**Development of Genetic Algorithm Based Machine Learning Model for Coal Characterization**

*A. K. Gorai1 and S. Mishra2*

1Professor and Head, Department of Mining Engineering National Institute of Technology Rourkela-769008, India, email: [amit\_gorai@yahoo.co.uk](mailto:amit_gorai@yahoo.co.uk)

2Ex-Student, Department of Mining Engineering National Institute of Technology Rourkela-769008, India

***Keywords:*** Automation;Coal characterization; Machine Learning; Genetic Algorithm

# ABSTRACT

Coal is a heterogeneous material with the major constituents of carbon, moisture, ash, volatile matter, oxygen, sulphur, nitrogen, and hydrogen and thus utmost important to make characterisations before specific use. Thus, the present study aims to develop a genetic algorithm (GA)-based machine learning algorithms for the prediction of coal compositions (ash, volatile matter, and moisture content). The work involves multiple steps, including collection of the heterogeneous coal samples, capturing images of coal samples in a controlled environment, laboratory analysis (proximate analysis) of coal samples corresponds to each captured image, pre-process (de-noising and segmentation) of the image to remove distortions that exist in the image, features extractions (colour and texture), data augmentation, feature optimization, and build different machine learning models [ridge cross-validation (RCV) regression and artificial neural network (ANN)]. The model hyper-parameter optimization was done using GridSearchCV. Finally, the model is exported for the preparation of a graphical user interface (GUI) for automatic prediction and visualisation of coal quality. The study results indicated that both the developed models offer good prediction accuracy on the testing dataset. The model performances were measured using different error indices like R2 value, mean squared error, etc. The RCV regression model gives an R2 value of 0.83, 0.98, and 0.97 respectively for ash content, volatile matter, and moisture content of the coal samples; whereas, the ANN model gives R2 -value of 0.87, 0.82, and 0.82 respectively for ash content, volatile matter and moisture content of coal. That is, a good correlation between the predicted and observed coal quality are observed in both the developed machine learning model. Therefore, either of the algorithms can be used for the prediction of coal composition.