Applied Soft Computing 64 (2018) 109–125

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A new Ensemble based multi-agent system for prediction problems: Case study of modeling coal free swelling index

Mehdi Golzadeh [a], Esmaeil Hadavandi [b,*], S. Chehreh Chelgani [c,*]

[a] Faculty of Engineering, Farabi Campus, University of Tehran, Tehran, Iran
[b] Department of Industrial Engineering, Birjand University of Technology, Birjand, Iran
[c] Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, United States

ARTICLE INFO

Article history:
Received 18 April 2017
Received in revised form 8 December 2017
Accepted 10 December 2017
Available online 14 December 2017

Keywords:
Multi-agent system
Neural network ensemble
Negative correlation learning
Free swelling index
Coke

ABSTRACT

In this article, a new ensemble based multi-agent system called "EMAS" is introduced for prediction of problems in data mining. The EMAS is constructed using a four-layer multi-agent system architecture to generate a data mining process based on the coordination of intelligent agents. The EMAS performance is based on data preprocessing and prediction. The first layer is dedicated to clean and normalize data. The second layer is designed for data preprocessing by using intelligent variable ranking to select the most effective agents (select the most important input variables to model an output variable). In the third layer, a negative correlation learning (NCL) algorithm is used to train a neural network ensemble (NNE). Fourth layer is dedicated to do three different subtasks including; knowledge discovery, prediction and data presentation. The ability of the EMAS is evaluated by using a robust coal database (3238 records) for prediction of Free Swelling Index (FSI) as an important problem in coke making industry, and comparing the outcomes with the results of other conventional modeling methods Coal particles have complex structures and EMAS can explore complicated relationships between their structural parameters and select the most important ones for FSI modeling. The results show that the EMAS outperforms all presented modeling methods; therefore, it can be considered as a suitable tool for prediction of problems. Moreover, the results indicated that the EMAS can be further employed as a reliable tool to select important variables, predict complicated problems, model, control, and optimize fuel consumption in iron making plants and other energy facilities.

1. Introduction

Over the last few decades, ironmaking industry has significantly growth and the role of coke as a fuel, a chemical reducing agent and a permeable support has become more important [1–5]. As further improvements in coke quality are demanded, it is critical to evaluate coking coal properties with accurate analyses. Free Swelling Index (FSI) as an ASTM standard test is the most robust analysis that can assess coke-ability of coal particles [6]. FSI measures the increase in volume of powdered coal when heated under certain conditions. Based on ASTM D 720 coking capability of coal particles can classify into three groups: weak (0–2), medium (>2-4) and strong (>4-9) as a standard FSI index [6]. Adjusting the proper temperature rate in the blast furnace, weathering coal samples (surface and deep oxidation), and the proper size for coal particles (the amount of fine particles has to be kept at a minimum level) are problems associated with FSI evaluation [7–9]. Intelligent models based on comprehensive data base of coal properties can be applied to study and predict coke quality more accurately and can overcome the mentioned analysis challenges. Moreover, a realistic accurate model can be used to scale up laboratory result into industrial level and better control the factors that may impact fuel consumption in a blast furnace.

Considering these critical challenges, several statistical models have been proposed by different investigations to predict FSI by using multivariable regressions [7,10]; however, having an expert system to assess coal properties and predict FSI would be quite applicable for fuel processing industry. Coal particles have complex structures and regression models cannot examine these complicated relationships between coal parameters and select the most effect ones for FSI modeling where pervious investigations just could take all input variables without considering possible interactions [7,10–12]. Thus, a data mining (DM) process with a

* Corresponding authors.
E-mail addresses: es.hadavandi@birjandut.ac.ir (E. Hadavandi), schehreh@umich.edu (S.C. Chelgani).

https://doi.org/10.1016/j.asoc.2017.12.013
1568-4946/

comprehensive database can model FSI. There are many DM models in different sectors of health, management, and engineering which are implemented as a multi-agent system (MAS) [13]. The characteristic of MAS and its distributed nature make it as one the most typical DM systems [14]. A MAS can use to reduce the complexity of different DM methods by creating modular components where each component is in charge of its own subtasks, and the whole system is going to achieve a common goal [15,16]. Each agent uses a suitable method to overcome its own task and applies a tuning process to select its appropriate parameters [17]. In other words, MAS is a common approach to implement decision support systems due to its flexibility and unique features [18,19].

Moreover, decentralized control of MAS implies that individual agent in this architecture operates autonomously and somehow self-deterministic [20]. This feature also makes the MAS robust because the system operation will continue when individual agent has crashed or does not respond in a reasonable period of time. Another feature of a decentralized control is extendibility which means that whenever a new functionality was needed, it would be possible simply by adding new agents to the system. In the context of sharing resources and expertise, the MAS also has some particular advantages to offer with respect to other DM methods [20]. All of these specifications lead us to conclude that this model would be beneficial for predicting the coke quality index (FSI).

