（19）国家知识产权局

（12）发明专利申请

（10）申请公布号 CN *******
（43）申请公布日 2024.**.**

（21）申请号 *********
（22）申请日 2023.11.30
（71）申请人 观源（上海）科技有限公司
地址 200336 上海市闵行区紫星路 588 号 2 号楼 5A、5 层
（72）发明人 梁慧强 陆海宁 **

（74）专利代理机构 ********
专利代理师 *********

（51）

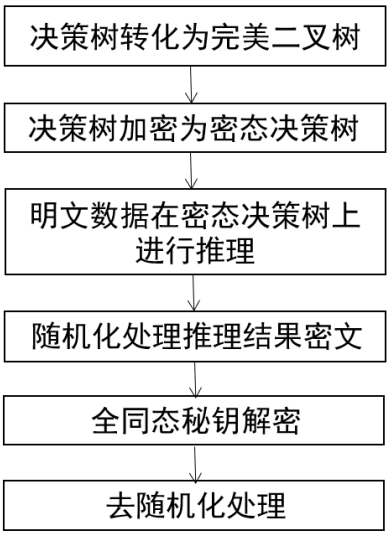
权利要求书 2 页 说明书 5 页 附图 4 页

（54）发明名称

一种基于全同态加密的隐私决策树推理方法

（57）摘要

本申请公开了一种基于全同态加密的隐私决策树推理方法，属于信息安全技术领域。在数据方-模型方的两方模型下，包括以下步骤：第一步，模型方将决策树转化为完美二叉决策树，然后加密为密态决策树后发送给数据方；第二步，数据方在密态决策树上使用数据明文进行推理得到推理结果密文，随机数掩盖得到分类结果密文，发送给模型方；第三步，模型方解密得到分类结果并发送给数据方；第四步，数据方根据分类结果和随机数得到推理结果。本发明提高了数据和决策树交互过程中的隐私性，数据方得到了推理结果而没有决策树的具体信息，而模型方没有数据的具体信息，解决了在两方模型的隐私决策树推理过程中，数据不出本地且决策树保持隐私的问题。



权 利 要 求 书

1. 一种基于全同态加密的隐私决策树推理方法，其特征在于，包括以下步骤：

将决策树模型转化为完美二叉决策树模型，并利用全同态加密算法对决策树模型进行加密，得到密态决策树模型；

根据所述决策树模型的层次，通过每一层内部节点的节点值密文计算出该层待比较的内部节点；

根据所述密文决策树模型所在层的待比较的内部节点的索引值密文、数据明文计算出该层待比较的内部节点位于索引的数据属性密文，利用全同态密-密比较算法对所述节点的门限值密文与数据属性密文进行比较，得到比较结果；

基于所述密文决策树模型所在层的待比较的内部节点的比较结果，计算该层节点的左右孩子节点的节点值密文；

根据叶子节点的节点值密文与分类标签值密文，计算出推理结果密文；

基于所述推理结果密文，随机数掩盖得到分类结果密文，使用全同态解密算法对所述分类结果密文进行解密得到分类结果，根据分类结果和随机数，得到推理结果，完成决策树推理过程。

2. 根据权利要求 1 所述的方法，其特征在于，将决策树模型转化为完美二叉决策树模型，包括：

对所述决策树模型的节点进行拆分得到二叉决策树模型，对二叉决策树模型的节点进行增添得到完美二叉树决策树模型。

3. 根据权利要求 1 所述的方法，其特征在于，利用全同态加密算法对决策树模型进行加密，得到密态决策树模型，包括：

使用全同态算法加密秘钥，对所述决策树模型中的内部节点和叶子节点的数据进行加密，得到所述密态决策树模型。

4. 根据权利要求 3 所述的方法，其特征在于，对所述决策树模型中的内部节点和叶子节点的数据进行加密，包括：

对所述决策树模型的内部节点的门限值和索引值进行加密，对所述决策树模型的叶子节点的分类标签值进行加密。

5. 根据权利要求 1 所述的方法，其特征在于，根据所述决策树模型的层次，通过每一层内部节点的节点值密文计算出该层待比较的内部节点，包括：

所述密态决策树模型所在层内部节点的节点值为$E(1)$或者$E(0)$，其中只有一个为$E(1)$，其余为$E(0)$，$E(\cdot)$代表全同态加密算法，该层所有内部节点共同计算得到节点值为$E(1)$的内部节点，作为所述密态决策树模型该层待比较的内部节点。

6. 根据权利要求 1 所述的方法，其特征在于，根据所述密文决策树模型所在层的待比较的内部节点的索引值密文、隐私数据明文计算出该层待比较的内部节点位于索引的数据属性密文，包括：

所述密态决策树模型所在层次待比较的内部节点的索引值密文为一个向量，其中位于索引的是$E(1)$，其余为$E(0)$，与数据明文进行计算得到索引值为$E(1)$的数据属性的全同态密文，作为该层待比较的内部节点位于索引的数据属性密文。

7. 根据权利要求 1 所述的方法，其特征在于，利用全同态密-密比较算法对所述节点的门限值密文与数据属性密文进行比较，得到比较结果，并用于计算该层节点的左右孩子节点的控制比特值，包括：

在所述节点中，其中$E(\vec{t})$代表所述节点的门限值密文，$E(\overrightarrow{x_a})$代表所述节点的数据属性密文，$CMP(*,*)$代表全同态密-密比较算法，若$x_a \geq t$，则b_l为$E(1)$，否则为$E(0)$。

