

# 基于循环神经网络与注意力机制的波动预测模型

李希今<sup>1</sup>, 王祥任<sup>2,3</sup>, 刘金石<sup>1</sup>

1. 吉林大学 人力资源处, 长春 130012;
2. 内蒙古大学 计算机学院(软件学院), 呼和浩特 010021;
3. 内蒙古大学 人工智能学院, 呼和浩特 010021)

**摘要:** 针对经典机器学习算法(如决策树、随机森林)在建模复杂隐式交互关系时预测准确率较低的问题, 提出一个基于循环神经网络与注意力机制的波动预测模型. 首先通过注意力机制计算各影响因素之间复杂的交互关系, 然后采用循环神经网络学习表示模型的隐变量, 从而实现精准预测. 与多个经典预测模型进行仿真对比实验的结果表明, 该模型的预测准确率显著高于其他机器学习模型, 从而为波动预测领域提供了一种更高效、精准的解决方案.

**关键词:** 循环神经网络; 注意力机制; 机器学习; 预测模型

**中图分类号:** TP183 **文献标志码:** A **文章编号:** 1671-5489(2025)05-1397-07

## Fluctuation Prediction Model Based on Recurrent Neural Network and Attention Mechanism

LI Xijin<sup>1</sup>, WANG Xiangren<sup>2,3</sup>, LIU Jinshi<sup>1</sup>

1. *Human Resource Department, Jilin University, Changchun 130012, China;*
2. *College of Computer Science (College of Software), Inner Mongolia University, Hohhot 010021, China;*
3. *College of Artificial Intelligence, Inner Mongolia University, Hohhot 010021, China)*

**Abstract:** Aiming at the problem of low prediction accuracy when classical machine learning algorithms (e. g., decision trees, random forests) modelled complex implicit interaction relationships, we proposed a fluctuation prediction model based on recurrent neural networks and an attention mechanism. We first calculated the complex interaction relationships among various influencing factors through the attention mechanism, and then used recurrent neural networks to learn the hidden variable representations of the model, thereby achieving precise prediction. The results of simulation and comparative experiments with multiple classical prediction models show that the prediction accuracy of proposed model is significantly higher than other machine learning models, providing a more efficient and accurate solution for the field of volatility prediction.

**Keywords:** recurrent neural network; attention mechanism; machine learning; prediction model

近年来, 随着计算能力和数据规模的提升, 机器学习在自然语言处理、计算机视觉等领域取得了一定进展<sup>[1]</sup>. 在自然语言处理领域, 通过引入稀疏激活技术和更高效的训练算法, 大语言模型在规模与效率之间实现了更好的平衡, 在保持模型性能的同时, 有效降低了计算成本, 并广泛应用于多领

收稿日期: 2024-09-23.

第一作者简介: 李希今(1989—), 女, 汉族, 硕士, 助理研究员, 从事智能算法、人力资源管理和薪酬模型算法的研究, E-mail: lxj0907@jlu.edu.cn.

基金项目: 吉林省自然科学基金(批准号: 20210101167JC).

域<sup>[2-4]</sup>. 在计算机视觉领域, 视频生成技术也取得了重要突破. 通过结合时序信息和空间信息, 模型能生成连续且自然的视频内容, 但也面临着高质量视频生成和多环境泛化能力等挑战. 此外, 时间序列预测也是机器学习的一个重要应用方向. 机器学习模型能对动态数据进行建模, 进而为相关决策提供有力支持. 例如, Hamlen 等<sup>[5]</sup>提出了基于分层线性回归进行预测的分析方法; Bansal 等<sup>[6]</sup>采用多元线性回归方法进行预测; Navyashree 等<sup>[7]</sup>采用包括随机森林、支持向量机、决策树等多种机器学习方法进行预测. 同时, Bayes 模型在该领域也展现了广泛的应用价值, 例如, 用 Bayes 分层平滑算法, 通过整合结果的分位数、平均值等关键统计信息, 精确建模预测分布情况. 但此类方法考虑的预测结果不仅影响因素较片面, 而且应用的传统机器学习模型也很难建模复杂高维数据, 使模型准确率和应用场景受限.

