E-Commerce Letters 电子商务评论, 2025, 14(9), 319-329
Published Online September 2025 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2025.1492916

Hans汉斯

融合注意力机制与图神经网络的电商虚假评论识别研究

王　静，肖　创

武汉科技大学管理学院，湖北　武汉

收稿日期：2025年7月25日；录用日期：2025年8月8日；发布日期：2025年9月2日

摘　要

针对现有图神经网络模型在电商虚假评论识别任务中对关键信息聚集不足的问题，本文通过对CARE-GNN模型引入节点基于点积的注意力聚合，提出一种融合注意力机制的Att-CARE-GNN改进模型。相较于原始模型，本文所改进的模型能够在聚合阶段为目标节点的邻居节点动态赋予聚合权重，从而优化特征选择过程，进而增强模型对电商等虚假评论的识别效果。本文基于YelpChi以及Amazon数据集进行实验，对比CARE-GNN等现有基准模型，结果表明，Att-CARE-GNN在F1值、准确率上表现优异，在YelpChi数据集上F1值与准确率分别提升1.7%、4.9%以上，在Amazon数据集上，F1值与准确率分别提升0.3%、1.1%以上，验证了注意力机制在抑制噪声干扰、提升关键特征权重分配方面的有效性。本文为电商平台的虚假评论识别提供了更具鲁棒性和可解释性的解决方案。

关键词

图神经网络，注意力机制，电商虚假评论识别

Research on the Identification of False Reviews in E-Commerce by Integrating Attention Mechanism and Graph Neural Network

Jing Wang, Chuang Xiao

School of Management, Wuhan University of Science and Technology, Wuhan Hubei

Received: Jul. 25th, 2025; accepted: Aug. 8th, 2025; published: Sep. 2nd, 2025

Abstract

Aiming at the problem of insufficient aggregation of key information in the task of identifying false

文章引用: 王静, 肖创. 融合注意力机制与图神经网络的电商虚假评论识别研究[J]. 电子商务评论, 2025, 14(9): 319-329. DOI: 10.12677/ecl.2025.1492916

reviews in e-commerce by the existing graph neural network models, this paper introduces node attention aggregation based on dot product to the CARE-GNN model and proposes an improved Att-CARE-GNN model integrating the attention mechanism. Compared with the original model, the improved model proposed in the paper can dynamically assign aggregation weights to the neighbor nodes of the target node in the aggregation stage, thereby optimizing the feature selection process and further enhancing the model's recognition effect on false reviews in e-commerce and other fields. This paper conducts experiments based on the YelpChi and Amazon datasets and compares existing benchmark models such as CARE-GNN. The results show that Att-CARE-GNN performs excellently in F1 value and accuracy. On the YelpChi dataset, the F1 value and accuracy increase by more than 1.7% and 4.9% respectively. On the Amazon dataset, the F1 value and accuracy rate have increased by more than 0.3% and 1.1% respectively, verifying the effectiveness of the attention mechanism in suppressing noise interference and improving the distribution of key feature weights. This article provides a more robust and interpretable solution for identifying false reviews on e-commerce platforms.

Keywords

Graph Neural Network, Attention Mechanism, Identification of False Reviews in E-Commerce

1. 引言

随着电子商务行业的发展，线上用户评论在消费者决策过程中扮演着越来越关键的角色，消费者在选购商品或服务时，会参考其他用户的评价来判断其质量、性价比等诸多方面。正是由于用户评论的重要影响力，一些不良商家、竞争对手或利益相关方为了谋取不正当利益，蓄意制造虚假评论。对于海量增长的评论数据，人工审核的方式效率低下，故而借助计算机技术达成自动化的虚假评论识别意义重大。

2008 年，Jindal 率先提出了虚假评论这一概念，此后，越来越多的学者围绕如何精准识别虚假评论展开了全方位的研究工作[1]。如今，虚假评论识别方法主要包括三种：机器学习方法、经典深度学习方法以及基于图神经网络的方法。

基于机器学习的虚假评论识别方法需要设计并提取文本特征，并将特征输入支持向量机、朴素贝叶斯等传统的分类器中对虚假评论进行分类。挖掘文本的语言及文本特征并结合机器学习的方法是虚假评论识别研究早期的重要领域，如 Li 等利用词袋特征、词性特征，结合 SVM 分类器对众包平台生成的数据集中的虚假评论进行检测，实现了良好的检测效果[2]。基于深度学习的方法能够自动提取多层次特征，受到更加广泛的应用。如 Hajek 利用 DFNN (Deep Feed-Forward Neural Network)与 CNN (Convolutional Neural Network)两种不同的深度学习模型整合词的语义特征与情感特征，对虚假评论进行检测[3]。曾致远等将评论文本拆分为首、中、尾三部分，使用自注意力机制将三个局部表示编码成一个全局特征表示，所建立模型识别的平均准确率和平均精度均有所提高[4]。

图神经网络(Graph Neural Networks, GNN)是深度学习的子领域，近年出现的图神经网络具有鲁棒性较好、可有效聚合邻域特征的优点，成为虚假评论识别任务的热点方法。学界学者根据不同的任务情景，设计图神经网络模型并取得良好效果。如曹东伟基于词和文档构建的图神经网络进行文本分类的基础上，提出基于融合语义相似度的图卷积网络(Sematicgraph Convolution Networks)的虚假评论检测方法，在公开

