

A Novel Semi-Automated Pipeline for Optimizing 3D-Printed Drug Formulations

Youssef Abdalla, Martin Ferienc, Haya Alfassam, Atheer Awad, Ruochen Qiao, Miguel Rodrigues, Mine Orlu,* Abdul W. Basit,* and David Shorthouse*

3D printing offers a promising approach to creating personalized medicines. However, costly, expertise-dependent trial-and-error methods hinder efficient drug formulation, posing challenges for tailoring treatments to individual patients. To address this, a novel pipeline is developed for 3D printing using selective laser sintering (SLS), replacing laborious steps with advanced computational methods. A differential evolution-based optimizer generates formulations for the desired drugs, while a deep learning ensemble predicts the optimal printing parameters along with associated confidence intervals. Manual handling is only required for the final formulation preparation and printing processes. The pipeline successfully generates diverse formulations, composed of a wide variety of materials and with high printability probabilities. This was validated by successfully printing 80% of the generated drug formulations and achieving 92% accuracy in predicting printing parameters. Notably, the time required to develop and print a new drug formulation is decreased to a single day. This study is the first to demonstrate a semiautomated, 3D printing drug formulation design and printing parameter selection pipeline. Furthermore, the pipeline is not limited to SLS printing but can also be adapted for the optimization of other 3D printing technologies or formulation platforms.

medications, something conventional technologies which were designed for large-scale manufacturing, struggle to achieve. Subsequently, recent clinical trials have been conducted to investigate personalized 3D-printed dosage forms.^[3]

Selective laser sintering (SLS) is a powder bed fusion 3D printing technology that primarily utilizes carbon dioxide lasers to fuse powder particles.^[4] SLS has proven highly effective for 3D printing medicines due to its simplicity, versatility in producing various drug delivery systems, and its suitability for large-scale production.^[5] Its success is further attributed to its ability to create complex 3D objects without the need for support structures and using powder feedstock materials without solvents.^[5] As a result, SLS has been successfully applied in the development of various drug delivery systems and has recently been trialed in humans for the first time.^[6] Despite its advantages and ability to outperform conventional drug manufacturing methods,^[7] trans-

lating 3D printing technologies into widespread pharmaceutical use has been slow. This is partly due to the challenges in formulating medicines compatible with 3D printing technologies, which were not initially designed for this purpose. Consequently, the current development of 3D-printable medicines relies on a trial-and-error approach, dependent on user expertise. This method is iterative, time-consuming, expensive, and wasteful.^[8]

Machine learning (ML), which leverages data for learning rather than relying on explicit programming, has gained significant interest in pharmaceutical manufacturing, enabling the

1. Introduction

3D printing, also known as additive manufacturing, encompasses a range of technologies used to fabricate 3D objects based on digital designs.^[1] Originally developed for engineering, this technology is now being explored for manufacturing medicinal products to address patient heterogeneity and the health disparities stemming from the current one-size-fits-all treatment approach.^[2] 3D printing enables the production of small batches of medicines tailored to individual patients, facilitating the production of personalized

Y. Abdalla, H. Alfassam, A. Awad, R. Qiao, M. Orlu, A. W. Basit, D. Shorthouse
UCL School of Pharmacy
University College London
29-39 Brunswick Square, London WC1N 1AX, UK
E-mail: m.orlu@ucl.ac.uk; a.basit@ucl.ac.uk; d.shorthouse@ucl.ac.uk

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202401112>.

© 2025 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202401112

M. Ferienc, M. Rodrigues
Department of Electronic and Electrical Engineering
University College London
Gower Street, London WC1E 6BT, UK

H. Alfassam
Advanced Diagnostics and Therapeutics Institute
King Abdulaziz City for Science and Technology (KACST)
Health Sector, Riyadh 11442, Saudi Arabia

A. Awad
Department of Clinical Pharmaceutical and Biological Sciences
University of Hertfordshire
Hatfield AL10 9AB, UK

development and optimization of complex drug delivery systems.^[9] In the context of 3D printing, ML has shown remarkable success, facilitating the prediction of whether drug formulations can be 3D printed,^[10] optimizing printing parameters,^[11] and determining the properties of 3D-printed medicines.^[10b,12] For SLS printing, previous studies have demonstrated the ability to predict whether formulations can be successfully SLS printed.^[13] Notably, we have developed a deep learning (DL) ensemble model capable of predicting printability with over 90% accuracy.^[13c] However, no studies have explored the prediction of optimal printing parameters for SLS printing of medicines. Two key parameters- the printing temperature and laser scanning speed - must be optimized during SLS printing. As a result, even if a formulation is printable, parameter optimization remains iterative and resource-intensive. Moreover, designing formulations to predict their printability relies heavily on scientific expertise. If a formulation is unprintable, scientists must use their heuristic knowledge to develop a new formulation, requiring significant experience to do so accurately and often leading to inefficient experimental loops of refinement and testing. Currently, no robust methods exist for generating drug formulations for 3D printing. Due to the complexity of drug formulations and the limitations of available data, only one study to date has attempted to generate 3D-printed drug formulations, though with limited exploration and success.^[14] These challenges must be addressed before the full potential of SLS printing in pharmaceutical applications can be realized.

To address these challenges, building on our previous work,^[13b] we developed a novel DL and differential evolution (DE)-based pipeline to automate the SLS drug formulation design and printing parameter selection process. To overcome formulation design and optimization issues, we propose a system that allows researchers to input a nonprintable formulation or an unformulated drug into the algorithm, which then generates an optimal formulation for SLS printing. To address the challenge of determining printing parameters, we developed DL models capable of predicting these parameters with associated confidence intervals. By eliminating human intervention in the initial trial-and-error loop and automating the iterative process of designing formulations, while only requiring human input in the final formulation preparation and printing stage, our pipeline demonstrates the ability to both generate new formulations and optimize nonprintable formulations to make them printable. We validated this approach by successfully printing 80% of the generated formulations and achieving 92% accuracy in predicting printing parameters. This demonstrates the first optimized and automated drug formulation process. This technology has broad applicability, extending to other 3D printing technologies and other drug formulation challenges, and presents the first step towards more efficient drug design and development.

2. Results and Discussion

2.1. Automating the Formulation Design and Printing Parameter Selection Processes

Currently, SLS printing follows a two-step process shown in Figure 1A. A formulation scientist designs a new formulation

for their drug, which is then inputted into a neural network (NN) to predict printability. If the formulation is unprintable, the researcher must iteratively design and trial new formulations until a printable one is identified. Once a formulation is considered printable, the printing parameters required for successful printing - the printing temperature (ranging from room temperature to 300 °C) and laser scanning speed (ranging from 0–500 mm s⁻¹)- are still unknown. Therefore, the researcher must conduct multiple printing trials, varying printing parameters until successful printing is achieved; this process can take multiple weeks. Both formulation design and parameter optimization rely heavily on user expertise, making this process time-consuming, wasteful, and costly. To circumvent this, our proposed pipeline, shown in Figure 1B, reduces the need for expertise and the trial-and-error approach. In this streamlined process, a new formulation, or an unformulated drug, is inputted into the NN to predict printability. If the formulation is unprintable, it is passed to our optimizer, which modifies the formulation to make it printable. Once the formulation is printable, it is inputted into NNs that predict the optimal printing temperature and laser scanning speed. The optimized formulation and its printing parameters can then be used by formulation scientists to successfully print the pharmaceutical drug product, eliminating the most time-consuming steps in the 3D printing process. This entire workflow can be completed within a day, with only the final formulation preparation and printing requiring manual input.

2.2. DE is the Best Formulation Optimizer

The formulation optimizer is an integral component of our pipeline (Figure 2). Briefly, an unknown formulation or drug to be formulated is inputted into the optimizer, which proposes a new formulation. The optimization problem is framed as a constrained maximization task to determine the most printable formulation. The optimizer generates a candidate formulation, which is then inputted into the printability-predicting NN that returns a printability probability score between 0 and 1. Guided by the mean-squared error (MSE) of this score, the optimizer iteratively generates and evaluates new formulations through the NN, refining the material proportions to minimize the MSE loss (maximize the printability prediction). This process is repeated *n* times, and the formulation with the highest printability score is returned as the optimal formulation. Additionally, this pipeline supports the personalization of formulations by allowing constraints on the presence or absence of certain materials. For instance, specific amounts of materials can be fixed if desirable, or particular materials can be excluded to account for factors such as allergies. A similar pipeline was proposed by Erps et al.,^[15] who used Bayesian optimisation (BO) to identify new photocurable inks for 3D printing. However, their process required manual printing at each iteration of the optimization loop, while the pipeline proposed herein is automated by replacing the printing step with an ensemble NN predicting printability and only requires printing at the final step—when the optimized formulation is outputted.