目前, 循环神经网络(RNN)及其变体, 如长短期记忆网络(LSTM)已成为序列建模的核心技术, 并在自然语言处理、时间序列预测等领域取得了显著优势. 通过引入残差连接、正则化技术及优化等训练策略, 循环神经网络的长距离依赖建模能力与训练稳定性得到了显著提升. 例如, 在语言生成任务中, LSTM 能有效捕捉文本中的跨句语义关联; 在价格预测中, 门控循环单元(GRU)通过简化门控结构实现了更高的计算效率. 然而, RNN 仍面临其固有挑战. 在长序列输入时, 尽管 LSTM 等循环神经网络变体缓解了梯度消失问题, 但堆叠多层网络仍易导致训练速度下降和内存占用激增, 此外, RNN 的序列依赖特性使其难以并行化, 从而限制了其在大规模数据场景下的可扩展性.

注意力机制通过动态分配计算资源, 已成为提高模型表达能力的关键技术. 其应用已从自然语言处理领域扩展至视频生成、多模态融合等多领域. 在技术改进方面, 稀疏注意力通过限制关注范围, 有效降低了计算复杂度, 而位置编码的引入则增强了模型对序列顺序的感知能力. 例如, 在时间序列预测中, 多头注意力机制可同时捕捉不同时间尺度的依赖关系, 显著提升预测精度. 但注意力机制仍存在明显局限: 首先, 自注意力计算的时间复杂度随序列长度呈平方级增长, 处理超长序列时效率较低; 其次, 注意力权重缺乏可解释性, 难以直观分析模型决策依据; 最后, 在多因素交互场景中, 传统注意力机制对特征间非线性关系的建模能力有限, 易忽略局部细节信息.

波动预测是一个典型的多因素时序依赖性建模任务, 其核心在于准确捕捉动态系统中各变量间的复杂非线性交互关系及长期演化规律. 现有的波动预测模型通常在简单场景下性能稳定, 但在处理高维异构数据时存在显著局限性. 首先, 这些模型难以有效建模多因素间的非线性隐式耦合关系, 导致关键交互信息丢失. 其次, 尽管一些研究通过引入 LSTM 和 GRU, 在一定程度上缓解了长序列建模中的梯度消失问题, 但在处理超长序列时仍可能面临捕捉全局依赖的挑战. 最后, 现有方法多采用静态特征融合策略, 无法动态调整不同时间步及特征维度的重要性权重. 为解决现有方法复杂交互关系表征不足、长程依赖捕捉低效和特征动态聚焦能力欠缺等问题, 本文提出一种基于循环神经网络与注意力机制的协同预测模型. 该模型通过注意力机制实现多粒度特征交互的动态感知, 精准量化各时间步及特征维度的权重; 同时依托门控循环单元捕获序列的深层演化规律. 最后, 通过二者的协同优化, 提升波动预测的精度和鲁棒性.

## 1 相关工作

在波动预测领域, 传统预测方法常使用统计模型和经典机器学习算法, 已有研究基于文档嵌入模型(Doc2vec)<sup>[8]</sup>从文本信息中计算特征向量, 再结合随机森林、支持向量机等机器学习模型建立预测模型. 例如, Jackman 等<sup>[9]</sup>采用词频-逆文本频率模型(TF-IDF)提取特征向量, 并分别使用逻辑回归、神经网络等方法进行预测. 张浩宇<sup>[10]</sup>采用潜在 Dirichlet 分配(LDA)主题模型与 Doc2vec 模型相结合的方法计算文本信息之间的相似度, 通过选取最优参数, 构建预测模型. 然而这类方法依赖于自然语言处理模型的特征提取能力, 提取出的文本特征也大多难以解释, 最终导致影响因素分析的偏差. 随着深度学习技术的发展, 研究人员因其强大的建模能力进一步将多种模型引入到预测任务中. 例如, Meng 等<sup>[11]</sup>提出了一种基于矩阵分解的预测模型; 孙莹等<sup>[12]</sup>提出了一种基于边际贡献的增量式集合效用模型网络. 近年来, 循环神经网络及其变体也逐渐被引入该领域中, 例如, Wang 等<sup>[13]</sup>集合了卷积

