Physicochem. Probl. Miner. Process., 55(5), 2019, 1311-1322

http://www.journalssystem.com/ppmp

Physicochemical Problems of Mineral Processing

ISSN 1643-1049

Received June 26, 2019; reviewed; accepted September 04, 2019

Selection of optimal coal blends in terms of ash fusion temperatures using Support Vector Machine (SVM) classifier - a case study for Polish coals

Alina Żogała, Maciej Rzychoń, Jacek M. Łączny, Leokadia Róg

Central Mining Institute, plac Gwarków 1, 40-166 Katowice, Poland

Corresponding author: azogala@gig.eu (Alina Żogała)

Abstract: One of the most important criteria for selecting coal for a given technology are the ash fusion temperatures (AFTs). An effective way to regulate the AFTs so that they meet the criteria for a given industrial application is to form blends of different coals. The values of the AFTs in the blends are non-additive, therefore they can't be calculated using the weighted average of the blend components. On the other hand, direct determination of ATFs values requires many additional time-consuming and expensive laboratory tests. Therefore, it is important to develop a solution that, in addition to the effective prediction of the values of AFTs, will also enable optimal selection of components of the blend in terms of its key parameters. The aim of the work was to develop an algorithm for the selection of the optimal coal blends in terms of AFTs for given industrial applications. This algorithm uses nonlinear classifying model which was built using machine learning method, support vector machine (SVM). To carry out the training samples of Polish hard coals from different mines of the Upper Silesian Coal Basin were used. The accuracy of the developed model is 92.3%. The results indicate the effectiveness of the proposed solution, which can find practical application in the form of an expert system used in the coal industry. The paper presents the concept of developed IT tool which has been tested for a selected case.

Keywords: coal blends, ash fusion temperature, support vector machine, principal component analysis, machine learning

1. Introduction

One of the most important factors to be considered in choosing a coal for a given application is the measurements of ash-fusion temperatures (AFTs) which affect behavior of coals during the combustion and gasification (Vassileva and Vassilev, 2002). The AFTs are a measure of the high temperature softening and melting characteristics of a coal ash. Together with the chemical composition of ash, it is one of the factors having a significant impact on the tendency of slagging and fouling of power boilers (Róg, 2003). The AFTs of coal give also detail information on the suitability of a coal source for gasification procedures, and specifically to which extent ash agglomeration or clinkering is likely to occur within the gasifier (Van dyk et al., 2001; Miao et al., 2016; Seggiani, 1999). For instance, in the case of fixed bed reactors, with the dry ash bottom handling system, the temperature of ash melting should be higher than 1200°C (1473.15K), and for the wet bottom version it should be lower than 1300°C (1573.15K) (Porada et al., 2017). In the case of solid-state slag-tap boilers generally pulverized coal with a high AFTs is required to prevent slagging, whereas the liquid-state slag-tap boilers require a low AFTs (Wang and Massoudi, 2013). Detailed guidelines related to the appropriate AFTs values for a specific industrial application are given in the literature (Luxsanayotin et al., 2010; Porada et al., 2017; Tillman et al., 2012; Patterson and Hurst, 2000; Collot, 2006; Thomas, 2002).

Fusion temperatures typically are measured at four defined points under both reducing and oxidizing conditions. These points are simply the temperature range in which the process of ash melting takes place. The ash fusion behavior is conventionally described by the temperature of initial

DOI: 10.5277/ppmp19059

deformation (IDT), spherical temperature (ST), hemispherical temperature (HT) and ash fluid temperature (FT) (Gray, 1987; Jak, 2002). The method of determining the characteristic AFTs of solid fuels, the shape and size of samples and the method of their preparation are defined by standard used in Poland - ISO 540:1995.

The most important indicator representing AFTs in the Polish industry to assess the suitability of coal for combustion as well as gasification purposes is temperature corresponding to hemisphere temperature (HT). Investigations of ash fusion from Polish coals allowed the separation of easily fusible ashes, which ash melting (HT) in the reducing atmosphere is lower than 1200°C (1473.15K), medium fusible with HT from 1200 to 1350°C (1473.15K to 1623.15K) and ashes difficult to melt, which the HT is greater than 1350°C (1623.15K). Fig. 1 shows the connection of selected type of boilers and reactors with the HT required for their performance (based on literature guidelines) expressed in predefined intervals (easily, medium, and difficult fusible).

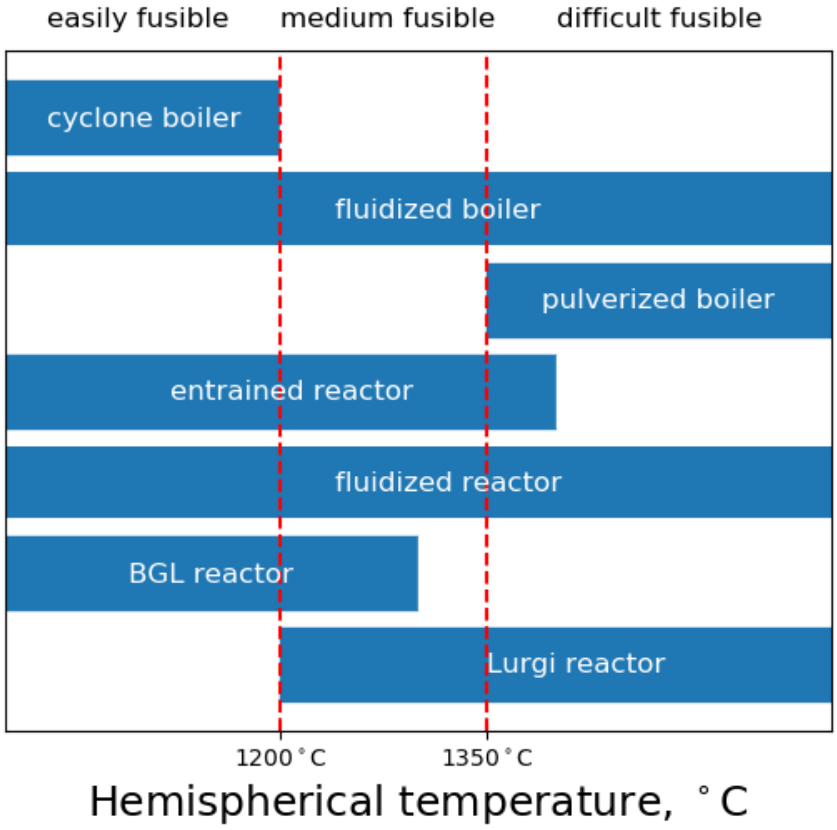

Fig. 1. Selected type of boilers and reactors with the HT required for their performance expressed in predefined intervals

