

Article

Biodiesel Production from Jatropha: A Computational Approach by Means of Artificial Intelligence and Genetic Algorithm

Abhirup Khanna ¹, Bhawna Yadav Lamba ^{2,*}, Sapna Jain ^{2,*}, Vadim Bolshev ^{3,4}, Dmitry Budnikov ³, Vladimir Panchenko ⁵ and Alexandr Smirnov ³

¹ Department of Systemics, School of Computer Science, University of Petroleum and Energy Studies, Dehradun 248007, India; abhirupkhanna@yahoo.com

² Department of Applied Sciences (Chemistry), University of Petroleum and Energy Studies, Dehradun 248007, India

³ Federal Scientific Agroengineering Center VIM, 109428 Moscow, Russia; vadimbolshev@gmail.com (V.B.)

⁴ Laboratory of Intelligent Agricultural Machines and Complexes, Don State Technical University, 344000 Rostov-on-Don, Russia

⁵ Department of Theoretical and Applied Mechanics, Russian University of Transport, 127994 Moscow, Russia

* Correspondence: bhawnalamba5@gmail.com (B.Y.L.); sapnaj22@gmail.com (S.J.)

Abstract: In the past couple of years, the world has come to realize the importance of renewable sources of energy and the disadvantages of excessive use of fossil fuels. Numerous studies have been conducted to implicate the benefits of artificial intelligence in areas of green energy production. Artificial intelligence (AI) and machine learning algorithms are believed to be the driving forces behind the fourth industrial revolution and possess capabilities for interpreting non-linear relationships that exist in complex problems. Sustainable biofuels are derived from renewable resources such as plants, crops, and waste materials other than food crops. Unlike traditional fossil fuels such as coal and oil, biofuels are considered to be more sustainable and environmentally friendly. The work discusses the transesterification of jatropha oil into biodiesel using KOH and NaOH as alkaline catalysts. This research aims to examine and optimize the nonlinear relationship between transesterification process parameters (molar ratio, temperature, reaction time, and catalyst concentration) and biodiesel properties. The methodology employed in this study utilizes AI and machine learning algorithms to predict biodiesel properties and improve the yield and quality of biodiesel. Deep neural networks, linear regression, polynomial regression, and K-nearest neighbors are the algorithms implemented for prediction purposes. The research comprehensively examines the impact of individual transesterification process parameters on biodiesel properties, including yield, viscosity, and density. Furthermore, this research introduces the use of genetic algorithms for optimizing biodiesel production. The genetic algorithm (GA) generates optimal values for transesterification process parameters based on the desired biodiesel properties, such as yield, viscosity, and density. The results section presents the transesterification process parameters required for obtaining 72%, 85%, and 98% biodiesel yields. By leveraging AI and machine learning, this research aims to enhance the efficiency and sustainability of biodiesel production processes.

Keywords: biodiesel; deep neural networks; genetic algorithms; machine learning transesterification process; kinetics optimization



Citation: Khanna, A.; Lamba, B.Y.; Jain, S.; Bolshev, V.; Budnikov, D.; Panchenko, V.; Smirnov, A. Biodiesel Production from Jatropha: A Computational Approach by Means of Artificial Intelligence and Genetic Algorithm. *Sustainability* **2023**, *15*, 9785. <https://doi.org/10.3390/su15129785>

Academic Editors: Pratibha Rani and Arunodaya Raj Mishra

Received: 24 May 2023

Revised: 13 June 2023

Accepted: 14 June 2023

Published: 19 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rise in economic progress for countries across the world has increased their need for energy, the majority of which is sourced from scarce fossil fuels such as coal, natural gas, and petroleum. This, together with increasing environmental concern and strict environmental rules, has prompted a plethora of scientists to investigate renewable feedstock as a source of alternative fuels [1,2]. Among various options, biodiesel is becoming an increasingly popular choice because of its environmental benefits. Biodiesel is a clean-burning

alternative fuel that can be synthesized from indigenous resources [3,4]. Basically, it is a methyl ester with long-chain fatty acids generated from vegetable oil and animal fats. When compared to petroleum-based diesel, it has several benefits, such as a reduction of exhaust emissions, viz., less carbon monoxide, sulfur oxides, nitrogen hydride, particulate matter, unburned hydrocarbons during combustion, improved biodegradability, non-toxicity, inherent lubricity, a higher flash point, and domestic origin. Therefore, it is advantageous to utilize biodiesel as an alternative fuel to replace conventional petroleum-based fuels [5]. Biodiesel can be directly used in the existing diesel engine, either in neat or blended form. The alternative fuel must be technically, ecologically, and commercially viable. These conditions make triglycerides and their derivatives viable diesel fuel replacements.

The major problems with triglycerides being used as diesel fuel replacements are high viscosity, poor volatility, and polyunsaturation. These problems, however, can be mitigated by developing vegetable oil derivatives through different processes such as dilution, microemulsion, pyrolysis, and transesterification. Among various methods, transesterification is the most prevalent technique for making biodiesel, where triglycerides react with alcohol in the presence of a catalyst to produce biodiesel and glycerol. The process depends on the free fatty acid (FFA) content of the feedstock. If the FFA content of the oil is less than 1 wt % FFA, an alkali-catalyzed method is used. However, for oil containing a higher level of FFA, a two-step transesterification process is suggested that involves acid esterification followed by alkali transesterification. This present study focuses on the transesterification of *Jatropha Curcas* with FFA less than 1% using potassium hydroxide (KOH) as a catalyst and the implementation of a DNN model along with multiple machine learning models over process parameters of the transesterification process for predicting multiple characteristics of the biodiesel produced.

Machine learning is a branch of AI that works on the concept of self-learning and enables computers to do so without explicit programming [6,7]. With the assistance of adequate training data, machine learning algorithms can detect patterns, learn from them, and make conclusive predictions [8]. Similar to machine learning algorithms, the application of neural networks for predictions is a subset of AI and is inspired by the workings of the human brain. Neural networks are widely applicable for solving supervised learning and reinforcement learning-related problems. A deep neural network (DNN) is a sophisticated version of an artificial neural network (ANN) comprising multiple densely connected hidden layers between the input and output layers [9,10]. Multiple studies have suggested the superiority of DNN over shallow ANN in modeling complex nonlinear relationships. A DNN comprises several connected nodes, also known as neurons. The neurons receive input signals and trigger computational processes, which subsequently generate outputs that act as inputs for other neurons in the subsequent layers of the neural network.

The richness of the data and learning time are directly proportional to the efficiency of a DNN, with certain caveats that are introduced with different problem areas. The reason for selecting a DNN or machine learning algorithm is to identify complex relationships that exist between different input variables; the relationships may be nonlinear or polynomial in nature. The ability of the model to understand the existing relationships will determine its efficiency in predicting outcomes. Similar to AI and machine learning algorithms, the use of genetic algorithms (GA) has started gaining popularity in terms of its implementation beyond traditional computer science [11]. GA is widely used for solving optimization problems across different areas of medical science, vehicle routing, financial markets, etc. A genetic algorithm is an optimization technique that works on the principle of natural selection and helps find optimal or near-optimal solutions. GA is a type of evolutionary algorithm that is inspired by Darwin's theory of evolution. It is a randomized approach that works towards finding the best solution by making slow and slight continuous changes in every set of solutions. GA and AI are complementary in nature to one another and, when used in combination, can prove significantly beneficial.

Alternative fuels must be technically, environmentally, and economically viable. These requirements make triglycerides (vegetable oils/animal fats) and their derivatives viable diesel fuel alternatives [12–15]. Substituting triglycerides is difficult due to their high viscosity, low volatility, and polyunsaturation in the catalytic process during the transesterification of oils [16,17]. The optimization of these parameters is required for any effort that seeks.

The work discusses the use of two regression algorithms, random forest and artificial neural networks, for predicting cetane numbers for biodiesel [18]. The authors collected experimental data from various literary works and created a consolidated dataset. The data set included 12 different FAME profiles from 131 different biodiesel types. The author is aiming at establishing the impact of the FAME composition on the cetane number for biodiesel. The root mean squared error (RMSE) and the coefficient of determination (R^2) are used as performance indicators for both regression algorithms. The experimental analysis depicted a better performance for ANN as compared to the random forest algorithm. The RMSE and R^2 values for ANN are 0.95 and 2.53, respectively. Higher values of stearic acid and myristic acid beyond 51.95 and 44.95%, respectively, resulted in a positive impact on the cetane number, whereas higher values beyond 68.4% for linolenic acid resulted in a negative impact on the same.

The authors propose the creation of an estimation model for the cetane number of biodiesels [19]. The work combines the use of least squares support vector machines, genetic algorithms (GA), particle swarm optimization (PSO), and a hybrid of GA and PSO. The dataset comprised 232 fuel samples that were collected by the authors from different literary works. The coefficient of determination (R^2) and mean relative errors (MREs) were used as performance indicators. The statistical analysis indicated LSSVMHGA-PSO as the most accurate model for the estimation of cetane numbers.

The work discusses the use of ANN for predicting the physical and chemical properties of biodiesel blends [20]. The authors described the production process of biodiesel as a result of blends of virgin castor oil (VCO) and waste frying oil (WFO). The following are the respective blend compositions: B1 (100% VCO), B2 (100% WFO), B3 (50% VCO), B4 (25% VCO), and B5 (75% VCO). Acidity level, saponification index, and density were considered independent variables of the ANN model. The statistical analysis indicated the influence of WFO on the physical and chemical properties of biodiesel. Depending on the blend composition and the subsequent transesterification process, the ANN model predicted different values for fuel acidity, cetene index, and kinematic viscosity.

The authors discuss the production of biodiesel from cotton oil. Subsequent blends for biodiesel were created by mixing diesel fuel with the produced cotton biodiesel. The work proposes the use of artificial neural networks and a linear regression model for predicting the viscosity and density of the fuel and its blends that were produced [21]. The authors used temperature and blend ratio as the two independent variables for the model. The temperatures ranged from 293 K to 373 K, whereas the blends were created by mixing 20, 30, 40, 50, and 75% of diesel fuel in volumetric ratios. MAPE was used as a performance indicator for both models. The ANN model performed better in comparison to the linear regression model with a 0.02% ANOVA for predicted density.

The authors propose a genetic algorithm-based back propagation neural network model for predicting biodiesel properties concerning its FAME compositions. The model comprises five input parameters or independent variables such as methyl palmitate, methyl stearate, methyl oleate, methyl linoleate, and methyl linolenate [22]. The proposed prediction model aims to identify the nonlinear relationship between FAME compositions and biodiesel properties. The root mean square error (RMSE) and mean absolute percentage error (MAPE) are the performance indicators chosen by the authors. The authors observed that combining a genetic algorithm with a back propagation neural network assists in parameter optimization and improves prediction accuracy. The authors consider cetene number, kinematic viscosity, iodine value, and cold filter plugging point as the four output parameters or dependent variables. The experiment analysis concluded that saturated

FAMES had a positive impact on cetane number, kinematic viscosity, and cold filter plugging point, whereas the iodine value for biodiesel was dependent on unsaturated FAMES.

The authors propose an ANN-based model for yield prediction of biodiesel blends. The work discusses the production of biodiesel blends from the *Jatropha* algae oil mixture [23]. Catalyst concentration, reaction time, temperature, and methanol/oil volumetric ratio were used as independent variables for the model. The coefficient of determination (R^2) is selected as a performance indicator for the model in terms of predicting blend yield.

The work discusses the use of linear regression algorithms for the prediction of biodiesel properties. The biodiesel samples used in the dataset were produced from 17 different blends [24]. The authors selected viscosity, density, flash point, higher heating value, and oxidative stability as dependent variables. Saponification value, iodine value, and polyunsaturated fatty acids were considered independent variables. Flashpoint and oxidative stability displayed no correlation during the statistical analysis. The authors implemented the ANOVA model after regression analysis to determine the significance of the model and the importance of individual independent variables. The proposed model was successful in predicting biodiesel properties such as viscosity and density with significantly higher accuracy as compared to the flashpoint. An increase in accuracy for the model was observed by the authors when they included polyunsaturated/monounsaturated fatty acid balance (PU/MU) as part of the independent variables.

The work discusses the implementation of a linear regression model and artificial neural networks to predict yield for biodiesel [25]. The authors produce biodiesel using the transesterification process of soybean oil at a constant temperature. Reaction time, catalyst percentage, and molar ratio were selected as independent variables for both models. R^2 and RMSE are the selected performance indicators for both models. The ANN model presented better results in terms of accuracy and establishing a correlation in comparison with the linear regression model. The R^2 values for linear regression and ANN are 0.41 and 0.98, respectively. The ANN model predicted a minimum of 70.97 and a maximum of 87.56% yield for biodiesel.

The study discusses publication distribution for the applicability of machine learning algorithms for biodiesel production during 2017 and 2022 from the Scopus database [26] mentioned in Appendix A. Table 1 illustrates a comparative study between selected works focusing on the application of machine learning algorithms and neural networks for biodiesel production.

Table 1. Applications of ML algorithms in biodiesel production.

Reference	Feedstock	Input Parameters	Output Parameters	Model Selected	Performance Parameters	Genetic Algo for Optimization
[27]	-	Composition of fatty acid methyl ester (FAME)	Cetane number (CN), Viscosity and density	ANN	MSE	NA
[18]	-	12 FAMES, 131 different biodiesel types	Cetane Number	Random Forest, ANN	RMSE, R^2	NA
[20]	Virgin castor oil and Waste frying oil	Fuel acidity, ketone index, and kinematic viscosity	Yield %	ANN	average absolute deviations	NA
[22]	-	FAME	CN, kinematic viscosity, iodine value and cold filter plugging point (CFPP)	GA-based BPNN	RMSE, MAPE	Yes
[21]	Cotton Oil	Temperature, Blend Ratio	Viscosity, density	LR, ANN	MAPE	NA
[28]	Soybean oil	Molar ratio of methanol/oil, Reaction temperature	Yield	ANN	Accuracy	NA

Table 1. Cont.

Reference	Feedstock	Input Parameters	Output Parameters	Model Selected	Performance Parameters	Genetic Algo for Optimization
[29]	Castor Oil	Catalyst concentration, methanol to oil molar ratio, temperature	% FAME Yield	ANN	MSE	NA
[30]	Thevetia peruviana seed oil	Catalyst loading, methanol/oil molar ratio, time	Yield	Adaptive neuro-fuzzy inference system	SEP, R ² , MAPD	Yes
[23]	Jatropha-algae oil	Flash point, FFA, Acid number	Yield	ANN	R ²	NA
[31]	Waste Cooking Oil	Dosage of catalyst, molar ratio, mixing speed, mixing time, temperature, humidity and impurities	Viscosity turbidity, density, high heating value, and yield	SVM	RMSE	Yes
[32]	Rice bran	FFA %, Methanol to rice bran ratio, Sulfuric acid to rice bran mass ratio, reaction time, 5 N NaOH content	Yield	SVM	RMSE	Yes
Proposed Work	Jatropha	molar ratio, temperature, reaction time, catalyst concentration	Yield, Density, and Viscosity	DNN, Linear Regression, Polynomial Progression, KNN	MSE, RMSE, MAE	Yes

In the context of the present paper, it is divided into two broad sections: prediction and optimization. In terms of prediction, a DNN model along with multiple machine learning models is created over the process parameters of the transesterification process for predicting multiple characteristics of the biodiesel produced. The aspect of optimization discusses the use of genetic algorithms for predicting optimized transesterification process parameter values with respect to the desired biodiesel characteristics. A detailed explanation of our work is discussed in the subsequent methodology section.

