International Journal of Innovative Research and Scientific Studies, 6(2) 2023, pages: 374-380

IJIRSS International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Coal rank data analytic for ASTM and PSDBMP classification

Mardhani Riasetiawan[1*], Ferian Anggara[2], Vanisa Syahra[3], Ahmad Ashari[4], Bambang Nurcahyo Prastowo[5], Inneke Chyntia Kusumawardani[6], Prabowo Wahyu[7]

[1,4,6]*Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Indonesia.*
[2]*Department of Geology, Faculty of Engineering, Universitas Gadjah Mada, Indonesia.*
[3,5,7]*Cloud and Grid Technology Working Group, Indonesia.*

Corresponding author: Mardhani Riasetiawan (*mardhani@ugm.ac.id*)

Abstract

The large amount of coal production in Indonesia generates a significant amount of data that can be used to understand the rank of the coal. To effectively process and interpret this data, our study employs the use of big data techniques, including big data management and big data analysis. Big data management allows us to organize and understand the data patterns, while big data analysis is used to gain insights and knowledge about the data, such as coal rank analysis and identifying the type of coal. Our study uses a python-based approach to define variables and automatically classify the coal rank based on the threshold values obtained from the two basic analyses described earlier. Our results show that this method is able to accurately classify the coal according to the given threshold. We found that according to the Indonesian Coal Standardization based on Pusat Sumber Daya Mineral Batubara dan Panas Bumi (PSDBMP) standard, the calorific value (in adb) is dominated in low to medium calories for 14 boreholes. The coal rank in American Standard Testing and Material (ASTM) analysis is dominated by Lignite A and B for 14 boreholes. The last analysis, according to the atomic ratio, shows that the coal can be classified as Lignite and Subbituminous Coal. Thus, by implementing the big data concept, we can easily analyze the coal classification with comprehensive and large amount of data.

Keywords: ASTM, Big data management, Coal rank analysis, Data analytic, PSDBMP.

DOI: 10.53894/ijirss.v6i2.1469
Funding: This study received no specific financial support.
History: Received: 5 January 2023/**Revised:** 7 March 2023/**Accepted:** 17 March 2023/**Published:** 31 March 2023

Authors' Contributions: All authors contributed equally to the conception and design of the study.
Competing Interests: The authors declare that they have no competing interests.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained.
Ethical Statement: This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

Natural resources have been the main commodity to supply and fulfill energy needs throughout the world, and Indonesia is no exception. In 2018, energy production in Indonesia was produced by oil and gas, coal, and several renewable energies reaching approximately 411.6 MTOE (Million Tonne of Oil Equivalent). Around 64% of total energy production (~261.4 MTOE) was produced by coal and Liquefied Natural Gas, which were also exported throughout the world. Coal production itself is predicted to be increased due to the domestic high demand, making Indonesia one of the biggest coal exporters in the world after Australia. Thus, this high demand needs the help of big data to manage and analyze the data itself which must contain a lot of variables. Big data in the mining industry can be used for predicting energy consumption and national production to support energy development in Indonesia [1]. Big data can be a helpful way to transform the data technology since it is quite impossible to do all analyses manually [2-4]. This transformation can help to collect, manage, map, and analyze the data comprehensively. The fundamental step is to manage the data using the concept of big data management (BDM) before applying the big data analysis (BDA). BDM works on the data collection, data structure, storage, and pre-processing of the data and then followed by application of BDA, which works on the interpretation, knowledge discovery, then summarization of the relevant information as well as relationship and pattern in the data [5-7].

In this study we aimed to break down all the data needed for coal rank interpretation. Coal mining data can either be well-structured or still unstructured. Tracing the source of the data is also an important way to deal with BDM, whether it is a direct source or an indirect source [2]. The most essential thing we need to highlight is the existence of the data called "proximate" and "ultimate" analysis, which are essential for supporting our basic analysis and examining the quality of coal. Moreover, the location of each borehole should be identified to support our advanced analysis, such as the initial guess for the Underground Coal Gasification development. In this research, we used the Pusat Sumber Daya Mineral Batubara dan Panas Bumi (PSDMBP) standard and American Standard Testing and Material (ASTM) as the basis of the analysis. The research also used the Mangunjaya (MJ) Area in Indonesia as a case study. Figure 1 shows the analysis structure used in this study, which includes PSDMBP, Atomic Ratio, and ASTM.

