International Journal of Heat and Mass Transfer 178 (2021) 121607

Contents lists available at ScienceDirect

International Journal of Heat and Mass Transfer

journal homepage: www.elsevier.com/locate/hmt

Machine learning algorithms to predict flow boiling pressure drop in mini/micro-channels based on universal consolidated data

Yue Qiu[a], Deepak Garg[b], Sung-Min Kim[c], Issam Mudawar[d], Chirag R. Kharangate[a,*]

[a] Department of Mechanical and Aerospace Engineering, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
[b] Mechanical and Aerospace Engineering Department, University of California, Los Angeles, CA 90095, USA
[c] School of Mechanical Engineering, Sungkyunkwan University, 300 Cheoncheon-dong, Suwon 16419, Republic of Korea
[d] School of Mechanical Engineering, 585 Purdue Mall, West Lafayette, IN 47907-2088, USA

ARTICLE INFO

Article history:
Received 9 March 2021
Revised 11 June 2021
Accepted 14 June 2021
Available online 9 July 2021

Keywords:
Machine learning
Neural networks
Ann
XGBoost
KNN
Light GBM
Flow boiling
Pressure drop

ABSTRACT

Two-phase flow in mini/micro-channels can meet the high heat dissipation requirements of many state-of-the-art cooling solutions. However, there is lack of accurate universal methods for predicting parameters like pressure drop in these configurations. Conventional ways of predicting pressure drop employ either Homogeneous Equilibrium Model (HEM) or semi-empirical correlations. This current study leverages the availability of data collected over the past few decades to build several machine learning models to demonstrate the efficacy and ease of building and deploying such models. A consolidated database of 2787 data points for flow boiling pressure drop in mini/micro-channels is amassed from 21 sources that includes 10 working fluid, reduced pressures of 0.0006 –0.7766, hydraulic diameters of 0.15–5.35 mm, mass velocities of $33.1 < G < 2738$ kg/m 2 s, liquid-only Reynolds numbers of 14–27,658, superficial vapor Reynolds number of 75.58–199,453 and flow qualities of 0 and 1. This consolidated database is utilized to develop four machine learning based regression models viz., Artificial Neural Networks (ANN), KNN regression, Extreme Gradient Boosting (XGBoost) and Light GBM. Both input parameters and hyperparameters are optimized for the individual models. The models with dimensionless input parameters: Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go},We_{tp} predict the test data for ANN model, XGBoost model, KNN model, and LightGBM model with MAEs of 9.58%, 10.38%, 13.52%, and 14.49%, respectively. The optimized machine-learning models performed better than highly reliable generalized pressure drop correlations plus showed good performance across individual datasets, flow regimes, and channel configurations.

1. Introduction

1.1. Flow boiling in mini/micro-channel

Flow boiling studies has received more and more interest in the last few decades, due to the increasing demand for innovative compact systems in aerospace, avionics, consumer electronics and electrical vehicles [1]. By utilizing coolant's latent heat, two-phase cooling is able to achieve orders-of-magnitude enhancement in the heat transfer coefficient. This attribute makes two-phase cooling a lucrative alternative to conventional single phase cooling systems [2]. Flow boiling in mini/micro-channels shows the most popularity among a variety of two-phase cooling configurations, due to simplicity in structures, high surface to volume ratio, low coolant content, and relative uniform surface temperature [3]. However, the high pressure drops caused by enhanced surface area per mass velocity seen in miniaturized channels is the major disadvantage of mini/micro channels. The high-pressure drop increases the power consumption, and can critically affect the system by flashing or even choking the two-phase flow system [4]. This necessitates an accurate predictive tool for pressure drop for reliable and robust two-phase flow based thermal system designs.

To investigate the physics behind flow boiling pressure drop, multitude of experiments have been conducted by researchers worldwide. Qu and Mudawar [5] measured flow boiling pressure drop in a two-phase micro-channel heat sink and found that pressure drop increases appreciably upon commencement of boiling in the micro-channels. The upstream flow pattern oscillates between the slug and annular flow, and downstream is predominantly annular. Copetti et al. [6] performed an experimental study of flow boiling with R134a in a horizontal tube of 2.6 mm inner diameter. They analyzed the pressure drop under the variation of dif-

* Corresponding author.
E-mail address: chirag.kharangate@case.edu (C.R. Kharangate).

https://doi.org/10.1016/j.ijheatmasstransfer.2021.121607
0017-9310/

Nomenclature

Bd	bond number, $Bd = g(\rho_f - \rho_g)D_h^2/\sigma$
Bo	boiling number, $Bo = q''_{;H}/Gh_{fg}$
Co	convection number, $Co = [(1-x)/x]^{0.8}(\rho_g/\rho_f)^{0.5}$
c_p	specific heat at constant pressure
c_v	specific heat at constant volume
D_h	hydraulic diameter of flow channel
Fr_f	saturated liquid Froude number, $Fr_f = [G(1-x)]^2/(\rho_f^2 g D_h)$
Fr_g	saturated vapor Froude number, $Fr_g = (Gx)^2/(\rho_g^2 g D_h)$
Fr_{fo}	liquid-only Froude number, $Fr_{fo} = G^2/(\rho_f^2 g D_h)$
Fr_{go}	vapor-only Froude number, $Fr_{go} = G^2/(\rho_g^2 g D_h)$
F_f	fluid-dependent parameter
$\hat{f}$	KNN regression function
G	mass velocity
g	gravity acceleration; activation function
h	heat transfer coefficient
h_{fg}	latent heat of vaporization
k	liquid conductivity
K	number of nearest neighbors
L	training error
MAE	mean absolute error
MSE	mean square error
N_o	set of K-nearest training observations
p	data point vector
P	pressure
P_c	critical pressure
Pe_f	saturated liquid Peclet number, $Pe_f = Re_f\ Pr_f$
Pe_g	saturated vapor Peclet number, $Pe_g = Re_g\ Pr_g$
P_F	wetted perimeter of channel
pH	heated perimeter of channel
P_R	reduced pressure, $P_R = P_c/P$
Pr_f	saturated liquid Prandtl number, $Pr_f = \mu_f c_{pf}/k_f$
Pr_g	saturated vapor Prandtl number, $Pr_g = \mu_g c_{pg}/k_g$
q	data point vector
q''	heat flux
q''_H	heat flux based on heated perimeter of channel
R	relative roughness, $R = e/D_h$
RSS	residual sum of squares
R^2	coefficient of determination
Re_f	saturated liquid Reynolds number, $Re_f = G(1-x)D_h/\mu_f$
Re_g	saturated vapor Reynolds number, $Re_g = GxD_h/\mu_g$
Re_{fo}	liquid-only Reynolds number, $Re_{fo} = GD_h/\mu_f$
Re_{go}	vapor-only Reynolds number, $Re_{go} = GD_h/\mu_g$
S	output, suppression factor
Su_f	saturated liquid Suratman number,
Su_g	saturated vapor Suratman number,
T	temperature
We_f	saturated liquid Weber number, $We_f = [G(1-x)]^2 D_h/(\rho_f \sigma)$
We_g	saturated vapor Weber number, $We_g = (Gx)^2 D_h/(\rho_g \sigma)$
We_{fo}	liquid-only Weber number, $We_{fo} = G^2 D_h/(\rho_f \sigma)$
We_{go}	vapor-only Weber number, $We_{go} = G^2 D_h/(\rho_g \sigma)$
Xtt	Lockhart-Martinelli parameter, $Xtt = (\mu_f/\mu_g)^{0.1}[(1-x)/x]^{0.9}(\rho_g/\rho_f)^{0.5}$
Y	response variable
$\hat{Y}$	fitted values

Greek Symbols

α	vapor void fraction; coefficient
β	aspect ratio, exponential decay rate
ε	percentage data predicted within ±50%,
θ	percentage data predicted within ±30%,
μ	dynamic viscosity, mean value
λ	learning rate
ρ	density
σ	surface tension; standard deviation

Subscripts

c	critical
exp	experimental
f	saturated liquid, fluid
fo	liquid only
g	saturated vapor
go	vapor only
l	liquid
n	number of data points
pred	predicted
tp	two phase
w	heated wall; inner wall

ferent parameters and reported that low vapor quality and mass velocity resulted in less frictional pressure drops while lower saturation temperatures slightly increased the pressure drop. Saturated flow boiling characteristic of deionized water in square microchannels were investigated by Markal et al. [7]. Their experiment was performed under different mass velocity and heat flux conditions. They concluded that total pressure drop increased with an increase in heat flux and exit vapor quality. Recently, Yin et al. [8] explored flow boiling pressure drop with deionized water in open microchannel. This innovative configuration can enhance microchannel performance with a small pressure drop penalty. They concurred that two-phase pressure drop increases with increase in heat flux. However, after the onset of stratified flow, pressure drop remains almost constant for large mass fluxes and decreases for smaller mass fluxes.

1.2. Predicting flow boiling pressure drop

The most widely used approaches of predicting flow boiling pressure drop can be classified into two categories, the homogeneous equilibrium models (HEM) and the separated flow models (SFM). The homogeneous model assumes that the two-phase mixture is well dispersed and behaves as a single pseudo fluid with mean properties and only latent heat transport between the phases [9–12]. The two-phase friction factor and viscosity are two key unknowns of homogeneous equilibrium model. The separated flow models, which permit differences between phase velocities, are mostly based on the model developed by Lockhart and Martinelli [13]. This model shows high accuracy for flow boiling in macrochannels, but is less accurate for mini/micro-channels. Based on Lockhart and Martinelli's pioneering work, some researchers have published their own models to improve prediction for mini/microchannels [14–17]. Lee and Lee [18] used the Lockhart-Martinelli type correlation with modified C parameter which took into account the flow rate and the channel size. Their correlations predicted the experiment data points with MAE less than 10%, and they also achieved MAEs of 15% and 20% in predicting data point from Wambsganss et al. [19] and Mishima and Hibiki [20], respectively. A modification of the Chisholm correlation [16], which is based on Lockhart-Martinelli but with an interfacial shear term, was developed by Yu et al. [21]. The comparison between their predictions and measured data points showed good agreement with root mean square error (RMSE) of 7%. Zhang and Webb [22] developed a new correlation by modifying Friedel correlation

[17], which is also an updated Lockhart-Martinelli type correlation. The frictional pressure drops of R134a, R22 and R404a in small diameter tubes were measured, and the new correlation predicted the 119 data points with a MAE of 11.5%. The drawback of this kind of predictive method is that it cannot always provide accurate prediction outside the tested range. The computational fluid dynamics (CFD) approach has also shown some success in predicting pressure drop. CFD approaches can simulate transient flow boiling with more detailed information including void fraction, phase velocities and local temperatures. Alizadehdakhel et al. [23] conducted a CFD simulation of gas–liquid two-phase flow in an experimental tube with Volume-of-Fluid (VOF) model. The CFD predictions showed good precision when comparing with the experimental measurements and the corresponding flow regimes. Bhramara et al. [24] applied homogenous equilibrium model in their CFD analysis. They compared CFD results with separated flow correlations, and it matched well with Muller-Steinhagen and Heck correlation [25]. Overall, CFD models have seen limited utilization, mainly due to low accuracy in capturing two-phase flows and the high computational cost.

Over the last decade, a few researchers have developed generalized predictive approaches which have shown great effectiveness. These approaches are based on generating a large database amassing numerous experiments conducted by different researchers worldwide with various working fluids covering an extensive range of geometries and operating conditions. Kim and Mudawar [1] developed a generalized universal correlation for saturated flow boiling pressure drop in mini/micro channels. They amassed a database of 2378 data points from 16 sources with 9 working fluids. The correlation showed excellent accuracy, with an overall MAE of 17.2% for the entire database. Tibirica et al. [26] developed a complete set of simple and optimized correlations for two phase pressure drop in microchannels. Their database included 1076 data points, with 14 working fluid, covering a wide range of operating conditions and channel dimensions. The overall MAE was 13.1%, with 92.1% of datapoints having an error less than 30%.

1.3. Machine learning in thermal analysis

In recent years, both statistical models and machine learning (ML) models, for example, artificial neural networks (ANN), K-nearest neighbors (KNN), support vector machines regression (SVMR) and decision trees, etc. have shown remarkable development and widespread applications in various data driven research fields. In the field of thermal analysis, some early research was conducted by Thibault and Grandjean [27] and Jambunathan et al. [28], who showed that ANNs can be used effectively in modeling practical heat conduction problems and predicting heat transfer coefficients, respectively. After these early studies, ML models have been implemented in analyzing more complex thermal systems including to estimate the pressure drops. Alizadehdakhel et al. [23] applied an ANN trained with three inputs from experiment data to predict average pressure drop across the tube. Their trained ANN model could predict two-phase pressure drop with mean square error (MSE) of 0.043 Pa/m. More recently, in order to predict two-phase pressure drop of refrigerants, Khosravi et al. [29] applied three kinds of ML models, namely, multilayer feed-forward neural network (MLFFNN), support vector regression (SVR) and group method of data handling (GMDH) type neural network. They concluded that the three models can predict the pressure drop with root mean square errors (RMSE) of 139.42 Pa/m, 703.71 Pa/m and 131.62 Pa/m, respectively. Khamehchi and Bemani [30] proposed two novel ML models in predicting two-phase pressure drop, namely gradient tree boosting (GTB) and extreme learning machine (ELM). A comprehensive dataset was implemented to train the models. By comparing with experiment measurements, GTB and ELM models showed excellent prediction with mean relative error of 0.72% and 3.43%, respectively. Longo et al. [31] presented a prediction of refrigerant two-phase frictional pressure drop with the GBM model. Their model was trained by an extensive database including 2549 data points and the MAE was 6.6%. A multiple-layer perceptron neural network (MLPNN) model is presented to predict the pressure drop in a gas/non-Newtonian liquid by Shadloo et al. [32]. The model was trained by 85% of 511 experimental data points and validated by remaining 15% data points. The model estimated pressure drop with a minimum MAE of 4.58%. Recently, Lee et al. [33] presented an ANN based tool for predicting frictional pressure drop in micro-pin fin heat sink. Their model is trained by 1651 data points from 22 studies. The model showed better prediction accuracy than existing correlations with an overall MAE of 14.49%.

1.4. Objective of study

Recently, the authors of present study applied ML models to develop predictive tools for mini/micro-channel flow boiling heat transfer coefficient [34] and mini/micro-channel condensation heat transfer coefficient [35], which showed excellent results. The present study follows a similar methodology. A consolidated database consisting of 2787 data points for flow boiling pressure drop in mini/micro-channels is amassed from 21 sources [5,6,36–54]. Table 1 provides key information on these individual databases incorporated into the consolidated database. The consolidated database includes a broad range of two-phase flow boiling pressure drop data points with the coverage summarized in Table 2.

The objective of this study is to predict flow boiling pressure drop encountered in mini/micro channels by using statistical and ML model like *K*-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Gradient Boosted Trees (XGBoost and LightGBM). These ML models are trained by a large set of input parameters from the consolidated experimental database discussed earlier. In addition, rather than taxonomizing different models that can be used for prediction, the current study also offers performance results with respect to the conventional physics based modeling to analyze how they fare. Finally, some of the statutory limitation of ML modeling will be highlighted by testing them on datasets (hold out data) held out from the training examples.

2. Machine learning models

Problems in which datasets comprises of independent variables along with dependent variables are known as supervised learning problems. Some ML methods rely on some distance measure of similarity between data points, expressed either implicitly or explicitly, which does not scale well in higher dimensions. Hence, when the dimensionality of the problem increases, the performance of several ML algorithms deteriorates - known as the “curse of dimensionality”.

A supervised learning algorithm receives a training data set S, sampled from an unknown distribution D and labeled by a target function f. The objective of the algorithm is to find a predictor that minimizes the error with respect to unknown distribution and target function. A training dataset is a sample available to the learning algorithm. The learning paradigm to find a predictor h_s, which minimizes the training error $L_s(h)$ defined below, is also called Empirical Risk Minimization (ERM) [55].

$$L_s(h)=\frac{|\{i \in \{1,.....,n\} : h(x_i) \neq y_i\}|}{n} \quad (1)$$

The above equation is based on learning theory discussed in [56]. One trivial way to achieve this is by memorization of the

Table 1
Flow boiling pressure drop data for mini/micro-channels in the consolidated database.

Author(s)	Channel geometry*	Channel material	D_h [mm]	Fluid(s)	G [kg/m^{2}s]	Data points
Lezzi et al. (1994)	C, single, H	Stainless Steel	1	Water	776~ 2738	86
Tran (1998)	C, single, H	Brass	2.46	R12, R134a	33.1~ 832.35	439
Yan & Lin (1998)	C, multi, H	–	2	R134a	50~ 200	120
Pettersen (2002)	C, multi, H	Aluminum	0.81	CO2	190~ 570	57
Qu & Mudawar (2003)	R, multi, H	Copper + Lexan cover	0.349	Water	134.93~ 401.94	164
Monroe et al. (2003)	R, multi, H	Aluminum	1.66	R134a	99~ 402	37
Huo (2005)	C, single, VU	Stainless Steel	2.01, 4.26	R134a	400, 500	74
Lee & Mudawar (2005)	R, multi, H	Copper + Lexan cover	0.349	R134a	128.16~ 656.72	87
Owhaib et al. (2008)	C, single, VU	Stainless Steel	1.7, 1.224, 0.826	R134a	100~ 400	53
Hu et al. (2009)	C, single, H	–	4.18, 2	R410A	200~ 620	48
Quiben et al. (2009)	R, multi, H	Copper	3.71, 5.35, 3.5, 4.88	R22, R410A	150~ 500	264
Ducoulombier (2010)	C, single, H	Stainless Steel	0.529	CO2	400, 1200	268
Copetti et al. (2011)	C, single, H	Stainless Steel	2.6	R134a	240, 440, 930	86
Tibirica et al. (2011)	C, single, H	Stainless Steel	2.32	R134a	99.87~ 600.3	49
Tibirica & Ribatski (2011)	C, single, H	Stainless Steel	2.32	R245fa	199.3~ 700.6	142
Wu et al. (2011)	C, single, H	Stainless Steel	1.42	CO2	300~ 600	254
Costa-Patry & John (2012)	R, multi, R	Copper	0.246	R134a	291~ 568	110
Kharangate et al. (2012)	R, single, VU	Copper + Lexan cover	3.3	FC72	340.27~ 351.45	121
Maqbool et al. (2012)	C, multi, H	Stainless Steel	1.7, 1.224	Ammonia	100~ 500	235
Anwar et al. (2015)	C, single, VU	Stainless Steel	1.6	R1234yf, R134a	400, 500	73
Markal et al. (2016)	R, multi, H	Silicone	0.15	Water	51~92.6	20
Total						2787

* C: circular, R: rectangular, H: horizontal, VU: vertical upward, VD: vertical downward;.

Table 2
Range of two-phase flow boiling pressure drop data points.

Parameter	Range	*Mean* (μ)	*Standard Deviation* (σ)
Working fluids	Ammonia, CO2, R12, R1234yf, R134a, R22, R245fa, R410a, FC72, Water	–	–
Reduced pressures	$0.0006 < P_R < 0.7766$	0.1932	0.1288
Hydraulic diameter	0.15 mm $< D_h <$ 5.35 mm	1.91 mm	1.28 mm
Mass velocity	33.1 kg/m^{2}s $< G <$ 2738 kg/m^{2}s	434.32 kg/m^{2}s	370.25 kg/m^{2}s
Liquid-only Reynolds number	$14 \leq Re_{fo} \leq 27{,}658$	4367.62	4082.10
Superficial vapor Reynolds number	$75.58 \leq Re_g \leq 199{,}453$	21,132	24,888
Flow quality	$0 < x < 1$	0.37	0.24

training data which however performs undesirably for unseen data [57]. A commonly adopted solution to the problem is to apply ERM rule over a restricted space by selecting a list of predictors known as hypothesis class. The next few sections briefly discuss some of the predictive models evolved from learning theory.

2.1. Artificial neural network (ANN)

ANN is the mix of several ML concepts like perceptrons, regression, ensemble and gradient descent. It extracts linear combination of input variables and maps the dependent variable as non-linear functions of derived features from the input variables. They are the simplest form of deep networks consisting of several layers of hidden neurons each fully connected to the layer below (from which they receive input) and those above (from which they influence). Any general functions can be constructed by considering networks having successive layers of interconnected units. Each layer processes the weighted linear combination of inputs and transforms it through an activation function. Weights (w_{ji}) are analogous to regression coefficients introduced earlier. For example, if w^l_{ji} is the weight connecting the unit j from unit i in layer l, g is the activation function, then output from jth unit is computed by:

$$z_j = g\left(\sum_{i=0}^{d} w^l_{ji} x_i\right) \tag{2}$$

There are many choices of activation function, and Rectified Linear Unit (ReLU) is used for the present work. ANNs are capable of learning the feature themselves and are trained using the technique known as “backpropagation”, which is a very popular neural network learning algorithm as it is conceptually simple and computationally efficient even though the solution is not a global solution. However, getting it to work well and sometimes to work at all can be complicated [58]. ANN have the drawback of overfitting, and recommended approaches to solve this are by early stopping and weight decay. For the present work early stopping was not used but exponential decay for first and second moment vector in Adam solver was set to 0.9 and 0.999, respectively.

2.2. Extreme gradient boosting (XGBoost)

Boosting is the one of the most powerful concepts introduced in ML in the last few decades. Originally developed for classification problems [59] it has been successfully implemented for regression problems. Boosting algorithms amplifies the strength of weak learners to approximate gradually good predictors for larger and harder classes to learn. Boosted trees have emerged as a versatile and adaptive approach for a wide array of problems. Tree boosting overcomes the problem of the curse of dimensionality introduced earlier by not relying on the distance metric rather the data points relationship is learnt through adaptive adjustment of neighborhoods [60]. This also makes the model invariant to data transformation, unlike neural networks, and scaling the features becomes irrelevant. Also, the deeper the trees constructed, higher the interactions between the features that can be captured [60].

Xgboost is a variant of Gradient Boosted Decision Trees (GBDT) with introduction of regularization term to prevent overfitting. Moreover, rather than using first order derivative in GBDT, a second order Taylor series of loss function is adopted in XGBoost [61].

Table 3
Model parameters selected in this study.

Model	Parameter	Value
ANN	Activation function	ReLu
	L2 Regularization Parameter, α	0.00075
	Solver	Adam
	Batch Size	200
	Learning rate, λ	0.00025
	Exponential decay rate for estimates of first moment vector, $\beta 1$	0.9
	Exponential decay rate for estimates of second moment vector, $\beta 2$	0.999
	Tolerance	0.001
	Hidden Layers	(120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10) *
XGBoost	η	0.1
	Υ	0.0
	Max depth	10
	Minimum child weight	1
	Maximum delta step	0
	Subsample	1
	Sampling Method	Uniform
	λ	1.3
	α	1.3
	Scale Pos Weight	1
	Refresh Leaf	1
	Grow Policy	Depthwise
	Max Leaves	0
	Max Bin	256
KNN	K	2
	Weights	Uniform
	Algorithm	Auto
	Leaf size	30
	Power parameter for the Minkowski metric	2
	Distance metric	Minkowski
LightGBM	Boosting type	gbdt
	Learning rate	0.1
	Max_depth	No limit
	Minimum split gain	0.0
	Minimum child weight	0.001
	Minimum child samples	20
	Subsample	1

*Based on optimization conducted in Table 5.

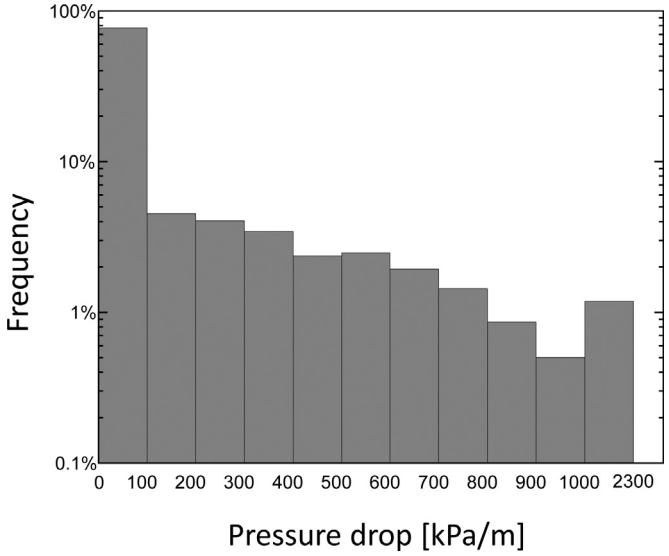

Fig. 1. Distribution of the target variable, pressure drop.

2.3. K-Nearest neighbors (KNN)

K-nearest neighbors regression (KNN-regression) is based on the principle of finding a set of points within a dataset which are closest to the "query" point as quantified by some distance metric (like Euclidean, Manhattan or Minskowski). For a given value of K, KNN regression identifies the K training observations that are closest to X_0, represented by N_0 and can be expressed as [62].

$$\hat{f}(x_0) = \frac{1}{K} \sum_{x_i \in N_0} Y_i \tag{3}$$

The optimal value of K depends on the bias-variance tradeoff with a small value of K enabling a low bias and high variance while larger value of K having high bias and low variance. KNN algorithm organizes the dataset into a hierarchical tree-based data structure [63]. Since KNN regression is memory based it does not require any model to be fit [64]. For distance between two vectors $p = (p_1,p_2,....p_n)$ and $q = (q_1,q_2,q_3.....q_n)$, plausible distance metrics are defined as [65]:

$$\text{Manhattan Distance} = \left[\sum_{i=1}^{n} |p_i - q_i|\right]$$

$$\text{Euclidean Distance} = \left[\sum_{i=1}^{n} |p_i - q_i|^2\right]^{1/2}$$

$$\text{Minkoski Distance} = \left[\sum_{i=1}^{n} |p_i - q_i|^{\alpha}\right]^{1/\alpha} \tag{4}$$

2.4. Light gradient boosting machine (LightGBM)

The objective of Gradient Boosted Decision Trees (GBDT) is to minimize the training error, L(y,F(x)), introduced earlier as ERM. An iterative criterion to minimize this training error using line search can be obtained, however, due to large number of data

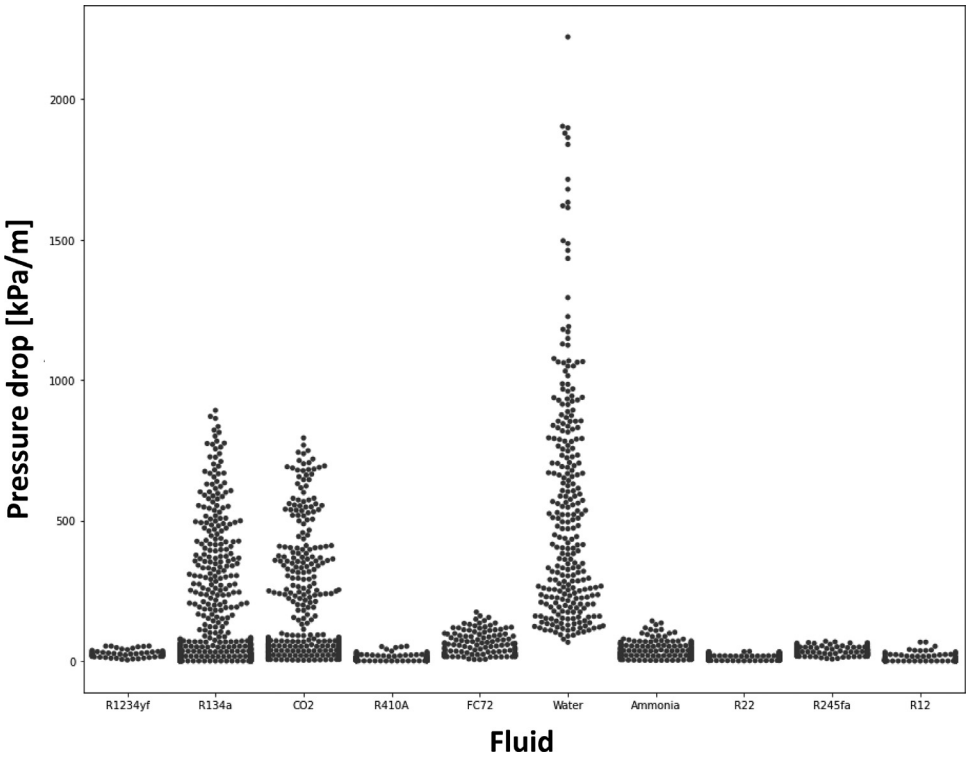

Fig. 2. Swarm plot of the target variable across different fluids.

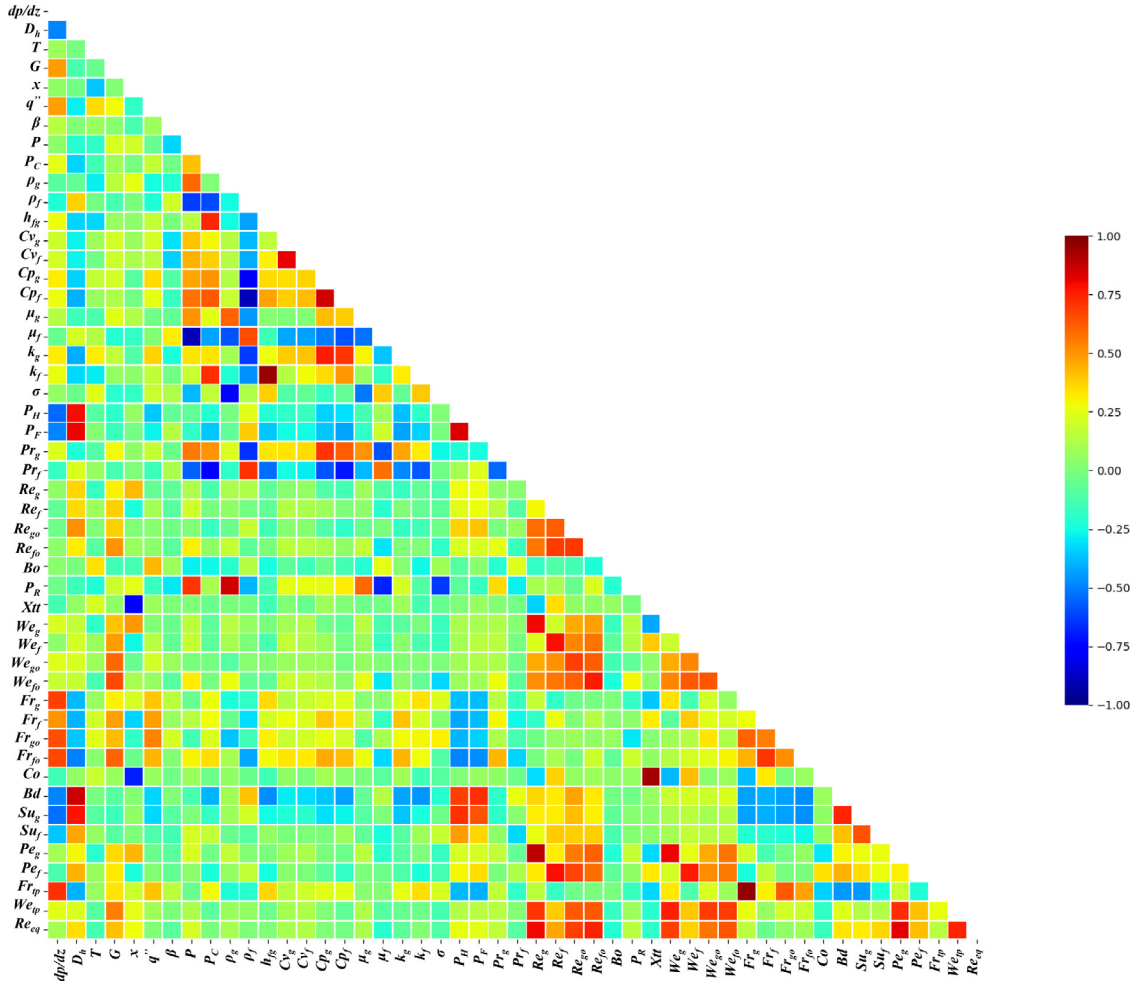

Fig. 3. Kendall rank coefficient heat map for all variables in databases.

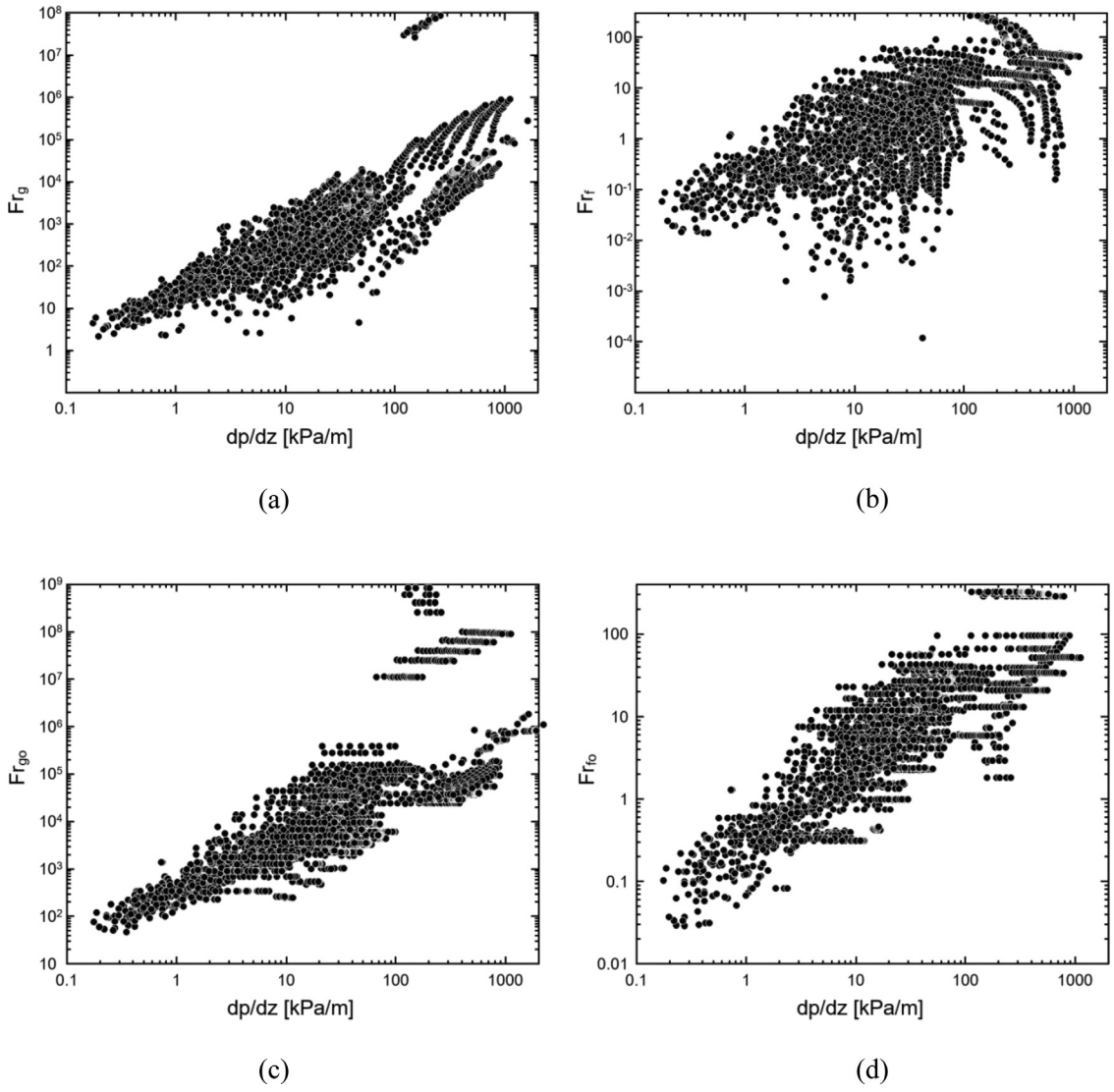

Fig. 4. Univariate distribution of (a) Fr_g, (b) Fr_f, (c) Fr_{go}, and (d) Fr_{fo} with pressure drop.

points and high number of features the conventional GBDT deteriorates. In the light of this, LightGBM is a relatively new algorithm, proposed by Nielsen [66] with the idea to make gradient boosting on decision trees faster. This is based on the concept that instead of checking all possible splits while creating leaves only some of them are selected. This is facilitated by two novel techniques which are the Gradient based One-Side Sampling (GOSS) and the Exclusive Feature Bundling (EFB). With GOSS, a significant proportion of data instances with small gradients is excluded and only the rest is used to estimate information gain. EFB is used to reduce the number of features by bundling mutually exclusive ones. With these two implementations LightGBM achieves the same accuracy as XGBoost but with a speed-up of 20 times. In depth discussion is provided in [66]. The LightGBM code is available at https://github.com/Microsoft/LightGBM.

2.5. Parametric optimization

For the ANN model, parametric optimization was done using the manual search technique similar to [34]. Optimization of neural network is not straightforward as it depends primarily on data distribution, batch size, learning rate and optimizer. Previous attempts to optimize ANN includes alternatives to backpropagation like tabu search [67]. The important parameters like the learning rate, λ, and L2 regularization parameter (αr) are set to 0.001, while the exponential decay rate of the first and second moments is specified to be 0.9 and 0.999, respectively. Mean Squared Error (*MSE)* was the chosen loss function [68].

KNN regression first identifies the '*k*' (user defined) closest points in the space and averages the target value for the nearest '*k*' points. For the present study the optimal value of '*k*' is 2. All the points in the neighborhood are weighted equally. The distance metric selected to compute distance between two points is Minkowski.

The final hyperparameters used in this study for the ML models are summarized in Table 3. Some optimization of the model parameters will be discussed in detail in the results and discussion section later.

2.6. Model evaluation

The model coding is performed in Python, and Pandas, Numpy and Scikit-learn [69] libraries are used to develop the ML algorithms. Because of the limited size of our dataset (2787 data points) when compared to traditional ML studies, the computational cost for training and testing the models is low and can

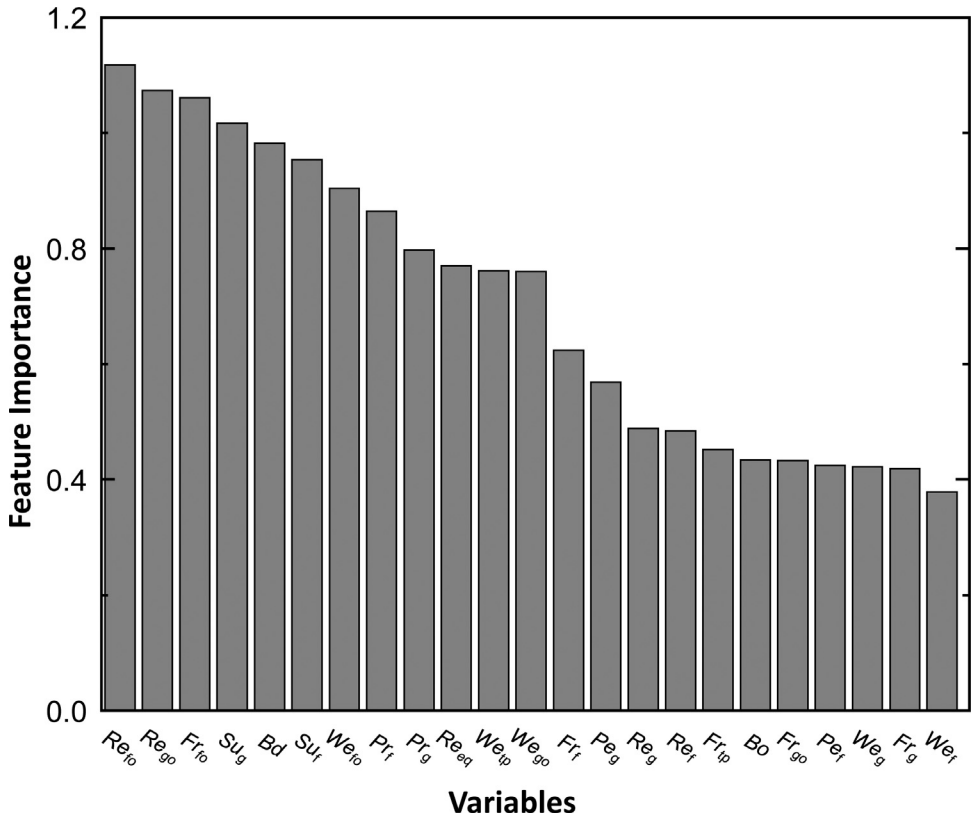

Fig. 5. Feature selection results obtained by mutual information (MI) feature selection method.

be completed within minutes even with desktop computers. The split ratio between training and test datasets is usually 70%−30%, 75%−25%, or 80%−20%. It is necessary that training dataset should be 3–4 times more than test data size as the bulk of the model performance depends on training dataset volume. Different combinations of test and train datasets have been investigated and 75%−25% split ratio is seen to be acceptable. Therefore, the consolidated dataset was split into training and test datasets with 75%–25% split ratio with 2090 training data points and 697 test data points. Model fit is evaluated using four metrics viz. R^2, Adjusted R^2, Mean Absolute Error (MAE) and Explained variance regression score. R^2 is the fraction of variance explained and equals $Corr(Y, \hat{Y})^2$, the square of correlation between response and fitted values. On the other hand, explained variance regression score equals $1 - (\text{var}\{Y - \hat{Y}\}/\text{var}\{Y\})$. R^2 and explained variance regression score equal each other only when mean of the error is zero. It turns out that R^2 monotonically increases upon addition of predictors regardless of the strength of the association between it and the response variable. Hence, the adjusted R^2 "adjusts" this monotonic increase of R^2 by incurring a penalty term for every variable brought into the model. It's preferable to evaluate the model on MAE for the output parameter, pressure drop, rather than mean squared error (MSE) as the penalty for latter is higher than former if the response variable distribution contains outliers. Two more parameters will be used to understand model performance, θ, which is the percentage data predicted within ± 30%, and ε, which is the percentage data predicted within ± 50%.

3. Results and discussion

3.1. Data distribution and correlations

Fig. 1 shows the distribution of the target variable (dependent or response variable), i.e., pressure drop for flow boiling in mini/micro-channels. The mean value (μ) is 122.19 kPa/m while the standard deviation (σ) is 238.18 kPa/m. Fig. 2 shows the box plot of the target variable across different fluids. The box shows the quartiles of the distribution while the whiskers extend to cover the remaining distribution, except the outliers which are shown as points. The supervised learning model is performed on features that qualify both the physics-based and data-driven criterion. Considering strong dependency on the regression models with the curse of dimensionality and the underlying assumptions in each of the ML models, only relevant and non-redundant features should be processed by the models [70]. Fig. 3 shows the Kendall rank coefficient heat map to visualize monotonic relationship between all variables in the dataset. It's a non-parametric method in the sense that it does not rely on any underlying distribution. It is usually preferred over Spearman's correlation map [71], especially when the sample size is small and there are outliers (as shown in Fig. 2) in the data. Interestingly, the heatmap shows strong correlation (> 0.6) with Fr_{tp}, Fr_g, Fr_{go} and Fr_{fo}. Such pattern of strong target variable correlation with these dimensionless numbers was also seen in previous studies [34,35]. The univariate distribution of these four variables with the target is shown in Fig. 4. This explains why these four variables are picked up as significant predictors and although none of the physics based models include them [1], yet there is evidence that these variable are candidates to be used in physics based correlations.

3.2. ML model performance comparison

The consolidated data from flow boiling in mini/micro-channels which is amassed in this study is utilized to develop and analyze the predicting performances for the four ML models. A combination of input parameters, including flow parameters, geometric parameters, fluid properties, and relevant dimensionless numbers are utilized to setup the model with the total pressure drop being the

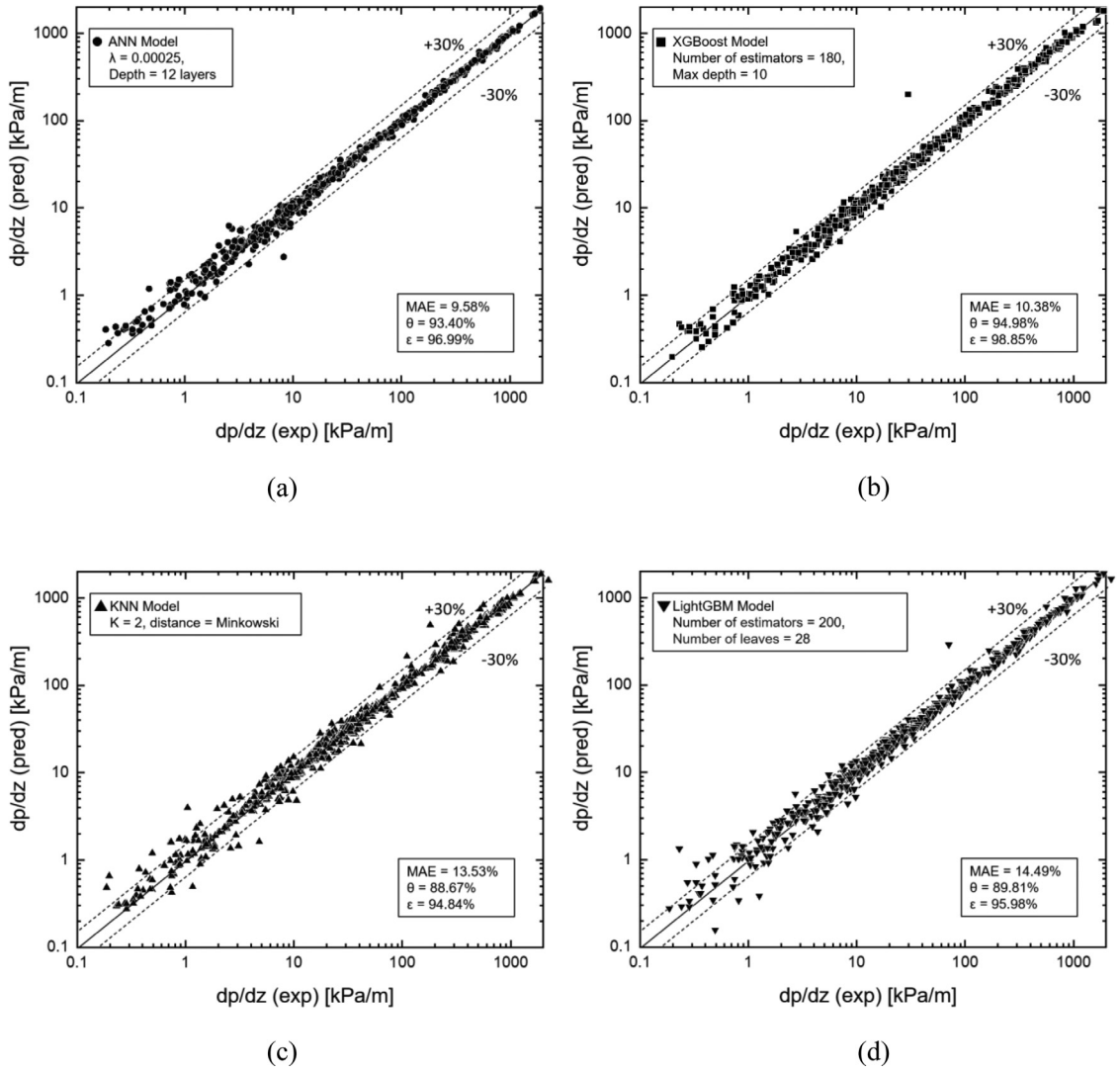

Fig. 6. Pressure drop predictions of (a) ANN model, (b) XGBoost model, (c) KNN model, and (d) LightGBM model with input parameters *Bd, Bo,* Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go} and We_{tp}.

required output parameter. The input parameter selection is made to achieve prediction accuracy as well as facilitate model scalability without unnecessary redundancy.

Table 4 shows the model performance for different combinations of predictor variables for various ML models. For the first two cases, developing the model with limited variables takes a heavy toll on model accuracy. Even though the model fit, measured by R^2 is more than 0.98, the residuals measured by MAE are quite unsatisfactory. For the next three cases, the model fit as well as residuals improve significantly as the complete list of predictors are included. However, bringing in redundant predictors (as shown in Case 5) doesn't yield optimal results as it can lead to model overfitting. This behavior is consistent between the four models. To aid in model scalability in the features, the authors would like to include only the set of dimensionless variables in the current model; hence, the final proposed list of variables are *Bd, Bo,* Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go} and We_{tp}. Use of dimensionless variables is consistent with developing generalized predicting tools for two-phase performance parameters that can be utilized to predict over a large range of fluid, geometric and operating parameters [34,35].

Fig. 5 shows the feature selection results obtained by mutual information (MI) feature selection method for the final proposed list of input variables. MI is a powerful statistical method to measure the extent of relatedness between datasets. The KNN method [72] is used to plot the feature importance. The absolute values are not of significance rather only the relative value of each feature shows the importance by this method. The results show that the five most important features selected are Re_{fo}, Re_{go}, Fr_{fo}, Su_g and *Bd*. This can be further corroborated from the universal correlation proposed by Kim and Mudawar [1] that shows pressure drop dependence on four out of these five features (except *Bd*).

The ANN architecture was optimized manually for the current study and results for hidden layer optimization are shown in Table 5. Fig. 6(a) shows the comparison between experimental and predictions made by the optimized ANN model with input parameters: *Bd, Bo,* Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go}, and We_{tp}, and hidden layers (120,110,100,90,80,70,60,50,40,30,20,10). Of all the models considered, ANN yields the lowest MAE of 9.6% with the percentage data predicted within ± 30% is 93.4% and percentage data predicted within ± 50% is 97%. The results show how well the model can capture the pressure drop behavior across the whole range of the consolidated data.

Fig. 6(b) shows the comparison between experimental and predicted data made by the XGBoost model with input parameters:

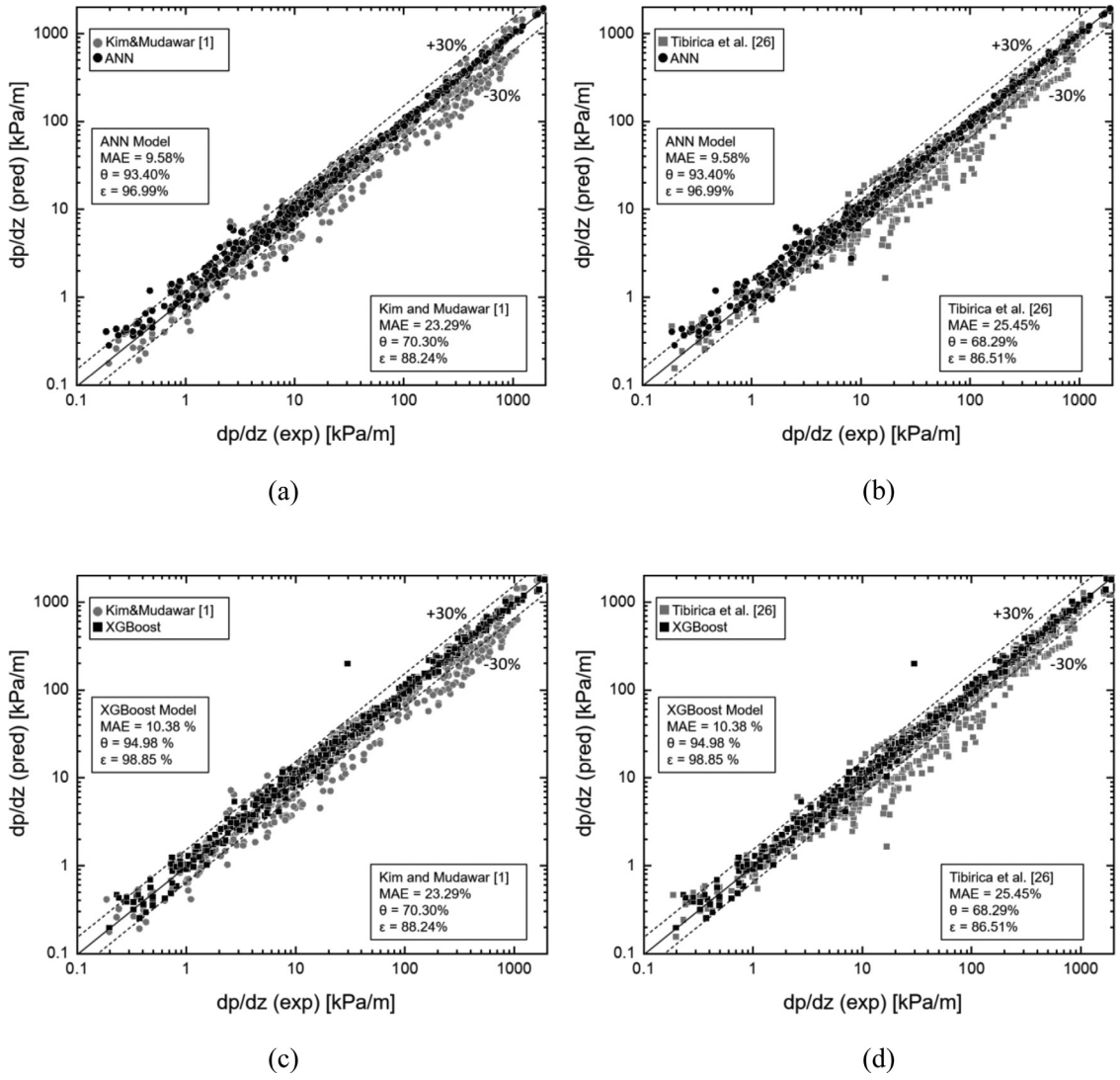

Fig. 7. Comparison between predictions of the generalized correction by Kim and Mudawar [1] with test data predictions based on the (a) ANN model and (b) XGBoost model. Comparison between predictions of the generalized correction by Tibirica et al. [26] with test data predictions based on the (c) ANN model and (d) XGBoost model.

Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go}, and *We_{tp}*. This model gives an *MAE* of 10.38% and percentage data predicted within ± 30% is 94.98% and percentage data predicted within ± 50% is 98.85%. This model performs slightly better than KNN model but not as good as the ANN model.

Table 6 compares the predicting performance for the KNN model based on different hyper parameters. The algorithm to construct the tree is the ball tree method where each node's point is assigned to the node's two closest children. The leaf size selected to construct the tree is 30 (optimal value depends on the nature of the problem). To find the distance between the points, the 'Minkowski distance' metric is used. Selecting the model with 2 nearest neighbors and distance as the weights gives the lowest MAE of 13.53%. Fig. 6(c) shows the comparison between experimental and predicted data made by this KNN model with input parameters: *Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go},* and *We_{tp}*. About 88.67% of the datapoints lie within ±30% of the predicted value while 94.84% of the points lie within ± 50% of the predicted value.

Fig. 6(d) shows the comparison between experimental and predictions made by Light GBM model with input parameters: *Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go},* and *We_{tp}*. To the best of our knowledge, this is the first time Light GBM is being used to predict pressure drop in thermal system modeling. LightGBM can be modeled with three types of gradient boosting methods, viz., gradient boosted decision trees (GBDT), dropout meet multiple additive regression trees (dart) [61], and gradient based one side sampling (GOSS). We noticed the best performance by *GBDT,* for which the MAE is 14.49% with the percentage data predicted within ± 30% is 89.81% and percentage data predicted within ± 50% is 95.98%.

ANN outperforms all models because of its ability to learn complex relationship between input and output parameters by virtue of nonlinear activation function manifested in the hidden layers of the network. They are robust even in the presence of noisy data and even when input parameters have different statistical distribution with respect to the response function [73]. Tree boosting is highly effective and versatile approach for ML problems. The second best performing model, XGBoost, usually can outperforms many other ML algorithms since it uses Newton's method which employs higher order alternative to optimization problem [60]. Additionally, XGBoost has extra randomization parameters which is used to de-correlate individual trees and reduce overall variance of the model. The aforementioned reasons make XGBoost a highly adaptive ML model which considers bias-variance trade-off into ev-

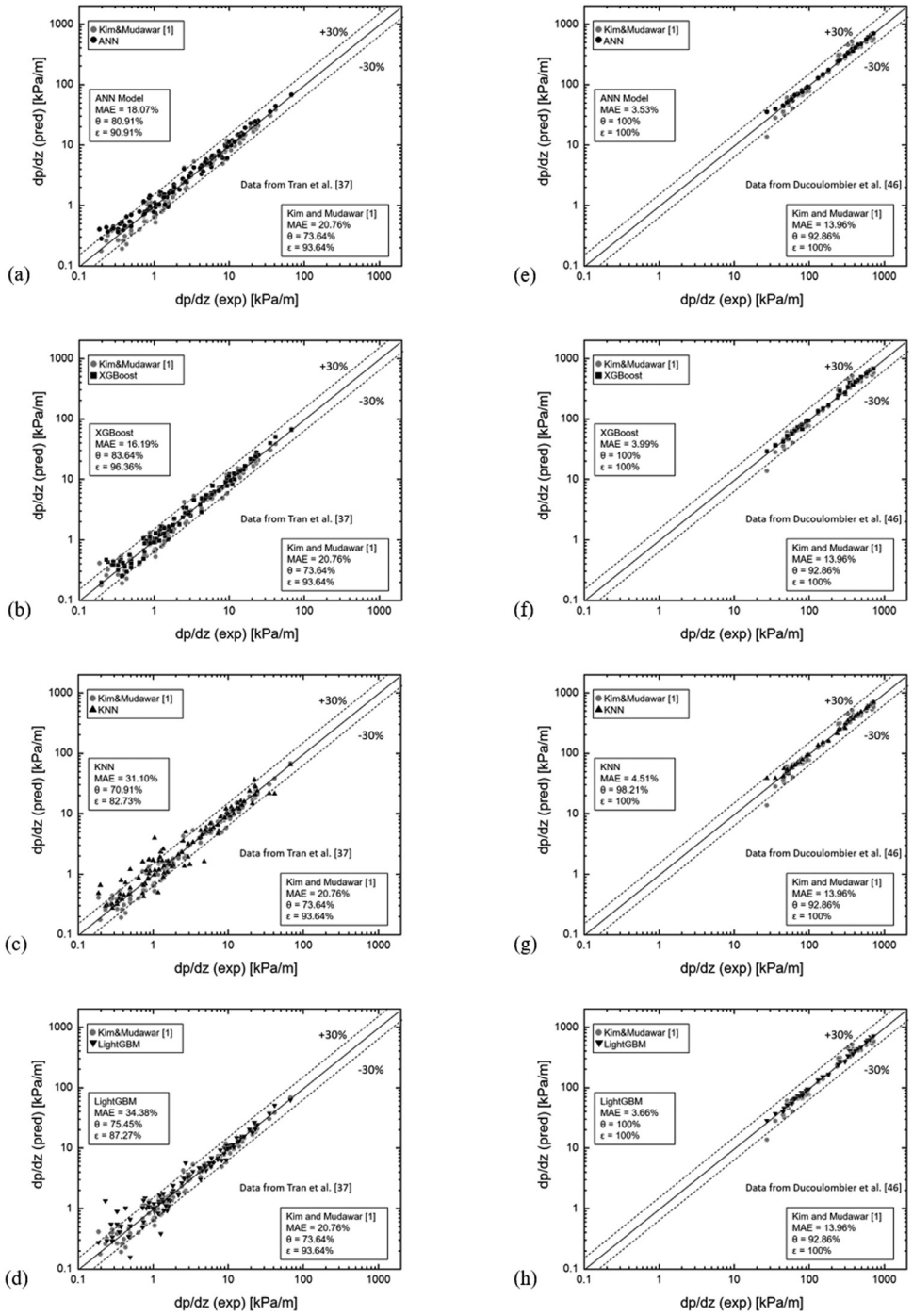

Fig. 8. Test data points from individual datasets of Tran et al. [47] predicted by (a) ANN model, (b) XGBoost model, (c) KNN model and (d) LightGBM model; Test data points from individual datasets of Ducoulombier et al. [46] predicted by (e) ANN model, (f) XGBoost model, (g) KNN model and (h) LightGBM mode.

ery aspect of the learning theory [60]. Prior work has shown that Light GBM performance is usually better than XGBoost [74]. However, we saw its MAE being higher than XGBoost as well as KNN. This could be possibly due to lack of best hyperparameters tuning, although both grid search and random search were investigated to find the optimal hyperparameters for this model.

3.3. Comparison with generalized flow boiling pressure drop correlations

In order to show the capability of the four ML models based on the consolidated database, their results were compared with predictions of the generalized corrections for flow boiling pressure drop proposed by Kim and Mudawar [1] and Tibirica et al. [26]. These two universal/generalized saturated flow boiling heat transfer correlations are presented in Table 7. In Fig. 7 (a–d), we can see that the MAEs for Kim and Mudawar [1] and Tibirica et al. [26] are 23.3% and 25.4%, respectively, while the MAEs obtained from ANN and XGBoost models are 9.6% and 10.4%, respectively. Also, the number of data points predicted within ± 30 and ± 50% of the actual data is in the range of 93–95% and 97,98% for the ML models. Comparatively, the correlations have a wider error band of 68–70% and 86–88% of the data points lying within ± 30 and

Table 4
Model predictions for different combinations of input parameters.

Case	Parameters	Models	MAE	R^2	Adjusted R^2	Explained variance ratio
1	D_h, T, G, x, q'', P	KNN	21.08	0.989	0.989	0.989
		ANN	12.88	0.998	0.998	0.998
		XGBoost	10.39	0.979	0.978	0.979
		LightGBM	17.61	0.996	0.996	0.996
2	Bd, Bo, Fr_f, Pr_f, Pe_f, Re_f, Su_f, We_f,	KNN	73.43	0.793	0.793	0.794
		ANN	43.60	0.986	0.986	0.986
		XGBoost	93.06	0.959	0.958	0.958
		LightGBM	62.97	0.981	0.981	0.981
3	D_h, T, G, x, q'', P, P_C, ρ_f, ρ_g, h_{fg}, Cp_g, Cp_f, μ_f, μg, k_f, k_g, σ	KNN	22.77	0.992	0.992	0.992
		ANN	11.35	0.997	0.997	0.997
		XGBoost	10.29	0.995	0.995	0.995
		LightGBM	16.16	0.996	0.996	0.996
4	Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go}, We_{tp}	KNN	13.53	0.988	0.988	0.988
		ANN	9.58	0.988	0.987	0.988
		XGBoost	10.38	0.987	0.986	0.987
		LightGBM	14.24	0.996	0.996	0.996
5	Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go}, We_{tp}, D_h, T, G, x, q'', P	KNN	13.53	0.988	0.988	0.988
		ANN	12.64	0.998	0.998	0.998
		XGBoost	10.12	0.996	0.996	0.996
		LightGBM	12.66	0.996	0.996	0.996

Table 5
ANN model predictions for fixed input parameters and different combinations of hidden layers.

Test Case	ANN Model Hidden Layers	Input Parameters	MAE	R^2	Adjusted R^2	Explained variance ratio
0	(180,170,160,150,140,130,120,110,100,90,80,70,60,50,40,30,20,10)	*Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go}, We_{tp}*	13.83	0.985	0.985	0.986
1	(170,160,150,140,130,120,110,100,90,80,70,60,50,40,30,20,10)		11.45	0.962	0.961	0.963
2	(160,150,140,130,120,110,100,90,80,70,60,50,40,30,20,10)		13.31	0.987	0.986	0.987
3	(150,140,130,120,110,100,90,80,70,60,50,40,30,20,10)		9.97	0.971	0.971	0.972
4	(140,130,120,110,100,90,80,70,60,50,40,30,20,10		11.56	0.977	0.976	0.978
5	(130,120,110,100,90,80,70,60,50,40,30,20,10)		12.43	0.987	0.986	0.987
6	**(120,110,100,90,80,70,60,50,40,30,20,10)***		**9.58**	**0.988**	**0.987**	**0.988**
7	(110,100,90,80,70,60,50,40,30,20,10)		12.33	0.982	0.981	0.982
8	(90,80,70,60,50,40,30,20,10)		12.39	0.984	0.983	0.984
9	(60,50,40,30,20,10)		15.25	0.987	0.987	0.987
10	(30,20,10)		36.20	0.983	0.983	0.983

*Final selected ANN model configuration.

Table 6
KNN predictions for different hyper parameter *K*.

K	MAE	R^2	explained variance ratio
2	13.53	0.988	0.988
4	17.57	0.988	0.988
6	20.91	0.985	0.985
8	24.97	0.979	0.979
10	29.45	0.971	0.971

± 50%. This shows that the generalized flow boiling pressure drop correlations are typically outperformed by the ML models.

3.4. Understanding sub-dataset predictions

In this section, we want to understand the ability of ML models in capturing various sub-divided datasets and check the predicting performance uniformity across the database. One way is dividing the database into individual datasets by the authors. Fig. 8 shows comparison between predictions of the 4 ML models and the generalized correlation developed by Kim and Mudawar [1] for the two largest datasets in the consolidated databases. The two datasets are by Tran [37] and Ducoulombier [46], with the latter having higher pressure drop values (> 10 kPa/m), while the former has mostly lower pressure drop values (80% of the data points are less than 10 kPa/m). These two databases span 3 working fluids: R134a, R12 and CO_2, mass velocity: 33.1–1200 kg/m²s, hydraulic diameter: 0.529, 2.46 mm, reduced pressure: 0.08 –0.47, quality: 0 and 1.0, and fluid only Reynolds number: 353 –11,517. The first database by Tran [37] has R12 and R134a as the working fluids. As shown in Fig. 8 (a~d), the four ML models have MAEs of 18.07%, 16.19%, 31.10% and 34.38%, respectively, while the MAE of the correlation by Kim and Mudawar is 20.76%. ANN predictions at the low pressure drop ranges (<1 kPa/m) are seen to be slightly overpredicted in comparison to XGBoost. The second database by Ducoulombier [46] has CO_2 as the working fluid. As shown in Fig. 8 (e~h), the four ML models have MAEs of 3.53%,

Table 7
Generalized condensation heat transfer correlations.

	Author(s)	Correlation Formulation
	Kim & Mudawar [1]	$.(\frac{dP}{dz})_F = (\frac{dP}{dz})_f \phi_f^2$ Where $\phi_F^2 = 1 + \frac{C}{X} + \frac{1}{X^2}$, $X^2 = \frac{(dP/dz)_f}{(dP/dz)_g}$, $-(\frac{dP}{dz})_f = \frac{2 f_f \upsilon_f G^2 (1-x)^2}{D_h}$, $-(\frac{dP}{dz})_g = \frac{2 f_g \upsilon_g G^2 x^2}{D_h}$, $f_k = 16\mathrm{Re}_k^{-1}$ for $\mathrm{Re}_k < 2.000$, $f_k = 0.07\mathrm{Re}_k^{-0.25}$ for $2,000 \le Re_k < 20,000$, $f_k = 0.046 Re_k^{-0.2}$ for $Re_k \ge 20,000$, For laminar flow in rectangular channel, $f_k Re_k = 24(1 - 1.3553\beta + 1.9467\beta^2 - 1.7012\beta^3 + 0.9564\beta^4 - 0.2537\beta^5)$, where subscript k denotes f or g for liquid and vapor phases, respectively, $\mathrm{Re}_f \frac{G(1-x)D_h}{\mu_f}$, $\mathrm{Re}_g = \frac{GxD_h}{\mu_g}$, $Re_{fo} = \frac{GD_h}{\mu_f}$, $Su_{go} = \frac{\rho_g \sigma D_h}{\mu_g^2}$, $We_{fo} = \frac{G^2 D_h}{\rho_f \sigma}$, $Bo = \frac{q''_H}{Gh_{fg}}$.
Liquid	Vapor (gas)	$C_{non\text{-}boiling}$
Turbulent	Turbulent	$0.39 Re_{fo}^{0.03} Su_{go}^{0.10} (\frac{\rho_f}{\rho_g})^{0.35}$ for $Re_f \ge 2000, \mathrm{Re}_g \ge 2000$ (tt)
Turbulent	Laminar	$8.7 \times 10^{-4} Re_{fo}^{0.17} Su_{go}^{0.50} (\frac{\rho_f}{\rho_g})^{0.14}$ for $Re_f \ge 2000, \mathrm{Re}_g < 2000$ (tv)
Laminar	Turbulent	$0.0015 Re_{fo}^{0.59} Su_{go}^{0.19} (\frac{\rho_f}{\rho_g})^{0.36}$ for $Re_f < 2000, \mathrm{Re}_g \ge 2000$ (vt)
Laminar	Laminar	$3.5 \times 10^{-5} \mathrm{Re}_{fo}^{0.44} Su_{go}^{0.50} (\frac{\rho_f}{\rho_g})^{0.48}$ for $Re_f < 2000, \mathrm{Re}_g < 2000$ (vv)
	C	
	$Re_f \ge 2000$	$C_{non-boiling}[1 + 60 We_{fo}^{0.32} (Bo \frac{\rho_H}{\rho_F})^{0.78}]$
	$Re_f < 2000$	$C_{non-boiling}[1 + 530 We_{fo}^{0.52} (Bo \frac{\rho_H}{\rho_F})^{1.09}]$
	Tibirica et al. [26]	Homogenous void fraction: $\alpha_{\mathrm{hom}} = (1 + \frac{1-x}{x} \frac{\rho_g}{\rho_f})^{-1}$ Homogenous density: $\rho_{\mathrm{hom}} = \alpha \cdot \rho_g + (1-\alpha) \cdot \rho_f$ Homogenous viscosity: $\mu_{\mathrm{hom}} = x \cdot \mu_g + (1-x) \cdot \mu_f$ Homogenous Reynolds number: $\mathrm{Re}_{\mathrm{hom}} = \frac{G D_h}{\mu_{\mathrm{hom}}}$ New two-phase frictional factor: $f_{new} = 1.415 \cdot (\frac{\rho_f}{\rho_g})^{-0.3263} \cdot \mathrm{Re}_{\mathrm{hom}}^{-0.2342} \cdot (\frac{\rho_f - \rho_g}{\rho_f})^{6.0858}$ Two-phase frictional pressure drop: $(\frac{dp}{dz}) = f_{new} \cdot \frac{G^2}{2 \cdot \rho_{\mathrm{hom}} \cdot D_h}$ Range of applicability $349 \le \mathrm{Re}_{\mathrm{hom}} \le 96862$ $0.02 \le f_{\mathrm{hom}} \le 0.08$ $7.8 \le \rho_f / \rho_g \le 840$ $0 < x \le 1$

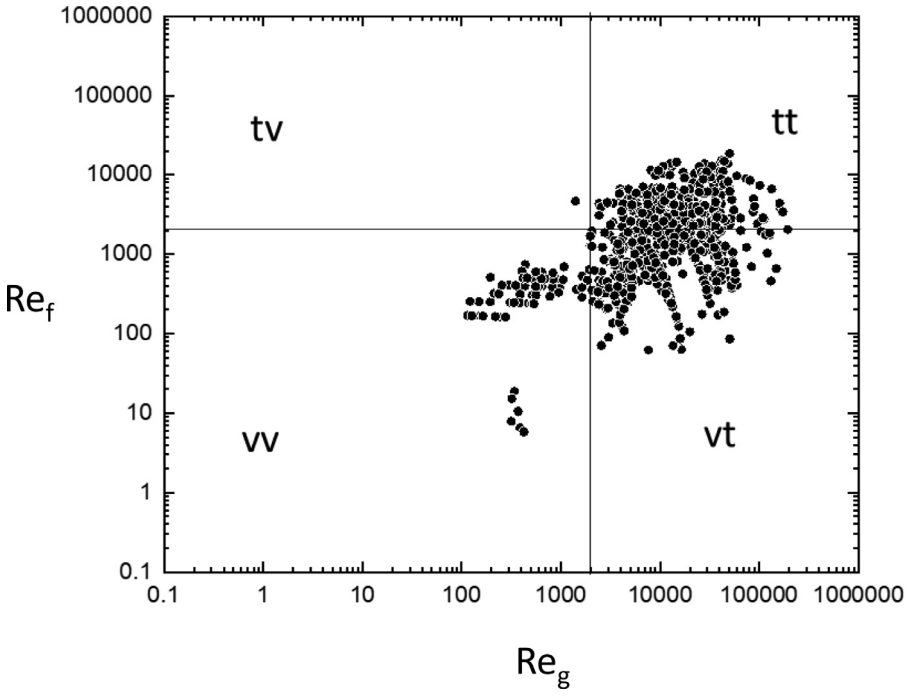

Fig. 9. Distribution of 697 test data points in the 4 flow regimes, namely, laminar-laminar (vv), turbulence-laminar (tv), laminar-turbulence (vt) and turbulence-turbulence (tt) with a transition Reynold number of 2000.

3.99%, 4.51% and 3.66%, respectively, whereas Kim and Mudawar's correlation predicted with a MAE of 13.96%. A peculiar aspect about this dataset is all the target values are high and the ML models developed in this study have shown accurate predictions at high pressure drop values in comparison to low ones. The better performance at the higher pressure drops can also be attributed to an overall reduction in uncertainty in measurement of the experimental data as the pressure drops values increase. Overall, the performance of ML models across all the individual databases is seen to be better than the universal correlation by Kim and Mudawar [1].

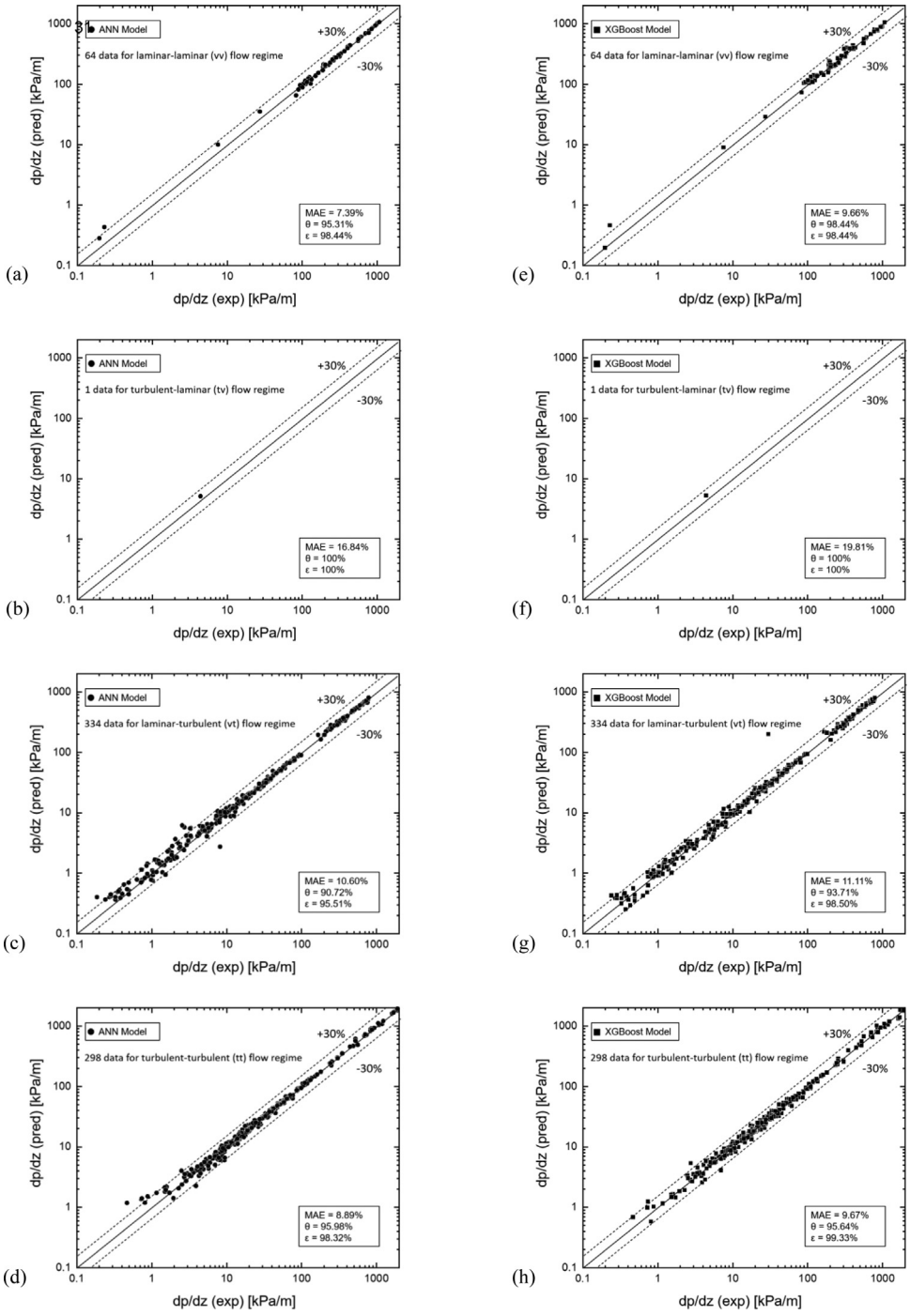

Fig. 10. ANN model predictions for each flow regime: (a) laminar-laminar, (b) turbulent-laminar, (c) laminar-turbulent, and (d) turbulent-turbulent; XGBoost model predictions for each flow regime: (e) laminar-laminar, (f) turbulent-laminar, (g) laminar-turbulent, and (h) turbulent-turbulent.

Another way to divide the database is into the flow regimes. Flow boiling in mini/micro-channels can be divided into 4 flow regimes, including laminar-laminar (vv), turbulence-laminar (tv), laminar-turbulence (vt) and turbulence-turbulence (tt). Fig. 9 shows the cluster of 697 data points subdivided into 4 flow regimes with a transition Reynold number of 2000. Fig. 10 (a~d) shows that the ANN model predicts the data points in each flow regime with MAEs of 7.39% (vv), 16.84% (tv), 10.60% (vt) and 8.89% (tt), respectively. Fig. 10 (e~h) shows that the XGBoost model predict the data points in each flow regime with MAEs of 9.66% (vv), 19.81% (tv), 11.11% (vt) and 9.67% (tt), respectively. For both models, the predictions for turbulence-laminar (tv) flow regime are done on a single data point, because our consolidated test database only has one data point located within the turbulence-laminar flow regime with the remaining 4 going into the training dataset. Overall, the performance is consistently good across the flow regimes with a slightly lower MAE for the dataset with the single data point.

One more way to divide the database is into channel configurations, single-channel vs. multi-channels. There are 416 test datapoints for single-channel, and 281 test datapoints for multi-channels configurarions. From Table 8, we can see that ANN and

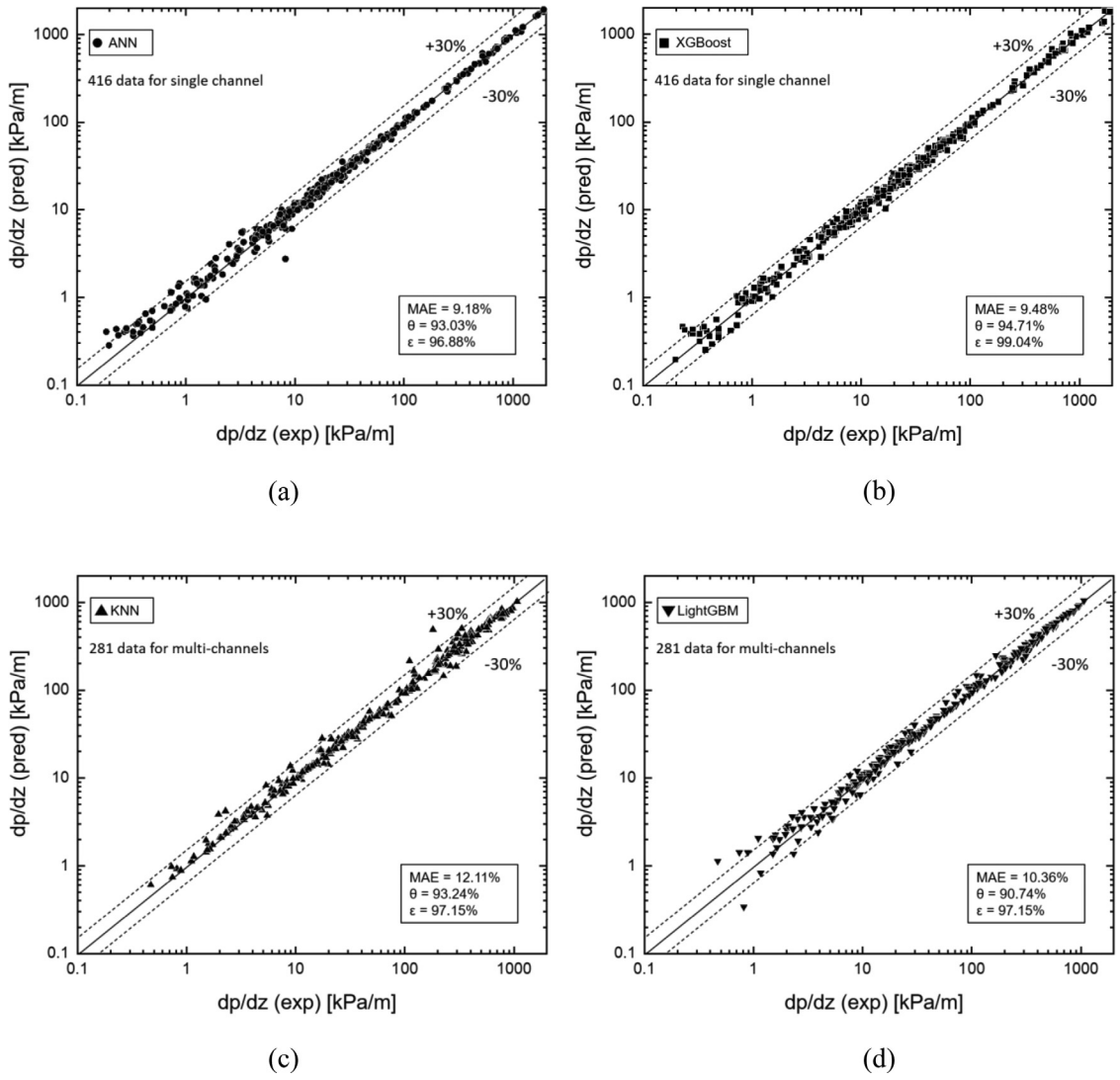

Fig. 11. (a) ANN model predictions for single-channel test data. (b) XGBoost model predictions for single-channel test data. (c) KNN model predictions for multi-channels flow test data. (d) LightGBM model predictions for multi-channels test data.

Table 8
Model predictions for different channel configurations.

Model	Single-channel	Multi-channel
ANN	9.18%	10.19%
XGBoost	9.48%	11.69%
KNN	16.10%	12.11%
LightGBM	15.67%	10.36%

XGBoost have slightly smaller MAEs for single channel, while KNN and LightGBM predict multi-channels better than single channel data points. Fig. 11(a) and (b) shows that the ANN and XGBoost models have MAEs of 9.18% and 9.48%, respectively for single-channel. Fig. 11(c) and (d) show that the KNN and LightGBM models have MAEs of 12.11% and 10.36%, respectively for multi-channel data points. In general, we can see that the ML models are able to perform effectively for the complete database.

3.5. Predicting excluded (hold out) datasets

As discussed in earlier sections, when we use the consolidated database to train the ML models, ANN, XGBoost, KNN, and LightGBM perform reasonably well with MAEs of 9.58%, 10.38%, 13.53% and 14.49%, respectively. However, the ability to predict dependent variables outside the range of the training data is also critical in evaluating the scalability and robustness of ML model for new, unseen data points. We select three datasets to exclude and understand the model performance. The three datasets excluded (hold out data) are Yan and Lin [38], Wu et al. [49–50] and Kharangate et al. [51].

Fig. 12(a–d) shows the predictions of ANN, XGBoost, KNN and LightGBM model when Yan and Lin [38] is excluded from the training dataset that included data only from the 20 remaining sources. For ANN model and XGBoost model, when dataset of Yan and Lin [38] was excluded, MAEs are 30.5%, and 34.27%, respectively. For KNN model and LightGBM model, the MAEs are higher at 42.14%, and 43.02%, respectively. The predictions when Wu et al. [49] is excluded are shown in Fig. 12 (e–h). For ANN model and XGBoost model, MAEs are 40.23%, and 34.23%, respectively. For KNN model and LightGBM model, the MAEs are 52.92%, and 38.25%, respectively. As shown in Table 1, R134a and CO_2 are the most common working fluids in our consolidated databases and we expected them to perform well. Better predicting capability in predicting hold out datasets was obtained in our past ML modeling efforts on heat transfer coefficient in min/micro-channel [34,35]. One rea-

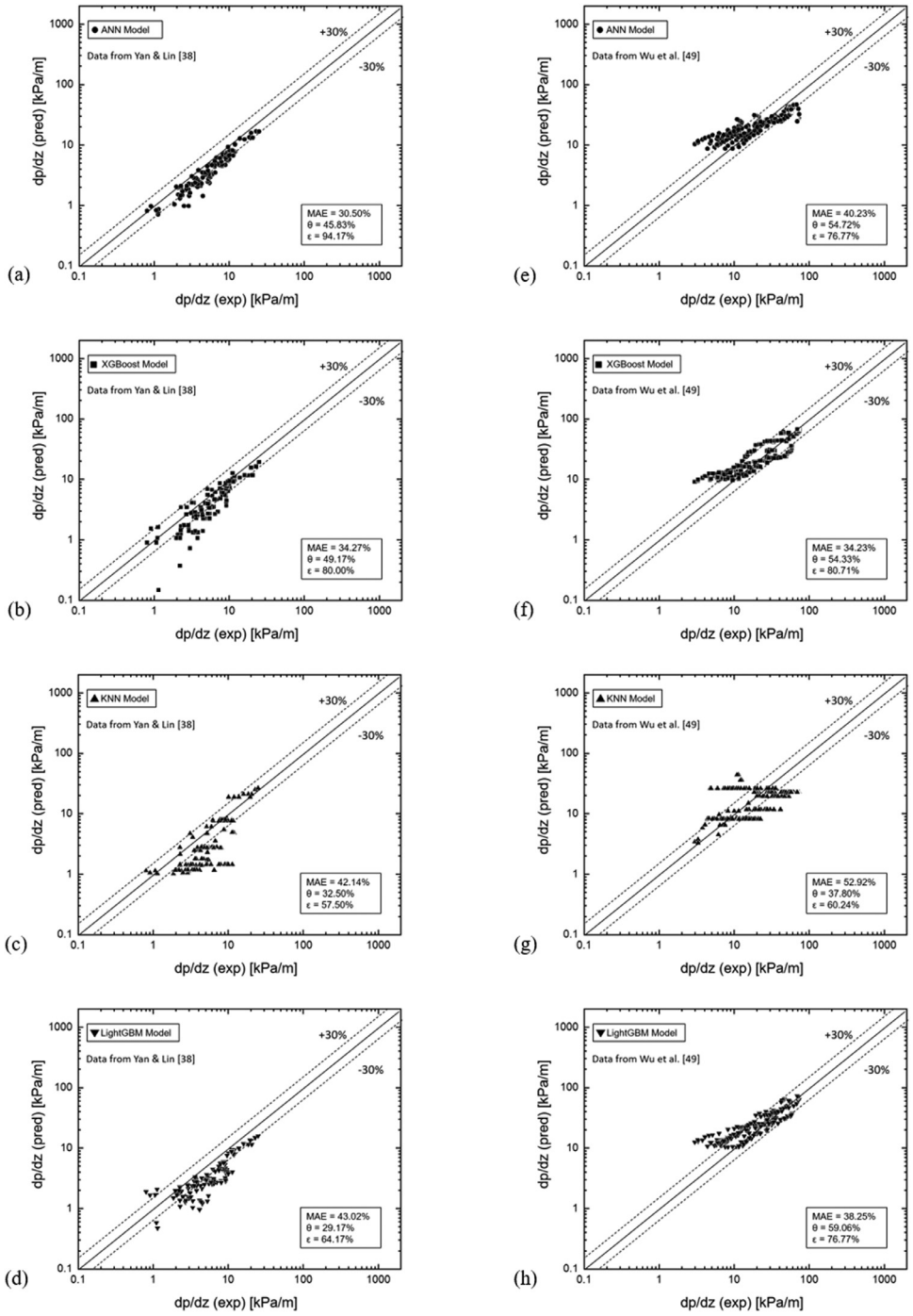

Fig. 12. Predictions of pressure drop of excluded data for Yan & Lin [38] based on the (a) ANN model, (b) XGBoost model, (c) KNN model, and (d) LightGBM model. Predictions of pressure drop of excluded for Wu et al. [49] based on the (e) ANN model, (f) XGBoost model, (g) KNN model, and (h) LightGBM model.

son is that the size of this dataset is much smaller than the earlier studies (2787 in the current study compared to 16,953 and 4882 in [34] and [35], respectively). Specifically, from Table 1, we can also see less overlap between other datasets and the two we are discussion here. For Yan and Lin [38], it is the only dataset in multi-circular channels configuration among all 10 datasets which utilize R134a as the working fluid. In addition, the mass velocity range of [38] is 50 ~ 200 kg/m^{2}s, smaller than most other R134a data points in the database. Similarly, Wu et al. [49] has the largest hydraulic diameter among all CO_2 datasets as seen in Table 1. This highlights the need for some overlapping data to build and deploy such ML models.

Fig. 13 (a–d) shows the predictions of ANN, XGBoost, KNN and LightGBM model when data from Kharangate et al. [51] is excluded from the training database that included data only from the 20 remaining sources. The ANN, XGBoost, KNN, and LightGBM models predict with MAEs of 117.40%, 65.74%, 93.43%, and 31.59%, respectively. The working fluid in Kharangate et al. [51] is FC72, and the lower performance can easily be attributed to no information on the working fluid being available in the training database.

A strong dependence of model performance on the size of consolidated database and the availability of fluid data in the training datasets can be observed from the results. The high reliability of ML models on availability of existing data is a major disadvantage

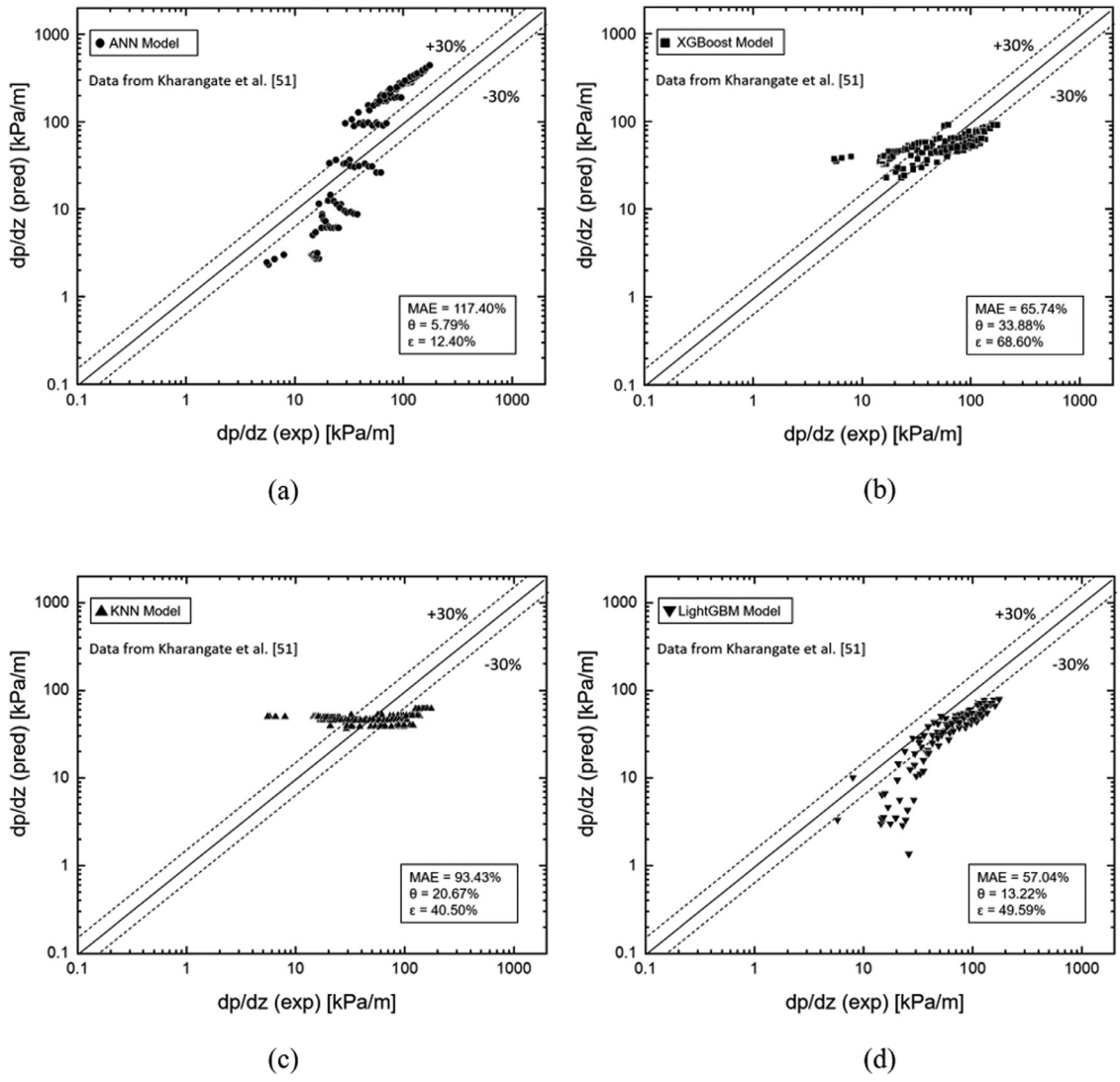

Fig. 13. Predictions of pressure drop for excluded data for Kharangate et al. [40–51] based on the (a) ANN model, (b) XGBoost model, (c) KNN model, and (d) LightGBM model.

of such tools. Despite that, the ML models can still predict pressure drop with a reasonable accuracy, when a reasonable amount of data points and fluid information are made available. Lack of data is also seen to impact performance of other traditional methods utilized for predicting pressure drop in flow boiling, including universal/generalized correlations (Kim and Mudawar [1] and Tibirica et al. [26]). As we amass more data and include more working fluids that cover a larger range of operating and geometric conditions than the current consolidated database, the shortcoming is possible to be addressed.

4. Conclusions

In this study, a new method for predicting pressure drop for saturated flow boiling in mini/micro channels is proposed. A consolidated database is amassed and used to develop ML models for predicting pressure drop. Key findings from this study are as follows:

(1) A consolidated database of 2787 data points for saturated flow boiling pressure drop in mini/micro-channels is amassed from 21 sources that includes 10 working fluid, reduced pressures of 0.0006–0.7766, hydraulic diameters of 0.15–5.35 mm, mass velocities of $33.1 < G < 2738$ kg/m^{2}s, liquid-only Reynolds numbers of 14–27,658, superficial vapor Reynolds number of 76–199,453 and flow qualities of 0 and 1.

(2) Four ML based models, namely, Artificial Neural Networks, Extreme Gradient Boosting, K-Nearest Neighbors and Light Gradient Boosting Machine were developed and parametrically optimized in this study. Based on dimensionless input parameters, Bd, Bo, Fr_f, Fr_{fo}, Fr_g, Fr_{go}, Fr_{tp}, Pr_f, Pr_g, Pe_g, Pe_f, Re_f, Re_{fo}, Re_g, Re_{go}, Re_{eq}, Su_f, Su_g, We_f, We_{fo}, We_g, We_{go} and We_{tp}, ANN model, XGBoost model, KNN model and LightGBM model predicted the test data with MAEs of 9.58%, 10.38%, 13.52% and 14.49%, respectively.

(3) The ANN and XGBoost models' predictions are compared with generalized correlations for saturated flow boiling pressure drop developed by Kim and Mudawar [1] and Tibirica et al. [26]. Both ANN and XGBoost models performed better than the highly reliable universal generalized flow boiling pressure drop correlations in mini/micro-channels, even predicting individual databases with high accuracy. The models showed comparable good performance across individual datasets, flow regimes, and channel configurations.

(4) In order to test the predicting capability of four ML models for datasets outside its training database, three datasets were indi-

vidually excluded from the training data, and the models developed based on the remaining data was used to train and predict the complete excluded dataset. None of the ML models were able to perform with the same accuracy as the original models. This is because of the smaller size of this database and less overlap between the datasets from the different sources.

Declaration of Competing Interest

None.

CRediT authorship contribution statement

Yue Qiu: Formal analysis, Investigation, Data curation, Validation. **Deepak Garg:** Conceptualization, Methodology, Software, Writing – original draft. **Sung-Min Kim:** Data curation, Methodology, Writing – review & editing. **Issam Mudawar:** Methodology, Writing – review & editing. **Chirag R. Kharangate:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Acknowledgment

This project was partly sponsored by the Office of Naval Research (ONR), under Grant Number N00014–21–1–2078. The views and conclusions contained herein are those of the authors only and should not be interpreted as representing those of ONR, the U.S. Navy or the U.S. Government.

References

[1] S.M. Kim, I. Mudawar, Universal approach to predicting two-phase frictional pressure drop for mini/micro-channel saturated flow boiling, Int. J. Heat Mass Transf. 58 (2013) 718–734, doi:10.1016/j.ijheatmasstransfer.2012.11.045.
[2] W. Qu, I. Mudawar, Flow boiling heat transfer in two-phase micro-channel heat sinks––I. Experimental investigation and assessment of correlation methods, Int. J. Heat Mass Transf. 46 (2003) 2755–2771, doi:10.1016/S0017-9310(03)00041-3.
[3] I. Mudawar, Two-phase microchannel heat sinks: theory, applications, and limitations, J. Electron. Packag. 133 (2011) 041002, doi:10.1115/1.4005300.
[4] S.M. Kim, I. Mudawar, Review of databases and predictive methods for pressure drop in adiabatic, condensing and boiling mini/micro-channel flows, Int. J. Heat Mass Transf. 77 (2014) 74–97, doi:10.1016/j.ijheatmasstransfer.2014.04.035.
[5] W. Qu, I. Mudawar, Measurement and prediction of pressure drop in two-phase micro-channel heat sinks, Int. J. Heat Mass Transf. 46 (2003) 2737–2753, doi:10.1016/S0017-9310(03)00044-9.
[6] J.B. Copetti, M.H. Macagnan, F. Zinani, N.L.F. Kunsler, Flow boiling heat transfer and pressure drop of *R*-134a in a mini tube: an experimental investigation, Exp. Therm. Fluid Sci. 35 (2011) 636–644, doi:10.1016/J.EXPTHERMFLUSCI.2010.12.013.
[7] B. Markal, O. Aydin, M. Avci, An experimental investigation of saturated flow boiling heat transfer and pressure drop in square microchannels, Int. J. Refrig. 65 (2016) 1–11, doi:10.1016/J.IJREFRIG.2015.12.013.
[8] L. Yin, P. Jiang, R. Xu, H. Hu, L. Jia, Heat transfer and pressure drop characteristics of water flow boiling in open microchannels, Int. J. Heat Mass Transf. 137 (2019) 204–215, doi:10.1016/j.ijheatmasstransfer.2019.03.108.
[9] W.H. McAdams, Vaporization inside horizontal tubes-II, benzene oil mixtures, Trans. ASME 64 (1942) 193–200.
[10] D.R.H. Beattie, P.B. Whalley, Simple two-phase frictional pressure drop calculation method, Int. J. Multiph. Flow 8 (1982) 83–87 (United Kingdom).
[11] S. Lin, C.C.K. Kwok, R.Y. Li, Z.H. Chen, Z.Y. Chen, Local frictional pressure drop during vaporization of R-12 through capillary tubes, Int. J. Multiph. Flow 17 (1991) 95–102, doi:10.1016/0301-9322(91)90072-B.
[12] M.M. Awad, Y.S. Muzychka, Effective property models for homogeneous two-phase flows, Exp. Therm. Fluid Sci. 33 (2008) 106–113, doi:10.1016/j.expthermflusci.2008.07.006.
[13] R.W. Lockhart, R.C. Martinelli, Proposed correlation of data for isothermal two-phase, two-component flow in pipes, Chem. Eng. Prog. 45 (1949) 39–48.
[14] J.R.S. Thom, Prediction of pressure drop during forced circulation boiling of water, Int. J. Heat Mass Transf. 7 (1964) 709–724, doi:10.1016/0017-9310(64)90002-X.
[15] C.J. Baroczy, Systematic correlation for two-phase pressure drop, Chem. Eng. Prog. Symp. Ser. 62 (No. 64) (1966) 232–249 Atomics International, Canoga Park, Calif., 1966.
[16] D. Chisholm, A theoretical basis for the Lockhart-Martinelli correlation for two-phase flow, Int. J. Heat Mass Transf. 10 (1967) 1767–1778, doi:10.1016/0017-9310(67)90047-6.
[17] L. Friedel, Improved friction pressure drop correlation for horizontal and vertical two-phase pipe flow, in: Proceedings of the European Two-phase Flow Group Meeting, 1979, Ispra,, Italy, 1979.
[18] H.J. Lee, S.Y. Lee, Pressure drop correlations for two-phase flow within horizontal rectangular channels with small heights, Int. J. Multiph. Flow 27 (2001) 783–796, doi:10.1016/S0301-9322(00)00050-1.
[19] M.W. Wambsganss, J.A. Jendrzejczyk, D.M. France, N.T. Obot, Frictional pressure gradients in two-phase flow in a small horizontal rectangular channel, Exp. Therm. Fluid Sci. 5 (1992) 40–56, doi:10.1016/0894-1777(92)90055-A.
[20] K. Mishima, T. Hibiki, Some characteristics of air-water two-phase flow in small diameter vertical tubes, Int. J. Multiph. Flow 22 (1996) 703–712, doi:10.1016/0301-9322(96)00010-9.
[21] W. Yu, D.M. France, M.W. Wambsganss, J.R. Hull, Two-phase pressure drop, boiling heat transfer, and critical heat flux to water in a small-diameter horizontal tube, Int. J. Multiph. Flow 28 (2002) 927–941, doi:10.1016/S0301-9322(02)00019-8.
[22] M. Zhang, R.L. Webb, Correlation of two-phase friction for refrigerants in small-diameter tubes, Exp. Therm. Fluid Sci. 25 (2001) 131–139, doi:10.1016/S0894-1777(01)00066-8.
[23] A. Alizadehdakhel, M. Rahimi, J. Sanjari, A.A. Alsairafi, CFD and artificial neural network modeling of two-phase flow pressure drop, Int. Commun. Heat Mass Transf. 36 (2009) 850–856, doi:10.1016/j.icheatmasstransfer.2009.05.005.
[24] P. Bhramara, K.V. Sharma, T.K.K. Reddy, Prediction of pressure drop of refrigerants for two-phase flow inside a horizontal tube using CFD analysis, ARPN J. Eng. Appl. Sci. 4 (9) (2009) 64–71.
[25] H. Müller-Steinhagen, K. Heck, A simple friction pressure drop correlation for two-phase flow in pipes, Chem. Eng. Process. 20 (1986) 297–308, doi:10.1016/0255-2701(86)80008-3.
[26] C.B. Tibiriçá, D.M. Rocha, I.L.S. Sueth, G. Bochio, G.K.K. Shimizu, M.C. Barbosa, S.D.S. Ferreira, A complete set of simple and optimized correlations for microchannel flow boiling and two-phase flow applications, Appl. Therm. Eng. 126 (2017) 774–795, doi:10.1016/j.applthermaleng.2017.07.161.
[27] J. Thibault, B.P.A. Grandjean, A neural network methodology for heat transfer data analysis, Int. J. Heat Mass Transf. 34 (1991) 2063–2070, doi:10.1016/0017-9310(91)90217-3.
[28] K. Jambunathan, S.L. Hartle, S. Ashforth-Frost, V.N. Fontama, Evaluating convective heat transfer coefficients using neural networks, Int. J. Heat Mass Transf. 39 (1996) 2329–2332, doi:10.1016/0017-9310(95)00332-0.
[29] A. Khosravi, J.J.G. Pabon, R.N.N. Koury, L. Machado, Using machine learning algorithms to predict the pressure drop during evaporation of R407C, Appl. Therm. Eng. 133 (2018) 361–370, doi:10.1016/j.applthermaleng.2018.01.084.
[30] E. Khamehchi, A. Bemani, Prediction of pressure in different two-phase flow conditions: machine learning applications, Meas. J. Int. Meas. Confed. (2020) 108665, doi:10.1016/j.measurement.2020.108665.
[31] G.A. Longo, S. Mancin, G. Righetti, C. Zilio, R. Ceccato, L. Salmaso, Machine learning approach for predicting refrigerant two-phase pressure drop inside brazed plate heat exchangers (BPHE), Int. J. Heat Mass Transf. 163 (2020) 120450, doi:10.1016/j.ijheatmasstransfer.2020.120450.
[32] M.S. Shadloo, A. Rahmat, A. Karimipour, S. Wongwises, Estimation of pressure drop of two-phase flow in horizontal long pipes using artificial neural networks, J. Energy Resour. Technol. (2020) 142, doi:10.1115/1.4047593.
[33] H. Lee, M. Kang, K. Wook, C.R. Kharangate, S. Lee, M. Iyengar, C. Malone, M. Asheghi, K.E. Goodson, H. Lee, An artificial neural network model for predicting frictional pressure drop in micro-pin fin heat sink, Appl. Therm. Eng. 194 (2021) 117012, doi:10.1016/j.applthermaleng.2021.117012.
[34] Y. Qiu, D. Garg, L. Zhou, C. Kharangate, S.M. Kim, I. Mudawar, An artificial neural network model to predict mini/micro-channels saturated flow boiling heat transfer coefficient based on universal consolidated data, Int. J. Heat Mass Transf. 149 (2020) 119–211 (n.d.), doi:10.1016/j.ijheatmasstransfer.2019.119211.
[35] L. Zhou, D. Garg, Y. Qiu, S.M. Kim, I. Mudawar, C.R. Kharangate, Machine learning algorithms to predict flow condensation heat transfer coefficient in mini/micro-channel utilizing universal data, Int. J. Heat Mass Transf. 162 (2020) 120351, doi:10.1016/j.ijheatmasstransfer.2020.120351.
[36] A.M. Lezzi, A. Niro, G.P. Beretta, Experimental data on CHF for forced convection water boiling in long horizontal capillary tubes, in: Proceedings of the International Heat Transfer Conference, Begel House Inc., 1994, pp. 14–18. Digit. Libr..
[37] T.N. Tran, Pressure drop and heat transfer study of two-phase flow in small channels [dissertation], Texas Tech University, TXLubbock, 1999.
[38] Y.Y. Yan, T.F. Lin, Evaporation heat transfer and pressure drop of refrigerant R-134a in a small pipe, Int. J. Heat Mass Transf. 41 (1998) 4183–4194, doi:10.1016/S0017-9310(98)00127-6.
[39] J. Pettersen, Flow vaporization of CO_2 in microchannel tubes, Exp. Therm. Fluid Sci. 28 (2004) 111–121, doi:10.1016/S0894-1777(03)00029-3.
[40] C.A. Monroe, T.A. Newell, J.C. Chato, An Experimental Investigation of Pressure Drop and Heat Transfer in Internally Enhanced Aluminum Microchannels, Air Conditioning and Refrigeration Center College of Engineering, 2003.
[41] X. Huo, Experimental study of flow boiling heat transfer in small diameter-tubes, Ph.D thesis, London South Bank University, UK, 2005.
[42] J. Lee, I. Mudawar, Two-phase flow in high-heat-flux micro-channel heat sink for refrigeration cooling applications: part II—heat transfer characteristics, Int. J. Heat Mass Transf. 48 (2005) 941–955, doi:10.1016/J.IJHEATMASSTRANSFER.2004.09.019.

[43] W. Owhaib, C. Martín-Callizo, B. Palm, Two-phase flow pressure drop of R-134A in a vertical circular mini/micro channel, in: ASME Sixth InternationalConference on Nanochannels, Microchannels, and Minichannels, Germany,ICNMM2008-62243 (2008) 343–353.
[44] H. Hu, G. Ding, X. Huang, B. Deng, Y. Gao, Pressure drop during horizontal flow boiling of R410A/oil mixture in 5mm and 3mm smooth tubes, Appl. Therm. Eng. 29 (2009) 3353–3365, doi:10.1016/j.applthermaleng.2009.05.011.
[45] J.M. Quibén, L. Cheng, R.J. Da Silva Lima, J.R. Thome, Flow boiling in horizontal flattened tubes: part I - two-phase frictional pressure drop results and model, Int. J. Heat Mass Transf. 52 (2009) 3634–3644, doi:10.1016/j.ijheatmasstransfer.2008.12.032.
[46] M. Ducoulombier, Ebullition convective du dioxyde de carbone - étudeexpérimentale en micro-canal, Ph.D thesis, Institut National des SciencesAppliquées (INSA) de Lyon, France, 2010.
[47] C.B. Tibiriçá, J.D. Da Silva, G. Ribatski, Experimental investigation of flow boiling pressure drop of R134A in a microscale horizontal smooth tube, J. Therm. Sci. Eng. Appl. 3 (2011), doi:10.1115/1.4003728.
[48] C.B. Tibiriçá, G. Ribatski, Two-phase frictional pressure drop and flow boiling heat transfer for R245fa in a 2.32-mm Tube, Heat Transf. Eng. 32 (2011) 1139–1149, doi:10.1080/01457632.2011.562725.
[49] J. Wu, T. Koettig, C. Franke, D. Helmer, T. Eisel, F. Haug, J. Bremer, Investigation of heat transfer and pressure drop of CO_2 two-phase flow in a horizontal minichannel, Int. J. Heat Mass Transf. 54 (2011) 2154–2162, doi:10.1016/J.IJHEATMASSTRANSFER.2010.12.009.
[50] E. Costa-Patry, J. Olivier, J. Thome, Heat transfer charcacteristics in a copper micro-evaporator and flow pattern-based prediction method for flow boiling in microchannels, Front. Heat Mass Transf. 3 (2012) 1–14, doi:10.5098/hmt.v3.1.3002.
[51] C.R. Kharangate, I. Mudawar, M.M. Hasan, Experimental and theoretical study of critical heat flux in vertical upflow with inlet vapor void, Int. J. Heat Mass Transf. 55 (2012) 360–374, doi:10.1016/j.ijheatmasstransfer.2011.09.028.
[52] M.H. Maqbool, B. Palm, R. Khodabandeh, Flow boiling of ammonia in vertical small diameter tubes: two phase frictional pressure drop results and assessment of prediction methods, Int. J. Therm. Sci. 54 (2012) 1–12, doi:10.1016/j.ijthermalsci.2011.11.018.
[53] Z. Anwar, B. Palm, R. Khodabandeh, Flow boiling heat transfer, pressure drop and dryout characteristics of R1234yf: experimental results and predictions, Exp. Therm. Fluid Sci. 66 (2015) 137–149, doi:10.1016/J.EXPTHERMFLUSCI.2015.03.021.
[54] B. Markal, O. Aydin, M. Avci, An experimental investigation of saturated flow boiling heat transfer and pressure drop in square microchannels, Int. J. Refrig. 65 (2016) 1–11, doi:10.1016/j.ijrefrig.2015.12.013.
[55] S. Shalev-Shwartz, S. Ben-David, Understanding Machine Learning: From theory to Algorithms, Cambridge University Press, 2014.
[56] V. Vapnik, Principles of risk minimization for learning theory, Adv. Neural Inf. Process. Syst. (1992) 831–838.
[57] H. Zhang, M. Cisse, Y. N. Dauphin, D. Lopez-Paz, mixup: beyond empirical risk minimization, ICLR, 2018.
[58] Y.A. LeCun, L. Bottou, G.B. Orr, K.R. Müller, Efficient backprop, in: Neural Networks: Tricks of the Trade, Springer, 2012, pp. 9–48.
[59] Y. Freund, R. Schapire, N. Abe, A short introduction to boosting, J. Jpn. Soc. Artif. Intell. 14 (1999) 1612.
[60] D. Nielsen, Tree boosting with XGBoost-why does XGBoost win "'every"' machine learning competition? (Master's thesis), NTNU (2016).
[61] R.K. Vinayak, R. Gilad-Bachrach, Dart: dropouts meet multiple additive regression trees, in: Proceedings of the Artificial Intelligence and Statistics, PMLR, 2015, pp. 489–497.
[62] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical learning: Data mining, inference, and Prediction, Springer Science & Business Media, 2009.
[63] N. Kumar, L. Zhang, S. Nayar, What is a good nearest neighbors algorithm for finding similar patches in images? in: Proceedings of the European Conference on Computer Vision, Springer, 2008, pp. 364–378.
[64] J. Friedman, T. Hastie, R. Tibshirani, The Elements of Statistical Learning, Springer Series in Statistics, New York, 2001.
[65] I.H. Witten, E. Frank, M.A. Hall, C.J. Pal, in: Practical machine learning tools and techniques, 578, Morgan Kaufmann, 2005, p. 1.
[66] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.Y. Liu, Lightgbm: a highly efficient gradient boosting decision tree, in: Proceedings of the Advances in Neural Information Processing Systems, 2017, pp. 3146–3154.
[67] R.S. Sexton, B. Alidaee, R.E. Dorsey, J.D. Johnson, Global optimization for artificial neural networks: a tabu search application, Eur. J. Oper. Res. 106 (1998) 570–584.
[68] H. Li, Z. Xu, G. Taylor, C. Studer, T. Goldstein, Visualizing the loss landscape of neural nets, in: Proceedings of the Advances in Neural Information Processing Systems, 2018, pp. 6389–6399.
[69] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[70] G. Van Dijck, M.M. Van Hulle, Speeding up the wrapper feature subset selection in regression by mutual information relevance and redundancy analysis BT - Proceedings of the Artificial Neural Networks-ICANN 2006, in: S.D. Kollias, A. Stafylopatis, W. Duch, E. Oja (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 2006: pp. 31–40.
[71] C. Spearman, The proof and measurement of association between two things, The American Journal of Psychology 15 (1) (1904) 72–101 https://doi.org/10.2307/1412159.
[72] B.C. Ross, Mutual information between discrete and continuous data sets, PLoS ONE 9 (2014) e87357.
[73] N. Memon, S.B. Patel, D.P. Patel, Comparative analysis of artificial neural network and XGBoost algorithm for PolSAR image classification, in: B. Deka, P. Maji, S. Mitra, D.K. Bhattacharyya, P.K. Bora, S.K. Pal (Eds.), Pattern Recognition and Machine Intelligence, Springer International Publishing, Cham, 2019, pp. 452–460.
[74] Y. Su, Prediction of air quality based on Gradient Boosting Machine Method, in: Proceedings of the International Conference on Big Data, IEEE, 2020, pp. 395–397. Informatiz. Educ..