

(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 2022/001887 A1

(43) 国际公布日
2022 年 1 月 6 日 (06.01.2022)

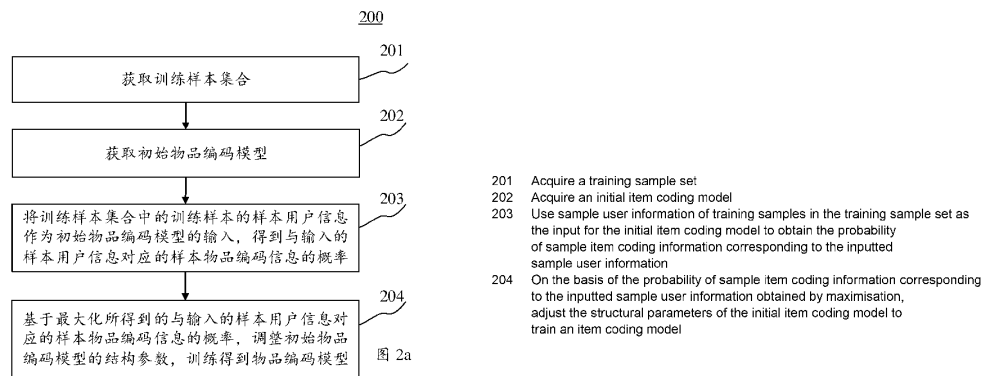
- (51) 国际专利分类号:
G06K 9/62 (2006.01) *G06N 20/00* (2019.01)
- (21) 国际申请号: PCT/CN2021/102486
- (22) 国际申请日: 2021 年 6 月 25 日 (25.06.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010613628.4 2020 年 6 月 30 日 (30.06.2020) CN
- (71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街 30 号院 3 号楼 2 层 B-0035 房间, Beijing 100041 (CN)。字节跳动有限公司 (BYTEDANCE INC.) [US/US]; 美国加利福尼亚州

亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。

- (72) 发明人: 高伟豪(GAO, Weihao); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。樊向军(FAN, Xiangjun); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。孙建凯(SUN, Jiankai); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。肖文之(XIAO, Wenzhi); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。王崇(WANG, Chong); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。刘小兵(LIU, Xiaobing); 美国加利福尼亚州洛杉矶西杰弗逊大道 12655 号六层第 137 室, California 90066 (US)。

(54) Title: METHOD AND APPARATUS FOR TRAINING ITEM CODING MODEL

(54) 发明名称: 用于训练物品编码模型的方法和装置



(57) Abstract: Disclosed in the embodiments of the present invention are a method and an apparatus for training an item coding model. A specific embodiment of the present method comprises: acquiring an initial item coding model and a training sample set; using sample user information of training samples in the training sample set as the input for the initial item coding model to obtain the probability of sample item coding information corresponding to the inputted sample user information; on the basis of the probability of sample item coding information corresponding to the inputted sample user information obtained by maximisation, adjusting the structural parameters of the initial item coding model to train an item coding model, the item coding model being used for characterising the correspondence between inputted sample user information and sample item coding information and the correspondence between sample item information and sample item coding information. The present embodiment can use the trained item coding model to implement item recommendation and can use the item coding information as an index to increase retrieval efficiency.

[见续页]

(74) 代理人: 北京天达共和律师事务所 (EAST & CONCORD PARTNERS); 中国北京市朝阳区东三环北路8号亮马河大厦2座19层刘璠, Beijing 100004 (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, 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: 本申请实施例公开了用于训练物品编码模型的方法和装置。该方法的一具体实施方式包括: 获取初始物品编码模型和训练样本集合; 将该训练样本集合中的训练样本的样本用户信息作为初始物品编码模型的输入, 得到与输入的样本用户信息对应的样本物品编码信息的概率; 基于最大化所得到的与输入的样本用户信息对应的样本物品编码信息的概率, 调整初始物品编码模型的结构参数, 训练得到物品编码模型, 其中, 物品编码模型用于表征输入的样本用户信息与样本物品编码信息之间的对应关系和样本物品信息与样本物品编码信息之间的对应关系。该实施方式可以利用训练好的物品编码模型实现物品推荐, 并可以将物品编码信息作为索引以提升检索效率。