

硕士学位论文

基于多任务深度学习的假新闻检测

**FAKE NEWS DETECTION BASED ON
MULTI-TASK LEARNING**

韩浩

哈尔滨工业大学

2019 年 7 月

国内图书分类号: TM301.2
国际图书分类号: 62-5

学校代码: 10213
密级: 公开

工学硕士学位论文

基于多任务深度学习的假新闻检测

硕士研究生: 韩浩

导 师: 王轩教授

副 导 师: 廖清副教授

申 请 学 位: 工学硕士

学 科: 计算机技术

所 在 单 位: 哈尔滨工业大学(深圳)

答 辩 日 期: 2019 年 7 月

授予学位单位: 哈尔滨工业大学

Classified Index: TM301.2

U.D.C: 62-5

Dissertation for the Master's Degree in Engineering

FAKE NEWS DETECTION BASED ON MULTI-TASK LEARNING

Candidate:	Hao Han
Supervisor:	Prof. Xuan Wang
Associate Supervisor:	Associate Prof. Qing Liao
Academic Degree Applied for:	Master of Engineering
Specialty:	Computer Technology
Affiliation:	Harbin Institute of Technology, Shenzhen
Date of Defence:	December, 2019
Degree-Confering-Institution:	Harbin Institute of Technology

摘 要

随着互联网的普及，更多的人选择在网络上消费新闻。互联网上的新闻消费有许多传统新闻消费方式所没有的特点，如消费成本低，时效性高等。然而，正是因为新闻在互联网上的传播速度快、成本低等原因，大量的假新闻得以肆意传播。假新闻造成的危害显而易见，轻则损害个人利益，重则引起社会恐慌。毋庸置疑的是，假新闻的大量涌入给互联网新闻消费这一领域带来了严重的负面影响。因此，如何识别假新闻，并及时停止其传播，对新闻消费领域至关重要。

假新闻一般是个人或团体故意撰写令人误导的信息以期获取政治或经济利益。也因为这个特点，假新闻的检测并非易事，其甄别过程往往需要领域知识的介入。现阶段对抗假新闻的方式主要是通过人工审核，随着人工智能的发展以及在各个领域获得的成功应用，研究者们期望通过人工智能技术自动检测假新闻。然而，考虑到当前互联网信息消费的特点，传播新闻的长度往往较短，这些方法的效果在其上表现不尽人意。

本文回顾了目前假新闻检测的相关方法。其中，基于特征构建的机器学习方法因为泛化能力较弱而逐渐被抛弃，取而代之使用深度神经网络自动学习特征。考虑到各个主题下产生假新闻的概率不同，新闻的真实性与新闻的主题存在一定联系，基于此，本文研究利用多任务深度学习来完成假新闻的自动检测这一任务。其次，当新闻的文本内容较短时，其提供的信息有限，仅依靠这些少量信息很难取得满意的效果，所以本文结合了信息的上下文信息，并同时用以提升假新闻检测和主题分类的性能。

综上，本文提出了用以检测假新闻和主题分类的多任务学习模型。根据调研结果，本课题首次提出将假新闻检测和主题分类两个任务通过多任务深度学习结合起来，并取得了目前最好的结果。

关键词：假新闻检测；多任务深度学习；主题分类

Abstract

With the popularity of the Internet, more people choose to consume news on the Internet. There are many features in context of news consumption on the Internet that traditional news consumption does not have, such as low cost of consumption and high timeliness. However, it is precisely because of the rapid spread of news on the Internet, low cost, etc., a large amount of fake news can be spread. The harm caused by fake news is obvious, It hurts personal interests, but causes social panic . Undoubtedly, the influx of fake news has had a serious negative impact on the field of Internet news consumption. Therefore, how to identify fake news and stop its dissemination in a timely manner is crucial to the field of news consumption.

Fake news is generally an individual or group deliberately writing misleading information in order to obtain political or economic benefits. Because of this feature, the detection of fake news is not an easy task, and its screening process often requires the involvement of domain knowledge. At this stage, the way to counter fake news is mainly through manual review. With the development of artificial intelligence and successful application in various fields, researchers hope to automatically detect fake news through artificial intelligence technology. However, considering the current characteristics of Internet information consumption, the length of dissemination of news is often short, and the effects of these methods are not satisfactory.

This article reviews the current methods of fake news detection. Among them, the machine learning method based on feature construction is gradually abandoned because of the weak generalization ability, and the deep neural network automatic learning feature is used instead. Considering the different probability of generating fake news under various themes, the authenticity of news has certain connection with the theme of news. Based on this, this paper studies the task of using multi-task deep learning to complete the automatic detection of fake news. Secondly, when the text content of the news is short, the information provided by it is limited, and it is difficult to obtain satisfactory results by relying on such a small amount of information. Therefore, this paper combines the context information of the information and at the same time enhances performance of fake news detection and topic classification. .

In summary, this paper proposes a multi-task learning model to detect fake news and

topic classification. According to the survey results, this is the first work that proposed combining the two tasks of fake news detection and topic classification through multi-task deep learning, and achieved the best results at present.

Keywords: fake news detection, multi-task learning, topic classification

目 录

摘 要	I
ABSTRACT	II
第 1 章 绪论	1
1.1 课题研究背景及意义	1
1.2 国内外研究现状	2
1.3 本文的内容	4
1.4 本文的结构安排	5
第 2 章 假新闻检测相关技术介绍	6
2.1 文本信息抽取	6
2.1.1 文本特征表示	6
2.1.2 文本特征抽取	8
2.2 注意力机制	10
2.3 分类学习算法	11
2.4 多任务深度学习	13
2.5 假新闻检测算法	14
2.5.1 基于特征构建的机器学习模型	14
2.5.2 基于新闻内容的假新闻检测	15
2.5.3 基于用户评论的假新闻检测	15
2.5.4 基于网络结构的假新闻检测	15
2.5.5 混合模型	16
第 3 章 多任务学习模型的设计与实现	17
3.1 嵌入层	17
3.2 表征层	18
3.3 多任务学习层	18
第 4 章 模型验证及分析	19
4.1 实验数据	19
4.2 实验设置	19
4.2.1 性能指标	19

4.2.2 假新闻检测对比方法	21
4.2.3 主题分类对比方法	21
4.3 性能对比	21
4.3.1 假新闻检测性能对比	21
4.3.2 主题分类性能对比	23
结 论	26
参考文献	27
攻读硕士学位期间发表的论文及其他成果	30
哈尔滨工业大学学位论文原创性声明和使用权限	31
致 谢	32

第 1 章 绪论

1.1 课题研究背景及意义

随着互联网的普及,越来越多的人选择在社交媒体上查看新闻。相比较传统的新闻消费方式,如报纸、电视等,互联网上的新闻消费有更多优势:用户不仅可以以几乎零成本的方式及时的获取新闻,而且更容易分享、评论,与他人一起讨论交流。一项 2016 年的调查显示,约有 62% 的美国成年人从互联网上获取新闻,而这一数字在 2012 年仅为 49%^[1]。假新闻几乎是伴随着新闻而生,其最早可以追溯到 1439 年^①。然而却始终没有一个严格统一的标准去定义假新闻。可以确定的是,假新闻有两点明显区别于真新闻的关键特征:真伪性与有意性。也就是说,

1. 假新闻中包含已被证实的虚假信息;
2. 假新闻是被有意撰写来误导读者以期获取利益。

基于此,本文采用文献^[2]给出的定义,也即:

定义 1.1 假新闻 假新闻是被作者有意撰写以期获取利益并被证实有虚假内容的信息。

得益于新闻在互联网上的传播速度快、成本低等原因,大量的假新闻得以肆意传播。假新闻造成的危害显而易见,轻则损害个人利益,重则引起社会恐慌。调查显示,在 2016 年的美国总统选举中,有关支持候选人的假新闻在美国著名社交媒体 Facebook 上转发量超过三千七百万次^[3]。为了及时遏制假新闻的传播,一些组织机构提供了揭露假新闻的平台,如国内的微博辟谣^②、国外的 FACTCHECK.ORG^③, POLITIFACT^④等。然而这些平台通常需要专业知识的介入,并且时效性较低,当这些平台将假新闻暴露于公众时,假新闻造成的危害往往已经发生。如图 1-1 所示,Google 工程师因为印度本地关于绑架孩童的谣言被当地暴民殴打致死^⑤。

基于以上情况,假新闻的自动检测在近年来得到了广泛的关注。不同于人工审核,假新闻的自动检测可以让机器根据输入自动判断新闻的真假。研究者们通过收集新闻的相关数据,如新闻内容、新闻评论等,从中设计一系列可以区分真假新闻的特征,然后使用机器学习方法训练分类器,进而可以及时停止假新闻的传播,尽可能的降低其带来的危害。

① <http://www.politico.com/magazine/story/2016/12/fakenews-history-long-violent-214535>

② <https://weibo.com/weibopiyao>

③ <http://www.factcheck.org>

④ <https://www.politifact.com>

⑤ <https://www.rt.com/news/433297-india-men-lynched-kidnappers-rumor/>



图 1-1 Google 工程师因谣言被印度暴民殴打致死

假新闻的检测是一项非常有意义，但却充满挑战的工作。由于假新闻常常伴随着政治或经济利益，经常被有目的用来误导读者，即使人类自身也很难判断其真假，尤其是在新闻内容很短时，很难从有限的文本信息中提取有用信息。此时，必须借助其他信息，如新闻作者、新闻主题等。因此，如何有效的利用新闻的上下文信息来帮助假新闻的自动检测，是一个亟待研究的问题。

1.2 国内外研究现状

国外关于假新闻检测的研究开始较早。鉴于假新闻在 Twitter、Facebook 等社交媒体上肆意传播，所以研究者首先从社交媒体上的虚假信息下手。Castillo 等人^[4]于 2011 年开始研究 Twitter 上信息的真假检测，手工设计了一系列特征，包括文本特征、用户特征、主题特征、传播特征等，然后以有监督的方式训练分类器，如支持向量机、决策树等传统机器学习方法。

基于各种网络结构的特征也被证实有助于甄别假新闻。Kwon 等人^[5]通过分析谣言的传播过程，发现谣言在传播过程中会出现周期性的波动，而真实信息在一定时间内仅有一个峰值。此外，一些假新闻制作者常常使用虚假的图片、视频等激起读者的愤怒情绪，以此吸引读者注意力，获得更广泛的传播。文献^[6,7]使用深度神经网络从文本和图片中抽取特征，实验表明，相对于单模态的假新闻检测（如仅针对文本信息），多模态信息的结合可以提高假新闻检测的准确率。

