The Journal of The Textile Institute

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjti20

Optimising the manufacturing of electrospun nanofibrous structures for textile applications: a machine learning approach

Elisa Roldán, Neil D. Reeves, Glen Cooper & Kirstie Andrews

To cite this article: Elisa Roldán, Neil D. Reeves, Glen Cooper & Kirstie Andrews (18 Mar 2025): Optimising the manufacturing of electrospun nanofibrous structures for textile applications: a machine learning approach, The Journal of The Textile Institute, DOI: 10.1080/00405000.2025.2472089

To link to this article: https://doi.org/10.1080/00405000.2025.2472089

Published online: 18 Mar 2025.

THE JOURNAL OF THE TEXTILE INSTITUTE
https://doi.org/10.1080/00405000.2025.2472089

RESEARCH ARTICLE

Optimising the manufacturing of electrospun nanofibrous structures for textile applications: a machine learning approach

Elisa Roldán[a], Neil D. Reeves[b], Glen Cooper[c] and Kirstie Andrews[a]

[a]Department of Engineering, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK; [b]Lancaster Medical School, Faculty of Health and Medicine, Lancaster University, Lancaster, UK; [c]School of Engineering, University of Manchester, Manchester, UK

ABSTRACT
Electrospun structures, known for their high porosity and surface area, can be tuned by optimising manufacturing parameters. These characteristics make them ideal for waterproof and breathable textiles, skin-like non-woven fabrics, and smart wearable bioelectronic textiles. This research aims to develop a manufacturing optimisation methodology using machine learning models to control fibre diameter and inter-fibre separation for textile applications. Polyvinyl alcohol (PVA) structures were produced with varying concentrations (10, 12, 14, 16 w/v) and different parameters such as flow rate (0.5–5 ml/h), voltage (18–25 kV), needle diameter (15–23 G), distance between needle and collector (5–11 cm), and mandrel revolution (500–3000 rpm). Data from 2560 observations of fibre diameter and inter-fibre separations were used to train 20 machine learning models. C5.0 Decision Trees and Rule-Based Models identified optimal setups, achieving high prediction accuracy for fibre diameter (0.868) and inter-fibre separation (0.861). This research advances the optimisation of electrospinning techniques for textile applications.

ARTICLE HISTORY
Received 26 July 2024
Accepted 20 February 2025

KEYWORDS
Electrospinning; machine learning models; decision trees; manufacturing optimisation; textiles morphology; non-woven fabrics

1. Introduction

Electrospinning is a highly adaptable manufacturing technique capable of producing micro- and nano-structured mats characterised by a large specific surface area, microporosity, and high porosity (Wang et al., 2023). These properties are particularly advantageous for applications in tissue engineering, breathable and waterproof textiles, and smart wearable bioelectronic textiles (Liu et al., 2020). The diameter of the fibres and the porous size (defined in this article as the inter-fibre separation, or the maximum horizontal distance between fibres forming the pore) can be tailored according to the intended use of the electrospun mats. For example, to create a fabric that mimics the morphology of the extracellular matrix (ECM) of human tissues such as skin, the diameter of fibres should range between 40 and 150 nm, replicating collagen fibrils, and they should exhibit high inter-fibre separation (Maleki et al., 2022). To manufacture breathable and waterproof textiles, the goal is to produce microporous structures smaller than the smallest raindrops (100 µm) but larger than water vapor molecules (40 nm) (Mukhopadhyay and Vinay Kumar, 2008). In the case of smart wearable bioelectronic textiles, reducing the fibre diameter increases the specific surface area, providing numerous interaction sites with the environment; additionally, high porosity offers high number of channels, facilitates transport and enhances sensitivity (Liu et al., 2020).

To tailor the fibre diameter and inter-fibre separation of the structures for specific applications, various electrospinning manufacturing and environmental parameters can be adjusted. Solution parameters such as viscosity, concentration, molecular weight, surface tension, and conductivity can be modified. Electrospinning process parameters like voltage, flow rate, needle gauge, type of collector, mandrel revolutions, and the distance between the needle and the collector can also be adjusted. Furthermore, environmental parameters such as relative humidity and temperature can be controlled to produce high-quality nanofibre mats with the desired morphology (Haider et al., 2018). However, determining the optimal combination of these parameters is challenging, time-consuming, and requires extensive experimentation to produce high-quality, non-beaded fibre mats with the ideal morphology and mechanical properties for specific applications (Roldán et al., 2023a).

Machine learning (ML) techniques offer highly accurate predictions in non-ideal situations involving poor experimental design, imbalanced data, complex nonlinear interactions, or non-parametric conditions (Bzdok et al., 2018). Given specific manufacturing parameters, ML techniques

CONTACT Elisa Roldán Elisa.Roldan-Ciudad@mmu.ac.uk Department of Engineering, Faculty of Science & Engineering, Manchester Metropolitan University, Manchester M1 5GD, UK.

can infer and predict the morphology, topography and mechanical properties of electrospun mats (Roldán et al., 2024a, 2023b), thus optimising the manufacturing setup for specific purposes. To the authors' knowledge, only two studies have focused on optimising electrospinning parameters through ML models to create ideal scaffolds for specific purposes (Roldán et al., 2023c; Roldán et al., 2023a). These studies concentrated on creating biomimetic vascular grafts and only explored Chi-squared Automatic Interaction Detection (CHAID) models, artificial neural networks (ANN), and traditional statistical approaches.

Some of the most popular classification ML techniques are: Multiple Logistic Regression Model (LM), Linear Discriminant Analysis (LDA), Lasso and Elastic-Net Regularized Generalized Linear Models (GLMNET), k-Nearest Neighbours (KNN), Classification and Regression Trees (CART), Random Forest (RF), C5.0 Decision Trees and Rule-Based Models (C50), Bagged Model (BAG), Stochastic Gradient Boosting Models (GBM) and Support Vector Machine (SVM).

Multiple Logistic Regression Model (LM) is a binary classifier able to predict if an event occurs or not, this model was used to predict the diameter of polyvinyl alcohol (PVA) electrospun fibres, their orientation and orientation, in a recent study (López-Flores et al., 2024). Linear Discriminant Analysis (LDA) are used in multi-class classification problems to reduce data dimensionality, recent studies explored this technique to assess the quality of the PVA electrospun fibres (Ieracitano et al., 2021). Lasso and Elastic-Net Regularized Generalized Linear Models (GLMNET) are indicated when there are extreme correlations between estimators, and employs a combination of the l1 component (lasso) and the l2 component (ridge regression) penalties in the algorithm (Tay et al., 2023), this technique was used to predict the water contact angle, oil absorption capacity, and mechanical strength of electrospun polystyrene/polyacrylonitrile (Wang et al., 2020). K-Nearest Neighbours (KNN), Classification and Regression Trees (CART), Random Forest (RF), Bagged Model (BAG), Stochastic Gradient Boosting Models (GBM) and Support Vector Machine (SVM), have been used, in a recent study, to predict the mechanical behaviour of 2D and 3D PVA electrospun scaffolds under a wide variety of testing conditions (Roldán et al., 2024b). C5.0 Decision Trees and Rule-Based Model (C50) is developed based on the Iterative Dichotomiser 3 (ID3) (Jin et al., 2009) and C4.5 (an extension of ID3) (Quinlan, 2014) decision trees models. Although CHAID models, based on decision trees algorithm, have demonstrated to be a powerful technique to identify the path to follow to obtain a specific objective (Roldán et al., 2023a), the accuracy of their predictions can be improved with C5.0 algorithm never explored for electrospun scaffolds. Even though some studies have been done with electrospun mats and popular classification ML techniques, none of them have comprehensively evaluated a broad spectrum of ML methods to predict both the fibre diameter and inter-fibre separation of the electrospun structures with the purpose of optimising the manufacturing parameters.