In the industry, the practice of blending coals with different physicochemical parameters is often used to obtain the fuel with the desired properties (Yörükoğlu, 2017; Alekhnovich and Bogomolov, 2010). Initially the practice of blending with an engineering and scientific base was dominated by environmental issues. A very important aspect was the control of SO_2, NOx and particulates. Today the practice is dominated by economic and technical consideration (Tillman et al., 2012). Coal blending may have various targets e.g. blending for cost savings, blending for diversifying supplies for security reasons, blending for enhancing fuel flexibility and extending the range of acceptable coals or blending to achieve a particular AFTs to meet different industrial requirements (Shen at al., 2015). In the latter case, after blending process, a comprehensive investigation is required to unravel the effect of elemental ash in coal on AFTs of the blended coals (Bahrin, 2009). The temperature at which coal ash melts is an important consideration for coal-fired boilers and gasification reactors. Therefore, it is important to know, or be able to predict, the ash fusibility characteristics of a coal or coals blend before it is fed to the boiler or reactor (Seggiani and Pannocchia, 2003; Carpenter, 1995). Determining the AFTs of the coal

blend without time-consuming and expensive experiments is difficult, because these parameters are considered as a non-additive values (Carpenter, 1995). To develop blending proportions, it is not sufficient to take weighted values of the measured AFTs of the individual coal since this ignores the underlying chemistry of the ash system (Jak, 2002).

There have been a number of investigations on the prediction of AFTs using variety of methods like statistical correlations between AFTs and ash composition (Seggiani, 1999; Shi et al. 2018), empirical methods (Huggins et al., 1981) and machine learning methods based on the black box modelling (Karimi et al., 2014; Tambe et al., 2018; Yazdani et al., 2018). Many of them are characterized by reasonably good prediction accuracies. However, they are not without drawbacks e.g. in some cases not enough amount of data, especially for statistical models. Next issue concerns models based on a large number of variables, which can lead to excessive complexity. Another issue is the lack of discussion on the practical application of research results in industry.

Many references in literature emphasize the importance of composing coal blends in order to influence the values of AFTs in the context of a specific industrial application (Tillman et al., 2012; Jak et al., 2002; Shen et al., 2015; Sasi et al., 2018; Li et al., 2017). However, no practical tool was found that would be able to accomplish this task in a completely automatic way.

Therefore, the authors of the paper developed an IT tool using a machine learning-based classifier that allows selection of the optimal composition of the blend meeting the requirements of a specific industrial application in terms of AFTs. The support vector machine (SVM) technique has been examined and was chosen as the basis of the algorithm used in the IT tool. This tool has been tested for a selected case - coal from a specific mine from the Upper Silesian Coal Basin intended for combustion in a pulverized coal-fired boiler. The AFTs of the coals used in the pulverized coal-fired boilers should be as high as possible in order to prevent slagging (Tillman et al., 2012). Therefore, the task was to select such oxides in the blend, so that the resulting AFTs would meet this requirement.

2. Materials and methods

2.1. Data set

The subject of analysis and the basis for development classification model were samples of Polish hard coals. These samples were acquired from various mines of the Upper Silesian Coal Basin, containing the largest hard coal resources in country.

For each sample the hemispherical temperature (HT) in a reduction atmosphere (mixture of CO:CO_2 in proportions 3:2) and ash oxide composition (content of SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, TiO_2, Na_2O, K_2O, SO_3) was determined. The analyses of the samples were carried out in the accredited laboratory of the Department of Solid Fuel Quality Assessment at the Central Mining Institute, Poland.

In Fig. 2 statistical distribution of ash components in examined coals is presented. The plot displays the average concentration of constituents as the dotted lines. The bars indicate lower and upper quartile of the distribution, while the whiskers indicate the maximal and minimal observed concentration of component. Figure 2 shows statistical distribution of parameters after removing of the most distant outliers (388 samples).

In general, the values presented above correspond to a typical composition of the ash of Polish coals: SiO_2 – 41.60%, Al_2O_3 – 23.74%, Fe_2O_3 – 17.52%, CaO – 8.39%, MgO – 1.75%, SO_3 – 6.05%, others – 0.95%.

Each of the analyzed samples was given one of three labels – ‘L’, ‘M’ or ‘H’, corresponding to the specific hemispherical temperature range (according to Fig. 1). Label 'L' means a sample with a low HT (easily fusible), the label 'M' - a sample with a medium HT (medium fusible), label 'H' - a sample with a high HT (difficult fusible).

In the analyzed data set, 45 coal samples were characterized by low hemispherical temperature of ash, 138 - high hemispherical temperature of ash and 205 - medium hemispherical temperature of ash, 388 in total.

The oxide composition was the input vector of the classification model and HT was the output.

The entire data set was divided in a random manner into a training set (80% of samples) and a test set (20% of samples). The samples were divided in such a way, that the proportions of the classes in the training set correspond to the proportions of the classes in the test set. The training set was used to create model. Its effectiveness was then evaluated on the test set.