神经网络和双向门控循环单元, 建模文本信息中存在的特征以实现端到端的预测模型. 这些模型通过引入循环结构, 能有效处理具有时间依赖性的数据, 广泛应用于自然语言处理、金融时间序列预测等领域. 循环神经网络是一种专门用于处理序列数据的神经网络结构, 它通过引入循环结构捕捉序列中的时间依赖关系. RNN 的基本形式可表示为

$$h_t = \sigma(W_{hh}h_{t-1} + W_{xh}x_t + b_h), \tag{1}$$

其中  $h_t$  为时间步  $t$  的隐藏状态,  $x_t$  为时间步的输入,  $\sigma$  为激活函数,  $W_{hh}$  和  $W_{xh}$  为权重矩阵,  $b_h$  为偏置项. 这种结构使得 RNN 能将前一时间步的信息传递到当前时间步, 从而捕捉序列中的时间依赖性. 但 RNN 在处理长序列数据时易出现梯度消失或梯度爆炸的问题, 导致模型难以学习到长距离的依赖关系. 为解决这些问题, 研究人员提出了长短期记忆网络和门控循环单元等变体. 这些变体通过引入门控机制, 有效缓解了梯度消失问题, 提高了模型的训练效率和预测精度. 例如, LSTM 通过遗忘门、输入门和输出门的机制控制信息的流动, 而 GRU 则通过更新门和重置门控制信息的更新和遗忘. 在实际应用中, RNN 及其变体的性能可通过以下损失函数进行优化:

$$L = \sum_{t=1}^T l(y_t, \hat{y}_t), \tag{2}$$

其中  $y_t$  为真实值,  $\hat{y}_t$  为预测值,  $l$  为损失函数(如均方误差或交叉熵),  $T$  为序列的长度. 通过最小化该损失函数, 模型可学习到最优的参数, 从而提高预测的准确性. 此外, 为进一步提高模型的泛化能力, 通常会在训练过程中加入正则化项, 如  $L_2$  正则化:

$$L_{reg} = L + \lambda \sum_i \|W_i\|^2, \tag{3}$$

其中  $L$  为原始损失函数,  $\lambda$  为正则化参数,  $W_i$  为模型的权重矩阵. 通过加入正则化项, 可以有效防止模型过拟合, 提高模型在测试集上的性能. 然而, 影响预测结果的因素是多方面的, 仅依赖循环神经网络的模型只关注提取独立的全局特征, 难以刻画多种影响因素之间复杂的依赖关系, 使模型的准确率受限.

注意力机制作为一种模拟人类注意力选择的机制, 解决了传统基于自然语言处理技术难以建模多种因素之间复杂关系的问题. 在自然语言处理、时间序列分析等领域得到了广泛应用. 注意力机制的核心思想是使模型在处理信息时能集中关注最重要的部分, 从而提高模型的性能. 传统注意力机制可表示为

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{QK}^T}{\sqrt{d_k}}\right)\mathbf{V}, \tag{4}$$

其中  $\mathbf{Q}, \mathbf{K}, \mathbf{V}$  分别为查询(Query)、键(Key)和值(Value),  $d_k$  为键向量的维度, Softmax 为归一化函数. 在机器翻译中, 注意力机制能使模型更好地对齐源语言和目标语言的单词, 生成更准确的翻译结果; 在时间序列预测中, 注意力机制能引导模型关注关键点, 提高预测的准确率. 例如, 通过计算每个时间步的注意力权重, 模型可动态地分配更多的计算资源处理更重要的信息:

$$\alpha_t = \frac{\exp\{e_t\}}{\sum_{k=1}^T \exp\{e_k\}}, \tag{5}$$

其中  $e_t$  为时间步  $t$  的未归一化注意力分数,  $\alpha_t$  为归一化后的注意力权重,  $T$  为序列的长度. 通过这种方式, 模型可更有效地利用有限的计算资源提高预测精度. 但传统注意力机制在处理多因素交互关系时, 难以有效融合不同因素之间的信息, 导致模型对数据的整体理解不全面. 此外, 注意力机制的计算复杂度较高, 尤其在处理大规模数据集时, 计算资源的消耗较大. 为缓解上述问题, 研究人员提出了多种改进的注意力机制. 例如, 引入位置编码增强模型对位置信息的敏感性:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{QK}^T + \mathbf{P}}{\sqrt{d_k}}\right)\mathbf{V}, \tag{6}$$

其中  $\mathbf{P}$  为位置编码矩阵, 用于引入位置信息, 帮助模型更好地理解序列中的位置关系.

