对于第四问的求解，我们使用了在论文*DeepTrader: A Deep Reinforcement Learning Approach for Risk-Return Balanced Portfolio Management with Market Conditions Embedding*  中提出的模型DeepTrader

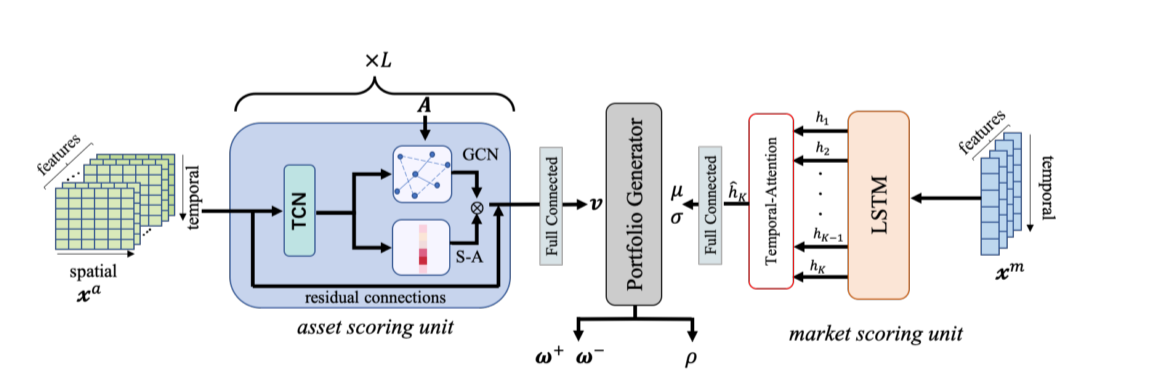
如下图所示，DeepTrader主要由三个部分组成：

1.Asset scoring unit。输入为股票指数 和构建的图结构 A，输出为赢家得分

2.Market scoring unit。输入为市场标准 ，输出一个高斯分布的参数（均值和方差）。

3.Portfolio generator。输入为赢家得分和高斯分布 ，输出为做空总资产比例

和投资组合 和。



对于中特估股票，我们将股票指数定义为中特估指标

在时间维度上的矩阵

将市场标准 定义为中特估指标

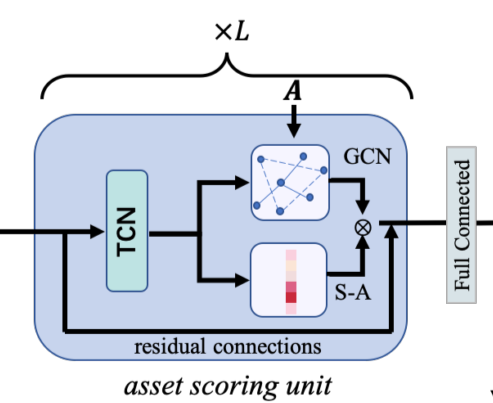
在时间维度上的矩阵

相应的，将LSTM层和TCN层接受的特征数量分别设置为6个，15个。

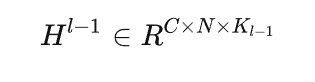
下面介绍网络中各个模型结构

1.Asset Scoring Unit

该单元由 L 个带有残差连接的 Spatial-TCN 块堆叠而成。Spatial-TCN 块主要由三部分组成：时间卷积层、空间注意力机制和 图卷积层。

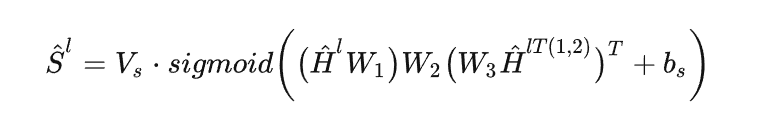


我们把第 块（block）的输入记作

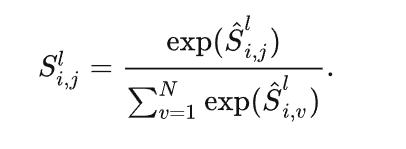


用TCN提取的股票时间特征分别被输入到spatial attention和GCN两个组件中产生股票短期相关性权重和长期相关性权重两者相乘作为股票的最终空间相关性权重

