Biomedical Physics &
Engineering Express

Endorsed by IPEM Institute of Physics and Engineering in Medicine
PURPOSE-LED PUBLISHING™

PAPER • **OPEN ACCESS**

Machine learning models to predict the relationship between printing parameters and tensile strength of 3D Poly (lactic acid) scaffolds for tissue engineering applications

To cite this article: Duygu Ege *et al* 2023 *Biomed. Phys. Eng. Express* **9** 065014

View the article online for updates and enhancements.

You may also like

- Multi-layer collaborative stacking optimization framework for sprocket stress prediction
Fangqiang Yan, Bo Li, Rui Xia et al.
- Hybrid framework of physics-inspired optimization and explainable ensemble learning for irrigation classification mapping in Morocco
Chiranjit Singha, Satiprasad Sahoo and Ajit Govind
- Effective cardiac disease classification using FS-XGB and GWO approach
Daphin Lilda S and Jayaparvathy R

This content was downloaded from IP address 193.140.199.113 on 19/03/2026 at 06:22

Biomedical Physics & Engineering Express

PAPER

Machine learning models to predict the relationship between printing parameters and tensile strength of 3D Poly (lactic acid) scaffolds for tissue engineering applications

Duygu Ege, Seda Sertturk, Berk Acarkan and Ahmet Ademoglu

Institute of Biomedical Engineering, Bogazici University, Kandilli Campus, Uskudar, 34684 Istanbul, Turkey

E-mail: duygu.ege@boun.edu.tr

Keywords: random forest, XG Boost, 3D printing, tensile strength, poly(lactic acid), neural networks

OPEN ACCESS

RECEIVED
4 May 2023

REVISED
19 August 2023

ACCEPTED FOR PUBLICATION
31 August 2023

PUBLISHED
9 October 2023

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Abstract

3D printing is an effective method to prepare 3D scaffolds for tissue engineering applications. However, optimization of printing conditions to obtain suitable mechanical properties for various tissue engineering applications is costly and time consuming. To address this problem, in this study, scikit-learn Python machine learning library was used to apply four machine learning-based approaches which are ordinary least squares (OLS) linear regression, random forest (RF), light gradient Boost (LGBM), extreme gradient boosting (XGB) and artificial neural network models to understand the relationship between 3D printing parameters and tensile strength of poly(lactic acid) (PLA). 68 combinations of process parameters for nozzle temperature, printing speed, layer height and tensile strength were used from investigated research papers. Then, datasets were divided as training (80%) and test (20%). After building the OLS linear regression, RF, LGBM, XGB and artificial neural network models, the correlation heatmap and feature importance of each printing parameter for tensile strength values were determined, respectively. Then, the tensile strength was predicted for real datasets to evaluate the performance of the models. The results demonstrate that XGB model was the most successful in predicting tensile strength among the studied models with an R^2 value of 0.98 and 0.94 for train and test values, respectively. A close R^2 value for the train and test also indicated that there was no overfitting of the data to the model. Finally, SHAP analysis shows significance of each feature on prediction of tensile strength. This study can be extended for independent variables including nozzle pressure, strut size and molecular weight of PLA and dependent variables such as elongation and elastic modulus of PLA which may be a powerful tool to predict the mechanical properties of scaffolds for tissue engineering applications.

1. Introduction

3D printing is a technique in tissue engineering for the production of 3D structures. This technique is divided into extrusion, material jetting, and vat polymerization. Extrusion-based printing is the most commonly studied technique in the tissue engineering field [1–3].

Machine learning is a rapidly rising field and has high potential in 3D printing in terms of developing materials and processes. In fact, several studies investigated models to predict various properties of the printed object from various printing parameters [4]. Ruberu *et al* [5] optimized 3D printing parameters of methacrylated hyaluronic acid/gelatin methacrylate composite ink by machine learning. Conev *et al* [6] utilized RF regression to determine effect of different 3D printing parameters on 3D printed poly (propylene fumarate). Menon *et al* [7] used a hierarchical machine learning framework to produce biomaterials with high print fidelity at a high printing speed.

With machine learning technique, mechanical properties may be predicted effectively before many trial and errors [2]. In this study, for the first time, ordinary least squares (OLS) linear regression, random forest (RF), extreme gradient boosting (XGB), light boosting (LGBM) and artificial neural network models were used to predict tensile strength from nozzle temperature, printing speed, infill density and layer

© 2023 © 2022 The Author(s). Published by IOP Publishing Ltd

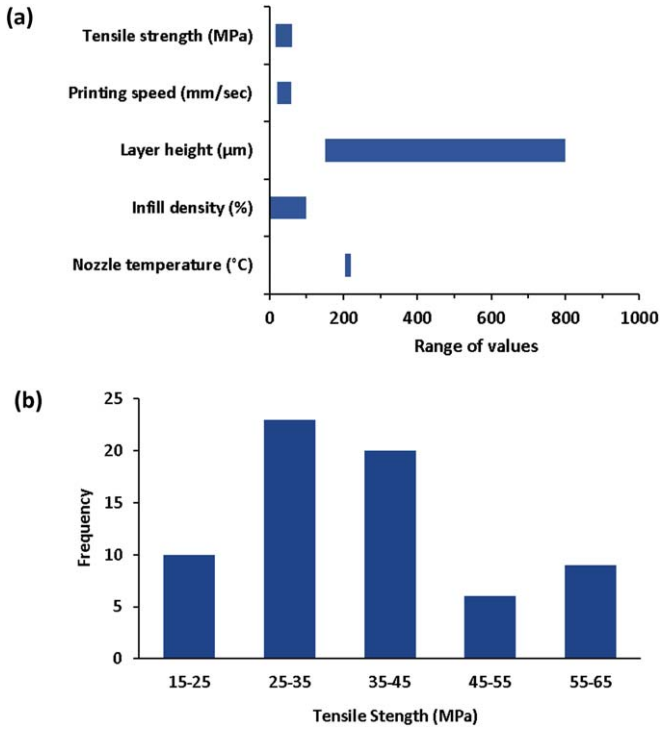

Figure 1. Shows (a) the variation of the dependent variables and the independent variable for the experiments and (b) distribution of tensile strength values.

Table 1. Average values, standard deviation, standard error of temperature, infill density, layer height, printing speed and tensile strength.

	Nozzle temperature (°C)	Infill density (%)	Layer height (mm)	Printing speed (mm/s)	Tensile strength (MPa)
Mean	214.2	80.5	0.28	47.7	36.2
Std.	5.6	29.1	0.18	14.0	12.0
Min.	205	0	0.15	20.0	15.9
25%	210	80	0.15	40.0	31.1
50%	215	100	0.25	50.0	35.1
75%	215	100	0.30	60.0	40.4
Max.	225	100	0.80	80.0	60.3
No. of values	5	11	8	5	68

height. The feature importance of the independent variables and correlation heatmap were determined to study the strength of the relationship between independent variables and tensile strength. Overall, XGB was found to be the most effective model to predict tensile strength among the studied algorithms.

2. Methodology

2.1. Data collection

Datasets of nozzle temperature, printing speed, layer height and tensile strength were obtained from four research papers [8–11]. Overall, this dataset covered 68 possible combinations of printing speed (40–80 mm s^{-1}), nozzle temperature (205 °C–220 °C), layer height (0.15–0.8 mm), infill density (2%–100%) and tensile strength of PLLA. Nozzle temperature is the temperature that the material inside the nozzle is held while printing. Layer height is the distance between the needle and the deposition layer. Printing speed is the distance deposited per second. Infill density is the percentage of the material being deposited [3]. Figures 1(a) and (b) show the variation of the dependent variables and the independent variable for the experiments and distribution of tensile strength values, respectively. Table 1 shows the average and standard deviation of the used parameters and the tensile strength and table 7 is presented in the supplementary file which shows all the processing conditions and the resultant tensile strength.

2.2. Computational modeling

All machine learning algorithms were implemented using Python with Pandas, Numpy, Scipy, Matplotlib, Seaborn and Scikit-learn. Five different models were

Table 2. Comparison of regression techniques [18, 23, 28–30, 43].

Model	Advantage	Disadvantage
Linear regression	Easy to run, fast, easy to interpret.	Inaccurate for polynomial relationships.
RF	Ensemble model, reduces overfitting, normalization is not required.	Resultant model is difficult to read slower performance, high GPU usage.
XGB	Ensemble model fast performance, very high accuracy, easy to use, prevents overfitting, works well with small dataset.	Relatively more complex, prone for overfitting if not well tuned, hypertuning may be time-consuming
LGBM	Ensemble model, fast performance.	Less accurate than XGB, More suitable for for larger datasets.
Neural networks	Suitable for complex	Not suitable for small data sets. non-linear relationships.

used in the study which were OLS linear regression, RF, XGB, LGBM and artificial neural network regression [12]. The Python codes used are provided at https://github.com/duyguege/machine-learning.git.

2.3. Linear regression

OLS is a linear regression technique to find the best-fitting line between the independent and dependent variables. In our model, there is a single dependent and four independent variables in which the OLS method minimizes the sum of squared differences between the actual data and their prediction values as given in equation (1) [13].

$$\min \sum_i (Y_i - y_i)^2 \quad (1)$$

where Y_i and y_i are actual data and its prediction, respectively.

2.4. RF regressor

RF was first developed by Breiman *et al* in 2001 which is a decision tree based ensemble learning algorithm [14–17]. RF is a bagging method (bootstrap aggregation) which builds independent estimators (trees) and the mean is found to make a prediction [18]. In each node of the tree, one feature is selected at each node and splitting is made to produce a sub-node which is called a child node. The splits are done by minimizing variance by searching through splits. When splitting cannot be performed any longer, the node is named a leaf node. In RF, instead, a set of decision trees are constructed. Then, the final prediction is made by the average of each prediction which improves accuracy compared to decision tree regression [13, 14, 19].

2.5. XGB regressor

This ensemble gradient boosting algorithm was devised in 2016 by Chen and Guestrin [20, 21]. In this model, trees are added sequentially after evaluation of the previous tree [22]. This way, weak learners are trained to obtain a strong learner [23]. Each leaf node is given a score and each score is added up to determine the predicted value [17]. XGB is a scalable tree boosting system with high efficiency and prediction accuracy consisting of many decision trees and is typically used in the field of regression. XGB uses second-order Taylor expansion to the loss function and normalization is utilized in the objective function to prevent over-fitting to outliers [20, 21, 23–27].

2.6. LGBM regressor

This is another gradient boosting ensemble of weak learners used to improve model efficiency technique which uses decision tree-based algorithm [13, 23]. It is a faster and more efficient algorithm than XGB as it leads to leafwise growth rather than level wise growth. Each output feature has been predicted by a separate LGBM regressor tree wherein, each leaf in such a tree ends in a regression [18, 23, 28, 29]. It has a higher tendency to overfitting which is hindered by regularization [23]. Figure 2 shows the difference of tree growth of XGB and LGBM boosting models and table 2 shows the comparison of the methods applied in this work.

2.7. Artificial neural networks

Artificial neural network is a model influenced by brain's neural networks which are organized in layers. Neurons obtain signals and signals are processed by Rectified linear unit (Relu) activation function. The output is then transmitted to the following layer of neurons. Training neural networks with data tunes weight of connections and refine prediction [30].

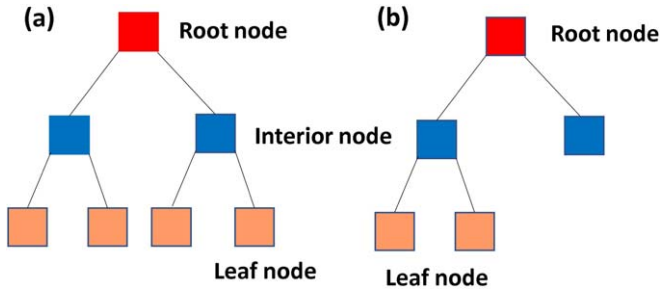

Figure 2. Difference of tree growth of (a) XGB and (b) LGBM boosting regression models.

Table 3. Hyperparameters and their best parameters for tuning by using RF, LGBM and XGB models (best parameters are given in bold).

Hyperparameters	Subsample ratio of columns	Number of estimators (number of trees)	Maximum depth	Learning rate
RF	—	30, **100**	**1**,6	—
LGBM	**0.4**, 0.5, 0.6	30, **100**	**2**, 3, 4	**0.1**, 0.01, 0.5
XGB	**0.4**, 0.5, 0.6	30,100, **200**	2, 3, **4**	**0.1**, 0.01, 0.5

2.8. Training, hyper tuning and validation processes

In this study, the dataset was split into training (80%) and test partitions (20%). The training dataset was used for model building and the test dataset was used for its evaluation. For application of the models, Scikit-learn library from Python was used. For decision tree and gradient boosting regression models (RF, XGB and LGBM), hyperparameter tuning was performed for subsample ratio of columns, number of estimators, maximum depth and learning rate (shrinkage factor) after determination of the best parameters which is shown in table 3. For artificial neural networks alpha and hidden layer sizes were selected from values of 0.1, 0.01, 0.02, 0.005 and (20,20), (100,50,150) and (300,200,150). The best parameters were found as 0.005 and (100,50,150) respectively for alpha and hidden layer sizes. The best parameter for both training and test sets was selected to prevent underfitting and overfitting [25]. 10-fold cross-validation was used for performance assessment which randomly splits the dataset into ten equal folds [31]. Then, the model was again applied in the final stage to determine the predicted values of the tensile strength [32].

2.9. Correlation heatmap

A correlation heatmap was produced by Seaborn module of Phyton to determine the strength of relationship between the independent variables (nozzle temperature, printing speed, layer height and infill density) and the dependent variable (tensile strength) [25, 33]. A higher correlation coefficient (r) indicates multicollinearity between independent variables [18].

2.10. Feature importance

The features importance was determined using an embedded function in the Scikit-learn implementation of the RF model. Then the features were ranked according to their importance [6, 34].

2.11. Determination of the model performance

Higher R^2 values with root mean square error (RMSE) and lower mean absolute error (MAE) values imply a higher success rate of the models [35, 36]. The coefficient of determination (R^2), calculation of MAE and RMSE were used to evaluate model performance in equation (1) [37]. The R^2, MAE, RMSE are given in equations (2), (3) and (4), respectively.

$$R^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - y_i| \tag{3}$$

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - y_i)^2\right)^{1/2} \tag{4}$$

where n is the total number of the samples, Y_i is the actual data, y_i is its prediction y_i is the test value and whose mean values are denoted by $\bar{Y}$ and $\bar{y}$, respectively. For R^2, the range of values for degree of success are given in table 4. The Y_i versus y_i curves are given in figure 3 for each model.

2.12. Shapley additive explanation

Lundberg and Lee [38] recently developed Shaply Addive explanation (SHAP) technique in 2017 which enables analysis of complex relationships in machine learning models. In this study, SHAP which is a model interpretation package in Phyton is utilized to further understand

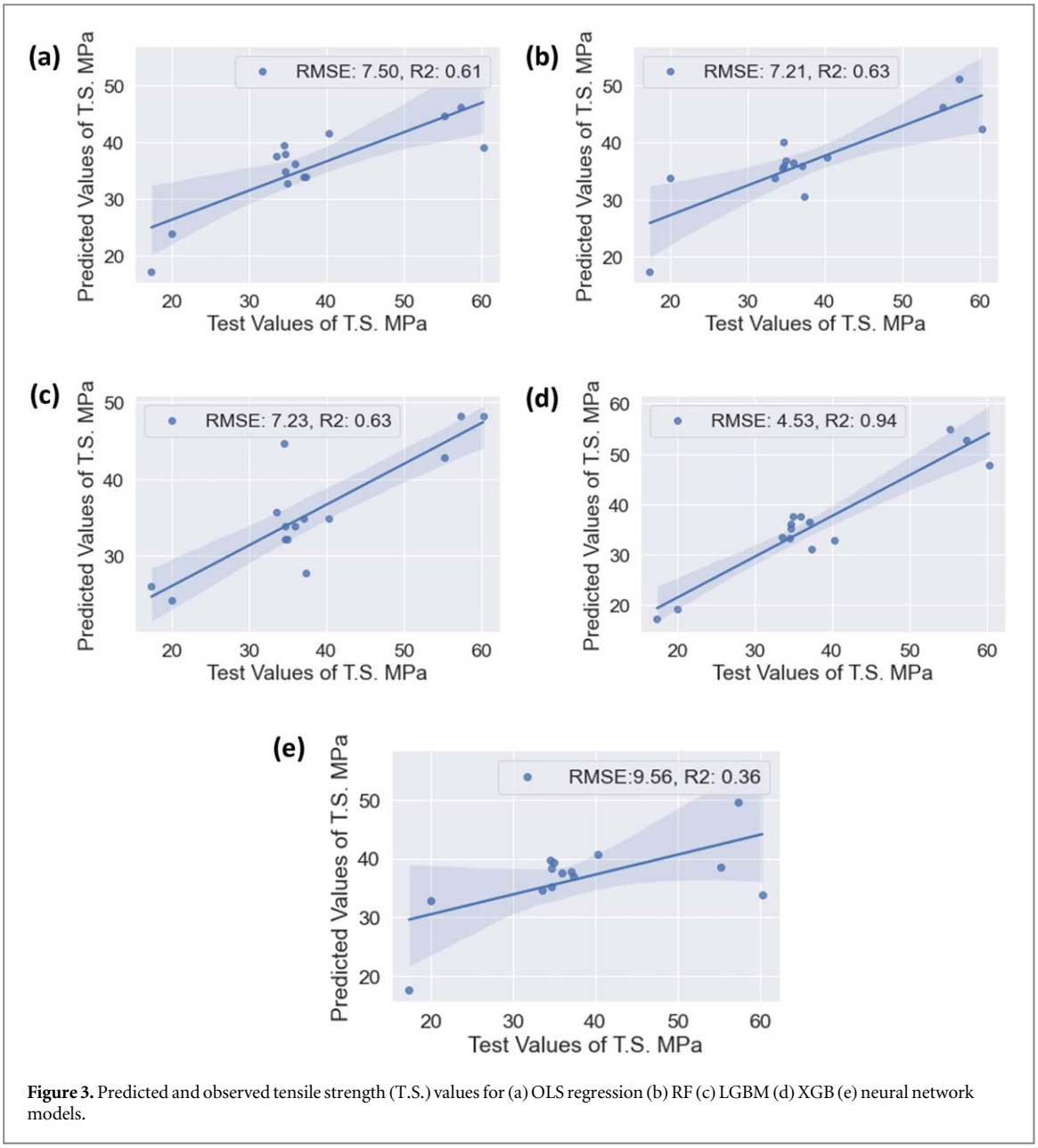

Figure 3. Predicted and observed tensile strength (T.S.) values for (a) OLS regression (b) RF (c) LGBM (d) XGB (e) neural network models.

Table 4. Indication of R2 values [18].

R^2 values	Indication
0.6	unsatisfactory prediction
0.6-0.75	satisfactory prediction
0.75-0.95	good prediction
0.95	excellent prediction

the marginal relationship between predicted tensile strength value and each feature value. SHAP value is calculated by average prediction made with contribution of each feature value minus prediction made without a feature value. A negative and positive SHAP value indicates negative and positive contribution of the feature value for prediction of tensile strength, respectively. SHAP summary plot is drawn to show contribution of each value on tensile strength. Then SHAP dependence plot is drawn to show effect of each feature value (infill density, layer height, printing speed and nozzle temperature), separately. Primary y-axis shows the SHAP value and secondary y-axis shows a color bar illustrating the high feature values in red and low feature values in blue.

2.13. Statistics

Quantitative data were reported as mean value ± standard error from at least three independent experiments. Statistical differences between groups were analyzed using the two-way analysis of variance (ANOVA) statistical test, with Tukey's pairwise post hoc test. Statistical significance was represented as $p < 0.05$ (in comparison to the test values).

3. Results and discussion

In this study, the data was fitted to OLS linear regression, RF, LGBM, XGB and artificial neural

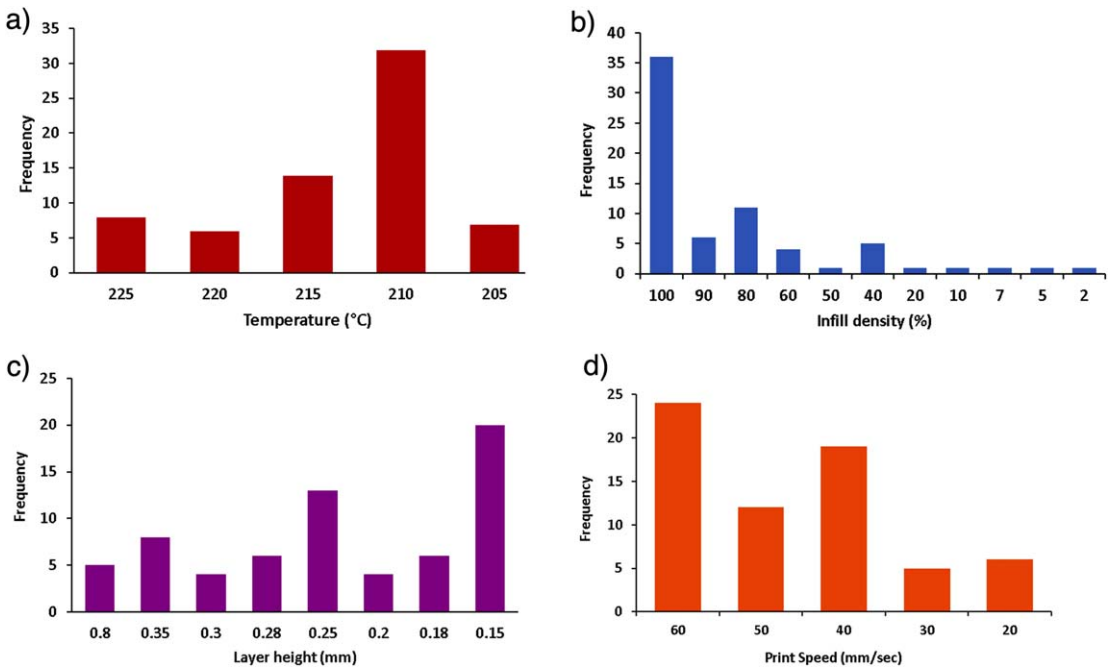

Figure 4. Variation of the independent variables (a) temperature, (b) infill density, (c) layer height and (d) print speed.

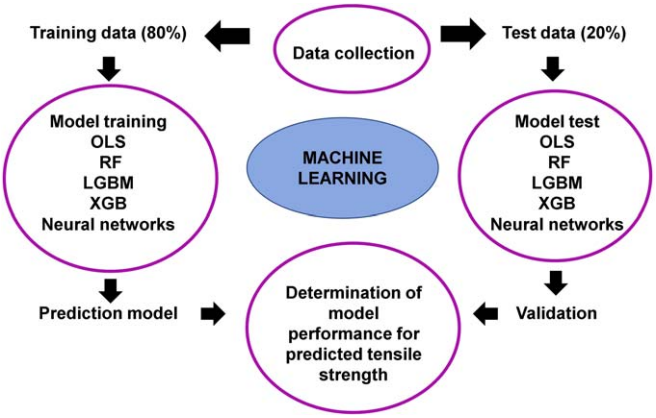

Figure 5. The summary of the study [37].

network models than the performance of the models were tested. Having homogeneous distribution of the independent variables is important to build a successful model. Figure 4 shows the variation of the independent variables and their frequencies.

According to figure 4, overall, the variables have quite a large variation. However, for infill densities, the variation is limited in the range lower than 60%. For printing speed, the data is larger for above 40 mm s^{-1}. We may assume that a more homogeneous distribution of the data may be better for more successful predictions. Figure 5 shows the summary of the study. Using Seaborn in Phyton, a correlation heatmap is produced which is used to evaluate the correlation between printing parameters and tensile strength. In figure 6, the correlation heatmap presents the correlation between the components in the matrix, in which the *r* value is the correlation coefficient. A larger correlation coefficient value indicates a stronger relation between these two components.

The correlation heatmap shows that printing speed had a negative correlation with tensile strength which indicated that as printing speed reduced, the tensile strength tended to increase. Among other parameters, infill density showed the highest correlation with the tensile strength following with the layer height and nozzle temperature. Additionally, correlation coefficient was lower than 0.8 between each independent variable which indicated that there was no collinearity between independent variables [18].

Feature importance was also determined from ensemble models. The ranking of the independent variables based on the feature importance function of Scikit-learn is shown in figure 7.

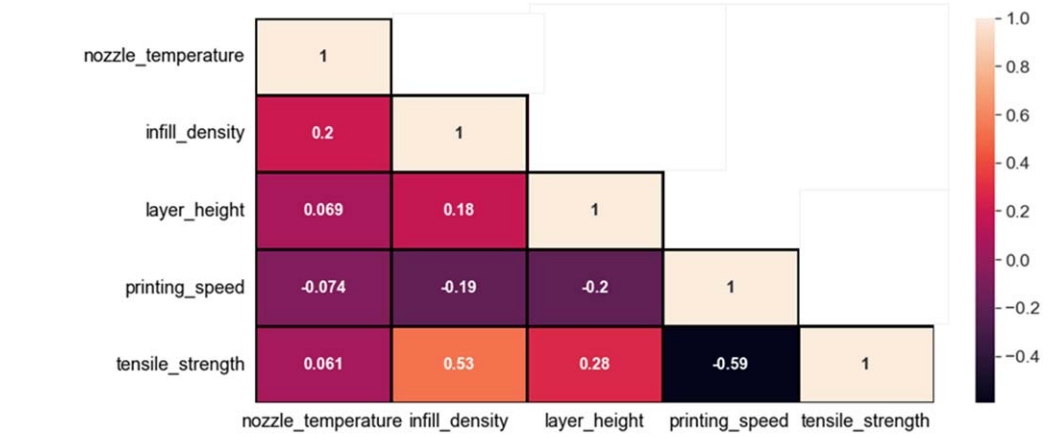

Figure 6. Correlation heatmap for 3D printing parameters and tensile strength.

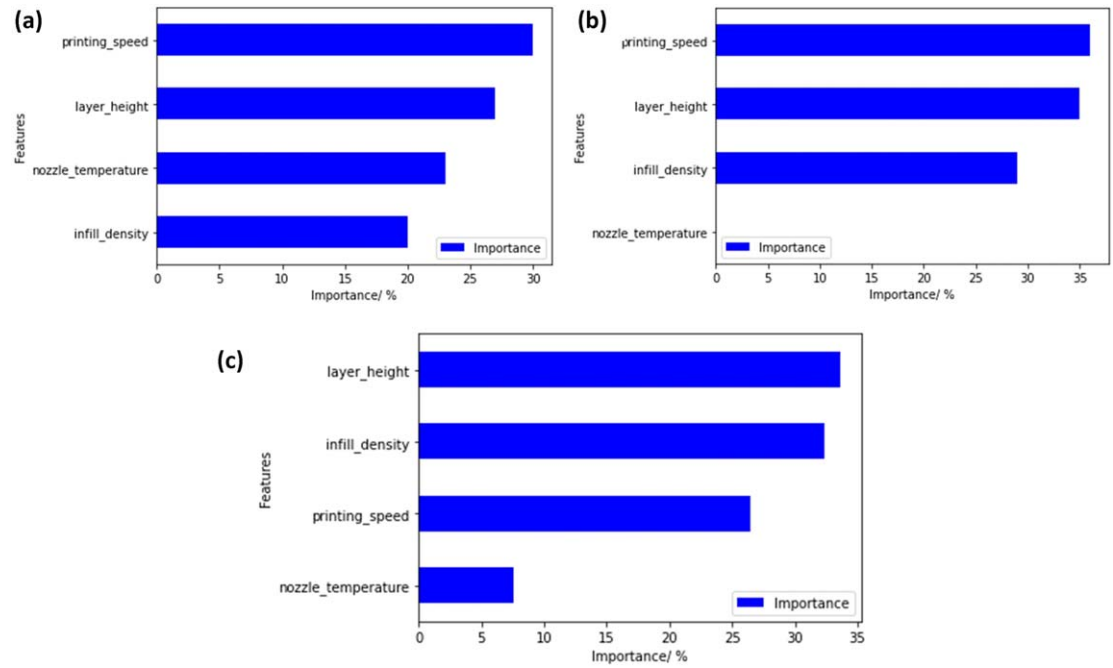

Figure 7. Feature importance which is implemented from (a) RF, (b) LGBM and (c) XGB models.

As shown in figures 6 and 7, a slower printing speed increases resolution and leads to better accuracy of the constructed scaffolds. A faster printing rate may lead to formation of artifacts which ultimately reduces quality of the printed constructs [39]. As layer height increased, a higher tensile strength was also observed [40]. A direct relationship between infill density and increase of tensile strength is expected [41]. Nozzle temperature also affects the tensile strength due to the relation of PLA temperature and its viscosity [42]. After running the models, R^2 train, R^2 test, MAE, and RMSE values were found for the OLS linear regression, RF, XGB and LGBM models as presented in table 5. Figure 8 shows (a) the comparison of R^2 train and test values and (a) MAE, RMSE for the test values. Since R^2 is more convenient for comparing suitability of the models, more focus was given to this parameter [23]. For training, R^2 value was highest for XGB model followed by RF, OLS and LGBM models. For testing, R^2 was highest for XGB followed by RF, LGBM and OLS linear regression models. According to table 5 and figure 8 (a), RF had R^2 of higher than 0.6 for both training and test data; however, R^2 for train was much higher than test which indicated overfitting. For linear regression and LGBM, R^2 values were lower than 0.6 for the training dataset which implied that the models were unsatisfactory [18]. Additionally, for OLS and LGBM regression model, R^2 for test was higher than training, which showed that test data fitted better to the model than the training data.

For XGB, R^2 for training and test dataset were above 0.94 and they were almost the same. This

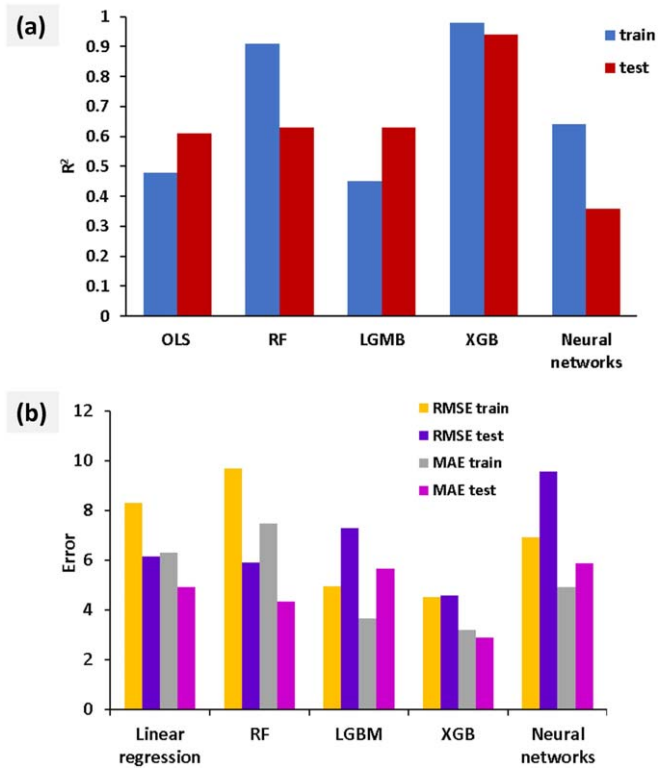

Figure 8. Comparison of (a) R^2 of train and test values and (b) MAE, RMSE for test values of OLS, RF, XGB, LGBM and neural networks models.

Table 5. MAE test , R^2 and RMSE test data for different models.

Model	R^2 training	R^2 test	MAE test	RMSE test	Success level
OLS Linear Regression	0.48	0.61	4.93	6.16	Unsatisfactory R^2 of train low
RF	0.91	0.63	4.34	5.91	Unsatisfactory R^2 of test low
LGBM	0.45	0.63	5.67	7.30	Unsatisfactory R^2 of train low
XGB	0.98	0.94	2.89	4.53	Excellent
Neural networks	0.64	0.36	5.87	9.56	Unsatisfactory

indicated that the XGB model worked very successfully for the dataset. XGB model also led to low RMSE and MAE values which further supported that the performance of XGB was highest for prediction of tensile strength. LGBM had higher RMSE and MAE values for test than train data which might further indicate that it led to overfitting. Literature also indicates that LGBM had a worse prediction than XGB as it is more prone to overfitting [23]. Furthermore, artificial neural networks had poor prediction of tensile strength which is indicated from the low R^2. Neural networks also led to overfitting as R^2 of train was much higher than R^2 of the test. This is because of relatively poor performance of artificial neural network models with small datasets. [31]. Figure 3 compares the predicted values and the experimental results for OLS, random forest, LGBM and XGB models. Artificial neural networks is eliminated at this stage due to its low R^2. Figure 3 shows that there are outliers in all models. For all models, most of the data points distribute around the diagonal, while some values deviate from the diagonal line. This deviation is most likely attributed to the limited datasets which can be improved with a larger variation of dataset. Also, data is gathered from four different papers. Although 3D printing technique is kept as exrusion-based and the same polymer is used for all studies, gathering data from a single study may lead to improved predictions. This way, molecular weight, nozzle diameter and the model of the extrusion based printer may be also controlled. However, currently, in the literature, it is not possible to find large enough dataset from a single source. Figure 9 shows the average test value and average respective predictions of tensile strength with the practiced models for low,

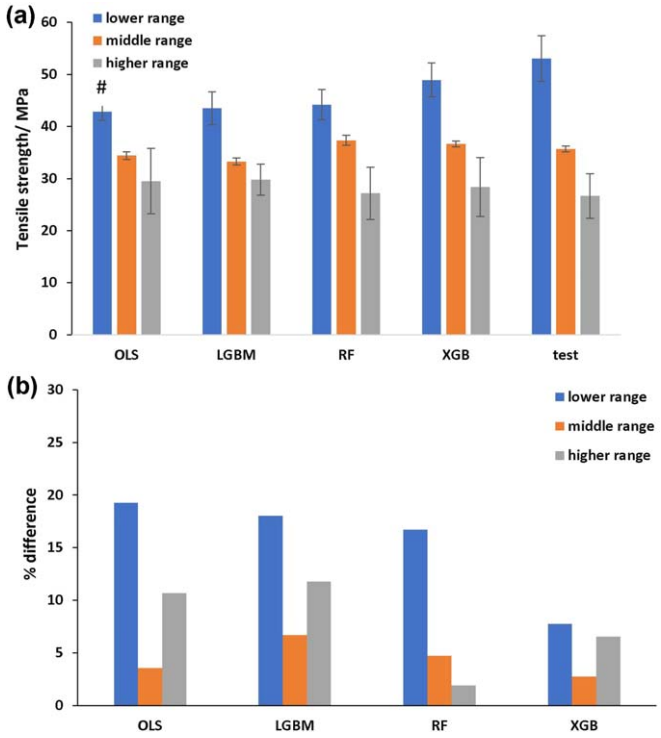

Figure 9. Test value and respective predictions of tensile strength with the practiced models and (b) % difference with tensile strength ($p < 0.05$ in comparison to the control group).

Table 6. % Difference with test value and the models' predicted tensile strength values for the lower, medium and higher range.

% Difference with test values	OLS	LGBM	RF	XGB
Lower range Index values (27, 8, 44, 39)	1.9	11.8	2.0	6.5
Medium range Index values (25, 4, 12, 5)	3.5	6.7	4.7	2.73
Higher range Index values (64, 60, 59, 36)	19.3	18.0	16.7	7.8

medium and high value of tensile strength. Table 6 shows the index number of the independent and dependent variables chosen from each (lower, medium and higher) range from table 7. Figure 9(a) shows that the average predicted values of the tensile strength had quite similar values with the tensile strength. Only significant difference was for predicted tensile strength from OLS regression from the lower range test values. From figure 9(b), it is clearly observed that XGB led to lower % difference compared with the test values of tensile strength compared to other models. For models, the prediction is worse for lower and higher values of tensile strength. This could be due to a lower number of data points for lower and higher tensile strength values compared to middle-range tensile strength values. Moreover, the dataset might not be suitable for linear regression. The improved prediction for XGB could be accounted for the loss function it employs for minimization and the second-order Taylor approximation normalizing the weak learners in ensemble modeling [23]. Table 6 shows the % difference between the test values and the model predictions. Figure 10 shows distribution of experimental and predicted values after executing OLS, RF, XGB and LGBM models depending on the index number in table 7.

To increase the accuracy of the models, further experimentation may be required focusing on collecting data with larger coverage of printing speed, layer height and nozzle temperature on examining a larger variety of material compositions. Such additional data would improve machine learning approaches. It is also important to note that our dataset was obtained from four different publications which might increase the variability than using a single source of dataset. If the dataset is generated from a single study, a higher accuracy may be achieved using the applied models. Figures 11 and 12 show the SHAP summary plot and dependence plot, respectively. In figure 11, SHAP summary plot shows impact of each feature value on prediction of tensile strength. The high values of the features are given as red and low values of the feature are given as blue.

According to shap summary plot, low values of infill density have negative contribution on the prediction. The high values of infill density have positive influence on the prediction. This is possibly because there is more data for high infill density values in the dataset. Layer height is observed to have negative influence for most of the values. For printing speed, high values have negative influence and high values

Table 7. 3D printing parameters and resultant tensile strength values of PLA [8, 11, 44].

Index	Nozzle temperature	Infill density	Layer height	Printing speed	Tensile strength	References
0	205	100	0.15	40	34.53	[8]
1	205	100	0.15	50	35.10	[8]
2	205	100	0.15	60	36.37	[8]
3	205	100	0.25	40	35.87	[8]
4	205	100	0.25	50	35.93	[8]
5	205	100	0.25	60	34.90	[8]
6	205	100	0.35	40	37.90	[8]
7	205	100	0.35	50	37.13	[8]
8	215	100	0.15	40	28.83	[8]
9	215	100	0.15	50	33.47	[8]
10	215	100	0.15	60	31.07	[8]
11	215	100	0.25	40	34.17	[8]
12	215	100	0.25	50	34.70	[8]
13	215	100	0.25	60	33.67	[8]
14	215	100	0.35	40	35.07	[8]
15	215	100	0.35	50	36.17	[8]
16	215	100	0.35	60	34.70	[8]
17	225	100	0.15	40	35.60	[8]
18	225	100	0.15	50	25.87	[8]
19	225	100	0.15	60	29.33	[8]
20	225	100	0.25	40	32.17	[8]
21	225	100	0.25	50	32.40	[8]
22	225	100	0.25	60	32.93	[8]
23	225	100	0.35	40	36.50	[8]
24	225	100	0.35	50	36.17	[8]
25	210	40	0.35	30	37.12	[11]
26	210	40	0.15	30	37.72	[11]
27	210	60	0.15	40	33.96	[11]
28	210	80	0.15	50	37.35	[11]
29	210	100	0.15	60	45.47	[11]
30	210	40	0.2	40	34.37	[11]
31	210	60	0.2	30	41.63	[11]
32	210	80	0.2	60	26.26	[11]
33	210	100	0.2	50	40.00	[11]

Table 7. (Continued.)

Index temperature	Nozzle density	Infill height	Layer speed	Printing strength	Tensile	References
34	210	40	0.25	50	27.22	[11]
35	210	60	0.25	60	18.45	[11]
36	210	80	0.25	30	40.36	[11]
37	210	100	0.25	40	39.57	[11]
38	210	40	0.30	60	15.98	[11]
39	210	60	0.30	50	17.27	[11]
40	210	80	0.30	40	19.94	[11]
41	210	100	0.30	30	34.47	[11]
42	210	2	0.15	60	16.10	[9]
43	210	5	0.15	60	17.49	[9]
44	210	7	0.15	60	17.18	[9]
45	210	10	0.15	60	17.37	[9]
46	210	20	0.15	60	19.02	[9]
47	210	50	0.15	60	20.14	[9]
48	210	80	0.15	60	25.12	[9]
49	210	100	0.15	60	32.00	[9]
50	210	80	0.8	20	31.26	[10]
51	215	90	0.18	40	46.62	[10]
52	220	100	0.28	40	54.97	[10]
53	210	90	0.8	60	39.51	[10]
54	215	100	0.18	20	52.55	[10]
55	220	80	0.28	40	43.92	[10]
56	215	80	0.8	40	34.82	[10]
57	220	90	0.18	60	41.08	[10]
58	210	100	0.28	20	59.38	[10]
59	220	100	0.8	40	55.18	[10]
60	210	80	0.18	60	56.39	[10]
61	210	90	0.28	20	57.39	[10]
62	215	100	0.8	60	53.91	[10]
63	220	80	0.18	20	58.44	[10]
64	210	90	0.28	40	60.29	[10]
65	220	90	0.8	20	56.12	[10]
66	210	100	0.18	40	58.98	[10]
67	215	80	0.28	60	56.50	[10]

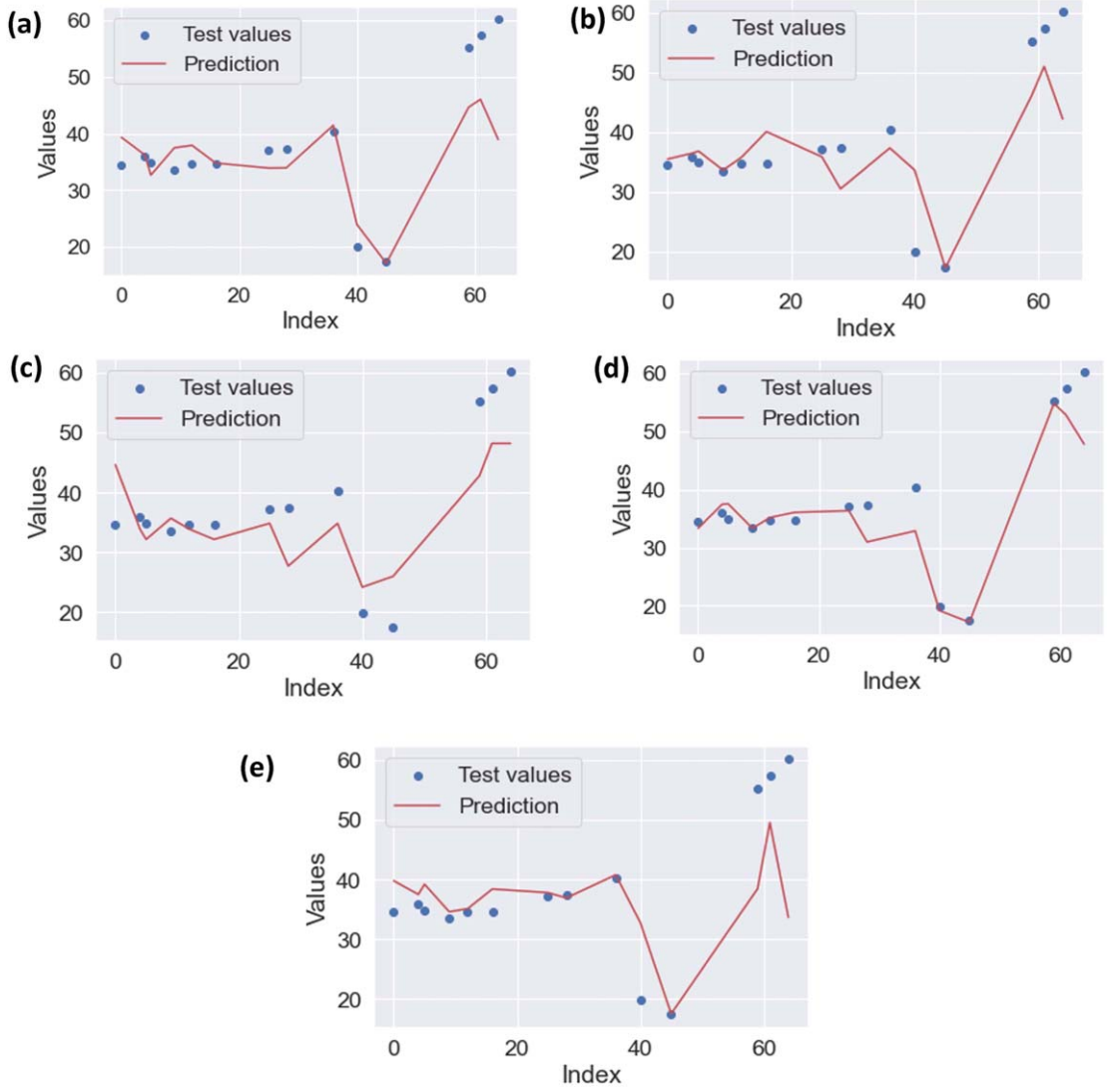

Figure 10. Distribution of experimental and predicted values (a) OLS (b) RF (c) XGB (d) LGBM (e) neural network models.

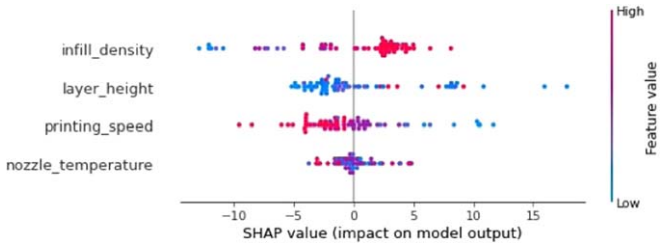

Figure 11. Shap summary plot based on XGB model.

have positive influence on prediction. Nozzle temperature has mostly positive influence on the prediction at some of the high values however there is no obvious trend observed for this feature.

In figure 13, tensile strength was predicted by using XGB model for different values of nozzle temperature for constant printing speed, layer height and infill density. According to figure 13, XGB model predicted that the tensile strength would increase if the processing temperature of PLA decreased to 220 °C from 225 °C. However, further reduction of processing temperature of PLA would lead to reduction of tensile strength. This information would be beneficial for the determination of printing parameters of PLA. For bone tissue engineering applications, 60% in fill density is suitable for osteogenic differentiation [3].

Therefore while keeping infill density as 60%, nozzle temperature as 220 °C and printing speed as 20 mm s^{-1}, tensile strength was predicted for different layer heights by XGB model as in figure 14. According

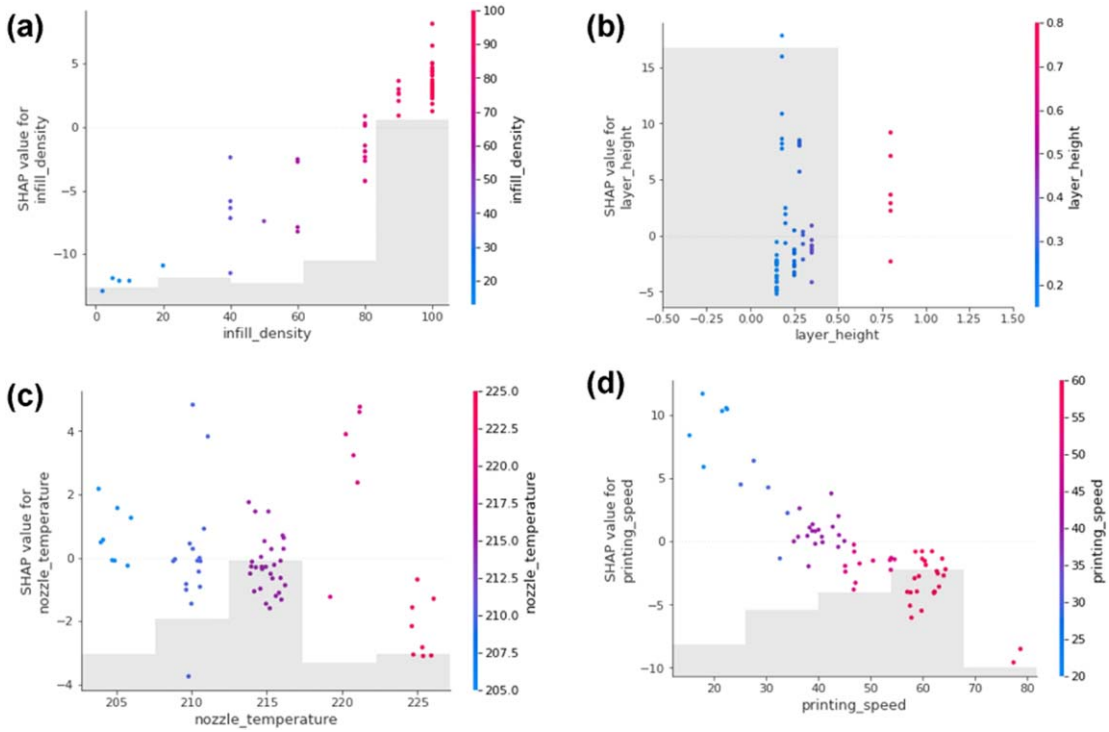

Figure 12. SHAP dependence plot based on XGB model.

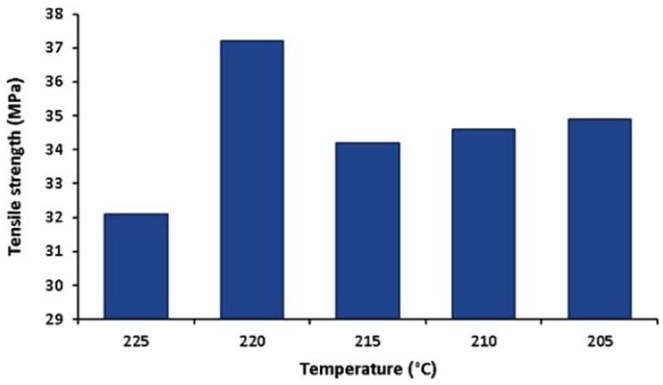

Figure 13. Prediction of tensile strength for different nozzle temperature at constant printing speed (20 mm sec^{-1}), infill density (60%) and layer height (0.25 mm).

to figure 14, the highest tensile strength was achieved for 0.6 mm and 0.8 mm layer height while printing speed, nozzle temperature and infill density were constant. To achieve the highest tensile strength, it is predicted that 3D printing of PLA with 60% infill density at 220 °C with a printing speed of 20 mm sec^{-1} at a layer height of 0.6 mm may be suitable for bone regeneration applications. In the future, adaptive Boosting (Adaboost) may also be applied as a promising approach for a prediction with higher accuracy. Moreover, deep learning may be also applied for larger datasets to make predictions of mechanical properties of biomaterials for different 3D printing parameters. This study demonstrates the potential benefit of machine learning to form the core of a guidance system for determining suitable printing parameters, thus reducing experimentation [6].

4. Conclusion

In this study, datasets were collected from four different publications on effect of nozzle temperature, layer height, printing speed and infill density on tensile strength. Then, OLS linear regression, RF, LGBM, XGB and artificial neural network models were applied to test their accuracy to predict tensile strength. R^2, RMSE and MAE values for all models indicated that XGB regression yielded a more

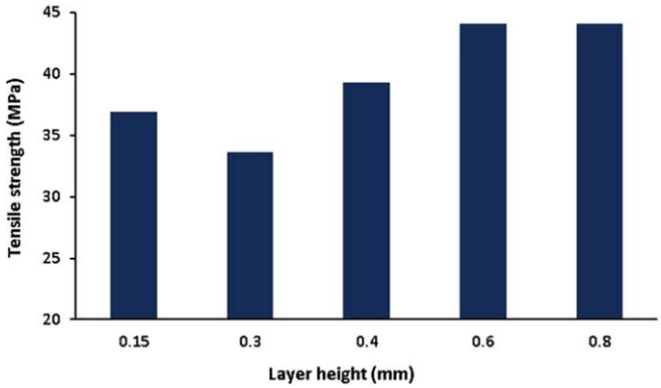

Figure 14. Prediction of tensile strength for different layer heights at constant printing speed (20 mm sec^{-1}), infill density (60%) and nozzle temperature (220 °C).

successful prediction of tensile strength values. Furthermore, SHAP analysis indicated that high infill densities and low printing speed had a positive influence on the prediction of tensile strength. A large dataset may improve the accuracy of the predictions. Also, a wider range of independent variables could be used to reveal the importance of each printing parameter such as infill pattern, gauge length, bed temperature, molecular weight of PLA. Finally, in addition to tensile strength, other important parameters qualifying the mechanical properties may be also included such as elongation and Youngs modulus. Overall, machine learning is a promising tool to predict tensile strength values and reduce the number of tests required for optimization of 3D printing parameters for tissue engineering applications.

Acknowledgments

We would like to thank Bogazici Research Fund (18001) for supporting this study.

Data availability statement

The data cannot be made publicly available upon publication because no suitable repository exists for hosting data in this field of study. The data that support the findings of this study are available upon reasonable request from the authors.

Supplementary data

Table 7 shows the parameters used, tensile strength and the papers used to extract this data.

ORCID iDs

Duygu Ege https://orcid.org/0000-0002-9922-6995
Seda Sertturk https://orcid.org/0000-0002-8853-7587

References

[1] Jain S, Fuoco T, Yassin M, Mustafa K and Finne-Wistrand A 2019 Printability and critical insight into polymer properties during direct-extrusion based 3d printing of medical grade polylactide and copolyesters *Biomacromolecules* **2019** 388–96
[2] Fu Z, Angeline V and Sun W 2021 Evaluation of printing parameters on 3d extrusion printing of pluronic hydrogels and machine learning guided parameter recommendation *International Journal of Bioprinting* **7** 343
[3] Ege D and Hasirci V 2023 Is 3D printing promising for osteochondral tissue regeneration? *ACS Applied Bio Materials* **6** 1431–44
[4] Yu C and Jiang J 2020 A perspective on using machine learning in 3d bioprinting *International Journal of Bioprinting* **6** 253
[5] Ruberu K, Senadeera M, Rana S, Gupta S, Chung J, Yue Z, Venkatesh S and Wallace G 2021 Coupling machine learning with 3d bioprinting to fast track optimisation of extrusion printing *Applied Materials Today* **22** 100914
[6] Conev A, Litsa E, Perez M, Diba M, Mikos A and Kavraki L 2020 Machine learning guided 3d printing of tissue engineering scaffolds *Tissue Engineering. Part* A **26**
[7] Menon A, Poczos B, Feinberg A and Washburn N 2019 Optimization of silicone 3d printing with hierarchical machine learning *3D Printing and Additive Manufacturing* **6** 181–9
[8] Gao G, Xu F and Xu J 2022 Parametric optimization of fdm process for improving mechanical strengths using taguchi method and response surface method: a comparative investigation *Machines* **10** 750
[9] Nguyen T, Tran V T, Hong N, Nguyen V T, Thanh N, Thi H, Duy N, Thanh D and Nguyen T 2023 Influences of material selection, infill ratio, and layer height in the 3d printing cavity process on the surface roughness of printed patterns and casted products in investment casting *Micromachines* **14** 395
[10] Hikmat M, Rostam S and Ahmed Y M 2021 Investigation of tensile property-based taguchi method of pla parts fabricated by fdm 3d printing technology *Results in Engineering* **11** 100264
[11] Yu W, Shi J, Sun L and Lei W 2022 Effects of printing parameters on properties of fdm 3d printed residue of astragalus/polylactic acid biomass composites *Molecules* **27** 7373
[12] Rickert C, Hayta E, Selle D, Kouroudis I, Harth M, Gagliardi A and Lieleg O 2021 Machine learning approach to analyze the surface properties of biological materials *ACS Biomaterials Science and Engineering* **7** 4614–25
[13] Singh B, Vijayvargiya A and Kumar R 2022 Kinematic modeling for biped robot gait trajectory using machine learning techniques *J. Bionic Eng.* **19** 355–69
[14] Zheng X, Wang W, Li L and Xie G 2020 Artificial lateral line based relative state estimation between an upstream oscillating fin and a downstream robotic fish *Bioinspiration and Biomimetics* **16** 016012

[15] Deng W, Huang Z, Zhang J and Xu J 2021 A data mining based system for transaction fraud detection *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)* pp 542–5
[16] Liu Y, Wang Y and Zhang J 2012 New machine learning algorithm: Random forest *International Conference on Information Computing and Applications* (https://doi.org/10.1007/978-3-642-34062-8)
[17] Guo L, Li Z, Tian Q, Guo L and Wang Q 2023 Prediction of csg splitting tensile strength based on xgboost-rf model *Materials Today Communications* **34** 105350
[18] de Prado Gil J, Palencia C, Monteiro N and Martínez García R 2022 To predict the compressive strength of self compacting concrete with recycled aggregates utilizing ensemble machine learning models *Case Studies in Construction Materials* **16** e01046
[19] Johansson U, Boström H, Löfström T and Linusson H 2014 Regression conformal prediction with random forests *Mach. Learn.* **97** 1–22
[20] Chen T and Guestrin C 2016 Xgboost: a scalable tree boosting system *Proceedings of the XXII ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* pp 785–94
[21] de Prado Gil J, Palencia C, Jagadesh P J and Martínez García R 2022 A comparison of machine learning tools that model the splitting tensile strength of self-compacting recycled aggregate concrete *Materials* **15** 4164
[22] Zhang Z, Shi J, Yu T, Santomauro A, Gordon A, Gou J and Wu D 2020 Predicting flexural strength of additively manufactured continuous carbon fiber-reinforced polymer composites using machine learning *J. Comput. Inf. Sci. Eng.* **20** 061015
[23] Minfei L, Chang Z, Wan Z, Gan Y, Schlangen E and Šavija B 2021 Interpretable ensemble-machine-learning models for predicting creep behavior of concrete *Cem. Concr. Compos.* **125** 104295
[24] Song K, Yan F, Ding T, Gao L and Lu S 2020 A steel property optimization model based on the xgboost algorithm and improved pso *Comput. Mater. Sci.* **174** 109472
[25] Chandrahas N S, Choudhary B, Teja M, S Venkataramayya M and Prasad N 2022 Xg boost algorithm to simultaneous prediction of rock fragmentation and induced ground vibration using unique blast data *Applied Sciences* **12** 5269
[26] Wang Q, Hussain A, Farooqi M and Deifalla A 2022 Artificial intelligence-based estimation of ultra-high-strength concrete's flexural property *Case Studies in Construction Materials* **17** e01243
[27] Pervaiz M, Shorfuzzaman M, Alsufyani A, Jalal A, Alsuhibany S and Park J 2023 Tracking and analysis of pedestrian's behavior in public places *Computers, Materials and Continua* **74** 841–53
[28] Ganguly S and Manna S 2023 Prediction of continuous cooling transformation diagrams in steels using light gradient boosting and rule-based optimization *Mater. Manuf. Processes* 1–16 pages
[29] Lee M, Kim M and Min K 2022 Evaluation of principal features for predicting bulk and shear modulus of inorganic solids with machine learning *Materials Today Communications* **33** 104208
[30] Alzubaidi L, Zhang J, Humaidi A J, Al-Dujaili A Q, Duan Y, Al-Shamma O, Santamaría J, Fadhel M A, Al-Amidie M and Farhan L 2021 Review of deep learning: concepts, cnn architectures, challenges, applications, future directions *Journal of Big Data* **8**
[31] Han T, Stone-Weiss N, Huang J, Goel A and Kumar A 2020 Machine learning as a tool to design glasses with controlled dissolution for healthcare applications *Acta Biomater.* **107** 286–98
[32] Khokhar S, Ahmed T, Khushnood R, Ali S M and Shahnawaz S 2021 A predictive mimicker of fracture behavior in fiber reinforced concrete using machine learning *Materials* **14** 7669
[33] Amin M, Ahmad W, Khan K, Ahmad A, Nazar S and Alabdullah A 2022 Use of artificial intelligence for predicting parameters of sustainable concrete and raw ingredient effects and interactions *Materials* **15** 5207
[34] Cai R, Wang K, Wen W, Peng Y, Baniassadi M and Ahzi S 2022 Application of machine learning methods on dynamic strength analysis for additive manufactured polypropylene-based composites *Polym. Test.* **110** 107580
[35] Jin S and Goh G 2021 Bayesian selection of best subsets via hybrid search *Comput. Stat.* **36**
[36] Anjum M, Khan K, Ahmad W, Ahmad A, Amin M and Nafees A 2022 New shapley additive explanations (shap) approach to evaluate the raw materials interactions of steel-fiber-reinforced concrete *Materials* **15** 6261
[37] Yuanxie S, Sun J and Liang S 2022 Interpretable machine learning models for punching shear strength estimation of frp reinforced concrete slabs *Crystals* **12** 259
[38] Lundberg S M and Lee S I 2017 A unified approach to interpreting model predictions *Advances in Neural Information Processing Systems* **30**
[39] Miazio Ł 2019 Impact of print speed on strength of samples printed in fdm technology *Agric. Eng.* **23** 33–8
[40] Druga C, Serban I, Braun B and Alexandru T 2020 Analysis of the influence of the layer height on the strength of 3D printed structures *Sciendo* **12** 177–81
[41] Aw Y Y, Yeoh C K, Idris M A, Teh P L, Hamzah K A and Sazali S A 2018 Effect of printing parameters on tensile, dynamic mechanical, and thermoelectric properties of fdm 3d printed cabs/zno composites *Materials* **11**
[42] Fan C, Shan Z, Zou G, Zhan L and Yan D 2021 Interfacial bonding mechanism and mechanical performance of continuous fiber reinforced composites in additive manufacturing *Chin. J. Mech. Eng.* **34** 21
[43] Goh G, Lee J M, Guo Liang G, Huang X, Lee S and Yeong W Y 2022 Machine learning for bioelectronics on wearable and implantable devices: Challenges and potential *Tissue Eng.* A **29** 20–46
[44] Nugroho A W and Budiantoro C 2019 Improving the tensile properties of 3d printed pla by optimizing the processing parameter *JEMMME (Journal of Energy, Mechanical, Material, and Manufacturing Engineering)* **4** 29–36