

PAPER

An ensemble stacking machine learning model for temperature prediction of extruder nozzle in an additive manufacturing process

To cite this article: Vijay Subramanian *et al* 2026 *Eng. Res. Express* **8** 075246

View the [article online](#) for updates and enhancements.

You may also like

- [Odd-harmonic repetitive controller with extreme learning machine approach for precise tracking and robust disturbance rejection in servo systems](#)
Edi Kurniawan, Enggar B Pratiwi, Samsul Hafiz et al.
- [A multi-feature fusion deep neural network with attention for multi-task crop monitoring](#)
Efrem Yohannes Obsie, Yun Zhang and Yongguo Liu
- [A quantitative damage diagnosis method for CFRTP/Ti6Al4V hybrid welded/bolted joints based on an attention-enhanced capsule network](#)
Zhi Li, Xingzhou Chen, Jun Zhang et al.

Engineering Research Express



PAPER

An ensemble stacking machine learning model for temperature prediction of extruder nozzle in an additive manufacturing process

RECEIVED
7 December 2025

REVISED
15 March 2026

ACCEPTED FOR PUBLICATION
30 March 2026

PUBLISHED
14 April 2026

Vijay Subramanian¹ , Aditya Nair H¹, V Madhusudanan Pillai^{1,*} , Sachin Salunkhe² and Chandrasekhar U³

¹ Department of Mechanical Engineering, National Institute of Technology Calicut, Kerala 673601, India

² Department of Mechanical Engineering, Gazi University Faculty of Engineering, Maltepe, Ankara, Turkey

³ NAFEMS India & Adjunct Faculty CPDM Indian Institute of Science, Bangalore, India

* Author to whom any correspondence should be addressed.

E-mail: vijaysubu14@gmail.com, adityanair1709@gmail.com, vmp@nitc.ac.in, Kashid32@gmail.com and rapidchandra@gmail.com

Keywords: machine learning, predictive maintenance, additive manufacturing, ensemble stacking, temperature prediction

Abstract

Additive Manufacturing (AM) technique is preferred over subtractive manufacturing because it creates complex geometric structures with minimal material waste. Even though it has many advantages, a few drawbacks limit its application in the industrial environment. One of the main disadvantages of the process is the extruder nozzle temperature fluctuation which causes nozzle clogging, overheating and failure. Sensorized AM machine control is the most adapted process monitoring strategy to avoid these issues. However, manually managing the process variables with sensorized control becomes complex due to a substantial amount of data and variables in AM. Machine Learning (ML) is a suitable approach for analyzing massive amounts of sensor data with many attributes. This paper considers a fused filament fabrication process and its extruder nozzle temperature prediction. We developed ensemble stacking regressor models using sensor data to predict the extruder nozzle temperature. As a precursor of the study, we applied ML models such as Naive Bayesian, Random Forest, XGBoost, CatBoost, Artificial Neural Network, and Support Vector Machine as stand-alone models. The stand-alone models with better performance measure values based on Root Mean Squared Error (RMSE) and R squared (R^2) are considered for developing the ensemble stacking models. The best-performing ensemble stacking model employs XGBoost, CatBoost and Random Forest in layer zero and XGBoost in the meta-model layer. It has the highest R^2 value (0.94), indicating the best fit than the stand-alone models.

1. Introduction

Additive Manufacturing (AM) is a 3D printing technique that builds objects layer by layer using computer-controlled material deposition. It has revolutionized various industries due to its flexibility and cost-effectiveness, with applications ranging from aerospace to biomedical engineering (Attaran 2017, Jerez-Mesa *et al* 2018). Popular AM methods include Fused Filament Fabrication (FFF), Stereolithography (SLA), and Selective Laser Sintering (SLS). FFF, in particular, is favored for its affordability and ease of use (Tofail *et al* 2018). In FFF, a continuous polymer strand is melted and extruded through a heated nozzle to form the desired shape. However, maintaining consistent nozzle temperature is crucial for quality and efficiency. Variations in nozzle temperature can lead to defects like warping, clogging, and poor adhesion (He *et al* 2018, Tlegenov *et al* 2018, Gonabadi *et al* 2020, Hsueh *et al* 2021). These issues increase the failure rate of prototypes, affecting part quality.

While several factors influence AM quality, nozzle temperature is arguably the most critical. It affects material flow, layer adhesion, and cooling rate, indirectly impacting factors like layer thickness, printing speed, and material properties (Tanabi 2021, Aslani *et al* 2020). Lewis and Moore (2021) developed an *in situ* infrared temperature sensing method for real-time defect detection in AM, effectively identifying defects through

statistical temperature deviation analysis. Heymann and Schmitt (2023) introduced a machine learning pipeline for predictive maintenance in polymer 3D printing, focusing on predicting nozzle remaining useful life and suggesting further refinement through automated machine learning and broader application. However, maintaining optimal nozzle temperature manually during printing is challenging.

To address this, accurate prediction of extruder nozzle temperature is essential. Ensemble methods, which combine multiple machine learning models, have shown promise in predicting complex parameters like nozzle temperature and bed temperature (Luo *et al* 2021). The MeltpoolINR framework forecasts continuous temperature fields, melt-pool geometry, and cooling-rate variations through an implicit neural representation utilizing Fourier encoding (Manav *et al* 2025). This study centers on Laser Powder Bed Fusion (LPBF) yet its methodology underscores a significant transition in additive manufacturing research towards continuous-field temperature modeling as opposed to discrete point-based measurements. These advancements indicate that temperature forecasting has evolved into a multidisciplinary challenge across additive manufacturing processes, necessitating the development of more resilient thermal prediction models, even in polymer FFF extrusion. Boosting, Bagging, and Stacking are three types of ensemble methods. In this study, we introduce a hybrid ensemble stacking method to predict extruder nozzle temperature, utilizing identified variables to train the machine learning models and improving prediction accuracy.

While deep learning approaches such as Long Short-Term Memory (LSTM) networks and Transformer-based models have demonstrated strong sequence modeling capabilities in time-series prediction tasks, they introduce significant computational overhead and require substantially larger datasets for effective training. Agron *et al* (2021) employed LSTM-based neural networks for nozzle thermal estimation, reporting competitive RMSE values but at the cost of higher training complexity and epoch-dependent convergence. In contrast, the proposed ensemble stacking approach leverages gradient-boosted tree models (XGBoost, CatBoost) and Random Forest, which are inherently better suited to tabular sensor data with nonlinear variable interactions, and achieve superior generalization without the vanishing gradient issues associated with recurrent architectures. The present work is anticipated to achieve better R^2 with significantly lower computational demand, making it more practical for industrial deployment on standard hardware.

In this paper, a comprehensive analysis is presented with various sections dedicated to different aspects of the study. Section 2 begins with a thorough literature review, highlighting the relevant prior research. The methodology adopted for the study is outlined in section 3. Details regarding data management, including collection and preprocessing techniques, are presented in section 4. In section 5, the identification of stand-alone models utilized for prediction is elaborated upon, along with hyperparameter tuning and training. The performance evaluation and comparison of these stand-alone models are discussed in section 6. Moving forward, section 7 covers filtering the best-performing models identified in the previous section and section 8 combines best performing stand alone models to form an ensemble stacking model. Section 9 offers an in-depth discussion of the results obtained from the study. Finally, section 10 concludes the paper by summarizing the key findings and presenting avenues for future research.

2. Literature review

Extruder nozzle temperature is one of the parameters influencing the quality of the printed parts. Therefore, it is crucial to comprehend the effects of extruder nozzle temperature and its characteristics. This section ultimately covers the vital literature on: (1) the impact of the extruder nozzle temperature on the AM machine and the printed parts, (2) stand-alone ML models available in the literature for the extruder nozzle temperature prediction and their performance measures, and (3) ensemble ML models available in the literature for the extruder nozzle temperature prediction and their performance measures.

Extruder nozzle temperature is an important feature that causes various imperfections in AM, such as blobs, warping, dimensional inaccuracy, etc Agron *et al* (2021), Liu *et al* (2017), and He *et al* (2018) conducted studies related to extruder nozzle temperature. Agron *et al* (2021) predicted the real-time voltage difference detected by a thermocouple connected to the extruder nozzle. Liu *et al* (2017) studied a fuzzy logic integrated with a Proportional, Integral and Derivative (PID) controller used to maintain temperature control across the extruder nozzle made of PEEK (Poly Ethylene Ethyl Ketone). The authors found a relationship between the voltage applied across the heater and the extruder nozzle temperature by varying the voltage across the heater using a microcontroller. He *et al* (2018) and Miao *et al* (2018) explained the influence of extruder nozzle temperature on the formation of various defects in AM parts. He and Guo (2019) developed an online defect monitoring system for FFF, considering the thermal variation in AM parts. The authors concluded that due to nozzle temperature non-stabilization, printed parts resulted in insufficient precision, warping, and nozzle failure. Miao *et al* (2018) predicted model distortion due to the input variables such as extruder nozzle temperature and feed rate. Their results concluded that the extruder nozzle temperature positively affects the distortion of

Table 1. Literature related to ML algorithms and their performance measures in the AM process.

Author (Year)	Application	Algorithms employed	Performance measures
Kammerer <i>et al</i> (2019)	Prediction of the temperature anomalies using sensors	Decision Tree Regressor, MLP, and RF	MSE
Rojek <i>et al</i> (2021)	Optimization of process parameters	Genetic Algorithm with ANN	MSE
Zhang <i>et al</i> (2020)	Prediction of printing parameters of polymer gears	Genetic Algorithm with ANN	Pearson product moment correlation coefficient R^2
Agron <i>et al</i> (2021)	Development of a nozzle monitoring system	Time delay Convolutional Neural Network (TCNN), Gated Recurrent Neural Network (GRNN), and Long Short-Term Memory (LSTM) based Neural Network	MSE, RMSE, and R^2
Hsueh <i>et al</i> (2021)	Prediction of model distortion using nozzle temperature and bed temperature	SVM, ANN, and Linear Regression (LR)	MSE
Naser <i>et al</i> (2023)	Machine Learning framework to automate lifecycle assessment	XGboost, and MLP	RMSE
Wei and Wu (2023)	Real-time prediction framework for surface roughness	Deep Learning	MSE, and RMSE

the printed parts. Gonabadi *et al* (2020) and Roy *et al* (2019) formulated relations of extruder nozzle temperature with different features of AM, and the relations were established through experimentation. The findings of Gonabadi *et al* (2020) show that the layer thickness, infill density, nozzle temperature, and feed rate affect the quality of the printed parts. Roy *et al* (2019) addressed relationships among feed/flow rate ratio (ratio of feed rate to flow rate), layer height, and extruder temperature. According to their findings, residual stresses in AM models were directly related to the FFF process's temperature field. The temperature of the nozzle must be kept constant for the smooth functioning of the equipment and better quality of the printed parts.

