

基于跨语言图神经网络模型的属性级情感分类

范震廉

三明医学科技职业学院 福建 三明 365000

[摘要]随着互联网的飞速发展,各种社交媒体平台上积累了大量的评论数据,其中包含大量关于商品、品牌、商家的评论,这些评论可以用来辅助用户进行产品购买决策。但是,由于不同语言之间的表达形式存在差异,导致了情感分类任务中语义信息的不匹配。本文提出一种跨语言图神经网络模型(MGNN),该模型在网络中引入了一种新的属性节点表示形式,将评论文本中的语义信息和图结构信息结合起来,构建了一个跨语言属性级情感分类模型。本文在多个公开数据集上进行了实验验证,结果表明该模型在跨语言属性级情感分类任务上取得了最好的效果。

[关键词]跨语言图神经网络;情感分类;跨语言属性级;图神经网络

[中图分类号] G641 **[文献标识码]** A **[文章编号]** 1647-9265 (2024)-0052-12 **[收稿日期]** 2024-02-09

一、引言

为了解决这一问题,本文提出了一种跨语言图神经网络模型(MGNN),该模型能够将评论文本中的语义信息和图结构信息结合起来,将评论文本中的情感属性信息和图结构信息进行融合。该模型利用两种不同的表示形式(图神经网络和图编码器)来获取用户对商品、品牌、商家的情感属性,并在网络中引入一种新的节点表示形式,使得情感属性可以被有效地表示出来。然后,根据不同语言之间的表达形式差异,使用不同的解码器来对用户的情感属性进行预测。实验结果表明,本文提出的跨语言图神经网络模型在情感分类任务中取得了最好效果。

在本文提出的 MGNN 模型中,首先将评论文本中的属性节点进行提取和处理,得到了一个用于表示评论文本中情感属性信息的图表示形式。然后将这种图化形式应用到网络中进行情感分类。为了使网络中不同语

言之间能有效地融合信息,本文使用一种新的节点表示形式来表示评论文本中情感属性信息。最后通过循环神经网络(RNN)将不同语言之间的情感属性进行融合预测。最后,使用训练好的模型进行训练和测试,实验结果表明本文提出的模型能够有效地将评论文本中情感属性信息进行融合预测。

二、相关工作

在情感分类任务中,目前已经有很多相关工作。Bengio 提出了一种基于 TF-IDF 的情感分类方法,该方法基于词对齐来表示情感词。Gu 提出了一种基于注意力机制的跨语言情感分类方法,该方法通过节点在文本中的情感极性来表示文本情感。这些方法虽然取得了一定的效果,但是都没有考虑到语言之间的差异。针对以上问题,许多研究人员提出了一些跨语言属性级情感分类模型,但是这些模型只考虑了一种语言的信息,忽略了其他语言信息。

在属性级情感分类任务中，当前主要的研究方向是利用词向量表示不同语言之间的相似性。例如，Chen 提出了一种基于双向 LSTM 网络和词嵌入相结合的属性级情感分类模型，该模型在整个任务中考虑了词向量和实体嵌入的信息。Lu 提出一种基于词向量表示的跨语言属性级情感分类模型，该模型使用了词向量和图嵌入来捕捉不同语言之间的相似性。在以上方法中，只考虑了一种语言中的信息，没有考虑到其他语言信息。

三、基于 MGNN 的跨语言情感分类模型

图神经网络是一种以节点为中心，边为节点属性的新型神经网络，在计算机视觉、自然语言处理、语音识别等领域中得到广泛的应用。在情感分类任务中，不同的语言表达方式存在差异，图神经网络通过将不同语言中的节点映射到对应的边上来实现跨语言信息的传递。因此，本文提出一种基于图神经网络的跨语言属性级情感分类模型

