



- (51) **International Patent Classification:**
Not classified
- (21) **International Application Number:**
PCT/US2024/035341
- (22) **International Filing Date:**
25 June 2024 (25.06.2024)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
63/523,419 27 June 2023 (27.06.2023) US
- (71) **Applicant:** **VISA INTERNATIONAL SERVICE ASSOCIATION** [US/US]; P.O. Box 8999, San Francisco, California 94128 (US).
- (72) **Inventors:** **HE, Runxin**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128

(US). **ZHAO, Yong**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **LOU, Mingji**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **CHETIA, Chiranjeev**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **KERSTING, Nicholas, Stephen**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **YU, Junjun**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **GU, Yu**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **LIU, Can**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US). **AGRAWAL, Shubham**; Visa International Service Association, P.O. Box 8999, San Francisco, California 94128 (US).

(54) **Title:** SYSTEM, METHOD, AND COMPUTER PROGRAM PRODUCT FOR MODEL MONITORING USING BATCHED INFERENCE

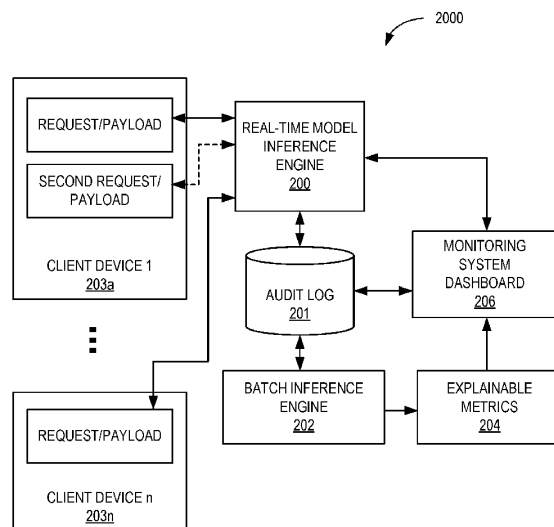


FIG. 2

(57) **Abstract:** Systems, methods, and computer program products for model monitoring using batched inferences are provided. A system includes at least one processor configured to receive processing requests to execute a machine-learning model; execute the at least one machine-learning model for each processing request to generate a decision; communicate the decision to a corresponding system; and further process the processing requests in a batch process by: generating a batch including processing requests; assigning each of the processing requests to a processor of a plurality of processors; and further processing each processing request in parallel with at least one other processing request by executing the machine-learning model for the batch to generate a second output for each processing request, the second output associated with a corresponding decision.

[Continued on next page]

(74) **Agent: PREPELKA, Nathan, J.** et al.; The Webb Law Firm, One Gateway Center, 420 Ft. Duquesne Blvd., Suite 1200, Pittsburgh, Pennsylvania 15222 (US).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*