



(12)发明专利申请

(10)申请公布号 CN 108305690 A

(43)申请公布日 2018.07.20

(21)申请号 201810122828.2

(22)申请日 2018.02.07

(71)申请人 王博冉

地址 100050 北京市东城区天坛西里6号天
坛医院医务处

(72)发明人 王博冉 王志斌 易鹏

(74)专利代理机构 北京金蓄专利代理有限公司
11544

代理人 孙巍

(51)Int.Cl.

G16H 50/70(2018.01)

G06F 17/30(2006.01)

G06N 3/12(2006.01)

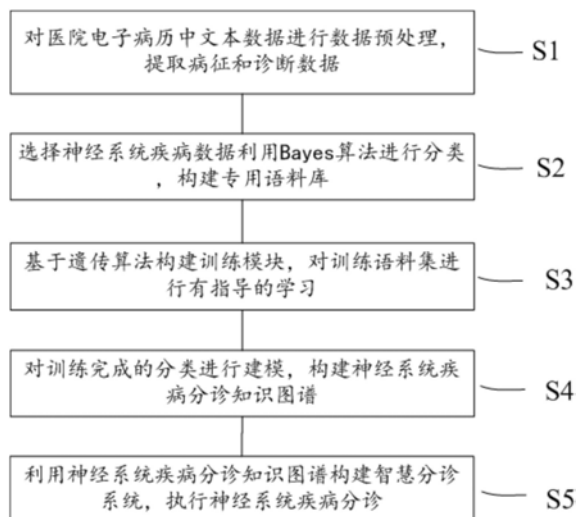
权利要求书1页 说明书6页 附图1页

(54)发明名称

神经系统疾病分诊方法及神经系统疾病的
分诊系统

(57)摘要

本发明提供了一种神经系统疾病的分诊方法及神经系统疾病的分诊系统。本发明的神经系统疾病分诊方法包括：第一步骤：对医院电子病历中文本数据进行数据预处理，提取病征和诊断数据；第二步骤：选择神经系统疾病数据利用Bayes算法进行分类，构建专用语料库；第三步骤：基于遗传算法构建训练模块，对训练语料集进行有指导的学习；第四步骤：对训练完成的分类进行建模，构建神经系统疾病分诊知识图谱。



1. 一种神经系统疾病分诊方法,其特征在于,所述神经系统疾病分诊方法包括:
第一步骤:对医院电子病历中文本数据进行数据预处理,提取病征和诊断数据;
第二步骤:选择神经系统疾病数据利用Bayes算法进行分类,构建专用语料库;
第三步骤:基于遗传算法构建训练模块,对训练语料集进行有指导的学习;
第四步骤:对训练完成的分类进行建模,构建神经系统疾病分诊知识图谱。
2. 根据权利要求1所述的神经系统疾病分诊方法,其特征在于,所述神经系统疾病分诊方法还包括:
第五步骤:利用神经系统疾病分诊知识图谱构建智慧分诊系统,执行神经系统疾病分诊。
3. 根据权利要求1或2所述的神经系统疾病分诊方法,其特征在于,第一步骤包括:采集特定时间段内的神经学科患者的门诊电子病历资料以完成神经系统疾病原始医疗文本信息采集,对原始医疗文本信息进行预处理以便将原始医疗文本中的特征值存入中间数据文件中,对中间数据文件进行概率计算以生成Bayes分类模型作为专用语料库。
4. 根据权利要求3所述的神经系统疾病分诊方法,其特征在于,特征值包括病征及诊断数据。
5. 根据权利要求1或2所述的神经系统疾病分诊方法,其特征在于,专用语料库包括:神经系统疾病的所有类别属性、每个类别属性对分类最有价值的特征属性、每个特征属性的权重、类别先验概率和类条件概率。
6. 根据权利要求1或2所述的神经系统疾病分诊方法,其特征在于,神经系统疾病分诊引导包括预约挂号流程引导。
7. 一种神经系统疾病分诊系统,其特征在于用于执行根据权利要求1至6之一所述的神经系统疾病分诊方法。
8. 根据权利要求7所述的神经系统疾病分诊系统,其特征在于包括:注册登录功能模块、基于症状的智能分诊功能模块、以及基于分诊结果的自助导诊功能模块。

