Article

Use of Data Mining Techniques for the Prediction of Surface Roughness of Printed Parts in Polylactic Acid (PLA) by Fused Deposition Modeling (FDM): A Practical Application in Frame Glasses Manufacturing

Esther Molero, Juan Jesús Fernández, Oscar Rodríguez-Alabanda [ID] **, Guillermo Guerrero-Vaca *** [ID] **and Pablo E. Romero** [ID]

Department of Mechanical Engineering, University of Cordoba, Medina Azahara Avenue, 14071 Cordoba, Spain; esther.molero@uco.es (E.M.); p82feruj@uco.es (J.J.F.); orodriguez@uco.es (O.R.-A.); p62rocap@uco.es (P.E.R.)

* Correspondence: guillermo.guerrero@uco.es; Tel.: +34-957-212-230

check for
updates

Abstract: In the present work, ten data mining algorithms have been used to generate models capable of predicting the surface roughness of parts printed on polylactic acid (PLA) by using fused deposition modeling (FDM). The models have been trained using experimental data measured on 27 horizontal (XY) and 27 vertical (XZ) specimens, printed using different values for the parameters studied (layer height, extrusion temperature, print speed, print acceleration and flow). The models generated by multilayer perceptron (MLP) and logistic model trees (LMT) have obtained the best results in a cross-validation. Although it does not obtain such optimal results, the J48 algorithm (C4.5) allows the generation of models in the form of a decision tree. These trees permit to determine which print parameters have an influence on the surface roughness. For XY specimens, the surface roughness measured in the direction parallel to the extrusion path ($R_{a,0,XY}$) depends on the flow, the print temperature and the layer height; in the direction perpendicular to the extrusion path, the surface roughness ($R_{a,90,XY}$) depends only on the flow. For XZ specimens, the surface roughness measured in the direction parallel to the extrusion path ($R_{a,0,XZ}$) depends only on the print speed; in the direction perpendicular to the extrusion path ($R_{a,90,XZ}$), it depends on the layer height and the extrusion temperature. According to the study carried out, the most suitable set up provides values of $R_{a,0,XY}$, $R_{a,90,XY}$, $R_{a,0,XZ}$ and $R_{a,90,XZ}$ equal to 0.46, 1.18, 0.45 and 11.54, respectively. A practical application of this work is the manufacture of PLA frame glasses using FDM.

Keywords: fused deposition modeling; FDM; FFF; data mining; machine learning; PLA; surface roughness; WEKA; decision trees; C4.5; neural networks; ANN; frame glasses

1. Introduction

Currently, the industry is suffering a profound revolution. Different technological tools are being used intensively in factories [1]: augmented reality, virtual twins, data mining, additive manufacturing, among others. This fact is known as industry 4.0.

Additive manufacturing (AM) was initially used to manufacture prototypes in the product development stage. Today, it is also used to manufacture: customized objects (rapid prototyping), tools for other processes (rapid tooling), small batches of fully functional parts (rapid manufacturing) [2].

There are different technologies of AM. However, fused deposition modeling (FDM) is the most extended technology [3–5]. The low cost of the equipment and the diversity of filaments in the market contribute to this. Although there are many materials available, polylactic acid (PLA) is still one of the most widely used in FDM [6,7]. There are several reasons for this [8]: it is biodegradable, easily

printed, does not give off any vapors during printing and there are PLA formulas with mechanical properties similar to acrylonitrile butadiene styrene (ABS) [9].

For some applications, 3D printed parts must achieve minimum mechanical properties [8,10–13]. In other cases, apart from a high tensile strength value, the parts need to have a proper aesthetic quality [14–16]. Some authors have evaluated which factors allow achieving the best rankings in all hedonic, tactile and visual assessments of FDM 3D printed parts [17]; in the tactile and visual evaluations, the authors conclude that the pieces with the highest scores are those with the lowest values of surface roughness. On the other hand, the surface finish is crucial when printed parts are in contact with human skin, as is the case with frame glasses [18].

One of the motivations for this work is to determine which printing factors are the most influential in the surface finish on PLA parts printed using FDM. In the literature, you can find previous works in which the influence of different printing parameters on the surface roughness of 3D printed parts has been studied (Table 1). However, there are printing parameters that have not yet been analyzed, such as print acceleration or flow. This is one of the main contributions of this work.

Traditionally, the experimental study of the quality of manufactured parts has been carried out with statistical tools, such as the Taguchi method and analysis of variance (ANOVA) [19]. However, the large amount of data currently generated by industry 4.0 sensors and the need to use algorithms capable of modeling non-linear problems has made it necessary to use data mining (also known as machine learning, ML) techniques [20–22]. The second motivation for the work is to use data mining (DM) techniques to predict the surface finish of 3D printed parts using FDM.

Razvi et al. [23] have reviewed the existing literature on the use of DM in additive manufacturing; the papers reviewed have been grouped around four different topics: design, process optimization, monitoring and control, inspection and testing. Amand et al. [24] have used DM techniques to predict possible defects during the configuration step in FDM 3D printing. Wu et al. [25] have used the random forest algorithm to generate a model to predict the surface roughness of a 3D printed part from data collected by different sensors placed in the FDM printer. Sohnius et al. [26] have employed DM techniques to predict the quality of printed parts using FDM from data obtained via the machine vision method. Mahapatra and Sood [27] proposed the use of artificial neural networks (ANN) to determine the relationship between five input FDM parameters such as layer thickness, orientation, raster angle, raster width, and air gap with surface roughness in the top, bottom, and side surface of the acrylic nitrile butadiene styrene (ABS) built part. Boschetto et al. [28] proposed a feed-forward neural network to fit experimental data and to determine surface roughness parameter models reliable over the entire part surface. Vahabli and Rahmati [29] have established a robust model using empirical data based on optimized ANN to estimate the surface roughness distribution in fused deposition modeling ABS parts; this work includes four medical case studies.

The aim of this work is to generate and validate models via data mining techniques that allow predicting the surface finish of PLA printed parts according to the selected values for the following printing parameters: layer height, extrusion temperature, print speed, print acceleration and flow. For this purpose, 27 horizontal and 27 vertical specimens have been manufactured, according to a fractional experiment design. The surface roughness of these specimens was measured. The results were used to generate and test different models via data mining algorithms (Bayes Net, naïve-Bayes, multilayer perceptron, simple logistics, sequential minimal optimization, IBk, Kstar, J48, logistic model tree and random forest). In addition, using the J48 algorithm, decision trees were generated to determine which printing parameters significantly influence the surface roughness of the parts.

As a practical application of the work, the manufacture of a frame glasses is proposed. There are standard frames on the market, which are adapted to the characteristics of an average person. However, there are people with such a facial morphology that they cannot find glasses in the optical shops. By means of 3D printing, it is possible to manufacture customized frames [30], in biodegradable materials such as PLA [18]. One of the most important specifications of a frame glasses is the surface finish, for aesthetic reasons and to reduce friction with the skin of the user.

Table 1. Previous works that have studied the surface roughness of polylactic acid (PLA) printed parts using fused deposition modeling (FDM).

Author	Layer Height	Tempe-Rature	Print Orienta-tion	Print Speed	Filling Density	Nozzle Diame-ter	Wall Thick-ness	Print Path
García-Plaza et al. [31]	√	-	√	√	-	-	-	-
Ramli et al. [32]	√	-	-	-	√	-	-	-
Alsoufi and Elsayed [33]	√	-	-	-	-	√	-	-
Kovan et al. [34]	√	√	-	-	-	-	-	-
Perez et al. [19]	√	√	-	√	-	-	√	√

2. Materials and Methods

2.1. Design of Experiments and Printing of Specimens

In this work, 54 test pieces with dimensions 25.0 mm × 25.0 mm × 2.4 mm were printed following a fractional experiment design, with five factors and three levels (Table 2). The factors studied were: layer height (LH), extrusion temperature (T), print speed (PS), print acceleration (PA) and flow (F). Table 3 shows the parameters set in each test. A total of 27 specimens were printed in the XY orientation, and others 27 in the XZ orientation (Figure 1).

Table 2. Factors and levels used in the design of experiments (DOE).

