

基于可见性图嵌入的沪深 300 深度学习增强策略

2023 年 10 月 13 日

➤ **可见性图可以刻画量价时序性结构特征。**可见性图是一种将时间序列数据转化成一张图结构的方法。抽取可见图的结构特征并训练分类器的方法被广泛用于时间序列数据的分类问题。可见图转化为复杂网络后，复杂网络的结构包含了时间序列中局部或者全局的波动率与趋势信息，多篇研究利用可见性图及其特征进行时序数据的预测并取得了显著效果。简单的可见性图结构特征可以构造弱有效的选股因子，日度 IC 均值 0.03。

➤ **利用结构向量法提取图结构特征，可以刻画股票的潜在属性。**尽管人为定义计算的可见性图结构特征已经被证实过对时序数据有一定的分类效果，我们选择用一种无监督学习的方式来抽取复杂网络的图结构特征——结构向量法。与传统的图嵌入方法相比，struc2vec 更加高效，且在多个数据集上分类更加精确，并可以应用于大规模图的嵌入。用结构向量法提取的图结构特征对沪深 300 的股票进行聚类，在多种风格因子上有显著分类效果，在周期行业上区分效果相对不明显。

➤ **利用二阶段循环神经网络与跨资产注意力网络构造预测模块，从而纳入潜在属性及股票间的相互关系。**两阶段注意力循环神经网络 DA-RNN 可以通过按照时间顺序上不断训练结构性信息的模型结构，提取图结构嵌入特征，在生成股票因子表示时纳入股票的潜在信息。跨资产注意力网络 CAAN 可以对股票间的相互关系进行建模，计算 batch 内因子表示的相似度，作为注意力机制的来源。最终通过 sigmoid 函数，以下一个交易日股票价格是否上涨为因变量进行建模，输出股票上涨概率。利用混淆矩阵来衡量模型准确率，验证集内准确率 68%，精确率达 70%。

➤ **日频模型选股和择时效果显著，合成股指信号依然有效。**将模型预测的个股日度上涨概率作为日频选股因子 struc_learning，因子日度 IC 均值 0.16，ICIR1.2；中性化后因子表现有所下降，但选股效果依然显著。日度调仓多头组合年化超额收益 95.6%，信息比率 8.0，但策略容量较低。增加持股数量构建沪深 300 增强组合，组合平均持股数量 137 只，策略容量显著上升，年化超额收益 46.7%，信息比率 5.87。用个股信号合成股指日度涨跌概率信号，并构建沪深 300 仓位择时策略，年化超额收益 13.3%，信息比率 1.47，预测涨跌准确率 64%。

➤ **风险提示：**量化模型基于历史数据，市场未来可能发生变化，策略模型有失效可能。



分析师 叶尔乐

执业证书：S0100522110002

邮箱：yeerle@mszq.com

相关研究

1. 量化专题报告：冷门猎手：权益基金提前布局能力分析-2023/10/12
2. 量化专题报告：驱动创新的资本力量：创新效率中的超额收益-2023/10/12
3. 量化专题报告：从传统策略到深度学习的可转债投资-2023/10/10
4. 资产配置月报：十月配置视点：能源板块当下的配置价值几何？-2023/10/08
5. 量化周报：景气度跳升，节后有望反弹-2023/10/08

目录

1 可见性图的构造以及图结构嵌入	3
1.1 可见性图.....	3
1.2 复杂网络的节点权重	5
1.3 从图结构中提取嵌入信息.....	5
2 模型结构与开箱分析	8
2.1 基于图结构嵌入的深度神经网络	8
2.2 DA-RNN 网络.....	9
2.3 CAAN 网络.....	10
2.4 图结构特征对股票的分类效果	11
3 模型效果实证	13
3.1 参数选择及准确率衡量	13
3.2 模型选股表现.....	14
3.3 模型择时表现.....	17
4 总结与思考	22
5 风险提示	23
6 参考文献	24
插图目录	25

1 可见性图的构造以及图结构嵌入

1.1 可见性图

可见性图 (Visibility Graph) 是一种将时间序列数据转化成图结构的方法。

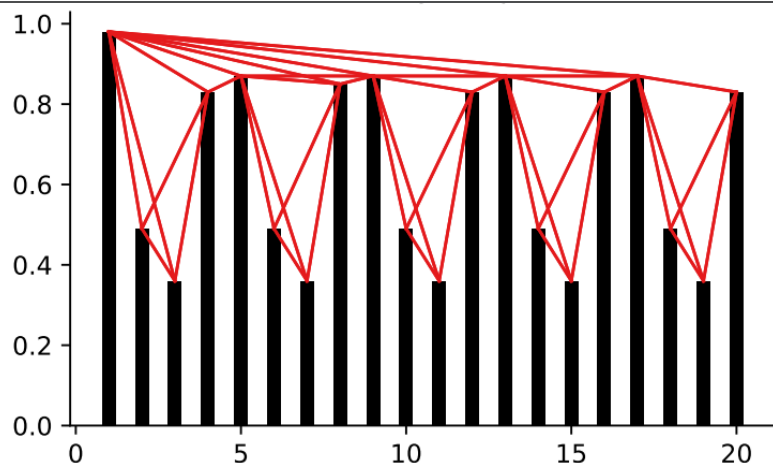
抽取可见图的结构特征并训练分类器的方法被广泛用于时间序列数据的分类问题。

可见性图的定义如下：对于一个时间序列 $P = (p_1, \dots, p_n)$ ，其可见图表示为：给定两个数据点 (t_i, p_i) 和 (t_j, p_j) ，若其中的所有数据点 (t_k, p_k) ， $t_i < t_k < t_j$ 满足：

$$p_k < p_i + \frac{t_k - t_i}{t_j - t_i} (p_j - p_i)$$

，则两点的连线视为一条“边” (edge)，时序中所有符合条件的数据点以及其“边”的集合即为这一时序的可见图。用更直观的角度解释：若将一个时间序列表达为一个柱状图，依次站在柱子的顶端“往前看”，如果可以“看到”任意一条柱子的顶端，则在这两条柱子间连接一条线，形成具有网络结构的边。

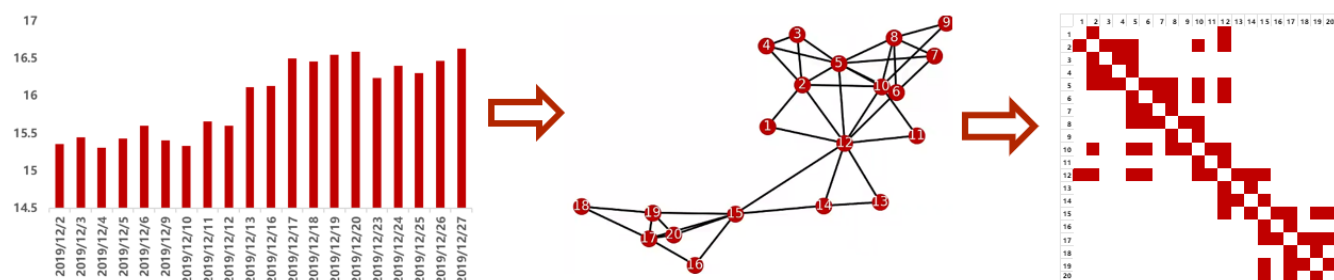
图1：可见性图连接规则示意



资料来源：Daoyuan Li et al. 2018 [1]

可见图方法通过结构图直观地展示时间序列映射后的复杂网络，此时的复杂网络继承了时间序列的属性。使用可见图方法转换时间序列为复杂网络的方法最早可以追溯到 2008 年 Lucas Lacasa 等人的论文 From time series to complex networks: The visibility graph [1]，我们可以把时间序列数据中的时间信息转换为复杂网络的节点位置信息，时间序列中的可见性关系转换为复杂网络中节点与节点间的连线关系，这种可见性的连接关系同样刻画了时间序列数据中数值的相对大小关系。但不同于时间序列的两个维度 t (时间)和 y (数值)维度，可见图的最终表示形式则为一个 $n \times n$ 的矩阵 Γ ，若点 i 与点 j 有连线，则矩阵中的元素 $\gamma_{ij} = 1$ ，否则为 0。下图展示了某一只股票在 2019 年底的 20 天收盘价构造的可见性图，可视化复杂网络，及矩阵化表示（红色为 1，白色为 0）。

图2：某只股票 20 天收盘价的可见性图及复杂网络

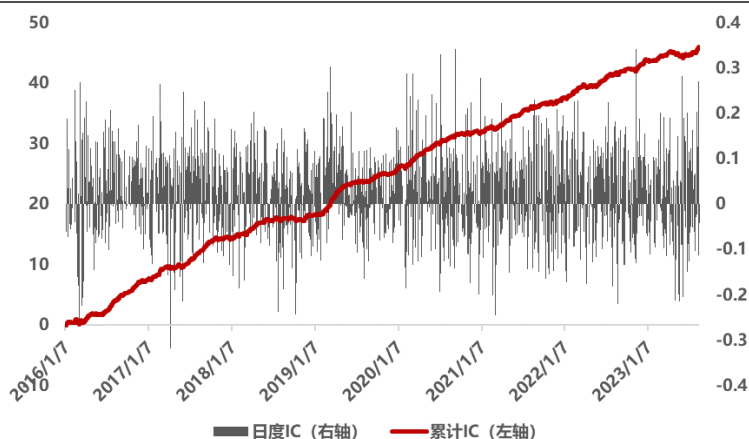


资料来源：Wind，民生证券研究院

可见性图及其特征被广泛应用于时序数据的预测中。比如，在 Daoyuan Li, et al 在 2018 年发表的论文 Extracting Statistical Graph Features for Accurate and Efficient Time Series Classification [3]中，作者将可见性图中提取出的多个特征输入 XGBoost 模型，对 UCR 大学创建的时间序列数据集中的多个时间序列数据进行了分类，并取得了显著的分类效果，证明了可见性图的结构特征可以对样本进行分类。从这一角度出发，我们可以将可见性图结构特征的分类效果应用到股票市场中，将可见图结构特征作为股票的潜在属性，输入到神经网络中，提升模型的预测效果。