Four different optimizers were trialed in the current pipeline. The first was a random search (RS) algorithm,^[16] which

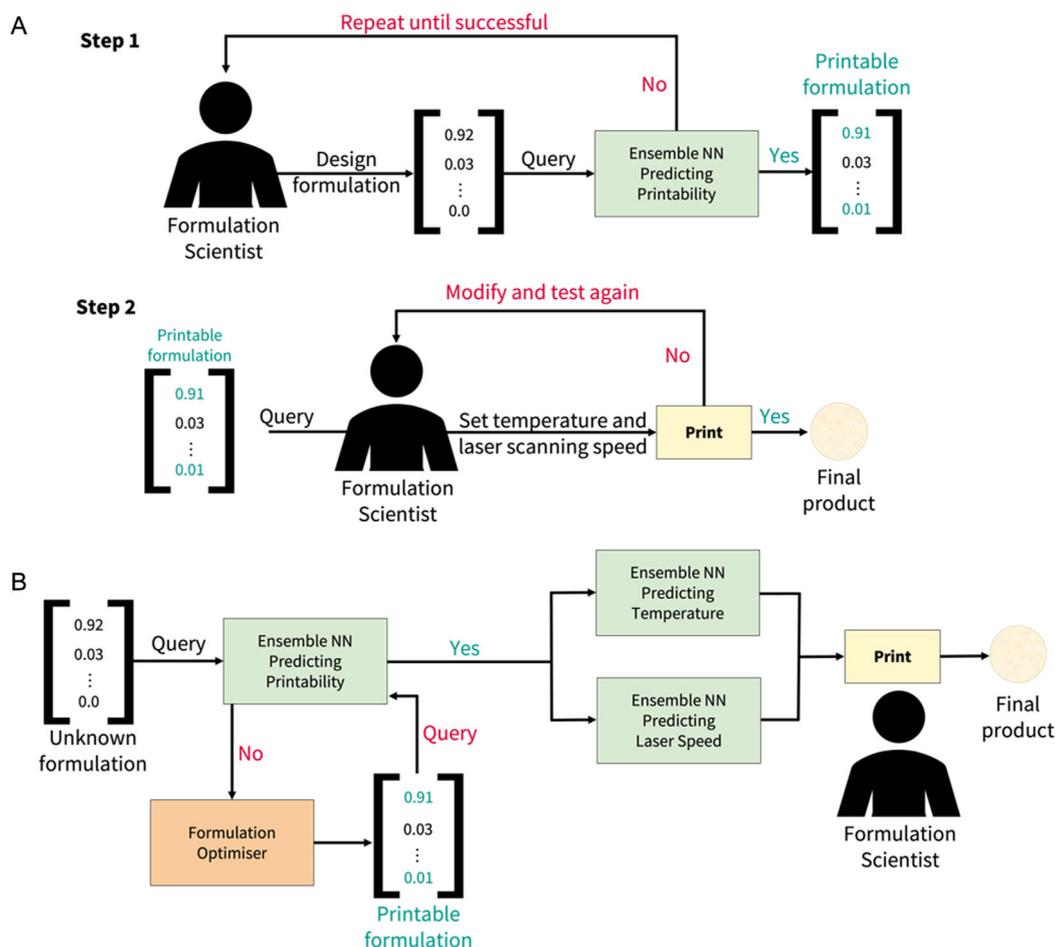


Figure 1. A) The current SLS 3D printing pipeline is a two-step process. A formulation scientist designs a drug formulation and inputs it into the NN. If it is unprintable, they modify it until it is printable. Once printable, they repeat the printing trials with various parameters. B) Our proposed novel pipeline for SLS 3D printing allows a scientist to input an unknown formulation into the ensemble NN. If unprintable, the optimizer provides a new optimized drug formulation. The parameters for printable formulations are then predicted for the final printing of the medicine.

randomly samples from Gaussian noise to generate material proportions to be added to the existing formulation. The second was BO,^[17] which uses a Gaussian process regressor to model the relationship between material proportions and MSE of the printability, incorporating uncertainty estimates to explore the search space and identify the optimal formulation. The third approach was gradient descent (GD),^[18] which computes the gradient of the printability MSE score for each material proportion and updates the proportions in the direction that minimizes loss. Finally, a DE algorithm,^[19] a genetic optimization technique, was used, which generates a population of potential formulations and produces new candidate configurations through vector differences and combinations between distinct members of the population.

One hundred and twenty-four unprintable formulations, including one hundred and fifteen materials from our previous publication,^[13c] were inputted into the four optimizers. All four optimizers led to a significant improvement in printability probability compared to the unoptimized formulations ($p < 0.001$, one-way analysis of variance (ANOVA) with post-hoc Tukey's

test), as determined by the deep ensemble. Furthermore, BO, GD, and DE performed significantly better than RS ($p < 0.05$, one-way ANOVA with post-hoc Tukey's test) (Figure 3A). Although there was no significant difference between BO, GD, and DE, DE slightly outperformed the others, with the highest average printability probability of 89.4%. DE also generated the most reasonable formulations and the fewest "hallucinations" (implausible formulations), as assessed by 3D printing experts. This aligns with the findings of Mendes et al.,^[20] who reported that DE outperformed other optimization algorithms in optimizing fed-batch fermentation in bioreactors. Figure 3B shows the distribution of printability scores before and after DE optimization. Before optimization, most predicted probabilities ranged between 0 and 5%. In contrast, after optimization, most scores were between 95% and 100%, as determined by our deep ensemble. This improvement was achieved as the optimizer systematically removed materials that negatively impacted printability and replaced them with polymers possessing desirable printing properties. For instance, in one formulation, triethyl citrate (a nonprintable liquid) was replaced with

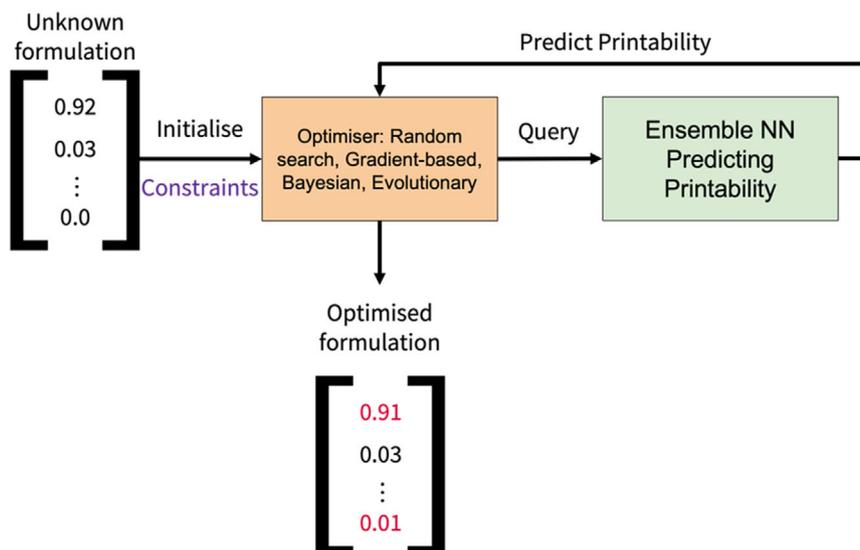


Figure 2. Optimizer pipeline for generating new formulations. The optimizer is initialized with an unknown formulation and any material constraints. It iteratively modifies this formulation, querying it with the NN predicting printability. After n loops, the formulation with the highest printability probability is output as the optimal formulation.

