

# Hyperband-LSTM

陈健 刘伟\*  
(上海财经大学, 上海 200030)

出于与, 一个于  
价, 以决前 任务上  
不 不 低下。  
, 以 价为例, 入 作为  
, 具  
, 且 优化 优化 一,  
出。  
优于其。使上,  
与 保一, 具 健。  
算法 神 票价格 测 检  
中 分 : F830.91 :A :

[投稿日期]: 2021-01-04; [修回日期]: 2022-02-10

[作者简介]陈健, 女, 1972.-, 经济学博士, 上海师范大学商学院副教授, 硕士生导师; 研究领域: 金融时间序列分析; 电子邮箱: chenjian@shnu.edu.cn; 联系电话: 021-64322339; 邮政编码: 200030; 通讯地址: 上海市徐汇区桂林路 100 号。

通讯作者: 刘伟基\* (通讯作者), 男, 1996.3-, 经济统计学专业硕士研究生, 上海师范大学商学院; 研究领域: 金融时间序列分析; 电子邮箱: vickylau@163.com; 联系电话: 13172907221; 邮政编码: 200030; 通讯地址: 上海市徐汇区桂林路 100 号。感谢匿名评审专家和编辑部的宝贵意见, 文责自负。

**R a c S c P c P c B a H b a -LSTM**

**M**

Jian chen, Weiji Liu

**Abstract:**

**Key Words:**

## 一、引言

产 其 ， 到了 。但 产 伴  
，一 价 。  
于 习 列中 关 ， 于 价 。但 于  
传 ， 习 ，但

关 别 力仍 ， 且 乏一 力。 中 出  
习 ， ，具 优 别 力 化  
力， 包 内 任务上 出优于 传 习 （ ， ）  
[6]。 ， 于 习 价 优于传 习  
。 （ ）[7] 使 价，  
值 低于

价 上 以 低 到 。

## 二、模型原理

分 先 做 一 个 分 ，  
价 列 ， 以 入  
， 以  
升 。

### (一) LSTM 神经网络原理

为了  
， 以 一 个  
为 例 。 ， 一 ( )  
入 值 ，  
元  $g_i^1(X) = W_i^1 X + b_i^1$  入  
值  $X = \{x_1, x_2, \dots, x_n\}$  了  $i$   
。  
函  $f^1(\cdot)$  出 出  $y_i^1$ ， 其中  
 $y_i^1 = f^1(g_i^1)$ ， 出  
 $i$  个 作 为 出 入 。

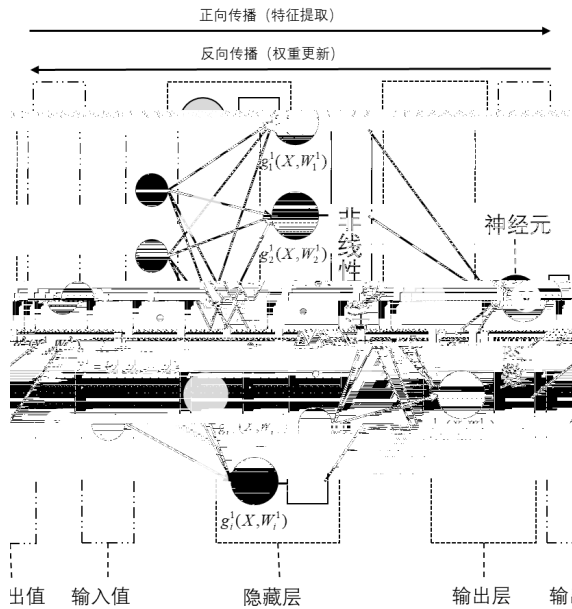


图1 双层神经网络

以  $Y^1 = \{y_1^1, y_2^1, \dots, y_i^1\}$  作为 二 ( 出 ) 入 ， 似 出  
出  $y$ 。 个 为 传 ， 其 作 入  
， 以 到 。 ， 以 为 依 ， 以 下  
为  $W = \{W_1^1, W_2^1, \dots, W_i^1; W_1^2\}$ ， 具 体 依 具 体 优 化  
， 为 传 ， 到 使 到  
， 具 一 力。