## 2 模型设计

波动预测旨在从众多影响预测结果的相关数据中学习到有用的特征, 并对其进行分类. 本文首先将要处理的数据分为多个类别, 然后将编码后的特征向量输入到注意力层中学习特征之间的复杂关系, 最后采用循环神经网络学习代表波动的隐变量, 从而得到更准确的预测结果, 模型结构如图 1 所示.

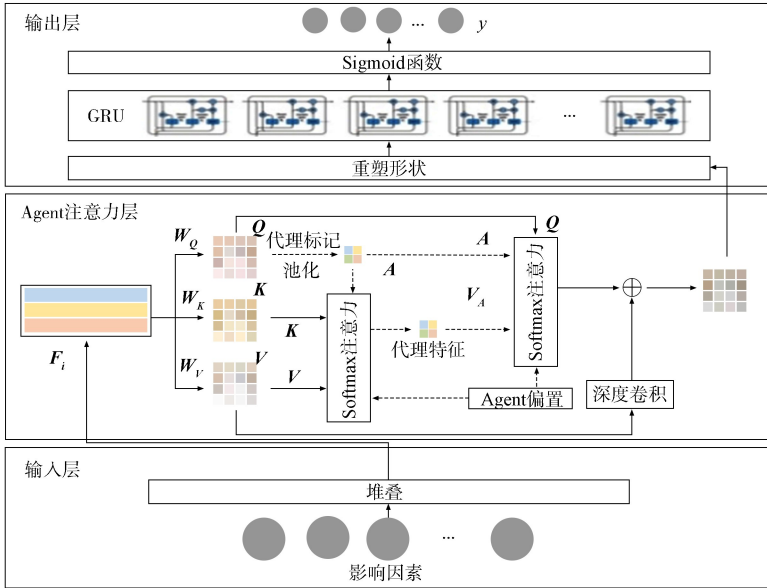


图 1 模型结构

Fig. 1 Model structure

### 2.1 波动预测影响因素

波动预测是分析和预测变量随时间变化趋势的重要工具, 广泛应用于多领域. 为实现准确的波动预测, 需要综合考虑多种影响因素. 本文首先将影响预测结果的相关数据分为多个类别, 并对这些类别中的特征向量进行编码, 获得特征的代表向量. 然后, 将这些分类后的因素通过 one-hot 编码技术转换成特征向量, 获得这些特征向量后, 将它们合并成一个表示整体特征的矩阵, 作为模型的输入. 最后, 模型通过学习这些输入数据之间的复杂关系, 实现对波动的精准预测.

### 2.2 注意力机制设计

表示整体特征的矩阵  $F_i$  通过多个特征融合得到, 具有多特征、高维度的特点, 因此采用一般性的注意力机制进行特征融合会降低模型的整体运算效率. 为解决该问题, 受 Agent 注意力机制在视觉、自然语言处理等任务中的成功启发, 模型采用兼具高计算效率与高表达能力的机制对嵌入矩阵进行处理, 计算公式如下:

$$Q = F_i W_Q, \quad K = F_i W_K, \quad V = F_i W_V, \tag{7}$$

其中  $W_Q$  和  $W_K$  均为可学习权重矩阵. 最终的输出可表示为

$$O^A = \text{Attn}^S(QA^T + B_2) \text{Attn}^S(AK^T + B_1)V + \text{DWC}(V), \tag{8}$$

其中:  $\text{Attn}^S$  表示 Softmax 注意力;  $B_1$  和  $B_2$  表示 Agent 偏置, 以更好地利用位置信息;  $A$  表示代理标记, 通过对  $Q$  的池化操作获得;  $\text{DWC}$  表示深度卷积, 以保持特征的多样性. 由图 1 可见, Agent 注意力模块由两个 Softmax 注意力操作组成, 分别称为 Agent 聚合和 Agent 传播. 首先将  $A$  视为查询, 并在  $A, K$  和  $V$  之间执行注意力计算, 以从所有值中聚合 Agent 特征  $V_A$ . 其次, 在具有查询矩阵  $Q$  的第二次注意力计算中, 使用  $A$  作为密钥,  $V_A$  作为值, 将来自 Agent 特征的全局信息广播到每个查询 token, 并获得最终的输出  $O^A$ . 通过这种方式, 避免了计算  $Q$  和  $K$  之间的成对相似性, 同时通过代理标记保持每个“查询-密钥”对之间的信息交换. 模型将代理标记的数量设置为一个小的超参数, 以降

低计算复杂度. 最后, 通过该层得到的特征融合向量  $\mathbf{O}^A$  送入到一个转换层中, 获得富含丰富特征信息的向量  $\mathbf{e}_i$  作为循环神经网络模型的输入.