Factors	Level 1	Level 2	Level 3
Layer height, LH (mm)	0.16	0.20	0.24
Temperature, T (°C)	200	210	220
Print speed, PS (mm/s)	40	50	60
Print acceleration, PA (mm/s^2)	500	1000	1500
Flow, F (%)	90	100	110

Table 3. Design of experiment L27 used in the present work.

| Layer Height, LH (mm)	Temperature, T (°C)	Print Speed, PS (mm/s)	Print Acceleration, PA (mm/s^2)	Flow, F (%)
0.16	200	40	500	900
0.16	200	40	500	100
0.16	200	40	500	110
0.16	210	50	1000	90
0.16	210	50	1000	100
0.16	210	50	1000	110
0.16	220	60	1500	90
0.16	220	60	1500	100
0.16	220	60	1500	110
0.20	200	50	1500	90
0.20	200	50	1500	100
0.20	200	50	1500	110
0.20	210	60	500	90
0.20	210	60	500	100
0.20	210	60	500	110
0.20	220	40	1000	90
0.20	220	40	1000	100
0.20	220	40	1000	110
0.24	200	60	1000	90
0.24	200	60	1000	100
0.24	200	60	1000	110
0.24	210	40	1500	90
0.24	210	40	1500	100
0.24	210	40	1500	110
0.24	220	50	500	90
0.24	220	50	500	100
0.24	220	50	500	110
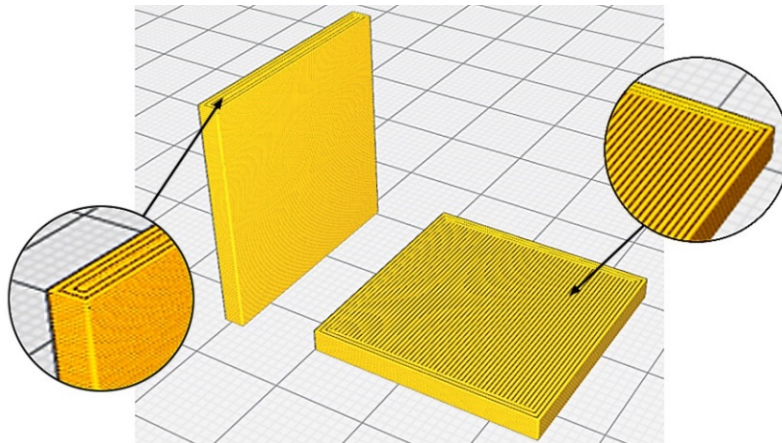
Figure 1. Print orientation: vertical (XZ) (**left**), and horizontal (XY) (**right**).

The specimens were designed using SolidWorks (Dassault Systemes, Vélizy-Villacoublay, France). The selection of print parameter values and the generation of the numerical code (CN) was performed using Ultimaker CURA software (version 4.0.0, Ultimaker, Utrecht, The Netherlands).

The specimens were produced on an Ender 3 printer (Creality 3D, Shenzhen, China), with a 220 × 220 × 250 mm^3 workspace and a hot bed (50 °C). An extrusion nozzle with a diameter of 0.4 mm was used in the tests.

2.2. Surface Roughness Measurement

Surface roughness (R_a) of the printed specimens was measured using a Mitutoyo SJ-201 profilometer (Mitutoyo, Kawasaki, Japan). R_a was measured five times in the direction parallel to the extrusion path ($R_{a,0}$) and five times in the direction perpendicular to the extrusion path ($R_{a,90}$) (Figure 2). The representative value in each direction for each specimen was calculated as the arithmetic mean of these five measurements.
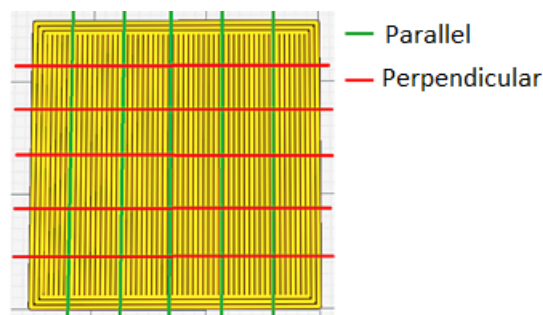

Figure 2. Directions of measurement taken in the roughness test.

2.3. Data Mining Algorithms

Data mining is a discipline that allows us to analyze data sets and extract knowledge from them. It is a key tool in the industry to process the data that is generated daily on machines and manufacturing lines [20].

There are many DM algorithms. These algorithms can be classified into two main groups [35]: supervised and unsupervised. Supervised algorithms are those that work with instances that a priori already belongs to a class. Unsupervised algorithms are used precisely to try to classify instances into groups that were not known a priori. Regression and classification models are the most well-known supervised models while clustering is the main technique of the unsupervised category [36].

The data obtained from the surface roughness measurements have been categorized into two classes, using the median value as the border: surface roughness values below the median value have been categorized as class 1; surface roughness values above the median value have been categorized as class 2.

From these data, a total of 40 models have been generated, using ten of the most common classification algorithms in data mining. These algorithms are briefly described in Table 4. Two parameters have been used to evaluate each model: the percentage of correctly classified instances and the kappa statistic. These parameters are calculated using the cross-validation procedure (10-folds).

Table 4. Description of the data mining algorithms used in the present work (elaborated from [37]).

Algorithm	Description	Ref.
Bayes Net (BN)	It is a Bayesian classification algorithm that provides joint conditional probability distributions. BN algorithm consists of a directed acyclic graph and a set of conditional probability tables. Each random variable is expressed by a node in the directed acyclic graph. The conditional probability table for the values of the variables indicate each possible combination of the values of its parent nodes.	[38]
Naïve-Bayes	It is a statistical classification algorithm which is based on Bayes' theorem. The suppositions of accepting that predictive attributes are conditionally independent given the class and no hidden or latent attributes influence the predictive process make the algorithm a suitable tool for classification and learning.	[39]
Multilayer Perceptron (MLP)	MLP is a feed-forward neural network with one or more hidden layers that uses back-propagation to classify instances. The structure of an MLP typically consists of an input layer, hidden layers and output layer, where the input signals are propagated in the forward direction.	[40]
Simple Logistics	Logistic regression is a statistical model that predicts the probability of some event occurring as a linear function of a set of predictor variables. Linear regression presents two problems: the membership values are not proper probability values and the least-squares regression takes errors as both statistically independent and normally distributed with the same standard deviation. In order to get rid of these problems, logistic regression generates a linear model based on a transformed target variable.	[35]
Sequential Minimal Optimization (SMO)	It is a support vector machine (SVM) classifier that employs sequential minimal optimization for training. SVM is a method for classification of linear and nonlinear data that uses a nonlinear mapping for transforming the original data into a higher dimension.	[38]
IBk	It is an instance-based learning algorithm which is a slightly modified version of the K-nearest neighbor (KNN) algorithm. The algorithm can determine the appropriate value for K based on cross-validation. It normalizes the ranges of attributes, processes instances incrementally and has a policy for tolerating missing values.	[41]
KStar	It is an instance-based learning algorithm which uses an entropy-based distance function. It handles with symbolic attributes, real-valued attributes and missing values properly owing to the use of entropy as a distance function. The technique of summing probabilities over all possible paths overcomes the problem of smoothness.	[42]
J48	J48 is a slightly modified version of C4.5 in WEKA. C4.5 is a successor of ID3 algorithm. The test attribute selection criteria of the algorithm is information gain to overcome the attribute bias problem of ID3. For a given set, each time the algorithm selects an attribute with the highest information gain.	[43]
Logistic Model Trees (LMT)	LMT is a classification algorithm that integrates decision tree induction with logistic regression. The tree structure of the algorithm is grown in a similar manner to the C4.5 algorithm. Here, an iterative training of additive logistic regression models is performed. By splitting, the logistic regressions of the parent node are passed to the child nodes. This provides to have all parent models and probability estimates for each class at the leaf nodes of the final model.	[44]
Random Forest	It is an ensemble of classification or regression trees, induced from bootstrap samples of the training data. In this model, the generalization error of the classifier depends on the power of the individual trees and the association between the trees. Random feature selection is used in the tree induction process. This enables the algorithm to perform comparable to the Adaboost algorithm and to be tolerable with noisy data.	[45]

2.4. Decision Trees

Decision trees are one of the most widely used supervised algorithms. In this category are [35]: C4.5, CART, random forest, random tree, among others. The C4.5 algorithm was developed by Quinlan [43]. This algorithm generates tree-shaped models that allow the classification of instances in a simple, visual and easy to understand way.