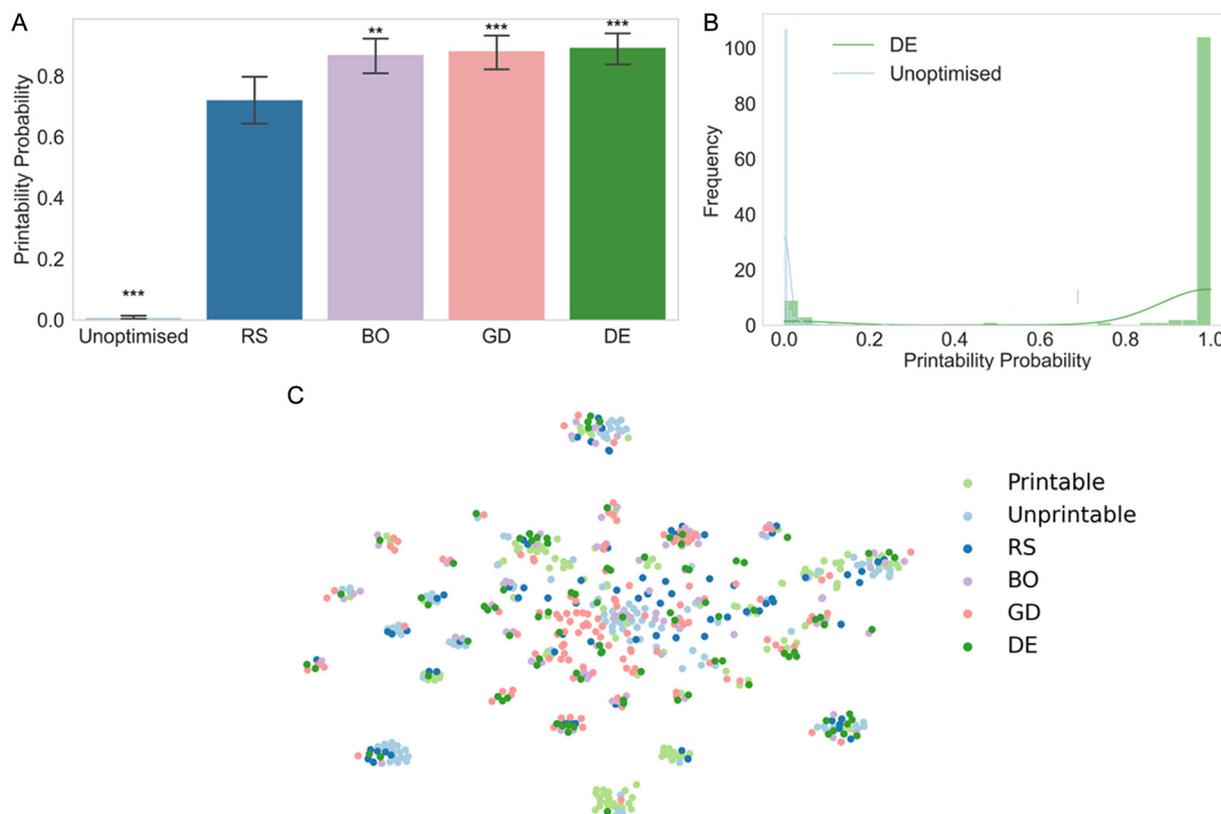


Figure 3. A) Average printability probability scores before optimization compared to the four optimization algorithms, presented as mean and 95% confidence interval. Statistical significance was assessed using a one-way ANOVA with post-hoc Tukey's test ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$). Statistical significance is shown only in comparison to RS. B) Distribution of printability scores (histogram with overlaid kernel density estimate plot) for unoptimized formulations and formulations optimized through DE. C) t-SNE visualization of printable and unprintable formulations alongside formulations optimized by each algorithm.

Eudragit RS PO, a highly printable polymer, while the rest of the formulation remained unchanged. Figure 3C presents a t-distributed stochastic neighbor embedding (t-SNE) plot of real printable formulations, nonprintable formulations, and formulations generated by the different algorithms. DE-generated formulations clustered around both the printable and nonprintable formulations, highlighting that the new formulations were diverse but minimally modified from the original nonprintable formulations. This is desirable, as the goal was to minimize changes to the initial formulation and prioritize materials already included. DE spanned the entire plot, demonstrating its ability to generate diverse formulations, but clustered most closely with real formulations, likely explaining why it had the most realistic formulations and generated the fewest hallucinations. In contrast, the greatest diversity was observed in GD, likely explaining why it had the most hallucinations among all models. Another advantage of this approach is that DE can handle multiobjective optimization problems.^[21] This adaptability means that our approach can be further extended to optimize additional properties, such as thermal stability, drug release profiles, or costs, making it a more versatile tool for pharmaceutical applications.

2.3. The Deep Ensemble is the Best-Performing SLS Parameter Predictor

Having established the optimal formulation for the drug, the next step is identifying the optimal parameters for SLS printing this formulation. To achieve this, we used the printing parameters from the printable formulations in our previous publication.^[13c] Five different ML models were trained to predict printing temperature and laser scanning speed: eXtreme gradient boosting (XGBoost), random forest (RF), support vector machine regressor (SVM), and the deep ensemble developed by us previously.^[13c] We tested two featurization methods: a one-hot-encoded vector of the drug formulation composition and the Morgan fingerprint (MFP) of the materials in the formulation, scaled by their material proportions. The latter was previously established as the best featurization method for predicting SLS printability.^[13c] Model performance was evaluated using 5-fold cross-validation (Table 1) with the mean absolute error (MAE) as the primary metric. An acceptable MAE threshold was set at below 10% of the prediction range, which corresponds to 50 mm s⁻¹ for laser scanning speed and 30 °C for surface temperature. The root mean squared error (RMSE) was also utilized; however, MAE provides a more balanced measure of overall performance, as it is not disproportionately influenced by outliers. Therefore, MAE was considered the primary error metric. Overall, all models performed within the acceptable range, and no significant differences were observed between models ($p > 0.05$, one-way ANOVA) or between the two featurization methods ($p > 0.05$, paired *t*-test). Since the MFP featurization added an extra processing step without providing a performance advantage, and the formulation vector is already used by the optimizer, we opted to use the formulation composition vector to simplify and streamline the process further. Furthermore, formulation composition has been previously identified as the best

Table 1. Cross validation scores for models trained to predict laser scanning speed or temperature, using the one-hot encoded formulation composition or MFP, presented as mean (standard deviation).

	Formulation		MFP ^{a)}	
	MAE	RMSE	MAE	RMSE
Laser scanning speed [mm s⁻¹]				
XGBoost	44.1 (7.9)	64.5 (15.5)	42.1 (12.0)	61.7 (18.2)
RF	42.4 (5.4)	60.4 (9.9)	39.2 (6.0)	59.3 (14.3)
KNN	51.4 (6.3)	67.4 (7.8)	41.2 (8.9)	57.2 (14.8)
SVM	41.7 (9.3)	65.2 (18.4)	41.7 (10.1)	65.2 (18.4)
Ensemble	45.6 (15.1)	70.2 (25.9)	43.0 (17.6)	67.9 (28.5)
Temperature [°C]				
XGBoost	19.9 (4.7)	26.1 (5.1)	15.8 (5.1)	23.4 (8.0)
RF	21.4 (3.2)	27.6 (4.6)	18.4 (3.7)	25.3 (4.0)
KNN	15.7 (4.3)	23.1 (5.8)	14.7 (2.9)	21.7 (3.9)
SVM	22.5 (4.0)	27.4 (5.0)	22.2 (3.6)	27.0 (4.4)
Ensemble	16.7 (2.4)	21.9 (4.3)	20.8 (2.8)	27.3 (4.3)

^{a)}Best scores are in bold.

feature set for predicting 3D printing parameters in other pharmaceutical printing technologies.^[10,11,22]

To further evaluate model performance, the trained models were tested on an external validation set of 22 formulations. The deep ensemble model performed the best, with percentage MAEs of approximately 8% and 6% for the laser scanning speed and printing temperature, respectively (Figure 4). This aligns with the literature, which shows that NNs outperform traditional ML models for predicting printing parameters.^[13c,22] While the error in this study is slightly higher than reported in similar research, for example, Elbadawi et al.^[11] achieved a 4% error for predicting printing temperatures, those studies utilized different printing technologies. Notably, there is no existing research predicting SLS printing parameters, likely due to the complexity of the process, where multiple combinations of printing temperatures and laser scanning speeds can lead to successful SLS printing, depending on the desired medicine properties. This complexity makes the deep ensemble model particularly valuable due to its ability to quantify uncertainty, as described previously.^[23] The model generates predictions by averaging the outputs of multiple NNs, yielding a mean value and a standard deviation. This standard deviation not only reflects model uncertainty but also defines a range of temperatures or laser scanning speeds that users can trial. An analysis of the marginal coverage (representing the proportion of actual values falling within the model's 98% confidence interval) showed coverage rates of 73% for laser scanning speed and 78% for temperature. The uncertainty interval is crucial for SLS printing, as it highlights a range of temperatures and laser scanning speeds under which printing is likely to succeed. Users can select the optimal values within this range to match their formulation needs. For example, higher laser scanning speeds may be selected to achieve a faster drug release profile.^[24] For these reasons, the NN ensemble was chosen as the optimal model for the current pipeline.

