Predictive Modelling for Quality Prediction and Assurance of Extrusion Blow Molding

Vongur Ramulu, E. V. Ramana, N. Kiran Kumar

Abstract: Extrusion Blow Molding process plays an important role in manufacturing of hollow products with wide variety of materials like polyethylene (PE), polypropylene (PP), polyvinylchloride (PVC). Extrusion blow molded products are rejected due to the occurrence of defects such as die lines, blowouts, shrinkage, over weight of part. The complex relationships that exist between the process variables, and causes of defects are investigated for 1 litre container made of high-density polyethylene (HDPE) using data mining techniques in order to reduce scrap. In this paper Data Mining approach is implemented by applying Decision Tree, k-Nearest Neighbors, Rule Induction and Vote techniques in RapidMiner for quality assurance and prediction of the quality of the extrusion blow molded product.

Keywords: Rule Induction, k-Nearest Neighbors, Decision Tree, Vote, Extrusion Blow Molding.

I. INTRODUCTION

Extrusion Blow Molding is one of the manufacturing processes in which plastic pellets of different kind of polymers are transferred into barrel heaters through hopper where plastic materials are melted and comes out as hollow tube from die head. This hollow tube is called parison or preform which is expanded in mold cavity by blowing air into it in order to produce different types of hollow products [1]. The part quality is affected by various process variables such as barrel temperatures, extrusion time, blow pressure, blow time, extrusion speed, cut off blade temperature, cooling time, ejection time, cycle time etc. These process variables have to be continuously monitored and controlled to achieve required quality and to minimize the rejections.

The techniques like Fuzzy iterative learning control algorithm, Neural Network approach, Grey Relational Analysis and Taguchi method and Genetic Algorithm are applied in optimizing Extrusion Blow Molding process [2]-[5]. Quality prediction of the plastic injection molding process is done by using Support vector machines, k-Nearest Neighbors and General classification and regression trees techniques [6].

In present work using the RapidMiner with Rule Induction, Decision Tree, k-Nearest Neighbors and Vote techniques are implemented for predicting quality and to find root causes of defects in Extrusion Blow Molded products.

Revised Manuscript Received on September 03, 2019

Vongur Ramulu, PG~~post graduating~~ student in Advanced Manufacturing Systems (AMS) from VNR VJIET, Hyderabad.

E. V. Ramana is working as a Professor, Department of Mechanical Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad.

Kiran Kumar Namala has done doctoral program from IIT Delhi, post-graduation in production engineering.

II. DATASET

The process dataset for Extrusion Blow Molding product (1 ~~–~~litre ~~–~~container) made of high-density polyethylene (HDPE) material with 200 records is used in the current work. Process attribute data: barrel temperatures at five zones; blow pressure, blow time, cut off blade temperature, cooling time, cycle time, extrusion speed, extrusion time, ejection time attributes are collected. The process dataset is divided into training and test datasets comprising of 160 and 40 examples respectively.

III. RULE INDUCTION

Rule induction technique create models by following the if-else-then rule type. The training dataset is given as input to Rule Induction operator as shown in the process presented in Fig. 1. Information gain is selected as criterion in which entropy of all the attributes is calculated and split is done based on attribute with minimum entropy. The sample ratio of training examples applied to grow and prune is set as 1.0. The pureness parameter is set to 0.5. The minimal prune benefit parameter is assigned the value of 0.25 [8]. The rule model is shown Fig. 2.

The model delivered through output (mod) port of Rule Induction operator is subsequently fed to Apply Model operator through input (mod) port. The test dataset is given as input to Apply Model operator through unlabelled (unl) port. The test dataset with known class labels is supplied as input to performance vector operator to evaluate the performance of Rule Induction model for classification task.

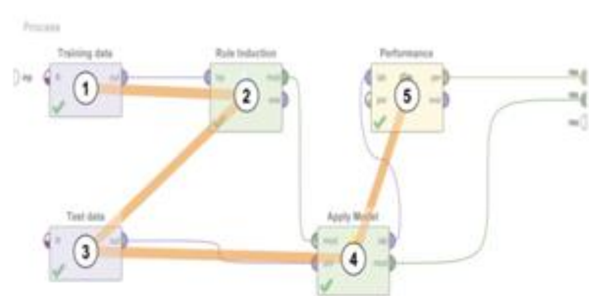


Fig. 1 RapidMiner process with Rule Induction operator

If Cycle Time (sec) > 22.750 and Recycled Material = yes then over weight of part (32 / 25 / 0 / 0 / 0)
if Blow Pressure (kg/cm²) > 6.350 then blowouts (0 / 0 / 32 / 0 / 0)
if Recycled Material = No then accepted (0 / 0 / 0 / 32 / 0)
else shrinkage (0 / 7 / 0 / 0 / 32)
correct: 128 out of 160 training examples.