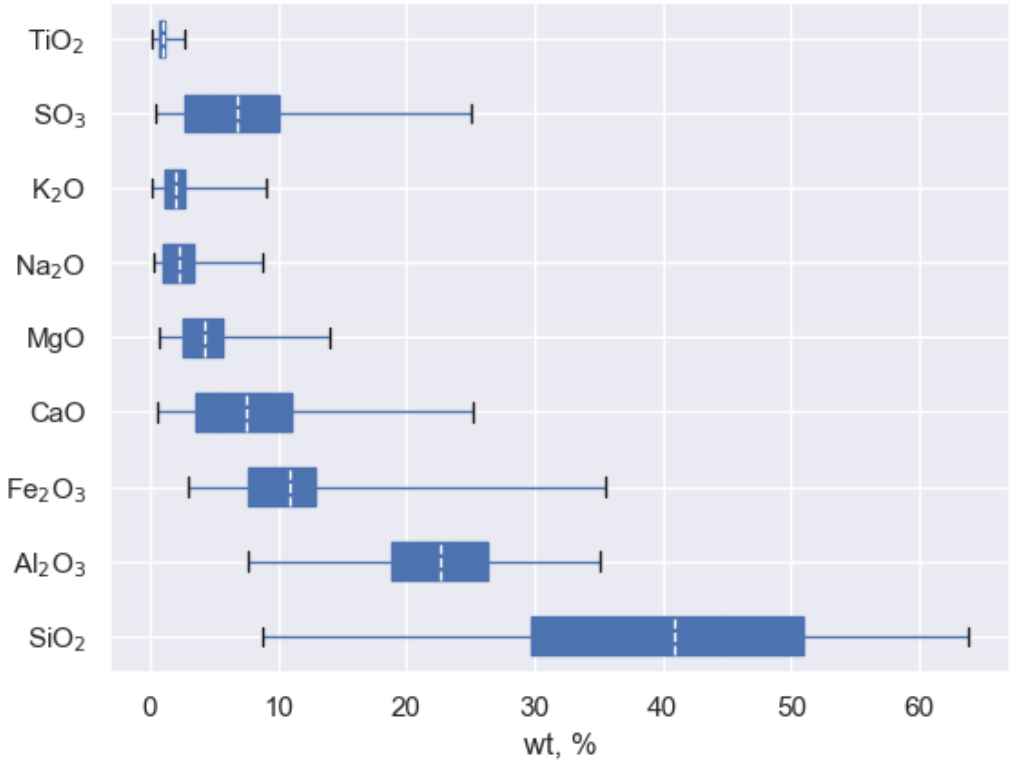

Fig. 2. Statistical distribution of ash components in examined coals

2.2. Principal component analysis

Principal component analysis (PCA) is a technique used to reduce the dimensionality of the set of variables while retaining the maximum variability of original data. PCA allows to remove multicollinearity from the data. Multicollinearity is unfavorable phenomenon during classification process as it cause redundancy and increase in the computational effort.

PCA is based on the transformation of observable input variables into new, unobservable and uncorrelated variables called principal components (PC). Each of the principal components is a linear function of the original input variables. The sum of variances of all input variables is equal to the sum of variances of the principal components. It means that the transformation of the original predictors into a set of principal components does not lead to any loss of information about analyzed phenomenon.

The principal components are ordered in such a way, that their variances are successively smaller. The first few principal components contain the vast majority of information about the analyzed phenomenon, provided by original input variables. It allows to reduce the number of principal components with the minimal loss of input information (Geron, 2017).

In the presented paper PCA analysis was applied before classification procedure in order to reduce nine-dimensional input data set.

In order to perform PCA analysis each attribute of the input vector has to be standardized:

$$X' = \frac{X - \text{mean}\,(X)}{std(X)}\,. \tag{1}$$

After that transformation features have the properties of a Gaussian distribution with mean = 0 and standard deviation = 1.

2.3. Balancing procedure

Due to the unequal number of samples belonging to individual classes of training set, the balancing procedure was performed. The number of samples of the training set belonging to classes ‘L’ and ‘M’ was balanced to the number of samples of the class ‘H’. Oversampling and undersampling in data analysis are techniques used to adjust the class distribution of a data set.

Oversampling aims to achieve an equal distribution by replicating the minority samples so that the distribution is balanced. There are a number of methods available to oversample a dataset used in a typical classification problems. The technique called Synthetic Minority Over-sampling Technique (SMOTE) was chosen to increase the number of samples of class 'L' so that it equals the number of samples of class 'H'. The SMOTE (Chawla et al., 2012) algorithm carries out an oversampling approach to rebalance the original training set. Instead of applying a simple replication of the minority class instances, the key idea of SMOTE is to introduce synthetic examples. This new data is created by interpolation based on the distances between the point and its nearest neighbors.

Undersampling is the process where some of the observations are deleted from the majority class in order to match the numbers with the minority class. Random undersampling technique was chosen to reduce the number of samples of class 'M' so that it equals the number of samples of class 'H'. This is a classical method in which the goal is to balance class distributions through the random elimination of majority class examples.

After applying the balancing procedure, the training set contained an equal number of samples of each class.

2.4. Support vector machine

Support vector machine (SVM) is supervised learning technique designed for performing binary classification, but it can be extended to the multi-class problems. Suppose we are given a set of training data $\{(x_i, d_i)$, i=1,2…,n$\}$ where n is the number of samples, x_i is input vector, d_i is response variable and can take on one of two possible values: −1 or 1. The objective of the algorithm is to find a hyperplane in an N-dimensional space of features that distinctly classifies data points.

There are many possible hyper-planes separating two classes of data points. The optimal hyperplane is considered to maximize the margin of separation between the two classes. Models of this type are more resistant to generalization errors compared to algorithms with narrow margins. Therefore, the issue of learning the structure of the SVM is defined as the following optimization problem:

$$\min_{w} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{n}\xi_i \tag{2}$$

with the following constraints (i=1,2…,n):

$$d_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \tag{3}$$

$$\xi_i \geq 0 \tag{4}$$

where $\mathbf{w}$ is N-dimensional vector of weights, b - the bias (offset for that hyperplane), C – the capacity constant (adjusted by the user), ξ_i non-negative slack variable used for handling non separable data, selected individually for each sample, ϕ - kernel function (mapping the original data is into the space of features with a higher number of dimension) (Vapnik, 1998; Cortes and Vapnik, 1995).

