



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
19.09.2018 Bulletin 2018/38

(51) Int Cl.:
G06F 3/06 (2006.01) G06F 12/02 (2006.01)

(21) Application number: **18187162.5**

(22) Date of filing: **02.08.2018**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(30) Priority: **19.10.2017 US 201762574534 P**
24.10.2017 US 201762576523 P
30.10.2017 US 201762579057 P
22.01.2018 US 201862620286 P
26.03.2018 US 201862648368 P
30.03.2018 US 201862650736 P

(71) Applicant: **Pure Storage, Inc.**
Mountain View, California 94041 (US)

(72) Inventors:
• **GOLD, Brian**
Mountain View, CA California 94041 (US)
• **WATKINS, Emily**
Mountain View, CA California 94041 (US)
• **JIBAJA, Ivan**
Mountain View, CA California 94041 (US)
• **OSTROVSKY, Igor**
Mountain View, CA California 94041 (US)
• **KIM, Roy**
Mountain View, CA California 94041 (US)

(74) Representative: **FRKelly**
27 Clyde Road
Dublin D04 F838 (IE)

(54) **ENSURING REPRODUCIBILITY IN AN ARTIFICIAL INTELLIGENCE INFRASTRUCTURE**

(57) Ensuring reproducibility in an artificial intelligence infrastructure that includes one or more storage systems and one or more graphical processing unit ('GPU') servers, including:

identifying, by a unified management plane, one or more transformations applied to a dataset by the artificial intelligence infrastructure, wherein applying the one or more transformations to the dataset causes the artificial intelligence infrastructure to generate a transformed dataset; storing, within the one or more storage systems, information describing the dataset, the one or more transformations applied to the dataset, and the transformed dataset; identifying, by the unified management plane, one or more machine learning models executed by the artificial intelligence infrastructure using the transformed dataset as input; and storing, within the one or more storage systems, information describing one or more machine learning models executed using the transformed dataset as input.

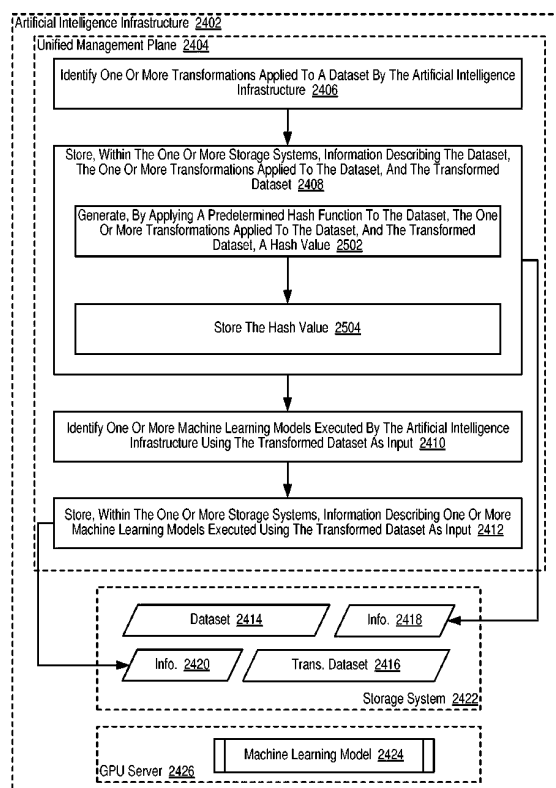


FIG. 25