Prediction of Dead Oil Viscosity: Machine Learning vs. Classical Correlations

Fahimeh Hadavimoghaddam [1], **Mehdi Ostadhassan** [2,*], **Ehsan Heidaryan** [3], **Mohammad Ali Sadri** [4], **Inna Chapanova** [4], **Evgeny Popov** [4], **Alexey Cheremisin** [4] **and Saeed Rafieepour** [5]

[1] Department of Oil Field Development and Operation, Faculty of Oil and Gas Field Development, 119991 Moscow, Russia; fahimemoghadam@gmail.com

[2] Key Laboratory of Continental Shale Hydrocarbon Accumulation and Efficient Development, Ministry of Education, Northeast Petroleum University, Daqing 163318, China

[3] Department of Chemical Engineering, Engineering School, University of São Paulo (USP), Caixa Postal 61548, São Paulo 05424-970, Brazil; heidaryan@usp.br

[4] Skolkovo Institute of Science and Technology (Skoltech), 143026 Moscow, Russia; MohammadAli.Sadri@Skoltech.ru (M.A.S.); i.chapanova@skoltech.ru (I.C.); E.Popov@skoltech.ru (E.P.); A.Cheremisin@skoltech.ru (A.C.)

[5] McDougall School of Petroleum Engineering, University of Tulsa, Tulsa, OK 74110, USA; saeed-rafieepour@utulsa.edu

* Correspondence: mehdi.ostadhassan@nepu.edu.cn

Abstract: Dead oil viscosity is a critical parameter to solve numerous reservoir engineering problems and one of the most unreliable properties to predict with classical black oil correlations. Determination of dead oil viscosity by experiments is expensive and time-consuming, which means developing an accurate and quick prediction model is required. This paper implements six machine learning models: random forest (RF), lightgbm, XGBoost, multilayer perceptron (MLP) neural network, stochastic real-valued (SRV) and SuperLearner to predict dead oil viscosity. More than 2000 pressure–volume–temperature (PVT) data were used for developing and testing these models. A huge range of viscosity data were used, from light intermediate to heavy oil. In this study, we give insight into the performance of different functional forms that have been used in the literature to formulate dead oil viscosity. The results show that the functional form $f(\gamma_{API}, T)$, has the best performance, and additional correlating parameters might be unnecessary. Furthermore, SuperLearner outperformed other machine learning (ML) algorithms as well as common correlations that are based on the metric analysis. The SuperLearner model can potentially replace the empirical models for viscosity predictions on a wide range of viscosities (any oil type). Ultimately, the proposed model is capable of simulating the true physical trend of the dead oil viscosity with variations of oil API gravity, temperature and shear rate.

Keywords: viscosity; PVT properties; dead oil viscosity; machine learning; SuperLearner

1. Introduction

Reservoir pressure–volume–temperature (PVT) properties are some of the most important ones for petroleum engineers and essential for different aspects of reservoir calculations. The precision of other measurements in reservoir engineering also relies primarily on the correctness of PVT data (e.g., calculations for material balance, reserve estimation, well test analyses, advanced data analysis, nodal analysis for a surface network, surface separation and numerical reservoir simulations) [1]. PVT data are obtained in an optimum situation from representative fluid samples collected from wellhead, surface or wellbore [2]. PVT reports the results of PVT tests usually at reservoir pressure (P) and temperature (T), whereas daily assessments are typically conducted as part of the oil and gas field monitoring programs at other P and Ts. These include standard tank oil gravity calculation of API, dead oil viscosity at ambient or different temperatures, gas gravity and composition

among others. Due to the necessity of having PVT properties at reservoir P and T, the bubble point and the dew point pressure can be determined from the production data. Therefore, pressure vs. depth measurements (repeat formation tester, RFT) will provide an average downhole density for reservoir fluids [2]. For other production engineering measurements, general hydrocarbon PVT properties include saturation pressure (oil and dew point pressure bubble); gas–oil solution ratio; vaporized oil–gas ratio; oil and gas formation volumes; oil and gas density; oil and gas viscosity, one- or two-phase Z factor for gases; and oil and gas isothermal compressibility and dew point pressure [3–6]. Incorrect measurement of PVT properties can lead to serious errors in the results of the calculations.

In the literature [7], the effect of uncertainty in PVT data on test results has been addressed, for example, in material balance equations [8] and reserves and output estimates for more volatile fluids [9]. Hence, in the past decades, different models were developed for the estimation of the fluid reservoir properties. These previously published models are empirical [10–25], compositional [26–29] and neural-network-based [30,31]. Among all PVT properties for the regulation of fluid movement through the porous media, viscosity plays an important role that should be quantified precisely [32,33]. Therefore the viscosity of the crude oil must be accurately estimated at different stages of oil exploration and production. This parameter is crucial for the assessment of porous fluid flow, well efficiency, well simulation, well testing and flow in pipes for the production and design of pipeline, transport equipment and the design of operations and developments of field [34]. The viscosity of crude oil is dependent on the specific physical and thermodynamic properties of the form and the characteristics of its chemical composition, density, bubble point pressure, temperature, oil gravity, gas and gas solubility [35,36] as well as the fluid molecules form and thickness [37]. In addition to the fluid density and fluid type (Newtonian or non-Newtonian) [2], fluid viscosity is a function of pressure (P) and temperature (T). A rise in temperature often induces a decrease in viscosity and a decrease in saturated pressure due to the release of dissolved gas below the bubble point to raise the viscosity [37]. Hence, the crude oil viscosity can be divided into three categories, namely saturated oil viscosity, undersaturated oil viscosity, and dead oil viscosity, depending on the reservoir pressure [38,39].

(1) Dead oil viscosity, μ_{od}, which is the crude oil that at atmospheric pressure is free of gas.
(2) Saturated, μ_{ob}, which is the oil viscosity at reservoir temperature and pressure (saturation).
(3) Undersaturated, μ_{oa}, is the viscosity of oil when its pressure and temperature is above the reservoir conditions (saturation).

Traditionally, crude oil viscosity is determined experimentally at the reservoir temperature and pressure on the subsurface or surface samples, but this is often costly and time-consuming and requires a strong technical specialty [40–45]. In this regard, a large number of empirical and semi-empirical relationships have been developed in the past decades, mainly from the corresponding equation of state to predict the crude oil viscosity. Most of the correlations presented have been established for a given region, so if used for other areas, erroneous results will be produced [34]. Empirical correlations are used to estimate dead oil, and saturated and undersaturated viscosities based on field data, but the output of these empirical correlations is typically unsatisfactory, and improvement is still sought [46]. The best-known ones are those developed by Beal 1946 [47], Beggs and Robinson (1975) [48], Glaso (1980) [13], Kaye (1985) [49], Al-Khafaji et al. (1987) [50], Petrosky (1990) [24], Egbogah and Ng (1990) [51], Labedi (1992) [52], Kartoatmodjo and Schmidt (1994) [53], De Ghetto (1994) [54], Bennison (1998) [55], Elsharkawy and Alikhan (1999) [56], Hossain et al. (2005) [57], Naseri et al. (2005) [58] and Alomair et al. (2011) [59], Hemmati et al. (2013) [38]. Correlations summarized in Table 1 relate dead oil viscosity. Some authors correlate crude oil viscosity to typically difficult-to-measure properties such as molar mass, critical temperature and acentric factor [60–62]. Furthermore, lack of availability of reservoir fluid samples may add another obstacle to reliable measurements; thus, reservoir engineers are encouraged to use established correlations to estimate crude oil viscosity. Most of the published correlations have been developed based on limited data

points and ranges [39]. The evaluation of most developed correlations in the literature have shown that high errors are generated when applied on other data sets rather than the one the correlation is generated for [63–65].

Table 1. The summary of the ranges and origins data used in some famous dead oil viscosity correlations.