在 A 股中，可见性图中的结构性特征对股票有一定的选择效果，其逻辑在于结构中蕴含的波动性与趋势。最简单且常见的图结构特征为平均最短步长，即每个节点连到其他点最短步长的平均。想象一种 U 型价格走势，按照可见性图的规则，图中的所有价格点都可以“看见”彼此，则类似的图结构有着最低的平均最短步长（一步就可以从任一点到任一点）；反之，若价格走势为倒 U 型，则图结构有着最高的平均最短步长。根据类似图结构中包含的波动率与趋势信息，我们可以构建简单的选股因子。例如，我们可以基于股票过去 20 个交易日收盘价可见性图的平均最短步长减去负收盘价可见性图的平均最短步长，构建日频选股因子，在全 A 上取得 0.03 的日度 IC。

图3：可见性图结构因子日度 IC



资料来源：Wind，民生证券研究院

1.2 复杂网络的节点权重

除了每个网络的图结构特征外，我们还需要考虑每一个网络节点的权重，因为每个节点蕴含的信息权重是不一样的。在传统的复杂网络分析框架中，有很多方法可以用来衡量节点权重，如度中心性，介数中心性，接近度中心性等。我们采用 Flaviano Morone and Hernan A. Makse 于 2015 年发表的论文 *Influence maximization in complex networks through optimal percolation* 中提到的 Collective Influence [3] 算法，根据 Morone 等人的研究，CI 算法计算高效，并且相比起传统权重计算方法，可以更好识别网络结构中的节点贡献。在 CI 算法中，要计算点 i 的 CI 指数，首先从点 i 出发，跳 l 步可以到达的所有点的集合（即距离点 i 的最短步数为 l 的所有点）构成以点 i 为中心，半径为 $l=2$ 的“球面”，表示为 $\partial Ball(i, l)$ 。则每个节点的 CI 指数计算方式定义如下：

$$CI_i = (d_i - 1) \sum_{j \in \partial Ball(i, l)} (d_j - 1)$$

其中 d 为节点的度数，即节点连的边数。我们以某只股票 20 天的量价指标为例，计算节点权重如下：

图4：某只股票 20 天量价指标的相应节点权重示例

close	30	100	168	50	50	77	90	80	48	72	27	30	29	68	130	16	40	33	24	8
vol	216	50	69	50	256	85	205	207	56	56	28	50	120	120	120	140	128	140	63	63
high	63	92	56	84	144	56	56	76	85	50	24	31	42	148	30	40	45	14	18	12
open	136	144	147	120	136	77	140	120	162	228	164	104	164	50	46	190	11	21	16	16
low	133	140	126	126	11	88	116	116	152	130	130	132	132	64	40	14	4	9	9	4
amount	216	50	69	50	256	85	205	207	56	56	28	50	120	120	120	140	128	140	63	63

资料来源：Wind，民生证券研究院

1.3 从图结构中提取嵌入信息

上一节中我们提到，传统的复杂网络用人为定义的结构特征，如平均最短步长，对时序数据进行分类。尽管这种传统的方法已经被证实过对时序数据有一定的分类效果，我们选择用一种更有效的训练方式来抽取复杂网络的图结构特征——结构向量法 (Struc2vec)，此方法由 2017 年 Leonardo F. R. Ribeiro 等人发表的研究 *struc2vec: Learning Node Representations from Structural Identity* 中提出[4]。结构向量法是一种将复杂网络中的节点表示为嵌入向量，并用其进行分类的无监督学习算法。与传统的图嵌入方法相比，struc2vec 更加高效，且在多个数据集上分类更加精确，并可以应用于大规模图的嵌入。需要注意的是，1.2 中提到的 CI 指数与结构向量法是两个独立的算法，CI 指数衡量的是复杂网络中每个节点的信息权重，而结构向量法则是从这些节点中提取潜在特征，将节点信息变为嵌入向量，2 种算法作用不同，并无关联。在后续模型构建中，

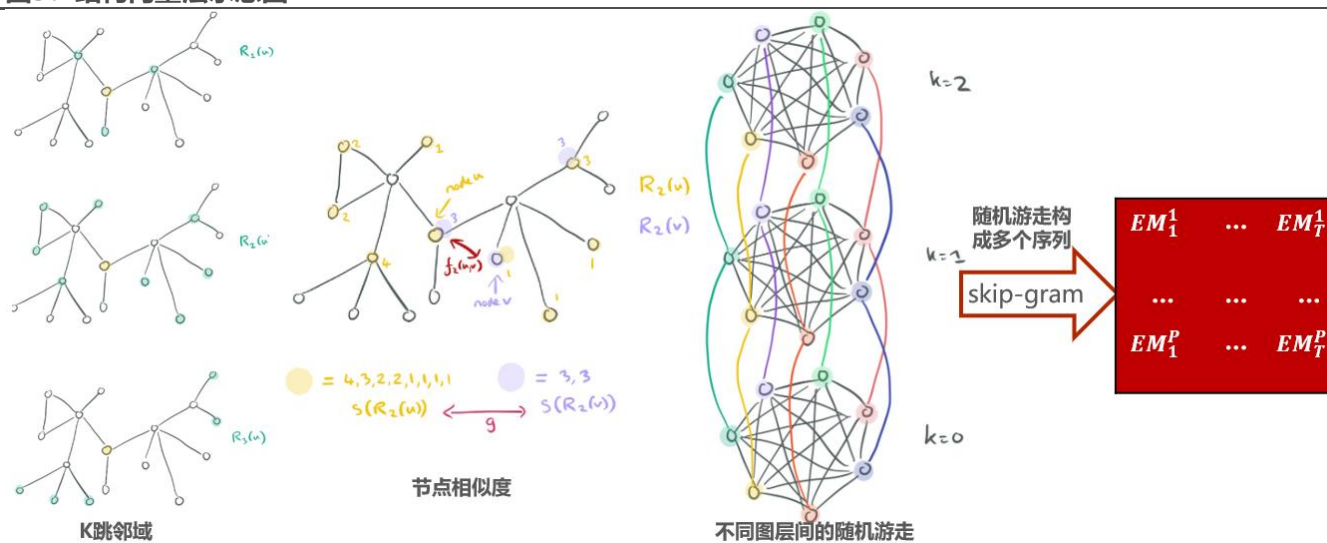
我们计算可见性图后将分别计算 CI 节点权重与结构向量嵌入，并输入到神经网络中进行预测。

Ribeiro 等人认为，一个好的可以反映节点结构特性的方法需要使得嵌入 embedding 向量间的距离可以反映节点之间的结构相似性，即两个局部结构相似节点的嵌入距离应该相近；并且节点表示不依赖于节点本身的属性，边的属性以及节点标签，即使两个节点离的很远，也能够识别出他们在结构上的相似性。基于上述目的，作者提出了结构向量算法，将每一个节点表示为嵌入向量。核心思想是：先根据节点相似度构建多个图层，在其中进行多次随机游走并构建上下文环境，再将上下文环境套入 Word2vec 算法中的 skip-gram 模型，学习每个节点的向量表示。skip-gram 模型是一种用于生成词向量的神经网络模型。它的目标是通过学习上下文信息来捕捉单词之间的语义关系。具体而言，skip-gram 模型试图预测给定一个中心单词周围出现的其他单词。举个例子来说，假设我们有一个句子：“北京的天空很蓝”。在 skip-gram 模型中，我们将选择一个中心单词比如“天空”，并尝试预测它周围出现的其他单词。通过这样的训练过程，skip-gram 模型可以学习到每个单词在不同上下文中的表示，对应到复杂网络中，即可以学习到每个节点在随机游走构建的上下文环境中的向量表示。

结构向量法的大致计算步骤如下：

1. 根据不同距离的相邻节点信息，对 K^* 个 k 跳邻域分别计算每两个节点的相似度，即计算 K^* 个层级的相似度信息，其中 K^* 为网络的直径，即网络中任意两点距离的最大值。计算相似度的核心思想是，如果两个节点的度相同，那么这两个节点结构相似；如果这两个节点的邻居度也相同，那么这两个节点的结构相似性比前者更高。图 5 中左边展示了一个网络中点 u （黄色点）的 1, 2, 3 跳邻域 $R_1(u)$, $R_2(u)$ 和 $R_3(u)$ （即绿色点的集合），以及节点 u 和 v 的相似度算法示意，其中 g 代表了 $R_2(u)$ 和 $R_2(v)$ 的距离函数，具体算法在本文中不再赘述。
2. 构建一个带权重的多层图，层数为 K ，每个图层中都包含网络中的所有节点，且每两点间均有边，边的权重与以这一层的层数 k 计算的两点的结构相似度有关，两个点结构相似度越高，这条边的权重越高。
3. 使用有偏的随机游走生成每一个节点的路径：对于每个节点，从基层开始进行随机游走，节点的每一次游走可以走到同一层的其他点，或是上下层的同一点，概率与步骤 2 中的边权重有关。图 5 中展示了 3 个图层的情况下随机游走的路径示意。
4. 根据多次随机游走获得的多个序列，将序列类比为上下文环境，使用 Skip-Gram 算法来生成节点表示。我们选择为 20 个节点的网络中的每个节点生成 $P=32$ 的向量表示。