由于新闻在社交媒体上的传播往往会伴随一系列的社交信息，如转发、分享、评论等，一些研究者试图通过新闻传播过程中产生的社交信息来检测假新闻。Ma 等人^[8]于 2016 年首次尝试使用循环神经网络（RNN）对用户的评论进行建模，发现假新闻下的评论往往带有质疑、否定等字眼；Liu 和 Wu^[9]通过对参与新闻传播的用户设计特征，如粉丝数、注册时间、是否认证等，并使用卷积神经网络（CNN）

和循环神经网络（RNN）分别抽取参与用户的局部特征和全局特征。由于这种方法仅使用基本的用户特征作为模型输入，相比其它基于语言特征、结构特征等方法更加稳健，在早期的假新闻检测中表现良好。

受到多任务学习的启发，Ma 等人^[10]将谣言检测（rumor detection）和立场检测（stance detection）一起训练，相比较于单任务训练，其设计的多任务模型能够捕捉更多有用的潜在模式，如能够发现更多的否定词，其背后的思想是：若用户对新闻事件更多的是持怀疑、否定态度，则该新闻更可能是假新闻。

国内对虚假信息检测的研究更多的是基于新浪微博。2012 年，Yang 等人^[11]研究新浪微博上的谣言检测，他们发现，除了 Castillo 等人提出的一系列特征，使用微博的客户端（web 端或者手机端）以及微博中提及事件所发生的地点（国内或国外）同样对真假信息有较强的区分力。2013 年，程亮等人^[12]设计了基于神经网络的谣言分类器，设计的特征包括相关微博的数量、微博话题等。Wu 等人^[13]通过分析微博上真假信息的传播过程，发现谣言与真实消息在传播过程中的差异：谣言通常被普通用户发布，经由一些加 V 用户转发后，接着被普通用户转发；而真实消息首先被加 V 用户发布，然后被许多普通用户转发。微博中经常出现配图与文本不符的情况，例如谣言发布者经常使用之前的真实图片来形容当前发生的事，故意夸大事实，博人眼球。为了检测文本与图片不符的谣言，Sun 等人^[14]通过查找微博配图在网络中出现的最早时间，用微博发布时间与配图出现的最早时间之差作为一类特征。2016 年，毛二松等人^[15]从微博情感倾向性、微博传播过程和微博用户历史信息等方面提取特征，并利用集成分类器对微博谣言进行检测。

通过以上对国内外研究现状的分析，可以看出，目前假新闻检测的研究都是将其看作是分类问题，即从新闻内容及新闻在传播过程中产生的社交信息中抽取特征，然后以有监督的方式训练分类器。早期工作集中于从新闻内容及社交信息中手工设计特征，例如新闻主体内容的文本特征、新闻主题、传播特征等。然而，手工设计特征不仅费时费力，通常需要领域知识的介入；更重要的是，新闻涉及的话题众多，并且来自于各种平台，手工设计出的特征往往仅能反应出某些特定数据集的特性，很难设计出一套覆盖全面、泛化能力好的特征。

鉴于深度学习强大的特征抽取能力，越来越多的研究者使用深度学习来代替手工特征的设计。基于此，一系列深度学习模型被用来检测假新闻。其中，使用 RNN 对用户评论进行建模的想法首先被证实有效；一些混合模型有效的结合了文本、图片等多模态的数据。

假新闻检测领域一直以来缺少一份基准数据集，Wang^[16]在 2017 年从 politi-fact.com 上收集了一万两千多条短文本新闻，每个新闻都有一个反应新闻真假性的

标签。由于假新闻中往往会掺杂一些真实信息，因此这些标签为以下更细粒度的六种：pants on fire, false, barely true, half true, mostly true, true。此外，这些新闻从政治辩论、电视广告、Facebook、Twitter 等上收集而来，并且涵盖了经济、医疗、教育、税收等多个话题。除了基本的新闻文本信息外，该数据集还提供了丰富的发言人信息，如发言人姓名、所属党派、工作、历史发言表现等信息。

由于发言人对不同领域的熟悉程度各不相同，因此他们在各个主题下发言的可信度也不尽相同。鉴于多任务学习（Multi-Task Learning）在计算机视觉、自然语言处理等方面有许多成功的应用案例，如何在抽取语义特征时，利用多任务学习融入主题信息，并有效的结合新闻的上下文信息，是一个有待深入研究的问题。

1.3 本文的内容

通过对国内外假新闻检测的研究分析和总结，可以看出，假新闻的检测逐步由原始的手工设计特征过渡到深度学习自动抽取特征。当新闻主体内容较短时，可提取的信息十分有限，新闻上下文信息的使用对提高假新闻检测的准确率至关重要；并且，由于不同作者对各个领域的熟悉程度不同，因此他们在各个话题下的话语可信度也不一样。

如图1-3所示，关于医疗的新闻相比较教育主题的新闻而言，假新闻占比更多；图1-3揭示了不同作者其发表真假新闻的概率分布也有很大差距。

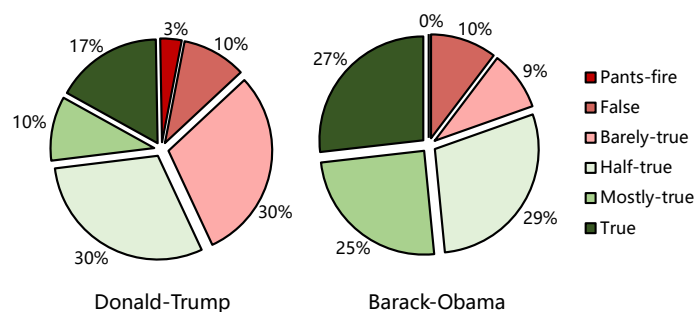


图 1-2 作者真假分布

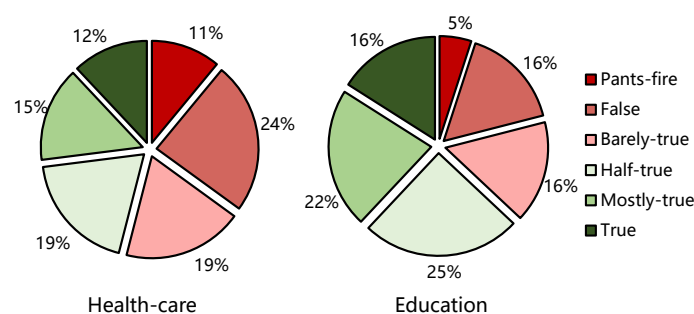


图 1-3 主题真假分布

为了解决上述两个问题，本文将假新闻检测和新闻主题分类一起训练，并且

有效的融合新闻的上下文信息，同时提升假新闻检测和主题分类的效果。本课题的最终目标可总结为：给定新闻及其新闻的上下文信息，模型不仅输出新闻更细粒度的真假信息，而且同时输出新闻讨论的主题。

具体工作内容如下：

1. 分析新闻的真实性与新闻的主题、新闻的作者之间的关系；
2. 基于多任务深度学习，设计并实现多任务深度学习模型，将新闻的真实性检测和新闻的主题分类同时训练，并提高二者的任务表现；
3. 研究如何利用新闻的上下文信息，以同时帮助假新闻的检测和新闻的主题分类，包括上下文信息的特征表示、特征抽取、特征融合等。

1.4 本文的结构安排

本文主要研究假新闻的自动检测问题，共包含四个章节，本篇论文的组织结构如下：

第一章为绪论，首先详细介绍了假新闻产生的背景以及相关定义、假新闻的危害、假新闻的自动检测的难点以及研究意义。然后针对上述问题的国内外研究现状进行了总结。接着概述了本文的研究内容，引出本文研究的三个主要问题，最后阐述了本文的组织结构安排。

第二章主要对本文用到的相关知识进行详细介绍。本章首先概述了假新闻检测的一般解决思路以及可能存在的不足。接着介绍了深度学习模型在文本处理上的应用，最后总结了假新闻检测的常见研究方法

第三章详细介绍了本文提出的多任务学习模型。首先介绍了通过数据分析得到的新闻主题及其真实性的潜在关系，然后介绍了模型的具体设计思路，并对比了其他主流方法。最后对提出的模型进行具体的实现并提供与当前主流算法的性能比较。

第四章对假新闻检测的实验设置及结果进行了详细的分析。首先简要的概述了对比方法的实现思路。然后针对实验提出若干衡量标准。最后提供了详细的对比结果及分析。

第 2 章 假新闻检测相关技术介绍

2.1 文本信息抽取

一般来说，新闻的组成部分通常包括新闻作者、新闻内容等，一些新闻为了丰富新闻内容，可能会在新闻中插入若干图片。本文讨论的新闻主要涉及文本内容。检测假新闻的第一步便是从新闻的文本内容中抽取相关特征表示。下面就文本特征表示、文本特征抽取、注意力机制三个方面做详细介绍。

2.1.1 文本特征表示

在处理自然语言处理领域的问题时，文本的特征表示一直是个老大难的问题。不同于图片其早已呈现成数值矩阵的形式，自然语言虽然对人类来说直观且易于理解，但对计算机来说只能将其看出不连续的一个个字符。相关研究者在自然语言的数学表示上付出了巨大努力并取得不凡的成绩。

在表示一段文档时，TFIDF 是一个不错的选择。在一份给定的文件里，词频（Term Frequency，TF）指的是某一个给定的词语在该文件中出现的频率。这个数字是对词数（Term Count）的归一化，以防止它偏向长的文件。（同一个词语在长文件里可能会比短文件有更高的词数，而不管该词语重要与否。）对于在某一特定文件里的词语来说，它的重要性可表示为：

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (2-1)$$

逆向文件频率（Inverse Document Frequency，IDF）是一个词语普遍重要性的度量。某一特定词语的 IDF，可以由总文件数目除以包含该词语之文件的数目，再将得到的商取以 10 为底的对数得到：

$$IDF_i = \lg \frac{|D|}{|\{j : t_i \in d_j\}|} \quad (2-2)$$

然后 TF-IDF 可以表示为

$$TFIDF_{i,j} = TF_{i,j} \times IDF_i \quad (2-3)$$

某一特定文件内的高词语频率，以及该词语在整个文件集合中的低文件频率，可以产生出高权重的 TF-IDF。因此，TF-IDF 倾向于过滤掉常见的词语，保留重要的词语。使用 TFIDF 便可以将一份文档表示为一个固定长度的向量，其中该向量的长度等于词典大小。TF-IDF 经常被用来计算两个文档的相似性。