Authors	Region of Data Source	Input	Data Points	T, °C	API	μ_{od}, cp
Beals 1946 [47]	US	γ_{API}, T	98	36.6–121.1	10–52	0.86–1550
Beggs and Robinson (1975) [48]	-	γ_{API}, T	460	21.1–146.1	16–58	-
Glaso (1980) [13]	North Sea	γ_{API}, T	38	10–148.8	20–48	0.60–39
Kaye (1985) [49]	Ofshore California	γ_{API}, T	-	61.6–138.8	7–41	-
Al-Khafaji et al. (1987) [50]	-	γ_{API}, T	-	15.5–148.8	15–51	-
Petrosky (1990) [24]	Gulf of Mexico	γ_{API}, T	118	45.5–142.2	25–46	0.72–10.25
Egbogah and Ng (1990) [51]	-	γ_{API}, T	394	15–80	5–58	-
Labedi (1992) [52]	Libiya	γ_{API}, T	91	37.7–152.2	32–48	0.66–4.79
Kartoamtmodjo and Schmidt (1994) [53]	Worldwide	γ_{API}, T	661	26.6–160	14–59	0.50–586
Bennison (1998) [54]	North sea	γ_{API}, T	16	3.8–148.8	11–20	6.40–8396
Elsharkawy and Alikhan (1999) [55]	Middle East	γ_{API}, T	254	37.7–148.8	20–48	0.60–33.7
Dindoruk and Christman (2004) [66]	Gulf of Mexico	$\gamma_{API}, T, P_b, R_s$	95	17.4–40	17.4–40	
Hossain et al. (2005) [56]	World wide	γ_{API}, T	184	0–101.6	7–22	12–451
Naseri et al. (2005) [57]	Iran	γ_{API}, T	472	40.5–147.7	17–44	0.75–54
Bergman and Sutton (2009) [67]	Worldwide	γ_{API}, T	9837			1.78–11360
Alomair et al. (2011) [58]	Kuwait	γ_{API}, T	374/118	20–160	10–20	0.39–70
Hemmati et al. (2013) [38]	Iran	γ_{API}, T	1000	10–143.3	17–44	-
El- hoshoudy et al. (2013) [34]	Egypt	γ_{API}, T	1000	-	-	-

As a result, there is a need to propose methods that can estimate dead oil viscosity at wider ranges of temperature intervals when a physical sample to perform laboratory measurements is missing or there are time constraints and rapid results are needed even though fluid samples are available. Ultimately, we need to develop a method that would be cheaper compared to laboratory studies when both situations are not an issue. Regarding what was said, this has graveled the way for modification and adoption of already existing empirical correlations over a period of time. Furthermore, to overcome these challenges, some machine learning (ML) and artificial intelligence techniques (AI) have also been used to improve the prediction of oil viscosity, including radial basis function neural network (RBFNN) [30], artificial neural network (ANN) [68–73], functional networks (FN) [46,71], genetic algorithm (GA) [74], support vector machine (SVM) [75], and group method of data handling (GMDH) [76] and Ensemble models [77]. The literature argues that the lowest average absolute relative error can be achieved when viscosity is predicted by AI models and the highest correlation coefficient as compared to existing empirical correlations [70].

Based on what was said above, the industry has an interest to develop models that can predict viscosity versus shear rate and temperature of a wide range of oils for both Newtonian and non-Newtonian ones. Considering these points, the main goal of this study is to develop a model to predict dead oil viscosity vs. temperature through recently established ML methods that make us independent of creating new correlations and can be applicable to all type of oils and regions.

2. Materials and Methods

2.1. Experiments

Measurements of viscosity of the samples were carried out on the Anton Paar Rheometer MCR 302, located at the Skolkovo Institute of Science and Technology at the Center for Hydrocarbon Recovery. Based on the manufacturer's recommendations and the nature of the substance, viscosity was measured in various systems: coaxial cylinders, cone-plate according to ASTM D 4402-06 (ASTM D4402-06, Standard Test Method for Viscosity De-

termination of Asphalt at Elevated Temperatures Using a Rotational Viscometer, ASTM International, West Conshohocken, PA, U.S., 2006). A large number of oil samples were tested in the C-ETD 300 system equipped with CC-17 coaxial cylinders for measurements of low viscous specimens. For high viscous samples, viscosity values were obtained in the P-ETD 400 system with cone-plate CP50-2 measuring system.

The elemental composition of oil samples was determined using a LECO Corporation CHN628 with an attachment for determination of the sulfur. The optional sulfur module for the 628 Series analyzers is designed to determine the sulfur content in a wide range of organic matrices. It allows expansion of the analytical capabilities of the devices CHN628, CN628 and FP628. The module can be used to work with samples of coal, coke, various types of fuels and some inorganic materials, for example, samples of soil, cement and limestone. A pre-weighed sample is placed in a ceramic boat and then in a horizontal furnace, the sample is burned in an oxygen stream, and the sulfur contained in it is converted to SO_2. After that, due to the concentric design of the furnace tubes, the gases pass through the hot zone twice, which allows them to stay as long as possible in the high temperature zone for their complete oxidation. In the next stage, out of the furnace, gases pass through the anhydrone to remove moisture and the flow controller, where it is stabilized. The sulfur content is determined in the measuring infrared cell SO_2 of the analyzer of the 628 series, where the sulfur module is connected. Figure 1A displays dead oil viscosity from 10 representative samples as a function of temperature. As seen from the image, as the temperature increases, the viscosity of all samples decreases exponentially. A typical behavior of one oil sample in various temperatures vs. shear rate is also shown in Figure 1B, where *y*-axis is logarithmic for better data presentation. Density measurements were performed on an Anton Paar DMA TM 4200M density meter in the temperature ranges 10–90 and atmospheric pressure. Measurements were performed in accordance with ASTM D4052 (ASTM D4052-09, Standard Test Method for Density, Relative Density, and API Gravity of Liquids by Digital Density Meter, ASTM International, West Conshohocken, PA, U.S., 2009). A typical behavior of an oil sample from those 10 representative ones in various temperatures is shown in Figure 1C below.
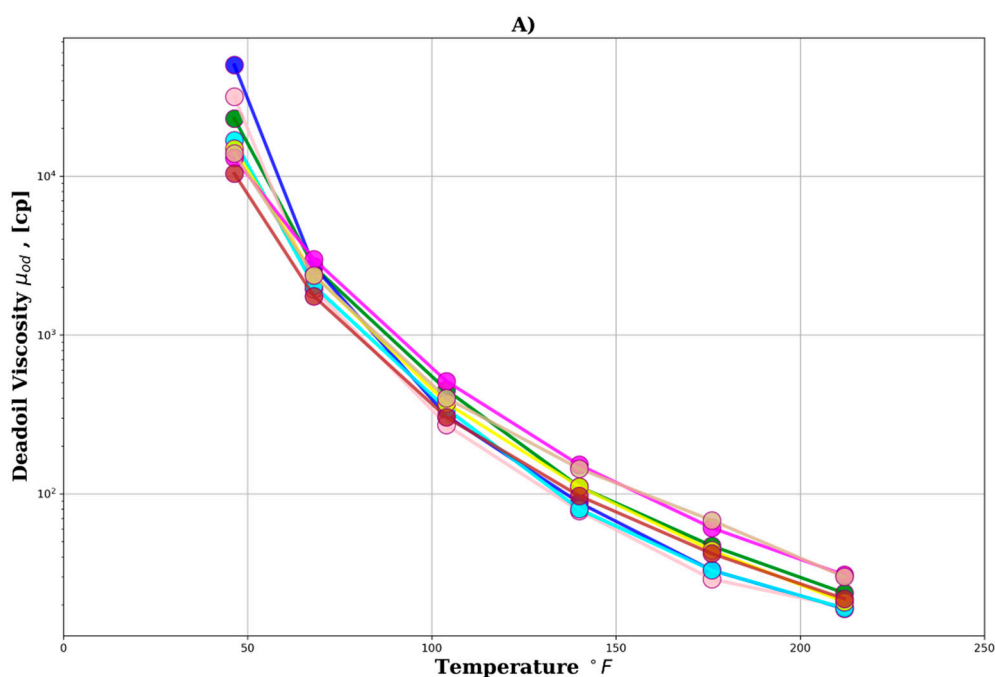

