Intelligent System for Diagnosis of a Three-Phase Separator

Irina Ioniţă
Informatics, Computer Science, Mathematics and Physics,
Petroleum-Gas University of Ploieşti, Ploiești, Romania
tirinelle@yahoo.com

Liviu Ioniţă
Informatics, Computer Science, Mathematics and Physics,
Petroleum-Gas University of Ploieşti, Ploiești, Romania
lionita@gmail.com

Abstract

Intelligent systems for diagnosis have been used in a variety of domains: financial evaluation, credit scoring problem, identification of software and hardware problems of mechanical and electronic equipment, medical diagnosis, fault detection in gas-oil production plants etc. The goal of diagnosis systems is to classify the observed symptoms as being caused by some diagnosis class while advising systems perform such a classification and offer corrective remedies (recommendations). The current paper discuss the opportunity to combine more intelligent techniques and methodologies (intelligent agents, data mining and expert systems) to increase the accuracy of results obtained from the diagnosis of a three-phase separator. The results indicate that the diagnosis hybrid system benefits from the advantages of each module component: intelligent agent module, data mining module and expert system module.

Keywords: diagnosis, intelligent agents, data mining, expert system

1. Introduction

The daily operations of an oil and gas production system require many decisions, at different levels of organization and making adequate decision is a difficult task because of the economic and environmental risks occurrence if abnormal situations are not detected and managed appropriately.

Building a diagnosis system requires many years of experience in domain, a consolidate team of workers (engineers, operators, developers etc.) and a significant volume of data for implementation and testing phase of the system, offline simulations.

A diagnosis system is one that is capable of identifying the nature of a problem by examining the observed symptoms (the input of the system) and to produce a diagnosis report with or without an explanation or justification (the output of the system). In many applications of interest it is desirable for the system and quite necessary, to not only identify the possible causes of the problem but also to offer suitable solutions (recommendations) to remediate the anomalies that occurred.

Identifying the probable cause associated to a fault is performed based on the adequate knowledge of the analyzed domain. The necessary knowledge may be represented in a standard form (i.e. diagnostic rules for expert systems) into the system by domain experts (heuristic knowledge). In other case, machine learning techniques are used for automated knowledge acquisition.

The authors proposed a hybrid diagnosis system used in oil industry (*DiagAgentExpertDM)*, as a result of their research work regarding the application on intelligent techniques in industrial field (Ioniță & Ionită, 2013, Ioniță, 2014, Ioniță & Ioniță, 2015). The data used for training and testing the system were collected from a gas-oil separation plant (GOSP). The task of data mining (DM) component is to provide the best classification model (as decision tree model). The expert system uses the induction rules acquired from the DM component and gives an evaluation report regarding the functioning of a gas-oil separator, as a group of three attributes: alarm, cause, solution. The output of expert system is processed by Diagnosis Agent which communicates with Operator Agent (the main agents from the multi-agent system GOSP-MAS (Ioniță, 2014)).The diagnosis system *DiagAgentExpertDM* possess the most important features of each "intelligent" component: adaptive learning from agents, knowledge from expert system module and predictive behavior from data mining component.

After a short description of the diagnosis problem and the motivation of the authoress in choosing this subject (in introduction), section 2 of this paper discusses about various approaches to the design of intelligent diagnosis systems, emphasizing their strengths and limitations. This analysis also suggests interesting possibilities for hybrid techniques that combine different approaches (intelligent agent, data mining and expert system), the proposed diagnosis system for a three-phase separator being presented in section 3 of the article. Section 4 contains the results and the remarks, and finally, conclusions regarding the research work are mentioned.

2. Intelligent techniques for diagnosis in oil industry

A definition for agent concept has been proposed by Wooldridge in (Wooldridge, 2002): "An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives". An intelligent agent has one or more goals and might communicate with other agents to achieve the proposed goals. He "knows" its environment by perceptions and takes these observations into account when it decides which actions to perform. The main characteristics of an agent are: autonomous (internal states and goals defines the autonomy), proactive (they have a goal to achieve), reactive (they react to changes in the environment when they find it necessary), social (they collaborate and negotiate with other agents), robust (agents can recover from failure), flexible (they have several ways to achieve a goal).

In recent years, intelligent agents were mentioned in several research papers for a wide area of application in industrial field: fault management of industrial processes (Cerrada, Cardillo, Aguilar, Faneite, 2007) (García, Leme, Pinto & Sanchez-Pi, 2012), intelligent control and optimization of production wells (Engmo & Hallen, 2007), intelligent supervision of petroleum processes (Atalla & Taylor, 2004) (Canelón, Dávila, Morles & C. La Hechicera, 2009) (Ioniță & Ioniță, 2014) (Ioniță, 2014), dynamic scheduling in petroleum process (Aissani & Beldjilali, 2009). The results indicate that agents represents an adequate alternative to solve complex problems due their capabilities mentioned above. Using this technology, the production costs were reduced and the response time of industrial system was minimized.

