Lithofacies With Sparse Well Log Data Combined With Calculated Attributes In A Gradational Fluvial Sequence

Mr.Sakthivel[1],Mr.S.Bharathselvan[2],Mr.R. Dhanush[3],Mr.V.Indirajith[4]
[1]Assistant Professor, Dept of Computer Science and Engineering
[2, 3, 4]Dept of Computer Science and Engineering
[1, 2, 3, 4] Erode Sengunthar Engineering college, Perundurai, Erode, Tamil Nadu, India.

Abstract- *LightGBM serves as the core element for accurate lithofacies prediction in complex reservoirs. By incorporating derivative attributes into well-log data, the system extends the feature space, resulting in improved prediction accuracy. Furthermore, a state-of-the-art multiple-optimizer technique is introduced to effectively identify the most significant log attributes for lithofacies prediction, thus addressing the specific challenges associated with complex and laterally heterogeneous geological formations. The system not only excels in lithofacies prediction but also offers valuable insights through feature importance analysis, enhancing reservoir characterization and providing a deeper understanding of geological behaviours in these intricate environments.This innovative approach is poised to significantly benefit the field of reservoir characterization, particularly in scenarios where traditional methods may fall short due to geological complexity, ensuring more precise predictions and increased insights into reservoir behaviour.*

I. LITERATURE SURVEY

1.1 A SYSTEMATIC REVIEW OF EMERGING FEATURE SELECTION OPTIMIZATION METHODS FOR OPTIMAL TEXT CLASSIFICATION: THE PRESENT STATE AND PROSPECTIVE OPPORTTUNITIES

In the domain of predictive modeling for classification tasks, the effectiveness of data preparation techniques holds a pivotal role in optimizing data and achieving optimal performance. Fundamental processes such as outlier detection, imputation of missing values, and feature selection (FS) are critical for enhancing model efficiency. FS, a crucial technique, contributes to faster performance, elimination of noise, removal of redundancy, reduction of overfitting, and improved precision, resulting in enhanced generalization on testing data. While traditional FS techniques have been employed for classification tasks, they often face challenges in handling the high dimensionality of text feature spaces, leading to inefficient predictive models. Emerging technologies, including metaheuristics and hyper-heuristics optimization methods, offer a novel approach to FS. These methods demonstrate efficiency in improving classification accuracy, addressing computational demands, optimizing storage, and swiftly solving complex optimization problems.
Despite the advantages of these emerging FS methods, the existing literature lacks detailed best practices for their case-by-case application, creating a landscape with both clear and unclear findings. This paper conducts a systematic review of over 200 articles to present the current state of FS, with a specific focus on metaheuristics and hyper-heuristic methods. The objective is to provide the latest insights and trends, guiding analysts, practitioners, and researchers in the field of data analytics toward a clearer understanding and effective implementation of FS optimization methods, particularly for enhancing text classification tasks.

In recent decades, the widespread adoption of the Internet has revolutionized the operations of individuals, organizations, and governments. Modern architectures such as the Internet of Things (IoT), Internet of Medical Things (IoMT), Industrial Internet of Things (IIoT), Internet of Flying Things (IoFT), among others, present incredible opportunities for intelligent living and the well-being of humanity. Consequently, a massive amount of digital data is generated daily in various forms, including texts, numbers, audios, videos, tapes, graphs, images, and more, serving as extensions of knowledge.

1.2 ADVANCES IN SINE COSINE ALGORITHM: A COMPREHENSIVE SURVEY

The Sine Cosine Algorithm (SCA) was introduced by Mirjalili in 2016 as a population-based optimization method inspired by trigonometric sine and cosine functions. This overview is followed by a survey of various SCA variants and applications found in the literature, and subsequently, the results of computational experiments are presented to validate the performance of SCA when compared to similar algorithms.

Optimization methods, particularly meta-heuristics, are typically categorized into two fundamental parts: local search methods and population search methods. Local search methods operate with a single solution during their optimization process, aiming to enhance their solution using a neighborhood mechanism such as tabu searches, simulated annealing, hill climbing, or -hill-climbing. Although these methods excel in exploitation search-ability, their main drawback is their emphasis on exploitation (local) search rather than exploration (global) searches. In contrast, population search methods utilize multisolutions (populations) in each run, continually improving candidate solutions by generating one or more better solutions at each iteration.

