***The problem of oil displacement was solved using neural networks and machine learning classifiers. The Buckley-Leverett model is selected, which describes the process of oil displacement by water. It consists of the equation of continuity of oil, water phases and Darcy's law. The challenge is to optimize the oil displacement problem. Optimization will be performed at three levels: vectorization of calculations; implementation of classical algorithms; implementation of the algorithm using neural networks. A feature of the method proposed in the work is the identification of the method with high accuracy and the smallest errors, comparing the results of machine learning classifiers and types of neural networks. The research paper is also one of the first papers in which a comparison was made with machine learning classifiers and neural and recurrent neural networks. The classification was carried out according to three classification algorithms, such as decision tree, support vector machine (SVM) and gradient boosting. As a result of the study, the Gradient Boosting classifier and the neural network showed high accuracy, respectively 99.99 % and 97.4 %. The recurrent neural network trained faster than the others. The SVM classifier has the lowest accuracy score. To achieve this goal, a dataset was created containing over 67,000 data for class 10. These data are important for the problems of oil displacement in porous media. The proposed methodology provides a simple and elegant way to instill oil knowledge into machine learning algorithms. This removes two of the most significant drawbacks of machine learning algorithms: the need for large datasets and the robustness of extrapolation. The presented principles can be generalized in countless ways in the future and should lead to a new class of algorithms for solving both forward and inverse oil problems***

***Keywords: Buckley-Leverett model, neural network, machine learning, architecture, metric, training***

UDC 004.8, UDC 519.62
DOI: 10.15587/1729-4061.2021.241858

# IMPLEMENTATION OF THE SOLUTION TO THE OIL DISPLACEMENT PROBLEM USING MACHINE LEARNING CLASSIFIERS AND NEURAL NETWORKS

**Beimbet Daribayev**
PhD*
**Aksultan Mukhanbet**
*Corresponding author*
Master of Engineering Science*
E-mail: mukhanbetaksultan0414@gmail.com
**Yedil Nurakhov**
Master of Computer Science*
**Timur Imankulov**
PhD*
*Department of Computer Science
Al-Farabi Kazakh National University
Al-Farabi ave., 71, Almaty,
Republic of Kazakhstan, 050040

*Received date 13.09.2021*
*Accepted date 19.10.2021*
*Published date 29.10.2021*

***How to Cite:*** *Daribayev, B., Mukhanbet, A., Nurakhov, Y., Imankulov, T. (2021). Implementation of the solution to the oil displacement problem using machine learning classifiers and neural networks. Eastern-European Journal of Enterprise Technologies, 5 (4 (113)), 55–63. doi: https://doi.org/10.15587/1729-4061.2021.241858*

## 1. Introduction

Machine learning and computer vision have long been one of the most popular areas of research. In recent years, machine learning methods have been widely used in various fields of science for processing data. The use of machine learning in the oil industry is also actively expanding. To solve the problems of oil recovery, it is necessary to use geological models of reservoir deposits. As the size of the reservoir model increases, the computation time increases. Therefore, it takes more time to predict oil recovery.

One of the approaches to solving the problem is to use machine learning methods and neural networks, which is the purpose of this paper. This paper discusses approaches to using effective machine learning methods for predicting oil recovery. To train the system, historical data from the oil field and synthetic data obtained from surrogate models based on two wells (injection and production) were used. Synthetic data were obtained on the basis of mathematical models (oil displacement models, enhanced oil recovery models) by varying various geological parameters. This problem relates to «supervised learning» – a type of machine learning. For supervised learning, a complete set of marked-up data is required for training the model at all stages of its construction. When implementing the algorithm, machine learning and classification methods were taken into account. As a result, it was found that compared to traditional computational experiments on a regular grid, calculations using machine learning methods are more productive.

Energy systems represent a unique set of problems for machine learning methods. They have large financial and environmental consequences, great uncertainty and strict physical restrictions. Recent advances in instrumentation, telemetry, and data storage have allowed operators to increasingly rely on data for decision-making. However, the integration of all these data is a problem, while the need to inform decisions in a timely manner has become an important competitive difference in most industries [1]. One of the main motivations of this work is that there is still no reliable methodology for applying machine learning to terrestrial systems.

After reviewing the literature on existing modern physics-informed machine learning methods, the method and

its application to the Buckley-Leverett problem are presented [2]. Variations of the original problem will continue, including several physical and dimensional ones.

Currently, calculations are one of the most important sections in the field of oil and gas production. Oil displacement and other processes can be simulated. Many leading oil and gas companies pay more attention to calculation, because without it, the future development of companies is impossible. The problem of oil displacement in a porous medium is complex, and using machine learning methods, it is possible to correctly predict data for oil production and displacement.

## 2. Literature review and problem statement

In [3], the technique of a physics informed neural network (PINN) is implemented, which includes information from the physics of fluid flow, as well as observed data for modeling the Buckley-Leverett problem. The classical problem of removing gas into a porous medium filled with water was used for checking the implementation. Several cases have been tested that show the importance of the connection between the observed data and physics-informed neural networks for various parameter spaces. The authors of the paper show that PINNs are able to capture the general trend of the solution even without the observed data, but the resolution and accuracy of the solution are significantly improved with the observed data.

The paper [4] explores the development of an artificial neural network predicting oil production in the State Oil Company of the Azerbaijan Republic (SOCAR). The multi-layer perceptron neural network is used to predict oil production. As a result, 99 % of the training accuracy is achieved. The authors compared the accuracy of the developed model with empirical correlations. The disadvantage is that the paper presents only a model of a neural network, the forecast with real data was not implemented. An excellent correspondence was found between the neural network and real data.

In [5], a new model for predicting oil production is proposed. The model is a recurrent neural network with deep synchronization, consisting of several hidden layers, each of which has several nodes. The proposed model has a simple architecture and the ability to track time series data sets with a large interval. To assess the reliability of the model, the proposed method was tested using various standard approaches. As a result of the work, the authors show that the proposed model is superior to existing approaches. The disadvantage is that the convergent networks that are presented in the papers are relatively slow, especially when capturing long-interval dependencies.