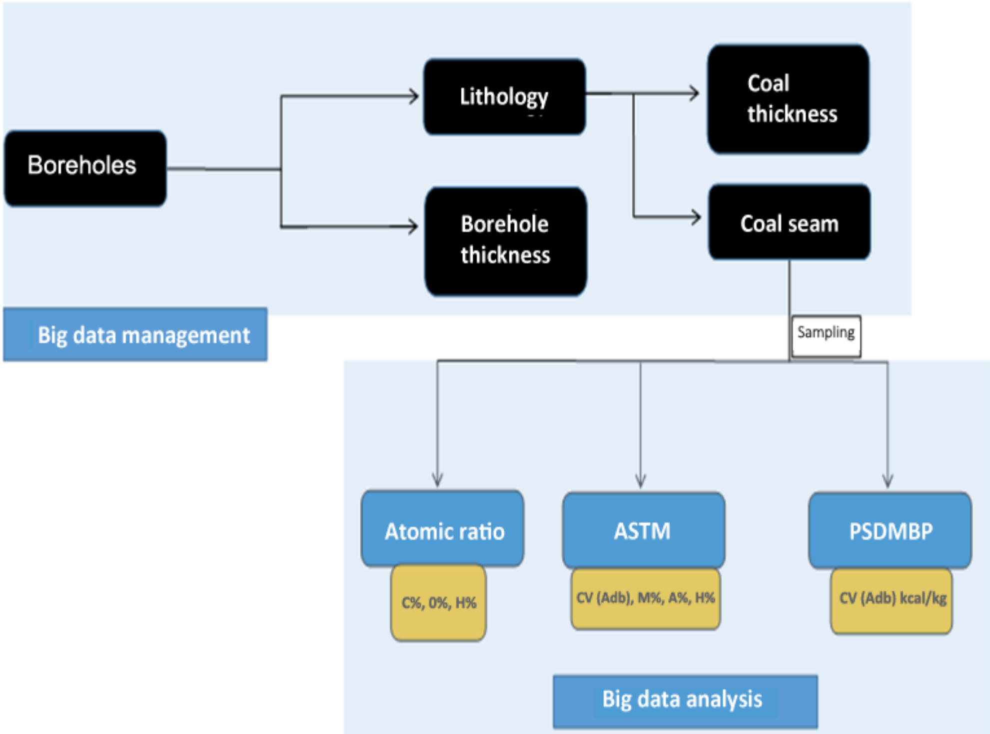

Figure 1.
Big data analysis structure.

2. Coal Data Analysis

The research has demonstrated that big data is something like 'Pacific Ocean', while traditional data is like 'Lake Toba' [8]. We can say that managing the data is 'fishing' and the data itself is the 'fish'. Therefore, big data is 'fishing for fish in the Pacific Ocean'. We need a special technique to catch the fish, don't we? Thus, big data can be a solution for us to use. Generally, big data has different criteria for some authors. The research suggests that big data management has three essential characteristics, namely, size, complexity, and technology [1]. The research has identified another way to analyze big data by using the attributes of volume, variety, velocity, veracity, and value [2, 5]. From these criteria, the research identified that big data application often consists of three main steps: collecting the source data, structuring the data, and knowledge discovery [7].

Coal data usually comes in a large quantity of unstructured data that needs to be reorganized to ease the analysis. Big data typically works on finding and comparing the algorithm that requires computing devices, especially in the processor and memory for performing queries on a large amount of data [9]. Structuring the data is not necessary if the data is already well-organized by the mining operation. However, in this study, we rearranged the data structure to facilitate our further data management and analysis.

As the first step in BDM, we collected the data from the Center of Mineral Coal and Geothermal Resource, known as Pusat Sumber Daya Mineral Batubara dan Panas Bumi (PSDBMP) in Indonesian and found out that the data was already well-organized. Data sources in the mining industry, as explained in his study, can be classified into direct and indirect sources [9, 10]. In our study, the direct source can be obtained from the Global Positioning System (GPS) which shows the exact location of the coal mining. The indirect source derived from laboratory experiments such as the gas content, fixed carbon, ash content, metallurgy, geophysical observation, etc. For the next step, we arranged the data structure Figure 1 starting from the 'borehole' information.