Figure 1. *Cont.*
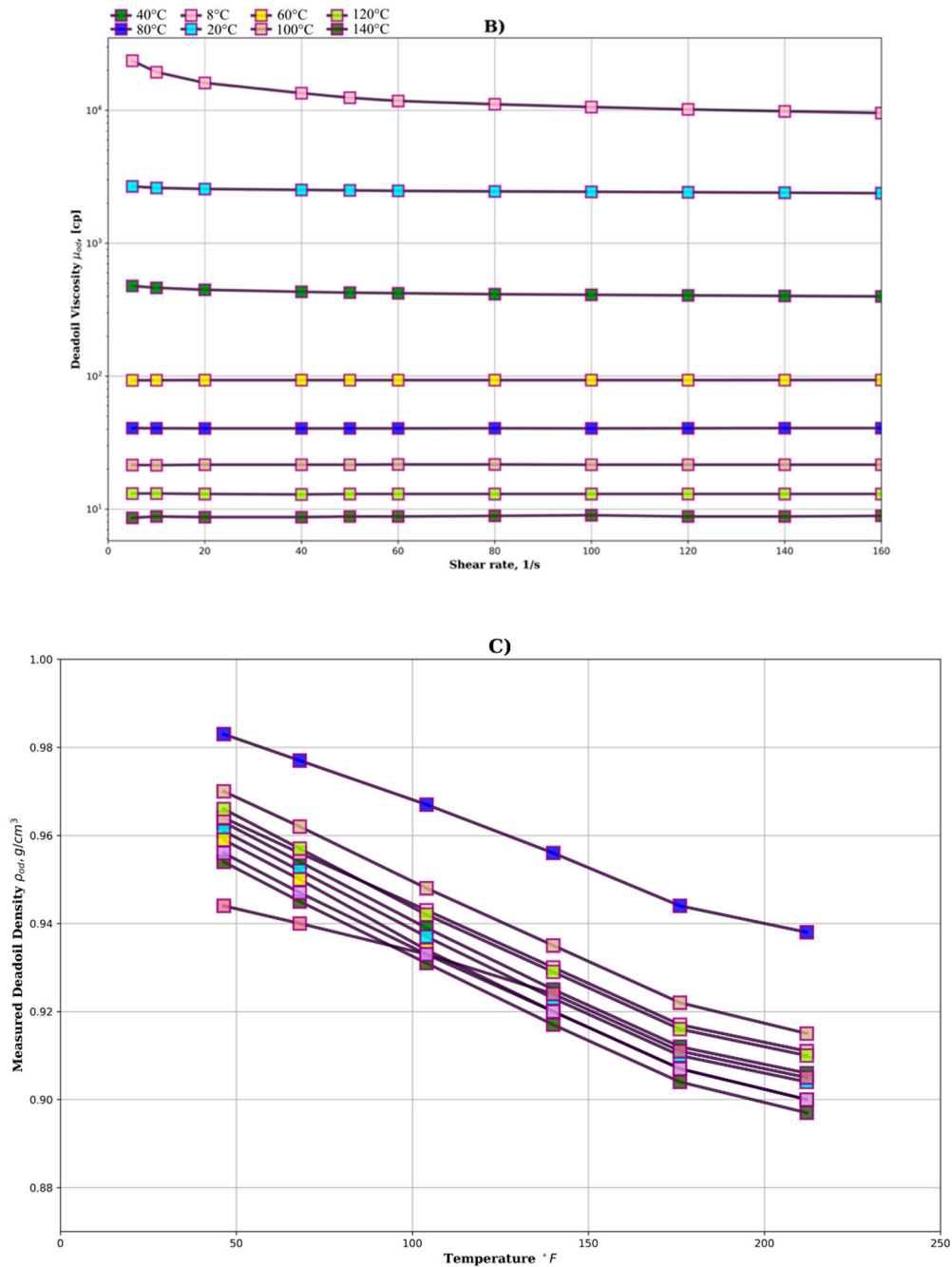
Figure 1. (**A**) Experimental results of 10 representative samples' viscosity vs. temperature, (**B**) experimental results of one representative sample dead oil viscosity (μ_{od}), *cp* in logarithmic format vs. shear rate (γ), 1/*c* and (**C**) experimental results of one sample dead oil density (ρ_{od}) vs. temperature °C.

2.2. Input Features and Data

A total of 2247 data points were collected from the literature [27,28,39,52,66,67,71,77–81] and added to the authors dataset for the viscosity simulations. A statistical summary of the data is presented in Table 2.

Table 2. Dataset statistical summary.

	API	Tr (F)	Rs (scf/STB)	Pb	μ_{od} **cp**
min	6.0	25.6	8.6	107.3	0.5
max	56.8	341.6	3298.7	6613.8	15,836.9
Average	29.73	201.08	605.16	2324.99	1066.28

In this study, in addition to the common correlating properties, γ_{API}, T, other important variables have been incorporated as well in the formulations. It is important to note that not all of these approaches use hydrocarbon components; detailed information as seen below:

$$\mu_{od} = f(\gamma_{API}, T, P_b, R_s) \tag{1}$$

$$\mu_{od} = f(\gamma_{API}, T, P_b) \tag{2}$$

$$\mu_{od} = f(\gamma_{API}, T, R_s) \tag{3}$$

$$\mu_{od} = f(\gamma_{API}, T) \tag{4}$$

In the above equations, μ_{od} is the dead oil viscosity in cp, *API* is oil API gravity, *T* is temperature in *F*, R_s is gas oil ratio in scf/STB and P_b is bubble point pressure in psia. Based on the mentioned functional forms (Equations (1)–(4)), six different ML models are developed through Python programming that include: random forest (RF), lightgbm, XGBoost, MLP neural network, SRV and SuperLearner. The code programs that are written rely on the python package library described in Table 3 and can become available to readers upon request. For developing dead oil viscosity model estimation approximately 60% of the data was used for training model while the remaining was used for blind tests.

Table 3. Machine learning methods and their packages.

Algorithms	Python Package	Package Version	Website of Package	Date Accessed
XGBoost	xgboost	0.9.0	https://xgboost.readthedocs.io/en/latest/index.html	Janurary 2020
Lightgbm	lightgbm	2.3.2	https://lightgbm.readthedocs.io/en/	February 2020
SVR	scikit-learn	scikit-learn 0.23.1	https://scikit-learn.org/dev/index.html	February 2020
MLP neural network	pytorch	9.2	https://pytorch.org	March 2020
Random Forest	scikit-learn	0.22.2	https://scikit-learn.org/dev/index.html	March 2020
SuperLearner	SuperLearner	0.9.0	http://ml-ensemble.com	April 2020

2.3. ML Model Development

In recent decades, different machine learning models have been used for PVT properties' estimation.

Among these, the artificial neural network has caught significant attention since the late 1990s. When we employ ML algorithms, one should keep in mind that theoretically, since dead oil viscosity of each type of oils will differ, various ML models should be tested, and a single algorithm will not suffice [77]. This means that one ML model might work properly and perform well on one dataset while others fail, and it becomes erroneous to exhibit their inferiority. For example, Van der Laan [78] developed SuperLearner where ensembles are utilized to stack base learners to reduce errors in forecasts. In this aspect, SuperLearner is superior compared to the base learner since systematic errors of base learners are found and delineated on the final prediction [79] to make it applicable in a variant of fields: biology domain including medicine, healthcare [79], biostatistics and genetics [80–84], and epidemiology [85]. It is important to note that this does not mean that each base learner and meta learner in a SuperLearner algorithm is independent of training for a precise prediction. In this study, the following five machine learning algorithms are implemented for dead oil viscosity prediction and results are compared across the board: XGBoost [86], LightGBM [87], random forest [88], an artificial neural network algorithm [89] and SVR [90]. Likewise, due to the superiority and robustness of SuperLearner that is

proven in other fields [78,80,91–93], for the first time, this method is also applied, and the output is compared to five other algorithms.

In order to estimate and accurately model, 6-fold cross-validation on all two sets of input was performed. For evaluation of the advanced machine learning algorithms for each data set we used cross-plots in order to evaluate measured and predicted values of μ_{od}. The robustness and accuracy of the models in this study have been evaluated using different statistical quality measures [94,95], including the coefficient of determination, mean squared error (MSE), mean absolute error (MAE), the percentage of accuracy–precision (PAP) [96] and root mean squared error (RMSE) through the following equations:

Root mean squared error (RMSE) metric is given by Equation (5):

$$
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\text{Actual}_i - \text{Predicted}_i)^2}{N}} \tag{5}
$$

where, N is total number of observation

Mean squared error (MSE) metric is given by Equation (6):

$$
\text{MSE} = \frac{\sum_{i=1}^{N}(\text{Actual}_i - \text{Predicted}_i)^2}{N} \tag{6}
$$

Mean absolute error (MAE) metric is given by Equation (7):

$$
\text{MAE} = \frac{\sum_{i=1}^{N}|\text{Actual}_i - \text{Predicted}_i|}{N} \tag{7}
$$

Coefficient of determination (R^2) metric is given by Equation (8):

$$
R^2 = 1 - \frac{(\text{Actual}_i - \text{Predicted}_i)^2}{(\text{Actual}_i - \text{mean of the observed data})^2} \tag{8}
$$

And finally, the percentage of accuracy–precision (PAP) is given by Equation (9):

$$
\text{PAP} = 100\left(1 - \frac{\sqrt{2}}{2}\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\text{Actual}_i - \text{Predicted}_i}{\text{Actual}_i}\right|\right)^2 + \left(\sum_{i=1}^{n}(\text{Actual}_i - \text{Predicted}_i)^2 / \sum_{i=1}^{n}(\text{Actual}_i - \text{Predicted}_i)^2\right)^2}\right) \tag{9}
$$

3. Results and Discussion

We should recall that most of the correlation methods (Table 4) are based on data from different origins, and this is why some correlations cannot give good accuracy. The main aim of this study was to develop advanced ML methods to predict dead oil viscosity in interested temperature interval $T = [40\,°F\text{–}233.6\,°F]$ for any oil type. We developed correlations that can predict the viscosity of any type of heavy as well as lighter oils at any temperature in the mentioned interval using the various functional forms. Comparing the performance of advanced ML algorithms with common empirical correlations from the literature, the results of all simulated by the advanced machine learning algorithms were better than the results of all the experiential relationships in Table 4. Figures 2 and 3 give a graphical comparison of the machine learning algorithms for μ_{od} prediction. Figure 2 shows the cross-validation results based on the testing dataset respectively. Based on the mentioned functional forms of μ_{od} (Equations (1)–(4)), different models for these considered machine learning algorithms were developed. Figures 2 and 3 exhibit the results of these approaches where experimental data on about 200 samples is plotted vs. predictions. It is important to note that the accuracy of our methods is determined based on the metrics (Equations (5)–(8)) by plotting experimental data vs. prediction results. Here, it is assumed that experimental data is the most reliable one and closeness of predicted values to that outcome can be the measure of algorithm/correlation performance. Comparing the metric and statistical error analysis results, it can be seen that SuperLearner might be

a promising tool in viscosity prediction. Table 4, below, shows the performance of some common empirical correlations that are applied to our dataset. Here, it is seen that the best results among these empirical correlations are given by Naseri [57] and the next one is the correlation by Beal [47]. In the correlation developed by Dindoruk [66] as it was mentioned in Table 4, although there are other additional parameters in the function such as bubble point pressure (P_b) and solution gas oil ratio (R_{so}), it has not improved the results compared to other methods. In the correlation by Dindoruk [66], the functional form for estimation of dead oil viscosity is $f(\mu_{od}) = \gamma_{API}, T, R_{sb}, P_b$ and we see that having additional parameters has not improved the results of the prediction. All other correlations have the same functional form $f(\mu_{od}) = \gamma_{API}, T$.

