

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE SCIENTIFIQUE
ÉCOLE NATIONALE POLYTECHNIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique

Département Génie Minier

End-of-study project dissertation

For obtaining the State Engineer's degree in Mining Engineering

Evaluation and Development of a Predictive Model for the Analysis
of Geophysical Well Log Data in Reservoir Characterization

HAOUES Doaa & KHALED Ihssane

Under the direction of **Pr. AKKAL Rezki** ENP

Presented and defended publicly on (13/09/2025)

Composition of the jury :

President: Mr. CHANANE Larouci ENP
Promoter: Pr. AKKAL Rezki ENP
Examiner : Pr. YAHYAOUI Sami ENP
Incubator : Mr. BOUSBAI M'hamed ENP

ENP 2025

الملخص

تستجيب هذه الدراسة للطلب المتزايد على التفسير البتروفيزيائي الفعّال والدقيق في صناعة النفط والغاز من خلال دمج تقنيات الذكاء الاصطناعي في سير عمل توصيف المكامن. فعلى الرغم من أهمية الطرق التقليدية الحتمية والتجريبية، إلا أنها غالبًا ما تكون مستهلكة للوقت و محدودة في قدرتها على التعامل مع التبايرية الجيولوجية.

وللتغلب على هذه التحديات، قمنا بتطوير نظام ذكي مصمم للتنبؤ بالخصائص الأساسية للمكامن وتصويرها مباشرةً انطلاقًا من بيانات السجلات البئرية. شمل خط العمل المنهجي معالجة البيانات الأولية، هندسة الخصائص، تدريب النماذج تحت إشراف، وتحليل قابلية التفسير.

وقد أظهرت خوارزميات التجميع أداءً تنبؤيًا قويًا: حيث حقق نموذج Gradient Boosting للتنبؤ بنسبة حجم الطفل (Vsh) قيمة $R^2 = 0.914$ و $MAE = 0.0642$ ، بينما أعطى نموذج XGBoost للنفاذية الفعّالة (PHIE) قيمة $R^2 = 0.801$ و $MAE = 0.0248$ ، وبلغ نموذج Extra Trees لتشبع الماء (Sw) قيمة $R^2 = 0.895$ و $MAE = 0.0127$. تؤكد هذه النتائج قدرة نماذج الذكاء الاصطناعي على التقاط العلاقات غير الخطية الكامنة في البيانات البتروفيزيائية والتفوق على المقاربات التقليدية. وقد جرى تنفيذ النظام في شكل تطبيق مكتبي سهل الاستخدام، مما يتيح للجيولوجيين والمهندسين الحصول بسرعة على رؤى موثوقة لتقييم المكامن.

بصورة عامة، يبرهن هذا العمل على جدوى وتأثير الطول المعتمدة على الذكاء الاصطناعي في تطوير تفسير المكامن، ودعم اتخاذ القرارات المبينة على أسس سليمة، وفتح آفاق جديدة نحو النشر المستقبلي في الزمن الحقيقي وعلى نطاق واسع.

الكلمات المفتاحية

التعلم الآلي، توصيف المكامن، بيانات التسجيل الجيولوجي، المعاملات البتروفيزيائية.

Résumé

Cette étude répond à la demande croissante d'une interprétation pétrophysique efficace et précise dans l'industrie pétrolière et gazière grâce à l'intégration de l'intelligence artificielle dans les workflows de caractérisation des réservoirs. Les méthodes déterministes et empiriques traditionnelles, bien que précieuses, s'avèrent souvent chronophages et limitées dans leur capacité à gérer l'hétérogénéité géologique. Pour surmonter ces défis, nous avons développé un système intelligent conçu pour prédire et visualiser les principales propriétés des réservoirs directement à partir des diagraphies de puits. Le pipeline méthodologique comprenait la préparation des données, l'ingénierie des variables, l'entraînement supervisé des modèles et l'analyse de l'interprétabilité. Les algorithmes d'ensemble ont démontré de solides performances prédictives : Gradient Boosting pour le volume d'argiles (Vsh) a atteint $R^2 = 0,914$, $MAE = 0,0642$; XGBoost pour la porosité (PHIE) a donné $R^2 = 0,801$, $MAE = 0,0248$; et Extra Trees pour la saturation en eau (Sw) a atteint $R^2 = 0,895$, $MAE = 0,0127$. Ces résultats confirment la capacité des modèles d'IA à capturer les relations non linéaires inhérentes aux données pétrophysiques et à surpasser les approches conventionnelles. Le système a été implémenté sous forme d'une application de bureau conviviale, permettant aux géoscientifiques et aux ingénieurs d'obtenir rapidement des informations fiables pour l'évaluation des réservoirs. Dans l'ensemble, ce travail démontre la faisabilité et l'impact des solutions basées sur l'IA dans l'avancement de l'interprétation des réservoirs, en soutenant la prise de décision éclairée et en

ouvrant des perspectives pour un déploiement futur en temps réel et à grande échelle.