以 了 到： 具 化

力。 ， 于 价 使 。但  
 价 列 列 关， 别出 列  
 ， （ ） 别出 列 列 关，  
 与 （ ） 、 （ ）  
 习 不 ：其 关 到不 入值与 值 关 ， 关 到 列中  
 列 关。 于 信 ， 列 关  
 列中， 信 。 分 列 不仅仅与 信  
 ， 与 信 关 ， 列，人们 不会 关 一  
 两 信 ， 于 列 势。 一个 ，  
 中加入一个 元， ， 个 元包 、 入 出 ， 其  
 中 决 信 、 入 决 信 保 、 出  
 信 信 出 前 出，使 决 列中 信  
 。

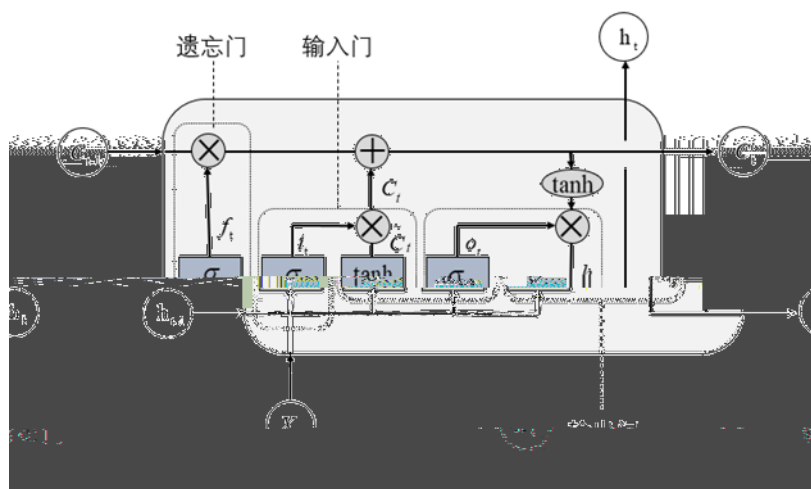


图2 LSTM单元

：  $x_t$  为  $t$  刻 入， 分别为上一 元  
 ， 也作为  $t$  刻 元 入，也 前 元 入包  $X_t$ 、  
 。 先  $X$  公 ( ) ， 出  $f_t$  决 。

入 :  $X_t$   $h_{t-1}$  公 ( ) ， 出  $i_t$  决 信 保 ，

$X$  公 ( ) , 出  $\tilde{C}_t$  为 前 元 入 信 。

$$i = \sigma(W_i h_{t-1} + X_t + b_i)$$

$$C_t = \sigma(W_c[h_{t-1}, X_t] + b_c)$$

出 : 公 ( ) 出 前 元 , 公 ( ) , 出 出  $o_t$  , 公 ( ) ,  $\tanh$  函  $C_t$  乘 出 元 出  $h_t$  。

$$C_t = f_t * C_{t-1} + i_t \otimes \tilde{C}_t$$

$$= \sigma(\quad)$$

$$h_t = o_t \otimes \tanh(C_t)$$

其中,  $\tanh(\cdot)$   $\tanh$  函 ,  $[\cdot]^2$  ,  $\otimes$  元 乘 ,  $W_f$ 、 $W_i$ 、 $W_o$ 、 $b_f$ 、 $b_i$  , 。

上 , 使 一 元代 了 元, 使 以 别出 列 列 关 , 使 价 。 , 于 中 前 , 元 函 , 前 以 准 判 , 一个 助 中 别 些 , 准 。

## (二) Hyperband-LSTM 模型原理

上 , 与 关 元 函 , 为 函 决 入 以何 以何 化为 出, 决 了 别 力。 决于 中 , 决于 元 、 决于优化 , 函 前 一 , 决 。

以 为 ， 入 为 ， 准 。 主 于 了  $n$  之 ， 其中  $B$  为 优化 中 。 中 ， 个 。 保  $1/\eta$ 。初 化 两个主 ：  $s_{\max} = \lfloor \log_{\eta}(R) \rfloor$ ，  $B = (s_{\max} + 1)R$ ， 其中  $\lfloor \cdot \rfloor$  下 ，  $s_{\max}$  决 ，  $!$  为 。 了初 ， 出

