



**Information Systems Department
Faculty of Computer & Information Sciences
Ain Shams University**

Intelligent System For Crude Oil Prediction

Thesis submitted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy in Faculty of Computer and Information Sciences, Information System dept.
Ain shams University

By

Senan Abdullah Ali Ghallab

Under Supervision of

Prof. Dr. Mohamed Fahmi Tolba

Professor of Scientific Computing
Department of Scientific Computing
Faculty of Computer and Information Sciences
Ain Shams University

Prof. Dr. Abdel-Badeeh M. Salem

Professor of Computer Science
Faculty of Computer and Information Sciences
Department of Computer Science
Ain Shams University

Dr. Nagwa Lotfi Badr

Assistant Professor of Information Systems
Faculty of Computer and Information Sciences
Department of Information Systems
Ain Shams University

Cairo-December 2014

قال تعالى:

"وَلَا يُنَبِّئُكَ هِئَلُ خَيْرٍ"

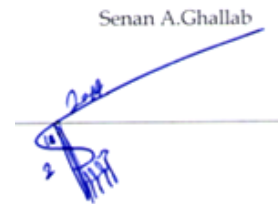
صدق الله العظيم

I certify that this work has not been accepted in substance for any academic degree and is not being concurrently submitted in candidature for any other degree.

Any portions of this thesis for which I am indebted to other sources are mentioned and explicit references are given.

Senan Abdullah Ali Ghallab

Senan A.Ghallab



ACKNOWLEDGMENT

To: My Mother, My Father soul

To: My Wife, My Coming Son

To: Sands of YEMEN

To: All Friends supporters

I would like to express my deepest gratitude to Prof. Dr. Mohamed Fahmi Tolba , Prof. Dr. Abdel Badeeh Salem and Dr. Nagwa Lotfi Badr. It was only through their meticulous supervision, support, and constructive advice, that this work has been completed.

I am greatly indebted to them for valuable input and continue support and direction.

This list of persons could be extended much further, therefore I would like to thank all those not mentioned above, and who have supported me.

ABSTRACT

Since the earliest ages, petroleum belongs to the minerals that have been used by humanity, earlier than metals and coal, and for numerous different purposes. Petroleum is a mixture of naturally occurring hydrocarbons that may exist in the solid, rocks, liquid or gaseous states, depending upon the conditions of pressure and temperature to which it is subjected.

Petroleum Engineers combine technology, chemistry, mathematics, crude oil properties, physics and geology with engineering methods to enhanced petroleum industry. They are concerned with finding deposits of oil and gas in quantities suitable for commercial use and with the economic extraction of these materials from the ground.

In this thesis, an intelligent prediction system based on computational intelligence technique (fuzzy) has been designed as Intelligent Petroleum Prediction System (IPPS). Multiple processes are used to achieve an accurate result. The required data for evaluation has been chosen from distinct and specialized sources of oilfields.

The proposed system utilizes a huge amount of petroleum dataset using fuzzy technique to apply prediction function on Daqing oilfield and other oilfields in Yemen. Another goal of the proposed system is to share petroleum knowledge among Chief Executive Manager and engineers to make drilling decision on oil wells or nullity. More than one module applied on the prediction system; such as petroleum dataset acquisition, classification, data mining, prediction, knowledge validation and other processes.

Recently, using artificial intelligence techniques for prediction is highly usage in prediction domain, precision is the criterion. The proposed work reviewing engineering's experiences, analysis oil wells dataset, reordering memberships, forming an intelligent system to utilize crude oil knowledge and build an accurate predicted result.

The proposed system results evaluation applied through compared system predicted result with measured or empirical values of different oilfields and other intelligent technique results. IPPS achieved approximated results; verified and give the oil engineers a range of prediction process before the drilling operation. More than one module approved in this thesis, from dataset collection within distinct sources until validation prediction result test by multiple evaluation functions, the proposed system modules are:

- Dataset acquisition, analysis, classification, clustering and mining using Weka utility.
- Interesting of analysis, classifying, data mining, prediction, test cases, and regression through applying time series statistics of vagueness petroleum data.
- Integrate crude oil ontology into chemical knowledge. The development of crude oil ontology architecture approved by PROLOG and Protégé-OWL utilities, respectively.
- Knowledge validation by expert's and scientific methods.
- Test a novel application for Oil and Gas ratio prediction using ANFIS technique.
- Building an intelligent prediction system based on fuzzy system called Intelligent Petroleum Prediction System (IPPS).
- Retest both historical and test cases statues of predicted results through validation processes. The computational intelligence technique (Fuzzy), test case and clustering functions are used to achieve overlap Strictness.
- Confirm predicted results through logistic regression model which terminates the controversial and system efficiency.
- Using Time series forecasting Model to retest predicted results for more efficiency and minimize error range.

