于理解和解释，只需少量的数据准备工作，其它许多机器学习模型需要数据归一化、虚拟变量创建以及移除空白值(blank values)等预处理操作，金融股票数据中各数据的单位、大小、空值等非常多样，若是数据的处理工作对结果的影响比较大。决策树能够同时处理数值数据(numerical data)和类别数据(categorical data)，金融数据中类别数据如股票的行业分类、概念分类等许多都是非数值型的数据。但是最优决策树的问题已经被证明是一个 NP-完全问题，即使是某些比较简单的情形下获得最优解也非常困难，实际应用的决策树学习算法一般都是基于启发式优化方法，本文将不涉及对决策树的最优化处理，仅使用决策模型说明预测问题。

模型的训练数据使用上文经过滑动窗口计算出的历史描述数据和股票的基本描述数据，其中股票基本描述数据如表 3，股票基本数据可以从各金融网上获取。

使用上一节中滑动窗口计算出历史的描述数据(历史数据时间与待预测的时间不能相隔过长，范围在 1~10 天，因为预测的天数不超过 10 天)，经过整理后的训练数据中的部分输入数据(Input data)如图 6 所

Table 3. Stock basic data description table

表 3. 股票基本数据描述表

字段	描述
pe	市盈率
outstanding	流通股本(亿)
reservedPerShare	每股公积金
rev	收入同比(%)
gpr	毛利率(%)
holders	持股人数

days	corr_datacode	std_datacode	vol	corr_history	maxdays	pe	totals	gpr	mean	std_precode
1	0.738502	0.01322852	0.4007602	0.6385335	7	0.8265802	2.134868	1.360486	1.122479	0.01891153
8	0.7815569	0.02116021	0.6639105	0.6580805	7	0.6453183	0.8506998	0.710453	1.260501	0.01321692
6	0.7170734	0.01657294	0.6153888	0.6351805	5	0.2462853	1.974729	1.850272	0.3735474	0.01321692
4	0.8531280	0.01831972	0.5670393	0.6368387	6	2.753322	0.4919065	1.085729	1.404006	0.01321692
8	0.7308267	0.02274805	0.621064	0.693367	7	0.6169738	0.3907143	1.022568	1.616475	0.01321692
8	0.8540108	0.02584213	0.7546852	0.6436461	5	0.08529745	0.9177853	7.551852	0.7811691	0.01321692
8	0.7154445	0.01347902	0.4851831	0.6271043	5	0.04523624	0.6904348	1.020801	1.250156	0.01430198
8	0.8803784	0.01990922	0.8232677	0.6308288	5	0.5273903	0.8519313	1.913172	1.204098	0.01430198
8	0.7750559	0.01563777	0.4045564	0.6276775	6	0.1845642	0.318109	0.4523111	1.017864	0.01430198
8	0.8353259	0.0151352	0.5487254	0.7241061	6	0.3115968	0.7576336	0.5708495	0.948728	0.01430198
8	0.9175682	0.02253975	0.4857111	0.6375023	6	0.1727929	0.3401885	0.421635	0.5776873	0.01430198

Figure 6. Part of the input data of the model training

图 6. 模型训练的部分输入数据

示, 图中包含的描述数据字段有: {滑动天数, 相似度, 预测数据方差, 交易量相似度, 历史相似度, 最大相似天数, 最大滑动相似天数, 原始数据方差, 均值, 市盈率, 毛利率}。

模型训练数据的目标数据(Target data), 使用预测数据和真实数据比较得到。首先获取预测数据和真实数据的股票价格走势的形状, 设有股票价格时间序列 $X = \{x_1, x_2, x_3, \dots, x_n\}$, 计算得到序列的变化率序列 $X_p = \left\{ \frac{x_2 - x_1}{x_1}, \frac{x_3 - x_2}{x_2}, \frac{x_4 - x_3}{x_3}, \dots, \frac{x_n - x_{n-1}}{x_{n-1}} \right\}$, 设置变化率阈值 $\beta (0.1 > \beta > 0.05)$, 使用下面的公式可以将变化率序列 X_p 变为包含 -1, 0 和 1 的序列, 用来表示股票价格时间序列 X 的走势。

$$\varphi(x_{n-1}, x_n) = \begin{cases} 1, & \frac{x_n - x_{n-1}}{x_{n-1}} \geq \beta \\ 0, & -\beta < \frac{x_n - x_{n-1}}{x_{n-1}} < \beta \\ -1, & \frac{x_n - x_{n-1}}{x_{n-1}} \leq -\beta \end{cases}$$