数据集上，模型识别准确率有较大提升[5]。Yao 等首次提出的基于图卷积神经网络文本分类模型 Text-GCN (Text-graph Convolution Networks)，以文档中的词为节点构建异构图，对基于评论内容的虚假评论检测展开研究[6]。除了评论文本信息，学者还利用其他辅助信息检测虚假评论，如 Wang 等首次提出利用评论、评论者、商品之间的关系构建异质网络图，通过挖掘异质网络图中的特征检测虚假评论[7]。而 Dou 等人创建的 CARE-GNN (CAmouflage-REsistant Graph Neural Network)模型，是首个考虑关系伪装和特征伪装的基于空间的 GNN 模型[8]。Song 等提出一种基于动态传播图的虚假评论检测方法，通过捕捉静态网络中缺失的动态传播信息，均取得较好识别效果[9]。此外学者们还对现有模型不断改进，如袁紫烟将 TrustRank 算法与 GraphSAGE 模型结合，以此改进 GraphSAGE 模型的随机采样策略，形成 TR-GraphSAGE 模型，模型性能取得较大提升[10]。其中，注意力机制可有效消除模型数据噪声影响，因此融合注意力机制也成为图神经网络模型改进的一个重要方法。如张蓉等构建了基于层次注意力机制与异构图注意力网络的层次异构注意力网络模型，在 Yelp 网站的酒店数据集上的检测效果优于 CNN 等图神经网络基准模型[11]。Shang 等在图卷积网络中使用多头注意力机制，使用注意力机制去计算异构节点的嵌入，节点在不同的关系链接下，每个注意力函数考虑由特定的链接类型定义的邻居，模型具有良好性能[12]。颜梦香等利用用户视图和产品试图的注意力机制对评论文本进行建模，建立出一种基于层次注意力机制的神将网络模型，在 Yelp 数据集上识别准确率相比于传统离散模型提高了一至四个百分点[13]。

基于图神经网络的虚假评论识别模型具有更出色的性能，但易受噪声数据干扰的特点，本文对现有模型进行改进，为现有 CARE-GNN 模型引进注意力机制消除节点邻域噪声干扰，由于 CARE-GNN 模型强化学习(Reinforcement Learning, RL)选择机制是基于节点间欧氏距离进行的，若在聚合阶段继续采用基于欧式距离的权重聚合，将会使得权重趋于平均，因此本文在聚合阶段采用基于节点点积的注意力聚合方式为目标节点的邻居节点分配聚合权重，形成 Att-CARE-GNN 模型，并在 Yelp 以及 Amazon 电商平台公开数据集上融合多种视图结构进行实验，以检验模型性能。

2. 基于注意力机制的虚假评论识别模型

2.1. 传统图神经网络

图神经网络其核心思想是通过消息传递(Message Passing)机制实现节点间拓扑关系与特征信息的协同学习。在图关系网络中，图通常使用二元数组 $G=(V,E)$ 表示，其中 V 是节点的集合，E 是边的集合，每个节点 $v\in V$，有其特征向量 X_i，这些特征向量包含了与该节点相关的各种信息，进行虚假评论识别的节点分类任务中，首先需要将图的拓扑结构：邻接矩阵 $A\in\{0,1\}^{N\times N}$ (N 为节点数)、节点特征矩阵 X，以及部分节点的标签数据输入至模型中，接着在模型的卷积层，聚合邻居信息更新节点表示，公式表达为(1)：

$$H^{(l)}=\sigma\left(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}H^{(l-1)}W^{(i)}\right) \tag{1}$$

其中 $\hat{A}=A+I$ (添加自环)，$\hat{D}$ 为度矩阵。$H^{(l)}$ 是第 l 层的节点表示，$W^{(i)}$ 是参数矩阵。σ 是非线性激活函数。通过多层图神经网络的堆叠，节点可以逐步聚合更远处邻居的信息，节点表示映射到更有利于分类的新特征空间，最后对每一个节点运用于分类函数(如 Softmax)，得出其类别预测标签 $\hat{y}_v$。

2.2. CARE-GNN 模型

在基于图神经网络的对抗性攻击场景下，欺诈者常采用特征伪装以及关系伪装的手段隐藏自己的欺诈行为，例如虚假买家(特征伪装：模仿真实用户购买记录)通过大量关注和虚假交易(关系伪装)提升群体信誉度，严重破坏图结构完整性，欺骗基于图的信誉模型，致使传统 GNN 受邻居噪声干扰而性能降低，

而作为新型可处理多关系网络的图神经网络欺诈检测模型 CARE-GNN，是专门针对当前欺诈检测领域的核心挑战而设计的。为解决这一挑战，CARE-GNN 创新性地构建了协同工作机制，设计了三大关键模块以识别欺诈者的伪装：标签感知的相似性度量、强化学习邻居选择以及层次化聚合模块。

标签感知的相似性度量模块针对特征伪装，通过结合节点特征和已知标签，计算节点间的相似性，避免仅依赖原始特征导致误判。模型利用稀疏已知标签作为监督信号，采用多层感知机(MLP)实现标签预测，同时结合特征向量的欧氏距离度量，构建相似性矩阵，对每种关系采用 top-p 采样，将邻居节点按照相似度降序排列，根据设定的邻居节点选择阈值过滤不相似节点，即实现了过滤与目标节点语义不一致的伪装邻居；在强化学习邻居选择模块针对关系伪装，通过强化学习动态选择信息量最大的邻居节点，过滤欺诈者构造的虚假关联(如异常边连接)。具体采用强化学习寻找最优邻居选择阈值，首先计算一个 epoch 内节点的平均距离，利用强化学习来优化这个阈值，奖赏定义为两个 epoch 间平均距离的变化，当平均距离增大时为负值(记为−1)，减小时为正值(记为+1)，当满足终止条件(10 个 epoch 内总波动小于 2)时，停止强化学习过程，此时得到的阈值即为最优邻居节点选择个数阈值，从而实现动态调整每种关系下的邻居采样数量；层次化聚合模块针对结构隐藏，通过分层聚合局部和全局信息，避免欺诈者通过局部伪装逃避检测，使节点在多层 GNN 中逐层筛选邻居并聚合信息，逐步增强节点表示，最终经多关系融合后，输出具有强判别性的节点嵌入。