The present study aims to manufacture a series of electrospun structures to inform the development of a number of popular ML prediction models mentioned above. The suitability of 20 classification ML models is compared for their ability to predict the fibre diameter and inter-fibre separation of electrospun structures and to optimise the manufacturing parameters used in electrospinning for specific textile manufacturing applications. Tree-like diagrams reported in this study provided a visual and interpretable route to determine optimal input parameters for textiles applications such as producing biomimetic tissue-engineered scaffolds, breathable and waterproof textiles or smarts wearable bioelectronic textiles. The novel approach of applying C5.0 algorithm to determine the ideal manufacturing set-up will substantially reduce the production time and cost of electrospun textiles while providing the highest accuracy among all the ML techniques analysed.

2. Materials and methods

2.1. Scaffold production and characterisation

PVA (#P8136, molecular weight 30,000–70,000, Sigma Aldrich, UK) and distilled water were used to prepare homogeneous solutions with concentrations of 10, 12, 14 and 16% by heating at 100 °C and stirring for 1 h. All the electrospun meshes were manufactured with an electrospinning device (Prefector, Spraybase®, Maynooth, Ireland) at 25 °C by systematically modifying the concentration (10–16%), flow rate (0.5–5 ml/h), voltage (18–25 kV), diameter of the needle (15–23 G), distance between needle and collector (5–18 cm), type of collector (flat, rotational 8 and 15 cm) and revolution of the mandrel (500–3000 rpm). Three replicates were manufactured for each different scaffold to ensure the repeatability of the results. Crosslinked samples were prepared using 25% glutaraldehyde (GTA), sourced from Sigma Aldrich (UK). The crosslinking process was conducted on the optimised electrospun scaffolds, found through machine learning models, *via* vapour deposition by pouring 25 ml of 25% GTA into a Petri dish positioned at the base of a sealed desiccator. The samples were placed on a metallic mesh above the Petri dish and exposed to GTA vapour for 24 h. Following the crosslinking process, the samples were dried under a fume hood for an additional 24 h to remove residual moisture and minimise the toxicity associated with GTA exposure.

Samples were coated with gold using a SC7640 sputter coater (Quorum Technologies Ltd., Kent, UK) before visualization with a field emission scanning electron microscope (Zeiss Supra 40, FE-SEM, Carl Zeiss SMT Ltd., Cambridge, UK). The coating was performed at an intensity of 20 mA, a voltage of 0.8 kV, and for a duration of 120 s. SEM images of each sample were taken at a working distance of approximately 6 mm, with a voltage of 2 kV, and at magnifications of 1000×, 2000×, and 3000×. Fibre diameters and inter-fibre separations were measured using AxioVision SE64 Rel. 4.9.1 (Carl Zeiss SMT Ltd., Cambridge, UK) by analysing 20 fibres per sample following a previous study (Roldán et al., 2024c). A total of 2560 observations of diameter of fibres and inter-fibre separations of the electrospun meshes were obtained to inform the ML models.

2.2. Statistical analysis and machine learning models

An initial exploratory analysis and a treatment of aberrant data was performed prior to the development of the models. The normality and homoscedasticity were assessed through Kolmogorov Smirnov and Breusch-Pagan tests respectively. The input and output variables were discretised (diameter of the fibre: 150 nm <, and >150 nm; inter-fibre separation: < 600 and >600 nm) to inform the classification machine learning (ML) models. The diameter of the fibres was discretised in two ranges, between 40 (minimum value observed) and 150 nm and values above 150 nm, to differentiate between diameters of fibres optimum for smart wearable bioelectronic textiles due to their greatest surface area and able to mimic the collagen fibrils of our ECM to create skin-like non-woven fabrics, and non-optimum fibres for those purposes. Inter-fibre separation was also discretised in two ranges, below 600 nm and above 600 nm with a maximum pore size of 6.21 µm, to be able to evaluate electrospun meshes with ideal pore size to promote breathable (with inter-fibre separation larger than a water vapor molecule 400 nm) and waterproof (with inter-fibre separation smaller than the smallest rain drop 100 µm) textiles as indicated in Figure 1. SPSS version 29.0.1.0 (IBM Inc, US) was used to conduct all statistical analyses.

Seven manufacturing variables (polymer concentration, flow rate, voltage, diameter of the needle, distance between needle and collector, type of collector and revolutions of the mandrel) were used to predict 2 endogenous variables (diameter of the fibres and inter-fibre separation). A total of 20 different ML models were used, including: Multiple Logistic Regression Model (LM), Linear Discriminant Analysis (LDA), Lasso and Elastic-Net Regularized Generalized Linear Models (GLMNET), k-Nearest Neighbours (KNN), Classification and Regression Trees (CART), Random Forest (RF), C5.0 Decision Trees and Rule-Based Models (C50), Bagged Model (BAG), Stochastic Gradient Boosting Models (GBM) and Support Vector Machine (SVM). The "caret()" library implemented in R-4.3.0 and RStudio 2023.03.1 was used to develop all the ML models. Figure 2 shows the methodology followed in this article.

The pre-processes were set up as "BoxCox" for all the models. The training set constituted 70% of the data and test set the 30% following previous studies (Kalantary et al., 2020; Roldán et al., 2023a; Sarma et al., 2022). All models were validated with a 10-fold cross-validation and 3 repeats created with the function "trainControl()", the method "repeatedcv" and metric "accuracy". For the SVM models, sigma values of 0.025, 0.05, 0.1, 0.15 and a sequence between 1 and 10 were used to optimise the hyperparameters; for the rest of the models, the default hyperparameters and settings of the "caret()" library were used. The predictions were calculated with the function "predict()". The accuracy and the confusion matrix were determined with the function "confusionMatrix()" from test data and after cross-validation. The exploratory analysis of the accuracy after cross-validation was performed for each model. Final accuracy of the cross-validation (Accuracy $_{cv}$) was calculated with the average of the accuracy results for each validation stage (Accuracy$_{cv1}$ until Accuracy$_{cv10}$). To determine the performance of the optimised model, the final accuracy, precision, recall and F-score (Accuracy $_{Test}$, Precision $_{Test}$, Recall$_{Test}$ and F-score$_{Test}$) were calculated with the test data, the "confusionMatrix()" function and Equations 1–4.

$$\text{Accuracy} = \frac{tp + tn}{tp + fp + tn + fn} \tag{1}$$

$$\text{Precision(PR)} = \frac{tp}{tp + fp} \tag{2}$$

$$\text{Recall(RC)} = \frac{tp}{tp + fn} \tag{3}$$

$$F\text{-score} = 2 \times \frac{PRxRC}{PR + RC} \tag{4}$$

where tp, tn, fp and fn are the acronyms of true positive, true negative, false positive and false negative. In this study, for the prediction of the diameter of fibres, true positive will be the diameter of fibres below 150 nm correctly identified, true negative will be the diameter of fibres above 150 nm correctly identified, and false positive and false negative will

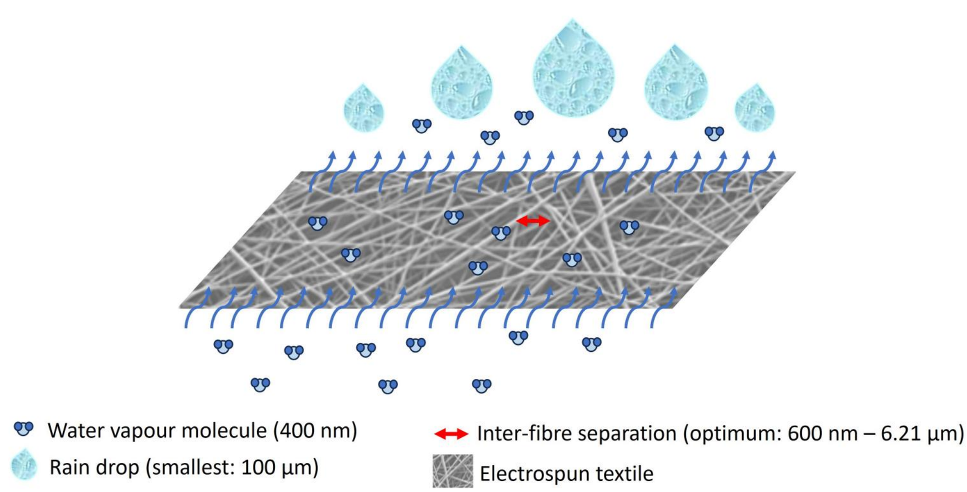

Figure 1. Representation of dimensions of water vapour molecules, rain drops and inter-fibre separations in electrospun textiles.