Data mining is considered the process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems (Piatetsky-Shapiro, Brachman, Khabaza, Kloesgen & Simoudis, 1996). In oil and gas industry everyday operation produces large and complex volume of data that it's difficult to handle using data management tools. Traditional sources, such as equipment monitoring and maintenance records provide data that is often stored in data warehouses and used as needed. To support rapid decision-making in key areas such as reservoir modelling, drilling and production optimization, gas and oil companies need tools that integrate and synthesize various data sources into a unified whole and to discover knowledge. It is obvious that an oilfield operator can create strategic advantages over its competitors applying data mining techniques to get important insights from the data it collects from measurements or records provided by industrial equipment (sensors, actuators, control valves, pumps etc.). Data mining get the answer to many questions such as: how to select the best candidate well for simulation, how to adjust the parameters of a three-phase separator to assure the proper condition for separation process, how to avoid the anomalies or how to remediate them in a gas-oil production plant, what are the decisions to be taken in order to increase the profit of a gas-oil company etc. (Piatetsky-Shapiro, Brachman, Khabaza, Kloesgen & Simoudis, 1996) (Rabiei, Gupta, Cheong & Sanchez Soto, 2010) (Rabiei, Gupta, Cheong, & Soto, 2009). Applications of data mining referring to production operation process in oil field, around production operation control, production operation parameter prediction, the production process fault monitoring and production operation cost control are discussed in (Li, Li & Wu, 2009) (Ccoicca, 2013) (Aulia, Keat, Maulut, El-Khatib & Jasamai, 2010) (Badr, 2014) (Wang, Gao & Li, 2012). Data mining techniques come to meet organization to benefit from their enormous data with valuable capacities, unknown patterns that can be used in decision-making.

An alternative to build intelligent diagnosis systems is expert system technology as is mentioned in (Balakrishnan & Honavar, 1998) (Noureldin & Ruveta, 2002). Expert systems are known as interactive computer-based decision tools that use both facts and rules to solve complex decision problems based on knowledge acquired from a human expert in a certain domain. The main components of an expert system are: a knowledge base, an inference engine for rationing task and a working memory to hold the facts provided by the user and conclusions obtained during inference procedure. A knowledge engineer has the role to transform the experience of the human personal in knowledge and store it in knowledge base as rules or other standard format for knowledge representation (logic, semantic frameworks etc.). This process (knowledge engineering) can be expansive and time consuming. Applications of expert systems in gas and oil field cover the following categories: predicting, diagnosing, planning, monitoring and controlling (Ghallab, Badr, Salem & Tolba, 2013; Lokko, 2012; Ioniţă & Ioniţă, 2013; Kopliku & Condanni, 1994). Some of expert systems advantages are: availability, reliability (consistency), cost effective. Compared with intelligent agents, expert system are limited domain and are not always up to date, do not have the learning capacity. Engineers have to maintain and configure permanently the expert system, which is considered a disadvantage.

The motivation for choosing a special design method for intelligent diagnosis system is influenced by the sources of knowledge, the available amount of data and the representation form. In the next section, the authoress proposed a hybrid system for diagnosis of a three-phase separator taking into account the analysis of intelligent diagnosis systems presented above.

The purpose of the GOSP is to process the well flow into clean, marketable products such as: oil, natural gas or condensates. A number of additional utility systems are included, which are not part of the actual process but provide energy, water, air or some other utility to the plant (Devon, 2013).

3. The hybrid diagnosis system

In recent years, large companies (ie. oil and gas companies) realized that can be competitive only if the knowledge hidden in enormous data (daily transactions) is discovered in real time and used properly. For example, end users spend most of their time sorting and filtering the data, and little time searching significance and interpreting the discovered patterns. As a result, volumes of data go unused and important patterns that may help the decision-making remain undiscovered. According to the Gartner Group (Kyte, 2002), the amount of data collected by large companies doubles every year and knowledge workers can analyze only a small part of data. First task of knowledge workers is to extract knowledge and then to find the meaning of the discovered patterns and used them to improve the decision-making capabilities.

Automation acquisition of the valuable knowledge from the experts and operators, pattern interpreting and real-time decision development are the desirable characteristics of an intelligent diagnosis system.

Motivated by the advantages of intelligent techniques presented above, the authors propose a method to increase the efficiency of monitoring and diagnosis system: combination of intelligent agents with data mining techniques and expert system technology. The hybrid system will benefit from autonomous agents and their reactive character, from the heuristic reasoning of expert system, and from the ability to extract knowledge and increase accuracy of results using data mining module.