Published papers

- 1- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem, M. F. Tolba “***Modeling an Application for Oil and Gas Ratio Prediction Using ANFIS***”. In the Egyptian Computer Science Journal, **Egypt**, Vol. 38. pp. 1-6, Sep-2012.
- 2- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem , M. F. Tolba “***Integration Web-Based Crude Oil Ontology***”. In the proceeding of International Conference on Computing and Informatics (ICCI), **Egypt**, Vol. 1, pp. 91-96, 2012.
- 3- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem , M. F. Tolba “***A Fuzzy Expert System For Petroleum Prediction***”. In the proceeding of World scientific and engineering Academy Society (WSEAS), **Croatia**, Vol. 3 , pp. 70-75, 2013.
- 4- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem , M. F. Tolba “***Computational Intelligence Approaches For Manipulating Vagueness Petroleum Data***”, In the International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), **India**, Vol. 3, pp. 284-289, 2014.
- 5- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem, M. F. Tolba “***Effectiveness Ontology-Based for Crude Oil Deduction Model***”. In the International Journal of Applications of Fuzzy Sets and Artificial Intelligence, **Greece**, Vol. 4, pp.127-139, 2014.
- 6- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem , M. F. Tolba “ ***Intelligence Test Case Based-Approach for Crude Oil Prediction System***“. In the proceeding of Intelligent Systems(IS’14), **Poland**, IEEE Press, Vol. 323, pp. 893-903 June- 2015.
- 7- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem , M. F. Tolba “***An Intelligent Prediction System for Crude Oil Presence Using Fuzzy Model*** ”. In the proceeding of Advanced Machine Learning Technologies and Applications (AMLTA14).Egypt, Vol. 488, pp.289-295 , **Egypt**, Springer, Nov-2014.
- 8- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem , M. F. Tolba “***Logistic Test Case Regression Model of Petroleum Availability Prediction System***”. In the proceeding of Advanced Machine Learning Technologies and Applications (AMLTA14).Vol. 488, pp. 248-257, **Egypt**, Springer, Sep-2014.
- 9- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem, M. F. Tolba, “***Strictness Petroleum Prediction System Based on Fuzzy Model***”. Paper submitted in the International Journal of Service Science, Management, Engineering, and Technology (IJSSMET), **USA**, 2014.

- 10- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem, M. F. Tolba, “**Intelligent Systems Utilization For Petroleum Prediction**”, Accepted as book chapter, Paper accepted in the proceeding of Advanced Machine Learning Technologies and Applications (AMLT14), **Egypt**, Nov-2014.
- 11- Senan A.Ghallab, N.L. Badr, Abdel-Badeeh Salem, M. F. Tolba, “**Fuzziness Petroleum Data Consolidation Using A Time Series Forecasting Model**”. Paper submitted in the 7th International Conference on Information Technology (ICIT), **Jordan**, June-2015.

Table of Contents

List of figures.....	...X
List of tables.....	XIII
Abbreviations.....	(XV-XVI)
Chapter 1: Introduction.....	
1.1. Introduction.....	1
1.2. Problem definition.....	5
1.3. Research Importance.....	7
1.4. Thesis Goals.....	8
1.5. Thesis contribution	9
1.6. Thesis Organization.....	9
Chapter 2: Literature Review.....	
2.1 Introduction.....	11
2.2 Petroleum domain.....	13
2.3 Prediction	25
2.3.1 Oil Prediction.....	26
2.4 Artificial intelligence.....	28
2.4.1 Intelligent Systems.....	32
2.4.2 Expert Systems.....	37
2.5 Geographic Information System.....	46
2.5.1 GIS components.....	47
2.5.2 GIS TASKS.....	48
2.6 Knowledge Base.....	49
2.6.1 Ontology.....	51
2.6.1.1 Types of Ontology.....	52
2.6.1.2 Usage of ontology	52
2.7 Knowledge Reasoning	55
2.8 Data mining.....	57

2.8.1 Data intelligence.....	57
2.9 Summary.....	63
Chapter 3: Related Work	
3.1 Introduction.....	64
3.2 Black Oil.....	64
3.3 Oil Predication Tools.....	66
3.4 Proposed System vs. related Work.....	90
3.5 Summary.....	100
Chapter 4: The Proposed System.....	
4.1 Introduction.....	102
4.2 Proposed System Overview	103
4.3 The Proposed System Architecture	109
4.3.1 Dataset Acquisition.....	112
4.3.2 Petroleum Knowledge Classification	113
4.3.3 Knowledge Validation.....	121
4.4 Summary.....	131
Chapter 5: Intelligent Petroleum Prediction System (IPPS).....	
5.1 Introduction.....	132
5.2 Prediction Module.....	133
5.2.1 ANFIS Model.....	134
5.2.2 Fuzzy System.....	139
5.2.3 Evaluation Module.....	152
5.3 Summary.....	168

