

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 1 月 2 日 (02.01.2020)



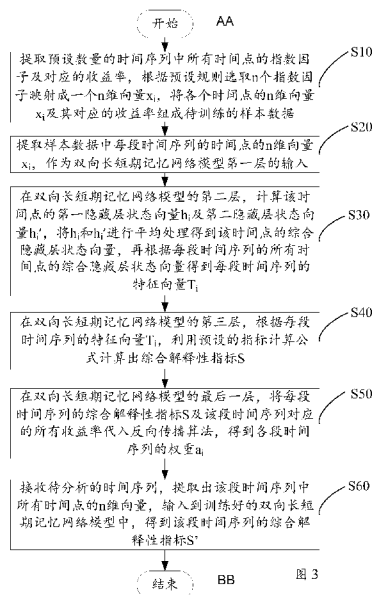
(10) 国际公布号
WO 2020/000715 A1

- (51) 国际专利分类号:
G06Q 10/04 (2012.01) **G06Q 40/04** (2012.01)
- (21) 国际申请号: PCT/CN2018/107484
- (22) 国际申请日: 2018 年 9 月 26 日 (26.09.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201810694893.2 2018年6月29日 (29.06.2018) CN
- (71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

- (72) 发明人: 李正洋(LI, Zhengyang); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。
- (74) 代理人: 深圳市沃德知识产权代理有限公司(普通合伙)(SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦四路 10 号中浩大厦 1528-1530 室于志光, Guangdong 518000 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) Title: INDEX FEATURE EXTRACTION-BASED STOCK INDEX PREDICTION METHOD, SERVER AND STORAGE MEDIUM

(54) 发明名称: 基于指数特征提取的股指预测方法、服务器及存储介质



- S10 Extract index factors of all time points in a preset number of time series and corresponding yields, select, according to a preset rule, a number n of index factors to map into an n-dimensional vector xi, and form the n-dimensional vector xi of each time point and the corresponding yield into sample data to be trained
- S20 Extract the n-dimensional vector xi of the time point of each time series in the sample data as an input of the first layer of a bidirectional long-term and short-term memory network model
- S30 Calculate, at the second layer of the bidirectional long-term and short-term memory network model, a first hidden layer state vector hi and a second hidden layer state vector hi' of the time point, average the hi and the hi' to obtain a comprehensive hidden layer state vector of the time point, and obtain a feature vector Ti of each time series according to the comprehensive hidden layer state vectors of all the time points of each time series
- S40 Calculate, at the third layer of the bidirectional long-term and short-term memory network model, a comprehensive explanatory index S by using a preset index calculation formula according to the feature vector Ti of each time series
- S50 Substitute, at the last layer of the bidirectional long-term and short-term memory network model, the comprehensive explanatory index S of each time series and all yields corresponding to the time series into a back-propagation algorithm to obtain the weight ai of each time series
- S60 Receive a time series to be analyzed, extract the n-dimensional vectors of all time points in the time series, input the n-dimensional vectors into the trained bidirectional long-term and short-term memory network model to obtain the comprehensive explanatory index S' of the time series
- AA Start
- BB End

(57) Abstract: The present application provides a index feature extraction-based stock index prediction method and device and a storage medium. The method comprise: extracting index factors of all time points in a preset number of time series and corresponding yields; selecting, according to a preset rule, a number n of index factors to form an n-dimensional vector; and respectively forming the n-dimensional vector of each time point and the corresponding yield into sample data to be trained. Then, by using the n-dimensional vector in the sample data and the corresponding yield, performing training on the bidirectional long-term and short-term memory network model to determine model parameters. Finally, receiving a time series to be analyzed, and extracting the n-dimensional vectors of all time points of the time series and inputting the n-dimensional vectors into the trained bidirectional long-term and short-term

MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

memory network model to obtain a comprehensive explanatory index of the time series. With the present application, the features of the index can be deeply extracted, and the accuracy of stock index prediction can be improved.

(57) 摘要: 本申请提供了一种基于指数特征提取的股指预测方法、装置及存储介质, 该方法提取预设数量的时间序列中所有时间点的指数因子及对应的收益率, 根据预设规则选取 n 个指数因子构成一个 n 维向量, 将各个时间点的 n 维向量及其对应的收益率分别组成一个待训练的样本数据。之后, 该方法利用样本数据中的 n 维向量及其对应的收益率对双向长短期记忆网络模型进行训练, 确定模型参数。最后, 该方法接收待分析的时间序列, 提取出该时间序列所有时间点的 n 维向量输入到训练好的双向长短期记忆网络模型中, 得到该段时间序列的综合解释性指标。利用本申请, 能够深层次的提取指数的特征, 提高股指预测的准确性。

基于指数特征提取的股指预测方法、服务器及存储介质

本申请要求于2018年06月29日提交中国专利局、申请号为201810694893.2, 名称为“基于指数特征提取的股指预测方法、服务器及存储介质”的中国专利申请的优先权, 该中国专利申请的整体内容以参考的方式结合本申请中。

技术领域

本申请涉及计算机技术领域, 尤其涉及一种基于指数特征提取的股指预测方法、服务器及计算机可读存储介质。

背景技术

因子一般是指对研究事物具有影响作用的特征或因素。指数相对于其他指标而言, 往往具有时序依赖的特征。目前的指数因子特征提取方法, 可以是提取指数位于某一时刻的特定维度的特征, 也可以是提取长时间段上指数的趋势性与周期性特征, 而对于相邻时序间的依赖关系特征较难提取, 从而不能够科学地衡量指数走势。

发明内容

鉴于以上内容, 本申请提供一种基于指数特征提取的股指预测方法、服务器及计算机可读存储介质, 其主要目的在于提取多重指数特征所包含的综合信息, 科学、准确地预测指数走势。

为实现上述目的, 本申请提供一种基于指数特征提取的股指预测方法, 该方法包括:

样本采集步骤: 提取预设数量的时间序列中所有时间点的指数因子及对应的收益率, 根据预设规则选取 n 个指数因子构成一个 n 维向量 x_i , $i>0$ 且 i 为整数, 将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据;

提取步骤: 提取样本数据中每段时间序列的所有时间点的 n 维向量 x_i , 作为双向长短期记忆网络模型第一层的输入;

处理步骤: 在双向长短期记忆网络模型的第三层, 根据某段时间序列的

某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i ，并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ，将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理，得到该时间点的综合隐藏层状态向量，直至算出所有时间点的综合隐藏层状态向量，再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i ；