The authors' interest in knowledge-based systems (KBSs) and intelligent agents applied in oil industry was found in previous work (Ioniţă & Ionită, 2013; Ioniţă, 2014; Ioniţă & Ioniţă, 2015).

An example of multi-agent system applied in oil industry for diagnosis of a gas-oil separation plant is given in (Ioniţă, 2014). This intelligent diagnosis system named GOSP-MAS uses several agents that collaborate and communicate with each other in order to achieve the main goal: the diagnosis of a three-phase separator. The supervisor agent (Diagnosis Agent) has the main role in the diagnosis system GOSP-MAS, being responsible with generation of the diagnosis report and with forwarding it to other agent (Operator Agent). The human operator receives through GUI the diagnosis report from Operator Agent and after analyzing the recommendations listed for the generated alarms can execute the adequate operation to remediate the faults. To improve the quality of diagnosis process, the author proposed in (Ioniţă, 2014) an expert system (Expert-GOSP). A similar intelligent diagnosis system based on knowledge-based system technology is described in (Ioniţă, Ioniţă, 2013). More than that, to increase the accuracy of result and to refine the solutions/advices generated after diagnosis process a data mining module can be added and integrated in the intelligent diagnosis system. Application of data mining technique for diagnosis in oil industry is given in (Serapião & Bannwart, 2013) (Rabiei, Gupta, Cheong & Sanchez Soto, 2010) (Rabiei, Gupta, Cheong, & Soto, 2009).

Table 1 presents the analysis of three diagnosis systems (KBS-Diag, Expert-GOSP and *DiagAgentExpertDM*). The input for each diagnosis system is represented by the operating parameters associated to a three-phase separator. *DiagAgentExpertDM* uses over twenty attributes (table 2), the other two systems use less than fifteen attributes. The output of the diagnosis system is presented as a message to inform the user about the system state and was improved with more information regarding the alarms, causes and solutions. *DiagAgentExpertDM* is characterized by adaptive learning, knowledge discovery features and predictive behaviour.

Table 1. Comparison of diagnosis systems

Diagnosis system	Input	Output	What techniques are used for diagnosis?	Features	Software used to develop the system
KBS-Diag	Operating parameters (oil temperature, oil tank pressure, oil tank temperature, water temperature etc.)	A message regarding to the alarms occurred and the solutions given	Induction rules	Knowledge discovery	Visual Prolog
Expert-GOSP	Operating parameters	The diagnosis report built as a group of alarms, causes and solutions	Decision tree (induction rules as IF THEN ELSE)	Knowledge discovery	Exsys Corvid
DiagAgent ExpertDM	Operating parameters	The diagnosis report preprocessed by Diagnosis Agent	Data mining model (Decision tree) Induction rule for expert system Intelligent agents	Adaptive learning Knowledge discovery Predictive behaviour	Exsys Corvid Weka Prometheus

For the current hybrid system *(DiagAgentExpertDM)*, data from different sources and in various forms are preprocessed, in order to represent them in a unified way to be able to upload them in a learning module. Simultaneously, inconsistency tests are made to eliminate the measurement errors caused by improper calibration of transducers etc.

In condition of using a diagnosis method based on process history, the next step in fault identification referring to the gas-oil separation process is to scan the historical data. Retrieved data will possess a label with a certain priority which will be used in the next retrieval process. The learning module is supplied with multiple preprocessed data samples, in order to extract knowledge from them. For each data samples, a data mining algorithm will be applied (figure 1).