The borehole provides us with more information about the further analysis, such as lithology, coal seam, coal thickness, etc. Lithology is used to identify the rock layer in the subsurface, such as at what depth of sandstone, coal, and claystone [11]. Therefore, we are also able to identify the thickness of the coal layer. In the laboratory, each coal layer would undergo testing to obtain parameters such as proximate and ultimate analysis, along with the caloric value (heating value). These results would serve as our indirect source data.

3. Methodology

Coal is a sedimentary rock that consists of various organic materials, including carbon, hydrogen, oxygen, nitrogen, and other elements and minerals [12-14]. It forms over time through the compaction, alteration, and metamorphism of organic matter, resulting in the formation of carboniferous deposits.

As a result, coal is composed of carbon, hydrogen, and oxygen. These elements are often used in coal rank analysis to determine the quality of the coal. Data on coal typically includes a calorific value, proximate analysis, and ultimate analysis which can be used for basic coal rank analysis. The calorific value is commonly used as a ranking parameter for coal, ranging from low to high-calorie coal [8, 14].

Proximate analysis is used to measure the content of ash, moisture, volatile matter, and fixed carbon and is typically conducted in laboratory experiments [15, 16]. The first step in our analysis is to classify the calorific value of coal according to the Indonesian Coal Classification, which is based on Presidential Decree No. 13 of 2000 and Presidential Regulation No. 45 of 2004, as well as modifications from the US System (ASTM), International System (United Nations-Economic Commission for Europe), and Standar Nasional Indonesia (SNI) 5015-2011. The Indonesian government uses this classification, which is based on calorific value and is shown in Table 2 [17]. The next step is to classify the coal rank according to the ASTM classification, which involves comparing the basis of both analyses and converting them to the same basis if they are not already.

In our study, we have used a dry-ash-free basis, which assumes that there is no moisture or mineral matter present in the sample. We then converted the calorific value from as-determined data to a dry-ash basis, using the percent moisture and percent ash from the proximate analysis data. Our coal rank classification, according to the ASTM system, is based on the degree of metamorphism and ranges from lignite (lowest rank) to anthracite (highest rank). This classification is determined by the fixed carbon and calorific value, which were both calculated on a dry-ash-free basis. Coal is classified according to the fixed carbon value if it is greater than 69%. If the calorific value is less than 14,000 Btu/lb, it is classified based on this value using Table 4.

However, our data is in gross value, so we calculated and converted it to net value using the percent hydrogen, percent moisture, and percent oxygen from the ultimate analysis data, for a more accurate representation of the calorific value in the boiler plant. Equation 1 explains dry-ash-free (daf) based on fixed carbon (FC) and calorific value (CV) with proximate analysis presented in Table 1. To calculate the net CV in Equation 2, we need the percent hydrogen of weight (H), percent moisture of weight (M), and percent oxygen of weight (O) that can be obtained from the ultimate analysis data (for O% and H%) with proximate analysis presented in Table 2.

$$daf = adb \times \frac{100}{100 - M_{adb} - A_{adb}} \quad (1)$$

$$Net\ CV = Gross\ CV - 91.2H - 10.5M - 0.34O \quad (2)$$

Table 1.
Proximate analysis.

Basis (P)	M (%)	A (%)	Vol (%)	FC (%)
Adb	11.8	3.47	48.79	35.94
Adb	11.98	2.9	48.62	36.5
Adb	12.61	3.94	47.55	35.9
Adb	11.8	3.47	48.79	35.94

Table 2.
Ultimate analysis.

Basis (P)	C (%)	H (%)	N (%)	O (%)	S (%)
Daf	69.1	5.09	0.61	25.08	0.12
Daf	69.59	5.06	0.72	24.44	0.19
Daf	68.54	4.91	0.78	25.63	0.13

