

Article

Development of GBRT Model as a Novel and Robust Mathematical Model to Predict and Optimize the Solubility of Decitabine as an Anti-Cancer Drug

Walid Kamal Abdelbasset ^{1,2,*}, Shereen H. Elsayed ³, Sameer Alshehri ⁴, Bader Huwaimel ⁵, Ahmed Alobaida ⁶, Amal M. Alsubaiyel ⁷, Abdulsalam A. Alqahtani ⁸, Mohamed A. El Hamd ^{9,10,*}, Kumar Venkatesan ¹¹, Kareem M. AboRas ¹² and Mohammed A. S. Abourehab ^{13,14}

- ¹ Department of Health and Rehabilitation Sciences, College of Applied Medical Sciences, Prince Sattam bin Abdulaziz University, P.O. Box 173, Al-Kharj 11942, Saudi Arabia
 - ² Department of Physical Therapy, Kasr Al-Aini Hospital, Cairo University, Giza 12613, Egypt
 - ³ Department of Rehabilitation Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
 - ⁴ Department of Pharmaceutics and Industrial Pharmacy, College of Pharmacy, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
 - ⁵ Department of Pharmaceutical Chemistry, College of Pharmacy, University of Hail, Hail 81442, Saudi Arabia
 - ⁶ Department of Pharmaceutics, College of Pharmacy, University of Hail, Hail 81442, Saudi Arabia
 - ⁷ Department of Pharmaceutics, College of Pharmacy, Qassim University, Buraidah 52571, Saudi Arabia
 - ⁸ Department of Pharmaceutics, College of Pharmacy, Najran University, Najran 11001, Saudi Arabia
 - ⁹ Department of Pharmaceutical Sciences, College of Pharmacy, Shaqra University, Al Dwadmi 11961, Saudi Arabia
 - ¹⁰ Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, South Valley University, Qena 83523, Egypt
 - ¹¹ Department of Pharmaceutical Chemistry, College of Pharmacy, King Khalid University, Abha 62529, Saudi Arabia
 - ¹² Department of Electrical Power and Machines, Faculty of Engineering, Alexandria University, Alexandria 21928, Egypt
 - ¹³ Department of Pharmaceutics, College of Pharmacy, Umm Al-Qura University, Makkah 21955, Saudi Arabia
 - ¹⁴ Department of Pharmaceutics and Industrial Pharmacy, Faculty of Pharmacy, Minia University, Minia 61519, Egypt
- * Correspondence: walidkamal.wr@gmail.com (W.K.A.); aboelhamdmohamed@su.edu.sa (M.A.E.H.)



Citation: Abdelbasset, W.K.; Elsayed, S.H.; Alshehri, S.; Huwaimel, B.; Alobaida, A.; Alsubaiyel, A.M.; Alqahtani, A.A.; El Hamd, M.A.; Venkatesan, K.; AboRas, K.M.; et al. Development of GBRT Model as a Novel and Robust Mathematical Model to Predict and Optimize the Solubility of Decitabine as an Anti-Cancer Drug. *Molecules* **2022**, *27*, 5676. <https://doi.org/10.3390/molecules27175676>

Academic Editor: Francisco Torrens

Received: 5 July 2022

Accepted: 17 August 2022

Published: 2 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The efficient production of solid-dosage oral formulations using eco-friendly supercritical solvents is known as a breakthrough technology towards developing cost-effective therapeutic drugs. Drug solubility is a significant parameter which must be measured before designing the process. Decitabine belongs to the antimetabolite class of chemotherapy agents applied for the treatment of patients with myelodysplastic syndrome (MDS). In recent years, the prediction of drug solubility by applying mathematical models through artificial intelligence (AI) has become known as an interesting topic due to the high cost of experimental investigations. The purpose of this study is to develop various machine-learning-based models to estimate the optimum solubility of the anti-cancer drug decitabine, to evaluate the effects of pressure and temperature on it. To make models on a small dataset in this research, we used three ensemble methods, Random Forest (RFR), Extra Tree (ETR), and Gradient Boosted Regression Trees (GBRT). Different configurations were tested, and optimal hyper-parameters were found. Then, the final models were assessed using standard metrics. RFR, ETR, and GBRT had R2 scores of 0.925, 0.999, and 0.999, respectively. Furthermore, the MAPE metric error rates were 1.423×10^{-1} , 7.573×10^{-2} , and 7.119×10^{-2} , respectively. According to these facts, GBRT was considered as the primary model in this paper. Using this method, the optimal amounts are calculated as: P = 380.88 bar, T = 333.01 K, Y = 0.001073.

Keywords: optimization; anti-cancer drug; simulation; artificial intelligence

1. Introduction

Recent studies in the area of clinical pharmacology have necessitated the invention of novel, promising, and environmentally friendly tools to increase the performance of therapeutic drugs [1,2]. In order to achieve this purpose, numerous efforts have been made to develop disparate approaches to reduce the application of potentially detrimental/deleterious organic solvents.

Decitabine (currently sold under the brand name DACOGEN[®]) is an intravenously administered chemotherapeutic drug, which acts as a nucleoside metabolic inhibitor [3–5]. Despite the emergence of some adverse events such as neutropenia, thrombocytopenia, and embryo–fetal toxicity in patients, the drug's great efficiency in ameliorating MDS has encouraged researchers to use it extensively [6–8]. Having long been considered, the most important duty of research and development (R&D) centers of pharmaceutical companies has been to concentrate on the development of supercritical fluid technology (SCFT) [9].