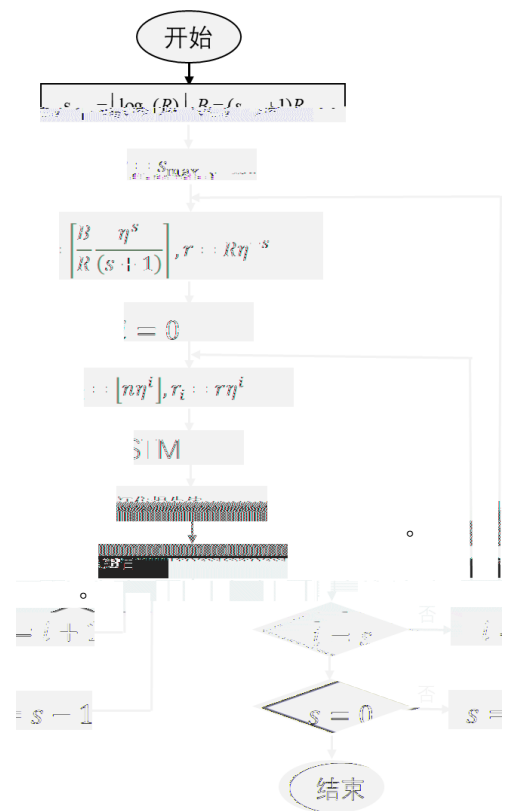


图3 Hyperband-LSTM原理

一 分 到  $r$ ； 初 化 内 减 ( ) ， 内 一 内 代  $n_0$  。 分 到  $r$ ； 内 做 作 到 中使 分别 ； 全 与 作 ， 不 ； 产 值 ， 保 其中  $1/\eta$  ； 下 一 内 ，  $i = s$ ， 内 二 ；  $s = !$  ， 一 ， 到 一 于 ， 使 其 力。 于 ， 别 一 ， 予 依 低 ， 上 优化 。 优化 低 优化 ， 以 丰 了 入



优化。 ， 为  
 一 优 ， 决了 动 不 ， 且 别 一  
 ， 优化 ， 了以 低  
 ， 决了 优化 优化 不 ，  
 优化 了 一 升。

### 三、实证研究

#### (一) 指标选取及数据说明

于万 ，包 了 分  
 ， 为 个。 为 五分 价，  
 个 。

表1 预测因子

<i>open</i>	<i>CCI</i>
<i>high</i>	<i>OSC</i>
<i>low</i>	<i>ROC</i>
<i>val</i>	<i>W&amp;R</i>
<i>vol</i>	<i>EMV</i>
<i>close</i>	<i>WVAD</i>
<i>BBI</i>	<i>CR</i>
<i>DDI</i>	<i>PSY</i>
<i>MACD</i>	<i>VR</i>
<i>MTM</i>	<i>BB</i>
<i>TRIX</i>	<i>BOLL</i>

划分为 、 三个 ，分别 三个  
 、一化 为[0,1 内 ， 于  
 于 于 于 价 。  
 于 充分利 则， 一个 从 中 ，  
 三个 一 交 ，但 不 信  
 前 ， 仍 于 信 信 做出 。 使 为 动  
 ， 动 中 使 价 。

表2 数据集划分

7656	2020/8/03 9:35	2021/3/30 11:30
2568	2021/3/30 9:35	2021/6/18 11:30
2544	2021/6/18 9:35	2021/8/31 15:00

中 个 ，会 别 ， 不  
 关 不 ， 会 关 ， 会  
 。为了 决 一 ， ，一化 ，  
 一 。 ，一化 以减 ，使  
 到 优 。 使 ，一化 为 ，一化，一化 公 ( )。

(7)

其中， 全 ， 为 值函 ， 为 值函 ， 为  
 值， 为、一化 值。  
 ， 于 假 ， 中 ，一化 使  
 中 值 值。 ，为了使 具 健 ，  
 ，一化 使 中 值 值。  
 ，一化 ， 值 化为 区 值， ，一化  
 ，一化 到 ， ，一化 公 ( )。

(8)