### 2.3 波动预测模型

循环神经网络目前已被成功应用于各类预测任务中, 例如股价预测、天气预测等. 给定特征向量, 应用该网络在时间步  $t$  进行波动预测可表示为

$$\mathbf{h}_t = \text{RNN}(\mathbf{e}_t, \mathbf{h}_{t-1}; \theta_h), \tag{9}$$

其中  $\theta_h$  为 RNN 参数集合, RNN 为循环神经网络函数.

门控递归单元作为循环神经网络的一种变体, 由于其结构简单、计算效率高的特点, 广泛应用于多种任务中. 受此启发, 本文采用 GRU 构建波动预测模型, 将式(2)转换为  $\mathbf{h}_t = \text{GRU}(\mathbf{e}_t, \mathbf{h}_{t-1}; \theta_h)$ . 对于时间步  $t$ , 学习者的隐藏知识状态计算公式为

$$\mathbf{z}_t = \sigma(\mathbf{W}_z \mathbf{e}_t + \mathbf{U}_z \mathbf{h}_{t-1}), \tag{10}$$

$$\mathbf{r}_t = \sigma(\mathbf{W}_r \mathbf{e}_t + \mathbf{U}_r \mathbf{h}_{t-1}), \tag{11}$$

$$\hat{\mathbf{h}}_t = \tanh(\mathbf{W}_h \mathbf{e}_t + \mathbf{r}_t \odot \mathbf{U}_h \mathbf{h}_{t-1}), \tag{12}$$

$$\mathbf{h}_t = \mathbf{z}_t \odot \mathbf{h}_{t-1} + (1 - \mathbf{z}_t) \odot \hat{\mathbf{h}}_t, \tag{13}$$

其中  $\sigma$  表示 Sigmoid 函数,  $\odot$  表示 Hadamard 积.  $\mathbf{W}_z, \mathbf{W}_r, \mathbf{W}_h, \mathbf{U}_z, \mathbf{U}_r, \mathbf{U}_h$  均为可学习的权重矩阵. 在时间步  $t$ , 模型输出波动结果可表示为

$$\hat{\mathbf{p}}_t = \text{Sigmoid}(\mathbf{W}_o \mathbf{h}_t + b_o), \tag{14}$$

其中  $\mathbf{W}_o$  和  $b_o$  分别为可学习的权重矩阵和偏置项. 最终, 采用预测标签与真实标签之间的交叉熵进行模型训练:

$$L = \sum_t [\mathbf{p}_t \log \hat{\mathbf{p}}_t + (1 - \mathbf{p}_t) \log(1 - \hat{\mathbf{p}}_t)]. \tag{15}$$

## 3 实验结果分析

### 3.1 实验环境与参数设置

实验采用 PyTorch 作为框架, Python3. 8. 10 作为编程语言. 实验设备采用 Intel Core i7-9750H @ 2.60 GHz CPU, GPU 型号为 RTX3070 8G V2. 实验采用吉林大学个人薪酬数据集作为测试模型的数据集, 该数据集通过数据分析, 综合考量了薪酬的影响因素. 例如, Khongchai 等<sup>[14]</sup>采用员工的职业培训、职业证书和平均绩点(GPA)等特征预测员工的薪酬水平, 最终吉林大学个人薪酬数据集将影响薪酬的因素归结为三方面: 企业因素、个人因素和外部因素. 企业因素主要包括经营状况、负担能力和发展时期. 个人因素主要包括教育水平、工作年限、工作表现和工作内容. 外部因素主要包括经济环境和政策、地理位置和市场供需. 实验采用的吉林大学个人薪酬数据集规模约为 10 000 条数据, 确保了有足够的数量验证模型的性能和有效性. 实验按 8 : 2 分为训练集和测试集, 采用分批次训练方法, 批次大小设定为 64. 网络初始学习率设定为 0.001, 衰减因子设定为 0.9, 衰减步长设定为 10.