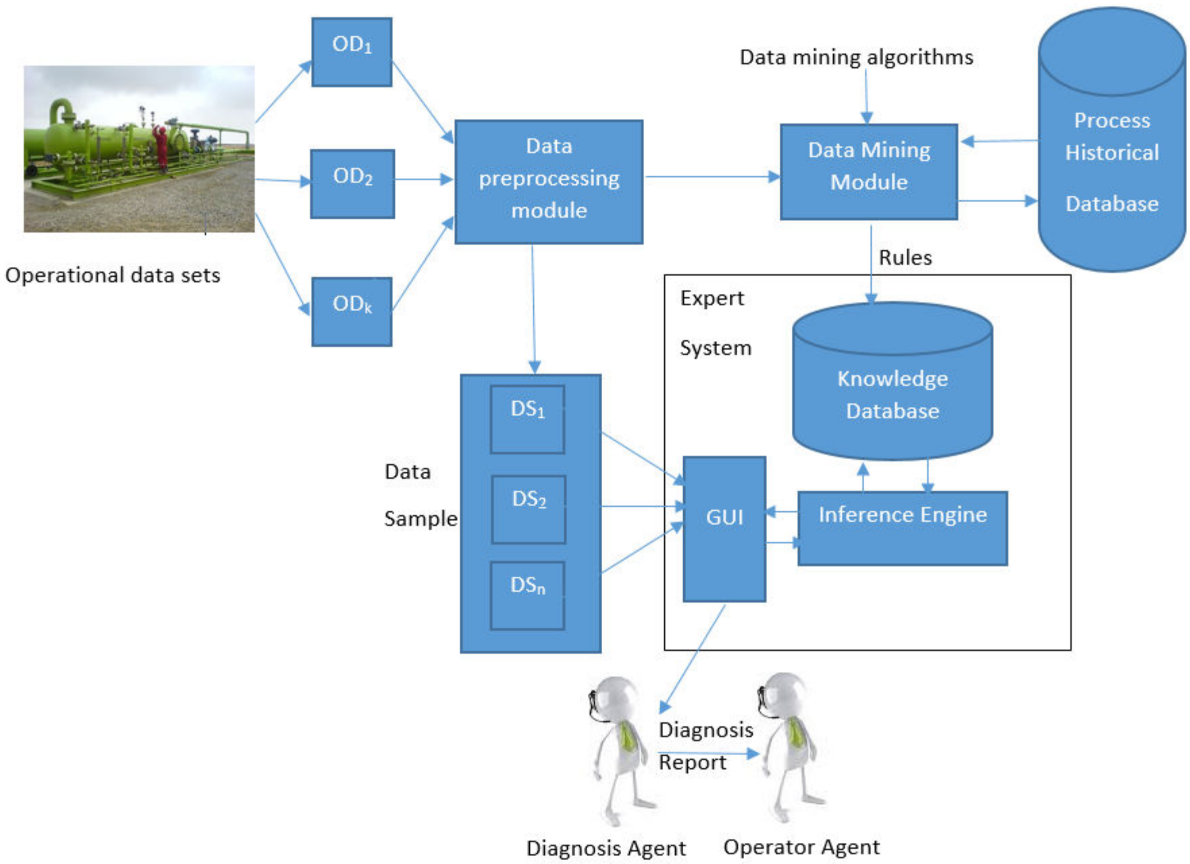

Figure 1. DiagAgentExpertDM architecture (adapted from Ioniță, 2014)

Decision trees were developed with Weka (Weka, 2016), and the obtained rules are implemented in the expert system knowledge base.

The induction rules are in the standard form IF THEN ELSE, as following:

```
IF measured_value_parameter1 >= accepted_value_min AND
Measured_value_parametru1<= accepted_value_max THEN decision=stable_system
[105/3];
```

The generated rules describe both normal state of the three-phase separator and the possible scenarios of faults. Each rule may have a confidence coefficient calculated as a ratio between the number of observation described by this rule (ie. 105) and the number of misclassified cases (ie. 3).

A concrete example, is given below:

```
IF:
[oil_temp_inside_separator] <=25
THEN:
[alarm] = "          Oil temperature inside the separator LOW"
[cause] = "          Cause: temperature control valve fault TCV-41002
or TCV-41003 or thermal insulation vessel damage"
[solution]  = "                  Solution:  verify  the  operating  state  of
temperature control valve TCV-41002/TCV-41003 or verify thermal isolation
vessel"
SET  [report.ADD]  "TEMPEARTURE ALARM"
SET  [report.ADD]  [alarm.VALUE]
SET  [report.ADD]  [cause.VALUE]
SET  [report.ADD]  [solution.VALUE]
```

More induction rules are presented in (Ioniță & Ioniță, 2015), as well as GUI for the expert system built with Exsys Corvid.

The designed system hold several knowledge bases, whose consistency is subsequently verified. The similar cases from these knowledge bases are analysed, merged, and the confidence coefficient is computed again based on confidence coefficients defined for individual knowledge base. The testing data sets were been collected from the analysed three-phase separator. The process history contains heuristic knowledge from domain experts and operators, technical data from specialized documentation and real-time data collected from field transducers.

The expert system generates two levels of process state (normal and abnormal state), being coded as following: red colour means abnormal functioning state, and green colour means a normal functioning state. The graphic interface provides the current condition that may cause alarms and possible occurred problems.

4. Results and Remarks

Five data mining models based on decision trees/induction rules were applied: J48, Simple CART, LMT, JRIP and REPTree (Witten, Frank & Hall, 2011).

J48 is an implementation of the well-known Quinlan algorithm (C4.5) (Quinlan, 1993), which is an improvement of the basic ID3 algorithm. This classification data mining model builds a decision tree for the given dataset, whose nodes represent discrimination rules acting on selective patterns by recursive partitioning of data, using depth-first strategy.