8. 根据权利要求 1 所述的方法，其特征在于，基于所述密文决策树模型所在层的待比较的内部节点的比较结果，计算该层节点的左右孩子节点的节点值密文，包括：

对于根节点所在层，根据比较结果b_l，赋予左孩子节点或右孩子节点的节点值为$E(1)$或者$E(0)$，对于其他节点所在层，父节点值为c_i的左右孩子节点值分别赋予$b_l\&c_i, !b_l\&c_i$，其中&代表 AND 操作，!代表 NOT 操作，使得下一层所有节点的节点值只有一个为$E(1)$，其余为$E(0)$。

9. 根据权利要求 1 所述的方法，其特征在于，根据叶子节点的节点值密文与分类标签值密文，计算出推理结果密文，包括：

所述密态决策树模型叶子节点所在层的节点值只有一个为$E(1)$，其余为$E(0)$，与节点的分类标签值密文共同计算得到节点值为$E(1)$的所在节点的分类标签值密文，作为所述密态决策树模型在所述数据明文上推理的分类结果密文。

10. 根据权利要求 1 所述的方法，其特征在于，基于所述推理结果密文，随机数掩盖得到分类结果密文，使用全同态解密算法对所述分类结果密文进行解密得到分类结果，根据分类结果和随机数，得到推理结果，完成决策树推理过程，包括：

对于推理结果密文，随机取反得到一个使用随机数掩盖的推理结果的密文作为分类结果密文，其中取反位置为该随机数，使用全同态秘钥对分类结果密文进行解密得到随机数掩盖的推理结果，根据取反位置得到推理结果。

一种基于全同态加密的隐私决策树推理方法

技术领域

[0001] 本申请涉及信息安全技术领域，特别涉及一种基于全同态加密的隐私决策树推理方法。

背景技术

[0002] 数据有多重定义，它是对事实、活动等现象的记录。是基于二进制编码的、按预先设置的规则汇聚的现象记录。也可以理解为汇聚起来用于认知的原材料。与其他生产要素相比，数据要素具有多种特点，包括虚拟性、低成本赋值、主体多元性、非竞争性、潜在的排他性、异质性（相同数据对不同使用者和不同应用场景的价值不同）等等，例如数据要素的来源广泛、内容参差不齐，也可以不断的复制、共享、再生。同时数据形态多样，难以定价，缺乏统一的价值评估体系，此外，数据要素涉及到个人隐私和所有权问题。这些特点使得数据要素的市场化建设过程中面临更多新的挑战。

[0003] 数据流通是数据要素市场化建设的核心环节，描述了数据从数据提供方按照一定规则到达数据需求方的过程。而数据共享和数据交易作为数据流通的主要路径和重要手段，实现安全的数据共享和数据交易能有效的促进企业间的数据流通，解决数据孤岛问题，进而释放出数据要素的强大生产力。在数据安全共享方面，已经形成了多条技术路线，如：以安全多方计算为代表的基于密码学的隐私计算技术、以联邦学习为代表的人工智能与隐私保护技术同和衍生技术、以可信执行环境为代表的基于可信硬件的隐私计算技术。而在当前大模型阶段，数据共享的含义更加丰富，其中基于安全多方技术往往需要多轮交互和较大的带宽用于通信，而同态加密需要较大的计算量，数据安全共享方案仍需要结合实际需求进一步研究和探索。

[0004] 当前数据共享主要有四种形式：原始数据直接共享、数据查询接口服务、数据处理隐私后共享（例如对数据脱敏、匿名化、差分隐私等）、数据与模型定制服务。其中，数据与模型定制服务将数据中的信息嵌入到学习模型中，提供高级别的数据服务，同时有助于数据交易与共享过程中的安全与可信管控。在数据与模型定制服务中，模型方利用大量数据创建相关模型，然后定义接口实现与数据方之间的信息共享。举例来说，银行希望确定某个申请者是否合适贷款，银行将该申请者的信息（如年龄、收入、婚姻状况等）提交给模型方，模型方返回一个建议，即是否批准贷款。在实际场景下，往往需要满足一定的安全标准才能投入生产。首先，银行不希望自己用户的数据被泄露，而模型方也不希望其训练的模型参数泄露。为满足这一安全需求，抽象出一个隐私模型推理方案，该方案由两个参与方组成，即数据方和模型方。数据方持有私有数据，模型方持有私有模型。数据方希望使用模型方的私有模型对自己的私有数据进行推理，同时不泄露数据的任何信息，模型方则希望数据方只能获取推理结果，而无法获取私有模型的任何信息。

[0005] 同态加密算法引申于具有乘法同态的 RSA 加密，在密文上按照一定规则的运行结果解密后和相应明文按照一定规则运算后的结果一致。半同态加密是指具有单一加法或者乘法运算性质的同态，如具有加法同态的 Paillier 加密已经被广泛应用，能在一定程度上有效实现一部分数据的“可用不可见”。部分同态加密是指支持有限次数的加法和乘法的同态算法，随着 Gentry 在 2009 年提出的 bootstrapping 技术，可以将部分同态加密变为支持任意次数的全同态加密。目前以 FHEW 和 TFHE 为代表的第三代全同态加密，使用 GSW 方法将 bootstrapping 技术的计算时间降低在 0.1s 内，全同态加密算法也可以在实际的数据共享方法中应用。尽管效率与明文上直接计算相差数个量级，但是全同态算法直接在密文上进行任意深度运算的优越性质而具有非常高的通用性，并且在深度已知的运算情况下，使用部分同态加密可以避免

繁重的 bootstrapping 运算得以平衡计算效率和隐私性，也有批处理技术减轻运算负载。相比与需要多轮交互和较高通信量的安全多方技术实现的隐私模型推理方案，各具优势，适用于不同的场景。在实际情况下，带宽比计算量更加的珍贵，同态加密也有更多潜在的应用场景，可以从多个角度保护参与方的隐私数据。