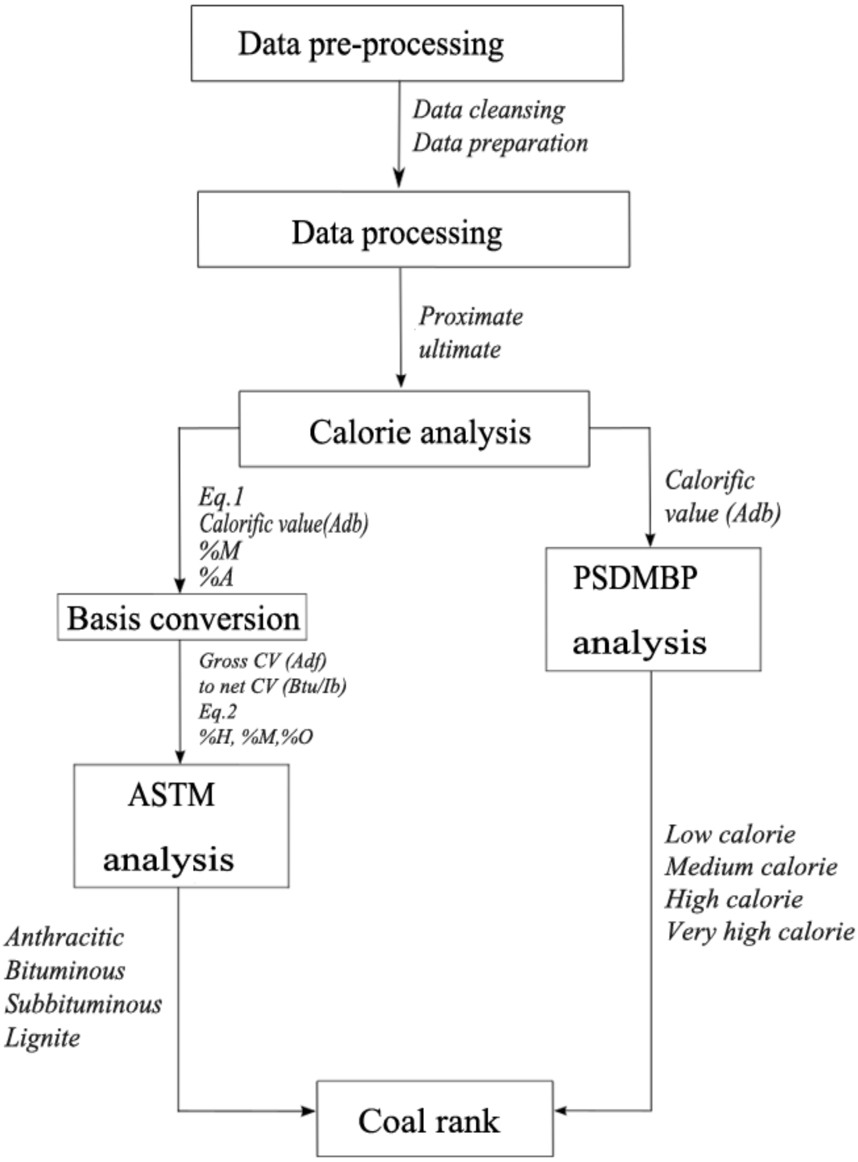

Figure 2.
Coal rank analysis flow.

4. Results and Discussion

Our goal is to classify the coal rank by utilizing the automatic process to make the interpretation easier while working on a big data Figure 2. Classification is very crucial in analysis in coal data implementation. Since we could do the further steps in processing the coal data or even implementing the coal itself (e.g., solid coal) after we know the classification of the coal. Classification in our study deals with the rank of coal. We examined the rank of coal classification using two basic analyses: Indonesian Coal Standard Classification (PSDMBP) and ASTM coal rank analysis. We implemented the automatic coding system using a python environment to classify the CV value based on Indonesian Coal Standard (PSDMBP) and coal rank ASTM from 14 boreholes. We set the system according to the value of the threshold that had been described before, as shown in Table 3 for magnetic properties and Table 4 for proximate analysis. The algorithm learns about the variables and assigns a coal rank classification to each borehole.

Table 3.
Units for magnetic properties.

Group	Calorific value (kal/g) (Adb)	Description
Low calorie	< 5100	Hard-soft, easy to squeeze, high water content, the wooden structure still can be seen
Medium calorie	5100 - 6100	Hard, easy too hard to squeeze, less water content, the wooden structure still can be seen
High calorie	6100 - 7100	Less water content, the wooden structure cannot be seen

Table 4.
Classification of coal rank based on proximate analysis [17].