Table 4. Performance of μ_{od} prediction.

Methods	RMSE	MAE	R^2	MSE	PAP
SuperLearner	0.3387	0.1154	0.9568	0.1147	87.3139
Lightgbm	0.3384	0.1147	0.9541	0.1145	85.7713
Random forest	0.3387	0.1158	0.9476	0.1147	82.6887
XGBoost	0.3372	0.1142	0.9465	0.1137	87.8166
MLP ANN	0.4794	0.1151	0.9329	0.2298	84.7704
SVR	0.4998	0.1068	0.9217	0.2498	86.3304
Sattarian [36]	0.3380	11.5140	0.9012	0.1142	83.6400
Naseri [57]	0.8864	57.3010	0.8940	0.7857	80.7704
Dindrouk [66]	0.4533	21.4530	0.8911	0.2055	77.7661
Hemmati [39]	0.4477	26.5440	0.8871	0.2004	77.8325
Ghorbani [44]	0.2480	10.8330	0.8435	0.0615	75.2955
Hossain [56]	0.4354	15.4320	0.8071	0.1896	76.5429
Lashkerani [43]	0.2760	13.6450	0.8057	0.0762	75.8557
Beal [47]	0.6609	15.6740	0.7600	0.4368	70.4494
Elsharkway [55]	0.7928	20.7810	0.7198	0.6285	70.3990
Ubong and Oyedeko [97]	0.6910	15.9180	0.7198	0.4775	71.3572
Glaso [13]	0.8756	27.1860	0.6630	0.7667	65.1807
Chew and Connally [98]	0.7829	28.5010	0.6340	0.6129	65.6894
Al Khafaiji [50]	0.8740	30.0140	0.6334	0.7639	65.2024
Petrosky [24]	0.8973	19.7160	0.6315	0.8051	67.8336
Khan [99]	1.1024	33.0050	0.6120	1.2153	69.2030
Beggs [48]	1.7703	55.6430	0.5543	3.1340	65.1991
Labedi [52]	1.8803	19.5610	0.5458	3.5355	65.3186

a) $\mu_{od} = f(\gamma_{API}, T)$

b) $\mu_{od} = f(\gamma_{API}, T, R_s)$

c) $\mu_{od} = f(\gamma_{API}, T, P_b)$

d) $\mu_{od} = f(\gamma_{API}, T, P_b, R_s)$
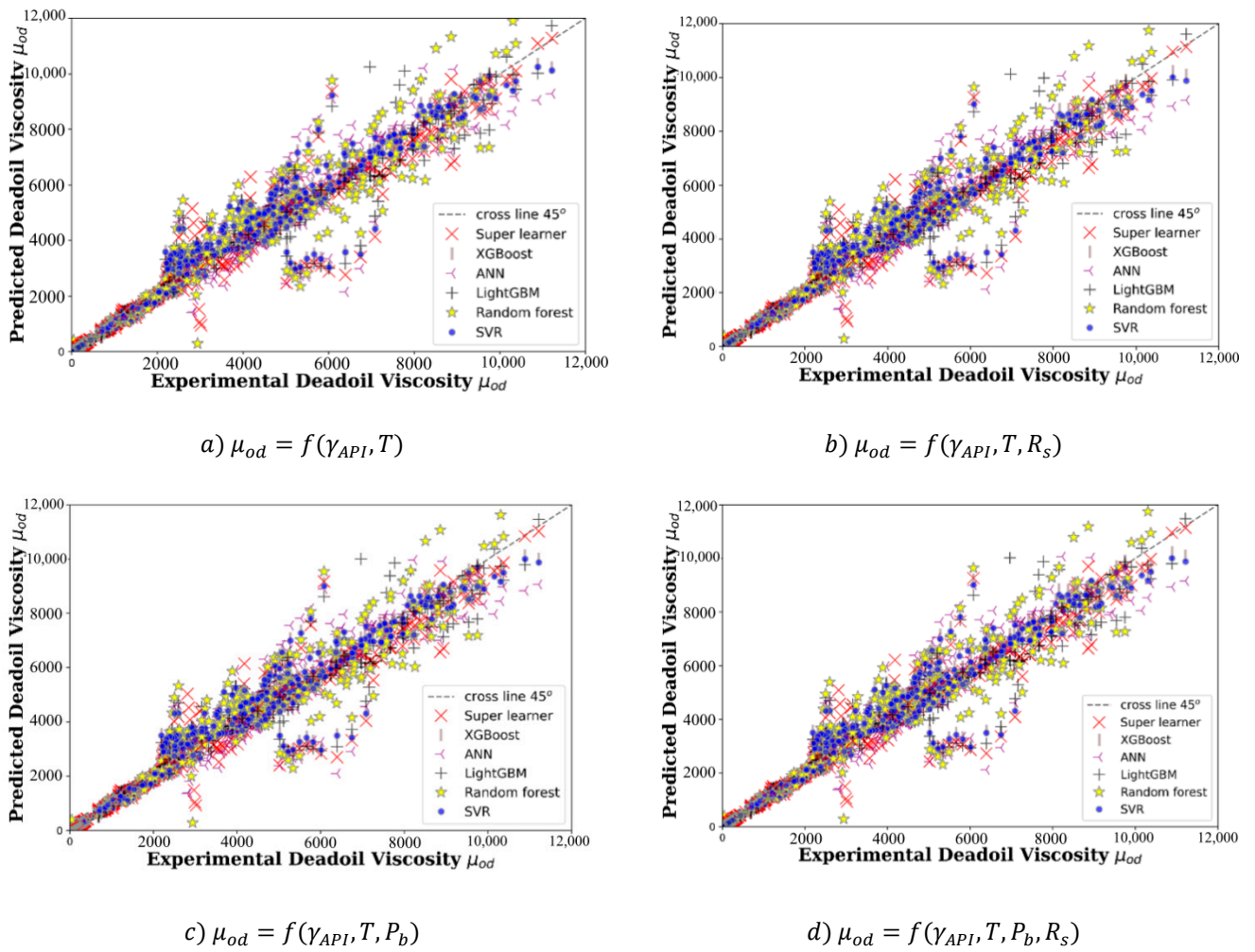
Figure 2. Experimental data vs. predictions on testing datasets for four different forms of μ_{od}.

