doi: 10.7690/bgzdh.2025.02.003

基于深度学习的电力客服技术研究

唐国亮，徐尤峰
（中国南方电网有限责任公司市场营销部，广州 510700）

摘要：为提高电力客服服务质量，提出一种电力智能客服问答系统。基于卷积神经网络（convolutional neural networks，CNN）和双向长短期记忆（bidirectional long short-term memory，BiLSTM）网络提取和表示重要信息和上下文信息；结合 BiLSTM 网络和协同注意机制，提取语义信息并进一步表示为特征向量，解决长语句中前后词之间的依赖问题，获得问题对之间的相关特征表示；提出一种将余弦相似性和欧氏距离进行调和的相似性计算函数，实现问题对的高效匹配；以某电力公司提供的电力数据为例，对所提模型进行实验验证。结果表明：所提模型性能最优，准确率和召回率分别为 90.96%和 88.63%，为电力客服智能服务的发展提供了一定借鉴作用。

关键词：电力系统；智能客服；深度学习；注意机制；相似性函数
中图分类号：TP393　　**文献标志码：**A

Research on Power Customer Service Technology Based on Deep Learning

Tang Guoliang, Xu Youfeng
(*Marketing Department*, *China Southern Power Grid Company Limited*, *Guangzhou* 510700, *China*)

Abstract: In order to improve the service quality of electric power customer service, an intelligent question answering system for electric power customer service is proposed. Extracting and representing important information and context information based on convolutional neural networks (CNN) and bidirectional long short-term memory (BiLSTM), extracting semantic information and further representing the semantic information as a feature vector by combining a BiLSTM network and a collaborative attention mechanism, solving the dependence problem between front and rear words in a long sentence, and obtaining the related feature representation between question pairs; A similarity calculation function is proposed to reconcile the cosine similarity and Euclidean distance to achieve efficient matching of the problem pairs. Taking the power data provided by a power company as an example, the proposed model is verified. The results show that the performance of the proposed model is the best, the precision and recall are 90.96% and 88.63%, respectively, which provides a reference for the development of electric power customer service intelligent service.

Keywords: power system; intelligent customer service; deep learning; attention mechanism; similarity function

0　引言

随着网络、大数据、物联网、通信技术[1-2]的不断发展，电力服务行业迎来了改革与创新。电力公司在不同场景下的相关业务逐渐增多，部分业务已从线下服务转向各种在线远程服务[3-4]。然而，无论是线下服务还是远程服务，都会存在电力业务资源受限问题，如客户服务人员短缺、业务能力差异，这将直接影响电力服务质量水平；因此，设计一种高质量的电力企业智能客户服务系统[5-6]已迫在眉睫。

传统智能答疑系统大多使用机器学习[7-9]来分析和检索文本，并应用知识库来回答常见问题，因此需要大量的注释数据，无法有效地处理语义结构复杂的问题；同时，机器学习对问题和答案的语义信息相关性表现不佳。随着人工智能的重新发展，深度学习已成为智能问答的主要研究方法。文献[10]提出了基于双向长短期记忆网络（BiLSTM）网络的智能客服语音识别系统，从而提高智能客服服务质量。文献[11]针对人工智能客服系统中的海量数据大数据分析需求，提出了一种基于 Hadoop 的智能客服系统框架。深度学习网络可以更好地基于输入问题构造答案的向量表示，但这类模型忽略了答案对问题向量表示的影响，导致结果存在一定偏差。另一方面，目前开放领域智能问答在深度学习研究领域取得了显著进展。然而，受限领域中的问题很难处理，例如针对电力智能客服，应设计特定语料库数据，从而保持电力专业领域具有较强的逻辑性和真实性。

为改善上述问题，笔者结合卷积神经网络（CNN）、BiLSTM 网络、协同注意机制等，提出了一个电力智能客服问答系统。首先，基于 CNN 和 BiLSTM 提取和表示重要的文本信息和上下文信息。

收稿日期：2024-07-07；**修回日期：**2024-08-03
第一作者：唐国亮（1969—），男，广西人。

其次，基于协同注意机制求解问题对的语义交互和特征表示，为后续计算提供完整信息。最后，调和欧氏距离以及余弦相似性，生成一个相似性计算函数，使其能够同时衡量以上 2 大要素，从而实现问题对的高效匹配。

1　电力智能客服问答系统

笔者构建了基于人工智能技术的电力企业场景智能客户服务问答系统，从而提高电力企业的客户服务质量。如图 1 所示。

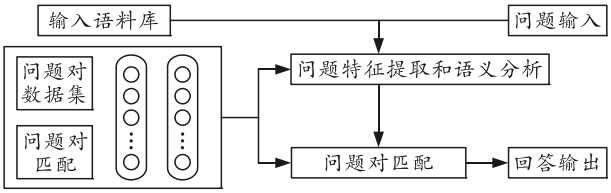

图 1　电力智能客户服务问答系统结构

从图 1 可以看出，系统可实现收集客户提出的各种问题，并对同一目标的不同表达进行编码和特征提取。首先，为实现有效的特征提取，模型需要一个底层问答语料库来训练问题对的语义表示和匹配模型；其次，需要分析客户提出问题的语义和结构，并构建特征向量模型；接着，使用匹配算法计算输入问题向量和现有问题向量之间的相似度，并找到一系列相关问题，同时得到匹配度最高的答案；最后，通过智能客服语言处理反馈给客户，实现与客户的互动问答。

2　问题特征提取和语义分析

2.1　问题特征提取

考虑到汉语句子的表达更加复杂多样，提出了一种有效理解中文语义的网络结构。客户问题语义特征提取模型如图 2 所示。网络中 CNN 用于捕获关键特征信息，而 BiLSTM 网络用于获取整个句子的语义分析。图 2 中网络执行过程如下：首先，将分词后的句子发送到自建语料库训练的 word2vec 模型中进行编码，利用 CNN，该模型可以充分考虑输入句子的词序和语义上下文关系；然后，提取语句的局部关键信息并压缩为固定长度，为后续网络中的语义交互和问题对表示提供基础数据；接着，结合 BiLSTM 网络和协同注意机制，提取语义信息并进一步表示为特征向量；该特征提取方法不仅可以解决长语句中前后词之间的依赖问题，还可以获得问题对之间的相关特征表示；最后，通过 softmax 函数预测输出。

2.2　语义分析

汉语表达中词的相关性非常密切，句子语义分析和特征提取包括每个词的序列编码和词间权重分配。因此，笔者在问题特征提取的框架下，利用 CNN 和 BiLSTM 的融合结构来计算词语的相关性，并映射句子中浅层特征的局部深度。

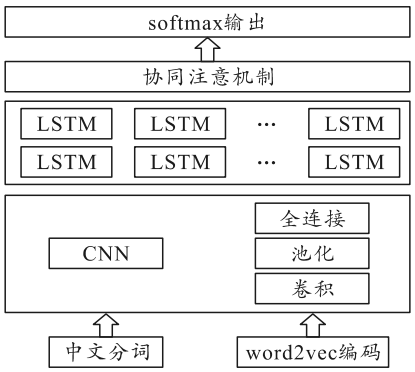

图 2　客户问题语义特征提取模型

在 CNN 网络层，初始输入主要通过卷积和池化 2 个过程获得文本的初步解析向量。CNN 网络层内部结构如图 3 所示。

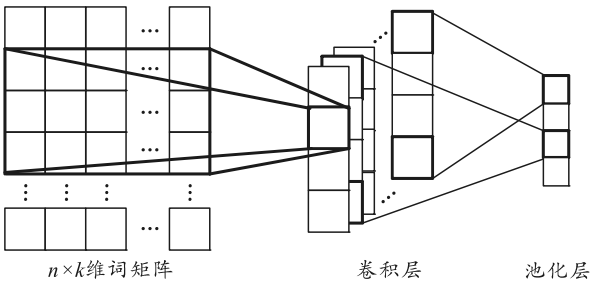

图 3　CNN 网络层内部结构

输入到网络中的是提前训练的 word2vec 模型产生的 $n\times k$ 维词矩阵，n 为句子中全部词汇的数量，也就是句子本身的最大长度，k 为对应于各个词汇的向量维数(令 k=300)。使 $\boldsymbol{x}_i \in \mathbb{R}^k$ 代表句子中第 i 个词的 k 维词向量，则长度为 n 的句子可由密集向量表示如下：

$$\boldsymbol{x}_{1:n} = \boldsymbol{x}_1 \oplus \boldsymbol{x}_2 \oplus, \cdots, \oplus \boldsymbol{x}_n 。 \tag{1}$$

式中：$\oplus$ 为词向量并行串联运算符；$\boldsymbol{x}_{a:b}$ 为由 $\{\boldsymbol{x}_a, \cdots, \boldsymbol{x}_b\}$ 组成的词向量矩阵。

在卷积神经网络处理自然语言的过程中，语句上下几行的字通常会被卷积核覆盖，所以文本自己的卷积核不存在宽度，只有长度。笔者设置卷积核窗口 $W \in \mathbb{R}^{h\times k}$ 时，存在着固定长度 h($h\times k$、横向 k 以及纵向 h 依次代表着卷积窗口大小、词向量自己的维数以及滑动窗口中的词数)，向 $\boldsymbol{x}_{i:i+1}$ 中涵盖的词实施卷积。句子中的第 i 个词上面，可以如下计

算卷积核窗口自身的特征操作。

$$c_i=\sigma\left(\boldsymbol{W}\bullet\boldsymbol{x}_{i:i+h-1}+b\right)。\tag{2}$$

式中：c_i，σ，$\boldsymbol{x}_{i:i+h-1}$ 和 b 分别为词 i 在卷积后的特征输出、激活函数、由 i 到 $i+h-1$ 个词组成的词向量矩阵和偏差系数；$\boldsymbol{W}$ 为卷积层的运算矩阵，即滑动窗口。以此类推，当纵向长度为 h 的卷积核窗口作用于长度为 n 的整个句子时，句子可获得 $n-h+1$ 个新的特征向量，且这些特征向量形成相应语句的 1 维特征映射。该过程具体计算为：

$$\boldsymbol{C}=[c_1, c_2, \cdots, c_{n-h+1}]。\tag{3}$$

接下来，池化层是卷积层本身的下采样过程，它能够降低应该处置的数据大小，同时有效地保留和提取语句蕴含的重要语义特征笔者借助 1 维最大池化层提取卷积后的相关卷积核向量本身的最大特征值，该最大特征值被当作 BiLSTM 网络的相应输入。可以如下计算池化层：

$$\boldsymbol{C}=\max[c_1, c_2, \cdots, c_{n-h+1}]。\tag{4}$$

应指出的是：在卷积核向量应用 1 维最大池化层的过程中，各特征映射输出的最大值只有一个。该方法可以降 DI 采样卷积层输出的相关特征图的全部维数。完成了相关合并后，最终能够获得一个向量矩阵，即组合输出向量 $\boldsymbol{C}$，其固定大小，其维数是从卷积层输出的全部特征映射数。而且，借助池化以及卷积操作，语句本身的向量矩阵能够实现特征提取和初始语义分析，从而获取语句的局部重要信息特征，并有效减少训练参数。

同理，令通过 CNN 网络提取问题对 $Q=(q_1, q_2, \cdots, q_n)$和 $Q'=(q'_1, q'_2, \cdots, q'_m)$ 句子中具有表征能力的结构化语义信息输出特征为 $\boldsymbol{C}_Q$ 和 $\boldsymbol{C}'_Q$，则有：

$$\boldsymbol{C}_Q=\max(c_1, c_2, \cdots, c_{n-h+1});\tag{5}$$

$$\boldsymbol{C}'_Q=\max(c_1, c_2, \cdots, c_{n-h+1})。\tag{6}$$

式中：m 与 n 为问题 Q' 以及问题 Q 中涵盖的词的数量，也就是句子长度；h 为卷积窗口大小。接下来，将 $\boldsymbol{C}_Q$ 和 $\boldsymbol{C}'_Q$ 传输到 BiLSTM 网络，从而提取句子中距离词的相互依赖性和影响。

考虑到单一的 LSTM 单元捕获的只是各词所在句子中前半部分句子的语义信息，缺乏捕获后面半部分蕴含的语义信息。为克服该缺点，笔者使用双层 LSTM 语义捕获模型，其包含 2 个相反方向的 LSTM 隐藏层：前向和后向 LSTM 单元。令句子序列为 $x=(x_1, x_2, \cdots, x_n)$，自始至终同时遍历整个语句，前向、后向神经元隐藏层的输出序列 $\vec{h}_t$、输出序列 $\overleftarrow{h}_t$ 能够描述如下：

$$\vec{h}_t=(\vec{h}_1, \vec{h}_2, \cdots, \vec{h}_n);\tag{7}$$

$$\overleftarrow{h}_t=(\overleftarrow{h}_1, \overleftarrow{h}_2, \cdots, \overleftarrow{h}_n)。\tag{8}$$

式中：n 为句子序列中所涵盖的词的数量。接下来，借助级联反向与正向输出，求出双向 LSTM 网络隐藏层的编码输出 y_t：

$$y_t=[\vec{h}_t, \overleftarrow{h}_t];\tag{9}$$

$$\vec{h}_t=\sigma(\boldsymbol{W}_{\vec{h}_t}x_t+\boldsymbol{W}_{\vec{h}_{t-1}}\vec{h}_{t-1}+b_{\vec{h}});\tag{10}$$

$$\overleftarrow{h}_t=\sigma(\boldsymbol{W}_{\overleftarrow{h}_t}x_t+\boldsymbol{W}_{\overleftarrow{h}_{t-1}}\overleftarrow{h}_{t-1}+b_{\overleftarrow{h}});\tag{11}$$

$$y_t=\boldsymbol{W}_{\vec{h}_y}\vec{h}_t+\boldsymbol{W}_{\overleftarrow{h}_y}\overleftarrow{h}_t+b_y。\tag{12}$$

式中：$\boldsymbol{W}$ 和 b 分别为与 LSTM 的 3 个门相对应的权重向量和偏置。

进一步，笔者在 LSTM 单元的基础上构建了一个双层 BiLSTM 网络，从而充分实现对单个语句的语义解析，有效提高模型的分类和回归能力。笔者将上层 BiLSTM 网络的输出 y_t 界定成了下一级 BiLSTM 网络本身的输入。双层 BiLSTM 网络本身的结构如图 4 所示。

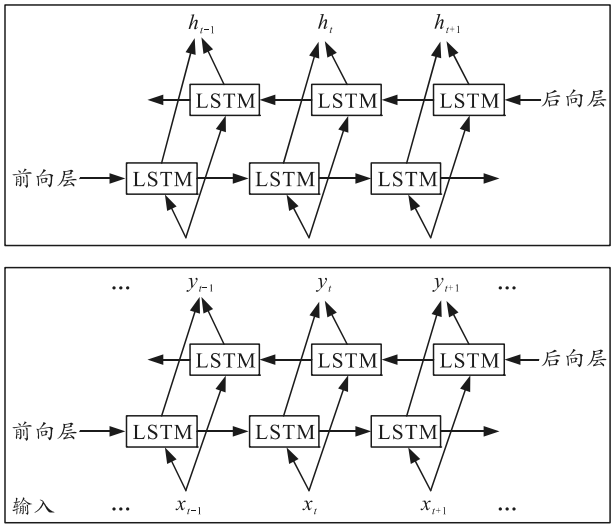

图 4 双层 BiLSTM 网络结构

以下是这个句子中各词的状态输出：

$$h_t=\boldsymbol{W}_{\vec{h}_y}\vec{h}_t+\boldsymbol{W}_{\overleftarrow{h}_y}\overleftarrow{h}_t+b_h。\tag{13}$$

同理，定义 $\boldsymbol{p}_t$ 和 $\boldsymbol{p}'_t$ 分别为问题对序列 $\boldsymbol{C}_Q$ 和 $\boldsymbol{C}'_Q$ 的第 t 个向量。CNN 网络之后的问题对序列将分别发送到双层 BiLSTM 网络，从而获得其状态矩阵 $\boldsymbol{H}_Q$ 和 $\boldsymbol{H}_{Q'}$，则有：

$$h_0^q=0, h_t^q=D_B(h_{t-1}^q, h_{t+1}^q, p_t);\tag{14}$$

$$h_0^{q'} = h_n^q, h_t^{q'} = D_B(h_{t-1}^{q'}, h_{t+1}^{q'}, p_t'); \quad (15)$$

$$\boldsymbol{H}_Q = \left[h_1^q, h_2^q, \cdots, h_n^q\right] \in R^{d\times n}; \quad (16)$$

$$\boldsymbol{H}_{Q'} = \left[h_1^{q'}, h_2^{q'}, \cdots, h_m^{q'}\right] \in R^{d\times m}。 \quad (17)$$

式中：d 为隐藏层的输出状态矩阵的维数，且 $d \in R$。

3　注意机制和问题对匹配

3.1　注意机制

笔者在协同注意机制中设计了一个关联矩阵来捕捉向量之间的相关性和相互作用，并使用 softmax 激活函数，把一系列神经元的相关输出映射到与之相对的区间内。借助双层 BiLSTM 网络的帮助，将问题对本身的状态矩阵 $\boldsymbol{H}_Q$ 和 $\boldsymbol{H}_{Q'}$ 当作输入，图 5 展示的是协同注意体制自身的内部结构。

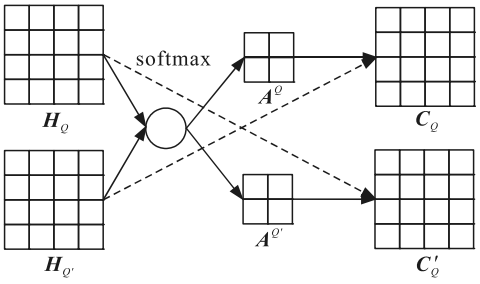

图 5　协同注意机制的内部结构

由图 5 可以看出：将状态矩阵 $\boldsymbol{H}_Q$ 和 $\boldsymbol{H}_{Q'}$ 进行矩阵乘法，从而计算关联矩阵 $\boldsymbol{L}$。关联矩阵中的每一项都是问题对语句中不同词之间的相关分数，即问题对语句之间的交互作用反映如下：

$$\boldsymbol{L} = \boldsymbol{H}_{Q'}^{\mathrm{T}} \boldsymbol{H}_Q \in R^{m\times n}。 \quad (18)$$

网络中 softmax 函数用来归一化向量元素；因此，问题对的隐藏层状态的注意权重为：

$$\boldsymbol{A}^Q = S(\boldsymbol{L}) \in R^{m\times n};\ \boldsymbol{A}^{Q'} = S(\boldsymbol{L}^{\mathrm{T}}) \in R^{n\times m}。 \quad (19)$$

式中：$S(\bullet)$为 softmax 函数；同时，交互后，问题对的相应特征输出为：

$$\boldsymbol{C}^Q = \boldsymbol{H}_{Q'} \boldsymbol{A}^Q \in R^{d\times n};\ \boldsymbol{C}^{Q'} = \boldsymbol{H}_Q \boldsymbol{A}^{Q'} \in R^{d\times m}。 \quad (20)$$

在双层 BiLSTM 网络和协同注意机制的数据传输以及计算中，可能问题对语句中会出现部分信息不匹配的现象。为消除这种现象，笔者立足于协同注意机制，向其增加了注意机制层，从而连接和整合前几步的语句向量信息；因此，模型输入为问题对语句的新特征向量表示 $\boldsymbol{C}^Q$ 和 $\boldsymbol{C}^{Q'}$，其中 $\boldsymbol{C}_t^Q$ 为输入问题语句的第 t 个注意特征向量。而且，笔者采取最大池化层，把输入转换成了长度固定的向量 $\boldsymbol{O}_q$，其为输入问题语句的最终特征向量输出。当使用注意机制表示现有问题陈述的最终向量时，可以根据 $\boldsymbol{O}_q$ 计算并获得现有问题陈述中所有特征向量 $\left\{\boldsymbol{C}_1^{Q'}, \boldsymbol{C}_2^{Q'}, \cdots, \boldsymbol{C}_m^{Q'}\right\}$ 的 softmax 权重。在分配权重后，基于注意机制模型可以得到现有问题陈述的最终特征向量输出 $\boldsymbol{O}_{q'}$。注意机制可以整合语句信息，利用语句信息可以获得问题对语义解析的最终输出，具体如下所示：

$$\boldsymbol{O}_q = \max_{0<t\leqslant n} \boldsymbol{C}_t^Q; \quad (21)$$

$$\boldsymbol{O}_{q'} = \sum_{t=1}^{m} \boldsymbol{C}_t^{Q'} S_{q'q}(t)。 \quad (22)$$

式中：$S_{q'q}$ 为由 softmax 函数标准化的注意权重，其与 $\boldsymbol{C}_t^{Q'}$ 成正比。$S_{q'q}$ 的值越高，$\boldsymbol{C}_t^{Q'}$ 与输入问题之间的相关性就越高，也就是说，现有问题陈述对输入问题的特征向量表示有更大的影响。

3.2　问题对匹配

在获得问题对的最终特征向量表示后，通过测量特征向量之间的相似性来确定输出，其中余弦相似性和欧氏距离是最常用的 2 种相似性计算方法。将 2 个向量间存在的空间角度当作度量余弦相似性的标准，欧氏距离代表的是求出 2 个向量间存在着的绝对空间距离。

通常情况下，期待问题对本身的特征向量间的距离最短，有足够小的角度，从而可最大限度地获取和计算问题对表达向量的相似性。笔者调和了欧氏距离以及余弦相似性，使其生成一个相似性计算函数，使其能够同时分析以上 2 大因素。余弦相似性度量 S_c 计算如下：

$$S_c\left(\boldsymbol{O}_q, \boldsymbol{O}_{q'}\right) = \boldsymbol{O}_q \bullet \boldsymbol{O}_{q'} \Big/ \left(\left|\boldsymbol{O}_q\right| \left|\boldsymbol{O}_{q'}\right|\right)。 \quad (23)$$

欧氏距离相似性度量 S_e 计算如下：

$$S_e(\boldsymbol{O}_q, \boldsymbol{O}_{q'}) = 1 \Big/ (1 + \left\|\boldsymbol{O}_q - \boldsymbol{O}_{q'}\right\|_2)。 \quad (24)$$

式中$\|\bullet\|_2$ 为欧氏距离。进一步，将余弦相似性度量化归一化为[0, 1]：

$$S_c(\boldsymbol{O}_q, \boldsymbol{O}_{q'}) = 0.5 S_c(\boldsymbol{O}_q, \boldsymbol{O}_{q'}) + 0.5。 \quad (25)$$

最终向量相似性的计算函数 S_t 如下：

$$S_t = \frac{2 \bullet s_c(\boldsymbol{O}_q, \boldsymbol{O}_{q'}) \bullet s_e(\boldsymbol{O}_q, \boldsymbol{O}_{q'})}{s_c(\boldsymbol{O}_q, \boldsymbol{O}_{q'}) + s_e(\boldsymbol{O}_q, \boldsymbol{O}_{q'})}。 \quad (26)$$

式中：$\bullet$为点乘运算；$|\boldsymbol{O}_q|$和$\left|\boldsymbol{O}_{q'}\right|$分别为相应特征向量的模长；$\left\|\boldsymbol{O}_q - \boldsymbol{O}_{q'}\right\|_2$ 为 2 个向量之间存在的欧式距离。

4 仿真与分析

4.1 数据集与实验环境

实验用到的数据集是由某电力企业 2018 至 2019 年的电力客服文本语料库。数据中大部分疑问句都是与客户相关的电力咨询信息，如电费查询、电费咨询、电表查询、电量异常查询等。实验过程是对 2 个输入问题陈述的语义与结构进行理解，而且对上述句子的相似度和相关性进行匹配，以明确其是否存在同样的目的、意图和倾向。数据集中，将具有相同意图的问题对设置为正例，且标记为 1；而不具有相同意图的问题对设置为负例，且标记为 0。最终，电力客服文本语料库共包含约 1 万个标记问题对，将其当作训练集，将其中的 0.1 万个未标记问题对作为验证集，将其中的 1 000 万个未标记问题对当作测试集。

仿真软件环境是 pycharm 创造的算法框架，由 python 以 Keras 以及 tensorflow 为基础建立学习算法。而且，算法运行硬件环境采取了 Intel Core i9−9280X CPU，内存达到了 32 G，操作系统和显卡分别是 Ubuntu 18.04 64 位以及 NVIDIA RTX2080Ti 11G 2 块。

4.2 实验设置

为做完相关的模型训练，可设置如下实验：整理电力客服文本语料库中的训练集数据，完成 word2vec 模型的预训练。从这次预训练来看，设置输出词向量达到 300 维数，设置各问题语句本身的输出长度为 35。而且，CNN 中能够设定不同的 2 个卷积窗口，使 h 依次为 3 与 2，步长为 1。特征映射得以生成，通过最大池化处理之后，获得了针对性的矩阵输出，而且最后变成了双层 BiLSTM 网络本身的特征向量输入。

设置模型的初始学习率以及正则化参数分别为 10^{-3} 和 10^{-5}，设置数据批量处理大小以及最大迭代次数分别为 50 和 25，设置学习率衰减倍数为 0.2。

此外，从训练过程来看，卷积层中加入了 dropout 层，其概率达到了 0.2。它有利于模型规避过度拟合的现象。训练过程中采取 Adam 算法，对模型进行优化，以更新和完善相关的网络参数。按照评估标准，数据集中的任务只有二分类任务，数量为 1，其标签集是{1，0}。在这个标签集中，1 代表着给定的 2 个问题对存在着相似或者相同的意图；0 代表着给定的 2 个问题陈述间不存在相似或者相同的意图。假如 2 个问题的实际陈述间的相似性大于 0.65，应出示标签 1，不然应出示标签 0。

4.3 结果与分析

为对所提模型固有的性能进行测试，基于上述 5 个模型开展验证对比：BiLSTM、叠加 BiLSTM（SBiLSTM）、叠加 BiLSTM+调和余弦相似性和欧氏距离（SBiLSTMCE）、叠加 BiLSTM+协同注意机制+调和余弦相似性和欧氏距离（SBILSMcoA）、CNN+叠加 BiLSTM+共注意机制+调和余弦相似性和欧几里德距离（所提模型）。对比指标分别选取准确率、召回率和 F 分数。其中 F 分数为综合考虑准确率和召回率的调和值。表 1 为各种模型综合对比的结果。

表 1 不同模型综合对比结果

模型名称	准确率	召回率	F 分数
BiLSTM	0.758 7	0.743 2	0.750 9
SBiLSTM	0.767 3	0.774 8	0.771 0
SBiLSTMCE	0.830 5	0.815 4	0.822 9
SBILSMcoA	0.851 9	0.817 0	0.834 1
所提模型	0.909 6	0.886 3	0.897 8

上表中可以看出，SBiLSTM 网络模型的实验结果优于单层 BiLSTM。分析原因，更深层次的 LSTM 网络有助于模型理解映射语句中词与词之间存在的某种关系。此外，相比于单纯的叠加 BiLSTM，能够在一定程度上提升协同注意机制模型的准确率和召回率。究其原因，协同注意机制能够将句子中蕴含的关键信息提取出来。所提模型（CNN+叠加 BiLSTM+共注意机制+调和余弦相似性和欧几里德距离）取得了最优性能，召回率以及准确率依次达到 88.63%以及 90.96%。实验结果证实了所提模型本身的有效性以及可行性，而且表明调和余弦相似度和欧氏距离可以平衡向量之间的角度和距离关系，从而有助于句子向量之间的匹配。

5 结论

笔者对电力行业智能客服进行了分析，建立了一种电力智能客服问答系统，可实现收集客户提出的各种问题，并对同一目标的不同表达进行编码和特征提取。首先，分析客户提出问题的语义和结构，并构建特征向量模型；其次，使用匹配算法计算输入问题向量和现有问题向量之间的相似度，并找到一系列相关问题，同时得到匹配度最高的答案；最后，通过智能客服语言处理反馈给客户，实现与客户的互动问答。该模型为电力客服智能服务的发展提供了一定借鉴作用。

（下转第 26 页）