The C4.5 algorithm is based on the concept of entropy, understanding entropy as a measure of data disorder [43]. The entropy of a vector and $\vec{y}$ can be calculated as shown in Equation (1) iterating over all possible values of $\vec{y}$. Conditional entropy is calculated as shown in Equation (2).

$$Entropy(\vec{y}) = -\sum_{j=1}^{n} \frac{|y_j|}{|\vec{y}|} log \frac{|y_j|}{|\vec{y}|} \tag{1}$$

$$Entropy(j|\vec{y}) = \frac{|y_j|}{|\vec{y}|} log \frac{|y_j|}{|\vec{y}|}. \tag{2}$$

Finally, the gain is defined as shown in Equation (3).

$$Gain(\vec{y}, j) = Entropy(\vec{y} - Entropy(j|\vec{y}) \tag{3}$$

For each node in the tree, the algorithm chooses the attribute that most effectively divides the original set into different smaller subsets. The attribute with the highest information gain is chosen as a decision parameter. The WEKA data mining software, developed by the University of Waikato (Hamilton, New Zealand), includes the J48 algorithm, based on the C4.5 [35].

2.5. Case Study: Printing a Frame Glasses

As a practical application of the work, the manufacture of a frame glasses is proposed. A frame is made up of two main elements: front and temples. Each of these elements has been personalized for the face/head of one of the authors, who cannot find glasses on the market that fit his morphology. In this case, the front has been printed in a horizontal orientation (XY) and the temples in a vertical orientation (XZ). The printing parameters chosen in each case have been selected after this study.

3. Results

3.1. Data from the Tests

The mean values and standard deviation of surface roughness in the direction parallel to extrusion path ($R_{a,0}$) and in the direction perpendicular to extrusion path ($R_{a,90}$) for both print orientation (XY and XZ) are shown in Tables 5 and 6. As expected, due to the deposition of the fused filament layer after layer, the surface roughness in the direction parallel to the extrusion path is lower than the surface roughness in the perpendicular direction.

3.2. Comparison of Models Generated via Data Mining Algorithms

To analyze the data, they have been classified into two groups (Table 7): class 1 (low surface roughness values) and class 2 (high surface roughness values). This data has been processed by the WEKA software. By means of this software, different classification algorithms have been used to predict whether an instance belongs to class 1 or class 2. Each algorithm has been used to generate 4 models: $R_{a,0XY}$, $R_{a,90XY}$, $R_{a,0XZ}$, $R_{a,90XZ}$. Each model has been evaluated using two criteria: on the one hand, the number of correctly classified instances; on the other, the value of the kappa statistic (Table 8).

Table 5. Mean and standard deviation of surface roughness for XY specimens.

Test	$R_{a,0,XY}$ (µm)		$R_{a,90,XY}$ (µm)	
	Mean	Std. Dev.	Mean	Std. Dev.
1	0.80	0.34	0.71	0.34
2	1.68	0.56	1.41	0.75
3	0.46	0.08	1.18	1.59
4	1.55	0.33	1.96	1.87
5	1.02	0.41	2.14	2.49
6	0.39	0.09	2.16	3.33
7	1.11	0.38	2.83	3.63
8	2.09	0.25	3.45	4.05
9	1.16	0.23	3.46	4.82
10	0.59	0.17	3.59	5.56
11	1.29	0.64	4.31	5.80
12	0.96	0.18	4.38	6.61
13	0.74	0.24	4.66	7.23
14	6.58	1.47	7.35	6.30
15	0.69	0.26	5.32	8.39
16	0.89	0.50	5.80	8.84
17	7.93	0.86	8.60	8.09
18	0.48	0.15	6.21	10.21
19	0.92	0.20	6.70	10.65
20	10.36	1.35	10.57	9.33
21	0.64	0.20	7.28	11.88
22	6.96	1.79	10.25	10.50
23	9.46	4.06	12.17	9.76
24	0.67	0.25	8.31	13.59
25	4.91	0.83	10.25	12.94
26	1.48	0.44	9.31	14.47
27	0.80	0.36	9.39	15.26

Table 6. Mean and standard deviation of surface roughness for XZ specimens.

Test	$R_{a,0,XZ}$ (µm)		$R_{a,90,XZ}$ (µm)	
	Mean	Std. Dev.	Mean	Std. Dev.
1	0.58	0.25	13.29	1.15
2	0.57	0.21	12.63	0.66
3	0.45	0.03	11.54	0.13
4	0.39	0.13	13.04	0.25
5	0.40	0.10	12.25	0.42
6	0.60	0.23	11.55	0.05
7	0.48	0.15	12.92	0.49
8	0.66	0.11	12.87	0.51
9	1.12	1.11	12.44	0.35
10	0.76	0.24	15.31	0.20
11	0.37	0.08	15.34	0.32
12	0.34	0.10	14.89	0.16
13	0.69	0.55	16.35	0.68
14	0.71	0.16	15.81	0.61
15	0.65	0.16	14.44	0.14
16	0.63	0.18	16.91	1.00
17	0.34	0.12	16.23	0.62
18	0.42	0.06	15.58	0.77
19	0.67	0.20	19.25	0.59
20	0.56	0.25	18.93	0.33
21	0.72	0.30	17.86	0.10
22	0.58	0.27	19.19	0.55
23	0.63	0.18	19.16	0.28
24	0.46	0.16	18.94	0.88
25	0.67	0.19	20.06	1.25
26	0.74	0.25	18.80	0.33
27	0.64	0.21	18.36	0.50

Table 7. Values used to establish the classes needed to perform data processing at WEKA.

Orientation	Direction	Class 1	Class 2
XY	0°	0.39–1.02 μm	1.03–10.36 μm
XY	90°	5.79–13.43 μm	13.44–18.69 μm
XZ	0°	0.34–0.60 μm	0.61–1.12 μm
XZ	90°	11.54–15.58 μm	15.59–20.06 μm

Table 8. Strenght of concordance for kappa statistic.

Kappa Statistic	Strength of Concordance
0.00	Poor
0.01–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substancial
0.81–1.00	Almost perfect