计算步骤：在双向长短期记忆网络模型的第三层，根据每段时间序列的特征向量 T_i ，利用预设的指标计算公式计算下一段时间序列的综合解释性指标 S ；

权重确定步骤：在双向长短期记忆网络模型的最后一层，将每段时间序列的综合解释性指标 S 及该段时间序列对应的所有收益率代入反向传播算法，得到各段时间序列的权重 a_i ；

预测步骤：接收待分析的时间序列，提取该段时间序列中所有时间点的 n 维向量，输入到训练好的双向长短期记忆网络模型中，得到该时间序列的综合解释性指标 S 。

此外，本申请还提供一种服务器，该服务器包括：存储器、处理器及显示器，所述存储器上存储股指预测程序，所述股指预测程序被所述处理器执行，可实现如下步骤：

样本采集步骤：提取预设数量的时间序列中所有时间点的指数因子及对应的收益率，根据预设规则选取 n 个指数因子构成一个 n 维向量 x_i ， $i>0$ 且 i 为整数，将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据；

提取步骤：提取样本数据中每段时间序列的所有时间点的 n 维向量 x_i ，作为双向长短期记忆网络模型第一层的输入；

处理步骤：在双向长短期记忆网络模型的第二层，根据某段时间序列的某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i ，并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ，将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理，得到该时间点的综合隐藏层状态向量，直至算出所有时间点的

的综合隐藏层状态向量，再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i ；

计算步骤：在双向长短期记忆网络模型的第三层，根据每段时间序列的特征向量 T_i ，利用预设的指标计算公式计算下一段时间序列的综合解释性指标
5 S；

权重确定步骤：在双向长短期记忆网络模型的最后一层，将每段时间序列的综合解释性指标S及该段时间序列对应的所有收益率代入反向传播算法，得到各段时间序列的权重 a_i ；

预测步骤：接收待分析的时间序列，提取该段时间序列中所有时间点的n
10 维向量，输入到训练好的双向长短期记忆网络模型中，得到该时间序列的综合解释性指标S。

此外，为实现上述目的，本申请还提供一种计算机可读存储介质，所述计算机可读存储介质中包括股指预测程序，所述股指预测程序被处理器执行时，可实现如上所述基于指数特征提取的股指预测方法中的任意步骤。

15 本申请提出的基于指数特征提取的股指预测方法、服务器及计算机可读存储介质，通过提取每段时间序列所有时间点的指数因子及对应的收益率，并选取n个指数因子构成n维向量 x_i ，接着，将n维向量 x_i 及收益率代入双向长短期记忆网络模型中训练，得到完整的模型，最后，提取待分析的时间序列的n维向量，输入到模型中，得到该时间序列的综合解释性指标，从而减少梯
20 度爆炸或梯度消失，准确的预测股指的走势。

附图说明

图1为本申请服务器较佳实施例的示意图；

图2为图1中股指预测程序较佳实施例的模块示意图；

25 图3为本申请基于指数特征提取的股指预测方法较佳实施例的流程图；

图4为本申请双向长短期记忆网络模型的示意图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

如图 1 所示，是本申请服务器 1 较佳实施例的示意图。

5 在本实施例中，服务器 1 是指产品服务平台，该服务器 1 可以是服务器、平板电脑、个人电脑、便携计算机以及其它具有运算功能的电子设备。

该服务器 1 包括：存储器 11、处理器 12 及网络接口 13。其中，网络接口 13 可选地可以包括标准的有线接口、无线接口（如 WI-FI 接口）。

存储器 11 至少包括一种类型的可读存储介质。所述至少一种类型的可读
10 存储介质可为如闪存、硬盘、多媒体卡、卡型存储器等的非易失性存储介质。在一些实施例中，所述存储器 11 可以是所述服务器 1 的内部存储单元，例如该服务器 1 的硬盘。在另一些实施例中，所述存储器 11 也可以是所述服务器 1 的外部存储单元，例如所述服务器 1 上配备的插接式硬盘，智能存储卡(Smart Media Card, SMC)，安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）
15 等。在本实施例中，所述存储器 11 不仅可以用于存储安装于所述服务器 1 的应用软件及各类数据，例如股指预测程序 10 及训练好的双向长短期记忆网络模型等。

处理器 12 在一些实施例中可以是一中央处理器（Central Processing Unit, CPU），微处理器或其它数据处理芯片，用于运行存储器 11 中存储的程序代
20 码或处理数据，例如执行股指预测程序 10 的计算机程序代码、执行双向长短期记忆网络模型的训练等。

图 2 仅示出了具有组件 11-13 以及股指预测程序 10 的服务器 1，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。

25 可选地，该服务器 1 还可以包括显示器，显示器可以称为显示屏或显示单元。在一些实施例中显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及有机发光二极管(Organic Light-Emitting Diode, OLED)触摸器等。显示器用于显示在服务器 1 中处理的信息以及用于显示可视化的工作界面，例如显示各个任务当前的执行状态。

30 可选地，该服务器 1 还可以包括用户接口，用户接口可以包括输入单元

比如键盘 (Keyboard)、语音输出装置比如音响、耳机等, 可选地用户接口还可以包括标准的有线接口、无线接口。

可选地, 该服务器 1 还包括触摸传感器。所述触摸传感器所提供的供用户进行触摸操作的区域称为触控区域。此外, 这里所述的触摸传感器可以为电阻式触摸传感器、电容式触摸传感器等。而且, 所述触摸传感器不仅包括接触式的触摸传感器, 也可包括接近式的触摸传感器等。此外, 所述触摸传感器可以为单个传感器, 也可以为例如阵列布置的多个传感器。用户可以通过触摸所述触控区域启动股指预测程序 10。

此外, 该电子装置 1 的显示器的面积可以与所述触摸传感器的面积相同, 也可以不同。可选地, 将显示器与所述触摸传感器层叠设置, 以形成触摸显示屏。该装置基于触摸显示屏侦测用户触发的触控操作。

该服务器 1 还可以包括射频 (Radio Frequency, RF) 电路、传感器和音频电路等等, 在此不再赘述。

如图 2 所示, 是图 1 中股指预测程序 10 较佳实施例的模块示意图。本申请所称的模块是指能够完成特定功能的一系列计算机程序指令段。

在本实施例中, 股指预测程序 10 包括: 样本采集模块 110、提取模块 120、处理模块 130、计算模块 140、权重确定模块 150 及预测模块 160, 结合图 3 中程序模块的应用环境示意图, 所述模块 110-160 所实现的功能或操作步骤如下:

样本采集模块 110, 用于从数据平台, 如金融数据和分析工具服务商 (Wind)、彭博 (Bloomberg) 等, 提取沪深 300 指数 2 个月内每段时间序列所有时间点的指数因子及对应的收益率, 并根据预设规则选取 n 个指数因子映射成一个 n 维向量 x_i , $i > 0$ 且 i 为整数。其中, 解释沪深 300 指数的常见指数因子有以下 13 种, 包括: 中债国债到期收益率: 10 年-中债企业债到期收益率 (AAA): 10 年、风险溢价、股息率、慢速 KD 指标 (SlowKD)、平滑异同移动平均线直方图 (Moving Average Convergence and Divergence Histogram, MACD Histogram)、布林线指标 (Bollinger Bands)、移动平均线-相对强弱指标 (MA of RSI(14) [m=22])、4-period MA of 4week MA of modified OBV-(MA4*4)、CR 指标、大小盘换手率比值、RSRS 指标、溢价率、主动买入额。所述预设规

则包括：分析用户的需求，根据用户需求确定用户意图，并根据预设的意图与指数因子的对应关系，选取对应的指数因子。所述用户意图是指用户希望综合指标所反映的结果，包括希望综合指标所反映的结果是时间序列的指数走势或时间点的具体的收益率。所述收益率是指日收益率，即以一天为时间间隔的相对收益。所述日收益率的计算公式为：第 n 日的收益率=(第 $n+1$ 日收盘价-第 n 日收盘价)/第 n 日收盘价。

进一步地，在指数因子特征提取前，还需要先判断每段时间序列的长度是否超过预设值，若时间序列的长度超过预设值，则以预设方式对时间序列进行截取，得到长度合适的时间序列。所述预设方式可以是指以步长为 t ($t=2,3,\dots$) 进行截取数据。假设，预设值为30，对于长度超过30个时间单位的时间序列，可能会造成梯度爆炸或梯度消失，待分析的时间序列为50个时间单位，采用步长为2的采样方法对序列进行重构，将时间序列长度归为25个时间单位。

提取模块120，用于提取样本数据中每段时间序列所有时间点的 n 维向量 x_i ，作为双向长短期记忆网络模型第一层的输入。例如，提取样本数据中每段时间序列的每日的 n 个指数因子及对应的收益率，根据 n 个指数因子构成 n 维向量，将 n 维向量和对应的收益率进行数据划分为训练集及验证集，将训练集中是 n 维向量输入到双向长短期记忆网络进行训练。

处理模块130，用于在双向长短期记忆网络模型的第二层，根据某段时间序列的某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i ，并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ，将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理，得到该段时间序列中每个时间点 n 维向量的综合隐藏层状态向量。以此类推，算出时间段上所有时间点的综合隐藏层状态向量，再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i 。

在另一个实施例中，还可以将该段时间序列中每个时间点 n 维向量的第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行拼接处理，得到每个时间点 n 维向量的综合隐藏层状态向量 T_i 。例如，将4月份每一日的 n 维向量的

第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行拼接处理, 得到每一日的综合隐藏层状态向量。

计算模块 140, 用于在双向长短期记忆网络模型的第三层, 根据每段时间序列的每个时间点的 n 维向量的综合隐藏层状态向量 T_i , 利用预设的指标计算公式计算每段时间序列的综合解释性指标 S 。其中, 所述预设的指标计算公式为: $S = a_1 T_1 + a_2 T_2 + a_3 T_3 + \dots + a_k T_k$, 其中, T_1 、 T_2 、.....、 T_k 分别代表每个时间点 n 维向量的综合隐藏层状态向量, a_1 、 a_2 、.....、 a_k 分别代表 T_1 、 T_2 、.....、 T_k 的权重。

权重确定模块150, 用于在双向长短期记忆网络模型的最后一层, 将每段时间序列的综合解释性指标 S 及该段时间序列对应的所有收益率代入反向传播 (Back Propagation, BP) 算法, 得到各个权重 a_i 。例如, 根据沪深300指数各个月份的综合解释性指标 S 及每日的收益率, 利用BP算法确定各个权重 a_i 。其中, 所述BP算法的表达式为:

$$a'_k = a_k - \eta * \left(\frac{\partial E}{\partial a_k} \right)$$

其中, 所述 a'_k 代表更新后的第 k 个权重, a_k 代表更新前的第 k 个权重, η 代表学习速率, $0.1 < \eta < 3$, $\frac{\partial E}{\partial a_k}$ 代表整体误差 E 对 a_k 的偏导值。具体地, 整体误差 $E = (S - S')^2$, S 代表综合解释性指标预测值, 即上述指标计算公式计算得到的综合解释性指标, S' 代表综合解释性指标的真实值, 即由每日的收益率得到的综合解释性指标。进一步地, 当输出的综合解释性指标为2维, 即综合解释性指标为 S_1 和 S_2 时, $E = 0.5(S_1 - S'_1)^2 + 0.5(S_2 - S'_2)^2$ 。对 E 求 a_k 的偏导乘以学习速率, 得到偏置值, a_k 减去偏置值得到 a_k 的更新值 a'_k 。持续循环训练, 直到偏置值 $\eta \left(\frac{\partial E}{\partial a_k} \right)$ 达到预设值, 模型为最优模型, 输出最终的各个权重值。

预测模块 160, 用于接收待分析的时间序列, 提取出时间序列中所有时间点的 n 维向量, 输入到训练好的双向长短期记忆网络模型中, 得到该时间序列的综合解释性指标。例如, 需要解释 2018 年 5 月指数走势时, 可以选取前几个月的逐日数据, 提取每日的 n 维向量, 输入到训练好的双向长短期记忆网络模型中, 得到该 5 月的指数走势。

如图 3 所示, 是本申请基于指数特征提取的股指预测方法较佳实施例的流程图。

下面以沪深 300 指数为例，但不仅限于沪深 300 指数，对本申请加以阐述。在本实施例中，处理器 12 执行存储器 11 中存储的股指预测程序 10 的计算机程序时实现基于指数特征提取的股指预测方法包括：步骤 S10-步骤 S60，其中，步骤 S10 为样本数据的采集步骤，步骤 S20-50 为双向长短期记忆网络模型的训练步骤，如图 4 所示，是本申请双向长短期记忆网络模型的示意图，步骤 S60 为股指趋势的预测步骤。