In [6], the application of an artificial neural network (ANN) for predicting oil recovery and storage capacity of $CO_2$ in the RE was presented. Uncertainty parameters, including geological factors and well operations, were used to create a training database. Then a total of 351 numerical samples were modeled and the total oil production, total storage of $CO_2$ and total accumulation of $CO_2$ were created. According to the results of the study, the developed ANN model had excellent forecasting characteristics with a high correlation coefficient ($R^2$) exceeding 0.98 compared to objective values, and a total standard error of less than 2 %. Results comparison with other papers gives the highest accuracy and speed of learning.

In [7], the vertical deformation of the Earth's surface was used to calculate changes in the reservoir pore pressure. A marker stand was installed to measure the displacement of the ground and the vertical deformation in place was measured. In addition, the authors of the paper provided an improved new convolutional neural network (CNN), which adopted the image-to-image mode, removed the union layers and full connection layers, and also used a new loss function that takes into account the matrix of the border influence coefficients. Then, the machine learning method in the study was used to invert the vertical deformation of the surface to change the pore pressure in the oil reservoir. A lot of time was spent in processing and training the model.

The paper [8] describes a new approach to managing large projects for pumping liquid into dense reservoirs with hydraulic fracturing. Neural networks were used to analyze the past results of flooding projects and predict future oil recovery, as well as water injection and production. Neural networks are useful because you do not need to specify a structural model between injection and production to predict performance. The neural network approach takes into account that the behavior of an individual well may depend on the history of the well and the injection/production conditions in the surrounding wells. In addition, rental production is the result of injection and production at many wells and their interaction. The authors of the paper envisage this approach as an injection policy that leads to a minimum amount of injected water and the best oil recovery.

In [9], a new hybrid approach to reservoir modeling based on machine learning and physics was presented. The model is a neural network that is jointly trained to compare with any available experimental data and respect the basic physical laws. In the paper, the approach is used as a new way of modeling and comparing flow and transport problems in porous media.

The paper [10] presents the application of deep learning taking into account physics to the problems of reservoir modeling. The model is a neural network that is jointly trained to respect the basic physical laws and boundary conditions. The present methodology is applied to the Buckley-Leverett model. The model is able to give an exact physical solution both from the point of view of the compaction jump and the rarefaction, and takes into account the basic partial differential equation together with the initial and boundary conditions. Various hypotheses (homogeneous and inhomogeneous initial conditions) were tested and show that with the correct implementation of physical constraints, a reliable solution can be trained in a reasonable amount of time and iterations. According to the authors, the presented principles can be generalized in countless ways in the future and should lead to a new class of algorithms for solving both direct and inverse physical problems.

## 3. The aim and objectives of the study

The aim of this study is to predict and improve the efficiency of solving the oil displacement problem with machine learning methods. The scientific novelty of this work lies in the identification of a method with high accuracy and the smallest errors for solving oil displacement problems using machine learning classifiers and neural networks. This will allow predicting oil displacement with a good and accurate model.

To achieve the aim, the following objectives were set:

– to collect a dataset using the Buckley-Leverett model for oil displacement;

– to implement oil displacement forecasting using machine learning methods;

– to build a three-dimensional model containing such indicators as accuracy, recall and F1-measure for all classes;

– to implement prediction of oil displacement problems using neural networks;

– to conduct a numerical assessment of the quality of algorithms in order to determine the best classifier in forecasting problems.

## 4. Materials and methods of research

### 4. 1. The Buckley-Leverett physical model

Consider the calculation of pumping water from an injection well at a certain temperature for a certain period of time. Pressures $P_{inj}$ and $P_{prod}$ ($P_{inj} > P_{prod}$) are given in injection and production wells. The pumped water displaces the oil in the reservoir, which, in turn, enters the production well. It is necessary to study this problem and develop a computer model for analyzing non-mass processes during oil displacement in the reservoir. Fig. 1 shows the layout of injection and production wells.

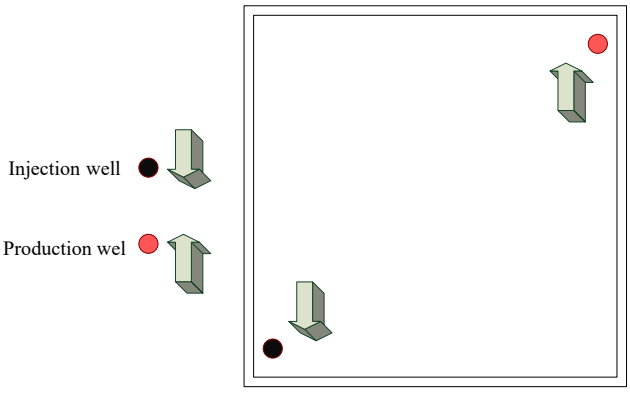



Fig. 1. Scheme of injection and production wells (2D)

The mathematical model of two-phase filtration consists of the equations of the balance of water and oil in the flow. The system of equations for the area Ω across the boundary interval ∂Ω is written as follows:

$$m\frac{\partial s_1}{\partial t}+\operatorname{div}\left(\vec{v}_1\right)=0, \quad (1)$$

$$m\frac{\partial s_2}{\partial t}+\operatorname{div}\left(\vec{v}_2\right)=0, \quad (2)$$

$$\vec{v}_1=-k\frac{f_1}{\mu_1}\nabla P, \quad (3)$$

$$\vec{v}_2=-k\frac{f_2}{\mu_2}\nabla P, \quad (4)$$

where $m$ is the porosity of the reservoir, $s_1$, $s_2$ are the saturation with water and oil, respectively, $0 \le s_1, s_2 \le 1$, $s_1+s_2=1$; $\vec{v}_1, \vec{v}_2$ – permeability rate; $k$ – absolute permeability; $f_1, f_2$ – relative phase permeability; $\mu_1, \mu_2$ – viscosity; $P$ – pressure.