While population search methods are effective in identifying and promoting areas in a broader search space, they may be inefficient in exploiting the widest regions of the search space, with a higher likelihood of getting trapped in local optima. Despite this, their key strength lies in their exploitation search-ability, while their primary weakness is the focus on exploitation (local) search methods over exploration searches.

1.3 SALP SWARM ALGORITHM: THEORY LITERATURE REVIEW AND APPILCATION IN EXTREME LEARNING MACHINES

The Salp Swarm Algorithm (SSA) is a recent metaheuristic inspired by the collective behavior of salps in oceans. Since its introduction, SSA has demonstrated efficiency across various applications. This chapter provides an in-depth exploration of the algorithm, its operators, and highlights notable works that have utilized SSA. Additionally, the chapter investigates the application of SSA in optimizing the Extreme Learning Machine (ELM) to enhance accuracy and address shortcomings in its conventional training method. To validate its effectiveness, the algorithm undergoes testing on 10 benchmark datasets and is compared to two well-known training methods. The comparative results reveal that SSA-based training methods outperform others in terms of accuracy and exhibit strong competitiveness in prediction stability.
Artificial Neural Networks (ANNs) are mathematical models extensively applied in machine learning for both supervised and unsupervised tasks. These networks can be categorized based on their architecture and learning algorithms. The architecture refers to the arrangement of basic processing elements across multiple layers, while the learning algorithm is responsible for optimizing connection weights to minimize or maximize predefined criteria. The Single Hidden Layer Feedforward Networks (SLFN) stand out as a popular type of ANN in the literature. Typically, SLFNs are trained using the Back-propagation algorithm, a gradient descent training method.

1.4 VALUE OF GEOLOGICALLY DERIVED FEATURES IN MACHINE LEARNINGFACIES CLASSIFICATION

The Purpose of enhance machine learning-based facies classification with uncertainty assessment by incorporating geologically interpretive features. Manual interpretation of lithofacies from wireline log data is traditionally a labor-intensive task conducted by specialists, and for large datasets, it can be time-consuming and prone to biases. Quantifying interpretational uncertainty in facies categorization is challenging but crucial, as decisions related to reservoir development may heavily rely on it. Utilizing machine learning to automate facies categorization offers an efficient and straightforward approach to interpreting facies from vast amounts of data. Additionally, it ensures that potential alternative lithological scenarios are considered, thereby improving the measurement of uncertainty in facies classification.

To enhance the classifier's geological insight and interpretability, this work proposes the inclusion of geological features and expert knowledge as additional inputs. This augmentation aims to improve the performance of purely data-driven classifiers by providing additional context. The study contrasts random forest classifiers with support vector machines to demonstrate the superiority of the former. Specifically, the research compares facies categorization using additional geological features against a scenario where only standard wireline log inputs are considered. The first experiment incorporates geological rule-based constraints as extra inputs, derived to capture the primary geological characteristics that a geologist or petrophysics specialist might use to identify and classify typical wireline log responses.

1.5 RECENT ADVANCES IN HARRIS HAWKS OPTIMIZATION: A COMPARATIVE STUDY AND APPLICATIONS

The Harris Hawk Optimizer (HHO) stands out as a recent metaheuristic algorithm in the field of population-based optimization, designed to emulate the hunting strategies of hawks. This swarm-based optimization approach distinguishes itself through its unique strategies for exploration, exploitation, and multiphase search. This review emphasizes the applications and progress of HHO, acknowledging its robust nature and recognition as one of the leading swarm-based techniques in 2020. Numerous experiments have showcased the potency and efficacy of HHO by comparing it with nine other cutting-edge algorithms, employing Congress on Evolutionary Computation (CEC) benchmarks such as CEC2005 and CEC2017. The literature review explores potential future directions, suggesting the exploration of new HHO algorithm variants and diverse applications.