为了模拟股票之间的短期空间属性。得到TCN输出的股票时间特征表示后，输入到空间注意力机制中用来产生一个权重，这个权重代表股票间的相关性关系:



接下来再把每个元素用softmax归一化，用来代表股票i和j的相关性



图卷积层用来对**股票间的长期空间关系**进行建模。

虽然个股的表现具有变化的波动性，但**行业（industry）**的整体表现通常更能反映未来的经济形势热点。这里用了图卷积网络GCN，通过消息传递来获取图中节点的依赖关系，将边和节点的信息集成到表示中。

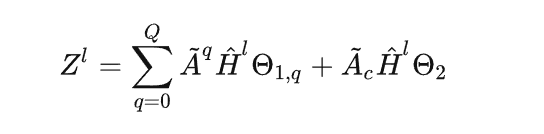
这里关于股票的**图结构**的构建，原作者实验了以下几种方式：

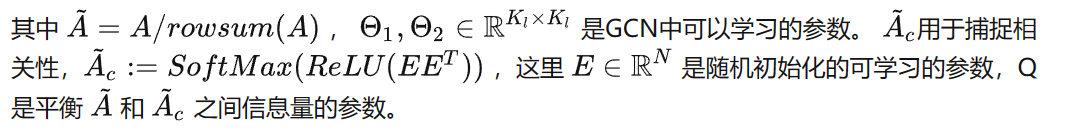
* 股票行业分类
* 股票收益的相关性
* 股票收益的偏相关性
* 因果关系（通过PC算法识别股票之间的因果结构）