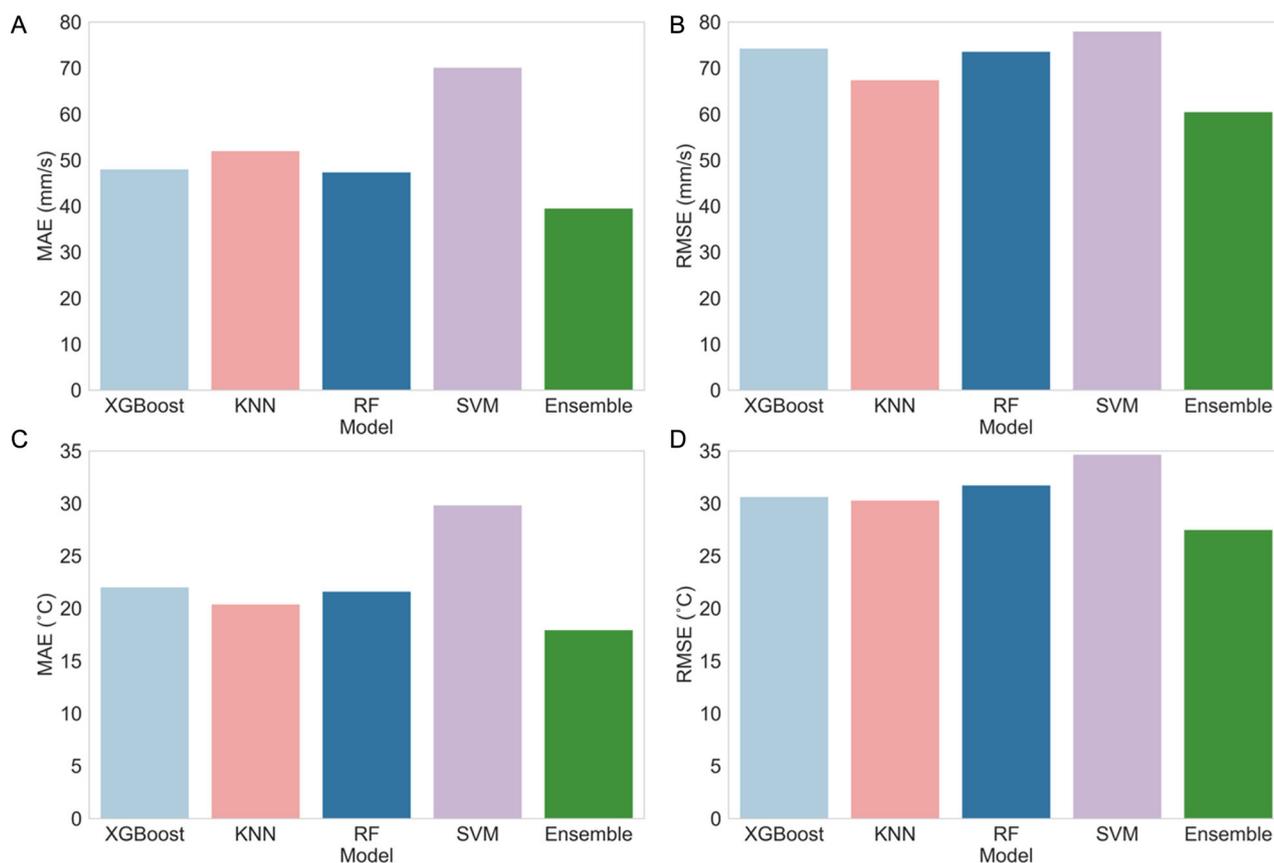


Figure 4. External validation scores for models trained on the one-hot encoded formulations. A) MAE for models that predict laser scanning speed. B) RMSE for models that predict laser scanning speed. C) MAE for models that predict temperature. D) RMSE for models that predict temperature.

2.4. Successful Printing of Generated SLS Formulations

Having validated the optimizer's and parameter predictors' success, their utility was tested in practical applications. Fifteen originally nonprintable formulations were inputted into the pipeline, and the DE algorithm modified their compositions to enhance their predicted printability as determined by the ensemble NN. Ensemble NNs then predicted the optimal laser scanning speed and printing temperatures. The newly optimized formulations, along with their predicted parameters (Table 2), were trialed using the SLS printer to assess printability. The pipeline demonstrated significant success, with 80% of the formulations printing successfully, exceeding the performance reported in a previous study exploring the generation of 3D-printed drug formulations,^[14] in which conditional generative adversarial networks (cGANs) were used to generate drug formulations for fused deposition modeling (FDM) 3D printing. This approach successfully printed 2 of 4 generated formulations out of 9 cGANs. Our superior performance likely results from the difficulty of training cGANs, particularly with small datasets.^[25] Notably, it only took us approximately one day to generate and print each formulation.

A closer analysis of the generated formulations revealed that the optimizer successfully produced heterogeneous formulations encompassing a variety of materials in different combinations.

This highlights the model's ability to generate meaningful and diverse formulations. However, the three nonprintable formulations were still identified as printable by the NN, indicating that the discrepancy arose from the NN rather than the DE optimizer.^[13c] Since the optimizer relies on the NN's predictions, the NN acts as a bottleneck in the pipeline, with the quality of the outputted formulations directly dependent on the NN's scoring accuracy. To address such discrepancies in the future, additional data is likely required to train a more robust NN ensemble.

Table 2 presents the suggested parameters and the range of trialed parameters for the nonprintable formulations, none of which led to successful printing. All three nonprintable formulations exhibited a common property not accounted for by the NN - poor powder flow. Good flow is essential for SLS formulations, as the powder bed must be uniformly distributed during printing.^[13a] In the first nonprintable formulation, the poor flow was likely due to the high proportion of the excipient mannitol, which also caused the burning of the medicine during printing due to repeated sintering of the same layer. The second formulation contained a large proportion of Opadry AMB II, a tablet coating material typically not used in high concentrations. This formulation exhibited burning at higher temperatures and crumbling at lower temperatures. Notably, only two formulations in the training set contained Opadry AMB II, which likely contributed to the model's inability to correctly classify it as

Table 2. Formulations generated through our DE optimizer, with corresponding predictions for printing parameters presented as mean (standard deviation).

Formulation ^{a)}	Temperature [°C]		Laser scanning speed [mm s ⁻¹]	
	Predicted	Real	Predicted	Real
Printable				
5% Paracetamol, 9% Kollidon SR, 83% Plasdone S-630	77 (8)	80	117 (19)	120
78.5% Ethyl cellulose CP 10, 18.5% PVA 87%–90% hydrolyzed	105 (8)	100	83 (22)	80
97% Kollicoat Protect	83 (7)	120	100 (16)	100
2% Magnesium stearate, 95% Ethyl cellulose (Aqualon N7)	95 (3)	100	57 (15)	100
5% Paracetamol, 92% Soluplus	82 (13)	60	60 (14)	50
32.07% Eudragit RL PO, 64.93% Ethyl cellulose CP 10	83 (11)	100	82 (22)	80
5% Paracetamol, 3% PEO 7M, 89% PEO 1M	36 (7)	50	142 (20)	150
20% Paracetamol, 77% Ethyl cellulose CP 10	96 (12)	120	84 (20)	80
5% Riboflavin, 6% Xylitol, 10% PVP 40000, 76% Mannitol	96 (15)	130	147 (15)	500
3% Chitosan, 28% Eudragit L100-55, 66% Hypromellose acetate succinate Aqoat AS-HG	92 (10)	100	81 (17)	100
40% Stearamide, 57% Ethyl cellulose (Aqualon N50)	78 (20)	100	91 (38)	70
23% Magnesium stearate, 74% PEO 7M	48 (7)	60	44 (15)	40
Nonprintable^{b)}				
5% Paracetamol, 92% Mannitol	144 (19)	100–140	263 (20)	80–300
42% Opadry AMB II: High Performance Moisture Barrier Film Coating Yellow, 55% HPC Klucel EF	101 (23)	50–100	75 (8)	80–100
10% Magnesium stearate, 87% Kollidon CL-M	96 (20)	90–150	77 (20)	80–100

^{a)}All formulations contained 3% Candurin. ^{b)}For nonprintable formulations, the range of temperatures and laser scanning speeds trialed are shown. “Predicted” refers to values predicted by the NNs. “Real” refers to the optimum parameter values which was determined through experimental validation.