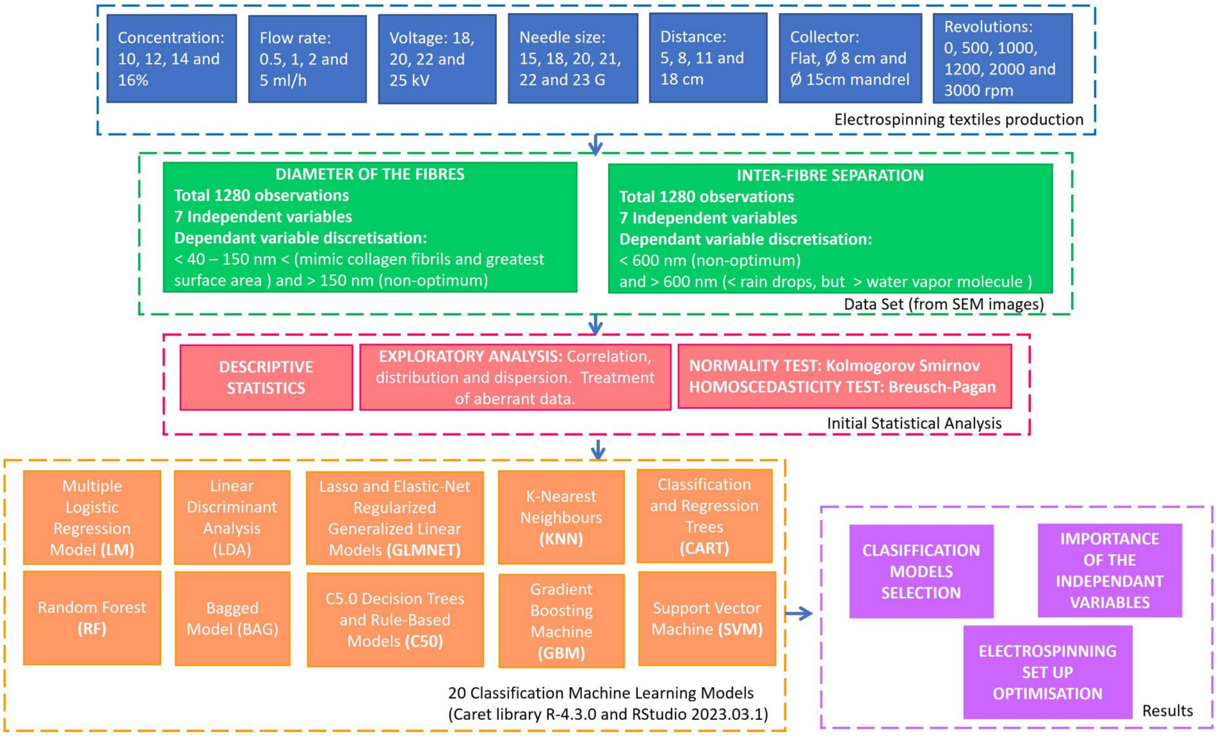

Figure 2. Outline of the followed methodology.

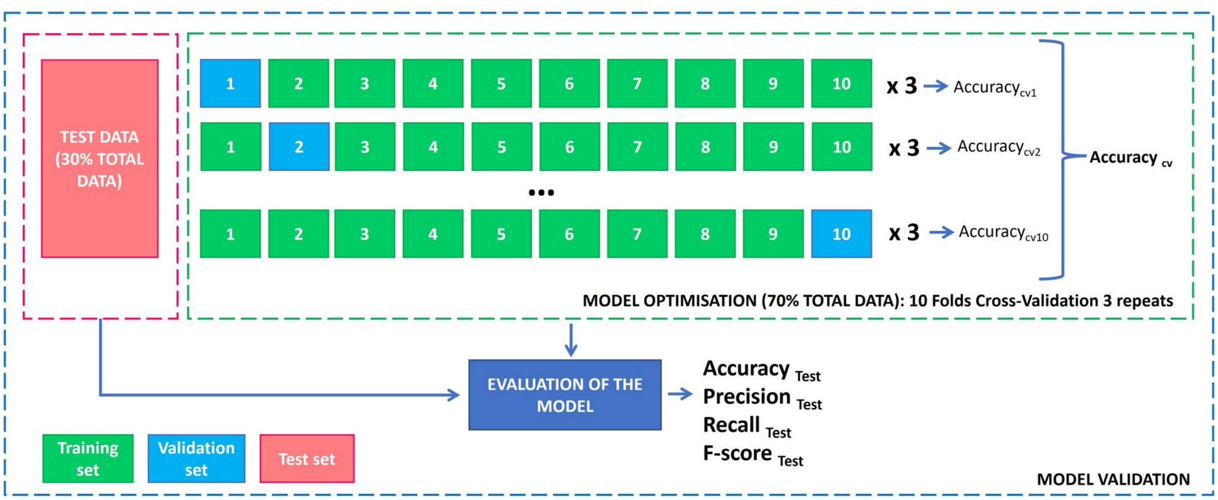

Figure 3. Validation process of the ML models.

be the diameter of fibres below 150 nm and above 150 nm respectively misclassified. For the prediction of the inter-fibre separation, the true positive and true negative will be the inter-fibre separation above 600 nm and below 600 nm respectively correctly identified, and false positive and false negative will be the inter-fibre separation above 600 nm and below 600 nm respectively misclassified. All these values were obtained from the confusion matrix.

Figure 3 shows the validation process implemented for each ML model.

For all the ML models, the importance of the exogenous variables on the endogenous variables was calculated with the function "varImp()" and training and validation data. A percentual average of the importance of the variables was calculated considering all popular ML models (10 models per independent variable) and all the repetitions for the cross-validation (30 repetitions). Therefore, the importance of the predictors on the predicted variable was computed with a total of 300 cases per independent variable (diameter of the fibres and inter-fibre separation).

The electrospinning set up optimisation was studied with the CART and C5.0 models due to their easy interpretability and their results were compared.

3. Results and discussion

3.1. Scaffolds characterisation

PVA is widely regarded as a promising polymer for tissue engineering due to its biodegradability, non-toxicity, and biocompatibility (Park et al., 2010; Roldán et al., 2023a;