The existence of noteworthy advantages such as negligible processing times, the manufacturing of no organic co-solvent, and its great capability of extracting bioactive molecules has encouraged the researchers to use SCFT for drug discovery from natural sources [10–12]. In recent years, CO₂ supercritical fluid (CO₂SCF) has received more attention within SCFT as an efficient solvent, due to significant privileges such as chemical inactivity, availability, cost-effectiveness, low critical temperature/pressure, and its approval as a food-grade solvent [13].

In recent years, artificial intelligence (AI) has found its place as a versatile tool, offering a high potential for applications in different industries such as separation, extraction, and nanotechnology, as well as drug development, including the identification, validation, and designation of novel drugs [14,15]. The various advantages of AI technology, such as robustness and time-effectiveness, have provided an appropriate chance to overcome the incompetence and discrepancies which may take place during conventional drug optimization and development techniques [16,17]. Machine learning (ML) is a predictive mathematical approach based on AI, which has paved the way to estimate the solubility of drugs in CO₂SCF. To increase the generalization and performance of a single model, an ensemble of models is used in ML. Because of the blending of diverse predictions, ensembles generate effective predictive algorithms with enormous generalization capability [18]. Some ML techniques, such as decision tree and linear regression, are inherently unstable, which means that changing the training dataset results in a significantly different estimator. Unstable estimators have a low bias and high variability. Ensemble approaches have been proposed to reduce generalization error, that is, to reduce variance, bias, or both. In these approaches, the training dataset is modified, and an ensemble of different base estimators is created. These estimators are then combined to create a single estimator [19]. This section provides a quick overview of three main ensemble algorithms: bagging, gradient boosting, and Extra Trees.

In this study, a few basic models were first studied. Given that the decision tree gave significantly better outcomes, but these findings were not general enough to be presented as a powerful model backend, it was decided to employ models that reinforce it. Bagging and boosting are known as the most efficient advanced methods with decision trees. Bagging (bootstrap aggregating), created by Breiman et al. [20], can be considered as both a principal approach and a straight ensemble approach, illustrating brilliant efficiency as long as it reduces variance and avoids overfitting. The bootstrap technique, which replicates training datasets and creates training data subsets, contributes to the diversity of the bagging algorithm. Each subset is utilized to fit a various basic estimator, and the ultimate prediction outcomes are compiled by applying a majority vote procedure.

The other ensemble technique which is introduced from the work of Freund and Schapire is boosting [21]. In contrast to bagging, that provided a variety of basic learners by gradually reweighting the training data. Each sample weakly estimated by the previous estimator is given a higher weight in the next training step. As a result, training samples weakly estimated by predecessors are more likely to occur in the following bootstrap

sample, and bias can be removed effectively. The final boosting algorithm model integrates all the underlying base estimators, which are weighted using their prediction performance.

There is a recently developed decision tree model called Extremely Random Tree (ExtraTree) that is an improved version of the traditional top-down decision tree (DT) model. There is an ensemble of DTs that has been trained in a recursive manner. The final model is built using a massive DT that is trained in a recursive manner. In each case, the tree must be expanded using the entire set of data, and the proper cut point for each split can be calculated through the amount of data gained from each split [22,23].

The three algorithms selected for this study are:

- Random Forest (Bagging of Regression Trees);
- Extra Trees (Bagging of Regression Trees);
- Gradient Boosting (Boosting of Regression Trees).

2. Dataset

Solubility models were created using a dataset with 32 input vectors, similar to [24]. The dataset is illustrated in Table 1. The distribution of features and output is shown in Figure 1. The diagonal subplots (when the x-axis and y-axis are identical) of this figure also show the kernel density estimate (KDE) plot. KDE plots visualize the distribution of observations in a dataset, like histograms. With KDEs, one or more dimensions of probability density curves are used to represent the data.

Table 1. The whole dataset.

No	X1 = P (Bar)	X2 = T (K)	Y = Solubility (Mole Fraction)
1	120	308	5.04×10^{-5}
2	120	318	4.51×10^{-5}
3	120	328	3.69×10^{-5}
4	120	338	2.84×10^{-5}
5	160	308	8.23×10^{-5}
6	160	318	9.37×10^{-5}
7	160	328	9.11×10^{-5}
8	160	338	7.79×10^{-5}
9	200	308	1.18×10^{-4}
10	200	318	1.55×10^{-4}
11	200	328	1.77×10^{-4}
12	200	338	2.05×10^{-4}
13	240	308	1.37×10^{-4}
14	240	318	1.87×10^{-4}
15	240	328	2.82×10^{-4}
16	240	338	3.71×10^{-4}
17	280	308	1.76×10^{-4}
18	280	318	2.40×10^{-4}
19	280	328	3.42×10^{-4}
20	280	338	4.90×10^{-4}
21	320	308	1.97×10^{-4}
22	320	318	2.69×10^{-4}
23	320	328	4.27×10^{-4}

Table 1. Cont.