The models generated from $R_{a,0,XY}$, $R_{a,90,XY}$, $R_{a0,XZ}$ and $R_{a90,XZ}$ data are shown in Figures 3–6, respectively. The algorithms that achieve better results for $Ra_{0,XY}$ are MLP, J48, LMT and random forest. The algorithms that achieve better for $Ra_{90,XY}$ results are Bayes Net and LMT. The algorithm that achieves better results for $Ra_{0,XZ}$ are MLP and LMT. The algorithms that achieve better results for $Ra_{90,XZ}$ are MLP, SMO, Kstar and random forest.
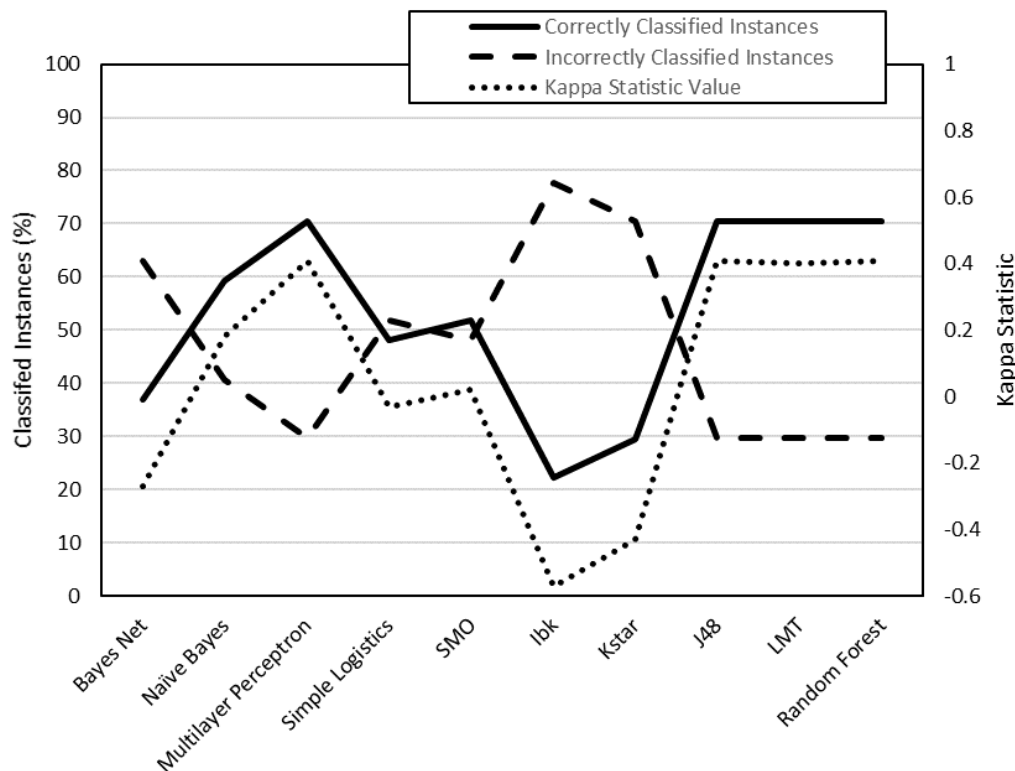
Figure 3. Evaluation of the models generated by WEKA for $R_{a,0,XY}$ data (cross-validation, 10 folds).
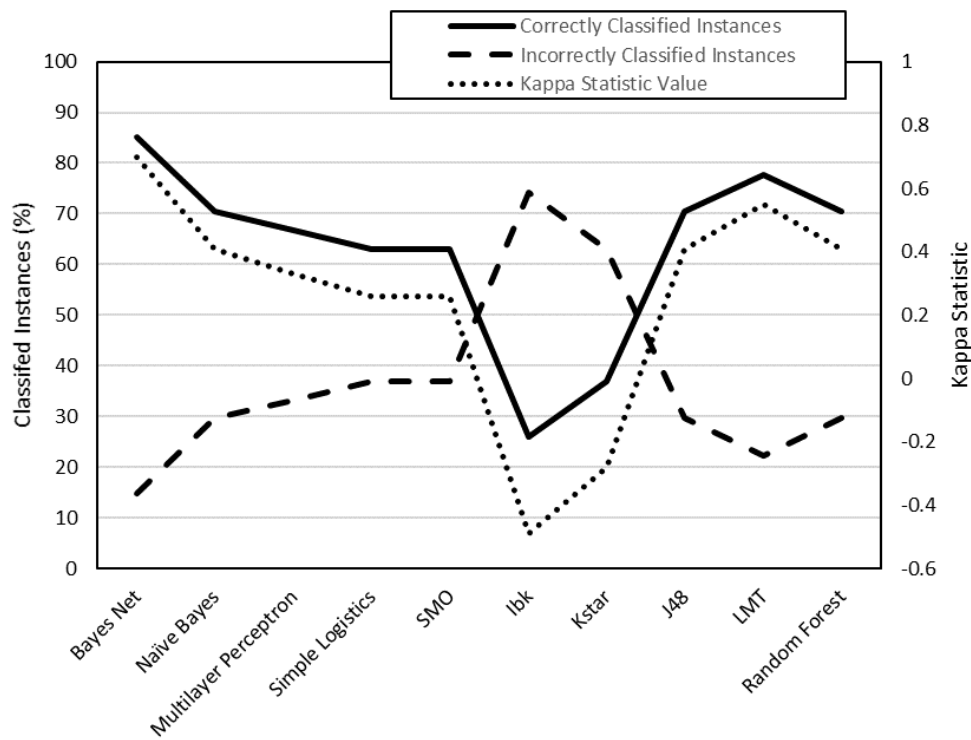
Figure 4. Evaluation of the models generated by WEKA for $R_{a,90,XY}$ data (cross-validation, 10 folds).
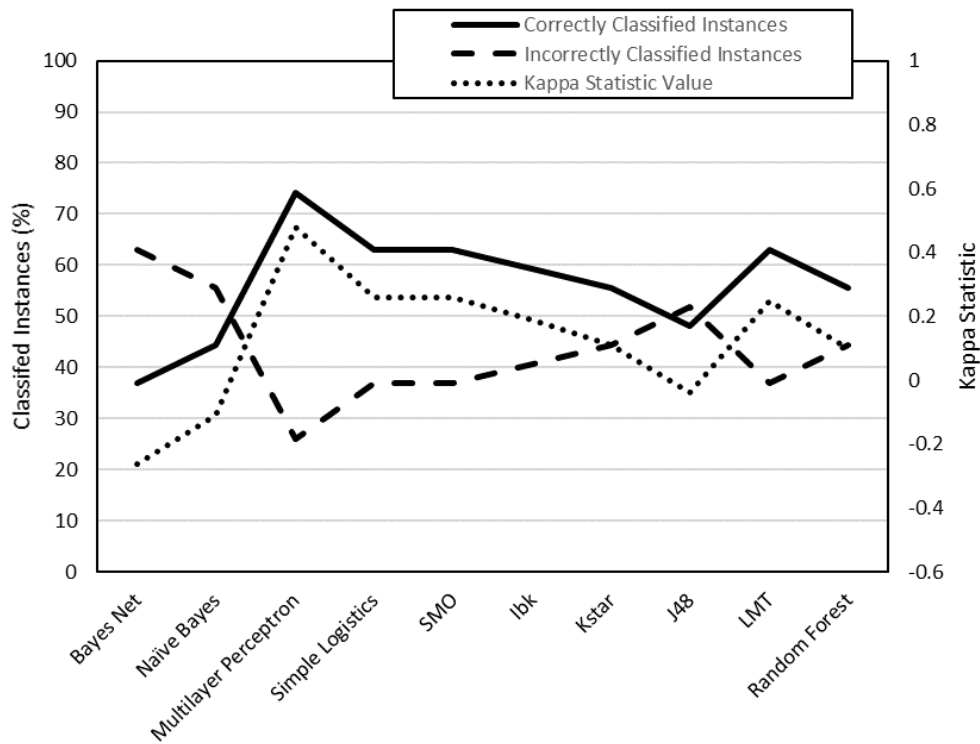
Figure 5. Evaluation of the models generated by WEKA for $R_{a,0,XZ}$ data (cross-validation, 10 folds).
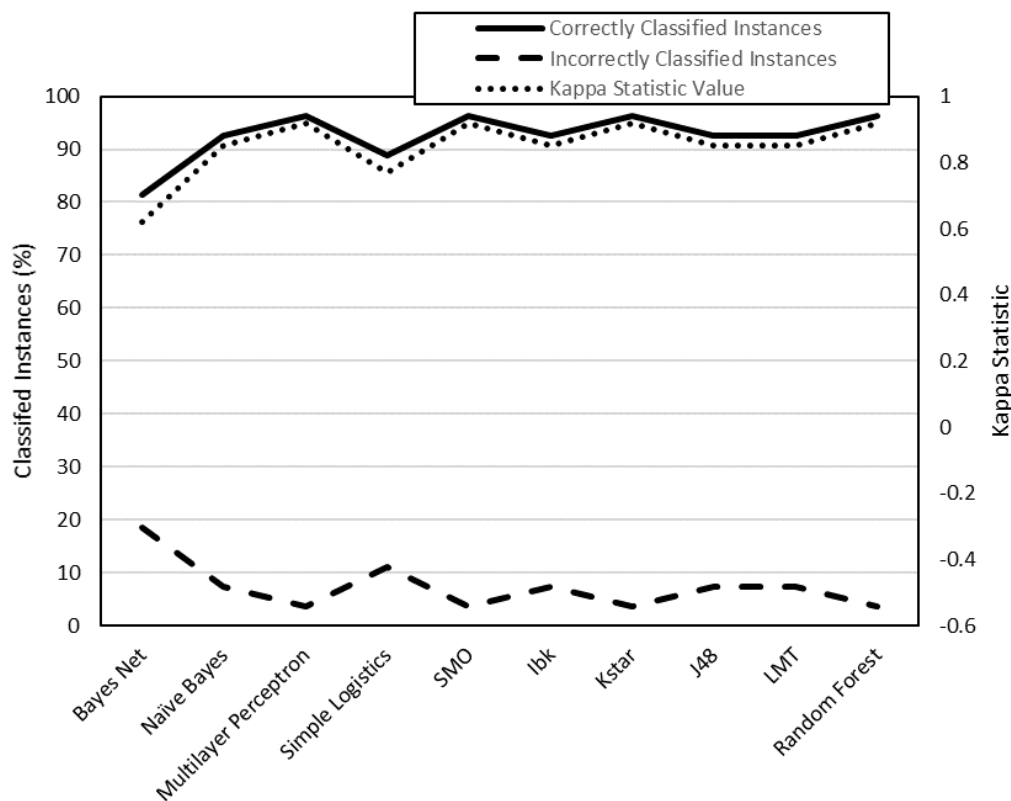
Figure 6. Evaluation of the models generated by WEKA for $R_{a,9,0XZ}$ data (cross-validation, 10 folds).