a) $\mu_{od} = f(\gamma_{API}, T)$

b) $\mu_{od} = f(\gamma_{API}, T, R_s)$

Figure 3. *Cont.*

c) $\mu_{od} = f(\gamma_{API}, T, P_b)$

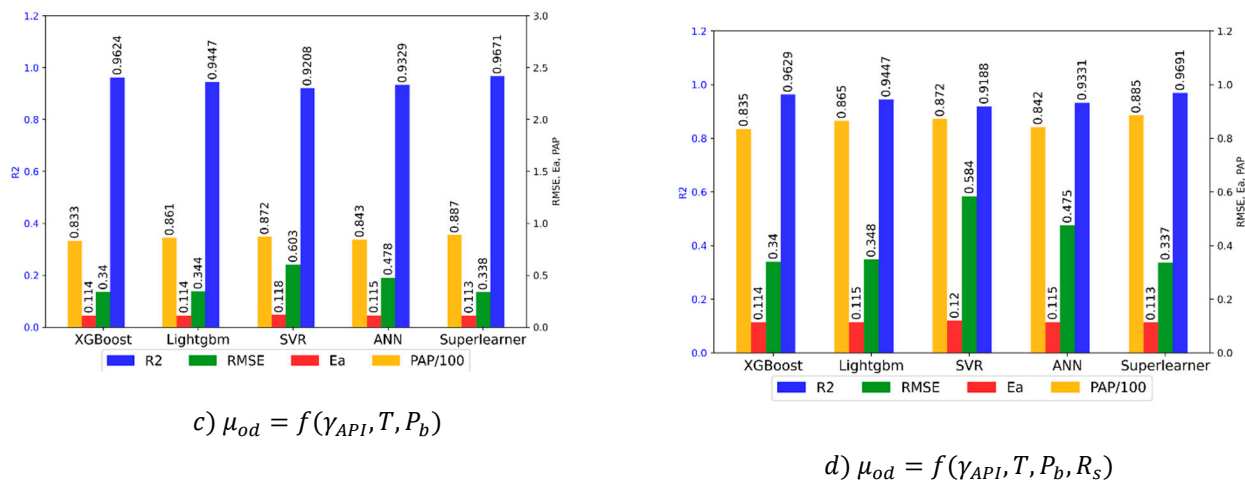d) $\mu_{od} = f(\gamma_{API}, T, P_b, R_s)$
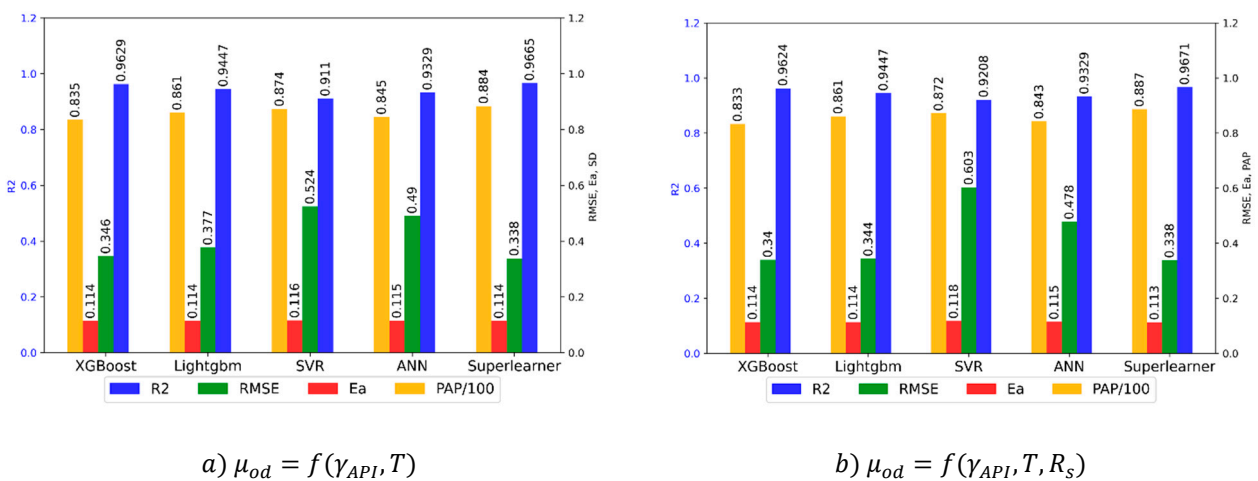
Figure 3. Statistical results (metric analysis) of different machine learning (ML) algorithms for μ_{od} of different functional for, prediction for the testing dataset.

In order to validate the objective of this study and evaluate the performance of developed models, ML results were compared based on various metrics parameters to the results obtained from the existing classical correlations. The results of Table 4 show that the developed SuperLearner outperforms all pre-existing models. It is evident that the proposed models in this study are more robust and reliable and accurate than other published correlations in terms of statistical parameters.

Considering the results from all functional forms of the viscosity (*f* (μ_{od}) = γ_{API}, *T*), the SuperLearner and random forest (RF), lightgbm, XGBoost, MLP neural network, SRV and SuperLearner exhibited very satisfactory patterns and behavior, which provided precise results and match with experimental data. This becomes more important when we consider that the majority of the empirical associations were formed based on these two variables only as the following functional form *f* (μ_{od}) = γ_{API}, *T*. In this regard, the SuperLearner has the least root mean square errors (RMSE) for all methods according to Figure 3. The best overall strategy lies among the determination coefficient's (R^2) SuperLearner. With the aid of this model, accurate dead oil viscosity can be calculated as one of the critical inputs for calculations and for the generation of live oil viscosity for fluids under different wellbore conditions.

In order to check the feasibility of the proposed SuperLearner model to see whether, with variance of input parameters, the model will catch the physically predicted patterns, we plotted the results of predicted value with classical correlations and SuperLearner, and compared them with experimental values.

Figure 4 demonstrated the general effect of temperature on dead oil viscosity and it can be seen that these effects are correctly predicted by the SuperLearner. As we see in Figure 4, dead oil viscosity decreases when temperature is increased. At lower temperatures the effect is more noticeable. The results provided in Figure 4 provided annotated lists of the most common dead oil viscosity correlations. The results show that the Bennison method overpredicted the value of dead oil viscosity in temperatures lower than 120 °F, while the Petrosky method underpredicted the value of dead oil viscosity in all temperature ranges. The developed SuperLearner method could predict dead oil viscosity variance between predicted value and experimental value.
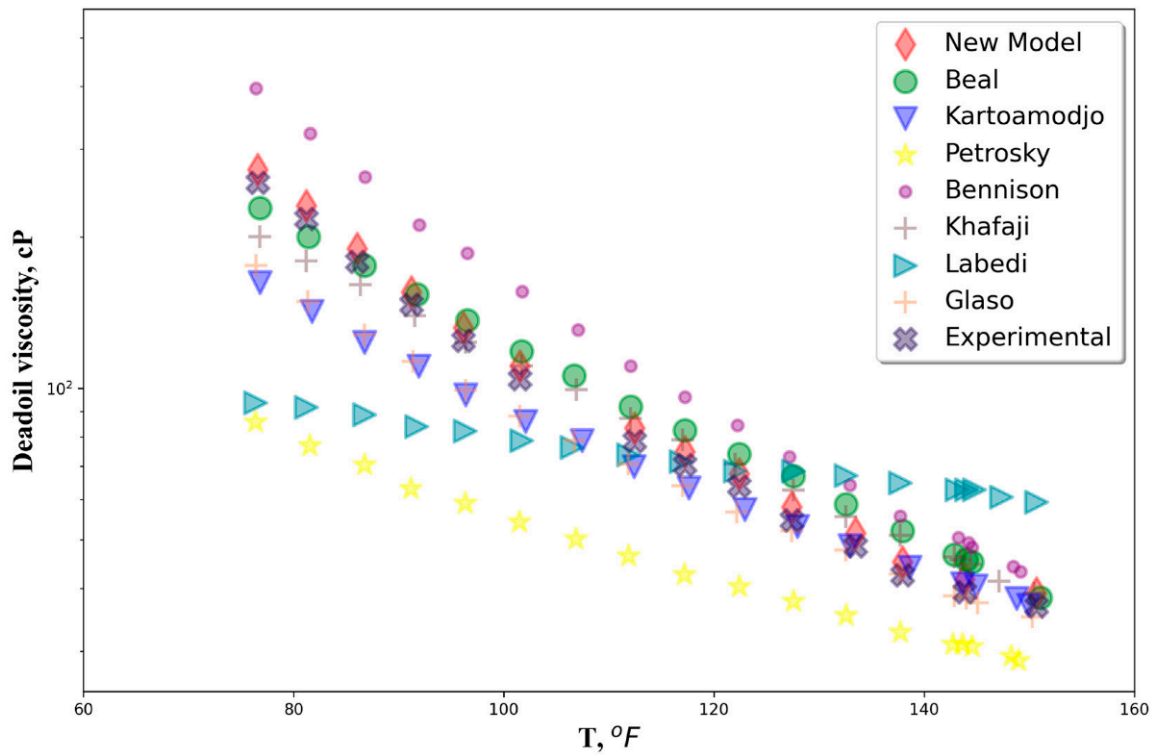
Figure 4. Dead oil viscosity from most leading correlations at various temperatures.