2. Methodology

The transesterification process was carried out in a reactor. The apparatus comprises an oil bath, a reaction flask fitted with a condenser, and a digitally controlled mechanical stirrer. The glass reactor had a capacity of 500 mL, and it had three necks: one for the stirrer, the other for the condenser, and the inlet for the reactants. To determine the temperature of the reaction, a thermometer indicator was utilized. At the very bottom of the batch reactor was a valve that could be used to collect the finished product. To keep the temperature stable, the flask used for the reaction is stored inside an oil bath. Before beginning the reaction, the oil sample (200 mL) was brought up to the desired temperature through preheating. To preserve the catalytic activity and avoid moisture absorption, the potassium hydroxide–methanol solution was freshly prepared. The methanolic solution was then added to the oil that was contained in the reaction flask, and at this point, the measurement of the time began. This laboratory setup was utilized to optimize the primary process conditions in relation to the transesterification of the virgin Jatropha oils.

In order to guarantee the accuracy of the measurements of all the raw materials for the chemical reactions, as well as the temperature measurement and the other safety measures, precautions were taken. The traditional approach to the production of biodiesel, which makes use of a homogenous alkali catalyst such as KOH and produces a high yield of methyl ester, was utilized in this experiment. In the first case, the reaction conditions were optimized for maximum yield while simultaneously keeping the quality of the methyl esters as close to the specification as possible.

The amounts of methanol and metal hydroxide, in addition to the temperature of the reaction and the amount of time it takes, are the most important factors that influence the transesterification reaction. It has been determined that the molecular weight of Jatropha curcus oil, including its primary chemical constituents, is 870.19. Since the oil also consists

of some other less significant components. It was estimated that the molecular weight of the oil extracted from *Jatropha curcus* was approximately 900. For the transesterification reaction to take place, there must be a total of three moles of methanol present for every mole of vegetable oil. Because the molecular weight of methanol is 32, 96 g of methanol was required for the transesterification of one mole (or 900 g) of *Jatropha curcus* oil, which resulted in a methanol concentration of 10.67 percent.

The process of systematic literature review has resulted in the creation of the following research questions. The research questions have been formulated concerning existing works and futuristic possibilities concerning the production of biodiesel. The authors have attempted to answer the following research questions through their respective findings in the results section:

RQ1 What are the key transesterification parameters that determine biodiesel properties?

RQ2 What is the relationship between different transesterification parameters?

RQ3 What is the impact of individual transesterification parameters on different properties of biodiesel?

RQ4 What methods can be adopted to produce biodiesel with the desired physical and chemical properties?

RQ5 How can we optimize transesterification parameters to yield the most sustainable and profitable biodiesel?

Tables 2 and 3 illustrate the training dataset for input and output variables, respectively. NaOH and KOH are the two alkaline catalysts that have been used during the training process.

Table 2. Training data input variables.

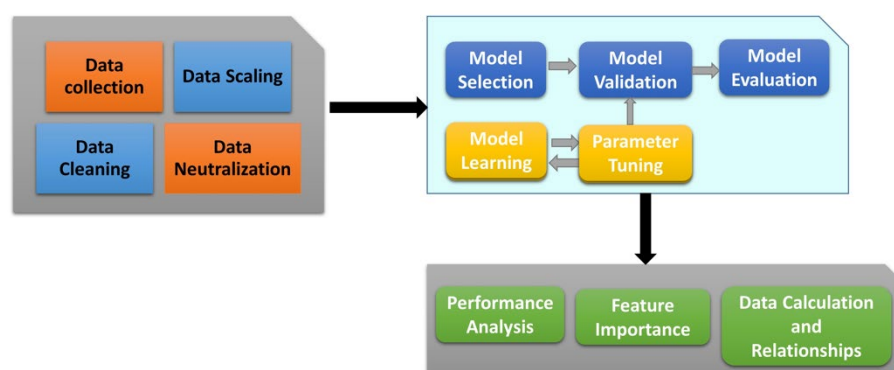
Methanol	Oil	Temp	Duration	NaOH/KOH
0.5	0.5	60	60	1
0.66666667	0.33333333	60	60	1
0.75	0.25	60	60	1
0.8	0.2	60	60	1
0.83333333	0.16666667	60	60	1
0.85714286	0.14285714	60	60	1
0.92307692	0.07692308	60	60	1
0.92307692	0.07692308	60	60	1
0.94736842	0.05263158	60	60	1
0.85714286	0.14285714	60	60	0.1
0.85714286	0.14285714	60	60	0.2
0.85714286	0.14285714	60	60	0.3
0.85714286	0.14285714	60	60	0.4
0.85714286	0.14285714	60	60	0.5
0.85714286	0.14285714	60	60	1
0.85714286	0.14285714	60	60	1.5
0.85714286	0.14285714	20	60	1
0.85714286	0.14285714	30	60	1
0.85714286	0.14285714	45	60	1
0.85714286	0.14285714	60	60	1
0.85714286	0.14285714	60	30	1
0.85714286	0.14285714	60	60	1
0.85714286	0.14285714	60	90	1
0.85714286	0.14285714	60	120	1

Table 3. Training data output variables.

	Yield	Density	Viscosity		Yield	Density	Viscosity
NaOH	32	0.8898	4.82	KOH	34	0.8876	4.64
	40	0.8898	4.84		46	0.8877	4.64
	51	0.8898	4.96		57	0.8877	4.65
	64	0.8895	4.89		69	0.8878	4.63
	70	0.889	4.89		78	0.888	4.63
	89	0.8891	4.98		93	0.8885	4.63
	92	0.8892	4.88		93.5	0.8885	4.63
	89	0.8886	4.89		94	0.8885	4.62
	88	0.8886	4.92		94	0.8885	4.62
	39	0.889	4.89		43	0.8891	4.71
	50	0.889	4.9		54	0.8891	4.69
	61	0.889	4.92		69	0.889	4.69
	75	0.889	4.98		80	0.8889	4.67
	82	0.8885	4.99		91	0.8886	4.63
	89	0.8886	4.99		93	0.8885	4.63
	82	0.8885	4.97		92.5	0.8885	4.63
	30	0.898	4.98		35	0.899	4.9
	60	0.899	4.97		65	0.8975	4.88
	75	0.8891	4.99		82	0.8885	4.76
	85	0.889	4.63		93	0.8885	4.63
	72	0.8898	4.84		78	0.8895	4.84
	82	0.8891	4.8		93	0.8891	4.8
	83	0.8898	4.73		95	0.8888	4.73
	88	0.8898	4.63		92	0.8885	4.63

3. Experimental Details

Biodiesel is a renewable and biodegradable form of diesel synthesized from renewable sources such as vegetable oils, waste oils, and animal fat oils. The process of biodiesel production involves multiple steps such as feedstock pre-treatment, transesterification reactions, purification of biodiesel, etc. Our current work focuses on the application of machine learning and neural network algorithms for optimizing the transesterification reaction. Figure 1 represents the model workflow. Figure 2 depicts the architecture for the DNN model implemented in our work. In this section, the authors discuss the four algorithms that have been implemented to comprehend the relationships between transesterification process parameters and predict biodiesel properties. The section talks about the role of genetic algorithms in optimizing the process parameters for transesterification. A flowchart has been presented that depicts the workings of the genetic algorithm.

**Figure 1.** Model workflow.

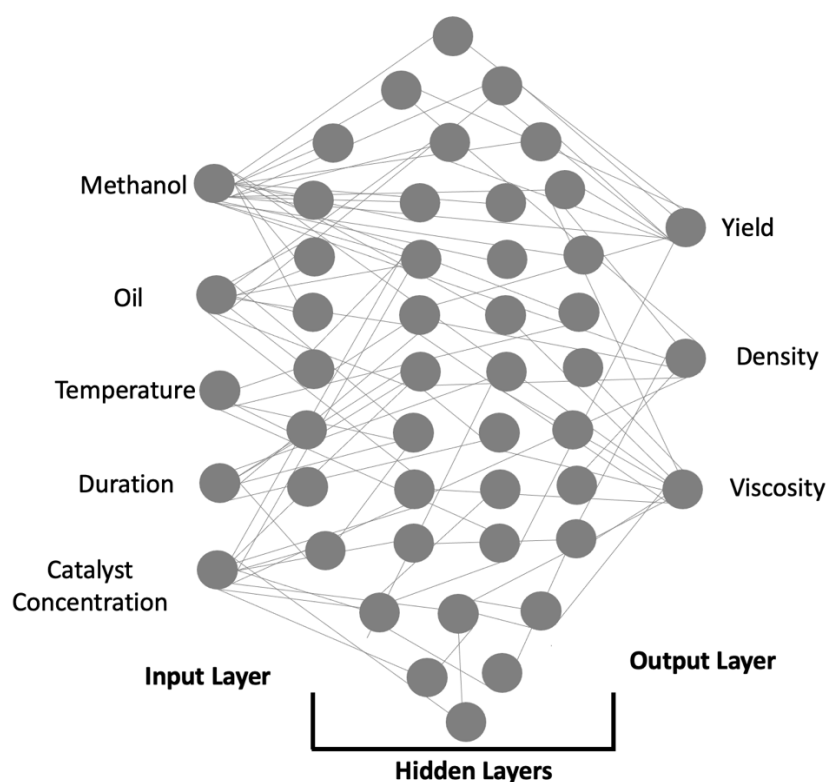


Figure 2. Deep neural network architecture.

Linear regression is one of the most widely implemented machine learning algorithms for predictive analytics. It falls under the category of supervised learning and performs regression tasks. The purpose of a regression model is to predict the target variable based on certain independent variables. The target variable may also be addressed as the output variable, whereas the independent variables may be addressed as input variables. The LR algorithm assesses the relationship between dependent and independent variables. It estimates the impact of independent variables on the magnitude of the output variable. The linear regression algorithm generates a straight line or a surface that minimizes the distance between predicted and actual output values. A simple linear regression model may comprise one input and one output variable, whereas in the case of multivariate regression, there are multiple independent variables involved. Multi-output linear regression involves predicting multiple output variables from a set of various independent variables. In this work, the multi-output linear regression algorithm has been implemented, which considers transesterification process parameters as input variables and properties of biodiesel as target or output variables. The model establishes a linear correlation between input and output variables. It predicts the behavior of output variables based on the values of multiple independent variables.

The linear regression algorithm works well in the case of a linear relationship between independent and dependent variables. In the case of a non-linear relationship between dependent and independent variables, certain polynomial expressions have been added to the linear regression to convert it into a polynomial regression problem. A polynomial regression model computes an n th-degree polynomial relationship between independent and dependent variables. In the current work, the polynomial regression algorithm has been implemented to ensure the prediction of biodiesel properties. The work computes the model on different polynomial degrees, evaluates the performance and accuracy of the model at each stage, and subsequently selects the best n th-degree option.

The K-Nearest Neighbor (KNN) regression model is non-parametric and intuitive in nature. It approximates the association between independent and dependent variables based on distance functions. KNN-based regression is most suited for problems with lower

data dimensionality. The KNN algorithm can be used for both classification and regression problems. The KNN model works on the concept of common intuition (feature similarity) and is adequate for handling nonlinear relationships in the data set without involving any complicated data engineering practices.

The final set of algorithms that have been implemented in this work of ours is the deep neural network (DNN). The DNN architecture comprises an input layer, multiple densely connected hidden layers, and an output layer. Each layer in the neural network is comprised of an activation function that is responsible for introducing nonlinearity in the data set. Hyperparameter tuning is one of the key aspects of implementing a DNN model. The efficiency, accuracy, and learning rate of the model are directly related to its hyperparameters. The proposed work discusses the implementation of a DNN model coupled with a genetic algorithm for its hyperparameter tuning. The deep net model considers transesterification process parameters as input variables, constructs generalizations based on input values, and stores them in the form of weights in the hidden layers. Table 4 describes the various hyperparameters that have been used to construct the deep neural network.

Table 4. DNN network parameters.

Parameters	Values
Number of neurons in the Input Layer	20
Total number of hidden layers	4
Number of Input Parameters	5
Number of neurons in hidden layer 1	20
Number of neurons in hidden layer 2	32
Number of neurons in hidden layer 3	48
Number of neurons in hidden layer 4	64
Number of neurons in the output layer	3
Activation function	ReLU
Optimization Algorithm	Adam
Learning Cycles	200, 250, 300 Epochs
Loss functions	MSE, Cross Entropy

Deep neural networks (DNNs) are a type of artificial neural network (ANN) with multiple layers of interconnected nodes, also known as artificial neurons. These networks are called “deep” because they have many layers, typically more than three. DNNs are modeled after the structure and function of the human brain and are used for a wide variety of tasks, such as image and speech recognition, natural language processing, and decision-making. They are trained to recognize patterns and make predictions by adjusting the weights and biases of the artificial neurons in each layer. A DNN typically has an input layer, one or more hidden layers, and an output layer. The input layer receives the raw data, and the output layer produces the final prediction or decision. The hidden layers process the data and extract features that are used to make the final decision. The training process of DNNs is conducted by providing them with a large set of labeled data and adjusting the weights and biases of the artificial neurons in each layer to minimize the error between the predicted output and the true output. The training process of a DNN is carried out by providing it with a large set of labeled data and adjusting the weights and biases of the artificial neurons in each layer to minimize the error between the predicted output and the true output.

The algorithm used to adjust the weights and biases is called backpropagation. During the forward pass, the input data is passed through the network, and the artificial neurons in each layer process the data by applying a set of mathematical operations to it. These operations include matrix multiplications, activation functions, and bias additions. The activation function is a non-linear function that is applied to the output of each neuron to introduce non-linearity into the network, allowing it to learn and represent more complex patterns. During the backward pass, the error is propagated back through the network,

and the weights and biases of the neurons are updated to reduce the error. This process is repeated multiple times until the error is minimized. After the training is complete, the DNN can be used to make predictions on new, unseen data by passing it through the trained network. The predictions are made by the output layer, which produces the final decision based on the features extracted by the hidden layers.

