

教育部高等学校大学计算机课程教学指导委员会

中国大学生计算机设计大赛



人工智能应用类作品文档简要要求

作品编号：_____

作品名称：_____ 基于HTM模型的股票智能分析平台 _____

作 者：_____ 王家宝、占圆梦、吴同旺 _____

版本编号：_____

填写日期：_____ 2022/3/10 _____

填写说明：

- 1、本文档适用于**所有**涉及软件开发的作品，包括：软件应用与开发、大数据、人工智能、物联网应用；
- 2、正文一律用五号宋体，一级标题为二号黑体，其他级别标题如有需要，可根据需要设置；
- 3、本文档为简要文档，不宜长篇大论简明扼要为上；
- 4、提交文档时，以 PDF 格式提交本文档；
- 5、本文档内容是正式参赛内容组成部分，务必真实填写。如不属实，将导致奖项等级降低甚至终止本作品参加比赛。

目 录

第一章 需求分析.....	1
1.1 开发背景.....	1
1.2 市场分析.....	1
1.3 作品简介.....	2
第二章 概要设计.....	3
2.1 模块设计.....	3
第三章 详细设计.....	4
3.1 界面设计.....	4
3.2 数据库设计.....	8
3.3 关键技术.....	8
3.4 技术创新.....	16
第四章 测试报告.....	18
4.1 前端样式测试.....	18
4.2 网页加载测试.....	18
4.3 实时获取数据测试.....	18
4.4 数据库连接验证测试.....	19
4.5 后端分配函数优化.....	19
第五章 安装及使用.....	20
第六章 项目总结.....	21
6.1 任务分配.....	21
6.2 面对困难.....	21
6.3 开发感悟.....	21
6.4 后续安排.....	21

第一章 需求分析

1.1 开发背景

1602年，股票伴随着东印度公司的成立而出现，多年以来一直备受金融行业人士的喜爱，并随着不断发展，成为一些人的生存之本。为了方便股民更加方便地炒股和分析股票，越来越多与股票有关的网站相继建立。随着股票行业不断走向大众化，繁琐复杂的分析方式不再适用于普通民众，人们迫切需要简单易懂且方便使用的分析平台以满足基本需求。目前大多数网站只提供基本技术指标图表和数据展示，且不提供在线分析平台，需要用户下载相应app，通过安装股票分析插件或者自己编写源码来分析股票，很少提供基于新手用户的相关建议和分析参考信息，本项目主要针对上述情况，致力于为普通民众提供友好的股票学习平台并给予相关决策意见。

1.2 市场分析

1.2.1 目标用户及用户需求

网站的主要目标用户群体是对股票感兴趣的普通民众以及炒股的股民。

目前，普通民众对于股票知识的缺乏了解，该网站设置了股票基础指标的说明介绍，普通民众可以通过该网站初步了解股票知识，从而进一步了解股票。

该网站除了基本的股票技术指标说明外，还设置了HTM模型分析预测窗口，帮助炒股的股民进行分析预测，从而做出更加慎重精准的操盘。

1.2.2 竞品分析

在对项目进行市场调查时发现一般的股票分析平台为应用app，需要用户下载安装，且设计繁琐，内容复杂，不方便普通民众使用且查看复杂，更甚者，一些基本的股票技术指标还需要用户自己编写源码计算。而该网站设计简约，内容丰富而不繁琐，不需要用户下载安装app，在线的设计方便普通民众浏览和股民分析使用。

本项目利用HTM层级实时记忆模块、Echarts图表、Django开发框架、tushare接口、mysql数据库等技术，实现用户在线登录浏览，实时更新数据等功能。

表 1.1 竞品分析表

应用名称	不需要安装下载	设计简约易懂	技术指标说明	源码设计
通达信	×	×	×	×
东方财富网	√	×	×	√

1.3 作品简介

1.3.1 主要功能

(1) 基本股票技术指标的说明

本网站为了方便普通民众的使用，在主页设置了基本股票技术指标的基本介绍、使用方法和操作心得，能够让刚刚接触股票的普通民众入手。

(2) 多维数据可视化

本网站的数据展示，除了基本的K线图外，还在一旁使用Echarts绘制了MACD, KDJ, RSI等趋势预测指标的图标，能够使用户更直观的了解股票趋势和更方便的分析股票。

(3) 提供实时有效的信息

本网站运用python的tushare接口实时获取万维网上的数据，让数据获取实时有效。

(4) HTM层级实时记忆模块分析预测功能

本网站使用HTM层级实时记忆模块，将近五年来股票的涨跌趋势进行了训练学习，并以此为基础，分析预测了近一个月的涨跌趋势，供用户参考分析。

1.3.2 流程图

首先用户进入首页，可通过点击更多查看相关股票技术指标的详细介绍说明；点击登录将会跳转到登录验证页面。在登陆验证页面，可通过返回键返回主页；可点击注册进入注册页面进行用户注册；可通过输入用户名和密码进行验证，验证成功则跳转到数据展示页面；注册页面可通过返回键返回登录页面；可通过home键返回首页，如图1.1所示。

