

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



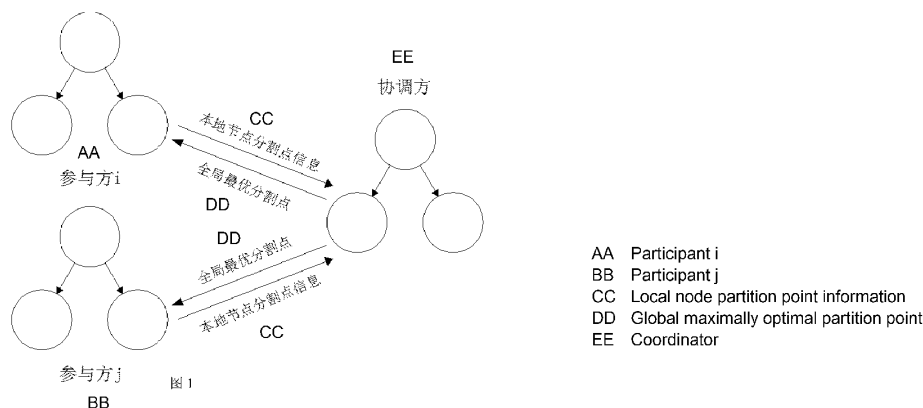
(10) 国际公布号
WO 2022/151654 A1

(43) 国际公布日
2022 年 7 月 21 日 (21.07.2022)

- (51) 国际专利分类号:
G06N 3/04 (2006.01) *G06K 9/62* (2006.01)
- (21) 国际申请号: PCT/CN2021/101319
- (22) 国际申请日: 2021 年 6 月 21 日 (21.06.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110046246.2 2021 年 1 月 14 日 (14.01.2021) CN
- (71) 申请人: 新智数字科技有限公司 (ENNEW DIGITAL TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市朝阳区望京东路 1 号摩托罗拉大厦 10 层, Beijing 100020 (CN)。
- (72) 发明人: 张金义 (ZHANG, Jinyi); 中国北京市朝阳区望京东路 1 号摩托罗拉大厦 10 层, Beijing 100020 (CN)。 李振飞 (LI, Zhenfei); 中国北京市朝阳区望京东路 1 号摩托罗拉大厦 10 层, Beijing 100020 (CN)。
- (74) 代理人: 北京嘉科知识产权代理事务所 (特殊普通合伙) (BEIJING JIAKE INTELLECTUAL PROPERTY LAW FIRM); 中国北京市朝阳区小关北里甲 2 号楼 B 座 6 层 B601 室, Beijing 100020 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: RANDOM GREEDY ALGORITHM-BASED HORIZONTAL FEDERATED GRADIENT BOOSTED TREE OPTIMIZATION METHOD

(54) 发明名称: 一种基于随机贪心算法的横向联邦梯度提升树优化方法



(57) Abstract: A random greedy algorithm-based horizontal federated gradient boosted tree optimization method, which comprises the following steps: a coordinator configures gradient boosted tree model relevant parameters, which comprise but are not limited to a maximum number of decision trees (T), a maximum tree depth (L), an initial prediction value (base), etc., and sends same to each participant (p_i); and each participant partitions a current node dataset according to a partition feature f and a partition value v, and allocates new partitioned data to a child node. In the present random greedy algorithm-based horizontal federated gradient boosted tree optimization method, horizontal federated learning comprises participants and a coordinator, the participants possess local data, the coordinator does not possess any data, and is a center that aggregates participant information, the participants respectively calculate a histogram and send the histogram to the coordinator, and after collecting all histogram information, the coordinator searches for a maximally optimal partition point according to a greedy algorithm, and then shares same with each participant, and operates in coordination with an internal algorithm.

[见续页]

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种基于随机贪心算法的横向联邦梯度提升树优化方法, 包括如下步骤: 协调方设置梯度提升树模型相关参数, 包括但不限于决策树最大数量 T 、树最大深度 L 、初始预测值 $base$ 等, 并下发到各个参与方 p_i , 各个参与方根据分割特征 f 和分割值 v 分割当前节点数据集, 并将新的分割数据分配给子节点, 该基于随机贪心算法的横向联邦梯度提升树优化方法, 横向联邦学习中包括参与方和协调方, 参与方拥有本地数据, 协调方不拥有任何数据, 进行参与方信息聚合的中心, 参与方分别计算直方图, 将直方图发送给协调方, 协调方汇总全部直方图信息后, 根据贪心算法寻找最优分割点, 然后分享给各个参与方, 配合内部的算法进行工作。