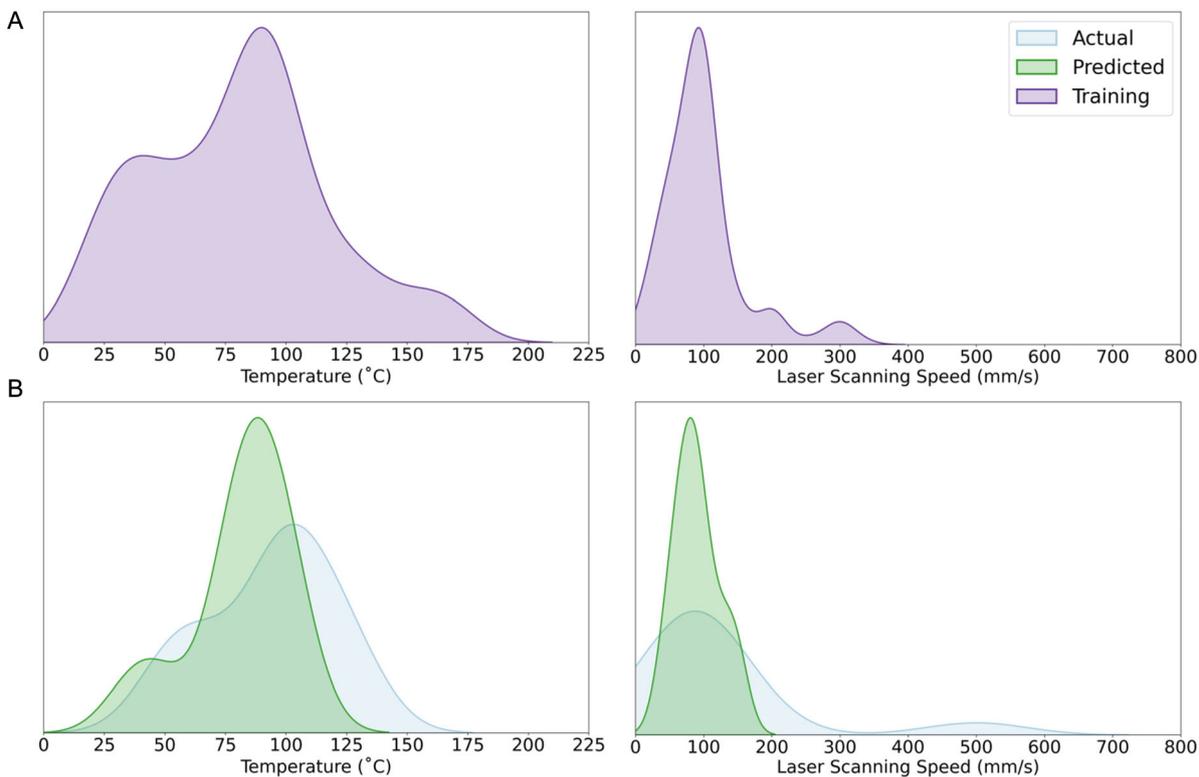


Figure 5. A) KDE plots of the training data’s distributions of printing temperatures and laser scanning speeds. B) KDE plots of the predicted and actual temperature distribution and laser scanning speeds for the DE-optimized formulations.

nonprintable. In the third formulation, the poor powder flow prevented the formation of a new powder layer with each iteration. As a result, the same layer was repeatedly sintered, preventing proper printlet formation.

For the regression tasks, the MAE remained at approximately 5.6% for temperature and 7.9% for laser scanning speed, highlighting the model's generalizability. In addition, only a single formulation fell outside the 98% confidence interval of the predicted values for both models, emphasizing that they are well-calibrated and demonstrating approximately 92% prediction accuracy. Model performance was slightly better for temperature predictions than laser scanning speed (Figure 5). This is anticipated due to the smaller range and lower temperature variability. Predictions were least accurate at the highest values for both parameters, reflecting limited data availability in these extremes. Additional data from these underrepresented regions is necessary to improve model performance further. The most challenging case for the model was the formulation containing riboflavin, with which we previously experienced significant difficulty printing. Although this formulation showed the poorest model performance, it represents a notable achievement for the optimizer by successfully predicting the composition of a formulation containing riboflavin, emphasizing its potential in addressing complex cases.

This work presents a novel approach establishing a robust pipeline for pharmaceutical formulation generation using AI. However, while the printability of medicines was assessed using a simple binary classification, pharmaceutical manufacturing processes are often more complex. Factors such as drug release, tablet hardness, friability, weight uniformity, and stability are essential pharmacopeial requirements. Therefore, future work should focus on generating new formulations and evaluating them based on their compliance with British Pharmacopoeial requirements. Furthermore, models can be trained to predict these factors and be integrated into the DE optimizer to ensure they are also accounted for in the formulation design process.

3. Conclusion

This work represents a significant step towards the optimization of 3D-printed medicines. We developed a novel pipeline for automating the design and selection of printing parameters of drug formulations for SLS 3D printing, addressing the iterative and wasteful nature of traditional drug formulation design. Amongst the various optimization algorithms and ML models tested, a DE optimizer and DL ensembles proved the most effective, forming the foundation of our pipeline. Through an in silico optimization loop, directed by a DE optimizer, the pipeline automatically generates formulations suitable for 3D printing, reducing the necessity for experimental trials. The DL ensemble predicts printability and optimal printing parameters for formulations, providing confidence intervals for these predictions. The effectiveness of this pipeline was demonstrated by successfully printing 80% of the generated formulations, with printing parameters being accurate for 92% of the formulations, and the entire process requiring approximately 1 day to generate and print a new drug formulation. This advancement significantly reduces time, waste generation, and resource demands,

as formulation scientists are only required for the final formulation preparation and printing step, rather than being involved in the entire formulation design and parameter selection processes. This allows experts to focus on specialized tasks rather than repetitive processes, increasing efficiency in pharmaceutical development. Importantly, the pipeline is not limited to SLS printing and could potentially be adapted for other 3D printing and formulation platforms, bringing us closer to a streamlined drug formulation process.

4. Experimental Section

Materials: The materials used as part of this study to make the drug formulations were AQOAT AS-HG (Shin-Etsu, Tokyo, Japan); Aqualon EC-N50, Aqualon EC-N7, Klucel hydroxypropyl cellulose EF, and Plasdone S-630 (Ashland, Schaffhausen, Switzerland); Candurin gold sheen and polyvinyl alcohol (PVA) 87%–90% hydrolyzed (Merck Life Science Limited, Dorset, UK); chitosan medium molecular weight, magnesium stearate, mannitol, paracetamol, polyethylene oxide (PEO) 1M, polyvinylpyrrolidone (PVP) 40000, stearamide (N,N-ethylenbis), and xylitol (Sigma Aldrich, Gillingham, UK); ethyl cellulose CP 10 (Fisher Scientific Ltd., Loughborough, UK); Eudragit L100-55 and Eudragit RL PO (Evonik, Darmstadt, Germany); Kollicoat Protect, Kollidon CL-M, Kollidon SR, and Soluplus (BASF, Ludwigshafen, Germany); Opadry AMB II: high-performance moisture barrier film coating yellow (Colorcon, Dartford, UK); and PEO 7M (The Dow Chemical Company, Midland, USA).

3D Printing: Drug Formulation Preparation Process: Formulations generated by our model (Table 3) were prepared by formulation scientists following the procedure outlined by Abdalla et al.^[13c] Briefly, materials were sieved using a 180 mm sieve and weighed to produce 20 g of the final product. They were mixed with a pestle and mortar until a uniform color was achieved. Before printing, the materials were sieved again using a 180 mm sieve.

SLS 3D Printing: SLS printing, done by formulation scientists, followed the procedure outlined by Abdalla et al.^[13c] Cylindrical discs (10 mm diameter × 3.6 mm height) were designed using Onshape (Version 1.160, Boston, MA, USA) and exported into the Sintratec Central program (Version 1.1, Sintratec Kit, AG, Brugg, Switzerland). The drug formulations were then loaded into the SLS printer (Sintratec Kit, AG, Brugg, Switzerland) to produce the discs, following the standard procedure described in the literature.^[26] Printing parameters were set as those predicted by our DL models. The surface temperature and laser scanning speed were set as determined by our DL model, and the chamber temperature was set at 20 °C below the surface temperature, as is the standard procedure. To ensure the stability of the final drug product, all drugs we used in this study - whether for printing, model training, or made available to optimizers - were selected based on their stability at the temperatures used for printing. Only drugs that had been previously successfully SLS printed and demonstrated to be thermally stable were included. Additionally, no printing was carried out at temperatures exceeding any drug's degradation temperature or melting/glass transition temperature.

Formulation printability was assessed as per the criteria established by Abdalla et al.^[13c] Formulations were considered printable if the resulting discs showed no deformations, visible material degradation, burning, or charring, exhibited good structural integrity and shape, and maintained integrity during post-printing processing. Each new formulation (Table 3) was allocated a single day for generation and printing, with the entire process of printing 15 formulations completed in 15 days.

Data Curation: Data was utilized from Abdalla et al.,^[13c] which comprised information on 278 pharma-inks made up of 115 materials with varying material compositions and whether they could be printed using a desktop SLS 3D printer (Sintratec Kit, AG, Brugg, Switzerland) into cylindrical discs (10 mm diameter × 3.6 mm height). This data was one-hot-encoded for further ML. Alternatively, the simplified molecular-input

Table 3. Unoptimized formulations inputted into the DE optimizer and the corresponding optimized formulations outputted.

