Using AI To Predict 3D Printing Attributes

Panth Korat
12th grade
Delhi Public School Surat, Dumas Surat, India

DOI:10.37648/ijrst.v14i04.006

[1]*Received: 28 September 2024; Accepted: 06 November 2024; Published: 14 November 2024*

ABSTRACT

Today, machine learning is being used in various fields, from quality control to medical diagnosis. Working in the same line, our paper focuses on a particular application of machine learning. In this paper, we have implemented machine learning algorithms to develop prediction for roughness and tension strength of 3D Printed object based on the 3D Printing Dataset. This paper first introduces the learning algorithms used. Then, it discusses multiple regression models and their parameter values. Based on the prediction results, the paper first compares the performance of ML (machine learning) models through quantitative metrics, evaluates the importance of attributes influencing the roughness and tension strength labels through SHAP values, and finally draws a conclusion on the labels using practical reasoning

Keywords: *3D Printer; Roughness and Tension Strength Prediction; Machine Learning; SHAP Values*

INTRODUCTION

3D printing, also known as additive manufacturing (AM), has emerged as a revolutionary technology that allows for the creation of complex three-dimensional objects by layering material based on a digital design. The versatility of 3D printing enables it to be used in a variety of industries, from aerospace and automotive to healthcare and consumer products. One of the key challenges in 3D printing is the ability to predict and optimize the mechanical properties of the printed parts, such as surface roughness (μm) and tensile strength(MPa), which are crucial for ensuring the performance and reliability of these components in real-world applications. Thus we were driven to develop a predictive model that could help predict the roughness and tension strength of the printed parts and the attributes to account for while evaluating it.

The machine learning approach employed in this study utilizes algorithms that "learn" from data to identify patterns and generate meaningful predictions. By analysing the data from 3D printers, along with key attributes such as roughness and tensile strength, machine learning models can be used to understand and predict important performance factors. These models can also identify the most significant features of 3D printers and offer a sophisticated method for valuing these features from the user's perspective. In our research, we have specifically applied regression and ensemble techniques to model and analyse the performance characteristics of 3D printers. By doing so, we evaluate the contribution of each feature to the overall performance, offering insights into which factors are most influential in determining the printer's output quality.

Our research project relies on machine learning models. These models use machine learning techniques to predict roughness and tension strength label of the 3D printed objects.

One of the key techniques utilized in this study is gradient boosting [1]. Unlike other machine learning methods, gradient boosting aggregates multiple weak learners to form a single, stronger model. In each iteration, the new model aims to correct the errors made by the previous one, with each model being weighted based on its accuracy in learning. As a result, subsequent learners focus specifically on the data points misclassified by the prior models. The experiments conducted in this research make use of libraries such as XGBoost (XGB) [2], a popular

[1] How to cite the article: Korat P.; November 2024; Using AI To Predict 3D Printing Attributes; *International Journal of Research in Science and Technology*, Vol 14, Issue 3, 54-63, DOI: http://doi.org/10.37648/ijrst.v14i04.006

implementation of gradient boosting. The choice to use XGBoost was crucial for this study, particularly because it addresses the issue of multicollinearity, where predictor variables exhibit strong correlations with each other.

RELATED WORK

We have selected the 3D Printing Dataset dataset to perform a study on the prediction of roughness and tension strength of the printed object from the 3D printer. The Dataset was taken from [3].

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

The author in [4] employed machine learning in the prediction of tensile strength(tension strength) of 3D Poly Scaffolds. In this work, machine learning algorithms were implemented using Python . It uses gradient boosting models along with multiple linear regression to train the data and further predicting the target variable. Similar to our work, this work also uses correlation heatmap and quantitative metrics for the analysis. But , this work uses a limited number of printing parameters. However, our model uses many more printing parameters then this work to reveal the importance of each printing parameter.

B. *Comparative analysis and experimental validation of statistical and machine learning-based regressors for modeling the surface roughness and mechanical properties of 316L stainless steel specimens produced by selective laser melting*

The work of [5] analyses the surface roughness and mechanical properties of 316L samples produced by Selective Laser Melting (SLM) through the application of statistical regression and Machine Learning techniques. While we used models like XGB, linear regression and RF, it uses many other models like Response Surface Methodology, Multi- layer Perceptron and Support Vector Regression. In this work, an exploratory data analysis was conducted to assess the behaviour of the studied responses. Then six prediction methods were fitted and compared to construct an ensemble model, and then each obtained model was validated.