其中， 为 出 值， 为 ，一化 。

(二) 参数说明

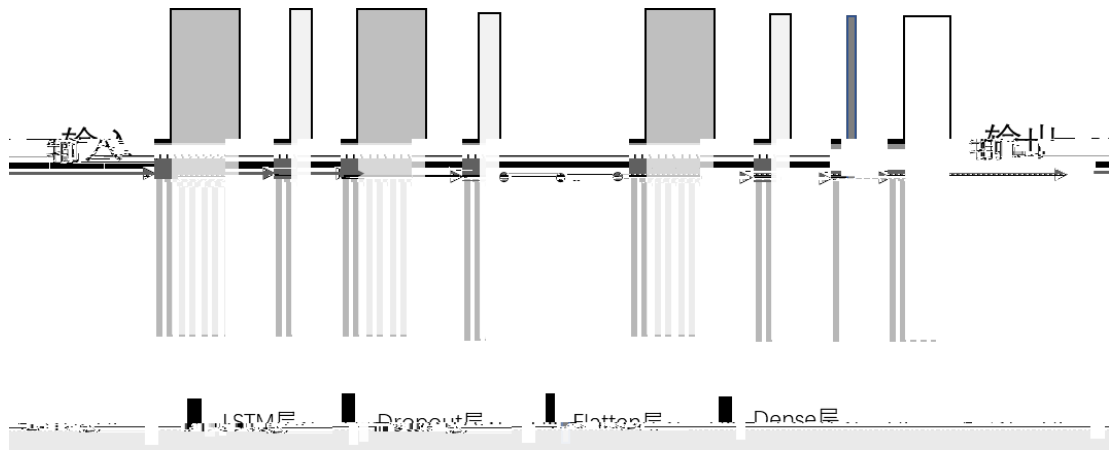


图4 LSTM神经网络的架构

万，中一，且元，力以任任何函。但，中，使；元以使任务（利·，）<sup>[13]</sup>。前价中中，元为2 3（，）<sup>[14]</sup>。上元上—，元，免元不。

关别力主体；函中，中函中到了不作。中，函；函中，，，函，函，但其不，到也具体。、为：初化，使—；，依下则，下一，作传；不停，到到，。

于优化决，前为优化包、、，但关中使优化到。、之前函优化前，函优化。

，以—于动，不任务中动化，前人关

中  $\eta=1/3$ 、 $R=$  内 会保 三分之一 ，  
代。

表3 超参数搜索空间

	2,5	1
	16,256	16
	0.1,0.6	0.1
		--
		--

### (三) 评价指标

入 价 估，分别为 ( )、  
、 分 、 关 。  
、 值与 值之 ， 分 值与  
值之 ( ， ) [15]。 ( )  
件 件) 下 优化 代 ，  
以 上体 。 、 、  
佳。 于 值 值之 关 ， 值 ， 值与 值  
(党 从 ， ) [16]。 、 、 分  
关 分别 公 ( )、公 ( )、公 ( ) 公 ( ) 。

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (9)$$

$$= \frac{1}{n} \sum_{i=1}^n \left| \hat{y}_i - y_i \right| \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (11)$$

$$PCCs = \frac{\sum_{i=1}^n y_i - \frac{1}{n} \sum_{i=1}^n y_i \quad \sum_{i=1}^n \hat{y}_i - \frac{1}{n} \sum_{i=1}^n \hat{y}_i}{\sqrt{\sum_{i=1}^n y_i - \frac{1}{n} \sum_{i=1}^n y_i} \sqrt{\sum_{i=1}^n (\hat{y}_i - \frac{1}{n} \sum_{i=1}^n \hat{y}_i)}} \quad (12)$$

其中 为 值， 为 值， 为 。

#### (四) 评价结论

信 判 假 。

### (五) 实证结果

为了体 价 上 ， 入了  
 作为 ， 使 优化 优化  
 。 于上 ， 出 ， 于  
 优化 ， 价 。

表4 沪深300指数实验结果

-	82.92	6.858	0.0013	0.9968	37 32
	107.6	7.607	0.0015	0.9975	
-	215.0	11.08	0.0022	0.9913	67 19
	322.8	12.89	0.0026	0.9916	
	1092	30.50	0.0060	0.9934	--
	864.0	26.12	0.0052	0.9945	

从 ， 优化 体上 优于 ， 优化  
 升 助 。 内 上， 上，  
 到 优于 ， 且  
 减 了 优化 ， 以  
 上 升 。 且， 中 低于  
 ， 具 优 化 。 了  
 上 ， 以 出 优于  
 。 ， 从 ， 价  
 任务上 出 低 ， 从 了 出  
 价 上 。