[0006] 决策树作为一种普遍的机器学习算法，其简单性、可解释性和易于训练性使得决策树模型在不同场合下广泛使用。将决策树作为隐私模型具有重要意义，一方面决策树结构简单，决策树由内部节点和叶子节点组成，内部节点主要包含索引值和门限值，叶子节点包含分类标签值。另一方面，决策树的工作流程较为固定，输入数据，从根节点开始，通过内部节点的索引值得到数据位于索引的属性值与该节点的门限值进行比较，根据比较结果移动到下一层内部节点，依次进行直到叶子节点，返回分类标签值，完成决策树推理过程。最后，决策树具有高效性和广泛性，决策树仅使用深度次数的判断即可完成分类工作而具备高效性，在医疗、金融、科技和教育行业中普遍使用，实现诸多分类功能。在隐私模型推理方案中，根据实际需求可采用密态决策树和明文数据、明文决策树和密态数据或者密态决策树和密态数据，在这三种可能的方案的比较中，使用同态加密的密态决策树在通信量上相对较低，计算量持平，与使用海量的密态数据相比，决策树模型包含的信息更加精炼，在实际计算中也更加高效。

[0007] 随着要素市场化过程的纵深发展，数据流通的地位也会不断地提高，数据共享和数据交易的方式也会呈现更加多样化。安全的数据共享和数据交易对数据要素有序流通的意义重大，这有助于打破信息孤岛，释放数据要素强大的生产潜力、实现数据生态的完整产业链。隐私模型推理作为数据与定制服务，能有效实现安全的数据共享。在数据方-模型方的使用场景下，模型方的私有模型和数据方的私有数据都具有很高的敏感性和价值属性，需要我们在技术上全面的保护两者的信息，防止模型和数据的非法泄露、分发和复制，实现安全的数据共享服务，为数据要素的有序流通提供新的解决方案，促进数据要素市场化建设过程。

发明内容

[0008] 本申请提供一种基于全同态加密的隐私决策树推理方法，用于保护决策树推理过程中数据与模型安全，避免模型和数据被非法泄露、分发和复制的问题。

[0009] 本申请提供了一种基于全同态加密的隐私决策树推理方法，包括以下步骤：将决策树模型转化为完美二叉决策树模型，并利用全同态加密算法对决策树模型进行加密，得到密态决策树模型；根据所述决策树模型的层次，通过每一层内部节点的节点值密文计算出该层待比较的内部节点；根据所述密文决策树模型所在层的待比较的内部节点的索引值密文、数据明文计算出该层待比较的内部节点位于索引的数据属性密文，利用全同态密

-密比较算法对所述节点的门限值密文与数据属性密文进行比较，得到比较结果；基于所述密文决策树模型所在层的待比较的内部节点的比较结果，计算该层节点的左右孩子节点的节点值密文；根据叶子节点的节点值密文与分类标签值密文，计算出推理结果密文；基于所述推理结果密文，随机数掩盖得到分类结果密文，使用全同态解密算法对所述分类结果密文进行解密得到分类结果，根据分类结果和随机数，得到推理结果，完成决策树推理过程。

[0010] 将决策树模型转化为完美二叉决策树模型，包括：对所述决策树模型的节点进行拆分得到二叉决策树模型，对二叉决策树模型的节点进行增添得到完美二叉树决策树模型。

[0011] 利用全同态加密算法对决策树模型进行加密，得到密态决策树模型，包括：

使用全同态算法加密秘钥，对所述决策树模型中的内部节点和叶子节点的数据进行加密，得到所述密态决策树模型。

[0012] 根据所述决策树模型的层次，通过每一层的内部节点的节点值密文计算出该层待比较的内部节点，包括：所述密态决策树模型所在层内部节点的节点值为$E(1)$或者$E(0)$，其中

只有一个为$E(1)$，其余为$E(0)$，$E(\cdot)$代表全同态加密算法，该层所有内部节点共同计算得到节点值为$E(1)$的内部节点，作为所述密态决策树模型该层待比较的内部节点。

[0013] 根据所述密文决策树模型所在层的待比较的内部节点的索引值密文、隐私数据明文计算出该层待比较的内部节点位于索引的数据属性密文，包括：所述密态决策树模型所在层次待比较的内部节点的索引值密文为一个向量，其中位于索引的是$E(1)$，其余为$E(0)$，与数据明文进行计算得到索引值为$E(1)$的数据属性的全同态密文，作为该层待比较的内部节点位于索引的数据属性密文。

[0014] 利用全同态密-密比较算法对所述节点的门限值密文与数据属性密文进行比较，得到比较结果，并用于计算该层节点的左右孩子节点的控制比特值，包括：在所述节点中, 其中$E(\vec{t})$代表所述节点的门限值密文，$E(\overrightarrow{x_a})$代表所述节点的数据属性密文，$CMP(*,*)$代表全同态密-密比较算法，若$x_a \geq t$，则b_l为$E(1)$，否则为$E(0)$。

[0015] 基于所述密文决策树模型所在层的待比较的内部节点的比较结果，计算该层节点的左右孩子节点的节点值密文，包括：对于根节点所在层，根据比较结果b_l，赋予左孩子节点或右孩子节点的节点值为$E(1)$或者$E(0)$，对于其他节点所在层，父节点值为c_i的左右孩子节点值分别赋予$b_l \& c_i, !b_l \& c_i$，其中&代表 AND 操作，!代表 NOT 操作，使得下一层所有节点的节点值只有一个为$E(1)$，其余为$E(0)$。

[0016] 根据叶子节点的节点值密文与分类标签值密文，计算出推理结果密文，包括：所述密态决策树模型叶子节点所在层的节点值只有一个为$E(1)$，其余为$E(0)$，与节点的分类标签值密文共同计算得到节点值为$E(1)$的所在节点的分类标签值密文，作为所述密态决策树模型在所述数据明文上推理的分类结果密文。