C. *Accurate Estimation of Tensile Strength of 3D Printed Parts Using Machine Learning Algorithms*

The work of [6] uses machine learning algorithms for the accurate estimation of tensile strength of 3D printed parts. The work involves machine learning models like XGB ,RF and linear regression similar to what we have used. But they have also used AdaBoost regression and gradient boosting regression, different from our approach. It also involves training and testing of data along with a correlation map for analysis. The study involves two case studies, each involving different attributes affecting tensile strength. However, this study did not use hyperparameter tuning which was used in our work.

After studying and evaluating these work , we were motivated to focus our research on advanced machine learning techniques for analysis and incorporated SHAP [7] to introduce the concept of 'explainability’. This would specifically help strengthen our evaluation by validating the results of the models.

We also concentrated on fine-tuning a different set of hyperparameters for XGBoost [8]. For instance, the 'min_child_weight' parameter specifies the minimum weight required in a leaf node before further partitioning can occur. The 'learning_rate' hyperparameter controls how quickly the model learns from the data, adjusting the step size at each iteration and influencing the model's rate of adaptation. The 'gamma' parameter was tuned to prevent overfitting by controlling the model's complexity, while the 'colsample_bytree' parameter determines the proportion of columns used for sampling at each tree.

For the Random Forest model [9], we focused on the 'n_estimators' parameter, which dictates the number of trees in the forest—more trees generally improve accuracy. In decision trees, internal nodes are those that perform further splits, while leaf nodes do not. The 'min_samples_split' parameter specifies the minimum number of samples needed in an internal node to allow a split, while the 'min_samples_leaf' parameter sets the minimum number of samples required at the leaf node (the final split). The 'max_features' parameter defines how many features are considered for each split or tree, with 'sqrt' typically selecting the square root of the total features per tree. The 'max_depth' parameter limits the maximum depth a decision tree can grow, with deeper trees increasing the model's complexity. Finally, the 'bootstrap' parameter controls whether bootstrap samples are used during tree creation.

While experimenting with linear regression, we used the OLS regression to determine the R-squared values. OLS regression is a method used to estimate the parameters (coefficients) of a linear regression model. Its aim is to find the line that best fits the observed data by minimizing the sum of the squared differences between the actual values and the values predicted by the model.

IMPLEMENTATION

We have taken 11 attributes for the study, which contain 9 features and 2 labels. We will be using the 9 features to predict on the 2 labels which are roughness and tension strength.

A. *The Dataset*

The Dataset include different attributes that influence the labels. The 3D printing dataset includes attributes such as:

TABLE I. ATTRIBUTES OF [3]

ATTRIBUTES	**MEANING**
layer_height	thickness of each individual layer of material
Wall_thickness	the thickness of the outer shell or walls of the printed object
Infill_density	the amount of internal structure within a printed object
Infill_pattern	the internal structure or the filling pattern within a printed object
nozzle_temperature	the temperature of the hotend nozzle, which is the part of the printer that heats and extrudes the filament through a small opening.
Bed_temperature	the temperature of the heated print bed, which is the surface on which the 3D print is built layer by layer.
Print_speed	the rate at which the printer's nozzle moves and deposits material during the printing process.
material	the substance or filament that is used by the 3D printer to create objects.
Fan_speed	the rate at which the cooling fans operate during the printing process.
roughness	the texture or surface finish of a printed object.
Tension strength	the maximum amount of tensile (pulling or stretching) force a material can withstand before breaking or failing.

A correlation matrix shows the relationship between different variables in a dataset. The values in the matrix range from -1 to 1. Values closer to 0 mean that the two variables are weakly correlated. Values closer to 1 mean that the two variables are strongly positively correlated, and values closer to -1 mean that the two variables are strongly negatively correlated. The diagonal of the matrix is always 1 because every variable is perfectly related to itself.