下图 1 为对于目标节点 V_0 在 CARE-GNN 模型中的流程图，其中 r_1、r_2、r_3 表示不同的连接关系。

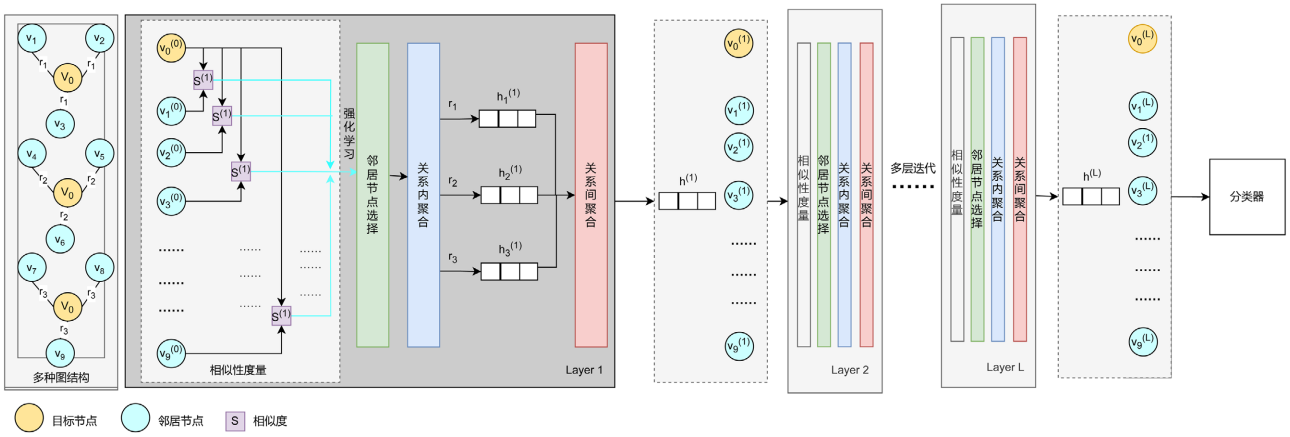

Figure 1. Flowchart of the CARE-GNN model
图 1. CARE-GNN 模型流程图

模型初始阶段，输入包含 V_0 节点的多关系图，目标节点首先经过相似性度量模块，计算与各关系图中邻居节点相似性，结合强化学习策略，为每种关系筛选出最合适的邻居节点，随后，通过关系内聚合模块处理，生成第一层对应于各关系的节点表示矩阵 $h_r^{(1)}$ $(r=1,2,3)$，接着，进行关系间聚合，得到各节点更新后的表示矩阵 $h^{(1)}$。经过多层迭代后，最终获得节点的表示向量 $h^{(L)}$。将此表示向量输入分类器，即可获得各节点的类别预测结果。

2.3. Att-CARE-GNN 模型的构建

注意力机制(Attention Mechanism)是一种模拟人类视觉和认知系统选择性聚焦的信息处理机制。其数学本质是通过动态权重分配，赋予模型对输入数据中不同部分的重要性差异进行建模的能力。相较于传统神经网络的静态参数分配，注意力机制的核心优势在于动态性与可解释性，模型能够根据具体任务需求，自主决定关注输入数据特征。而 CARE-GNN 模型在消息传递阶段默认使用基于设定阈值的加权聚合

或者是平均聚合，这种静态的聚合方式对所有的邻居节点特征维度平等地处理，无法区分不同邻居节点对目标节点的贡献差异，本文对强化学习模块动态选择出的邻居节点，设计多关系类型的基于节点点积的注意力聚合模块，实现 Att-CARE-GNN 模型。

注意力聚合的核心是通过计算节点之间的注意力系数来动态分配权重，在本文建立的注意力聚合模块，首先提取各关系图中的节点特征，构建中心节点与邻居节点的联合特征矩阵，将中心节点特征根据邻居节点数量进行复制，使其与邻居特征维度匹配，以便拼接以及后续作线性映射。拼接操作公式表达为(2)：

$$\text{combined}_{i,r} = \left[h_i \,\|\, h_{i,r}^{\text{neigh}} \right] \tag{2}$$

其中 h_i 表示中心节点 i 的嵌入特征，$h_{i,r}^{\text{neigh}}$ 表示节点 i 在关系 r 下的邻居特征，$\text{combined}_{i,r}$ 是节点 i 在关系 r 下的拼接特征，对拼接特征作一个可学习的线性组合，经激活函数得出注意力得分，公式为(3)：

$$e_{i,r} = \text{LeakyReLU}\left(\boldsymbol{a}^{\top} \cdot \text{combined}_{i,r} \right) \tag{3}$$

$\boldsymbol{a}$ 是可学习的注意力向量，最后将所有关系的原始注意力得分进行 Softmax 归一化，即可为邻居节点动态分配权重，实现节点关系内聚合以及关系间聚合，公式为(4)。在训练模型时对注意力系数应用 Dropout 防止过拟合，总公式为(5)：

$$\alpha_{i,r}^{\text{norm}} = \frac{\exp\left(e_{i,r} \right)}{\sum_{r'=1}^{R} \exp\left(e_{i,r'} \right)} \tag{4}$$

$$\alpha_{i,r}^{\text{norm}} = \text{Dropout}\left(\text{Softmax}_{\text{r}} \left(\text{LeakyReLU}\left(\boldsymbol{a}^{\top} \cdot \left[h_i \,\|\, h_{i,r}^{\text{neigh}} \right] \right) \right) \right) \tag{5}$$