Class	Group	Fixed carbon (%)		Volatile matter (%)		Calorific value (Btu/lb)	
		Equal to or greater than	Less than	Greater than	Equal to or greater than	Equal to or greater than	Less than
Anthracitic	Metaanthracite	98			2	NA	NA
	Anthracite	92	98	2	8	NA	NA
	Semianthracite	86	92	8	14	NA	NA
Bituminous	Low-volatile bituminous coal	78	86	14	22	NA	NA
	Medium-volatile bituminous coal	69	78	22	31	NA	NA
	High-volatile A bituminous coal	NA	69	31	NA	14,000	NA
	High-volatile B bituminous coal	NA	NA	NA	NA	13,000	14,000
	High-volatile C bituminous coal	NA	NA	NA	NA	11,500	13,000
Subbituminous	Subbituminous A coal	NA	NA	NA	NA	10,500	11,500
	Subbituminous B coal	NA	NA	NA	NA	9,500	10,500
	Subbituminous C coal	NA	NA	NA	NA	8,300	9,500
Lignite	Lignite A	NA	NA	NA	NA	6,300	8,300
	Lignite B	NA	NA	NA	NA	NA	6,300

The system firstly classifies the coal rank according to the Indonesian Coal Standard. To implement this code, we require a data standardization as shown in Table 3 and Table 4. This code will then help the user to see the coal distribution in each borehole. We provided an option to choose which borehole the user wants to see for either PSDMBP or ASTM coal rank analysis. According to the system, we figured out that our coal data range is dominant as medium calorie (5100 – 6100 kcal/kg), while for the ASTM, we identified that the coal ranks were distributed as Lignite A and B. We also developed a graph. If the user wants to see the coal rank and calorific value (in Adb) distribution in each borehole, just by filtering the borehole name, the user can view the coal rank distribution in the selected borehole Figure 3. It could be specified that the graph represents the distribution of coal classification and caloric value as Lignite A and Lignite B. In Figure 3, the value of Net CV is presented with * mark which shows the range of value.

Table 5.
Data dictionary of coal rank analysis features.

Variable	Explanation	Example
Borehole	Borehole code	MJ01
CV	Calorific value in Adb (Unit kcal/kg) in proximate analysis	5540
A	Ash content in Adb in proximate analysis	3.47
M	Moisture content in Adb in proximate analysis	11.80
C	Carbon content in Daf in the ultimate analysis	69.10
O	Oxygen content in Daf in the ultimate analysis	25.08

Table 6.
Data dictionary of boreholes mapping.

Variable	Explanation	Example
Area name	The name of the area where the borehole is located	Mangunjaya
Borehole name	Borehole code	MJ01
Calories	The calorie content in the borehole	5540
From	Initial depth of borehole contents	75.5
To	Initial depth of borehole contents	85
Variable	Explanation	Example

Table 7.
Data standardization for coal rank analysis.

	Borehole	CV (adb)	A (%)	M(%)	C(%)	O(%)	H(%)
0	MJ01	5540.0	3.47	11.80	69.10	25.08	5.09
1	MJ01	5574.0	2.90	11.98	69.59	24.44	5.06
2	MJ01	5349.0	3.94	12.61	68.54	25.63	4.91
3	MJ01	5360.0	4.47	12.17	68.75	25.30	5.00
4	MJ01	5330.0	6.62	11.96	69.30	24.36	5.28