Figure 5 displays the proposed SuperLearner model in this work that has the same trend as other famous correlations to predict the API vs. viscosity. Figure 5 shows that dead oil viscosity decreases when oil API gravity increases and shows the trend of the developed model and some other correlation at 120 °F.
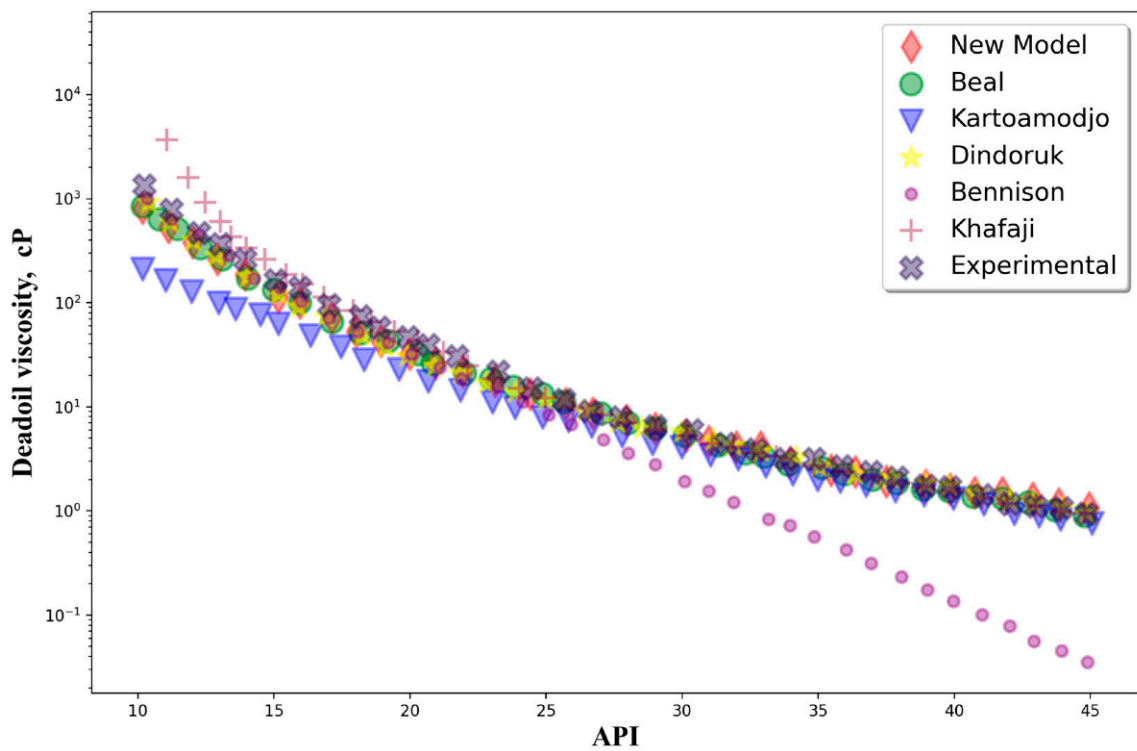

Figure 5. Dead oil viscosity from various correlations at various API at 120 °F.

From this figure, it is seen that results predicted by the SuperLearner are matching perfectly with the experimental data.

The results provided at Figure 5 show that the method proposed by the authors of this study are suited for all API ranges and the developed SuperLearner model can predict dead oil viscosity with the highest accuracy ($R^2 = 0.96$). The Al-Kafaji method is unfit for crude oil gravities below 15° API, while the method developed by Bennison, which was based on North Sea crude oil's low API gravity, is not appropriate for gravities greater than 30° API.

The most common method for getting dead oil viscosity data is viscosity correlation, which is very useful and effective in estimation of dead oil viscosity at different temperatures for different oil types. The critical point in the application of common correlations is their limitation on the parameters from which these correlations have derived. This means these correlations are regionally-dependent and cannot be applied universally. Ultimately, in this study we developed a model based on the newly introduced ML SuperLearner algorithm to estimate viscosity regardless of temperature, oil type and other critical parameters that are hard to obtain experimentally. As it was said in the literature, measuring heavy oil viscosity can be challenging at low temperatures because many times the expected values of the viscosity can exceed the upper limit of the device.

The developed model is applicable for a wide range of dead oil viscosity and variety of oils and temperatures. One main advantage of the proposed model is that there is no need for compositional analysis of the oil or asphalt content for the proposed dead oil viscosity correlation. The developed model has better predictive potential than the leading correlation in the whole considered range of viscosity. The proposed SuperLearner model can be used as a fast tool to verify the quality of the experimental data and/or the validity and accuracy of different viscosity models, especially where there are discrepancies and uncertainties in the datasets. The proposed SuperLearner model, which has improved crude oil viscosity accuracy and efficiency compared to previously published ones, can be applied in any reservoir simulator program.

4. Summary and Conclusions

This article demonstrates the idea of application of several machine learning algorithms, such as random forest (RF), lightgbm, XGBoost, MLP neural network, SRV and SuperLearner to predict dead oil viscosity based on different functional forms. The study revealed that XGBoost and SuperLearner might be promising tools in dead oil viscosity prediction from these methods and can be applied to reduce the cost of laboratory measurements. It was found that since SuperLearner integrates the merits of base machine learning algorithms, its capability in forecasting a more accurate dead oil viscosity is improved significantly. In addition, it has a potential to increase the accuracy of viscosity modeling of crude oil where some data are not available.

The main advantage of using machine learning and intelligence methods in the estimation of dead oil viscosity is that with knowing examples of previous patterns they can be easily trained and put to effective solving of unknown or untrained instances of the problem as was the case here. The results confirm the performance of the proposed SuperLearner model in estimation of dead oil viscosity. The results show that the simplest functional form to obtain viscosity would be adequate to provide statistically valid and acceptable results. This denotes that additional correlation variables might not be necessary to improve the performance of a model. From the empirical correlations, it can be noticed that the errors are higher than the ensemble models compared to the ML techniques, which are XGBoost, lightgbm, MLP ANN and SVR. The developed SuperLearner model in this study can be used when the dead oil viscosity should be predicted with limited input data when it is hard to perform an experiment with high accuracy, or when there is no physical sample to do additional experiments, especially for enhanced oil recovery (EOR) processes in the industry where they need viscosity at elevated temperatures. Ultimately,

the proposed SuperLearner model can be employed in other static and dynamic reservoir modeling tasks.

Data Availability Statement: Data can be provided upon request and direct email to authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khoukhi, A.; Albukhitan, S. PVT properties prediction using hybrid genetic-neuro-fuzzy systems. *Int. J. Oil Gas Coal Technol.* **2011**, *4*, 47–63. [CrossRef]
2. McCain, W.D.; Spivey, J.P.; Lenn, C.P. *Petroleum Reservoir Fluid Property Correlations*; PennWell Books: Tusla, OK, USA, 2011.
3. Ahmed, T. *Equations of State and PVT Analysis*; Gulf Publishin Company: Houston, TX, USA, 2013.
4. Arabloo, M.; Amooie, M.-A.; Hemmati-Sarapardeh, A.; Ghazanfari, M.-H.; Mohammadi, A.H. Application of constrained multi-variable search methods for prediction of PVT properties of crude oil systems. *Fluid Phase Equilib.* **2014**, *363*, 121–130. [CrossRef]
5. Ikiensikimama, S.S.; Ajienka, J.A. Impact of PVT correlations development on hydrocarbon accounting: The case of the niger delta. *J. Pet. Sci. Eng.* **2012**, *81*, 80–85. [CrossRef]
6. Rafiee-Taghanaki, S.; Arabloo, M.; Chamkalani, A.; Amani, M.; Zargari, M.H.; Adelzadeh, M.R. Implementation of svm framework to estimate PVT properties of reservoir oil. *Fluid Phase Equilib.* **2013**, *346*, 25–32. [CrossRef]
7. Spivey, J.P.; Pursell, D.A. Errors in Input Data and the Effect on Well-Test Interpretation Results. In Proceedings of the SPE Permian Basin Oil and Gas Recovery Conference, Midland, TX, USA, 23–26 March 1998.
8. Ambastha, A.K.; van Kruysdijk, J.W. Effects of Input Data Errors on Material Balance Analysis for Volumetric, Gas and Gas-Condensate Reservoirs. In Proceedings of the PETSOC the Annual Technical Meeting, Calgary, AB, Canada, 8–11 May 1993.
9. Baker, R.O.; Regier, C.; Sinclair, R. PVT Error Analysis for Material Balance Calculations. In Proceedings of the Canadian International Petroleum Conference, Calgary, AB, Canada, 10–12 June 2003.
10. Standing, M.B. A Pressure-Volume-Temperature Correlation for Mixtures of California Oils and Gases. In Proceedings of the Drilling and Production Practice, New York, NY, USA, 1 January 1947.
11. Lasater, J.A. Bubble point pressure correlation. *J. Pet. Technol.* **1958**, *10*, 65–67. [CrossRef]
12. Vazquez, M.; Beggs, H. Correlations for fluid physical property prediction. *J. Pet. Technol.* **1980**, *32*, 968–970. [CrossRef]
13. Glaso, O. Generalized pressure-volume-temperature correlations. *J. Pet. Technol.* **1980**, *32*, 785–795. [CrossRef]
14. Pedersen, K.S.; Fredenslund, A. An improved corresponding states model for the prediction of oil and gas viscosities and thermal conductivities. *Chem. Eng. Sci.* **1987**, *42*, 182–186. [CrossRef]
15. Al-Marhoun, M.A. PVT correlations for Middle East crude oils. *J. Pet. Technol.* **1988**, *40*, 650–666. [CrossRef]
16. Al-Marhoun, M.A. New correlations for formation volume factors of oil and gas mixtures. *J. Can. Pet. Technol.* **1992**, *31*. [CrossRef]
17. Abdul-Majeed, G.H.; Salman, N.H. An empirical correlation for oil FVF prediction. *J. Can. Pet. Technol.* **1988**, *27*. [CrossRef]
18. Asgarpour, S.; McLauchlin, L.; Wong, D.; Cheung, V. Pressure-volume-temperature correlations for western Canadian gases and oils. *J. Can. Pet. Technol.* **1989**, *28*. [CrossRef]
19. Labedi, R. Use of production data to estimate volume factor. density and compressibility of reservoir fluids. *J. Pet. Sci. Eng.* **1990**, *4*, 375–390. [CrossRef]
20. Dokla, M.; Osman, M. Correlation of PVT properties for UAE crudes (includes associated papers 26135 and 26316). *SPE Form. Eval.* **1992**, *7*, 41–46. [CrossRef]
21. Frashad, F.; LeBlanc, J.L.; Garber, J.D.; Osorio, J.G. Empirical PVT Correlations for Colombian Crude Oils. In Proceedings of the SPE Latin America/Caribbean Petroleum Engineering Conference, Port-of-Spain, Trinidad, 23–26 April 1996.
22. Elsharkawy, A.M.; Elgibaly, A.A.; Alikhan, A.A. Assessment of the PVT correlations for predicting the properties of Kuwaiti crude oils. *J. Pet. Sci. Eng.* **1995**, *13*, 219–232. [CrossRef]
23. Almehaideb, R.A. Improved PVT Correlations for UAE Crude Oils. In Proceedings of the Middle East Oil Show and Conference, Manama, Bahrain, 15–18 March 1997.
24. Petrosky, G.E. PVT Correlations for Gulf of Mexico Crude Oils. Ph.D. Thesis, University of Southwestern Louisiana, Lafayette, LA, USA, 1990.
25. Özdoğan, S.; Yücel, H.G. Correlations towards prediction of petroleum fraction viscosities: An empirical approach. *Fuel* **2001**, *80*, 447–449. [CrossRef]