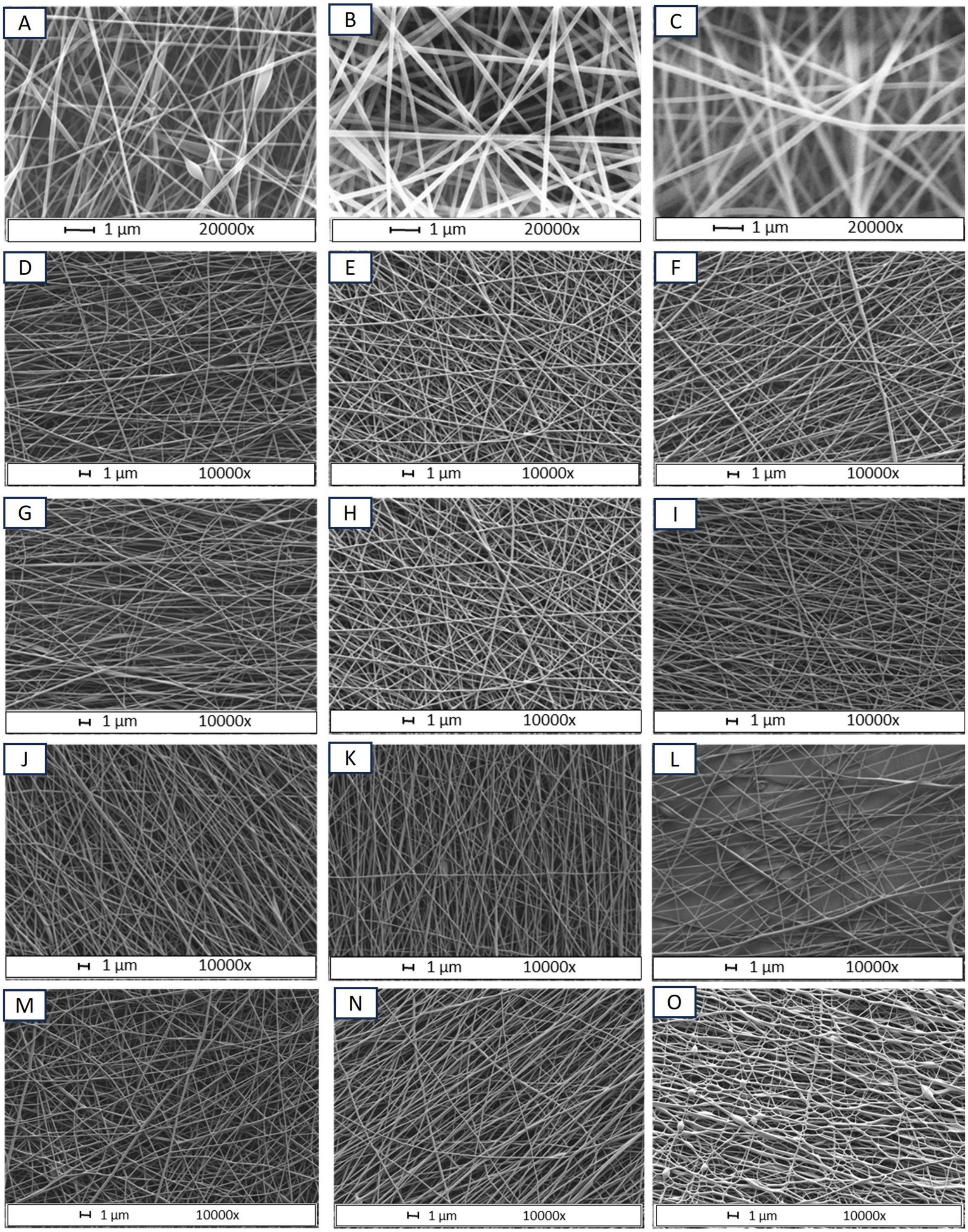

Figure 4. PVA electrospun samples produced with: (A) 10%, (B) 12% and (C) 14% polymer concentration; (D) 50 mm, (E) 80 mm and (F) 110 mm distance between needle and collector; (G) (H) 18G, (I) 20G, (J) 21G, (K) 22G and and (L) 23 G gauge of the needle; (M) 500 rpm, (N) 2000 rpm, and (O) 3000 rpm revolutions of the mandrel..

Supaphol & Chuangchote, 2008), as well as its suitability for textile applications owing to its favourable mechanical properties (Hudson et al., 1993; Jain et al., 2017). Its solubility in water makes PVA particularly advantageous for electrospinning processes, as it is easily electrospun and avoids the need for handling toxic solvent typically required by

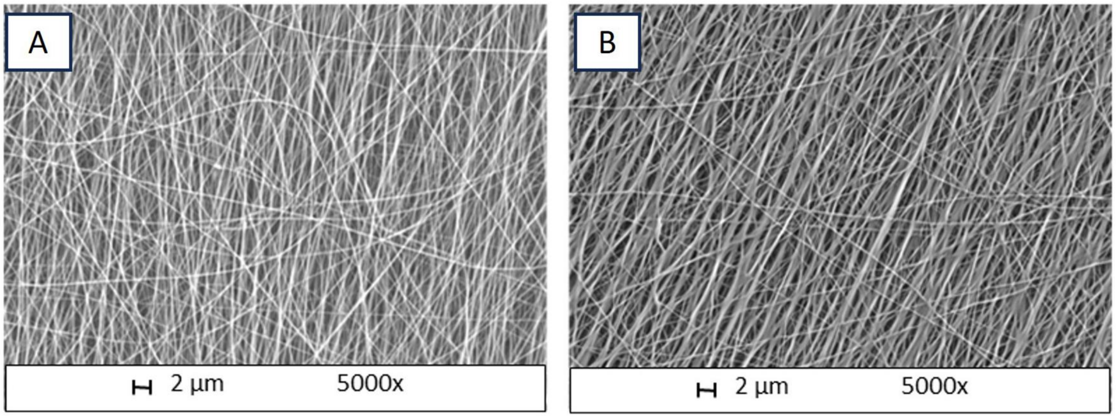

Figure 5. SEM images of PVA scaffolds manufactured with the ideal set up found with C5.0 and CART models (A) non-crosslinked, and (B) crosslinked.

other polymers (Roldán et al., 2024c). Figure 4 shows representative PVA electrospun samples created varying the manufacturing set up.

Despite its benefits, the biodegradable nature of PVA necessitates the use of a crosslinking agent to enhance the stability of scaffolds by reducing their degradation rate. While the machine learning investigations were conducted using non-crosslinked samples, aligning with the study's focus on developing a methodology to optimise manufacturing through machine learning models by controlling fibre diameter and inter-fibre separation in electrospun structures, crosslinking was applied to samples produced using the optimal electrospinning parameters identified by the C5.0 and CART models (section 3.4). A comparative analysis of fibre diameter and inter-fibre separation between crosslinked and non-crosslinked samples (Figure 5) revealed no statistically significant differences (p value $> .05$) between the two groups.

3.2. Statistical analysis

Kolmogorov Smirnov (p value $< .001$) and Breusch-Pagan (p value $< .001$) tests proved that none of the 2 output variables (diameter of the fibres and inter-fibre separation) followed a normal distribution, and they did not meet the homoscedasticity. Therefore, traditional parametrical statistical analysis such as linear regression models or Multivariate Analysis of Variance (MANOVA) models were not suitable for this study, and hence more appropriate data analysis such as classification ML models were applied in this research.

3.3. Machine learning models selection

3.3.1. Diameter of the fibres

Ten machine learning models were used to predict the diameter of the fibres using seven discrete input variables (polymer concentration, flow rate, voltage, diameter of the needle, distance between needle and collector, type of collector and revolutions of the mandrel). The discretisation of the diameter of the fibres enables the models assess if a certain manufacturing set up produces scaffolds with optimum diameter of the fibres and inter-fibre separation for textiles applications.

To select the most accurate models, the accuracy for each model have been calculated after cross-validation. Figure 6 shows the accuracy of each ML model to predict the diameter of the fibres. As 30 repetitions were produced for each ML model, the minimum, first quartile, median, mean, third quartile and maximum were determined per model.

The highest mean of accuracy to predict the diameter of the fibres was obtained with the C5.0 model, with a mean value of 0.869. And the lowest mean of accuracy corresponded to the models LM, LDA and GLMNET with values of accuracy of 0.848 for the three models.

The optimised model was evaluated calculating the accuracy, precision, recall and F-score, with the test data set and the confusion matrix (Table 1), obtaining values of 0.862, 0.854, 0.986 and 0.915 respectively.

3.3.2. Inter-fibre separation

The same 10 ML techniques with the same 7 discrete input variables indicated in the above subsection were used to predict the inter-fibre separation. The accuracy of each model was calculated after cross-validation to determine the best model to predict the inter-fibre separation. Minimum, first quartile, median, mean, third quartile and maximum values of the accuracy obtained per ML model were determined and are shown in Figure 7.

