**专利技术交底书**

提出部门：上海交通大学

全部发明人：曹健、沈正飞

第一发明人身份证号码：320422197209242817

联系人：曹健

联系电话：13651963843

电子邮箱：cao-jian@sjtu.edu.cn

1. **发明名称：**

基于专家权威性L确信算法与信度结合的治疗方案推荐方法

1. **所属技术领域：**

本发明主要涉及机器学习与医疗决策领域相关技术。本发明同时提出了专家权威性L确信算法用以量化每个历史病例的专家意见的不确定性。本发明最后利用历史病例的信度与不确定性对目标病例进行治疗方案上的推荐。

1. **技术背景**

对于医疗方案的推荐，现有的方法主要是结合历史病例数据与相应的机器学习算法进行最终的决策。不直接考虑病人的方案，其本身具有不确定性，所以要考虑专家的意见，要结合专家权威性；同时又要注意专家在不同病例上的权威性有差异，跟据病人背景和病例特点有不同。针对上述观点一个切实可行的思路是：对于一个容量给定的历史近似病例集合，对每一个历史病例的实际方案对最终决策支持程度（以下称为信度）进行量化，并利用DS信度结合理论进行融合，进而进行最终的决策。

但是，考虑到历史病例的治疗方案往往由多位医生共同投票给出，因此其自身本身具有一定的不确定性。而这种不确定性与医生们的投票情况存在密不可分的关系，直观上这种关系体现为如下两点：

1.医生们投票结果越一致，不确定性应该越小；反之亦然。

2.权威的多位医生投票结果越一致，或者存在某位非常权威的一位医生，不确定性应该越小；反之亦然。

本发明综合考量上述两个方面给出估计样本点不确定性的方法。

然后本发明利用该不确定性对历史病例的信度进行了优化，最后利用DS信度结合模型将其结合，并给出最后的决策结果。

1. **发明目的**

针对上述现有技术中的缺陷，本发明要解决的技术问题体现在以下几点：

1）量化专家投票结果与样本点不确定性之间的关系，给出计算样本点不确定性的算法

2）给出计算历史病例信度的方法，利用1）中量化的不确定性对其进行优化

3）利用DS信度结合理论将所有信度结合，给出最终推荐方案

1. **技术方案**

本发明中的方法包括以下步骤：

步骤1：确定目标病例的K个相似历史样本点。获取每个样本点医生对治疗方案投票的集合与该病例的最终治疗方案（即获得票数最多的方案），然后使用下述方法计算每名医生在这K个历史样本点所构成的集合内的准确率：1.若医生投票次数大于等于0.7\*K，则用其投票准确率表示其在这个集合中投票的准确率；2.若医生的投票次数小于0.7\*K，则将其准确率置为0。

步骤2:将所有医生的准确率按照由高到低的顺序排序获得权威专家列表。然后使用下述方法计算第i个样本点的不确定性UCi：1.权威专家列表中前L个医生投票结果一致，则令UCi的值为1；2.若权威专家列表中前L个医生投票结果不一致，则须首先计算投票的信息熵Hi，再将其带入到某个单调递减的函数中，这里推荐使用负指数函数，用函数值表示不确定度，即UCi=exp（-1\*Hi）

步骤3:对K个历史样本点中的每一个一次确定其mass函数的值：1.计算每个样本点对于目标病例的距离r，例如欧几里得距离。2.将该距离r代入到某个单调递减函数中以表示该样本点的mass 函数数值。

步骤4，对于K个历史病例中的每个样本点，将步骤2中的mass函数数值与步骤1中计算所得的样本点不确定性相乘，得到该点最终的信度。

步骤5，利用DS将步骤三中各个点的信度结合，选择信度最大的治疗方案作为最终的推荐治疗方案。

1. **本专利申请的有益效果**

由于采用了上述技术方案，本发明的有益效果是：

充分考虑了医生决策过程对样本点不确定性的影响，首次将医生意见的分歧度整合到DS理论之中，使得决策过程更加合理且更具有说服力，同时提升了推荐结果的准确率。

1. **可以实现本发明的方式与途径**（结合交底书第五项的技术方案来详细描述实施举例实现的过程）

下面结合具体实施例对本发明进行详细说明。以下实施例将有助于本领域的技术人员进一步理解本发明，但不以任何形式限制本发明。应当指出的是，对本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进。这些都属于本发明的保护范围。

