Thermal conductivity of Water-based nanofluids: Prediction and comparison of models using machine learning

ABSTRACT

Statistical methods, and especially machine learning, have been increasingly used in nanofluid modeling. This paper presents some of the interesting and applicable methods for thermal conductivity prediction and compares them with each other according to results and errors that are defined. The thermal conductivity of nanofluids increases with the volume fraction and temperature. Machine learning models were proposed to represent the thermal conductivity as a function based on the temperature, nanoparticles volume fraction and the thermal conductivity of the nanoparticles. The results of models were in appropriate agreement with the experimental data. This work represents 8 machine learning models for the predicting the thermal conductivity of water-based nanofluids. The models have been trained and tested on two separate sets of data. Three metrics have been employed to evaluate the performance of the models. The best method for each system is selected using results.

Keywords: Nanofluids; Modeling; Machine learning; Thermal conductivity; Prediction.

INTRODUCTION

Interest in nonmaterial has increased in recent years because of their unique physical and chemical properties [1]. Heating or cooling fluids is significant for many industrial segments, including energy supply and generation, transportation and electronics. The thermal conductivity of these fluids plays a critical role in the improvement of energy-efficient heat transfer apparatus [2]. The plan of escalating thermal conductivity of fluids with conducting particles suspended on them is not novel [3]. The idea was first initiated by a series of research works at the Argonne National Laboratory and probably Choi [4] was the first to call the fluids with particles of nanometer size suspended in them as “nano-fluids,” which has obtained amicability. The word nanofluid refers to the suspension of nanometric size (the limit is usually established in 100 nm) particles in any base fluid [5].
Suspensions containing a small amount of nanoparticles have been shown to have higher thermal conductivities compared with the base fluids [2, 6-9]. The increased thermal conductivity of nanoparticles suspension appears to be highly dependent on the densities of the particle and the thermal conductivity of the base fluid [6, 7]. Recently, researchers have shown that nanofluids including CuO or Al₂O₃ nanoparticles in water or ethylene glycol demonstrate improved thermal conductivity [6]. There are two different approaches to investigate the enhanced heat transfer of the suspensions: the two-phase and the single-phase one. The first provides the possibility of understanding the functions of both the fluid phase and the solid particle in the heat transfer process, but needs much computation time and computer capacity [9]. By combining Lagrangian statistics and direct numerical simulation (DNS) this approach applied to analyze the mechanism of two-phase heat and turbulent transport by solid particles (on the micrometer order) suspended in a gas flow [10]. The mixture model, based on a single fluid two phase approach assumes that the coupling between phases is strong, and particles closely follow the flow. The two phases are assumed to be interpenetrating, meaning that each phase has its own velocity vector field and slip velocity (relative velocity) is defined as the velocity of the nanoparticles phase in relation to the velocity of the base fluid phase. Within any control volume there is a volume fraction of primary phase and also a volume fraction of the secondary phase. This model is employed in the simulation of nanofluids [11]. This model is more consistent with the experimental results. The second approach assumes that both the fluid phase and particles are in a thermal equilibrium condition and flow at the same velocity. This approach which is based on the single phase flow is simpler and takes less computation time. In cases that the main interest is focused on heat transfer calculations, this approach may be more suitable [9]. Homogenous model is one of the nanofluid single phase models. This model differs from conventional pure fluid model only in the effective properties. It means that the continuity, Navier–Stokes and the energy equations are used with the nanofluid effective properties. According to this model, usual correlations of flow and heat transfer feature of a single phase fluid can be generalized on the nanofluid. Nanofluid studied based on this model which it is not able to predict the heat transfer features of nanofluids [12-14]. The present study aims to drive the mathematical model of the dispersion mechanism, and shows that this model has the ability to predict the nanofluid heat transfer more accurately. Microscopic models are interesting to describe the effects of interactions between the nanoparticles and liquid particles. One such model, based on the Monte Carlo method, was successfully applied, in combination with fractal theory, to predict the effective thermal conductivity of nanofluids. A model proposed which is a combination of Monte Carlo simulations and the fractal geometry theory [15]. The predictions have shown appropriate agreement with the existing experimental data. The comparison of different approaches for numerical modeling on heat transfer of nanofluids was presented [16, 17]. The correlations for temperature dependent effective conductivity of water-based nanofluids were developed [18, 19]. A new model for assessing the effective viscosity of water-based nanofluids was developed [20]. Also, by knowing the inter-atomic potentials, computational tools consisting of the molecular dynamic simulation (MD) and the Monte Carlo (MC) methods have been employed to model the transport properties of CNTs [21-27]. MD simulation revealed that isolated SWCNTs had a very similar thermal conductivity as those of a hypothetical isolated graphene sheet with the same number of atoms at certain temperatures [28]. In another study, the dependence of the thermal conductivity of a nanotube on its structure, defects, diameter and chirality was investigated [29, 30]. Much work has been devoted to the investigation on the enhanced thermal properties of nanofluid and its enhancement mechanisms. The thermal conductivity of copper nanoparticles measured in ethylene glycol and found that the increase in the thermal conductivity was twice the value predicted by the Maxwell effective medium theory [31]. The enhancement of the water–platinum nanofluids thermal conductivity estimated based on molecular dynamics simulation model and compared the results with the existing experimental results that indicated great enhancement in the thermal conductivity [32]. Four possible explanations suggested for the anomalous enhancement in the thermal conductivity of nanofluid and showed that the key factor in the understanding of the thermal
The property of nanofluid is ballistic, rather than diffusive [33].