下图 7 的示例图, 横坐标为时间, 纵坐标为股票价格, 折线图为股票走势图。计算下图时间序列可以得到走势序列 $\{-1, 0, -1, 1, 0, 1\}$ 。通过比较两个走势序列, 可以得知两个股票价格时间序列是否相等。

使用以上的训练数据建立的决策树模型, 设置决策树最大深度为 5, 节点分裂使用信息增益熵 (entropy), 模型拟合后使用 Graphviz 下的 dot 工具绘制出生成的决策树结果如图 8 所示。

从上图中可以看出决策树模型只对已出现数据的决策, 不能用来预测超出预测数据范围以外的值, 且建树的大小和数据本身数量相关性比较大, 很容易过拟合或欠拟合。本文使用决策树来判断上文滑动窗口计算出的预测值的准确性, 其中数据经过处理后都有一定的范围, 即用决策树来验证本文预测方法正确性可以满足需求。

4. 算法描述

本文提出的基于时间序列滑动窗口金融数据分析算法, 分为三个部分: 1、利用时间序列滑动窗口从历史数据获取模型训练数据; 2、训练数据输入决策树 CART 训练出预测模型; 3、再次使用滑动时间序列窗口获取最新的模型输入数据, 用已经训练好的模型判断数据预测的正确性。为了得到有效的模型, 数据集的范围取股票市场的有效股票。时间序列滑动窗口预测算法具体描述如图 9 所示。

