ICS 35.240.01
CCS L 80

T/CUPC

团　　体　　标　　准

T/CUPC XXX-XXXX

基于同态加密的隐私保护机器学习算法技术规范

Technical specification for privacy-preserving machine learning algorithms based on homomorphic encryption

（征求意见稿）

20XX-XX-XX 发布　　20XX-XX-XX 实施

中国城镇化促进会　发布

目　　次

前言 II
1 范围 1
2 规范性引用文件 1
3 术语和定义 1
4 缩略语 2
5 总体框架 2
6 同态加密基础算法选型与通用要求 3
7 基于同态加密的隐私保护机器学习训练算法要求 4
8 基于同态加密的隐私保护机器学习推理算法要求 5
9 全流程安全防护要求 5
10 性能与精度指标 6
11 实现与集成要求 6
12 测试方法 7
附录 A （规范性） 面向隐私保护机器学习场景的主流同态加密算法标准参数配置表 8

前　　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由天津仁爱学院提出。

本文件由中国城镇化促进会归口。

本文件起草单位：。

本文件主要起草人：×××

基于同态加密的隐私保护机器学习算法技术规范

1 范围

本文件规定了基于同态加密的隐私保护机器学习算法的系统总体框架、同态加密基础算法选型与通用要求、基于同态加密的隐私保护机器学习训练算法要求、基于同态加密的隐私保护机器学习推理算法要求、全流程安全防护要求、性能与精度指标、实现与集成要求及测试方法等要求。

本文件适用于：

——　从事基于同态加密的隐私保护机器学习算法、底层密码计算软件库、分布式密态计算平台研发开发工作的企业单位、科研事业单位；

——　面向金融、医疗、政务、互联网、工业互联网行业，开展密态机器学习业务系统方案设计、线上部署、长期运维管理工作的建设单位与运营单位；

——　具备第三方检测评估资质，开展同态加密隐私保护机器学习算法安全合规性、功能完整性、综合性能专项验证工作的检测机构。

本文件不适用于：

——　不采用任何密码隐私保护手段的明文机器学习算法系统、未结合同态加密技术的单纯联邦学习系统、仅依靠可信执行环境或者多方安全计算单一技术实现隐私保护的计算方案；

——　仅用于数据完整性校验的轻量化对称同态哈希组件，不支持完整数值矩阵运算的简易同态密码模块；

——　集成电路芯片底层同态加密硬件加速电路设计、芯片制造相关技术规范。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 25069-2022　信息安全技术 术语