The main advantage of the SVM algorithm is low sensitivity to the number of learning samples. This is particularly important in solving problems for which the amount of data is limited. This technique also scales relatively well to the high dimensional data and captures the nonlinear relationships between variables (Mohamed, 2017). These advantages make the SVM algorithm a good tool for analyzing data sets related to coal parameters, in which the number of records is limited, dimensionality is high and relationships between variables are complex and non-linear.

2.5. Software

Data preparation and exploration, development of classification model and its evaluation were carried out using the Python programming language and its libraries dedicated to data analysis, machine learning and visualization (Pandas, Scikit-Learn, Matplotlib, Seaborn).

3. Results and discussion

3.1. Principal component analysis

Nine inputs (content of SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, TiO_2, Na_2O, K_2O, SO_3 in ash) of data set were subjected to the principal component analysis (PCA). After transformation nine PCs were extracted. In

Fig. 3 variance ratio explained by particular principal components is presented. First component captures 66.1% percentage of variance, second - 12.2%, third – 7.9%, fourth – 6.4%.

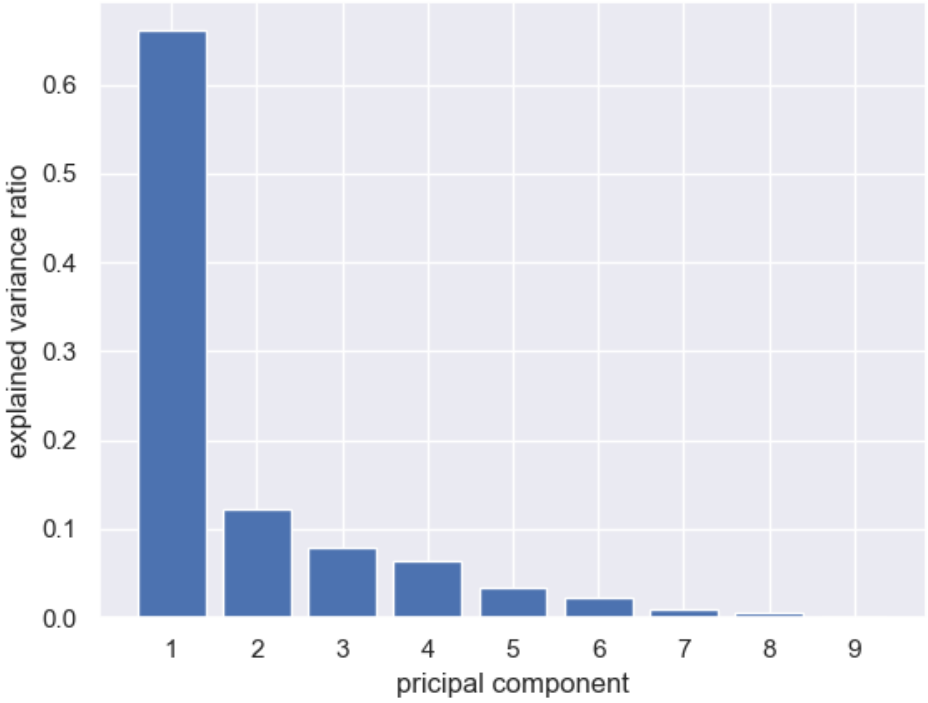

Fig. 3. Ratio of variance explained by particular principal components (PC)

Fig. 4-6 show the projections of the examined coal samples onto the planes determined by particular pairs of principal components. Generally, the classes are quite well separated from each other.

The projection of records onto the plane determined by PC1-PC2 allows the observation that all the coal samples with low hemispherical temperature of ash are grouped together. Samples with medium and high hemispherical temperature of ash are more difficult to distinguish. However, most records of 'H' class are placed on the right side of the coordinate system.