Input formulation	Output formulation
3% Candurin, 5% Paracetamol, 47% Kollicoat IR, 20% Kollidon SR, 25% Methyl paraben	3% Candurin, 5% Paracetamol, 9% Kollidon SR, 83% Plasdone S-630
3% Candurin, 82% PVA 87%–90% hydrolyzed, 15% Citric acid monohydrate	3% Candurin, 78.5% Ethyl cellulose CP 10, 18.5% PVA 87%–90% hydrolyzed
3% Candurin, 48.5% Benecel K4M CR, 48.5% Kollicoat Protect	3% Candurin, 97% Kollicoat Protect
3% Candurin, 85% Ethyl cellulose (Aqualon N7), 12% Triethyl citrate	3% Candurin, 2% Magnesium stearate, 95% Ethyl cellulose (Aqualon N7)
3% Candurin, 5% Paracetamol, 92% Benecel A15LV PH	3% Candurin, 5% Paracetamol, 92% Soluplus
3% Candurin, 17% Tween 80, 80% Ethylcellulose CP10	3% Candurin, 32.07% Eudragit RL PO, 64.93% Ethyl cellulose CP 10
3% Candurin, 5% Paracetamol, 44% PEO 1M, 48% PVA Nippon Goshei	3% Candurin, 5% Paracetamol, 3% PEO 7M, 89% PEO 1M
20% Paracetamol only	3% Candurin, 20% Paracetamol, 77% Ethyl cellulose CP 10
5% Riboflavin only	3% Candurin, 5% Riboflavin, 6% Xylitol, 10% PVP 40000, 76% Mannitol
3% Candurin, 97% Chitosan	3% Candurin, 3% Chitosan, 28% Eudragit L100-55, 66% Hypromellose acetate succinate Aqoat AS-HG
3% Candurin, 97% Stearamide	3% Candurin, 40% Stearamide, 57% Ethyl cellulose (Aqualon N50)
3% Candurin, 20% Magnesium stearate, 77% PPG 7M	3% Candurin, 23% Magnesium stearate, 74% PEO 7M
3% Candurin, 5% Paracetamol, 5% Talc, 38% Benecel K4M, 24% Mannitol, 5% Magnesium stearate, 20% Methyl paraben	3% Candurin, 5% Paracetamol, 92% Mannitol
3% Candurin, 97% Opadry AMB II: High Performance Moisture Barrier Film Coating Yellow	3% Candurin, 42% Opadry AMB II: High Performance Moisture Barrier Film Coating Yellow, 55% HPC Klucel EF
3% Candurin, 10% Tween 80, 87% Kollidon CL-M	3% Candurin, 10% Magnesium stearate, 87% Kollidon CL-M

line-entry system (SMILES) notation for each material was obtained from PubChem, and the MFP (2048 bits, radius 2) was computed using Rdkit (Version 2022.9.5).

ML Models: All ML models were run on a Windows desktop (Operating System: Windows 11; Processor: AMD Ryzen Threadripper 7960X 24-core 4.2 GHz; RAM Memory: 128 GB, GPU: RTX 4090 24 GB). Python (Version 3.10.4) was used to run the ML models, using the Scikit-learn (Version 1.1.1) package except for XGBoost, which was run through its library (Xgboost Version 1.6.1). The deep ensemble was built and run using PyTorch Lightning (Version 2.0.4). All code is available at: https://github.com/y-babdalla/sls_optimisation.

Deep Ensemble: The ensemble NN model proposed by Abdalla et al.^[13c] was utilized for the printability classification and modified for the regression tasks. Each NN in the ensemble was a residual feed-forward network with N layers, 1D batch normalization after each layer, a hidden size of H , and a rectified linear unit (ReLU) activation function. The networks were trained for the classification tasks using a binary cross entropy loss function, while a MSE loss function was applied for regression tasks. Each NN was trained independently on the entire dataset, using the Adam optimizer with early stopping to prevent overfitting. The networks were initialized with different random weights to ensure diversity among the ensemble members. The final predictions were obtained by averaging the outputs of the individual ensemble members.

The ensemble NN took the formulation \mathbb{F} as a one-hot-encoded vector of material proportions, where the proportions sum to 1. The model was trained on experimental data with known printability labels, learning to map material proportions to printability predictions through a set of learnable parameters θ . By representing a single formulation as $\mathbb{F} = \sum (i = 1)^S f_i = 1$, where f_i represents the proportion of the i -th material out of S materials, the model can systematically learn how different material compositions affect printability and printing parameters.

For the classification task, the ensemble NN model was designed to predict the printability of a given formulation $M_\theta(\mathbb{F})$.^[13c] This predictive model took the formulation F and generated a sigmoid probability between 0 and 1 indicating the likelihood of printability. The output probability $p(\text{printable}|M_\theta(\mathbb{F}))$ provided a quantitative assessment of the formulation's potential success in printing processes. For the regression task, the model was designed to predict key printing parameters of a given

formulation. Instead of returning a probability, the model generated continuous outputs corresponding to the printing temperature $\mathbb{T}_\theta(\mathbb{F})$ or laser scanning speed $\mathbb{L}_\theta(\mathbb{F})$ along with their associated standard deviation errors, representing the prediction's uncertainty.

Model Performance: To evaluate model performance, the dataset was split into a training set and a test set. The training set's hyperparameters (Table 4) were optimized using Bayesian optimization, and model performance was assessed through fivefold cross validation. The optimal set of hyperparameters was then used to train the model on the entire training dataset, and its performance was tested on the external test set.

Model performance was evaluated using the MAE and RMSE

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (1)$$

Equation (1): MAE equation. Where N is the number of samples, y_i is the actual values and \hat{y}_i is the predicted value.

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

Equation (2): RMSE Equation. Where N is the number of samples, y_i is the actual values and \hat{y}_i is the predicted value.

Model Calibration: The marginal coverage of the ensemble models (defined as the proportion of samples whose true values fall within the 98% confidence interval of the predictions) was calculated to evaluate the model's calibration. Assuming a normal distribution of predictions, the 98% confidence interval for predictions, $\hat{C}(X)$ is

$$\hat{C}(X) = [\mu(X) - z_{\{0.99\}} \times \sigma(X), \mu(X) + z_{\{0.99\}} \times \sigma(X)] \quad (3)$$

Equation (3): Confidence interval equation, where $\mu(X)$ is the prediction mean, $\sigma(X)$ is the prediction standard deviation, and $z_{\{0.99\}}$ is the z-score for 98% confidence, The marginal coverage over a dataset with n samples is calculated as

Table 4. Model hyperparameter search space.

Model	Hyperparameter	Search space
XGBoost	max_depth	1–10
	learning_rate	0.01–1.0, log-uniform
	n_estimators	10–1000
	min_child_weight	1–10
	subsample	0.2–1.0
	colsample_bytree	0.2–1.0
RF	n_estimators	10–1000
	max_depth	1–20
	min_samples_split	2–20
	min_samples_leaf	1–20
KNN	n_neighbors	1–20
	weights	Uniform, distance
	p	1, 2
SVM	C	0.1–100, log-uniform
	kernel	Linear, RBF, poly
	gamma	0.0001–1.0, log-uniform
	epsilon	0.01–1.0, log-uniform
Deep ensemble	ensemble_members	5–10
	learning_rate	0.0001–0.1
	depth/width	32, 64, 128
	hidden_size	1, 2, 3
	weight_decay	0.01–0.00001

$$\hat{C}_{\text{marginal}} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}\{Y_i \in \hat{C}(X_i)\} \quad (4)$$

Equation (4): Marginal coverage equation, where Y_i is the true value for input X_i and $\mathbb{I}\{\cdot\}$ is the indicator function (1 if true, 0 if false).

Optimization Algorithms: Our methodology consists of two interconnected components aimed at predicting and optimizing the printability of formulations.

The first component is the aforementioned ensemble NN model designed to predict the printability of a given formulation $M_\theta(\mathbb{F})$.^[13c] The output probability $p(\text{printable}|M_\theta(\mathbb{F}))$ is used to compute the MSE loss $(1 - p(\text{printable}|M_\theta(\mathbb{F}^*)))^2$, which serves as the optimization objective for the optimizer.