Simple CART (Classification and Regression Trees) (Timofeev, 2004) is a classification method which uses historical data to construct decision trees, when the number of classes are a priori known. The resulted decision tree are used to classify new data. An important practical characteristic of Simple CART is that the structure of its classification or regression trees is invariant with respect to monotone transformations of independent variables.

Logistic Model Tree (LMT) represents a combination between a standard tree structures with logistic regression functions at the leaves using posterior class probabilities in order to produce a single decision tree (Witten, Frank & Hall, 2011). LMT consists of a tree structure that is built up of a set of inner nodes and a set of leaves or terminal nodes in an instance space.

JRIP is a data mining model that implements repeated incremental pruning to produce error reduction (RIPPER) in Java, a prepositional rule learner, as proposed in (Cohen, 1995) and is based on the construction of a rule set in which all positive examples are covered. In this algorithm, the discovered knowledge is represented in the form of IF THEN ELSE prediction rules.

Reduced Error Pruning Tree (REPTree) is a simple procedure for learning and pruning decision trees and the main goal is to build a decision or regression tree using information gain as the splitting criterion and prunes trees using reduced-error pruning. This data mining model only sorts values for numeric attributes once.

The attributes used for this experiment are presented in table 2.

Table 2. Operating parameters of the three-phase separator

Attribute (operating parameter)	Type
oil_temp_inside_separator	NUMERIC
oil_temp_outlet_heat_exchanger_41HE002	NUMERIC
oil_temp_outlet_heat_exchanger_41HE003	NUMERIC
water_pressure_outlet_separator	NUMERIC
gas_pressure_outlet_separator	NUMERIC
oil_pressure_outlet_separator	NUMERIC
oil_level_separator	NUMERIC
water_level_separator	NUMERIC
alarm_temp_separator	NOMINAL

Attribute (operating parameter)	Type
alarm_temp_41HE002	NOMINAL
alarm_temp_41HE003	NOMINAL
alarm_pressure_water (APW)	NOMINAL
alarm_pressure_gas (APG)	NOMINAL
alarm_pressure_oil (APO)	NOMINAL
alarm_level_oil (ALO)	NOMINAL
alarm_level_water(ALW)	NOMINAL
cause_AT	NOMINAL
cause_APW	NOMINAL
cause_APG	NOMINAL
cause_APO	NOMINAL
cause_ALO	NOMINAL
cause_ALW	NOMINAL
solution_temp	NOMINAL
solution_PW (PW – pressure water)	NOMINAL
solution_PG (PG – pressure gas)	NOMINAL
solution_PO (PO – pressure oil)	NOMINAL
solution_LO (LO – level oil)	NOMINAL
solution_LW (LW – level water)	NOMINAL
diagnosis_report	NOMINAL

The source ARFF file (Attribute-Relation File Format) contains 530 records (only 292 of records are used to build the models, the rest of records are used in evaluation phase), the target variable of the model being *diagnosis_report* with two possible values (*normal_state/abnormal_state).*

After source file loading, the classifier is chosen (J48) and the model is run according to testing options. As a result, a decision tree is built and the statistics are calculated.

The procedure is the same in the case of the rest of data mining models (*Simple CART, JRIP, LMT* and *REPTree).*

The interpretation of evaluation measures (correctly classified instances, misclassified instances, Kappa Statistic, mean absolute error, root mean squared error) offer the analysis of performance associated to the data mining models (table 3). For example, J48 model classified 270 instances as correct classified which means that instances considered initial with the normal_state value for target attribute (65 instances) were classified as normal_state and the instances considered initial with abnormal_state value for target attribute (205 instances) were classified by model as abnormal_state. The rest of 22 instances were classified as incorrect (the initial value for the target attribute is not identic with the classified one).

The comparison of the results indicates a better classification in case of LMT, J48 and REPTree models (the Root mean squared erorr is respectively 0.23, 0.25 and 027).

Unlike JRIP, which is another classifier based on discovering rules, J48 presented the best performance for unseen data (test set). JRIP produced only five simple rules for classifying the flow patterns. The method based on Simple CART had only 90.06% performance.