其中 Dropout 仅在模型训练阶段使用，$\alpha_{i,r}^{\text{norm}}$ 是归一化后的注意力系数，R 是关系总数。

3. 基于图神经网络的虚假评论识别实验设置

3.1. 实验环境配置

实验环境为 Windows 11 家庭中文版，开发语言为 Python3.7，编程环境为 Pycharm，实验主要使用 Python 库及版本为：Pandas = 2.0.3；torch = 1.4.0；Numpy = 1.16.4；scipy = 1.2.1；scikit_learn = 1.2.2。

3.2. 数据集来源

本文采用的数据集为公共基准数据集 YelpChi 以及 Amazon。

YelpChi 数据集包含了酒店和饭店两大领域的评论。YelpChi 数据集以具有 100 维特征的评论为节点设计了三种关系网络结构：(1) net_rur，以同一用户发表的不同评论为联系建立三种关系网络结构。(2) net_rtr，以同一商品获得相同评分的评论建立关系网络结构。(3) net_rsr，以同一商品在特定时间间隔内获得的评论建立关系网络结构。

Amazon 数据集包含 Amazon 平台上乐器类商品的用户评论。Amazon 数据集以具有 100 维特征的用户为节点建立三种关系网络结构：(1) net_upu，以对至少一种相同商品评论的用户建立关系网络结构。(2) net_usu，以一定时间内对商品有相同评分的用户建立关系网络结构。(3) net_uvu，以相似性排名靠前的用户建议关系网络结构。

YelpChi 与 Amazon 数据集均为不平衡数据集，YelpChi 数据集共有 45,954 个标注节点，其中包含虚假节点数为 6677，真实节点数为 39,277，虚假节点占比约为 14.53%，Amazon 数据集共有 8639 个标注节点，其中包含虚假节点数为 821，真实节点数为 7818，虚假节点占比约为 9.50%。两数据集基本信息如表

1 所示，特征相似度是边两端节点特征的平均相似性，在本文中是由节点之间的欧氏距离得出，标签相似度是边两端节点标签一致的比例。

Table 1. Dataset information table
表 1. 数据集信息表

数据集	关系网络	节点数量	边数量	特征相似度	标签相似度
YelpChi	net_rur	23,831	98,630	0.99060	0.90890
	net_rtr	45,432	1,147,232	0.98795	0.17636
	net_rsr	45,914	6,805,416	0.98783	0.18574
Amazon	net_upu	10,244	351,216	0.71066	0.16731
	net_usu	11,854	7,132,958	0.68663	0.05576
	net_uvu	11,863	2073474	0.69691	0.05316

3.3. 模型性能评价指标

在论文所进行的虚假评论识别任务中，主要运用了一下模型评价指标：

(1) 准确率

准确率衡量的是模型正确预测的样本数占总样本数的比例，计算公式为(6)：

$$\text{Accuracy} = \frac{\text{正确预测节点数}}{\text{总节点数}} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{6}$$

其中 TP 是正确预测为正类的样本数，TN 是正确预测为负类的样本数，FP 是错误预测为正类的样本数，FN 是错误预测为负类的样本数。

(2) F1 值

F1 分数是一种综合评估分类模型精确率和召回率的指标。计算公式为(7)：

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

(3) AUC 值

AUC 值是衡量分类模型性能的重要指标之一。计算公式为(8)：

$$\text{AUC} = \frac{\sum_{i \in M} rank_i - \frac{|M| \times (1 + |M|)}{2}}{|M| \times |N|} \tag{8}$$

其中 M 为正样本集合，N 为负样本集合，$|M|$为正样本数量，$|N|$为负样本数量，$rank_i$ 表示第 i 个正样本在所有样本集合中的排序值。

3.4. 模型参数设置

为了更快地收敛到全局最优解，论文基于交叉熵损失采用 adam 算法进行优化训练，在模型训练开始时的学习率为 0.01，同时相似性损失权重为 2，节点嵌入大小 64。指定每训练 3 轮后在测试集上评估模型性能，以减少计算成本并及时检测模型性能。处理不平衡数据集时使用多数类欠采样的方法，将多数类样本与少数类样本数量调整至一致，强化学习中动作的步长大小为 2e−2。YelpChi 数据集的 batch 的大小设置为 1024，Amazon 数据集的 batch 的大小设置为 256。模型训练的总轮数 60 个 epoch。为消除随

机因素对实验的影响，保存评估模型的稳定性和一致性，此次将实验重复进行 5 次，取模型评价指标的平均值，得到一个综合的性能度量。考虑数据集划分对实验的影响，将数据集按照 7:2:1、8:1:1 的比例划分训练集、验证集、测试集。

3.5. 消融实验设计

论文在 CARE-GNN 框架中引入注意力聚合模块，设计系统性消融实验评估其有效性。将改进后的模型与原始 CARE-GNN 模型(默认聚合方式)以及基于均值聚合 CARE-GNN(mean)进行对比分析，以验证注意力机制对模型性能的影响。

3.6. 对比实验设计

论文设置传统的图神经网络模型 GCN、GAT，以及具有代表性邻居采样策略的 GraphSAGE 模型作为对比实验。

GCN：图卷积神经网络(Graph Convolutional Network)，扩展了传统卷积神经网络(CNN)的思想，使其能够对非欧几里得空间中的图数据进行特征学习和表示。

GAT：图注意网络(Graph Attention Network)，基于注意力机制的图神经网络，在聚合阶段使用注意力机制，模型自动学习图结构中节点间的动态依赖关系，从而更有效处理图数据。