No	X1 = P (Bar)	X2 = T (K)	Y = Solubility (Mole Fraction)
24	320	338	7.15×10^{-4}
25	360	308	2.18×10^{-4}
26	360	318	3.40×10^{-4}
27	360	328	5.60×10^{-4}
28	360	338	8.74×10^{-4}
29	400	308	2.83×10^{-4}
30	400	318	5.06×10^{-4}
31	400	328	7.88×10^{-4}
32	400	338	1.07×10^{-3}

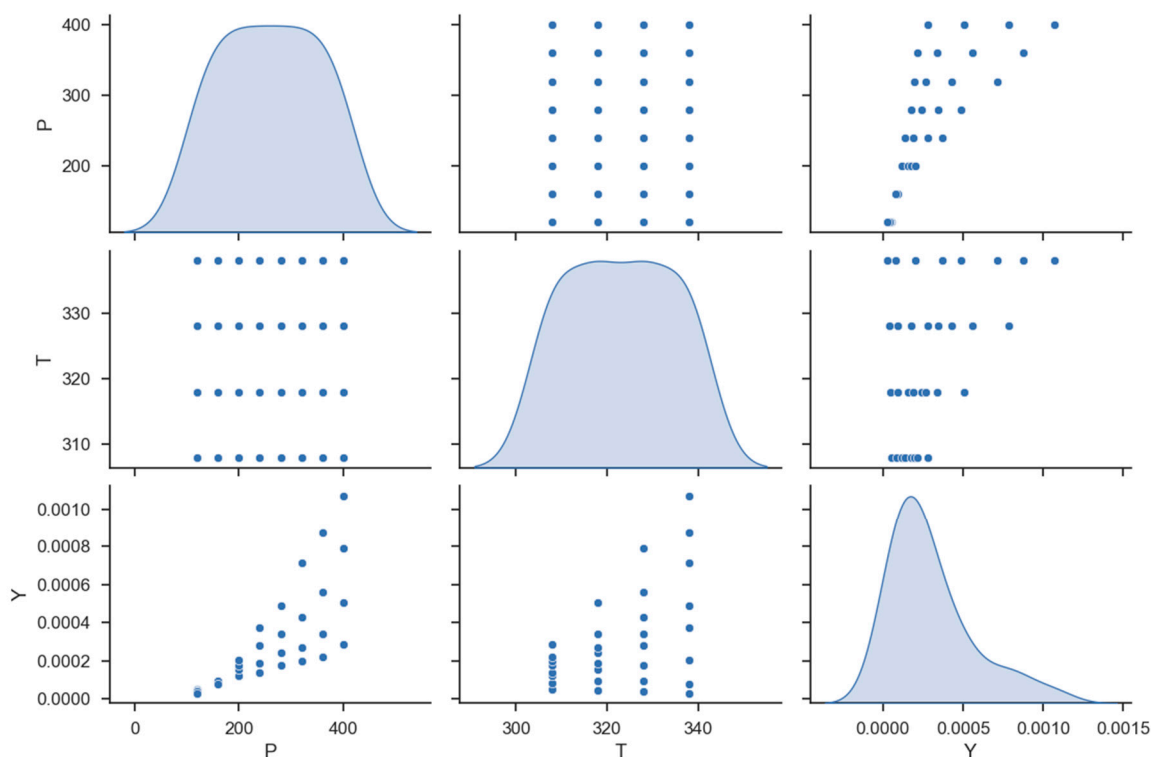


Figure 1. Data distribution, P (pressure), T (temperature), and Y (solubility).

3. Methodology

3.1. Random Forest Regression (RFR)

This regression method is an ML procedure which estimates the targeted output by combining the results of several DT learners [25,26]. When Random Forest receives an (x) input data point, it includes the amounts of the various input features probed for a given training area, which creates K regression trees and the averages of their results. The RF regression predictor, after such K -trees $\{T(x)\}_1^K$ have been trained, is [27]:

$$\hat{f}_{rf}^K(x) = \frac{1}{K} \sum_{k=1}^K T(x)$$

For bypass tree correlation, Random Forest makes the trees grow using different training subsets generated across a routine called bagging. This process is a subset creation

technique that involves resampling the original dataset randomly with replacement, in order to generate the next training subset $\{h(x, \Theta_k), k = 1, 2, \dots, K\}$, where each Θ_k is the same distributed independent random vector. Accordingly, some data can be applied many times during training, while some data points may never be used. By creating a tree using RF, the best split point will be created through a set of input light [28–30].

Moreover, the data which are not used in training step in the k -th tree model in the bagging method, they have been utilized in an out-of-bag subset (oob). k -th tree model be able to utilize the oob items to calculate accuracy [29]. Moreover, the non-selected items in the training step of the k -th tree along the bagging routine are considered in the out-of-bag subset (oob). On the other hand, the k -th tree must be able to utilize these oob items in order to evaluate the efficiency. Increasing the quantity of trees results in a reduction in error, which is illustrated by the fact that the Random Forest does not have an overfitting issue. The relative importance of the features is likewise determined using RF. In order to select the best features in multi-source investigations, it is critical to find the relationship between each item and predicted procedure, and this feature can assist with that understanding [30,31].

3.2. Extra Tree Regression (ETR)

Geurts et al. [32] introduced the extremely randomized tree (ExtraTree), which is an improved version of traditional top-down decision tree (DT) models. The ExtraTree is an ensemble of DTs that has been trained in a recursive manner, and the final model was built employing a massive DT. Each develops the tree utilizing the whole dataset, and the suitable cut point for each split can be decided through achieved information [22,23].

This model is very close to the Random Forest and the Extra Tree model's primary innovations in that (i) the nodes are separated randomly when applying cut points, and (ii) the whole training dataset was used for developing the decision tree instead of subset generation using the bootstrap in Random Forest method [23,33].