SVM resulted the model with the highest mean of accuracy to predict the inter-fibre separation, with a mean value of 0.875, followed by the KNN and the C5.0 models with mean values of 0.870 and 0.869 respectively. And the lowest mean of accuracy was obtained with the models BAG, LM, LDA and GLMNET with values of accuracy of 0.859 for the four models.

The accuracy, precision, recall and F-score obtained from the optimised model and assessed with test data and confusion matrix (Table 2) were 0.875, 0.877, 0.969 and 0.921 respectively for the prediction of the inter-fibre separation.

It is worth noting that recall, or also called sensitivity, is the proportion of true positive predictions to the total number of actual positive instances (both true positives and false negatives). Whereas the precision, or confidence, is called to

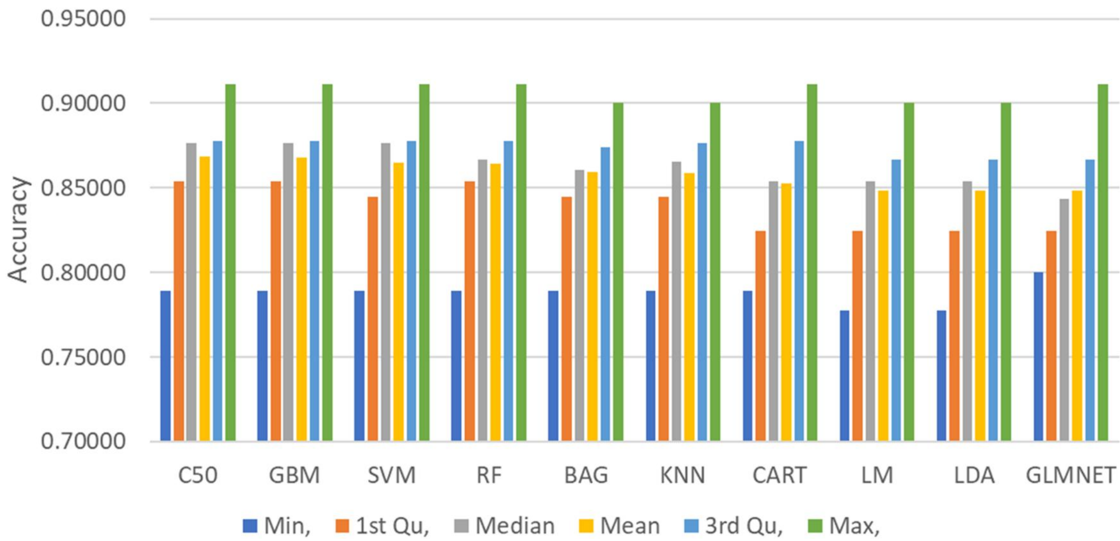

Figure 6. Accuracy of ML models to predict the diameter of the fibres after cross-validation.

Table 1. Confusion matrix from the prediction of the diameter of the fibres.

	≤150 nm	>150 nm
Confusion matrix diameters	Prediction	
Observation	1	2
1	286 (tp)	49 (fp)
2	4 (fn)	44 (tn)

the ratio of true positive predictions to the total number of positive predictions (both true positives and false positives) (Powers, 2020). F-score exhibits few limitations, the first limitation is that it is designed for binary classification (such as the present article), other limitation is that assumes equal importance of precision and recall (which in our case is true, therefore it will be not a limitation), and requires a threshold to assign observations to classes and this threshold can influence the performance of the model (Flach, 2019). Accuracy is the most intuitive performance metric, representing the proportion of accurately predicted observations out of the total observations. It is ideal when the classes are evenly distributed, and the impact of false positives and false negatives is roughly equal, like in the case of the present study (Yin et al., 2019).

In terms of the ML models, linear models (LDA and GLMNET) and the logistic model (LM) resulted the less accurate models in both cases. This performance was also observed in other studies, for instance, Ieracitano et al. reported values of accuracy of 0.648 for LDA, 0.667 for SVM and 0.925 for ANN to predict defects on the PVA electrospun mats (Ieracitano et al., 2021). GLMNET exhibited poor performance in the prediction of mechanical properties of PVA electrospun compared to KNN, CART, RF or SVM models (Roldán et al., 2024b). Sarma et al. reported values of R^2 of 0.8 for GBM, 0.75 for RF, 0.48 for KNN and 0.32 for linear models, for regression models predicting the diameter of the fibres of electrospun polyvinylidene fluoride mats (Sarma et al., 2022). Although R^2 and accuracy values are not comparable itself; however, how well performed each model can be contrasted, the R^2 values obtained by Sarma et al. were aligned to the results of accuracy observed for our models predicting the diameter of the PVA fibres.

3.4. Electrospinning set up optimisation

Decision trees were demonstrated to be a useful tool to determine the optimum set up to produce biomimetic electropun vascular grafts (Roldán et al., 2023a). In this article, we compare the predictability and interpretability of two tree-based algorithms, CART and C5.0, and we discuss the suitability of being used for textiles applications.

3.4.1. Diameter of the fibres

As it was explained in the materials and methods section, the diameter of the fibres was discretised in two ranges, values lower or equal to 150 nm and values above 150 nm. The number "1" was assigned for the first range (<150 nm) and the number "2" for the second range (>150 nm), being "1" the diameters of fibres suitable for smart wearable bioelectronic textiles, due to their greatest surface area and therefore interaction sites, and able to mimic the collagen fibrils of our ECM to create skin-like non-woven fabrics and being "2" the non-optimum diameter of fibres for those purposes.

Figure 8 shows the CART model to predict the diameter of the fibres. It can be observed that the majority of the fibres (75.5%) registered a diameter between 40 (minimum value observed) and 150 nm, ideal for textiles and tissue engineering applications. The path to follow to obtain those values was through two different routes, or with concentration of polymers lower than 11% or with concentration of PVA of 12%, a rotational collector of 15 cm in diameter, a gauge of the needle lower than 22 G and a voltage lower than 21 kV.

C5.0 algorithm, based on decision trees, classification rules and boosting, also provided a tree-like diagram that allowed determining the path to follow to obtain meshes with a desired morphology. As with the CART models, C5.0 algorithm provides results easy to interpret and understand.

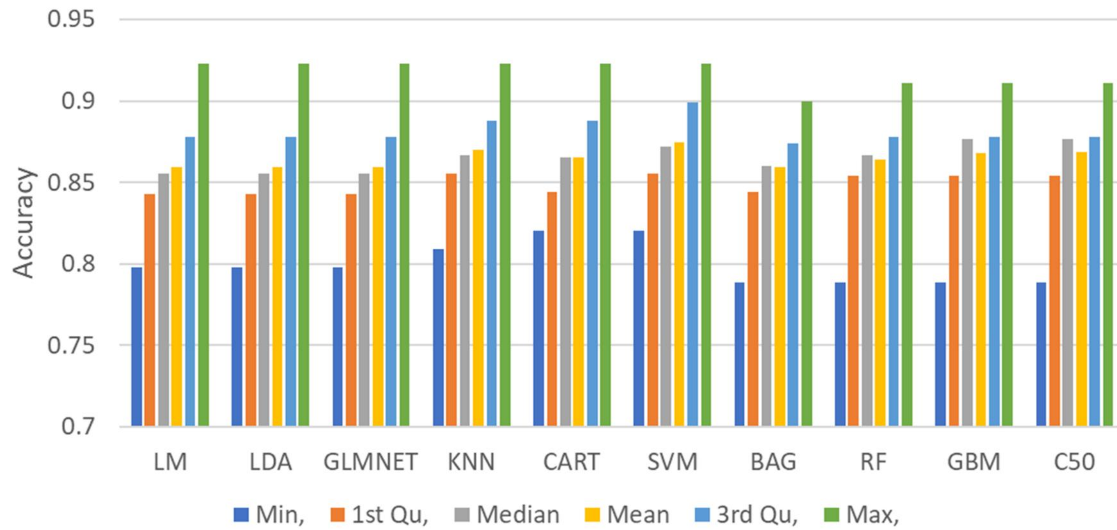

Figure 7. Accuracy of ML models to predict the inter-fibre separation after cross-validation.