Fig. 2 Rule model

The first rule is applicable when cycle time is greater than 22.750 seconds and recycled material to classify the parts that belong to *over weight of part* class and 32 cases are supporting the rule. This rule misclassified 25 die line cases as over weight class. The second rule implies that when blow pressure is greater than 6.350 kg/cm^2 then it results in *blowouts* with the support of 32 cases. It can be derived from the third rule that if there is no usage of recycled material the parts belongs to *accepted* class with support of 32 cases and it misclassified 7 cases as belonging to *die lines* class and 32 cases to *shrinkage* class.

The confusion matrix showing the class precision and recall, prediction accuracy of Rule Induction model on test dataset is presented in Table I.

Table- I: Confusion matrix for Rule Induction model

Accuracy: 80%						
	True over weight of part	**True die lines**	**True blowouts**	**True accepted**	**True shrinkage**	**Class precision**
Pred. Over weight of part	8	8	0	0	0	50.00%
Pred. Die lines	0	0	0	0	0	0.00%
Pred. Blowouts	0	0	8	0	0	100.00 %
Pred. Accepted	0	0	0	8	0	100.00 %
Pred. Shrinkage	0	0	0	0	8	100.00 %
Class recall	100.00 %	0.00 %	100.00 %	100.00 %	100.00%	

The rule induction model predicted the products having blowouts, shrinkage and accepted ones as true positives for all cases. Prediction of over weight of part class resulted in 8 true positive cases and 8 cases of false negatives (die lines).

IV. DECISION TREE

Decision tree comprises of root node, leaf node and branches. The tests on attributes are represented by internal nodes, branches represent outcome of the tests and leaf nodes represent the class labels. The training dataset is given as input to Decision Tree operator as shown in the process presented in Fig. 3. The gain ratio is a criterion which is to be selected for splitting the attributes. Maximum depth parameter set to value of 3 for controlling size of decision tree. Pruning and prepruning parameters are to be enabled to calculate pessimistic error. Minimal gain parameter is set to 0.25 for node split to control the size of tree. Minimum leaf size is set to 2 to ensure that each leaf node has at least two examples in its subset. The parameter of minimal size for split is set to 3 to ensure minimum number of examples in nodes for splitting. Number of Prepruning alternatives are set to 4 in case splitting is prevented by pre-pruning.

The model delivered through output (mod) port of decision tree operator is subsequently fed to Apply Model operator through input (mod) port. The test dataset is given as input to Apply Model operator through unlabelled (unl) port. The test dataset with known class labels is supplied as input to performance vector operator which evaluate the performance of Decision Tree model for classification task.

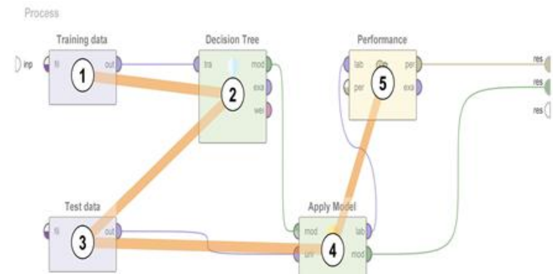
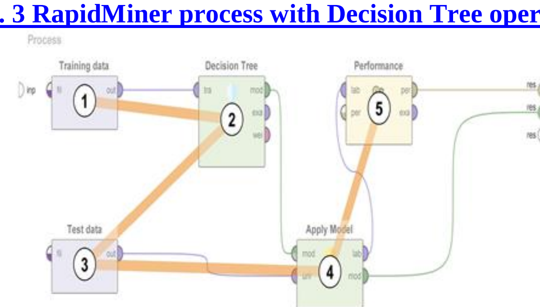


Fig. 3 RapidMiner process with Decision Tree operator



The Decision Tree (DT) built on training dataset by the DT model is presented in Fig. 4. The following rules are resulted from the decision tree.

Rule 1: If Trace of Extraneous Material=No and Recycled Material=No then class=accepted (support: 32 cases).

Rule 2: If Trace of Extraneous Material=No and Recycled Material=Yes then class=over weight of part (support: 32 cases).