A genetic algorithm (GA) is a type of optimization algorithm that is inspired by the process of natural selection. It is used to find an optimal solution to a problem by mimicking the process of evolution. A GA is used to find an optimal solution in a large search space by simulating the process of natural selection. Table 5 describes the operating parameters for the implemented GA. Figure 3 explains the workflow for a GA. The basic components of the GA are a population of candidate solutions, a fitness function, a selection mechanism, and genetic operators such as crossover and mutation. The population of candidate solutions is a set of potential solutions to the problem, represented as a set of chromosomes. Each chromosome is a string of bits or numbers that encodes a possible solution. The fitness function is used to evaluate the quality of each candidate solution by assigning it a fitness value. The selection mechanism is used to select the fittest individuals from the population, which will be used to generate the next generation of solutions.

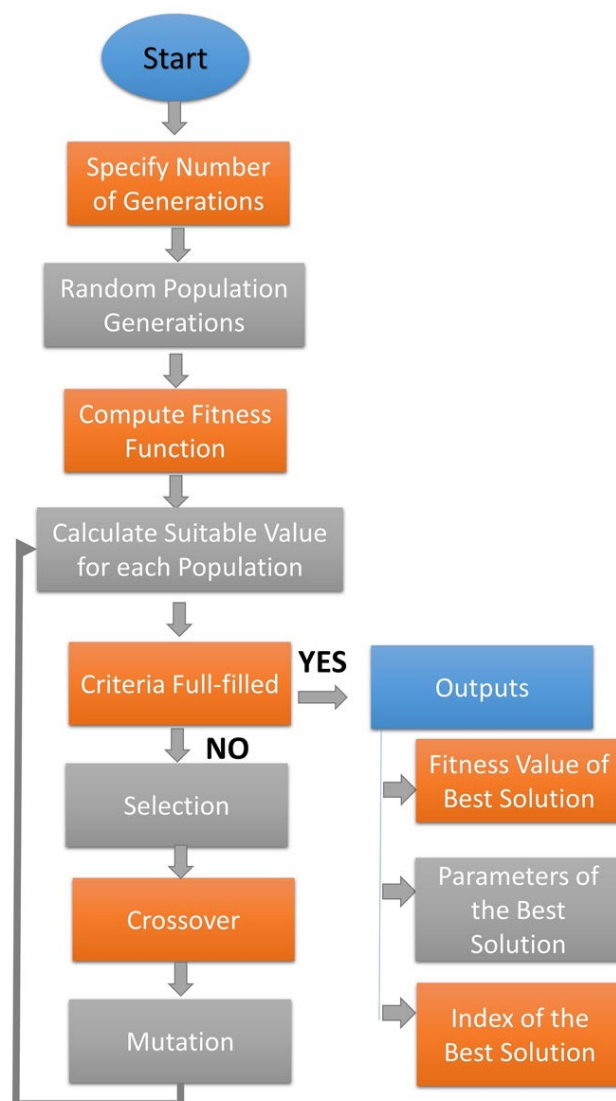


Figure 3. Working of genetic algorithm.

Table 5. Genetic algorithm operating parameters.

Parameters	Values
Number of generations	100
Population size	50
Crossover Rate	0.8
Mutation Rate	0.2
Selection Scheme	Tournament selection

The genetic operators are used to create new solutions by combining the genetic information of the selected individuals. The process starts with a random population of candidate solutions, and then it repeatedly applies the selection, crossover, and mutation operators to create new generations of solutions. The algorithm stops when a satisfactory solution is found or when a stopping criterion is met, such as a maximum number of iterations. Tables 6 and 7 represent the results obtained after implementing GA.

Table 6. Genetic algorithm optimization data KOH.

Biodiesel Properties	Values	Methanol	Oil	Temperature	Duration	Catalyst
Yield	72	0.83	0.17	60	60	0.5
	85	0.86	0.14	40	60	1.2
	98	0.85	0.15	43	98	0.7
Density	0.89	0.85	0.15	30	60	0.9
	0.88	0.83	0.17	60	60	1.2
Viscosity	4.8	0.86	0.14	60	60	0.9
	4.99	0.85	0.15	30	60	1
	4.63	0.87	0.13	60	60	1

Table 7. Genetic algorithm optimization data NaOH.

Biodiesel Properties	Values	Methanol	Oil	Temperature	Duration	Catalyst
Yield	72	0.86	0.14	43	60	0.5
	85	0.86	0.14	48	60	1
	98	0.93	0.07	60	60	1.2
Density	0.89	0.85	0.15	60	60	0.3
	0.88	0.88	0.12	60	60	1
Viscosity	4.8	0.86	0.14	60	60	1
	4.99	0.85	0.15	43	60	0.7
	4.63	0.85	0.15	60	60	1

4. Results and Discussion

In this section, a clear and concise summary of the experimental results has been added, including quantitative evaluations of the performance of the proposed model. These results provide insight into the behavior of different regression models and discuss patterns that have emerged from the experiments. Tables 8–15 represent prediction values for yield, density, and viscosity obtained from the implementation of DNN along with different regression models.

Table 8. Deep neural network prediction NaOH.

S No.	Predicted Yield	Predicted Density	Predicted Viscosity
1	27.2	0.863106	4.6272
2	34	0.863106	4.6464
3	43.35	0.863106	4.7616
4	54.4	0.862815	4.6944
5	59.5	0.86233	4.6944
6	75.65	0.862427	4.7808
7	78.2	0.862524	4.6848
8	75.65	0.861942	4.6944
9	74.8	0.861942	4.7232
10	33.15	0.86233	4.6944
11	42.5	0.86233	4.704
12	51.85	0.86233	4.7232
13	63.75	0.86233	4.7808
14	69.7	0.861845	4.7904
15	75.65	0.861942	4.7904

Table 9. Deep neural network prediction KOH.

S No.	Predicted Yield	Predicted Density	Predicted Viscosity
1	30.26	0.878724	4.5472
2	40.94	0.878823	4.5472
3	50.73	0.878823	4.557
4	61.41	0.878922	4.5374
5	69.42	0.87912	4.5374
6	82.77	0.879615	4.5374
7	83.215	0.879615	4.5374
8	83.66	0.879615	4.5276
9	83.66	0.879615	4.5276
10	38.27	0.880209	4.6158
11	48.06	0.880209	4.5962
12	61.41	0.88011	4.5962
13	71.2	0.880011	4.5766
14	80.99	0.879714	4.5374
15	82.77	0.879615	4.5374

Table 10. Linear regression observed and predicted data KOH.

Yield	Density	Viscosity	Predicted Yield	Predicted Density	Predicted Viscosity
91	0.8886	4.63	70.30959831	0.88879447	4.67993182
82	0.8885	4.76	66.14170564	0.89243599	4.746798
69	0.8878	4.63	75.74625353	0.8883315	4.654406
93	0.8885	4.63	84.82672903	0.88849759	4.654877
78	0.8895	4.84	79.92276644	0.88843614	4.649803
65	0.8975	4.88	47.45668226	0.8963744	4.838719
54	0.8891	4.69	61.59931988	0.8889726	4.694965
78	0.888	4.63	81.04319753	0.88842838	4.654681

Table 11. Linear regression observed and predicted data NaOH.

Yield	Density	Viscosity	Predicted Yield	Predicted Density	Predicted Viscosity
82	0.8885	4.99	64.79293	0.889073	4.895656
75	0.8891	4.99	60.10289	0.892384	4.919801
64	0.8895	4.89	69.48634	0.889182	4.875849
89	0.8886	4.99	78.16462	0.889014	4.883681
72	0.8898	4.84	73.2634	0.888542	5.01589
60	0.899	4.97	42.04116	0.895753	4.95592
50	0.889	4.9	56.76992	0.889109	4.90284
70	0.889	4.89	74.54867	0.889084	4.880418

Table 12. Poly regression observed and predicted data KOH.

Yield	Density	Viscosity	Predicted Yield	Predicted Density	Predicted Viscosity
34	0.8876	4.64	29.39312	0.887523	4.62954
46	0.8877	4.64	54.28247	0.887761	4.660547
57	0.8877	4.65	68.04869	0.887977	4.663683
69	0.8878	4.63	76.73132	0.888137	4.661608
78	0.888	4.63	82.69595	0.888257	4.658575
93	0.8885	4.63	87.0427	0.888349	4.655601
93.5	0.8885	4.63	99.45522	0.888631	4.643853
94	0.8885	4.62	99.45521731	0.888631	4.643853

Table 13. Poly regression observed and predicted data NaOH.

Yield	Density	Viscosity	Predicted Yield	Predicted Density	Predicted Viscosity
32	0.8898	4.82	27.8818	0.889884	4.821446
40	0.8898	4.84	47.63148	0.889698	4.867874
51	0.8898	4.96	60.52953	0.889464	4.88095
64	0.8895	4.89	69.23578	0.889279	4.88555
70	0.889	4.89	75.44304	0.889136	4.887265
89	0.8891	4.98	80.07423	0.889025	4.887828
92	0.8892	4.88	93.75772	0.888678	4.886507
89	0.8886	4.89	93.75772	0.888678	4.886507

Table 14. K nearest neighbor observed and predicted data KOH.

Yield	Density	Viscosity	Predicted Yield	Predicted Density	Predicted Viscosity
91	0.8886	4.63	67.4	0.88894	4.678
82	0.8885	4.76	85.2	0.8903	4.706
69	0.8878	4.63	78	0.8881	4.634
93	0.8885	4.63	90	0.88852	4.664
78	0.8895	4.84	90	0.88882	4.706
65	0.8975	4.88	73.6	0.8924	4.76
54	0.8891	4.69	67.4	0.88894	4.678
78	0.888	4.63	90	0.88852	4.664

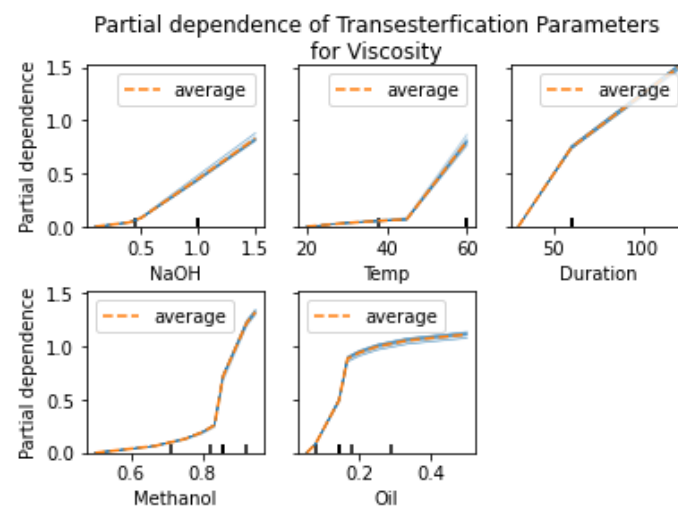
Table 15. K nearest neighbor observed and predicted data NaOH.

Yield	Density	Viscosity	Predicted Yield	Predicted Density	Predicted Viscosity
82	0.8885	4.99	61.4	0.8889	4.936
75	0.8891	4.99	79.6	0.89096	4.912
64	0.8895	4.89	72.6	0.8892	4.942
89	0.8886	4.99	83	0.88896	4.858
72	0.8898	4.84	83.4	0.88912	4.848
60	0.899	4.97	68.6	0.89276	4.982
50	0.889	4.9	61.4	0.8889	4.936
70	0.889	4.89	83	0.88896	4.858

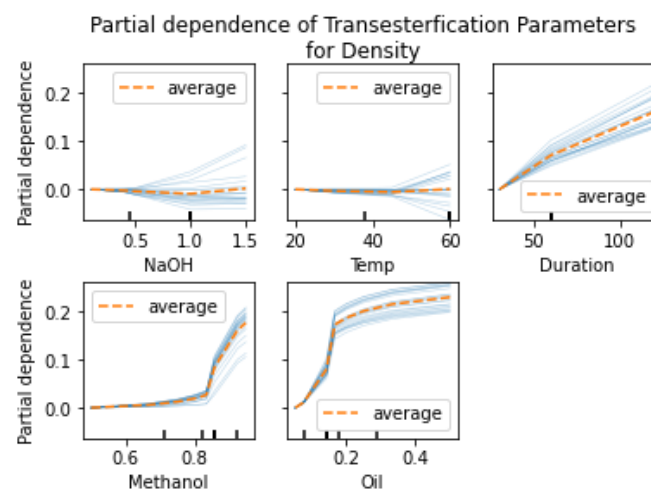
PDP (partial dependence plot) graphs are a way to visualize the relationship between a single feature of a dataset and the outcome variable in a machine learning model. The graph shows how the outcome variable changes as the features of interest are varied while keeping all other features constant. PDP graphs are often used to interpret the results of complex models and identify which features have the greatest impact on the outcome. They are also useful for identifying non-linear relationships between features and the outcome variable. Figures 4 and 5 illustrate the PDP graphs between transesterification process parameters and biodiesel properties for NaOH and KOH as catalysts, respectively.

4.1. Influence of Methanol/Oil Ratio

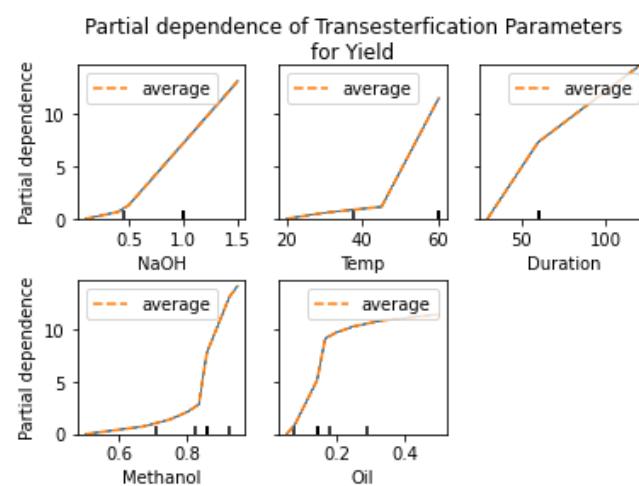
One of the crucial aspects that determines the outcome of the transesterification reaction is the proportion of alcohol to oil that is used. In theory, transesterification requires three moles of alcohol for every mole of oil that is being converted. In actuality, however, the molar ratio ought to be on the higher side to move the reaction forward in accordance with Le Chatelier's principle, which states that an increase in the concentration of the reactant shifts the equilibrium toward the formation of the product. This is the case to move the reaction forward in accordance with Le Chatelier's principle. Experiments were run at 60 °C and 400 revolutions per minute with varying proportions of methanol to oil to determine the optimal amount of methanol needed for the reaction (6:1, 12:1, 15:1, 18:1). When using a ratio of 6:1, the yield that could be obtained in two hours was a maximum of 88% weight. There is not much of an increase in the % yield with higher molar ratios, but the process takes less time.