图5：结构向量法示意图



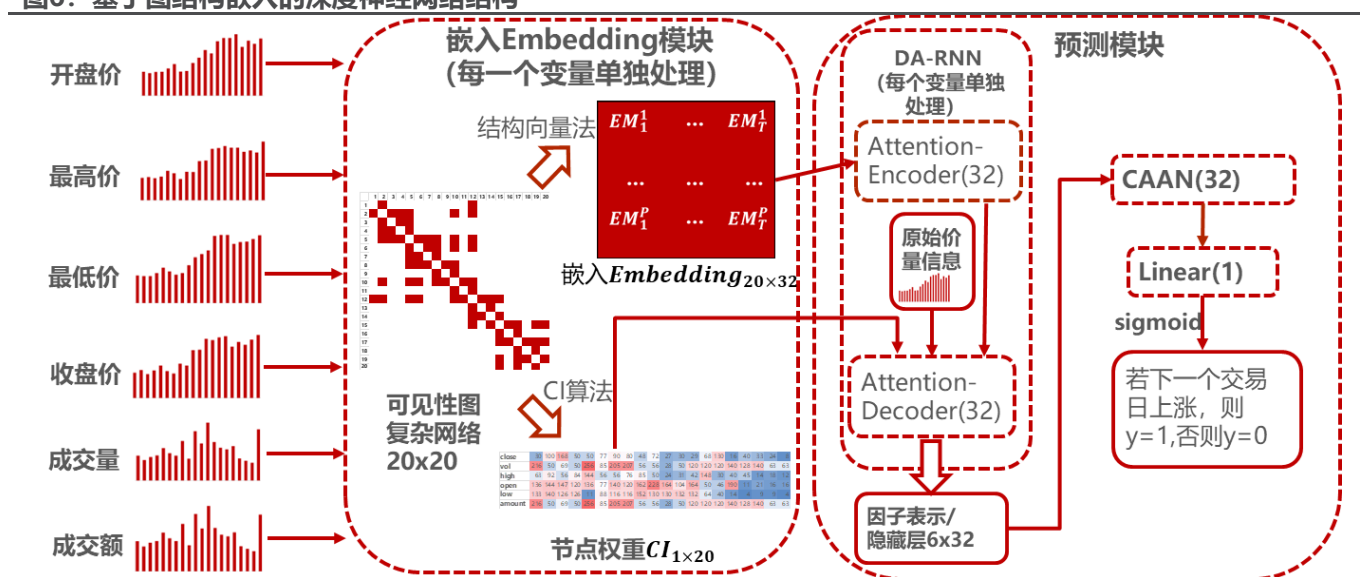
资料来源：Adrian Colyer, 2017 [5], 民生证券研究院

2 模型结构与开箱分析

2.1 基于图结构嵌入的深度学习神经网络

基于可见性图的图结构嵌入可以应用到深度神经网络中。我们在上一篇深度学习的专题报告《深度学习如何利用公募持仓网络优化选股效果？》中，论证了神经网络中加入股票潜在属性嵌入的可行性，提出了用股票动态的潜在属性影响因素表示并预测股票收益的方法（即将 embedding 与 LSTM 输出的隐藏层一起拼接到 MLP 中）。在本篇报告中，我们参考 Wu, J. et al., 2022 [6]中提出的图结构嵌入的神经网络预测模型，采用类似的思路：我们在 1.3 中提出了用结构向量法构建每一个可见性图的嵌入表示，以此来表示股票的潜在属性。接下来，我们在预测模块中将图结构嵌入输入到深度神经网络中，进行股票涨跌预测。总体模型结构如下：

图6：基于图结构嵌入的深度学习神经网络结构



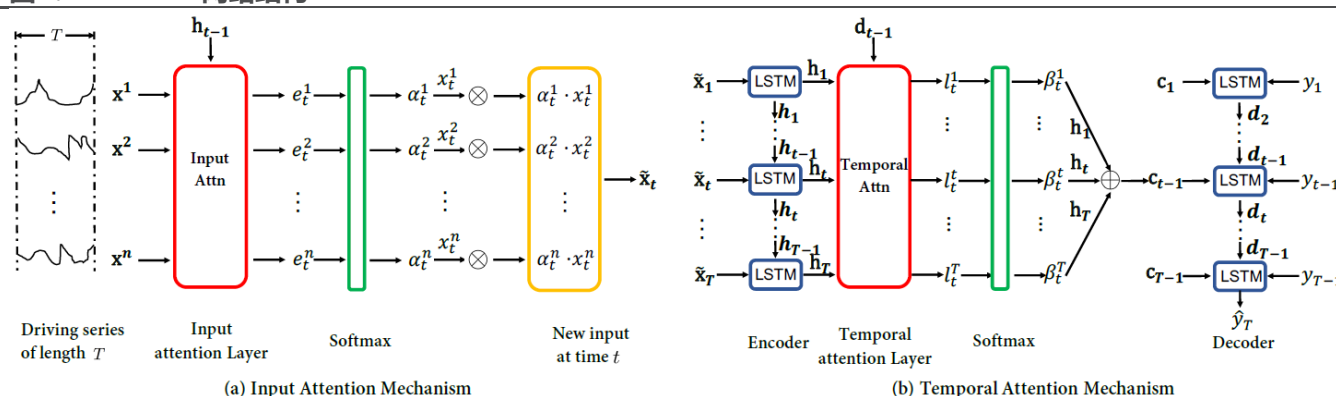
资料来源：民生证券研究院绘制

在上图的网络结构中，我们对于每一个样本，即单天单只股票的单个量价指标过去 20 天的数据，计算可见性图；随后分别计算节点权重 CI 以及结构向量嵌入矩阵。我们的预测模块由 2 部分构成，第一部分是 **DA-RNN**，一个双步的深度神经网络，来对嵌入信息进行编码，并将编码后的股票潜在属性与原始量价信息和节点权重一起输入到 DA-RNN 的第二步中，输出股票个股当天单变量的因子表示，再将 6 个价量指标的隐藏层相加，输入到**跨资产注意力网络 CAAN** 中，最后通过一个**全链接神经网络**，即上图中的 Linear(1)，以及 **sigmoid 函数**输出一个 0-1 的预测变量，即个股在下一个交易日的上涨概率。下面我们将介绍预测模块中用到的 DA-RNN 网络和 CAAN 网络。

2.2 DA-RNN 网络

两阶段注意力循环神经网络 Dual-stage Attention-based RNN，即 DA-RNN，可以通过按照时间顺序上不断训练结构性信息，用双步的模型结构提取图结构嵌入特征并输入到特征处理中，影响原始价量数据的建模。DA-RNN 由 Yao Qin 等人在 2017 年发表的论文 *A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction* [7] 中提出。DA-RNN 是一个双阶段的注意力机制循环神经网络模型，主要思想就是将前 t 个时间步的信息利用编码器 encoder 编码到一个融合吸纳前 t 个时间步信息的定长向量，然后使用解码器 decoder 进行解码对后续的时间步骤进行预测。与传统的将 embedding 特征与隐藏层进行拼接不同，DA-RNN 用 Encoder-Decoder 的结构纳入嵌入信息，充分利用了之前提取的结构特征。在 Encoder 中，我们在每一个时间步输入图嵌入以用一个 LSTM 提取隐藏特征，并将上一个时间步的隐藏特征输入到这个时间步，并使用输入注意力机制来选择相关特征（对应下图中的 Input Attention 机制），赋予每个特征一定权重；随后我们在模型的第二步 Decoder 中加入原始量价数据以及 CI 算法计算的节点权重，并加入时间注意力机制来选择整个时间步长中相应的隐藏层状态（对应下图中的 Temporal Attention 机制），并计算最终的隐状态，即因子表示。

图7：DA-RNN 网络结构



资料来源：Yao Qin et al. 2017 [7]

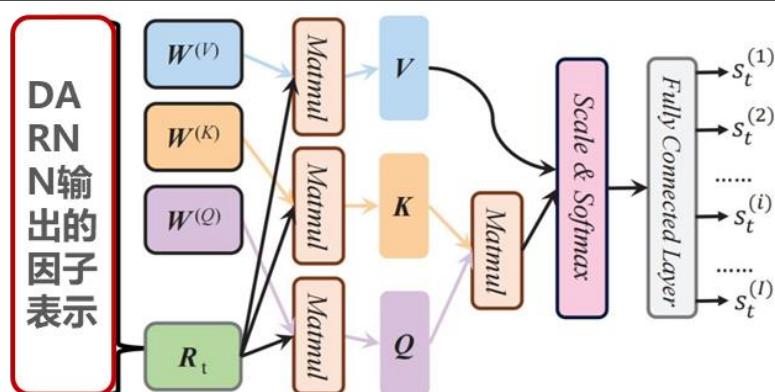
图 7(a)中展示了 DA-RNN 中特有的输入注意力机制编码器。首先构建一个输入注意力层以输出初始隐藏层 h_0 ，随后在每一个时间步上，使用当前时刻的输入 x_t 以及上一个时刻编码器的隐藏层 h_{t-1} 输入 LSTM，并更新当前时间步上的隐状态 $h_t = f_1(h_{t-1}, x_t)$ ，这里 $n = 32$ ，即为结构向量法中的特征维度。随后，把第 k 个特征与前一时刻的隐状态与细胞状态进行线性组合，再用 \tanh 激活得到 $e_t^k, k = 1, \dots, n$ ，将 e_t^k 用 softmax 函数进行归一化便得到了 t 时刻的注意力权重 $\alpha_t^k = \frac{\exp(e_t^k)}{\sum_{i=1}^n \exp(e_t^i)}$ ，最后更新下一步 t 时刻的输入 $\tilde{x}_t = (\alpha_t^1 e_t^1, \dots, \alpha_t^n e_t^n)$ 。**输入注意力机制使得编码器能够关注其中输入特征中重要的特征，而不是对所有特征一视同仁，这也是所有注意力机制的本质作用。**

图 7(b)中介绍了第二阶段的解码器注意力机制。需要注意的是，图 7(b)中已经包含了(a)中的过程，而图 7(a)可以看作是得到 h_t 与 \tilde{x}_t 的过程细节。第二阶段解码器中采用的 Temporal Attention 类似于传统的注意力机制，在不同时刻采用不同的背景向量 (context vector) 以达到只选取相关时刻的隐藏层的目的，使得解码器能够关注时序中重要的隐状态，而不是对每一天的隐状态一视同仁。其中背景向量 $c_t = \sum_{t'=1}^T \beta_t^{t'} h_{t'}$ 表示为所有时刻隐藏层的加权平均，权重 $\beta_t^{t'}$ 即为注意力权重。对于注意力权重 $\beta_t^{t'}$ 的计算，作者基于前一个时刻解码器的隐状态 d_{t-1} 与细胞状态进行线性组合，再用 tanh 激活得到 $l_t^{t'}$, $t' = 1, \dots, T$ ，将 $l_t^{t'}$ 用 softmax 函数进行归一化便得到了 t 时刻的注意力权重 $\beta_t^{t'} = \frac{\exp(l_t^{t'})}{\sum_{j=1}^T \exp(l_t^j)}$ ，更新下一步的背景向量 $c_t = \sum_{t'=1}^T \beta_t^{t'} h_{t'}$ ，最终更新当前 t 时刻解码器的隐状态 $d_t = f_2(y_{t-1}, d_{t-1}, c_{t-1})$ 。而当前时刻的隐状态 d_t 即为我们模型中输出的最终因子表示。

