云计算环境下朴素贝叶斯安全分类外包方案研究

陈 思
(南京理工大学信息化建设与管理处 江苏 南京 210094)

摘 要 当前基于大数据环境的机器学习模型训练和使用模式正饱受争议,尤其在用户针对已训练模型输入特征实例得到分类结果的模型使用阶段。一方面用户不愿意在使用过程中暴露自己的输入数据及最终结果,另一方面模型拥有者迫切需要将分类业务外包给云服务器,同时不暴露模型的明文参数。基于此应用场景,提出一种基于同态加密技术及盲化技术的朴素贝叶斯安全分类外包方法,并在云计算环境下实现仿真。整个系统允许模型拥有者加密上传模型,用户与云服务器利用同态性质完成安全多方计算。在多个朴素贝叶斯分类实例上进行仿真,结果表明该方案在不降低分类准确率的前提下实现了针对训练模型、输入数据及分类结果的隐私保护。

关键词 云计算环境 朴素贝叶斯 同态加密 外包

中图分类号 TP391.9 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2020.07.045

NAIVE BAYESIAN SECURITY CLASSIFICATION OUTSOURCING SCHEME IN CLOUD COMPUTING ENVIRONMENT

Chen Si
(*Division of Informationization Construction and Management, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, China*)

Abstract The current training and usage patterns of machine learning models based on big data environments are controversial, especially in the model using phase of the classification results obtained by inputting feature instances for trained models. On the one hand, users are reluctant to expose their input data and final results. On the other hand, model owners urgently need to outsource the classification business to the cloud server without exposing the plaintext parameters of the model. Based on this application scenario, this paper proposes a Naive Bayesian security classification outsourcing system based on homomorphic encryption technology and model blinding technology, and implements simulation in cloud computing environment. The entire system allows the model owner to upload the encrypted model, and the user together with the cloud server use the homomorphic nature to perform secure multiparty computing. We simulate on a number of Naive Bayesian classification instances. The results show that the scheme achieves privacy protection for training models, input data and classification results without reducing the classification accuracy.

Keywords Cloud computing environment Naive Bayes Homomorphic encryption Outsourcing

0 引 言

人类、机器、物理世界三元的高度融合引发了数据规模的急速式增长和数据模式的高度复杂化,我们已进入了大数据时代[1]。与此同时,在处理、分析海量数据方面表现良好的机器学习已经成为人工智能领域中的一个重要分支。现机器学习算法主要包括分类、聚类算法,其中分类算法主要包括支持向量机(SVM)、朴素贝叶斯、决策树等,甚至部分神经网络也能完成分类任务。在一个机器学习实例中,主要包括模型训练和模型使用两个阶段,模型训练即利用本地训练集完成模型初始化并进行参数优化,模型使用则是利用训练完成的模型,通过输入特征向量得到分类预测结果。

收稿日期:2019-04-23。国家自然科学基金项目(61572255);信息安全国家重点实验室开放项目(2017-ZD-01)。**陈思**,博士生,主研领域:网络与信息系统建设,大数据隐私保护。

无论是机器学习模型训练还是模型使用都可以看作是一种特殊的计算,且伴随着应用场景的拓展及使用数据量的扩大,这些计算的规模也会剧增,这对于一些本地资源受限的个人用户来说,很难高质量地独立完成。为此我们需要借助云服务器的计算存储能力,实现机器学习模型训练和模型使用的外包。

云计算是一种基于互联网的计算方式,通过这种方式,共享的软硬件资源和信息可以按需求提供给计算机和其他设备,云是网络、互联网的一种比喻说法[2]。外包计算是云计算中最重要的应用之一,指的是一个计算能力有限的客户将任务外包给云中的一个或者多个服务器[3]。这一基于云计算的应用场景正好契合由大规模数据驱动的机器学习模型训练及预测任务。当结合外包计算技术和机器学习实例时,我们常需要借助云服务器的计算存储能力代替用户的本地计算与存储,而这一过程必须考虑数据安全性问题。第一,用于训练机器学习模型的训练数据常常包含其大量个人隐私信息,因此其持有者不希望将明文暴露给其他人;第二,基于云服务器的外包环境具有不确定性,在外包方案中常以半可信状态假设(会推测敏感信息),即云服务器会正确执行用户预设的计算任务,但其会想方设法推测用户的隐私信息而不易被察觉。

