

# 一种量子概率启发的对话讽刺识别网络模型

张亚洲, 俞 洋, 朱少林, 陈 锐, 戎 璐, 梁 辉

(郑州轻工业大学软件学院, 河南郑州 450002)

**摘要:** 对话讽刺识别已经成为人工智能领域中一项极具挑战性的课题,其目的是辨别互动对话中晦涩难懂的诸如讽刺、轻蔑、嘲笑等隐喻性情感。从语言哲学分析,目前的对话讽刺识别方法难以衡量人类语言在讽刺表达与理解方面固有的不确定性。鉴于量子概率在建模不确定性方面的优势,本文探索量子概率在讽刺识别领域的潜力并提出一种量子概率启发式网络。该网络主要包含复值嵌入层、量子复合层、量子测量层以及全连接层。本文将互动对话中每句话语视作是一组单词的类量子叠加,表征为复数向量。相邻话语之间的上下文交互被建模为量子系统与其周围环境的复合,表示为密度矩阵。本文对每句话语进行量子测量,提取讽刺特征,并将讽刺特征输入到全连接层预测得到讽刺识别结果。本文在两个基准数据集上进行实验,结果表明本文提出的模型优于先进讽刺识别模型,讽刺识别准确率分别提升 5.2% 与 2.38%。

**关键词:** 讽刺识别; 互动对话; 深度学习; 人工智能; 量子理论

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 0372-2112(2022)08-1885-09

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20211075

## A Quantum Probability Inspired Network for Dialogue Sarcasm Recognition

ZHANG Ya-zhou, YU Yang, ZHU Shao-lin, CHEN Rui, RONG Lu, LIANG Hui

(Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou, Henan 450002, China)

**Abstract:** Dialogue sarcasm recognition has been a challenging artificial intelligence(AI) research topic, aiming to discover elusive ironic, contemptuous and metaphoric information implied in daily dialogue. From the perspective of emotional logic, most existing works are insufficient to measure the intrinsic uncertainty in emotional expression and understanding. In view of the advantages of quantum probability(QP) in modeling the uncertainty, this paper explores the potential of QP in dialogue sarcasm recognition and proposes a quantum probability inspired network(QPIN). Specially, QPIN consists of a complex-valued embedding layer, a quantum composition layer, a quantum measurement layer and a dense layer. Each utterance is treated as a quantum superposition-like of a set of basis words, using a complex-valued representation. The contextual interaction between adjacent utterances is described as the composition system between a quantum system and its surrounding environment, which is represented by the density matrix. A quantum measurement is performed on the density matrix of each utterance to extract sarcastic features, and thus feeds these features to a dense layer to yield the probabilistic outcomes. Extensive experiments are conducted on two benchmark datasets, and the results show that our model outperforms the state-of-the-art baselines, with accuracy scores enhanced by 5.2% and 2.38%, respectively.

**Key words:** sarcasm recognition; interaction dialogue; deep learning; artificial intelligence; quantum theory

### 1 引言

现代汉语词典将“讽刺”解释为“讽刺是一种修辞策略,通过夸张、比喻等手法对人或事进行揭露、批评或嘲笑”。讽刺表达的字面含义往往与真实意图截然相

反,从而可以完全扭转作者的情感极性,传递出批评或嘲弄的隐含情绪。随着互联网与社交网络的迅速发展,越来越多的用户参与网上聊天、评论互动等,讽刺语言的使用日趋频繁,在日常对话中扮演着重要的角色。例

收稿日期:2021-08-11;修回日期:2022-04-11;责任编辑:李勇锋

基金项目:国家自然科学基金(No.62006212);河南省重点研发与推广专项(科技攻关)(No.212102210418)

如,“你看这位大哥面相如何?”,“我觉得他长得真有意,太抽象了”。

识别文本中的讽刺情感在许多领域都有广泛的应用,例如帮助企业预测消费者对其产品的态度,根据用户偏好进行个性化推荐,或帮助政府机关了解民众的舆论态势等<sup>[1]</sup>。因此,学术界与工业界对识别文本中的讽刺产生了浓厚的兴趣<sup>[2,3]</sup>。一般而言,讽刺识别是指利用自然语言处理技术、统计知识、机器或深度学习等,对语句、文档、对话等不同粒度的文本的讽刺极性进行鉴别。讽刺识别也属于情感分类的子任务。传统的讽刺识别方法主要集中在叙述式文本,例如产品评论,微博等,没有涉及到用户之间的互动对话。

对话讽刺识别正在成为该领域一个崭新且更具挑战性的研究课题,主要是因为:(1)互动对话中,每位谈话者并不独立,而是持续受到其他谈话者的影响,导致其讽刺情绪前后发生变化;(2)谈话者间的交互,默认隐藏了许多信息,例如他们的性别、周围环境、文化背景等<sup>[4,5]</sup>。目前的对话讽刺识别方法主要探讨上下文作用或学习上下文依赖,难以考虑自然语言固有的不确定性。

**量子的必要性解释** 在语言哲学范畴上,根据已有的研究成果<sup>[6,7]</sup>,自然语言固有的不确定性是指人类情感活动的自发性,不经过任何先前的知觉,而由身体的组织、精力或由对象接触外部感官而发生于自身的原始感官印象,即主观情感无需经过任何理性推理过程而自动生成,且情感活动的变化无需任何符合逻辑的理由。即使已经收集了全部先验知识,也可能无法提前预知人类的情感起伏。例1:小明发现他的手机坏了,小明可能会感到难过,生气,也可能感到高兴。如果小明生性节俭,那么小明会自然感到伤心;如果小明正好想换一部新手机,那么小明会有一种找到借口的高兴。因此,情感规律具备这样的内在不确定性。反映到书面语言(文本)上,指的是情感表达的不完备性与上下文性,使得其无法孤立地表达确切的情感<sup>[7]</sup>。已有的方法都是建立自经典概率和经典逻辑基础上,认为任何时刻(即使在决策判断之前),建模对象的状态也是确定的。然而一旦面对情感活动的不可确定性时,经典概率有时很难发挥作用。国内外的科学家们已经证实人类的情感与决策并不总是遵循经典概率,譬如琳达问题<sup>[8]</sup>,次序选择(即人们对先听好消息后听坏消息与先听坏消息后听好消息两种顺序有不同的情感态度)<sup>[9]</sup>等。