The second component is an optimization algorithm designed to minimize the MSE of the printability probability of formulations. The optimization problem is structured as a constrained maximization task that seeks to find the optimal material proportions that maximize the printability prediction. The optimization is subject to six primary constraints: first, the total material proportions must sum to 1; second, each material proportion must be between 0 and 1; third, the drug proportion is kept constant to ensure personalization for the individual; fourth, the proportion of Candurin was fixed at 3%, as this is essential for successful printing; fifth, the deviation of any material proportion from its original formulation per iteration is limited by a predefined epsilon value; and sixth, the optimizer was limited to modifying up to n materials per iteration, prioritizing materials already present in the formulation. Mathematically, this can be expressed as a constrained optimization problem that aims to minimize the MSE loss of the printability score (and hence maximize printability) while maintaining these strict constraints. The user provides the initial

formulation F , which could also only contain the desired drug proportion. To address this optimization challenge, we propose four distinct approaches that offer different strategies for exploring and refining the solution space.

RS: The RS algorithm provides a straightforward approach to exploring potential formulation configurations. Given a number of iterations, the algorithm randomly samples Gaussian noise within the defined constraints to generate new material proportions. These proportions are added to the original formulation and renormalized to ensure the total sum remains 1. The best-performing formulation is retained, and the process is repeated for the specified number of iterations. The algorithm is outlined in **Algorithm 1**.

Algorithm 1: RS optimization.

Input: Initial formulation \mathbb{F} , number of iterations N , allowed noise bounds ϵ

Output: Optimized formulation \mathbb{F}^*

$\mathbb{F}^* \leftarrow \mathbb{F}$

best_score $\leftarrow (1 - p(\text{printable}|M_\theta(\mathbb{F}^*)))^2$

For $i \leftarrow 1$ to N :

a. noise \approx Gaussian($-\epsilon$, ϵ)

b. $\mathbb{F}_{\text{candidate}} \leftarrow \mathbb{F}^* + \text{noise}$

c. $\mathbb{F}_{\text{candidate}} \leftarrow \mathbb{F}_{\text{candidate}} / \sum \mathbb{F}_{\text{candidate}}$

d. $\mathbb{F}_{\text{candidate}} \leftarrow \text{ensure_constraints}(\mathbb{F}, \mathbb{F}_{\text{candidate}})$

e. candidate_score $\leftarrow (1 - p(\text{printable}|M_\theta(\mathbb{F}_{\text{candidate}})))^2$

f. **If** candidate_score < best_score:

i. $\mathbb{F}^* \leftarrow \mathbb{F}_{\text{candidate}}$

ii. best_score \leftarrow candidate_score

end

end

Return \mathbb{F}^*

Algorithm 2: GD optimization.

Input: Initial formulation \mathbb{F} , learning rate α , number of iterations N , constraint parameters ϵ

Output: Optimized formulation \mathbb{F}^*

$\mathbb{F}^* \leftarrow \mathbb{F}$

best_score $\leftarrow (1 - p(\text{printable}|M_\theta(\mathbb{F}^*)))^2$

For $i \leftarrow 1$ to N :

a. $\nabla \leftarrow \nabla_{\mathbb{F}} p(\text{printable}|M_\theta(\mathbb{F})) + \text{noise}$

b. $\mathbb{F}_{\text{candidate}} \leftarrow \mathbb{F}^* + \alpha \nabla$

c. $\mathbb{F}_{\text{candidate}} \leftarrow \mathbb{F}_{\text{candidate}} / \sum \mathbb{F}_{\text{candidate}}$

d. $\mathbb{F}_{\text{candidate}} \leftarrow \text{ensure_constraints}(\mathbb{F}, \mathbb{F}_{\text{candidate}})$

e. candidate_score $\leftarrow (1 - p(\text{printable}|M_\theta(\mathbb{F}_{\text{candidate}})))^2$

f. **If** candidate_score < best_score:

i. $\mathbb{F}^* \leftarrow \mathbb{F}_{\text{candidate}}$

ii. best_score \leftarrow candidate_score

end

end

Return \mathbb{F}^*

GD: The GD^[18] approach offers a more targeted optimization strategy. This method iteratively refines the formulation by computing the gradient of the printability score with respect to material proportions and adding Gaussian noise (to promote exploration). The algorithm updates material proportions in the direction that minimizes MSE loss (maximizes printability) while carefully projecting the solutions to ensure they remain within the defined constraints. This approach utilizes first-order optimization techniques to progressively improve the formulation's predicted printability. The algorithm is outlined in **Algorithm 2**.

Bayesian Optimization: Bayesian optimization provides a probabilistic and adaptive approach to formulation optimization by leveraging Gaussian process regression with a Matérn kernel to model the relationship between material proportions and MSE of printability. Unlike deterministic methods, this technique constructs a surrogate probabilistic

Algorithm 3: Bayesian optimization.

Input: Initial formulation \mathbb{F} , number of iterations N , acquisition function $\alpha(\mathbb{F})$

Output: Optimized formulation \mathbb{F}^*

$\mathcal{D} \leftarrow \{(\mathbb{F}, (1 - p(\text{printable}|M_\theta(\mathbb{F}^*)))^2)\}$ // Initialize dataset

GP \leftarrow Train Gaussian Process on \mathcal{D}

For $i \leftarrow 1$ to N :

a. $\mathbb{F}_{\text{next}} \leftarrow \text{argmax}_{\mathbb{F}} \alpha(\mathbb{F}, \text{GP})$ // Optimize acquisition function

b. $\text{score} \leftarrow (1 - p(\text{printable}|M_\theta(\mathbb{F}_{\text{next}})))^2$

c. $\mathcal{D} \leftarrow \mathcal{D} \cup \{(\mathbb{F}_{\text{next}}, \text{score})\}$

d. GP \leftarrow Update Gaussian Process with \mathcal{D}

end

$\mathbb{F}^* \leftarrow \text{argmin}_{(\mathbb{F}, \text{score}) \in \mathcal{D}^{\text{core}}}$

Return \mathbb{F}^*

Algorithm 4: DE optimization.

Input: Initial formulation \mathbb{F} , population size P , mutation factor MU , crossover rate CR , number of iterations N

Output: Optimized formulation \mathbb{F}^*

population \leftarrow {randomly generate P formulations near \mathbb{F} }

fitness $\leftarrow \{(1 - p(\text{printable}|M_\theta(F_{\text{candidate}})))^2 | F_i \in \text{population}\}$

For $i \leftarrow 1$ to N :

For $j \leftarrow 1$ to P :

i. $\mathbb{F}_a, \mathbb{F}_b, \mathbb{F}_c \leftarrow$ random distinct vectors from j from population

ii. $\mathbb{F}_{\text{donor}} \leftarrow \mathbb{F}_a + MU \cdot (\mathbb{F}_b - \mathbb{F}_c)$

iii. $\mathbb{F}_{\text{trial}} \leftarrow$ crossover between \mathbb{F}_j and $\mathbb{F}_{\text{donor}}$ with rate CR

iv. $\mathbb{F}_{\text{trial}} \leftarrow \mathbb{F}_{\text{trial}} / \sum F_{\text{trial}}$

v. $\mathbb{F}_{\text{trial}} \leftarrow \text{ensure_constraints}(\mathbb{F}, \mathbb{F}_{\text{trial}})$

vi. $\text{trial_fitness} \leftarrow (1 - p(\text{printable}|M_\theta(\mathbb{F}_{\text{trial}})))^2$

vii. **If** $\text{trial_fitness} < \text{fitness}[j]$:

A. $\text{population}[j] \leftarrow \mathbb{F}_{\text{trial}}$

B. $\text{fitness}[j] \leftarrow \text{trial_fitness}$

end

end

$\mathbb{F}^* \leftarrow \text{argmin}_{F_i \in \text{population}} \text{fitness}[F_i]$

Return \mathbb{F}^*

Table 5. Optimizer hyperparameter search space.

Model	Hyperparameter	Search space
RS	Number of iterations	3000
	Noise scale factor	0.03
	Number of materials	6
GD	Learning rate	0.03
	Number of iterations	4000
	Noise scale factor	0.001
Bayesian optimization	Number of materials	6
	Number of iterations	300
	Number of materials	6
DE	Population size	100
	Mutation Factor	0.4
	Crossover rate	0.8
	Number of iterations	50
	Number of materials	6

model that captures both the MSE of the predicted printability and the uncertainty associated with each candidate formulation. The primary components of this approach include a Gaussian process prior, a likelihood function based on the MSE obtained from the printability-predicting model M_θ , and an acquisition function, the expected improvement (EI), which manages the exploration-exploitation trade-off. At each iteration, the acquisition function $\alpha(F)$ determines the next candidate formulation. The algorithm is outlined in **Algorithm 3**.