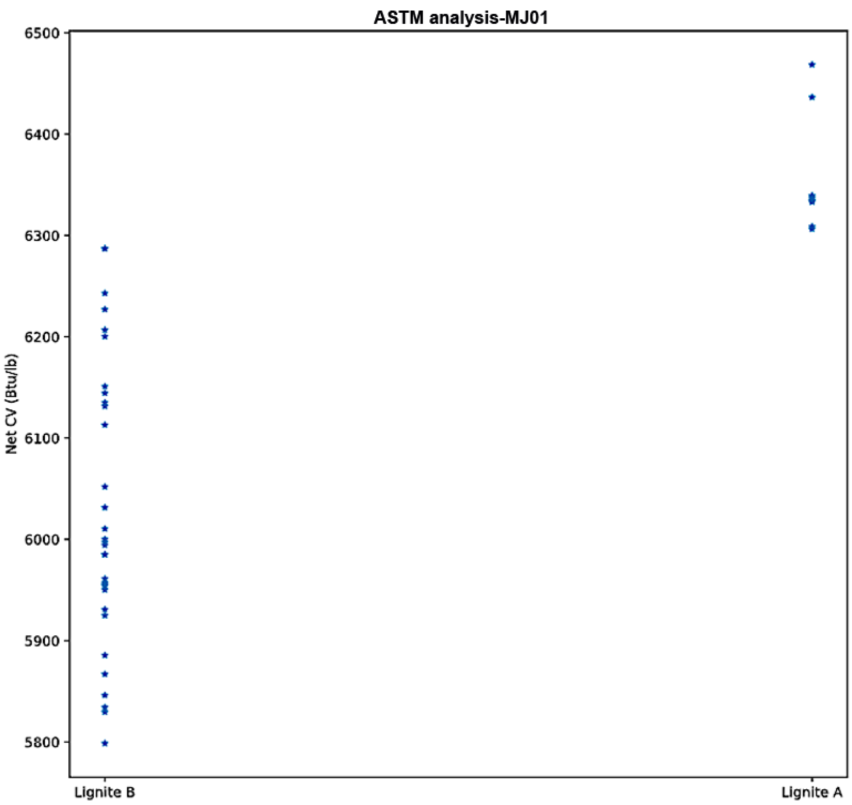

Figure 3.
A graphic of coal rank distribution in borehole MJ01.

We have also developed a system called "Borehole Mapping" that that can deploy the borehole distribution in the study area which allows the user to identify the distribution of boreholes in a map as displayed by the system according to their calorific value (adb) – the PSDMBP classification as shown in Figure 4. To deploy the system, we require a dataset standardization that is used in this feature as presented in Table 6. For example, if the user selects "low calorie" distribution, the system will display the distribution of boreholes that contain low-calorie coal. This feature can assist the user for further coal development. To classify the use of coal itself, the term 'coal rank' is often used. Coal rank is used to describe the degree of metamorphism that has occurred due to the pressure and heat during the formation of coal. Since coal is formed from plant material that undergoes the burial and metamorphism, its composition and properties change accordingly [6]. Table 7 presents the percentage of each parameter in Mangunjaya Area (MJ01) data set.

Figure 4.
Illustration of borehole mapping.

Various studies involving the coal data have been done. Xue, et al. [8] conducted a coal dilatation energy experiment to find a gas that is similar to methane through dilatation energy. They leveraged big data through big data analytic technology and a series of computing obtaining the result of the combination of 45% CO_2 and 55% N_2. Coal data can also be implemented as a cost-saving for the coal enterprise. Every time the company starts digging and drilling, the cost is huge for the company. Big data can be implemented to manage these expenses. The geologists can easily access the data provided by the cloud computing server and can make good decisions for the upcoming projects. Coal data can also be implemented in Warning System in coal mining Sun [4]. Sun [4] identified four kinds of accidents: gas explosion, rock burst accident, spontaneous coal combustion, and water disaster. It is something that deals with the gas emission, water pressure, hydrological aspect, microseismic, geophysical data, gas content, electromagnetic radiation, borehole location, geological structure, and others.

Larose and Larose [3] classified data mining into several groups according to the task: description, estimation, prediction, classification, clustering, and association. The first term "description" tends to depict the pattern and trends of the data. Applying this concept to our data, we have obtained information about each borehole that contains more data for further analysis, such as the borehole location and the thickness of the borehole used in the mining company. In estimation, researchers aimed to approximate the value of a numeric target variable from the coal data. We built a model that provides the value of the target variable to ease the user. The target variables were set into the model which can easily be used by the users if they might be interested in estimating the variables, such as the rank of coal according to ASTM analysis and Indonesian Coal Standard Classification. Implementing coal data into big data is an interactive way, especially to predict the prospecting area to be explored in the future, the Underground Coal Gasification for example. Working on coal, big data is possibly able to conduct the cost estimation for several months into the future of the coal power plant, even the coal plant production and financial planning, etc. In prediction, it should work on statistical methods such as simple linear regression, confidence interval, k-nearest neighbor, and even neural networks. However, up to this point, we have not yet built the system for the prediction part.

5. Conclusion and Future Work

The massive coal production in Indonesia generates various data that can be used for coal rank interpretation. To facilitate the interpretation, the use of big data is necessary, and it can be summarized as follows: big data management and big data analysis. Big data management eases us to read and know the data pattern to do the next step which is big data analysis. Our study managed the big data management according to the borehole information which contains a lot of information such as lithology, coal seam, and coal thickness. Big data analysis helps us in knowledge discovery about the data such as the coal rank analysis, and what kind of coal it is, and what can be used for further analysis. Our study is based on a python and defines all the used variables to automatically trace, describe, and classify the rank of coal. We found that our source code can classify the coal well according to the given threshold as we used the two basic analyses mentioned above.