2.3 CAAN 网络

跨资产注意力网络 CAAN(Cross-Asset Attention Network) 可以对股票间的相互关系进行建模，作为注意力机制的来源。CAAN 最早由 Jingyuan Wang 等人于 2019 年发表的论文：AlphaStock: A Buying-Winners-and-Selling-Losers Investment Strategy using Interpretable Deep Reinforcement Attention Networks [8]中提出。传统的 self-attention 注意力机制着重学习输入特征在时序上相关性较高的部分，而跨资产注意力网络则以股票间因子的相似度作为注意力机制的来源。我们将上一步中得到的因子表示输入到 CAAN 中，给定资产表征 r ，我们计算出股票 i 的查询向量 q 、密钥向量 k 和价值向量 v ， $q^{(i)} = W^{(Q)}r^{(i)}$, $k^{(i)} = W^{(K)}r^{(i)}$, $v^{(i)} = W^{(V)}r^{(i)}$ 。其中 $W^{(Q)}$, $W^{(K)}$ 和 $W^{(V)}$ 是待学习的参数矩阵。资产 j 与资产 i 的相互关系被建模为使用资产 i 的 $q^{(i)}$ 来查询资产 j 的密钥 $k^{(j)}$ ，即 $q^{(i)}$ 与 $k^{(j)}$ 的内积 β_{ij} ，然后，将相互关系 $\{\beta_{ij}\}$ 作为权重，将其他资产的值向量 $\{v^{(j)}\}$ 加起来，成为一个归一化加权得分 $a^{(i)} = \frac{\exp(\beta_{ii})}{\sum_{j'=1}^I \exp(\beta_{ij'})} \cdot v^{(j)}$ ，作为个股的得分。最后用全连接层，即下图中的 Fully Connected Layer (Linear(1)) 通过 sigmoid 函数将得分转化为个股的上涨概率，即 0-1 的标准化输出。

图8: CAAN 网络结构示意图



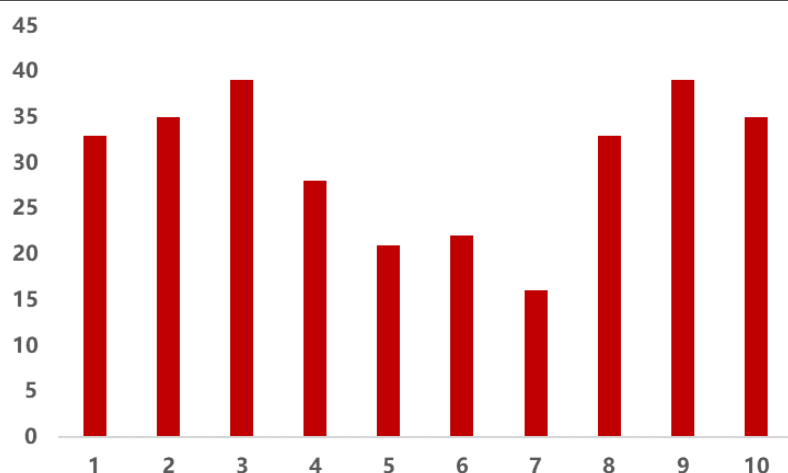
资料来源: Jingyuan Wang et al. 2019 [8], 民生证券研究院

整个模型的训练分为两部分。我们将嵌入模块与预测模块分开训练，先得到可见性图，随后分别计算节点 CI 权重，和训练结构向量法模型得到节点嵌入 embedding 信息。预测模块包含了上文中提到的 DA-RNN 与 CAAN，即按照模型结构进行端到端的训练，输入为原始量价信息，以及嵌入模块中得到的节点权重 CI 与节点嵌入 embedding，最终输出股票的上涨概率。

2.4 图结构特征对股票的分类效果

结构向量法提取的图结构特征刻画了股票的潜在属性，因而对股票有一定的分类效果。在师野等作者于 2022 年发表的论文《基于可见图方法的股票行业分析》中，作者用收盘价的平均度，聚类系数，紧密中心性等特征对美国道琼斯指数中的 30 只股票进行了聚类，得到了近似于行业的分类结果。这表明相似属性的股票理应具有相似的量价图结构特征。我们取 2019 年年底截面上沪深 300 的所有股票，用训练好的模型提取出 DA-RNN 编码器的输出，即个股的 32 维隐状态。这个隐状态表示了股票在这一截面上的“潜在属性”。潜在属性应是无法直接量化的属性，它可以是多种风格和行业属性的叠加，或者是更深层次，暂时无从发掘的股票属性。我们对每支股票的 6 个价量指标输出的隐状态取平均，并用 k-means 法对 300 支股票进行聚类，聚类结果如下：

图9：k-means 法聚类各类别股票个数



资料来源：Wind，民生证券研究院

我们将沪深 300 内的股票聚为 10 类，可以看出 10 类的数量分布没有明显不均。我们进一步计算每一类股票相对沪深 300 的主动风格偏离，解析图结构特征对股票的分类效果：

图10: 10 类股票的风格偏离

类别	beta	盈利	成长	杠杆	流动性	动量	非线性市值	市值	价值	波动率
1	-0.51	0.09	-0.22	0.12	-0.48	-0.28	0.12	-0.09	0.42	0.02
2	0.40	0.25	0.13	-0.01	0.35	0.41	0.00	0.02	-0.14	0.37
3	0.11	-0.34	0.29	-0.37	0.25	0.77	-0.24	0.22	-0.73	0.34
4	-0.27	0.31	-0.05	0.18	-0.22	-0.12	-0.01	0.03	0.29	0.16
5	-0.83	-0.25	-0.08	0.17	-0.43	-0.36	-0.12	0.09	0.03	-0.24
6	-0.38	0.41	0.12	0.05	0.15	-0.32	0.05	-0.02	-0.10	0.13
7	-0.36	-0.49	0.46	-0.80	0.00	0.06	0.12	-0.14	-0.93	0.53
8	0.44	0.24	-0.24	0.03	0.09	0.17	-0.03	0.03	0.13	-0.32
9	0.38	-0.16	-0.11	-0.04	0.20	-0.14	0.15	-0.09	-0.01	-0.21
10	0.16	-0.14	-0.08	0.39	-0.16	-0.56	0.01	-0.12	0.65	-0.51
标准差	0.42	0.29	0.21	0.32	0.27	0.38	0.11	0.11	0.46	0.32

资料来源: Wind, 民生证券研究院

从结果可以看出, 每一类股票都有较为明显的多种风格偏离, 且不尽相同。类别 1 以高贝塔, 低流动性, 低动量, 高价值为主; 类别 2 以高贝塔, 高流动性, 高动量, 低价值为主.....以此类推, 发现每一类股票都有较为明显的多种风格。纵向来看, 在 beta, 动量, 价值等风格因子上, 10 类股票的标准差较大, 证明在这些风格上, 结构特征对股票的区分度较高, 或者说股票在这几种风格上对图结构特征较为敏感。

图11: 10 类股票的行业分布

类别	汽车	银行	基础化工	通信	计算机	建筑	国防军工	电子	电力设备 及新能源	食品饮料	电力及公 用事业	机械	传媒	医药	有色金属	石油石化	能源	房地产	证券	交通运输
1	1	5	1	0	1	3	1	2	2	0	1	0	1	3	0	4	1	0	0	2
2	2	2	0	0	0	0	1	1	0	4	3	2	0	0	6	1	0	2	0	0
3	2	1	1	4	2	0	0	5	1	3	0	0	2	8	0	0	0	0	0	5
4	2	1	2	0	4	4	0	0	0	0	0	0	0	2	1	2	2	2	1	0
5	0	0	1	0	0	0	0	0	0	0	0	0	0	2	1	0	1	1	8	0
6	1	1	0	1	0	0	1	0	0	0	0	1	0	2	1	0	0	5	0	4
7	0	0	1	0	0	0	1	1	0	1	0	0	2	2	1	0	0	2	0	0
8	1	2	1	0	0	1	1	1	2	0	0	3	1	2	2	0	1	2	2	2
9	0	6	0	2	2	1	3	3	0	1	3	1	0	3	1	0	2	0	1	2
10	2	2	0	0	1	1	0	2	1	0	1	0	2	1	0	0	2	1	11	1
标准差	0.83	1.90	0.64	1.27	1.26	1.34	0.87	1.50	0.80	1.37	1.17	1.00	0.87	2.01	1.68	1.27	0.83	1.43	3.72	1.69

资料来源: Wind, 民生证券研究院

观察行业在各个类别上的分布, 可以看到图结构特征在证券, 医药, 银行等行业的区分效果比较明显。而在一些周期性行业如基础化工, 国防军工, 能源, 汽车等行业上, 图结构特征的分类效果相对较弱。总的来说, 可以认为图结构特征根据股票的行业, 风格和潜在属性, 对股票进行分类, 从而在 DA-RNN 解码器中发挥作用, 输出更加动态, 包含更多股票潜在属性的因子表示。

3 模型效果实证

3.1 参数选择及准确率衡量

模型数据输入与参数选择具体如下表。大多数参数选择均与参考论文中披露的一致。建模时，我们选择了沪深 300 成分股作为股票池，主要考虑到沪深 300 的流动性较高，以及结构向量法的中间结果较大，无法一次训练过多数据等原因。训练采用 80G 内存，RTX4090 显卡的服务器。Struc2vec 训练单个量价指标中间结果需要 10 个小时左右，预测模块训练单次 epoch 需要 5 分钟左右，整个训练过程需要运行 3 天左右。故全历史训练单次需要一周左右。