Rule 3: If Trace of Extraneous Material=No and Recycled Material=Yes then class=shrinkage (support: 32 cases).

Rule 4: If Trace of Extraneous Material=Yes when Blow Pressure (kg/cm^2) ≥6.350 then class=blowouts (support: 32 cases).

Rule 5: If Trace of Extraneous Material=Yes when Blow Pressure (kg/cm^2) ≤ 6.350 then class=die lines (support: 32 cases) ~~cases).~~

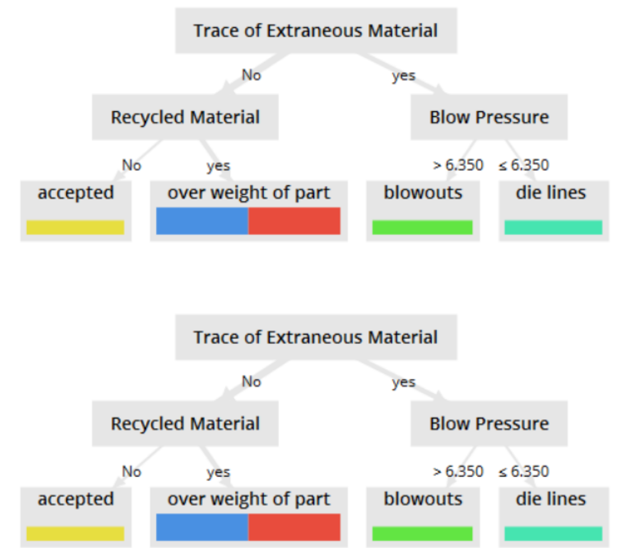


Fig. 4 Decision Tree model

The confusion matrix representing class precision and recall, prediction accuracy of Decision Tree model on test dataset is shown in Table II.

Table- II: Confusion matrix for Decision Tree model

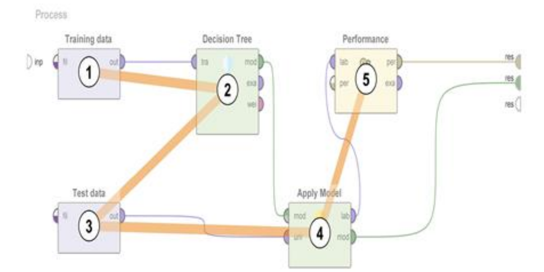
Accuracy: 80%						
	True over weight of part	**True die lines**	**True blowouts**	**True accepted**	**True shrinkage**	**Class precision**
Pred. Over weight of part	8	0	0	0	8	50.00%
Pred. Die lines	0	8	0	0	0	100.00 %
Pred. Blowouts	0	0	8	0	0	100.00 %
Pred. Accepted	0	0	0	8	0	100.00 %
Pred. Shrinkage	0	0	0	0	0	0.00%
Class recall	100.00 %	100.00%	100.00%	100.00%	0.00%	

The decision tree models predicted the products having die lines, blowouts and accepted ones as true positives for all cases. Prediction of over weight of part class resulted in 8 true positives cases and 8 cases of false positives (shrinkage). Prediction of shrinkage class has outcome of 8 false negatives (overweight of part).

V. k-NEAREST NEIGHBORS

k-Nearest Neighbors (k-NN) is based on analogy learning where it compares a given test example with training examples which are similar to it [7]. The training dataset is given as input to k-NN operator as shown in the process presented in Fig. 5. The parameter k is set to 4 due to the performance of the model found to be the highest among other values in the range of 1-5 as depicted in Fig.6. The class of four nearest neighbors is assigned as the class of unknown example by the model. Weighted vote is considered so that nearest neighbors have more weightage than the distant ones. Mixed measure type is chosen to calculate distances for both nominal and numerical attributes and Mixed Euclidean Distance measure for calculation of numerical values in constructing the model.

The model delivered through output (mod) port of k-NN operator is subsequently fed to Apply Model operator through input (mod) port. The test dataset is given as input to Apply Model operator through unlabelled (unl) port. The test dataset with known class labels is supplied as input to performance vector operator to evaluate the performance of k-NN model for classification task.