Thus, the function $P$, that is, the pressure, must be found under the following initial and final conditions:

$$\left.\frac{\partial P}{\partial n}\right|_{\partial\Omega}=0, \quad (5)$$

where ∂Ω – the border of the territory [11].

For the calculation by the linear method, a pressure determination algorithm is used. In the $n$-th time layer of this pressure, the pressure is in this time layer, after which the calculations are repeated in the same sequence. To check the reliability of the results, the compliance of the flow rates of the injection and production wells is checked and monitored.

Calculations are made in the following order:

– initial data required for the calculation are provided;

– the pressure distribution is calculated before the condition is met;

– the next calculation takes place in time [12].

In this work, the obtained synthetic data of the mathematical model were divided into training and test samples. Four parameters were taken as the input parameters of the machine learning model, and the oil recovery factor was taken as the output parameter.

Various combinations of parameters of the oil production problem (porosity, viscosity of the oil phase and absolute permeability of the rock) were taken as input parameters. The value of the oil recovery factor was chosen as the output parameter. Using the Buckley-Leverett model, 6 synthetic datasets were generated for different permeability indices. Each dataset contains values for viscosity, porosity and oil recovery factor. Oil viscosity ranges from 0.1–0.5, porosity ranges from 0.1–0.3 and various permeability options. The Buckley-Leverett mathematical model from which the data were taken can be seen in [12].

### 4. 2. Machine Learning Algorithms

The first machine learning method used in this research paper is decision trees [13]. The entire training set is considered at the root of the tree. If it is executed, a forecast is selected that will be issued for the node, which can be done in several ways. Otherwise, you need to split the set into several disjoint ones. As a rule, a decision rule is set at the vertex, which takes a certain range of values. This range is divided into disjoint sets of objects, where the number of descendants is at the top, and each is a set of objects that fall into the descendant. The set at a node is divided according to the selected rule, the algorithm is run recursively for each node. The next machine learning algorithm chosen to predict data for oil displacement problems is a support vector machine [14]. The idea behind a support vector machine is to construct a hyperplane that acts as a solution surface, separating the positive and negative examples from the training set as much as possible. In particular, a support vector machine is an approximate implementation of the structural risk minimization method, which is based on the fact that the error rate of a machine trained on a test set can be represented as the sum of learning errors.

The next research method is the gradient improvement algorithm [15], one of the most powerful algorithms in the field of machine learning.

The next machine learning research method is a neural network. In the study, consisting of 5 input data, 128 hidden layers and one output. In the neural network architecture [16], as shown in Fig. 2, the input parameter is the parameters from the Buckley-Leverett model, these are porosity, viscosity, absolute permeability, time iteration, oil recovery coefficient. The architecture of the neural network can be seen in Fig. 2.

As shown in the architecture of the algorithm (Fig. 3), data from the Buckley-Leverett equation are taken as input. Porosity, oil phase viscosity, absolute rock permeability and

time iteration were set as input parameters. If there is raw data, then these data are preprocessed, and the missing values are filled using the linear interpolation function. It is then processed to validate outliers, scaled to a given range, and then split into training and testing subsets.

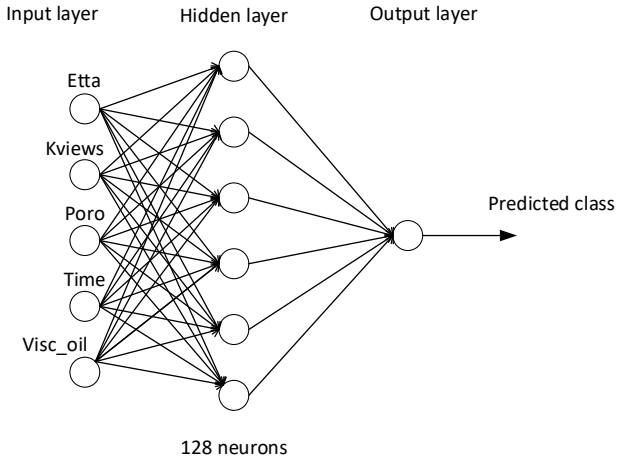



Fig. 2. Neural network architecture for the oil displacement problem

After dividing the data, this data were trained using neural networks and classifiers. The data consist of 9 classes, but 8 classes were taken for high accuracy of the model. When the data were trained with 9 classes, our models were retrained and the accuracy of the models showed a lower result.

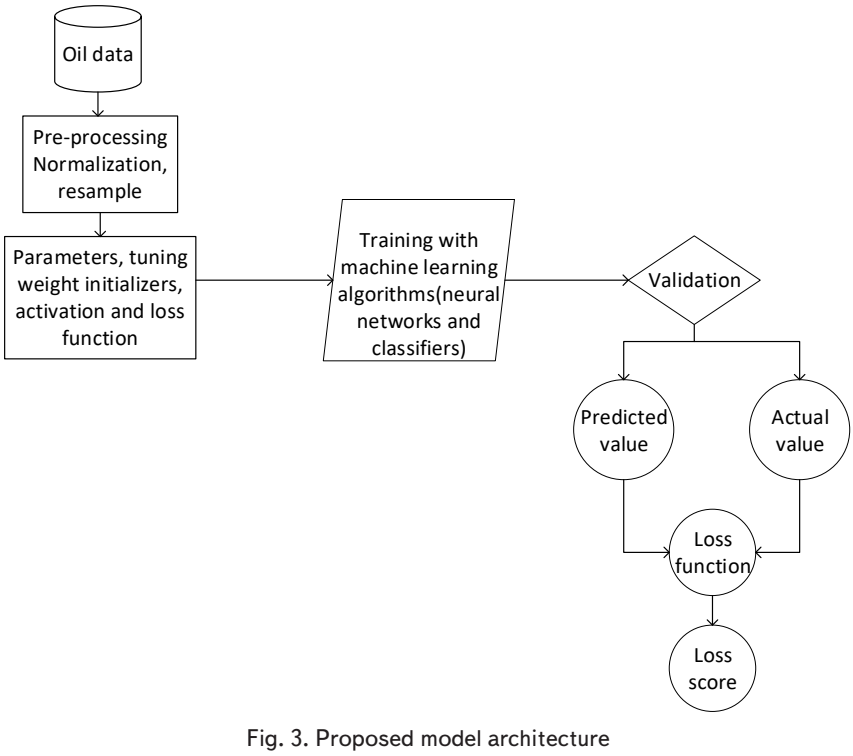



Fig. 3. Proposed model architecture

The optimal number of NN layers and the window size are selected using a genetic algorithm. The best parameters are modeled to compare the performance of the NN and RNN models.