(a)

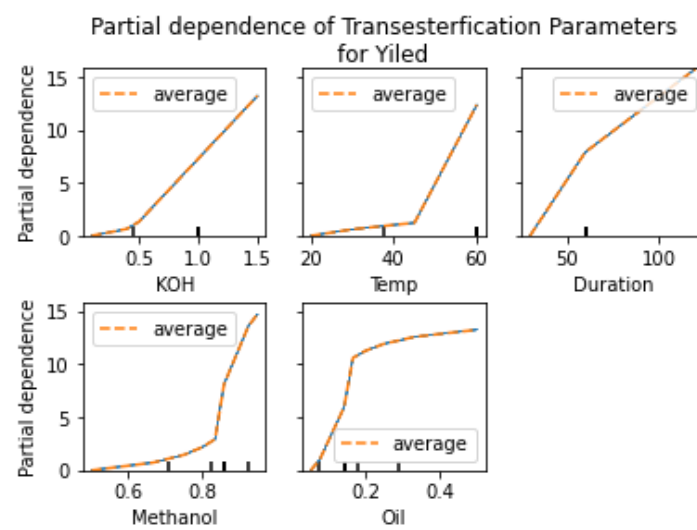


(b)

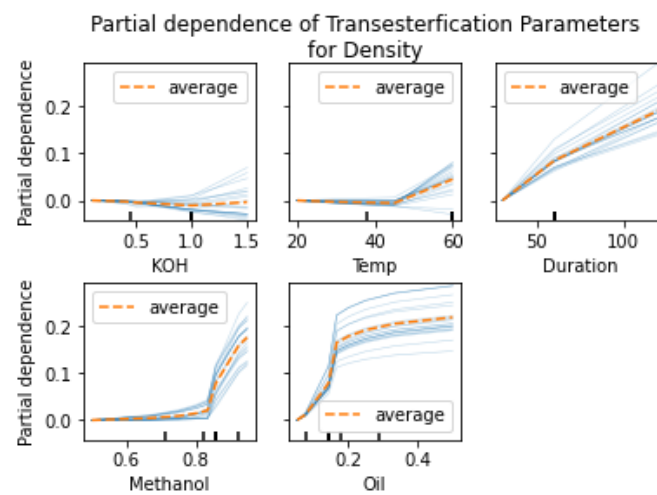


(c)

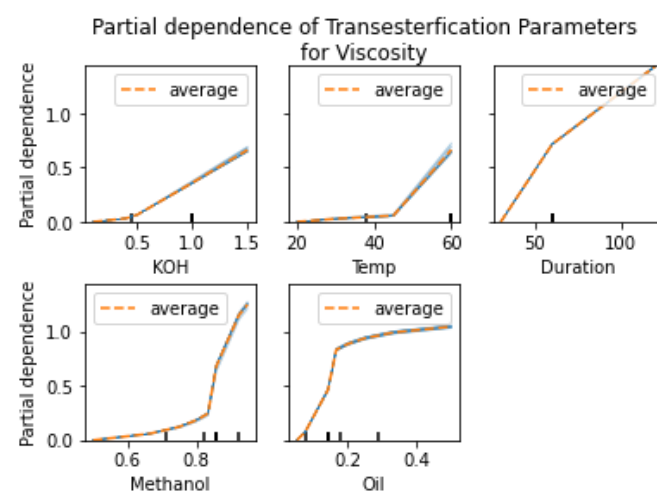
Figure 4. PDP Graphs NaOH. (a) Viscosity; (b) density; (c) yield.



(a)



(b)



(c)

Figure 5. PDP Graphs KOH: (a) Viscosity; (b) density; (c) yield.

4.2. Influence of Catalyst Concentration

KOH was used as a catalyst for the transesterification of *Jatropha curcus* oil with methanol. The reaction was carried out at a temperature of 60 °C, with a rotational speed of 400 revolutions per minute and a ratio of six parts methanol to one part oil. It was determined through observation that a concentration of KOH of 1.0% is necessary for the most efficient transesterification possible. It was observed that although there was not a significant increase in the ester yield when the concentration of KOH was decreased below or increased above 1%, there was an increase in the formation of glycerol. The highest possible ester yield of 97% was achieved by using a KOH concentration of 1.0% by weight.

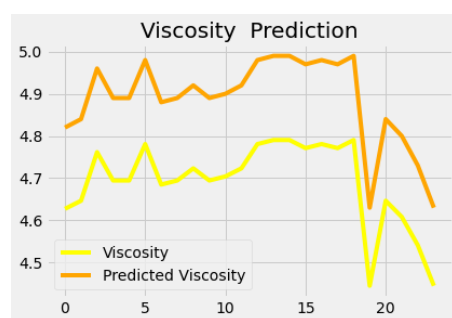
4.3. Influence of Reaction Temperature

The temperature ranges of 30, 45, and 60 °C were used in this investigation. The reaction was carried out with a constant reaction time of 120 min, a constant methanol-to-oil ratio, and constant KOH concentrations of 6:1 and 1.0%, respectively. The highest ester yield of 97% was achieved at a temperature of 60 degrees Celsius. It is evident from this that the ester yield increases in direct proportion to the degree to which the reaction temperature is raised. The temperature of the reaction should never rise above methanol's boiling point, which is 65 degrees Celsius. As a result, the temperature of the reaction was maintained at 60 degrees Celsius.

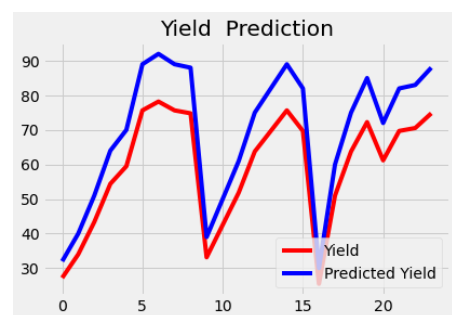
4.4. Influence of Reaction Time

In order to determine the optimal time for the reaction, it was carried out for 30, 60, 90, and 120 min at a temperature of 60 °C with a rotational speed of 400 revolutions per minute and a ratio of 6:1 methanol to oil. According to the findings, it is abundantly clear that the ester yield rises as the reaction time increases. At 90 and 120 min of reaction time, the ester yields were almost identical. It is abundantly clear that a reaction time of 120 min resulted in the highest ester yield possible, which was 89% by weight.

Figures 6 and 7 describe the performance evaluation graphs for the DNN with respect to NaOH and KOH as catalysts. The graphs indicate the comparison between actual values of biodiesel properties and predicted values.

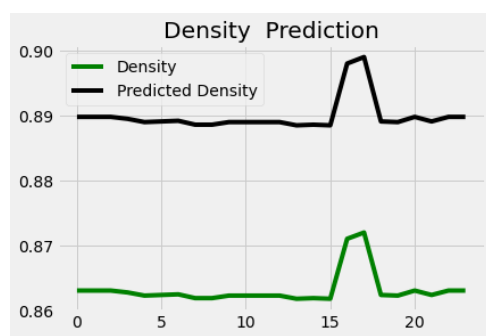


(a)

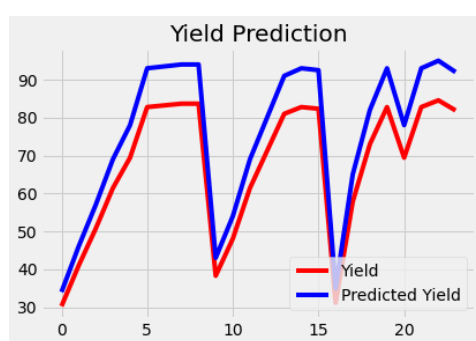


(b)

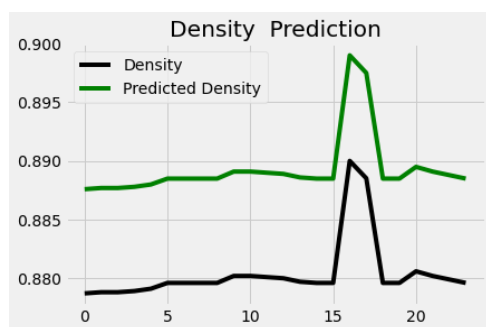
Figure 6. Cont.



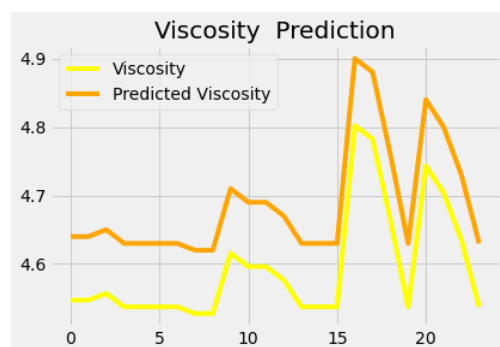
(c)

Figure 6. DNN prediction graphs NaOH: (a) Viscosity; (b) yield; (c) density.

(a)



(b)



(c)

Figure 7. DNN prediction graphs KOH: (a) Yield; (b) density; (c) viscosity.

Figures 8 and 9 discuss the functioning of polynomial regression in terms of calculating the degree of the polynomial equation that describes the best non-linear relationship between transesterification process parameters and individual biodiesel properties.

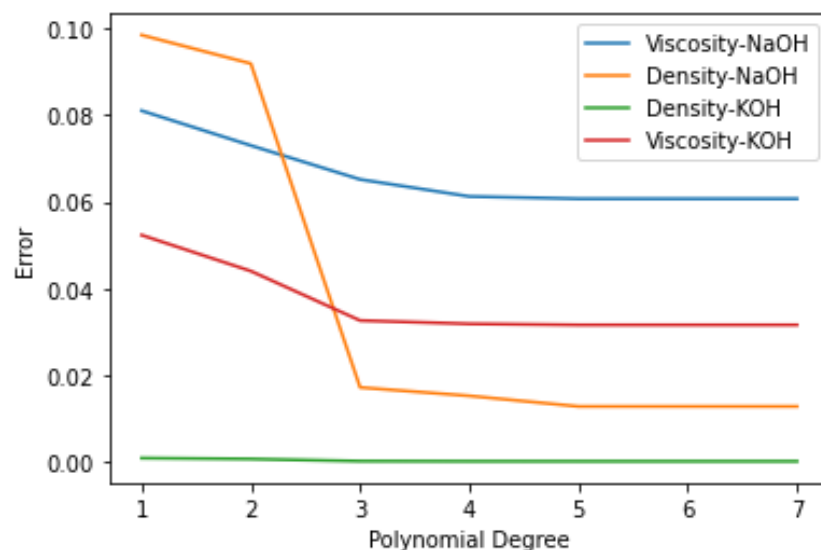


Figure 8. Polynomial regression NaOH and KOH: Density vs. Viscosity.

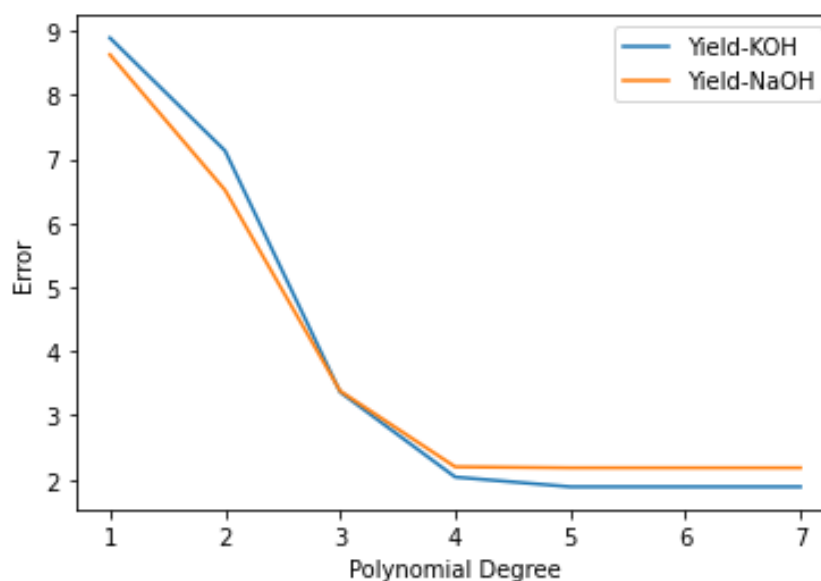
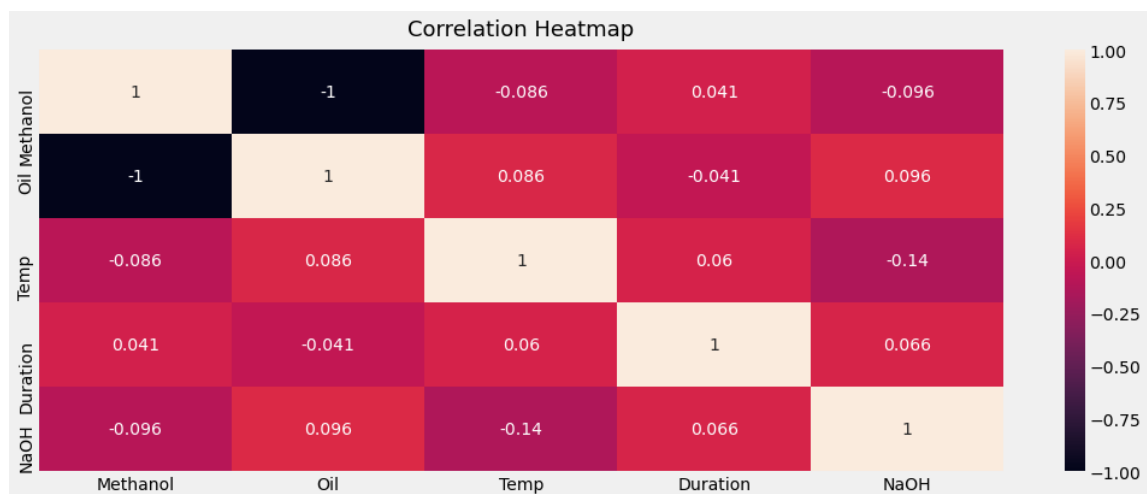


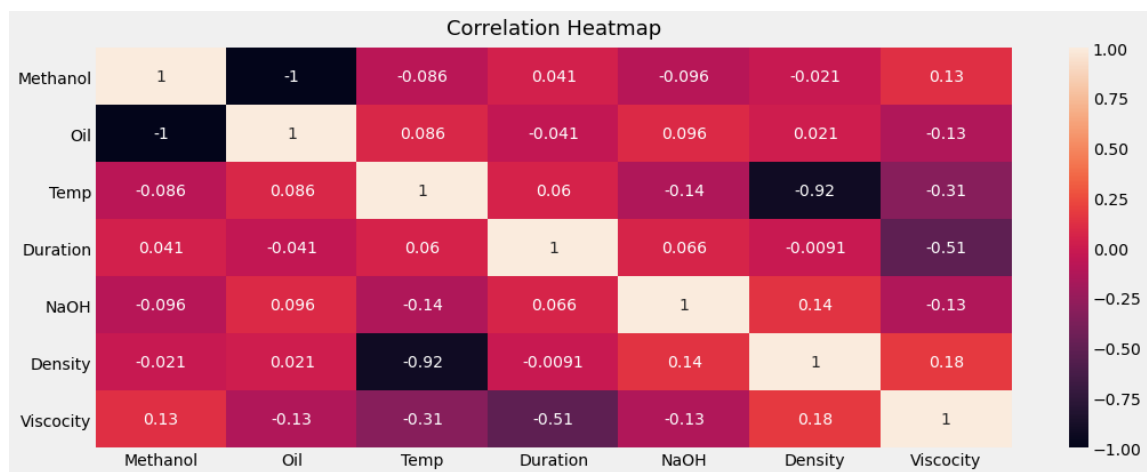
Figure 9. Polynomial regression NaOH and KOH yield.