3.3. Generation of Decision Trees (J48)

As seen in the previous section, except in the case of $Ra_{0,XZ}$, the models generated by the J48 algorithm present adequate results. On the other hand, this algorithm is not a black box: it allows the generation of decision trees that are easy to interpret and that provide additional information.

Figure 7 shows the decision tree generated for the surface roughness in the direction parallel to the extrusion path for the data obtained from specimens printed in the XY orientation ($R_{a,0,XY}$). As it can be seen in the figure, the most important parameters are flow (F), extrusion temperature (T) and layer height (LH). According to this tree, a lower surface roughness (class 1) can be obtained in the following cases:

- By selecting a flow value (F) higher than 100 % (directly).
- By selecting a flow value (F) equal to 90 % and a print temperature (T) equal to 200 °C.
- By selecting a flow (F) value equal to 90 %, a print temperature (T) higher than 200 °C and a layer height (LH) equal to 0.2 mm.

Figure 8 shows the decision tree generated for the surface roughness in the direction perpendicular to the extrusion path, for the data obtained from specimens printed in XY orientation ($R_{a,90,XY}$). In this case, the most important parameter is the flow (F). According to this tree, a lower surface roughness (class 1) can be obtained simply by selecting a flow value (F) higher than 100%.

Figure 9 shows the decision tree generated for surface roughness in the direction parallel to the extrusion path from the data obtained from specimens printed in XZ position ($R_{a,0,XZ}$). As can be seen in this figure, the most important parameter, in this case, is the print speed (PS). According to this tree, a lower surface roughness (class 1) can be obtained simply by selecting a value for the printing speed (PS) equal to 40 mm/s.
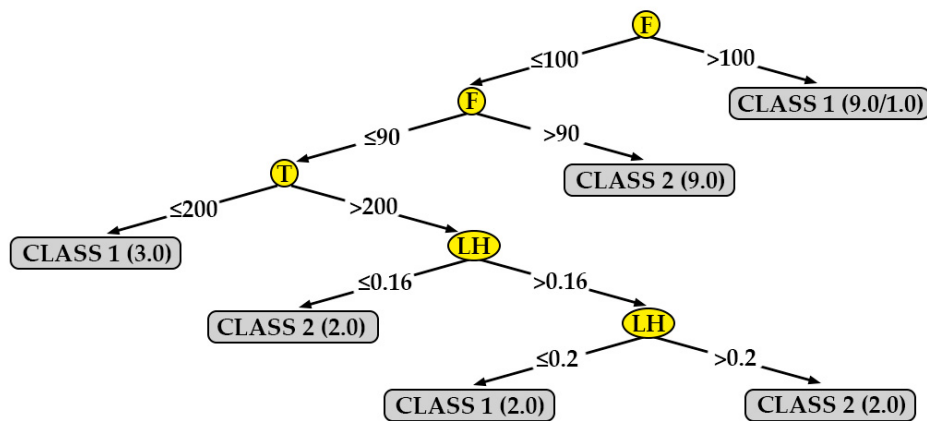
Figure 7. Decision tree for surface roughness measured in the direction parallel to the extrusion path on printed specimens in the XY orientation ($R_{a,0,XY}$).
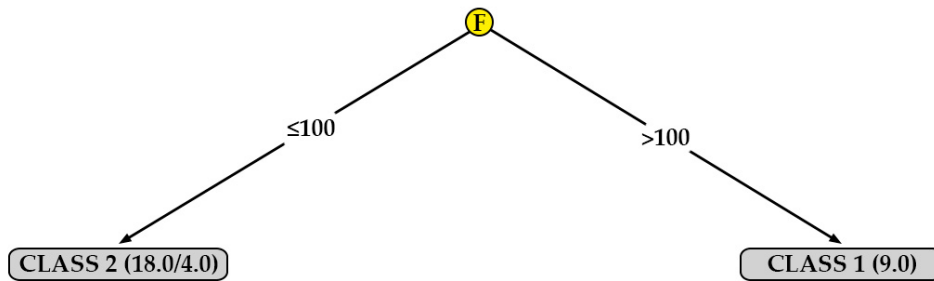

Figure 8. Decision tree for surface roughness measured in the direction perpendicular to the extrusion path on printed specimens in the XY orientation ($R_{a,90,XY}$).
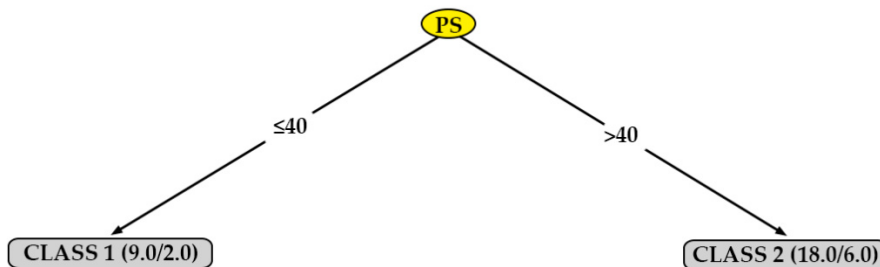

Figure 9. Decision tree for surface roughness measured in the direction parallel to the extrusion path on printed specimens in the XZ orientation ($R_{a,0,XZ}$).

Figure 10 shows the decision tree generated for roughness in the direction perpendicular to the extrusion path, for the data obtained from specimens printed in XZ position ($R_{a,90,XZ}$). In this case, the most important parameters are the layer height (LH) and the extrusion temperature (T). According to this tree, a lower surface roughness (class 1) can be obtained in the following cases:

- By selecting a layer height (LH) equal to 0.16 mm (directly).
- By selecting a layer height (LH) equal to 0.20 mm and a print temperature (T) equal to 200 °C.

Table 9 shows the percentages of correctly and incorrectly classified instances, as well as the kappa statistic for each model. According to Table 8, the models generated for $R_{a,0,XY}$ and $R_{a,90,XY}$ is rated as 'moderate'; on the other hand, the model generated for $R_{a,90,XZ}$ is rated as 'almost perfect'. Finally, the model obtained for $R_{a,0,XZ}$ obtains very poor results. The results obtained from the first three models are supported. This is not the case with the last model.
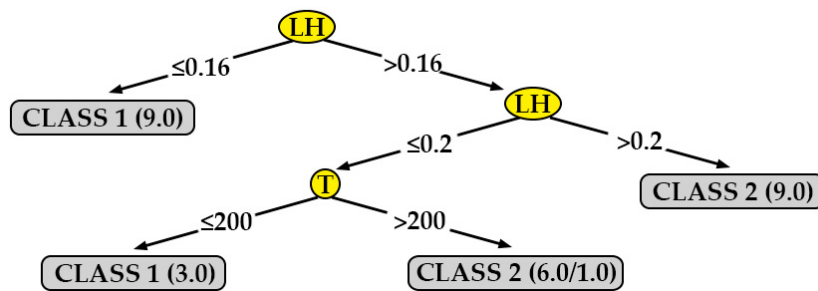
Figure 10. Decision tree for surface roughness measured in the direction perpendicular to the extrusion path on printed specimens in the XZ orientation ($R_{a,90,XZ}$).