在模型训练阶段,主要考虑训练数据集的数据安全问题,例如以疾病预测为目标的机器学习医疗系统,常用的训练数据集是与病患直接关联的诊断数据,极有可能带有大量的隐私信息,这类数据在训练外包过程中往往不能以明文的形式传递。在模型使用阶段,用户通常注重输入实例及最终分类结果的隐私保护,而对于提供分类服务的模型拥有者,由于高精度高鲁棒性的模型常需要大量人力、物力进行训练,因此执行分类外包时模型的明文不能直接发布给远程服务器。如何在保证模型训练及模型使用可行的情况下完成关键数据隔离,是机器学习安全外包主要的研究内容,也是亟需解决的关键问题。

为了解决这个问题,文献[4]针对 SVM、朴素贝叶斯和决策树三种分类器的模型训练阶段,通过与 AdaBoost 迭代算法结合,构建弱分类器并最终组合的方式提出训练过程外包模型,但没有考虑模型使用阶段的应对方法。文献[5]针对支持向量机分类器模型训练及模型使用阶段的安全问题设计了安全外包方案,但该方案不能很好地迁移到朴素贝叶斯分类外包方案中。文献[6]利用差分隐私的思想解决多数据源情况下的朴素贝叶斯模型训练外包的场景,也没有考虑模型使用阶段的外包安全。文献[7]针对朴素贝叶斯分类外包任务,利用私有信息检索技术及同态加密技术实现安全外包,但该方案时空开销较大,不能很好地实际部署与应用。

现如今大多数朴素贝叶斯外包方案多注重模型训练阶段的安全外包,尚没有人考虑模型预测阶段的安全外包。本文首次针对朴素贝叶斯模型分类预测场景,设计实现了一套基于同态加密技术及盲化技术的朴素贝叶斯安全分类外包系统,该系统借助云服务器高效计算存储能力以及随时在线提供分类预测服务的特性,实现了模型的高效、准确分类外包。本文创新点如下:

(1) 允许模型拥有者在本地不限编程语言地训练朴素贝叶斯分类模型,将模型加密委托给半可信的云服务器后可以离线,之后的分类任务将不需要模型拥有者的参与;

(2) 首次针对朴素贝叶斯分类预测阶段,利用同态加密方法及盲化性质设计了朴素贝叶斯安全分类外包方案并基于 Java 编程语言实现了系统可视化;

(3) 允许用户登陆安全外包系统与云服务器进行交互并在确保隐私的情况下得到安全可靠的分类结果,且支持各类朴素贝叶斯分类实例。

1 朴素贝叶斯安全分类外包原理

1.1 朴素贝叶斯分类器及分类外包

朴素贝叶斯分类是分类预测样本标签的有效算法,朴素贝叶斯分类器的思想原理很简单:给出待分类样本,求出该样本属于某个类别的后验概率,哪个概率最大,就认为此样本属于哪个类别。简要描述朴素贝叶斯分类过程如下:

(1) 设特征向量 $\boldsymbol{x}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_d\}$ 为一个待分类项,每一个 $\boldsymbol{x}_i$ 代表 $\boldsymbol{x}$ 的一个特征属性。

(2) 有类别 $y=\{y_1,y_2,\cdots,y_n\}$。

(3) 计算 $P(y_1\mid\boldsymbol{x})$、$P(y_2\mid\boldsymbol{x})$、$\cdots$、$P(y_n\mid\boldsymbol{x})$,即 $\boldsymbol{x}$ 属于每个类的后验概率。

(4) 若 $P(y_k\mid\boldsymbol{x})=\max\{P(y_1\mid\boldsymbol{x}),\cdots,P(y_n\mid\boldsymbol{x})\}$,就认为 $\boldsymbol{x}$ 属于第 k 类。

现在主要目标就是求得第 3 步的各后验概率,并依据其中最大的值判断分类结果。分类器以 $\boldsymbol{X}$ 作为输入,分类函数计算 $i_0\leftarrow\arg\max_{i\in[n]}P(Y=y_i,X=\boldsymbol{x})$,假设每个特征之间的属性都是独立的,不相互产生影响,根据贝叶斯公式可以得到:

$$P(y \mid \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n) = \frac{P(\boldsymbol{x}_1 \mid y)P(\boldsymbol{x}_2 \mid y)\cdots P(\boldsymbol{x}_n \mid y)P(y)}{P(\boldsymbol{x}_1)P(\boldsymbol{x}_2)\cdots P(\boldsymbol{x}_n)}$$

