INTISARI

Penelitian ini menganalisis karakteristik geokimia batubara di Indonesia menggunakan pendekatan *unsupervised machine learning* guna mengidentifikasi pola tersembunyi dalam data multidimensi. Penelitian ini bertujuan mengidentifikasi karakteristik geokimia batubara berdasarkan perbedaan cekungan dan umur geologi. Data geokimia batubara dari lima cekungan utama di Indonesia (Sumatera Selatan, Tarakan, Barito, Pasir, dan Ombilin) yang mencakup variasi umur Eosen dan Miosen digunakan dalam analisis ini. Metode yang diterapkan meliputi eliminasi data *outlier* menggunakan *boxplot*, normalisasi data dengan *Isometric Log-Ratio Transformation* (ILR), serta analisis perbedaan signifikansi menggunakan *Kruskal-Wallis Test and Dunn Test*. Karakteristik geokimia diidentifikasi lebih lanjut melalui *Principal Component* Analysis (PCA), *Hierarchical Clustering Analysis* (HCA), dan *Self-Organizing* Map (SOM).

Hasil penelitian menunjukkan perbedaan signifikan dalam konsentrasi unsur jejak antar cekungan dan umur geologi. Batubara Miosen, khususnya di Cekungan Sumatera Selatan dan Tarakan, didominasi oleh unsur REY dan Zr yang berkaitan dengan masukan material vulkaniklastik dan alterasi abu vulkanik. Sementara itu, batubara Eosen di Cekungan Barito (Formasi Tanjung), Pasir, dan Ombilin dicirikan oleh dominasi Ba, Sr, dan V, yang merefleksikan pengaruh lingkungan marin. Analisis oksida utama dan parameter proksimat tidak menunjukkan karakteristik spesifik antar cekungan atau berdasarkan umur pembentukan batubara.

Dengan demikian, unsur jejak terbukti lebih efektif dalam mencerminkan perbedaan genetik batubara berdasarkan cekungan dan umur geologi. Penelitian ini memberikan wawasan komprehensif mengenai variasi geokimia batubara antar cekungan dan umur di Indonesia. Temuan ini mengindikasikan bahwa kondisi geologi lokal berperan penting dalam distribusi unsur jejak.

Kata Kunci: Geokimia Batubara, *Unsupervised Machine Learning*, Karakteristik, Cekungan, Umur

ABSTRACT

This study analyzes the geochemical characteristics of coal in Indonesia using an unsupervised machine learning approach to uncover hidden patterns in multidimensional data. The research aims to identify coal geochemical characteristics based on differences in basins and geological age. Geochemical data from coal samples in five major Indonesian basins (South Sumatra, Tarakan, Barito, Pasir, and Ombilin), spanning ages from Eocene and Miocene, were used in this analysis. The applied methods include outlier removal using boxplots, data normalization with Isometric Log-Ratio Transformation (ILR), and significance analysis using the Kruskal-Wallis Test and Dunn Test. Geochemical characteristics were further identified through Principal Component Analysis (PCA), Hierarchical Clustering Analysis (HCA), and Self-Organizing Map (SOM).

The results show significant differences in trace element concentrations among basins and geological ages. Miocene coal, particularly in the South Sumatra and Tarakan Basins, is dominated by REY dan Zr, which are associated with volcaniclastic input and volcanic ash alteration. Meanwhile, Eocene coal from the Barito (Tanjung Formation), Pasir, and Ombilin Basins is characterized by the dominance of Ba, Sr, and V, reflecting marine environmental influences. Analysis of major oxides and proximate parameters did not reveal specific characteristics across basins or coal formation ages.

Thus, trace elements prove more effective in reflecting the genetic differences of coal based on basin and geological age. This study provides comprehensive insights into the geochemical variations of coal across Indonesian basins and ages. The findings indicate that local geological conditions play a crucial role in trace element distribution.

Keywords: *Coal Geochemistry, Unsupervised Machine Learning, Characteristics, Basin, Age*