The correlation matrix that we created for the 3D printing dataset is shown in Fig.1. This visual would give us a better understanding of the relation of all features with the labels, making it easier to hypothesize the most important attributes.

(IJRST) 2024, Vol. No. 14, Issue No. 4, Oct-Dec **e-ISSN: 2249-0604, p-ISSN: 2454-180X**

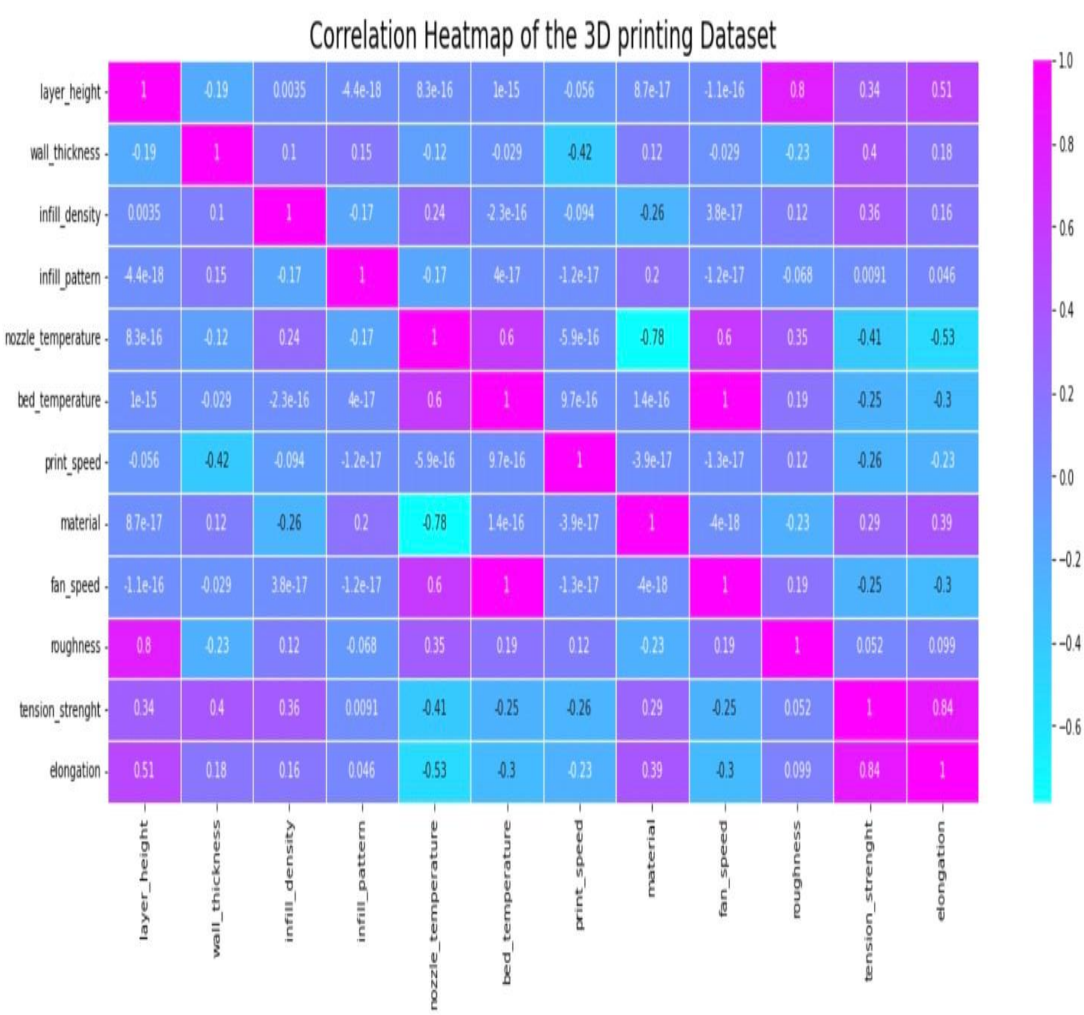

Fig.1 Correlation Matrix of [3]

On the basis of Fig.1 , we see that layer height is the higher scoring attribute when correlated with roughness, implying the strongest relationship with the target label. Similarly, tension strength and elongation shows a strong positive correlation with each other.