由于分母一般都是常量,则 X 分到某个类的最大后验概率 $P(Y=y_i, X=\boldsymbol{x})$ 等于 $P(Y=y_i)\prod_{j=1}^{d} P(X_j = x_j) \mid Y=y_i)$,计算得到最大的后验概率 P 同样可以表达成对数形式 $\log P(Y=y_i) + \sum_{j=1}^{d} \log P(X_j = \boldsymbol{x}_j \mid Y = y_i)$。取对数的意义是乘法运算转化为了加法运算,避免了乘法计算结果下溢的问题。我们用已知变量 y 的所有可能值来计算后验概率,选择最大的概率值作为分类结果。因此,朴素贝叶斯分类器的模型 W 可以由下列概率的集合组成:

先验概率:$\{P(Y=y_1), P(Y=y_2), \cdots, P(Y=y_n)\}$,其中第 i 个元素表示 $\boldsymbol{x}$ 属于 y_i 类的概率。

类条件概率:$\{P(X_j = v \mid Y = y_i)\}$,它表示 $\boldsymbol{x}$ 属于 y_i 类时,$\boldsymbol{x}$ 的第 j 个分量为 v 的概率,其中 v 属于 X_j 的值域 S_j,$i \in [n]$,$j \in [d]$。

当有其他用户需要借助训练好的朴素贝叶斯模型进行分类预测问题时,一方面,模型拥有者不希望直接暴露模型明文供其他人任意使用,另一方面,用户不愿暴露自己待预测的特征向量以及最终分类结果。另外,模型拥有者也无法始终保持在线以及同时应对大量用户的预测任务,因此我们考虑基于云计算环境实现朴素贝叶斯分类模型的分类预测外包任务。借助云服务器的计算存储能力,将训练好的模型上传至云服务器后模型拥有者可以离线,使用模型的大量用户可以同时与云服务器交互,输入其特征向量 $\boldsymbol{x}$,服务器结合模型 W 返回具体分类结果 $Y_W(\boldsymbol{x})$。这一过程看似简单,实际上需要考虑多方的数据安全。当引入第三方云服务器代替模型拥有者参与计算时,外包方案需要保证上传的模型以及用户输入的特征向量对云服务器不可见。这一过程中,我们选择引入同态加密这一常用的密码学方案,允许云服务器在密文条件下实现正确分类。

1.2　同态加密技术

同态加密使明文的计算能够在相应密文上执行,且不暴露明文信息。一个非对称同态加密方案 $AH\varepsilon$ 支持一般加法及数乘的密文操作。给定两个使用了同一公钥加密的消息 $AH\varepsilon.Enc(a)$ 和 $AH\varepsilon.Enc(b)$,存在一个加法操作 $\oplus$ 使得 $AH\varepsilon.Enc(a) \oplus AH\varepsilon.Enc(b)$ 的结果解密就是明文 $a+b$ 的结果。$AH\varepsilon.Enc(a)$ 表示明文 a 加密的结果,c 是一个固定的值,数乘 $AH\varepsilon.Enc(ca)$ 满足以下等式:

$$AH\varepsilon.Enc(a) \odot \cdots \odot AH\varepsilon.Enc(a) = AH\varepsilon.Enc(a)^c$$

由于效率问题,本文使用 Paillier 同态加密系统,简单介绍如下:

(1) 随机选取两个素数 p 和 q,满足 $\gcd(pq, (p-1)(q-1)) = 1$,$\gcd(\cdot)$ 求取最大公约数。

(2) 计算 $n = pq$ 和 $\lambda = \mathrm{lcm}(p-1, q-1)$,$\mathrm{lcm}(\cdot)$ 求取最小公倍数。

(3) 保证 $u = (L(g^{\lambda} \bmod n^2))^{-1} \bmod n$ 存在,其中 $L = \frac{u-1}{n}$,随机选取 $g \in Z_{n^2}^*$。