量子概率(Quantum Probability, QP)作为量子物理中建模不确定粒子行为的数学框架,已被用于描述人工智能中各种自然语言处理任务<sup>[10-13]</sup>。作为量子力学背后的抽象数学与统计解释,量子概率不应只被用于

描述微观物理世界的规律,而同样可以脱离原始的物理背景,作为一种数学框架应用于信息科学等宏观领域。注意,这种应用并不是把宏观系统还原为微观粒子的量子效应,而是将它看作是一个整体系统。鉴于基于量子理论的对话讽刺识别研究几近空白,本文计划将量子几何的哲学思想与数学主义应用于对话讽刺识别领域,从量子视角重新探讨情感表达与演化的本质属性,构建量子启发的对话讽刺识别网络模型,为自然语言处理与人工智能领域提供一种新思路。

本文提出一种量子概率启发式对话讽刺识别网络(Quantum Probability Inspired Network, QPIN)。具体而言,QPIN包含一个复值话语嵌入层,一个量子复合层,一个量子测量层以及一个全连接层。首先,本文将对话中每句话语视作是一组单词的类量子叠加,表示为复值向量。其次,我们将相邻话语之间的上下文交互建模为量子系统与其周围环境的交互,构成一个量子复合系统,由密度矩阵表示。再次,鉴于量子系统的信息与性质可以由量子测量结果的概率分布描述,我们对每句话语进行量子测量,进一步提取讽刺特征,将其输入到全连接层和softmax函数获得讽刺识别结果。

本文在MUSARD与2020 Sarcasm Detection Reddit Track两个基准数据集上进行实验评价,以验证QPIN模型的有效性。通过与众多前沿模型比较,例如卷积神经网络(Convolutional Neural Network, CNN),双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU),多头注意力长短期记忆网络(Multi-Head Attention-based Bidirectional Long Short Term Memory, MHA-BiLSTM),上下文LSTM(Contextual Long Short Term Memory, C-LSTM),双向自注意力变换网络(Bidirectional Encoder Representations from Transformers, BERT),上下文交互网络(Contextual Network, C-Net)以及多任务学习框架(Multi-Task Learning, MTL),表明本文的方法在F1与Acc指标上取得更高的分类结果。

本文的主要创新贡献有以下几点。(1)首次利用量子概率,提出面向对话讽刺识别的量子概率启发式网络。它能够自然地将量子理论引入到宏观人工智能领域,是一套从文本表征到上下文交互再到特征识别的端到端式理论结构,探索量子理论的潜力。(2)引入复数概率幅,提出一种复值话语表示方法。该方法将文本话语表征为指数形式的复值向量,将振幅与语义知识联系起来,将隐藏的相位角与情感知识联系起来,既能够同时捕捉语义与情感,也能够借助复数概率幅描述不确定性。(3)提出一种建模话语上下文性的量子复合表示方法。量子复合将能够以“全局到局部”方式建模上下文交互。

## 2 量子概率基础

### 2.1 量子概率

量子概率理论是由冯·诺依曼发展建立的一种基于线性代数的一般化概率理论,目的是解释量子理论的数学基础<sup>[14,15]</sup>.量子概率理论是量子力学背后的抽象数学与统计解释,它更关心的是符号之间抽象的关系与结构,而非符号对应的实物(例如物理量).因此,量子概率并不是只能描述微观粒子,同样可以脱离于原始物理背景,而去描述宏观系统中的类量子现象,例如人类决策判断、次序效应、认知干涉、不确定性等.

在量子概率中,量子概率空间封装于复数希尔伯特空间 $\mathbb{H}$ .希尔伯特空间是欧式空间的直接推广,是一个无限维的内积空间,被广泛应用于数学分析与量子力学中.量子概率中,假设一个量子状态向量 $u = (u_1 u_2 \dots u_n)^T \in \mathbb{H}$ ,记作左矢 $|u\rangle$ .它的转置向量,记作右矢 $\langle u|$ .两个状态向量 $|u\rangle$ 和 $|v\rangle$ 的外积构成一个矩阵,记作 $|u\rangle\langle v|$ .对于状态向量 $|u\rangle$ ,在 $u$ 方向上的投影算符可以写作 $\Pi = |u\rangle\langle u|$ ,代表量子概率空间的基本事件.

### 2.2 量子叠加

量子叠加是指一个量子系统可以同时处于多个互斥基态的叠加态,直到它被测量.测量之后,该系统从叠加态塌缩到其中一个基态上.假设 $\{|w_1\rangle, |w_2\rangle, \dots, |w_n\rangle\}$ 构成量子概率空间上的一组正交基,那么 $|u\rangle = \sum_{j=1}^n z_j |w_j\rangle$ ,其中 $z_j$ 是复数权值.

量子概率中,量子态可以处于纯态,也可以处于混合态.纯态对应于希尔伯特空间中的状态向量 $|u\rangle$ ,而混合态是由几种纯态依照概率组成的量子态,由密度矩阵 $\rho$ 表示.假设量子系统处于纯态 $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$

的混合中,对应的密度矩阵定义为: $\rho = \sum_j p_j |u_j\rangle\langle u_j|$ ,

其中 $p_j$ 代表每个纯态的经典概率,全部概率的总和为1.密度矩阵是经典理论中位置和状态概率分布的量子扩展,将量子态与经典不确定性容纳进同一体系下,描述了系统全部信息与性质.

## 3 量子概率启发式讽刺识别网络

本文提出的QPIN模型,如图1所示. QPIN模型包含一个复值嵌入层、一个量子复合层、一个量子测量层以及一个全连接层.