GB/T 32915-2026　网络安全技术 二元序列随机性检测方法

GB/T 45230-2025　数据安全技术 机密计算通用框架

GB/T 46284-2025　人工智能 联邦学习技术规范

GM/Z 0001-2013　密码术语

YD/T 4690-2024　隐私计算 多方安全计算产品安全要求和测试方法

T/SHCCIA 002-2023　云计算中同态加密技术应用要求

3 术语和定义

GB/T 25069-2022、GM/Z 0001-2013 界定的以及下列术语和定义适用于本文件。

3.1

同态加密　homomorphic encryption；HE

具备同态运算特性的公钥密码算法，对密文执行加法或乘法后解密，所得明文结果等价于对原始明文执行相同运算；分为部分同态加密、层次全同态加密、全同态加密三类。

3.2

隐私保护机器学习 privacy-preserving machine learning；PPML

在不泄露原始明文数据、模型敏感参数的前提下，完成模型训练、预测推理的机器学习技术体系。

3.3

密态计算 ciphertext computation

不解密密文，借助同态求值算子直接完成数值运算、矩阵运算、神经网络前向传播计算的过程。

3.4

求值密钥 evaluation key；evk

同态加密专用辅助密钥，用于实现密文乘法、重线性化、模数切换、密文引导等复杂密态运算。

3.5

密文引导 ciphertext bootstrapping

全同态加密核心处理流程，对噪声溢出阈值的密文重新加密刷新噪声，延长密文可连续运算深度。

4 缩略语

下列缩略语适用于本文件。

PHE 部分同态加密（partially homomorphic encryption）

BFV BFV 整数全同态方案（brakerski-fan-vercauteren）

CKKS CKKS 浮点近似全同态方案（cheon-kim-kim-song）

FHEW 环域轻量化全同态方案（fully homomorphic encryption over the torus）

5 总体框架

5.1 参与方角色划分

基于同态加密的隐私保护机器学习完整计算系统包含四类独立基础参与主体，在满足安全约束的前提下允许多个角色合并部署至同一台物理服务器设备，四类参与主体完整定义如下：

a) 数据持有主体：原始业务样本数据、业务特征数据的所有权归属主体，完整执行原始数据明文预处理操作、本地加密转换操作，仅向远端计算主体传输加密完成后的密文数据，原始明文数据全程不离开本地安全管控边界；

b) 模型计算主体：部署同态加密求值运算引擎、机器学习标准算子模块，仅接收、处理各类密文数据，不持有任何解密私钥文件，完整承担密态模型离线训练、密态模型线上推理全部数值计算工作；

c) 密钥管理主体：统一完成同态加密公私密钥、求值密钥的生成、分发、定期轮换、安全销毁全流程管理工作，完整承担密钥操作行为安全审计工作；金融、政务等高安全等级业务场景要求密钥管理主体独立部署为第三方可信机构，不与数据持有主体、模型计算主体合并；

d) 结果使用主体：持有整套系统解密私钥文件，接收模型计算主体传输的密态计算输出结果，执行解密操作后开展业务数据分析工作，禁止依托解密完成的业务结果反向推导还原任意一条原始业务样本数据。

5.2 技术架构

整套密态机器学习系统自上而下划分为五层独立技术层级，不同层级之间通过标准化接口完成数据交互，五层技术层级完整说明如下：

a) 第一层为业务应用层级，承载全部机器学习业务计算任务，包含数据分类计算、数据回归计算、数据聚类计算、神经网络预测计算等各类业务功能；
b) 第二层为隐私保护机器学习算法调度层级，完整承担模型训练任务、模型推理任务流程编排、底层算子动态适配、全局隐私预算统一管控全部工作；
c) 第三层为同态加密密文计算层级，完整覆盖同态加密密钥全生命周期管理、原始数据加解密操作、密文求值运算、密文噪声引导刷新全部底层密码功能；
d) 第四层为基础密码支撑层级，集成哈希运算、数字签名、安全随机数生成、TLS 安全传输全套基础密码功能组件；
e) 第五层为硬件资源支撑层级，包含中央处理器、图形处理器密码加速硬件、可信安全硬件、分布式持久化存储设备全部物理资源。

5.3 安全模型基础假设条件

5.3.1 通用推荐、普通工业预测场景采用半诚实安全模型。
5.3.2 金融、医疗、政务高敏感场景应采用恶意安全模型，并配套零知识证明校验密文合法性、协议防篡改机制。
5.3.3 底层格密码应满足标准格困难假设，安全强度不应低于 128 bit。

6 同态加密基础算法选型与通用要求

6.1 同态加密算法分类与场景选型

技术人员应当依据机器学习业务使用的数值数据类型、模型计算允许的连续乘法运算深度完成标准化算法选型工作，严禁无技术依据选用安全强度不满足标准要求的自研密码算法。完整场景选型约束如下：

a) 业务数据为整数离散特征、业务模型为浅层线性模型（线性回归模型、逻辑回归模型）的业务场景，优先选用 BFV 全同态加密算法方案开展系统开发工作；技术人员可选择 Paillier 部分加法同态加密算法作为备选技术方案；
b) 业务数据为浮点连续特征、业务模型为深度神经网络，需要开展完整线上推理、离线训练计算的业务场景，强制选用 CKKS 浮点近似全同态加密算法方案，该算法原生支持浮点数数值处理、支持非线性激活函数多项式近似计算工作；
c) 业务场景为移动端、嵌入式边缘设备轻量化线上推理，神经网络总层数不超过三层的业务场景，技术人员可选用 FHEW 轻量化全同态加密算法方案开展系统开发工作；
d) 强制性禁止条款：安全强度低于 128 bit、未经过商用密码标准化验证的自主研发同态加密算法不允许直接应用于正式投产运行的隐私保护机器学习业务系统。