3.3. Gradient Boosting Regression Trees (GBRT)

To improve prediction accuracy, boosting uses a series of base estimator compare to a single predictor to get an average between them. Base estimators/models (such as decision trees) are coordinated to clear bias in a stage-wise process. In order to modify the loss function, a new learner is introduced at each phase. Using training data, the first learner decreases the loss function to the lowest amount [34–36]. The following estimators make use of the previous estimators' residuals. The following Algorithm 1 demonstrates the gradient boosting procedure [35–37]:

Algorithm 1

Initialize $F_0(x) = \operatorname{argmin}_p \sum_{i=1}^N L(y_i, P)$

For $m \in \{1, 2, \dots, M\}$:

1. Compute the negative gradient

$$\bar{y}_i = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F_{x_i}} \right]$$

2. Create a model

$$a_m = \operatorname{argmin}_{a, \beta} \sum_{i=1}^N [\bar{y}_i - \beta h(x_i, a_m)]^2$$

3. Select a gradient descent step size as

$$p_k = \operatorname{argmin}_p \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + p h(x_i, a))$$

4. Modify the estimation of $F(x)$

$$F_m(x) = F_{m-1}(x) + p_k h(x, a_m)$$

Output: the aggregated regression function $F_m(x)$

Here, x demonstrates the feature vector and y demonstrates the corresponding class label. $\{x_i, y_i\}_{N_i=1}$, as training data and the aim is to calculate $F_*(x)$, be able to design x to y graph, a specific loss function $L(y, F(x))$ could be reduced to the lowest amount.

4. Results

After tweaking the hyper-parameters of models by testing various combinations of them, we employed *MAPE* and R^2 [38] to verify the accuracy and generality of the models.

The performance success of estimating findings is frequently measured using *R*-squared, which is without a doubt the most often utilized criterion. Using this metric, you can see how closely the predicted results match up with the observed data [39].

$$R^2 = 1 - \frac{\sum(e_i - \bar{o}_i)^2}{\sum(o_i - \bar{o}_i)^2}$$

MAPE is also one of the most used evaluation metrics. *MAPE* illustrates error size, which is between 0 and 1 [40].

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{o_i - e_i}{o_i} \right|$$

e_i and o_i are predicted and actual (observed) values. \bar{o}_i is the average of the actual data. n indicates the size of the dataset. Comparisons of the actual and estimated values using RFR, ETR, and GBRT ML-based mathematical models are depicted in Figures 2–4. In these figures, points are estimated amount (black is used as training and the other one, red can be used as the test) and the line indicates the actual values. The results imply that the GBRT method is the most general and precise model. Consideration of the R^2 and RMSE amounts through Table 2 confirms the greater accuracy of the GBRT model with the best generality.

Table 2. Outputs.

Models	R^2 Score	MAPE
RFR	0.925	1.423×10^{-1}
ETR	0.999	7.573×10^{-2}
GBRT	0.999	7.119×10^{-2}

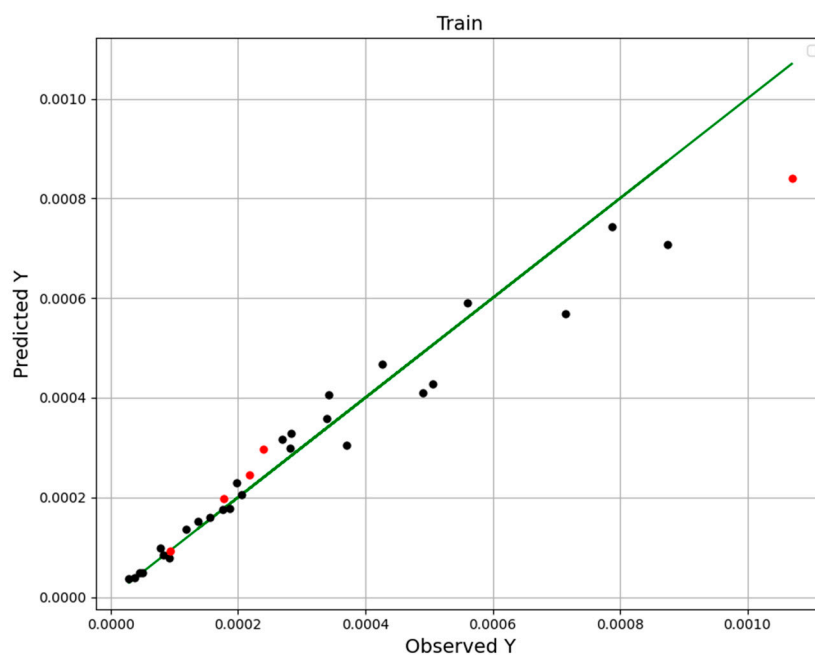


Figure 2. RFR Model: test and train data predictions.

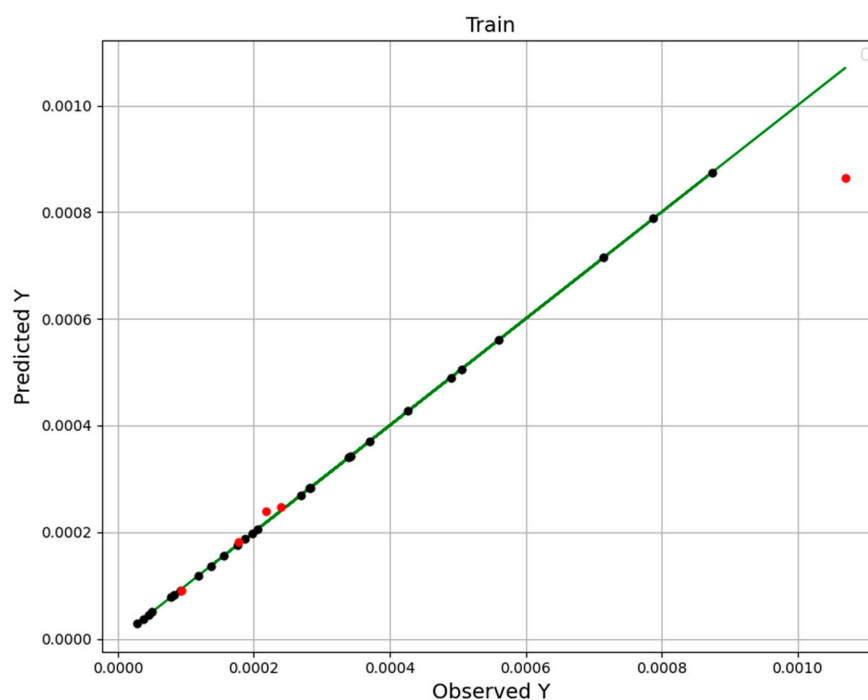


Figure 3. ETR Model: test and train data predictions.