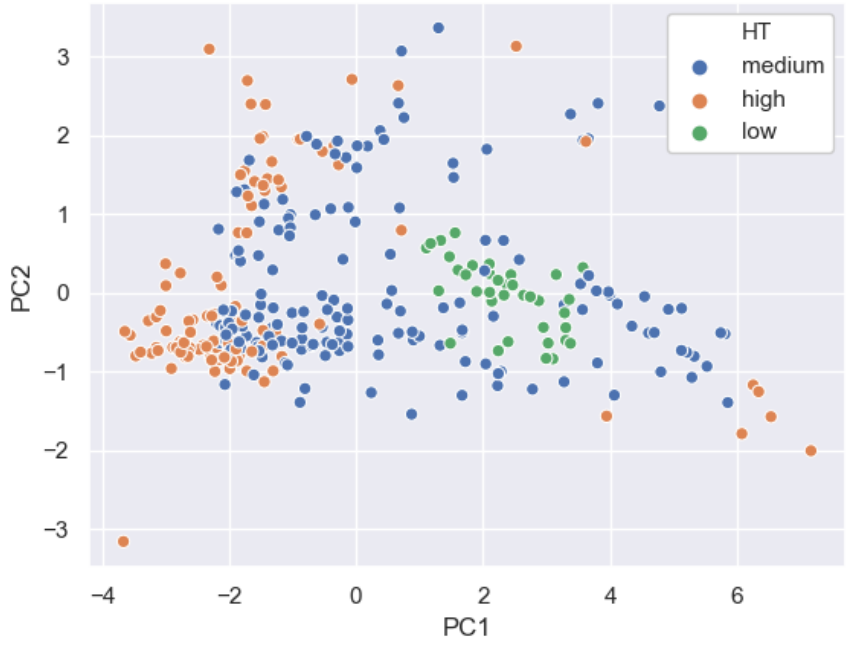

Fig. 4. Projection of samples onto the plane defined by PC1-PC2

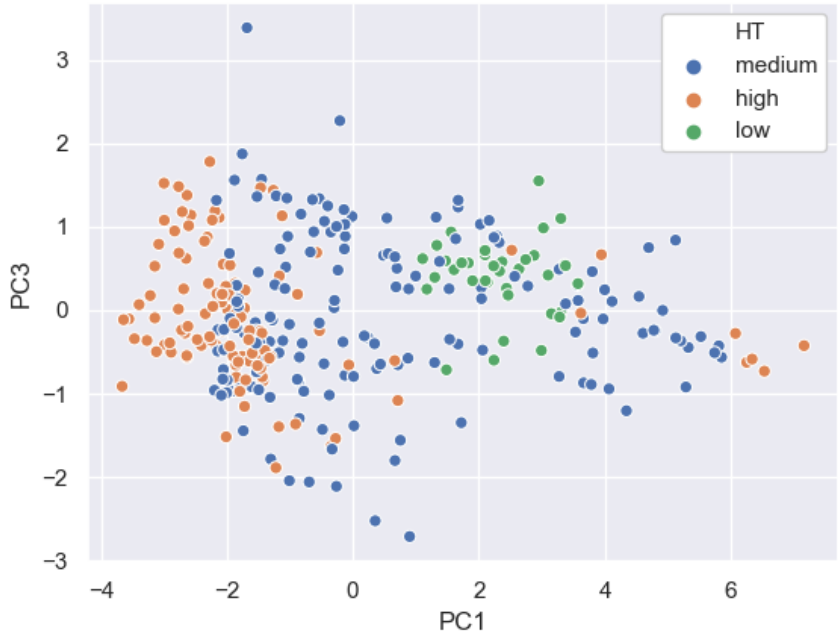

Fig. 5. Projection of samples onto the plane defined by PC1-PC3

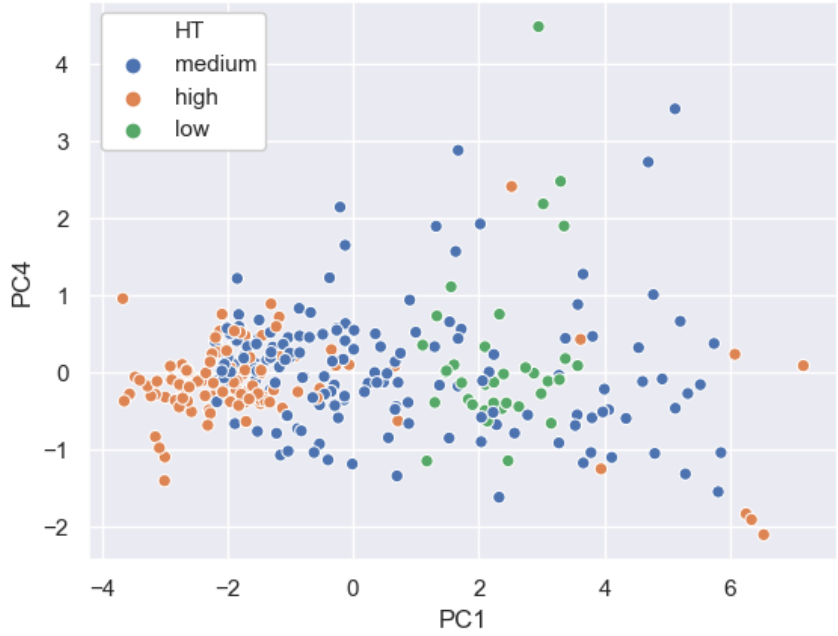

Fig. 6. Projection of samples onto the plane defined by PC1-PC4

The first four components explain 92.6% of the total variance. Therefore, only these components have been included as input variables of the classification model.

3.2. Support vector machine

Many SVM models with different kernel functions have been tested. Finally, SVM algorithm with RBF kernel function was applied in this work. The model employs two hyperparameters C and γ, which have a strong influence on its performance.

C is the penalty parameter of the error term. It controls the trade-off between errors on training data set and margin maximization. Models with larger values of C are better at classifying all training points correctly but have a smaller margin of separation (are more prone to overfitting). Model with smaller C have a larger margin of separation, but allow for a higher classification error.

γ parameter determines the shape of the decision boundary. With the increase of γ, the decision boundary becomes more irregular and bends at each example (the classifier tries to perfectly fit the training data). In the case of smaller values of the gamma coefficient, the decision boundary is characterized by a smoother course (Geron, 2017).

Model hyperparameters *C* and γ were estimated using grid search procedure (simple grid search). This technique consists in preparing a list of values of various parameters. Then the algorithm assesses the effectiveness of the model for each combination of these values, so as to determine the optimal configuration.

The parameters of the SVM model were optimized by cross-validated grid-search with 10 folds. It means that overall train set was divided into 10 subsets. During the 10 iterations, 9 subsets were used in the learning process, 1 was applied to assess the model efficiency. The overall efficiency of the model was estimated by determining the mean value for 10 iterations. Then, the process was repeated for all C and γ values of the grid. It was proved that when C=1 and γ=1 the model performance was the best.

The efficiency of the determined classifier is assessed by comparing the compliance of the estimated labels and the actual labels of individual samples. Table 1 presents the confusion matrix of SVM model (determined on test set). Correct classifications are located in the diagonal of the table.

Table 1. Confusion matrix of SVM model

n=78		Predicted label		
		'L'	'M'	'H'
Actual label	'L'	9	0	0
	'M'	0	38	3
	'H'	0	3	25

Based on the values given in the confusion matrix, the overall accuracy of the SVM model was determined as the ratio of correctly qualified samples to the total number of samples. The accuracy of the developed model was 92.3%, which means a satisfactory result.

