



- (51) **International Patent Classification:**
G06N 3/08 (2006.01)
- (21) **International Application Number:**
PCT/US2021/063895
- (22) **International Filing Date:**
16 December 2021 (16.12.2021)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
17/137,720 30 December 2020 (30.12.2020) US
- (71) **Applicant:** ALTERYX, INC. [US/US]; 3345 Michelson Drive, Suite 400, Irvine, CA 92612 (US).
- (72) **Inventors:** FIRMIN, Sydney, Marie; c/o Alteryx, Inc., 3345 Michelson Drive, Suite 400, Irvine, CA 92612 (US). KANTER, James, Max; c/o Alteryx, Inc., 3345 Michelson Drive, Suite 400, Irvine, CA 92612 (US). VEERA-MACHANENI, Kalyan, Kumar; c/o Alteryx, Inc., 3345 Michelson Drive, Suite 400, Irvine, CA 92612 (US).
- (74) **Agent:** HOFFMAN, Brian, M. et al.; Fenwick & West LLP, Silicon Valley Center, 801 California Street, Mountain View, CA 94041 (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, 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))

- (54) **Title:** INTEGRATED FEATURE ENGINEERING

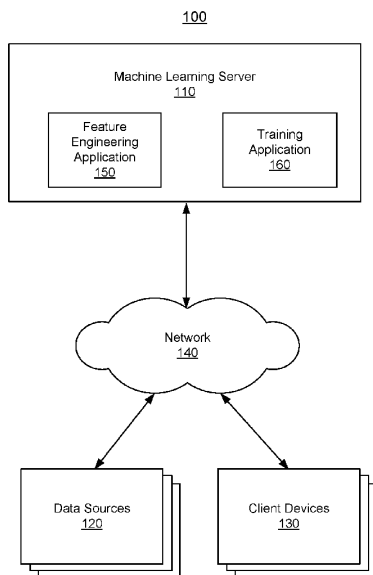


FIG. 1

(57) **Abstract:** A feature engineering application receives a plurality of data sets from different data sources for training a model for making a prediction based on new data. The feature engineering application generates primitives based on the data sets. A primitive is to be applied to a variable in the data sets to synthesize a feature. The feature engineering application also receives a temporal parameter that specifies a temporal value for generating time-based features. After the primitives are generated and the temporal parameter is received, the feature engineering application aggregates the plurality of data entities based on primary variables in the plurality of data entities and generate an entity set based on the aggregation. The feature engineering application then synthesize features, including the time-based features, based on the entity set, at least some of the primitives, and the temporal parameter.