Table 9. Information regarding decision tree models generated in this work.

Tree	Correctly Classified Instances (%)	Incorrectly Classified Instances (%)	Kappa Statistics
$R_{a,0,XY}$	70.37	29.63	0.41
$R_{a,90,XY}$	70.37	29.63	0.41
$R_{a,0,XZ}$	44.44	55.56	−0.11
$R_{a,90,XZ}$	92.59	7.41	0.85

In order to check that the information extracted from the decision trees is correct, the mean value and the standard deviation obtained for surface roughness in each of the 27 tests carried out have been represented. As an example, Figures 11 and 12 show the results for R_{a0XY} and R_{a90XZ}, respectively. Figure 11 clearly shows the instances belonging to class 1 (F > 100; F ≤ 90 and T ≤ 200; F ≤ 90 and T > 200 and LH = 0.2) and the instances belonging to class 2 (the rest). Likewise, Figure 12 shows that the instances that meet certain criteria (LH ≤ 0.16; 0.16 < LH ≤ 0.20 and T < 200) have a lower surface roughness than the rest.
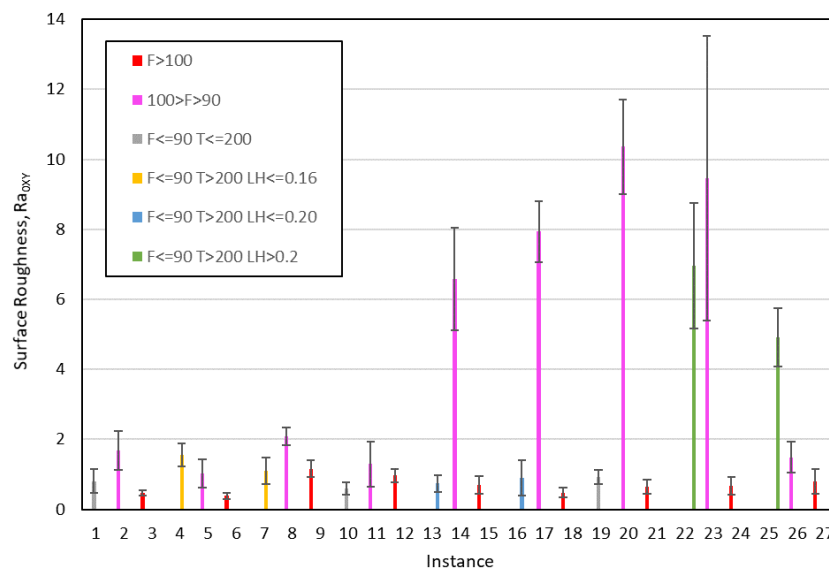

Figure 11. Mean values and standard deviation for R_{a0XY}; each color represents a leaf of the decision tree: pink, yellow and green instances belong to class 2; the rest of color belong to class 1 (F: Flow; T: Temperature; LH: layer height).
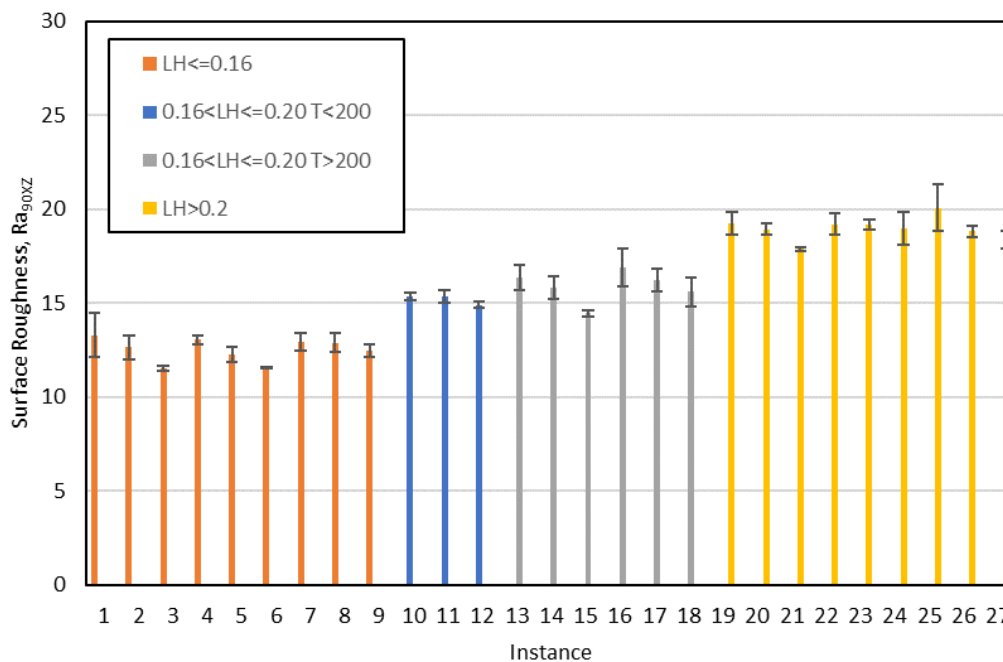
Figure 12. Mean values and standard deviation for R_{a90XZ}; each color represents a leaf of the decision tree: orange and blue instances belong to class 1; grey and yellow ones belong to class 2 (LH: layer height; T: temperature).

3.4. Practical Application: Reduction of Surface Roughness in FDM 3D Printed Frame Glasses

A practical application of the results obtained in this work is the production of frame glasses printed in PLA using FDM (Figure 7). Frames glasses are a perfect example of a product that should be custom-made: each user has a face width, a nose width and a face-to-ear distance (Figure 13, left). The most important dimensions of a frame glasses can be parameterized in a CAD design and adjusted according to the needs of each customer. Once the model was customized, it was printed using an F equal to 110% and an LH equal to 0.16 mm (Figure 13, right).
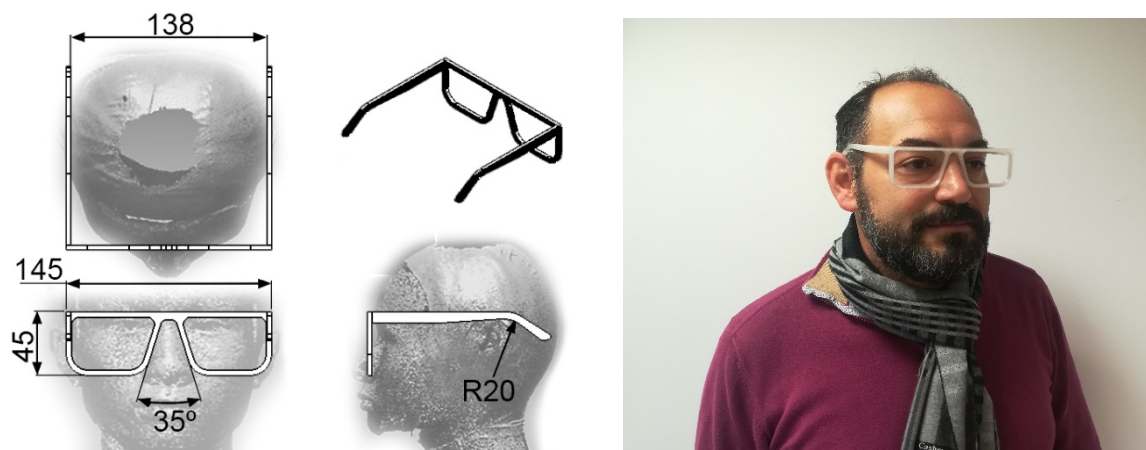

Figure 13. Customized frame glasses: parameterized distances (**left**); real frame glasses produced in PLA via FDM (F = 110%; LH = 0.16 mm) (**right**).