In practice, besides accuracy, other measures are used to assess classification models, as (Geron, 2017):

- precision

$$PPV = \frac{TP}{TP+FP} \tag{5}$$

- recall:

$$TPR = \frac{TP}{TP+FN} \tag{6}$$

and *F*1, being combination of precision and recall:

$$F1 = 2 \cdot \frac{PPV \cdot TPR}{PPV+TPR} \tag{7}$$

where: *TP* - the number of cases actually belonging to a determined group and classified to it correctly, *FP* – the number of cases classified to a determined group but not belonging to it, *FN* - the number of cases belonging to a determined group, but classified into other groups.

Values of the above-described measures for particular classes ('L', 'M', 'H'), determined for the test set, were collected in Table 2.

Table 2. Evaluation of SVM model

Class	Precision	Recall	F1-score
'L'	1.00	1.00	1.00
'M'	0.93	0.93	0.93
'H'	0.89	0.89	0.89

All records belonging to the 'L' class have been identified as samples with a low hemispherical temperature of ash. None of the samples that have 'M' or 'H' label was classified in 'L' category. The values of precision and recall statistics (and consequently F1-score) for the 'L' class are therefore

equal to 1. It means that the presented algorithm very effectively identifies coal samples with the lowest HT.

Nevertheless, few samples with a medium hemispherical temperature of ash were incorrectly classified in category 'H'. Also, several records characterized by high HT were identified by algorithm as samples belonging to category 'M'. The values of precision and recall statistics for 'M' and 'H' classes were therefore less than 1.

4. IT tool for classifying coal blends for specific industrial applications

4.1. Description of IT tool

The presented system allows to analyze the behavior of AFTs in terms of creating blends in various proportions with different coals. In addition, the system indicates the most optimal blend combinations for a given industrial application. The diagram in Fig. 7 shows the idea of a proposed solution.

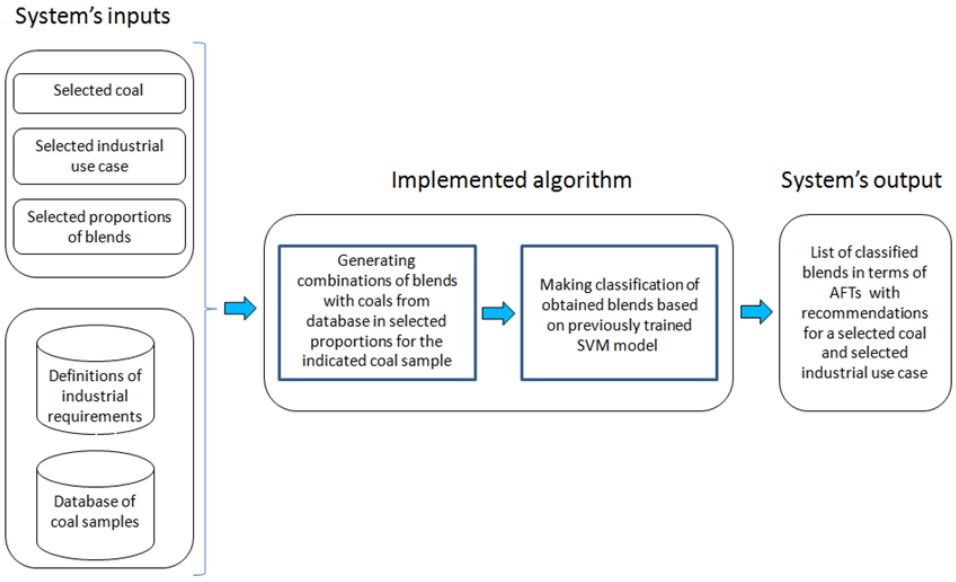

Fig. 7. The idea of a system for automatic coal blends analysis in terms of AFTs

The system uses a coal samples database and a knowledge base related to industrial applications and their requirements for AFTs. Before starting the program, the user can select a coal sample for which the algorithm will look for the components of the blends, as well as the proportions of blends and the specific industrial application for which the recommendation is issued by the system. The algorithm in the first step generates all combinations of blends of selected coal with other coals from the database in selected proportions. The oxide composition of the ash of the blend is calculated from the weighted average of the components of the blend. The algorithm then uses the previously trained SVM model to give blends appropriate labels. The final step of the algorithm is to select coal samples with labels compatible with the industrial application selected prior to running the algorithm and issue a recommendation on this basis. The system at the output indicates the recommended blend compositions. In practice, when selecting coal for blends, other factors, such as price or other key parameters of coal, are often taken into account. The system allows the analysis of how AFTs change in combinations of blends with individual coals. In this case, it is possible to ignore the automatic recommendation and focus on the analysis of the formation of AFTs in various proportions in blends with the coal that we are interested in (e.g. having the desired price). In the next chapter the practical use of the tool developed by the authors is described.

4.2. Use case - selection of blends to the pulverized boiler

The following practical problem has been defined: The particular coal was selected for combustion in a pulverized boiler. This coal is characterized by the medium fusion temperature of ash (determined in

laboratory test), whereas the AFTs of the coal used for combustion in pulverized boilers should be the highest possible (these samples should belong to the 'H' class). Therefore, it is necessary to compose blends with selected coal which properties meet the assumed requirements – high ash fusion temperature. The content of the sample database is presented in Table 3. Coals from 19 active mines in the Upper Silesian Coal Basin and two samples of low-quality coals were selected as an example. The selected samples were the most representative for the mines concerned. Low quality coals have been selected due to the often-used practice of blending coals of poor quality, but low price to reduce fuel costs.