The aim of this paper is to develop some models using machine learning to predict the thermal conductivity of water-based nano-fluids, based on the data coming from other studies. Then comparison between methods can be made based on errors that are defined.

EXPERIMENTAL

Machine Learning

In this work, functions are used for predicting thermal conductivity of nanofluid. Some functions are compatible with our data type, with default parameters. In the case of this problem, the compatible functions for predictions are used.

Algorithms

- Zero Regression (ZeroR)

  Pseudo-regression technique always makes models with cross-validation coefficient. In the structure of this method the value of a property is usually predicted to be equivalent to its average value on the training set. This method predicts the mean for a numeric class and the mode for a nominal class. ZeroR is typically employed as a reference point for comparison with other regression techniques.

- Linear regression

  Linear regression is a type of regression method in which experimental data is used and modeling is performed by a function which is a linear combination of the model parameters. This combination depends on independent variables, the linear regression model that we have used is based on the Akaike criterion for model selection (AIC) [34], based on the Kullback–Leibler information between two densities, related to the true model and fitted model.

- Least median square regression

  One type of regression method is a least median squared linear regression algorithm that uses the linear regression class to form predictions. The functions of least squared regression method are produced from random samples of the data. The least squared regression with the minimum median squared error is selected as the ultimate model. The algorithm is based on the work of Rousseeuw and Leroy [35].

- Support vector machine regression

  Support vector machines (SVMs) are a system of related managed learning methods used for regression and classification. Data is viewed as two sets of vectors in an n-dimensional space; an SVM will build a segregated hyperplane in that space, one which makes the most of margin between the two data sets. For calculating of the margin, two parallel hyperplanes are constructed, one on each side of the segregated hyperplane, which are “pushed up against” the two data sets. Naturally, on excellent separation is achieved by the hyperplane that has the greatest distance from the adjacent data points of both sets, because in general the greater the margin the better the generalization error of the classifier. The parameters can be learned by means of different algorithms. The algorithm is chosen by setting the RegOptimizer. The most interesting algorithm (RegSMOImproved), is due to Shevade, et al. and used as the default RegOptimizer [36]. The benefit of SVM regression models is their admirable general prediction accuracy.

- IBK

  The k-nearest neighbor’s algorithm (KNN) is a regression technique that classifies objects according to closest training examples in the feature space. It is a kind of instance-based learning, or lazy learning where the function is only approximated locally and all computational operation is delayed until regression. The KNN method is used for regression by simply assigning the property value for the item to be the average of the values of its k nearest neighbors. It is helpful to weigh the contributions of the neighbors; therefore the nearer neighbors give more to the average than the neighbors with more distance. The objects are characterized by position vectors in a multidimensional feature space, In order to identify neighbors. In the testing phase, the test example is represented like a vector in the feature space. Distances from this vector from all stored vectors are calculated and the k closest samples are chosen to find out the actual magnitude of the test case. This algorithm is sensitive to the organization of the data. The best choice of k is related to data;
usually, larger values of $k$ decrease the influence of noise. Heuristic techniques like cross-validation can help to set a good $k$.

- **Multilayered perceptron**
  The multilayered Perceptron Artificial Neural Network is a type of machine learning methods. We used the Back propagation algorithm with a learning rate equal to 0.3[37, 40]. All the neurons had a sigmoid activation function. A momentum of 0.1 progressively decreasing until 0.0001 has been used to escape local minima on the error surface.

- **M5P**
  M5P [38] is a method of regression that combines a conventional decision tree with the possibility of linear regression functions at the nodes. A decision-tree induction algorithm is used to build a tree initially; a splitting criterion is used instead of maximizing the information gain at each inner node. This procedure minimizes the intra-subset variation in the class values down each branch. Sharp discontinuities between the subtrees are harmful, so a smoothing procedure is used. This method combines the leaf model prediction with every node along the path back to the root, smoothing it at each of nodes by combining it with the predicted value from the linear model. Methods developed by Breiman et al. [39] for their CART system are adopted. All enumerated attributes are turned into binary variables therefore all the splits in M5P are binary. As to missing values, M5P applies a method of “surrogate splitting” that gets another attribute to split instead of place of the original one and employs it instead. In training part, M5P applies as surrogate attribute the class magnitude believing that this is the attribute that should be associated with the one used for splitting. At the end of splitting procedure, all missing values are changed by the average values of the corresponding attributes of the training example. In the testing part the average value of that attribute for all training instances that reach the node are used instead of an unknown attribute value. M5P produces compact and relatively comprehensible models.

- **Regression by discretization**
  Regression by Discretization is a regression scheme that uses each classifier on a copy of the data that has the class attribute discredited. The expected value of the mean class value for each discredited interval is the predicted value. This class supports conditional density estimation by constructing a univariate density estimator from the target values in the training data. Weight of training data is determined by the class probabilities. Some of metrics are employed to evaluate the performance of the models explained.

**Evaluation Metrics**

Some of metrics are employed to evaluate the performance of the models. In the next part these metrics are explained.

- **Correlation Coefficient**
  The correlation coefficient is a measure of how trends in actual values are followed by well trends in the predicted values. It is an evaluation of how well the predicted values from a predicted model fit the real-life data. The correlation coefficient is a magnitude in the range of -1 and 1. If the predicted values and the actual values are independent and no relationship is between them, the correlation coefficient is close to 0. If the strength of the relationship between the actual values and predicted values increases, so does the correlation coefficient. An ideal fit gives a coefficient of 1.0. Opposite but correlated trends result in a correlation coefficient magnitude limit to -1. Negative correlation values are not typically expected in the learning of a predictive model.

- **Mean absolute error**
  The mean absolute error averages the value of the every error without considers their sign. Mean-squared error tends to exaggerate the effect of outliers, but absolute error does not have this performance: all values of error are treated evenly according to their magnitude.

**Root Mean Squared Error**

A type of predictive regression model is the mean squared error (MSE) that is a different way to quantify the distinction between set of actual (target) values, $x_t$ and set of predicted values, $x_p$. The root mean squared error (RMSE) is defined as: the mean absolute error averages the value of the every error without considering their sign. Mean-squared error tends to exaggerate the effect
of outliers, but absolute error does not have this performance: all values of error are treated evenly according to their magnitude.

$$\text{RSME}(\mathbf{x}_p, \mathbf{x}_i) = \sqrt{\text{MSE}(\mathbf{x}_p, \mathbf{x}_i)} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{pi} - x_{ii})^2}$$

RESULTS AND DISCUSSION

Error analysis

The experimental data used for the training of models has been previously achieved and fully described in Ref. [40]. The error analysis for different methods is listed in Tables 1-3. Table 1 shows the analysis of correlation coefficient for methods. Ideal value of correlation coefficient is equal to 1. Nearer to 1 shows better agreement between predicted and experimental data and nearer to 0 shows less agreement between them. As Table 1 show, best value of correlation coefficient for Al₂O₃ data set is obtained from KNN method. Other method also indicates good prediction and negligible error. For CuO data set SMOreg shows the best result in comparison to other methods. In general correlation coefficient for Al₂O₃ data set is better than CuO data set. This difference can be a consequent of data features like number of data and accuracy of them.

Table 1(a, b). Analyzing of Correlation Coefficient

<table>
<thead>
<tr>
<th>Dataset</th>
<th>ZeroR</th>
<th>Linear Regression</th>
<th>LeastMedSq</th>
<th>SMOreg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al₂O₃</td>
<td>0.00</td>
<td>0.94</td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td>CuO</td>
<td>0.00</td>
<td>0.89</td>
<td>0.89</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Table 2(a, b) indicates the mean absolute error for methods based on two data set of CuO and Al₂O₃. This error is the type like thermal conductivity. Zero is the best value for mean absolute error and more limits to zero indicates more accuracy of predictions. Table 2(a, b) shows good magnitudes of this error for all methods generally. Among these methods most convenient value is derived from KNN method like correlation coefficient. Best value of mean absolute error for Cu data set is analyzed and KNN method has most accurate prediction.

Table 2(a, b). Analyzing of Mean Absolute Error

<table>
<thead>
<tr>
<th>Dataset</th>
<th>ZeroR</th>
<th>Linear Regression</th>
<th>LeastMedSq</th>
<th>SMOreg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al₂O₃</td>
<td>0.0281</td>
<td>0.0098</td>
<td>0.0103</td>
<td>0.0100</td>
</tr>
<tr>
<td>CuO</td>
<td>0.0179</td>
<td>0.0088</td>
<td>0.0088</td>
<td>0.0086</td>
</tr>
</tbody>
</table>

Table 3(a, b) shows the results of root mean squared error for methods based on two data set of CuO and Al₂O₃. Root mean squared error is the type of thermal conductivity like mean absolute error and ideal value for this error zero. KNN method has the best prediction for Al₂O₃ dataset according to this error and results that were shown in the Table 2(a, b). These are interesting results by agreement among three error analyses because KNN is the best method in three error analysis. M5P shows the best result for CuO dataset according to Table 3(a, b).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>IBK</th>
<th>Multilayer Perceptron</th>
<th>M5P</th>
<th>Regression By Discretization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al₂O₃</td>
<td>0.0085</td>
<td>0.0115</td>
<td>0.0096</td>
<td>0.0091</td>
</tr>
<tr>
<td>CuO</td>
<td>0.0082</td>
<td>0.0120</td>
<td>0.0086</td>
<td>0.0095</td>
</tr>
</tbody>
</table>
Table 3(a, b). Analyzing of Root Mean Squared Error

<table>
<thead>
<tr>
<th>Dataset</th>
<th>ZeroR</th>
<th>Linear Regression</th>
<th>LeastMedS</th>
<th>SMOreg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al₂O₃</td>
<td>0.0344</td>
<td>0.0121</td>
<td>0.0127</td>
<td>0.0123</td>
</tr>
<tr>
<td>CuO</td>
<td>0.0231</td>
<td>0.0112</td>
<td>0.0111</td>
<td>0.0109</td>
</tr>
</tbody>
</table>

Result of Al₂O₃ dataset

Figure 1(a-d) presents the results predicted for the Al₂O₃ as a function of the volume fraction of nanoparticles and temperature based on the KNN model. The KNN model is able to account for the increase in thermal conductivity for low nanoparticles volume fractions and the higher nanoparticles volume fractions. The KNN model also accounts very well for the change of temperature. The predicted values were calculated at several volume fractions to ensure the model was able to interpolate well between experimental points. In addition, the training of the model was performed with experimental data obtained at temperatures between 20 and 45°C and the network was validated for some experimental data obtained at this range. Notably, Phi is volume fraction, K is thermal conductivity experimental data and K’ is thermal conductivity prediction.

Result of CuO dataset

Figure 2(a, b) present the results of prediction using KNN model. The KNN model is employed to show the results as the example of machine learning methods. The model is able to account for the variation in thermal conductivity for low or high nanoparticles volume fractions. The KNN model also accounts very well for the temperature variations. The modeling was calculated at several volume fractions to ensure the model was able to predict the experimental data. The results are shown in two volume fractions as examples.
CONCLUSIONS

Machine Learning methods; have been employed for predicting the value of the thermal conductivity of water-based nanofluids. This application is important because the ability of correctly predicting this value could help to select the best model. Thermal conductivity of Al₂O₃ and CuO nanofluids was studied statistically. The thermal conductivity of nanofluids shows significant enhancement with nanoparticles volume fraction. At higher nanoparticles volume fractions, the thermal conductivities of CuO and Al₂O₃ nanofluids are higher than the base fluid. The reason of this phenomenon can be the reduction of interparticle distances and the creation of chain-like structures of nanoparticles. The thermal conductivity of the base fluid also increases with increase in temperature.

REFERENCES

containing nanosized sic particles.Journal of Thermophysics, 23, 571.


properties of crystalline b-SiC II. Thermal conductivity and effects of point defects. *Journal of Nuclear Materials*, 255, 139-152.