GraphSAGE：图采样和聚合方法(Graph Sample and Aggregate)，通计设计采样和聚合的策略，突破了传统 GNN 对全图邻接矩阵的依赖，属于一种归纳学习算法，它通过学习一种聚合函数，聚合节点邻居的特征信息来学习目标节点本身的嵌入表达。

4. 实验结果及分析

4.1. 损失值结果与分析

损失值是量化模型预测的节点类别概率分布与真实标签的差异，反映分类准确率。如图 2、图 3 分别展示了 Att-CARE-GNN 模型在 YelpChi、Amazon 数据集上不同训练集比例划分的训练损失随 epoch 变化情况，其中图 2、图 3 中(a)、(b)子图分别为模型在各自数据集 70%、80%的训练集比例上损失值变化图。

由图 2、图 3 可见，Att-CARE-GNN 模型在 YelpChi 以及 Amazon 数据集上损失值总体下降并趋于收敛，表明模型在训练集上稳定学习。在图 2(a)中模型损失值在前 15 个 epoch 稳定下降，幅度约为 5.0%，之后进入动态调整阶段，在最后 12 个 epoch 逐渐收敛于 0.378 左右，在图 2(b)中损失值在前 15 个 epoch

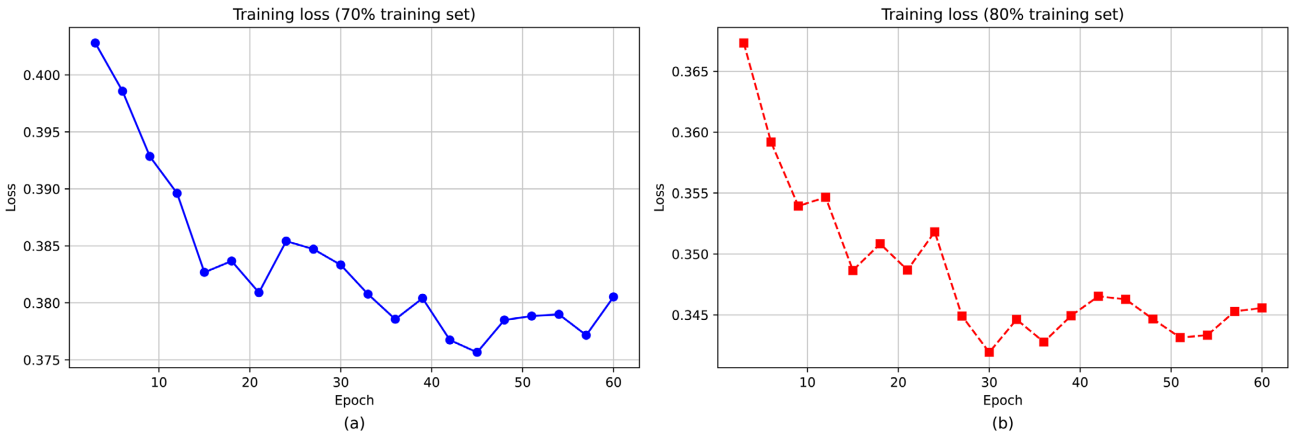

Figure 2. Loss value change graph (YelpCHi)
图 2. 损失值变化图(YelpCHi)

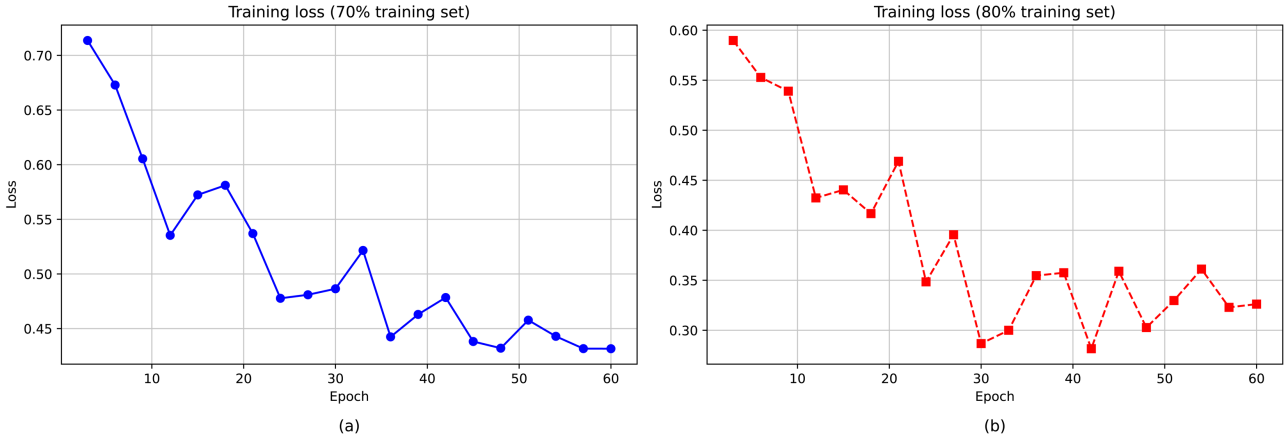

Figure 3. Loss value change graph (Amazon)
图 3. 损失值变化图(Amazon)

快速下降，下降幅度约为 5.6%，之后进入动态调整，最后趋于 0.345。在图 3(a)与图 3(b)中损失值同样经过快速下降期、动态调整期并最终分别趋于 0.453、0.303。对比两数据集损失值变化情况，可知随着训练量增大，模型损失降幅更大，训练过程更鲁棒。