Table 3. Coals forming the basis for creating blends (L.Q.C – low quality coal)

sample	SiO_2, %	Al_2O_3, %	Fe_2O_3,%	CaO, %	MgO, %	Na_2O, %	K_2O, %	SO_3, %	TiO_2, %
Mine 1	36.56	28.3	8.72	7.56	4.08	1.38	2.04	6.44	1.2
Mine 2	40.94	24.18	9.62	7.62	4.45	1.16	1.72	7.55	1.2
Mine 3	38.15	25.99	8.67	7.85	4.34	1.28	1.91	7.99	0.94
Mine 4	43.88	24.49	9.83	5.97	3.48	1.85	2.19	4.56	0.8
Mine 5	48.95	24.74	13.14	2.31	2.02	0.93	3.03	2.28	0.97
Mine 6	46.34	32.96	4.34	3.48	2.08	1.52	2.76	2.23	1.38
Mine 7	34.07	25.1	7.42	11.05	6.08	2.08	1.31	9.11	0.97
Mine 8	32.88	24.93	10.29	9.49	5.26	1.37	1.44	11.31	0.8
Mine 9	46.25	28.21	6.43	4.42	2.96	0.64	2.8	4.55	1.11
Mine 10	48.46	27.74	6.98	3.49	2.85	0.65	2.82	4.15	1.04
Mine 11	39.23	28.57	8.27	5.92	3.58	4.64	1.23	6.06	1.3
Mine 12	41.77	29.07	7.02	5.81	3.24	2.77	1.42	6.31	1.22
Mine 13	49.66	26.21	11.17	3.29	2.27	0.53	2.12	2.51	1.22
Mine 14	47.9	27.7	10.76	2.83	1.81	2.78	2.63	1.45	1.03
Mine 15	18.94	14.27	13.54	19.71	11.17	2.68	0.83	16.72	0.56
Mine 16	15.11	12.8	17.78	17.99	10.98	3.25	0.88	18.74	0.36
Mine 17	29.34	16.37	13.84	16.85	8.91	2.56	1.43	8.63	0.59
Mine 18	39.2	29.69	9.78	5.45	3.25	2.49	1.55	5.5	1.27
Mine 19	28.37	19.18	23.59	7.69	5.28	2.74	1.38	9.59	0.88
L.Q.C 1	53.86	23.97	6.46	3.16	3.03	1.55	3.03	2.89	1
L.Q.C 2	56.93	26.6	5.01	1.49	1.99	0.79	3.48	1.43	1.09

The selected coal comes from mine 16. Blends consist of two components. The proportions of blends have been set as 90/10, 10/90, 70/30, 30/70, 50/50.

The application of the described IT tool allowed to select the following blends with a high AFTs and containing coal from mine 16. Table 4 shows coal blends selected by proposed algorithm.

Table 4. Selected coal blends for pulverized boilers

No	Component 1		Component 2	
	Mine	Proportion, %	Mine	Proportion, %
1	16	10	6	90
2	16	30	6	70
3	16	10	10	90
4	16	30	10	70
5	16	10	L.Q.C 2	90
6	16	30	L.Q.C 2	70
7	16	50	L.Q.C 2	50

The algorithm combined coal from mine 16 with coal from mines 6 and 10 and low quality coal 2. In the selected blends, coal from mine 16 was rather an additive than a basic ingredient.

Coal from mine 16 is characterized by high content of SO_3 and CaO in ash. These components affect the lowering of the ash fusion temperature, which is an unfavorable phenomenon in the case of

pulverized boilers. To enable the combustion of coal from mine 16 in this type of boilers, this coal should be blended with samples containing small amounts of CaO and SO_3 in the ash. Moreover, since the content of CaO and SO_3 in the ash of the sample from the mine 16 is very high, this coal should not be the primary component of the blend.

The results of the optimization algorithm are consistent with the knowledge contained in the literature. It is generally known that Fe_2O_3, CaO, MgO, SO_3, Mn_3O_4 reduce the ash fusion temperature (Róg, 2003; Song et al., 2010).

5. Conclusions

- Chosen machine learning technique Support Vector Machine (SVM) was used to develop a tool which task was to classify coal samples due to the hemispherical temperature (HT). Three categories of coal have been distinguished - samples with low, medium and high HT. Each of these classes has a specific suitability for various industrial applications (gasification and combustion processes).
- Contents of nine different oxides (SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, TiO_2, Na_2O, K_2O, SO_3) in ash were the input variables of analyzed problem. In order to reduce dimensionality of data set PCA analysis was applied. It was observed, that the first four principal components explained 92.6% of total variance. Therefore, it was decided to keep these components and apply them in next analysis steps.
- Projections of the analyzed coal samples onto the planes determined by particular pairs of principal components indicated that the classes are quite well separated from each other.
- Data after PCA transformation was used to develop the SVM classification model. The accuracy of this model was 92.3%.
- IT tool that enable selection of coal blends for a specific type of boiler/reactor in terms of ash hemispherical temperature was developed. The SVM classifier was the core of the tool. It allowed to determine which coals and in what proportions should be blended with a given sample in order to obtain an oxide composition ensuring the desired HT.
- The presented tool was used to analyze the behavior of HT in blends consisting of coals from the Upper Silesian Coal Basin. However, it can also be adapted to other coal basins.

Acknowledgments

This work was supported by the Ministry of Science and Higher Education, Poland [grant number 11326019].

References

ALEKHNOVICH, A. N., BOGOMOLOV, V. V., 2010. *Use of coal blends at thermal power plants*. Power Technology and Engineering 44(3), 213–219, DOI: 10.1007/s10749-010-0167-3

BAHRIN, D., 2009. *Predicting Coal Ash Fusion Temperature of South Sumatera's Blended Coal with Coal Blending Simulation and Laboratory Analysis*. SISEST - RUSNAS PEBT

CARPENTER, A., 1995. *Coal blending for power stations*. IEA Coal research, United Kingdom

CHAWLA, N., BOWYER, K., HALL, L., KEGELMEYER, W., 2002. *SMOTE: synthetic minority oversampling technique*. J. Artif. Intell. Res. 16, 321–357

COLLOT, A.G., 2006. *Matching gasification technologies to coal properties*. International Journal of Coal Geology 65(3), 191-212, DOI: 10.1016/j.coal.2005.05.003

CORTES, C., VAPNIK, V., 1995. *Support-Vector Networks. Machine Learning*, 273–297, DOI: 10.1023/A:1022627411411