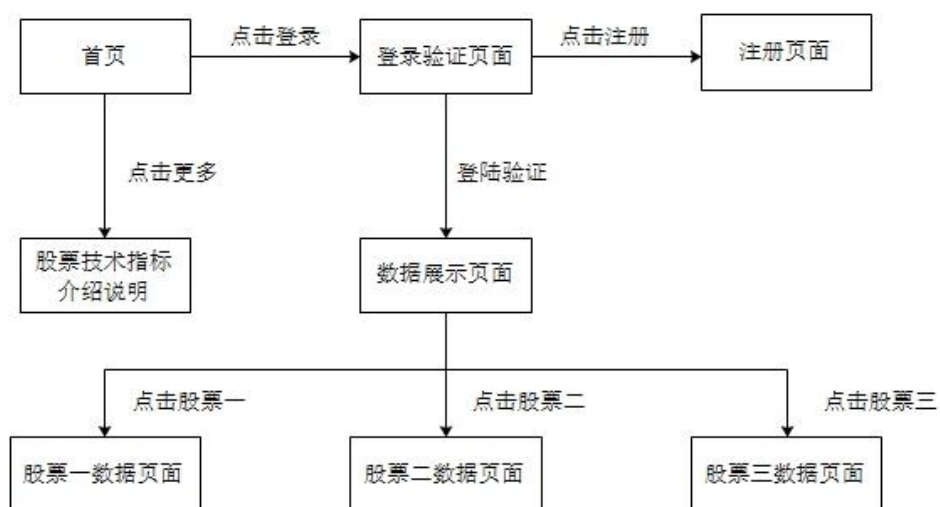


图1.1 页面浏览流程图

第二章 概要设计

2.1 模块设计

2.1.1 功能结构图

“基于HTM模块的股票智能分析平台”分为三个主要模块，分别为首页、登录注册页面以及数据展示页面。其中首页的功能包含主视图展示、股票技术指标介绍以及HTM模块技术简介；登录注册页面包含用户信息验证和用户注册；数据展示页面包含K线图、技术指标图表、HTM分析预测结果、使用介绍以及项目文档说明。其功能结构图如下图2.1所示。

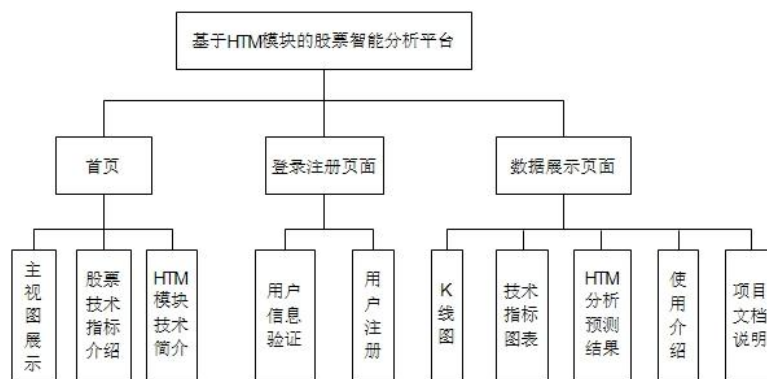


图2.1 功能结构图

2.1.2 模块结构图

本网站结构，首先Python从万维网上获取实时股票数据，在经过Python的数据处理和指标计算后，将HTM模块训练所用的数据保存在文件中供Java端读取，并将K线数据和技术指标数据部署到相关的json文件当中，Java端开始对HTM模型进行建立，并从之前Python操作的文件中获取训练数据和测试样本，开始对股票进行训练和分析预测，并将分析预测结果存放到Django服务器的相关文件中，最后用户通过浏览器查看网站时就可以看到我们所展示的数据展示页面。网站的各个模块结构可如下图2.2所示

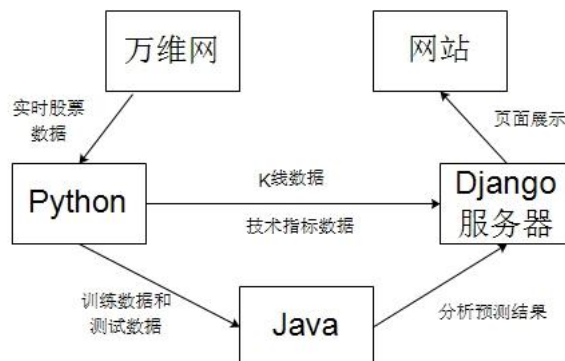


图2.2 模块结构图

第三章 详细设计

3.1 界面设计

3.1.1 首页

“基于HTM模块的股票智能分析平台”网站首页采用上中下布局。其上包含转接登录界面和主视图展示，如图3.1所示。其中主要用于简单介绍HTM模块的技术要点，如图3.2所示。其下为股票技术指标列表，用户可点击更多，查看相关技术指标的详细介绍，如图3.3所示。



图3.1 主页上部分展示图



图3.2 主页中部分展示图

股票十大指标详解

 MACD 异同移动平均线。是从双指数移动平均线发展而来的。由快的指数移动平均线..... 更多>>	 DMA 平行线差(DMA)指标是利用两条不同期间的平均线。来判断当前买卖能量的大小..... 更多>>
 RSI 相对强弱指标(RSI)是根据一定时期内上涨点数和下跌点数之和的比率制作出的一种技术曲线。 更多>>	 KDJ KDJ指标中文名叫随机指标。是一种相当新颖、实用的技术分析指标。它起先用于期货..... 更多>>
 VR VR(Volatility Volume Ratio)中文名称:成交量异率,主要的作用在于以成交量的角度测量和观察股价。..... 更多>>	 OBV OBV的英文全称是:On Balance Volume。是由美国的投资分析家Joe Granville所创..... 更多>>
 BRAR 布林带指标(BRAR)也称为人气意愿指标。其英文缩写亦可表示为ARBR..... 更多>>	 CCI 顺势指标又叫CCI指标。CCI指标是美国股市技术分析家霍华德·道伯特(Donald Lambert)于20世纪80年代提出的..... 更多>>
 PSY 心理线(PSY)指标是研究投资者对股市涨跌产生心理波动的情绪指标..... 更多>>	 BOLL 布林线(BOLL)指标是股市技术分析的常用工具之一。通过计算股价的“标准差”，再求股价的..... 更多>>