在使用深度神经网络时，往往将自然语言处理相关问题看做是一个序列问题：对于文档中的每个词，将其表示为一个向量，整个句子或文档则是一个由若干向量组成的序列。如何将每个词表示为向量？一个简单直观的方式是针对每个字给与一个独热编码（one-hot encoding），独热编码下，每个词可以表示为如下向量：

$$c_i = [0, 0, 0, \dots, i, \dots, 0, 0, 0]^T \quad (2-4)$$

其中，向量长度为词典大小，仅该词对应的位置 i 为 1，其余位置均为 0。不过这种表示存在以下问题

- 词典较大，导致向量维度急剧增长，存储和计算都不高效。
- 词与词之间没有关系，各个词之间的表示无法表示其逻辑上的关系。

2013 年，由 Mikolov 等人^[17] 提出的分布式词向量在很大程度上解决了独热编码存在的弊端，并在接下来的几年被自然语言处理领域接受，其提出的 word2vec 模型获得了广泛应用。

word2vec 的基本思想是利用大量无监督的语料库去学习一个神经网络语言模型，其训练模型按照训练目标可分为以下两种：

- **CBOW: Continuous Bag of Words**, 其任务是根据词的上下文信息 $context(w)$ 预测当前是哪个词

- **Skip-Gram**: 其任务是根据当前词 w 来预测当前词的上下文语境 $context(w)$

下图2-1展示了 word2vec 的两种训练方式。

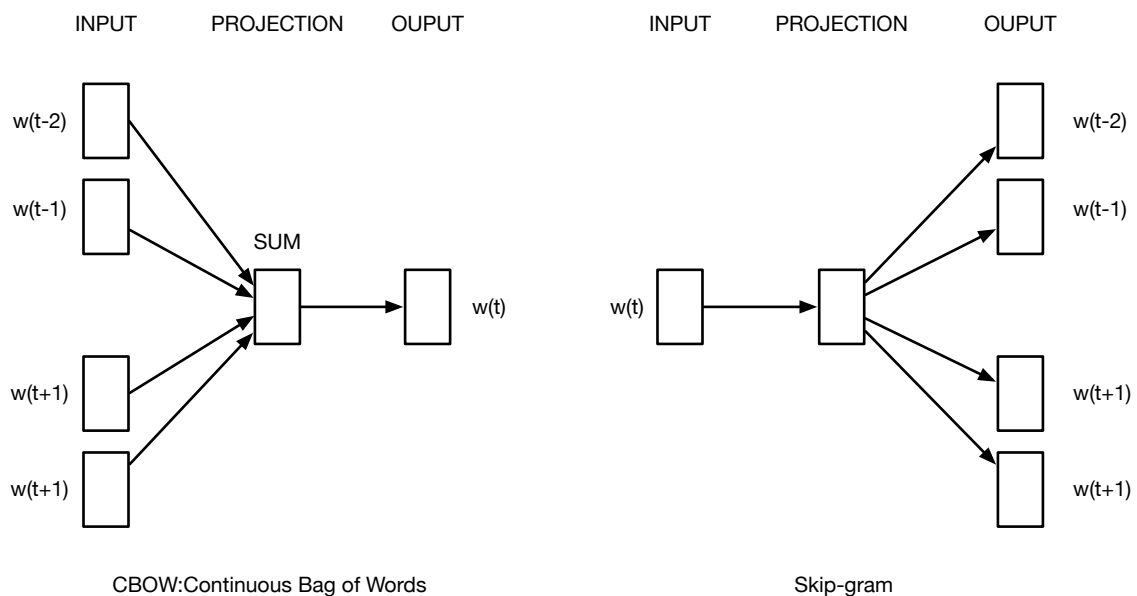


图 2-1 word2vec 两种训练方式

由于语料库中词表可能巨大，一般对中文来说，词表大小可达 200,000，英语

可达 30,000。为了加速模型训练, word2vec 使用了 Hierarchical Softmax 和 Negative Sampling 来加速计算。由 word2vec 得到的词向量是一个富含了语义信息的低纬度稠密向量, 如此一来便解决了独热编码所引起的问题。

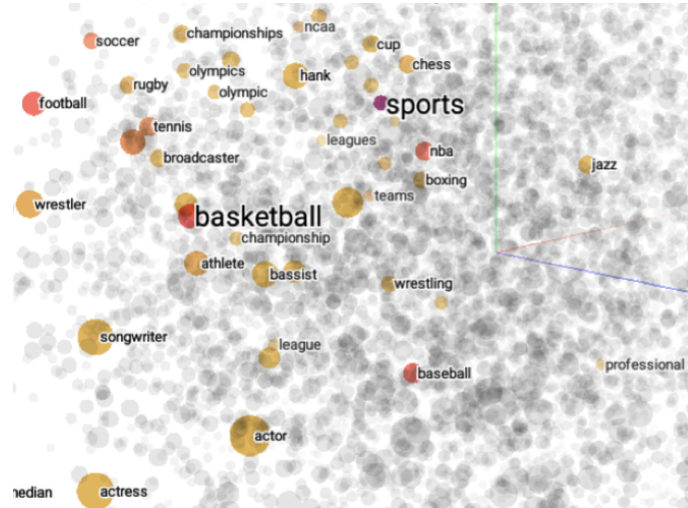


图 2-2 word2vec 在三维空间的可视化

经由 word2vec 训练得到的词向量蕴含了语义信息, 其直观体现为词义相近的词其词向量距离较近, 图2-2直观的表现出这种特性。

但是, word2vec 在一词多义这个问题上并没有办法识别, 由于每个词唯一对应一个词向量, 所以如苹果(水果)和苹果(苹果公司), word2vec 将其视为同一个词。有没有办法解决呢? 近期推出了模型 ELMo^[18], GPT^[19], BERT^[20] 等模型都不同程度的致力于解决这个问题, 由于这些不是本文重点, 因此不做详细介绍, 更多细节可阅读相关论文。

2.1.2 文本特征抽取

在获得文本的信息表示后, 如使用 word2vec 获取每个词的词向量, 则句子可以表示为一个序列

$$\mathcal{X} = x_1, x_2, \dots, x_T \quad (2-5)$$

其中 T 表示文本的长度, 为了获取一个固定长度的向量表示, 需要使用一些特殊结构的神经网络来获取。一个简单的方法是直接在词向量上施加 Pooling 即可。这种方式看似简单, 但是已有文献证明其有效性^[21]。此外, 还有一些网络结构来专门处理这种问题, 如接下来要介绍的循环神经网络和卷积神经网络。

2.1.2.1 循环神经网络

循环神经网络(RNN)最早由 Elman^[22] 于 1990 年提出, 可以对具有时间序列的对象进行建模, 在 t 时刻的神经元状态受它前面 $t-1$ 时刻神经元状态的影响。通

过不断的输入时间序列的数据，RNN 能够记忆前面的信息并将其参与到后续的计算当中，从而获得整个时间序列对象的高层特征表示。RNN 模型结构及其展开图如图2-3所示。

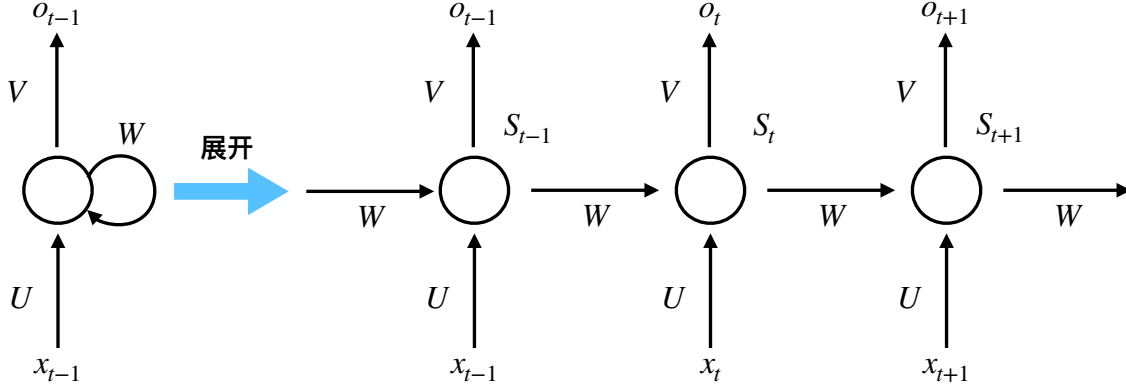


图 2-3 循环神经网络计算方式

RNN 结构简单，但在处理较长序列时，会出现梯度消失或梯度爆炸的问题，这使得 RNN 很难去学习长期依赖。GRU (Gated Recurrent Unit) 由 Chung 等人^[23] 于 2014 年提出，通过引入门结构，很大程度上解决了 RNN 的梯度消失问题。GRU 内部的转换公式如公式2-6所示：

$$\begin{aligned}
 z_t &= \sigma(W_z x_t + U_z h_{t-1}) \\
 r_t &= \sigma(W_r x_t + U_r h_{t-1}) \\
 h'_t &= \tanh(W_h x_t + U_h (h_{t-1} \cdot r_t)) \\
 h_t &= (1 - z_t) \cdot h_{t-1} + z_t \cdot h'_t
 \end{aligned} \tag{2-6}$$

其中， x_t 为 t 时间步的词向量，重置门 r_t 与之前的记忆结合，更新门 z_t 决定了是否将之前的记忆参与到后面的计算。通过 GRU 的计算，我们可以得到每个时间步的隐含层输出 h_1, h_2, \dots, h_n ，为了获得整个序列的语义表示，可以对其进行 average pooling 或 max pooling，或直接取最后一个时间步的输出 h_n 作为最终的特征表示。与 GRU 类似的长短期记忆网络 (Long Short Term Memory, LSTM) 也是为了解决 RNN 无法学习长期依赖而提出的，其也是通过门结构控制信息流入流出来达到这一目的，详细原理不再赘述。

2.1.2.2 卷积神经网络

卷积神经网络 (Convolutional Neural Networks, CNN) 在计算机视觉方面取得了巨大成功。2014 年，Yoon Kim 将 CNN 成功的运用于文本的特征抽取，并取得了很好的效果。CNN 通过使用卷积核去扫描文本的词向量矩阵，以获得特征图。具体地，特征图 c_i 可由下式获得：

$$c_i = f(W \cdot x_{i:i+h-1} + b) \quad (2-7)$$

其中 W 为权值矩阵, b 为对应的偏置项。F 为非线性激活函数, 如 *ReLU*、*Tanh* 等。 $x_{i:i+h-1}$ 表示长度为 h 的词向量矩阵, 通过扫描整个文本的词向量矩阵, 该卷积核可生成特征图:

$$c = [c_1, c_2, \dots, c_{n-h+1}] \quad (2-8)$$

接着通过最大池化 (Max-Pooling) 后获得最具表示能力的特征: $c' = \max c$ 。以上过程描述了一个卷积核的计算过程。通常情况下, 会使用不同大小的卷积窗口, 并且每个卷积窗口使用多个卷积核来捕捉不同层次的特征。如图2-4展示了使用 2, 3, 4 卷积窗口, 并且每个卷积窗口有 2 个卷积核的卷积操作。