Figures 10 and 11 illustrate the correlations that exist between individual transesterification process parameters and biodiesel properties and transesterification process parameters, respectively. The following heatmaps indicate the level of importance of a transesterification process parameter in terms of its impact on biodiesel properties.

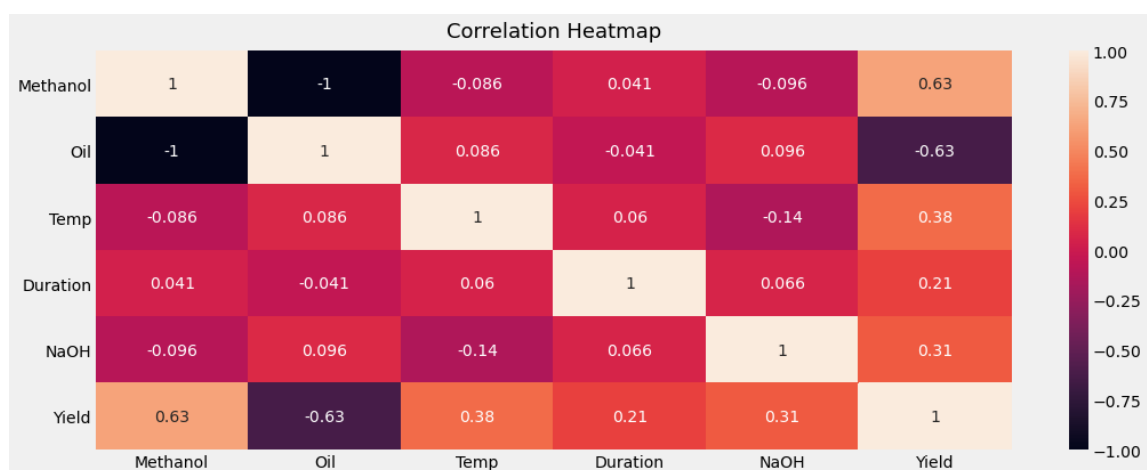
Mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) have been used as performance indicators for comparing the performance of different regression models and the DNN. Figures 12–17 represent the same for individual biodiesel properties of yield, density, and viscosity.



(a)

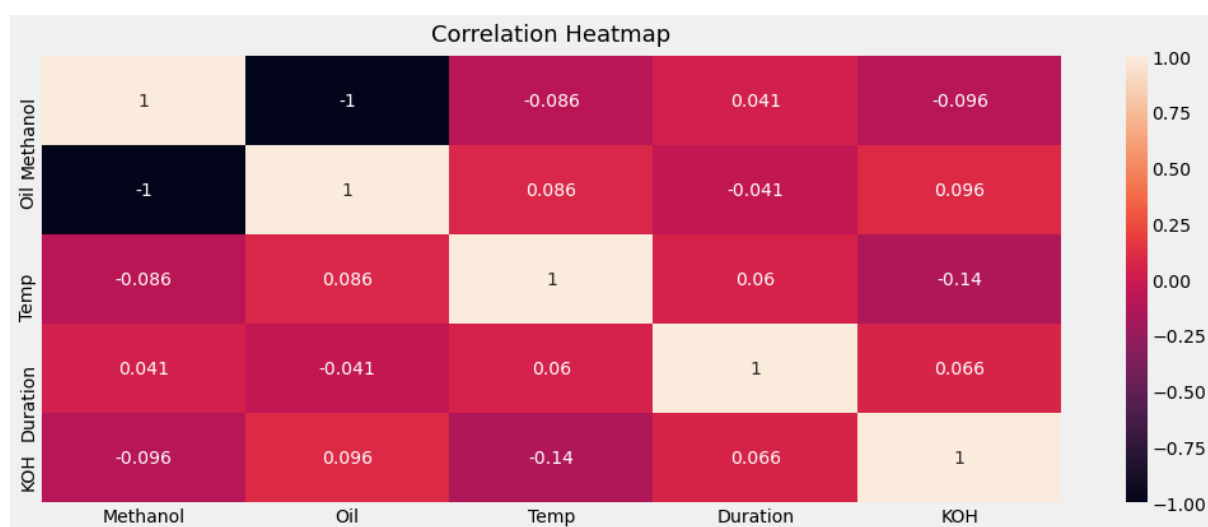


(b)

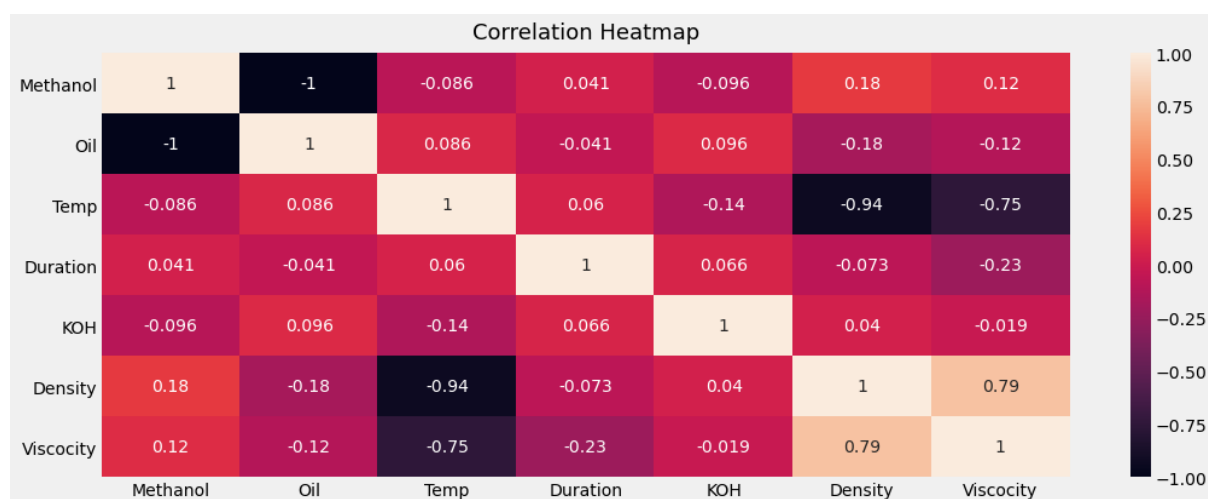


(c)

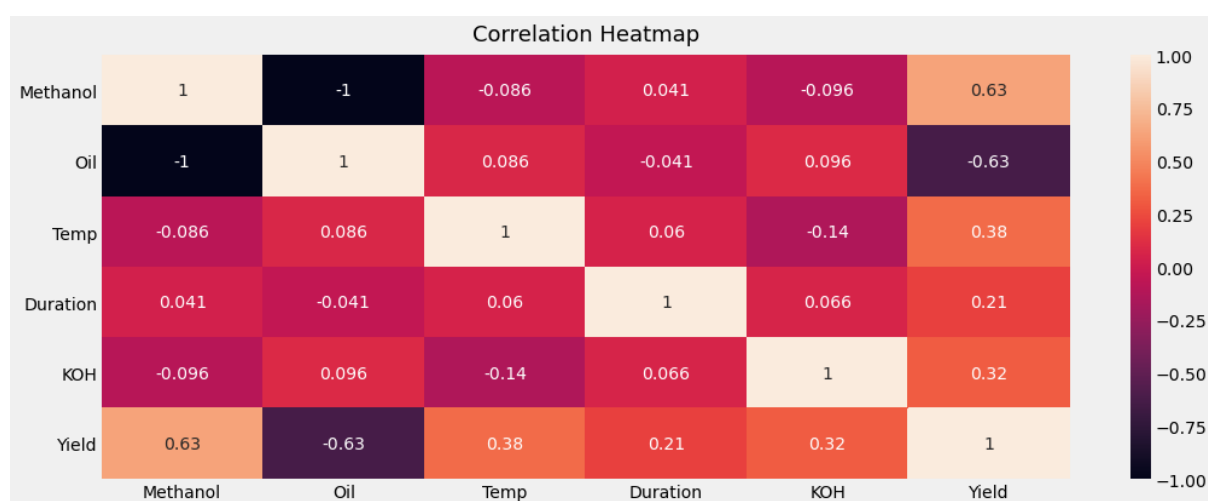
Figure 10. Heatmap NaOH: (a) Input parameters; (b) input parameters vs. density and viscosity; (c) input parameters vs. yield.



(a)

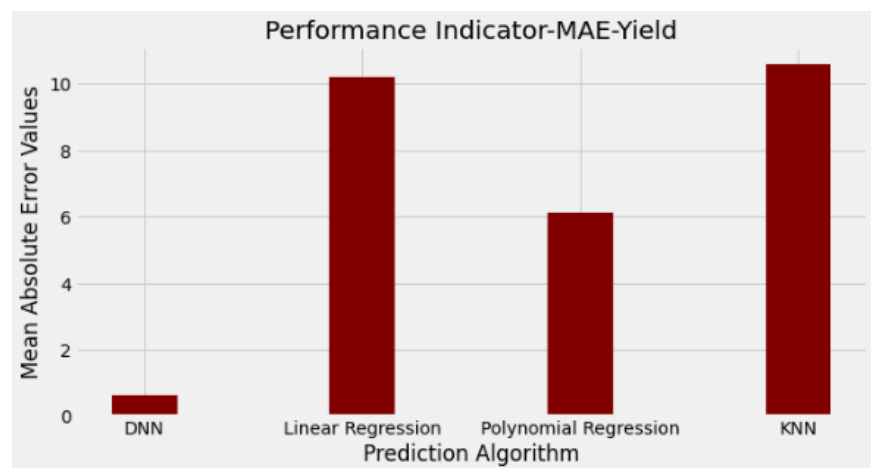


(b)

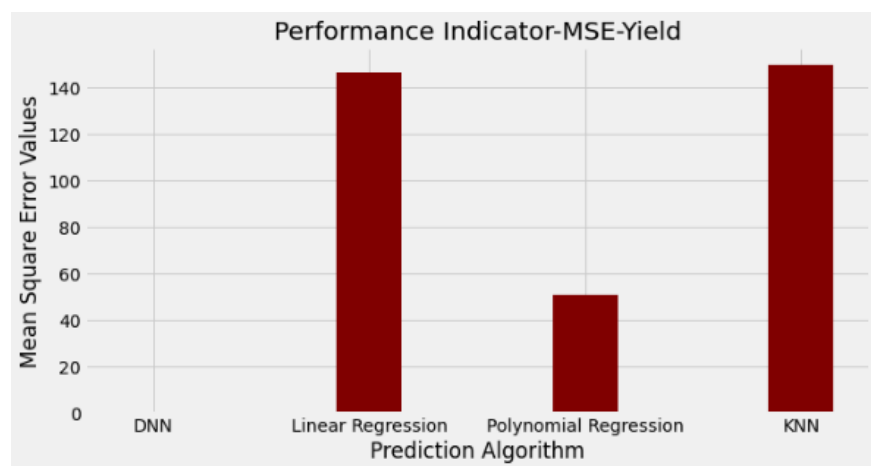


(c)

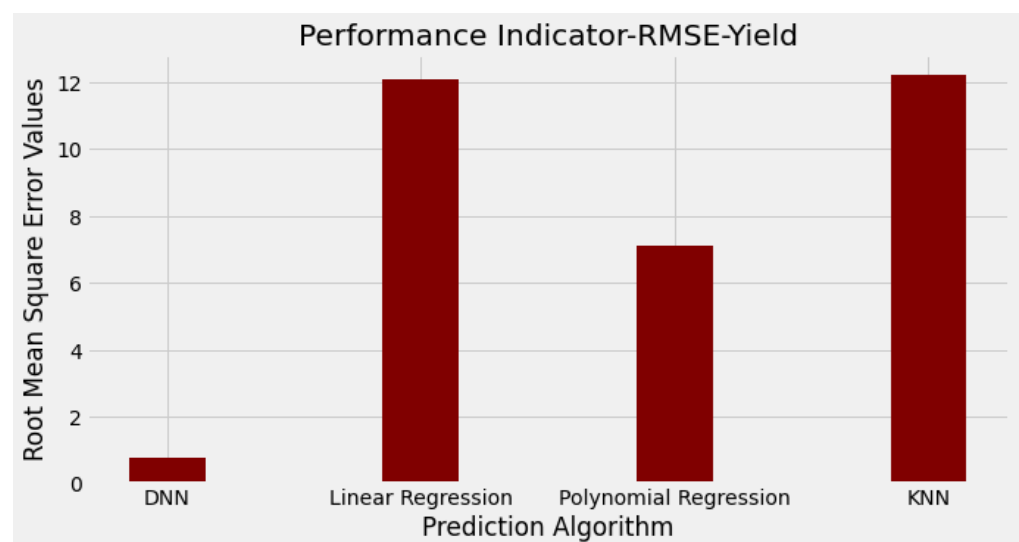
Figure 11. Heatmap KOH: (a) Input parameters; (b) input parameters versus density and viscosity; (c) input parameters versus yield.