输入: 股票数据时间序列集合 $Y = \{C_1, C_2, C_3, \dots, C_n\}$ 。

输出: 股票时间序列 $C_k (C_k \in Y)$ 预测值。

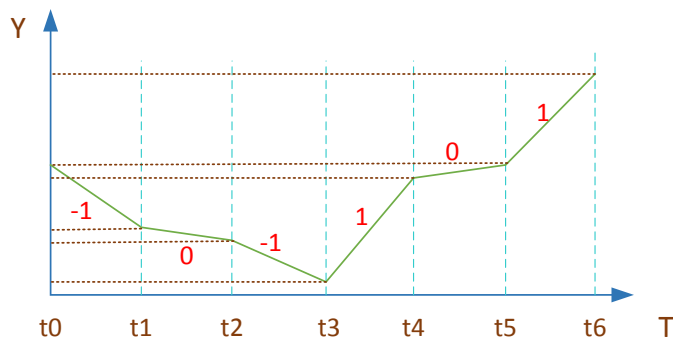


Figure 7. An example of a stock trend

图 7. 股票走势示例图

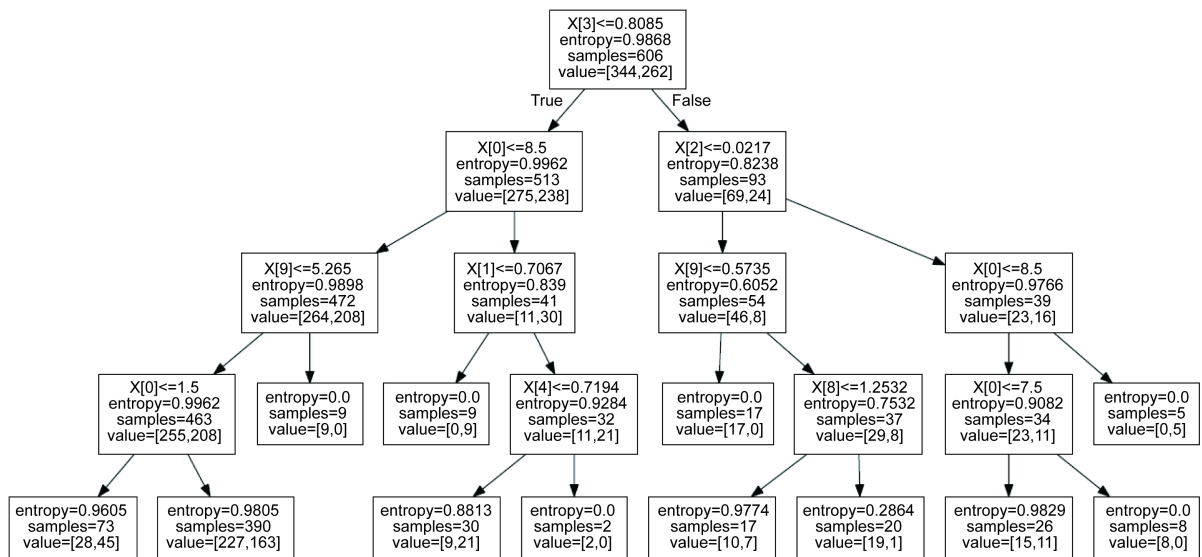


Figure 8. Decision tree model graphical result display

图 8. 决策树模型图形化结果显示

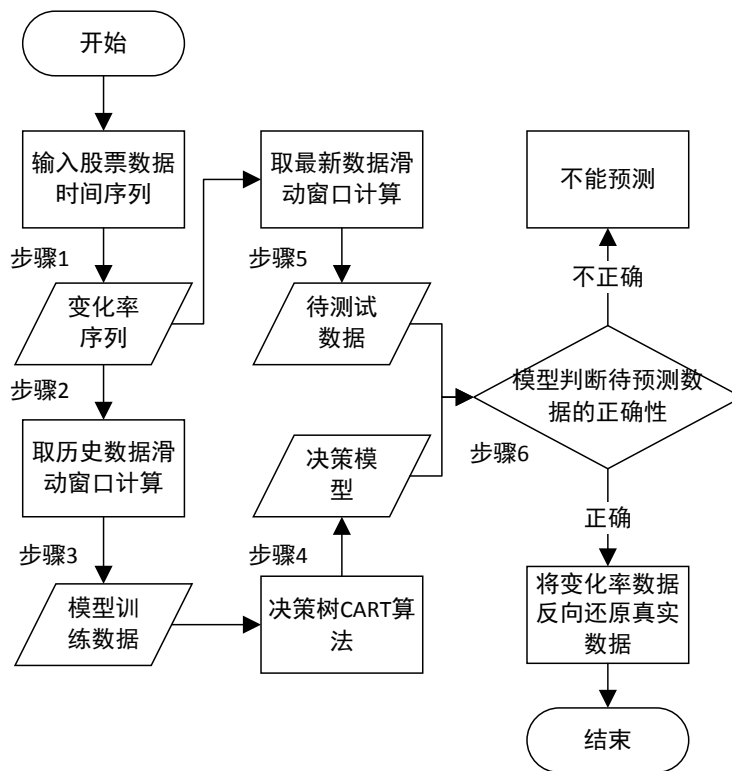


Figure 9. Sliding window prediction algorithm flow chart

图 9. 时间序列滑动窗口预测算法流程图

步骤 1: 将时间序列集合 $Y = \{C_1, C_2, C_3, \dots, C_n\}$ 中的序列转化成变化率序列 $Y_p = \{C_{p1}, C_{p2}, C_{p3}, \dots, C_{pn}\}$ 。

步骤 2: 取集合 Y_p 中序列的历史数据, 设 $C_{pk} = \{x_1, x_2, x_3, \dots, x_m\} (C_{pk} \in Y_p)$, 历史相隔天数 $\beta (1 \leq \beta \leq 10)$, 则股票历史数据序列 $C'_{pk} = \{x_1, x_2, x_3, \dots, x_{m-\beta}\}$ 。

步骤 3: 使用历史数据序列做滑动窗口计算, 得到描述数据集 S_k (上文 1.2 中), 获取股票的基本描述数据(市盈率、流通股本等), 组成模型训练数据中的输入数据。取 $C_{pk} - C'_{pk} = \{x_{m-\beta}, x_{m-\beta+1}, \dots, x_m\}$ 与描述数据中的预测数据比较, 得到模型训练的结果数据。

步骤 4: 将得到的训练数据输入决策树 CART 模型, 得到训练好的决策模型。

步骤 5: 使用整个 C_{pk} 序列进行滑动窗口计算, 同样可以得到一个描述数据集 S'_k , 获取股票的基本描述数据, 我们可以得到模型测试数据。

步骤 6: 将通过测试数据输入训练好的模型可以得到测试数据是否用于适合预测, 我们将适合预测的数据的变化率序列转化为预测值序列 $C_k = \{\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_l\}$, 可用 $R_{MSE} = \sqrt{\frac{1}{n} \sum_{t=1}^m (\gamma_t - x_t)}$ 检验预测的准确性。

5. 实验分析

5.1. 实验说明

为了确保算法的真实有效, 实验数据来源于真实交易数据(主要来源于凤凰财经 <http://finance.ifeng.com/>), 以天为单位获取股票的历史前复权数据, 包括开市价格、闭市价格、成交量。同时获取股票基本信息包括: 所属行业、地区、市盈率、流通股本(亿)、总股本(亿)、总资产(万)、流动资产、固定资产、公积金、每股公积金、每股收益、每股净资产、市净率、上市日期、未分利润、每股未分配、收入同比(%), 利润同比(%), 毛利率(%), 净利润率(%), 股东人数。选取 2015-01-04 到 2016-10-27 一年多的股票前复权数据。