Chapter 6: Case Studies & Evaluation.....	
6.1 Introduction	169
6.2 Prediction Experimental	170
6.2.1 Case study 1.....	173
6.2.1.1 Petroleun Dataset.....	173
6.2.1.2 IPPS Graphical User Interface.....	177
6.2.2 Case study 2.....	184
6.3 IPPS Evaluation.....	186
6.3.1 Experimental Evaluation.....	187
6.3.1.1 Test Cases Evaluation.....	187
6.3.1.2 Regression Model Validation.....	188
6.3.1.3 Time series.....	189
6.4 Summary.....	191
Chapter 7: Conclusion & Future Work.....	
7.1 Introduction.....	192
7.2 Future Work.....	197
Petroleum Glossary	(XVII)
References.....	(XIX-XXIV)
Arabic Abstract.....	

List of Figures

Figure 2.1 Sequences of Oil Deposits	17
Figure 2.2 Anticline Trap	17
Figure 2.3 Fault Trap	18
Figure 2.4 Combination Trap	18
Figure 2.5 Petroleum Products	19
Figure 2.6 Oil areas distribution	20
Figure 2.7 Seismic Waves	24
Figure 2.8 Artificial Intelligence and Computational Intelligence	31
Figure 2.9 Flowchart for genetic programming	35
Figure 2.10 Proposed system Methodology	36
Figure 2.11 Components of AI	37
Figure 2.12 Components of an expert system	38
Figure 2.13 The architecture of expert system	39
Figure 2.14 The main players of expert system development team	40
Figure 2.15 Knowledge validation process by expert	46
Figure 2.16 GIS components	47
Figure 2.17 GIS Utilization	48
Figure 2.18 knowledge base methodologies	49
Figure 2.19 Knowledge engineering process	50
Figure 2.20 Ontology types	52
Figure 2.21 Knowledge sharing process	53
Figure 2.22 Knowledge engineering and Expert systems	56
Figure 2.23 Knowledge base ontology.....	56
Figure 2.24 AQ algorithm.....	59
Figure 2.25 A simplified skeleton of the apriori algorithm and a simple example	61
Figure 2.26 Bagging algorithm.....	62
Figure 3.1 Oilfields areas around the world	66
Figure 3.2 Related Petroleum Prediction applications	67
Figure 3.3 Structure of Functional and its Simplifications	69
Figure 3.4 Structure of a Type-2 FLS	69
Figure 3.5 Design Framework of the Hybrid Model	71
Figure 3.6 Correlation Coefficient for Porosity and permeability training.	72
Figure 3.7 RBF Fuzzy-neural networks	73
Figure 3.8 Conceptual Framework of the FN-SVM Hybrid Model	76
Figure 3.9 Conceptual Framework of the FN-T2FL Hybrid Model	76
Figure 3.10 Optimal Radii for ANFIS-SC with site 1 Well 1 Dataset	78
Figure 3.11 Optimal Radii for ANFIS-SC with site 1 Well 3 Dataset	78
Figure 3.12 Optimal Radii for ANFIS-SC with site 2 Well 2 Dataset	78
Figure 3.13 Testing CC comparison for porosity Datasets	79
Figure 3.14 Testing RMSE comparison for porosity Datasets	79
Figure 3.15 Training and Testing run time comparison for porosity Datasets.....	97
Figure 3.16 Training CC comparison for Permeability Datasets	80
Figure 3.17 Training RMSE comparison for Permeability Datasets	80
Figure 3.18 Training and Testing run time comparison for Permeability Datasets.....	80
Figure 3.19 Multi layer perceptron	83
Figure 3.20 Experimental values compared with calculated values by each correction.....	83
Figure 3.21 Experimental values compared with calculated values by ANN model.....	84
Figure 3.22 Logistic regression curve	87