### 3.1 复值嵌入层

受李秋池的工作启发<sup>[11]</sup>,本文采用复值嵌入表示方法.鉴于单词是组成人类语言的基本单元,本文将对话中的每个单词 $w$ 视作一个基态 $|w\rangle$ ,假设 $\{|w_1\rangle, |w_2\rangle, \dots, |w_n\rangle\}$ 构成对话希尔伯特空间 $\mathbb{H}_{\text{dig}}$ 的正交基向量.本文采用独热编码(one-hot encoding)去表示每一个单词基向量例如第 $j$ 个基向量: $|w_j\rangle = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix}^T$ .

为了捕捉讽刺话语中的不确定性,本文将每句话语视作是一组单词基向量 $\{|w_1\rangle, |w_2\rangle, \dots, |w_n\rangle\}$ 的量子叠加.那么目标话语 $u_i$ 可以表示为:

$$|u_i\rangle = \sum_{j=1}^n z_j |w_j\rangle, \quad (1)$$

$$z_j = r_j e^{i\theta_j}.$$

其中, $z_j$ 是第 $j$ 个单词的复数概率幅,满足 $\sum z_j^2 = 1$ . $i$ 称为虚数单位, $r$ 是概率幅的振幅, $\theta$ 是相位角.本文赋予振幅与相位具体的含义,将振幅与语义信息关联,将相位角置为情感倾向程度,从而同时建模了语义与情感信息.

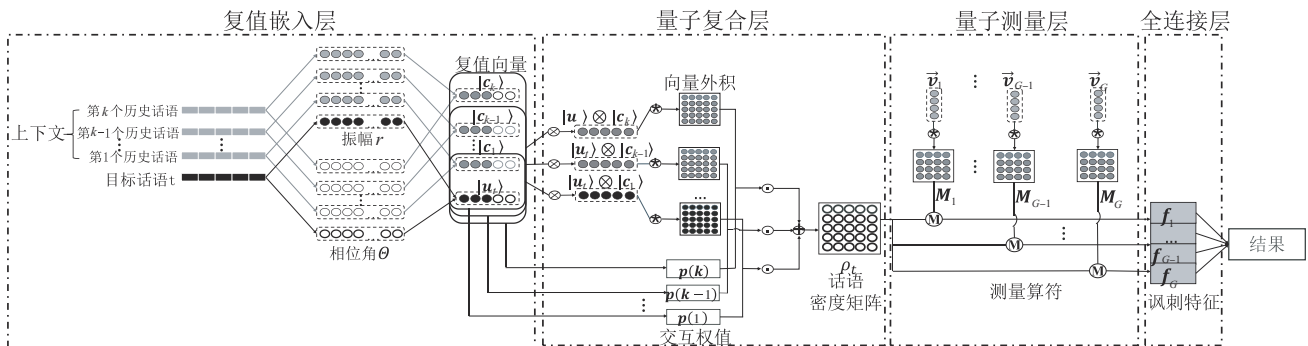


图1 类量子交互网络结构示意图. ⊗表示张量积, ⊙表示向量外积操作, ⊙表示逐位相乘, ⊕表示矩阵相加, ⊞表示量子测量操作

### 3.2 量子复合层

在量子概率中,理想的量子测量描述了被测系统,测量装置以及周围环境(例如临近系统)的完全交互.

但是,在实际测量中,我们认为测量装置与周围环境并不会同等地参与被测系统的交互,即它们的参与程度不是相等的,例如距离远的系统与距离近的系统对被

测系统的影响是不同的. 这种交互类似于对话中不同话语之间的交互. 不同的上下文话语表达着不同强度的人际交互.

本文将目标话语  $|u_t\rangle$  视作被测系统, 将其上下文  $\{|c_1\rangle|c_2\rangle\cdots|c_\lambda\rangle\cdots|c_k\rangle\}$  视作周围环境. 两者的交互构成一个量子复合系统, 例如目标话语  $|u_t\rangle$  与第  $\lambda$  个上下文话语  $|c_\lambda\rangle$  之间的交互构成了一个量子复合系统  $|\Psi_t^\lambda\rangle$ . 考虑到所有上下文的影响, 本文建模每一个上下文与目标话语的交互, 构造出  $k$  个不同的复合系统. 其中, 第  $\lambda$  个复合系统形式化为:

$$|\Psi_t^\lambda\rangle = |u_t\rangle \otimes |c_\lambda\rangle = (\Psi_{t,1}^\lambda, \Psi_{t,2}^\lambda, \dots, \Psi_{t,d}^\lambda)^T \quad (2)$$

其中,  $\lambda \in [1, 2, \dots, k]$ . 根据式(2), 可以得到  $k$  个复合系统, 即  $|\Psi_k\rangle = \{|\Psi_t^1\rangle| \Psi_t^2\rangle \cdots |\Psi_t^k\rangle\}$ , 从而学习到全部的上下文交互. 接下来, 目标话语的最终表示将被视作是  $k$  个对话中每一个目标话语被表示成一个复值向量  $|u_t\rangle = (r_1 e^{i\theta_1} r_2 e^{i\theta_2} \cdots r_n e^{i\theta_n})^T$ . 那么上下文话语  $\{c_1, c_2, \dots, c_\lambda, \dots, c_k\}$  均使用相同方法, 即式(1)表示成复值向量, 即  $\{|c_1\rangle|c_2\rangle \cdots |c_\lambda\rangle \cdots |c_k\rangle\}$ . 复值向量将作为每句话语的初始表示参与隐藏层的交互建模.

复合系统上的量子混合态, 表示为密度矩阵  $\rho_t$ :

$$\begin{aligned} \rho_t &= \sum_{\lambda=1}^k p(\lambda) |\Psi_t^\lambda\rangle \langle \Psi_t^\lambda| \\ &= \begin{bmatrix} \sum_{\lambda=1}^k p(\lambda) (\Psi_{t,1}^\lambda)^2 & \cdots & \sum_{\lambda=1}^k p(\lambda) \Psi_{t,1}^\lambda \Psi_{t,d}^\lambda \\ \vdots & \ddots & \vdots \\ \sum_{\lambda=1}^k p(\lambda) \Psi_{t,d}^\lambda \Psi_{t,1}^\lambda & \cdots & \sum_{\lambda=1}^k p(\lambda) (\Psi_{t,d}^\lambda)^2 \end{bmatrix} \end{aligned} \quad (3)$$

其中,  $p(\lambda)$  是第  $\lambda$  个复合系统的交互概率, 衡量第  $\lambda$  个上下文话语的交互程度, 我们在模型训练过程中自动更新它的值.