References

[1] J. Ward and A. Barker, "Undefined by data: A survey of big data definitions," ArXiv Preprint ArXiv, Art. No. 1309.5821, 2013.

[2] M. Bilal *et al.*, "Big Data in the construction industry: A review of present status, opportunities, and future trends," *Advanced Engineering Informatics,* vol. 30, no. 3, pp. 500-521, 2016. https://doi.org/10.1016/j.aei.2016.07.001

[3] D. Larose and C. Larose, *Discovering knowledge in data: An introduction to data mining*. New Jersey: John Wiley & Sons, Inc, 2014.

[4] J. Sun, "Coal mine accident analysis and coal mine data and Internet of things," *Industry and Mine Automation,* vol. 16, no. 1, pp. 1-2, 2015. https://doi.org/10.5152/ttd.2015.008

[5] O. Kapliński, N. Košeleva, and G. Ropaitė, "Big Data in civil engineering: A state-of-the-art survey," *Engineering Structures and Technologies,* vol. 8, no. 4, pp. 165-175, 2016. https://doi.org/10.3846/2029882X.2016.1257373

[6] H. M. King, "Coal: What is coal and how does it form?," Retrieved: https://geology.com/rocks/coal.shtml. [Accessed 2021.

[7] C.-C. Qi, "Big data management in the mining industry," *International Journal of Minerals, Metallurgy, and Materials,* vol. 27, no. 2, pp. 131-139, 2020.

[8] Y. Xue, J. Liu, P. G. Ranjith, Z. Zhang, F. Gao, and S. Wang, "Experimental investigation on the nonlinear characteristics of energy evolution and failure characteristics of coal under different gas pressures," *Bulletin of Engineering Geology and the Environment,* vol. 81, no. 1, p. 38, 2021. https://doi.org/10.1007/s10064-021-02544-4

[9] Y. Xing, "Chinese coal big data platform," *China Land and Resource, 002,* 2015.

[10] D. van Krevelen, "Graphical-statistical method for the study of structure and reaction processes of coal fuel," *Japanese Journal of Soil Science and Plant Nutrition,* vol. 81, no. 2, pp. 162-167, 1950.

[11] N. H. Bostick and T. A. Daws, "Relationships between data from rock-eval pyrolysis and proximate, ultimate, petrographic, and physical analyses of 142 diverse U.S. coal samples," *Organic Geochemistry,* vol. 21, no. 1, pp. 35-49, 1994. https://doi.org/10.1016/0146-6380(94)90086-8

[12] S. O. Akande, M. D. Lewan, S. Egenhoff, O. Adekeye, O. J. Ojo, and A. Peterhansel, "Source rock potential of lignite and interbedded coaly shale of the Ogwashi–Asaba Formation, Anambra basin as determined by sequential hydrous pyrolysis," *International Journal of Coal Geology,* vol. 150-151, pp. 224-237, 2015. https://doi.org/10.1016/j.coal.2015.09.005

[13] J. G. Speight, *Handbook of coal analysis*. New Jersey: John Wiley & Sons, Inc, 2005.

[14] I. Suárez-Ruiz and C. Ward, "Basic factors controlling coal quality and technological behavior of coal in I. Suárez-ruiz, & J. Crelling, applied coal petrology: The role of petrology in coal utilization." London: Academic Press, 2008, pp. 19-59.

[15] R. C. Davis, S. W. Noon, and J. Harrington, "The petroleum potential of tertiary coals from Western Indonesia: Relationship to mire type and sequence stratigraphic setting," *International Journal of Coal Geology,* vol. 70, no. 1, pp. 35-52, 2007. https://doi.org/10.1016/j.coal.2006.02.008

[16] S. Indonesia, "Proximate and ultimate analysis," Retrieved: https://www.sgs.co.id/en/mining/analytical-services/coal-and-coke/proximate-and-ultimate-analysis. [Accessed 2021.

[17] ASTM, *Annual book of ASTM standards*. West Conshohocken PA American Society for Testing and Materials, 2004.