Recent studies in thermal modeling, especially regarding metal additive manufacturing systems, point to the urgent need for better thermal prediction accuracy. For instance, Kim *et al* (2024) have introduced the idea of using physics-informed generative models to improve our understanding of melt pool dynamics and thermal gradient behavior. They also applied thermal-dose modeling to determine the depth of the melt pool, showing that precise control over the local thermal field plays a significant role in the final shape and stability of printed structures (Correa-Gómez *et al* 2024). Additionally, machine learning has become essential for thermal monitoring. In real-time applications, ML models can pinpoint 'hot spots' and areas where heat accumulates, helping to prevent defects and structural issues caused by heat (Guirguis *et al* 2024). Furthermore, data-driven forecasting of melt-pool advancement from previous builds has demonstrated that machine learning can predict future temperature trends and process irregularities (Xiao *et al* 2024). Together, these insights emphasize that temperature stability in both polymer Fused Filament Fabrication (FFF) and metal Additive Manufacturing (AM) is a key factor influencing part integrity, defect generation, and the long-term reliability of printers.

Stand-alone ML techniques reported in the literature for temperature prediction are generally linear and non-linear models, regression trees, Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), time-delay neural networks, and vector autoregressive models. These models were assessed based on their ability to handle continuous data, dimensional accuracy, and other criteria. ML methods, namely Naïve Bayesian (NB), Support Vector Regressor (SVR), Random Forest (RF), Gradient Boost (GB), XtremeGB (XGB), Artificial Neural Network (ANN) and CatBoost (CB), were suitable for establishing logical relationships in non-linear data (Breiman 2001, Sheridan *et al* 2016, Kammerer *et al* 2019, Malinin *et al* 2020, Goh *et al* 2021). The performance indicators used were R squared (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). These models are equally effective in regression as well as classification problems. A summary of stand-alone ML models and their performance measures in the AM process are presented in table 1.

Banihashemi *et al* (2017) focused on developing a hybrid ensemble model using ANN and decision trees. Bartz-Beielstein *et al* (2018) explained the Optimization via Multi-Model Simulation (OMMS) approach, combining results from different models. Li *et al* (2023) proposed a data-driven model leveraging multiple machine learning algorithms to predict dimensional deviations in fused deposition modeling (FDM) processes, with the residual attention neural network demonstrating superior quality assurance capabilities. Zhu *et al* (2024) proposed a transfer learning-based feature extractor combined with gradient-boosting decision trees for surface quality prediction in FDM, achieving high precision and accuracy, and addressing data scarcity and feature extraction challenges. The results demonstrated that ensemble models could improve the robustness of the

Table 2. Summary of ensemble models reviewed.

Author (Year)	Application	ML model used	Performance Measures
Banihashemi <i>et al</i> (2017)	Prediction of energy consumption	ANN + Decision tree	MSE
Zhang <i>et al</i> (2019)	Optimal operation of aerosol jet process	K means clustering + SVM	Modelling error rate
Waqas Khan <i>et al</i> (2020)	Energy prediction	CB+RF+XGB	MAE, R ²
Shafqat <i>et al</i> (2021)	Energy Forecast	CB+RF+Light Gradient Boosting +GBM+XGB	MSE, RMSE, MAE, Accuracy, and Precision
Nalajam and Varadarajan (2021)	Prediction of melt pool temperature	Convolutional Neural Network (CNN)+LSTM	MSE, MSE, and MAPE
Zhao <i>et al</i> (2023)	Thermal field prediction	Recurrent Neural Network + Deep Neural Network (RNN-DNN), and CNN	Accuracy, MAE

predictions by reducing the impact of a poor surrogate model. Moreover, it showcased that ensemble models could perform far better than traditional stand-alone ML models. Various articles reported in the literature related to ensemble models are consolidated and are provided in table 2.

The primary motivation for this research work stems from the challenges associated with managing temperature fluctuations in AM processes, which can result in significant defects such as nozzle clogging, overheating, and part failure. Existing studies have highlighted the importance of precise temperature control but have not fully explored the potential of advanced ML techniques in this context. By addressing these challenges through an innovative ensemble ML approach, this research aims to enhance the reliability and efficiency of AM processes, making it more suitable for industrial applications (Gkertzos *et al* 2024). The successful implementation of the proposed approach in a real-world system would demonstrate its potential to significantly reduce downtime, improve print consistency, and enhance the mechanical properties of printed parts, addressing critical challenges in the field of AM.

The literature reviewed above reveals three key gaps that motivate the present work. First, while standalone ML models such as SVR, ANN, and gradient-boosting methods have been applied individually to AM temperature prediction, their comparative performance under a unified experimental framework using real FFF sensor data remains underexplored. Second, ensemble stacking, which hierarchically combines diverse model types to reduce both bias and variance has not been systematically applied to extruder nozzle temperature prediction in polymer FFF. Third, existing studies rarely incorporate a structured model filtering step to ensure that only well-performing base models contribute to the ensemble. The methodology presented in section 3 directly addresses these gaps by proposing a four-stage pipeline that integrates rigorous data management, standalone model benchmarking, performance-based filtering, and ensemble stacking construction.

3. Methodology

This work is based on the four-stage methodology that Ashmore *et al* (2021) proposed for establishing an ML-based solution in an industrial setting. For building the ensemble stacking model, the methodology was appropriately modified, and it consists of the following stages:

1. Data management, which involves the collection and preparation of data for building an ML model;
2. Model learning, which involves the identification of ML models, hyperparameter tuning and training;
3. Model verification, which is a precautionary check to see whether the model achieves specific functional and performance needs; and
4. Ensemble stacking combines the verified stand-alone models to predict the extruder nozzle temperature.

The various steps in different stages of the methodology are illustrated in figure 1.

In the data management stage, the data from the AM machine is collected using an array of sensors. The sensor data are transferred to a client using an Open Platform Communication Unified Architecture (OPCUA) server. OPCUA is a communication standard for manufacturing and permits data exchange control and management across distant industrial automation systems entities. The data is collected and stored in a cloud database. The raw data collected is further pre-processed by adopting suitable data-cleaning techniques. Feature extraction is carried out to identify and eliminate redundancy. Data within the appropriate temperature range

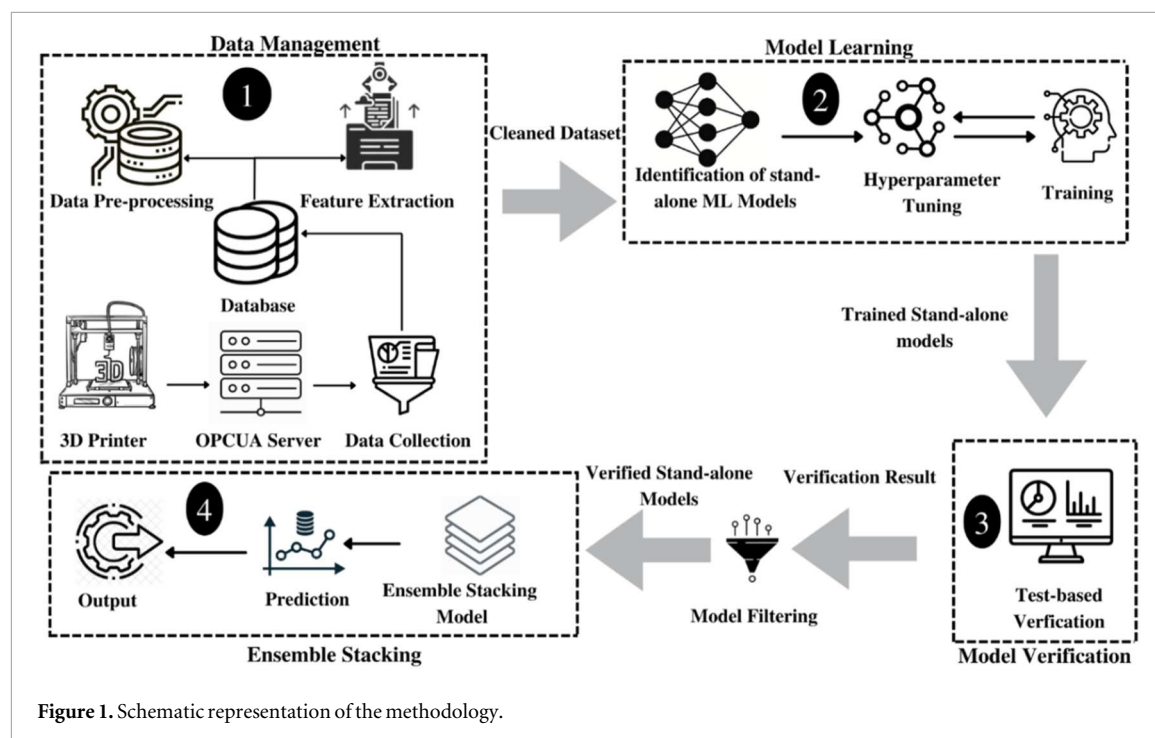


Figure 1. Schematic representation of the methodology.

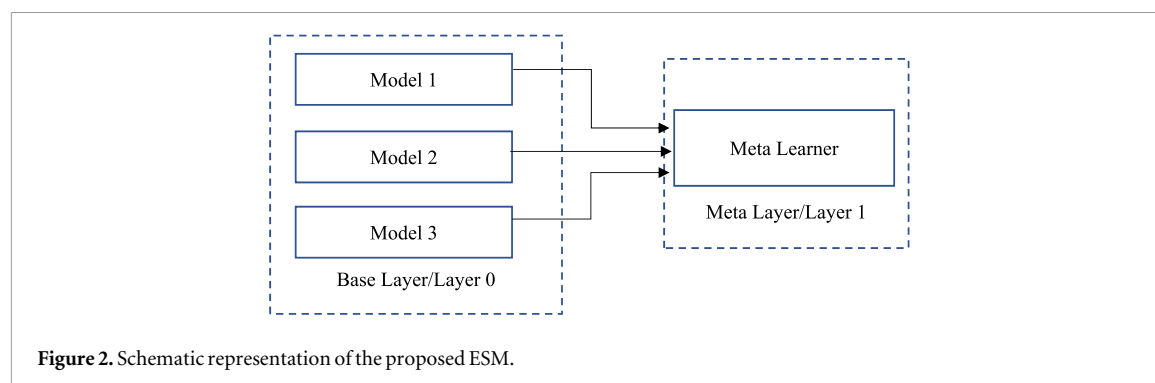


Figure 2. Schematic representation of the proposed ESM.

is filtered based on the AM machine operating conditions. The study utilizes the cleaned data for training ML algorithms to estimate the extruder temperature based on the input settings.

The model learning stage includes steps such as identification of ML algorithms, hyperparameter tuning, and training. The stand-alone ML model identification is based on the type of problem (regression or classification), data structure, and previous literature. Hyperparameter selection involves choosing the parameters of the training algorithm. Hyperparameters control the key ML model characteristics such as overfitting, underfitting and model complexity. Ogunsanya *et al* (2023) investigated the impact of grid search hyperparameter tuning on MLP models for FFF, identifying optimal configurations that enhance performance. The training activity aims to produce an ML model that minimizes loss function. When the resulting ML model achieves satisfactory performance levels, the third stage of the ML workflow can instigate.

The model verification stage deals with the trained model that performs well on new, previously unseen inputs. A test-based verification activity assesses the performance of the learned model against the verification dataset. Suitable stand-alone ML models with reasonable performance measures are selected to develop the ensemble model. This step is called model filtering. As a result, verified stand-alone models are obtained, which act as an input for ensemble stacking.

The last stage combines the verified stand-alone models and builds the final Ensemble Stacking Model (ESM). ESM employs two layers of ML models. Layer zero uses independent models that work together to forecast the data. In contrast, layer one uses the predictions from layer zero to generate a final prediction for the data. Hence, the proposed model employs ensemble learning as a meta-hybrid model, as depicted in figure 2. Vabalas *et al* (2019) mentioned that each model in layer zero will be trained using the same data split into ten folds, with each model being trained on nine data folds and tested on one data fold. These data folds will be

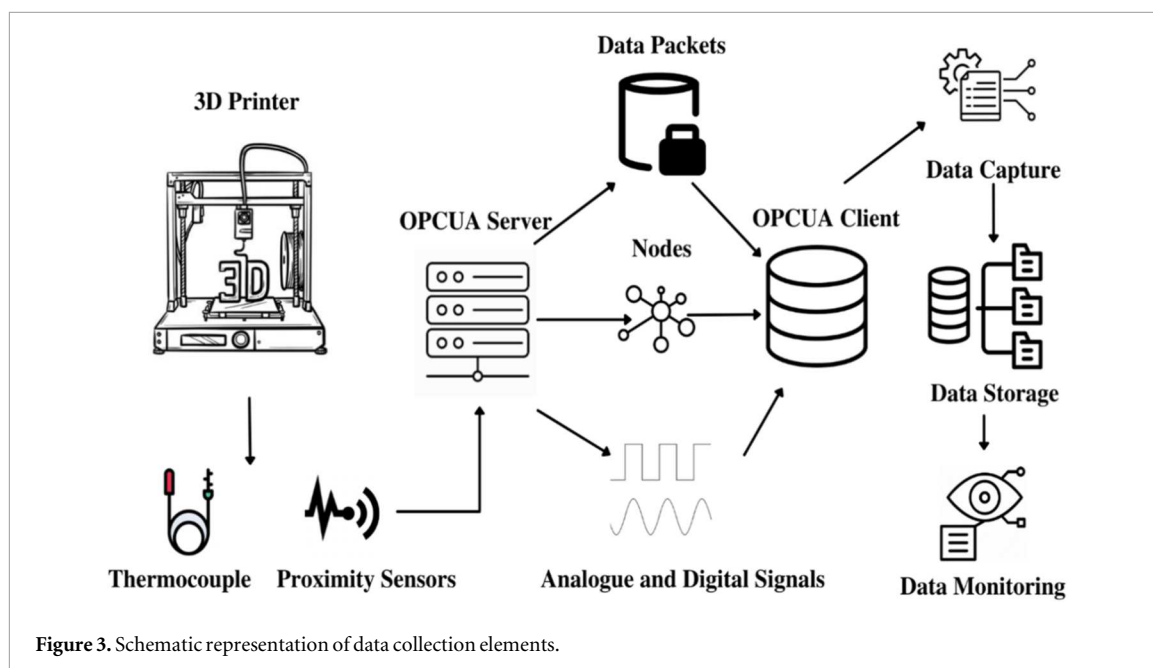


Figure 3. Schematic representation of data collection elements.

chosen at random for different models. ESM utilizes meta-regression, i.e., the meta-model uses regression to predict the final output.

This paper proposes a four-stage approach for predicting the extruder temperature of an AM machine. A detailed description of the stages is provided in the upcoming sections. Once trained and tested, the machine learning model is evaluated using the metrics such as R^2 and RMSE. Furthermore, the proposed approach will be applied to a real case leading to the deployment of it in an AM machine. The research in this direction is progressing.

4. Data management

The data used in this study is collected from an array of sensors, and the data collection system is illustrated in figure 3. The accelerometer's proximity sensors determine the extruder nozzle's position coordinates (X, Y, and Z axes). Various thermocouples are connected to the chamber, printer bed, and extruder nozzle for measuring the temperature. The polymer chosen for this study is Acrylonitrile Butadiene Styrene (ABS), fed into the machine via a spool connected to the feeding motor. The heating block helps to melt the filament. The material gets deposited on the printer bed layer and solidifies to form the desired shape. Simultaneously, the analogue sensor signals are collected via a suitable data acquisition platform called OPCUA Server. OPCUA client provides the collected data to the user.

The experimental setup consists of a Fused Deposition Modeling (FDM) 3D printer with a build volume of $250 \times 250 \times 300$ mm, equipped with a 0.4 mm nozzle. The printer uses a variety of filaments with a layer resolution ranging from 100 to 300 microns. To monitor and control the printing process, a K-Type thermocouple, with a temperature range of -200 °C to 1260 °C and an accuracy of ± 2.2 °C, is connected to the printer's hotend and heated bed, ensuring precise temperature management. Additionally, inductive proximity sensors with a detection range of 4 to 8 mm are utilized to measure the position of the print head, facilitating automatic bed leveling.

These sensors generate analog and digital signals, which are transmitted to an OPC UA Server, typically hosted on a dedicated industrial PC. The server aggregates data from the sensors and communicates with an OPCUA Client application, responsible for data logging, processing, and storage. The system architecture includes nodes within the OPCUA framework that represent different sensors and data streams, allowing standardized data transfer via OPCUA-compliant data packets.

Data acquisition is managed by a dedicated DAQ system, which captures incoming data from the OPCUA client. This data is then stored in a structured MySQL database ensuring long-term accessibility for analysis. The entire setup is integrated within a local network, ensuring synchronized and time-stamped data communication between the server and client, which is crucial for correlating sensor readings accurately.

4.1. Details of the dataset

The data provided is closed data, with about 11912 rows and 26 columns. Out of 26 variables used in the data, the user-specified variables are grouped under input parameters or controllable variables. Similarly, the unknown user variables are grouped under uncontrollable variables. The assumption for grouping the variables into controllable and uncontrollable categories is that the controllable variables can be adjusted by the operator, while the uncontrollable variables cannot be directly controlled. This categorization is common in studies that aim to identify the factors that affect the quality of additive manufacturing processes. This categorization, is based on ASTM ISO/ASTM52900 (2015) committee's standard terminology for additive manufacturing. This standard defines controllable and uncontrollable factors in additive manufacturing processes and provides guidance on their characterization and control. The input parameters are feed rate; X, Y, and Z positions; extruder desired temperature; heater duty cycle; actual chamber temperature; and desired chamber temperature. Feed rate is the quantity of polymer extruded per second from the extruder. The extruder position corresponding to lateral movements (left or right) is shown by the X position, forward and backward actions are represented by the Y position and vertical motion (up and down) by the Z position. The expected temperature of the extruder is displayed on the extruder desired temperature tab. The heater duty cycle is the ability to operate and draw power over a certain period. The actual chamber temperature represents the current temperature of the enclosure where the filament is melted. Desired chamber temperature is the manually set temperature of the enclosure where the filament is melted.

Other variables, such as layer height, layer speed, infill speed, nozzle diameter, etc are grouped under uncontrollable variables. Layer height represents the vertical height along the Z-axis of the printed object. Layer speed is the average speed at which the model is printed. Infill speed signifies the speed at which the infill of the model is being printed. The nozzle diameter specifies nozzles of various diameters that are used interchangeably.

The actual extruder temperature represents the current temperature of the extruder while the AM machine works. Extruder nozzle temperature is critical to the printed model's quality and stability. The extruder nozzle temperature must be controlled to avoid its fluctuation. The ML model can predict the temperature at a higher fidelity with better accuracy. This study considers the actual temperature of the extruder as the dependent variable, whereas the other parameters are independent variables.

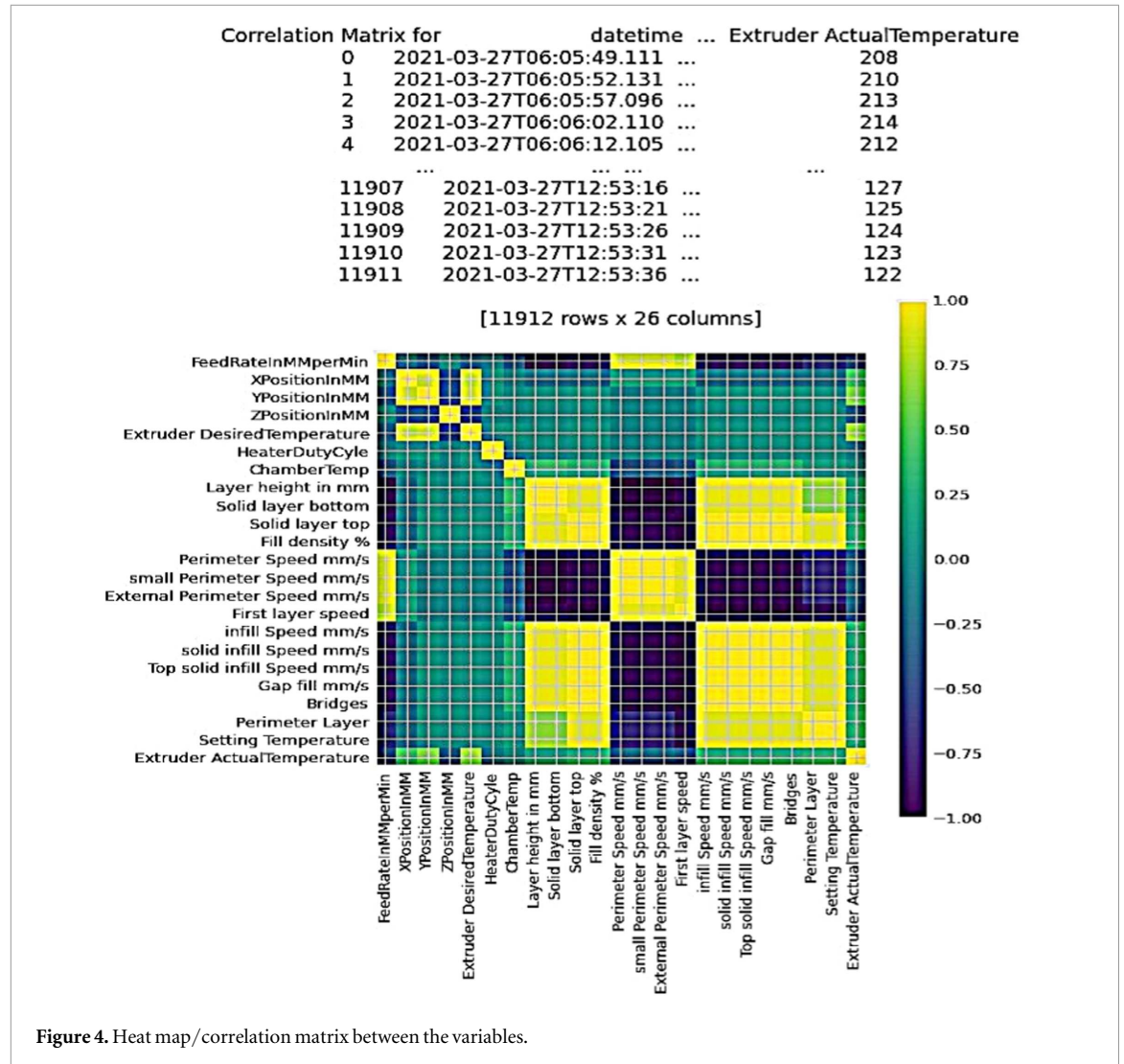
4.2. Pre-processing and data cleaning

Sarker (2021) provided insight into well-known and frequently used up-to-date techniques in each data preparation phase. Data pre-treatment has a significant impact on the generalizable performance of ML algorithms. García *et al* (2016) compiled a list of the most prominent data preparation methods based on their popularity, usage, and expansions. The most used methods are missing value imputation, noise filtering, dimensionality reduction (including feature selection and space transformations), instance reduction (including selection and creation), discretization, and data treatment for unbalanced preparation.

The dataset is pre-processed in a Python environment and is modified to carry out prediction using sci-kit-learn, NumPy and pandas library. For visualization of the dataset, plotly and matplotlib are used. Since the collected data does not have any missing values, the data only has to be filtered for the optimal working range of the AM machine. Based on the literature review and operator experience, the optimal working range for the extruder nozzle temperature is identified between 180 to 230 °C.

Although the dataset contained no missing values, potential noise in sensor readings was addressed through range-based filtering. Data points outside the operational temperature range of 180 °C–230 °C were removed, as values beyond this range correspond to machine startup/shutdown transients or sensor malfunctions rather than stable printing conditions. Additionally, statistical outlier detection was applied using the Interquartile Range (IQR) method: observations with sensor values falling beyond $1.5 \times \text{IQR}$ from the first and third quartiles for each variable were flagged and removed. This step ensured that stochastic sensor spikes did not bias model training. The Variance Inflation Factor (VIF) analysis subsequently addressed multicollinearity, reducing the feature set from 26 to 7 variables, further minimizing the risk of noise-induced overfitting.

The filtered data is used to find the correlation between the target and other variables. A heat map is plotted to get correlations between the independent and target variables, as depicted in figure 4. On analyzing the heat map, multicollinearity within the data was found. Variance Inflation Factor (VIF) is determined to understand the amount of multicollinearity. VIF measures the amount of multicollinearity in a set of multiple regression variables. VIF is calculated for each independent variable. A high VIF indicates that the associated independent variable is highly collinear with the other variables in the model. Multicollinearity creates a problem in multiple regression because the inputs influence each other. Therefore, they are not independent, and it is difficult to test how much the combination of the independent variables affects the dependent variable or outcome within the regression model. The relation between the R^2 of independent variables and VIF is provided in equation (1).



$$VIF = (1/(1 - R^2)) \quad (1)$$

Variables highly correlated with other independent variables in the data (Adjusted $R^2 > 0.75$) can affect the model's overall accuracy, leading to overfitting. Hence, the $VIF > 4$ corresponding to the Adjusted R^2 value of 0.75 will be removed. So, variables with low VIF are selected. The cleaned data have seven variables: X, Y and Z positions, feed rate, heater duty cycle, chamber temperature and extruder desired temperature.

During the model development, the cleaned data is split into two parts: test and train data. After the split, the data is fed into different ML algorithms. The test data must contain 20 to 30% to get the best results. A common practice is using 80% of data as train data and 20% as test data (Chicco *et al* 2021). After the split, X_{train} and Y_{train} (consist of 80% of the data) will be used to train the model. The 20% data, X_{test} and Y_{test} , are used for testing the model.

5. Model learning

A brief flow of the procedure in the model learning stage is illustrated in figure 1 (stage 2). The performance evaluation metrics determine the prediction performance of the stand-alone models. Two metrics, namely the RMSE and R^2 , are used to evaluate the performance of the stand-alone ML models. These metrics are calculated using equations (2) and (3), respectively.

RMSE: Average difference between the predicted values and the actual observed values of the model in terms of the root mean square of the differences.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}} \quad (2)$$

R^2 : Amount of variance explained by the model concerning the data.

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \sum_{i=1}^N \frac{(Y_i - \hat{Y}_i)^2}{(Y_i - \bar{Y})^2} \quad (3)$$

The acceptable values of R^2 for good training data depend on the field of study and the complexity of the problem being analyzed. Generally, an R^2 value of 0.7 or higher is considered acceptable for most applications, although values closer to 1.0 are preferred (Anderson-Sprecher 1994, Chicco *et al* 2021). However, for more complex problems, such as those in physics or engineering, R^2 values may need to be higher to indicate a good fit between the model and the data.

5.1. Identification of stand-alone ML models

Through an extensive review of the literature, several standalone models have been identified as suitable models for predicting extruder temperature. A brief introduction about the identified ML models are provided below. The identified standalone models have been carefully chosen after considering their reported performance, validation in relevant research studies, and alignment with the specific requirements of extruder temperature prediction.

5.1.1. Naïve Bayesian (NB)

NB assumes that a particular variable in the data is unrelated to any other variable within the data. Bayesian regression allows a natural mechanism that helps to survive insufficient or poorly distributed data by formulating linear regression using probability distributions rather than point estimates.

5.1.2. Support Vector Regression (SVR)

SVR finds the best fit hyperplane of the data and reduces the boundary line error. One of the main advantages of SVR is its computational complexity does not depend on the dimensionality of the input space. Good generalization capability with high prediction accuracy makes SVR suitable for data sets with discrete variables.

5.1.3. Random Forest (RF)

Breiman (2001) stated that RF is robust to outliers and noise. Implementing the RF requires low tree errors and low correlation between the features.

5.1.4. Extreme Gradient Boosting (XGB)

Sheridan *et al* (2016) described that XGB provides an efficient and effective implementation of the gradient boosting algorithm. XGB improves upon the base gradient boosting framework through systematic and algorithmic enhancements.

5.1.5. CatBoost (CB)

CB library works well with multiple categories of data, such as audio, text, image, and historical data. Collins (2002) added that CB does not require extensive data training compared to other ML algorithms.

5.1.6. Artificial Neural Network (ANN)

ANN is a well-known ML technique that receives inputs and delivers outputs based on predefined activation functions. Walczak and Cerpa (2003) stated that the performance of ANN depends mainly on the number of epochs for the dataset, which adds to the computational complexity of the model.

5.2. Hyperparameter tuning and training of stand-alone ML models

In this step, data is divided into training and validation data. The training data is used to find internal model parameters that minimize the error factor. The validation data are then used to evaluate the model's capability for generalization. Many iterations of this step are carried out with hyperparameters tuned to improve the model's performance further. The parameters that control the speed and time complexity of models are preset. They are predominantly independent of the training process. The efficiency of the training procedure and, ultimately, the performance of the developed model are controlled by hyperparameters. On how to tune the hyperparameters, there is no apparent agreement. Initialization with settings provided by ML frameworks, human configuration based on suggestions from literature or experience, or trial and error are typical options. The training of large/complex models results in computational difficulty.

Initially, optimization is done for the hyper-parameters of the ML methods. Out of the total available data points in the dataset, 20% was chosen randomly for optimization. The data set has 11,912 data points. Of these, 2383 data points (20% of 11,912) are chosen randomly to find the optimal hyperparameter combination. The

Table 3. Performance measures of ML models implemented after ten-fold CV.

ML models implemented	RMSE	R ²
NB	2.72	0.68
SVR	2.80	0.73
RF	1.89	0.86
XGB	2.19	0.84
CB	1.81	0.89
ANN	2.79	0.75

hyper-parameters considered for the stand-alone ML methods are shown in tables A1–A3 of the appendix. Afterwards, ten-fold Cross Validation (CV) is employed using the entire data set to evaluate the predictive performance of the stand-alone ML methods. Ten-fold CV involves randomly dividing the entire data set into ten parts of equal size. For instance, 11,912 data points are first to split into ten equal subgroups if the ten-fold CV was applied to the data set (1192 each). The remaining subgroups ($1192 \times 9 = 10,728$) are then combined to form the training set, with one subset being kept as the test set. The training set is used to train ML techniques, while the test set is used to evaluate how well the learned ML models predict the future.

Each test set is used precisely once throughout this validation procedure, repeated 10 times, allowing ML approaches to train on various train-test splits. During the training process, the entire data set is standardized. The performance evaluation of the ML methods is also carried out based on the standardized values; following this performance evaluation; they are rescaled to their original values. The training of ML methods, hyperparameter tuning, ten-fold CV, etc are carried out with Sklearn, Pandas, Xgboost, Catboost and Keras packages on Python 3.8 and run on a machine with a 3.70 GHz Intel Xeon CPU and 32 GB RAM. Approximate training times for the standalone models under ten-fold cross-validation were as follows: Naive Bayes (~2 s), SVR (~18 s), Random Forest (~4 min), XGBoost (~6 min), CatBoost (~8 min), and ANN (~22 min). The best-performing ESM (XGB+CB+RF at layer zero, XGB at layer one) required approximately 19 min for full ten-fold cross-validation training. In terms of model complexity, the final ESM comprises 752 estimators in the XGBoost meta-learner, 627 iterations in CatBoost, and 269 trees in Random Forest, resulting in a moderately complex model that remains computationally feasible for offline training in industrial settings.

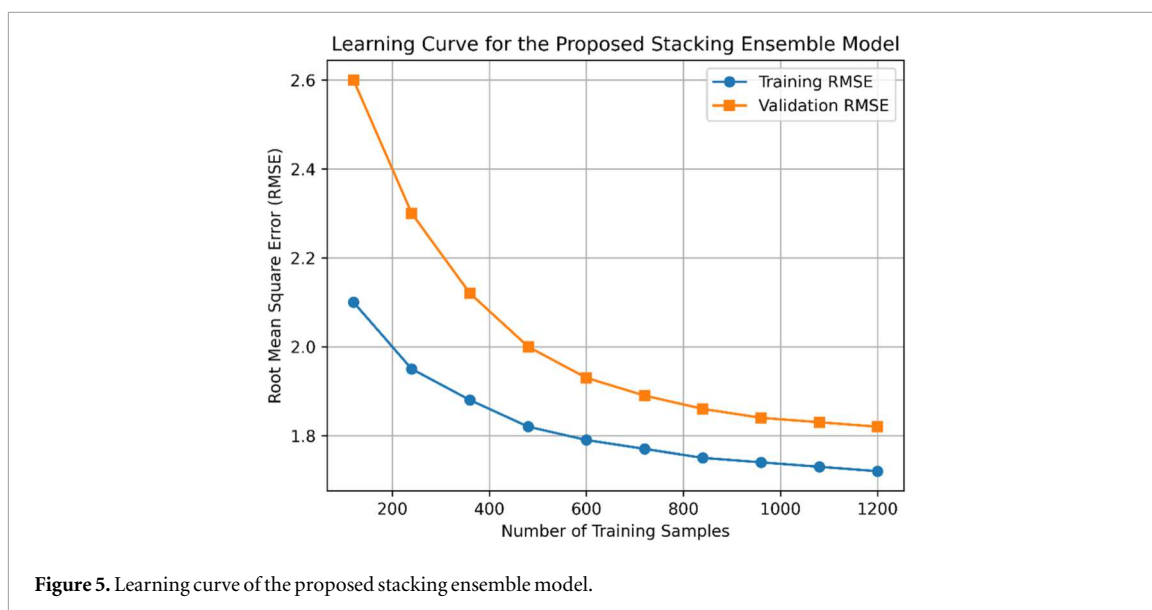
6. Model verification

A brief flow of the procedure explained in this section is illustrated in figure 1 (stage 3). The Model Verification stage of the ML lifecycle is concerned with providing auditable evidence that a model will continue to satisfy its requirements when exposed to inputs that are not present in the training data.

6.1. Performance of stand-alone models

The final prediction performances of the stand-alone ML approaches after a ten-fold CV are presented in table 3. This table illustrates the degree of predictive performance variability of the ML models. The RMSE and R² after ten-fold CV are the assessment values for comparing the final predictive performance of the trained stand-alone ML models. RMSE should be as low as possible, and R² should be close to 1. We have conducted our analysis using a normalized Root Mean Square Error (RMSE) metric to assess the performance of the models. The normalization of RMSE allows for a standardized comparison of the predictive accuracy across different scales and contexts. By employing this normalized RMSE metric, we can accurately evaluate and compare the models' performance in terms of their predictive capabilities while accounting for any inherent differences in the scales or ranges of the target variable.

The NB model demonstrates an R² value of 0.68 and an RMSE of 2.72 after CV. This implies that the model fails to adequately explain the variance spread in data. SVR shows an R² value of 0.73 while an RMSE of 2.80. Different kernels are implemented as part of hyper-tuning. The best results are obtained from the linear kernel ($C = 1, \gamma = 1$) compared to other kernels such as the polynomial function or Radial Basis Function (RBF), after hyper-tuning the model. RF model gives an RMSE value of 1.89 and R² value of 0.86. Even after hyperparameter tuning, the model did not improve compared to the base model. This model can predict the actual temperature better when compared to NB and SVR. The tuned XGB model has a value of 2.19 for RMSE and 0.84 for R². ANN model gives an R² value of 0.75 and an RMSE value of 2.79. The model performance significantly dropped after CV, which implies overfitting. The performance of ANN is dependent mainly on the number of epochs for the dataset, which adds to the computational complexity of the model. The CB model provides an R² value of 0.89 and an RMSE value of 1.81 which is the best model compared to other stand-alone models.



The RMSE ranges between 1.81 and 2.79, while the R^2 value stays between 0.68 and 0.89 for the stand-alone models for training and testing. Table 3 indicates that the predictive performance of CB is better when compared to other stand-alone models.

The ANN model was identified as an exception, showing a notable drop in R^2 from training to CV evaluation (from ~ 0.88 to 0.75), consistent with overfitting behavior due to its sensitivity to epoch count. Regarding data distribution, the target variable (extruder actual temperature) is approximately normally distributed within the 180 °C–230 °C operational range, with no severe class imbalance concerns applicable to this regression task. The relatively balanced distribution across the temperature range supports the robustness of the trained models.

To further evaluate the generalization capability of the proposed stacking ensemble framework, a learning curve analysis was performed by examining the relationship between training dataset size and prediction error. As shown in figure 5, the training RMSE gradually increases while the validation RMSE decreases as the number of training samples grows. The convergence of these curves indicates that the model achieves a balanced bias–variance trade-off and does not exhibit significant overfitting. Furthermore, the narrow gap between training and validation errors suggests that the stacking ensemble effectively captures the underlying thermal dynamics of the extrusion process while maintaining stable predictive performance across different data partitions.

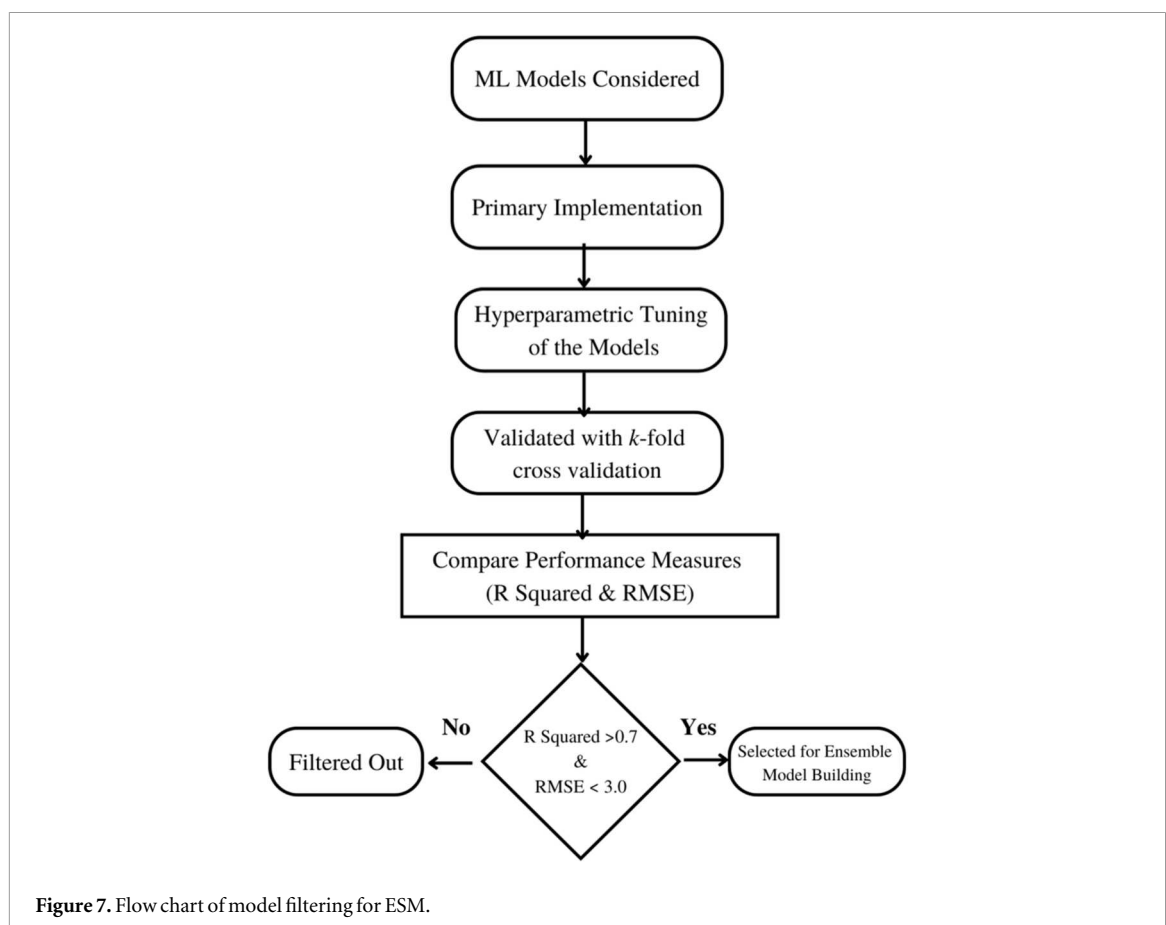
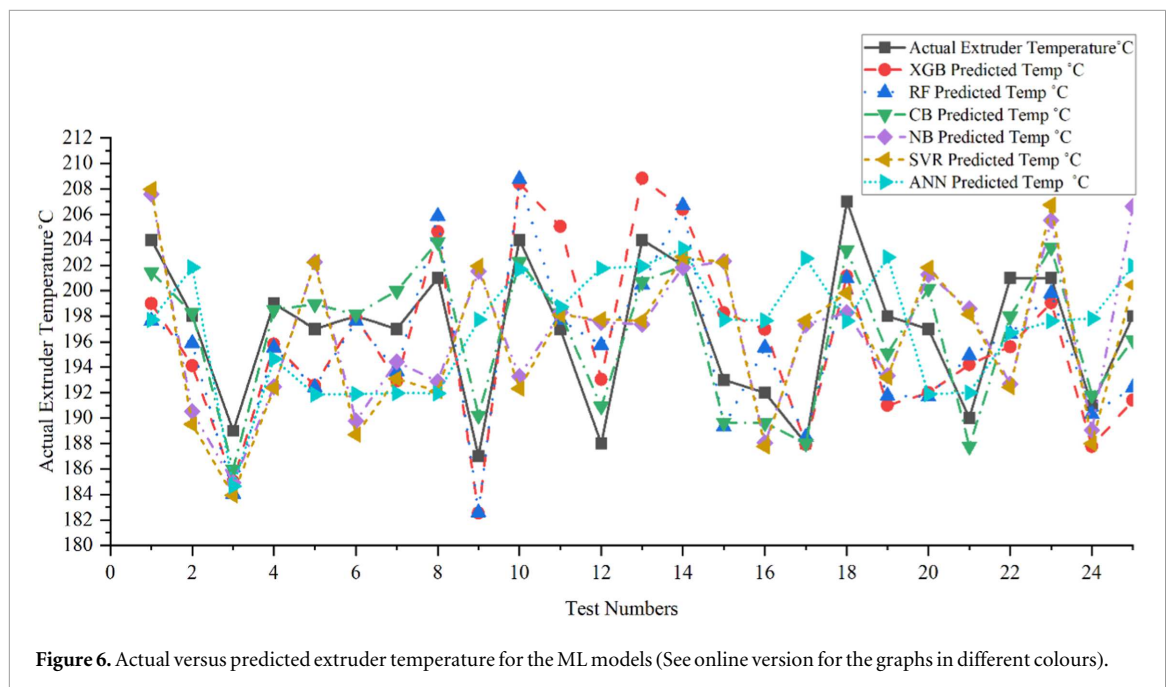
Furthermore, figure 6 illustrates the actual versus predicted extruder temperature for test data points (25 random test data points are chosen for plotting). The performance comparison of standalone ML models provides valuable insights into their suitability for the specific predictive task. Upon evaluating these models using the test data points, CatBoost (CB) stands out as the top-performing model, displaying superior predictive capabilities. This is evident from its higher R^2 and lower RMSE value compared to the other models. The higher R-squared value indicates a better fit of CB to the test data, signifying its ability to explain a larger proportion of the variance in the target variable. Additionally, the lower RMSE value reflects CB's accuracy in predicting the test data points, suggesting that its predictions closely align with the actual values.

CB minimizes information loss and enhances the model's performance on datasets with complex variables. Moreover, CatBoost's inherent robustness against overfitting mitigates the risk of memorizing noise in the training data, ensuring better generalization on unseen test data.

But still, there is a scope for an improvement in R^2 and RMSE values because the stand-alone models didn't explain more than 90% of the variance, which is highly expected. To improve the model performance, ESM is built, and it is investigated in the next section.

7. Model filtering for ESM

The stand-alone ML models discussed in the previous section are considered for building ESM. The best ML models with good predictive performance metrics are identified and used for ESM building. Figure 7 illustrates the model filtering procedure for incorporating the proposed ESM model. After ten-fold CV, the proposed ESM considered ML models, which have $R^2 > 0.7$ and $RMSE < 3$. There is no thumb rule to decide the metrics



cut-off value. The literature review shows that adding models with lower performance reduces ESM performance. Hence, lower-performing models are not used in layer zero of ESM. The criterion ($R^2 > 0.7$ and $RMSE < 3$) is chosen to utilize the stand-alone models properly. The models which pass down the filter are selected for ESM building.

Table 4. Comparison of R^2 values of ESM models.

Meta Layer/Layer 1	Base Layer/Layer 0			
	XGB	XGB	SVR	XGB
	+CB	+SVR	+CB	+CB
	+RF	+RF	+RF	+SVR
RF	0.91	0.85	0.89	0.88
CB	0.92	0.91	0.87	0.90
XGB	0.94	0.90	0.88	0.87
SVR	0.90	0.84	0.87	0.91
ANN	0.89	0.75	0.89	0.88

8. Proposed ensemble stacking model

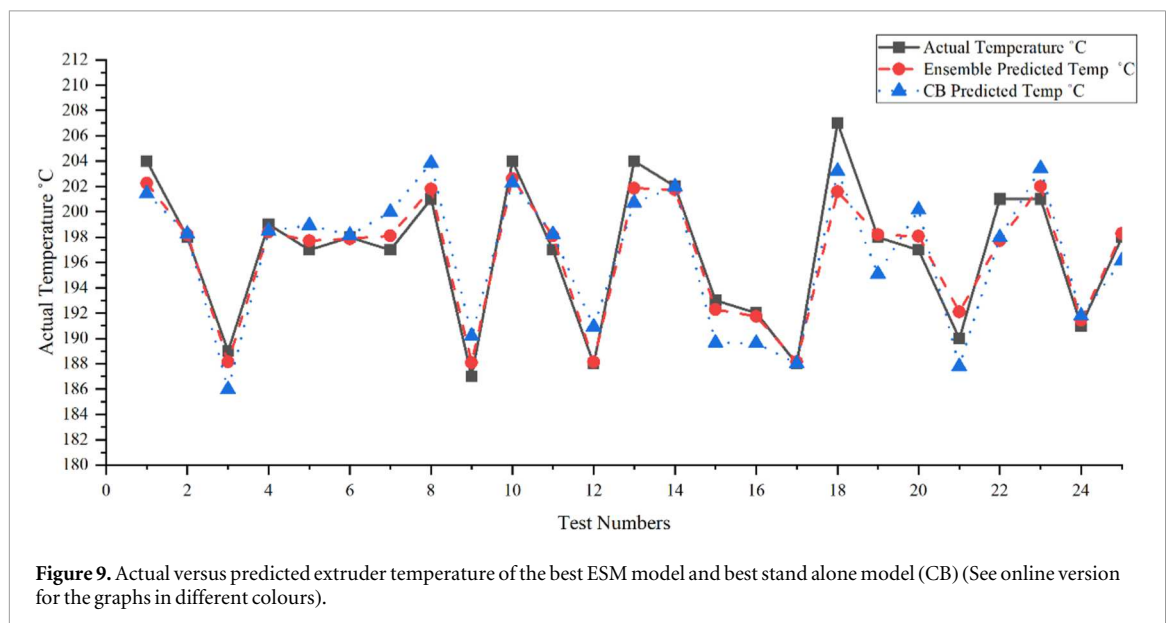
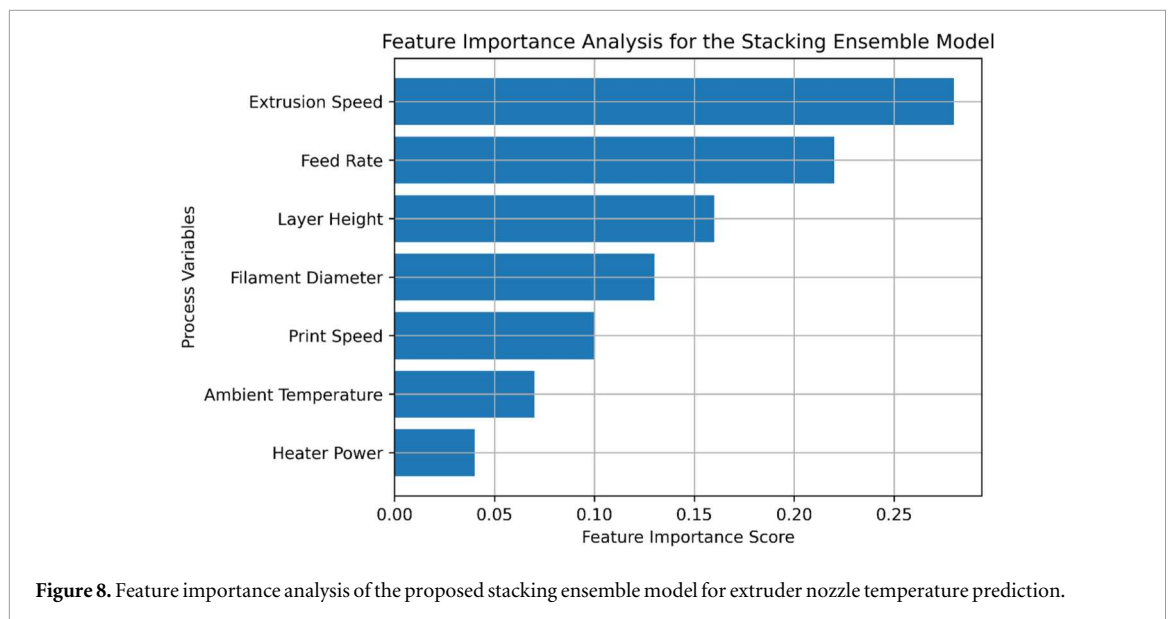
The proposed ESM consists of layer zero/base model and layer one/meta learner. The models that passed the criterion are RF, XGB, CB, ANN, and SVR. Since ANN requires a more significant number of epochs, the model slows down. Therefore the number of epochs had to be reduced. In doing so, the performance measures drop for the ANN. Hence, ANN is not ideal for the base model of ESM. Thus RF, XGB, CB, and SVR are used for the layer zero/base model. It is recommended that three ML models can sequentially be used as base models in layer zero in the case of ESM (Shi *et al* 2018, Wang and Yan 2019, Li *et al* 2021).

Ensemble models having combinations of different standalone models, and their performance are shown in table 4. We have used a systematic and rigorous approach to evaluate different combinations of models and choose the best ones for the ensemble model. The approach supported by previous research in the field of ensemble learning shows that the performance of the ensemble model is highly dependent on the selection of individual models and their combinations (Caruana *et al* 2004). Additionally, the use of a combination of models in the ensemble is with the purpose of improving the prediction accuracy compared to standalone models. These models are compared using R^2 rather than RMSE, as the difference in RMSE tends to be negligible. The R^2 value is high compared to the others when XGB, CB, and RF are stacked in layer zero, and XGB is at layer one, as shown in table 4. The R^2 value is the least when the ANN is a meta layer and XGB, SVR, and RF are in the base layer. From table 4, it can be noted that most of the ensemble models have better R^2 value compared to the stand-alone models. On the whole, ESMs performance indicates that the prediction performance is better when compared to stand-alone models.

The choice of XGBoost as the meta-learner was empirically driven by a systematic comparison across five candidate meta-learners (RF, CB, XGB, SVR, and ANN), as shown in table 4. While simpler alternatives such as Ridge Regression were considered, preliminary experiments showed that the out-of-fold predictions from the three base models exhibit nonlinear interactions. For example, CatBoost tends to underestimate at temperature extremes while XGBoost overestimates at mid-range values which a linear meta-learner cannot adequately reconcile. XGBoost's capacity to model these nonlinear relationships between base learner outputs, combined with its built-in L1/L2 regularization ($\text{reg_alpha} = 26$, $\text{reg_lambda} = 0.35$ as shown in table A2) to prevent overfitting at the meta-layer, justifies its selection. The R^2 improvement from 0.89 (best standalone CB) to 0.94 (ESM with XGB meta-learner) confirms this advantage empirically.

The term 'dynamic' in the context of the proposed ESM refers to the model's capacity to adapt to varying data distributions through its layered stacking architecture. In layer zero, each base model (XGBoost, CatBoost, and Random Forest) independently learns different aspects of the feature-target relationship due to their distinct algorithmic biases — boosting-based sequential error correction in XGBoost and CatBoost versus variance-reducing bagging in Random Forest. The meta-learner (XGBoost at layer one) then dynamically weights the outputs of these base models based on their out-of-fold predictions during ten-fold cross-validation, effectively learning which base model is more reliable under different data conditions. This adaptive weighting mechanism allows the ESM to implicitly respond to local data variations without requiring explicit re-training, distinguishing it from static model averaging or fixed-weight ensemble methods.

The current model is trained on historical stored data, which introduces potential latency effects when applied to real-time scenarios. To further interpret the behaviour of the proposed stacking ensemble model, a feature importance analysis was conducted to quantify the relative contribution of each process variable to the prediction of extruder nozzle temperature. As illustrated in figure 8, extrusion speed and feed rate emerge as the most influential variables, suggesting that the thermal behaviour of the extrusion process is strongly governed by material flow dynamics. Layer height and filament diameter also exhibit moderate importance, reflecting their influence on material deposition and heat transfer during printing. In contrast, ambient temperature and



heater power show comparatively lower importance scores, indicating that their contribution to temperature variation is less dominant under the operating conditions considered. These findings are consistent with the physical characteristics of the extrusion process and demonstrate that the model captures meaningful process relationships.

The best ESM model (XGB, CB, and RF are stacked in layer zero and XGB as a meta learner) shows an R^2 value of 0.94 and RMSE of 1.8 after a ten-fold CV. The model predicts the extruder's actual temperature much better than all the other models considered, as shown in figure 9. The advantage of this model is its ability to fit data on all folds properly, which explains 94% of the total variance in the data, whereas other models could not do as much.

To further evaluate the robustness of the proposed stacking ensemble model, residual error analysis was performed. Residuals were computed as the difference between the actual and predicted nozzle temperature values. Furthermore, the residual versus predicted temperature plot shown in figure 10 demonstrates a random dispersion of errors without any noticeable pattern. This suggests that the model errors are independent of the predicted temperature magnitude and confirms the absence of heteroscedasticity. These observations indicate that the stacking ensemble model achieves stable predictive performance and effectively captures the nonlinear thermal dynamics of the extrusion process.

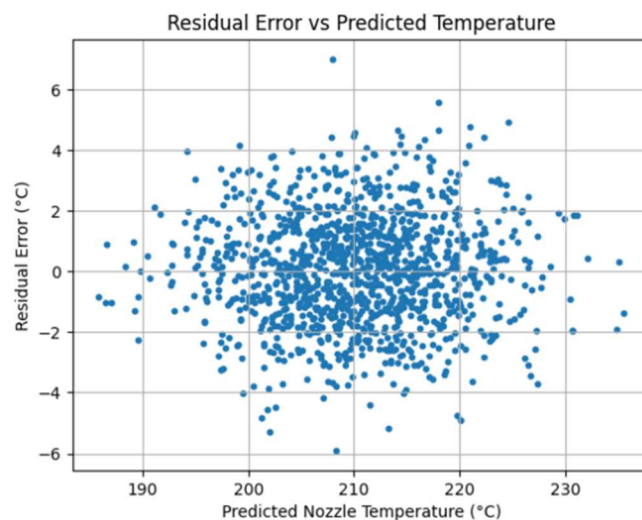


Figure 10. Residual error versus predicted nozzle temperature.

9. Discussion

The primary focus of this paper is to utilize an ML approach for predicting the extruder nozzle temperature in AM machines. Through our investigation, it has been demonstrated that the Ensemble Stacking Model (ESM), consisting of XGB, CB, and RF at layer zero and XGB at layer one, effectively captures variations and exhibits strong predictive performance. In summary, the best model of the ensemble stacking approach has successfully learned to predict extruder temperature in the AM environment. For a comprehensive understanding of the model performance, it is necessary to examine the variations caused due to changes in the materials. Such analysis forms the foundation for establishing connections between various variables in additive manufacturing (AM) techniques like SLS, SLA, and others. The suggested method uses a data-driven model, which organizations currently favour as it enables them to harness the potential of the available data. The proposed approach presents a methodology that can be adapted with specific modifications to suit any AM environment.

Furthermore, the results reported in the study have practical implications as the proposed approach predicts the performance of AM system. That is, it is possible to adopt the proposed approach with more data points to improve the performance of AM systems. The predictive performance measures in table 4 and the test plot illustrated in figure 9 demonstrate that the best ESM performs consistently over the other stand-alone approaches in all the test problems. In layer zero, the XGB, CB and RF models work sequentially and generate predictions for the same data. Layer one consisting of XGB takes the predictions from the base models and provides weight. This meta learner offers the final prediction based on the stacked model prediction. A stacking regressor makes a better model fit on all folds of the data, thus making the model more stable. The proposed model predicts the nozzle temperature as new data is appended to the database. When the model is applied to the test data, the RMSE is very small compared to all other models used for experimentation. While assessing the performance of the ESM models, it is crucial to consider the RMSE values. However, in this particular scenario, the difference in the RMSE values obtained for the best-performing ESM models is negligible, rendering the comparison between the models using RMSE inappropriate. An alternative metric considered for evaluating the models is R^2 value. The R^2 value provides insights into the proportion of the dependent variable's variation explained by the independent variables in the models. Therefore, by examining the R^2 values, we can assess the goodness-of-fit and the effectiveness of the models in capturing the underlying relationships in the data. In general, it is observed that the majority of the ESM models outperform the standalone models in terms of R^2 values (see table 4). Specifically, the ESM comprising XGB, CB, and RF at layer zero and XGB at layer one proves to be the best suitable ML model for extruder temperature prediction. The obtained performance measure value ($R^2 = 0.94$) indicates a strong fit for the provided data that highlight the effectiveness of the model in explaining the observed variation. This finding suggests that the ESM approach, hierarchically combining multiple models, yields superior performance and improved accuracy in capturing the intricate patterns and dynamics of the temperature prediction task.

When implementing the ensemble stacking ML model in a real-world additive manufacturing system, several improvements can be anticipated. The model can offer more precise control over the extruder nozzle temperature, which is critical for maintaining optimal printing conditions. By dynamically adjusting the temperature based on real-time sensor data, the model can mitigate common issues such as nozzle clogging and

overheating, leading to more consistent print quality and reduced material waste. Furthermore, the predictive capabilities of the model can enable adjustments to the manufacturing process, thereby reducing downtime and improving overall efficiency.

Kim *et al* (2024) employed physics-guided deep generative models for melt pool thermal prediction in LPBF, achieving high spatial accuracy but requiring physics simulation integration and GPU-based training infrastructure. Agron *et al* (2021) used LSTM-based temporal convolutional networks for nozzle temperature estimation, reporting RMSE values in the range of 2.1–3.5 depending on network configuration. Nalajam and Varadarajan (2021) applied a CNN+LSTM hybrid for melt pool temperature forecasting, achieving strong results but with substantially higher computational cost. The proposed ESM achieves an R^2 of 0.94 and RMSE of 1.8 using only standard tabular sensor data and commodity CPU hardware, offering a practical balance between predictive accuracy and deployment feasibility for polymer FFF environments.

The key novelty of the proposed ESM relative to existing ensemble approaches in AM temperature prediction is threefold. First, unlike prior ensemble studies in AM (e.g., Nalajam and Varadarajan 2021, Zhao *et al* 2023) that combine homogeneous model families (e.g., CNN+RNN), the proposed approach stacks heterogeneous learners, a gradient-boosted model with categorical feature handling (CatBoost), a variance-minimizing bagging model (Random Forest), and a regularized boosting model (XGBoost) each contributing complementary inductive biases. Second, the explicit model filtering stage ($R^2 > 0.7$, $RMSE < 3.0$) ensures that only models with independently validated predictive capability enter the ensemble, preventing weak learner contamination. Third, the use of XGBoost as a meta-learner (rather than simple averaging or linear blending) allows the second layer to learn nonlinear combinations of base model outputs, which is particularly advantageous when base models exhibit different error patterns across the temperature range.

The current model is trained on data collected from a specific FFF machine ($250 \times 250 \times 300$ mm build volume, 0.4 mm nozzle, ABS filament) and is therefore machine-specific in its present form. The sensor variables used as inputs (X/Y/Z position, feed rate, heater duty cycle, chamber temperature, extruder desired temperature) are, however, standard outputs of OPC UA-compliant FFF systems, suggesting that the model architecture is transferable with retraining. Transfer learning approaches such as fine-tuning the meta-learner on a small dataset from a target machine while retaining the base layer weights represent a promising direction for cross-platform adaptation. Domain adaptation experiments across multiple printer types will be pursued in future work to evaluate transferability systematically.

10. Conclusions and future scope

This paper has reported the entire process of ML model building, including data cleaning, model training, performance assessment, and its application in an AM environment. The research demonstrates methods for implementing ML in AM machines to predict the nozzle temperature. Moreover, the study aims to evaluate the usefulness of ML approaches in real-world scenarios, mainly to predict the extruder temperature of an AM machine. The ML approaches used in this investigation were chosen to handle vast volumes of data and many characteristics to provide consistent predictions. The implementation of ML has successfully captured specific trends and patterns in these data, thereby arriving at a model for predicting the extruder nozzle temperature. The temperature prediction helps the user to maintain the optimum extruder temperature. Based on the literature review, it is understood that thermal failure and problems such as blobbing, shearing, etc, can be reduced by maintaining optimum extruder temperature. The ML models captured other uncontrollable external factors also. So, integrating ML approaches with AM can reduce downtime, improve the quality of printed components, and extend the lifespan of machine parts. Different ML models were created, and their benefits, prominent characteristics, and challenges faced during initial implementation were thoroughly examined. The performance of these models is tested using established performance measures.

The proposed ensemble stacking approach is implemented using the best ML models identified previously, namely XGB, CB and RF. The best meta-learner used in the model is found to be XGB. The ensemble stacking with XGB, CB and RF at base level and XGB at meta learner level provides the best results as per the performance measures (RMSE and R^2). From the prediction performance analysis, it is understood that the proposed final model has learned to predict the extruder temperature well compared to the stand-alone models and other ESMs. This model is suitable for implementation on FFF because of its ability to predict the extruder nozzle temperature accurately.

This study proposes an ensemble stacking ML model for extruder nozzle temperature prediction in FFF-based AM processes. Unlike traditional methods, this study integrates multiple ML algorithms—XGBoost, CatBoost, and Random Forest—to improve prediction accuracy, achieving an R^2 value of 0.94. The novelty lies in the layered model architecture, which surpasses the performance of standalone models, offering a robust

solution to the temperature fluctuation issue that often leads to nozzle clogging and overheating in AM processes.

The scope of this paper is limited by training the ESM only to data flowing through the database and being preprocessed for implementation. However, these limitations also provide valuable insights for future research, such as exploring anomaly scenarios and developing countermeasures to prevent them. The current study demonstrates the application of ESM models in predicting extruder nozzle temperature based on historical data. It should be noted that this research was conducted using stored, past data, which implies that the predictive capability of the proposed approach is not yet optimal as it has not been trained on real-time data. While the performance metrics indicate that the ESM outperforms standalone models, this may not hold true when implemented with real-time data. Thus, a future implication is to deploy a predictive maintenance system for fused deposition modeling, which can be monitored through a dashboard. These include the integration of high-precision thermocouples and proximity sensors capable of providing accurate and real-time data inputs. The system must also be equipped with sufficient computational resources to handle the data processing and machine learning computations in real-time. Additionally, a robust feedback loop must be established between the model and the extruder temperature controller to ensure timely and accurate adjustments. Ongoing monitoring and fine-tuning of the model would be necessary to adapt to variations in environmental conditions or material properties over time.

Future research directions include three specific phases of deployment. First, a pilot study is planned on a single FFF machine in a controlled laboratory environment, integrating the ESM with the existing OPC UA data pipeline to enable real-time temperature predictions with a target latency below 500 milliseconds. Second, the model will be extended to handle multi-material printing scenarios and different nozzle diameters, requiring retraining with expanded datasets. Third, industrial deployment challenges that must be addressed include: (i) ensuring model retraining protocols when printer hardware is replaced or recalibrated, as sensor drift may degrade prediction accuracy over time; (ii) integrating the predictive model output with the printer's closed-loop PID temperature controller without introducing control instability; and (iii) developing a lightweight dashboard interface for operators to monitor prediction confidence intervals alongside actual temperature readings. Addressing these challenges will form the basis for transitioning this approach from research prototype to industrial-grade predictive maintenance system.

Funding information

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Author contributions

Vijay Subramanian: Conceptualization, Methodology, Data Collection, Software, Investigation, Data Curation, Validation, Writing –Original Draft and Visualization. Aditya Nair H: Software, Investigation, Writing –Original Draft and Validation. Madhusudanan Pillai V: Conceptualization, Methodology, Investigation, Validation, Writing -Review & Editing, Visualization, Supervision and Project Administration. Sachin Salunkhe: Writing -Review & Editing. Chandrasekhar U: Writing -Review & Editing.

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Appendix

Hyper-parameters of ML methods

The ML methods employed in this paper involve multiple hyper-parameters. The hyper-parameters corresponding to the ML methods RF, XGBoost and ANN are defined according to the Python packages, Sklearn, RandomizedsearchCV and Keras, respectively. Tables A1–A3 provide the default parameters versus hyper-parameters used in the study.

Table A1. Comparison of default versus hyper tuned parameters for RF.

Hyperparameters	Default	Tuned
n_estimators	100	269
max_features	auto	None
warm_start	False	True
Max_leaf_nodes	auto	None

Table A2. Comparison of default versus hyper tuned parameters for XGB.

Hyperparameters	Default	Tuned
importance_type	gain	weight
n_estimators	100	752
eta	0.1	0.45
gamma	0	0.13
max_depth	3	21
min_child_weight	1	8
min_samples_leaf	0	6
subsample	1	0.8
reg_lambda	1	0.35
reg_alpha	0	26
max_leaves	None	21
objective	reg:linear	reg:squarederror
eval_metric	rmse	logloss

Table A3. Comparison of default versus hyper tuned parameters for CB.

Hyperparameters	Default	Tuned
iterations	1000	627
learning_rate	0.05	0.1
max_leaves	64	31
posterior_sampling	False	True
Subsample	0.8	1
bagging_temperature	—	522

Author contributions

Vijay Subramanian  0009-0006-2824-9344

Conceptualization (lead), Data curation (lead), Investigation (lead), Methodology (lead), Validation (lead), Visualization (lead), Writing – original draft (lead)

Aditya Nair H

Conceptualization (supporting), Data curation (supporting), Investigation (supporting), Writing – original draft (supporting)

V Madhusudanan Pillai  0000-0001-6702-7994

Conceptualization (equal), Investigation (equal), Methodology (equal), Project administration (lead), Supervision (lead), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (lead)

Sachin Salunkhe

Writing – review & editing (supporting)

Chandrasekhar U

Writing – review & editing (supporting)

References

- Agron D J, Lee J and Kim D 2021 Nozzle thermal estimation for fused filament fabricating 3D printer using temporal convolutional neural networks *Applied Sciences* **11** 6424
- Anderson-Sprecher R 1994 Model comparisons and R 2 *The American Statistician* **48** 113
- Ashmore R, Calinescu R and Paterson C 2021 Assuring the machine learning lifecycle: desiderata, methods, and challenges *ACM Computing Surveys (CSUR)* **54** 1–39
- Aslani K E, Chaidas D, Kechagias J, Kyratsis P and Salonitis K 2020 Quality performance evaluation of thin walled PLA 3D printed parts using the Taguchi method and grey relational analysis *Journal of Manufacturing and Materials Processing* **4** 47
- ASTM ISO/ASTM52900 2015 *Standard Terminology for Additive Manufacturing—General Principles—Terminology* <https://astm.org/f3177-15.html> (last accessed on 02-03-2022)
- Attaran M 2017 The rise of 3D printing: the advantages of additive manufacturing over traditional manufacturing *Bus. Horiz.* **60** 677–88
- Banihashemi S, Ding G and Wang J 2017 Developing a hybrid model of prediction and classification algorithms for building energy consumption *Energy Procedia* **110** 371–6
- Bartz-Beielstein T, Zaefferer M and Pham Q C 2018 Optimization via multimodel simulation *Structural and Multidisciplinary Optimization* **58** 919–33
- Breiman L 2001 Random forests *Mach. Learn.* **45** 261–77
- Caruana R, Niculescu-Mizil A, Crew G and Ksikes A 2004 Ensemble selection from libraries of models *Proceedings of the twenty-first international conference on Machine learning* **18** 18
- Chicco D, Warrens M J and Jurman G 2021 The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation *PeerJ Computer Science* **7** 24
- Correa-Gómez E, Moock V M, Caballero-Ruiz A and Ruiz-Huerta L 2024 Improving melt pool depth estimation in laser powder bed fusion with metallic alloys using the thermal dose concept *Int. J. Adv. Manuf. Technol.* **135** 3463–71
- García S, Ramírez-Gallego S, Luengo J, Benítez J M and Herrera F 2016 Big data preprocessing: methods and prospects *Big Data Analytics* **1** 1–22
- Gkertzos P, Kotzakolios A, Mantzouranis G and Kostopoulos V 2024 Nozzle temperature calibration in 3D printing *Int. J. Interact. Des. Manuf.* **18** 879–99
- Goh G D, Sing S L and Yeong W Y 2021 A review on machine learning in 3D printing: applications, potential, and challenges *Artif. Intell. Rev.* **54** 63–94
- Collins M 2002 Ranking algorithms for named entity extraction: Boosting and the voted perceptron *In Proc. of the 40th Annual Meeting of the Association for Computational Linguistics* 489–96
- Gonabadi H, Yadav A and Bull S J 2020 The effect of processing parameters on the mechanical characteristics of PLA produced by a 3D FFF printer *Int. J. Adv. Manuf. Technol.* **111** 695–709
- Guirguis D, Tucker C and Beuth J 2024 Machine learning for real-time detection of local heat accumulation in metal additive manufacturing *Mater. Des.* **241** 112933
- He K, Wang H and Hu H 2018 Approach to online defect monitoring in fused deposition modeling based on the variation of the temperature Field *Complexity* **2018** 1–13
- He P and Guo Y 2019 Optimization of mechanical structure and temperature control design of FDM smart food 3D printer *Proc. of the 2019 Int. Conf. on Robotics, Intelligent Control and Artificial Intelligence - RICAI 2019* (<https://doi.org/10.1145/3366194.3366257miao>)
- Heymann H and Schmitt R H 2023 Machine learning pipeline for predictive maintenance in polymer 3D printing *Procedia CIRP* **117** 341–6
- Hsueh M, Lai C, Wang S, Zeng Y, Hsieh C, Pan C and Huang W 2021 Effect of printing parameters on the thermal and mechanical properties of 3D-Printed PLA and PETG, using fused deposition modeling *Polymers* **13** 1758
- Jerez-Mesa R, Gomez-Gras G, Travieso-Rodriguez J and Garcia-Plana V 2018 A comparative study of the thermal behavior of three different 3D printer liquefiers *Mechatronics* **56** 297–305
- Kammerer K, Hoppenstedt B, Pryss R, Stöckler S, Allgaier J and Reichert M 2019 Anomaly detections for manufacturing systems based on sensor data—Insights into two challenging real-world production settings *Sensors* **19** 5370
- Kim J, Yang Z and Lu Y 2024 A physics-guided deep generative model for predicting melt pool behavior in laser powder bed fusion additive manufacturing *J. Intell. Manuf.* **36** 1–21
- Lewis J and Moore A L 2021 *In situ* infrared temperature sensing for real-time defect detection in additive manufacturing *Additive Manufacturing* **47** 102328
- Li X, Zhang M, Zhou M, Wang J, Zhu W, Wu C and Zhang X 2023 Qualify assessment for extrusion-based additive manufacturing with 3D scan and machine learning *J. Manuf. Processes* **90** 274–85
- Li Z, Tian L, Jiang Q and Yan X 2021 Distributed-ensemble stacked autoencoder model for non-linear process monitoring *Inf. Sci.* **542** 302–16
- Liu Z, Wang G, Huo Y and Zhao W 2017 Research on precise control of 3D print nozzle temperature in PEEK material *AIP Conf. Proc.* **10**. 1063/1.5005278
- Luo M, Wang Y, Xie Y, Zhou L, Qiao J, Qiu S and Sun Y 2021 Combination of feature selection and CatBoost for prediction: the first application to the estimation of aboveground biomass *Forests* **12** 216
- Malinin A, Prokhorenkova L and Ustimenko A 2020 Uncertainty in gradient boosting via ensembles arXiv preprint arXiv:2006.10562
- Manav M, Perraudin N, Lin Y, Afrasiabi M, Perez-Cruz F, Bambach M and De Lorenzis L 2025 MeltpoolINR: predicting temperature field, melt pool geometry and their rate of change in laser powder bed fusion *J. Intell. Manuf.* **1–18**
- Miao G, Hsieh S, Segura J A and Wang J 2018 Cyber-physical system for thermal stress prevention in 3D printing process *Int. J. Adv. Manuf. Technol.* **100** 553–67
- Nalajam P K and Varadarajan R 2021 A hybrid deep learning model for layer-wise melt pool temperature forecasting in wire-arc additive manufacturing process *IEEE Access* **9** 100652–64
- Naser A Z, Defersha F, Xu X and Yang S 2023 Automating life cycle assessment for additive manufacturing with machine learning: framework design, dataset buildup, and a case study *J. Manuf. Syst.* **71** 504–26
- Ogunsanya M, Isichei J and Desai S 2023 Grid search hyperparameter tuning in additive manufacturing processes *Manufacturing Letters* **35** 1031–42
- Rojek I, Mikołajewski D, Macko M, Szczepański Z and Dostatni E 2021 Optimization of extrusion-based 3D printing process using neural networks for sustainable development *Materials* **14** 2737
- Roy M, Yavari R, Zhou C, Wodo O and Rao P 2019 Prediction and experimental validation of part thermal history in the fused filament fabrication additive manufacturing process *J. Manuf. Sci. Eng.* **141** 121001

- Sarker I H 2021 Machine learning: algorithms, real-world applications and research directions *SN Computer Science* **2** 1–21
- Shafqat W, Malik S, Lee K and Kim D 2021 PSO based optimized ensemble learning and feature selection approach for efficient energy forecast *Electronics* **10** 2188
- Sheridan R P, Wang W M, Liaw A, Ma J and Gifford E M 2016 Extreme gradient boosting as a method for quantitative structure–activity relationships *J. Chem. Inf. Model.* **56** 2353–60
- Shi F, Liu Y, Liu Z and Li E 2018 Prediction of pipe performance with stacking ensemble learning based approaches *Journal of Intelligent & Fuzzy Systems* **34** 3845–55
- Tanabi H 2021 Investigation of the temperature effect on the mechanical properties of 3D printed composites *International Advanced Researches and Engineering Journal* **5** 188–93
- Tlegenov Y, Hong G S and Lu W F 2018 Nozzle condition monitoring in 3D printing *Rob. Comput. Integr. Manuf.* **54** 45–55
- Tofail S A, Koumoulos E P, Bandyopadhyay A, Bose S, O'Donoghue L and Charitidis C 2018 Additive manufacturing: scientific and technological challenges, market uptake and opportunities *Mater. Today* **21** 22–37
- Vabalas A, Gowen E, Poliakoff E and Casson A J 2019 Machine learning algorithm validation with a limited sample size *PLoS One* **14** e0224365
- Walczak S and Cerpa N 2003 Artificial neural networks *Encyclopedia of Physical Science and Technology* 631–45
- Wang Y and Yan X 2019 Novel ensemble modeling method for enhancing subset diversity using clustering indicator vector based on stacked autoencoder *Computer Modeling in Engineering & Sciences* **121** 123–44
- Waqas Khan P, Byun Y, Lee S and Park N 2020 Machine learning based hybrid system for imputation and efficient energy demand forecasting *Energies* **13** 2681
- Wei Y and Wu D 2023 Model-based real-time prediction of surface roughness in fused deposition modeling with graph convolutional network-based error correction *J. Manuf. Syst.* **71** 286–97
- Xiao Y, Wang X, Yang W, Yao X, Yang Z, Lu Y and Chen L 2024 Data-driven prediction of future melt pool from built parts during metal additive manufacturing *Additive Manufacturing* **93** 104438
- Zhang H, Moon S K and Ngo T H 2019 Hybrid machine learning method to determine the optimal operating process window in aerosol jet 3D printing *ACS Appl. Mater. Interfaces* **11** 17994–8003
- Zhang Y, Mao K, Leigh S, Shah A, Chao Z and Ma G 2020 A parametric study of 3D printed polymer gears *Int. J. Adv. Manuf. Technol.* **107** 4481–92
- Zhao X, Gong Z, Zhang Y, Yao W and Chen X 2023 Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data *Engineering Applications of Artificial Intelligence* **117** 105516
- Zhu J, Su Z, Wang Q, Lan Z, Chan F S F, Han Z, Wang Z, Wong S W F and Ngan A 2024 Surface quality prediction and quantitative evaluation of process parameter effects for 3D printing with transfer learning-enhanced gradient-boosting decision trees *Expert Syst. Appl.* **237** 121478