(4) 公钥为 (n, g),私钥为 (λ, u)。

(5) 加密:选择一个随机数 $r \in (0, n-1]$,密文即为 $c = g^m \cdot r^n \bmod n^2$。

(6) 解密:计算 $m = L(c^{\lambda} \bmod n^2) \cdot u \bmod n$,验证 $m < n^2$ 通过,则 m 为解密后的明文。

作为一种加密工具,同态加密是云计算外包领域最常使用的方法之一。一方面,其同态性质允许我们针对不同的计算场景设计对应的计算方案并确保计算结果的正确性;另一方面,在不知道解密密钥的情况下,加密数据的安全性有严格的保障。本文系统通过引入两套同态加密系统并结合其他的一些密码学工具设计、实现了针对模型、特征向量以及分类结果安全性的保障,真正意义上实现了安全外包。

2　朴素贝叶斯安全分类外包方案

本文安全外包方案包含模型拥有者、远程服务器、用户(模型使用者)三方实体,方案的整体结构如图 1 所示。

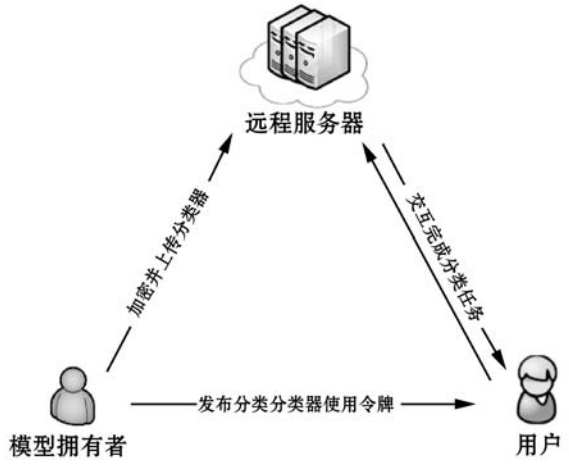

图 1　朴素贝叶斯安全分类外包方案总体结构

2.1　用户本地训练并加密上传模型

由于 Paillier 加密系统只能对整数进行操作,为了

Paillier 加密系统能对模型 W 进行加密,我们通过乘以一个事先给定的大整数 L 的方式把浮点数表示的概率转化为一个整数(向上取整),转化得到的整数在 Z_N 域中,即 Paillier 系统的明文空间。整个外包方案使用到两套 Paillier 同态加密系统 $<P_1,P_2>$,N_1、N_2 分别是 P_1、P_2 的模数,应保证 $N_1<N_2$。P_1 用来加密模型 W 以保证其隐私性,P_2 用于保证最终结果的隐私性,生成两对公私钥 $<pk_1,sk_1>$、$<pk_2,sk_2>$,公钥 pk_1、pk_2 对外公布,私钥 sk_1 由模型拥有者持有且作为令牌传输给用户,私钥 sk_2 由云服务器保管。

记 x_j 的值域为 S_j,其中 x_j 为实例向量的第 j 个属性。为了在朴素贝叶斯分类器模型 W 上使用 Paillier 加密,我们对提取出来的概率模型进行预处理,每种概率的对数用整数表示。

对每一个 $i\in[n]$,第 i 个先验概率为:

$$P(i)=\lceil L\log(KP(Y=y_i))\rceil$$

对每一个 $j\in[d]$,$v\in[S_j]$,类条件概率表示为:

$$P(i,j,v)=|L\log(KP(X_j=v\mid Y=y_i))|$$

式中:K 是一个用来把原始概率化为整数的整数,本文选取 $K=1\ 000\ 000$ 作为整数进行转换,并取以 1.002 为底的对数,最后取整以便使用 Paillier 系统对模型加密。经过测试,转换后的模型精度不变。Paillier 明文空间上最大正数的大小是 $l_{\max}=\log(d+1)(L+\log K)$,远小于 Paillier 系统的模数大小。因此,明文能够进行正确处理,而且分类结果不会受到影响。最终模型 W 可以表示成 $\{P(i),\{\{P(i,j,v)\}_{v\in S_j}\}_{j=1}^{d}\}_{i=1}^{n}$,用公钥 pk_1 加密模型 W 中的各个概率,得到密文概率 $\{E_1(P(i)),\{\{E_1(P(i,j,v))\}_{v\in S_j}\}_{j=1}^{d}\}_{i=1}^{n}$,以 <属性、密文概率> 的键值形式存储在 .txt 中上传至远程服务器。

