

摘要

深度学习已经成为一种强大的机器学习技术，它可以学习数据的多层表示或特征，并产生最先进的预测结果。随着深度学习在许多其他应用领域的成功，近年来深度学习也被广泛应用于情感分析。本文首先对深度学习进行了概述，然后对其在情感分析中的应用进行了全面的综述。

情感分析

情感分析或意见挖掘是对人们对产品、服务、组织、个人、问题、事件、主题及其属性等实体的意见、情绪、情感、评价和态度的计算研究。

研究者主要从三个粒度层次来研究情感分析:文档层次、句子层次和方面层次。

- **文档级情感分类** 将有观点 的文档(例如产品评论)分类为表达总体积极或消极的观点。它将整个文档视为基本信息单元，并假设已知该文档是固执己见的，并且包含对单个实体(例如，特定电话)的意见。

- **句子级情感分类** 对文档中的单个句子进行分类。

然而，不能假设每个句子都是固执己见的。传统上，人们通常首先将一个句子归类为自以为是或不自以为是，这被称为主观性分类。然后，由此产生的固执己见的句子被分类为表达积极或消极 的意见。句子层次情感分类也可以表述为一个三类分类问题，即把句子分为中性、肯定或否定三类。

- **方面级情感分析** 或基于方面的情感分析更细粒度。它的任务是提取和总结人们对实体和实体的方面/特征表达的意见，这些也称为目标。

例如，在产品评论中，它旨在分别总结 对产品不同方面的正面和负面意见，尽管对产品的总体看法可能是正面的或负面的。基于方面的情感分析的整个任务由几个子任务组成，如方面提取、实体提取和方面情感分类。例如，从句子“iPhone 的 声音质量很好，但电池没电”来看，实体提取应该将“iPhone”识别为实体，而方面提取应该识别“声音质 量”和“电池”是两个方面。方面情感分类应该将在 iPhone 语音质量上表达的情感分类为正面，在 iPhone 电池上表达的情感分类为负面。请注意，为简单起见，在大多数算法中，方面提取和实体提取是 结合在一起的，称为方面提取或情感/观点目标提取。

方面级情感分类

- 与文档级别和句子级别的情感分类不同，方面级别的情感分类会同时考虑情感和目标信息，因为情感始终具有目标。

如前所述，目标通常是实体或实体方面。为了简单起见，实体和方面通常都被称为方面。

- 给定一个句子和一个目标体，方面级别的情感分类旨在推断句子对目标体的情感极性/取向。例如，在句子“屏幕非常清晰，但电池寿命太短。”如果目标方面是“屏幕”，则情绪是积极的，而如果目标方面是“电池寿命”，则情绪是消极的。我们将在下一节讨论自动方面或目标提取。

- 方面级情感分类具有挑战性，因为建立目标与其周围上下文词的语义相关性模型是困难的。不同的语境词对句子对目标的情感极性有不同的影响。因此，在使用神经网络建立学习模型时，有必要捕捉目标词和上下文词之间的语义联系。

- 使用神经网络进行方面级情感分类有三个重要任务。

a. 第一个任务是表示目标的上下文

其中上下文是指句子或文档中的上下文单词。这个问题可以用上面两节中提到的文本表示方法来解决。

b. 第二个任务是生成一个目标表示

它可以与其上下文进行适当的交互。一般的解决方法是学习目标嵌入，这类似于单词嵌入。

c. 第三个任务是为指定的目标识别重要的情感上下文(单词)

例如，在句子“iPhone 的屏幕清晰， 但电池寿命短”中，“清晰”是“屏幕”的重要上下文，而“短”是“电池寿命”的重要上下文。

这任务最近由注意机制处理。尽管已经提出了许多深度学习技术来处理方面级情感分类，但是就我们所知， 在文献中仍然没有主要的技术。