根据式(3), 目标话语已经由密度矩阵  $\rho_t$  表示. 本文使用密度矩阵表示的原因是: 密度矩阵能够统一目标话语的全部信息与性质, 例如语义知识, 情感信息, 上下文交互, 概率分布信息等.

### 3.3 量子测量层

在量子概率中, 量子系统的信息或性质, 将由测量结果的概率分布描述. 受此启发, 本文构造一组不同测量角度的测量算符  $\{M_\gamma\}_{\gamma=1}^G$  对目标话语的密度矩阵  $\rho_t$  测量, 以进一步抽取讽刺特征  $f_i = (f_1, f_2, \dots, f_G)$ .

测量算符  $\{M_\gamma\}_{\gamma=1}^G$  中,  $M_\gamma$  表示第  $\gamma$  个测量算符, 由第  $\gamma$  个测量向量  $|v_\gamma\rangle$  通过外积计算得到, 即  $M_\gamma = |v_\gamma\rangle \langle v_\gamma|$ .

第  $\gamma$  个特征分量提取过程为:

$$\begin{aligned} f_\gamma &= \text{tr} \left( M_\gamma (M_\gamma)^\dagger \rho_t \right) \\ &= \text{tr} \left( M_\gamma (M_\gamma)^\dagger \sum_{\lambda=1}^k p(\lambda) |\Psi_t^\lambda\rangle \langle \Psi_t^\lambda| \right) \\ &= \sum_{\lambda=1}^k p(\lambda) \text{tr} \left( M_\gamma (M_\gamma)^\dagger |\Psi_t^\lambda\rangle \langle \Psi_t^\lambda| \right) \end{aligned} \quad (4)$$

其中,  $\text{tr}$  是迹操作,  $\gamma \in [1, 2, \dots, G]$ .  $f_\gamma$  是特征向量  $f_i$  的第  $\gamma$  个特征分量, 得到  $f_i = (f_1, f_2, \dots, f_\gamma, \dots, f_G)$ .

### 3.4 卷积层

本文同样设计一个卷积层针对密度矩阵提取特征. 目的是调查量子测量与卷积层的特征提取效率对比, 旨在理解量子测量对宏观信息提取的潜力, 如图2所示.

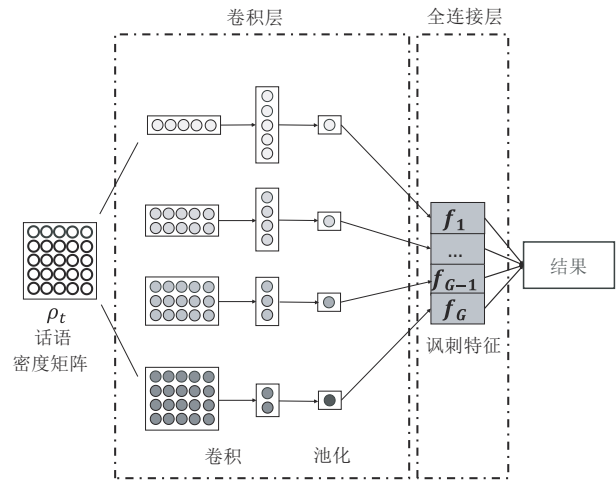


图2 卷积层对话语密度矩阵提取特征框架

本文尝试了不同的卷积核, 并根据最优实验结果, 设置了4个卷积核, 卷积核大小分别是  $\{1, 2, 3, 4\} \times d$ , 对目标话语的密度矩阵进行卷积操作, 对卷积后的特征最大池化, 将池化后的信息连接到一起构成讽刺特征  $f_i = (f_1, f_2, \dots, f_G)$ .

### 3.5 全连接层

将讽刺特征  $f_i$  输入到全连接层得到输出  $x_i = (x_1, x_2, \dots, x_\delta)$ , 通过 softmax 分类器产生讽刺标签  $\hat{y} \in \{y_{\text{sar}}, y_{\text{non}}\}$ :

$$\hat{y}_i = \text{softmax}(Wx_i + b) \quad (5)$$

$W$  和  $b$  表示权值矩阵与偏置向量. 本文采用交叉熵作为损失函数, 形式化为:

$$J = -\frac{1}{N} \sum_i \sum_c y_i^c \log \hat{y}_i^c + \mu \|\phi\|^2 \quad (6)$$

其中,  $N$  是训练集样本量,  $y_i^c$  表示真值标签,  $\hat{y}_i^c$  表示预测

标签,  $t$  是话语索引,  $c$  是类别索引,  $\mu \|\phi\|^2$  是正则项. 本文使用反向传播算法训练网络模型并更新参数. 为了避免训练过程中出现过拟合现象, 本文使用随机剪枝策略.

### 3.6 方法讨论

介绍量子概率启发式对话讽刺识别网络(下面简称为 QPIN)之后, 本文详细阐述并讨论其与现有深度神经网络方法的区别与相似.

**相似之处** 从模型构建角度分析, QPIN 与深度神经网络, 譬如卷积神经网络、全连接神经网络及长短期记忆网络等类似, 仍然采用深度学习与逐层训练的思路, 包含了输入层、隐藏层、输出层等基本组件, 通过深度多层抽象, 逐渐将初始的“低层”特征表示转化为“高层”特征表示. 整个训练是一个端到端式监督学习与拟合过程.

**区别之处** 从模型构建角度分析, QPIN 与已有的神经网络存在五点不同: (1) 整体架构不同, QPIN 是量子概率驱动的架构, 由量子理论中的核心组件自下而上搭建而成, 具备量子概率的数学支撑, 每一个组件都有物理解释, 不是作为“黑盒子”使用; (2) 输入层不同, 首次将复数带入到讽刺识别任务中, 将每句话表征为复值表示, 不再是实数向量; (3) 隐藏层不同, QPIN 采用量子符合与混合态构建隐藏层. 目标话语与上下文的交互, 被视作是  $k$  个复合系统上的量子混合态, 表示为密度矩阵; (4) 特征提取方式不同, 已有方法通常直接采用全连接层提取特征和降维, 而 QPIN 以一种测量的视角, 采用  $G$  个测量算符对目标话语的密度矩阵表示进行量子测量, 提取最终特征. 从研究目标分析, QPIN 作为量子讽刺识别领域的有效尝试, 旨在推动量子人工智能与量子信息处理的发展.