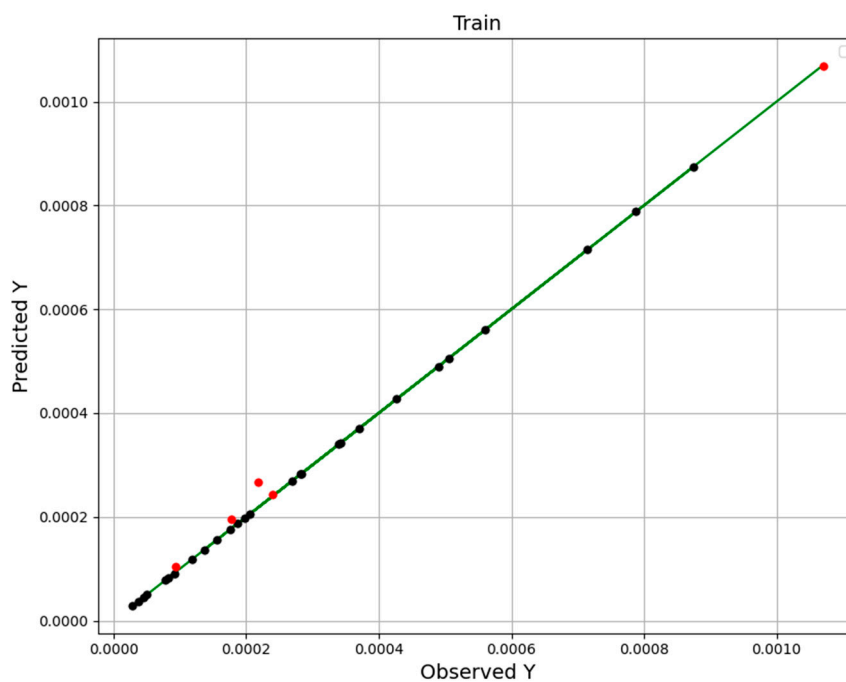


Figure 4. GBRT Model: test and train data predictions.

The reason why GBRT is superior to the other two models is because the dataset is small. Therefore, every data point can have a profound impact on the final model. In order to improve air performance, it has been shown that the boosting method, by which the points that are incorrectly predicted are corrected by weighting, is more effective than the conventional method.

Figure 5 illustrates the three-dimensional result based on the GBRT predictive model to simultaneously evaluate the effect of those two parameters, pressure and temperature as input parameters on the solubility of the anti-cancer drug decitabine as the only output.

Additionally, Figures 6 and 7 schematically demonstrate two-dimensional variations in pressure and temperature versus decitabine solubility. For all evaluated isotherms, an increase in pressure considerably improves the density of CO₂SCF due to an increase in the compaction of molecules. The increment in density resulted in an enhancement in the efficiency of the solvent and, therefore, the solubility value of decitabine in CO₂SCF increases. Despite the correlation between pressure and drug solubility, the influence of temperature is not straightforward, and a reverse alteration is seen after the 200-bar pressure point. It is worth pointing out that the solvent density and the sublimation pressure are considered as two competing parameters which entirely affect the effect of temperature on the solubility of the drug. By increasing the temperature, the density of the solvent significantly reduces, since greater molecular energy eventuates in the free movement of solvent molecules. Moreover, enhancement in the temperature of the system can improve the sublimation pressure, which presents a positive influence on drug solubility in a supercritical system. Considering the description, the net effect of the abovementioned competing factors can determine the positive or negative role of temperature on drug solubility. Analysis of the graphs illustrate a threshold pressure, the temperature increment positively encourages drug solubility owing to the significant role of sublimation pressure in comparison to density. Further, in pressure below the cross-over pressure, an increase in the temperature results in a substantial decrement in the decitabine solubility because of the decline in the solvent density. According to Table 3, 380.88 bar and 333.01 K, can be mentioned as the optimum values at highest decitabine solubility.

Table 3. Optimal values (GBRT Model).