4.2. 消融结果及分析

经多次实验，得出 CARE-GNN 与 CARE-GNN(mean)消融实验模型平均性能指标，与本文所建立的 Att-CARE-GNN 模型在不同数据集上对比结果如图 4、图 5 所示，其中图 4 与图 5(a)、图 5(b)子图分别为各消融实验模型与 Att-CARE-GNN 模型在 YelpChi 以及 Amazon 数据集 70%、80%的训练集划分下模型性能评价指标图。

由图 4、图 5 可得，本文所建立的 Att-CARE-GNN 模型在 YelpChi 与 Amazon 数据集，三大模型评价指标 F1、Accuracy、AUC 值上均有优异表现。在图 4(a)中，本文所建立的 Att-CARE-GNN 模型相较于使用默认聚合方式的 CARE-GNN 模型，F1 值提升 1.7%；Accuracy 提升 4.9%，准确率得到了较大的提升，这受益于注意力聚合使中心节点聚合了更多邻居节点重要信息；AUC 值略微降低，这可能是因为

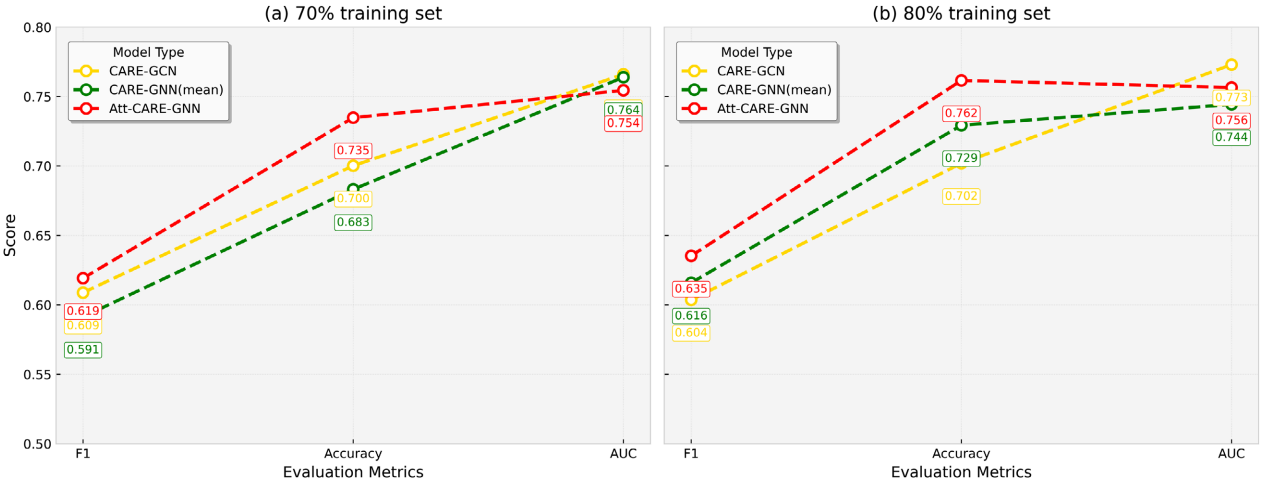

Figure 4. Comparison chart of evaluation indicators of ablation experiment model (YelpChi)
图 4. 消融实验模型评价指标对比图(YelpChi)

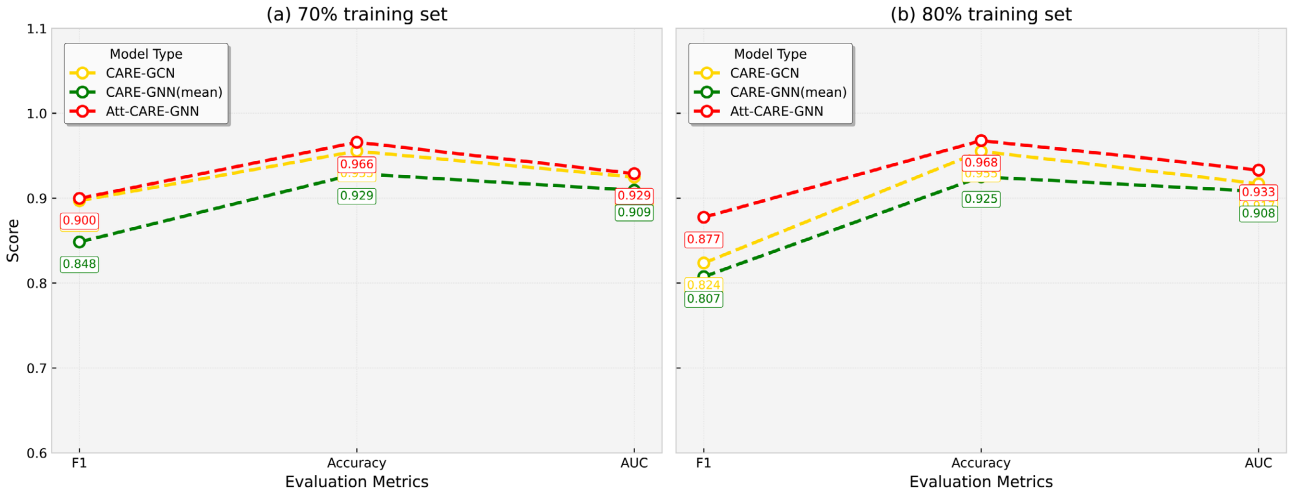

Figure 5. Comparison chart of evaluation indicators of ablation experiment model (Amazon)
图 5. 消融实验模型评价指标对比图(Amazon)

数据集不平衡的原因，虚假评论数量较少，AUC 值对正样本的敏感性被放大，图 4(b)中展现出模型更高的性能提升。图 5 则展示了 Att-CARE-GNN 模型在三模型评价指标上对比消融实验模型均具有一定提升，其中子图(a)展示模型 F1 值提升 0.3%，准确率提高 1.1%，AUC 值也有略微提升，子图(b)展示出模型性能有较高提升。