[0017] 基于所述推理结果密文，随机数掩盖得到分类结果密文，使用全同态解密算法对所述分类结果密文进行解密得到分类结果，根据分类结果和随机数，得到推理结果，完成决策树推理过程，包括：对于推理结果密文，随机取反得到一个使用随机数掩盖的推理结果的密文作为分类结果密文，其中取反位置为该随机数，使用全同态秘钥对分类结果密文进行解密得到随机数掩盖的推理结果，根据取反位置得到推理结果。

附图说明

[0018] 本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中奖变得明显和容易理解，其中：

[0019] 图 1 为一种基于全同态加密的隐私决策树推理方法流程图：

[0020] 图 2 为一种基于全同态加密的隐私决策树推理方法交互图：

[0021] 图 3 为一种具体地基于全同态加密的隐私决策树推理方法流程图：

[0022] 图 4 为决策树模型转化为完美二叉决策树模型示例图：

[0023] 图 5 为决策树模型加密为密态决策树模型流程图：

[0024] 图 6 为密态决策树内部节点计算待比较中间节点流程图：

[0025] 图 7 为密态决策树内部节点进行全同态密-密比较算法得到节点值流程图：

[0026] 图 8 为密态决策树从第 0 层节点值得到第 1 层节点的节点值流程图：

[0027] 图 9 为密态决策树从第l层待比较节点和节点值后得到第$l+1$层节点的节点值流程图：

[0028] 图 10 为密态决策树计算第d层叶子节点的推理结果密文流程图：

具体实施方式

[0029] 下面详细描述本申请的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本申请，而不能理解为对本申请的限制。

[0030] 数据表示为一行n维向量，记为$\vec{x} = \{x_0, x_1, \ldots, x_{n-1}\}$，其中$x_i \in [0, 2^k - 1]$代表第$i$位的

数据属性。

[0031] 决策树模型是一树状结构，由内部节点和叶子节点组成，内部节点对应着在某个属性上的划分，根据样本数据在该属性上的不同取值将其划分为若干这个子集，叶子节点对应着一个分类。将决策树的节点按层编号，每一层从 0 开始从左往右依次编号。不失一般性，记决策树的内部节点共有m个，索引范围为$[0, n-1]$，深度为d，叶子节点有s个。

[0032] 用$T = (\vec{t}, \vec{a}, \vec{w})$表示决策树模型：

[0033] $\vec{t} = \{t_0, t_1, \dots, t_{m-1}\}$，$t$代表内部节点的门限值集合，$t_i \in [0, 2^k - 1]$代表第$i$个内部节点的门限值。

[0034] $\vec{a} = \{a_0, a_1, \dots, a_{m-1}\}$，$a$代表内部节点的索引值集合，$a_i \in [0, n-1]$代表第$i$个内部节点的索引值。

[0035] $\vec{w} = \{w_0, w_1, \dots, w_{s-1}\}$，$w$代表分类标签值集合，$w_i \in [0, s-1]$代表第$i$个叶子节点的分类标签值。

[0036] 用一个映射表示决策树推理过程：$w^* \leftarrow T(\vec{x})$，向决策树模型中输入待推理数据$\vec{x}$，从根节点开始，进入第$i$个内部节点，根据$a_i$的索引值找到数据属性$x_{a_i}$，然后与$t_i$进行比较，判断下一个需要比较的内部节点，直到叶子节点，返回该条数据的分类标签值w^*。

[0037] 密态决策树模型是决策树模型的一种密文形式，通过对上述决策树模型的内部节点和叶子节点的内容使用全同态算法进行加密，达到保护决策树模型的目的。其中节点编号和上述决策树模型保持一致，内部节点包含全同态加密算法加密的门限值密文和索引值密文。叶子节点包含全同态加密算法加密的分类标签值密文。

[0038] 密态决策树以明文数据或者密文数据作为输入，以全同态加密算法加密的一个分类标签值或者多个分类标签值作为输出。

[0039] 本申请实施例的一种基于全同态加密的隐私决策树推理方法，其交互式模型如图 2 所示，模型方预先将密态决策树模型发送给数据方，数据方根据数据进行密态推理得到推理结果密文，使用随机数掩盖得到分类结果密文，交由模型方解密，再用分类结果和随机数得到该条数据的推理结果。

[0040] 图 3 为一种具体地基于全同态加密的隐私决策树推理方法流程图，包括以下步骤：

[0041] 在步骤 S110 中，将决策树模型转化为二叉树模型，将二叉树模型转化为完美二叉树决策树模型。

[0042] 转化决策树的目的是为了得到相同形状二叉决策树以便于我们后续的推理任务，相同形状的决策树加密后的密态决策树形状也想通，不会泄露更多信息。

[0043] 具体地，如图 4 所示，其中，圆形代表内部节点，方形代表叶子节点，虚线部分代表添加的节点；通过对内部节点判断条件的分解，添加相应节点使得每一个内部节点只有左右两个孩子节点，对于所得到的二叉决策树，在保持分类结果不变的情况下，添加相应节点使其变为完美二叉树。

[0044] 在步骤 S120 中，将决策树模型加密，得到密态决策树模型。

[0045] 决策树加密的目的是保护决策树的关键信息，并且确保数据可以在密态决策树上进行推理。

[0046] 具体地，如图 5 所示，利用全同态加密算法对决策树的内部节点的门限值和索引值进行加密，其中，在内部节点门限值密文为门限值的二进制展开后的加密，索引值加密为n维的全同态密文，位于索引的为$E(1)$，其余为$E(0)$，在叶子节点分类标签值密文为分类标签值的二进制展开后的加密。