2.2 用户与服务器交互得到加密后验概率

用户输入 x 的分类结果由分类函数 $\arg\max P(i)+\sum_{j=1}^{d}P(i,j,\boldsymbol{xj})$ 决定。我们需要把隐私保护分类问题转化为隐私保护下的 argmax 问题。用户应首先根据 $\boldsymbol{x}$ 从远程服务器得到相应先验概率和类条件概率,每一个分量 $\boldsymbol{x}_j$ 的条件概率都根据 $\boldsymbol{x}_j$ 的值存储在服务器端,我们需要设计算法在不暴露用户输入及整体模型的情况下求得加密条件下的后验概率。算法 1 详细描述了本文的解决方案,利用盲化技术保证了用户即使拥有第一套同态加密系统的私钥,解密得到的仍是盲化后的数据,并不会暴露模型明文。盲化技术是外包计算领域的关键技术,盲化过程包括两步:生成盲化因子及解除盲化因子,利用同态性质添加盲化及解除盲化。最终得到的加密条件下的后验概率将作为加密数据求取最大值算法的输入。

算法 1 加密后验概率求取算法

用户 **U** 输入:特征向量 $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_d)$,私钥 sk_1,公钥 pk_1、pk_2

远程服务器 **RS** 输入:加密模型 $\{E_1(P(i))\}$ 和 $\{E_1(P(i,j,v))\}$,私钥 sk_2,公钥 pk_1、pk_2

输出:$\{E_2(P_i)\}=\{E_2(P(Y=y_i,X=\boldsymbol{x}))\}$

for $i\in[n]$:

RS:从 Z_{N1} 上选择盲化因子 $O_{i,0}$

for $j\in[d]$:

RS:从 Z_{N1} 上选择盲化因子 $O_{i,j}$

满足 $O_{i,0}+\sum_{j=1}^{d}O_{i,j}<N_1-2^{l\max}(\bmod N_1)$

end for

end for

for $i\in[n]$:

RS:盲化 $E_1(P'(i))=E_1(P(i))\oplus E_1(O_{i,0})$

for $j\in[d]$,$v\in[S_j]$:

RS:盲化 $E_1(P'(i,j,v))=E_1(P(i,j,v))\oplus E_1(O_{i,j})$

end for

end for

RS:计算 $O_i=-O_{i,0}-\sum_{j=1}^{d}O_{i,j}\bmod N_1$

用 pk_2 加密 O_i-N_1 为 $E_2(O_i-N_1)$

发送 $E_1(P'(i))$、$E_1(P'(i,j,x_j))$ 和 $E_2(O_i-N_1)$ 给 U

U:解密 $E_1(P'(i))$ 得到 $P'(i)$

解密 $E_1(P'(i,j,x_j))$ 得到 $P'(i,j,x_j)$

for $i\in[n]$:

U:根据特征向量找到对应盲化后的先验概率及类条件概率,累加得到:$P'_i=(P'(i)+\sum_{j=1}^{d}P'(i,j,\boldsymbol{xj}))\bmod N_1$

使用 pk_2 加密 P'_i,并利用同态性质解除盲化

$E_2(P_i)=E_2(P'_i)\oplus E_2(O_i-N_1)$

end for

输出加密后验概率 $\{E_2(P_i)\}$

2.3 密文后验概率中求取最终分类结果

通过算法 1 我们已将朴素贝叶斯安全分类问题转化为密文条件下的 argmax 问题,只需要与云服务器交互得到满足 $\arg\max\{E_2\{P_i\}\}$ 的 i 即可。但在这一步骤中不能直接将 $E_2\{P_i\}$ 发送给服务器,需要保证拥有解密私钥的远程服务器对最终分类结果不可见,因此我们将再次利用同态性质及盲化技术实现安全双方计算,具体流程如算法 2 所示。

算法 2 加密数据求取最大值算法

用户 **U** 输入:n 个加密后验概率 $\{E_2\{P_i\}\}$,公钥 pk_2

远程服务器 **RS** 输入:私钥 sk_2

输出:$i \leftarrow \arg\max\{E_2\{P_i\}\}$ **U**:在$[n]$上选择一个随机排列$\pi(\cdot)$

将$\{E_2\{P_i\}\}$按随机排列顺序存放(第i个元素放在$\pi(\cdot)$中i所在的位置)