Figure 3.23 Simplified geological map of Yemen	89
Figure 3.34 Oil prediction and algorithms framework	96
Figure 3.25 Comparative study of previous intelligent applications and IPPS	99
Figure 3.26 Comparative study of previous intelligent applications and IPPS	99
Figure 4.1 dataset sources	103
Figure 4.2 The proposed System Methodology	105
Figure 4.3 The proposed system basic functions	106
Figure 4.4 System Cores	107
Figure 4.5 System entities relations	108
Figure 4.6 System Processes	109
Figure 4.7 The Proposed System architecture	111
Figure 4.8 System resources dataset.....	113
Figure 4.9 Dataset Acquisition.....	113
Figure 4.10 Dataset Acquisition.....	114
Figure 4.11 Weka processes.....	115
Figure 4.12 Dataset Acquisition.....	116
Figure 4.13 Knowledge manipulation.....	117
Figure 4.14 Knowledge manipulation.....	118
Figure 4.15 Petroleum Standard parameters.....	119
Figure 4.16 Petroleum Standard parameters.....	120
Figure 4.17 Crude Oil Ontology Structure.....	121
Figure 4.18 Crude Oil Ontology Structure.....	123
Figure 4.19 Cubic rule bases.....	125
Figure 4.20 Cubic rule bases.....	126
Figure 4.21 System rules.....	127
Figure 4.22 Data mining structure.....	128
Figure 5.1 ANFIS training values	141
Figure 5.2 Temperature membership	141
Figure 5.3 Pressure membership	142
Figure 5.4 Crude oil Density membership	142
Figure 5.5 Gravity membership	143
Figure 5.6 Gas-Density membership	143
Figure 5.7 Output membership	144
Figure 5.8 Rules editors.....	145
Figure 5.9 Prediction Result	148
Figure 5. 10 Defuzzification representationresults.....	148
Figure 5.11 Prediction cubic Results	149
Figure 5.12 Differnt IPPS Functions	149
Figure 5.13 IPPS GUI	150
Figure 5.14 Evaluation module Architecture	153
Figure 5.15 Test Case process architecture	156
Figure 5.16 Test Case Prediction Methodology	158
Figure 5.17 Regression model Methodology	162
Figure 5.18 Regression model methodology	163
Figure 5.19 Slope and intercept curve	164
Figure 6.1 Oil Blocks and Yemen map	171
Figure 6.2 Previous blockfields studied in Yemen	172
Figure 6.3 Daqing oilfields in China	172
Figure 6.4 Petroleum dataset relations	174
Figure 6.5 Development IPPS GUI	178
Figure 6.6 Enhanced IPPS GUI	179

Figure 6.7 IPPS within High percentage	180
Figure 6.8 IPPS within pressure membership	181
Figure 6.9 PP and CP results of ten wells in Daqing oilfields.....	182
Figure 6.10 PP, CP and Tc results of ten wells in Daqing oilfields	182
Figure 6.11 Results Comparison with previous measured values of Daqing	183
Figure 6.12 Prediction results within empirical values	185
Figure 6.13 Techniques prediction results	189
Figure 6.14 Petroleum dataset relations	189
Figure 6.15 Error range of prediction	190

List of Tables

Table (2.1) Elemental Analysis of Typical Crude	13
Table (2.2) Typical crude oil fractions	13
Table (2.3) The constitutions and the characteristics hard and soft computing	32
Table (2.4) Knowledge Acquisition procedures for developing expert system	44
Table (2.5) Association Rule Transactions	60
Table (3.1) Predicator variables for site 1 core data for porosity	70
Table (3.2) Predicator variables for site 2 WELL LOGS for Permeability.....	70
Table (3.3) Design of Dataset into training and testing subsets.....	71
Table (3.4) Data of crude oil properties of thirty well in Daqing oilfield	74
Table (3.5) Measured Value and Prediction Value of Gas-Oil Volume ratio	75
Table (3.6) Optimal number of clusters for dataset	76
Table (3.7) Statistical Experimental data of sample oils	82
Table (3.8) Summary of under saturated oil viscosity correlations	82
Table (3.9) Test Cases Example	86
Table (3.10) Fuzzy Rule Base	86
Table (3.11) TI test parameters with their values	87
Table (3.12) Prediction analysis outcome	88
Table (3.13) Related petroleum researches (2000-2014).....	91
Table (3.14) Comparison study (IPPS and other intelligent tools.....	98
Table (4.1) Standard petroleum parameters.....	119
Table (4.2) Crude oil classes	120
Table (4.3) Cubic model template	124
Table (5.1) Prediction result compared with related previous researches values.....	137
Table (5.2) Fuzzy rule base	145
Table (5.3) Wells predicted results	147
Table (5.4) CP & PP results	159
Table (5.5) Sample of test cases regression dataset model	162
Table (5.6) CP, PP and Tc results	164
Table (5.7) Logistic regression sample	165
Table (5.8) Oilfields data with three time levels	166
Table (5.9) SPSS Error time.....	166
Table (5.10) Comparison explanation within computational intelligent.....	167
Table (6.1) Data of thirty well on Daqing oilfield	173

Table (6.2) Measured value and prediction result	184
Table (6.3) Binon oilfields	185
Table (6.4) Twilah oilfield	174
Table (6.5) Twilah oilfield dataset	175
Table (6.6) Sample of Yemen oilfields	175
Table (6.7) Results comparison sample in Yemen	177
Table (6.8) Test Cases and PA	183
Table (6.9) Model Validation results	187
Table (6.10) Comparison explanation within computational intelligent	188