6.2 密钥生成与分发

6.2.1 密钥生成全部约束条款

6.2.1.1 同态加密公钥文件、私钥文件、求值密钥文件全部依托符合安全要求的密码安全随机数发生器生成，随机数输出序列应完整通过 GB/T 32915-2026 规定的全部二元序列随机性检测项目。
6.2.1.2 密钥对应的多项式环维度、模数链参数完整匹配本文件附录 A 规定的标准化配置参数，整套密码体系安全强度数值最低保证 128 bit。

6.2.1.3　解密私钥文件仅允许存储于数据持有主体、结果使用主体本地安全存储区域，模型计算主体服务器内部禁止留存任何私钥文件。

6.2.2　密钥分发全部约束条款

6.2.2.1　公钥文件、求值密钥文件全部依托 TLS 一点三加密传输协议完成跨主体传输操作，配套数字签名机制完成文件防篡改校验工作。
6.2.2.2　多机构联合分布式模型训练业务场景中，每一家参与的数据持有主体独立生成自身完整密钥套件，不同机构之间禁止共享解密私钥文件。
6.2.2.3　密钥定期轮换周期硬性规定：正式投产运行的业务系统密钥轮换周期最长不得超过 90 d；金融、政务等高安全等级业务场景密钥轮换周期最长不得超过 30 d。

6.3　明文空间与密文空间参数设置规范

6.3.1　CKKS 方案多项式次数不应小于 4096，缩放系数、模数链长度匹配模型精度需求。
6.3.2　BFV 方案明文模数匹配特征取值范围，模数链预留不少于 5 层乘法噪声预算。
6.3.3　密文累积噪声不应超出模数阈值，防止解密失真与隐私泄露。

6.4　密文求值运算通用执行规则

6.4.1　系统应完整实现密文加法、密文乘法、常数乘、数值平移基础算子。
6.4.2　单次前向传播密文乘法总深度不应超过模数链承载上限，超出时应执行密文引导。
6.4.3　每次密文乘法完成后应执行重线性化，控制密文尺寸、抑制噪声累积。

6.5　密文解密与计算结果数值还原

6.5.1　仅持有私钥主体可执行解密，全部解密操作应留存完整不可篡改审计日志。
6.5.2　CKKS 浮点密文解密后应依据缩放系数完成数值校正，消除多项式近似误差。
6.5.3　批量输出结果应添加差分隐私噪声，全流程累计隐私预算ε不应大于 8。

7　基于同态加密的隐私保护机器学习训练算法要求

7.1　原始数据预处理与密态转换完整规范

7.1.1　标准化、归一化、缺失值填充应在数据持有方本地明文环境完成。
7.1.2　特征数值映射至明文合法区间后方可加密，原始明文特征不应传输至计算方预处理。
7.1.3　类别特征独热编码应本地完成，仅加密编码向量向外传输。

7.2　线性机器学习模型同态加密训练算法

7.2.1　梯度下降矩阵乘、梯度聚合、参数迭代更新应全程密态执行，不解密。
7.2.2　多方联合训练梯度密文求和流程宜参照 GB/T 46284-2025 相关条款执行，不应传输明文梯度。
7.2.3　损失收敛判定仅可由结果使用方解密校验，损失明文不应下发至计算方。

7.3　人工神经网络同态加密训练算法

7.3.1　全连接、卷积线性计算复用密文矩阵乘法算子。
7.3.2　ReLU、Sigmoid、Tanh 等激活函数应采用不高于 15 阶切比雪夫多项式近似，不应明文计算激活。
7.3.3　池化层采用密文求和、最大值多项式近似，严控噪声累积。
7.3.4　反向传播梯度全程密态，梯度裁剪在加密前本地明文执行。