In reality, the most of industrial problems are too complicated to be solved with a simple prediction model. Therefore, various investigations were conducted to combine different types of prediction models to solve a specific problem [21,22]. A combination of prediction models and generation of an ensemble of experts would assist to model every single aspect of a problem and cover all behaviors of a system which leads to a better overall prediction and a high performance [23]. The main idea behind ensemble models is to simplify prediction of complicated problems (to generate relative simple tasks and to solve those tasks with its own expert) [24]. Several studies indicated that an ensemble learning method is an approach to enhance prediction accuracy for complicated problems (having highly complex trends with high dimensional feature sets)[25,26]. Tang et al. [27] proposed an ensemble model to crude oil price forecasting. The results of their investigation showed that the ensemble model outperformed other conventional methods based on prediction accuracy.

In an ensemble learning method, creating experts and adding diversity are two essential issues [28]. Different mechanisms are reported in various investigations for development and differentiate experts in prediction of problems [29]. The main purpose of diversifying experts for an ensemble model is to cover different parts of a problem space [30]. A common approach to build an ensemble model is altering the learning algorithm [31]. One of the common methods in this area is a neural network ensemble (NNE) that has been developed by Hansen and Salamon [32]. Various studies presented different developments of NNE method [21,28,33]. One of the most widely used NNE models is the negative correlation learning (NCL) [34] that reduces error correlation of experts to the lowest level [35]. In the NCL, a penalty term is added to objective function of the training process of experts and negatively generates correlated experts; therefore, NCL directly optimizes the diversity in experts of ensemble model. Weiguo et al. [36] proposed a niching evolutionary algorithm with an adaptive negative correlation learning for training of a NNE. Their proposed model was evaluated with a number of classification problems and compared with related ensemble learning algorithms. Their results showed that the NCL based method could be used to design a satisfactory NNE and outperform related works.

In this article, a FSI prediction model is developed by a wide range of American coal samples (3238 samples from 17 different states (Fig. 1)) using a new ensemble based multi-agent system (EMAS). To design and develop an agent-based system, different methodologies has been used based on various methods; Gaia [37], MaSE [2], Agent ULM [38], Prometheus [3], Tropos [39] etc. The generated EMAS has been modeled through Prometheus methodology and Prometheus Design Tool [8]. The generated toolkit provides some artifacts in eclipse development environment that apply to control consistency of the model and also some tools to export reports as an HTML page. Moreover, an implementation of the

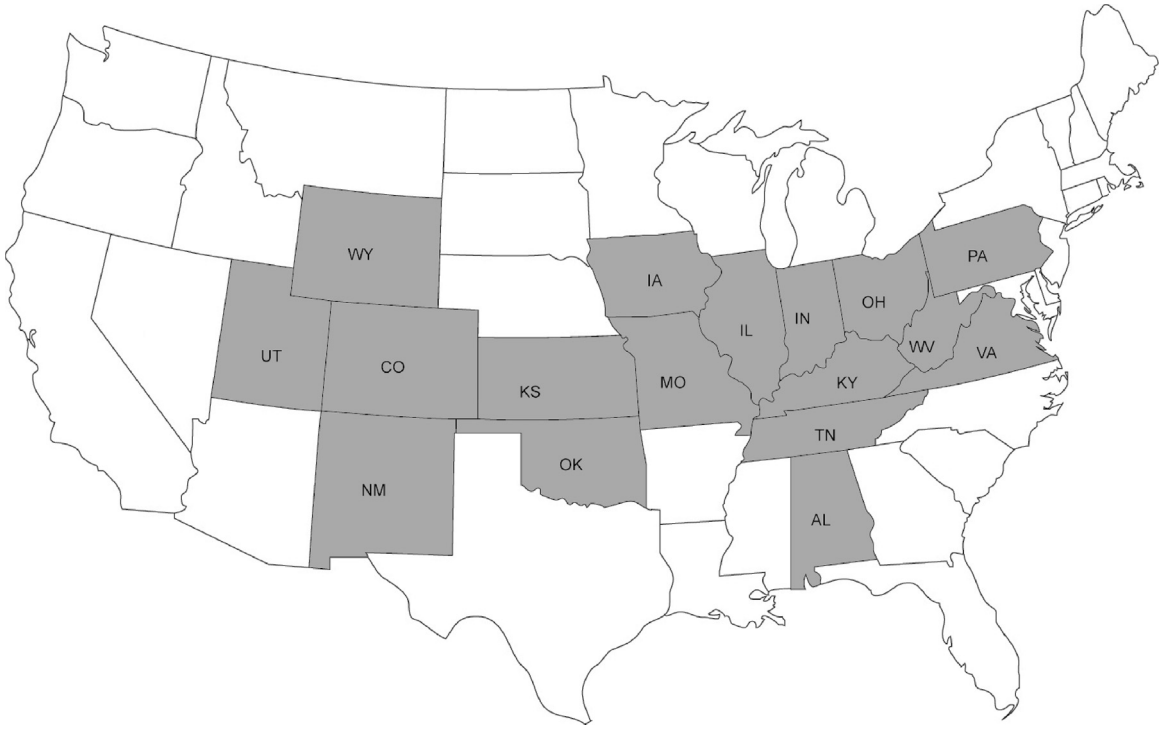

Fig. 1. American coal samples from 17 different states.