Table 3. Evaluation measures for a data set containing 292 instances

DM Model / Evaluation measures		J48	Simple CART	LMT	JRIP	REPTree
		Total number of instances=292				
Correctly Classified Instances		92,46% (270 instances)	90,06% (263 instances)	93,83% (274 instances)	88,69% (259 instances)	91,78% (268 instances)
Incorrect Classified Instances		7,53% (22 instances)	9,93% (29 instances)	6,16% (18 instances)	11,30% (33 instances)	8,21% (24 instances)
Kappa Statistic		0,80	0,73	0,84	0,70	0,78
Mean absolute error		0,09	0,12	0,08	0,14	0,13
Root mean squared error		0,25	0,28	0,23	0,31	0,27
Precision	*normal_state*	0,84	0,83	0,86	0,76	0,83
	abnormal_state	0,95	0,92	0,96	0,93	0,94
Recall	*normal_state*	0,86	0,76	0,90	0,80	0,85
	abnormal_state	0,94	0,94	0,94	0,91	0,94
F-measure	*normal_state*	0,85	0,79	0,88	0,78	0,84
	abnormal_state	0,94	0,93	0,95	0,92	0,94
ROC Area *(Weighted Avg.)*		0,95	0,93	0,94	0,87	0,91

The experimental evaluation has been extended on multiple data sets and a statistical analysis has been performed on the obtained results.

The evaluation measures correspond to the data mining models J48, SimpleCART, LMT, JRIP and REPTree applied for 424 instances, respectively 968 instances and are presented in table 4 and table 5.

Table 4. Evaluation measures for a data set containing 424 instances

DM Model / Evaluation measures		J48	Simple CART	LMT	JRIP	REPTree
		Total number of instances=424				
Correctly Classified Instances		93,16%	90,09%	94,57%	87,73%	89,85%
Incorrect Classified Instances		6,83%	9,90%	5,42%	12,26%	10,14%
Kappa Statistic		0,84	0,76	0,87	0,71	0,76
Mean absolute error		0,09	0,14	0,06	0,17	0,15
Root mean squared error		0,24	0,29	0,20	0,33	0,29
Precision	*normal_state*	0,91	0,87	0,92	0,82	0,84
	abnormal_state	0,93	0,91	0,95	0,89	0,92
Recall	*normal_state*	0,86	0,80	0,90	0,77	0,83
	abnormal_state	0,96	0,94	0,96	0,92	0,92
F-measure	*normal_state*	0,89	0,83	0,91	0,80	0,84
	abnormal_state	0,95	0,92	0,96	0,91	0,92
ROC Area *(Weighted Avg.)*		0,95	0,92	0,96	0,84	0,92

Table 5. Evaluation measures for a data set containing 968 instances

DM Model / Evaluation measures	J48	Simple CART	LMT	JRIP	REPTree
	Total number of instances=968				
Correctly Classified Instances	96,79%	95,86%	98,86%	95,35%	94,00%
Incorrect Classified Instances	3,20%	4,13%	1,13%	4,64%	5,99%
Kappa Statistic	0,92	0,90	0,97	0,88	0,85

Mean absolute error		0,03	0,05	0,01	0,06	0,09
Root mean squared error		0,16	0,19	0,10	0,20	0,22
Precision	*normal_state*	0,94	0,95	0,97	0,93	0,90
	abnormal_state	0,97	0,96	0,99	0,96	0,95
Recall	*normal_state*	0,95	0,90	0,99	0,91	0,90
	abnormal_state	0,97	0,98	0,98	0,97	0,95
F-measure	*normal_state*	0,94	0,92	0,98	0,92	0,90
	abnormal_state	0,97	0,97	0,99	0,96	0,95
ROC Area *(Weighted Avg.)*		0,99	0,97	0,99	0,96	0,96

Corresponding to figure 2, the performance of the data mining models increases with the number of instances from the data set.

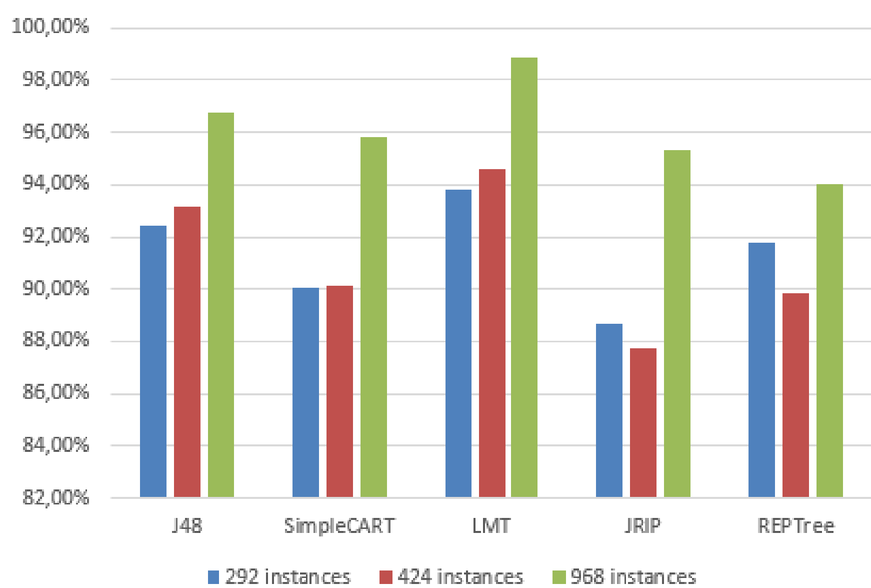

Figure 2. Correct Classified Instances for different data sets

The data mining models may be considered a superior technique that may be successful applied in diagnosis and may be develop in the future on the base of more training data to increase the accuracy of results.