## 4 实验与结果

### 4.1 实验数据

本文采用 MUSARD<sup>[16]</sup> 和 2020 Sarcasm Detection Reddit Track<sup>[17]</sup> 两个基准数据集进行实验. MUSARD 数据集收集自“生活大爆炸”、“老友记”等情景喜剧, 共包含 690 个视频对话. 每个对话记录了目标话语以及对话上下文话语, 其中目标话语被标注为“讽刺”或“非讽刺”.

2020 SarcasmDetection Reddit Track (下面简称为 Reddit) 收集自 Reddit 论坛, 仅包含文本模态. 它共有 3 100 个讽刺博文, 3 100 个非讽刺博文以及 18 618 个上下文博文. 实验数据统计如表 1 所示.

### 4.2 评估指标与参数设置

本文采用精确率 (Precision, P)、召回率 (Recall, R)、微观 F1 (Micro-F1, Mi-F) 及准确率 (Accuracy, Acc)

表 1 实验数据统计

数据集	话语数目		上下文话语数目
	讽刺	非讽刺	
MUSARD	345	345	2 261
Reddit	3 100	3 100	18 618

作为性能评估指标, 详细参数设置如表 2 所示.

表 2 参数设置

超参数	数据集	
	MUSARD	Reddit
Batch	64	
激活函数	Relu	
学习率	0.001	
复值嵌入维度	768	
正则化系数	$10^{-5}$	
测量算符数目 $G$	1 400	800
Dropout	0.6	0.5

### 4.3 对比模型

为了评估 QPIN 的有效性与实际性能, 本文对比了一系列前沿基线模型. 它们分别是:

(1) CNN<sup>[18]</sup>: 它包含两个卷积层, 一个全连接层. 它采用预训练的 GloVe 词向量对话语文本执行讽刺分类.

(2) BiGRU: 它采用一个双向的门控循环网络去学习目标话语的隐藏状态, 输入到 softmax 函数执行讽刺识别. 它考虑了单词的历史和未来的上下文信息.

(3) MHA-BiLSTM<sup>[19]</sup>: 它使用自然语言处理领域内热门的多头注意力机制, 并与双向 LSTM 融合去抽取目标话语中更突出的特征, 学习更优的话语表示.

(4) C-LSTM<sup>[4]</sup>: 它首先利用 CNN 去提取话语特征, 然后将历史话语特征与目标话语特征拼接, 输入到 LSTM 中执行讽刺识别.

(5) SVM+BERT<sup>[16]</sup>: 它首先利用 BERT 得到目标话语的向量表示, 然后输入 SVM 分类器执行讽刺分类. 本文将 SVM 的核函数设置为“高斯核”. 此外, 本文也将上下文特征与目标话语特征拼接, 以考虑上下文的影响.

(6) C-Net: 它利用目标话语的标签去标注上下文话语, 然后利用目标话语与上下文训练 BERT, 学习话语上下文表示.

(7) MTL: 它是最前沿的多模态多任务学习框架, 首先提出段内与段外两种注意力机制去学习段内与段外信息, 拼接这两种信息得到话语表示. 其次, 利用情感知识帮助提升讽刺识别的性能. 为了公平比较, 本文只用文本与图像模态的结果.

(8) QMSA<sup>[12]</sup>: 为了验证本文提出 QPIN 的有效性, 本文对比了其他量子启发的多模态情感分类模型. 与 QPIN 不同, QMSA 采用的是实数 BERT 向量构建每一句

文本与图像文档的密度矩阵,并不考虑对话上下文,最后将密度矩阵输入到SVM分类器中执行讽刺分类。

(9)QSR<sup>[10]</sup>:它是一个基于量子理论的文本情感分类模型,利用量子语言模型与word2vec将其表征为量子叠加态,利用最大似然估计训练为密度矩阵,输入到随机森林分类器中执行分类。

(10)QMN<sup>[4]</sup>:QMN使用密度矩阵去表征文本与图像特征,通过LSTM提取上下文特征后,采用量子干涉对文本与图像特征进行融合。

此外,本文也设计了QPIN的三种变体,分别是QPIN-QM, QPIN-CNN和QPIN-QM-CNN。其中,QPIN-QM只使用量子测量层提取特征, QPIN-CNN只使用卷积层提取特征,而QPIN-QM-CNN将量子测量与卷积层提取的特征拼接到一起,组成的新特征输入到全连接层执行分类。

#### 4.4 MUStARD数据集结果分析

各个模型在MUStARD数据集上的实验结果如表3所示。MUStARD数据集采集自情景喜剧,对话中逻辑跳跃性强,由于是演绎性样本,台词隐喻性、刻意性强,比较考验模型的上下文理解与文本表征能力。因此,在该数据集上模型的上下文建模与语义捕捉能力影响其性能。BiGRU、MHA-BiLSTM、C-LSTM、SVM+BERT、C-Net、MTL以及QPIN模型全部优于CNN。原因是它们均考虑到了上下文信息,表明在对话讽刺识别中上下文的重要性。BiGRU、MHA-BiLSTM与C-LSTM三个RNN变体模型中性能处于同一水平,BiGRU获得最高F1结果,而C-LSTM获得最高准确率。相比于这三种模型,SVM+BERT在微观F1指标上表现更好,这依赖于预训练BERT能够提供更优良的话语表示。SVM+BERT显著地超越CNN,在F1与Acc指标上分别提升13.3%与9.7%。这展示了预训练语言模型的特征抽象能力。此外,BERT作为预训练模型,已经在庞大的数据集上训练完成。但是BERT计算量过于庞大且参数量远高于CNN。相比于BiGRU、MHA-BiLSTM与C-LSTM,SVM+BERT在F1指标上获得显著改进,分别提升了3.7%、5.4%与4.4%。但是在Acc指标上没有显著改进,甚至落后于MHA-BiLSTM与C-LSTM。原因可能是MUStARD数据集数量很少,且角色话语分布极度不平衡,导致MHA-BiLSTM与C-LSTM获得很低的F1分值,却拥有较高的Acc分值。LSTM的优势在于建模短文本上下文,一定程度上缓解了梯度消失的缺点。对于这种不平衡数据集,机器学习领域通用做法是更加注重F1指标上的表现。SVM+BERT(+上下文)通过将上下文特征与目标话语拼接,提升了微弱的性能。