P (Bar)	T (K)	Y
380.88	333.01	0.001073

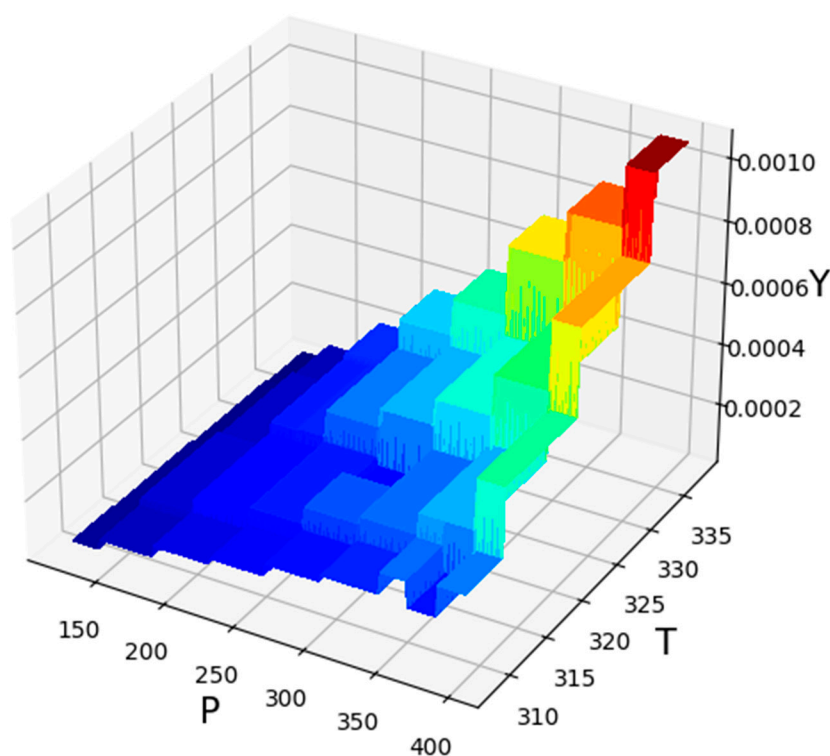


Figure 5. Three-dimensional illustration of inputs/outputs (GBRT Model).

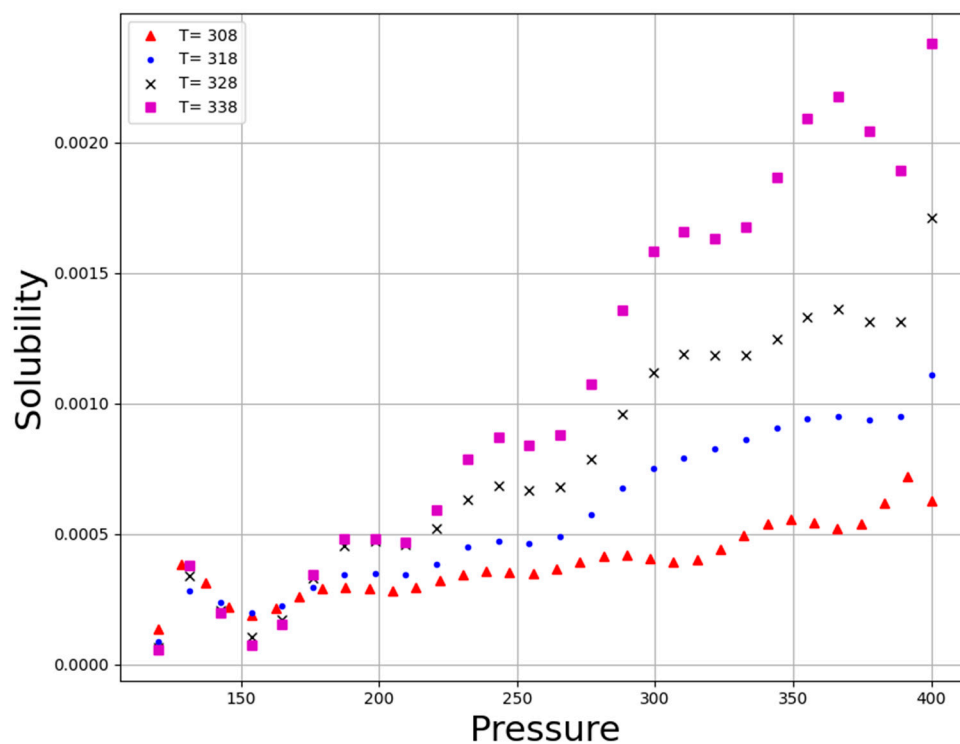


Figure 6. Tendency of P.

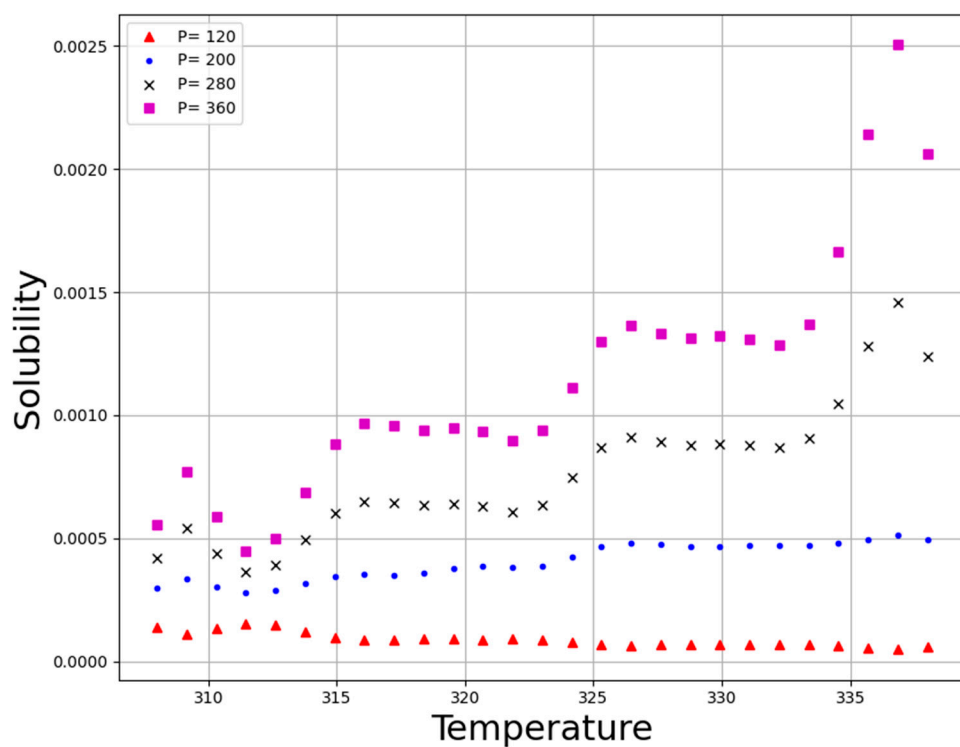


Figure 7. Tendency of T.