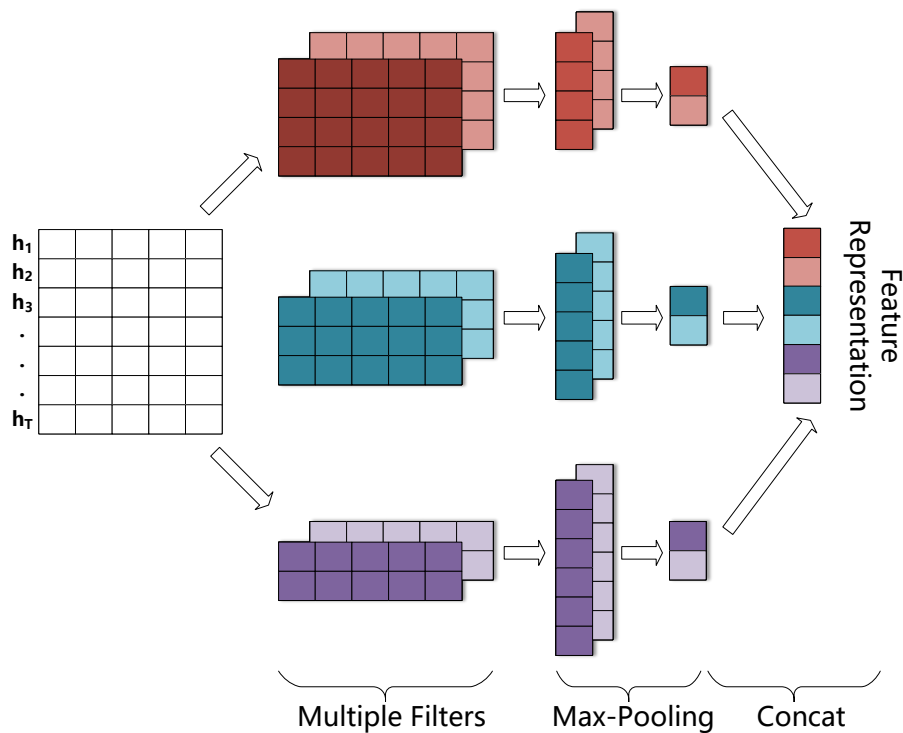


图 2-4 卷积神经网络在文本上的应用

2.2 注意力机制

注意力机制 (Attention) 最早是应用于图像领域的。2014 年, Google mind 团队在 RNN 模型上使用了 attention 机制来进行图像分类, 取得了很好的性能^[24]。随后 Bahdanau 等人^[25] 在论文《Neural Machine Translation by Jointly Learning to Align and Translate》中, 使用类似 attention 的机制在机器翻译任务上将翻译和对齐同时进行, 他们的工作算是第一个将 attention 机制应用到 NLP 领域中。接着 attention 机制就被广泛应用在基于 RNN/CNN 等神经网络模型的各种 NLP 任务中去了。2017

年，google 机器翻译团队发表^[26]的《Attention is all you need》中大量使用了自注意力（self-attention）机制来学习文本表示。由于其优异的表现吸引了研究者的目光，自注意力机制也自然而然的成为了大家近期的研究热点，并在各种 NLP 任务上进行探索，纷纷都取得了很好的性能。其提出的 Transformer 模型也被证实其强大的特征抽取能力。

注意力机制的本质是从人类视觉注意力机制中获得灵感。人类视觉在感知东西的时候，一般不会是一个场景从到头看到尾每次都看，而往往是根据需求观察注意特定的一部分。而且当我们发现一个场景经常在某部分出现自己想观察的东西时，我们就会进行学习在将来再出现类似场景时把注意力放到该部分上。

Attention 的计算可以总结为公式2-9：

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2-9)$$

其中， Q ， K ， V 分别称为 Query，Key，Value。

下图可以直观的理解 Attention 的作用。可以发现，对比传统的网络结构，使用 Self-Attention 机制可以在编码一个词时综合考虑其上下文，从而得到该词在该语境下的信息表示。

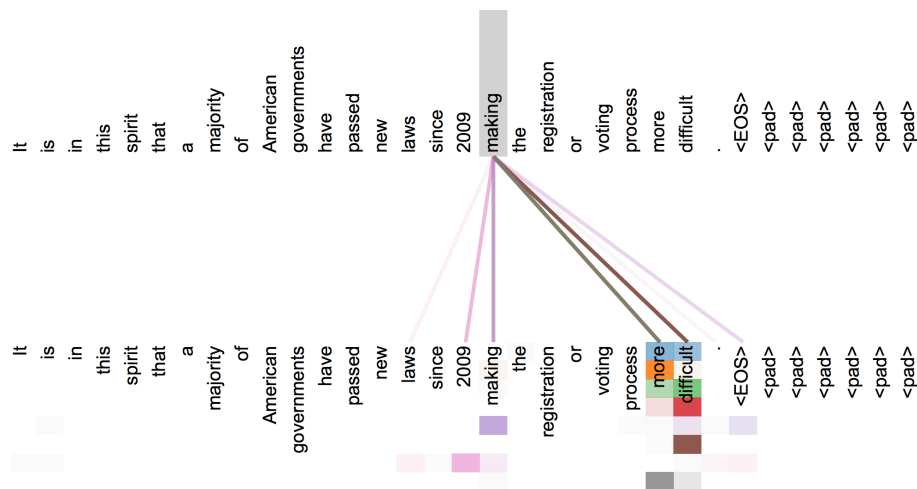


图 2-5 Attention 机制的可视化

2.3 分类学习算法

根据数据类型的不同，对一个问题的建模有不同的方式。在机器学习或者人工智能领域，人们首先会考虑算法的学习方式。在机器学习领域，有几种主要的学习方式。将算法按照学习方式分类是一个不错的想法，这样可以让人们在建模和算法选择的时候考虑能根据输入数据来选择最合适的算法来获得最好的结果。

其中，分类学习算法属于监督学习算法。分类算法实际应用广泛，如垃圾邮

件分类、评论情感分析、糖尿病检测等，因而受到了研究者们的青睐，大量的分类学习算法运用而生。

分类学习算法的基本思想可以总结为如下：给定训练数据 $[X, Y]$ ，其中 X 为输入， Y 对应输出，其可以是二分类，对应 Y 的值域则为 $\{0, 1\}$ ；也可以是多分类，对应的值域则为 $\{0, 1, 2, \dots, N-1\}$ (N 为分类个数)。分类学习算法的目的是根据训练数据 $[X, Y]$ ，训练一个分类学习模型 F ，其可以针对未预见的样本 x 进行预测，从而得到预测值 $y^{predict}$ 。常见的机器学习分类算法有：逻辑斯蒂回归 (Logistic regression)、决策树 (Decision Tree)、朴素贝叶斯 (Naive Bayes) 等，还有一些常见的集成学习算法如 AdaBoost、GBDT、XGBoost、随机森林 (Random Forest) 等。

随着深度学习的兴起，基于神经网络的分类器也得到了广泛的研究与应用。神经网络模型是一个非常强大的模型，起源于尝试让机器模仿大脑的算法，在 80 年代和 90 年代早期非常流行。同时它又是一个十分复杂的模型，导致其计算量非常巨大，所以在 90 年代后期逐渐衰落。近年来得益于计算机硬件能力，又开始流行起来。人类的大脑是一个十分神奇的东西，尽管当今人工智能科技已经十分发达，但很大程度上，无论建立一个多么完美的模型，其学习能力目前仍然逊色于大脑。因此神经网络是人工智能领域的一个热门研究方向。

为了描述神经网络模型，我们先从最简单的神经网络讲起，这个神经网络仅由一个“神经元”构成，图2-6 a)展示了具有一个神经元的神经网络。其中 x_1, x_2, x_3 称为输入（来自与其他神经元的输入信号）， x_0 称为偏置单元 (Bias Unit)， θ 称为权重或参数， $h_{\theta}(x)$ 称为激活函数 (Activation Function)。这里，使用 Sigmoid 激活函数可以直接输出待分类的概率，Sigmoid 函数的计算公式如公式2-10所示：

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (2-10)$$

使用 Sigmoid 函数对输出单元进行非线性变化。最终输出值即为输入样本为正例的概率。如图2-6 b)展示了其变化曲线。可以看到，Sigmoid 激活函数将输入 $z \in \mathbb{R}$ 压缩到 $\sigma(z) \in [0, 1]$

面对多分类问题，可以直接改变输出单元个数，并在输出上施加 Softmax 函数，如此一来便可得到输入样本属于各个类别的概率，取输出概率最大的类别即为预测类别。

Softmax 函数，或称归一化指数函数，是逻辑函数的一种推广。它能将一个含任意实数的 K 维向量“压缩”到另一个 K 维实向量中，使得每一个元素的范围都在 $(0, 1)$ 之间，并且所有元素的和为 1。该函数的形式通常按下面的式子给出：