符号定义：由医生决定，选取病人的m个已量化特征（包括病理类型，身高，体重等）构成维度为m的属性集合u={u1,u2...um},

步骤1：对于某个待预测治疗方案的病人，首先确定K个与之距离最近的历史病例，这里的距离可以用欧式距离。设当前待推荐病例与某历史病例的属性集分别为u'和u,则两个病例的欧式距离为：

将距离按照从小到大排序，选取前K个病例作为待预测病例的最相似病例，构成待预测方案病例的相似病例集合PNei

在这K个历史病例构成的集合PNei中，每位医生可能曾经对某些病例的治疗方案提出过意见。

设第i位医生曾经对这K个历史病例中的a个病例发表过意见。这a个病例中又有b个病例的实际治疗方案与该医生的意见一致，有b-a个病例与该医生的意见不一致，于是可以得到该医生在这个历史病例集合PNei中的准确率Acci：

其中factor是有效性因子。

上述规则的文字表述为：度量每一位医生投票的准确率应遵循如下规则：1.若医生投票次数大于等于f\*N，则用其投票准确率表示其在这个集合中投票的准确率；2.若医生的投票次数小于f\*N，则将其准确率置为0。

步骤2:将所有医生的准确率按照由高到低的顺序排序获得权威专家列表。然后使用下述方法计算第i个病例的不确定性UCi：

1.权威专家列表中前L个医生投票结果一致，则令UCi的值为1；

2.若权威专家列表中前L个医生投票结果不一致，则须首先计算投票的信息熵Hi，计算公式：

其中，n表示第i个病例的不同投票结果种类数，Pj表示第j种结果出现频率。

Hi计算完毕后，将其带入到某个单调递减的函数中，这里推荐使用负指数函数，用函数值表示不确定度，即UCi=exp（-1\*Hi）。

步骤3:利用如下算法确定K个相似病例每个点的mass函数数值：

对于第i位历史相似病人,预测方案为q的mass函数为：

其中，

这里，也可以被替换为其他形式，只要保证是的单调递减函数即可。

步骤4，对于K个历史病例中的每个样本点，将步骤2中的mass函数数值与步骤1中计算所得的样本点不确定性相乘，得到该点最终的信度：

步骤5，利用DS将步骤4中各个点的信度结合，选择信度最大的治疗方案作为最终的推荐治疗方案，公式如下：

1. **相关专利、公知技术检索结果及其与本发明/实用新型的异同点**

（检索到的专利和公知技术文献摘要以及这些专利和公知技术中存在的不足。这些文件可以是专利文件，也可以是期刊、杂志、手册和书籍等非专利文件。前者要写明国别、公开号和公开日期，后者应写明文件的详细出处。发明人应客观地指出这些专利和公知技术文献中的技术存在的问题和缺陷，这些问题和缺陷仅限于本申请的技术方案所解决的问题和缺点，可能的话，应说明产生这些问题和缺点的原因以及解决这些问题的难度。）

**关键词：** (填写专利与公知技术检索需要的关键词)

**关键词：** (填写专利与公知技术检索需要的关键词)

**相关检索结果1：**

申请（专利）号: xxxxxxxxxxxx 名称:xxxxxxxxxxxxxxxxx

摘要: xxxxxxxxxxxxxx

技术要点比较：

1．XXXXXXXXX：xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

2．XXXXXXXXXXXXXXXXXXXXXXXXX：xxxxxxxxxxxxxxxxxxxx

**相关检索结果2：**

申请（专利）号: xxxxxx 名称: xxxxxxxxx

摘要: xxxxxxxxxxxxxx

技术要点比较：

xxxxxxxxxxxxxxxxxxxxxxxxx

**相关检索结果3：**

现有公知技术名称，出处，

技术要点比较