### 3.2 评价指标

实验采用准确率(ACC)和 ROC(receiver operating characteristic)曲线下的面积(AUC)作为评价指标. 假设 TP, TN, FP, FN 分别表示真正例、真反例、假正例、假反例的数量, 真正例率(TPR)和假正例率(FPR)可表示为

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{16}$$

$$\text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}}. \tag{17}$$

二分类问题的 AUC 可通过以 FPR 为横坐标、TPR 为纵坐标绘制出的 ROC 曲线下的面积计算得出; 多分类的 AUC 通过计算多个种类的二分类 AUC 的均值得出. ACC 通过计算所有预测结果中正确分类的比例得出, 可表示为

$$ACC = \frac{TP + TN}{TP + FN + FP + TN}. \quad (18)$$

### 3.3 基线模型设置

实验的对比模型主要分为传统机器学习模型、自然语言处理模型和深度学习模型 3 类: 1) 传统机器学习模型包括决策树、随机森林(RF)和支持向量机(SVM), 本文决策树采用梯度提升决策树(GDBT)实现, 随机森林将生成树的数量设为 100; 2) 自然语言处理模型包括文档嵌入模型(Doc2vec)和双向编码器表示模型(BERT), 本文采用文本模型作为特征提取器, 将提取后的特征输入到多层感知机中进行最终预测; 3) 深度学习模型包括深度神经网络(DNN)和文本卷积神经网络(TextCNN).

### 3.4 对比实验

表 1 列出了本文模型与基线模型的对比实验结果. 实验中每个模型均采用十折交叉验证, ACC 和 AUC 的结果均为交叉验证的平均值.

表 1 本文模型与基线模型对比实验结果

Table 1 Comparison of experimental results between proposed model and baseline models

模型类别	模型名称	AUC	ACC
传统机器学习模型	梯度提升决策树	0.785 2	0.735 2
	随机森林	0.798 6	0.758 3
	支持向量机	0.777 8	0.738 4
自然语言处理模型	文档嵌入模型	0.824 9	0.776 2
	双向编码器表示模型	0.839 5	0.798 5
深度学习模型	深度神经网络	0.855 6	0.803 4
	文本卷积神经网络	0.867 1	0.801 1
本文模型	本文模型	0.894 7	0.837 5

由表 1 可见, 自然语言处理模型的预测结果普遍优于传统机器学习模型, 验证了在文本信息中提取特征的有效性. 同时, 深度学习模型的预测结果普遍优于其他两类模型, 验证了深度学习建模多种复杂特征的能力. 其中, 本文模型的预测结果最佳, 在 AUC 方面, 对比 TextCNN 模型提高约 3.1%, 对比 SVM 模型提高约 15.0%; 在 ACC 方面, 对比 DNN 模型提高约 4.2%, 对比 GDBT 模型提高约 13.9%.

此外, 设置两组消融实验验证本文模型注意力机制与 GRU 的有效性. 消融模型 1 去除模型的注意力层, 仅保留输入层和输出层; 消融模型 2 将输出层的 GRU 替换为 LSTM. 实验结果列于表 2.

表 2 消融实验结果

Table 2 Results of ablation experiment

模型	AUC	ACC
消融模型 1	0.852 5	0.809 3
消融模型 2	0.877 2	0.813 4
本文	0.894 7	0.837 5

由表 2 可见, 消融模型 1 可达到与深度学习基线模型相当的预测性能, 但模型精度远低于包含注意力机制的模型. 消融模型 2 可达到远超其他基线模型的预测性能, 从而证明了本文模型注意力层的有效性. 但其性能低于采用 GRU 作为输出层的模型, 从而证明了输出层的有效性. 两组消融实验结果证明了本文采用的注意力机制与 GRU 模型的有效性.

综上所述, 针对经典机器学习算法(如决策树、随机森林)在建模复杂隐式交互关系时预测准确率较低的问题, 本文提出了一个基于循环神经网络与注意力机制的波动预测模型, 通过引入注意力机制和 GRU 模型, 有效捕捉了影响因素之间的复杂交互关系和时间依赖性. 通过与多组基线模型的对比实验验证了该模型准确率较高. 在波动预测领域, 本文模型展示了其较好的性能和应用潜力.