4.3. 对比实验结果与分析

经多次实验，得出各对比实验模型模型平均性能指标，与本文所建立的 Att-CARE-GNN 模型在不同数据集上对比结果如图 6、图 7 所示，其中图 6 与图 7(a)、图 7(b)子图分别为各对比实验模型与 Att-CARE-GNN 模型在 YelpChi 以及 Amazon 数据集 70%、80%的训练集划分下模型性能评价指标图。

由图 6、图 7 可得，Att-CARE-GNN 模型在 F1、Accuracy、AUC 值上大幅度超越对比模型，其中 GNN 与 GAT 模型各项评价指标均有较低表现，这可能是因为 GNN 与 GAT 模型面对不平衡数据集时，邻居采样不均匀，少数类节点的信息被忽略，而 GraphSAGE 模型具有合理的采样策略，从而表现良好。

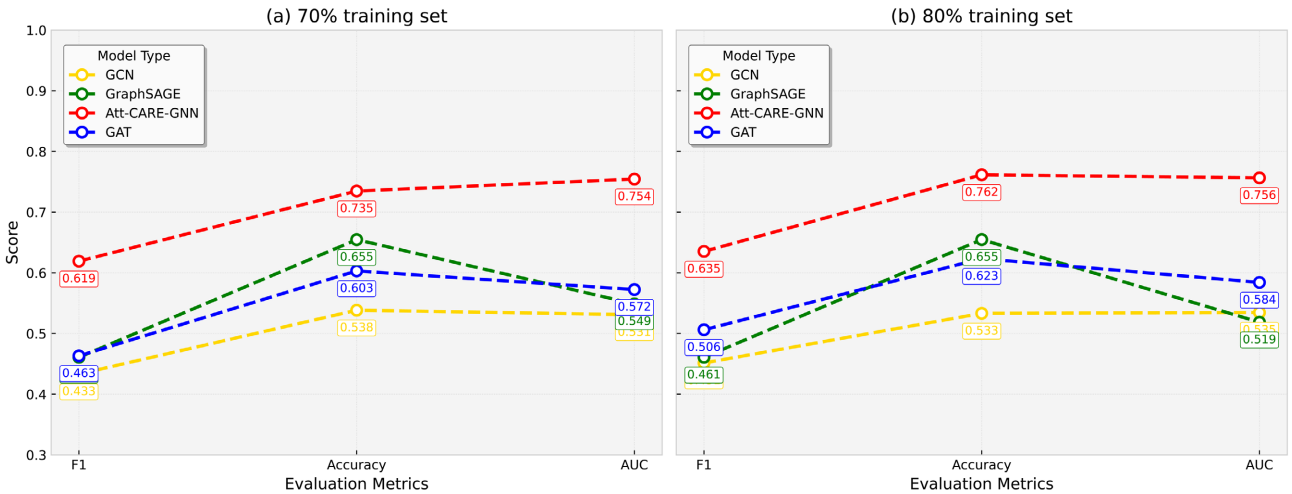

Figure 6. Comparison chart of evaluation indicators of comparative experimental models (YelpChi)
图 6. 对比实验模型评价指标对比图(YelpChi)

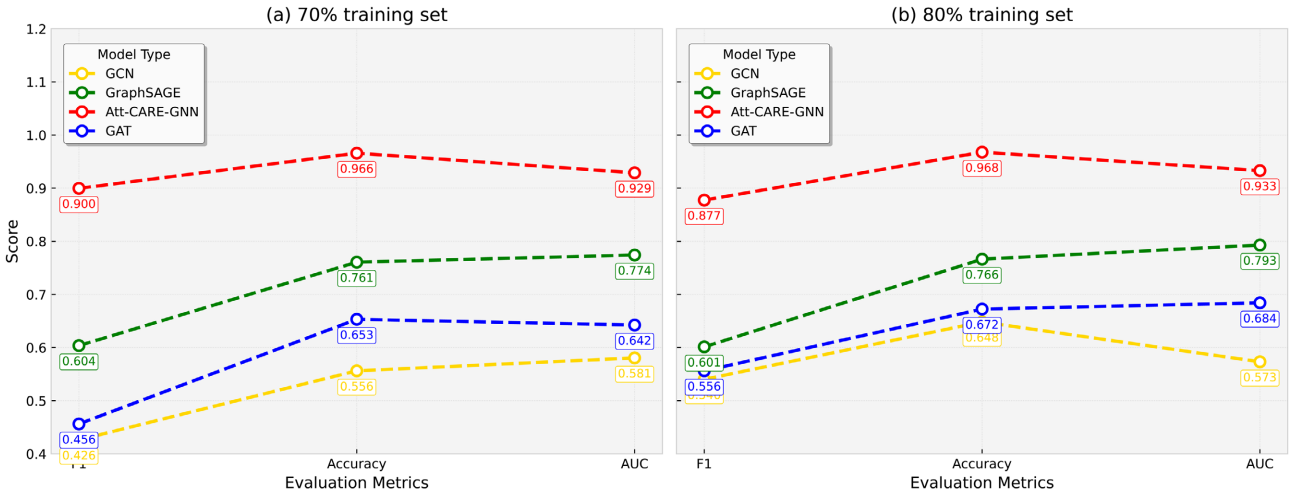

Figure 7. Comparison chart of evaluation indicators of comparative experimental models (Amazon)
图 7. 对比实验模型评价指标对比图(Amazon)