[0047] 在步骤 S130 中，根据所述决策树模型的层次，通过每一层内部节点的节点值密文计算出该层待比较的内部节点。

[0048] 计算该层待比较节点的目的是使得决策树的每一层只需要进行一次全同态密-密比较。

[0049] 具体地，如图 6 所示，所述密态决策树模型所在层内部节点的节点值为$E(1)$或者$E(0)$，

其中只有一个为$E(1)$，其余为$E(0)$，该层所有内部节点共同计算得到节点值为$E(1)$的内部节点，作为所述密态决策树模型该层待比较的内部节点。

[0050] 在步骤 S140 中，根据所述密文决策树模型所在层的待比较的内部节点的索引值密文、数据明文计算出该层待比较的内部节点位于索引的数据属性密文，利用全同态密-密比较算法对所述节点的门限值密文与数据属性密文进行比较，得到比较结果。

[0051] 计算待比较中间节点的比较结果的目的是为了下一层节点的节点值的计算。

[0052] 具体地，如图 7 所示，数据方的明文数据为$\vec{x} = \{x_0, x_1, \dots, x_{n-1}\}$，对每一个数据属性进行二进制展开得到$\vec{x} = \{\overrightarrow{x_0}, \overrightarrow{x_1}, \dots, \overrightarrow{x_{n-1}}\}$，其中$\vec{x}_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,n}\}, x_{i,j} \in \{0,1\}$表示$x_i$的二进制向量；密态决策树的内部节点有$E(\vec{t})$和$E(\vec{v})$，分别表示门限值密文和索引值密文，其中索引值密文为只有位于索引才为$E(1)$，其余为$E(0)$；与$\vec{x}$进行点乘，其中$x_{i,j} = 1$的位置替换为$E(v_i)$，$x_{i,j} = 0$的位置替换为空，然后累加得到每一个节点位于索引的数据属性密文$E(\overrightarrow{x_a})$；然后利用全同态密-密比较算法得到$b \in \{E(1), E(0)\}$。

[0053] 在步骤 S150 中，基于所述密文决策树模型所在层的待比较的内部节点的比较结果，计算该层节点的左右孩子节点的节点值密文。

[0054] 计算该层节点的左右孩子节点的节点值是为了下一层选择待比较的内部节点做准备。

[0055] 具体地，对于根节点所在层，如图 8 所示，根据比较结果b_0，赋予左孩子节点或右孩子节点的节点值为$E(1)$或者$E(0)$，对于其他节点所在层，如图 9 所示，根据比较结果b_l，父节点值为c_i的左右孩子节点值分别赋予$b_l \& c_i, !b_l \& c_i$，其中&代表 AND 操作，!代表 NOT 操作，使得下一层所有节点的节点值只有一个为$E(1)$，其余为$E(0)$。

[0056] 在步骤 S160 中，根据叶子节点的节点值密文与分类标签值密文，计算出推理结果密文。

[0057] 根据叶子节点的节点值密文与分类标签值密文，计算推理结果密文的目的是发送节点值为$E(1)$的叶子节点的分类标签值密文。

[0058] 具体地，如图 10 所示，所述密态决策树模型叶子节点所在层的节点值只有一个为$E(1)$，其余为$E(0)$，与节点的分类标签值密文进行点乘后累加得到节点值为$E(1)$的所在节点的分类标签值密文，作为所述密态决策树模型在所述数据明文上推理的分类结果密文。

[0059] 在步骤 S170 中，基于所述推理结果密文，随机数掩盖得到分类结果密文，使用全同态解密算法对所述分类结果密文进行解密得到分类结果，根据分类结果和随机数，得到推理结果，完成决策树推理过程。

[0060] 使用随机数掩盖推理结果密文得到分类结果密文，再从分类结果到推理结果的目的是数据方得到推理结果，而且确保模型方只进行了解密，而没有推理结果。

[0061] 具体地，对于推理结果密文，随机取反得到一个使用随机数掩盖的推理结果的密文作为分类结果密文，其中取反位置为该随机数，使用全同态秘钥对分类结果密文进行解密得到随机数掩盖的推理结果，根据取反位置得到推理结果。