## 5. Results of the study of machine learning methods and neural networks

### 5. 1. Data collection using the Buckley-Leverett model

In general, these data are synthetic data of the Buckley-Leverett mathematical model [12]. The total amount of data is 67240.

The parameters in the Buckley-Leverett model were obtained (Fig. 4), namely: porosity, viscosity, absolute permeability, time iteration, oil recovery coefficient.

| | Class | etta | Kviews | Poro | Time | Visc_oil |
|---|---|---|---|---|---|---|
| 0 | 1.0 | 0.004099 | 2.0 | 0.13089 | 0.0 | 0.414938 |
| 1 | 1.0 | 0.028301 | 2.1 | 0.13090 | 200.0 | 0.414939 |
| 2 | 1.1 | 0.049026 | 2.2 | 0.13091 | 400.0 | 0.414940 |
| 3 | 1.2 | 0.068843 | 2.3 | 0.13092 | 600.0 | 0.414941 |
| 4 | 1.3 | 0.087731 | 2.4 | 0.13093 | 800.0 | 0.414942 |

Fig. 4. Synthetic data of the Buckley-Leverett mathematical model

By studying the data set and visualizing these binned counts as columns, you can get a very direct and intuitive idea of the distribution of values within a variable.

### 5. 2. Results of the machine learning classifier in the solution

At the next stage, the quality of classifiers is evaluated using the indicators of accuracy, recall and F1-measure [17]. Heatmaps and classifier plots were generated using these metrics.

To calculate these results, a Jupyter Notebook and an Intel Core i7-10750H computer, NVIDIA GeForce GTX 1660 Ti were used.

The process of constructing decision trees is a sequential, recursive partition of the training set into subsets using decision rules at the nodes. The splitting process continues until all nodes at the end of all branches are declared leaves. Declaring a node as a leaf can occur naturally, or upon reaching a certain stopping condition. The paper used the minimum number of examples per node and the maximum tree depth. To display the forecast results and metrics, a classification report is used. It is needed to measure the quality of forecasts of the classification algorithm. How many predictions are correct and how many are false. In particular, true positive, false positive, true negative and false negative results are used to predict the indicators of the classification report, as shown in Fig. 5. With the help of this data, 3 classification metrics were taken. These metrics are precision, recall, F1-score. Using the Seaborn heatmap, visualization data were generated.

The report shows the main classification indicators: accuracy, recall and F1-score for each class. Metrics are calculated using true and false positives, true and false negative results. For the Decision Tree classifier, the forecast accuracy is 96 %.

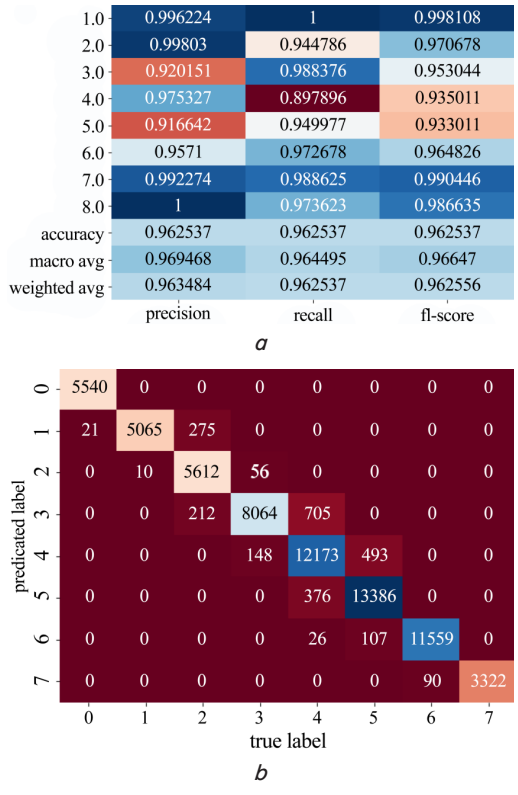



Fig. 5. Classification metrics:
*a* – Classification report for Decision Tree; *b* – Confusion matrix for the Decision Tree model

On the ROC (Receiver Operating Characteristic) curve, a higher value on the X-axis indicates a greater number of false positives than true negative ones [18]. While a higher value on the Y-axis indicates a greater number of true positive results than false negative ones. So, the choice of the threshold depends on the ability to balance between false positives and false negatives. A quantitative interpretation of this curve is given by an indicator of AUC (Area Under Curve) (Fig. 6), bounded by the ROC curve and the axis of the proportion of false-positive classifications. The higher the AUC result, the better the classifier works. ROC curve is simply a TPR relationship to FPR. True Positive Rate (TPR) indicates what percentage among all positively predicted models. False Positive Rate (FPR): What percentage among all negative is incorrectly predicted by the model.

Fig. 6 shows the worst and random sequence of labels. The ideal corresponds to the ROC curve passing through the point (0, 1), the area under it is equal to 1. The worst is the ROC curve passing through the point (1, 0), the area under it is 0.