重排密文加入比较队列

队列长度$k=n$

While$(k>1)$:

for $i \in [k]$:

从队列中依次选取元素$E_2\{P_k\}$和$E_2\{P_{k+1}\}$,在Z_{N^2}上随机选择一个整数r,计算

$$E_2(P'_k)=E_2(P_k)\oplus E_2(r)$$

$$E_2(P'_{k+1})=E_2(P_{k+1})\oplus E_2(r)$$

将$E_2(P'_k)$和$E_2(P'_{k+1})$发给 **RS**

RS:解密得到明文P'_k和P'_{k+1}返回比较结果

U 将较大元素的密文原文加入新比较队列,重新统计队列长度

end for

当比较队列中只有一个元素时停止交互,得到该元素所在随机排列中的初始数i

输出:最终分类结果i

3　系统可视化仿真

本文采用面向对象的 JavaScript 网络脚本语言,内嵌基于 Java 编程的具体算法,实现了朴素贝叶斯安全分类外包功能及可视化。为了仿真基于云服务器的外包计算环境,我们以多台 8 GB 内存的 Lenovo y400 笔记本及一台 64 GB 内存、Intel(R) Xeon(R) CPU E5-2640 2.60 GHz 处理器 Windows 8 操作系统的 ThinkServer 服务器共同组成交互式外包分类环境,借助 ThinkServer 强大的计算存储能力,实现了单服务器与多用户的一对多访问模式,允许模型拥有者本地训练并加密上传朴素贝叶斯模型至服务器,同时允许多个用户登录系统同时加密访问服务器使用模型完成分类。

本文系统主界面如图 2 所示,用户注册登录后可以查看使用已发布模型,模型拥有者在图 3 模型发布界面可以添加模型描述,加密上传模型。图 4 展示了用户使用模型得到分类预测结果的界面。为了证明本文外包方案适用于多种不同的朴素贝叶斯分类实例,我们选取了三类可使用朴素贝叶斯分类器进行分类预测的数据集:(1) 鸢尾花数据集:通过花萼长度、花萼宽度、花瓣长度、花瓣宽度 4 个属性预测花卉属于 3 个种类中哪一类;(2) 红酒数据集:根据红酒的 13 种成分判断其属于 3 个种类中哪一类;(3) 巴赫和弦数据集:根据 14 个和弦属性判断其具体为巴赫哪一个作品。

图 2　系统主界面图

图 3　模型上传界面图

图 4　分类结果显示界面图

本文利用十折交叉验证法对三类数据集分别进行分类预测统计,对比经由密文安全外包分类及明文直接分类两种方式的模型平均准确率,比较结果如表 1 所示。可以看出,本文外包方案在确保模型隐私、用户输入向量及分类结果隐私的前提下并没有损失过多模型分类准确率(小于 1%),且没有影响模型的实用性。分析与明文下测试的模型准确率的差距,主要来源于利用 Paillier 同态加密方法时,为确保加密成功需要有小数转换为整数的放缩过程,这一过程中的精度损失可以通过调整大整数 K 的大小进行控制。另外,本文系统允许模型拥有者上传加密模型后离线,服务器的高吞吐量也极大提高了多用户同时进行分类预测的效率。

表 1 模型准确率对比

数据集	分类准确率/%	
	明文分类	安全外包分类
鸢尾花(iris)	90.12	89.96
红酒(wine)	76.22	75.76
巴赫和弦(Bach Harmony)	66.97	66.82

4 结 语

随着云计算的快速发展,外包计算和机器分类学习的结合得到了广泛的关注和应用,现迫切需要设计和实现考虑用户隐私及模型安全的外包分类系统。本文提出的朴素贝叶斯分类系统由分类器模型拥有者、远程服务器以及用户组成,模型拥有者通过远程服务器为用户提供分类服务。本文仿真实现了服务器和用户的交互,得到了相应的分类结果,并对该系统中各个实体所关心的隐私数据通过同态加密或盲化技术实现隔离。结果表明,本文系统在不影响模型分类准确率的情况下实现了对分类器模型隐私性、用户特征向量及分类结果隐私性的保护。

参 考 文 献

[1] 王元卓,靳小龙,程学旗. 网络大数据:现状与展望[J]. 计算机学报,2013,36(6):1125 - 1138.

