(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2020 年 11 月 12 日 (12.11.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/224099 A1

(51) 国际专利分类号:
G06F 16/332 (2019.01) *G06N 3/04* (2006.01)
G06F 16/35 (2019.01) *G06N 3/08* (2006.01)
G06F 16/95 (2019.01)

(21) 国际申请号: PCT/CN2019/102194

(22) 国际申请日: 2019 年 8 月 23 日 (23.08.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910386282.6 2019年5月9日 (09.05.2019) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 侯丽 (HOU, Li); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所（普通合伙） (SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦四路10号中浩大厦1528-1530室于志光, Guangdong 518000 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** INTELLIGENT EMOTIONAL QUESTION ANSWERING METHOD AND DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM

(54) **发明名称**: 智能化情感问答方法、装置及计算机可读存储介质

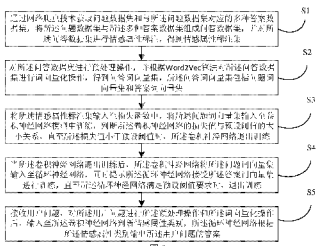


图 1

S1 Acquire, by means of web crawler technology, a question data set and multiple answer data sets corresponding to the question data set, constitute a question and answer data set using the question data set and the multiple answer data sets, and carry out emotional attribute annotation on the question and answer data set to obtain an emotional attribute annotation set

S2 Carry out a pre-processing operation on the question and answer data set, and carry out a word vectorization operation on the question and answer data set according to a Word2Vec algorithm to obtain a question and answer word vector set, wherein the question and answer word vector set comprises a question word vector set and an answer word vector set

S3 Input the emotional attribute annotation set into a loss function, input the question word vector set into a convolutional neural network model for training, determine the size relationship between the loss value of the convolutional neural network and a pre-set threshold, and the convolutional neural network exits training when the loss value is less than the pre-set threshold

S4 After the convolutional neural network exits training, the convolutional neural network inputs the question word vector set into a recurrent neural network and prompts the recurrent neural network to accept the answer word vector set for training, and exits training until the recurrent neural network meets pre-set threshold requirements

S5 Receive a user question, and carry out a pre-processing operation and the word vectorization operation on the user question and then input same into the convolutional neural network to determine an emotional attribute category, and the recurrent neural network outputs the answer to the user question according to the emotional attribute category

(57) Abstract: An intelligent emotional question answering method and device, and a computer-readable storage medium. The intelligent emotional question answering method comprises: receiving a question and answer data set, and carrying out emotional attribute annotation on the question and answer data set to obtain an emotional attribute annotation set (S1); carrying out pre-processing and word vectorization on the question and answer data set to obtain a question and answer word vector set (S2); first inputting the emotional attribute annotation set and the question and answer word vector set into a convolutional neural network model for training (S3); the convolutional neural network exiting training and inputting the question and answer word vector set into a recurrent neural network, and exiting training when the recurrent neural network meets pre-set threshold requirements (S4); and receiving a user question and inputting same into the convolutional neural network to determine an emotional attribute, and outputting an answer to the user question on the basis of the recurrent neural network (S5). Accordingly, an accurate intelligent emotional question answering function can be realized.

(57) 摘要: 一种智能化情感问答方法、装置及计算机可读存储介质，其中智能化情感问答方法包括：接收问答数据集，并对所述问答数据集进行情感属性标注得到情感属性标注集（S1）；对所述问答数据集进行预处理和词向量化得到问答词向量集（S2）；将所述情感属性标注集和所述问答词向量集先输入至卷积神经网络模型中训练（S3）；所述卷积神经网络退出训练并将所述问答词向量集输入至循环神经网络，直至所述循环神经网络满足预设阈值要求时退出训练（S4）；接收用户问题并输入至所述卷积神经网络判断情感属性，基于所述循环神经网络输出所述用户问题的答案（S5）。从而可以实现精准的智能化的情感问答功能。

WO 2020/224099 A1

[见续页]

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) **指定国**(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

智能化情感问答方法、装置及计算机可读存储介质

本申请基于巴黎公约申明享有2019年5月9日递交的申请号为CN201910386282.6、名称为“智能化情感问答方法、装置及计算机可读存储介质”的中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

技术领域

本申请涉及人工智能技术领域，尤其涉及一种接收用户问题后智能化的给予所述问题答案的智能化情感问答方法、装置及计算机可读存储介质。

背景技术

目前大部分问答系统的研究重点都关注在生成语句的语法和语义是否合理上，且所述问答系统的答案生成方式大都基于上下文或结合主题，很少考虑到对话者的情感，如用户输入：昨天的考试我没有及格，多数问答系统给予的回复一般如：竟然不是倒数，好神奇；用户输入：我养的狗今天去世了。回复：宠物狗特别容易死。但在现实生活中，如果对方通过语言表达出开心的情绪，通常情况下，对方的回复也应该是积极的情绪。如果对方通过语言表达出悲伤的情绪，往往另一方回应的内容也应是安慰等。因此，问答系统带有情感的回复往往更能受到用户的欢迎。

发明内容

本申请提供一种智能化情感问答方法、装置及计算机可读存储介质，其主要目的在于当用户输入问题时，给用户呈现出具有情感倾向的答案结果。

为实现上述目的，本申请提供的一种智能化情感问答方法，包括：

通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集；

对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据

Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集；

将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练；

当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练；

接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

此外，为实现上述目的，本申请还提供一种智能化的情感问答装置，该装置包括存储器和处理器，所述存储器中存储有可在所述处理器上运行的智能化的情感问答程序，所述智能化的情感问答程序被所述处理器执行时实现如下步骤：

通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集；

对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集；

将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练；

当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练；

接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

此外，为实现上述目的，本申请还提供一种计算机可读存储介质，所述计算机可读存储介质上存储有智能化的情感问答程序，所述智能化的情感问答程序可被一个或者多个处理器执行，以实现如上所述的智能化情感问答方法的步骤。

由于卷积神经网络的多层网络结构能自动提取数据的深层特征，可学习不同层次的特征，从而大大提高对文本处理的准确率，同时循环神经网络会联系数据的前后时序状态进行高效判断，因此本申请提出的智能化情感问答方法、装置及计算机可读存储介质，可以实现精准的智能化的情感问答功能。

附图说明

图1为本申请一实施例提供的智能化情感问答方法的流程示意图；

图2为本申请一实施例提供的智能化的情感问答装置的内部结构示意图；

图3为本申请一实施例提供的智能化的情感问答装置中智能化的情感问答程序的模块示意图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请提供一种智能化情感问答方法。参照图 1 所示，为本申请一实施例提供的智能化情感问答方法的流程示意图。该方法可以由一个装置执行，

该装置可以由软件和/或硬件实现。

在本实施例中，智能化情感问答方法包括：

S1、通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集。

本申请较佳实施例，根据所述网络爬虫(Web crawler)技术从 URL 页面内爬取以文本形式提问的问题，将所述以文本形式提问的问题组成问题数据集。遍历所述问题数据集内的问题，使用所述网络爬虫技术从 URL 页面内爬取与所述问题对应的多种不同的答案，直至所述问题数据集遍历结束，得到与所述问题数据集对应的多种答案数据集；

本申请较佳实施例中，根据所述问题数据集，将所述多种答案数据集中的答案按照情感倾向进行情感属性标注，得到所述情感属性标注集，所述情感倾向包括幽默、诚恳、建议等。

S2、所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据 Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集。

本申请较佳实施例中，所述分词根据所述所述问答数据集建立分词概率模型P(S)，并最大化所述分词概率模型，完成分词操作，所述分词概率模型P(S)为：

$$\mathrm{P(S)} = \mathrm{P}(W_1, W_2, \ldots, W_m) = \prod_{i=1}^{m} p(W_i \mid W_{i-1})$$

其中，$W_1, W_2, \ldots, W_m$为所述问答数据集内数据的词，m为所述问答数据集的数量；

所述关键字抽取包括构建所述词的相关度并基于所述相关度提取关键字，所述相关度为：

$$f(W_i, W_j) = \frac{tfidf(W_i) * tfidf(W_j)}{d^2}$$

其中，$f(W_i, W_j)$为词W_i和词W_j的相关度，$tfidf(W_i)$为词W_i的词频与逆向频率值，$tfidf(W_j)$为词W_j的词频与逆向频率值，d为词W_i和词W_j关于词向量的欧式距离；

本申请较佳实施例所述 Word2Vec 算法为 CBOW 模型，所述 CBOW 模型

包括输入层、投影层和输出层，所述投影层$\zeta(\omega,j)$为：

$$\zeta(\omega,j)=\left(1-d_j^{\omega}\right)log\left[\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]+d_j^{\omega}\,log\left[1-\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]$$

其中，$d_j^{\omega}\epsilon\{d_2^{\omega},d_3^{\omega},\dots,d_l^{\omega}\}$，表示在路径$\omega$内，第$j$个结点对应的霍夫曼编码，$\theta$为所述 CBOW 模型的迭代因子，σ表示 sigmoid 函数，X_{ω}为所述预处理操作完成的问答数据集。

S3、将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练。

本申请较佳实施例中，所述卷积神经网络包括卷积层、池化层、Flatten 层、Dropout 层和全连接层，且所述问题词向量集在时间维度上是一维向量的形式，所以所述卷积层与所述池化层的过滤器(filter)也是一维向量，同时为了防止过拟合现象，在所述卷积层与所述池化层中加入所述 Dropout 层；当多次卷积与池化操作后再扁平化数据，即所述 Flatten 层的作用，最后通过所述全连接层(Dense)输出所述训练值.

本申请较佳实施例所述损失值E为：

$$E=\sum_{j=1}^{m}\sum_{x\in m}|x-\mu_j|^2$$

其中，x为所述训练值，μ_j为所述情感属性标注集，m为所述问答数据集的数量，所述预设阈值一般设定为 0.01。

S4、当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练。

本申请较佳实施例，所述循环神经网络为长短期记忆网络，所述长短期记忆网络包括遗忘门、输入门、输出门，所述遗忘门为：

$$f_t=\sigma(w_t[h_{t-1},x_t]+b_t)$$

其中，f_t为所述遗忘门的输出数据，x_t为所述遗忘门的输入数据，t为所述问答词向量集的当前时间，$t-1$为所述问答词向量集当前时间的前一个时间，h_{t-1}为所述输出门在所述问答词向量集当前时间的前一个时间的输出数据，w_t为所述当前时间的权重，b_t为所述当前时间的偏置，[]为矩阵乘法操作，

σ表示所述 sigmoid 函数。

S5、接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

发明还提供一种智能化的情感问答装置。参照图 2 所示，为本申请一实施例提供的智能化的情感问答装置的内部结构示意图。

在本实施例中，所述智能化的情感问答装置 1 可以是 PC（Personal Computer，个人电脑），或者是智能手机、平板电脑、便携计算机等终端设备，也可以是一种服务器等。该智能化的情感问答装置 1 至少包括存储器 11、处理器 12，通信总线 13，以及网络接口 14。

其中，存储器 11 至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器（例如，SD 或 DX 存储器等）、磁性存储器、磁盘、光盘等。存储器 11 在一些实施例中可以是智能化的情感问答装置 1 的内部存储单元，例如该智能化的情感问答装置 1 的硬盘。存储器 11 在另一些实施例中也可以是智能化的情感问答装置 1 的外部存储设备，例如智能化的情感问答装置 1 上配备的插接式硬盘，智能存储卡（Smart Media Card， SMC），安全数字（Secure Digital， SD）卡，闪存卡（Flash Card）等。进一步地，存储器 11 还可以既包括智能化的情感问答装置 1 的内部存储单元也包括外部存储设备。存储器 11 不仅可以用于存储安装于智能化的情感问答装置 1 的应用软件及各类数据 ,例如智能化的情感问答程序 01 的代码等，还可以用于暂时地存储已经输出或者将要输出的数据。

处理器 12 在一些实施例中可以是一中央处理器（Central Processing Unit，CPU）、控制器、微控制器、微处理器或其他数据处理芯片，用于运行存储器 11 中存储的程序代码或处理数据，例如执行智能化的情感问答程序 01 等。

通信总线 13 用于实现这些组件之间的连接通信。

网络接口 14 可选的可以包括标准的有线接口、无线接口(如 WI-FI 接口），通常用于在该装置 1 与其他电子设备之间建立通信连接。

可选地，该装置 1 还可以包括用户接口，用户接口可以包括显示器（Display）、输入单元比如键盘（Keyboard），可选的用户接口还可以包括

标准的有线接口、无线接口。可选地，在一些实施例中，显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及 OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以适当的称为显示屏或显示单元，用于显示在智能化的情感问答装置 1 中处理的信息以及用于显示可视化的用户界面。

图 2 仅示出了具有组件 11-14 以及智能化的情感问答程序 01 的智能化的情感问答装置 1，本领域技术人员可以理解的是，图 1 示出的结构并不构成对智能化的情感问答装置 1 的限定，可以包括比图示更少或者更多的部件，或者组合某些部件，或者不同的部件布置。

在图 2 所示的装置 1 实施例中，存储器 11 中存储有智能化的情感问答程序 01；处理器 12 执行存储器 11 中存储的智能化的情感问答程序 01 时实现如下步骤：

步骤一、通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集。

本申请较佳实施例，根据所述网络爬虫(Web crawler)技术从 URL 页面内爬取以文本形式提问的问题，将所述以文本形式提问的问题组成问题数据集。遍历所述问题数据集内的问题，使用所述网络爬虫技术从 URL 页面内爬取与所述问题对应的多种不同的答案，直至所述问题数据集遍历结束，得到与所述问题数据集对应的多种答案数据集；

本申请较佳实施例中，根据所述问题数据集，将所述多种答案数据集中的答案按照情感倾向进行情感属性标注，得到所述情感属性标注集，所述情感倾向包括幽默、诚恳、建议等。

步骤二、对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据 Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集。

本申请较佳实施例中，所述分词根据所述所述问答数据集建立分词概率模型P(S)，并最大化所述分词概率模型，完成分词操作，所述分词概率模型P(S)

为：

$$\mathrm{P(S)} = \mathrm{P}(W_1, W_2, \ldots, W_m) = \prod_{i=1}^{m} p(W_i \mid W_{i-1})$$

其中，$W_1, W_2, \ldots, W_m$为所述问答数据集内数据的词，m为所述问答数据集的数量；

所述关键字抽取包括构建所述词的相关度并基于所述相关度提取关键字，所述相关度为：

$$f(W_i, W_j) = \frac{tfidf(W_i) * tfidf(W_j)}{d^2}$$

其中，$f(W_i, W_j)$为词W_i和词W_j的相关度，$tfidf(W_i)$为词W_i的词频与逆向频率值，$tfidf(W_j)$为词W_j的词频与逆向频率值，d为词W_i和词W_j关于词向量的欧式距离；

本申请较佳实施例所述 Word2Vec 算法为 CBOW 模型，所述 CBOW 模型包括输入层、投影层和输出层，所述投影层$\zeta(\omega, j)$为：

$$\zeta(\omega, j) = \left(1 - d_j^{\omega}\right) log\left[\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right] + d_j^{\omega}\, log\left[1 - \sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]$$

其中，$d_j^{\omega} \epsilon \{d_2^{\omega}, d_3^{\omega}, \ldots, d_{l^{\omega}}^{\omega}\}$，表示在路径$\omega$内，第$j$个结点对应的霍夫曼编码，$\theta$为所述 CBOW 模型的迭代因子，σ表示 sigmoid 函数，X_{ω}为所述预处理操作完成的问答数据集。

步骤三、将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练。

本申请较佳实施例中，所述卷积神经网络包括卷积层、池化层、Flatten 层、Dropout 层和全连接层，且所述问题词向量集在时间维度上是一维向量的形式，所以所述卷积层与所述池化层的过滤器(filter)也是一维向量，同时为了防止过拟合现象，在所述卷积层与所述池化层中加入所述 Dropout 层；当多次卷积与池化操作后再扁平化数据，即所述 Flatten 层的作用，最后通过所述全连接层(Dense)输出所述训练值.

本申请较佳实施例所述损失值E为：

$$E = \sum_{j=1}^{m} \sum_{x \in m} |x - \mu_j|^2$$

其中，x为所述训练值，μ_j为所述情感属性标注集，m为所述问答数据集的数量，所述预设阈值一般设定为 0.01。

步骤四、当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练。

本申请较佳实施例，所述循环神经网络为长短期记忆网络，所述长短期记忆网络包括遗忘门、输入门、输出门，所述遗忘门为：

$$f_t = \sigma(w_t[h_{t-1}, x_t] + b_t)$$

其中，f_t为所述遗忘门的输出数据，x_t为所述遗忘门的输入数据，t为所述问答词向量集的当前时间，$t-1$为所述问答词向量集当前时间的前一个时间，h_{t-1}为所述输出门在所述问答词向量集当前时间的前一个时间的输出数据，w_t为所述当前时间的权重，b_t为所述当前时间的偏置，[]为矩阵乘法操作，σ表示所述 sigmoid 函数。

步骤五、接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

可选地，在其他实施例中，智能化的情感问答程序还可以被分割为一个或者多个模块，一个或者多个模块被存储于存储器 11 中，并由一个或多个处理器（本实施例为处理器 12）所执行以完成本申请，本申请所称的模块是指能够完成特定功能的一系列计算机程序指令段，用于描述智能化的情感问答程序在智能化的情感问答装置中的执行过程。

例如，参照图 3 所示，为本申请智能化的情感问答装置一实施例中的智能化的情感问答程序的程序模块示意图，该实施例中，所述智能化的情感问答程序可以被分割为数据接收模块 10、数据处理模块 20、模型训练模块 30、问答结果输出模块 40 示例性地：

所述数据接收模块 10 用于：从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集。

所述数据处理模块 20 用于：对所述问答数据集进行包括分词和关键字抽

取的预处理操作，并根据 Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集。

所述模型训练模块 30 用于：将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中训练，所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，所述损失函数根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练，当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练。

所述问答结果输出模块 40 用于：接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

上述数据接收模块 10、数据处理模块 20、模型训练模块 30、问答结果输出模块 40 等程序模块被执行时所实现的功能或操作步骤与上述实施例大体相同，在此不再赘述。

此外，本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质上存储有智能化的情感问答程序，所述智能化的情感问答程序可被一个或多个处理器执行，以实现如下操作：

从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集。

对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据 Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集。

将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中训练，所述卷积神经网络模型训练得到训练值，并将所

述训练值输入至损失函数，所述损失函数根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练，当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练。

接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

需要说明的是，上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。并且本文中的术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、装置、物品或者方法不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、装置、物品或者方法所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在如上所述的一个存储介质(如 ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，或者网络设备等)执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

权 利 要 求 书

1、一种智能化情感问答方法，其特征在于，所述方法包括：

通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集；

对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据Word2Vec算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集；

将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练；

当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练；

接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

2、如权利要求1所述的智能化情感问答方法，其特征在于，所述损失值E为：

$$E=\sum_{j=1}^{m}\sum_{x\in m}|x-\mu_j|^2$$

其中，x为所述训练值，μ_j为所述情感属性标注集，m为所述问答数据集的数量。

3、如权利要求1所述的智能化情感问答方法，其特征在于，通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，包括：

根据所述网络爬虫技术从 URL 页面内爬取以文本形式提问的问题，将所述以文本形式提问的问题组成问题数据集；

遍历所述问题数据集内的问题，使用所述网络爬虫技术从所述 URL 页面内爬取与所述问题对应的多种答案，直至所述问题数据集遍历结束，得到与所述问题数据集对应的多种答案数据集。

4、如权利要求 3 所述的智能化情感问答方法，其特征在于，所述分词操作包括：

根据所述问答数据集建立分词概率模型P(S)，并最大化所述分词概率模型，完成分词操作，所述分词概率模型P(S)为：

$$\mathrm{P(S)} = \mathrm{P}(W_1, W_2, \ldots, W_m) = \prod_{i=1}^{m} p(W_i \mid W_{i-1})$$

其中，$W_1, W_2, \ldots, W_m$为所述问答数据集内数据的词，m为所述问答数据集的数量；

所述关键字抽取操作包括：

构建所述词的相关度，并基于所述相关度提取关键字，所述相关度为：

$$f(W_i, W_j) = \frac{tfidf(W_i) * tfidf(W_j)}{d^2}$$

其中，$f(W_i, W_j)$为词W_i和词W_j的相关度，$tfidf(W_i)$为词W_i的词频与逆向频率值，$tfidf(W_j)$为词W_j的词频与逆向频率值，d为词W_i和词W_j关于词向量的欧式距离。

5、如权利要求 4 中的智能化情感问答方法，其特征在于，所述 Word2Vec 算法为 CBOW 模型；

所述 CBOW 模型包括输入层、投影层和输出层；

所述投影层$\zeta(\omega, j)$为：

$$\zeta(\omega, j) = \left(1 - d_j^{\omega}\right) log\left[\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right] + d_j^{\omega}\, log\left[1 - \sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]$$

其中，$d_j^{\omega} \epsilon \{d_2^{\omega}, d_3^{\omega}, \ldots, d_{l^{\omega}}^{\omega}\}$，表示在路径$\omega$内，第$j$个结点对应的霍夫曼编码，$\theta$为所述 CBOW 模型的迭代因子，σ表示 sigmoid 函数，X_{ω}为所述预处理操作完成的问答数据集。

6、如权利要求 1 所述的智能化情感问答方法，其特征在于，所述卷积神经网络包括卷积层、池化层、Flatten 层、Dropout 层和全连接层。

7、如权利要求 5 所述的智能化情感问答方法，其特征在于，所述循环神经网络为长短期记忆网络，所述长短期记忆网络包括遗忘门、输入门、输出

门；

所述遗忘门为：

$$f_t = \sigma(w_t[h_{t-1}, x_t] + b_t)$$

其中，f_t为所述遗忘门的输出数据，x_t为所述遗忘门的输入数据，t为所述问答词向量集的当前时间，$t-1$为所述问答词向量集当前时间的前一个时间，h_{t-1}为所述输出门在所述问答词向量集当前时间的前一个时间的输出数据，w_t为所述当前时间的权重，b_t为所述当前时间的偏置，[]为矩阵乘法操作，σ表示所述 sigmoid 函数。

8、一种智能化的情感问答装置，其特征在于，所述装置包括存储器和处理器，所述存储器上存储有可在所述处理器上运行的智能化的情感问答程序，所述智能化的情感问答程序被所述处理器执行时实现如下步骤：

通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集；

对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据 Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集；

将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练；

当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练；

接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

9、如权利要求 8 所述的智能化的情感问答装置，其特征在于，所述损失值E为：

$$E = \sum_{j=1}^{m} \sum_{x \in m} |x - \mu_j|^2$$

其中，x为所述训练值，μ_j为所述情感属性标注集，m为所述问答数据集的数量。

10、如权利要求 8 所述的智能化的情感问答装置，其特征在于，通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，包括：

根据所述网络爬虫技术从 URL 页面内爬取以文本形式提问的问题，将所述以文本形式提问的问题组成问题数据集；

遍历所述问题数据集内的问题，使用所述网络爬虫技术从所述 URL 页面内爬取与所述问题对应的多种答案，直至所述问题数据集遍历结束，得到与所述问题数据集对应的多种答案数据集。

11、如权利要求 10 所述的智能化的情感问答装置，其特征在于，所述分词操作包括：

根据所述所述问答数据集建立分词概率模型P(S)，并最大化所述分词概率模型，完成分词操作，所述分词概率模型P(S)为：

$$\mathrm{P(S)} = \mathrm{P}(W_1, W_2, \ldots, W_m) = \prod_{i=1}^{m} p(W_i \mid W_{i-1})$$

其中，$W_1, W_2, \ldots, W_m$为所述问答数据集内数据的词，m为所述问答数据集的数量；

所述关键字抽取操作包括：

构建所述词的相关度并基于所述相关度提取关键字，所述相关度为：

$$f(W_i, W_j) = \frac{tfidf(W_i) * tfidf(W_j)}{d^2}$$

其中，$f(W_i, W_j)$为词W_i和词W_j的相关度，$tfidf(W_i)$为词W_i的词频与逆向频率值，$tfidf(W_j)$为词W_j的词频与逆向频率值，d为词W_i和词W_j关于词向量的欧式距离。

12、如权利要求 11 所述的智能化的情感问答装置，其特征在于，所述 Word2Vec 算法为 CBOW 模型；

所述 CBOW 模型包括输入层、投影层和输出层；

所述投影层$\zeta(\omega, j)$为：

$$\zeta(\omega, j) = \left(1 - d_j^{\omega}\right) log\left[\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right] + d_j^{\omega}\, log\left[1 - \sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]$$

其中，$d_j^{\omega}\epsilon\{d_2^{\omega}, d_3^{\omega}, \ldots, d_{l^{\omega}}^{\omega}\}$，表示在路径$\omega$内，第$j$个结点对应的霍夫曼编码，$\theta$为所述 CBOW 模型的迭代因子，σ表示 sigmoid 函数，X_{ω}为所述预处理操作完成的问答数据集。

13、如权利要求 8 所述的智能化的情感问答装置，其特征在于，所述卷积神经网络包括卷积层、池化层、Flatten 层、Dropout 层和全连接层。

14、如权利要求 12 所述的智能化的情感问答装置，其特征在于，所述循环神经网络为长短期记忆网络，所述长短期记忆网络包括遗忘门、输入门、输出门；

所述遗忘门为：

$$f_t = \sigma(w_t[h_{t-1}, x_t] + b_t)$$

其中，f_t为所述遗忘门的输出数据，x_t为所述遗忘门的输入数据，t为所述问答词向量集的当前时间，$t-1$为所述问答词向量集当前时间的前一个时间，h_{t-1}为所述输出门在所述问答词向量集当前时间的前一个时间的输出数据，w_t为所述当前时间的权重，b_t为所述当前时间的偏置，[]为矩阵乘法操作，σ表示所述 sigmoid 函数。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有智能化的情感问答程序，所述智能化的情感问答程序可被一个或者多个处理器执行，以实现如下步骤：

通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到与所述问答数据集对应的情感属性标注集；

对所述问答数据集进行包括分词和关键字抽取的预处理操作，并根据 Word2Vec 算法对所述预处理操作完成的问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集；

将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中，利用所述卷积神经网络模型训练得到训练值，并将所述训练值输入至损失函数，利用所述损失函数并根据所述情感属性标注集与所述训练值计算得到损失值，判断所述损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练；

当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量

集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练；

接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案。

16、如权利要求 15 所述的计算机可读存储介质，其特征在于，所述损失值E为：

$$E = \sum_{j=1}^{m} \sum_{x \in m} |x - \mu_j|^2$$

其中，x为所述训练值，μ_j为所述情感属性标注集，m为所述问答数据集的数量。

17、如权利要求 15 所述的计算机可读存储介质，其特征在于，通过网络爬虫技术从互联网中获取问题数据集和与所述问题数据集对应的多种答案数据集，包括：

根据所述网络爬虫技术从 URL 页面内爬取以文本形式提问的问题，将所述以文本形式提问的问题组成问题数据集；

遍历所述问题数据集内的问题，使用所述网络爬虫技术从所述 URL 页面内爬取与所述问题对应的多种答案，直至所述问题数据集遍历结束，得到与所述问题数据集对应的多种答案数据集。

18、如权利要求 17 所述的计算机可读存储介质，其特征在于，所述分词操作包括：

根据所述所述问答数据集建立分词概率模型P(S)，并最大化所述分词概率模型，完成分词操作，所述分词概率模型P(S)为：

$$\mathrm{P(S)} = \mathrm{P}(W_1, W_2, \ldots, W_m) = \prod_{i=1}^{m} p(W_i | W_{i-1})$$

其中，$W_1, W_2, \ldots, W_m$为所述问答数据集内数据的词，m为所述问答数据集的数量；

所述关键字抽取操作包括：

构建所述词的相关度并基于所述相关度提取关键字，所述相关度为：

$$f(W_i, W_j) = \frac{tfidf(W_i) * tfidf(W_j)}{d^2}$$

其中，$f(W_i, W_j)$为词W_i和词W_j的相关度，$tfidf(W_i)$为词W_i的词频与逆向

频率值，$tfidf(W_j)$为词W_j的词频与逆向频率值，d为词W_i和词W_j关于词向量的欧式距离。

19、如权利要求 18 所述的计算机可读存储介质，其特征在于，所述 Word2Vec 算法为 CBOW 模型；

所述 CBOW 模型包括输入层、投影层和输出层；

所述投影层$\zeta(\omega,j)$为：

$$\zeta(\omega,j)=\left(1-d_j^{\omega}\right)log\left[\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]+d_j^{\omega}\,log\left[1-\sigma\left(X_{\omega}\theta_{j-1}^{\omega}\right)\right]$$

其中，$d_j^{\omega}\epsilon\{d_2^{\omega},d_3^{\omega},\dots,d_l^{\omega}\}$，表示在路径$\omega$内，第$j$个结点对应的霍夫曼编码，$\theta$为所述 CBOW 模型的迭代因子，σ表示 sigmoid 函数，X_{ω}为所述预处理操作完成的问答数据集。

20、如权利要求 19 所述的计算机可读存储介质，其特征在于，所述循环神经网络为长短期记忆网络，所述长短期记忆网络包括遗忘门、输入门、输出门；

所述遗忘门为：

$$f_t=\sigma(w_t[h_{t-1},x_t]+b_t)$$

其中，f_t为所述遗忘门的输出数据，x_t为所述遗忘门的输入数据，t为所述问答词向量集的当前时间，$t-1$为所述问答词向量集当前时间的前一个时间，h_{t-1}为所述输出门在所述问答词向量集当前时间的前一个时间的输出数据，w_t为所述当前时间的权重，b_t为所述当前时间的偏置，[]为矩阵乘法操作，σ表示所述 sigmoid 函数。

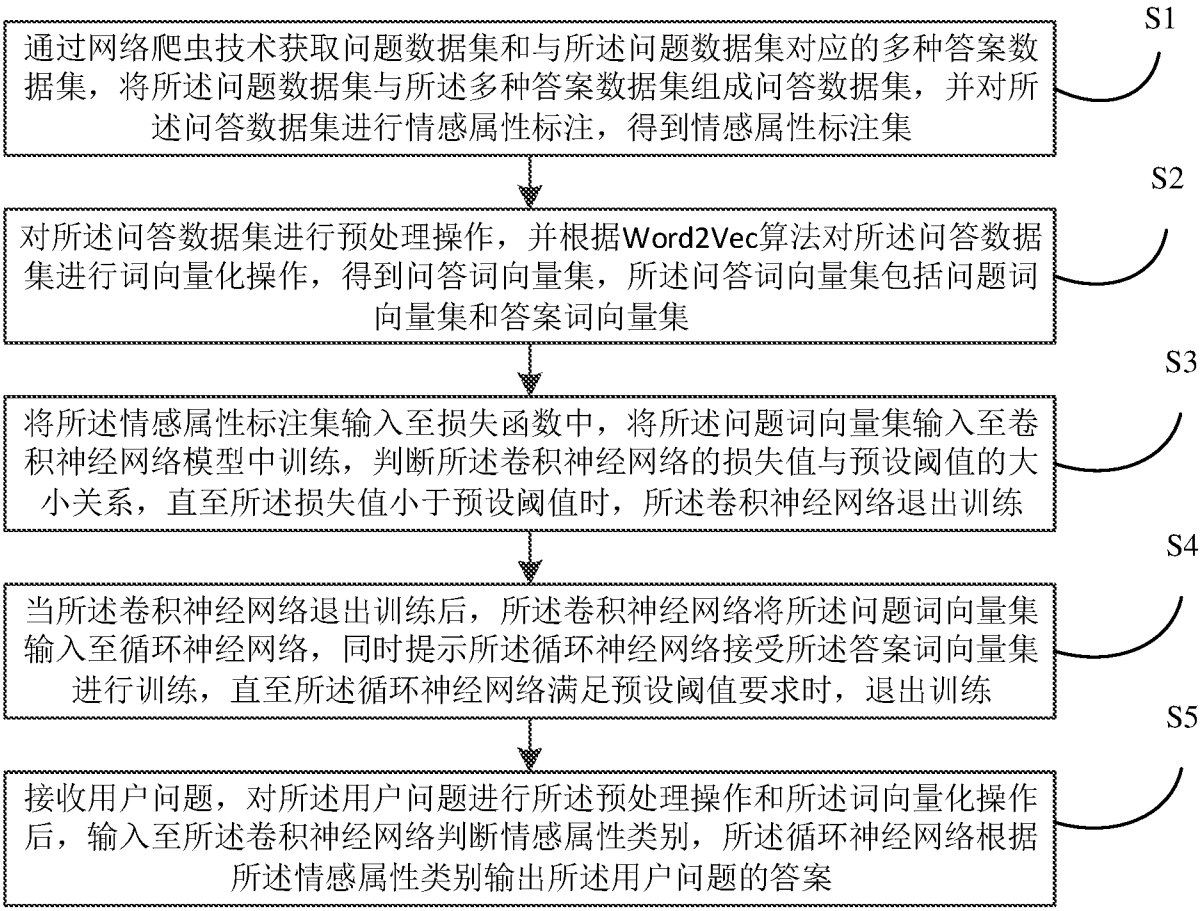

S1
通过网络爬虫技术获取问题数据集和与所述问题数据集对应的多种答案数据集，将所述问题数据集与所述多种答案数据集组成问答数据集，并对所述问答数据集进行情感属性标注，得到情感属性标注集
S2
对所述问答数据集进行预处理操作，并根据Word2Vec算法对所述问答数据集进行词向量化操作，得到问答词向量集，所述问答词向量集包括问题词向量集和答案词向量集
S3
将所述情感属性标注集输入至损失函数中，将所述问题词向量集输入至卷积神经网络模型中训练，判断所述卷积神经网络的损失值与预设阈值的大小关系，直至所述损失值小于预设阈值时，所述卷积神经网络退出训练
S4
当所述卷积神经网络退出训练后，所述卷积神经网络将所述问题词向量集输入至循环神经网络，同时提示所述循环神经网络接受所述答案词向量集进行训练，直至所述循环神经网络满足预设阈值要求时，退出训练
S5
接收用户问题，对所述用户问题进行所述预处理操作和所述词向量化操作后，输入至所述卷积神经网络判断情感属性类别，所述循环神经网络根据所述情感属性类别输出所述用户问题的答案

图 1

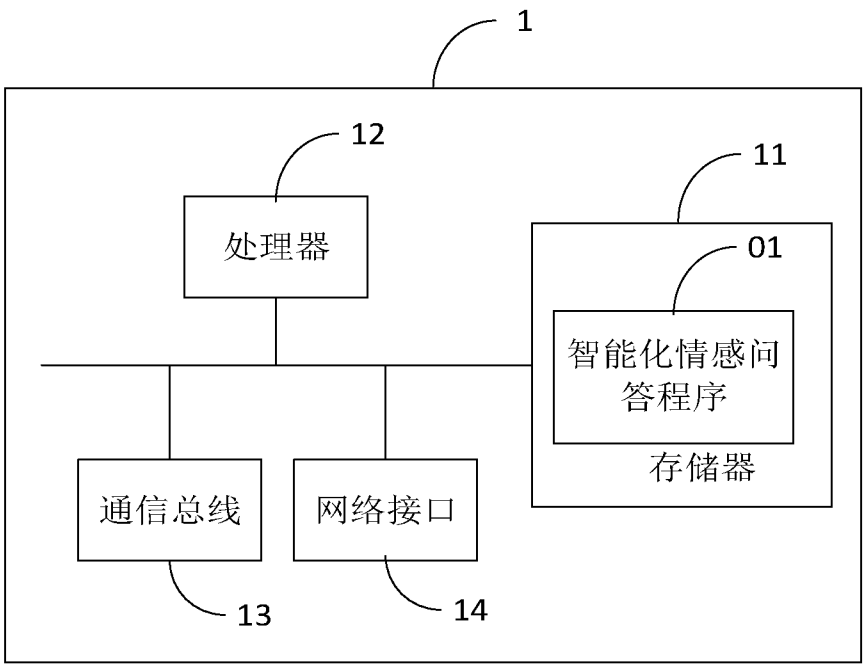

1
12
处理器
11
01
智能化情感问答程序
存储器
通信总线
网络接口
13
14

图 2

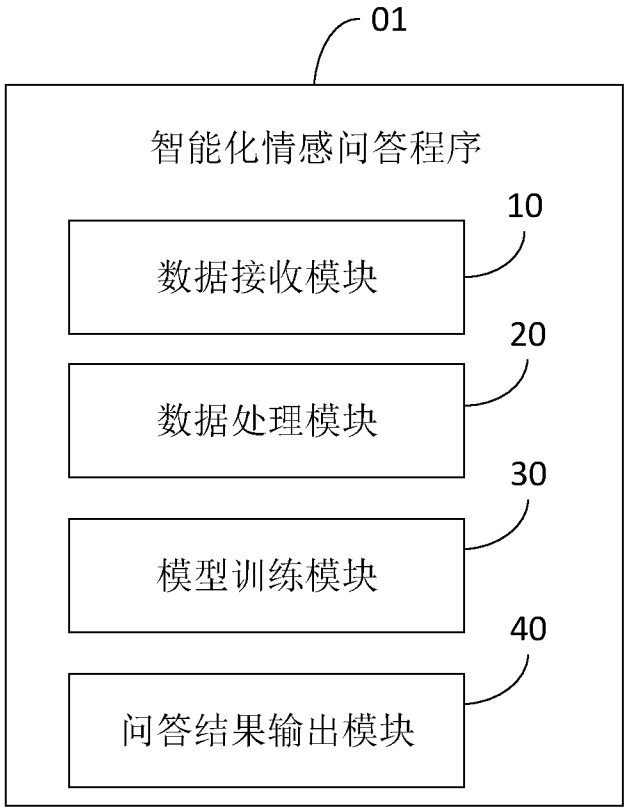


图 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/102194

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/332(2019.01)i; G06F 16/35(2019.01)i; G06F 16/95(2019.01)i; G06N 3/04(2006.01)i; G06N 3/08(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F; G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNKI, CNPAT, 情感, 问题, 答案, 问答, 应答, 回答, 训练, 损失, 卷积, 长短期记忆, 递归, 循环, 神经网络, 输出, 分词, 函数, emotion+, query, question, answer, training, loss, CNN, RNN, LSTM, GRV, function, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108932342 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 04 December 2018 (2018-12-04) description, paragraphs [0057]-[0128], and figures 1-9	1-20
A	CN 108875074 A (BEIJING HUIWEN TECHNOLOGY DEVELOPMENT CO., LTD.) 23 November 2018 (2018-11-23) entire document	1-20
A	CN 108427670 A (CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 21 August 2018 (2018-08-21) entire document	1-20
A	CN 107066446 A (GUANGDONG UNIVERSITY OF TECHNOLOGY) 18 August 2017 (2017-08-18) entire document	1-20
A	CN 109408633 A (SUN YAT-SEN UNIVERSITY et al.) 01 March 2019 (2019-03-01) entire document	1-20
A	US 2014229417 A1 (LELA, INC.) 14 August 2014 (2014-08-14) entire document	1-20

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 January 2020	**20 February 2020**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/102194

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2019000326 A1 (MICROSOFT TECHNOLOGY LICENSING LLC et al.) 03 January 2019 (2019-01-03) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/102194

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 108932342 A	04 December 2018	None	
CN 108875074 A	23 November 2018	None	
CN 108427670 A	21 August 2018	None	
CN 107066446 A	18 August 2017	None	
CN 109408633 A	01 March 2019	None	
US 2014229417 A1	14 August 2014	WO 2014127069 A1	21 August 2014
WO 2019000326 A1	03 January 2019	CN 109844741 A	04 June 2019

国际检索报告

国际申请号
PCT/CN2019/102194

A. 主题的分类

G06F 16/332(2019.01)i; G06F 16/35(2019.01)i; G06F 16/95(2019.01)i; G06N 3/04(2006.01)i; G06N 3/08(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F; G06N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

WPI,EPODOC,CNKI,CNPAT, 情感，问题，答案，问答，应答，回答，训练，损失，卷积，长短期记忆，递归，循环，神经网络，输出，分词，函数，emotion+, query, question, answer, training, loss, CNN, RNN, LSTM, GRV, function, model

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 108932342 A (腾讯科技深圳有限公司) 2018年 12月 4日 (2018-12-04) 说明书57至128段及图1至9	1-20
A	CN 108875074 A (北京慧闻科技发展有限公司) 2018年 11月 23日 (2018-11-23) 全文	1-20
A	CN 108427670 A (重庆邮电大学) 2018年 8月 21日 (2018-08-21) 全文	1-20
A	CN 107066446 A (广东工业大学) 2017年 8月 18日 (2017-08-18) 全文	1-20
A	CN 109408633 A (中山大学 等) 2019年 3月 1日 (2019-03-01) 全文	1-20
A	US 2014229417 A1 (LELA INC.) 2014年 8月 14日 (2014-08-14) 全文	1-20
A	WO 2019000326 A1 (MICROSOFT TECHNOLOGY LICENSING LLC等) 2019年 1月 3日 (2019-01-03) 全文	1-20

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：
“A” 认为不特别相关的表示了现有技术一般状态的文件
“E” 在国际申请日的当天或之后公布的在先申请或专利
“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）
“O” 涉及口头公开、使用、展览或其他方式公开的文件
“P” 公布日先于国际申请日但迟于所要求的优先权日的文件
“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件
“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性
“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性
“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2020年 1月 19日	2020年 2月 20日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 沈敏洁 电话号码 86-(10)-53961293

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/102194

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 108932342 A	2018年 12月 4日	无	
CN 108875074 A	2018年 11月 23日	无	
CN 108427670 A	2018年 8月 21日	无	
CN 107066446 A	2017年 8月 18日	无	
CN 109408633 A	2019年 3月 1日	无	
US 2014229417 A1	2014年 8月 14日	WO 2014127069 A1	2014年 8月 21日
WO 2019000326 A1	2019年 1月 3日	CN 109844741 A	2019年 6月 4日