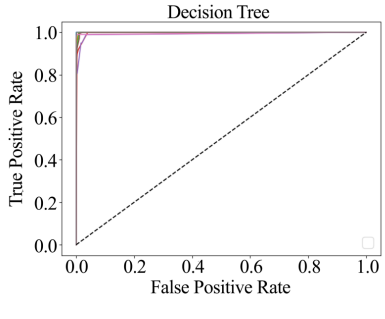



Fig. 6. Receiver Operating Characteristic curves for each class of the Decision Tree classifier

**5. 3. Building a three-dimensional model containing indicators such as accuracy, recall and F1-measure for all classes**

Fig. 7 shows a diagram of accuracy and completeness (*x*- and *y*-axes, respectively) and their corresponding indicator F1 (*z*-axis) for the classifier in the solution. When the accuracy value reaches one, and the completeness is zero, the F1-measure remains equal to 0, ignoring the accuracy. If one parameter is small, then the second parameter does not matter, since the F1-measure emphasizes the smallest value. Using the color indicator shown on the right side of the figure, you can see the ratio of accuracy and completeness for each class.

The SVM result shows less accuracy than other classifiers. Conditions in the construction of the SVM algorithm consist in the dividing hyperplane. The problem of constructing an optimal separating hyperplane is reduced to minimizing $w$, provided $c_i(w^*x_i-b)\geq 1$, $1\leq i\leq n$ [14]. This can be seen in the following results. When roc_auc_score predict results is called, an ROC curve is generated with only three points: bottom left, top right, and one point representing the model's solution function. This may be useful, but it is not a traditional ROC. A quantitative interpretation of this curve is given by an indicator of AUC (Fig. 8), bounded by the ROC curve and the axis of the proportion of false-positive classifications. The higher the AUC result, the better the classifier works.

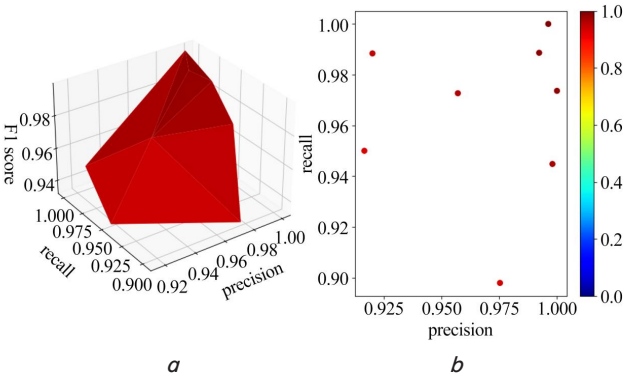



Fig. 7. The ratio of the quality indicators of the Decision Tree algorithm: *a* – the ratio of accuracy, completeness and F1 indicators; *b* – the ratio of accuracy and memorization indicators

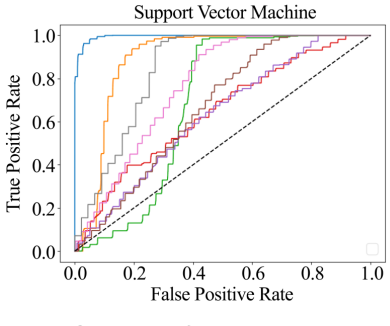
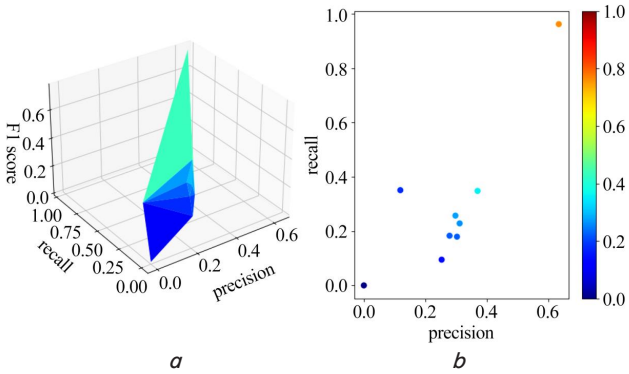



Fig. 8. Receiver Operating Characteristic curves for each class of the Support Vector Machine classifier:
– Class:0 ROC area=0.9951;
– Class:1 ROC area=0.8841;
– Class:2 ROC area=0.5898;
– Class:3 ROC area=0.6569;
– Class:4 ROC area=0.6434;
– Class:5 ROC area=0.6863;
– Class:6 ROC area=0.7673;
– Class:7 ROC area=0.8460

The ratio of the quality indicators of the SVM algorithm is shown in Fig. 9. Using the color indicator, you can see the ratio of accuracy and completeness for each class.




*a* *b*

Fig. 9. The ratio of the quality indicators of the Support Vector Machine algorithm: *a* — the ratio of accuracy, completeness and F1 indicators; *b* — the ratio of accuracy and memorization indicators

The idea of «gradient boosting» is to take a weak hypothesis or a weak learning algorithm and make a series of adjustments to it that will improve the strength of the hypothesis/trainee. This type of hypothesis confirmation is based on the idea of probably-approximately correct learning (PAC).

A quantitative interpretation of this curve is given by an indicator of AUC (Fig. 10), bounded by the ROC curve and the axis of the proportion of false-positive classifications. The higher the AUC result, the better the classifier works.

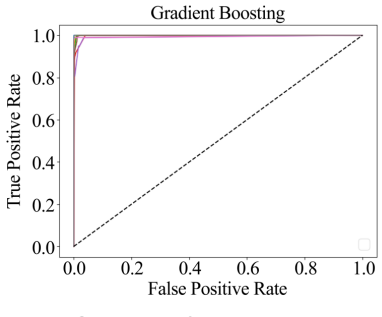
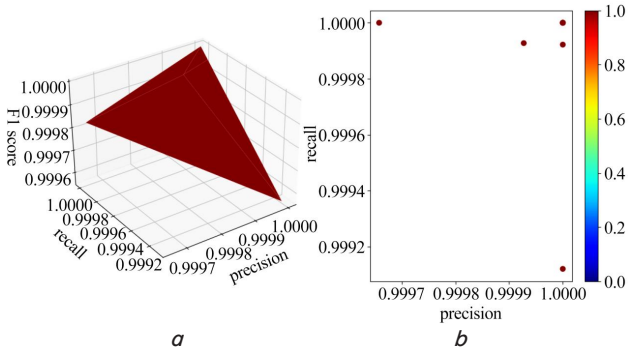



Fig. 10. Receiver Operating Characteristic curves for each class of the Gradient Boosting classifier:
— Class:0 ROC area=1.0000;
— Class:1 ROC area=0.9993;
— Class:2 ROC area=0.9994;
— Class:3 ROC area=0.9968;
— Class:4 ROC area=0.9915;
— Class:5 ROC area=0.9984;
— Class:6 ROC area=0.9947;
— Class:7 ROC area=1.0000

The ratio of the quality indicators of the Gradient Boosting algorithm is shown in Fig. 11. Using a color indicator, you can see the ratio of accuracy and completeness for each class.