The optimization field has witnessed extensive applications in addressing emerging challenges brought about by the rapid advancements in industrial technologies and artificial intelligence. This accelerated progress underscores the need for technologies capable of efficiently addressing complex problems within reasonable timeframes. Researchers, drawing from years of experience, have proposed solutions based on two primary classes of optimization methods: deterministic and stochastic. The former relies on gradient information and specific details of the search space, while the latter operates without such information, managing black-box optimization through sensing and interacting with problem surfaces. Swarm and evolutionary methods have emerged prominently within the stochastic optimizer class, and this paper enumerates various metaheuristic algorithms within this category, elucidating their algorithmic behaviors.

1.6 FACIES AND LITHPFACIES CLASSIFICATIONS FROM WELL LOGS

Lithofacies, representing distinct categories of rock quality with two or more states, introduce small to intermediate-scale variabilities in the geological analysis of subsurface formations. As cores are often limited, lithofacies data is frequently derived from well logs for reservoir characterization. Traditionally, automatic lithofacies classifications have been performed using statistical and neural network methods. This chapter underscores the importance of an integrated approach through data analytics, especially as automatic techniques prove effective mainly in simple cases or with extensive training data. By adopting a holistic integrated approach, consistent lithofacies classifications can be achieved even with limited training data. Machine learning methods can further enhance lithofacies classifications through synergies with data analytics.

The dominance of lithofacies can lead to multimodalities in well log histograms, and sometimes just two well logs are sufficient for classification. However, selecting input logs for lithofacies classification can be challenging due to interdependencies among available logs. While some argue against choosing highly correlated logs, emphasizing that they may not provide enough additional information, lower correlations align with classical criteria of minimizing within variance and maximizing between variance. The goal is to observe a natural separation between classes, with selected attributes distinctly outlining lithofacies. While correlation is a consideration, it should not be the sole criterion, as moderate to high correlations between input attributes often do not diminish their discriminability. Numerous examples, including GR-resistivity porosity-GR, and neutron-density, are provided to illustrate these principles.

1.7 3-D WELL PATH DESIGN USING A MULTI OBJECTIVE GENERIC ALGORITHM

Efficiently navigating wellbore trajectories to reach specific subsurface locations, incorporating a mix of vertical, deviated, and horizontal components, necessitates minimizing both wellbore length and frictional torque on the drill string. This becomes crucial for shallow horizontal wells limited by torque constraints. The simultaneous minimization of wellbore length and torque can lead to quicker and more cost-effective drilling compared to alternative trajectories. Balancing these conflicting objectives is inherently nonlinear. To address this, a multi-objective genetic algorithm (MOGA) methodology is developed, employing two objective functions—wellbore length and torque—to generate a set of Pareto optimal solutions. The MOGA's performance surpasses that of single-objective function approaches, demonstrating swift convergence to a set of Pareto optimal solutions. Adopting an adaptive approach, allowing genetic algorithm parameters to evolve with iterations, further accelerates convergence, outperforming scenarios with constant parameters. The appendix includes algorithm code listings for both MOGA and genetic algorithm (GA) applications in the presented analysis.
Drilling wellbores is a costly aspect of oil and gas exploration and development. In the current market conditions with high costs and low oil and gas prices globally, minimizing drilling expenses is a priority for companies. Studies show that drilling a horizontal well is approximately 1.4 times more expensive than drilling a vertical well. Horizontal and directional wells offer advantages, allowing access to a larger reservoir volume and effectively intersecting high-quality zones, leading to increased production and recovery rates. Consequently, reducing wellbore length within defined constraints can decrease the time to reach the target and, subsequently, reduce total drilling costs. Optimization, extensively applied in the petroleum industry, aims to enhance various aspects of drilling operations, including wellbore trajectory planning. Previous studies utilized genetic algorithms and particle swarm optimization with a single objective function—minimizing wellbore length—subject to defined constraints.