5. Conclusions

In this study, various predictive models were developed using AI approaches to estimate the optimum solubility value of the anti-cancer drug decitabine inside carbon dioxide supercritical fluid (CO₂SCF). We used three ensemble methods to build models on a small dataset: Random Forest (RFR), Extra Tree (ETR), and Gradient Boosted Regression

Trees (GBRT). Various configurations were tested, and optimal hyper-parameters were discovered. The final models were then evaluated using industry-standard metrics. RFR, ETR, and GBRT all had R2 scores of 0.925, 0.999, and 0.999, respectively. Furthermore, the MAPE metric error rates were 1.423×10^{-2} , 7.573×10^{-2} , and 7.119×10^{-2} , respectively. GBRT was selected as the primary method for this study through these facts and other visual considerations. The optimal values were calculated using this model as P = 380.88, T = 333.01, and Y = 0.001073.

Author Contributions: W.K.A.: conceptualization, writing, supervision, methodology, S.H.E.: writing, visualization, resources, analysis, S.A.: validation, writing, software, methodology, B.H.: editing, analysis, resources, investigation, A.A.: editing, methodology, investigation, resources, analysis, A.M.A.: supervision, writing, analysis, conceptualization, methodology, A.A.A.: software, analysis, methodology, editing, M.A.E.H.: writing, editing, analysis, visualization, resources, K.V.: analysis, editing, resources, formal analysis, K.M.A.: writing, editing, analysis, visualization, resources, M.A.S.A.: writing, editing, analysis, visualization, resources, investigation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R99), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The author would like to thank the Deanship of Scientific Research at Shaqra University for supporting this work. The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Large Groups Project under grant number (RGP.2/50/43). The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: 22UQU4290565DSR61.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available within the published paper.

Acknowledgments: This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R99), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The author would like to thank the Deanship of Scientific Research at Shaqra University for supporting this work. The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Large Groups Project under grant number (RGP.2/50/43). The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: 22UQU4290565DSR61.

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

References

1. Liu, Q.; Gilbert, J.A.; Zhu, H.; Huang, S.-M.; Kunkoski, E.; Das, P.; Bergman, K.; Buschmann, M.; ElZarrad, M.K. Emerging clinical pharmacology topics in drug development and precision medicine. In *Atkinson's Principles of Clinical Pharmacology*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 691–708.
2. Docherty, J.R.; Alsufyani, H.A. Pharmacology of drugs used as stimulants. *J. Clin. Pharmacol.* **2021**, *61*, S53–S69. [[CrossRef](#)] [[PubMed](#)]
3. Gore, S.D.; Jones, C.; Kirkpatrick, P. Decitabine. *Nat. Rev. Drug Discov.* **2006**, *5*, 891–893. [[CrossRef](#)] [[PubMed](#)]
4. Jabbour, E.; Issa, J.P.; Garcia-Manero, G.; Kantarjian, H. Evolution of decitabine development: Accomplishments, ongoing investigations, and future strategies. *Cancer Interdiscip. Int. J. Am. Cancer Soc.* **2008**, *112*, 2341–2351. [[CrossRef](#)]
5. Saba, H.I. Decitabine in the treatment of myelodysplastic syndromes. *Ther. Clin. Risk Manag.* **2007**, *3*, 807. [[PubMed](#)]
6. Xiao, J.; Liu, P.; Wang, Y.; Zhu, Y.; Zeng, Q.; Hu, X.; Ren, Z.; Wang, Y. A Novel Cognition of Decitabine: Insights into Immunomodulation and Antiviral Effects. *Molecules* **2022**, *27*, 1973. [[CrossRef](#)]
7. Senapati, J.; Shoukier, M.; Garcia-Manero, G.; Wang, X.; Patel, K.; Kadia, T.; Ravandi, F.; Pemmaraju, N.; Ohanian, M.; Daver, N. Activity of decitabine as maintenance therapy in core binding factor acute myeloid leukemia. *Am. J. Hematol.* **2022**, *97*, 574–582. [[CrossRef](#)]
8. Issa, J.-P. Decitabine. *Curr. Opin. Oncol.* **2003**, *15*, 446–451. [[CrossRef](#)]
9. Bhusnure, O.; Gholve, S.; Giram, P.; Borsure, V.; Jadhav, P.; Satpute, V.; Sangshetti, J. Importance of supercritical fluid extraction techniques in pharmaceutical industry: A Review. *Indo Am. J. Pharm. Res.* **2015**, *5*, 3785–3801.
10. Khaw, K.-Y.; Parat, M.-O.; Shaw, P.N.; Falconer, J.R. Solvent supercritical fluid technologies to extract bioactive compounds from natural sources: A review. *Molecules* **2017**, *22*, 1186. [[CrossRef](#)]