步骤 S10，样本采集模块 110 从数据平台，如金融数据和分析工具服务商（Wind）、彭博（Bloomberg）等，提取沪深 300 指数 2 个月内每段时间序列所有时间点的指数因子及对应的收益率，并根据预设规则选取 n 个指数因子映射成一个 n 维向量 x_i ， $i>0$ 且 i 为整数。其中，解释沪深 300 指数的常见指数因子有以下 13 种，包括：中债国债到期收益率：10 年-中债企业债到期收益率（AAA）：10 年、风险溢价、股息率、慢速 KD 指标（SlowKD）、平滑异同移动平均线直方图（Moving Average Convergence and Divergence Histogram, MACD Histogram）、布林线指标（Bollinger Bands）、移动平均线-相对强弱指标（MA of RSI(14) [m=22]）、4-period MA of 4week MA of modified OBV-(MA4*4)、CR 指标、大小盘换手率比值、RSRS 指标、溢价率、主动买入额。所述预设规则包括：分析用户的需求，根据用户需求确定用户意图，并根据预设的意图与指数因子的对应关系，选取对应的指数因子。所述用户意图是指用户希望综合指标所反映的结果，包括希望综合指标所反映的结果是时间序列的指数走势或时间点的具体的收益率。例如，当用户希望从宏观基本角度对沪深 300 走势进行刻画时，可以选取专注于宏观基本角度的 5 个指数因子，包括：中债国债到期收益率：10 年-中债企业债到期收益率（AAA）：10 年、风险溢价、股息率、溢价率及主动买入额，构成一个 5 维的向量。将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据。所述收益率是指日收益率，即以一天为时间间隔的相对收益。所述日收益率的计算公式为：第 n 日的收益率=(第 $n+1$ 日收盘价-第 n 日收盘价)/第 n 日收盘价。

进一步地，在指数因子特征提取前，还需要先判断每段时间序列的长度是否超过预设值，若时间序列的长度超过预设值，则以预设方式对时间序列进行截取，得到长度合适的时间序列。所述预设方式可以是指以步长为 t （ $t=2,3,\dots$ ）进行截取数据。假设，预设值为 30，对于长度超过 30 个时间

单位, 如 30 天, 的时间序列, 可能会造成梯度爆炸或梯度消失, 待分析的时间序列为 50 个时间单位, 采用步长为 2 的采样方法对序列进行重构, 将时间序列长度归为 25 个时间单位。

步骤 S20, 提取模块 120 提取样本数据中每段时间序列所有时间点的 n 维向量 x_i , 作为双向长短期记忆网络模型第一层的输入。例如, 提取样本数据中每段时间序列的每日的 n 个指数因子及对应的收益率, 根据 n 个指数因子构成 n 维向量, 将 n 维向量和对应的收益率进行数据划分为训练集及验证集, 将训练集中是 n 维向量输入到双向长短期记忆网络进行训练。

步骤 S30, 在双向长短期记忆网络模型的第二层, 处理模块 130 根据某段时间序列的某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i , 并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' , 将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理, 得到该段时间序列中每个时间点 n 维向量的综合隐藏层状态向量。以此类推, 算出时间段上所有时间点的综合隐藏层状态向量, 再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i 。例如, 根据沪深 300 指数 4 月 15 日的 n 维向量 x_i 和 4 月 14 日的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算 15 日的 n 维向量 x_i 的第一隐藏层状态向量 h_i , 并根据沪深 300 指数 4 月 15 日的 n 维向量 x_i 和 4 月 16 日的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该日的 n 维向量 x_i 的第二隐藏层状态向量 h_i' , 将 4 月 15 日的第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理, 得到 4 月 15 日的综合隐藏层状态向量。以此类推, 算出时序中所有时间点的综合隐藏层状态向量, 得到 4 月份的特征向量 T_i 。

在另一个实施例中, 还可以将该段时间序列中每个时间点 n 维向量的第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行拼接处理, 得到每个时间点 n 维向量的综合隐藏层状态向量 T_i 。例如, 将 4 月份每一日的 n 维向量的第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行拼接处理, 得到每一日的综合隐藏层状态向量。

步骤 S40, 在双向长短期记忆网络模型的第三层, 计算模块 140 根据每段时间序列的特征向量 T_i , 利用预设的指标计算公式计算下一段时间序列的综合

解释性指标S。其中，所述预设的指标计算公式为： $S = a_1T_1 + a_2T_2 + a_3T_3 + \dots + a_kT_k$ ，其中， T_1 、 T_2 、.....、 T_k 分别代表每个时间点n维向量的综合隐藏层状态向量， a_1 、 a_2 、.....、 a_k 分别代表 T_1 、 T_2 、.....、 T_k 的权重。例如，将1-4月份的特征向量 T_1 - T_4 输入预设的指标计算公式，得到5月份的综合解释性指标S。

步骤S50，在双向长短期记忆网络模型的最后一层，权重确定模块150将每段时间序列的综合解释性指标S及该段时间序列对应的所有收益率代入BP算法，得到各个权重 a_i 。例如，根据沪深300指数各个月份的综合解释性指标S及每日的收益率，利用BP算法确定各个权重 a_i 。其中，所述BP算法的表达式为：

$$a'_k = a_k - \eta \left(\frac{\partial E}{\partial a_k} \right)$$

其中，所述 a'_k 代表更新后的第k个权重， a_k 代表更新前的第k个权重， η 代表学习速率， $0.1 < \eta < 3$ ， $\frac{\partial E}{\partial a_k}$ 代表整体误差E对 a_k 的偏导值。具体地，整体误差 $E = (S - S')^2$ ，S代表综合解释性指标预测值，即上述指标计算公式计算得到的综合解释性指标， S' 代表综合解释性指标的真实值，即由每日的收益率得到的综合解释性指标。进一步地，当输出的综合解释性指标为2维，即综合解释性指标为 S_1 和 S_2 时， $E = 0.5(S_1 - S'_1)^2 + 0.5(S_2 - S'_2)^2$ 。对E求 a_k 的偏导乘以学习速率，得到偏置值， a_k 减去偏置值得到 a_k 的更新值 a'_k 。持续循环训练，直到偏置值 $\eta \left(\frac{\partial E}{\partial a_k} \right)$ 达到预设值，模型为最优模型，输出最终的各个权重值。

步骤S60，预测模块160接收待分析的时间序列，提取出时间序列中所有时间点的n维向量，输入到训练好的双向长短期记忆网络模型中，得到该时间序列的综合解释性指标。例如，需要解释2018年5月指数走势时，可以选取前几个月的逐日数据，提取每日的n维向量，输入到训练好的双向长短期记忆网络模型中，得到该5月的指数走势。

上述实施例提出的基于指数特征提取的股指预测方法，通过提取每段时间序列所有时间点的指数因子及对应的收益率，并选取n个指数因子构成n维向量 x_i 。之后，该方法将n维向量 x_i 及收益率代入双向长短期记忆网络模型中训练，得到完整的模型。最后，该方法提取待分析的时间序列的n维向量，输入到模型中，得到该时间序列的综合解释性指标，从而避免梯度消失和梯度爆炸，准确的分析股指的走势。