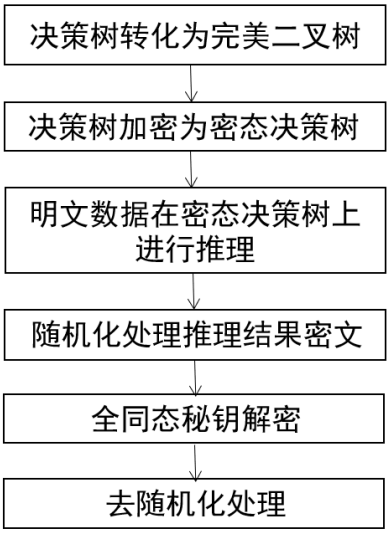

决策树转化为完美二叉树
决策树加密为密态决策树
明文数据在密态决策树上进行推理
随机化处理推理结果密文
全同态秘钥解密
去随机化处理

图 1

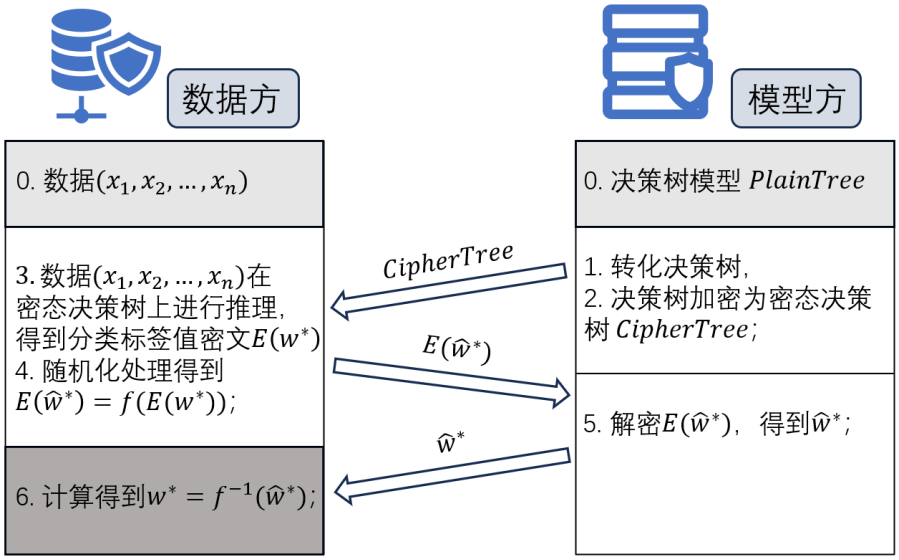

数据方
模型方
0. 数据$(x_1, x_2, \dots, x_n)$
0. 决策树模型 $PlainTree$
3. 数据$(x_1, x_2, \dots, x_n)$在密态决策树上进行推理，得到分类标签值密文$E(w^*)$
4. 随机化处理得到$E(\hat{w}^*) = f(E(w^*))$;
1. 转化决策树，
2. 决策树加密为密态决策树 $CipherTree$;
5. 解密$E(\hat{w}^*)$，得到$\hat{w}^*$;
6. 计算得到$w^* = f^{-1}(\hat{w}^*)$;
$CipherTree$
$E(\hat{w}^*)$
$\hat{w}^*$

图 2

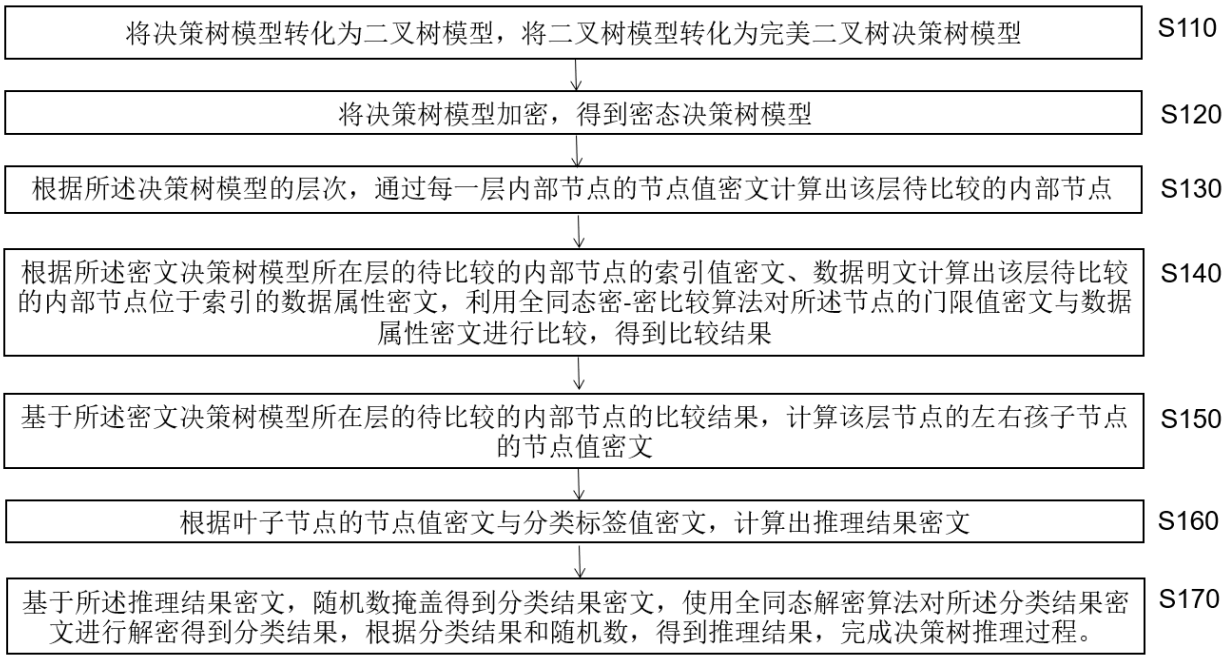

将决策树模型转化为二叉树模型，将二叉树模型转化为完美二叉树决策树模型
S110
将决策树模型加密，得到密态决策树模型
S120
根据所述决策树模型的层次，通过每一层内部节点的节点值密文计算出该层待比较的内部节点
S130
根据所述密文决策树模型所在层的待比较的内部节点的索引值密文、数据明文计算出该层待比较的内部节点位于索引的数据属性密文，利用全同态密-密比较算法对所述节点的门限值密文与数据属性密文进行比较，得到比较结果
S140
基于所述密文决策树模型所在层的待比较的内部节点的比较结果，计算该层节点的左右孩子节点的节点值密文
S150
根据叶子节点的节点值密文与分类标签值密文，计算出推理结果密文
S160
基于所述推理结果密文，随机数掩盖得到分类结果密文，使用全同态解密算法对所述分类结果密文进行解密得到分类结果，根据分类结果和随机数，得到推理结果，完成决策树推理过程。
S170