(a)

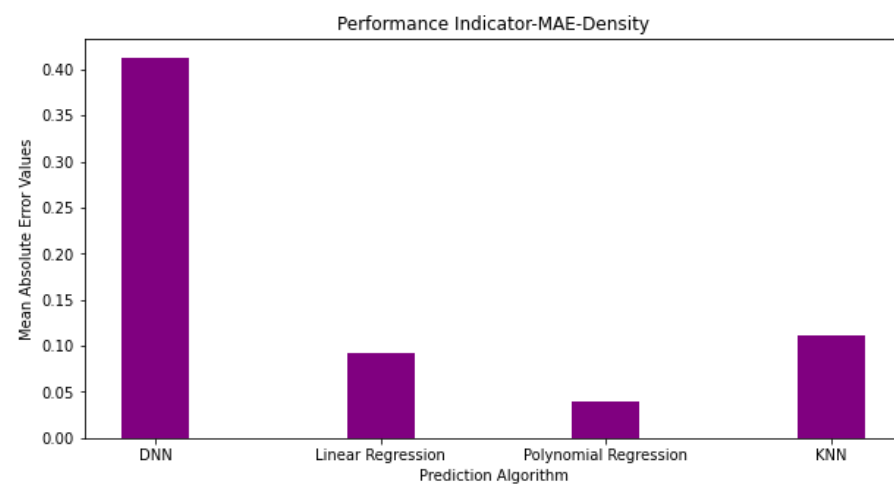


(b)

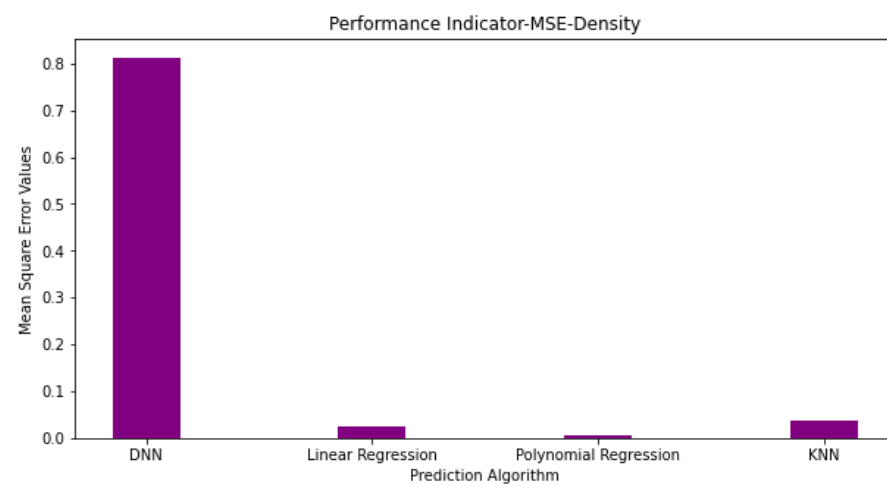


(c)

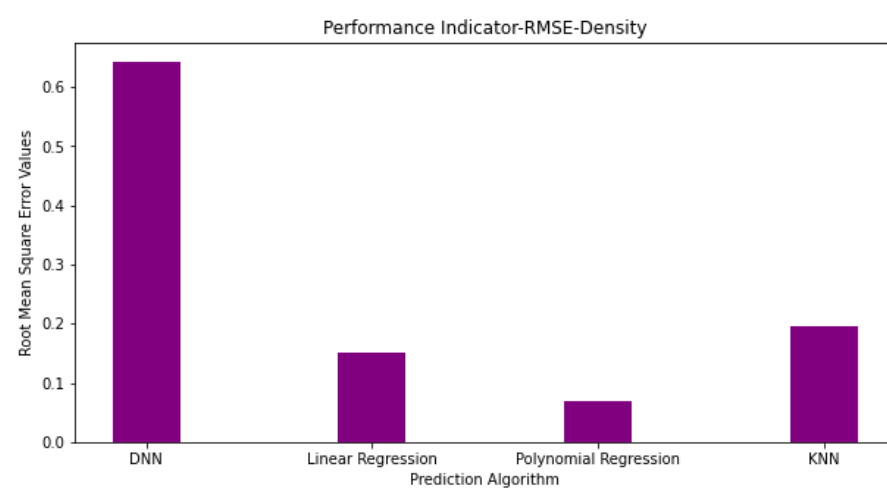
Figure 12. Performance indicator graph KOH: (a) Mean absolute error values of yield; (b) mean square error values of yield; (c) root mean square error values of yield.



(a)

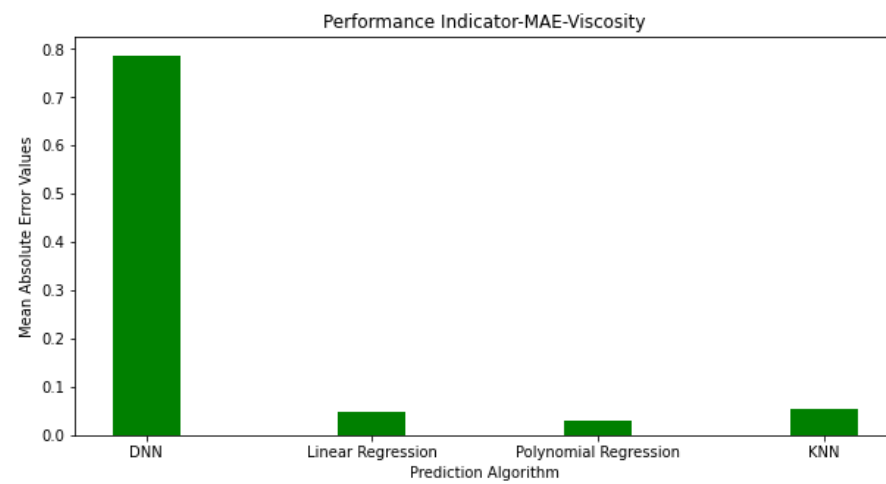


(b)

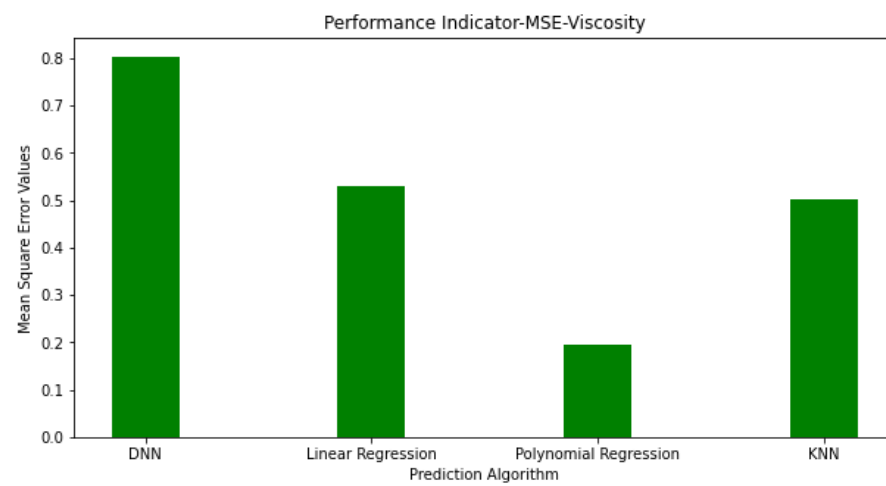


(c)

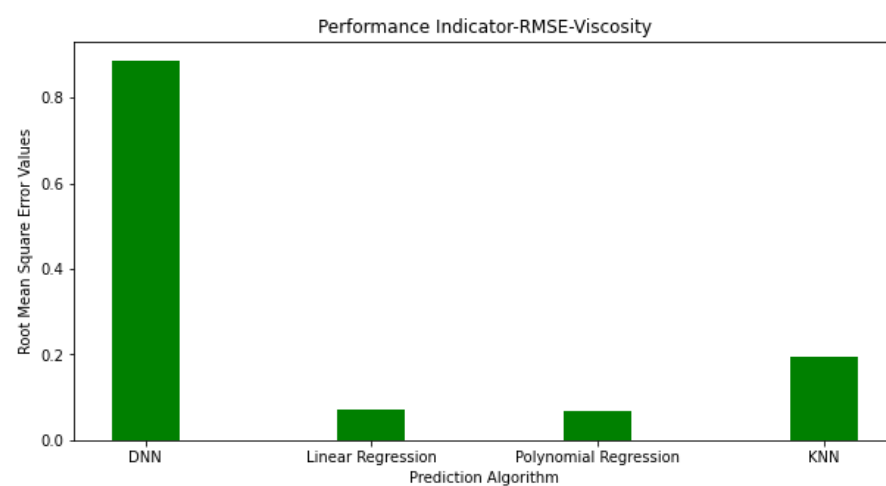
Figure 13. Performance indicator graph KOH: (a) Mean absolute error values of density; (b) mean square error values of density; (c) root mean square error values of density.



(a)

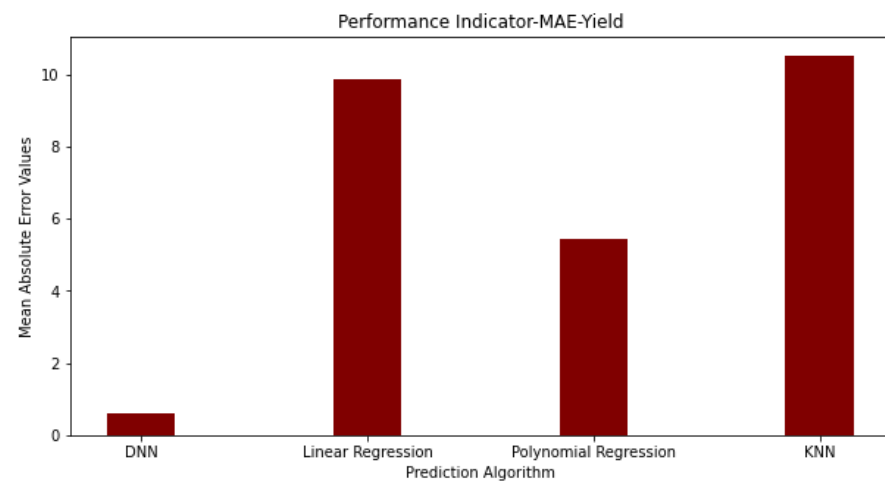


(b)

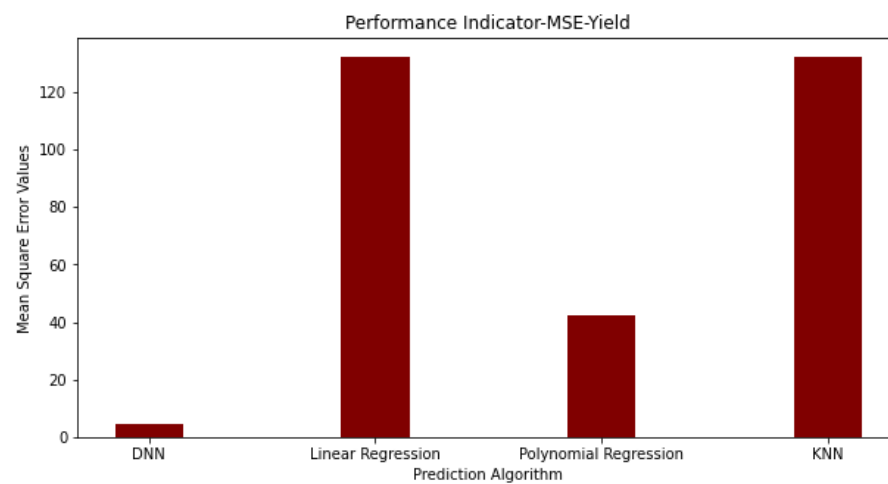


(c)

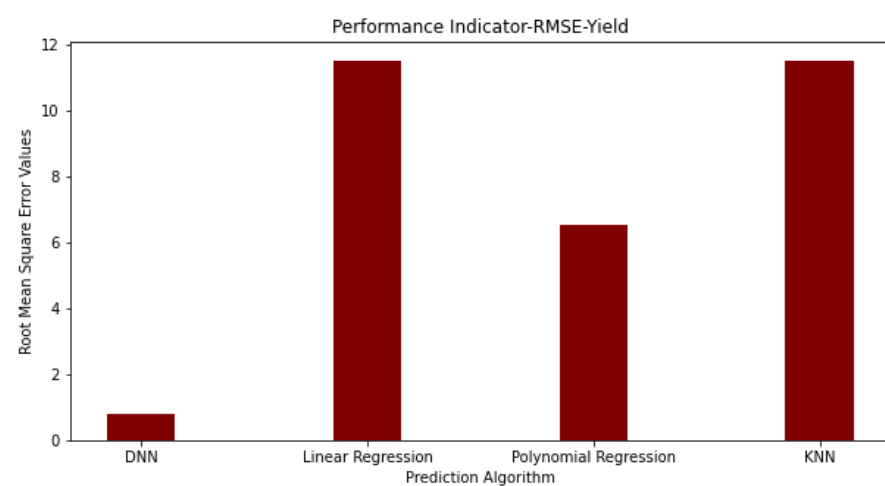
Figure 14. Performance indicator graph KOH: (a) Mean absolute error values of viscosity; (b) mean square error values of viscosity; (c) root mean square error values of viscosity.



(a)

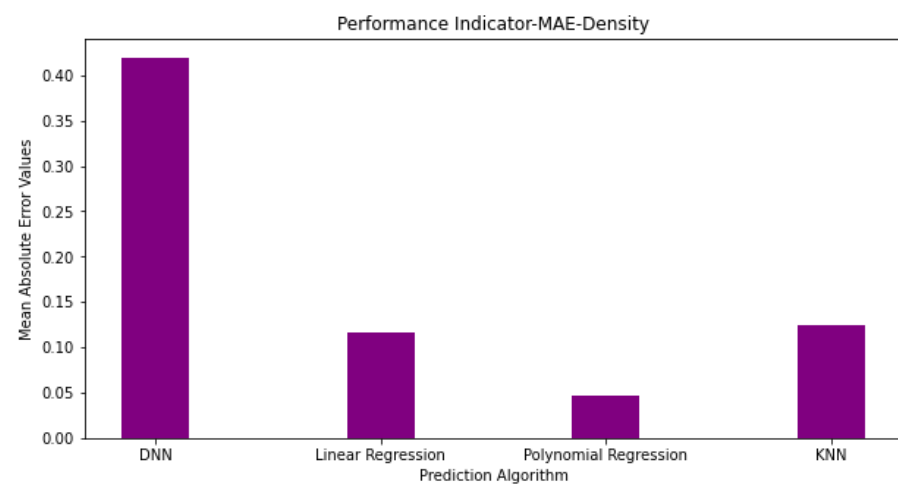


(b)

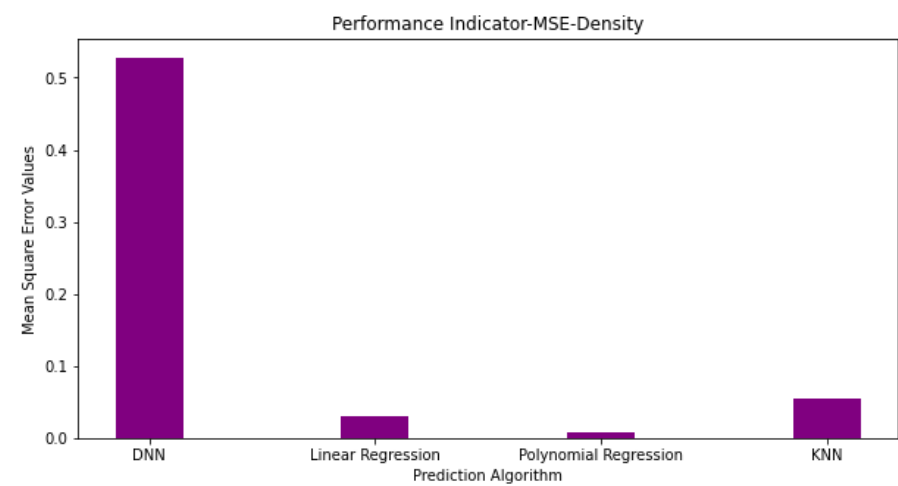


(c)

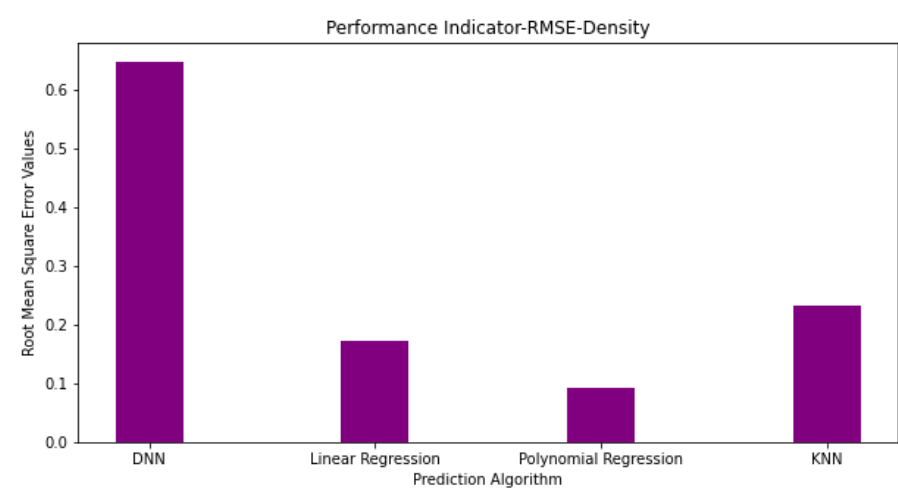
Figure 15. Performance indicator graph NaOH: (a) Mean absolute error values of yield; (b) mean square error values of yield; (c) root mean square error values of yield.