此外，本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质中包括股指预测程序 10，所述股指预测程序 10 被处理器执行时实现如下操作：

5 样本采集步骤：提取预设数量的时间序列中所有时间点的指数因子及对应的收益率，根据预设规则选取 n 个指数因子构成一个 n 维向量 x_i ， $i>0$ 且 i 为整数，将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据；

提取步骤：提取样本数据中每段时间序列的所有时间点的 n 维向量 x_i ，作为双向长短期记忆网络模型第一层的输入；

10 处理步骤：在双向长短期记忆网络模型的第二层，根据某段时间序列的某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i ，并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ，将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理，得到该时间点的综合隐藏层状态向量，直至算出所有时间点的综合隐藏层状态向量，再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i ；

15 计算步骤：在双向长短期记忆网络模型的第三层，根据每段时间序列的特征向量 T_i ，利用预设的指标计算公式计算下一段时间序列的综合解释性指标
20 S ；

权重确定步骤：在双向长短期记忆网络模型的最后一层，将每段时间序列的综合解释性指标 S 及该段时间序列对应的所有收益率代入反向传播算法，得到各段时间序列的权重 a_i ；

25 预测步骤：接收待分析的时间序列，提取该段时间序列中所有时间点的 n 维向量，输入到训练好的双向长短期记忆网络模型中，得到该时间序列的综合解释性指标 S' 。

优选地，所述预设规则包括：

分析用户的需求，根据用户需求确定用户意图；及

根据预设的意图与指数因子的对应关系，选取对应的指数因子。

30 优选地，所述预设的指标计算公式为：

$$S = a_1 T_1 + a_2 T_2 + a_3 T_3 + \cdots + a_k T_k$$

其中， T_1 、 T_2 、.....、 T_k 分别代表每段时间序列的特征向量， a_1 、 a_2 、.....、 a_k 分别代表 T_1 、 T_2 、.....、 T_k 的权重。

优选地，所述反向传播算法的公式为：

$$a'_k = a_k - \eta \left(\frac{\partial E}{\partial a_k} \right)$$

其中，所述 a'_k 代表更新后的第k个权重， a_k 代表更新前的第k个权重， η 代表学习速率， $0.1 < \eta < 3$ ， $\frac{\partial E}{\partial a_k}$ 代表整体误差E对 a_k 的偏导值。

优选地，该方法还包括：

判断预设数量的时间序列的长度是否超过预设值，若时间序列的长度超过预设值，则以预设方式对时间序列进行截取，得到长度合适的时间序列。

本申请之计算机可读存储介质的具体实施方式与上述基于指数特征提取的股指预测方法的具体实施方式大致相同，在此不再赘述。

上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。

需要说明的是，在本文中，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、装置、物品或者方法不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、装置、物品或者方法所固有的要素。在没有更多限制的情况下，由语句“包括一个.....”限定的要素，并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在如上所述的一个存储介质（如ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，或者网络设备等）执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其它相关的技术领域，均同理包括在本申请的专利保护范围内。

权利要求书

1、一种基于指数特征提取的股指预测方法，应用于服务器，其特征在于，所述方法包括：

样本采集步骤：提取预设数量的时间序列中所有时间点的指数因子及对应的收益率，根据预设规则选取 n 个指数因子构成一个 n 维向量 x_i ， $i>0$ 且 i 为整数，将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据；

提取步骤：提取样本数据中每段时间序列的所有时间点的 n 维向量 x_i ，作为双向长短期记忆网络模型第一层的输入；

处理步骤：在双向长短期记忆网络模型的第二层，根据某段时间序列的某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i ，并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ，将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理，得到该时间点的综合隐藏层状态向量，直至算出所有时间点的综合隐藏层状态向量，再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i ；

计算步骤：在双向长短期记忆网络模型的第三层，根据每段时间序列的特征向量 T_i ，利用预设的指标计算公式计算下一段时间序列的综合解释性指标 S ；

权重确定步骤：在双向长短期记忆网络模型的最后一层，将每段时间序列的综合解释性指标 S 及该段时间序列对应的所有收益率代入反向传播算法，得到各段时间序列的权重 a_i ；

预测步骤：接收待分析的时间序列，提取该段时间序列中所有时间点的 n 维向量，输入到训练好的双向长短期记忆网络模型中，得到该时间序列的综合解释性指标 S 。

2、根据权利要求1所述的基于指数特征提取的股指预测方法，其特征在于，所述预设规则包括：

分析用户的需求，根据用户需求确定用户意图；及

根据预设的意图与指数因子的对应关系，选取对应的指数因子。

3、根据权利要求1所述的基于指数特征提取的股指预测方法，其特征在

于, 所述预设的指标计算公式为:

$$S = a_1 T_1 + a_2 T_2 + a_3 T_3 + \cdots + a_k T_k$$

其中, T_1 、 T_2 、.....、 T_k 分别代表每段时间序列的特征向量, a_1 、 a_2 、.....、 a_k 分别代表 T_1 、 T_2 、.....、 T_k 的权重。

4、根据权利要求1或3所述的基于指数特征提取的股指预测方法, 其特征
5 在于, 所述反向传播算法的公式为:

$$a'_k = a_k - \eta \left(\frac{\partial E}{\partial a_k} \right)$$

其中, 所述 a'_k 代表更新后的第k个权重, a_k 代表更新前的第k个权重, η 代表学习速率, $0.1 < \eta < 3$, $\frac{\partial E}{\partial a_k}$ 代表整体误差E对 a_k 的偏导值。

5、根据权利要求1所述的基于指数特征提取的股指预测方法, 其特征在
10 于, 该方法还包括:

判断预设数量的时间序列的长度是否超过预设值, 若时间序列的长度超
过预设值, 则以预设方式对时间序列进行截取, 得到长度合适的时间序列。

6、根据权利要求5所述的基于指数特征提取的股指预测方法, 其特征在
于, 所述预设方式可以是指以步长为 t , t 为整数且 $t > 1$, 进行截取数据。

7、根据权利要求1所述的基于指数特征提取的股指预测方法, 其特征在
15 于, 所述收益率可以是指日收益率, 所述日收益率的计算公式为:

$$\text{第}n\text{日的收益率} = (\text{第}n+1\text{日收盘价} - \text{第}n\text{日收盘价}) / \text{第}n\text{日收盘价}$$

其中, n 为整数且 $n > 0$ 。

8、一种服务器, 其特征在於, 所述服务器包括: 存储器、处理器及显示
器, 所述存储器上存储有股指预测程序, 所述股指预测程序被所述处理器执
20 行, 可实现如下步骤:

样本采集步骤: 提取预设数量的时间序列中所有时间点的指数因子及对
应的收益率, 根据预设规则选取 n 个指数因子构成一个 n 维向量 x_i , $i > 0$ 且 i 为整
数, 将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据;

提取步骤: 提取样本数据中每段时间序列的所有时间点的 n 维向量 x_i , 作
25 为双向长短期记忆网络模型第一层的输入;

处理步骤: 在双向长短期记忆网络模型的第二层, 根据某段时间序列的
某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计
算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i , 并根据该时间点的 n 维向量

x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ,将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i' 进行平均处理,得到该时间点的综合隐藏层状态向量,直至算出所有时间点的综合隐藏层状态向量,再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i ;

计算步骤:在双向长短期记忆网络模型的第三层,根据每段时间序列的特征向量 T_i ,利用预设的指标计算公式计算下一段时间序列的综合解释性指标 S ;

权重确定步骤:在双向长短期记忆网络模型的最后一层,将每段时间序列的综合解释性指标 S 及该段时间序列对应的所有收益率代入反向传播算法,得到各段时间序列的权重 a_i ;

预测步骤:接收待分析的时间序列,提取该段时间序列中所有时间点的 n 维向量,输入到训练好的双向长短期记忆网络模型中,得到该时间序列的综合解释性指标 S 。

9、根据权利要求8所述的服务器,其特征在于,所述预设规则包括:

分析用户的需求,根据用户需求确定用户意图;及

根据预设的意图与指数因子的对应关系,选取对应的指数因子。

10、根据权利要求8所述的服务器,其特征在于,所述预设的指标计算公式为:

$$S = a_1 T_1 + a_2 T_2 + a_3 T_3 + \dots + a_k T_k$$

其中, T_1 、 T_2 、.....、 T_k 分别代表每段时间序列的特征向量, a_1 、 a_2 、.....、 a_k 分别代表 T_1 、 T_2 、.....、 T_k 的权重。

11、根据权利要求8或10所述的服务器,其特征在于,所述反向传播算法的公式为:

$$a'_k = a_k - \eta \left(\frac{\partial E}{\partial a_k} \right)$$

其中,所述 a'_k 代表更新后的第 k 个权重, a_k 代表更新前的第 k 个权重, η 代表学习速率, $0.1 < \eta < 3$, $\frac{\partial E}{\partial a_k}$ 代表整体误差 E 对 a_k 的偏导值。

12、根据权利要求8所述的服务器,其特征在于,所述股指预测程序被所述处理器执行时,还包括步骤:

判断预设数量的时间序列的长度是否超过预设值,若时间序列的长度超

过预设值，则以预设方式对时间序列进行截取，得到长度合适的时间序列。

13、根据权利要求12所述的服务器，其特征在于，所述预设方式可以是指以步长为 t ， t 为整数且 $t>1$ ，进行截取数据。

14、根据权利要求8所述的服务器，其特征在于，所述收益率可以是指日
5 收益率，所述日收益率的计算公式为：

$$\text{第}n\text{日的收益率}=(\text{第}n+1\text{日收盘价}-\text{第}n\text{日收盘价})/\text{第}n\text{日收盘价}$$

其中， n 为整数且 $n>0$ 。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中包括股指预测程序，所述股指预测程序被处理器执行时，可实现如下步骤：

10 样本采集步骤：提取预设数量的时间序列中所有时间点的指数因子及对应的收益率，根据预设规则选取 n 个指数因子构成一个 n 维向量 x_i ， $i>0$ 且 i 为整数，将各个时间点的 n 维向量 x_i 及其对应的收益率组成待训练的样本数据；

提取步骤：提取样本数据中每段时间序列的所有时间点的 n 维向量 x_i ，作为双向长短期记忆网络模型第一层的输入；

15 处理步骤：在双向长短期记忆网络模型的第二层，根据某段时间序列的某个时间点的 n 维向量 x_i 及前一个时间点的 n 维向量 x_{i-1} 的隐藏层状态向量 h_{i-1} 计算该时间点的 n 维向量 x_i 的第一隐藏层状态向量 h_i ，并根据该时间点的 n 维向量 x_i 及后一个时间点的 n 维向量 x_{i+1} 的隐藏层状态向量 h_{i+1} 计算该时间点 n 维向量 x_i 的第二隐藏层状态向量 h_i' ，将第一隐藏层状态向量 h_i 和第二隐藏层状态向量 h_i'
20 进行平均处理，得到该时间点的综合隐藏层状态向量，直至算出所有时间点的综合隐藏层状态向量，再根据每段时间序列的所有时间点的综合隐藏层状态向量得到每段时间序列的特征向量 T_i ；

计算步骤：在双向长短期记忆网络模型的第三层，根据每段时间序列的特征向量 T_i ，利用预设的指标计算公式计算下一段时间序列的综合解释性指标
25 S；

权重确定步骤：在双向长短期记忆网络模型的最后一层，将每段时间序列的综合解释性指标 S 及该段时间序列对应的所有收益率代入反向传播算法，得到各段时间序列的权重 a_i ；

预测步骤：接收待分析的时间序列，提取该段时间序列中所有时间点的 n
30 维向量，输入到训练好的双向长短期记忆网络模型中，得到该时间序列的综

合解释性指标S。

16、根据权利要求15所述的计算机可读存储介质，其特征在于，所述预设规则包括：

分析用户的需求，根据用户需求确定用户意图；及

5 根据预设的意图与指数因子的对应关系，选取对应的指数因子。

17、根据权利要求15所述的计算机可读存储介质，其特征在于，所述预设的指标计算公式为：

$$S = a_1 T_1 + a_2 T_2 + a_3 T_3 + \cdots + a_k T_k$$

其中， T_1 、 T_2 、……、 T_k 分别代表每段时间序列的特征向量， a_1 、 a_2 、……、 a_k 分别代表 T_1 、 T_2 、……、 T_k 的权重。

10 18、根据权利要求15或17所述的计算机可读存储介质，其特征在于，所述反向传播算法的公式为：

$$a'_k = a_k - \eta \left(\frac{\partial E}{\partial a_k} \right)$$

其中，所述 a'_k 代表更新后的第k个权重， a_k 代表更新前的第k个权重， η 代表学习速率， $0.1 < \eta < 3$ ， $\frac{\partial E}{\partial a_k}$ 代表整体误差E对 a_k 的偏导值。