图12：模型数据与参数细节

- **数据输入：**2010-2018 年沪深 300 成分股的日频量价数据，高开低收+成交量与成交额。
- **数据窗口：**个股过去 20 个交易日量价，图结构特征及 CI 节点权重，预测下一个交易日收盘价是否上涨。
- **数据处理：**对于输入到 DA-RNN 解码器的原始量价数据及 CI 节点权重做 z-score 标准化。
- **参数选择：**在 CI 算法中，取 $l=2$ ，在结构向量法中，取嵌入特征个数为 32 个，进行 80 次随机游走，每次随机游走进行 10 步，Skip-Gram 窗口取 3；
- **模型训练：**神经网络最大训练轮数为 800 轮，最小为 50 轮；每一年训练一次，预测下一年的个股日度涨跌。样本外从 2019 年开始，每次训练采用九年数据，即已 2010-2018 年的训练模型预测 2019 年的日度涨跌，2011-2019 年的训练模型预测 2020 年日度涨跌，以此类推，单次训练的训练集与验证集为打乱的 8: 2；损失函数取二元交叉熵 (Binary Cross Entropy)。

资料来源：民生证券研究院

我们采用混淆矩阵来衡量模型预测表现。混淆矩阵是衡量分类型模型准确度中最简单，直观的计算方法。它分别统计了分类模型归错类，归对类的观测值个数。我们取下一日股票上涨为 1，即 positive，下跌为 0，即 negative。本质上，模型在最后全链接层加入 sigmoid，决定了模型的预测值为下一个交易日个股的上涨概率，故若预测值大于 0.5，我们就认为个股会上涨，预测值取 1，反之则取 0，故混淆矩阵如下图所示：

图13：混淆矩阵的四分类

混淆矩阵	预测值为1 (上涨)	预测值为0 (下跌)
真实值为1 (上涨)	TP(True Positive) 预测上涨，真实上涨	FN(False Negative) 预测下跌，真实上涨
真实值为0 (下跌)	FP(False Positive) 预测上涨，真实下跌	TN(True Negative) 预测下跌，真实下跌

资料来源：民生证券研究院绘制

我们采用混淆矩阵中常用的 4 个指标来衡量模型的准确度。4 个指标的计算方式如下：

$$1. \text{ 准确率: } Accuracy = \frac{TN+TP}{TN+TP+FN+FP}$$

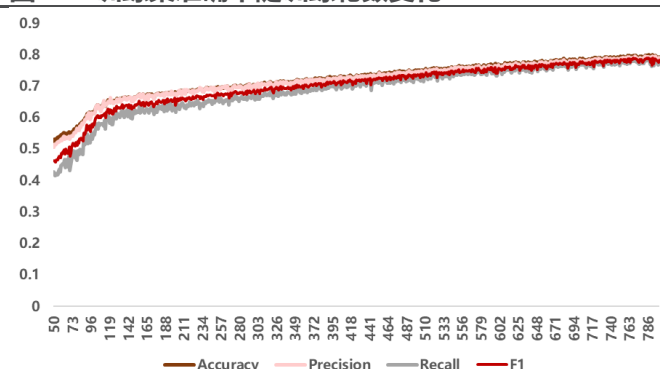
2. 精确率: $Precision = \frac{TP}{TP+FP}$

3. 召回率: $Recall = \frac{TP}{TP+FN}$

4. F1 得分: $F1score = 2 \frac{Precision \times Recall}{Precision + Recall}$

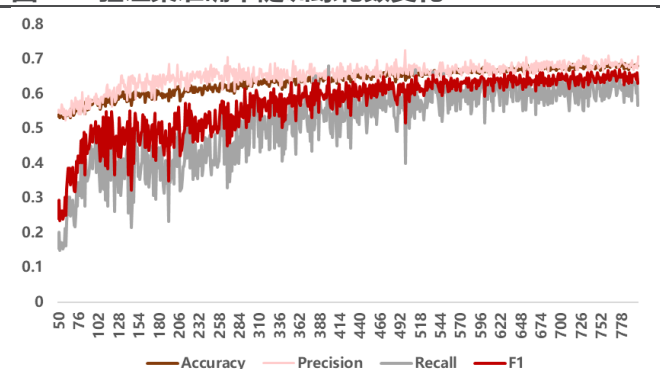
模型在训练集和预测集上的预测效果如下图所示：两个数据集上的表现都随训练轮数 epoch 增加稳步上升，验证集表现波动稍大，最终正确率稳定在 68% 左右，训练集正确率接近 80%。验证集最高精确率超过 70%，体现了模型在多头段的预测能力。

图14：训练集准确率随训练轮数变化



资料来源：wind，民生证券研究院

图15：验证集准确率随训练轮数变化



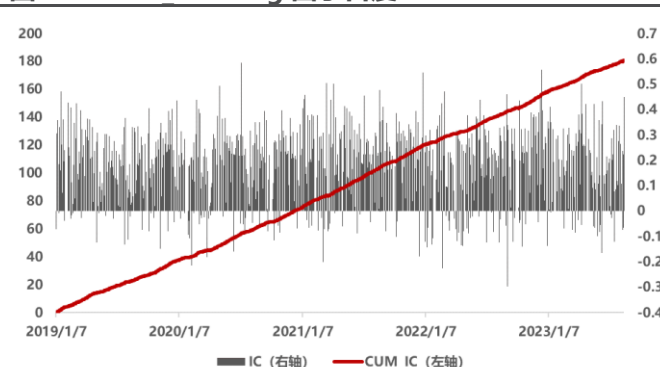
资料来源：wind，民生证券研究院

3.2 模型选股表现

我们取每次训练验证集表现最好的一次训练作为最终模型，将模型输出的涨跌概率作为个股 struc_learning 日频因子，在沪深 300 上回测日频选股效果。

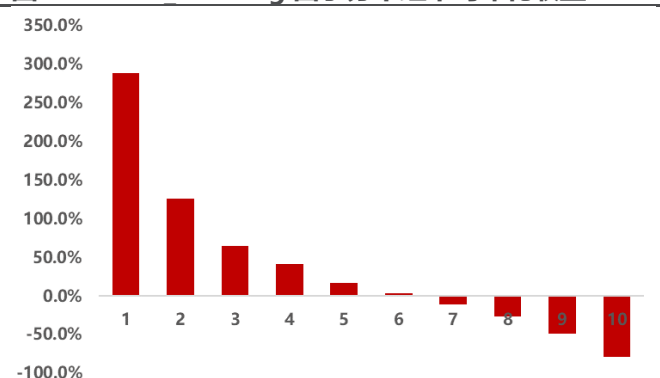
回测日期 2019 年 1 月 4 日至 2023 年 1 月 18 日。用当天因子值对应第二天开盘价到第三天开盘价？计算因子日度 IC 与分十组年化收益，因子表现出色，日度 IC 均值 0.16，ICIR1.2，分十组收益单调性强。

图16：struc learning 因子日度 IC



资料来源：wind，民生证券研究院

图17：struc learning 因子分十组平均年化收益

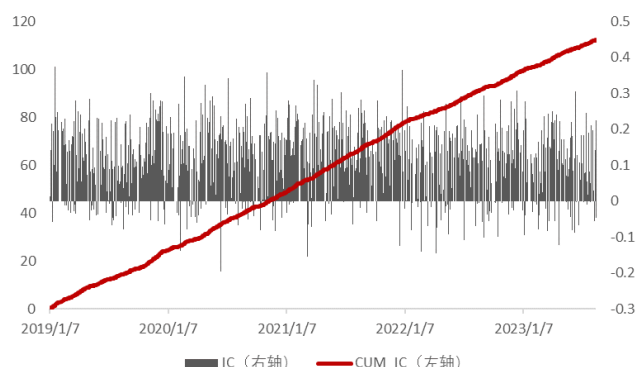


资料来源：wind，民生证券研究院

对 struc_learning 因子做市值行业中性化，中性化后因子 IC 均值下降到

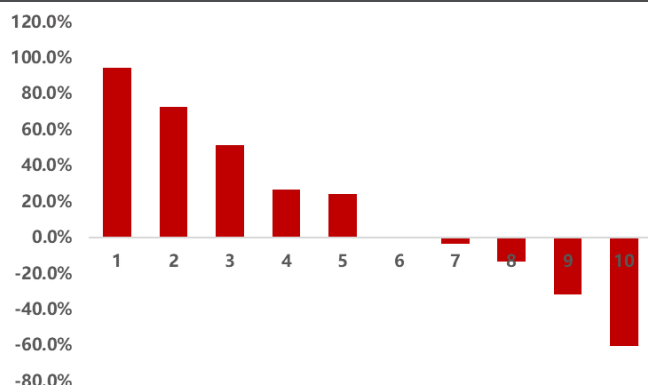
0.1, ICIR1.13, 分组收益单调性依然显著, 但收益下降明显。

图18: 中性化 struc learning 因子日度 IC



资料来源: wind, 民生证券研究院

图19: 中性化 struc learning 因子分十组平均年化收益



资料来源: wind, 民生证券研究院

计算因子多头组合收益时, 考虑到每日的开盘价竞争较为激烈, 我们采用每日 9:35 的价格作为成交价格, 避免股价高开, 以至于因子表现虚高。我们进行日度调仓, 取双边千三的手续费进行扣费, 计算因子 struc_learning 因子分年度表现如下图所示, 所有年份都取得显著收益, 表现优异。因子中性化后表现下降明显, 但整体仍能取得显著超额收益。