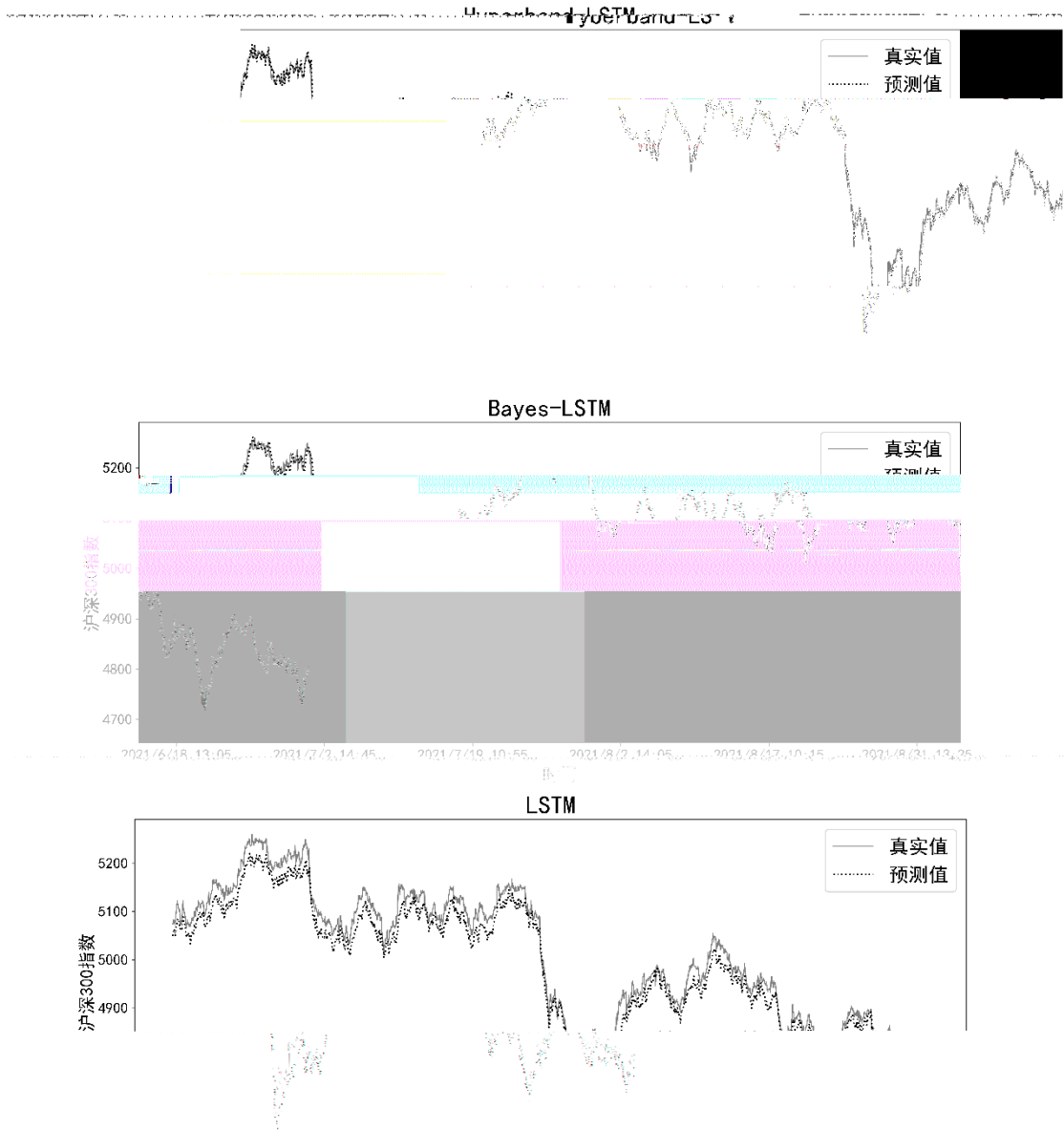


图5 预测效果图

使上，中出列分别两两，为  
 ，于，了。从，两  
 价不，值，以为三个  
 ，且与其两为，  
 价于其两个，  
 于其两个，一了力  
 升。

表5 DM检验结果-沪深300

		-	-	
-	--	-25.01***	-68.19***	
-	25.25***			



。 ，两个 不 列 中  
 为 ， 上 一 ，  
 于其 两 。 ， 一 了 健 。

表7 DM检验结果-上证50、深证100

指数
上 50
深 100

注：表中结果为 DM 统计量，\*\*\*、\*\*、\* 分别表示 1%、5%、10%的显著性水平。

#### 四、结语

先 前人 ， 前 价 中 到 一  
 使 。 中， 体上 优于其  
 ， 。 ， 也  
 中 不 ； 到 化  
 ； 为 低 。 出 与  
 于 价 中， 不 。 ，  
 ，从 上 两 其 于  
 价 列 上 。 ， 一 列  
 不 价 上 优于 ， 价  
 上 切 。 于 ， 出以下 ：  
 低 价 ； 价 ；  
 具 健 ， 于不 价 任务。

参考文献

习 动 化 中 业

:

党 从 于

与

: