

(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 2019/174419 A1

(43) 国际公布日
2019 年 9 月 19 日 (19.09.2019)

(51) 国际专利分类号:
G06F 16/35 (2019.01)

(21) 国际申请号: PCT/CN2019/073411

(22) 国际申请日: 2019 年 1 月 28 日 (28.01.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810215700.0 2018 年 3 月 15 日 (15.03.2018) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 张雅淋 (ZHANG, Yalin); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里

巴巴集团法务部, Zhejiang 311121 (CN)。 李龙飞 (LI, Longfei); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (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, 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,

(54) Title: METHOD AND DEVICE FOR PREDICTING ABNORMAL SAMPLE

(54) 发明名称: 预测异常样本的方法和装置

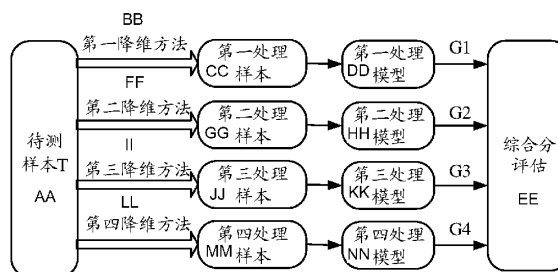


图 5

AA Samples T to be predicted
BB First dimensionality reduction method
CC First processed sample
DD First processing model
EE Comprehensive score assessment
FF Second dimensionality reduction method
GG Second processed sample
HH Second processing model
II Third dimensionality reduction method
JJ Third processed sample
KK Third processing model
LL Fourth dimensionality reduction method
MM Fourth processed sample
NN Fourth processing model

(57) Abstract: A method and device for predicting an abnormal sample, the method comprising: firstly, acquiring samples to be predicted (41); then, carrying out dimensionality reduction on the samples to be predicted respectively using a plurality of dimensionality reduction methods so as to obtain a plurality of processed samples (42); next, inputting the plurality of processed samples into a plurality of processing models respectively to obtain the score of each processed sample, wherein the ith processing model Mi scores a corresponding processed sample on the basis of a hypersphere Qi determined by an SVDD mode (43); next, determining, according to the score of each processed sample, the comprehensive score of the samples to be predicted (44); and finally, determining whether the samples to be predicted are abnormal samples or not according to the comprehensive score (45). Whether unknown samples are abnormal samples or not may be more effectively predicted by employing the described method.

[见续页]

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,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, 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: 一种预测异常样本的方法和装置, 方法包括: 首先获取待测样本 (41), 然后采用多个降维方法, 分别对待测样本进行降维处理, 以获得多个处理样本 (42); 接着将多个处理样本分别输入多个处理模型, 以获得各个处理样本的打分, 其中第*i*处理模型*M_i*, 基于SVDD方式确定的超球面*Q_i*, 为对应的处理样本打分 (43); 然后根据各个处理样本的打分确定待测样本的综合分 (44), 最后根据该综合分, 确定待测样本是否为异常样本 (45)。该方法可以更加有效地预测未知样本是否为异常样本。