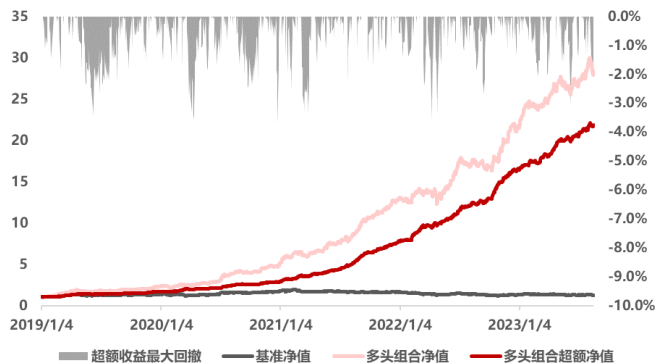
图20: struc learning 因子及中性化因子分年度表现

年份	费后多头收益	多头组合超额收益	多头组合信息比率	多空收益	日度胜率	中性化因子费后多头收益	中性化因子费后超额收益	中性化因子费后信息比率	中性化因子多空收益	中性化因子日度胜率
2019	130.0%	64.3%	5.96	990.7%	59.8%	67.7%	19.6%	1.98	333.1%	50.4%
2020	120.3%	74.4%	6.03	1402.5%	61.7%	59.1%	25.8%	2.60	370.7%	56.8%
2021	153.5%	165.1%	13.01	3924.7%	68.4%	39.4%	45.5%	4.90	773.6%	63.1%
2022	65.0%	111.5%	8.79	1884.9%	63.0%	-13.4%	10.8%	1.04	300.4%	49.4%
2023/8/18	31.8%	32.8%	3.28	460.6%	55.6%	-6.5%	-5.8%	-0.69	217.5%	44.4%
平均年化	106.7%	95.6%	8.00	1721.5%	62.2%	26.9%	19.9%	2.04	395.0%	53.5%

资料来源: wind, 民生证券研究院

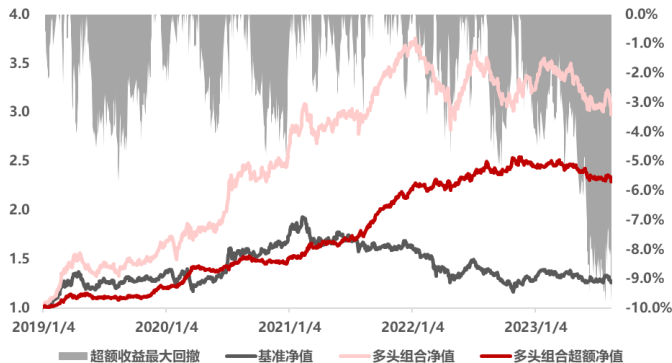
因子多头组合换手率较高, 年化单边换手率在 220 倍左右。

图21: struc learning 因子多头组合净值曲线



资料来源: wind, 民生证券研究院

图22: 中性化 struc learning 因子多头组合净值曲线

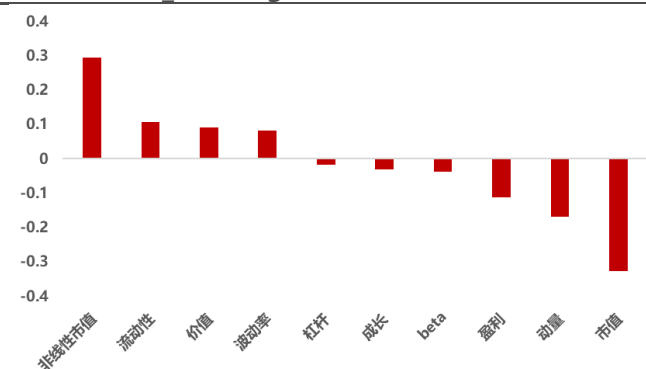


资料来源: wind, 民生证券研究院

对于中性化因子表现下降较多的原因, 我们对多头组合持仓计算主动风格暴

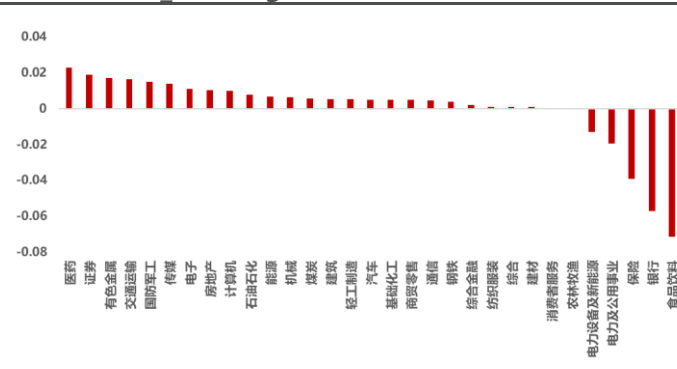
露及行业暴露，我们计算每一期的因子暴露并统计均值，发现组合的持股并没有在风格和行业上有过大偏离，相对较为明显的风格是沪深 300 内的相对中小市值，以及低配了食品饮料，银行等行业。在深度学习中，我们预测的是绝对收益，预测结果中本身就包含了对风格和行业收益的预测，对因子做中性化则剥离掉了这部分收益，得不偿失。

图23: struc_learning 因子多头组合平均风格偏离



资料来源：wind，民生证券研究院

图24: struc_learning 因子多头组合平均行业偏离

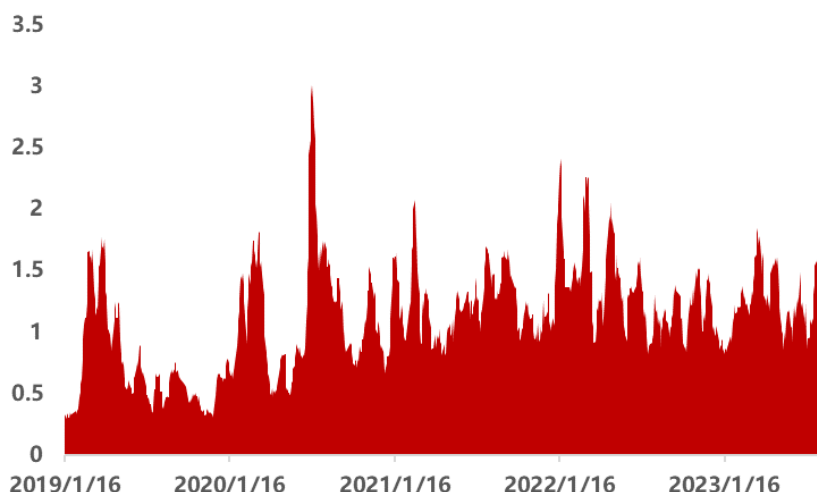


资料来源：wind，民生证券研究院

近年来，传统机器学习因子表现波动较大的环境下，struc_learning 因子仍然有着优异的表现，对比我们之前发布的报告《深度学习如何利用公募持仓网络优化选股效果？》中的 baseline LSTM 模型，效果有着明显提升。这也证明了基于可见性图结构嵌入的神经网络在选股上的可行性。

对因子多头组合做策略容量大约在 1 千万左右。我们取每日 9:35-9:40 的成交量最小股票的成交量乘以当天的持股数量，得到理想情况下，9:35-9:40 全部成交情况下的当日策略容量测算，进行 10 日的移动平均，struc_learning 因子多头组合平均每期持有 28-30 支股票，理想情况下策略 4 年平均容量为 **1.1 亿**。若假设这部分成交量中实际情况下可以成交 10%，则策略容量在千万左右。

图25: struc_learning 因子多头组合策略理想容量 (亿)

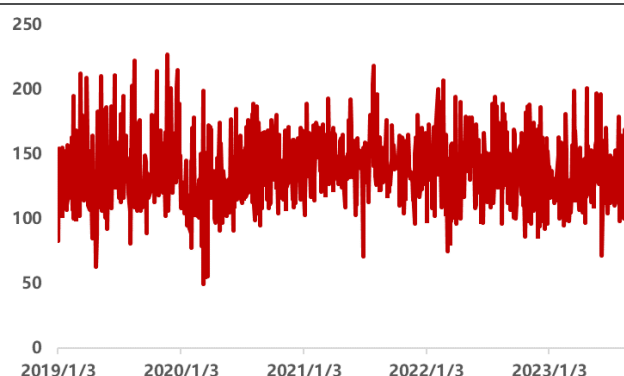


资料来源：Wind，民生证券研究院

3.3 模型择时表现

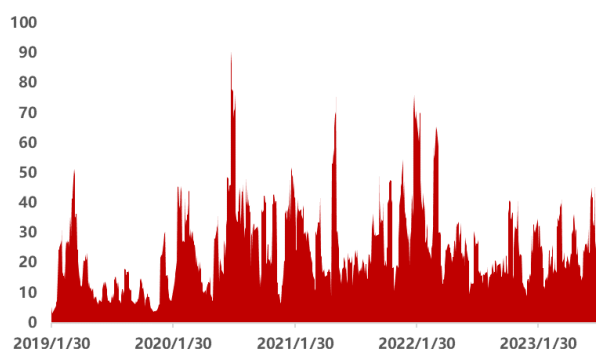
由于因子多头组合策略容量较低，我们考虑提升组合的持股数量来增加策略容量，即从截面选股策略改为类择时策略。我们对 struc_learning 因子做标准化，将因子值减去 0.5 后除以当天因子值标准差，得到标准化 struc_learning 因子，并按照因子值按比例买入因子值大于 0 的股票，构建 struc_learning 因子沪深 300 增强组合，组合平均持股 137 支，用类似的方法估算策略容量，增强组合策略容量平均为 24.7 亿，有明显提升。

图26: struc_learning 沪深 300 增强组合持股数量



资料来源：wind，民生证券研究院

图27: struc_learning 沪深 300 增强组合策略容量 (亿)