作为最新的对话讽刺识别模型之一,C-Net通过建模谈话者间的依赖信息,超越SVM+BERT而获得不错

的结果。在众多基线模型中,MTL取得最优实验结果。相比于C-Net,MTL分别在F1与准确率方面提升了1.8%和1.7%。原因在于MTL既考虑到上下文信息,也利用了情感知识对讽刺识别的帮助。QMSA作为经典的量子启发的多模态情感分类模型,其性能表现不佳。原因是QMSA是浅层模型,仅仅利用定义构建密度矩阵,也并不考虑对话上下文,限制了文档的表征能力。QMSA不是端到端模型结构,密度矩阵自定义之后不会根据训练集而自适应学习。通过与QPIN-QM的结构对比,QPIN-QM在文档表示上引入了复值,包含了BERT实数部分与虚数部分,该表示相比于QMSA加入新的补充知识,即相位角。此外,引入了上下文复合表示与密度矩阵的可学习设置,使得能够根据上下文不同,自适应学习话语密度矩阵。这两个组件对QPIN模型性能的提升贡献度最大,也是QPIN-QM显著超越QMSA模型的主要原因,体现了QPIN的先进性与创新性。QSR性能结果相比于QMSA更差,原因是其既不考虑多模态信息,不涉及多模态表示,也不考虑对话上下文交互,严重限制QSR模型分类能力。QMN作为目前量子对话情感分类领域中最前沿的方法之一,性能结果非常良好,超越了C-Net,仅次于MTL,证明了量子对话情感分类方法的潜力。

表3 各个模型在MUStARD数据集上的实验结果

模型	评估指标			
	P	R	Mi-F	Acc
CNN	0.579	0.569	0.571	0.589
BiGRU	0.631	0.643	0.624	0.646
MHA-BiLSTM	0.673	0.652	0.614	0.659
C-LSTM	0.642	0.627	0.620	0.662
SVM+BERT	0.651	0.646	0.647	0.646
SVM+BERT(+上下文)	0.655	0.651	0.650	0.651
C-Net	0.679	0.683	0.681	0.683
MTL	0.727	0.719	0.716	0.721
QMSA	0.629	0.623	0.627	0.633
QSR	0.604	0.607	0.607	0.609
QMN	0.680	0.685	0.684	0.691
QPIN-CNN	0.744	0.739	0.740	0.737
QPIN-QM	0.753	0.755	0.754	0.753
QPIN-QM-CNN	<b>0.759</b>	<b>0.763</b>	<b>0.760</b>	<b>0.758</b>
ΔSOTA	4.40%	6.25%	6.26%	5.25%

#### 4.5 Reddit数据集结果分析

相比于MUStARD,Reddit数据集不仅样本量更大,每条话语也更长,且仅包含文本模态。但是Reddit数据集是论坛回复性样本,上下文的时效性较长,交互性较差。同时,每条话语都是长文本,对于捕捉关键信息,譬如表达

讽刺的单词或短语,更加困难.长文本内单词间的上下文性也需要考虑在内.因此从样本角度,Reddit数据集更加考验模型的语义、情感表征能力.各个模型在Reddit数据集上的实验结果如表4所示.CNN、BiGRU、C-LSTM性能最差,而MHA-BiLSTM表现相对较好.因为Reddit数据集中上下文话语篇幅较长,上下文建模更为困难,使得拼接上下文话语特征入LSTM效果不明显.SVM+BERT显著地超越CNN、BiGRU、C-LSTM,在F1与Acc分值上分别提升了1.4%、2.7%、6.3%、6.3%、4.5%、4.1%.这表明在数据量更大、数据分布平衡的数据集上,BERT凭借预训练模型显露出明显的优势.类似地,SVM+BERT(+上下文)凭借BERT的特征学习能力,超越了传统的神经网络结构,原因是LSTM抽取特征的能力远弱于Transformer.

C-Net与MTL作为最前沿的对话讽刺识别框架在各项指标上胜过其他基线模型,且两者获得近乎相等的识别结果.原因是两者均是基于预训练语言模型设计文本表征方法,如BERT和ALBERT,均可以建模长文本内部的单词上下文以及利用多头注意力机制捕捉长文本中的关键信息.而作为量子启发的多模态情感分析模型QMSA表现非常差,其准确率位列最后一位.QMSA模型的两大核心组件分别是密度矩阵与多模态干涉特征融合.鉴于Reddit数据集只包含文本数据,并不涉及多模态交互.QMSA中的量子启发的决策融合完全失效,只剩下利用量子理论定义文本密度矩阵,且不会随着训练过程而改进.初始密度矩阵表征能力有限且固定,这严重限制了QMSA在Reddit数据集上的性能.相比而言,QPIN的四个核心组件:复值表示、上下文复合表示、端到端式训练以及量子测量均没有受到数据集的影响.因此QPIN远远超过QMSA框架,获得最佳了性能.

本文提出的QPIN-CNN与QPIN-QM仍然优于C-Net与MTL,取得优良的识别结果,F1分数分别达到67.5%与68.0%.但是由于本文在Reddit数据集上仅选择了800个量子测量算符,远远少于在MUSTARD数据集中的1400个测量算符,使得QPIN-QM提取的特征维度减少,降低了信息承载与描述能力,导致QPIN-CNN与QPIN-QM相差无几,处于同一水平.这表明,测量算符的数目将会直接影响QPIN的性能.QPIN-QM-CNN以微弱优势超越了QPIN-CNN与QPIN-QM,取得最佳识别性能.原因是在测量算符数目较少的情况下,结合CNN以补充提取特征是一种性能补偿手段.相比于MTL,QPIN-QM-CNN分别在F1与准确率方面提升了2.3%和2.3%.