Mots clés : Apprentissage automatique, Caractérisation des réservoirs, Données de diagraphies, Paramètres pétrophysiques.

Abstract

This study addresses the growing demand for efficient and accurate petrophysical interpretation in the oil and gas industry through the integration of artificial intelligence into reservoir characterization workflows. Traditional deterministic and empirical methods, while valuable, often prove time-consuming and limited in their ability to handle geological heterogeneity. To overcome these challenges, we developed an intelligent system designed to predict and visualize key reservoir properties directly from well log data. The methodological pipeline encompassed data preprocessing, feature engineering, supervised model training, and interpretability analysis. Ensemble learning algorithms demonstrated strong predictive performance: Gradient Boosting for shale volume (Vsh) achieved $R^2 = 0.914$, $MAE = 0.0642$; XGBoost for porosity (PHIE) yielded $R^2 = 0.801$, $MAE = 0.0248$; and Extra Trees for water saturation (Sw) reached $R^2 = 0.895$, $MAE = 0.0127$. These results confirm the ability of AI models to capture non-linear relationships inherent in petrophysical data and to outperform conventional approaches. The system was implemented as a user-friendly desktop application, enabling geoscientists and engineers to obtain rapid, reliable insights for reservoir evaluation. Overall, this work demonstrates the feasibility and impact of AI-powered solutions in advancing reservoir interpretation, supporting informed decision-making, and opening perspectives for future real-time and scalable deployment.

Keywords : Machine Learning , Reservoir Characterization, Well Log Data, Petrophysical parameters.

Contents

List of Tables

List of Figures

Acronyms

Introduction	24
1 Well Logging Measurements and Petrophysical Parameters	25
1.1 Introduction to Well Logging	26
1.2 Definition of Well Logging	26
1.3 Technical Methods of Well Logging Acquisition	27
1.4 Presentation of a Well Log	27
1.5 Wireline Logging	28
1.5.1 Definition:	28
1.5.2 Types of Wireline Logging	29
1.5.2.1 Electrical Logging	29
1.5.2.1.1 Spontaneous Potential	29
1.5.2.1.2 Resistivity	30
1.5.2.2 Nuclear Logging	32
1.5.2.2.1 Gamma Ray Logging	32
1.5.2.2.2 Neutron Logging	33
1.5.2.2.3 Density Logging	35
1.5.2.3 Acoustic Logging (Sonic)	36
1.5.2.4 Auxiliary Logging (Caliper)	36
1.6 Petrophysical Parameters of the Reservoir	38

1.6.1	Shale Volume (V_{sh})	38
1.6.1.1	Definition	38
1.6.1.2	Calculation Method	39
1.6.2	Porosity	39
1.6.2.1	Definition	39
1.6.2.2	Types of Porosity	40
1.6.2.3	Calculation Method:	41
1.6.3	Permeability (K)	42
1.6.3.1	Definition	42
1.6.3.2	Darcy's Law	43
1.6.3.3	Reservoir Classification Based on Permeability	44
1.6.3.4	Types of permeability	44
1.6.3.5	Calculation Method	44
1.6.4	Saturation (S)	45
1.6.4.1	Definition	45
1.6.4.2	Saturation types	45
1.6.4.3	Calculation Method	46
1.6.5	Density	47
1.6.5.1	Definition	47
1.6.5.2	Methods for Determining Rock Density	47
1.7	Conclusion	48
2	Artificial Intelligence in Well Logging: Theoretical Foundations	49
2.1	Introduction to Artificial Intelligence:	50
2.1.1	AI Technologies in Oil and Gas	50
2.2	Overview of Machine Learning	51
2.2.1	Machine Learning as a Subfield of AI	51
2.2.2	Types of Machine Learning	51
2.3	Machine Learning in Well Logging	53
2.3.1	Role of Machine Learning in Petrophysical Analysis	53
2.3.2	The Intersection of Petrophysics and Machine Learning	53