资料来源：wind，民生证券研究院

增加持股数量后，策略收益有所下降，但跟踪误差也有所降低，仍能取得良好表现。样本外仍能有 60%以上的胜率，证明了模型的有效性。

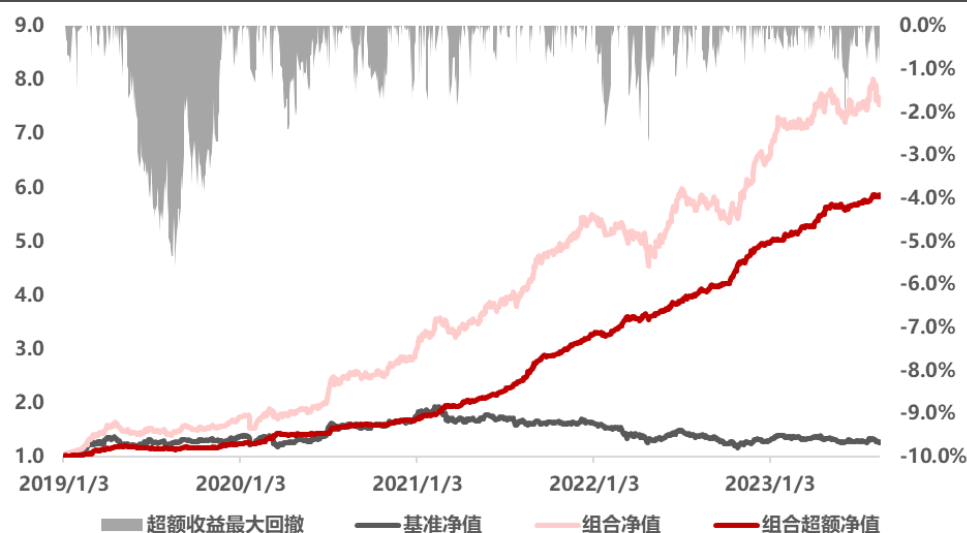
图28: struc_learning 因子沪深 300 增强组合分年度表现

年份	费后多头收益	多头组合超额收益	多头组合信息比率	最大回撤	日度胜率
2019	71.1%	22.0%	2.99	-5.6%	52.9%
2020	73.3%	37.1%	4.37	-2.4%	60.5%
2021	84.7%	93.0%	10.30	-1.8%	70.5%
2022	18.7%	52.0%	6.72	-2.7%	64.2%
2023/8/18	17.1%	18.0%	3.02	-2.0%	55.6%
平均年化	55.4%	46.9%	5.87	-5.6%	61.1%

资料来源：Wind，民生证券研究院

struc_learning 沪深 300 增强组合净值曲线如下。

图29：struc_learning 沪深 300 增强组合净值曲线

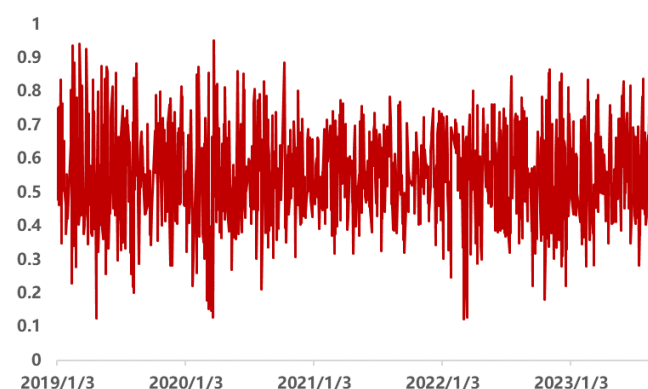


资料来源：wind，民生证券研究院

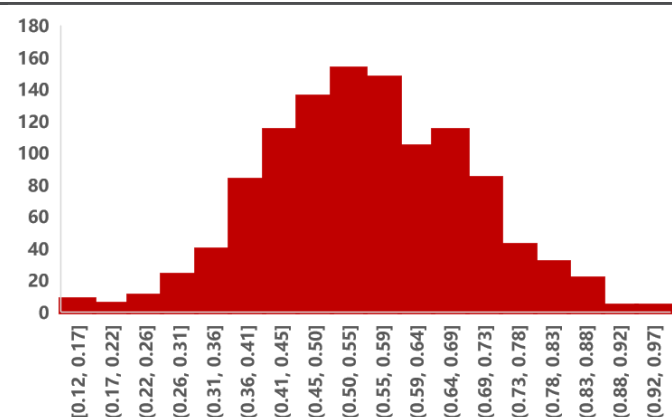
由于模型在沪深 300 成分股中优秀的选股与择时能力，我们考虑用个股信号合成股指信号，构建沪深 300 指数仓位择时策略，以探索模型在股指期货策略上的潜在应用意义。具体地，我们取模型预测的个股日度涨跌概率，以市值加权至沪深 300 股指信号，生成沪深 300 股指日度涨跌概率信号，信号波动较大，分布略微左偏。

图30：沪深 300 日度涨跌概率信号

图31：沪深 300 日度涨跌概率信号分布



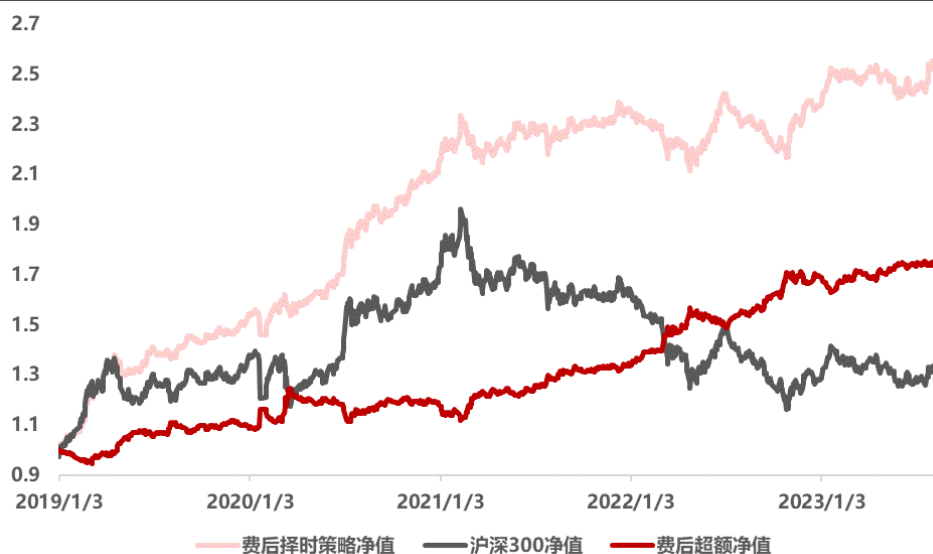
资料来源：wind，民生证券研究院



资料来源：wind，民生证券研究院

我们直接取当日沪深 300 日度涨跌概率信号作为持有沪深 300 指数的仓位比例。依然按照 9:35 分开盘价进行日度调仓，并且取双边千分之一作为回测费率日频调仓，总体预测正确率 64%，正确率较高。超额收益在历史上回撤原因主要为在市场行情较好时，没有满仓而导致的回撤。最终**年化超额收益 13.3%，信息比率 1.47。**

图32: struc learning 沪深 300 仓位择时策略净值



资料来源：wind，民生证券研究院

仓位择时策略的分年度表现如下：

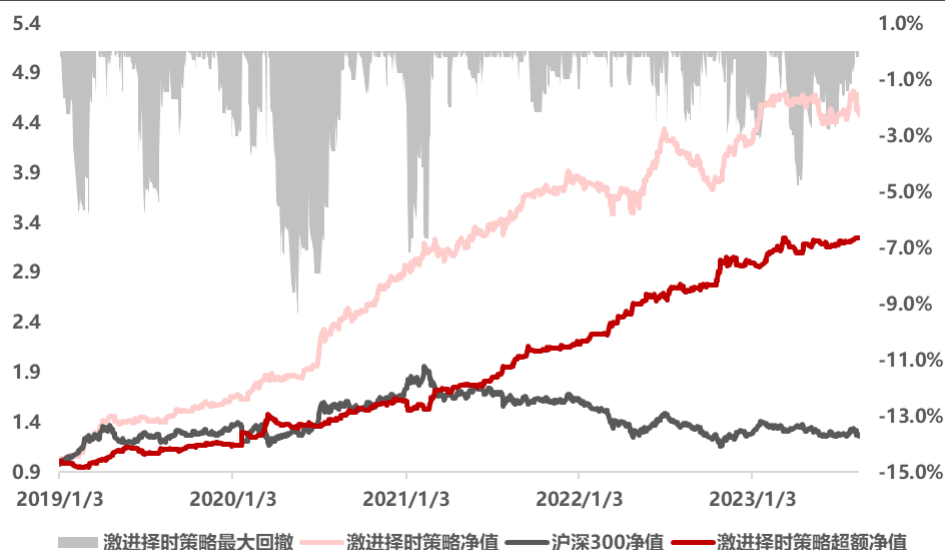
图33: struc learning 沪深 300 仓位择时策略分年表现

年份	策略绝对收益	策略超额收益	信息比率	夏普比率	日度胜率
2019	51.6%	9.7%	1.26	3.81	65.2%
2020	41.1%	7.5%	0.62	2.82	65.0%
2021	10.7%	13.0%	1.55	0.68	64.8%
2022	0.8%	26.5%	2.95	-0.20	63.0%
2023/8/18	3.3%	4.8%	0.97	0.16	60.3%
平均年化	21.6%	13.3%	1.47	1.56	64.0%

资料来源：wind，民生证券研究院

若我们采用更为激进的策略，保持胜率不变的情况下提升赔率，则可以取得更好的策略表现。我们规定若预测上涨概率为 50%以上，则全仓持有沪深 300 指数，否则在第二天上午 9:35 分卖出指数并空仓。这样一来，策略的波动率有显著上升，但收益可以获得更显著的提升。

图34: struc learning 沪深 300 激进择时策略净值



资料来源：wind，民生证券研究院

激进择时策略的分年度表现如下:

图35: struc learning 沪深 300 激进择时策略分年表现

年份	策略绝对收益	策略超额收益	信息比率	夏普比率	最大回撤
2019	61.7%	16.9%	1.53	4.06	-5.8%
2020	80.9%	37.6%	2.20	4.73	-9.3%
2021	32.0%	35.0%	3.00	1.93	-7.2%
2022	9.1%	37.2%	3.25	0.42	-3.0%
2023/8/18	6.5%	8.2%	1.37	0.36	-4.8%
平均年化	38.7%	29.3%	2.36	2.41	-9.3%

资料来源: wind, 民生证券研究院

若可以做空，则策略可以取得进一步收益。我们考虑不同信号阈值下的多空策略，分别尝试以 $[0.5, 0.5]$, $[0.45, 0.55]$, $[0.4, 0.6]$, $[0.35, 0.65]$, $[0.3, 0.7]$ 五组

阈值为多空信号下，模型的胜率和多空收益，例如[0.4, 0.6]代表若预测指数上涨概率小于 0.4，则卖空；大于 0.6，则买多，若介于 0.4 与 0.6 之间，则空仓的策略收益与胜率（若空仓则不计入胜率统计）。随着阈值增加，信号变窄，模型胜率有明显提升，然而长时间的空仓导致了收益有一定下滑，在[0.45, 0.55]的阈值信号上多空策略有最大的夏普比率。