(MGNN)，将节点表示为图中不同节点之间边的数量、类型和方向等信息，通过图卷积神经网络 (GCN) 和图循环网络

(GCN) 分别提取不同语言之间的特征和不同语言之间的情感特征。

根据评论文本中情感词和属性词之间存在的语义关联，本文构建了一个包含多个情感词、属性短语和属性词序列等节点表示形式的图神经网络。在图神经网络中，每个节点对应一个节点表示，所有节点的边可以通过图神经网络来表示。考虑到评论文本中存在着多个属性节点和多个情感词、属性短语

和属性词序列等信息，为了更好地表达这些信息，本文将这些信息转化为具有相同类型、数量、长度和方向等特性的图结构。在构建好的图卷积神经网络 (GCN) 中引入一个门控机制来选择节点类型、位置和边信息等信息，然后将这些信息传递给一个多层全连接神经网络进行编码。在每个隐藏层中，输入一组特征表示后，将这些表示发送给下一个隐藏层进行计算，得到最终输出结果。

四、实验

在本文中，我们使用了两种数据集进行实验验证。数据集 1:来自 Tokenization 网站上的用户评论数据。在这个数据集中，包含了用户对商品或服务的评价。由于该数据集是中文数据集，因此本文使用了中文情感词典对该数据集中的评论进行情感分类，其中包含了大量的负面评价。我们使用了一个预训练的 GCNs 模型进行训练，并在该数据集中使用了三个不同大小的隐藏层。为了测试该模型的性能，我们使用了与其他基线模型相同的损失函数。

在训练过程中，我们将我们的基线模型与 MGNN 进行对比，并使用交叉验证方法来评估我们的模型性能。为了评估 MGNN 模型在不同语言上的性能，我们选择了五个不同语言的中文情感词典进行训练和测试。同时，为了避免语言之间潜在信息的相互影响，我们还选择了一个与其他基线模型相同大小、但不同大小的隐藏层。

为了测试 MGNN 模型在不同语言上的性能，我们采用了两个公开数据集：

Tokenization 网站上用户评论数据集和 Tokenization 网站上用户评论数据集。为了训练和测试我们的模型，我们使用了四种不同大小的隐藏层。此外，为了避免语言之间潜在信息的影响，我们使用了三个不同大小、三个不同语言进行训练和测试。

五、结束语

随着互联网的快速发展，产生了大量的文本评论数据，如何从这些海量评论中提取有用的信息是当前研究的重点和难点。实验结果表明，本文提出的模型在多个公开数据集上均取得了最好的性能，验证了本文提出的跨语言图神经网络模型在属性级情感分类任务上具有一定的优越性。但是，本文提出的跨语言图神经网络模型是基于图卷积网络

提取图结构信息，并没有进一步挖掘文本中不同类型节点之间的关系，因此未来需要进一步挖掘文本中不同类型节点之间的关系来提高模型性能。

参考文献：

[1]王光,李鸿宇,邱云飞,等.基于图卷积记忆网络的方面级情感分类[J].中文信息学报.2021,(8).DOI:10.3969/j.issn.1003-0077.2021.08.013.

[2]王星,单力秋,侯磊,等.基于百科语料的中英文双语词典提取[J].中文信息学报.2021,(1).

[3]彭晓娅,周栋.跨语言词向量研究综述[J].中文信息学报.2020,(2).

Attribute-level emotion classification based on a cross-language graph neural network model

Fan Zhenlian

Sanming Vocational College of Medical Science and Technology, Fujian Sanming 365000

Abstract: With the rapid development of the Internet, a large number of comment data have been accumulated on various social media platforms, including a large number of comments about goods, brands and merchants, which can be used to assist users to make product purchase decisions. However, the differences in expression forms between different languages lead to a mismatch of semantic information in the emotion classification task. This paper proposes a cross-language graph neural network model (MGNN), which introduces a new form of attribute node representation in the network, and combines the semantic information and graph structure information in the comment text to build a cross-language attribute-level emotion classification model. In this paper, the experimental validation on multiple publicly available datasets shows that the model achieves the best results on cross-language attribute-level emotion classification tasks.

Key words: cross-language graph neural network; emotion classification; cross-language attribute level; graph neural network.