#### 4.6 量子概率与经典概率实验对比

本文已经分别从动机、理论层面详细描述了量子

表4 各个模型在Reddit数据集上的实验结果

模型	评估指标			
	P	R	Mi-F	Acc
CNN	0.631	0.623	0.629	0.623
BiGRU	0.600	0.598	0.600	0.600
MHA-BiLSTM	0.629	0.632	0.630	0.636
C-LSTM	0.613	0.609	0.610	0.615
SVM+BERT	0.640	0.642	0.638	0.640
SVM+BERT(+上下文)	0.649	0.647	0.643	0.651
C-Net	0.663	0.664	0.663	0.664
MTL	0.667	0.668	0.667	0.668
QMSA	0.610	0.614	0.611	0.617
QSR	0.603	0.607	0.607	0.614
QMN	0.654	0.659	0.657	0.663
QPIN-CNN	0.677	0.674	0.675	0.674
QPIN-QM	0.677	0.680	0.680	0.679
QPIN-QM-CNN	<b>0.683</b>	<b>0.688</b>	<b>0.686</b>	<b>0.684</b>
$\Delta$ SOTA	2.39%	2.42%	2.36%	2.38%

概率的潜力.为了支撑这一观点,本节将从实验角度出发,通过对比经典(贝叶斯)概率与量子概率的实验结果,验证量子概率的有效性.对于贝叶斯概率方法,本文将设计与QPIN相似的网络结构,以保证两者公平比较.鉴于贝叶斯概率并无复数、密度矩阵、量子复合与量子测量等概念,本文首先利用BERT得到每句话语的语义向量表示,并将每句话语的情感极性作为补充特征与之拼接,成为新特征(该设置对比复值嵌入表示).其次,对目标话语与其上下文的特征进行线性相加,用于捕捉上下文信息(该设置对比量子复合).再次,将上下文特征输入到朴素贝叶斯分类器中获得目标话语的预测概率,作为进一步提取的概率特征(该设置对比量子测量层).最后,将这些预测概率输入到全连接层得到讽刺识别结果.对比结果如表5所示.

可以观察到,在MUSTARD与Reddit数据集上,量子概率启发的网络QPIN在F1值与准确率方面均显著超越经典概率方法NB+BERT.该实验结果支撑了本文的理论论点,表明量子概率从动机、理论与实验都有较大潜力,可以作为另外一种更加一般化的概率体系去解决自然语言处理难题.

表5 量子概率与经典概率实验结果

数据集	方法	评估指标	
		Mi-F	Acc
MUSTARD	经典:NB+BERT	0.651	0.647
	量子:QPIN-QM	0.754	0.753
Reddit	经典:NB+BERT	0.656	0.652
	量子:QPIN-QM	0.680	0.679

#### 4.7 谈话者角色重要性分析

鉴于对话中谈话者角色也会影响讽刺极性的判断,例如情景喜剧“生活大爆炸”中“谢尔顿”这一角色相较于其他角色表达更多的讽刺,本文将谈话者的角色信息作为一种补充知识考虑到QPIN模型内,分析角色对模型性能的影响.此外,鉴于Reddit数据集都是论坛用户,用户名各种各样,并不体现特定讽刺信息,其角色信息并没有记录,因此我们只建模MUStARD数据集中的角色信息.我们采用两种方法引入角色信息以及角色对讽刺的影响:(1)利用BERT将每个角色名表征为向量,与对应的话语向量拼接成一个新向量,再执行后续的量子复合等操作;(2)受Transformer模型启发,将每位角色与讽刺表达的关联度作为缩放因子 $\gamma$ ,例如“谢尔顿”与讽刺表达的关联度较大,那么可以设置 $\gamma=1.2$ ，“拉杰什”与讽刺表达关联度较低,可以设置 $\gamma=0.8$ .该缩放因子 $\gamma$ 与讽刺特征 $f_i=(f_1 f_2 \cdots f_\gamma \cdots f_G)$ 结合组成新的讽刺特征 $f_{\text{newt}}=\gamma f_i$ .然后,我们将新讽刺特征输入到全连接层获得讽刺识别结果.为了寻求最佳性能,我们将每位角色的缩放因子 $\gamma$ 随机初始化,并设置为可训练.详细的实验结果如表6所示.

表6 角色信息实验结果

数据集	方法	评估指标	
		Mi-F	Acc
MUStARD	(1)仅利用BERT将每个角色名表征为向量	0.752	0.756
	(2)只采取缩放因子	0.759	0.764
	(3)综合(1)和(2)两种方法	0.757	0.763
	QPIN-QM	0.754	0.753

可以观察到,仅仅利用BERT将每位角色表征为向量的做法并未取得模型性能的提升,与之前的QPIN-QM处于同一性能水平,表明这种融入角色信息的方法过于朴素,需要更加详细深入的角色建模方法.第二种只采用缩放因子的方式以0.66%的微弱优势超过了当前的QPIN-QM,表明了角色引入对模型性能的帮助与必要性.第三种同时利用缩放因子和角色向量的方法并未超越第二种,但是胜过QPIN-QM.这印证了我们之前的解释,即仅仅将角色名表征为BERT向量对模型提升并无实质性帮助.综上,缩放因子的效果仍然有待提升,需要进一步深入研究,这些将留给我们下一步工作.

## 5 结论

对话讽刺识别是一项崭新且具有挑战性的人工智能任务.本文尝试了将量子概率与复数体系引入到经典讽刺识别中.基于此,本文提出了一种量子概率启发的对话讽刺识别网络模型,旨在建模人类讽刺语言中

固有的不确定性问题.本文在两个基准数据集上进行大规模实验,验证了本文提出方法的有效性.本文也进行了一系列模型分析,例如量子概率与经典概率方法对比分析、参数分析等,全方位剖析QPIN模型的优缺点,探索了量子概率在讽刺识别任务的潜力.