图36：不同阈值信号下年化多空收益与胜率

多空信号阈值	年化多空收益	夏普比率	预测胜率
[0.5, 0.5]	120.1%	6.22	64.0%
[0.45, 0.55]	112.8%	6.31	66.9%
[0.4, 0.6]	76.6%	5.06	69.8%
[0.35, 0.65]	44.4%	3.30	71.8%
[0.3, 0.7]	32.4%	2.88	74.9%

资料来源：wind，民生证券研究院

图37：不同阈值信号下多空收益净值曲线


资料来源：wind，民生证券研究院

4 总结与思考

在本篇研究中，我们利用沪深 300 内股票的价格信息构造可见性图复杂网络，使用结构向量法提取出图结构嵌入，输入到预测模块 DA-RNN+CAAN 中，预测下一日的股票上涨概率，取得显著效果，在选股和择时方面都有较好效果，证明了可见性图嵌入深度神经网络的可行性。

可见性图是一种将时间序列数据转化成一张图结构的方法。抽取可见图的结构特征并训练分类器的方法被广泛用于时间序列数据的分类问题。复杂网络每个节点的权重不同，我们采用 CI 算法定义节点权重，衡量节点及其局部结构在整个网络中的重要性。

利用结构向量法提取图结构特征，可以刻画股票的潜在属性。与传统的图嵌入方法相比，struc2vec 更加高效，且在多个数据集上分类更加精确，并可以应用于大规模图的嵌入。最后，利用二阶段循环神经网络与跨资产注意力网络构造预测模块，从而纳入潜在属性及股票间的相互关系。两阶段注意力循环神经网络 DA-RNN 可以通过按照时间顺序上不断训练结构性信息的模型结构，提取图结构嵌入特征，在生成股票因子表示时纳入股票的潜在信息。跨资产注意力网络 CAAN 可以对股票间的相互关系进行建模，计算 batch 内因子表示的相似度，作为注意力机制的来源。

利用混淆矩阵来衡量模型准确率，验证集内准确率 68%，精确率达 70%。日频模型选股和择时效果显著，合成股指信号依然有效。将模型预测的个股日度上涨概率作为日频选股因子 struc_learning，因子日度 IC 均值 0.16，ICIR1.2；中性化后因子表现有所下降，但选股效果依然显著。日度调仓多头组合年化超额收益 95.6%，信息比率 8.0，但策略容量较低。增加持股数量构建沪深 300 增强组合，组合平均持股数量 137 只，策略容量显著上升，年化超额收益 46.7%，信息比率 5.87。用个股信号合成股指日度涨跌概率信号，并构建沪深 300 仓位择时策略，年化超额收益 13.3%，信息比率 1.47，预测涨跌准确率 64%。

基于可见性图的时序预测有效窗口期较短，较为适合日频策略，若增加自变量窗口期或者用周 k 建模，效果有较大衰减。深度学习策略可解释性相对较弱，若未来策略失效，无法归因，重新训练参数耗时较长，这也是深度学习策略的通病。

5 风险提示

量化模型基于历史数据，市场未来可能发生变化，策略模型有失效可能。

6 参考文献

- [1] Lacasa, L., Luque, B., Ballesteros, F., Luque, J., & Nuño, J. C. (2008, April). From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences*, 105(13), 4972–4975.
- [2] Daoyuan Li; Jessica Lin; Tegawendé F. Bissyandé; Jacques Klein; Yves Le Traon. (2018). Extracting Statistical Graph Features for Accurate and Efficient Time Series Classification. *Proceedings of the 21st International Conference on Extending Database Technology (EDBT)*, ISBN 978-3-89318-078-3.
- [3] F. Morone, H.A. Makse, Influence maximization in complex networks through optimal percolation, *Nature* 524 (7563) (2015) 65–68.
- [4] Ribeiro, L. F., Saverese, P. H., & Figueiredo, D. R. (2017, August 4). struc2vec. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- [5] Adrian Colyer, September 2017, the morning paper: Struc2vec: learning node representations from structural identity, blog.acolyer.org
- [6] Wu, J., Xu, K., Chen, X., Li, S., & Zhao, J. (2022, April). Price graphs: Utilizing the structural information of financial time series for stock prediction. *Information Sciences*, 588, 405–424.
- [7] Y. Qin, D. Song, H. Cheng, W. Cheng, G. Jiang, G.W. Cottrell, A dual-stage attention-based recurrent neural network for time series prediction, in: *Proceedings of the 26th IJCAI*, 2017, pp. 2627–2633.
- [8] J. Wang, Y. Zhang, K. Tang, J. Wu, Z. Xiong, Alphastock, A buying-winners-and-selling-losers investment strategy using interpretable deep reinforcement attention networks, in: *Proceedings of the 25th SIGKDD*, 2019, pp. 1900–1908.

插图目录

图 1: 可见性图连接规则示意	3
图 2: 某只股票 20 天收盘价的可见性图及复杂网络	4
图 3: 可见性图结构因子日度 IC	4
图 4: 某只股票 20 天量价指标的相应节点权重示例	5
图 5: 结构向量法示意图	7
图 6: 基于图结构嵌入的深度神经网络结构	8
图 7: DA-RNN 网络结构	9
图 8: CAAN 网络结构示意图	10
图 9: k-means 法聚类各类别股票个数	11
图 10: 10 类股票的风格偏离	12
图 11: 10 类股票的行业分布	12
图 12: 模型数据与参数细节	13
图 13: 混淆矩阵的四分类	13
图 14: 训练集准确率随训练轮数变化	14
图 15: 验证集准确率随训练轮数变化	14
图 16: struc_learning 因子日度 IC	14
图 17: struc_learning 因子分十组平均年化收益	14
图 18: 中性化 struc_learning 因子日度 IC	15
图 19: 中性化 struc_learning 因子分十组平均年化收益	15
图 20: struc_learning 因子及中性化因子分年度表现	15
图 21: struc_learning 因子多头组合净值曲线	15
图 22: 中性化 struc_learning 因子多头组合净值曲线	15
图 23: struc_learning 因子多头组合平均风格偏离	16
图 24: struc_learning 因子多头组合平均行业偏离	16
图 25: struc_learning 因子多头组合策略理想容量 (亿)	16
图 26: struc_learning 沪深 300 增强组合持股数量	17
图 27: struc_learning 沪深 300 增强组合策略容量 (亿)	17
图 28: struc_learning 因子沪深 300 增强组合分年度表现	17
图 29: struc_learning 沪深 300 增强组合净值曲线	18
图 30: 沪深 300 日度涨跌概率信号	18
图 31: 沪深 300 日度涨跌概率信号分布	18
图 32: struc_learning 沪深 300 仓位择时策略净值	19
图 33: struc_learning 沪深 300 仓位择时策略分年表现	19
图 34: struc_learning 沪深 300 激进择时策略净值	20
图 35: struc_learning 沪深 300 激进择时策略分年表现	20
图 36: 不同阈值信号下年化多空收益与胜率	21
图 37: 不同阈值信号下多空收益净值曲线	21

分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并登记为注册分析师，基于认真审慎的工作态度、专业严谨的研究方法与分析逻辑得出研究结论，独立、客观地出具本报告，并对本报告的内容和观点负责。本报告清晰地反映了研究人员的研究观点，结论不受任何第三方的授意、影响，研究人员不曾因、不因、也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

评级说明

投资建议评级标准		评级	说明
以报告发布日后的 12 个月内公司股价（或行业指数）相对同期基准指数的涨跌幅为基准。其中：A 股以沪深 300 指数为基准；新三板以三板成指或三板做市指数为基准；港股以恒生指数为基准；美股以纳斯达克综合指数或标普 500 指数为基准。	公司评级	推荐	相对基准指数涨幅 15%以上
		谨慎推荐	相对基准指数涨幅 5% ~ 15%之间
		中性	相对基准指数涨幅-5% ~ 5%之间
		回避	相对基准指数跌幅 5%以上
	行业评级	推荐	相对基准指数涨幅 5%以上
		中性	相对基准指数涨幅-5% ~ 5%之间
		回避	相对基准指数跌幅 5%以上

免责声明

民生证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。

本报告仅供本公司境内客户使用。本公司不会因接收人收到本报告而视其为客户。本报告仅为参考之用，并不构成对客户的投资建议，不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，客户应当充分考虑自身特定状况，不应单纯依靠本报告所载的内容而取代个人的独立判断。在任何情况下，本公司不对任何人因使用本报告中的任何内容而导致的任何可能的损失负任何责任。

本报告是基于已公开信息撰写，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、意见及预测仅反映本公司于发布本报告当日的判断，且预测方法及结果存在一定程度局限性。在不同时期，本公司可发出与本报告所刊载的意见、预测不一致的报告，但本公司没有义务和责任及时更新本报告所涉及的内容并通知客户。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问、咨询服务等相关服务，本公司的员工可能担任本报告所提及的公司的董事。客户应充分考虑可能存在的利益冲突，勿将本报告作为投资决策的唯一参考依据。

若本公司以外的金融机构发送本报告，则由该金融机构独自为此发送行为负责。该机构的客户应联系该机构以交易本报告提及的证券或要求获悉更详细的信息。本报告不构成本公司向发送本报告金融机构之客户提供的投资建议。本公司不会因任何机构或个人从其他机构获得本报告而将其视为本公司客户。

本报告的版权仅归本公司所有，未经书面许可，任何机构或个人不得以任何形式、任何目的进行翻版、转载、发表、篡改或引用。所有在本报告中使用的商标、服务标识及标记，除非另有说明，均为本公司的商标、服务标识及标记。本公司版权所有并保留一切权利。

民生证券研究院：

上海：上海市浦东新区浦明路 8 号财富金融广场 1 幢 5F； 200120

北京：北京市东城区建国门内大街 28 号民生金融中心 A 座 18 层； 100005

深圳：广东省深圳市福田区益田路 6001 号太平金融大厦 32 层 05 单元； 518026