Industrial processes are dynamic processes with random behavior and whose evolution over time cannot be predicted unless it is well known the process model and use advanced predictive techniques. Consequently, design and implement an automated online monitoring and diagnosis three-phase separator remains a future direction of research conducted so far.

Conceptually, this system should have permanent access to data collected from field transducers, to be able to identify the type of fault occurred, to locate the fault and provide recommendations to remedy abnormal operating condition. Also, updating the database defects with new types of defects occurred and the adequate solutions adopted for eliminating errors in the operating mode is an important feature to be considered during the design of the online diagnosis system. This is possible if the system would have self-learning capabilities. To acquire this "skill", the automatic online diagnosis system may contain a diagnosis module based on artificial neural networks.

The hybrid intelligent diagnosis system for a three-phase separator proposed in this paper was developed based on expert system technology, intelligent agents and data mining techniques. The end user (the operator) visualizes the diagnosis report on the GUI of diagnosis system

(DiagAgentExpertDM) and may act according to the recommendation given in the report. A disadvantage of the diagnosis method based on expert system is that induction rules must be performed manually and cannot be adjusted for the new samples of data. To reduce the effect of this disadvantage, it was proposed integration of data mining module able to automatically generate induction rules for the expert system knowledge base. An important task of data mining component is to provide the most efficient classification model in order to quickly identify the alarms and to give the adequate solution.

5. Conclusions

Building monitoring and diagnosis systems in gas and oil industry is a difficult task due the complexity and dynamic of industrial processes. Also, fault detection and identification, alarms generation, recommendation editing in a diagnosis report in real-time are key issues for an intelligent diagnosis system based on Artificial Intelligent technology. This research paper suggests interesting possibilities for hybrid techniques that combine different approaches (intelligent agent, data mining and expert system) for diagnosis of a three-phase separator from a gas-oil separation plant from Romania.

Future work will focus to increase the accuracy of diagnosis process by system testing in various functioning conditions, and to convert the offline diagnosis system to online diagnosis system, adding a neural network module, with learning capabilities.

References

N. Aissani, N. & Beldjilali, N. (2009). *Dynamic Scheduling in Petroleum Process using Reinforcement Learning*, In CIIA.

Serapião, A.B.S. & Bannwart, A.C (2013). *Knowledge Discovery for Classification of Three-Phase Vertical Flow Patterns of Heavy Oil from Pressure Drop and Flow Rate Data,* Journal of Petroleum Engineering, vol. 2013, Article ID 746315, 8 pages, doi:10.1155/2013/746315.

Atalla, F., & Taylor, J. (2004). *Modeling and Control of Three-Phase Gravity Separators in Oil Production Facilities*, Department of Electrical & Computer Engineering, University of New Brunswick, Canada, pp. 1-6.

Aulia, A., Keat, T.B., Maulut, M.S.B., El-Khatib, N., & Jasamai, M. (2010), *Smart oilfield data mining for reservoir analysis*, International Journal of Engineering & Technology, 10(06), pp.78-88.

Badr, N.L. (2014), *Effectiveness Logistic Regression Approach for Petroleum Production Ratio*, International Journal of Advanced Research in Computer Science, *5*(8).

Balakrishnan, K. & Honavar, V. (1998). *Intelligent diagnosis systems*, Journal of Intelligent Systems, 8(3-4), pp.239-290.

Canelón, M. A. R., Dávila, J., Morles, E.C., & La Hechicera, C. (2009). *Intelligent supervision of petroleum processes based on multi-agent systems*, WSEAS Transactions On Systems and Control, 4(9), pp. 435-444.

Ccoicca, Y.J. (2013). *Applications of support vector machines in the exploratory phase of petroleum and natural gas: a survey*, International Journal of Engineering & Technology, 2(2), pp.113-125.

Cerrada, M., Cardillo, J., Aguilar, J., & Faneite, R. (2007). *Agents-based design for fault management systems in industrial processes*, Computers in Industry , 58, pp. 313–328.

Cohen, W. (1995), *Fast effective rule induction*" in Proceedings of the 12th International Conference on Machine Learning, Lake Tahoe, Calif, USA, pp. 115–123.