神经系统疾病分诊方法及神经系统疾病的分诊系统

技术领域

[0001] 本发明涉及神经系统疾病智慧分诊领域,特别涉及一种基于机器学习和认知计算的神经系统疾病分诊方法及神经系统疾病分诊系统。

背景技术

[0002] 机器诊断 (Computer Aided Diagnosis, CAD) 这一概念最早Ledley在1966年提出。1972年,Willcox尝试将贝叶斯理论运用到计算机鉴别细菌疾病中。2007年, Saeys在牛津大学出版社出版的生物信息学杂志 (Bioinformatics) 中回顾了文本挖掘中的特征选择技术在生物信息学中的应用。他指出,虽然特征选择 (Feature Selection) 技术在分类领域应用得十分广泛,但在生物医学领域的应用才刚起步,医疗文本信息具有维度大且样本长度短的特点,特征选择技术需要针对医疗领域数据的特点进行改进和优化。2010年,在同一杂志中, Abeel对癌症诊断中生化特征 (Biomarker) 识别的特征选择算法进行了研究,采用支持向量机分类算法,将整合的特征选择技术用在了疾病诊断中。近几年来,文本分类技术在医疗领域的研究日益成熟, Rupali将贝叶斯模型及Jelinek-Mercer平滑方法应用与心脏疾病诊断和预测中, Ahmed将改进的全贝叶斯分类器和人工蜂群结合起来应用于心脏疾病和神经疾病的诊断中,进行混合估计,整合模型在心脏疾病预测准确度接近100%。

[0003] 相比于国外,国内学者在医疗领域的研究不多,左春荣,余本宫将贝叶斯网络应用于大规模医疗数据,利用粗糙集理论进行剪枝,不断的约减属性,从而能够有效的降低网络结构的复杂性,从而获得最为精简的特征项。胡碧松,冯丹基于贝叶斯算法、嵌入式开发环境和面向对象程序设计语言,开发了一个移动端的诊断系统,根据疾病病征、检测和化验信息进行疾病智能诊断。柳秋云提出了一种改进的贝叶斯分类器,应用于医疗诊断中,代雪莲,陈正伟开展了粗糙集和遗传算法在医疗诊断中的应用研究。

[0004] 当前国内智慧分诊系统主要基于3D人体模型,按照人体性别差异进行导诊。患者可通过点选3D人体模型,选择病征部位后,系统进入模拟医生问诊流程,显示出此人体部位易出现的症状,并评估出导致该症状的病症,患者只需选择系统评估出的最大可能性疾病,即可查看疾病的详细介绍以及建议科室等相关就诊信息。系统核心主要为应用功能与疾病知识库之间数据信息的读取与交互,各导诊系统疾病知识图谱的建立需要将不同患者的常见病征进行分类整合,经过深入调研和大量数据收集,再通过反复数据建模完成。当前浙江大学医学院附属第一医院、温州医科大学附属第一医院等国内知名医院,已基于该模式完成其“掌上医院”智慧分诊功能建设,率先开展了智慧分诊应用。