Table 2. Confusion matrix from the prediction of the inter-fibre separation.

	> 600 nm	≤ 600 nm
Confusion matrix inter-fibre separation	Prediction	
Observation	1	2
1	279 (tp)	39 (fp)
2	9 (fn)	57 (tn)

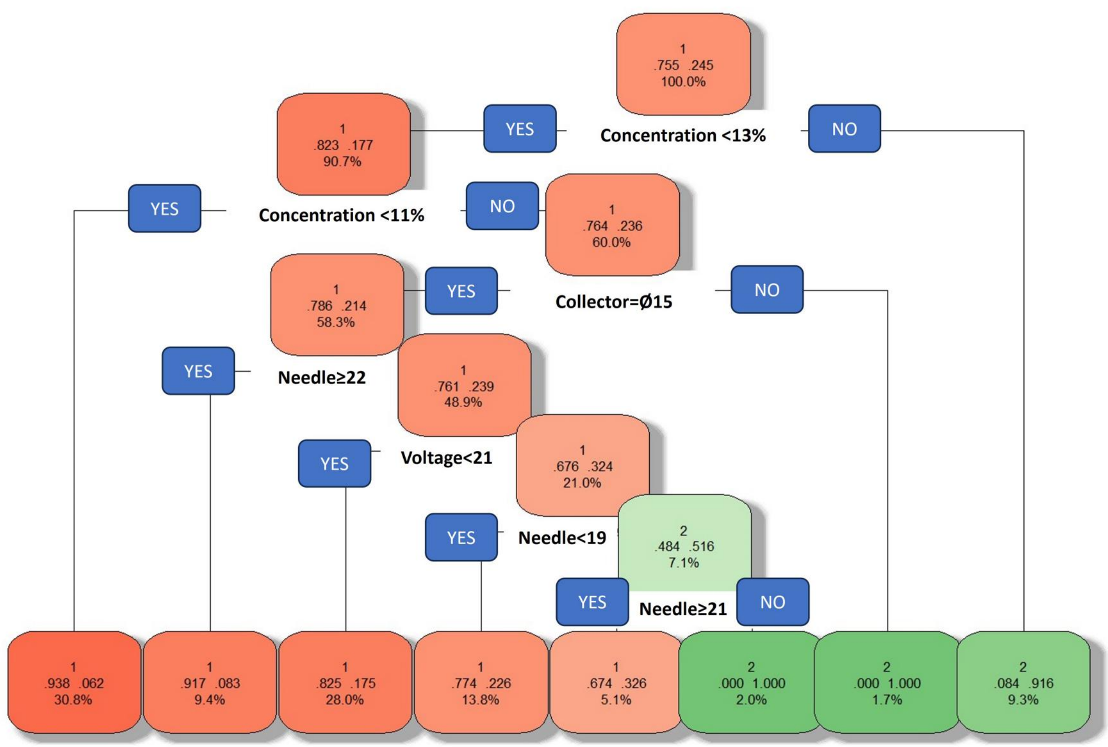

Figure 8. CART model to predict diameter of the fibres and determine the optimum parameters.

This model supports non-parametric and parametric conditions and tree pruning is done by default avoiding overfitting caused when a model fits training data very closely but has poor performance with test data. Figure 9 shows the tree-like diagram to predict the diameter of the fibres following the different routes.

Figure 9 shows that node with higher probability of obtaining diameter of the fibres between 50 and 150 nm,

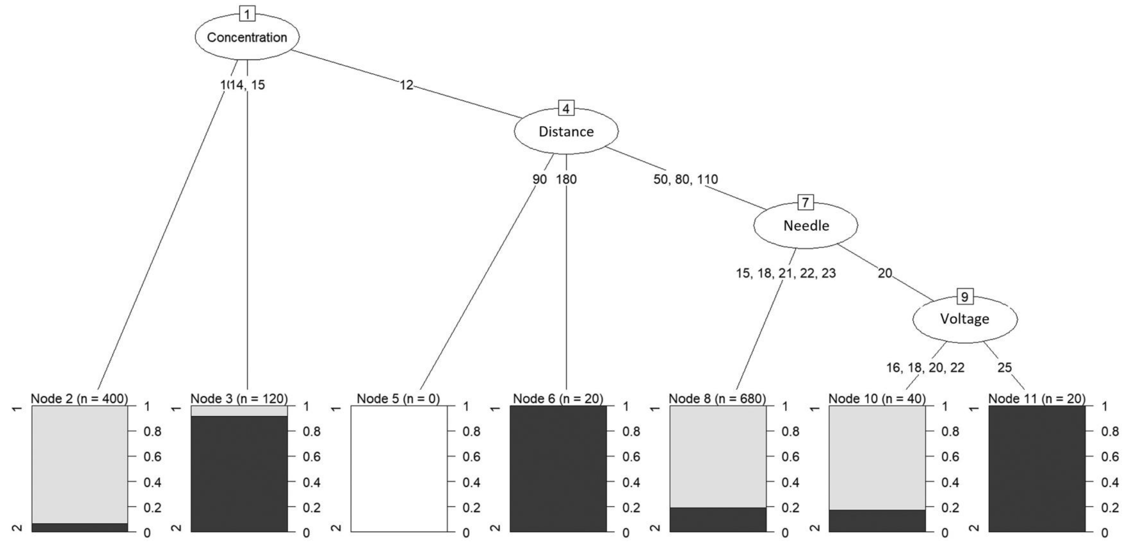

Figure 9. C5.0 model to predict diameter of the fibres and determine the optimum parameters.

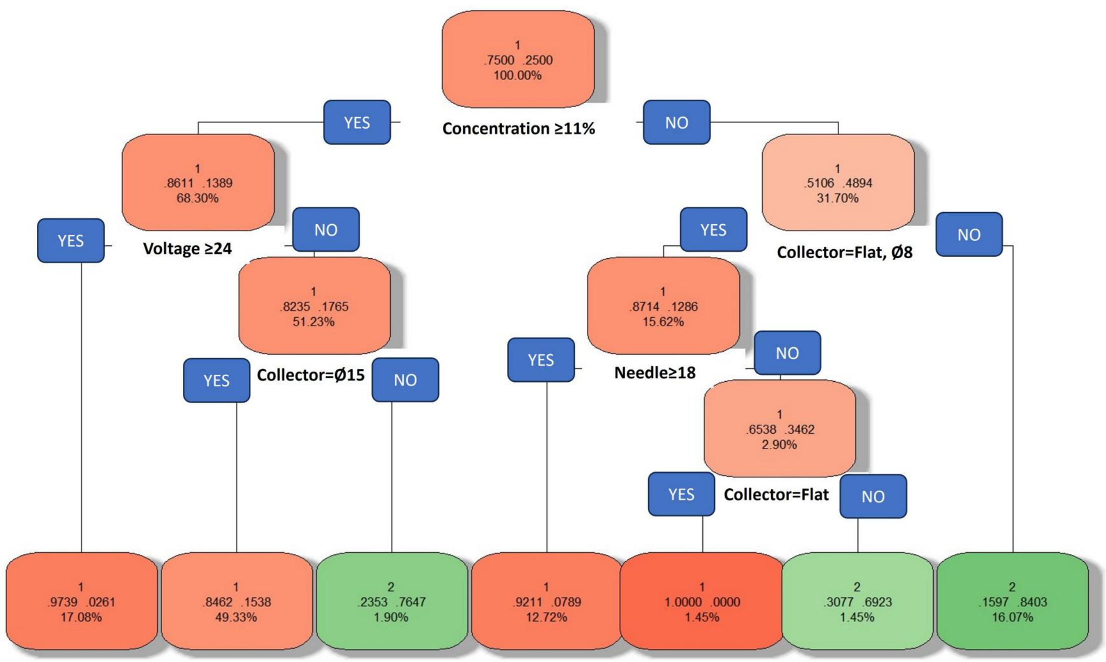

Figure 10. CART model to predict inter-fibre and determine the optimum parameters.

comparable to the collagen fibrils of soft tissue and ideal for breathable and waterproof textiles and smart wearable bioelectronic textiles, was the node 8 that can be obtained with polymer's concentration of 12%, gauge of the needle different to 20 G and voltage lower than 25 kV.