7.4 分布式训练场景下模型梯度密态聚合执行规则

7.4.1 多方梯度聚合仅采用密文加法模式，多方交互安全应符合 YD/T 4690-2024 规定。
7.4.2 梯度裁剪阈值本地预设，抑制异常梯度放大噪声、泄漏样本特征。
7.4.3 多轮迭代过程中定期执行密文引导，保障解密精度满足业务需求。

7.5 完成训练后的模型参数加密存储管理规范

7.5.1 训练完成的权重、偏置持久化存储应保持密文形态；明文模型仅可临时存放于结果使用方本地。
7.5.2 密文模型应配套哈希校验值存储，防止篡改与投毒。

8 基于同态加密的隐私保护机器学习推理算法要求

8.1 推理阶段原始输入数据密态预处理要求

用户原始输入本地预处理加密，特征维度、缩放系数应与训练阶段完全一致；维度不匹配时系统应拒绝推理任务。

8.2 神经网络线性层密态前向计算完整规范

卷积层、全连接层矩阵乘法计算统一复用同态加密密文乘法求值算子完成；批量多样本同步推理业务场景采用密文打包技术降低系统跨主体数据传输通信开销。

8.3 非线性激活函数近似同态计算技术约束

8.3.1 ReLU 采用分段一次多项式；Sigmoid、Tanh 采用 5 阶～15 阶切比雪夫多项式。
8.3.2 多项式系数预编码为明文常数，降低乘法深度。
8.3.3 单层激活近似计算误差不应大于 0.01。

8.4 批量归一化模块密态计算适配执行规则

批量归一化模块对应的均值参数、方差参数、缩放参数预先加密为常数密文文件存储；线上推理前向计算流程依托密文线性变换算子完成归一化数值计算，全程不解密归一化统计参数。

8.5 推理计算结果解密操作与隐私脱敏处理要求

8.5.1 单条业务样本独立推理计算任务可直接执行解密操作输出完整数值结果；一次性批量输出多条业务样本推理结果时，系统统一添加差分隐私噪声完成脱敏处理。
8.5.2 系统禁止输出能够反向定位至单一业务用户的细粒度数值结果；聚类统计、批量汇总类业务计算结果完成多样本聚合操作之后再执行解密流程。

9 全流程安全防护要求

9.1 数据隐私安全要求

9.1.1 原始明文数据留存于数据持有方本地安全域，不对外传输。
9.1.2 交互仅传输密文，攻击者无法通过密文推导原始样本数值。
9.1.3 数据安全管控框架应符合 GB/T 45230-2025 规定，全局隐私预算累计值不应大于 8。

9.2 训练模型知识产权保护要求

9.2.1　神经网络模型权重参数全程以密文形态完成传输、持久化存储，远端模型计算主体无法获取完整明文模型文件。
9.2.2　系统配套梯度数值扰动、推理结果噪声添加机制，禁止外部主体依托密态梯度文件、批量推理输出结果反向推导还原完整模型核心参数。

9.3　抗攻击安全要求

9.3.1　全部业务特征统一约束标准化缩放区间，消除不同样本密文数据分布差异带来的隐私泄漏风险。
9.3.2　采用恶意安全模型部署的业务系统强制配套零知识证明模块校验密态梯度数据合法性，防攻击检测方法参照 YD/T 4690-2024 执行。
9.3.3　系统实时监控每一组密文内部累积噪声数值，当噪声数值超出安全阈值时自动执行密文引导操作或者直接终止当前计算任务

9.4　密钥全生命周期安全管理要求

密钥生成、分发、存储、使用、备份、轮换、销毁全流程管理应满足 GB/T 45230-2025；私钥不应明文持久存储，由加密硬件或专用 KMS 托管。

10　性能与精度指标

10.1　推理时延

测试环境统一采用多项式次数 4096 配置的 CKKS 同态加密算法开展单样本推理时延测试工作：

a)　浅层模型（层数≤3）单样本推理时延不应大于 200 ms；
b)　中等网络（层数≤8）单样本推理时延不应大于 1 000 ms；
c)　分布式批量推理吞吐量不应低于 50 样本/min。

