

读书笔记 2: Fake News Detection on Social Media:

A Data Mining Perspective

文章从数据挖掘的角度讲述了对于社交媒体虚假新闻的检测（限制点：社交媒体）；

全文按照以下顺序展开：虚假新闻的定义、虚假新闻的检测、评价指标、相关领域的工作与未来方向。

(1) 对于虚假新闻的定义：文章给出了广义和狭义的定义，并最终选取狭义作为文章对于虚假新闻的定义：即真实性和意图性，并说明了从狭义的成果也适用于广义定义的检测；在定义这一块，文章分别从传统媒体的心理基础和社会基础、社交媒体的恶意用户和回音室效应对虚假新闻进行了特征分析。对于心理基础：主要源于读者倾向于相信自己的认知是唯一准确的，而认为不同意见的人是不合理的，同时他们也更容易接受支持他们观点的信息；对于社会基础：主要源于社会认同的安全感，在发布者谋求短期效用或者消费者谋求心理效用时虚假新闻就出现了；对于恶意用户：文章介绍了机器人、半机器人和 trolls 他们是虚假新闻的主要制造者和传播者；对于回音室效应：主要分为社会信用、频率启发，当人们觉得来源可靠时倾向于相信信息，同时如果是一个经常听见的消息即使是虚假新闻，也会相信它。

(2) 对于虚假新闻的检测：首先文章给出了问题的定义：包括发布者特征、内容特征、新闻特征（用户列表、帖子列表、时间）、预测函数（二分类任务）；在文章给出的虚假新闻检测方法中主要分为：特征提取和模型构建。

对于特征提取：分为内容特征和社会特征：对于内容特征：包含了源、标题、正文、图像视频，基于这些内容又可以构建基于语言和视觉的特征提取；对于语言特征：主要分为通用领域以及特定领域，通用领域分为词汇以及句法，特定领域则是包括引文、外链、图的数量和长度等等；对于视觉特征：包括图片、视频，比较常见的特征有清晰度、相关性、数量图像比等等。对于社会特征：分为用户、发帖、网络，考虑到用户如之前所定义的可能是机器人账户，因此需要获取用户的信息和特征，在这里又可以划分为个人层次（可信度）和组层次（关联性），同时对于发帖的特征提取可以分为 post、group、temporal；而对于网络特征则是根据回音室效应来考虑态度网络、相关网络、朋友圈网络、传播网络的特征。

对于模型构建：对应于特征提取，分为内容模型和社会模型；对于内容模型：分为知识型和风格型；对于知识型：为了检测到主要观点的真实性，主流的方法是人工检查、读者投票、计算机算法检测，在计算机算法检测中，首先是需要对观点陈述进行特征提取，然后根据开放网络或者知识图谱判断观点的虚实；对于风格型：又分为了欺骗性导向和客观性导向，欺骗性导向是通过 PCFG 或者 CNN 等算法检测欺骗性短语，客观性导向则是根据语言特征提取来检测新闻的客观性。对于社会模型：分为了基于态度和基于传播，而态度的立场一般会分为显式（明确的点赞或不同意）和暗式（含蓄的表达，需要特征学习，如 LDA、Tacchini 等），基于传播则是根据相关性预测新闻的可信度，分为同质（单一事件、发帖）、异构（多事件、子事件、发帖）。

(3) 对于评估指标，文章首先比较了现有数据集的优缺点，并指出没有一个数据集能够提供需要的所有种类的特征，所以构建一个合适的数据集是一个更好的发展方向。而在给出二分类任务的指标时，也给出了最经典的评估标准，TP、TN、FN、FP；以及由他们定义的召回率、准确率、精确率、F1、ROC、TPR、FPR、AUC，（因为这些指标是十分常见的二分类评估指标，所以就不展开了）。

(4) 在相关领域上，文章讲述了谣言分类（检测、跟踪、态度分类、真实度）对于虚

假新闻检测的意义，也点出它们所受因素的不同；还讲述了事实分辨、标题党、垃圾邮件对于虚假新闻的意义。在未来领域的方向上，文章提出了四个大方向：数据导向、特征导向、模型导向、应用导向；在数据导向上又提出了建立数据库、早期预警、捕捉回音室效应的基于数据的意图检测；在特征和模型导向上则提出了一些改进方法；在应用导向上提出可以将虚假新闻的检测推广到传播等方面。

感想：

- (1) 文章对于虚假新闻的检测确实事无巨细的讲述了每一个可能影响的因素，让我对于虚假新闻这一块有了更加深入的理解，特别是对于一些新闻事件的特征提取有了比较清晰地认识，涵盖了词汇、句法、风格、社会语境、网络等特征的考虑确实对于如今越来越多质量参差不齐的新闻提取提供了一种有效的参考。
- (2) 文章所考虑的角度确实很独到，不仅仅是从心理学、社会学，还考虑了回音室效应，给了我们研究问题时多角度分析的启发，特别是一些跨专业领域的应用。
- (3) 文章所提出的一些现象，包括机器人账户、虚假新闻的特征让我意识到虚假新闻确实渐渐成为我们日常生活中一种常见的现象，对于这一方面的研究确实十分有必要，并且有很大的研究空间。

改进方案：

- (1) 文章考虑了文本内容特征和社会特征，并在之后分别进行了模型建立，如果能够对这些特征进行综合加权表示，融合成一个模型会有更加完善的体系模型。
- (2) 二分类的模型可能不是对于新闻预测最好的方法，因为新闻本身会存在真假并存的现象，因此采用更加有效的模型可能是优化的一种方法。
- (3) 因为真实新闻的证据会比较多一些，模型可以从检测新闻是否为真实新闻的角度，反推出虚假新闻的可能性。
- (4) 在数据的特征提取时，可以采用像事件结构以及事件嵌入的方式进行类型的整合。
- (5) 正如文章提到的：建立合适的数据集也是很重要的、对于 **visual-based** 的方面需要进一步先进的特征学习方法。