Following the tree-like diagrams provided for both models, CART and C5.0, it can be concluded that PVA concentration of 12%, a rotational collector of 15 cm in diameter, a gauge of the needle lower than 22 G and different to 20 G, and a voltage lower than 21 kV, resulted in electrospun meshes with diameter of fibres between 40 and 150 nm which are ideal for smart wearable bioelectronic textiles due to their greatest surface area and able to mimic the collagen fibrils of our ECM to create skin-like non-woven fabrics.

In terms of the accuracy of the models, C5.0 provided a higher accuracy than the CART model with 0.869 and 0.852 respectively.

3.4.2. Inter-fibre separation

The inter-fibre separation variable was discretised in two ranges, below 600 nm and above 600 nm with a maximum pore size of 6.21 µm. In this case, "1" was assigned to inter-fibre separation above 600 nm, which is the ideal pore size to promote both breathable and waterproof textiles, and "2" was assigned to non-optimum inter-fibre separation for those purposes.

Figure 10 shows the CART model to predict the inter-fibre separation. It can be observed that the inter-fibre

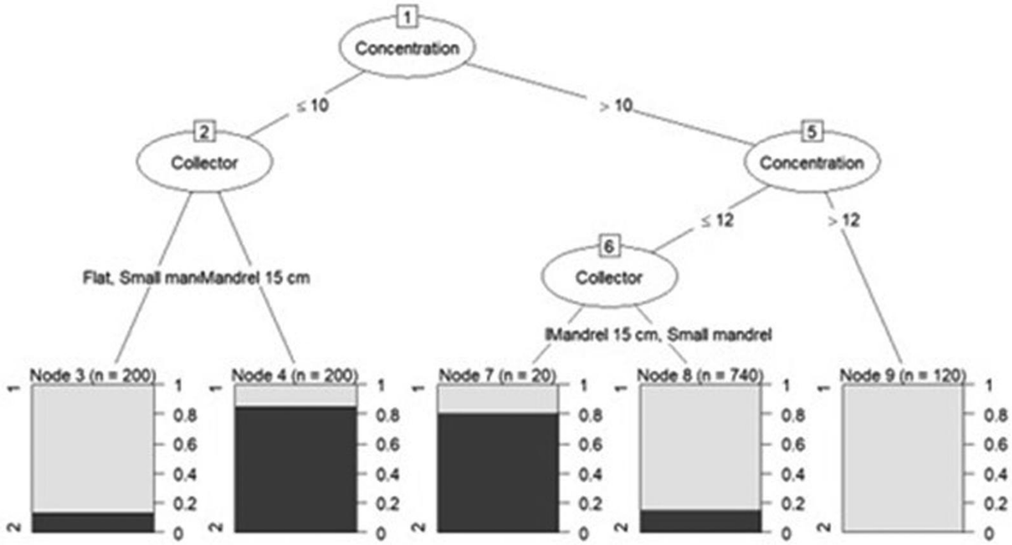

Figure 11. C5.0 model to predict inter-fibre separation and determine the optimum parameters.

separation was above 600 nm with a maximum inter-fibre separation of 6.21 µm in most of the cases (75%). These values are ideal for breathable and waterproof textiles since they are larger than a water vapor molecule 400 nm, therefore the fabric can perspire, but smaller than the smallest rain drop (100 µm), therefore the fabric can be impermeable. As it can be observed in Figure 10, the highest probability of obtaining those values (with a 49.33%) was with a concentration of polymer higher or equal than 11%, a voltage lower than 24 kV and a rotational collector of 15 cm in diameter.

Figure 11 shows the tree-like diagram to predict the inter-fibre separation following the different routes. In this graph, we can observe that node with higher probability of obtaining inter-fibre separation higher than 600 nm was the node 8 that can be obtained with polymer's concentration of 11 or 12%, and a mandrel of 8 cm.

Following the tree-like diagrams provided for both models, CART and C5.0, it can be concluded that PVA concentration of 11 or 12%, a voltage lower than 24 kV and a mandrel as a rotational collector resulted into electrospun meshes with inter-fibre separation between 600 nm and 6.21 µm ideal for which are ideal for breathable and waterproof textiles.

In terms of the accuracy of the models, as observed with the prediction of the diameter of the fibres, C5.0 provided a higher accuracy than the CART model with 0.861 and 0.859 respectively.

Summarising the results obtained for the C5.0 and the CART models, we can conclude that a set-up of the electrospinner with a rotational collector of diameter of 15 cm, a voltage between 16 and 21 kV, gauge of the needle lower than 22 G but different to 20 G and concentration of polymer equal to 12 or 11% would result into electrospun meshes with diameter of fibres between 40 and 159 nm, and inter-fibre separations between 600 nm and 6 µm. These results considerably narrow the manufacturing parameters to obtain the optimum fibres for the desired purpose.

Table 3. Importance of the independent variable on the dependent variable.

Independent variable	Diameter	Inter-fibre separation
Concentration	42.056	33.088
Collector	15.291	31.908
Voltage	14.759	9.066
Distance	11.724	8.669
Revolution	8.883	8.439
Needle	5.852	8.126
Flow rate	1.435	0.700

3.5. Importance of the independent variables

The importance of each factor in shaping the predictions of the diameter of the fibres and the inter-fibre separation, were found from the average of the 20 ML models.

Table 3 shows the importance of each independent variable in the prediction of the dependent variables, diameter of the fibres and inter-fibre separation.

The concentration factor has the greatest participation in the prediction of both, diameter of the fibres and inter-fibre separation, with a 42 and 33% of the total importance respectively. This fact is in agreement to the decision trees presented above (CART and C5.0), where the concentration is the root node for both predictions, and previous research (Kalantary et al., 2020; Roldán et al., 2023a). The shape of the collector contributed to the predictions a 15 and a 32% respectively. The flow rate exhibited the lowest influence on the prediction of these endogenous variables with 1.5 and 1% respectively.

Conclusions

In this research, 20 popular classification ML models were used to predict the diameter of the fibres and the inter-fibre separation of PVA electrospun scaffolds. The production settings necessary for customised electrospun structures can be optimised to create breathable and waterproof fabrics, skin-like non-woven fabrics, and smart wearable bioelectronic textiles by employing C5.0 machine learning techniques, being the ML model that combined both a good interpretability and high accuracy. It is accepted that the ML approaches explored within this article will have some errors in their prediction, but utilising this virtual prototyping tool will reduce the design space. Thus, enabling faster development and fewer physical prototypes leading to lower development costs for electrospun textiles. Future research will focus on three main goals: first, using this design approach to forecast the mechanical properties of electrospun structures based on specific machine configurations to improve their manufacturing process, second, analysing the fibre orientation to produce isotropic fabrics or anisotropic textiles depending on the textile purpose, and third, exploring various polymers to enhance the functionality of these fabrics.

Acknowledgments

Some experimental work included in this study was partially performed in the Manchester Institute of Biotechnology (University of Manchester). The authors would like to acknowledge Professor Paulo Jorge Da Silva Bartolo for his support allowing the use of the University of Manchester facilities. The authors would also like to thank the technical staff of the Manchester Metropolitan University, Hayley Andrews for training ER on the use of the SEM.