4.4. 实验结果汇总

本文设计消融实验以及对比实验，通过多次实验并取各模型平均评价指标验证本文建立的 Att-CARE-GNN 模型性能，各模型实验指标结果汇总于表 2 中。

Table 2. Summary table of evaluation indicators for each model
表 2. 各模型评价指标汇总表

Model	YelpChi (70%)			YelpChi (80%)			Amazon (70%)			Amazon (80%)		
	F_1	Accuracy	AUC	F_1	Accuracy	AUC	F_1	Accuracy	AUC	F_1	Accuracy	AUC
Att-CARE-GNN	0.6192	0.7348	0.7545	0.6353	0.7616	0.7565	0.8997	0.9659	0.9289	0.8775	0.9678	0.9330
CARE-GNN	0.6088	0.7002	0.7660	0.6036	0.7020	0.7731	0.8970	0.9554	0.9248	0.8236	0.9552	0.9170
CARE-GNN (mean)	0.5914	0.6832	0.7639	0.6159	0.7293	0.7444	0.8484	0.9288	0.9093	0.8075	0.9252	0.9080
GCN	0.4332	0.5384	0.5310	0.4514	0.5332	0.5346	0.4262	0.5562	0.5805	0.5401	0.6482	0.5732
GAT	0.4632	0.6032	0.5724	0.5062	0.6234	0.5842	0.4562	0.6532	0.6424	0.5562	0.6724	0.6842
GraphSAGE	0.4608	0.6547	0.5488	0.4608	0.6548	0.5187	0.6037	0.7607	0.7742	0.6013	0.7664	0.7929

由实验可得，本文所建立的模型 Att-CARE-GNN 在虚假评论识别任务上具有良好的效果，模型验证了注意力机制在抑制噪声干扰、提升关键特征权重分配方面的有效性。本研究为在线评论平台的虚假内容检测提供了更具鲁棒性和可解释性的解决方案。

5. 总结与展望

本文致力于构建一种高效的虚假评论检测方法，以助力消费者进行理性消费决策。论文基于 CARE-GNN 模型创新性地提出基于注意力机制的 Att-CARE-GNN 模型，该模型实现节点间注意力权重分配，从而有效捕获图数据中的高阶结构特征与复杂交互模式。实验环节设计消融实验和对比实验，在虚假评论检测任务中进行多维度性能评估。实证结果表明，提出的 Att-CARE-GNN 模型在关键评估指标上表现突出，对比原始 CARE-GNN 模型，在 YelpChi 数据集上 F1 值与准确率分别提升 1.7%、4.9%以上，在 Amazon

数据集上，F1 值与准确率分别提升 0.3%、1.1%以上。

然而，本研究也存在一定局限性，如本文所构建的模型属于有监督学习，实验数据标签依赖于人工标注，不仅标注成本较高，同时存在人工误判的风险，因此，未来重要的研究方向可能是建立使用少量标签数据的半监督学习模型或无标签数据的无监督模型，以进一步提升模型在实际应用场景中的适用性。

参考文献

[1] Jindal, N. and Liu, B. (2008) Opinion Spam and Analysis. *Proceedings of the International Conference on Web Search and Web Data Mining*, Palo Alto, 11-12 February 2008, 219-230. https://doi.org/10.1145/1341531.1341560

[2] Li, J., Ott, M., Cardie, C. and Hovy, E. (2014) Towards a General Rule for Identifying Deceptive Opinion Spam. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Volume 1, 1566-1576. https://doi.org/10.3115/v1/p14-1147

[3] Hajek, P., Barushka, A. and Munk, M. (2020) Fake Consumer Review Detection Using Deep Neural Networks Integrating Word Embeddings and Emotion Mining. *Neural Computing and Applications*, **32**, 17259-17274. https://doi.org/10.1007/s00521-020-04757-2

[4] 曾致远, 卢晓勇, 徐盛剑, 等. 基于多层注意力机制深度学习模型的虚假评论检测[J]. 计算机应用与软件, 2020, 37(5): 177-182.

[5] 曹东伟, 李邵梅, 陈鸿昶. 基于 GCN 的虚假评论检测方法[J]. 计算机工程与应用, 2022, 58(3): 181-186.

[6] Yao, L., Mao, C. and Luo, Y. (2019) Graph Convolutional Networks for Text Classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 7370-7377. https://doi.org/10.1609/aaai.v33i01.33017370

[7] Wang, G., Xie, S., Liu, B. and Yu, P.S. (2011) Review Graph Based Online Store Review Spammer Detection. 2011 *IEEE 11th International Conference on Data Mining*, Vancouver, 11-14 December 2011, 1242-1247. https://doi.org/10.1109/icdm.2011.124

[8] Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H. and Yu, P.S. (2020). Enhancing Graph Neural Network-Based Fraud Detectors against Camouflaged Fraudsters. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 19-23 October 2020, 315-324. https://doi.org/10.1145/3340531.3411903

[9] Song, C., Teng, Y., Zhu, Y., Wei, S. and Wu, B. (2022) Dynamic Graph Neural Network for Fake News Detection. *Neurocomputing*, **505**, 362-374. https://doi.org/10.1016/j.neucom.2022.07.057

[10] 袁紫烟, 任勋益, 黄家铭. 一种改进图神经网络的虚假评论检测方法[J]. 软件导刊, 2024, 23(3): 27-33.

[11] 张蓉, 张献国. 基于层次异构图注意力网络的虚假评论检测[J]. 计算机应用, 2021, 41(5): 1275-1281.

[12] Shang, C., Liu, Q., Chen, K.S., *et al*. (2018) Edge Attention-Based Multi-Relational Graph Convolutional Networks.

[13] 颜梦香, 姬东鸿, 任亚峰. 基于层次注意力机制神经网络模型的虚假评论识别[J]. 计算机应用, 2019, 39(7): 1925-1930.