Devold, H. (2013). *Oil and gas production handbook. An introduction to oil and gas production, transport, refining and petrochemical industry*, ABB Oil and gas, Oslo, ISBN 978-82-997886-3-2.

Engmo, L. & Hallen, L. (2007). *Software agents applied in oil production*, https://daim.idi.ntnu.no/masteroppgaver/003/3368/masteroppgave.pdf.

García, A., Leme, L., Pinto, F. & Sanchez-Pi, N. (2012). *MAS for Alarm Management System in Emergencies*, Advances in Artificial Intelligence–IBERAMIA 2012, pp.712-721.

Ghallab, S.A., Badr, N., Salem, A.B. & Tolba, M.F. (2013), *A Fuzzy Expert System For Petroleum Prediction*, WSEAS, Croatia, 2, pp. 77-82.

Ioniță, I. & Ioniță, L. (2013). *A Knowledge-Based System for Diagnosis in Oil Industry,* The 19th International Conference The Knowledge-Based Organization KBO 2013, Conference Proceedings 3 Applied Technical Sciences and Advanced Military Technologies, pp. 277-280.

Ioniță, L. & Ioniță, I. (2014). *A Methodology for Agent-Oriented Systems Development Applied in Oil Industry*, Studies in Informatics and Control, Volume 23, Issue 3, ISSN 1220-1766, pp. 239-248.

Ioniță, L. (2014). *Multi-agent System Used for Technical Diagnosis in Gas-Oil Separation Plant,* The 20th International Conference The Knowledge-Based Organization, Applied Technical Sciences and Advanced Military Technologies, Conference proceedings 3, „Nicolae Bălcescu" Land Forces Academy Publishing House, Sibiu, pp. 42-47.

Ioniță, L. (2014). *Sistem multi-agent pentru diagnoză tehnică aplicat în industria de petrol*, teză de doctorat, Universitatea Politehnica din București, Școala doctorală Automatică și Calculatoare, 2014.

Ioniță, L. & Ioniță I. (2015). *Expert-GOSP – Expert System for Three-Phase Separator Diagnosis*, Studies in Informatics and Control, ISSN 1220-1766, Vol. 24(3), pp. 293-300

Kopliku, A. & Condanni, D.L. (1994). *Expert system for corrosion evaluation and material selection In oil and gas wells*, In European Petroleum Computer Conference. Society of Petroleum Engineers.

Kyte, A. (2002). *Measure and Improve Knowledge Worker Productivity,* Garnter Group, gartner.com/1_researchanalysis/research_overview.html.

Li, X., Li, H. & Wu, Z. (2009). *Model-driven data mining in the oil & gas exploration and production*, In Knowledge Acquisition and Modeling, KAM'09. Second International Symposium on (Vol. 3, pp. 20-24). IEEE.

Lokko, N.N.B.C. (2012). *The role and use of expert systems for offshore assets on the Norwegian continental shelf: status quo and value creating measures*.

Noureldin, H. & Ruveta, F. (2002). *Using Expert System and Object Technology for Abnormal Condition Management*, In BIAS 2002 International Conference Milano. Milan, Italia.

Piatetsky-Shapiro, G., Brachman, R.J., Khabaza, T., Kloesgen, W., & Simoudis, E. (1996). *An Overview of Issues in Developing Industrial Data Mining and Knowledge Discovery Applications*, In KDD, 96, pp. 89-95.

Rabiei, M., Gupta, R., Cheong, Y.P. & Sanchez Soto, G.A. (2010). *A Novel Approach in Extracting Predictive Information From Water-Oil Ratio For Enhanced Water Production Mechanism, Diagnosis*, APPEA Journal, pp.567-579.

Rabiei, M., Gupta, R., Cheong, Y. P., & Soto, G.A.S. (2009). *Excess water production diagnosis in oil fields using ensemble classifiers*, In Computational Intelligence and Software Engineering, CiSE 2009. International Conference on (pp. 1-4), IEEE.

Timofeev, R. (2004). *Classification and regression trees (CART) theory and applications*.

Wang, T., Gao, X.W., & Li, K. (2012). *Application of data mining to production operation and control system in oil field*, In Control and Decision Conference (CCDC), 24th Chinese IEEE, pp. 642-647.

Witten, I.H., Frank, M., & Hall, M.A. (2011). *Data Mining: Practical Machine Learning Tool and Technique with Java Implementation,* Morgan Kaufmann, San Francisco, USA, 3rd edition.

Wooldridge, M. (1993). *An Introduction to MultiAgent Systems*, John Wiley & Sons, Ltd.

Quinlan, R.R. (1993). *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Francisco, USA.

Weka Documentation. (2016) Retrieved from http://www.cs.waikato.ac.nz/ml/weka/documentation.html, accessed on the 5th Jan. 2016.