26. Elsharkwy, A.; Gharbi, R. Comparing classical and neural regression techniques in modeling crude oil viscosity. *Adv. Eng. Softw.* **2001**, *32*, 215–224. [CrossRef]
27. Elsharkawy, A.M.; Hassan, S.A.; Hashim, Y.S.K.; Fahim, M.A. New compositional models for calculating the viscosity of crude oils. *Ind. Eng. Chem. Res.* **2003**, *42*, 4132–4142. [CrossRef]
28. Gharbi, R.B.; Elsharkawy, A.M.; Karkoub, M. Universal neural-network-based model for estimating the PVT properties of crude oil systems. *Energy Fuels* **1999**, *13*, 454–458. [CrossRef]
29. Elsharkawy, A.M. Predicting the dew point pressure for gas condensate reservoirs: Empirical models and equations of state. *Fluid Phase Equilib.* **2002**, *193*, 147–165. [CrossRef]
30. Gharbi, R.B.; Elsharkawy, A.M. Neural Network Model for Estimating the PVT Properties of Middle East Crude Oils. In Proceedings of the Middle East Oil Show and Conference, Manama, Bahrain, 15–18 March 1997.
31. Gharbi, R. Estimating the isothermal compressibility coefficient of undersaturated Middle East crudes using neural networks. *Energy Fuels* **1997**, *11*, 372–378. [CrossRef]
32. Ahmed, T. Fundamentals of Reservoir Fluid Behavior. In *Reservoir Engineering Handbook*; Ahmed, T., Ed.; Gulf Professional Publishing: Boston, MA, USA, 2010.
33. Xu, D.-H.; Khurana, A.K. A Simple and Efficient Approach for Improving the Prediction of Reservoir Fluid Viscosity. In Proceedings of the SPE Asia Pacific Oil and Gas Conference, Adelaide, Australia, 28–31 October 1996.
34. El-Hoshoudy, A.; Farag, A.; Ali, O.; El-Batanoney, M.; Desouky, S.; Ramzi, S. New correlations for prediction of viscosity and density of egyptian oil reservoirs. *Fuel* **2013**, *112*, 277–282. [CrossRef]
35. Makinde, F.; Ako, C.; Orodu, O.; Asuquo, I. Prediction of crude oil viscosity using feed-forward back-propagation neural network (FFBPNN). *Pet. Coal* **2012**, *54*, 120–131.
36. Sattarin, M.; Modarresi, H.; Bayat, M.; Teymori, M. New viscosity correlations for dead crude oils. *Pet. Coal* **2007**, *49*, 33–39.
37. McCain, W.D., Jr. Reservoir-Fluid Property Correlations; State of the Art. *SPE Reserv. Eng.* **1991**, *6*. [CrossRef]
38. Hemmati-Sarapardeh, A.; Shokrollahi, A.; Tatar, A.; Gharagheizi, F.; Mohammadi, A.H.; Naseri, A. Reservoir oil viscosity determination using a rigorous approach. *Fuel* **2014**, *116*, 39–48. [CrossRef]
39. Hemmati-Sarapardeh, A.; Khishvand, M.; Naseri, A.; Mohammadi, A.H. Toward reservoir oil viscosity correlation. *Chem. Eng. Sci.* **2013**, *90*, 53–68. [CrossRef]
40. Dimri, V.; Srivastava, R.P.; Vedanti, N. *Fractal Models in Exploration Geophysics: Applications to Hydrocarbon Reservoirs*; Elsevier: Oxford, UK, 2012; Volume 41.
41. De Ghetto, G.; Villa, M. Reliability analysis on PVT correlations. In Proceedings of the European Petroleum Conference, London, UK, 25–27 October 1994.
42. Isehunwa, O.S.; Olamigoke, O.; Makinde, A.A. A Correlation to Predict the Viscosity of Light Crude Oils. In Proceedings of the Nigeria Annual International Conference and Exhibition, Abuja, Nigeria, 31 July–2 August 2006.
43. Lashkenari, M.S.; Taghizadeh, M.; Mehdizadeh, B. Viscosity prediction in selected iranian light oil reservoirs: Artificial neural network versus empirical correlations. *Pet. Sci.* **2013**, *10*, 126–133. [CrossRef]
44. Ghorbani, B.; Ziabasharhagh, M.; Amidpour, M. A hybrid artificial neural network and genetic algorithm for predicting viscosity of iranian crude oils. *J. Nat. Gas Sci. Eng.* **2014**, *18*, 312–323. [CrossRef]
45. Al-Gathe, A.A.; Abd-El Fattah, K.A.; El-Banbi, A.H.; El-Metwally, K. A hybrid neuro-fuzzy approach for black oil viscosity prediction. *Ternational J. Innov. Appl. Stud.* **2015**, *13*, 946.
46. Al-Marhoun, M.; Nizamuddin, S.; Raheem, A.A.; Ali, S.S.; Muhammadain, A. Prediction of crude oil viscosity curve using artificial intelligence techniques. *J. Pet. Sci. Eng.* **2012**, *86*, 111–117. [CrossRef]
47. Beal, C. The viscosity of air. water. natural gas. crude oil and its associated gases at oil field temperatures and pressures. *Trans. AIME* **1946**, *165*, 94–115. [CrossRef]
48. Beggs, H.D.; Robinson, J. Estimating the viscosity of crude oil systems. *J. Pet. Technol.* **1975**, *27*, 1–140. [CrossRef]
49. Kaye, S.E. *Offshore California Viscosity Correlations*; Technical Report, No. TS85000940; Chevron Oil Field Research Co. (COFRC): La Habra, CA, USA, August, 1985
50. Al-Khafaji, A.H.; Abdul-Majeed, G.H.; Hassoon, S.F. Viscosity correlation for dead. Live and undersaturated crude oils. *J. Pet. Res.* **1987**, *6*, 1–16.
51. Egbogah, E.O.; Ng, J.T. An improved temperature-viscosity correlation for crude oil systems. *J. Pet. Sci. Eng.* **1990**, *4*, 197–200. [CrossRef]
52. Labedi, R. Improved correlations for predicting the viscosity of light crudes. *J. Pet. Sci. Eng.* **1992**, *8*, 221–234. [CrossRef]
53. Kartoatmodjo, T.; Schmidt, Z. Large data bank improves crude physical property correlations. *Oil Gas J.* **1994**, *92*, 7013088.
54. Bennison, T. Prediction of heavy oil viscosity. In Proceedings of the IBC Heavy Oil Field Development Conference, London, UK, 2–4 December 1998; Volume 2, p. 4.
55. Elsharkawy, A.; Alikhan, A. Models for predicting the viscosity of Middle East crude oils. *Fuel* **1999**, *78*, 891–903. [CrossRef]
56. Hossain, M.S.; Sarica, C.; Zhang, H.-Q.; Rhyne, L.; Greenhill, K. Assessment and development of heavy oil viscosity correlations. In Proceedings of the SPE International Thermal Operations and Heavy Oil Symposium, Calgary, AB, Canada, 1–3 November 2005.
57. Naseri, A.; Nikazar, M.; Dehghani, S.M. A correlation approach for prediction of crude oil viscosities. *J. Pet. Sci. Eng.* **2005**, *47*, 163–174. [CrossRef]

58. Alomair, O.A.; Elsharkawy, A.M.; Alkandari, H.A. Viscosity prediction of Kuwaiti heavy crudes at elevated temperatures. In Proceedings of the SPE Heavy Oil Conference and Exhibition, Society of Petroleum Engineers, Kuwait City, Kuwait, 12–14 December 2011.

