

(12) PATENT APPLICATION PUBLICATION

(21) Application No.202311058524 A

(19) INDIA

(22) Date of filing of Application :31/08/2023

(43) Publication Date : 29/09/2023

(54) Title of the invention : PARALLEL COLLECTIVE MATRIX FACTORIZATION FRAMEWORK FOR BIG DATA

(51) International classification :G06F 169535, G06F 171600, G06N 200000, G06Q 300200, G09G 032000
(86) International Application No :NA
Filing Date :NA
(87) International Publication No : NA
(61) Patent of Addition to Application Number :NA
Filing Date :NA
(62) Divisional to Application Number :NA
Filing Date :NA

(71)Name of Applicant :

1)Chitkara University

Address of Applicant :Chitkara University, Chandigarh-Patiala National Highway, Village Jhansla, Rajpura, Punjab - 140401, India. Patiala -----

2)Bluest Mettle Solutions Private Limited

Name of Applicant : NA

Address of Applicant : NA

(72)Name of Inventor :

1)MISHRA, Rahul

Address of Applicant :ODC-4, Panchshil Tech Park, inside Courtyard by Marriott premises, Hinjewadi Phase - 1, Pune - 411057, Maharashtra, India. Pune -----

2)PANDEY, Sakshi

Address of Applicant :ODC-4, Panchshil Tech Park, inside Courtyard by Marriott premises, Hinjewadi Phase - 1, Pune - 411057, Maharashtra, India. Pune -----

3)MANTRI, Archana

Address of Applicant :Chitkara University, Chandigarh-Patiala National Highway, Village Jhansla, Rajpura, Punjab - 140401, India Patiala -----

(57) Abstract :

The present invention introduces a Parallel Collective Matrix Factorization Framework for big data, offering a highly efficient and scalable solution for processing large-scale data matrices in distributed computing environments. The framework comprises a Data Partitioning Module that divides the input data matrix into smaller submatrices, a Computation Module performing matrix factorization operations in parallel using algorithms like alternating least squares (ALS) or stochastic gradient descent (SGD), and a Communication Module facilitating seamless information exchange among computing nodes. The framework finds applications in recommendation systems, machine learning, data mining, and real-time processing, enabling the analysis of vast high-dimensional data sets previously challenging to handle. Its adaptability, parallelization capabilities, and communication efficiency make it a valuable tool in various domains, enhancing data processing speed and accuracy on distributed computing clusters.

No. of Pages : 24 No. of Claims : 10