DE: DE represents a genetic optimization technique that provides a nuanced approach to formulation optimization. This method maintains a population of potential solutions, dynamically generating new candidate configurations through vector differences between distinct members of the population. The algorithm selects solutions based on their fitness, which in this case is determined by the MSE of the predicted printability. By balancing exploration of new solution regions with exploitation of promising areas, DE can effectively navigate complex, nonlinear optimization landscapes. The exploration-exploitation trade-off is managed through a set of hyperparameters that control the algorithm's behavior, allowing users to fine-tune the optimization process to suit their specific requirements, consisting of the population size, mutation factor, and crossover rate. The algorithm is outlined in **Algorithm 4**.

Optimizer Hyperparameters: The hyperparameters used for the different optimizers are listed in **Table 5**.

Acknowledgements

We thank Drs. Alvaro Goyanes, Fabrizio Fina, and Sarah J. Trenfield and Ms Stephanie Chung for developing the in-house formulations used. The table of contents figure was created on BioRender. This work was funded by University College London Research Excellence Scholarship and the University College London Centre for Digital Innovation Scholarship (YA).

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Youssef Abdalla: conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (lead); investigation (lead);

methodology (lead); project administration (lead); resources (lead); software (lead); validation (lead); visualization (lead); writing—original draft (lead); writing—review & editing (lead). **Martin Ferienc:** conceptualization (supporting); formal analysis (supporting); investigation (supporting); methodology (supporting); software (supporting); writing—review & editing (supporting). **Haya Alfassam:** data curation (supporting); writing—review & editing (supporting). **Atheer Awad:** supervision (supporting); writing—review & editing (supporting). **Ruochen Qiao:** investigation (supporting). **Miguel Rodrigues:** supervision (supporting). **Mine Orlu:** supervision (equal); writing—review & editing (supporting). **Abdul W. Basit:** supervision (equal); writing—review & editing (equal). **David Shorthouse:** supervision (equal); writing—review & editing (equal).

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Keywords

automation, autonomous labs, machine learning and artificial intelligence, Pharma 4.0, powder bed fusion 3D printing, precision medications, uncertainty estimations

Received: December 18, 2024

Revised: April 14, 2025

Published online:

- [1] L. Krueger, A. Awad, A. W. Basit, A. Goyanes, J. A. Miles, A. Popat, *Nat. Rev. Bioeng.* **2024**, *2*, 801.
- [2] a) K. Englezos, L. Wang, E. C. K. Tan, L. Kang, *Int. J. Pharm.* **2023**, *635*, 122785; b) *Lancet* **2018**, *391*. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(17\)33261-0/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)33261-0/fulltext).
- [3] a) A. Goyanes, C. M. Madla, A. Umerji, G. Duran Piñeiro, J. M. Giraldez Montero, M. J. Lamas Diaz, M. Gonzalez Barcia, F. Taherali, P. Sánchez-Pintos, M.-L. Couce, S. Gaisford, A. W. Basit, *Int. J. Pharm.* **2019**, *567*, 118497; b) L. Denis, A. K. Jørgensen, B. Do, I. Vaz-Luis, B. Pistilli, A. Rieutord, A. W. Basit, A. Goyanes, M. Annereau, *Int. J. Pharm.* **2024**, *661*, 124306.
- [4] A. Awad, F. Fina, A. Goyanes, S. Gaisford, A. W. Basit, *Int. J. Pharm.* **2020**, *586*, 119594.
- [5] A. Awad, F. Fina, A. Goyanes, S. Gaisford, A. W. Basit, *Adv. Drug Deliv. Rev.* **2021**, *174*, 406.
- [6] I. Seoane-Viaño, T. Pérez-Ramos, J. Liu, P. Januskaite, E. Guerra-Baamonde, J. González-Ramírez, M. Vázquez-Caruncho, A. W. Basit, A. Goyanes, *J. Control Release* **2024**, *365*, 348.
- [7] A. Awad, A. Goyanes, A. W. Basit, A. S. Zidan, C. Xu, W. Li, R. J. Narayan, R. K. Chen, *J. Manuf. Sci. Eng.* **2022**, *145*, 010802.
- [8] A. Awad, S. J. Trenfield, T. D. Pollard, J. J. Ong, M. Elbadawi, L. E. McCoubrey, A. Goyanes, S. Gaisford, A. W. Basit, *Adv. Drug Delivery Rev.* **2021**, *178*, 113958.
- [9] a) N. Seegobin, Y. Abdalla, G. Li, S. Murdan, D. Shorthouse, A. W. Basit, *Int. J. Pharm.* **2024**, *667*, 124905; b) Y. Abdalla, L. E. McCoubrey, F. Ferraro, L. M. Sonnleitner, Y. Guinet, F. Siepmann, A. Hédoux, J. Siepmann, A. W. Basit, M. Orlu, D. Shorthouse, *J. Controlled Release* **2024**, *374*, 103; c) A. Favaron, Y. Abdalla, A. W. Basit, M. Orlu, *Expert Opin. Drug Delivery*, *1*.
- [10] a) J. J. Ong, B. M. Castro, S. Gaisford, P. Cabalar, A. W. Basit, G. Pérez, A. Goyanes, *Int. J. Pharm.: X* **2022**, *4*, 100120; b) P. Carou-Senra, J. J. Ong, B. M. Castro, I. Seoane-Viaño, L. Rodríguez-Pombo, P. Cabalar, C. Alvarez-Lorenzo, A. W. Basit, G. Pérez, A. Goyanes, *Int. J. Pharm.: X* **2023**, *5*, 100181.
- [11] M. Elbadawi, B. Muñiz Castro, F. K. H. Gavins, J. J. Ong, S. Gaisford, G. Pérez, A. W. Basit, P. Cabalar, A. Goyanes, *Int. J. Pharma.* **2020**, *590*, 119837.
- [12] a) S. Obeid, M. Madžarević, M. Krkobabić, S. Ibrić, *Int. J. Pharma.* **2021**, *601*, 120507; b) M. Elbadawi, T. Gustaffson, S. Gaisford, A. W. Basit, *Int. J. Pharma.* **2020**, *590*, 119868.
- [13] a) M. Madžarević, Đ. Medarević, S. Pavlović, B. Ivković, J. Đuriš, S. Ibrić, *Pharmaceutics* **2021**, *13*, 1969; b) Y. Abdalla, M. Elbadawi, M. Ji, M. Alkahtani, A. Awad, M. Orlu, S. Gaisford, A. W. Basit, *Int. J. Pharma.* **2023**, *633*, 122628; c) Y. Abdalla, M. Ferienc, A. Awad, J. Kim, M. Elbadawi, A. W. Basit, M. Orlu, M. Rodrigues, *Int. J. Pharma.* **2024**, *661*, 124440.
- [14] M. Elbadawi, H. Li, S. Sun, M. E. Alkahtani, A. W. Basit, S. Gaisford, *Appl. Mater. Today* **2024**, *36*, 102061.
- [15] T. Erps, M. Foshey, M. K. Luković, W. Shou, H. H. Goetzke, H. Dietsch, K. Stoll, B. von Vacano, W. Matusik, *Sci. Adv.* **2021**, *7*, eabf7435.
- [16] J. Bergstra, Y. Bengio, *J. Mach. Learn. Res.* **2012**, *13*, 281.
- [17] W. Chen, T. Paraschivescu, X. Can, *Adv. Neural Inf. Process. Syst.* **2012**, *4*, 2951.
- [18] R. Sebastian, *arXiv preprint arXiv:1609.04747* **2016**, 1065.
- [19] R. Storn, K. Price, *J. Global Optim.* **1997**, *11*, 341.
- [20] R. Mendes, I. Rocha, J. P. Pinto, E. C. Ferreira, M. Rocha, *Adv. Differ. Evol.* **2008**, 299. https://link.springer.com/chapter/10.1007/978-3-540-68830-3_13.
- [21] B. V. Babu, R. Angira, *Comput. Chem. Eng.* **2006**, *30*, 989.
- [22] B. Muñiz Castro, M. Elbadawi, J. J. Ong, T. Pollard, Z. Song, S. Gaisford, G. Pérez, A. W. Basit, P. Cabalar, A. Goyanes, *J. Controlled Release* **2021**, *337*, 530.
- [23] B. Lakshminarayanan, A. Pritzel, C. Blundell, *Advances in Neural Information Processing Systems* **2017**, p. 30.
- [24] S. J. Trenfield, X. Xu, A. Goyanes, M. Rowland, D. Wilsdon, S. Gaisford, A. W. Basit, *Int. J. Pharm. X* **2023**, *5*, 100148.
- [25] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, T. Aila, *ArXiv* **2020**, abs/2006.06676.
- [26] F. Fina, A. Goyanes, S. Gaisford, A. W. Basit, *Int. J. Pharm.* **2017**, *529*, 285.