~~Fig. 5 RapidMiner process with k-NN operator~~

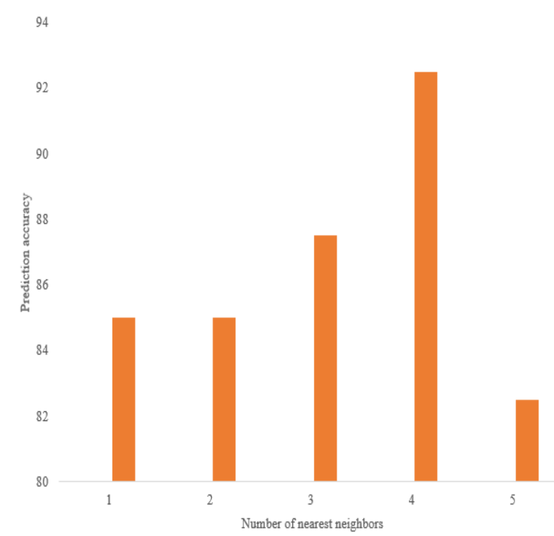


Fig. 6 Number of nearest neighbors vs Prediction accuracy

The confusion matrix representing class precision and recall, prediction accuracy of k-NN model on test dataset is shown in Table III and prediction accuracy is found to be 92.50.

Table- III: Confusion matrix for k-NN model

Accuracy: 92.5%						
	True over weight of part	**True die lines**	**True blow outs**	**True accepted**	**True shrinkage**	**Class precision**
Pred. Over weight of part	8	0	0	0	0	100.00%
Pred. Die lines	0	6	1	0	0	85.71%
Pred. Blowouts	0	2	7	0	0	77.78%
Pred. Accepted	0	0	0	8	0	100.00%
Pred. Shrinkage	0	0	0	0	8	100.00%
Class recall	100.00%	75.00 %	87.50 %	100.00%	100.00%	

k-NN models predicted the products having over weight of part, shrinkage and accepted ones as true positives for all cases. Prediction of die lines class resulted in 6 true positives, 1 false positive (die lines), 2 false negatives (blowouts). Whereas prediction of blowouts class has the outcome of 7 true positives, 1 false negative (blowouts), 2 false positives (die lines).

VI. VOTE

Vote is an ensemble method that makes use of major vote of inner models (learners) in its subprocess for classification. The training dataset is given as input to vote operator as shown in the process presented in Fig. 7. In the subprocess of vote operator three base learners (Rule Induction, k-Nearest Neighbors, Decision Tree) are trained on training dataset as shown in Fig.8. The vote model relies on major vote of base learners for classification significantly improved the prediction accuracy. Combination of these three base learners gives greater predictive performance of the model compared with combination of two base learners. The prediction accuracies of models for combination of three and two learners are presented in Table IV.

The model delivered through output (mod) port of vote operator is subsequently fed to Apply Model operator through input (mod) port. The test dataset is given as input to Apply Model operator through unlabelled (unl) port. The test dataset with known class labels is supplied as input to performance vector operator to evaluate the performance of Vote method.

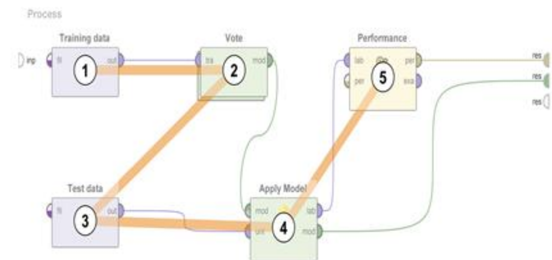


Fig. 7 RapidMiner process with VOTE method

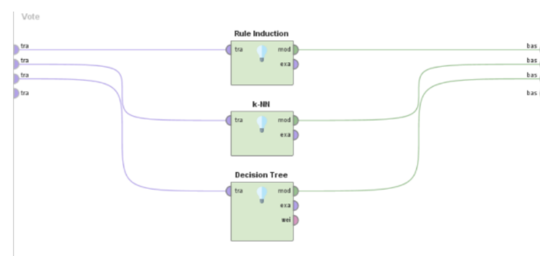


Fig. 8 Sub process of Vote

Table- IV: Prediction Accuracies for different combinations of base learners

S.No	Base learners			Prediction Accuracy
1	Rule Induction	K-Nearest Neighbors	Decision Tree	95%
2	Rule Induction	K-Nearest Neighbors	–	90%
3	–	K-Nearest Neighbors	Decision Tree	92.5%
4	Rule Induction	–	Decision Tree	87.5%

The confusion matrix representing class precision and recall, prediction accuracy of Vote method on test dataset is shown in Table V.