[0005] 然而,对于上述现有技术,对于减少挂错号、减少找错科室/专家的概率等方面,仍然不是特别有效。

发明内容

[0006] 本发明所要解决的技术问题是针对现有技术中存在上述缺陷,提供一种能够减少挂错号、减少找错科室/专家的概率的神经系统疾病分诊方法及系统。

- [0007] 根据本发明,提供了一种神经系统疾病分诊方法,包括:
- [0008] 第一步骤:对医院电子病历中文本数据进行数据预处理,提取病征和诊断数据;
- [0009] 第二步骤:选择神经系统疾病数据利用Bayes算法进行分类,构建专用语料库;
- [0010] 第三步骤:基于遗传算法构建训练模块,对训练语料集进行有指导的学习;
- [0011] 第四步骤:对训练完成的分类进行建模,构建神经系统疾病分诊知识图谱。
- [0012] 进一步地,神经系统疾病分诊方法,包括第五步骤:利用神经系统疾病分诊知识图谱构建智慧分诊系统,执行神经系统疾病分诊。
- [0013] 优选地,第一步骤包括:采集特定时间段内的神经学科患者的门诊电子病历资料以完成神经系统疾病原始医疗文本信息采集,对原始医疗文本信息进行预处理以便将原始医疗文本中的特征值存入中间数据文件中,对中间数据文件进行概率计算以生成Bayes分类模型作为专用语料库。
- [0014] 优选地,特征值包括病征及诊断数据。
- [0015] 优选地,专用语料库包括:神经系统疾病的所有类别属性、每个类别属性对分类最有价值的特征属性、每个特征属性的权重、类别先验概率和类条件概率。
- [0016] 优选地,神经系统疾病分诊引导包括预约挂号流程引导。
- [0017] 根据本发明,还提供了一种神经系统疾病分诊系统,用于执行上述神经系统疾病分诊方法。
- [0018] 本发明针对从医院电子病历中提取的病征数据,应用Bayes算法构建专用语料库,同时基于遗传算法构建训练模块,通过机器学习对分类模型进行有指导的训练,完成将训练完成的分类模型应用于智慧分诊中。本发明构建了一种根据医疗文本特征进行适应性改进的、基于遗传算法的神经系统疾病智慧分诊分类模型和神经系统疾病知识图谱,并对该分类模型进行实证研究,将效果最优的神经系统分类模型应用于智慧分诊系统中,完成智慧分诊系统的搭建,使患者可以通过输入病征进行神经系统疾病自助导诊。

附图说明

- [0019] 结合附图,并通过参考下面的详细描述,将会更容易地对本发明有更完整的理解并且更容易地理解其伴随的优点和特征,其中:
- [0020] 图1示意性地示出了根据本发明优选实施例的神经系统疾病分诊方法的流程图。
- [0021] 图2示意性地示出了根据本发明优选实施例的神经系统疾病分诊系统的功能模块示意图。
- [0022] 需要说明的是,附图用于说明本发明,而非限制本发明。注意,表示结构的附图可能并非按比例绘制。并且,附图中,相同或者类似的元件标有相同或者类似的标号。

具体实施方式

- [0023] 为了使本发明的内容更加清楚和易懂,下面结合具体实施例和附图对本发明的内容进行详细描述。
- [0024] 随着社会经济的不断发展和人们生活水平的提高,患者对医院的服务水平与质量要求也越来越高。与此同时,医院规模逐步扩大,学科发展不断壮大,各科室的分工也越来越细致,设置越来越复杂,患者就诊时挂错号或是不能对症就诊的情况经常出现,导致门诊

效率降低,患者时间和精力浪费。

[0025] 不能对症的就诊主要有两方面原因,一方面是因为患者自身对疾病知识的了解有限,使得在预约挂号时不能准确地根据自己的病征,选择对症的专科专家;另一方面,是患者对医院的学科设置不够了解,对各科医师(专家)的研究方向以及所擅长的领域也不够了解,导致无法找到病症所需对应的医师。

[0026] 现代化的医院应该可以通过智慧分诊系统,利用相对成熟和应用价值高的人工神经网络技术构建分类模型,模拟医生问诊流程,帮助患者进行自助导诊,并根据患者的病征,对其可能患有的疾病进行初步判断,再根据疾病推荐对症的专科专家,从而解决上述两方面的问题,使得门诊效率提高的同时,患者的时间和精力也能得到控制。

[0027] 智慧分诊系统核心主要为应用功能与疾病知识库之间数据信息的读取与交互。疾病分类模型和知识图谱的建立需要将不同患者的常见疾病进行分类整合,经过深入调研和大量数据收集,再通过反复数据建模完成。