Authors' Contributions

Conceptualisation, E.R.; methodology, E.R.; experimental design and characterisation design, E.R., K.A, G.C, N.R.; experimental performance, E.R; validation, E.R.; formal analysis, E.R.; investigation, E.R.; writing—original draft preparation, E.R.; writing—review and editing, E.R., K.A, G.C, N.R. All authors have read and agreed to the published version of the manuscript.

Disclosure statement

The authors declare no conflict of interest.

Data availability

The data supporting this article will be made available on request to the correspondence author Elisa.Roldan-Ciudad@mmu.ac.uk.

ORCID

Elisa Roldán http://orcid.org/0000-0002-7793-7542

References

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods*, *15*(4), 233–234. https://doi.org/10.1038/nmeth.4642

Flach, P. (2019). Performance evaluation in machine learning: The good, the bad, the ugly, and the way forward. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 9808–9814). https://doi.org/10.1609/aaai.v33i01.33019808

Haider, A., Haider, S., & Kang, I.-K. (2018). A comprehensive review summarizing the effect of electrospinning parameters and potential applications of nanofibers in biomedical and biotechnology. *Arabian Journal of Chemistry*, *11*(8), 1165–1188. https://doi.org/10.1016/j.arabjc.2015.11.015

Hudson, P. B., Clapp, A. C., & Joseph's, K. D. (1993). *Introductory textile science*. Harcourt Brace Jovanovich College Publishers.

Ieracitano, C., Paviglianiti, A., Campolo, M., Hussain, A., Pasero, E., & Morabito, F. C. (2021). A novel automatic classification system based on hybrid unsupervised and supervised machine learning for electrospun nanofibers. *IEEE/CAA Journal of Automatica Sinica*, *8*(1), 64–76. https://doi.org/10.1109/JAS.2020.1003387

Jain, N., Singh, V. K., & Chauhan, S. (2017). A review on mechanical and water absorption properties of polyvinyl alcohol based composites/films. *Journal of the Mechanical Behavior of Materials*, *26*(5–6), 213–222. https://doi.org/10.1515/jmbm-2017-0027

Jin, C., De-Lin, L., & Fen-Xiang, M. (2009). An improved ID3 decision tree algorithm. In *2009 4th International Conference on Computer Science & Education* (pp. 127–130). https://doi.org/10.1109/ICCSE.2009.5228509

Kalantary, S., Jahani, A., & Jahani, R. (2020). MLR and ANN approaches for prediction of synthetic/natural nanofibers diameter in the environmental and medical applications. *Scientific Reports*, *10*(1), 8117. https://doi.org/10.1038/s41598-020-65121-x

Liu, L., Xu, W., Ding, Y., Agarwal, S., Greiner, A., & Duan, G. (2020). A review of smart electrospun fibers toward textiles. *Composites Communications*, *22*, 100506. https://doi.org/10.1016/j.coco.2020.100506

López-Flores, F. J., Ornelas-Guillén, J. A., Pérez-Nava, A., González-Campos, J. B., & Ponce-Ortega, J. M. (2024). Data-driven machine learning approach for modeling the production and predicting the characteristics of aligned electrospun nanofibers. *Industrial & Engineering Chemistry Research*, *63*(22), 9904–9913. https://doi.org/10.1021/acs.iecr.4c00075

Maleki, S., Shamloo, A., & Kalantarnia, F. (2022). Tubular TPU/SF nanofibers covered with chitosan-based hydrogels as small-diameter vascular grafts with enhanced mechanical properties. *Scientific Reports*, *12*(1), 6179. https://doi.org/10.1038/s41598-022-10264-2

Mukhopadhyay, A., Vinay Kumar, M., 2008. A review on designing the waterproof breathable fabrics part I: Fundamental principles and designing aspects of breathable fabrics. *Journal of Industrial Textiles*, 33*7*, 225–262. https://doi.org/10.1177/1528083707082164

Park, J.-C., Ito, T., Kim, K.-O., Kim, K.-W., Kim, B.-S., Khil, M.-S., Kim, H.-Y., & Kim, I.-S. (2010). Electrospun poly(vinyl alcohol) nanofibers: Effects of degree of hydrolysis and enhanced water stability. *Polymer Journal*, *42*(3), 273–276. https://doi.org/10.1038/pj.2009.340

Powers, D. M. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. https://arxiv.org/abs/2010.16061

Quinlan, J. R. (2014). *C4.5: Programs for machine learning*. Elsevier.

Roldán, E., Reeves, N. D., Cooper, G., & Andrews, K. (2023a). Towards the ideal vascular implant: Use of machine learning and statistical approaches to optimise manufacturing parameters. *Frontiers in Physics*, *11*. https://doi.org/10.3389/fphy.2023.1112218

Roldán, E., Reeves, N. D., Cooper, G., & Andrews, K. (2023b). Can we achieve biomimetic electrospun scaffolds with gelatin alone? *Frontiers in Bioengineering and Biotechnology*, *11*, 1160760. https://doi.org/10.3389/fbioe.2023.1160760

Roldán, E., Reeves, N., Cooper, G., & Andrews, K. (2023c). Optimization of manufacturing parameters through machine learning techniques to create biomimetic vascular implants. *Tissue Engineering Part A*, *29*. https://doi.org/10.1089/ten.tea.2023.29043.abstracts

Roldán, E., Reeves, N. D., Cooper, G., & Andrews, K. (2024a). Machine learning to predict morphology, topography and mechanical properties of sustainable gelatin-based electrospun scaffolds. *Scientific Reports*, *14*(1), 21017. https://doi.org/10.1038/s41598-024-71824-2

Roldán, E., Reeves, N. D., Cooper, G., & Andrews, K. (2024b). Machine learning to mechanically assess 2D and 3D biomimetic electrospun scaffolds for tissue engineering applications: Between the predictability and the interpretability. *Journal of the Mechanical Behavior of Biomedical Materials*, *157*, 106630. https://doi.org/10.1016/j.jmbbm.2024.106630

Roldán, E., Reeves, N. D., Cooper, G., & Andrews, K. (2024c). 2D and 3D PVA electrospun scaffold evaluation for ligament implant replacement: A mechanical testing, modelling and experimental biomechanics approach. *Materialia*, *33*, 102042. https://doi.org/10.1016/j.mtla.2024.102042

Sarma, S., Verma, A. K., Phadkule, S. S., & Saharia, M. (2022). Towards an interpretable machine learning model for electrospun polyvinylidene fluoride (PVDF) fiber properties. *Computational Materials Science*, *213*, 111661. https://doi.org/10.1016/j.commatsci.2022.111661

Supaphol, P., & Chuangchote, S. (2008). On the electrospinning of poly(vinyl alcohol) nanofiber mats: A revisit. *Journal of Applied Polymer Science*, *108*(2), 969–978. https://doi.org/10.1002/app.27664

Tay, J. K., Narasimhan, B., & Hastie, T. (2023). Elastic net regularization paths for all generalized linear models. *Journal of Statistical Software*, *106*(1), 1. https://doi.org/10.18637/jss.v106.i01

Wang, B., Cai, J., Liu, C., Yang, J., & Ding, X. (2020). Harnessing a novel machine-learning-assisted evolutionary algorithm to co-optimize three characteristics of an electrospun oil sorbent. *ACS Applied Materials & Interfaces*, *12*(38), 42842–42849. https://doi.org/10.1021/acsami.0c11667

Wang, P., Lv, H., Cao, X., Liu, Y., & Yu, D.-G. (2023). Recent progress of the preparation and application of electrospun porous nanofibers. *Polymers*, *15*(4), 921. https://doi.org/10.3390/polym15040921

Yin, M., Wortman Vaughan, J., & Wallach, H. (2019). Understanding the effect of accuracy on trust in machine learning models. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI ‘19. (pp. 1–12). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300509