So far, unsurprisingly, gradient boosting performs better than decision tree, which in turn performs better than SVM.




*a* *b*

Fig. 11. The ratio of the quality indicators of the Support Vector Machine algorithm: *a* — the ratio of accuracy, completeness and F1 indicators; *b* — the ratio of accuracy and memorization indicators

### 5. 4. Implementation of the task using neural networks

The neural network architecture used a total of 21658 parameters. In this study, porosity, oil phase viscosity, absolute rock permeability and time iteration were specified as input parameters for the NN. The oil recovery factor was specified as an output parameter of the neural network. The neural network consists of 5 hidden layers. The activation function is relu, and the output layer of the activation function was softmax. Mean squared error was used for the loss function. 5 input parameters were taken. It was found that five neurons in the hidden layer are the optimal number for neural networks [19].

There is a work in which many input parameters were used to train a neural network. This led to overfitting and underfitting of the neural network.

The standard method for generating ROC curves for neural networks is to change the threshold of the output node for classification. Fig. 12 shows that it generates a higher ROC curve result in the sense that it consists of a better distribution of operating points.

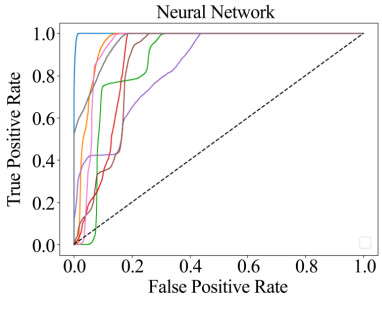
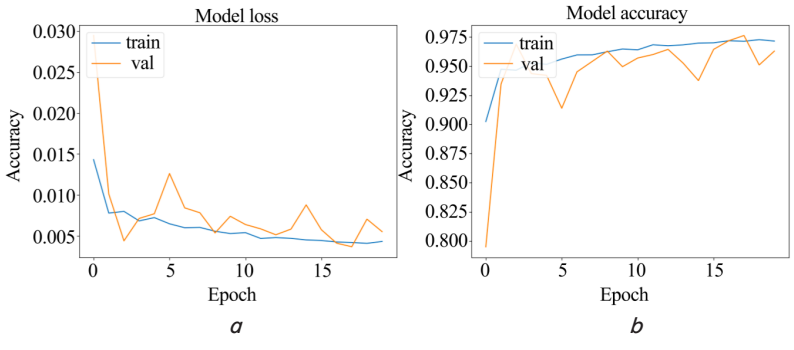



Fig. 12. Receiver Operating Characteristic curves for each class of neural networks:
— Class:0 ROC area=0.9995;
— Class:1 ROC area=0.9686;
— Class:2 ROC area=0.8771;
— Class:3 ROC area=0.7488;
— Class:4 ROC area=0.6806;
— Class:5 ROC area=0.7288;
— Class:6 ROC area=0.9353;
— Class:7 ROC area=0.9407

The smaller the loss, the better the model. The losses are calculated during training and validation, and their interaction shows how well the model copes with these two sets. Unlike accuracy, the loss is not a percentage. This is the sum of the errors made for each example in the training or test sets.

In the case of neural networks, the losses are usually a negative logarithmic probability and the residual sum of squares for classification and regression, respectively. Then, naturally, the main goal of the training model is to reduce (minimize) the value of the loss function with respect to the model parameters by changing the values of the weight vector using various optimization methods, such as back propagation in neural networks. The amount of loss shows how well or poorly a certain model behaves after each iteration of optimization. The interpretation of loss and accuracy for a neural network can be seen in Fig. 13.




*a* *b*

Fig. 13. Neural network metrics: *a* — loss function value; *b* — accuracy value

The next model for the study is a recurrent neural network. The summary of the created model contains the following parameters: layers and their order in the model, the output form of each layer, the number of parameters (weights) in each layer, the total number of parameters (weights) in the model [20].

A quantitative interpretation of this curve is given by an indicator of AUC (Fig. 14), bounded by the ROC curve and the axis of the proportion of false-positive classifications. The higher the AUC result, the better the classifier works.

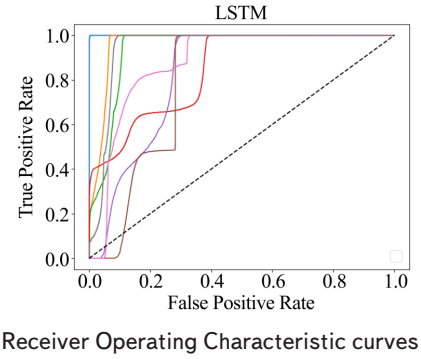
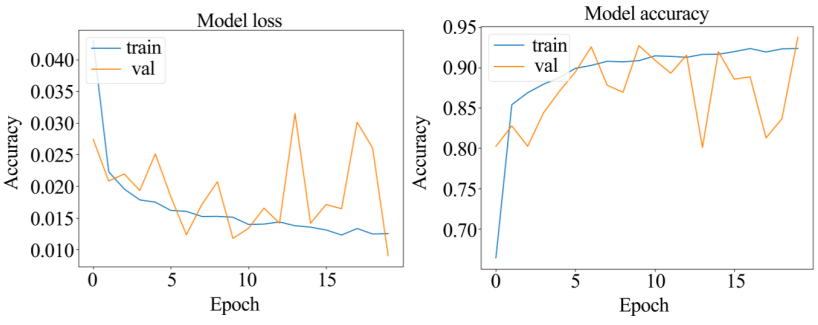



Fig. 14. Receiver Operating Characteristic curves for each class of recurrent neural networks:
— Class:0 ROC area=0.9979;
— Class:1 ROC area=0.8882;
— Class:2 ROC area=0.8795;
— Class:3 ROC area=0.8387;
— Class:4 ROC area=0.6737;
— Class:5 ROC area=0.7567;
— Class:6 ROC area=0.8350;
— Class:7 ROC area=0.9725

In Fig. 15, you can see that the error function of the training data is greater than that of the validation data.