4. Discussion

In the present work, data mining techniques have been used to generate models that allow predicting the surface roughness of parts printed in PLA using 3D FDM printing. The classifier algorithms used to generate the models have been: Bayes net, naïve-Bayes, multilayer perceptron

(ANN), simple logistic, SMO (SVM), IBk (KNN), KStar, J48 (C4.5), LMT y random forest. The models have been tested by means of a cross-validation (10-folds). The percentage of correctly classified instances and the kappa statistic were used to compare these models with each other. The input variables of the study have been: print orientation, layer height (LH), extrusion temperature (T), print speed (PS), print acceleration (PA) and flow (F). The output variables have been the surface roughness measured in the direction parallel to the extrusion path ($R_{a,0}$) and in the direction perpendicular to the extrusion path ($R_{a,90}$). A total of 54 specimens have been printed for this purpose: 27 specimens were printed with horizontal orientation (XY) and the other 27 with vertical orientation (XZ).

From the results shown in Figures 3–6, it can be stated that one of the models that best classify the instances according to their surface roughness (both for horizontal and vertical specimens) is the one generated by the multilayer perceptron (artificial neural network, ANN) algorithm. In the literature there are works that also use different types of ANN to predict the surface roughness of printed parts using FDM: Boschetto et al. [28] propose the use of a feed-forward neural network to predict surface roughness on surfaces that form different angles to the vertical, obtaining errors of less than 5%; Vahabli and Rahmati [29] propose a similar model, using another type of ANN, capable of predicting surface roughness as a function of build angle. However, both works do not take into account the influence of printing parameters on the surface roughness obtained.

Another algorithm that has generated in the present work models with positive results has been the LMT. This fact is consistent with other works of literature: Landwehr et al. [46] concluded that LMT produces more accurate classifiers than J48 (C4.5), CART, logistic regression, models tree, functional trees, naïve Bayes trees and LOTUS. The problem with this algorithm, like ANN, is that it does not generate a model that can be visualized or easily understood.

One algorithm that generates a model that can be represented graphically is J48. Figures 7–10 show the decision trees generated by this algorithm for the different measured surface roughness. From the models obtained, Table 10 has been drawn up. This table summarises the printing variables that must be taken into account in order to obtain a lower surface roughness (class 1). From these decision trees, a basic configuration can be established to obtain a good surface finish simultaneously on XY and XZ orientation, for perpendicular and parallel direction to extrusion path: F equal to 110 % and LH equal to 0.16 mm.

Table 10. Parameters involved in each data tree models.

	XY	**XZ**
$R_{a,0}$	**F**, T, LH	PS
$R_{a,90}$	**F**	**LH**, T

NOTE: The most important parameters have been highlighted in bold.

In addition to the above, the following statements can be made:

- The decision tree models can be easily interpreted by any 3D printer operator (they are not black-box models) [47]. For this reason, they are used for modeling other manufacturing processes in literature [48]. This is important in the current industry that is very concerned with visual management.

- Two parameters that had not been previously studied in the literature on PLA specimens were included in this work: print acceleration and flow. While print acceleration seems not to have an influence in any of the cases studied, the flow is revealed as an important parameter in the surface roughness obtained in XY orientation. This result is consistent with those obtained by the authors in similar tests performed on PETG [47].

- In $R_{a,90,XZ}$, the most influential parameters are layer height. These results coincide with those obtained by other authors: García-Plaza et al. [31] state that the layer height parameter is the most influential in upright and on-edge positions (similar to the XZ orientation in the present paper).

- Three of the four models generated have achieved a 'moderate' or 'almost perfect' rating, according to their kappa statistic values. This fact supports the validity of the results obtained in the present work, as shown in Figures 11 and 12. The model generated for $R_{a0,XZ}$ did not pass the evaluation.
- The results obtained in the present work have been used to print a frame glasses in PLA via FDM. A-frame glasses is the perfect example of a customized part/assembly. There are previous works in the literature focused on this topic [18], although they have not studied the printing parameters that allow to achieve a better surface finish.

5. Conclusions

In the present work, data mining algorithms have been used to generate from experimental tests models capable of predicting the surface roughness of horizontal and vertical parts printed in PLA using FDM. MLP and LMT are the algorithms that obtain the best results in cross validation tests.

The J48 algorithm obtains almost perfect results in one of the models and moderate results in two others. However, unlike MLP and LMT, this algorithm generates decision trees in which it is possible to see which print parameters influence surface roughness.

Of the five parameters studied, two have proved to be the most important for obtaining a better surface finish in XY and XZ specimens: flow and layer height. A flow equal to 110% and a layer height equal to 0.16 mm provides the lowest values of surface roughness (R_a) in PLA printed specimens in XY and XZ orientation. Using these values, a frame of glasses has been printed, as a practical and direct application of this work.

References

1. Lasi, H.; Fettke, P.; Kemper, H.G.; Feld, T.; Hoffmann, M. Industry 4.0. *Bus. Inf. Syst. Eng.* **2014**, *6*, 239–242. [CrossRef]
2. Singh, R. Some investigations for small-sized product fabrication with FDM for plastic components. *Rapid Prototyp. J.* **2013**, *19*, 58–63. [CrossRef]
3. Ahn, S.; Montero, M.; Odel, D.; Roundy, S.; Wright, P.K. Anisotropic material properties of fused deposition modeling ABS. *Rapid Prototyp. J.* **2002**, *8*, 248–257. [CrossRef]
4. Ngo, T.; Kashani, A.; Imbalzano, G.; Nguyen, Q.T.; Hui, D. Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. *Compos. Part B Eng.* **2018**, *143*, 172–196. [CrossRef]
5. Turner, B.N.; Strong, R.; Gold, S.A. A review of melt extrusion additive manufacturing processes: I. Process design and modeling. *Rapid Prototyp. J.* **2014**, *20*, 192–204. [CrossRef]
6. Klippstein, H.; Diaz de Cerio Sanchez, A.; Hassanin, H.; Zweiri, Y.; Seneviratne, L. Fused Deposition Modeling for Unmanned Aerial Vehicles (UAVs): A Review. *Adv. Eng. Mater.* **2017**, *1700552*, 1–17. [CrossRef]
7. Popescu, D.; Zapciu, A.; Amza, C.; Baciu, F.; Marinescu, R. FDM process parameters influence over the mechanical properties of polymer specimens: A review. *Polym. Test.* **2018**, *69*, 157–166. [CrossRef]
8. Afrose, M.F.; Masood, S.H.; Iovenitti, P.; Nikzad, M.; Sbarski, I. Effects of part build orientations on fatigue behaviour of FDM-processed PLA material. *Prog. Addit. Manuf.* **2016**, *1*, 21–28. [CrossRef]
9. Smart Materials 3D-PLA3D870. Available online: https://www.smartmaterials3d.com/en/inicio/172-1611-pla-3d870.html#/2-size-750g/25-colour-natural/62-shoes_size-175mm (accessed on 13 November 2019).

10. Torres, J.; Cotelo, J.; Karl, J.; Gordon, A.P. Mechanical property optimization of FDM PLA in shear with multiple objectives. *JOM* **2015**, *67*, 1183–1193. [CrossRef]

11. Song, Y.; Li, Y.; Song, W.; Yee, K.; Lee, K.Y.; Tagarielli, V.L. Measurements of the mechanical response of unidirectional 3D-printed PLA. *Mater. Des.* **2017**, *123*, 154–164. [CrossRef]

12. Chacón, J.M.; Caminero, M.A.; García-Plaza, E.; Núñez, P.J. Additive manufacturing of PLA structures using fused deposition modelling: Effect of process parameters on mechanical properties and their optimal selection. *Mater. Des.* **2017**, *124*, 143–157. [CrossRef]

13. Caminero, M.Á.; Chacón, J.M.; García-Plaza, E.; Núñez, P.J.; Reverte, J.M.; Becar, J.P. Additive manufacturing of PLA-based composites using fused filament fabrication: Effect of graphene nanoplatelet reinforcement on mechanical properties, dimensional accuracy and texture. *Polymers* **2019**, *11*, 799. [CrossRef] [PubMed]

14. Galati, M.; Minetola, P. On the measure of the aesthetic quality of 3D printed plastic parts. *Int. J. Interact. Des. Manuf.* **2019**. [CrossRef]

15. Boschetto, A.; Giordano, V.; Veniali, F. 3D roughness profile model in fused deposition modelling. *Rapid Prototyp. J.* **2013**, *19*, 240–252. [CrossRef]

16. Turner, B.N.; Gold, S.A. A review of melt extrusion additive manufacturing processes: II. Materials, dimensional accuracy, and surface roughness. *Rapid Prototyp. J.* **2015**, *21*, 250–261. [CrossRef]

17. Li, Y.; Linke, B.S.; Voet, H.; Falk, B.; Schmitt, R.; Lam, M. Cost, sustainability and surface roughness quality–A comprehensive analysis of products made with personal 3D printers. *CIRP J. Manuf. Sci. Technol.* **2017**, *16*, 1–11. [CrossRef]

18. Besnea, D.; Rizescu, D.; Rizescu, C.I.; Constantin, V.; Spanu, A. Additive technologies and materials used for making glasses frames. *Int. J. Mechatron. Appl. Mech.* **2018**, *3*, 13–17.