Table- V: Confusion matrix for VOTE method

Accuracy: 95%						
	True over weight of part	**True die lines**	**True blowouts**	**True accepted**	**True shrinkage**	**Class precision**
Pred. Over weight of part	8	2	0	0	0	80.00%
Pred. Die lines	0	6	0	0	0	100.00 %
Pred. Blowouts	0	0	8	0	0	100.00 %
Pred. Accepted	0	0	0	8	0	100.00 %
Pred. Shrinkage	0	0	0	0	8	100.00 %
Class recall	100.00%	75.00 %	100.00 %	100.00 %	100.00 %	

Retrieval Number J96760881019/2019©BEIESP
DOI: 10.35940/ijitee.J9676.0981119

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication

The vote method predicted the products having blowouts, shrinkage and accepted ones as true positives for all cases. Prediction of over weight of part class resulted in 8 true positives and 2 false positives whereas in case of die lines class outcome is 6 true positives and 2 false negatives.

VII. CONCLUSION

Data mining models are built on Extrusion Blow Molding dataset representing specific process conditions and the product. k-Nearest Neighbors, Rule Induction and Decision Tree techniques are used in building the models and predicting the quality of products. The predictive accuracies for the above techniques are evaluated as 92.5%, 80% and 80% respectively on the test dataset. k-NN model gives better predictive performance than Decision Tree and Rule Induction models. Among all the models, *Vote* gives higher predictive performance of 95% . Ensemble method (*Vote*) exhibited improved predictive accuracy than prediction accuracies of individual models. These models are applied in predicting product quality for different process configurations as well as finding the root causes of manufacturing defects. Ensemble method (Vote) shall be extended for other manufacturing processes to improve the performance in quality prediction.

REFERENCES

1. POLYMER TECHNOLOGY Reference Book 2016.
2. Huang, Geng-Qun, and Han-Xiong Huang. "Optimization of Parison Thickness in Extrusion Blow Molding Using Fuzzy Iterative Learning Control Algorithm." In ASME 2006 International Mechanical Engineering Congress and Exposition, pp. 691-696. American Society of Mechanical Engineers, 2006.
3. Huang, Han-Xiong, and Dong Li. "Parison Dimension Prediction in Extrusion Blow Molding Using Neural Network Approach: A New Strategy." In ASME 2007 International Mechanical Engineering Congress and Exposition, pp. 1953-1958. American Society of Mechanical Engineers, 2007.
4. Agarwal, Vikash, Jyoti Vimal, and Vedansh Chaturvedi. "Optimization of extrusion blow molding process parameters by grey relational analysis and taguchi method." (2012).
5. Yu, Jyh-Cheng, Tsung-Ren Hung, and Francis Thibault. "Performance optimization of extrusion blow molded parts using fuzzy neural-Taguchi method and genetic algorithm." In Proceedings of the 2002 ASME Design Engineering Technical Conferences, vol. 2,pp. 133-139. 2002.
6. Ramana, Dr EV, S. Sapthagiri, and P. Srinivas. "Data Mining Approach for Quality Prediction of Injection Molding Process Through Statistica SVM, KNN and GC & RT Techniques." International Journal of Mechanical Engineering and Technology 7, no. 6 (2016).
7. Han, Jiawei, Micheline Kamber, and Jian Pei. "Data mining concepts and techniques third edition." The Morgan Kaufmann Series in Data Management Systems (2011): 83-124.
   RapidMiner 8 Operator Reference Manual, RapidMiner GmbH, 2019.

ACKNOWLEDGEMENT

Authors would like to thank Mr. P. Srinivas, Associate Professor, Nalla Narasimha Reddy Group of Institutions Integrated Campus, Hyderabad for his support and valuable advice~~ses~~ in carrying out this work.

AUTHORS PROFILE

Vongur Ramulu is a PG~~post graduating~~ student in Advanced Manufacturing Systems (AMS) from VNR VJIET, Hyderabad. His area of research work is in data mining for Extrusion Blow Molding.

E. V. Ramana is working as a Professor, Department of Mechanical Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad. He received his first M.Tech degree in Energy Systems and second M.Tech degree in CAD/CAM in the years 1992, 1997 from JNT University, Hyderabad and Ph.D. from JNTUH. He possesses 28 years of experience in teaching, industry and research and published research papers in international Journals, national and international conferences. He has considerable experience in the development of software for mechanical engineering applications and guided many research projects. His active areas of research are Data Mining in Manufacturing and Machine Learning applications.

Kiran Kumar Namala has done doctoral program from IIT Delhi, post-graduation in production engineering. He published 28 research articles in the journals related to composite materials, energy, smart irrigation etc. He is a life member of ISTE.