图3.3 主页下部分展示图

3.1.2 股票技术指标详细介绍页面

该网站设置了十个股票常用技术指标菜单，用户可根据自己的需求自行选择想要了解的信息并阅读，可通过“返回”按钮返回主页，如图3.4所示。



图3.4 股票技术指标详细介绍页面展示图

3.1.3 登录界面

用户可在该界面输入用户名和密码进入数据展示页面，可通过“注册”进入注册页面进行用户注册，可通过左上角的home键返回首页，可通过“忘记密码”进入修改密码页面，如图3.5所示。

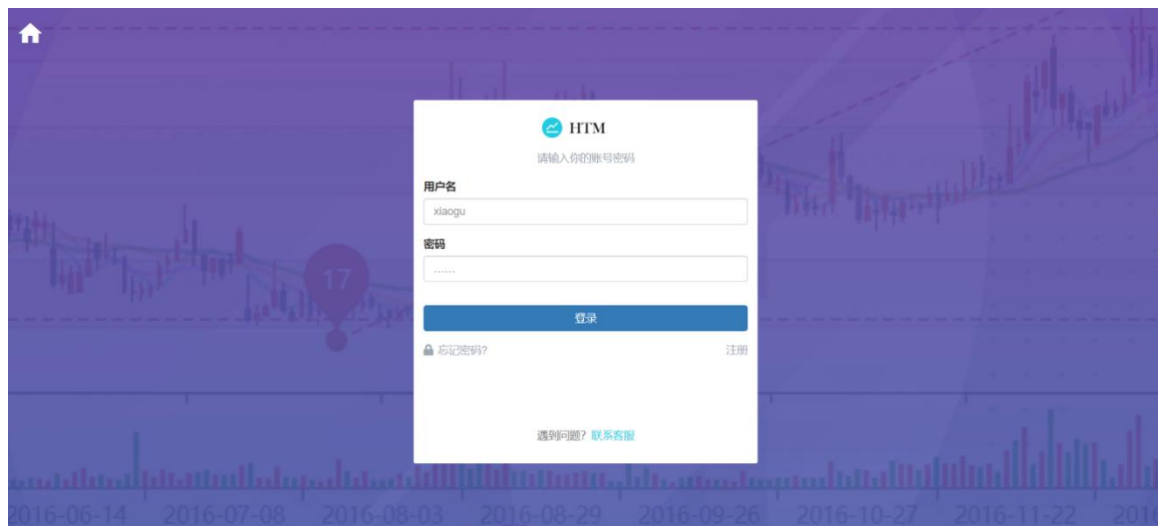


图3.5 登录页面展示图

3.1.4 注册页面

用户可在该页面输入用户名、密码以及邮箱(用于修改密码)来完成注册，可通过“登录”

返回登录界面，可通过左上角的home键返回首页，如图3.6所示。



图3.6 注册页面展示图

3.1.5 修改密码界面

用户可在该页面输入用户名和注册时设置的邮箱来重新设置密码，可通过“登录”来返回登录界面，可通过左上角的home键来返回首页，如图3.7所示。

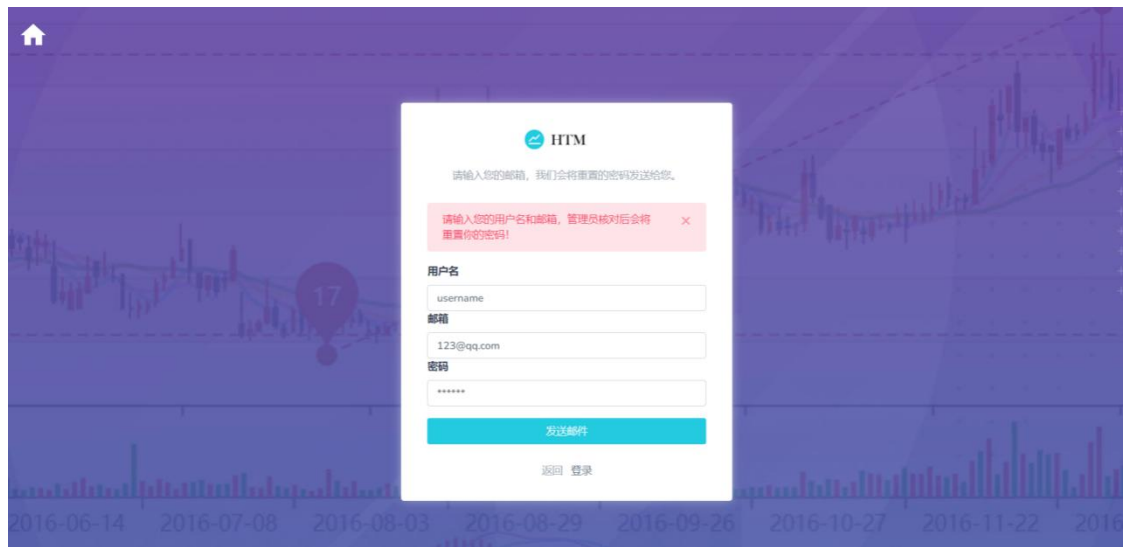


图3.7 修改密码页面展示图

3.1.6 数据展示页面

在通过登录验证后，用户即可进入数据展示页面，该页面展示了我们已经训练好的三支股票信息，用户可通过左上角的“股票一”、“股票二”、“股票三”来切换股票。其主要内容为K线图、股票技术指标图、分析预测比对、网页使用说明以及HTM模块简介，如图3.8所示。



图3.8 数据展示页面展示图

其中：

关于股票技术指标展示页面,为了方便用户使用,我们将十个指标图做了联动,当用户在查看MACD图时,日期为从2019/10/1-2019/12/1时,如果用户转为其它指标,展示的数据也是从2019/10/1-2019/12/1.