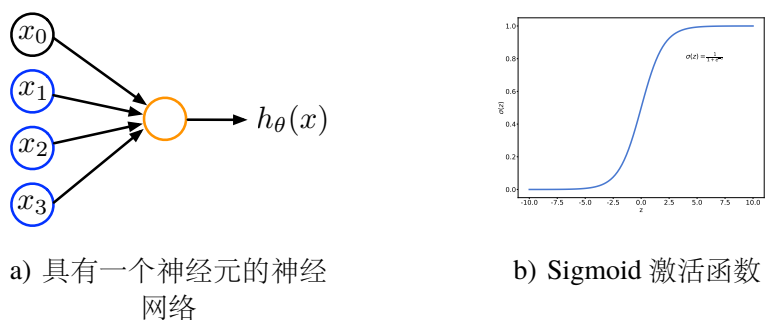


图 2-6 简单神经网络与激活函数示例

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K \quad (2-11)$$

为了增加神经网络的拟合能力，通常使用多层的神经网络来应对较复杂的分类问题，如图2-7展示了使用三层神经网络处理六分类的问题。

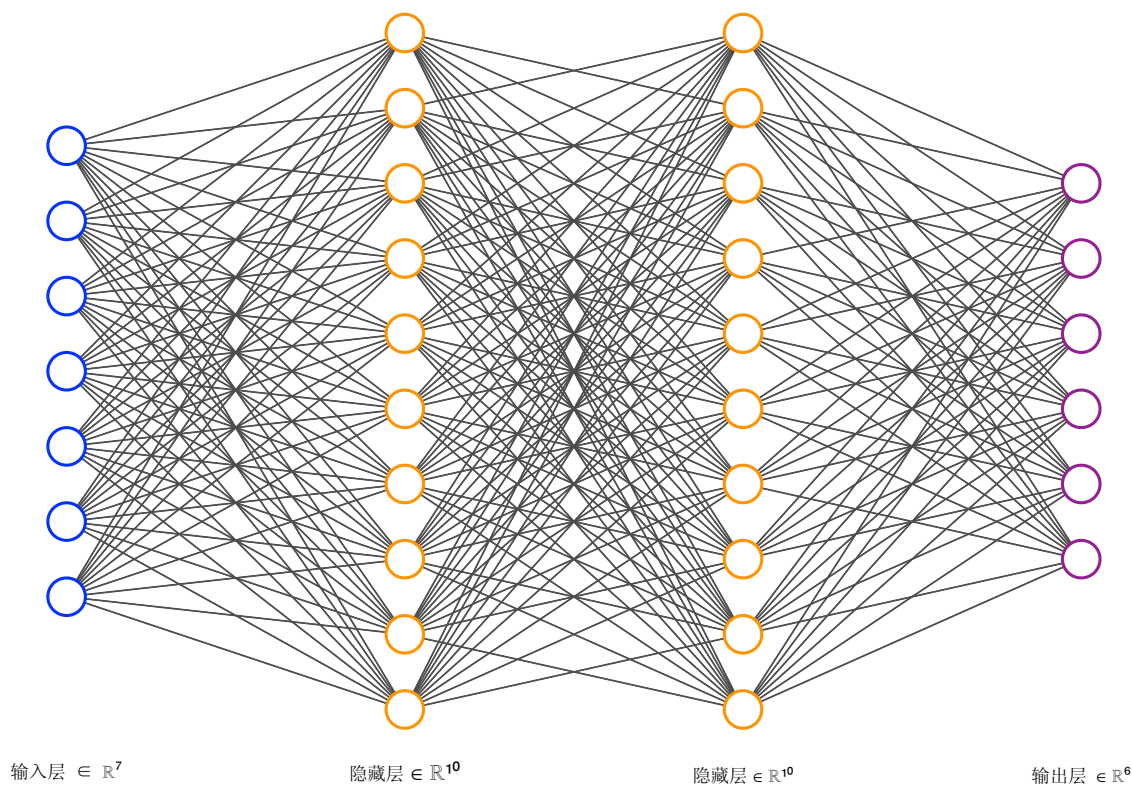


图 2-7 三层神经网络处理六分类问题

2.4 多任务深度学习

相较于传统的单任务学习，多任务学习的目标是通过利用多个相关学习任务之间的有用信息来提升它们的表现。基于深度神经网络的多任务学习通常有两种方法：隐层参数的硬共享与软共享。

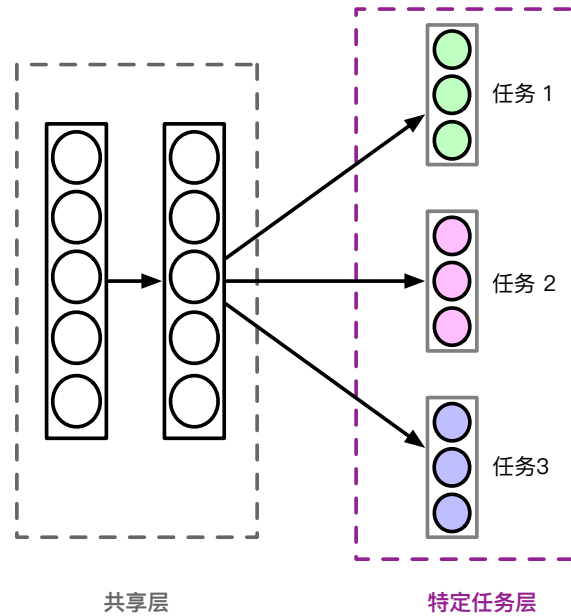


图 2-8 多任务学习的硬参数共享

硬参数共享是多任务深度学习中最常用的方法。通常通过在所有任务之间共享隐藏层，同时保留几个特定任务的输出层来实现。这种方式大大降低了过拟合的风险，同时学习的任务越多，模型找到一个含有所有任务的表征就越困难，而过拟合原始任务的可能性就越小。图2-8展示了基于深度神经网络的硬参数共享方式的多任务学习方法。

另一方面，当使用软参数共享时，每个任务都有自己的参数和模型。模型参数之间的距离是正则化的，以便鼓励参数相似化。

本课题采用硬参数共享的方式，设计多任务深度学习模型，同时训练新闻的真实性检测任务与新闻的主题分类任务。

2.5 假新闻检测算法

近年来，大量研究者提出了有效的模型及算法用以检测假新闻，本章对现存的模型及算法做一个总结，并对各个模型及算法做简单介绍。

2.5.1 基于特征构建的机器学习模型

通过对已有数据进行特征抽取，然后以此为特征集去训练机器学习模型，是一个简单直接的想法。这类方法的主要工作在于从数据中构建大量可能有用的特征集，然后以有监督的方式训练机器学习模型，如支持向量机、决策树等。如 Castillo 等人^[4]于 2011 年开始研究 Twitter 上信息的真假检测，手工设计了一系列特征，包括文本特征、用户特征、主题特征、传播特征等。这里的文本特征可以使用上述介绍的 TF-IDF 来表示，用户特征则主要包括用户性别、注册时长等。这一类方法的

缺点是基于手工设计的特征其泛化能力较弱，鉴于假新闻通常涵盖主题广，因此，很难设计一个通用的特征集。

所以，近年来的方法更多的是基于深度学习的模型，通过构建有效的深度神经网络，特征抽取的工作由神经网络自动完成。

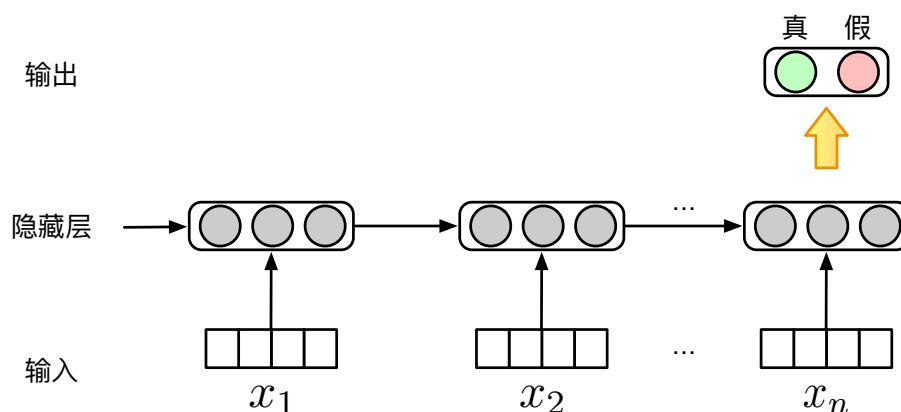


图 2-9 基于循环神经网络的假新闻检测

2.5.2 基于新闻内容的假新闻检测

基于以上介绍的文本信息抽取，可以使用 word2vec 等类似方法将每个词表示为一个词向量，对于新闻内容，则是一个由词向量组成的一个序列。此时我们可以使用基于 RNN 或 CNN 类似的方法进行特征抽取，然后送入分类器进行分类。如 Wang^[16] 在抽取文本特征时使用了 CNN 模型；而 Long 等人^[27] 使用了 LSTM 进行文本特征的抽取；Karimi 等人^[28] 通过组合 CNN 和 LSTM 对文本特征进行抽取。图2-9展示了使用循环神经网络检测假新闻。

2.5.3 基于用户评论的假新闻检测

如国内外研究现状所述，一些研究者通过对用户对新闻的回应（评论、转发等）进行建模，基于抽取的特征建立分类器。针对图2-9中的输入， x_i 则可以表示为每个用户的评论，如 Ma 等人^[8] 使用 LSTM 对用户评论进行建模，取得了不错的效果。再如^[9] 分别使用 LSTM 和 CNN 对参与评论用户的特征进行特征抽取，主要用到的是用户的基本信息，如用户性别、注册时长等。针对图2-9中的输入， x_i 则可以表示为每个用户的基本信息。

这类方法的主要缺点是需要新闻传播过程产生的大量数据，缺少时效性。

2.5.4 基于网络结构的假新闻检测

待补充

2.5.5 混合模型

一些研究者通过组合多种特征来构建更强大的混合模型，从而提升假新闻检测的性能。如 Ruchansky 等人^[29] 将用户信息与新闻文本内容结合；Wang 等人^[7] 将新闻的文本内容与图像内容结合；Karimi 等人^[28] 将新闻内容与新闻的上下文结合等。

由于本文提出的模型也结合了新闻的文本内容与新闻的上下文内容，所以，本文所提及的模型也属于混合模型。与之前工作不同的是，本文是首次将假新闻检测和新闻的主题分类这两个任务结合起来，并取得了目前最优的性能。

第3章 多任务学习模型的设计与实现

如图3-1是本文提出的多任务学习模型框架：FDML。FDML 可分为三层：

- **嵌入层 (Embedding Layer)**：本层的主要目的是将文本内容以及上下文信息嵌入到一个低维空间，对于文本内容中的每个词，将其嵌入为一个词向量；对于上下文内容，将每个信息表示为一个低维向量，并在后续得训练中得以更新。
- **表征层 (Representation Layer)**：本层的主要目的是接受嵌入层的输入，使用深度神经网络抽取高阶特征，针对不同任务设计各自所需的特征，供多任务学习层使用。
- **多任务学习层 (Multi-Task Learning Layer)**：本层的主要目的是接受表征层的高阶特征作为输入，使用神经网络分类器同时对两个任务进行优化。