GERON, A., 2017. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly Media

GRAY, V.R., 1987. *Prediction of ash fusion temperature from ash composition for some New Zealand coals*. Fuel 66, 1230-1239, DOI: 10.1016/0016-2361(87)90061-5

HUGGINS, F. E., KOSMACK, D. A., HUFFMAN, G. P., 1981. *Correlation between ash-fusion temperatures and ternary equilibrium phase diagrams*. Fuel 60, 577-584, DOI: 10.1016/0016-2361(81)90157-5

JAK, E., 2002. *Prediction of coal ash fusion temperatures with the F*A*C*T thermodynamic computer package*. Fuel 81, 1655–1668, DOI: 10.1016/S0016-2361(02)00091-1

KARIMI, S., JORJANI, E., CHELGANI, S. C., 2014. *Multivariable regression and adaptive neurofuzzy inference system predictions of ash fusion temperatures using ash chemical composition of us coals*. Journal of Fuels, DOI: 10.1155/2014/392698

LI, F., MA, X., XU, M., FANG, Y., 2017. *Regulation of ash-fusion behaviors for high ash-fusion-temperature coal by coal blending*. Fuel Processing Technology 166, 131-139, DOI: 10.1016/j.fuproc.2017.05.012

LUXSANAYOTIN, A., PIPATMANOMAI, S., BHATTACHARYA, S., 2010. *Effect of mineral oxides on slag formation tendency of Mae Mohlignites Songklanakarin*. Journal of Science and Technology,32(4), 403-412

MIAO, S., JIANG, Q., ZHOU, H., 2016. *Modelling and prediction of coal ash fusion temperature based on BP neural network*. MATEC web of conferences, 40, 5-10

MOHAMED, A., 2017. *Comparative Study of Four Supervised Machine Learning Techniques for Classification*. International Journal of Applied Science and Technology, 7(2), 5-18

PATTERSON, J.H., HURST, H.J., 2000. *Ash and Slag Qualities of Australian Bituminous Coals for Use in Slagging Gasifiers*. Fuel 79, 1671-1678, DOI: 10.1016/S0016-2361(00)00032-6

PORADA, S., DZIOK, T., CZERSKI, G., GRZYWACZ, P., STRUGAŁA, A., 2017. *Examinations of Polish brown and hard coals in terms of their use in the steam gasification process*. Mineral resources management, 33, 15-34, DOI 10.1515/gospo-2017-0007

RÓG, L., 2003. *Effects of petrographic and chemical structure of coal on the fusion temperature of ash.* Research Reports of Central Mining Institute. Mining & Environment, 1, 73-96

SASI, T., MIGHANI, M., ORS, E., TAWANI, R., GRABNER, M., 2018. *Prediction of ash fusion behavior from coal ash composition for entrained-flow gasification.* Fuel 176, 64-75, DOI: 10.1016/j.fuproc.2018.03.018

SEGGIANI, M., 1999. *Empirical correlations of the ash fusion temperatures and temperature of critical viscosity for coal and biomass ashes*. Fuel 78, 1121–1125, DOI: 10.1016/S0016-2361(99)00031-9

SEGGIANI, M., PANNOCCHIA, G., 2003. *Prediction of coal ash thermal properties using partial least-squares regression*. Industrial & engineering chemistry research, 42(20), 4919-4926, DOI:10.1021/ie030074u

SHEN, M., QIU, K., ZHANG, L., HUANG, Z., WANG, Z., LIU, J., 2015. *Influence of Coal Blending on Ash Fusibility in Reducing Atmosphere*. Energies 8, 4735-4754, DOI: 10.3390/en8064735

SHI, W.-J., KONG, L.-X., BAI, J., XU, J., LI, W.-C., BAI, Z.-Q. et al., 2018. *Effect of CaO/Fe2O3 on fusion behaviors of coal ash at high temperatures*. Fuel Processing Technology 181, 18-24, DOI: 10.1016/j.fuproc.2018.09.007

SONG, W. J., TANG, L. H., ZHU, X. D., WU, Y. Q., ZHU, Z. B., 2010. *Effect of Coal Ash Composition on Ash Fusion Temperatures*. Energy Fuels 24, 182-189, DOI: 10.1021/ef900537m

TAMBE, S. S., NANIWADEKAR, M., TIWARY, S., MUKHERJEE, A., DAS, T. B., 2018. *Prediction of coal ash fusion temperatures using computational intelligence based models*. International Journal of Coal Science & Technology 5(4), 486-507, DOI: 10.1007/s40789-018-0213-6

THOMAS, L., 2002. Coal Geology. John Wiley & Sons

TILLMAN, D. A., DUONG, D., STANLEY, N., 2012. *Harding Solid Fuel Blending Principles, Practices, and Problems*. Butterworth-Heinemann/Elsevier, United States

VAN DYK, J.C., KEYSER, M.J., VAN ZYL, J.W., 2001. *Suitability of feedstocks for the Sasol–Lurgi fixed bed dry bottom gasification process*. Gasification technology conference, Gasification Technologies Council, Arlington, VA, USA, Paper 10-8

VAPNIK, V., 1998.*Statistical learning theory.* Willey, New York

VASSILEVA, CH., VASSILEV, S., 2002. *Relations between Ash-Fusion Temperatures and Chemical and Mineral Composition of Some Bulgarian Coals*. Comptes Rendus de l'Academie Bulgare des Sciences, 55 (6), 61-66

WANG, P., MASSOUDI, M., 2013. Slag B*ehavior in Gasifiers. Part I: Influence of Coal Properties and Gasification Conditions*. Energies 6, 784–806, DOI: 10.3390/en6020784

YAZDANI, S., HADAVANDI, E., CHELGANI, S. C., 2018. *Rule-Based Intelligent System for Variable Importance Measurement and Prediction of Ash Fusion Indexes*. Energy & Fuels 32(1), 329-335, DOI: 10.1021/acs.energyfuels.7b03280

YÖRÜKOĞLU, M., 2017. *Coal blending for thermal power stations*. Madencilik 56(3), 109-116