关于HTM模块分析预测比对结果,如果我们认为HTM模块预测信息不准确,我们将给出“我们认为结果不可靠”的提示信息,避免不准确的分析预测给用户带来误判。该模块的功能是将近一个月以来我们所作的所有预测结果与现实情况做比对,以展示HTM模块的预测效果。

3.1.7 预测结果展示页面

左下角运行结果框中的更多,将会跳转到预测结果展示页面,该页面记录了近一个月以来的所有预测信息。如图3.9所示。

日期	实际情况	预测结果	准确度	总数	涨数	跌数
2022-02-14	跌	预测准确度高	0	0	0	0
2022-02-15	涨	预测准确度高	0	0	0	0
2022-02-16	跌	预测准确度高	0.5	655	323	332
2022-02-17	跌	预测准确度高	0	0	0	0
2022-02-18	涨	预测准确度高	0.5	655	323	332
2022-02-21	涨	涨	0.6666666666666666	525	304	221
2022-02-22	跌	预测准确度高	0	0	0	0
2022-02-23	涨	预测准确度高	0.5	655	323	332
2022-02-24	跌	涨	0.7142857142857143	424	265	159
2022-02-25	涨	预测准确度高	0	0	0	0
2022-02-28	跌	预测准确度高	0.5	655	323	332
2022-03-01	涨	预测准确度高	0.5555555555555556	655	323	332
2022-03-02	涨	预测准确度高	0	0	0	0

图3.9 预测结果展示页面图

3.2 数据库设计

我们使用数据库来管理用户的信息，使用user_table表来存储用户 的用户名、密码以及邮箱。

- user_table表：存储用户验证信息

数据表名：user_table 中文描述：用户表						
字段名称	字段描述	数据类型	长度	是否允许为空	缺省值	备注
usernmae	用户名	varchar	20	否	NULL	主键
password	密码	varchar	16	否	NULL	
email	邮箱	varchar	20	否	NULL	

3.3 关键技术

3.3.1 HTM模块概述

层级实时记忆，简称HTM。

层级实时记忆是一种以捕捉新大脑皮层的结构与算法特性为目标的机器学习技术。使用HTM编程不同于传统的电脑编程。现今的程序员针对特定的问题编写特定的程序。相对的，HTM通过感知数据流来训练自身。HTM的机能很大程度上是由训练它的数据所决定的。从根本上来讲，HTM是一个基于记忆的系统，也可被视作一个神经网络。HTM网络被大量具有时间性的数据训练而成，并依赖于存储的大量模式序列。

3.3.2 原理

从四个问题可以讲解HTM的原理：HTM的区域是如何构成的，为什么层级组织结构是重要的，为什么数据要以稀疏离散表征的形式储存，为什么基于时间的信息是关键的。

1) HTM的区域是如何构成的

区域的概念来自于生物学。新大脑皮层是一张约2mm厚的巨大的神经细胞层。生物学家主要依据这些区域互相的连接方式来把新大脑皮层划分成不同的“区域”，是许多细胞以一定规律排列而成。HTM区域也是仿照着这样的构造，由一层成列的高度内联的细胞组成的，如图3.10所示。

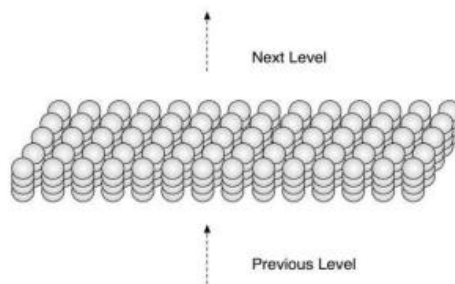


图3.10 HTM区域结构展示图

其中，一列细胞称为一个柱状区域。

2) 为什么层级组织结构是重要的

一个HTM网络是由按层级排列的区域所组成的。具有代表性的是每个区域代表层级中的一个等级。层级会随着等级的上升不断聚合，下级多个子元素将被聚合到一个上级元素中，如图3.11所示。

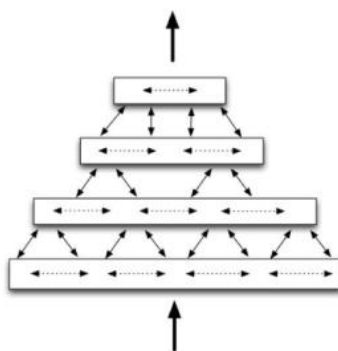


图3.11 HTM层级结构展示图

层级组织结构的好处就是高效率和重用。它显著减少了训练时间和使用的记忆空间，因为每一层级已学会的模式将以一种新奇的方式在更高的层级中得到重用。例如，你的大脑学会了边和角。这些低级的结构可以用来组成中级的结构，例如正方形和圆。再然后，可以用来组成立方体等三维图形。而在用低级的结构来组成更高级的结构时，你不需要再学习低级的结构。

3) 为什么数据要以稀疏离散表征的形式储存

虽然新大脑皮层中的神经元高度连通，抑制神经元保证同一时间只有很少一部分神经元处于活跃状态。因此，大脑中的信息总是通过巨大数量的神经细胞中的一小部分活跃细胞表示。这种编码方式就叫做“稀疏离散表征”。

4) 为什么基于时间的信息是关键

仅通过一个例子来讲解为什么时间的角色在HTM模块中是关键。

当我们听歌识曲时，我们只听一个音符是无法猜出歌名的，但是如果是一段连续的音符，如果我们很熟悉这首歌，我们便可以通过这段音乐猜出这首歌的歌名，并且可以唱出接下来的

部分。

这段音乐即是由一些音符信息按基于时间排列成的序列，才能够让我们识别出这段音乐，并给出分析结果。

3.3.3 基本概念

1) 细胞的状态

细胞的输出状态分为三种：由前馈输入激活、由横向输入激活以及非活跃状态。第一种输出状态相当于神经元中的动作电位的一个短脉冲。第二种输出状态相当于神经元中较缓慢、较平稳的动作电位。