59. Monnery, W.D.; Svrcek, W.Y.; Mehrotra, A.K. Viscosity: A critical review of practical predictive and correlative methods. *Can. J. Chem. Eng.* **1995**, *73*, 3–40. [CrossRef]

60. Mehrotra, A.K. Generalized one-parameter viscosity equation for light and medium liquid hydrocarbons. *Ind. Eng. Chem. Res.* **1991**, *30*, 1367–1372. [CrossRef]

61. Svrcek, W.; Mehrotra, A. One parameter correlation for bitumen viscosity. *Chem. Eng. Res. Des.* **1988**, *66*, 323–327.

62. Teja, A.; Rice, P. Generalized corresponding states method for the viscosities of liquid mixtures. *Ind. Eng. Chem. Fundam.* **1981**, *20*, 77–81. [CrossRef]

63. Johnson, S.E.; Svrcek, W.Y.; Mehrotra, A.K. Viscosity prediction of athabasca bitumen using the extended principle of corresponding states. *Ind. Eng. Chem. Res.* **1987**, *26*, 2290–2298. [CrossRef]

64. Alboudwarej, H.; Svrcek, W.Y.; Yarranton, H.W. PVT Investigation of Asphaltene Precipitation and Redissolution from Bitumens. In Proceedings of the Canadian International Petroleum Conference, Calgary, AB, Canada, 11–13 June 2002.

65. Al-Marhoun, M.A. Evaluation of empirically derived PVT properties for Middle East crude oils. *J. Pet. Sci. Eng.* **2004**, *42*, 209–221. [CrossRef]

66. Dindoruk, B.; Christman, P.G. PVT properties and viscosity correlations for Gulf of Mexico oils. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 30 September–3 October 2001.

67. Bergman, D.F.; Sutton, R.P. A consistent and accurate deadoil-viscosity method. *Spe Reserv. Eval. Eng.* **2009**, *12*, 815–840. [CrossRef]

68. Hajizadeh, Y. Intelligent prediction of reservoir fluid viscosity. In Proceedings of the Production and Operations Symposium SPE Annual Conference, Oklahoma City, OK, USA, 31 March–3 April 2007.

69. Omole, O.; Falode, O.; Deng, A.D. Prediction of nigerian crude oil viscosity using artificial neural network. *Pet. Coal* **2009**, *51*, 181–188.

70. Oloso, M.A.; Khoukhi, A.; Abdulraheem, A.; Elshafei, M. Prediction of crude oil viscosity and gas/oil ratio curves using recent advances to neural networks. In Proceedings of the SPE/EAGE Reservoir Characterization & Simulation Conference, Abu Dhabi, UAE, 19–21 October 2009.

71. Naseri, A.; Yousefi, S.H.; Sanaei, A.; Gharesheikhlou, A.A. A neural network model and an updated correlation for estimation of dead crude oil viscosity. *Braz. J. Pet. Gas* **2012**, *6*. [CrossRef]

72. Torabi, F.; Abedini, A.; Abedini, R. The development of an artificial neural network model for prediction of crude oil viscosities. *Pet. Sci. Technol.* **2011**, *29*, 804–816. [CrossRef]

73. Hajizadeh, Y. Viscosity prediction of crude oils with genetic algorithms. In Proceedings of the Latin American & Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina, 15–18 April 2007.

74. Hemmati-Sarapardeh, A.; Aminshahidy, B.; Pajouhandeh, A.; Yousefi, S.H.; Hosseini-Kaldozakh, S.A. A soft computing approach for the determination of crude oil viscosity: Light and intermediate crude oil systems. *J. Taiwan Inst. Chem. Eng.* **2016**, *59*, 1–10. [CrossRef]

75. Ghorbani, B.; Hamedi, M.; Shirmohammadi, R. A novel multi-hybrid model for estimating optimal viscosity correlations of Iranian crude oil. *J. Pet. Sci. Eng.* **2016**, *142*, 68–76. [CrossRef]

76. Oloso, M.A.; Hassan, M.G.; Bader-El-Den, M.B.; Buick, J.M. Ensemble SVM for characterisation of crude oil viscosity. *J. Pet. Explor. Prod. Technol.* **2017**, *8*, 531–546. [CrossRef]

77. Van der Laan, M.J.; Polley, E.C.; Hubbard, A.E. SuperLearner. UC Berkeley Division of Biostatistics Working Paper Series. Available online: Biostats.bepress.com/ucbbiostat/paper222 (accessed on 1 November 2014).

78. Ju, C.; Combs, M.; Lendle, S.D.; Franklin, J.M.; Wyss, R.; Schneeweiss, S.; van der Laan, M.J. Propensity score prediction for electronic healthcare databases using SuperLearner and high-dimensional propensity score methods. *J. Appl. Stat.* **2019**, *46*, 2216–2236. [CrossRef]

79. Polley, E.C.; Van Der Laan, M.J. *Super Learner in Prediction*; U.C. Berkeley Division of Biostatistics: Berkeley, CA, USA, 2010; p. 266.

80. Wang, H.; Rose, S.; van der Laan, M.J. Finding quantitative trait loci genes with collaborative targeted maximum likelihood learning. *Stat. Probab. Lett.* **2011**, *81*, 792–796. [CrossRef]

81. Sinisi, S.E.; Polley, E.C.; Petersen, M.L.; Rhee, S.-Y.; Van Der Laan, M.J. Super learning: An application to the prediction of HIV-1 drug resistance. *Stat. Appl. Genet. Mol. Biol.* **2007**, *6*, 7. [CrossRef]

82. Rose, S. A machine learning framel8 work for plan payment risk adjustment. *Health Serv. Res.* **2016**, *51*, 2358–2374. [CrossRef] [PubMed]

83. Sweeney, E.M.; Vogelstein, J.T.; Cuzzocreo, J.L.; Calabresi, P.A.; Reich, D.S.; Crainiceanu, C.M.; Shinohara, R.T. A Comparison of Supervised Machine Learning Algorithms and Feature Vectors for MS Lesion Segmentation Using Multimodal Structural MRI. *PLoS ONE* **2014**, *9*, e95753. [CrossRef]

84. Almudevar, A.; Oakes, D.; Hall, J. (Eds.) Statistical Modeling for Biological Systems. In *Memory of Andrei Yakovlev*; Springer Nature: New York, NY, USA, 2020.

85. Wyss, R.; Schneeweiss, S.; Van Der Laan, M.; Lendle, S.D.; Ju, C.; Franklin, J.M. Using SuperLearner Prediction Modeling to Improve High-dimensional Propensity Score Estimation. *Epidemiology* **2018**, *29*, 96–106. [CrossRef] [PubMed]

86. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 13–16 August 2016.

87. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. In Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 3146–3154.

88. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

89. Da Silva, I.N.; Spatti, D.H.; Flauzino, R.A.; Liboni, L.H.B.; dos Reis Alves, S.F. *Artificial Neural Networks*; Springer International Publishing: Cham, Switzerland, 2017.

90. Awad, M.; Khanna, R. Support vector regression. In *Efficient Learning Machines*; Springer: Berkeley, CA, USA, 2015; pp. 67–80.

91. Van Der Maaten, L. Accelerating t-sne using tree-based algorithms. *J. Mach. Learn. Res.* **2014**, *15*, 3221–3245.

92. Van der Vaart, A.W.; Dudoit, S.; van der Laan, M.J. Oracle inequalities for multi-fold cross validation. *Stat. Decis.* **2006**, *24*, 351–371.

93. Yang, X.; Dindoruk, B.; Lu, L. A comparative analysis of bubble point pressure prediction using advanced machine learning algorithms and classical correlations. *J. Pet. Sci. Eng.* **2020**, *185*, 106598. [CrossRef]

94. Sen, A.; Srivastava, M. *Regression Analysis: Theory, Methods, and Applications*; Springer Science & Business Media: Berlin, Germany, 2012.

95. Massaron, L.; Boschetti, A. *Regression Analysis with Python*; Packt Publishing: Birmingham, UK, 2016.

96. Heidaryan, E. A note on model selection based on the percentage of accuracy-precision. *J. Energy Resour. Technol.* **2019**, *141*. [CrossRef]

97. Chew, J.N.; Connally, C.A., Jr. A viscosity correlation for gas-saturated crude oils. *Trans. AIME* **1959**, *216*, 23–25. [CrossRef]

98. Ubong, U.; Oyedeko, K. Improved dead oil viscosity model. *J. Energy Technol Policy.* **2014**, *4*, 32–41.

99. Khan, S.A.; Al-Marhoun, M.A.; Duffuaa, S.O.; Abu-Khamsin, S.A. Viscosity correlations for Saudi Arabian crude oils. In Proceedings of the Middle East Oil Show, Manama, Bahrain, 7–10 March 1987.