图 3

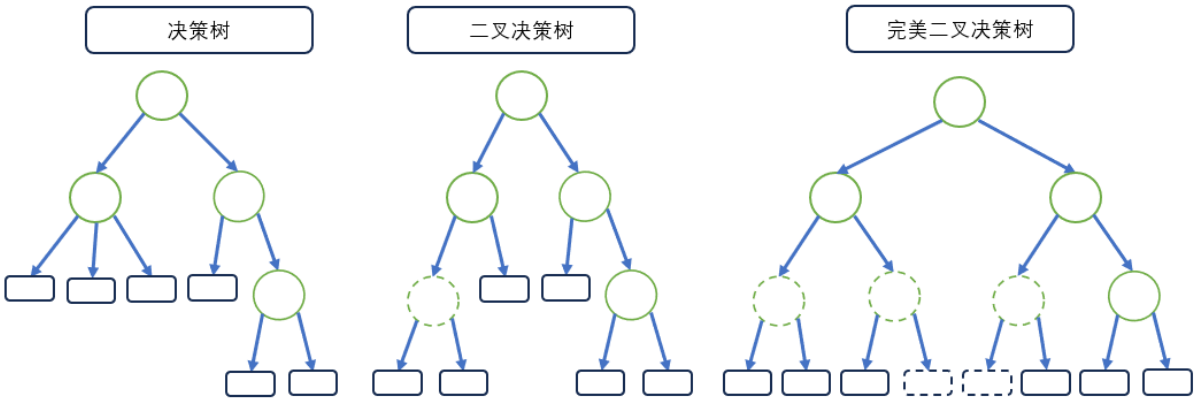

决策树
二叉决策树
完美二叉决策树

图 4

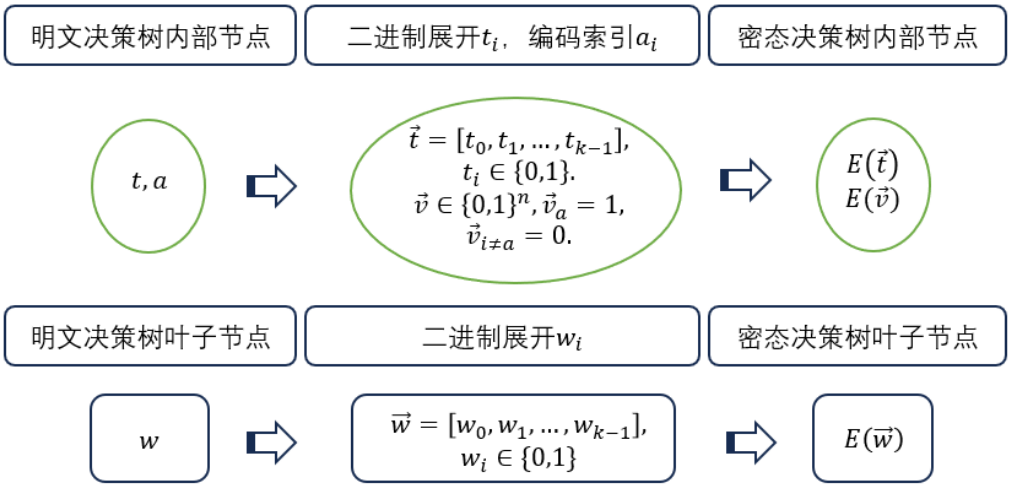

明文决策树内部节点
二进制展开t_i，编码索引a_i
密态决策树内部节点
t, a
$\vec{t} = [t_0, t_1, \ldots, t_{k-1}],$
$t_i \in \{0,1\}.$
$\vec{v} \in \{0,1\}^n, \vec{v}_a = 1,$
$\vec{v}_{i \neq a} = 0.$
$E(\vec{t})$
$E(\vec{v})$
明文决策树叶子节点
二进制展开w_i
密态决策树叶子节点
w
$\vec{w} = [w_0, w_1, \ldots, w_{k-1}],$
$w_i \in \{0,1\}$
$E(\vec{w})$

图 5

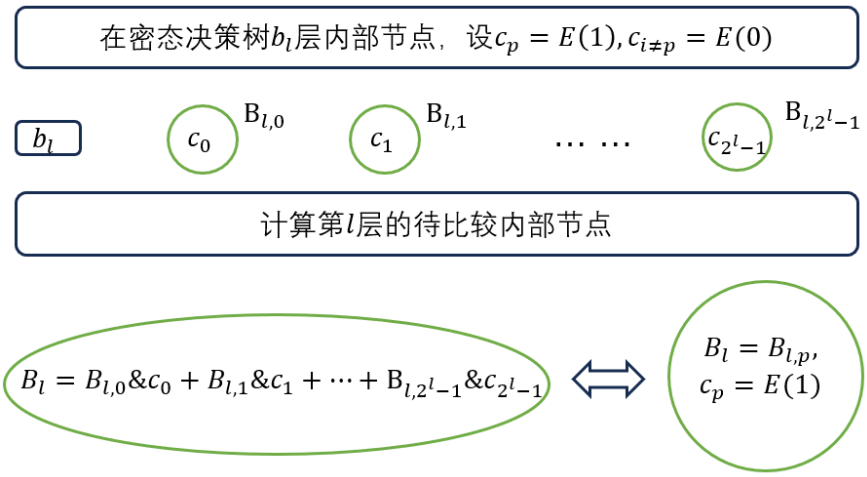

在密态决策树b_l层内部节点，设$c_p = E(1), c_{i\neq p} = E(0)$
b_l
c_0 $B_{l,0}$
c_1 $B_{l,1}$
… …
c_{2^l-1} $B_{l,2^l-1}$
计算第l层的待比较内部节点
$B_l = B_{l,0}\&c_0 + B_{l,1}\&c_1 + \cdots + B_{l,2^l-1}\&c_{2^l-1}$
$B_l = B_{l,p}$,
$c_p = E(1)$