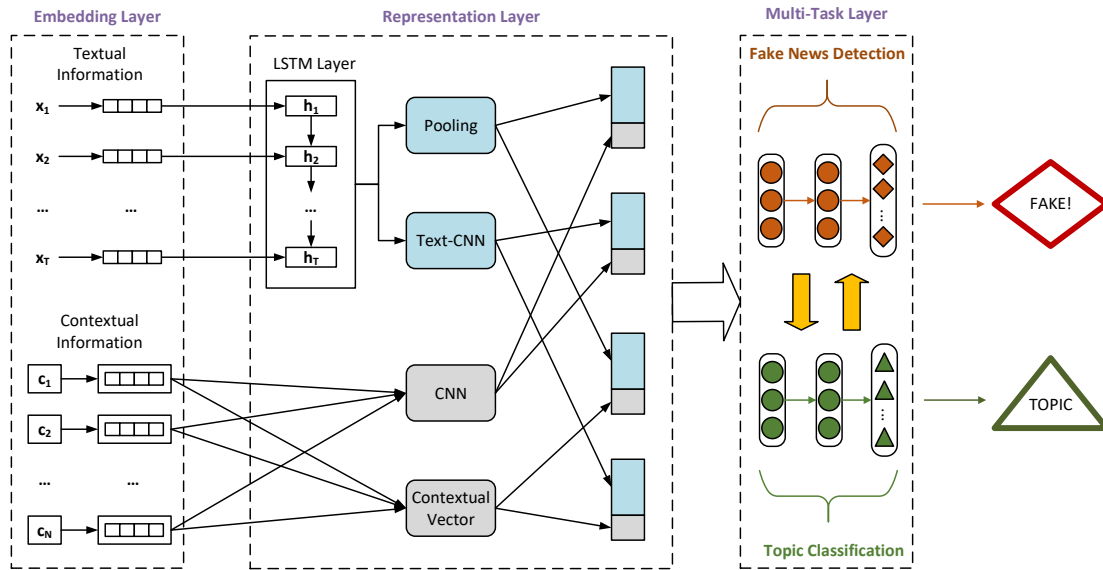


图 3-1 多任务学习模型：FDML

下面就这三层结构及其作用做详细介绍。

3.1 嵌入层

令 $X = \{x_1, x_2, \dots, x_T\}$ 表示新闻的文本内容， x_i 表示一个词， T 表示文本的长度，词嵌入可以表示为：

$$x'_i = E \cdot x_i \quad (3-1)$$

其中 x'_i 即为词向量 ($x'_i \in \mathbb{R}^d$)， d 为嵌入的词向量的维度， E 为词向量矩阵。

可以随机初始化，在整个训练过程中进行学习；也可以由其他训练好的词向量做初始化，如 word2vec^[17]，Glove^[30]，fasttext^[31] 等。

同文本的特征表示一样，对于上下文信息，同样以嵌入（Embedding）的方式获取对应的特征表示，考虑到不同的上下文信息所代表的领域不同，因此各个上下文信息分别嵌入到不同的向量空间，通过随机初始化来获得上下文信息的初始向量表示，在训练过程中更新此向量。

3.2 表征层

表征层的作用便是从嵌入层的输入中抽取高阶特征。如图3-1所示，本论文使用 LSTM 和 CNN 对文本内容进行特征抽取。将抽取完的特征组合成各个人物所需的任务送入多任务学习层进行分类器的训练。

3.3 多任务学习层

在表征层通过深度神经网络进行特征抽取后，获得了来自文本和上下文信息的高阶特征表示。在多任务学习层，通过对特征的融合，对于新闻的真实性检测以及新闻的主题分类这两个任务，可以分别获得相应的特征表示。然后，通过全连接神经网络得到最终的预测结果：

$$y'_m = \text{Softmax} (W_m^2 \cdot f(W_m^1 r_m + b_m^1) + b_m^2) \quad (3-2)$$

r_m 为任务 m 的最终特征表示， y'_m 即为任务 m 的预测结果。这里，令 $m = 1$ 表示新闻的真实性检测任务， $m = 2$ 表示新闻的主题分类任务。

模型的参数通过随机梯度下降进行更新，损失函数定义为：

$$\mathcal{L} = \sum_{m=1}^M \alpha_m L(y'_m, y_m) + \lambda \|\Theta\|_2 \quad (3-3)$$

其中 y_m 为任务 m 的真实标签， y'_m 为分类器预测的标签； λ 为 L_2 正则化系数， α_m 为任务 m 关于损失函数的权重；损失函数 L 采用交叉熵（Cross Entropy）。 Θ 为整个模型的参数。

由于在学习过程中两个任务是相互影响的，因此，本课题设计的多任务深度学习模型可以学习到两个任务的潜在关系，并提高这两个任务的性能表现。

第 4 章 模型验证及分析

4.1 实验数据

假新闻检测领域一直以来缺少一份公共数据集，2017 年，Wang^[16] 公布了一份较大的数据集，其中供包含了 12.8k 条新闻。除了新闻的文本内容外，此数据集还提供了丰富的上下文信息，如作者、党派等，每条新闻均有一个人工审核的真实性标签，其按照从假到真的程度，分为以下六种：Pants-Fire, False, Barely-True, Half-True, Mostly-True, True。此外，每条新闻均提供了详细的鉴定报告以供分析。本论文基于此数据集来验证提出模型的有效性。

表4-1给出了 LIAR^[16] 数据集的一个简单示例。

4.2 实验设置

在处理新闻的主题标签时，本课题取前二十个出现次数最多的主题作为新闻的主题标签，其余的新闻则赋予 others 标签。因此，对假新闻检测任务而言，是一个六分类问题；对新闻的主题分类任务而言，是一个二十一分类问题（前二十个主题 + others 标签）。前 20 个主题如表4-2所示。

实验相关参数设置如表4-3所示。

论文提及到的模型实现全部基于深度学习框架 pytorch，代码使用 python 编写。为了减小随机种子带来的误差，所有实验结果都是取随机十次实验的平均结果。显卡使用 NVIDIA Telsa V100。

4.2.1 性能指标

为了更全面的描述实验性能，本课题使用了如下的评价指标：

表 4-1 LIAR 数据集举例

新闻	“The number of illegal immigrants could be 3 million. It could be 30 million.”
作者	Donald Trump
工作职称	President-Elect
家乡	New York
政党	Republican
历史信用	(63, 114, 51, 37, 61, 14)
演讲地	Phoenix, Arizona
新闻主题	Immigration
真实性	Pants-fire

表 4-2 前 20 个主题

economy	health-care	immigration	federal-budget	crime
taxes	energy	abortion	state-budget	campaign-finance
children	deficit	elections	foreign-policy	guns
education	candidates-biography	jobs	congress	corrections-and-updates

表 4-3 实验参数设置

参数	设置
批大小 (Batch size)	128
学习率	0.01
优化器	批随机梯度下降
Dropout	0.8
动量 (momentum)	0.8
L_2 正则系数	1e-3
文本抽取器 CNN 卷积核	[2,3,4]
文本抽取器 CNN 卷积核数量	20
文本抽取器 LSTM 维度	300
上下文信息初始化	均匀分布 $\in [-0.1,0.1]$
未登录词初始化	均匀分布 $\in [-0.1,0.1]$

- 准确率 (Accuracy): 反应模型正确预测的比率;

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4-1)$$

- 精准率 (Precision): 计算预测出来的某类样本中, 有多少是被正确预测的;

$$Precision = \frac{TP}{TP + FP} \quad (4-2)$$

- 召回率 (Recall): 针对原先实际样本而言, 有多少样本被正确的预测出来了;

$$Recall = \frac{TP}{T + FN} \quad (4-3)$$

- F1 分数 (F1 Score): 对精准率与召回率的加权平均:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4-4)$$

表4-4给出了关于 TP, FP, FN, TN 的相关解释。

表 4-4 关键词解释

关键字	评价指标	解释
TP	True Positive	真阳性: 预测为正, 实际为正
FP	False Positive	假阳性: 预测为正, 实际为负
FN	False Negative	假阴性: 预测为负, 实际为正
TN	True Negative	真阴性: 预测为负, 实际为负

4.2.2 假新闻检测对比方法

为了更好的对比实验性能，本课题实现了若干对比方法，现对如下方法做简要介绍：

- Hybrid-CNN^[16]：Hybrid-CNN 使用 CNN 对新闻的文本内容进行特征抽取，使用 LSTM 编码上下文信息，经过 Max-pooling 后使用 CNN 进行高阶特征抽取；
- LSTM-Attention^[27]：新闻的上下文信息被用在两处：1. 在使用 LSTM 对文本进行特征抽取时作为注意力因子赋予词不同权重 2. 使用另一个 LSTM 对上下文进行编码作为另一个特征集；
- MMFD^[28]：MMFD 使用 CNN-LSTM 分别对文本、上下文信息、真实性检测报告等数据进行特征抽取，然后使用注意力机制将各个特征结合起来；
- Memory-Network^[32]：使用 Memory Network 检测假新闻，新闻的上下文信息作为注意力因子。

4.2.3 主题分类对比方法

本论文实现了如下主题分类模型作为对比：

- RF：使用 TF-IDF 作为特征，随机森林（Random Forest）作为分类器；
- Bi-LSTM^[33]：使用双向 LSTM 对新闻的文本内容进行特征提取，全连接神经网络做分类器；
- CNN^[34]：使用 CNN 作为文本特征抽取器，全连接神经网络做分类器；
- SWEM^[21]：Simple Word-Embedding based Model 直接对词向量进行 Pooling，后接全连接神经网络做分类器；
- LEAM^[35]：Label Embedding Attentive Model 通过计算文本中的词语与标签的相似度，给与每个词权重，然后对权重后的词向量进行 Pooling，后接全连接神经网络做分类器。

4.3 性能对比

4.3.1 假新闻检测性能对比

表4-5列出了各个模型在主题个数为二十一个时假新闻检测的性能表现。可以看到，除了 Precision 外，本课题提出的模型 FDML 在其他评价标准下均较其他方法有较大的提升。得益于多任务学习机制以及对上下文信息的运用，FDML 模型性能达到了最优。具体地，在所有的对比方法中，William Yang Wang 最早提出的 Hybrid-CNN 性能最差，Macro F1 Score 仅有 0.260；Memory-Network 取得了最好的结果，其 Accuracy 达到了 0.440，Macro F1 Score 达到了 0.432。相较于

Hybrid-CNN 模型, Memory-Network 在 Accuracy 和 Macro F1 Score 上分别获得了 12.7% 和 17.2% 的提升。本课题提出的模型在 Accuracy、Macro Recall 和 Macro F1 Score 上均大幅超过对比方法。比如, FDML 在整体性能指标 Macro F1 Score 上分别提升 25.7%, 15.7%, 9.1%, 8.5% (相较于 Hybrid-CNN、LSTM-Attention、MMFD、Memory-Network)。为了探究假新闻检测在不同数量主题下的性能, 本课题