算法的具体参数设置如表 4 所示。

5.2. 实验结果

实验结果与现在常用的时间序列分析模型 ARMA(Auto regressive Moving Average Model)(方法参考文献[11] [12] [13])对比, 模型 ARMA(p, q)中 p 和 q 的选取使用采用 BIC 准则精确得到, 模型拟合使用静态方式。图 10 和图 11 是股票代码为 600512 的股票在 2016-9-30, ARMA 模型(其中 BIC = -877.684926648, $p = 0, q = 1$)的历史价格拟合结果和预测结果。

从图 11 的拟合结果(绿线为真实值, 红线为预测值)可以看出使用 ARMA 模型并不能很好的拟合出原始数据曲线, 原因在于 ARMA 模型要求整个序列的变化确有一定的规律性, 即序列本身具有一定的稳定性, 序列历史数据对尚未产生的数据有一定的约束性。使用本文提出的方法预测出来的结果如图 12 所示, 结果虽有偏差但预测出来的走势是比较准确的。

Table 4. Experimental parameter setting table

表 4. 实验参数设置表

参数	描述	值
m	历史年数	1
θ	历史相似度阈值	0.62
φ	滑动窗口长度	6
ω	时间序列滑动长度	6
μ	滑动窗口相似度阈值	0.65



Figure 10. ARMA fitting effect diagram
图 10. ARMA 拟合效果图

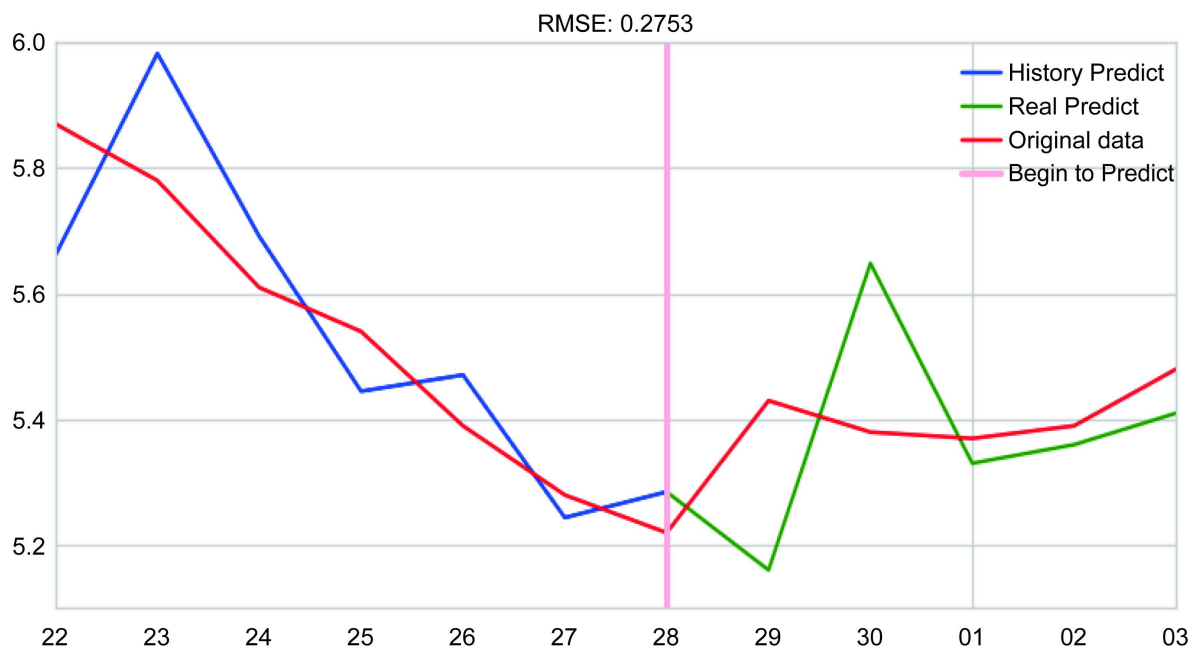


Figure 11. ARMA predictive effect
图 11. ARMA 预测效果图

测试使用 1500 支股票 2014-03-27 到 2016-09-26 的数据用本文滑动窗口计算出模型的训练数据，使用训练数据训练出的 CART 模型准确度仅为 0.62(训练数据中的 80%作为测试的训练数据，20%作为验证数据)。随机选取 500 支股票 2014-03-27 到 2016-09-30 的数据用滑动窗口计算得到预测的描述数据，计算出能预测的股票为 128 支(得到的数量与设置的相似度阈值有关)，然后使用拟合出来的 CART 模型判断预测准确的股票 64 支，计算真实值与预测值得走势形状准确度为 67.18%。