15 19、根据权利要求15所述的计算机可读存储介质，其特征在于，所述股指预测程序被所述处理器执行时，还包括步骤：

判断预设数量的时间序列的长度是否超过预设值，若时间序列的长度超过预设值，则以预设方式对时间序列进行截取，得到长度合适的时间序列。

20、根据权利要求19所述的计算机可读存储介质，其特征在于，所述预设方式可以是指以步长为t，t为整数且 $t > 1$ ，进行截取数据。

20

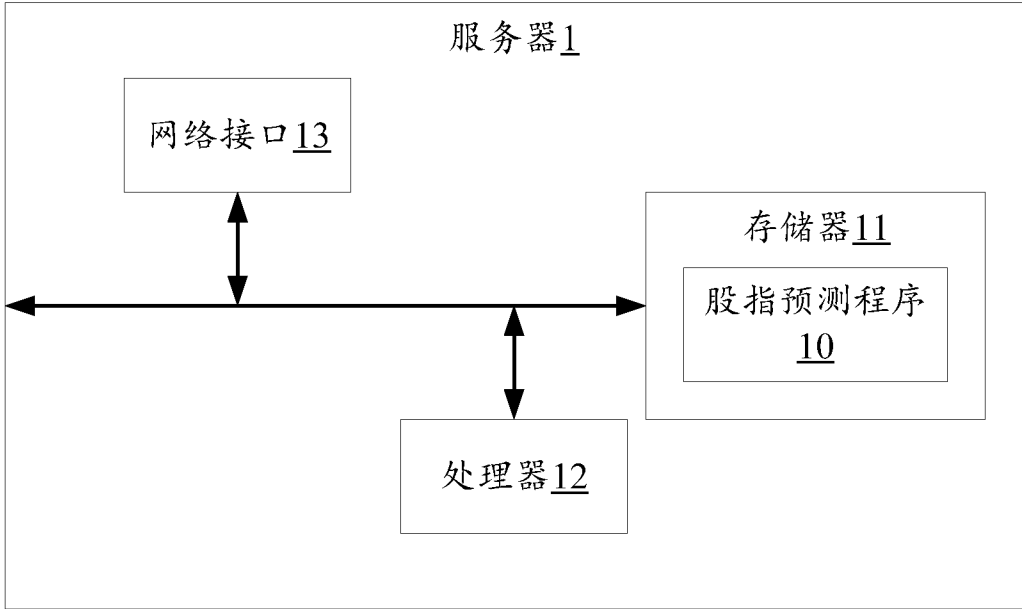


图 1



图 2

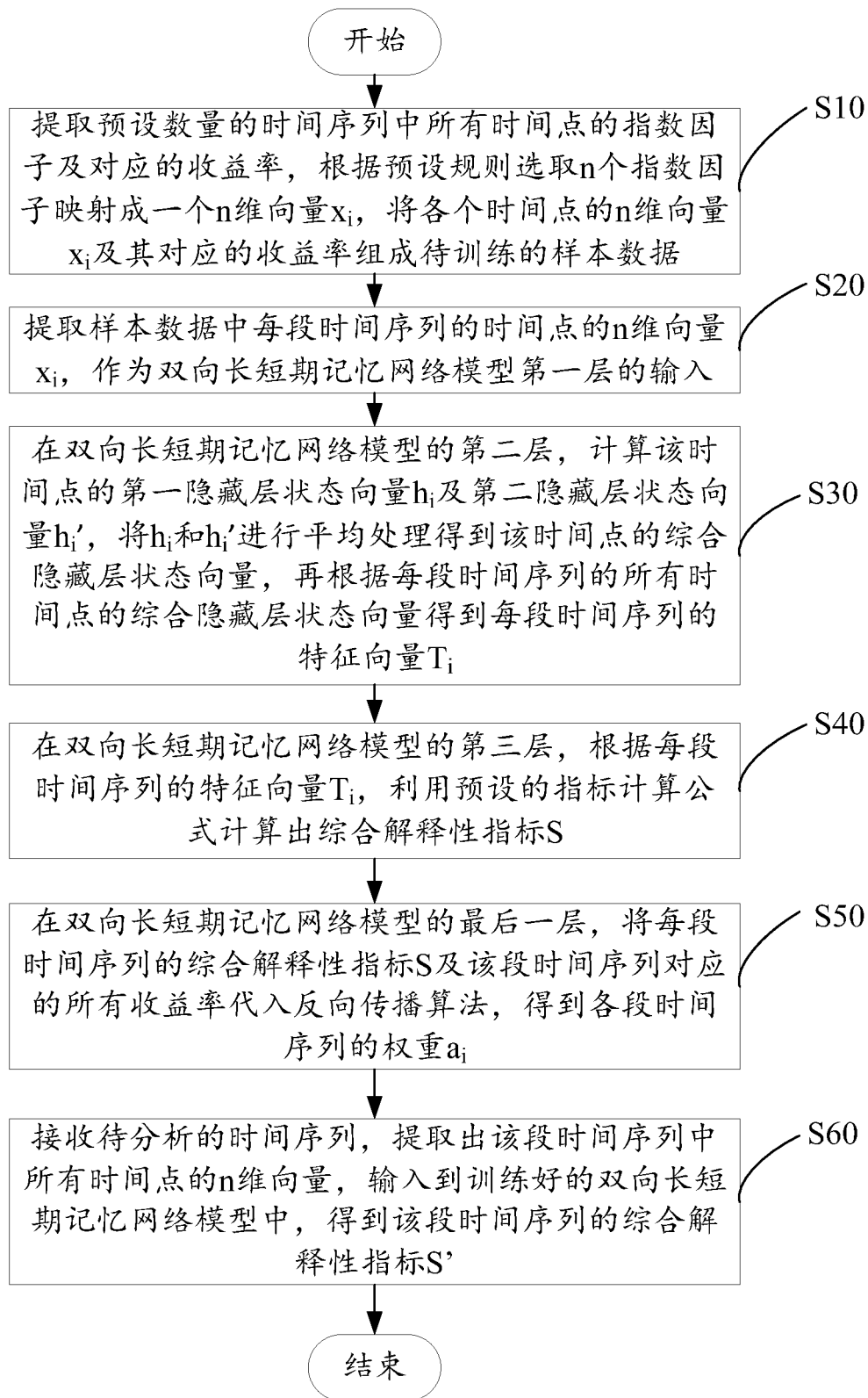


图 3

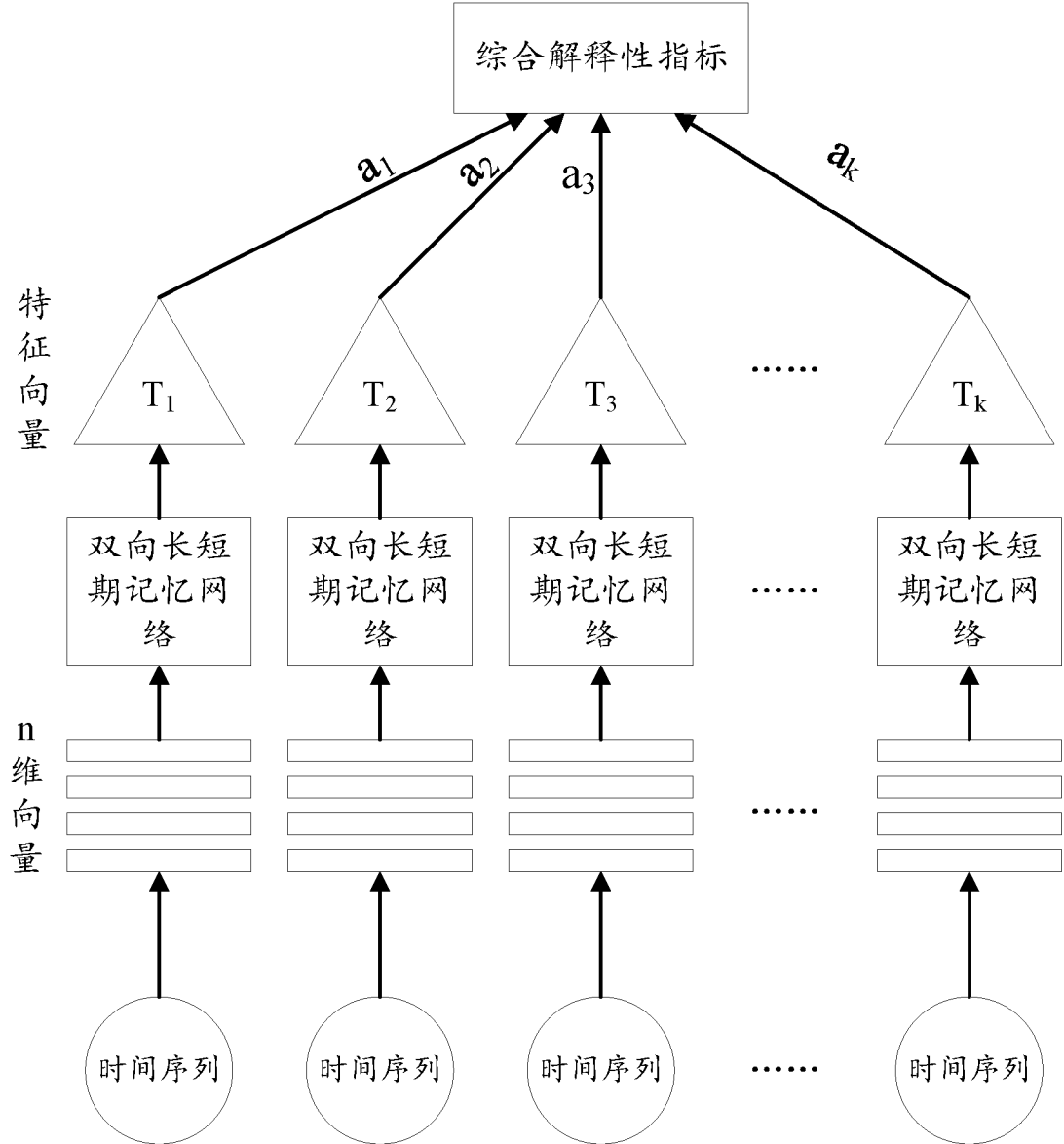


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/107484

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/04(2012.01)i; G06Q 40/04(2012.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNPAT, WPI, EPODOC, IEEE: 股票, 股指, 预测, 双向, 长短期记忆, 双向长短期记忆, 反向传播, 收益, 权重, 价格, stock, prediction, bidirectional, long 1w short 1w term 1w memory, BI 1W LSTM, weight, income, reverse, propagation, price

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 106952161 A (HONG, ZHILING) 14 July 2017 (2017-07-14) description, paragraphs 0010-0019	1-20
A	US 2017032243 A1 (GOOGLE INC.) 02 February 2017 (2017-02-02) entire document	1-20
A	CN 105787582 A (TSINGHUA UNIVERSITY) 20 July 2016 (2016-07-20) entire document	1-20
A	CN 107832897 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY) 23 March 2018 (2018-03-23) entire document	1-20

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

01 March 2019

Date of mailing of the international search report

27 March 2019

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/107484

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106952161	A	14 July 2017	None			
US	2017032243	A1	02 February 2017	EP	3274888	A1	31 January 2018
				JP	2018527636	A	20 September 2018