表 4-5 假新闻检测模型性能对比

Model Name	Label	Precision	Recall	F1 Score	Accuracy	Macro-F1
Hybrid-CNN	Pants-fire	0.600	0.045	0.085	0.313	0.260
	False	0.327	0.518	0.401		
	Barely-True	0.266	0.155	0.196		
	Half-True	0.311	0.3411	0.325		
	Mostly-True	0.344	0.4780	0.400		
	True	0.216	0.111	0.152		
LSTM-Attention	Pants-fire	0.409	0.274	0.328	0.375	0.360
	False	0.406	0.422	0.413		
	Barely-True	0.382	0.325	0.351		
	Half-True	0.356	0.535	0.427		
	Mostly-True	0.421	0.436	0.428		
	True	0.399	0.143	0.210		
MMFD	Pants-fire	0.706	0.416	0.524	0.432	0.426
	False	0.394	0.586	0.471		
	Barely-True	0.611	0.279	0.383		
	Half-True	0.374	0.547	0.444		
	Mostly-True	0.418	0.522	0.464		
	True	0.867	0.159	0.269		
Memory-Network	Pants-fire	0.668	0.434	0.526	0.440	0.432
	False	0.468	0.507	0.487		
	Barely-True	0.468	0.351	0.402		
	Half-True	0.403	0.532	0.458		
	Mostly-True	0.366	0.588	0.451		
	True	0.905	0.157	0.267		
FDML	Pants-fire	0.629	0.590	0.609	0.507	0.517
	False	0.468	0.568	0.513		
	Barely-True	0.503	0.406	0.449		
	Half-True	0.487	0.472	0.479		
	Mostly-True	0.503	0.555	0.527		
	True	0.557	0.489	0.521		

题完成了在不同主题数量下各个模型的实验。如图4-1所示, 横轴表示不同的主题数量, 如 2 表示仅含有两个主题, 21 则表示包含所有主题, 也即对应表4-5列出的结果。同表4-5一样, 为了更全面的探究各个模型的性能, 图 2-6 画出了各个模型在 Accuracy、Macro Precision、Macro Recall 和 Macro F1 Score 上的表现。图4-1中

在各个子图上方标记了各个子图所表示的性能指标。

从表4-5中可以看到，FDML 在 Macro Precision 这一指标下性能不如 MMFD 和 Memory-Network。其中，Memory-Network 在 Pants-fire 和 True 类上的 Precision 较高。然而，其 Recall 都相对较低。比如在类别 True 上，虽然 Precision 达到了 0.905，但其 Recall 非常低，仅有 0.157。而 FDML 虽然 Precision 仅有 0.557，但由于其较高的 Recall 0.489，所以 F1 Score 相较于 Memory-Network 获得了 25.4% 的提升。

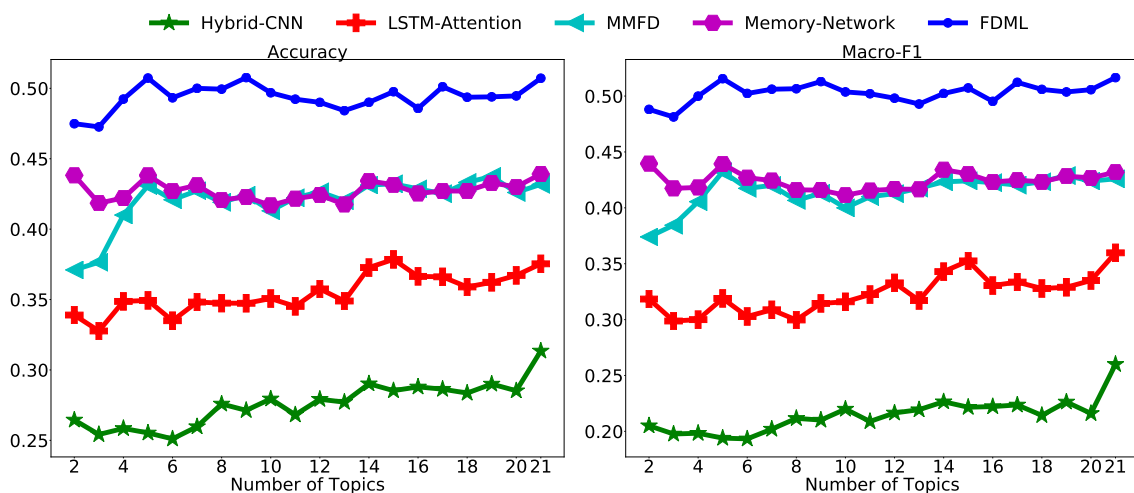


图 4-1 假新闻检测性能比较：Topic 2 至 Topic21

4.3.2 主题分类性能对比

表4-6展示了详细的性能对比结果。可以看出，随着主题数量的增多，各个模型的性能均呈下降趋势，因为任务变得相对困难。基于 TF-IDF 的方法 RF 表现最差，因为其离散的词表达方式无法编码语义信息，而这对主题分类是十分重要的。本课题提出的模型 FDML 在各个 Topic 数目上都表现较佳，特别是当任务变得困难时（Topic 数目变多），FDML 与其他模型的性能进一步拉大。

图4-2展示了在不同主题数量下主题分类的性能比较。

为了探究本论文提出多任务学习模型的有效性，本论文基于 FDML 构建了一个新的模型：FDML-。FDML-是 FDML 的一个变形，相较于 FDML，FDML-是一个单任务模型，即仅优化假新闻检测这一个任务，主题分类的相关部分被去除。图4-3展示了 FDML-和 FDML 在 Accuracy 和 Macro F1 上的相应表现。得益于有效的特征抽取与特征融合，FDML 和 FDML-均达到了较好的性能，在 Accuracy 和 Macro F1 Score 下，FDML 在不同主题数量下均优于 FDML-。得益于多任务学习机制，FDML 能够捕捉新闻主题和新闻真实性的关系，从而进一步提升实验效果。

表 4-6 主题分类性能比较: Topic2 至 Topic 21

K	Accuracy						Macro-F1					
	RF	Bi-LSTM	CNN	SWEM	LEAM	FDML	RF	Bi-LSTM	CNN	SWEM	LEAM	FDML
2	0.911	0.942	0.925	0.929	0.944	0.945	0.911	0.942	0.925	0.929	0.944	0.945
3	0.798	0.879	0.861	0.845	0.831	0.876	0.786	0.871	0.853	0.836	0.818	0.867
4	0.778	0.865	0.849	0.840	0.835	0.860	0.776	0.862	0.846	0.836	0.831	0.855
5	0.717	0.818	0.810	0.792	0.797	0.827	0.707	0.806	0.799	0.783	0.786	0.816
6	0.670	0.768	0.767	0.744	0.760	0.766	0.664	0.763	0.762	0.739	0.753	0.761
7	0.667	0.757	0.751	0.735	0.752	0.763	0.666	0.752	0.746	0.732	0.750	0.757
8	0.662	0.742	0.741	0.731	0.728	0.743	0.665	0.740	0.738	0.730	0.726	0.740
9	0.670	0.741	0.744	0.731	0.735	0.743	0.681	0.740	0.748	0.739	0.741	0.743
10	0.655	0.717	0.720	0.707	0.718	0.730	0.658	0.714	0.717	0.709	0.722	0.727
11	0.647	0.712	0.713	0.705	0.715	0.725	0.649	0.709	0.709	0.707	0.717	0.726
12	0.637	0.713	0.712	0.703	0.714	0.718	0.653	0.719	0.719	0.713	0.724	0.727
13	0.618	0.692	0.685	0.673	0.687	0.710	0.628	0.695	0.689	0.679	0.690	0.716
14	0.590	0.658	0.654	0.643	0.658	0.681	0.580	0.641	0.638	0.631	0.642	0.679
15	0.581	0.654	0.653	0.639	0.646	0.680	0.572	0.642	0.642	0.630	0.634	0.679
16	0.577	0.644	0.647	0.630	0.640	0.668	0.566	0.621	0.630	0.613	0.622	0.661
17	0.568	0.633	0.637	0.618	0.629	0.653	0.545	0.605	0.615	0.590	0.604	0.634
18	0.560	0.625	0.630	0.612	0.615	0.649	0.538	0.593	0.601	0.586	0.584	0.626
19	0.553	0.611	0.615	0.597	0.603	0.632	0.521	0.559	0.568	0.550	0.560	0.587
20	0.534	0.604	0.601	0.583	0.593	0.624	0.485	0.531	0.532	0.514	0.526	0.557
21	0.525	0.555	0.559	0.548	0.552	0.580	0.430	0.481	0.475	0.448	0.482	0.514
Avg.	0.646	0.716	0.714	0.700	0.708	0.729	0.634	0.699	0.698	0.685	0.693	0.716