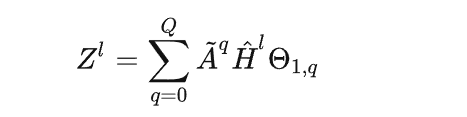
仍然是用TCN的输出作为输入

(2)对于使用股票行业分类方式构建的图结构：

由于只使用行业分类信息可能会忽略某些依赖关系，为了避免这个问题，这里GCN使用复杂一点的结构：

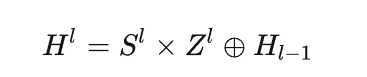




对于上面说的用后三种方式（收益相关性、偏相关性和因果关系）构建的图结构来说，使用

这样得到了股票长期相关性权重

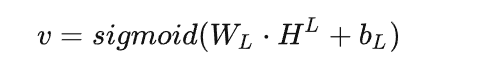
使用残差网络结构，第个block的输出为



为残差连接

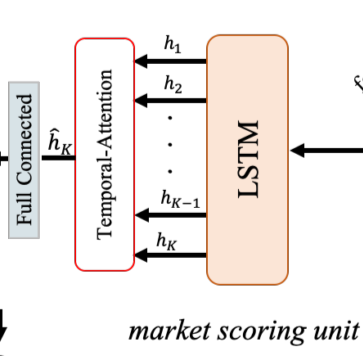
将这个block结构堆叠L次获得第L块的输出 就是最后的股票的时空特征

将使用一个全连接层提取特征，得到的最后股票的涨跌潜力打分



2 Market scoring unit

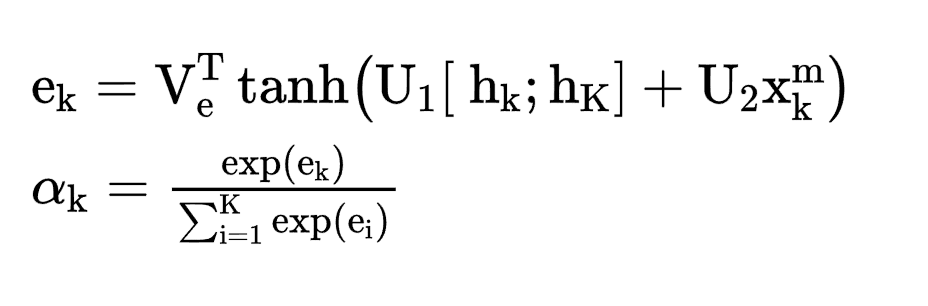
金融数据包含大量不可预测的不确定性。根据历史观察来准确判断中特估股票的涨跌是不可行的。在以往的投资模型中，投资策略仅仅是基于对每支股票的分析，而忽略了市场的变化。而顺应市场是一个更好的投资策略。当股市下跌时，有经验的投资者倾向于在卖空上花更多的钱。为了平衡收益和风险，本文提出了市场评分单元。以中特估特征指标为输入，动态调整做空资金的占比。



使用 LSTM + attention 来提取输入数据的序列化表示。计算过程如下：

为第k步的状态编码。

然后使用attention加权捕获信息关系



再次计算记为

计算出正态分布的均值和方差

3.Portfolio generator

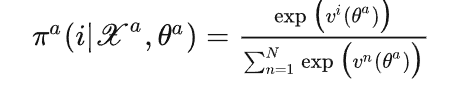
该模块选取前 G 只股票做多，后 G 只股票做空，分配股票占比计算如下：



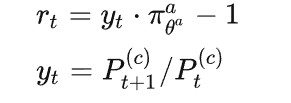
4. Optimization via RL

以上过程用RL来优化，策略由选股分配组合权重和空仓比例两部分组成。

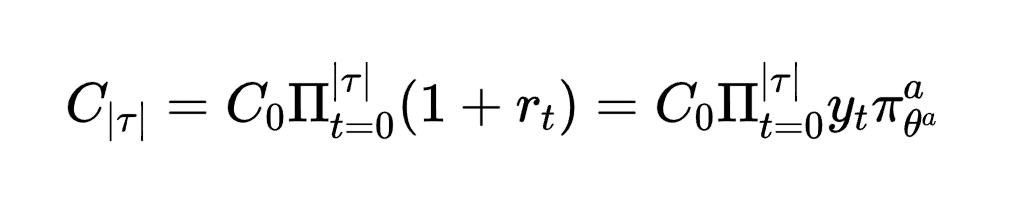
选股分配组合权重：



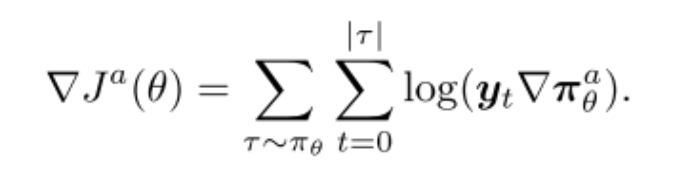
return rate：



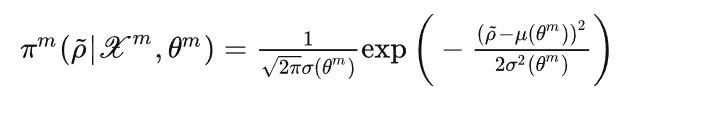
初始投资金额是 ，一个轨迹的累积金额：



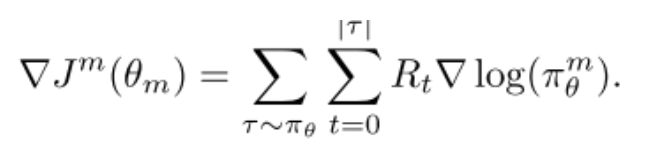
股票评分单元的优化目标为迹的对数累积财富最大化：



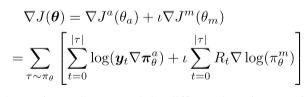
市场打分的策略使用高斯策略



给定reward ， 优化目标为



最终的优化目标：



使用梯度下降算法更新

评估：由于团队短时间内缺乏相关可用数据集以及训练模型可用算力，我们无法将该模型应用在中特估股票模型上。

下列为原文中使用CSI100数据集，DJIA数据集，HSI数据集对模型所作的评估。