(a)

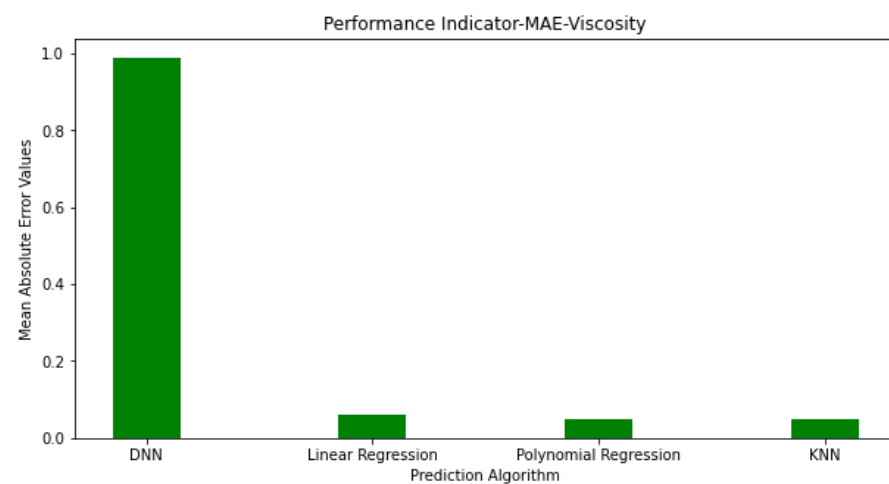


(b)

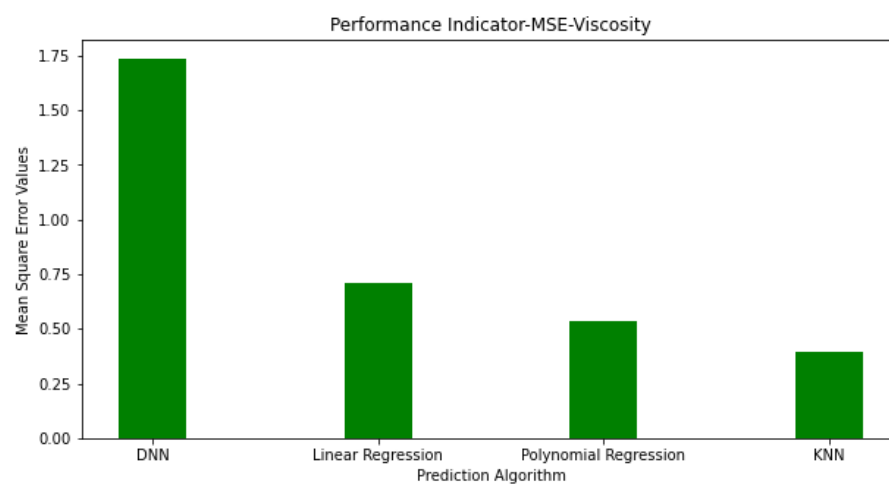


(c)

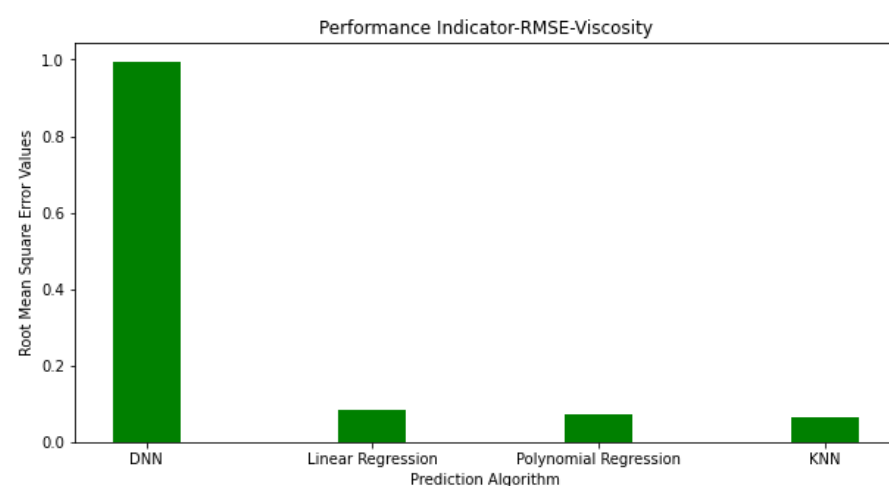
Figure 16. Performance indicator graph NaOH: (a) Mean absolute error values of density; (b) mean square error values of density; (c) root mean square error values of density.



(a)



(b)



(c)

Figure 17. Performance indicator graph NaOH: (a) Mean absolute error values of viscosity; (b) mean square error values of viscosity; (c) root mean square error values of viscosity.

4.5. Physicochemical Characterization of Biodiesel

ASTM D-6584 is the most important biodiesel analysis and standard test technique for monitoring the refining process. This outlines a GC-FID (gas chromatography with flame ionization detection) method for quantifying free and total glycerol in 100% methyl esters (B-100) biodiesel. The G.C. accurately determines the percent mass of free glycerol, mono, di, and triglycerides in methyl esters. Measuring the level of free glycerol and any unreacted mono-, di-, or triglycerides in biodiesel will indicate how efficiently the transesterification reaction is proceeding.

A detailed analysis giving the % composition of various components is given in Table 16.

Table 16. Concentration levels of the analyte components as seen in the chromatographic display.

Components Concentration	%
Glycerol	0.21
Butanetriol	1.11
C14 Tetradecanoic acid	0.56
C16 Hexadecanoic acid	22.94
C18 + 20, Octadecanoic acid + 9-Cis, 12-Cis Octadecadienoic acid + 9-Octadecenoic acid, (E)- (E)-9-Octadecenoic acid (E)-Oleic acid + Eicosanoic acid	61.5
C22 Docosonic (Behnic/Erusic)	1.19
Monopalmitin	1.58
Monolein + Monolinolenin + Monolinolein(Monglycerides)	2.48
Monostearin	0.39
Other: Diglycerides and Triglycerides	

GC analysis (Table 16) of the jatropha biodiesel shows that jatropha curcus methyl ester contains mostly C18 + C20 carbon-number fatty acids (stearic and arachidic) and C16 hexadecanoic acid. Other components: monglycerides at 2.48%, di and triglycerides.

4.6. Future Work and Recommendation

The developed model can be used to predict the optimal conditions for biodiesel synthesis from various feedstocks. Moreover, the sustainability of biodiesel production can be evaluated with a number of different tools. These include techno-economic, life cycle assessment, energy, and exergy assessments [33,34].

5. Conclusions and Prospects

The study investigates various parameters to obtain optimized conditions for the production of Jatropha biodiesel using NaOH and KOH as catalysts. A group of multi-objective regression models coupled with a deep neural network were implemented for the production of biodiesel properties. The statistical analysis enabled us to categorize the most relevant transesterification process parameters in terms of having the maximum impact on biodiesel properties. The empirical analysis depicted a positive correlation between temperature and duration and a negative correlation between methanol content and catalyst concentration. Moreover, the analysis stated that methanol concentration had the maximum positive correlation with the density and viscosity of the biodiesel produced. The PDP graphs depicted an almost linear curve between viscosity and duration, whereas the S-shaped curve between viscosity and amount of oil content. The analysis depicted the nature and extent of the correlation that exists between different transesterification process parameters. Mean square error (MSE), mean absolute error (MAE), and root mean square error (RMSE) were used as performance indicators for different prediction models. The deep neural network model proved to be more accurate in predicting yield and viscosity in comparison with linear regression, polynomial regression, and the KNN model. Whereas,

on the contrary, linear regression and polynomial regression proved to be better in terms of accuracy when predicting the density of the biodiesel produced. The same can be confirmed from Tables 10–13. Following the prediction phase, an evolutionary optimization technique in the form of a genetic algorithm is applied to optimize the transesterification process parameters based on desired biodiesel properties. The work presented in the paper signifies the applicability of AI in producing biodiesel with the desired set of properties and aligns itself with its contemporaries [35,36]. The GA-based optimization algorithm was able to predict transesterification process parameters for three different sets of values for yield, density, and viscosity of the biodiesel. The deep neural network exhibits favorable performance in simulating the biodiesel production process. The sustainability aspect of biodiesel production, which involves using environmentally friendly feedstocks and minimizing resource usage, is emphasized. The authors look forward to extending the existing work by incorporating data regarding biodiesel blends and simulating the biodiesel production process to produce more sustainable and economically viable forms of biodiesel.

Author Contributions: Conceptualization, A.K. and S.J.; methodology, A.K. and B.Y.L.; software, A.K., S.J., V.B. and D.B.; validation, A.K., B.Y.L., V.B. and D.B.; formal analysis, A.K., B.Y.L., V.B. and D.B.; investigation, A.K., B.Y.L., S.J. and V.B.; resources, A.K. and B.Y.L.; data curation, A.K., S.J., V.B. and D.B.; writing—original draft preparation, A.K., B.Y.L., S.J. and V.B.; writing—review and editing, V.B., D.B., V.P. and A.S.; visualization, S.J., V.B., V.P. and A.S.; supervision, V.B., D.B., S.J. and V.P.; project administration, A.K.; funding acquisition, A.K., V.B., D.B., V.P. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data can be made available on request.

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

Abbreviations

FFA	Free Fatty Acid
NaOH	Sodium Hydroxide
KOH	Potassium Hydroxide
FAME	Fatty Acid Methyl Ester
VCO	Virgin Castor Oil
WFO	Waste Frying Oil
CN	Cetane Number
CFPP	Cold Filter Plugging Point
GC	Gas Chromatography
GC-FID	Gas Chromatography with Flame Ionization Detection
AI	Artificial Intelligence
DNN	Deep Neural Network
GA	Genetic Algorithm
ANN	Artificial Neural Network
RMSE	Root Mean Squared Error
MAPE	Mean absolute percentage error
ANOVA	Analysis of variance
LR	Linear Regression
SVM	Support Vector Machine
KNN	k-nearest neighbours
PDP	Partial Dependence Plot
MSE	Mean Square Error
MAE	Mean Absolute Error
R ²	Coefficient of determination
PSO	Particle swarm optimisation

Appendix A

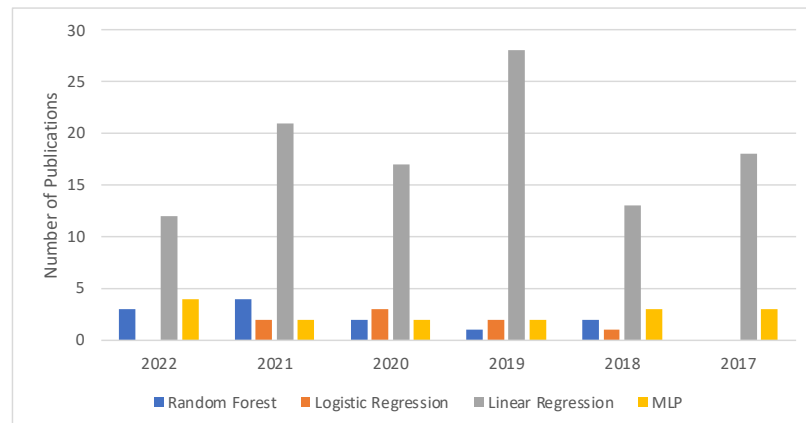


Figure A1. Year specific publication stats.

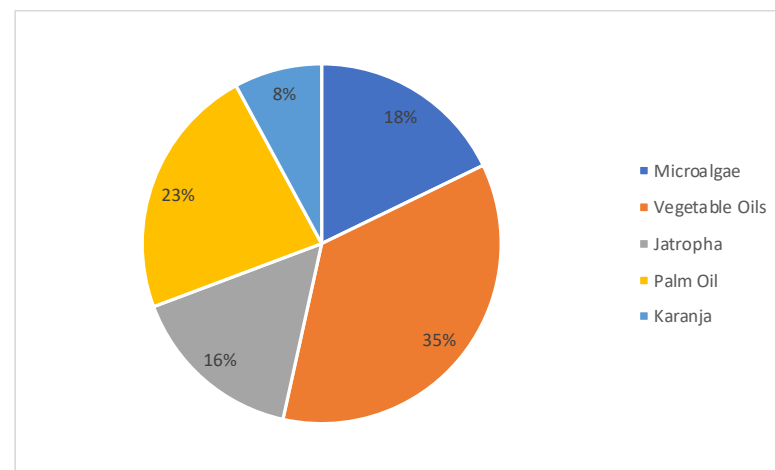


Figure A2. Feedstock based publication distribution.