#### 参考文献

- [1] ZHANG Y, LI X, RONG L, TIWARI P. Multi-task learning for jointly detecting depression and emotion[C]//2021 IEEE International Conference on Bioinformatics and Biomedicine. Houston, Texas, USA: IEEE, 2021: 3142-3149.
- [2] 张仰森,周炜翔,张禹尧,吴云芳.一种基于情感计算与层次化多头注意力机制的负面新闻识别方法[J].电子学报,2020,48(9):1720-1728.  
ZHANG Yang-sen, ZHOU Wei-xiang, ZHANG Yu-yao, WU Yun-fang. A negative news recognition method based on emotional computing and hierarchical multi-head attention mechanism[J]. Acta Electronica Sinica, 2020, 48(9): 1720-1728. (in Chinese)
- [3] 康世泽,马宏,黄瑞阳.一种基于Opinosis图和马尔科夫随机游走模型的多文本情感摘要框架[J].电子学报,2017,45(12):3005-3011.  
KANG Shi-ze, MA Hong, HUANG Rui-yang. An Opinosis and MRW based sentiment summarization framework [J]. Acta Electronica Sinica, 2017, 45(12): 3005-3011. (in Chinese)
- [4] ZHANG Y, SONG D, LI X, ZHANG P, et al. A quantum-like multimodal network framework for modeling interaction dynamics in multiparty conversational sentiment analysis[J]. Information Fusion, 2020, 62(1): 14-31.
- [5] ZHANG Y, TIWARI P, SONG D, et al. Learning interaction dynamics with an interactive LSTM for conversational sentiment analysis[J]. Neural Networks, 2021, 133: 40-56.
- [6] HUME D. A Treatise of Human Nature[M]. Chicago, USA: Courier Corporation, 2003.
- [7] 谭光辉.情感分析的几条基本规律[J].内蒙古社会科学(汉文版),2018,39(1):8-14.  
TAN Guang-hui. Basic rules of sentiment analysis[J]. Inner Mongolia Social Sciences, 2018, 39(1): 8-14. (in Chinese)
- [8] TVERSKY A, KAHNEMAN D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment[J]. Psychological Review, 1983, 90(4): 293-315.
- [9] WANG B, ZHANG P, LI J, et al. Exploration of quantum interference in document relevance judgement discrepancy [J]. Entropy, 2016, 18(4): 144-150.
- [10] ZHANG Y, SONG D, ZHANG P, et al. A quantum-in-

spired sentiment representation model for twitter sentiment analysis[J]. Applied Intelligence, 2019, 49(8): 3093-3108.

- [11] WANG B, LI Q, MELUCCI M, SONG D. Semantic Hilbert space for text representation learning[C]//Proceedings of the World Wide Web conference. San Francisco, USA: WWW, 2019: 3293-3299.
- [12] ZHANG Y, SONG D, ZHANG P, et al. A quantum-inspired multimodal sentiment analysis framework[J]. Theoretical Computer Science, 2018, 752: 21-40.
- [13] LIU Y, ZHANG Y, LI Q, et al. What does your smile mean? jointly detecting multi-modal sarcasm and sentiment using quantum probability[C]//Findings of the Association for Computational Linguistics: EMNLP. Virtual Conference: ACL, 2021: 871-880.
- [14] WANG P, HOU Y, LI Z, ZHANG Y. QIRM: A quantum interactive retrieval model for session search[J]. Neurocomputing, 2021, 451:57-66.
- [15] OCHS W. Gleason Measures and Quantum Comparative Probability[M]. Quantum Probability and Applications II. Heidelberg, Germany: Springer, 1985: 388-396.
- [16] CASTRO S, HAZARIKA D, PEREZ-ROSAS V, et al. Towards multimodal sarcasm detection (an obviously perfect paper)[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: ACL, 2020: 4619-4629.
- [17] GHOSH D, VAJPAYEE A, MURESAN S. A report on the 2020 sarcasm detection shared task[C]//Proceedings of the Second Workshop on Figurative Language Processing. Saarbrücken, Germany: ACL, 2020: 1-11.
- [18] KIM Y. Convolutional neural networks for sentence classification[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: ACL, 2014: 1746-1755.
- [19] KUMAR A, NARAPAREDDY VT, SRIKANTH VA, et al. Sarcasm detection using multi-head attention based bi-directional LSTM[J]. IEEE Access, 2020, 8(1): 6388-6397.

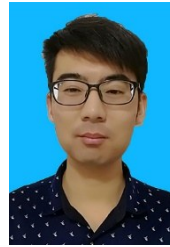
#### 作者简介



张亚洲 男,1990年11月出生于河南省安阳市. 现为郑州轻工业大学软件学院讲师、硕士生导师. 主要研究方向为多模态情感分析.  
E-mail: yzzhang@zzuli.edu.cn



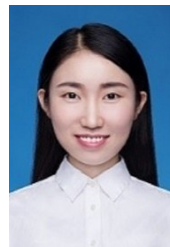
俞洋 男,1998年8月出生于河南省信阳市. 现为郑州轻工业大学软件学院硕士生. 主要研究方向为多模态情感分析.  
E-mail: yuyang19980818@outlook.com



朱少林 男,1989年11月出生于河南省新乡市. 现为郑州轻工业大学软件学院讲师. 主要研究方向为多模态机器翻译.  
E-mail: zhushaolin003@163.com



陈锐 男,1979年12月出生,现为郑州轻工业大学软件学院讲师. 主要研究方向为社交网络分析、个性化推荐.  
E-mail: ruichen@zzuli.edu.cn



戎璐 女,1995年1月出生于山西省忻州市. 现为郑州轻工业大学助理讲师. 主要研究方向为自然语言理解.  
E-mail: lurong2013@outlook.com



梁辉(通讯作者) 男,1977年1月生于河南省平顶山市. 现为郑州轻工业大学软件学院教授. 主要研究方向为数字媒体、虚拟现实、计算机动画技术.  
E-mail: huiliang@zzuli.edu.cn