10.2　存储开销

10.2.1　完整公私钥和求值密钥存储容量不应大于 300 MB。
10.2.2　单层网络密文权重相对明文膨胀倍数不应大于 80 倍。

10.3　通信带宽

单样本密文特征单次上传占用带宽不应大于 5 MB。

10.4　模型精度损失限制

10.4.1　分类模型密态训练准确率下降幅度不应大于 3%。
10.4.2　回归模型均方误差增幅不应大于 5%。

11　实现与集成要求

11.1　底层密码库适配要求

系统底层应兼容 OpenFHE、Microsoft SEAL、TFHE 开源同态库，不应仅绑定单一闭源商用库；云计算环境密码适配宜参照 T/SHCCIA 002-2023。

11.2　机器学习框架适配要求

系统提供标准化适配层，兼容 PyTorch、TensorFlow、Scikit-learn，上层业务代码改造量最小化。

11.3 跨硬件平台兼容要求

支持 x86 服务器、ARM 边缘设备、GPU 密码加速硬件，各平台接口行为保持一致。

11.4 标准化接口要求

系统提供 REST、gRPC 标准化接口，覆盖密钥管理、训练任务、推理任务、审计日志查询。

12 测试方法

12.1 算法基础功能正确性

按下列测试流程进行：

a) 加解密数值一致性验证测试：将原始明文数据加密转换为密文数据，依托密态算子完成完整数值运算，执行解密操作之后对比解密数值与直接明文运算数值，二者差值应当处于本文件规定误差阈值区间之内；

b) 模型训练收敛效果验证测试：依托同态密态流程完成完整模型迭代训练工作，最终模型预测准确率与纯明文训练模型准确率差值满足本文件第九章第四条款精度损失约束要求；

c) 神经网络前向传播全算子覆盖验证测试：遍历系统全部密态计算算子完成完整推理流程，不存在算子缺失、计算逻辑异常问题。

12.2 数据隐私安全性

按下列测试流程进行：

a) 密文信息泄露风险专项测试：模拟外部攻击者捕获全部传输密文数据，攻击者无法通过数学推导还原任意一条原始业务样本特征数值，隐私安全测试流程符合 YD/T 4690-2024 相关规定；

b) 全局隐私预算核算校验测试：完整统计整套业务流程累计消耗隐私预算数值，总数值不超出本文件规定上限；

c) 抵御投毒攻击、数据推断攻击渗透测试：按照标准化渗透测试流程模拟各类恶意操作行为，系统能够完整识别、阻断全部违规攻击行为。

12.3 性能指标

测试环境统一部署本文件附录 A 列明的标准化同态加密算法参数配置，完整采集单样本推理时延、批量推理吞吐量、密文文件存储占用空间、密文传输带宽消耗全部性能数据，采集得到的全部性能数值与第九章列明最低阈值逐项比对，全部指标达标判定为性能测试合格。

12.4 精度损失验证

测试过程统一采用国际公开标准数据集（手写数字识别 MNIST 数据集、鸢尾花分类 IRIS 数据集、公开金融风控数据集）分别完成明文模型训练、同态密态模型训练两项试验，对比两组模型预测准确率、回归均方误差核心指标，精度损失幅度满足 9.4 条件判定为精度验证合格。

附 录 A
（规范性）
面向隐私保护机器学习场景的主流同态加密算法标准参数配置表

本附录完整规定 128 bit 安全等级下，隐私保护机器学习业务场景 BFV、CKKS、FHEW 三类同态加密算法强制基础配置参数，所有正式投产业务系统底层密码参数不得低于表 A.1 配置标准。

表 A.1 面向隐私保护机器学习场景的主流同态加密算法标准参数配置

算法	多项式次数，N	logQ 模数总对数	明文缩放/模数	最大乘法深度	适用场景
CKKS	4096	450	2^30	5	深度学习、浮点特征
BFV	8192	600	65536	6	线性模型、整数离散特征
FHEW	2048	128	2^16	2	边缘轻量化推理