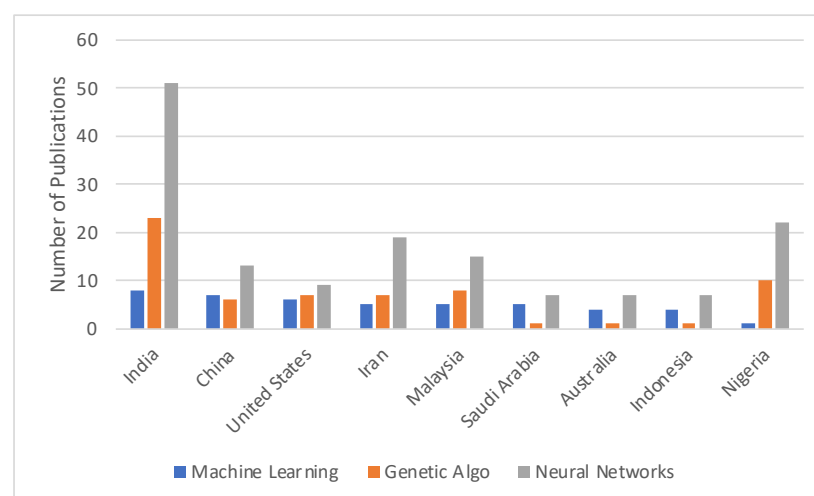


Figure A3. Region specific publication distribution.

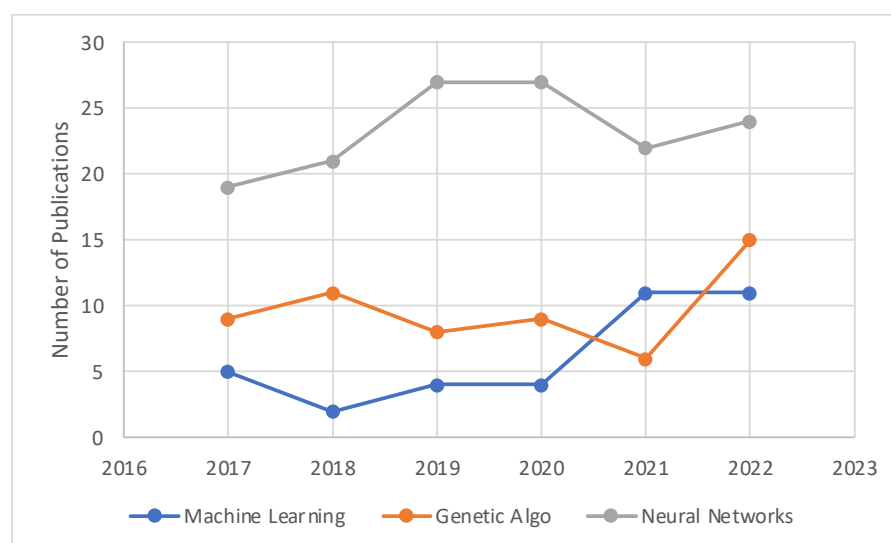


Figure A4. Publication trends for biodiesel production.

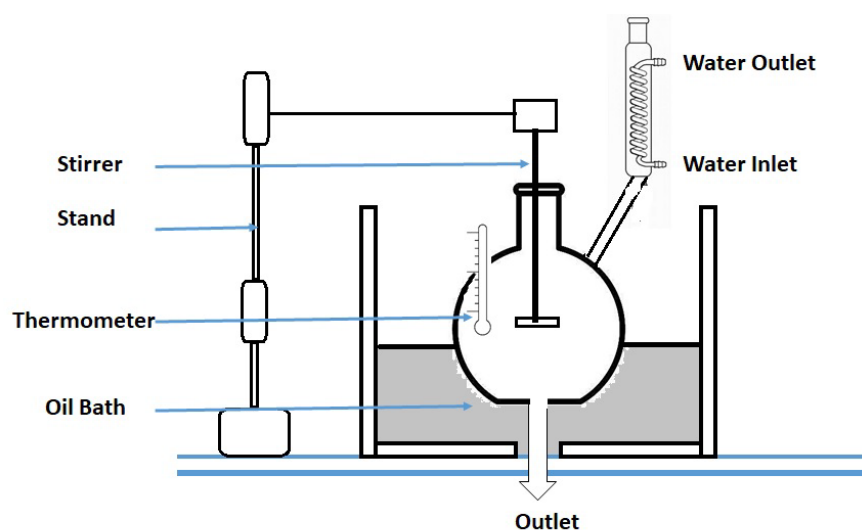


Figure A5. Schematic diagram of transesterification lab reactor.

References

1. Tan, D.; Wu, Y.; Lv, J.; Li, J.; Ou, X.; Meng, Y.; Lan, G.; Chen, Y.; Zhang, Z. Performance optimization of a diesel engine fueled with hydrogen/biodiesel with water addition based on the response surface methodology. *Energy* **2023**, *263*, 125869. [\[CrossRef\]](#)
2. Tan, D.; Meng, Y.; Tian, J.; Zhang, C.; Zhang, Z.; Yang, G.; Cui, S.; Hu, J.; Zhao, Z. Utilization of renewable and sustainable diesel/methanol/n-butanol (DMB) blends for reducing the engine emissions in a diesel engine with different pre-injection strategies. *Energy* **2023**, *269*, 126785. [\[CrossRef\]](#)
3. Wang, J.; Xia, A.; Deng, Z.; Huang, Y.; Zhu, X.; Zhu, X.; Liao, Q. Intensifying biofuel production using a novel bionic flow-induced peristaltic reactor: Biodiesel production as a case study. *Biofuel Res. J.* **2022**, *9*, 1721–1735. [\[CrossRef\]](#)
4. Liu, J.; Tao, B. Fractionation of fatty acid methyl esters via urea inclusion and its application to improve the low-temperature performance of biodiesel. *Biofuel Res. J.* **2022**, *9*, 1617–1629. [\[CrossRef\]](#)
5. Lamba, B.Y.; Joshi, G.; Rawat, D.S.; Jain, S.; Kumar, S. Study of oxidation behavior of Jatropha oil methyl esters and Karanja oil methyl esters blends with EURO-IV high speed diesel. *Renew. Energy Focus* **2018**, *27*, 59–66. [\[CrossRef\]](#)
6. Greener, J.G.; Kandathil, S.M.; Moffat, L.; Jones, D.T. A guide to machine learning for biologists. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 40–55. [\[CrossRef\]](#)
7. Aghbashlo, M.; Peng, W.; Tabatabaei, M.; Kalogirou, S.A.; Soltanian, S.; Hosseinzadeh-Bandbafha, H.; Mahian, O.; Lam, S.S. Machine learning technology in biodiesel research: A review. *Prog. Energy Combust. Sci.* **2021**, *85*, 100904. [\[CrossRef\]](#)

8. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [\[CrossRef\]](#)
9. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999–7019. [\[CrossRef\]](#)
10. Gawlikowski, J.; Tassi, C.R.; Ali, M.; Lee, J.; Humt, M.; Feng, J.; Kruspe, A.; Triebel, R.; Jung, P.; Roscher, R.; et al. A survey of uncertainty in deep neural networks. *arXiv* **2021**, arXiv:2107.03342.
11. Katoch, S.; Chauhan, S.S.; Kumar, V. A review on genetic algorithm: Past, present, and future. *Multimed. Tools Appl.* **2021**, *80*, 8091–8126. [\[CrossRef\]](#)
12. Lamba, B.Y.; Joshi, G.; Tiwari, A.K.; Rawat, D.S.; Mallick, S. Effect of antioxidants on physico-chemical properties of EURO-III HSD (high speed diesel) and Jatropha biodiesel blends. *Energy* **2013**, *60*, 222–229. [\[CrossRef\]](#)
13. Sun, K.; Lu, J.; Ma, L.; Han, Y.; Fu, Z.; Ding, J. Optimization of biodiesel synthesis from Jatropha curcas oil using kaolin derived zeolite Na-X as a catalyst. *Fuel* **2015**, *158*, 848–854. [\[CrossRef\]](#)
14. Joshi, G.; Lamba, B.Y.; Rawat, D.S.; Mallick, S.; Murthy, K.S.R. Evaluation of Additive Effects on Oxidation Stability of Jatropha Curcas Biodiesel Blends with Conventional Diesel Sold at Retail Outlets. *Ind. Eng. Chem. Res.* **2013**, *52*, 7586–7592. [\[CrossRef\]](#)
15. Peña, R.; Romero, R.; Martínez, S.L.; Natividad, R.; Ramírez, A. Characterization of KNO₃/NaX catalyst for sunflower oil transesterification. *Fuel* **2013**, *110*, 63–69. [\[CrossRef\]](#)
16. Neha, P.; Chintan, B.; Pallavi, D.; Neha, T. Use of Sunflower and Cottonseed Oil to prepare Biodiesel by catalyst assisted Transesterification. *Res. J. Chem. Sci.* **2013**, *3*, 42–47.
17. Lamba, B.Y.; Rawat, M.S.M.; Singh, H. Bio-diesel production from Jatropha curcas & Pongamia pinnata oil sea comparative study. *Dew* **2009**, *18*, 7.
18. Miraboutalebi, S.M.; Kazemi, P.; Bahrami, P. Fatty Acid Methyl Ester (FAME) composition used for estimation of biodiesel cetane number employing random forest and artificial neural networks: A new approach. *Fuel* **2016**, *166*, 143–151. [\[CrossRef\]](#)
19. Bemani, A.; Xiong, Q.; Baghban, A.; Habibzadeh, S.; Mohammadi, A.H.; Doranehgard, M.H. Modeling of cetane number of biodiesel from fatty acid methyl ester (FAME) information using GA-, PSO-, and HGAPSO- LSSVM models. *Renew. Energy* **2020**, *150*, 924–934. [\[CrossRef\]](#)
20. Arce, P.F.; Guimarães, D.H.P.; de Aguirre, L.R. Experimental data and prediction of the physical and chemical properties of biodiesel. *Chem. Eng. Commun.* **2019**, *206*, 1273–1285. [\[CrossRef\]](#)
21. Özgür, C.; Tosun, E. Prediction of density and kinematic viscosity of biodiesel by artificial neural networks. *Energy Sources Part A Recover. Util. Environ. Eff.* **2017**, *39*, 985–991. [\[CrossRef\]](#)
22. Yu, W.; Zhao, F. Prediction of critical properties of biodiesel fuels from FAMES compositions using intelligent genetic algorithm-based back propagation neural network. *Energy Sources Part A Recover. Util. Environ. Eff.* **2021**, *43*, 2063–2076. [\[CrossRef\]](#)
23. Kumar, S.; Jain, S.; Kumar, H. Prediction of jatropha-algae biodiesel blend oil yield with the application of artificial neural networks technique. *Energy Sources Part A Recover. Util. Environ. Eff.* **2019**, *41*, 1285–1295. [\[CrossRef\]](#)
24. Mairizal, A.Q.; Awad, S.; Priadi, C.R.; Hartono, D.M.; Moersidik, S.S.; Tazerout, M.; Andres, Y. Experimental study on the effects of feedstock on the properties of biodiesel using multiple linear regressions. *Renew. Energy* **2020**, *145*, 375–381. [\[CrossRef\]](#)
25. Kumar, S. Comparison of linear regression and artificial neural network technique for prediction of a soybean bio-diesel yield. *Energy Sources Part A Recover. Util. Environ. Eff.* **2020**, *42*, 1425–1435. [\[CrossRef\]](#)
26. Scopus Preview—Scopus—Welcome to Scopus. (n.d.). Available online: <https://www.scopus.com/> (accessed on 10 January 2023).
27. Rocabrundo-Valdés, C.; Ramírez-Verduzco, L.; Hernández, J. Artificial neural network models to predict density, dynamic viscosity, and cetane number of biodiesel. *Fuel* **2015**, *147*, 9–17. [\[CrossRef\]](#)
28. Moradi, G.; Dehghani, S.; Khosravian, F.; Arjmandzadeh, A. The optimized operational conditions for biodiesel production from soybean oil and application of artificial neural networks for estimation of the biodiesel yield. *Renew. Energy* **2013**, *50*, 915–920. [\[CrossRef\]](#)
29. Banerjee, A.; Varshney, D.; Kumar, S.; Chaudhary, P.; Gupta, V.K. Biodiesel production from castor oil: ANN modeling and kinetic parameter estimation. *Int. J. Ind. Chem.* **2017**, *8*, 253–262. [\[CrossRef\]](#)
30. Ighose, B.O.; Adeleke, I.A.; Damos, M.; Junaid, H.A.; Okpalaeke, K.E.; Betiku, E. Optimization of biodiesel production from Thevetia peruviana seed oil by adaptive neuro-fuzzy inference system coupled with genetic algorithm and response surface methodology. *Energy Convers. Manag.* **2017**, *132*, 231–240. [\[CrossRef\]](#)
31. Corral Bobadilla, M.; Fernández Martínez, R.; Lostado Lorza, R.; Somovilla Gómez, F.; Vergara González, E.P. Optimizing Biodiesel Production from Waste Cooking Oil Using Genetic Algorithm-Based Support Vector Machines. *Energies* **2018**, *11*, 2995. [\[CrossRef\]](#)
32. Cheng, M.-Y.; Prayogo, D.; Ju, Y.-H.; Wu, Y.-W.; Sutanto, S. Optimizing Mixture Properties of Biodiesel Production Using Genetic Algorithm-Based Evolutionary Support Vector Machine. *Int. J. Green Energy* **2016**, *13*, 1599–1607. [\[CrossRef\]](#)
33. Aghbashlo, M.; Hosseinzadeh-Bandbafha, H.; Shahbeik, H.; Tabatabaei, M. The role of sustainability assessment tools in realizing bioenergy and bioproduct systems. *Biofuel Res. J.* **2022**, *35*, 1697–1706. [\[CrossRef\]](#)
34. Hosseinzadeh-Bandbafha, H.; Nizami, A.-S.; Kalogirou, S.A.; Gupta, V.K.; Park, Y.-K.; Fallahi, A.; Sulaiman, A.; Ranjbari, M.; Rahnama, H.; Aghbashlo, M.; et al. Environmental life cycle assessment of biodiesel production from waste cooking oil: A systematic review. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112411. [\[CrossRef\]](#)

35. Ahmad, J.; Awais, M.; Rashid, U.; Ngamcharussrivichai, C.; Naqvi, S.R.; Ali, I. A systematic and critical review on effective utilization of artificial intelligence for bio-diesel production techniques. *Fuel* **2023**, *338*, 127379. [[CrossRef](#)]
36. Pahwa, M.S.; Dadhich, M.; Saini, J.S.; Saini, D.K. Use of Artificial Intelligence (AI) in the Optimization of Production of Biodiesel Energy. In *Artificial Intelligence for Renewable Energy Systems*; Scrivener Publishing LLC.: Salem, MA, USA, 2022; pp. 229–238. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.