1.8 A COMPARISON OF MACHINE LEARNING ALGORITHMS IN PREDICTING LKITHOFACIES: CASE STUDIES FROM NORWAY AND KAZAKHSTAN

Accurately delineating the distinctive physical properties of rocks assumes a critical role in evaluating reservoirs and optimizing hydrocarbon production, particularly given the challenges associated with acquiring core data from all wells. This study specifically focuses on evaluating lithofacies values through the application of machine learning algorithms to well log data collected from both Kazakhstan and Norway. Additionally, the incorporation of wavelet-transformed data is intended to unveil essential geological properties. The research involves a comprehensive numerical analysis of over 90 released wells from Norway and 10 wells from Kazakhstan, employing various machine learning algorithms such as KNN, Decision Tree, Random Forest, XGBoost, and LightGBM. Model evaluation utilizes metrics like accuracy, Hamming loss, and penalty matrix, while also exploring the impact of dataset features on predictions. The results underscore that the Random Forest model surpasses other considered algorithms, aligning with the SHapley Additive exPlanations (SHAP) framework. The understanding of geological structure heavily relies on data, with geophysicists contributing by defining key features in the subsurface. Although leveraging their expertise demands extensive hours of work and additional data sources like seismic surveys and cores, machine learning algorithms offer a systematic approach to expedite accurate predictions. Achieving precision in predictions requires a balanced utilization of a substantial amount of data in the training process, where geophysicists identify common regional features. Additional well log data is then employed to estimate uncommon features, taking into account relationships among lithotypes such as PS, RHOB, and NPHI.

1.9IMPROVED WHALE OPTIMIZATION ALGORITHM FOR SOLVINGCONSTAINED OPTIMIZATION PROBLEMS

To address the limitations of the whale optimization algorithm (WOA), including issues such as slow convergence speed, low accuracy, and vulnerability to local optima, an enhanced version known as the Improved Whale Optimization Algorithm (IWOA) is introduced. This upgraded iteration refines the standard WOA by adjusting parameters related to the initial population, convergence factor, and mutation operation, while also incorporating Gaussian mutation. The approach employs the no-fixed-penalty function method to transform constrained problems into unconstrained ones. Through extensive evaluation on 13 benchmark problems, the proposed IWOA demonstrates notable advantages, including enhanced global search capability, improved stability, faster convergence speed, and higher accuracy. These attributes collectively make the IWOA particularly effective in addressing complex constrained optimization problems. Originally introduced by Mirjalili and Lewis in 2016, the WOA emulates humpback whale hunting behavior, seeking to solve target problems by mimicking the predatory actions of whales. Despite its popularity in various applications, swarm intelligence optimization algorithms, including WOA, face common challenges such as slow convergence and susceptibility to local optima. In response to these challenges, recent years have witnessed scholars implementing numerous enhancements to improve the convergence speed and optimization accuracy of WOA, as well as similar algorithms like DE, PSO, and ACO.

1.10 THERMAL MATURITY AND BURIAL HISTORY MODELLING OF SHALE IS ENHANCED BY USE OF ARRHENIUS TIME-TEMPARATURE INDEX AND MEMETIC OPTIMIZER

In the realm of oil and gas exploration, thermal maturity indices and modeling grounded in Arrhenius-equation reaction kinetics have proven indispensable for understanding petroleum generation in kerogen-rich source rocks. The optimal integration of the Arrhenius equation, along with determining suitable values for activation energies (E) and frequency factors (A), remains a topic of discussion. The Time-Temperature Index (PTTIARR) method, introduced in 1998, is advocated for its solid theoretical foundation and practical benefits. This method utilizes a carefully chosen set of E-A values (E = 218 kJ/mol (52.1 kcal/mol); A = 5.45E+26/my) from the established A-E trend for published kerogen kinetics. The PTTIARR method, which spans approximately eighteen orders of magnitude, is easily calculated using spreadsheets, and Visual Basic can further optimize application code.

The optimization process aids in identifying potential geothermal gradients and erosion intervals, covering various burial intervals and aligning with measured Ro data and computed thermal maturities. A memetic optimizer, incorporating firefly and dynamic local search memes, is employed to explore and exploit the multi-dimensional thermal history solution space, yielding high-performing solutions for complex burial and thermal histories. An illustrative example featuring a sophisticated deep burial history, encompassing multiple uplift and erosion periods and varying heat flow, highlights the capabilities of the memetic optimizer.

By carefully layering constraints, unique insights into specific episodes in the models' thermal histories can be uncovered, enhancing the characterization of the time of petroleum formation. The mean square error (MSE) of several burial periods for the discrepancy between measured and calculated Ro is identified as the most efficient objective function for this optimization. Basin-wide thermal maturity modeling with numerous pseudo-wells, utilizing the sensitively-scaled PTTIARR methodology and the memetic optimizer, allows for rapid attainment of thermal maturity analysis at fine granularities.