2) 树突区域

每个HTM细胞有一个接近细胞核的树突和许多末端树突。靠近细胞中心的树突接收前馈输入，末端突触接收附件细胞的横向输入。用一个柱状区域共享的树突来代替该柱状区域中的细胞的近核树突。空间沉积池和时间沉积池(3.6和3.7介绍)分别在柱状区域的水平上操作共享的树突和每个细胞的水平上操作末端树突。

3) 突触

突触代表神经元的输出。HTM模型中包含两个概念：潜在突触和连通值。潜在突触值所有足够接近树突的轴突可能形成突触。连通值指树突与轴突之间的连通程度。当连通值达到一定的数值，我们才认为突触存在(有效)。

3.3.4 HTM的基本功能

HTM的四个基本功能为：学习、识别、预测以及行为。每一个HTM区域执行前的三项功能：学习、识别和预测。

1) 学习的特点：

更多的学习是发生在更高层级的区域当中。

2) 识别的特点：

当一个HTM学会了模式之后，它就可以对前所未见的输入进行识别。HTM会在接收输入的时候与以前学习过的空间实时模式进行匹配。新输入与旧序列的成功匹配时识别和模式匹配的关键。

3) 预测的特点：

(1) 预测时连续的。

(2) 预测发生在层级中的每一层。

- (3) 预测对上下文环境时敏感的。
- (4) 预测时区域稳定的。
- (5) 预测可以告诉我们一个新输入是否时期望的。
- (6) 预测有助于增强系统的鲁棒性。

3.3.5 HTM模块工作流程

一个HTM区域的输入由成千上万的二进制位组成，如何操作这些输入？

我们将从以下三点探讨：

1) 构建输入的稀疏离散表征

HTM区域所作的第一件事是把输入转化成新的表征。而输入的稀疏表征就是对哪些柱状区域活跃，哪些在抑制后不活跃的一个编码。抑制机制主要是为了实现将柱状区域的活跃百分比控制在一个相对稳定的范围内，即便是有大量变化显著的输入的时候，如图3.12所示。

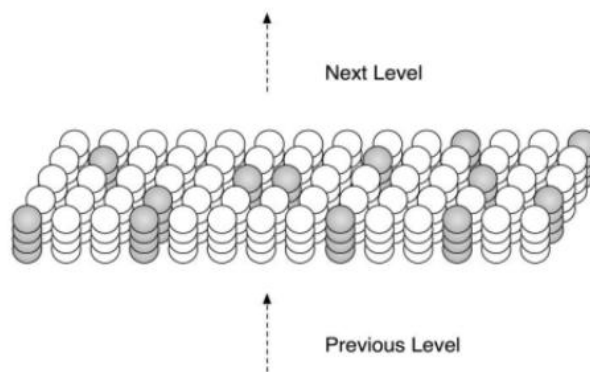


图3.12 输入的系数离散表征展示图

这一步的工作主要是由空间沉积池来完成。

2) 在有先前输入的环境中构建输入的代表

首先举一个例子：

当我们让HTM学会了abc和dbe。

情况一，当我们输入b时，HTM做出的预测将是c或者e；情况二，如果我们先输入a再输入b，HTM将会预测下一个将会c。

两种情况的差别在于是否再b之前输入a，而情况二能准确预测c的原因便是a为接下来的预测构建了表征。

而在HTM中，我们知道，在一个HTM区域当中，一个输入是由一些柱状区域来表示的，好比代表的数字4的柱状区域集合为{柱状区域1，柱状区域2，柱状区域5}，当输入4的时候，这些柱状区域的上处于预测状态(稍后将讲解)的细胞将会被激活，如果某个柱状区域上没有处于预

测状态的细胞，则激活该柱状区域上的所有细胞(表示HTM认为所有情况都有可能)，而这些被激活的细胞与横向上的一些细胞之间存在连通突触，这些细胞将会被激活。

但每个柱状区域是由一些细胞组成的，当{柱状区域1：细胞1，柱状区域2：细胞1，柱状区域5：细胞1}被激活和{柱状区域1：细胞2，柱状区域2：细胞2，柱状区域5：细胞2}两个都表示数字4，但这两种情况的上下文环境不相同，比如前者表示的是3后面的4，后者表示的是5后面的4。

在有先前输入的环境中构建输入的特征，即是根据当前激活细胞，来将横向的一些细胞变为预测状态。

3) 在有先前输入的环境中根据当前输入构建表征

一个区域怎样做出预测？当输入模式随着时间变化时，不同的柱状区域以及细胞的集合按着一定的顺序变得活跃。当一个细胞变得活跃起来，它会优先与附近马上能激活的细胞子集建立连接。当有激励从连接传来时，细胞会意识到它将要被激活并进入预测状态。因此一个细胞集合的前馈激励通常会导致另一个细胞集合的预测激励，如图3.13所示。

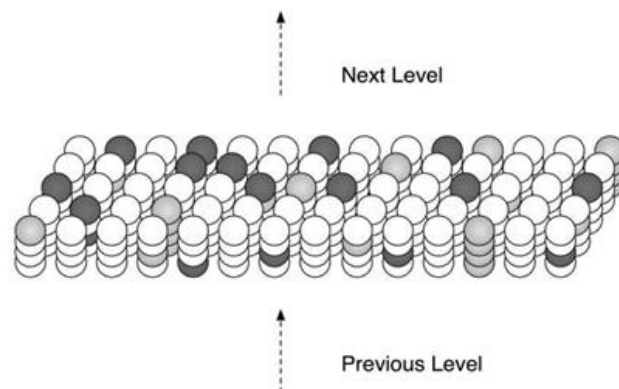


图3.13 激活细胞与横向细胞建立连接并将其激励为预测状态
(浅灰色为活跃状态，深灰色为预测状态)