## 参 考 文 献

- 2025, 55(2): 670-685. (CAO Z M, DING L, HAN J, et al. Super-resolution of Logging Curves with Large-Scale Differences Based on Ensemble Machine Learning [J]. Journal of Jilin University (Earth Science Edition), 2025, 55(2): 670-685.)
- [2] 罗焕坤, 葛一烽, 刘帅. 大语言模型在数学推理中的研究进展 [J]. 计算机工程, 2024, 50(9): 1-17. (LUO H K, GE Y F, LIU S. Research Progress of Large Language Models in Mathematical Reasoning [J]. Computer Engineering, 2024, 50(9): 1-17.)
- [3] 孟繁瑞, 王翔, 俄文娟, 等. 基于 GCN-CS-LSTM 的车辆多模态行驶轨迹预测 [J]. 江苏大学学报(自然科学版), 2024, 45(5): 506-512. (MENG F R, WANG X, E W J, et al. Multi-modal Trajectory Prediction of Vehicles Based on GCN-CS-LSTM [J]. Journal of Jiangsu University (Natural Science Edition), 2024, 45(5): 506-512.)
- [4] 曹瑞阳, 李诗雨, 刘擎超, 等. 基于 CrossFormer 的自动驾驶车辆周边行人轨迹预测 [J]. 江苏大学学报(自然科学版), 2025, 46(3): 276-283. (CAO R Y, LI S Y, LIU Q C, et al. Trajectory Prediction of Pedestrians Around Autonomous Vehicles Based on CrossFormer [J]. Journal of Jiangsu University (Natural Science Edition), 2025, 46(3): 276-283.)
- [5] HAMLIN K R, HAMLIN W A. Faculty Salary as a Predictor of Student Outgoing Salaries from MBA Programs [J]. Journal of Education for Business, 2016, 91(1): 38-44.
- [6] BANSAL U, NARANG A, SACHDEVA A, et al. Empirical Analysis of Regression Techniques by House Price and Salary Prediction [C]//Proceedings of the IOP Conference Series: Materials Science and Engineering. [S. l.]: IOP Publishing, 2021: 012045-1-012045-13.
- [7] NAVYASHREE M, NAVYASHREE M K, NEETU M, et al. Salary Prediction in IT Job Market [J]. International Journal of Computer Science and Engineering, 2019, 7(15): 1-5.
- [8] 潘博, 张青川, 于重重, 等. Doc2vec 在薪水预测中的应用研究 [J]. 计算机应用研究, 2018, 35(1): 155-157. (PAN B, ZHANG Q C, YU C C, et al. Research on the Application of Doc2Vec in Salary Forecast [J]. Application Research of Computers, 2018, 35(1): 155-157.)
- [9] JACKMAN S, REID G. Predicting Job Salaries from Text Descriptions [D]. Vancouver: University of British Columbia, 2013.
- [10] 张浩宇. 基于文本相似度与协同过滤的岗位薪资预测 [D]. 广州: 中南财经政法大学, 2018. (ZHANG H Y. Job Salary Prediction Based on Text Similarity and Collaborative Filtering [D]. Guangzhou: Zhongnan University of Economics and Law, 2018.)
- [11] MENG Q X, ZHU H S, XIAO K L, et al. Intelligent Salary Benchmarking for Talent Recruitment: A Holistic Matrix Factorization Approach [C]//Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM). Piscataway, NJ: IEEE, 2018: 337-346.
- [12] 孙莹, 章玉婷, 庄福振, 等. 基于集合效用边际贡献学习的可解释薪酬预测算法 [J]. 计算机研究与发展, 2024, 61(5): 1276-1289. (SUN Y, ZHANG Y T, ZHUANG F Z, et al. Interpretable Salary Prediction Algorithm Based on Marginal Contribution Learning of Collective Utility [J]. Journal of Computer Research and Development, 2024, 61(5): 1276-1289.)
- [13] WANG Z, SUGAYA S, NGUYEN D P T. Salary Prediction Using Bidirectional-GRU-CNN Model [C]//Proceedings of the 25th Annual Meeting of the Association for Natural Language Processing. [S. l.]: ANLP, 2019: 231-236.
- [14] KHONGCHAI P, SONGMUANG P. Implement of Salary Prediction System to Improve Student Motivation Using Data Mining Techniques [C]//Proceedings of the 2016 International Conference on Knowledge, Information and Creativity Support Systems (KICSS). Piscataway, NJ: IEEE, 2016: 1-6.

(责任编辑: 韩 啸)