B. *Data Refining and Standardizing*

For the analysis of data to be good, the data itself needs to be good. Thus, the data must be refined before we apply the machine learning algorithms. Categorical features were one such problem in the dataset used. So we had to convert the categorical features into numerical ones. After cleaning the data, it was important to standardize it to ensure uniformity across all values. We used StandardScaler as the standardization technique to standardize all input values. Under this technique, the mean of every column was subtracted from the value of every data point in that column and then divided by the column's standard deviation, thus 'standardizing' each datapoint onto one common range. The data was then ready to be put through the algorithms for multiple iterations and experiments.

C. *Model Training & Parameter Tuning*

It is crucial to train our data progressively to enhance the model's ability to predict the two labels based on specific attributes. The dataset was divided into an 80% training set and a 20% test set. We initially applied common regression techniques such as Simple Linear Regression[10]. However, to improve predictive accuracy, we also implemented ensemble learning methods like Random Forest (RF) and XGBoost (XGB).

For hyperparameter optimization, we utilized Scikit-Learn [11,12], allowing us to fine-tune each model's parameters to maximize accuracy. While some models may underperform, it is vital to identify at least one model that accurately captures the relationship between the attributes and the labels.

EXPERIMENTS AND RESULTS

For each new experiment with a regression model, the model's 'predict' function was applied to forecast the results for the test set. Additionally, the parameters of both the Random Forest and XGBoost models were fine-tuned.

Tables II, III, IV and V show the parameter values of the learning model after hyperparameter tuning for the dataset. For the roughness label of 3D printing dataset, the parameter configuration of XGBoost (XGB) obtained was:

TABLE II. XGBOOST PARAMETERS FOR ROUGHNESS

Parameters:	**Value:**
min_child_weight	3
max_depth	9
learning_rate	0.2
gamma	0
colsample_bytree	1

Similarly, for the tension strength label of the 3D printing dataset, the parameter configuration of XGB was:

TABLE III. XGBOOST PARAMETERS FOR TENSION STRENGTH

Parameters:	**Value:**
min_child_weight	3
max_depth	6
learning_rate	0.5
gamma	0
colsample_bytree	1

Then, the parameter configuration of RF for roughness label of the 3D printing dataset was:

TABLE IV. RANDOM FOREST PARAMETERS FOR ROUGHNESS

Parameters:	**Value:**
n_estimators	60
min_samples_split	8
min_samples_leaf	2
max_features	Auto
max_depth	10
Bootstrap	True

Similarly, for the tension strength label of the 3D printing dataset, the parameter configuration of RF was:

TABLE V. RANDOM FOREST PARAMETERS FOR TENSION STRENGTH

Parameters:	**Value:**
n_estimators	10
min_samples_split	2
min_samples_leaf	1
max_features	Auto
max_depth	6
Bootstrap	True

A. *Quantitative Results*

We collected and compared the *R*-squared regression scores, mean squared error (MSE) [13] values and mean absolute percentage error values (MAPE) of the various models for both our labels.

The *R*-squared score, mean squared error and mean absolute percentage error values for the roughness label of every model trained with the 3D printing dataset are:

TABLE VI. QUANTITATIVE METRICS OF MODELS FOR ROUGHNESS

Model	R 2 ↑	MSE ↓	MAPE ↓
Linear Regression	0.875	0.0716	38.8858
XGB with parameters	**0.9335**	**0.04677**	**22.87**
RF with parameters	0.83728	0.11438	47.0433

Table VI. shows that XGB generated the most accurate prediction of roughness based on the given dataset. In contrast the lower scoring model is RF. We also observed that the model with highest scores had the lowest MSE values. This confirmed the regularity of the quantitative evaluation and assures us that all performances have been assessed appropriately. While it is true that these models gave good results with the standardized data, we performed the same experiments with a normalized (dividing difference of every value and minimum value of that column by the range of the values in the column) data to compare the performance. We used MinMaxScaler to normalize the data. There was not a large variance from the scores in Table VI. XGB was still the best performing model for the dataset and RF was still the model having the least accuracy, thus validating the results in Table VI.