*a* *b*

Fig. 15. Recurrent neural network metrics: *a* — loss function value; *b* — accuracy value

Based on this, the accuracy of the validation data is greater than that of the training data.

### 5. 5. Numerical evaluation of the quality of algorithms in order to determine the best model

So far, unsurprisingly, gradient boosting performs better than decision tree, which in turn performs better than SVM. The metrics shown in Table 1 were taken from the results of the classifiers. These results are shown in Fig. 7, 9, 11. According to Table 1, the Gradient Boosting classifier showed a more accurate result and a higher learning rate than other classifiers.

Table 1

Metrics for evaluating the quality of the Decision Tree, SVM, Gradient Boosting classifiers

| Classification algorithms | Precision | Recall | F1-score |
|---|---|---|---|
| Decision Tree | 0.962537 | 0.962537 | 0.962537 |
| SVM | 0.278971 | 0.28 | 0.278971 |
| Gradient Boosting | 0.999926 | 0.999926 | 0.999926 |

A comparative indicator of two neural networks shows that a neural network predicts more accurately than a recurrent neural network. These loss and accuracy function data shown in Table 2 were taken from the results of the neural networks shown in Fig. 13, 15. As shown in Table 2, the difference is small.

Table 2

Loss indicators and accuracy for the neural network models

| Neural Networks | val_loss | train_loss | val_accuracy | train_accuracy |
|---|---|---|---|---|
| NN | 0.026 | 0.021 | 0.971 | 0.974 |
| RNN-LSTM | 0.048 | 0.057 | 0.93 | 0.92 |

As shown in Table 2, the learning loss function is almost the same as the validation loss function. Compared to the loss function of recurrent neural networks, the loss function of neural networks shows a higher result.

## 6. Discussion of the results of comparative analysis of algorithms

This paper develops a solution to the problem of predicting oil displacement using neural networks and classifiers. The dataset for the solution was taken from the Buckley-Leverett model [12]. The parameters were obtained within the Buckley-Leverett model, which can be seen in Fig. 4, namely: porosity, viscosity, absolute permeability, iteration over time, oil recovery factor. A total of 67,000 synthetic data was collected, consisting of 5 parameters. It was possible to take many parameters during oil production, but it will take a lot of time to process and train these data and there may be retraining of neural networks. When solving this problem mathematically, some parameters of the medium and liquid were omitted or taken as constant values, since the inclusion of such results in a manifold increase in the calculation time. As the computational domain, a two-dimensional square section of the tank was chosen and the symmetry condition was chosen to find the boundary values. The calculations show fluid movement in the simulation, but when choosing the parameters of the existing reservoir, it can be used to predict oil production and displacement in real conditions.

If look at Table 1, the machine learning classifier results are respectively 96 %, 27 %, 99 %.

Fig. 5 shows the results of the metrics for each class, the average result of each metric. In the SVM classifier for class 9, the metrics are 0. This means that only 8 classes can be taken during training. The total number of neural network parameters is 21658, and parameters of the recurrent neural network – 1642. The speed of training the model of the recurrent neural network showed a faster result than the rest of the models.

The ROC curves of each classifier and neural networks can be seen in Fig. 6, 8, 10, 12, 14. The lowest result of ROC curves is in the 4th class for the SVM classifier and for the recurrent neural network. The indicators are 0.6434 and 0.6737, respectively. Full Gradient Boosting confusion matrix and the number of values for each combination of matrices [17]. The indicators for true positive decisions (TP) are 11692 and indicators for true negative decisions (TN) are 3409. The classification errors for false positive (FP) and false negative (FN) decisions are 3 each. Support vector machines were also used to predict oil displacement, which was also used in the work, but the result is much less accurate than for gradient boosting. But in comparison with previous works, the accuracy of the classifier and neural networks is high. A feature of the proposed method is the identification of the method with high accuracy and the smallest errors, comparison of the results of machine learning classifiers and types of neural networks. The advantage of this work is the high forecasting accuracy, which is very important for use in forecasting problems.

In the future, it is planned to parallelize heterogeneous parallel computations in order to improve the performance and increase the speed of algorithms. To do this, the expected to solve problems with the training time that arise when working with a large number of training examples are considered.

## 7. Conclusions

1. The Buckley-Leverett model was studied to solve the problem of oil displacement. Oil displacement prediction was implemented using machine learning methods such as Decision Tree, SVM and Gradient Boosting. Prediction of the oil displacement problem was implemented using neural and recurrent networks. A numerical assessment of the quality of the algorithms was carried out in order to determine the best classifier in forecasting problems. A three-dimensional model containing metrics such as accuracy, recall, and F1-measure for all classes was built. The presented research work is aimed at the correct prediction of oil displacement. To achieve this goal, a dataset was created containing more than 67,000 data for Class 10.

2. The classification was carried out according to three classification algorithms. The average accuracy of the Decision Tree classifier was 96 %, the SVM algorithm showed an accuracy of 28 %, the Gradient Boosting algorithm showed an accuracy of 100 %. In addition, the quality of the classifier is evaluated by the speed of execution and the performance of the algorithm. As for the training time, the Decision Tree was faster than the support vector machine and Gradient Boosting.