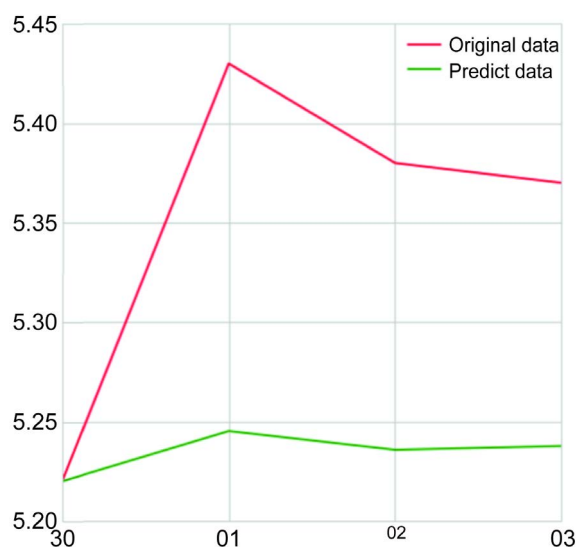


Figure 12. Sliding window prediction
图 12. 滑动窗口预测效果图

6. 结论

经过测试验证了本文提出的滑动窗口计算方法的有效性, 与 ARMA 模型相比预测的准确度更高。但本文选取的 CART 模型并非最适合的学习模型(还可选用 Stochastic Gradient Descent (SGD)、Logistic regression、Neural network 等), 且用于模型拟合的数据选取偏少或不够专业, 导致 CART 模型的拟合程度不高, 这些问题有待更加专业的领域知识的介入分析, 提高模型的拟合程度将极大的增加最终预测的准确度。

参考文献 (References)

- [1] 杜建卫, 王超峰. 小波分析方法在金融股票数据预测中的应用[J]. 数学的实践与认识, 2008, 38(7): 68-75.
- [2] 邓凯旭, 宋宝瑞. 小波变换在金融数据分析中的应用[J]. 数理统计与管理, 2006, 25(2): 215-219.
- [3] Das G., Lin, D., Mannila, H., Renganathan, G. and Smyth, P. (1998) Rule Discovery from Time Series. *Fourth Annual Conference on Knowledge Discovery and Data Mining (KDD)*, New York, 27-31 August, 23-29.
- [4] 李斌. 金融时间序列挖掘关键算法研究[D]: [博士学位论文]. 合肥: 中国科学技术大学, 2002.
- [5] 辛治运, 顾明. 基于最小二乘支持向量机的复杂金融时间序列预测[J]. 清华大学学报:自然科学版, 2008, 48(7): 1147-1149.
- [6] Cao, L. and Tay, F.E.H. (2001) Financial Forecasting Using Support Vector Machines. *Neural Computing and Applications*, **10**, 184-192. <https://doi.org/10.1007/s005210170010>
- [7] Kimoto, T., Asakawa, K., Yoda, M. and Takeoka, M. (1990) Stock Market Prediction System with Modular Neural Networks. *International Joint Conference on Neural*, **1**, 1-6. <https://doi.org/10.1109/ijcnn.1990.137535>
- [8] 易东云, 张维明. 金融数据挖掘中的非线性相关跟踪技术 Nonlinear Correlation Tracking Technique in Data Mining of Financial Markets [J]. 软件学报, 2000, 11(12): 1581-1586.
- [9] 汤胤. 时间序列相似性分析方法研究[J]. 计算机工程与应用, 2006, 42(1): 68-71.
- [10] Breiman, L.I., Friedman, J.H., Olshen, R.A. and Stone .C.J. (1984) Classification and Regression Trees (CART). *Biometrics*, **40**, 358.
- [11] 王丽娜, 肖冬荣. 基于 ARMA 模型的经济非平稳时间序列的预测分析[J]. 武汉理工大学学报交通科学与工程版, 2004, 28(1): 133-136.
- [12] 张海勇, 马孝江, 盖强. 非平稳信号的一种 ARMA 模型分析方法[J]. 电子与信息学报, 2002, 24(7): 992-996.
- [13] 吴朝阳. 改进的灰色模型与 ARMA 模型的股指预测[J]. 智能系统学报, 2010, 5(3): 277-281.

期刊投稿者将享受如下服务：

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：csa@hanspub.org