The R-squared score, mean squared error and mean absolute percentage error values for the tension strength label of every model trained with the 3D printing dataset are:

TABLE VII. QUANTITATIVE METRICS OF MODELS FOR TENSION STRENGTH

Model	R^2 ↑	MSE ↓	MAPE ↓
Linear Regression	0.673	0.9532	51.6186
XGB with parameters	**0.9147**	**0.167**	**20.47**
RF with parameters	0.6643	0.6571	42.7253

Table VII. consists of the quantitative metrics of tension strength in the 3D printing dataset . It identifies XGB as the model with highest R-squared values, implying that XGB was able to most accurately predict the tension strength. Conversely, RF was the model with the least accuracy in predicting. Again, the model with the highest R-squared score reported the lowest MSE value and vice versa. This shows consistency in the models' performance.

Similar to the approach adopted with roughness label of 3D printing dataset, we repeated these experiments with the tension strength label. Just like before, the performance of model with normalized data remained consistent with that of standardized data.

B. *Qualitative Results*

Finally, we used SHAP visualizations to understand how each model made its predictions. These visualizations show the importance of each feature in the prediction. Features with larger absolute SHAP values are more important. While each model ranked the features differently, we only focused on the best- performing model. This is because the best model is the one that fits the training data the best, so its feature importance results are the most reliable.

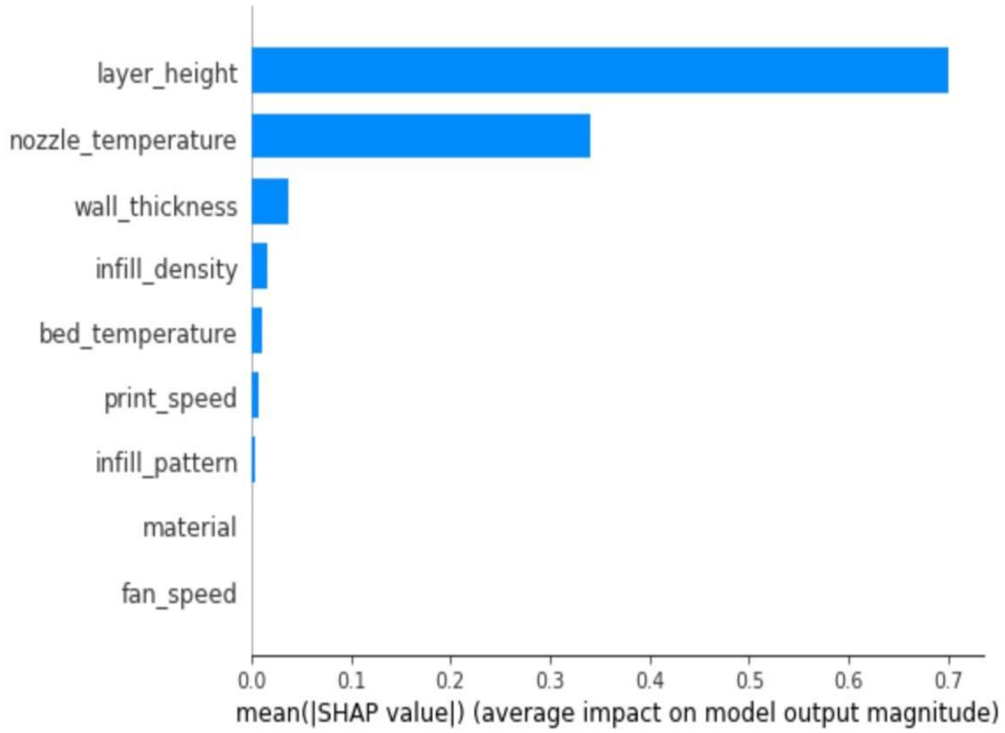

Fig.2. Visual Representation of SHAP Values for Roughness

From Fig.2 , we see that Layer Height and Nozzle Temperature are the most important features , identified by the explainer function of SHAP library, that influences the roughness on the basis of the given data. On the contrary, attributes such as print speed and infill pattern have very less influence on the roughness label. Also, attributes like fan speed has no effect on roughness. Then, the same method was applied for the tension strength label. The results of the SHAP explainer function were:

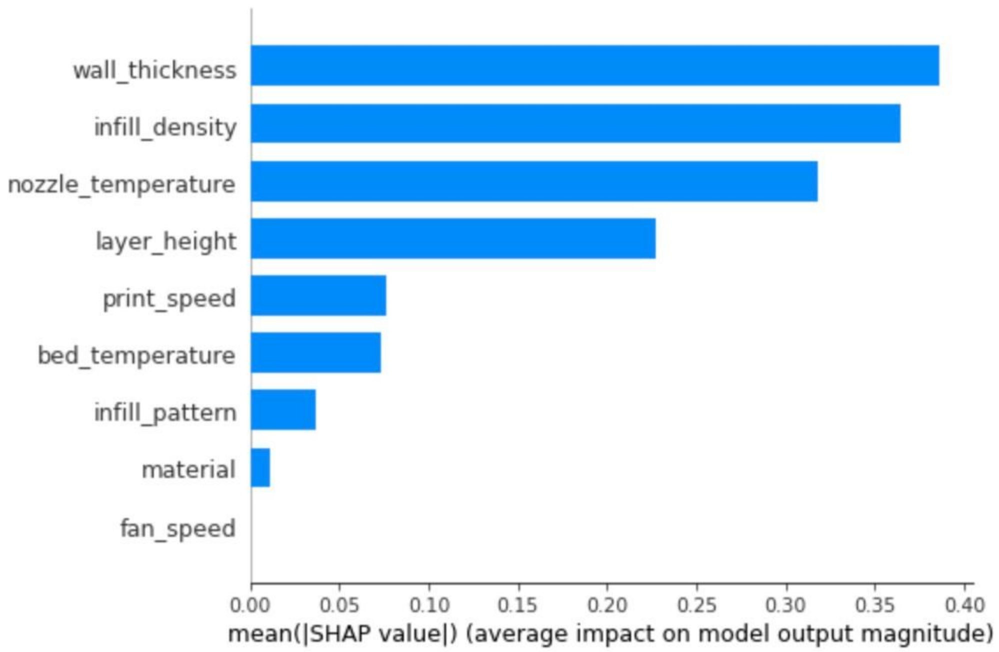

Fig.3. Visual Representation of SHAP Values for Tension Strength

Through Fig.3, Wall Thickness, Infill Density and Nozzle Temperature were the three most important factors. While infill pattern and material have much less influence , fan speed has no influence on tension strength.

When we applied the same method and evaluated the predictions after removing the columns for Wall Thickness, Infill Density, and Nozzle Temperature, the scores dropped significantly. The scores for XGBoost (XGB) and Random Forest (RF) were 0.0443 and 0.1521, respectively. These results showed a substantial decrease compared to the accuracies obtained using the original dataset. This drop is significant because it highlights the importance of Wall Thickness, Infill Density, and Nozzle Temperature as key factors that strongly influence the target label. The low scores are understandable; if the scores for the dataset without these attributes had remained high, it would have raised concerns about the validity of the explanation provided by SHAP. However, that was not the case.

CONCLUSION

Considering the experiments we have conducted with the two labels, we see that XGB, with tuning, has consistently performed best with both the labels of the dataset. However, the main goal of this research was to identify the features that most strongly affect the two labels of the 3D printing dataset. Layer Height and Nozzle temperature were the key attributes for the roughness label, and Wall Thickness, Infill Density and Nozzle Temperature were the strongest factors for tension strength label of the 3D Printer dataset.

If we see the result in practical terms, then we can safely establish a positive correlation between layer height and roughness. For a desired smooth surface by the user, the layer height should be small resulting in thinner layers being deposited on the object. This leads to more detailed surface where transition between layers are less noticeable. Meanwhile, for a desired rougher surface by the user, layer height should be large resulting in visible transition and hence a rougher surface. On the other hand, the nozzle temperature controls the viscosity and flow characteristics of the filament[14]. If the user keeps the nozzle temperature too low, the filament may not flow smoothly resulting in rough surfaces. On the contrary , if the user keeps higher nozzle temperatures, the filament becomes more fluid, improving the bonding between layers. This results in better layer adhesion and reduced surface roughness.