3. To check the accuracy, cross-validation was performed, where the data were divided into five blocks. As for the speed of forecasting in the task, although the Decision Tree has won in the speed of learning, it is inferior in the speed of execution. Thus, the accuracy of forecasts in the Decision Tree and Gradient Boosting methods is approximately the same. However, SVM and Gradient Boosting performed better due to the speed of execution when working in real time. The conducted research allowed us to draw the following conclusions based on the estimates of the algorithms: the average accuracy for the DT algorithm was 0.96, for the SVM algorithm – 0.28 and for Gradient Boosting – 1.00. The average recall value for DT was 0.96, for the continuation of the SVM algorithm – 0.28, and for Gradient Boosting – 1.00. For most classes, these metrics showed good results.

4. The results of neural networks are almost the same as the results of classifiers. The accuracy of the neural network is 97.1 %, and the accuracy of the recurrent neural network is 93 %.

5. The complete Gradient Boosting confusion matrix shows 11692 for true positive decisions (TP) and 3409 for true negative decisions (TN). Classification errors for false positive (FP) and false negative (FN) decisions are 3 each. In the future, it is planned to improve the performance of these classifiers by parallelizing them using CUDA and FPGA technologies.

## Acknowledgments

The research is funded by a grant from the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan under the project No. AP09563558.

References

1. De Bézenac, E., Pajot, A., Gallinari, P. (2019). Deep learning for physical processes: incorporating prior scientific knowledge. Journal of Statistical Mechanics: Theory and Experiment, 2019 (12), 124009. doi: https://doi.org/10.1088/1742-5468/ab3195
2. Buckley, S. E., Leverett, M. C. (1942). Mechanism of Fluid Displacement in Sands. Transactions of the AIME, 146 (01), 107–116. doi: https://doi.org/10.2118/942107-g

3. Almajid, M. M., Abu-Al-Saud, M. O. (2022). Prediction of porous media fluid flow using physics informed neural networks. Journal of Petroleum Science and Engineering, 208, 109205. doi: https://doi.org/10.1016/j.petrol.2021.109205
4. Muradkhanli, L. (2018). Neural Networks for Prediction of Oil Production. IFAC-PapersOnLine, 51 (30), 415–417. doi: https://doi.org/10.1016/j.ifacol.2018.11.339
5. Al-Shabandar, R., Jaddoa, A., Liatsis, P., Hussain, A. J. (2021). A deep gated recurrent neural network for petroleum production forecasting. Machine Learning with Applications, 3, 100013. doi: https://doi.org/10.1016/j.mlwa.2020.100013
6. Vo Thanh, H., Sugai, Y., Sasaki, K. (2020). Application of artificial neural network for predicting the performance of $CO_2$ enhanced oil recovery and storage in residual oil zones. Scientific Reports, 10 (1). doi: https://doi.org/10.1038/s41598-020-73931-2
7. Hu, C., Wang, F., Ai, C., Wang, X., Wang, Y. (2021). An Improved New Convolutional Neural Network Method for Inverting the Pore Pressure in Oil Reservoir by Surface Vertical Deformation. Lithosphere, 2021. doi: https://doi.org/10.2113/2021/5597238
8. Nikravesh, M., Kovscek, A. R., Murer, A. S., Patzek, T. W. (1996). Neural Networks for Field-Wise Waterflood Management in Low Permeability, Fractured Oil Reservoirs. All Days. doi: https://doi.org/10.2118/35721-ms
9. Fraces, C. G., Papaioannou, A., Tchelepi, H. (2020). Physics informed deep learning for transport in porous media. Buckley Leverett problem. arXiv.org. Available at: https://arxiv.org/pdf/2001.05172.pdf
10. Gasmi, C. F., Tchelepi, H. (2021). Physics informed deep learning for flow and transport in porous media. arXiv.org. Available at: https://arxiv.org/pdf/2104.02629.pdf
11. Kovenya, V. M., Chirkov, D. V. (2013). Metody konechnyh raznostey i konechnyh obemov dlya resheniya zadach matematicheskoy fiziki. Novosibirsk: NGU, 16–27. Available at: http://www.ict.nsc.ru/matmod/files/textbooks/KovenyaChirkov.pdf
12. Daribayev, B., Akhmed-Zaki, D., Imankulov, T., Nurakhov, Y., Kenzhebek, Y. (2020). Using Machine Learning Methods for Oil Recovery Prediction. ECMOR XVII. doi: https://doi.org/10.3997/2214-4609.202035233
13. Charbuty, B., Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. Journal of Applied Science and Technology Trends, 2 (01), 20–28. doi: https://doi.org/10.38094/jastt20165
14. Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 408, 189–215. doi: https://doi.org/10.1016/j.neucom.2019.10.118
15. Babajide Mustapha, I., Saeed, F. (2016). Bioactive Molecule Prediction Using Extreme Gradient Boosting. Molecules, 21 (8), 983. doi: https://doi.org/10.3390/molecules21080983
16. Cordoni, F. (2020). A comparison of modern deep neural network architectures for energy spot price forecasting. Digital Finance, 2 (3-4), 189–210. doi: https://doi.org/10.1007/s42521-020-00022-2
17. Yacouby, R., Axman, D. (2020). Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models. Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems. doi: https://doi.org/10.18653/v1/2020.eval4nlp-1.9
18. Janssens, A. C. J. W., Martens, F. K. (2020). Reflection on modern methods: Revisiting the area under the ROC Curve. International Journal of Epidemiology, 49 (4), 1397–1403. doi: https://doi.org/10.1093/ije/dyz274
19. Gupta, N., Nigam, S. (2020). Crude Oil Price Prediction using Artificial Neural Network. Procedia Computer Science, 170, 642–647. doi: https://doi.org/10.1016/j.procs.2020.03.136
20. Li, W., Kiaghadi, A., Dawson, C. (2020). Exploring the best sequence LSTM modeling architecture for flood prediction. Neural Computing and Applications, 33 (11), 5571–5580. doi: https://doi.org/10.1007/s00521-020-05334-3