11. Tran, P.; Park, J.-S. Application of supercritical fluid technology for solid dispersion to enhance solubility and bioavailability of poorly water-soluble drugs. *Int. J. Pharm.* **2021**, *610*, 121247. [[CrossRef](#)]
12. Zhuang, W.; Hachem, K.; Bokov, D.; Javed Ansari, M.; Taghvaie Nakhjiri, A. Ionic liquids in pharmaceutical industry: A systematic review on applications and future perspectives. *J. Mol. Liq.* **2022**, *349*, 118145. [[CrossRef](#)]
13. Zhang, Q.-W.; Lin, L.-G.; Ye, W.-C. Techniques for extraction and isolation of natural products: A comprehensive review. *Chin. Med.* **2018**, *13*, 20. [[CrossRef](#)] [[PubMed](#)]
14. Huang, Z.; Juarez, J.M.; Li, X. *Data Mining for Biomedicine and Healthcare*; Hindawi: London, UK, 2017.
15. Zhang, Y.; Zhang, G.; Shang, Q. Computer-aided clinical trial recruitment based on domain-specific language translation: A case study of retinopathy of prematurity. *J. Healthc. Eng.* **2017**, *2017*, 7862672. [[CrossRef](#)] [[PubMed](#)]
16. Seddon, G.; Lounnas, V.; McGuire, R.; van den Bergh, T.; Bywater, R.P.; Oliveira, L.; Vriend, G. Drug design for ever, from hype to hope. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 137–150. [[CrossRef](#)] [[PubMed](#)]
17. Mak, K.-K.; Pichika, M.R. Artificial intelligence in drug development: Present status and future prospects. *Drug Discov. Today* **2019**, *24*, 773–780. [[CrossRef](#)]
18. Dietterich, T.G. Ensemble methods in machine learning. In Proceedings of the International Workshop on Multiple Classifier Systems, Cagliari, Italy, 21–23 June 2000.
19. Izenman, A.J. *Modern Multivariate Statistical Techniques—Regression, Classification, and Manifold Learning*; Springer: New York, NY, USA, 2008; Volume 10.
20. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
21. Freund, Y.; Schapire, R.E. Experiments with a new boosting algorithm. In Proceedings of the 13th International Conference on Machine Learning (ICML'96), Bari, Italy, 3–6 July 1996; pp. 148–156.
22. Kocev, D.; Ceci, M. Ensembles of extremely randomized trees for multi-target regression. In Proceedings of the International Conference on Discovery Science, Banff, AB, Canada, 4–6 October 2015.
23. Wehenkel, L.; Ernst, D.; Geurts, P. Ensembles of extremely randomized trees and some generic applications. In *Proceedings of the Workshop “Robust Methods for Power System State Estimation and Load Forecasting”*; Paris, France, 29–30 May 2006.
24. Pishnamazi, M.; Zabihi, S.; Jamshidian, S.; Borousan, F.; Hezave, A.Z.; Marjani, A.; Shirazian, S. Experimental and thermodynamic modeling decitabine anti cancer drug solubility in supercritical carbon dioxide. *Sci. Rep.* **2021**, *11*, 1075. [[CrossRef](#)]
25. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
26. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [[CrossRef](#)]
27. Nguyen Duc, M.; Ho Sy, A.; Nguyen Ngoc, T.; Hoang Thi, T.L. An Artificial Intelligence Approach Based on Multi-layer Perceptron Neural Network and Random Forest for Predicting Maximum Dry Density and Optimum Moisture Content of Soil Material in Quang Ninh Province, Vietnam. In *CIGOS 2021, Emerging Technologies and Applications for Green Infrastructure, Proceedings of the 6th International Conference on Geotechnics, Civil Engineering and Structures, Ha Long, Vietnam, 28–29 October 2021*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1745–1754.
28. Li, Y.; Zou, C.; Berecibar, M.; Nanini-Maury, E.; Chan, J.C.-W.; Van den Bossche, P.; Van Mierlo, J.; Omar, N. Random forest regression for online capacity estimation of lithium-ion batteries. *Appl. Energy* **2018**, *232*, 197–210. [[CrossRef](#)]
29. Peters, J.; De Baets, B.; Verhoest, N.E.; Samson, R.; Degroeve, S.; De Becker, P.; Huybrechts, W. Random forests as a tool for ecohydrological distribution modelling. *Ecol. Model.* **2007**, *207*, 304–318. [[CrossRef](#)]
30. Verikas, A.; Gelzinis, A.; Bacauskiene, M. Mining data with random forests: A survey and results of new tests. *Pattern Recognit.* **2011**, *44*, 330–349. [[CrossRef](#)]
31. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [[CrossRef](#)]
32. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [[CrossRef](#)]
33. Ahmad, M.W.; Mourshed, M.; Rezgui, Y. Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression. *Energy* **2018**, *164*, 465–474. [[CrossRef](#)]
34. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
35. Mason, L.; Baxter, J.; Bartlett, P.; Frean, M. Boosting algorithms as gradient descent. In *Advances in Neural Information Processing Systems 12*; MIT Press: Cambridge, MA, USA, 1999; Volume 12.
36. Truong, V.-H.; Vu, Q.-V.; Thai, H.-T.; Ha, M.-H. A robust method for safety evaluation of steel trusses using Gradient Tree Boosting algorithm. *Adv. Eng. Softw.* **2020**, *147*, 102825. [[CrossRef](#)]
37. Xu, Q.; Xiong, Y.; Dai, H.; Kumari, K.M.; Xu, Q.; Ou, H.-Y.; Wei, D.-Q. PDC-SGB: Prediction of effective drug combinations using a stochastic gradient boosting algorithm. *J. Theor. Biol.* **2017**, *417*, 1–7. [[CrossRef](#)]
38. Willmott, C.J. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **1982**, *63*, 1309–1313. [[CrossRef](#)]
39. Gouda, S.G.; Hussein, Z.; Luo, S.; Yuan, Q. Model selection for accurate daily global solar radiation prediction in China. *J. Clean. Prod.* **2019**, *221*, 132–144. [[CrossRef](#)]
40. Kim, S.; Kim, H. A new metric of absolute percentage error for intermittent demand forecasts. *Int. J. Forecast.* **2016**, *32*, 669–679. [[CrossRef](#)]