2.4	Key Machine Learning Techniques for Well Logging	54
2.4.1	Tree-Based Models (DTs)	54
2.4.2	Linear Regression:	54
2.4.3	Gradient boosting	54
2.5	Petrophysical Data and Its Characteristics	55
2.5.1	Data Types in Petrophysics	55
2.5.2	Data Quality and Preprocessing	55
2.6	Model Evaluation and Validation	56
2.6.1	Performance Metrics	56
2.6.2	Cross-Validation Techniques:	57
2.6.3	Model Selection and Hyperparameter Tuning	57
2.7	Conclusion	59
3	Development Foundations of the AI-Driven Petrophysical Analysis Solution	60
3.1	General Context of the Project	61
3.1.1	Challenges of Petrophysical Interpretation in the Oil Industry	61
3.1.2	Limitations of Traditional Methods	61
3.1.3	The Potential Role of Artificial Intelligence	61
3.2	Problem Statement	62
3.2.1	Identified Problem in the Field	62
3.2.2	Specific Needs of End Users	62
3.3	Project Idea	63
3.3.1	General Overview of the Solution	63
3.3.2	Core Functionality	63
3.3.3	Product Vision and Startup Orientation	64
3.3.4	Project team	65
3.4	Value Proposition	65
3.4.1	Automation and Time Efficiency	65
3.4.2	Accessibility and Ease of Use	65
3.4.3	Objectivity and Consistency in Interpretation	65
3.4.4	Adaptability to Client Needs	66

3.4.5	Cost Efficiency and Economic Accessibility	66
3.4.6	Locally Developed, Strategic Innovation	66
3.5	Technical and Scientific Objectives	66
3.5.1	Technical Objectives	66
3.5.2	Scientific Objectives	67
3.6	Conclusion	68
4	Development and Interpretation of DIAA Results	69
4.1	Introduction	70
4.2	Development Stages of the DIAA Platform	70
4.2.1	Data Acquisition	70
4.2.1.1	Sources and Formats of Well Log Data	71
4.2.1.2	Depth Alignment and Unit Standardization	71
4.2.1.3	Dataset Overview and Summary	71
4.2.2	Data Cleaning and Preprocessing	72
4.2.2.1	Handling Missing Values and Outliers	72
4.2.2.2	Data Normalization and Standardization	72
4.2.2.3	Dataset Splitting	73
4.2.3	Feature Selection and Target Variables	73
4.2.3.1	Selection of Input Features	74
4.2.3.2	Target Parameters: V_{sh} , $PHIE$ and S_w	74
4.2.4	Machine Learning Models and Training Strategy	75
4.2.4.1	Overview of Selected Algorithms	75
4.2.4.2	Hyperparameter Tuning Techniques	76
4.2.4.3	Training Pipeline and Automation	77
4.2.5	Model Testing	77
4.2.5.1	Evaluation on the Test Set	77
4.2.5.2	Performance Metrics Used	78
4.2.6	Model Validation	78
4.2.6.1	Cross-Validation Methodology	78
4.2.6.2	Stability and Robustness Assessment	79

4.2.6.3	Error Analysis Across Wells/Zones	79
4.3	Interpretation of Obtained Results	80
4.3.1	Justifying Feature Inputs Using Heatmap Correlations	80
4.3.1.1	Prediction of Volume of Shale (V_{sh})	81
4.3.1.2	Prediction of Effective Porosity (PHIE)	81
4.3.1.3	Prediction of Water Saturation (S_w)	82
4.3.2	Comparative Evaluation of Model Performance per Target	83
4.3.2.1	V_{sh} Prediction – Model Performance Comparison	83
4.3.2.2	PHIE Prediction – Model Performance Comparison	88
4.3.2.3	(S_w) Prediction – Model Performance Comparison	93
4.3.3	Justification of Final Model Selection	98
4.3.3.1	Justification of the Final Predictive Model for V_{sh}	98
4.3.3.2	Justification of the Final Predictive Model for PHIE	100
4.3.3.3	Justification of the Final Predictive Model for S_w	101
4.3.4	Log-Based Visualization and Interpretation of Predictions	101
4.3.4.1	Visualization of Predicted vs. Actual V_{sh} Along Depth	101
4.3.4.2	Visualization of Predicted vs. Actual PHIE Along Depth	103
4.3.4.3	Visualization of Predicted vs. Actual S_w Along Depth	106
4.3.5	Synthesis of Prediction Results and Model Behavior	109
4.3.5.1	Variation in Model Performance Across Petrophysical Targets	109
4.3.5.2	Geological Influence on Prediction Accuracy	110
4.3.5.3	Interpretation of Model Limitations and Strengths	110
4.4	Predictive Report Generation and Results Structuring	110
4.5	Conclusion	111
5	Prototype Description and Functional Demonstrations	113
5.1	General Overview of the Prototype	114
5.1.1	Current Version (Minimum Viable Product “MVP”)	114
5.1.2	Type of Application (Desktop-Based, Offline)	115
5.1.3	Technologies Used	116
5.2	Functional Architecture and User Interface (UI)	117