[2] 杨延嵩,张宁,郑举,等. 基于云计算的呼叫中心系统应用研究[J]. 计算机科学,2012,39(z2):119 - 122.

[3] 陈振华,李顺东,黄琼,等. 云外包计算中空间位置关系的保密判定[J]. 计算机学报,2017,40(2):351 - 363.

[4] Bost R,Popa R A,Tu S, et al. Machine learning classification over encrypted data[C]//Network and Distributed System Security Symposium, 2015.

[5] Liu H, Ning H S, Xiong Q X, et al. Shared authority based privacy-preserving authentication protocol in cloud computing [J]. IEEE Transactions on Parallel & Distributed Systems, 2015, 26(1):241 - 251.

[6] Li T, Li J, Liu Z L, et al. Differentially private naive bayes learning over multiple data sources[J]. Information Sciences, 2018,444:89 - 104.

[7] Li T, Huang Z G, Li P, et al. Outsourced privacy-preserving classification service over encrypted data[J]. Journal of Network and Computer Applications, 2018,106:100 - 110.

[8] Li L C, Lu R X, Huang C. EPLQ: efficient privacy-preserving location-based query over outsourced encrypted data [J]. Internet of Things Journal, IEEE, 2016, 3(2):206 - 218.

[9] Alberto H C, Gines D T, Felix G M, et al. Resolving privacy-preserving relationships over outsourced encrypted data storages[J]. International Journal of Information Security, 2016, 15(2):195 - 209.

(上接第 215 页)

3 结 语

广义回归神经网络因其善于处理非线性问题已经应用在了诸多领域。为解决传统测距定位模型因外界干扰而产生的非线性问题,同时为改善 GRNN 的平滑参数需要人为选择的问题,本文提出一种 MFOA-GRNN 三维定位模型。通过仿真可以看出,MFOA-GRNN 定位模型与 FOA-GRNN 定位模型和 PSO-BP 定位模型相比定位误差更低,整体性能更好,定位误差在 30 mm 以内,基本满足定位要求,同时为以后的定位技术提供了一种新的方法。

参 考 文 献

[1] 高腾飞. 无线传感器网络节点室内定位技术研究[D]. 苏州:苏州大学,2016.

[2] 孙宝山,刘晟源,刘骁骁. 基于 HDV-Hop 的无线传感器网络大型室内定位算法研究[J]. 计算机应用与软件,2018,35(3):114 - 119,186.

[3] 张文安,陈国庆,杨旭升. UHF-RFID 环境下的移动机器人定位方法[J]. 控制与决策,2018,33(10):1807 - 1812.

[4] 段亚青,王华倩,乔学工. 基于测距和灰狼优化的无线传感器网络定位算法[J]. 传感技术学报,2018,31(12):1894 - 1899.

[5] 马振宇,张威,毕学军,等. 基于优化 PSO-BP 算法的软件缺陷预测模型[J]. 计算机工程与设计,2016,37(2):413 - 417.

[6] Specht D F. A genera regression neural network[J]. IEEE Transactions on Neural Networks,1991,2(6):568 - 576.

[7] 皮骏,马圣,张奇奇,等. 基于改进果蝇算法优化的 GRNN 航空发动机排气温度预测模型[J]. 航空动力学报,2019,34(1):8 - 17.

[8] Yuan X F,Dai X S,Zhao J Y. On a novel multi-swarm fruit fiy optimization algorithm and its application[J]. Applied Mathematics and Computation,2014,233(2):260 - 271.

[9] 李冬辉,尹海燕,郑博文. 基于 MFOA-GRNN 模型的年电力负荷预测[J]. 电网技术,2018,42(2):585 - 590.

[10] 王楚柯,陆安江,吴意乐. 自适应果蝇优化算法在 WSN 节点覆盖优化中的应用[J]. 微电子学与计算机,2019,36(2):11 - 15.

[11] 杨明,陈玲玲,尹忠科. 基于改进 ACFOA 的图像一维 OMP 稀疏分解[J]. 计算机应用与软件,2016,33(4):208 - 211,272.

[12] 洪波,刘龙,王涛. 修正型果蝇算法优化 GRNN 的大梁自动焊障碍预测[J]. 焊接学报,2017,38(1):73 - 76,132.