19. Pérez, M.; Medina-Sánchez, G.; García-Collado, A.; Gupta, M.; Carou, D. Surface quality enhancement of fused deposition modeling (FDM) printed samples based on the selection of critical printing parameters. *Materials* **2018**, *11*, 1382. [CrossRef]

20. Köksal, G.; Batmaz, I.; Testik, M.C. A review of data mining applications for quality improvement in manufacturing industry. *Expert Syst. Appl.* **2011**, *38*, 13448–13467. [CrossRef]

21. Harding, J.A.; Shahbaz, M.; Srinivas, S.; Kusiak, A. Data Mining in Manufacturing: A review. *J. Manuf. Sci. Eng.* **2006**, *128*, 969–976. [CrossRef]

22. Choudhary, A.K.; Harding, J.A.; Tiwari, M.K. Data mining in manufacturing: A review based on the kind of knowledge. *J. Intell. Manuf.* **2009**, *20*, 501–521. [CrossRef]

23. Razvi, S.; Feng, S.; Narayana, A.; Lee, Y.-T.; Witherell, P. A review of machine learning applications in additive manufacturing. In Proceedings of the ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Anaheim, CA, USA, 18–21 August 2019; pp. 1–10.

24. Amand, B.; Cordy, M.; Heymans, P.; Acher, M.; Temple, P.; Jézéquel, J. Towards learning-aided configuration in 3D printing: Feasibility study and application to defect prediction. In *Proceedings of the 13th International Workshop on Variability Modelling of Software Intensive Systems (VaMoS'19), Leuven, Belgium, 6–8 February 2019*; Perrouin, G., Weyns, D., Eds.; ACM: New York, NY, USA; Leuven, Belgium, 2019; p. 9.

25. Wu, D.; Wei, Y.; Terpenny, J. Surface roughness prediction in additive manufacturing using machine learning. In Proceedings of the ASME 2018 13th International Manufacturing Science and Engineering Conference (MSEC2018), College Station, TX, USA, 18–22 June 2018; pp. 1–6.

26. Sohnius, F.; Schlegel, P.; Ellerich, M.; Schmitt, R.H. Data-driven Prediction of Surface Quality in Fused Deposition Modeling using Machine Learning. In *Proceedings of the Production at the Leading Edge of Technology, Hamburg, Germany, 30 September–2 October 2019*; Wulfsberg, J., Hintze, W., Behrens, B., Eds.; Springer: Berlin, Germany, 2019; pp. 473–481.

27. Mahapatra, S.S.; Sood, A.K. Bayesian regularization-based Levenberg-Marquardt neural model combined with BFOA for improving surface finish of FDM processed part. *Int. J. Adv. Manuf. Technol.* **2012**, *60*, 1223–1235. [CrossRef]

28. Boschetto, A.; Giordano, V.; Veniali, F. Surface roughness prediction in fused deposition modelling by neural networks. *Int. J. Adv. Manuf. Technol.* **2013**, *67*, 2727–2742. [CrossRef]

29. Vahabli, E.; Rahmati, S. Improvement of FDM parts' surface quality using optimized neural networks-Medical case studies. *Rapid Prototyp. J.* **2017**, *23*, 825–842. [CrossRef]

30. Youreyewear. Available online: https://www.youreyewear.com/ (accessed on 18 January 2020).

31.　García Plaza, E.; Núñez López, P.J.; Caminero Torija, M.Á.; Chacón Plaza, J.M. Analysis of PLA geometric properties processed by FFF additive manufacturing: Effects of process parameters and plate-extruder precision motion. *Polymers* **2019**, *11*, 1581. [CrossRef]

32.　Ramli, F.R.; Faudzie, M.S.M.; Nazan, M.A.; Alkahari, M.R.; Sudin, M.N.; Mat, S.; Khalil, S.N. Dimensional accuracy and surface roughness of part features manufactured by open source 3D printer. *ARPN J. Eng. Appl. Sci.* **2018**, *13*, 1139–1144.

33.　Alsoufi, M.S.; El-sayed, A. How Surface Roughness Performance of Printed Parts Manufactured by How Surface Roughness Performance of Printed Parts Manufactured by Desktop FDM 3D Printer with PLA + is Influenced by Measuring Direction. *Am. J. Mech. Eng.* **2017**, *5*, 211–222.

34.　Kovan, V.; Tezel, T.; Topal, E.S.; Camurlu, H.E. Printing Parameters Effect on Surface Characteristics of 3D Printed Pla Materials. *Mach. Technol. Mater.* **2018**, *269*, 266–269.

35.　Witten, I.H.; Frank, E.; Hall, M.A. *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed.; Morgan Kaufmann: Burlington, MA, USA, 2011.

36.　Amini, M.; Chang, S. MLCPM: A process monitoring framework for 3D metal printing in industrial scale. *Comput. Ind. Eng.* **2018**, *124*, 322–330. [CrossRef]

37.　Onan, A. *On the Performance of Ensemble Learning for Automated Diagnosis of Breast Cancer*; Silhavy, R., Senkerik, R., Oplatkova, Z., Prokopova, Z., Silhavy, P., Eds.; Springer: Cham, Switzerland, 2015; Volume 347, ISBN 9783319184753.

38.　Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*; Morgan Kaufmann Publishers, Inc.: San Francisco, CA, USA, 2011.

39.　John, G.; Langley, P. Estimating continuous distributions in Bayesian classifiers. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Montreal, QU, Canada, 18–20 August 1995*; Morgan Kaufmann Publishers, Inc.: San Francisco, CA, USA, 1995; pp. 338–345.

40.　Negnevitsky, M. *Artificial Intelligence: A Guide to Intelligent Systems*; Addison-Wesley: Reading, UK, 2005.

41.　Aha, D.W.; Kibler, D.; Albert, M. Instance-Based Learning Algorithms. *Mach. Learn.* **1991**, *6*, 37–66. [CrossRef]

42.　Clearly, J.G.; Trigg, L.E. An Instance-based learner using and entropic distance measure. In *Proceedings of the Twelfth International Conference on Machine Learning, Tahoe City, CA, USA, 9–12 July 1995*; Morgan Kaufmann Publishers, Inc.: San Francisco, CA, USA, 1995; pp. 108–114.

43.　Quinlan, J.R. *C4.5: Programs for Machine Learning*; Morgan Kaufmann Publishers, Inc.: San Mateo, CA, USA, 2014.

44.　Landwehr, N.; Hall, M.; Frank, E. Logistic Model Trees. *Mach. Learn.* **2005**, *59*, 161–205. [CrossRef]

45.　Breiman, L. Random forest. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

46.　Landwehr, N.; Hall, M.; Frank, E. Logistic model trees. *Lect. Notes Artif. Intell. Subser. Lect. Notes Comput. Sci.* **2003**, *2837*, 241–252.

47.　Barrios, J.M.; Romero, P.E. Decision tree methods for predicting surface roughness in fused deposition modeling parts. *Materials* **2019**, *12*, 2574. [CrossRef] [PubMed]

48.　Kuriakose, S.; Mohan, K.; Shunmugam, M.S. Data mining applied to wire-EDM process. *J. Mater. Process. Technol.* **2003**, *142*, 182–189. [CrossRef]