				KR	20170132853	A	04 December 2017
				CN	107851462	A	27 March 2018
				WO	2017019707	A1	02 February 2017
				US	2017316313	A1	02 November 2017
CN	105787582	A	20 July 2016	None			
CN	107832897	A	23 March 2018	None			

国际检索报告

国际申请号

PCT/CN2018/107484

A. 主题的分类

G06Q 10/04(2012.01)i; G06Q 40/04(2012.01)n

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06Q

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNKI, CNPAT, WPI, EPODOC, IEEE: 股票, 股指, 预测, 双向, 长短期记忆, 双向长短期记忆, 反向传播, 收益, 权重, 价格, stock, prediction, bidirectional, long 1w short 1w term 1w memory, BI 1W LSTM, weight, income, reverse, propagation, price

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 106952161 A (洪志令) 2017年 7月 14日 (2017 - 07 - 14) 说明书第0010-0019段	1-20
A	US 2017032243 A1 (GOOGLE INC.) 2017年 2月 2日 (2017 - 02 - 02) 全文	1-20
A	CN 105787582 A (清华大学) 2016年 7月 20日 (2016 - 07 - 20) 全文	1-20
A	CN 107832897 A (浙江工业大学) 2018年 3月 23日 (2018 - 03 - 23) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 3月 1日

国际检索报告邮寄日期

2019年 3月 27日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

王朝英

电话号码 86-(10)-53961618

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/107484

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	106952161	A	2017年 7月 14日	无			
US	2017032243	A1	2017年 2月 2日	EP	3274888	A1	2018年 1月 31日
				JP	2018527636	A	2018年 9月 20日
				KR	20170132853	A	2017年 12月 4日
				CN	107851462	A	2018年 3月 27日
				WO	2017019707	A1	2017年 2月 2日
				US	2017316313	A1	2017年 11月 2日
CN	105787582	A	2016年 7月 20日	无			
CN	107832897	A	2018年 3月 23日	无			