Comparative Table:

No	Authors	Title	Year	Methodology	Demerits
1	Esther Omolara Abdulatif Abiodun Alawida Abdullah Alabdulati Rami S. Alkhawaldeh	A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities	2021	Data Preparation Techniques Feature Selection	1.Lack of Best Practices 2.Limited Discussion on Overfitting
2	LaithAbualigahAli Diabat	Advances in Sine Cosine Algorithm: A comprehensive survey	2021	SCA Algorithm Computational Experiments Comparison with Similar Algorithms	1.Lack of Specifics 2.Absence of Methodological Details
3	Hossam Faris, Seyedali Mirjalili, Ibrahim Aljarah, Majdi Mafarja Ali Asg	Salp Swarm Algorithm: Theory, Literature Review, and Application in Extreme Learning Machines	2020	Benchmark Datasets Data Collection and Analysis	1.Limited Methods 2.Data Source and Benchmark Datasets

II. CONCLUSION

The proposed optimization methodologies within machine learning, data analytics, and oil and gas exploration. It begins with a systematic review of emerging feature selection methods for text classification, emphasizing the limitations of traditional techniques and the potential of metaheuristics. Subsequent sections cover algorithms such as the Sine Cosine Algorithm and Salp Swarm Algorithm, showcasing their efficiency in optimization tasks. The text also explores the application of machine learning in lithofacies classification, advocates for integrated approaches, and introduces multi-objective genetic algorithms for wellbore trajectory optimization. Furthermore, it discusses the Harris Hawk Optimizer, an algorithm favored in 2020, and presents an improved whale optimization algorithm addressing its drawbacks. The content concludes with the integration of the Arrhenius equation and a time-temperature index in oil and gas exploration, showcasing a memetic optimizer's role in uncovering insights from complex burial histories. Overall, these discussions offer a comprehensive overview of cutting-edge optimization methods and their applications in diverse domains.

REFFERENCES

[1] Abiodun, E.O., Alabdulatif, A., Abiodun, O.I., Alawida, M., Abdullah Alabdulatif, A., Alkhawaldeh, R.S., 2021. A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. Neural Comput. Appl. 33, 15091–15118.

[2] Abualigah, L., Diabat, A., 2021. Advances in sine cosine algorithm: a comprehensive survey. Artif. Intell. Rev. 54, 2567.

[3] Faris, H., Mirjalili, S., Aljarah, I., Mafarja, M., Heidari, A.A., 2020. Salp swarm algorithm: theory, literature review, and application in extreme learning machines. In: Mirjalili, S., Song Dong, J., Lewis, A. (Eds.), Nature-inspired Optimizers, Studies in Computational Intelligence, vol. 811. Springer, Cham.

[4] Halotel, J., Demyanov, V., Gardiner, A., 2020. Value of geologically derived features in machine learning facies classification. Math. Geosci. 52, 5–29.

[5] Hussien, A.G., Abualigah, L., Abu Zitar, R., Hashim, F.A., Amin, M., Saber, A., Almotairi, K.H., Gandomi, A.H., 2022. Recent advances in Harris Hawks optimization: a comparative study and applications. Electronics 11, 1919.

[6] Ma, Y.Z., 2019. Facies and lithofacies classifications from well logs. In: Quantitative Geosciences: Data Analytics, Geostatistics, Reservoir Characterization and Modeling.

[7] Mansouri, V., Khosravanian, R., Wood, D.A., Aadnoy, B.S., 2015. 3-D well path design using a multi- objective genetic algorithm. J. Nat. Gas Sci. Eng. 27 (1), 219–235.

[8] Merembayev, T., Kurmangaliyev, D., Bekbauov, B., Amanbek, Y.A., 2021. Comparison of machine learning algorithms in predicting lithofacies: case studies from Norway and Kazakhstan. Energies 1896.

[9] Ning, G.-Y., Cao, D.-Q., 2021. Improved whale optimization algorithm for solving constrained optimization problems. Discrete Dynam Nat. Soc. 2021, 8832251

[10] Wood, D.A., 2018. Thermal maturity and burial history modelling of shale is enhanced by use of Arrhenius time-temperature index and memetic optimizer. Petroleum 4, 25–42.