我们用术语“时间沉积池”来表示为表征添加上下文环境和预测这两步骤。通过对模式序列构建变化缓慢的输出，我们实际上将时间上连续的不同模式“沉积”在一起。

3.3.6 沉积池的基本概念

空间沉积池和时间沉积池中的学习时相似的。两种学习都包括在细胞间建立连接或者说突触。空间沉积池负责输入与柱状区域的前馈连接，时间沉积池负责同一区域内细胞键的连接。

1) 二值权

HTM突触只有0和1两种结果。它们的“权值”是二元的，这一特性不同于其它神经网络模型用0到1的标量变量来表示。

2) 连通值

突触在学习的过程中不停地形成和分解。

我们为每一个突触设置一个标量（0.0到1.0）来反映连接的连通性。当一个连接得到加强时，它的连通值会增加。在相反的环境下连通值会减少。当连通值高于阈值（比如0.2）时，可以认为突触已建立。如果连通值低于阈值则突触失效。

3) 树突区域

突触连接到树突区域。有两种树突区域，靠近细胞体的和处于末端的。

靠近细胞体的树突与前馈输入构成突触。

末端树突区域与同区域内的其它细胞构成突触。每个细胞有多个末端树突区域。如果一个末端树突区域的突触活跃度之和超过了阈值，那么该细胞将会进入预测状态。由于每个细胞存在多个末端树突区域，所以细胞进入预测状态是几个胞内阈值探测器进行逻辑“或”操作的结果。

4) 潜在突触

每个树突区域有一份潜在突触的列表。所有的潜在突触被赋予了一个连通值并且在该值超过阈值时变成起作用的突触。

5) 学习

学习过程涉及树突区域的潜在突触的连通值的增加和减少。突触连通性增减的规则与“赫布（Hebbian）”定律相似。

举个例子，如果一个突触后的细胞因树突区域接收到的输入超过阈值而激活，那么这个树突区域中突触的连通值将被修改。

活跃的突触，并且因此而有助于细胞被激活的，它们的连通性会增加。

不活跃的突触，并且因此而对活跃无所贡献的，它们的连通性会下降。

3.3.7 空间沉积池

空间沉积池的主要功能是将一个区域的输入转换成稀疏模式。这一功能是十分重要的，因为在学习序列和做出预测的机制中要求从稀疏离散表征开始。

(1) 使用所有的柱状区域

一个HTM拥有固定数量的柱状区域用来学习如何表示输入中的共有模式。有一个原则是保证所有的柱状区域都学着去表示些有用的事物。如果一个柱状区域不是很活跃，它会变得更具“侵略性”。

(2) 维持一定的密度

一个区域需要对它的输入构建一个稀疏表征。柱状区域的大多数输入会抑制周围其它的柱状区域。而抑制半径与柱状区域的大小成比例，正因有抑制半径，我们只允许一定百分比的柱状区域在大多数输入中成为“赢家”。

(3) 避免琐碎模式

通过设置激活柱状区域的输入的最小阈值来实现所有柱状区域都表示输入中较重要的模式。

(4) 避免多余的链接

如果一个柱状区域构成了数量巨大的有效突触就会导致它许多不相关的输入模式都有强烈的响应。为了避免这一问题，我们减少对柱状区域当前活跃无贡献的突触的连通值。通过对无贡献的突触进行有效惩罚，我们可以保证一个柱状区域表示有限数量的输入模式，有时只有一个。

(5) 自适应的感受域

新大脑皮层的区域可以学会表示完全不同的事物并对各种变化做出反应。

如果新大脑皮层中的一部分被损坏，其它部分会调整自身并表示损坏部分所表示事物。如果一个感受器损坏或发生改变，新大脑皮层中与之相关联的部分会表示其它的事物（即放弃原来所表示的）。

这一系统称为自适应。

3.3.8 空间时间池

时间沉积池学习序列和做出预测的基本方法是，当细胞被激活时，与其先活跃起来的细胞建立连接。然后这些HTM细胞关注这些连接来预测他们何时会被激活。当所有的细胞都拥有这样的功能，那么它们便能做到储存序列和回忆序列。由此可见，HTM没有所谓的储存中枢，取而代之的是信息离散的存放在每一个细胞中，这样就让HTM系统有了很好的鲁棒性，即使单个细胞遭到破坏或者故障也不会影响整体的预测。

时间沉积池需要完成的工作：

(1) 对于每个活跃的柱状区域来说，查找柱状区域中处于预测状态的细胞并将其激活。如果没有细胞处于预测状态，则激活该柱状区域的所有细胞。结果得到的活跃细胞集合是当前输入在先前的输入环境下的表征。

(2) 对于每个细胞的每个树突区域突触来说，统计有多少已建立的突触连接到了活跃的细

胞。如果数量超过了阈值，那么该树突区域就会被标记为活跃。有活跃树突区域的细胞被置为预测状态。

(3) 当一个树突区域变得活跃时，修改与之相关的所有突触的连通值。对于活跃树突区域中的每个潜在的突触，增加那些与活跃细胞相连接的突触的连通值并减少那些与不活跃细胞相连的突触的连通值。突触连通值的这一变化被标记为临时的。

