

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
30 September 2021 (30.09.2021)



(10) International Publication Number
WO 2021/191703 A1

(51) International Patent Classification:
G06N 3/08 (2006.01)

(21) International Application Number:
PCT/IB2021/051532

(22) International Filing Date:
24 February 2021 (24.02.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/830,905 26 March 2020 (26.03.2020) US

(71) Applicant: **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York 10504 (US).

(71) Applicants (*for MG only*): **IBM (CHINA) INVESTMENT COMPANY LTD.** [CN/CN]; 25/F, Pangu Plaza, No. 27, Central North 4th Ring Road, Chaoyang District, Beijing 100101 (CN). **IBM DEUTSCHLAND GMBH** [DE/DE]; IBM-Allee 1, 71139 Ehningen (DE).

(72) Inventors: **BIGAJ, Rafal**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **CMIELOWSKI, Lukasz**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **SLOWIKOWSKI, Pawel**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL). **SOBALA, Wojciech**; c/o IBM Polska Sp. z o.o., ul. Armii Krajowej 18, 30-150 Krakow (PL).

(74) Agent: **VETTER, Svenja**; c/o IBM Deutschland GmbH, Patentwesen und Urheberrecht, IBM-Allee 1, 71139 Ehningen (DE).

(54) Title: METHOD FOR SELECTING DATASETS FOR UPDATING ARTIFICIAL INTELLIGENCE MODULE

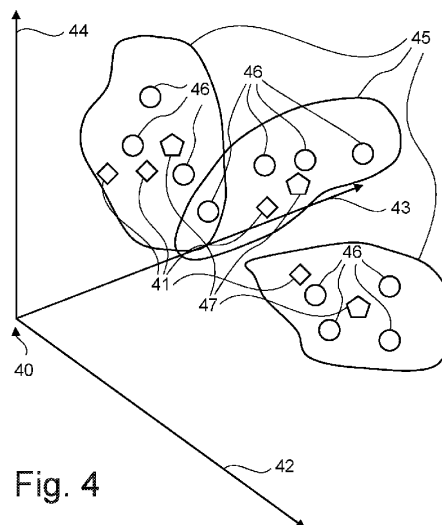


Fig. 4

(57) Abstract: A computer-implemented method for selecting a dataset from given datasets for updating an artificial intelligence module (AI-module). The given datasets (14) each comprise an input dataset (11) and a corresponding output dataset (12). The computer-implemented method comprises: obtaining values of parameters for defining different clusters (45) of the given datasets (14) (301), determining a metric of each given dataset (14), the metric of each given dataset (14) being dependent on a level of membership of the respective given dataset (14) to one of the clusters (45) and a distance of the respective given dataset (14) to a centroid (47) of the same one of the clusters (45) (302), and selecting at least one of the given datasets (14) from the given datasets (14) for updating the AI-module (1) on the basis of a comparison of the metrics of the given datasets (14) (303).

[Continued on next page]

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): 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) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (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).

Published:

— with international search report (Art. 21(3))