

(12) PATENT APPLICATION PUBLICATION

(21) Application No.202331075388 A

(19) INDIA

(22) Date of filing of Application :04/11/2023

(43) Publication Date : 10/11/2023

(54) Title of the invention : A Machine learning tool for selecting feature from high dimensional data

(51) International classification :G06K0009620000, G06N0020000000, G16B0025100000, G16B0020000000, G16H0050700000

(86) International Application No :PCT//
Filing Date :01/01/1900

(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)Indian Institute of Technology (ISM) Dhanbad

Address of Applicant :Department of Mathematics and Computing, Science Block, IIT(ISM) Dhanbad, Jharkhand-826004 Dhanbad -----

Name of Applicant : NA

Address of Applicant : NA

(72)Name of Inventor :

1)Dr. Atanu Bhattacharjee

Address of Applicant :C2-106 Arihant Anshula, Ghot Gaon, Taloja, Navi Mumbai-410206, Maharashtra, India. Mumbai -----

2)Dr. Gajendra K. Vishwakarma

Address of Applicant :Department of Mathematics and Computing, Indian Institute of Technology (ISM) Dhanbad, Dhanbad-846004, India Dhanbad -----

(57) Abstract :

The present invention relates to a method and system for high-dimensional time to event data analysis, addressing the challenges associated with intricate gene expressions in biological research. Machine learning techniques have become pivotal tools for dimensionality reduction and pattern recognition in data mining. In the realm of survival analysis, where the temporal aspect of events is fundamental, the identification of pertinent features is crucial for building accurate prediction models. However, the utilization of machine learning in high-dimensional survival data has been hampered by the lack of specialized programming functions and packages. This invention introduces a pioneering solution – the highMLR system. highMLR is a comprehensive and efficient machine learning procedure meticulously designed for the analysis of high-dimensional survival data linked with gene expressions. Unlike existing solutions, highMLR not only bridges the gap in handling high-dimensional data but also integrates a logarithmic loss function for advanced feature selection. Additionally, it incorporates a diverse array of statistical methods tailored specifically for survival analysis, providing unparalleled accuracy in predictions.

No. of Pages : 11 No. of Claims : 10