[0028] 本发明拟提出一种根据医疗文本特征进行适应性改进的,基于机器学习和认知计算的医学专科疾病智慧分诊分类模型和神经系统疾病知识图谱,并对该分类模型进行实证研究,将效果最优的分类模型应用于智慧分诊系统中,完成智慧分诊系统的建设,使患者可以通过输入病征进行神经系统疾病自助导诊,从而降低错误分诊率,减少退换号数量,提高看诊效率,提升管理能力,具有广泛的社会与经济效益。

[0029] 图1示意性地示出了根据本发明优选实施例的神经系统疾病分诊方法的流程图。

[0030] 具体地,如图1所示,根据本发明优选实施例的神经系统疾病分诊方法包括:

[0031] 第一步骤S1:对医院电子病历中文本数据进行数据预处理,提取病征和诊断数据

[0032] 第二步骤S2:选择神经系统疾病数据利用Bayes算法进行分类,构建专用语料库;

[0033] 具体地,第一步骤S1通过中文医学文本自然语义处理从医院电子病历中提取病征和诊断数据,其中执行数据预处理,通过读取原始医疗数据信息,并对信息进行预处理,生成中间文件(即语料库),将原始医疗文本数据中各项特征值存入中间数据文件中,使得在机器学习训练阶段只需读取中间数据文件中的信息即可完成类条件概率的计算。针对中文医学文本有医疗垂直行业文本特性,不同于其他领域或者是通用领域的独有特征,进行分词、词性分析、文体分析、句法分析、语义分析、脱敏,从而实现将原始医学文本中的关键信息全部提取。

[0034] 第三步骤S3:基于遗传算法构建训练模块,对训练语料集进行有指导的学习;

[0035] 具体地,例如,机器学习主要包括训练和测试评估两个模块,主要功能是对预处理模块输出的中间数据文件进行概率计算,生成分类模型,该模型主要包括所有类别属性、每个类别对分类最有价值的特征属性、每个特征属性的权重、类别先验概率和类条件概率等数据信息;测试和评估模块:主要是读取测试样本中的数据,完成类别判定、计算评估指标,获得评估结果报告。专业人员能够很方便的对平台各个环节进行控制和调整,如训练文档集和测试文档集的选择和装入,平滑处理子模块的增减,分类器的选择,对算法相关参数进行调整,查看测试结果及各个评估指标的值。通过机器学习实现中间库各关键信息间的语义关联关系提取,不规范用法、错误用法自动纠正、自动标准化。

[0036] 第四步骤S4:对训练完成的分类进行建模,构建神经系统疾病分诊知识图谱;

[0037] 优选地,分诊建模基于认知计算,其中基于自然语义处理和机器学习成果,通过结

构化和非结构化信息中学习和构建知识图谱,理解自然语言并更自然地与人类互动,同时持续捕获专业人员的知识和经验,提高人员认知过程,改善决策,提高组织内的决策质量和一致性。

[0038] 第五步骤S5:利用神经系统疾病分诊知识图谱构建智慧分诊系统,执行神经系统疾病分诊。

[0039] 例如,在具体实施例中,可具体执行下述步骤:

[0040] 1) 回顾性研究:采集例如2013.01.01至2017.12.31神经学科的患者门诊(含特需门诊)电子病历资料,完成神经系统疾病原始医疗文本信息采集。

[0041] 2) 文本挖掘(文本特征表示):对原始医疗文本信息进行预处理,将原始医疗文本中的各项特征值存入中间数据文件中。

[0042] 文本挖掘是文本分类的第一步,目的是将原始医疗文本中相关的患者病征及诊断数据进行转换,对原始文本量化程度会对后续的分类学习算法造成影响。

[0043] 3) 文本分类(文本特征加权):对中间数据文件进行概率计算,生成Bayes分类模型。

[0044] 该模型主要针对患者病征及诊断,包括神经系统疾病的所有类别属性、每个类别属性对分类最有价值的特征属性、每个特征属性的权重、类别先验概率和类条件概率等信息。