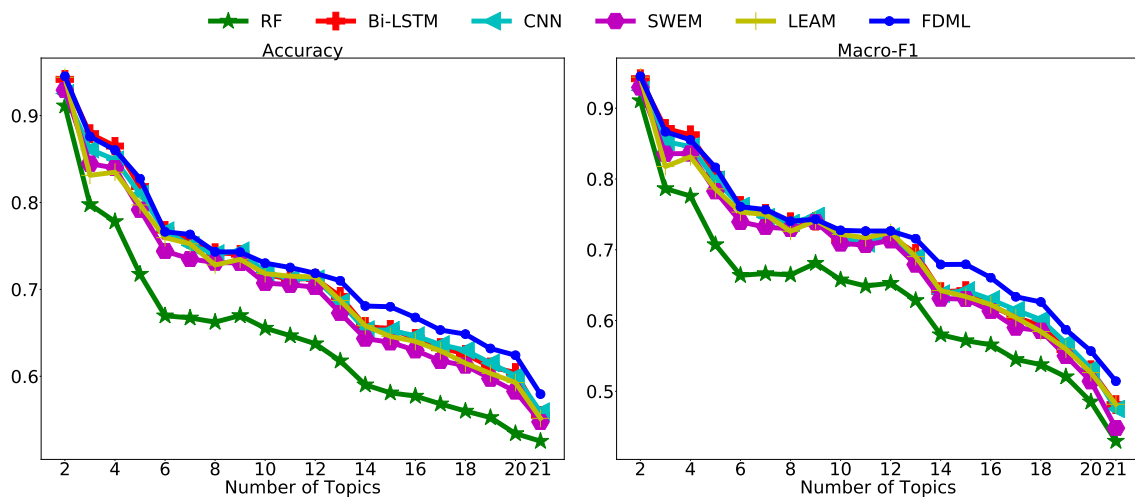


图 4-2 主题分类性能比较: Topic 2 至 Topic21

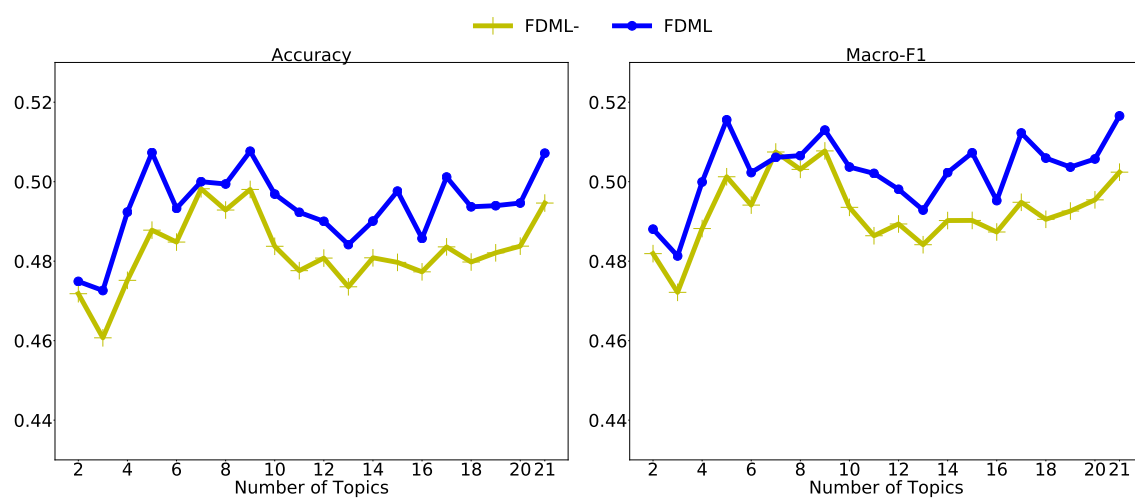


图 4-3 FDML 和 FDML-的性能比较

结 论

学位论文的结论作为论文正文的最后一章单独排写，但不加章标题序号。

结论应是作者在学位论文研究过程中所取得的创新性成果的概要总结，不能与摘要混为一谈。博士学位论文结论应包括论文的主要结果、创新点、展望三部分，在结论中应概括论文的核心观点，明确、客观地指出本研究内容的创新性成果（含新见解、新观点、方法创新、技术创新、理论创新），并指出今后进一步在本研究方向进行研究工作的展望与设想。对所取得的创新性成果应注意从定性和定量两方面给出科学、准确的评价，分（1）、（2）、（3）…条列出，宜用“提出了”、“建立了”等词叙述。

参考文献

- [1] Gottfried J, Shearer E. News Use Across Social Media Platforms 2016[M]. [S.l.] : Pew Research Center, 2016.
- [2] Shu K, Sliva A, Wang S, et al. Fake news detection on social media: A data mining perspective[J]. ACM SIGKDD Explorations Newsletter, 2017, 19(1) : 22-36.
- [3] Allcott H, Gentzkow M. Social media and fake news in the 2016 election[J]. Journal of economic perspectives, 2017, 31(2) : 211-36.
- [4] Castillo C, Mendoza M, Poblete B. Information credibility on twitter[C] //Proceedings of the 20th international conference on World wide web. 2011 : 675-684.
- [5] Kwon S, Cha M, Jung K, et al. Prominent features of rumor propagation in online social media[C] // 2013 IEEE 13th International Conference on Data Mining. 2013 : 1103-1108.
- [6] Jin Z, Cao J, Guo H, et al. Multimodal fusion with recurrent neural networks for rumor detection on microblogs[C] //Proceedings of the 25th ACM international conference on Multimedia. 2017 : 795-816.
- [7] Wang Y, Ma F, Jin Z, et al. Eann: Event adversarial neural networks for multi-modal fake news detection[C] //Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018 : 849-857.
- [8] Ma J, Gao W, Mitra P, et al. Detecting rumors from microblogs with recurrent neural networks.[C] //Ijcai. 2016 : 3818-3824.
- [9] Liu Y, Wu Y-F B. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks[C] //Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
- [10] Ma J, Gao W, Wong K-F. Detect rumor and stance jointly by neural multi-task learning[C] //Companion of the The Web Conference 2018 on The Web Conference 2018. 2018 : 585-593.
- [11] Yang F, Liu Y, Yu X, et al. Automatic detection of rumor on Sina Weibo[C] //Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. 2012 : 13.

- [12] 程亮, 邱云飞, 孙鲁. 微博谣言检测方法研究[D]. [S.l.]: [s.n.], 2013.
- [13] Wu K, Yang S, Zhu K Q. False rumors detection on sina weibo by propagation structures[C] // 2015 IEEE 31st international conference on data engineering. 2015 : 651-662.
- [14] Sun S, Liu H, He J, et al. Detecting event rumors on sina weibo automatically[C] // Asia-Pacific Web Conference. 2013 : 120-131.
- [15] 毛二松, 陈刚, 刘欣, et al. 基于深层特征和集成分类器的微博谣言检测研究[J]. 计算机应用研究, 2016, 33(11): 3369-3373.
- [16] Wang W Y. " liar, liar pants on fire": A new benchmark dataset for fake news detection[J]. arXiv preprint arXiv:1705.00648, 2017.
- [17] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C] // Advances in neural information processing systems. 2013 : 3111-3119.
- [18] Peters M E, Neumann M, Iyyer M, et al. Deep contextualized word representations[J]. arXiv preprint arXiv:1802.05365, 2018.
- [19] Radford A, Narasimhan K, Salimans T, et al. Improving language understanding by generative pre-training[J]. URL https://s3-us-west-2.amazonaws.com/openai-assets/researchcovers/languageunsupervised/language_understanding_paper.pdf, 2018.
- [20] Devlin J, Chang M-W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.
- [21] Shen D, Wang G, Wang W, et al. Baseline needs more love: On simple word-embedding-based models and associated pooling mechanisms[J]. arXiv preprint arXiv:1805.09843, 2018.
- [22] Elman J L. Finding structure in time[J]. Cognitive science, 1990, 14(2): 179-211.
- [23] Chung J, Gulcehre C, Cho K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint arXiv:1412.3555, 2014.
- [24] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention[C] // Advances in neural information processing systems. 2014 : 2204-2212.
- [25] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv:1409.0473, 2014.
- [26] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C] // Advances in neural information processing systems. 2017 : 5998-6008.

- [27] Long Y, Lu Q, Xiang R, et al. Fake news detection through multi-perspective speaker profiles[C] //Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers). 2017 : 252-256.
- [28] Karimi H, Roy P, Saba-Sadiya S, et al. Multi-source multi-class fake news detection[C] //Proceedings of the 27th International Conference on Computational Linguistics. 2018 : 1546-1557.
- [29] Ruchansky N, Seo S, Liu Y. Csi: A hybrid deep model for fake news detection[C] //Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017 : 797-806.
- [30] Pennington J, Socher R, Manning C. Glove: Global vectors for word representation[C] //Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014 : 1532-1543.
- [31] Bojanowski P, Grave E, Joulin A, et al. Enriching word vectors with subword information[J]. Transactions of the Association for Computational Linguistics, 2017, 5 : 135-146.
- [32] Pham T T. A Study on Deep Learning for Fake News Detection[J], 2018.
- [33] Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural networks, 2005, 18(5-6) : 602-610.
- [34] Kim Y. Convolutional neural networks for sentence classification[J]. arXiv preprint arXiv:1408.5882, 2014.
- [35] Wang G, Li C, Wang W, et al. Joint embedding of words and labels for text classification[J]. arXiv preprint arXiv:1805.04174, 2018.

攻读硕士学位期间发表的论文及其他成果

(一) 发表的学术论文

- [1] XXX, XXX. Static Oxidation Model of Al-Mg/C Dissipation Thermal Protection Materials[J]. Rare Metal Materials and Engineering, 2010, 39(Suppl. 1): 520-524. (SCI 收录, IDS 号为 669JS, IF=0.16)
- [2] XXX, XXX. 精密超声振动切削单晶铜的计算机仿真研究 [J]. 系统仿真学报, 2007, 19 (4): 738-741, 753. (EI 收录号: 20071310514841)
- [3] XXX, XXX. 局部多孔质气体静压轴向轴承静态特性的数值求解 [J]. 摩擦学学报, 2007 (1): 68-72. (EI 收录号: 20071510544816)
- [4] XXX, XXX. 硬脆光学晶体材料超精密切削理论研究综述 [J]. 机械工程学报, 2003, 39 (8): 15-22. (EI 收录号: 2004088028875)
- [5] XXX, XXX. 基于遗传算法的超精密切削加工表面粗糙度预测模型的参数辨识以及切削参数优化 [J]. 机械工程学报, 2005, 41 (11): 158-162. (EI 收录号: 2006039650087)
- [6] XXX, XXX. Discrete Sliding Mode Control with Fuzzy Adaptive Reaching Law on 6-PEES Parallel Robot[C]. Intelligent System Design and Applications, Jinan, 2006: 649-652. (EI 收录号: 20073210746529)

(二) 申请及已获得的专利 (无专利时此项不必列出)

- [1] XXX, XXX. 一种温热外敷药制备方案: 中国, 88105607.3[P]. 1989-07-26.

(三) 参与的科研项目及获奖情况

- [1] XXX, XXX. XX 气体静压轴承技术研究, XX 省自然科学基金项目. 课题编号: XXXX.
- [2] XXX, XXX. XX 静载下预应力混凝土房屋结构设计统一理论. 黑龙江省科学技术二等奖, 2007.

本人郑重声明：此处所提交的学位论文《基于多任务深度学习的假新闻检测》，是本人在导师指导下，在哈尔滨工业大学攻读学位期间独立进行研究工作所取得的成果，且学位论文中除已标注引用文献的部分外不包含他人完成或已发表的研究成果。对本学位论文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。

作者签名: _____ 日期: _____ 年 _____ 月 _____ 日

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果，知识产权归属哈尔滨工业大学。学位论文的使用权限如下：

(1) 学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文，并向国家图书馆报送学位论文；(2) 学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务；(3) 研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时，应征得导师同意，且第一署各单位为哈尔滨工业大学。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。

本人知悉学位论文的使用权限，并将遵守有关规定。

作者签名: _____ 日期: _____ 年 _____ 月 _____ 日

导师签名: _____ 日期: _____ 年 _____ 月 _____ 日

致 谢

衷心感谢导师 XXX 教授对本人的精心指导。他的言传身教将使我终生受益。

.....

感谢哈工大 L^AT_EX 论文模板 hiThesis !