5.2.1	General System Workflow	117
5.2.2	Interface Description	119
5.2.2.1	Home Tab	119
5.2.2.2	Data Management and Prediction Tab	120
5.2.2.3	Log-Style Visualization Tab	120
5.2.2.4	Report Tab	122
5.3	Implemented Functionalities (with Use Case Example)	123
5.3.1	File Import	123
5.3.2	Graphical Visualization of Results	125
5.3.2.1	Export of Results (Automated Report Generation)	127
5.4	Revenue Architecture and Deployment Strategy	128
5.4.1	Licensing Model: Perpetual License with Annual Maintenance Fees . . .	128
5.4.2	Deployment as a Standalone Desktop Application	129
5.4.3	Future SaaS Development Potential	129
5.5	Limitations of the Current Prototype	130
5.5.1	Missing Functionalities	130
5.5.2	Planned Improvements	130
5.5.3	Development Constraints	131
5.6	Conclusion	131
6	Commercialization and Growth Strategy	132
6.1	Market Segmentation	133
6.1.1	Target Clients	133
6.1.1.1	Oil and Gas Companies	133
6.1.2	Objectivity and Consistency in Interpretation	133
6.1.2.1	Oilfield Service Providers	134
6.1.2.2	Research Institutions and Universities	134
6.1.3	Economic Value Proposition	134
6.1.3.1	Cost Savings Through Automation	134
6.1.3.2	Lower Operational Expenses	135
6.1.3.3	Reduced Training and Adoption Costs	135

6.1.3.4	Improved Decision-Making & Risk Mitigation	135
6.1.3.5	Strategic Economic Benefits	135
6.2	Target Market	136
6.2.1	Market Selection Justification	136
6.2.2	Market Opportunity and Unmet Needs	136
6.3	Competitive Analysis	137
6.3.1	Direct and Indirect Competitors	137
6.3.1.1	Direct Competitors	137
6.3.1.2	Indirect Competitors	137
6.3.2	SWOT Analysis	138
6.3.2.1	Strengths	138
6.3.2.2	Weaknesses	138
6.3.2.3	Opportunities	138
6.3.2.4	Threats	138
6.4	Market Share Overview	139
6.4.0.1	Algeria and North Africa Context	139
6.5	Market Share and Software Access	139
6.6	Marketing and Sales Strategy	139
6.7	Strategic Roadmap	140
6.7.1	Implementation Timeline	140
6.7.1.1	Phase 1 : Prototype Development (0-6 months)	141
6.7.1.2	Phase 2 : Beta Testing and Pre-market Positioning (6-12 months)	141
6.7.1.3	Phase 3 : Market Launch and Early Adoption (12-24 months) .	141
6.7.1.4	Phase 4 : Scaling and Technological Expansion (24-36 months)	141
6.8	Conclusion	142
7	Financial Analysis and Strategic Perspectives	143
7.1	Development Costs (Software, Hardware, and Human Resources)	144
7.1.1	Human Resources	144
7.1.2	Operational Cost Analysis	144
7.1.2.1	Equipment and Infrastructure Investment	145

7.2	Revenue Models and Financial Projections	147
7.2.0.1	Pricing Strategy: Perpetual Licensing with Maintenance Support	147
7.2.0.2	Projected Sales Based on Market Segments	147
7.3	Financial Return Metrics and Investment Recovery	148
7.3.1	Annual Cost Breakdown	148
7.3.2	Revenue Projections and Sales Assumptions	148
7.3.3	Profit and Tax Analysis	149
7.3.4	Cumulative Net Profit and ROI	149
7.3.5	Payback Period	150
7.3.6	Interpretation and Financial Outlook	150
7.4	Risk Management and Mitigation Strategies	151
7.4.1	Technical Risks and Backup Plans	151
7.4.2	Market Risks and Competitive Pressure	151
7.4.3	Legal and Licensing Considerations	152
7.5	Business Model Canvas	152
7.5.1	Value Proposition	153
7.5.2	Channels	153
7.5.3	Customer Segments	153
7.5.4	Customer Relationships	153
7.5.5	Revenue Streams	154
7.5.6	Key Resources	154
7.5.7	Key Partners	154
7.5.8	Key Activities	154
7.5.9	Cost Structure	154
	General Conclusion	156
	Bibliography	159

Confidentiel