[0045] 4) 模型训练(机器学习):以例如近5年医疗文本作为测试样本,基于遗传算法构建训练模块,通过机器学习,读取测试样本中的数据,

[0046] 完成类别判定,计算评估指标,训练完成的神经系统疾病智慧分诊分类模型可应用于智慧分诊系统中。

[0047] 在本发明的另一优选实施例中,本发明还提供了一种采用根据本发明优选实施例的神经系统疾病分诊方法的神经系统疾病分诊系统。

[0048] 在具体应用中,可以基于分类模型构建神经系统疾病智慧分诊应用系统,优化医院预约挂号流程(官网、APP、微信公众号等)。

[0049] 系统核心主要为基于神经系统疾病分类模型及ICD-10疾病编码体系,构建神经系统疾病知识图谱,面向神经学科下各亚专科,提供模拟问诊流程,引导患者选择系统评估出的最大可能性疾病,即可查看疾病的详细介绍以及建议学科、专家等相关就诊信息。例如,系统应用功能主要包括:

[0050] (1) 病症引导:患者挂号预约时,根据分类模型选择病征,完成自助导诊过程;

[0051] (2) 疾病分诊:根据患者病征选择,以及疾病知识图谱,查看对症科室名称、科室详细介绍以及对症医师或专家基本信息、详细介绍等,此外还可以查看疾病相关防治知识。

[0052] 优选地,可以对神经系统疾病智慧分诊模型应用的效果进行调研,例如,对系统上线后神经学科门诊(含特需门诊)退换号数据进行更具针对性的影响因素研究分析,系统上线后结合文献资料和以往调研经验制定调查问卷,选择通过自助导诊的神经系统疾病患者进行发放。拟对收集数据做描述性、分析性及Logistic回归分析;或者,从问卷调查的患者中筛选出100名患者典型病例(如反映:病征引导分类、描述不清晰,疾病分诊内容不对症的患者)进行深度访谈,认真做好访谈记录,在患者知情、同意的情况下做好录音。

[0053] 根据调研及数据分析结果,形成神经系统疾病智慧分诊模型应用效果研究报告。

分类模型训练完成后,进行应用效果预估,也是分类模型优化的依据。例如,分析智慧分诊系统应用数据,持续优化分类模型,基于整合医学理论持续完善神经系统疾病的知识图谱(包含病症、学科、专家等)。建立智慧分诊系统与医院HIS系统、电子病历系统间数据互通,面向医生患者提供患者诊前病征信息,提高看诊效率。基于系统应用过程数据分析,结合整合医学理论,持续完善神经系统疾病知识图谱。通过患者病征与诊疗结果数据对比分析,进一步提升门诊管理能力。根据患者复杂病症看诊需求,提出新增或调整相应的多学科联合会诊项目设置。

[0054] 基于遗传算法的神经系统疾病智慧分诊分类模型,能够通过机器学习,在应用过程中随着样本量的增加进行自动的适应性改进,持续完善,降低错误分诊率,减少退换号数量的同时,提高门诊效率,有效提升患者满意度,提升患者获得感。

[0055] 随着社会健康需求的改变,医学服务流程也应随之而改变,在智慧分诊过程中,将尝试以整体观来对待患者,传统模式下把患者分在不同科室往往不能从根本上解决患者需求。分类模型及知识图谱优化过程中,将纳入整合医学理论,以解决疑难患者的多学科综合诊治问题,并应用于智慧分诊系统中,复杂情况患者将获得多学科诊治建议。

[0056] 本发明以神经系统疾病为研究对象,利用人工神经网络技术,从医院电子病历中提取的病征等数据,应用Bayes算法构建专用语料库,同时基于遗传算法构建训练模块,通过机器学习对分类模型进行有指导的训练,完成将训练完成的文本分类模型应用于智慧分诊系统中。同时构建专用知识图谱,通过认知计算识别用户症状,根据疾病进行自助导诊,向用户提供可能患有的疾病、对症专科专家基本信息和对症专科专家详细信息等。

[0057] 总之,现代化医院需要一种可智能分诊的系统,能够帮助患者进行自助导诊,对所患疾病进行初步判断,从而减少挂错号、找错科室/专家的概率。相应地,本发明利用现有的医疗文本信息和文本分类技术来提高医院门诊效率,具有重要意义,主要体现在以下3个方面:

[0058] (1) 患者角度--智慧分诊过程能够有效帮助患者了解自身病症,并且能够根据分诊结果了解医院学科设置、特色门诊、专家擅长领域等信息,节省患者及家属的时间和精力,减少不能对症就诊现象;

[0059] (2) 医生角度--接收患者诊前智慧分诊信息,相对专业的病征描述能够有效提高看诊效率,精准看诊;

[0060] (3) 医院管理角度--节省人工导诊成本、降低错误分诊率,减少退换号数量的同时,利用大数据和人工智能等理论作为依据,能够根据患者看诊需求,提出新增或调整相应的多学科联合会诊项目设置,进一步提升管理能力,为整合医学信息平台构建提供数据支撑。

[0061] 具体地,图2示意性地示出了根据本发明优选实施例的神经系统疾病分诊系统的功能模块示意图。

[0062] 如图2所示,基于机器学习和认知计算的智慧分诊系统,利用高性能分类器构建神经系统疾病智慧分诊系统专用知识图谱,并基于该知识图谱为用户提供分诊服务,分诊应用系统核心功能应包括:注册登录功能模块、基于症状的智能分诊功能模块、以及基于分诊结果的自助导诊功能模块(建议亚科、疾病知识、治疗方案等)。

[0063] 注册登录功能模块:主要功能包括注册、登录、登出、找回密码及注册信息修改查

看,主体为面向用户(患者),注册过程中需填写个人注册信息,可查看注册条款,通过提交进行数据验证,通过后用户注册信息提交至服务器,存入数据库。

[0064] 基于症状的智能分诊功能模块:主要功能包括识别用户输入的症状数据,引导并帮助用户自助分诊,该模块植入机器学习和认知计算平台的疾病专用分类模型,对用户输入的症状信息进行疾病类别判定。

[0065] 基于分诊结果的自助导诊功能模块:用户可以根据分诊结果查看对症的专科专家信息,该功能来源于已根据用户症状识别疾病进行自助导诊,由于用户对医院的专科设计和专家信息不了解,导致就诊困难,用户可以通过此模块查看可能患有的疾病、对症专科专家基本信息和对症专科专家详细信息等,信息来源基于医院专科知识图谱。

[0066] 需要说明的是,除非特别指出,否则说明书中的术语“第一”、“第二”、“第三”等描述仅仅用于区分说明书中的各个组件、元素、步骤等,而不是用于表示各个组件、元素、步骤之间的逻辑关系或者顺序关系等。

[0067] 可以理解的是,虽然本发明已以较佳实施例披露如上,然而上述实施例并非用以限定本发明。对于任何熟悉本领域的技术人员而言,在不脱离本发明技术方案范围情况下,都可利用上述揭示的技术内容对本发明技术方案作出许多可能的变动和修饰,或修改为等同变化的等效实施例。因此,凡是未脱离本发明技术方案的内容,依据本发明的技术实质对以上实施例所做的任何简单修改、等同变化及修饰,均仍属于本发明技术方案保护的范围