(4) 当一个细胞因前馈输入而从不活跃变得活跃，我们会遍历每个与该细胞相连的潜在突出并且去掉一切临时标记。

(5) 当一个细胞因前馈输入而从不活跃的时候，撤销该细胞的潜在突触中标记为临时的连通值变化。

3.3.9 tushare

本项目使用Python进行获取数据和数据处理，使用tushare第三方库来实时获取股票数据，将其与计划程序相结合，实时采集数据以确保数据的实时性和准确性。

3.3.10 Django

Django是一种基于Python语言的开源免费的Web应用开发框架。它定义了服务发布、路由映射、模板编程以及数据处理的一整套功能，具有极高的完整性。Django使用path函数和正则表达式来管理URL映射，因此给开发者带来极高的灵活性。

3.3.11 Echarts

Echarts是一个使用JavaScript实现的可视化库，能够提供直观、交互丰富、高度个性化定制的数据可视化图报表、Echarts支持折线图、散点图、地图、仪表盘等12类图标，利用它能够轻松实现商业级的数据图表，帮助用户对数据进行挖掘和整合。

3.3.12 Bootstrap

Bootstrap 是基于 HTML、CSS、JAVASCRIPT 的前端开发框架，它在jquery的基础上进行了更人性化和个性化的完善，其最大的特点就是可实现响应式布局，开发者可以方便地将网页部署到台式机，平板，手机等任何尺寸的终端，使得用户获得最佳地浏览体验。

3.3.13 JQuery

JQuery是一个便捷的，简洁的JavaScript框架，封装了Js常用的功能代码，提供一种简便的Js设计模式，优化Html文档操作、事件处理、动画设计和Ajax交互。它具有独特的链式语法

和短小清晰的多功能接口和高效灵活的CSS选择器，可对CSS选择器进行扩展；同时拥有便捷的插件扩展机制和丰富的插件。

3.4 技术创新

3.4.1 编码改进

HTM模块由空间沉积池和时间沉积池两个部分组成。其中空间沉积池所进行的工作是编码和构建系数离散表征，传统的HTM模型是不对数据进行预处理，而是编写一个编码器，直接将原始数据交给编码器处理，形成HTM模块的内部表征，我们对编码的过程进行了改进。改进如下：

对于股票数据的处理，为了让预测更加准确，我们已十天的涨跌情况为一个数据单元，以1表示涨，0表示跌。例如：0110000010。

然后对该序列进行扩展，0使用“10000000000”来替换。而1使用以下规则来替换：

原本序列中的第*i*位为1，则

`str = "0" * 10`

`str[i] = "1"`

`1 → "0" + str`

例如，第二个1则被替换成为“00100000000”

对每一位都进行扩展，10位的二进制序列则被扩展为一个110位的01序列，扩展后的序列即恰恰满足了空间沉积池的功能。

于是，我们在数据获取阶段，利用python直接将数据进行编码处理，省略HTM模块中的空间沉积池部分，不仅处理简单，易于理解，并且让HTM模块的学习效率获得提升。

3.4.2 数据处理

关于空间沉积池中的编码工作，在进行编码之前需要对数据项的取值范围和编码位数进行确定。而股价的数值往往是并不是在一个确切的区间以内的，并且取值范围较大，编码位数也是无法确定的。除此以外，影响股价的因素还有很多，比如通货膨胀，汇率，市场流动资金、市场情绪以及宏观经济发展，这样有可能会因为某些因素而导致股价浮动，导致已经训练好的HTM模型预测不准甚至无法使用的问题。

为了消除这些因素对股价带来的印象，我们将绝对股价转换为涨跌幅度，这样有利于增加HTM模型的安全性以及可持续性。并且中国股价的涨跌幅度上限为10%，取值范围小，这样可以确保HTM模型因为股价大幅动荡所带来的影响。

关于涨跌幅度的计算，我们记涨跌幅度为 $extent$ ，当天的收盘价为 $close_{today}$ ，前一天的收盘价为 $close_{yesterday}$ ，以如下式子计算涨跌幅度：

$$extent = \frac{close_{today} - close_{yesterday}}{close_{yesterday}} \times 100\%$$

为了提高预测精度，我们还将 $extent$ 的计算结果进行乘10处理，将取值范围扩大为 $[-100, 100]$ ，以确保更加精准的预测。

3.4.3 精准度改进

HTM模型的预测结果，不止一个结果，而是预测所有可能的结果，而每一个结果在HTM模型的内部表征表现为一些柱状区域。我们将所有的结果的柱状区域存放在一个集合当中，记这个集合为 $result_list$ ，而当前输入的所激活的柱状区域的集合记为 $input_list$ ，记 $len(list)$ 为集合 $list$ 中元素的个数。对于准度 $quasi_degree$ ，我们定义其计算公式为：

$$quasi_degree = \frac{len(result_list \cap input_list)}{len(input_list)}$$

而对于精度 $accuracy$ ，我们定义其计算公式为：

$$accuracy = \frac{len(result_list \cap input_list)}{len(result_list)}$$

由此可见，当预测的准度高时，精度不一定高。即虽然HTM模型预测到了接下来可能发生的情况，但这种情况只是众多情况中的一种，这并不代表HTM模型给出了既定的未来。因此我们需要对预测结果进行评价。

本次项目中，经过多次调试，为了达到更加理想的预测结果，设定每种情况激活的柱状区域个数为10，设定 $len(result_list \cap input_list) \geq 7$ 为准度达标条件，即准度到达0.7。并在此基础上，设定精度 $accuracy \geq 0.6$ 为达标条件。当两项评测指标都达标时，再给出预测分析结果，避免用户做出错误的决策。

3.4.4 结果改进

HTM模型的预测结果是一个柱状区域的集合，对于操作员和用户来说都是可读性极低的。为了更加清晰简单的描述预测结果，我们将预测结果的柱状区域集合进行处理。

本项目中的数据项为一个十位二进制数，并且用1和0分别表示涨和跌，为了更加清晰地表述结果，我们将十位二进制数转换为一个十进制数，取值范围 $[0, 1023]$ 。我们将预测的所有可能结果全部转换为一个十进制数，而预测出来的结果，为明天开始往后十天的所有情况，而所需要的数据只是十位二进制数中的最高位。