For the tension strength label, wall thickness can be looked at as thickness of outer shell of printed object, hence showing that tension strength varies as thickness of printed object varies. Logically, thicker wall helps in evenly distributing the applied load across the part and increases the surface area available to distribute tensile forces, resulting in absorbing more tensile force.

Another factor users should pay attention is infill density. This is because higher infill density means more material is used inside the printed object and fewer voids making it highly resistant to tension. Moreover, nozzle temperature mainly affects tension strength due to layer adhesion[15]. Users may neglect this factor, but it could also affect the cooling rate[16]. Hence , a higher nozzle temperature can lead to a stronger object.

FUTURE WORK

In this research we focused on only three machine learning models for the analysis, but other machine learning models like SVM could also be used for a wider scope of analysis. We also do not have an external validation regime which can make the analysis more generalisable. At last, many other datasets could be used for the same study. Since this project predicts only roughness and tension strength of the 3D printed part, many other features could be predicted using the same experiment. Then, we could compare most important attributes of those features and make relevant conclusions.

ACKNOWLEDGEMENT

This research project was undertaken independently in pursuit of higher education. It holds no connection with the school and nor is it the part of a school curriculum.

REFERENCES

[1] Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in neurorobotics, 7, 21.

[2] Chen, T. "Xgboost: extreme gradient boosting." *R package version 0.4- 2* 1.4 (2015).

[3]3D Printing Dataset Available: https://www.kaggle.com/datasets/afumetto/3dprinter

[4] Ege, D., Sertturk, S., Acarkan, B. and Ademoglu, A., 2023. Machine Learning models to predict the relationship between printing parameters and tensile strength of 3D Poly (lactic acid) scaffolds for tissue engineering applications. *Biomedical Physics & Engineering Express*, *9*(6), p.065014.

[5] La Fé-Perdomo, Iván, Jorge A. Ramos-Grez, Ignacio Jeria, Carolina Guerra, and Germán Omar Barrionuevo. "Comparative analysis and experimental validation of statistical and machine learning-based regressors for modeling the surface roughness and mechanical properties of 316L stainless steel specimens produced by selective laser melting." *Journal of Manufacturing Processes* 80 (2022): 666-682.

[6] Jayasudha, M., Elangovan, M., Mahdal, M. and Priyadarshini, J., 2022. Accurate estimation of tensile strength of 3D printed parts using machine learning algorithms. *Processes*, *10*(6), p.1158.

[7] Lundberg, S. M., & Lee, S. I. (2017, December). A unified approach to interpreting model predictions. In Proceedings of the 31st international conference on neural information processing systems (pp. 4768-4777).

[8] Chen, T., He, T., Benesty, M., & Khotilovich, V., (2019). Package 'xgboost'. R version, 90.

[9] Rigatti, Steven J. "Random Forest" *Journal of Insurance Medicine* 47, no 1 (2017): 31-39

[10] Maulud, D., & Abdulazeez, A. M. (2020) A review on Linear Regression Comprehensive in Machine Learning. Journal of Applied Science and Technology Trends, 1(4), 140-147

[11] Pedregosa, F., Varoquaux, Ga"el, Gramfort, A., Michel, V., Thirion, B., Grisel, O., … others. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825–2830.

[12] A. Ghatnekar and A. D. Shanbhag, "Explainable, Multi-Region Price Prediction," *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Cape Town, South Africa, 2021, pp. 1-7, doi: 10.1109/ICECET52533.2021.9698641

[13] Botchkarev, Alexei. "Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology." arXiv preprint arXiv:1809.03006(2018)

[14] Tanikella, Nagendra G., Ben Wittbrodt, and Joshua M. Pearce. "Tensile strength of commercial polymer materials for fused filament fabrication 3D printing." *Additive Manufacturing* 15 (2017): 40-47.

[15] Thumsorn, Supaphorn, et al. "Rheological behavior and dynamic mechanical properties for interpretation of layer adhesion in FDM 3D printing." *Polymers* 14.13 (2022): 2721.

[16] Akhoundi, Behnam, et al. "An experimental study of nozzle temperature and heat treatment (annealing) effects on mechanical properties of high-temperature polylactic acid in fused deposition modeling." *Polymer Engineering & Science* 60.5 (2020): 979-987