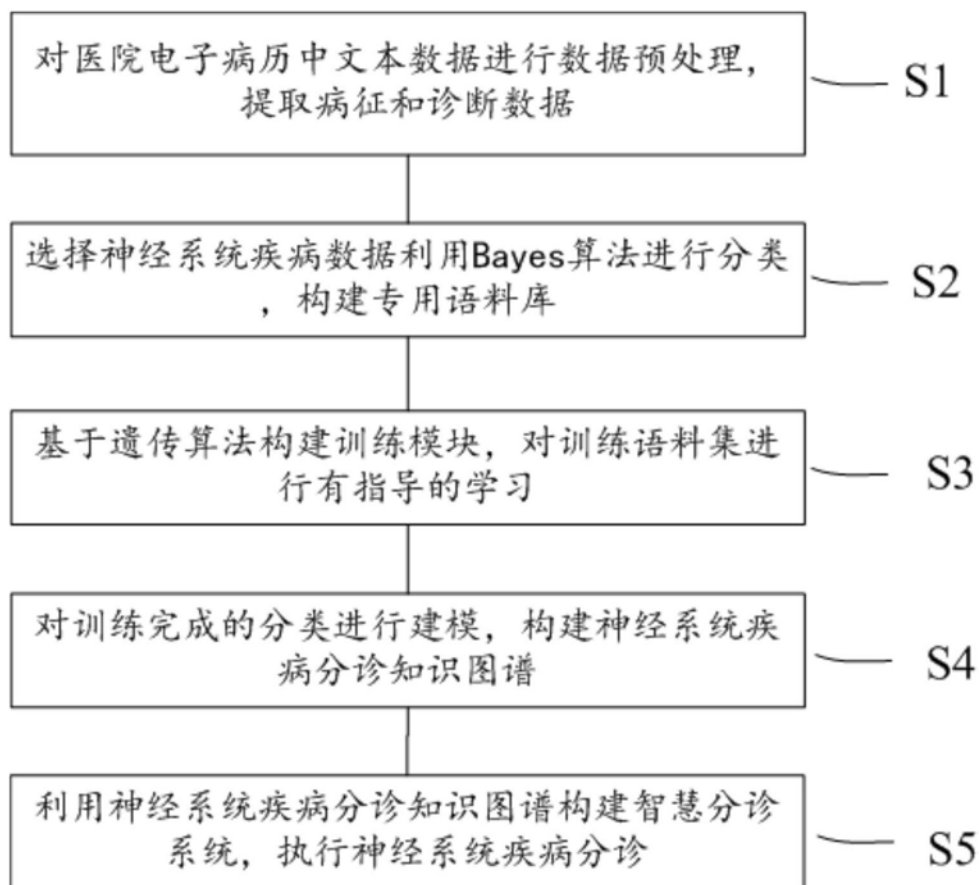


图1

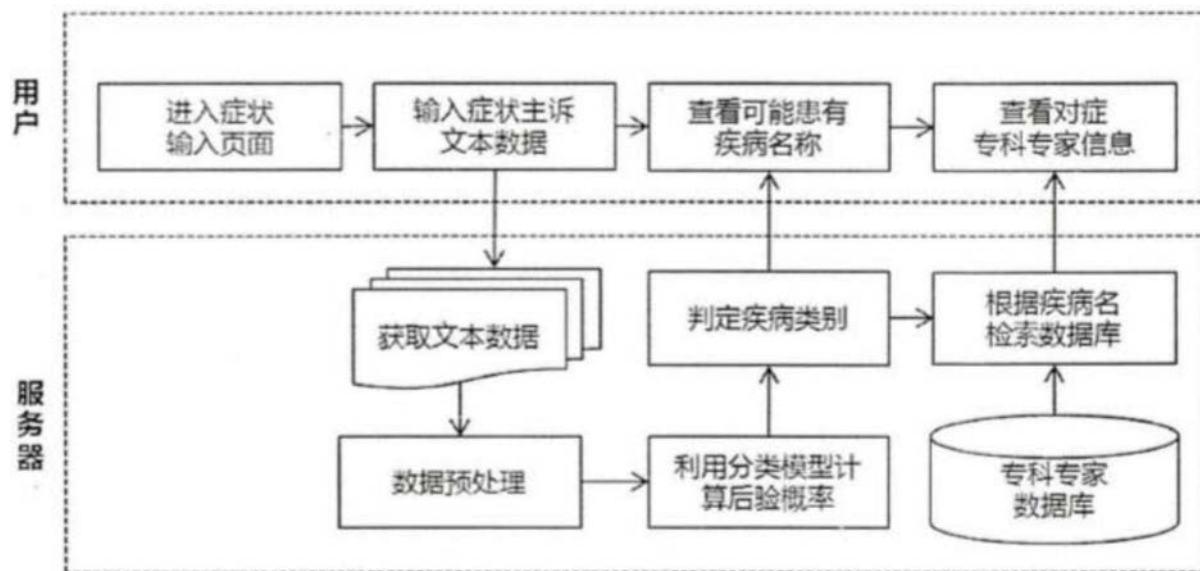


图2