当所转换得到的二进制数，大于等于512即是第二天股价会涨，而小于512则是第二天股价会跌。我们将所有预测的结果进行统计，并计算其中的预测为涨和跌的个数，并在网页中展示，给予用户更多的决策参考。

第四章 测试报告

4.1 前端样式测试

技术指标:

运行速度: 平均0.05s

可用性: 良→优

安全性: 优

测试以及修正过程:

在最初的测试中, 由于股票K线图和技术指标的数据量太大, 将其直接输入到html文件中并不现实。在学习了Django后, 我们使用json替换的方式, 将数据写到json文件中, 并将其与html文件中的字段进行替换, 完成了大量数据写入html文件的操作。

4.2 网页加载测试

技术指标:

运行速度: 平均0.2s

可用性: 良→优

安全性: 优

测试以及修正过程:

在测试中, 技术指标图各自独立, 这样给用户再查看指标时需要反复调试, 体验不佳。我们将其修改成了图标联动, 保证所有图表的数据显示同步, 在没有降低运行速度的前提下提升了用户体验。

4.3 实时数据获取测试

技术指标:

运行速度: 平均1.1s(包含数据处理)

可用性: 良→优

安全性: 良→优

测试以及修正过程:

本项目实时爬取通过设置任务计划程序定时启动 python 运行脚本。再做数据处理时出现过java端训练HTM模型的数据与实际日期不同步的情况, 经过处理后, 将错误解决。

4.4 数据库连接验证测试

技术指标：

运行速度：平均0.02s

可用性：良→优

安全性：优

测试以及修正过程：

连接验证无错误，响应良好。

4.5 后端分配函数优化

技术指标：

运行速度：平均0.11 s→平均0.04s

可用性：良→优

安全性：优

测试以及修正过程：

后台程序包含将json文件中的数据替换到html文件的数据集合，原本包含数据处理，后将数据处理移交到数据获取的预处理工作中，提升了运行速度。

第五章 安装及使用

开发时使用的 Windows 10 操作系统，硬件环境是处理器：Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz，内存：7.8GB，磁盘：KBG40ZNS128G NVMe TOSHIBA 128GB。Python使用Anaconda的Spyder运行环境，需要安装tushare第三方库、django第三方库，使用Sql server管理mysql数据库，Python使用pymysql库进行数据库操作，HTM模块的编程和训练使用Java语言，使用Eclipse编程软件。

后端流程：使用Python实时获取股票数据，获取数据后使用Java进行HTM模型的建立和训练学习，并给出分析预测结果，并将结果存储到相应文件，用户在通过浏览器访问平台查看分析预测结果和相关股票数据。

系统运行时，Python Django需要一直处于运行状态。

第六章 项目总结

6.1 任务分配

在任务开始前，我们就根据小组成员的又是特长进行了明确的分工。占圆梦同学具有很好的Php编程和网页制作的工地，审美能力不俗且心细专注，主要负责网站前端的页面设计。实现以及数据可视化。吴同旺同学的逻辑思维能力较强，对数据的敏感度较高，主要负责股票数据的获取和对数据的预处理。王家宝同学对于关键技术了解程度最深，以及具备一些Web应用的基础，因此，HTM模型编程和训练，Django框架编程以及文稿的大部分撰写由他负责。

在项目进行过程中，我们各司其职，秉持着重交流的工作方式，即时向组长汇报工作进度，以及商讨如何解决困难，无形中也起到了相互监督的作用，推动了项目的稳步发展。

6.2 面对困难

在整个网站开发过程中，我们也遇到了不少难题。例如对于股票技术知识的缺乏，而该项目开发成果更涉及用户的资金变动和商务决策，我们花费了大量的时间用于股票技术知识的学习，避免因开发人员的知识匮乏而导致产品成果的不正确和不专业。除此之外，我们组的成员来自不同专业，对于对方专业的知识了解甚少，特别是负责前端开发的同学不了解后端业务，负责后端业务的同学不了解前端开发，为了避免让前端开发和后端业务更好的衔接，我们相互讨论，学习对方的长处经验，并在努力后很好地实现了前端和后端的链接。

6.3 开发感悟

“基于HTM模型的股票智能分析平台”网站是我们夜以继日的心血成果。我们经历了它的每一次功能拓展、每一次报错调试，伴随了它从无到有的整个成长过程，也在培育它的过程中的收获颇多。

通过本次项目开发，我们将理论应用于实践，发现了自身存在的不足以及能力的匮乏，让我们深刻体验到与跻身于领域前线不断开发的专业人员之间差距，我们将不断精进自身能力，力求成为更加专业的开发人员。除此以外，我们通过本次项目也学习到了本专业以外的知识，以及一些金融的基础知识，受益良多。

6.4 后续安排

未来我们计划在以下几个方面继续改进

- 精进技术

一个训练完成的HTM股票模型并不一定适用于所有股票，为了让分析预测的结果更加准确，需要对每一支股票都建立一个全新的HTM模型。当前的HTM模型的训练参数仍需人为调整，我们力求在后续发展中让HTM模型实现自我调整，寻找最好的训练参数。

- 面向大众推广

当前我们的网站只展示了三支股票，原因在于从数据获取到HTM模型训练学习需要花费大量的时间，为了让用户有更好的浏览体验，我们目前提供个股的分析预测结果。往后，在精进技术的前提下，我们力求将实现由用户自主选择股票进行分析预测。

- 与相关单位合作

本网站的开发人员对于股票知识终究不及该领域的专业人员，我们希望能够与金融领域的人员合作，更加完备专业知识，使得网站更加功能完善以贴近用户的需求。