图 6

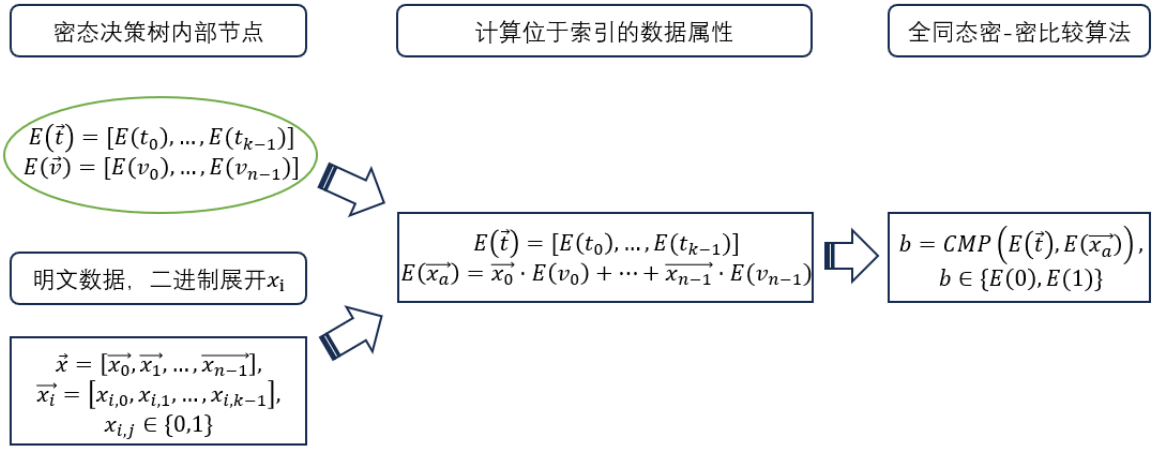

密态决策树内部节点
计算位于索引的数据属性
全同态密-密比较算法
$E(\vec{t}) = [E(t_0), \ldots, E(t_{k-1})]$
$E(\vec{v}) = [E(v_0), \ldots, E(v_{n-1})]$
明文数据，二进制展开x_i
$\vec{x} = [\overrightarrow{x_0}, \overrightarrow{x_1}, \ldots, \overrightarrow{x_{n-1}}]$,
$\overrightarrow{x_i} = [x_{i,0}, x_{i,1}, \ldots, x_{i,k-1}]$,
$x_{i,j} \in \{0,1\}$
$E(\vec{t}) = [E(t_0), \ldots, E(t_{k-1})]$
$E(\overrightarrow{x_a}) = \overrightarrow{x_0} \cdot E(v_0) + \cdots + \overrightarrow{x_{n-1}} \cdot E(v_{n-1})$
$b = CMP\left(E(\vec{t}), E(\overrightarrow{x_a})\right)$,
$b \in \{E(0), E(1)\}$

图 7

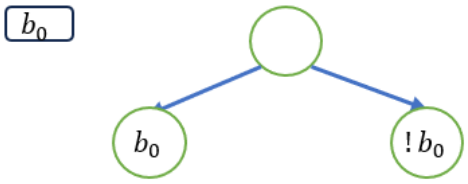

b_0
b_0
$!\,b_0$

图 8

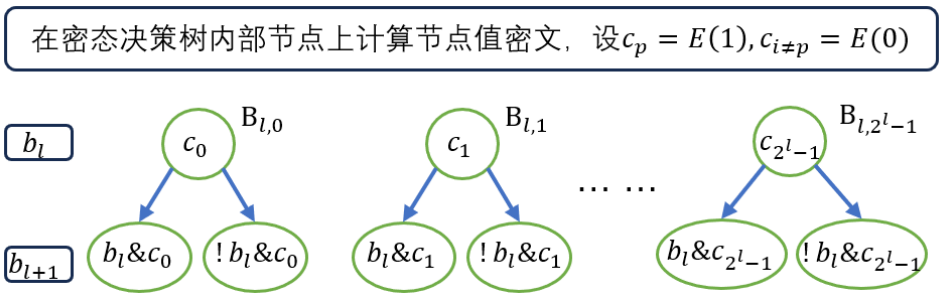

在密态决策树内部节点上计算节点值密文，设$c_p = E(1), c_{i\neq p} = E(0)$
b_l
b_{l+1}
c_0 $B_{l,0}$
c_1 $B_{l,1}$
… …
c_{2^l-1} $B_{l,2^l-1}$
$b_l\&c_0$
$!\,b_l\&c_0$
$b_l\&c_1$
$!\,b_l\&c_1$
$b_l\&c_{2^l-1}$
$!\,b_l\&c_{2^l-1}$

图 9

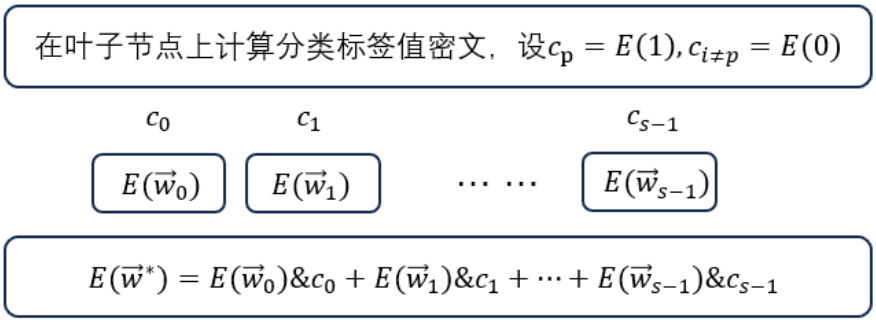

在叶子节点上计算分类标签值密文，设$c_{\mathrm{p}} = E(1), c_{i\neq p} = E(0)$
c_0
c_1
c_{s-1}
$E(\vec{w}_0)$
$E(\vec{w}_1)$
$\cdots\ \cdots$
$E(\vec{w}_{s-1})$
$E(\vec{w}^*) = E(\vec{w}_0)\&c_0 + E(\vec{w}_1)\&c_1 + \cdots + E(\vec{w}_{s-1})\&c_{s-1}$

图 10