Data Mining Approach for Quality Prediction and Improvement of Injection Molding Process

Dr. E.V.Ramana

Professor, Department of Mechanical Engineering
VNR Vignana Jyothi Institute of Engineering &Technology, Hyderabad, India
email:egvramana@gmail.com

Abstract-**Data mining techniques are gaining importance in extracting hidden relationships, associations and patterns from manufacturing process data for prediction of quality of the products. Data mining models are built on historical injection molding process dataset using Decision Tree, k-Nearest Neighbor (KNN), Support Vector Machine (SVM) and Probabilistic Neural Network (PNN) techniques to predict the quality of product for a specific setting of process parameters. These models are evaluated against test cases, and found that Decision Tree and k-NN models have less rate of misclassification than other models. The key factors that are causing flash, sink marks, short shot and burn marks in the product are identified by Decision Tree and presented as explicit rules.**

Keywords: Data Mining, SVM, PNN, k-NN, Injection Molding

I.   INTRODUCTION

Injection molding is the most popular mass production method of manufacturing plastic products. In this process, the plastic material fed in granular or solid form is heated to the molten state and forced into the mold under pressure. The plastic material takes the shape imparted by the mold after cooling and hardening and then ejected .The quality and productivity of products manufactured by injection molding depends on real time and close control of process variables such as injection pressure, barrel temperatures, mould temperature, hold pressure, injection time etc.

Multi-objective optimization, Design of Experiments (DOE), Principal Component Analysis (PCA) and Artificial Neural Network (ANN), Injection Molding Parameter Optimization and Simulation(IMPOS), Statistica Automated Neural Networks (SAAN), General Chi-square Automatic Interaction Detector (GCHAID), Association Rules, Statistica Support Vector Machine (SVM), General Classification and Regression Trees (GCRT) techniques are tried to optimization of injection molding process, quality prediction and improve the quality of products [1-7]. This paper proposes a data mining approach in the prediction of quality of product for a set of input process parameters by data mining models built through Decision Tree, k-Nearest Neighbor (k-NN), Support Vector Machine (SVM) and Probabilistic Neural Network (PNN) techniques and detecting key factors responsible for flash, sink marks, short shot and burn marks in the product using KNIME software.

II. DATASET

Injection molding process dataset related to the product - Cap made of Polypropylene (PP) comprises of 90 records has been collected. The data related to barrel temperatures at five zones; injection pressure and speed; injection, hold, refill, cooling and cycle times; mould temperature, clamping force, material grade and screw speed are collected. Dataset has been divided into training and test datasets by randomly assigning 70 records to training set and remaining 20 records to test set. Test dataset meant for evaluating prediction accuracies of the models is not used in training the models.

III. DECISION TREE

This technique has been chosen for classification task in this work. The training dataset is provided through XLS Reader Node to Decision Tree Learner Node to generate classification tree in the main memory. "Gini index" has been chosen as quality measure based on which split is calculated. Minimum numbers of records are selected as 2, which relates to stopping criteria. The work flow and options selected for decision tree are shown in Fig.1 and Fig.2 respectively. Decision Tree Predictor Node takes decision tree model and test dataset as inputs to predict label of the class as shown in the work flow in Fig.1. The simple decision tree view has been displayed in Fig.3.
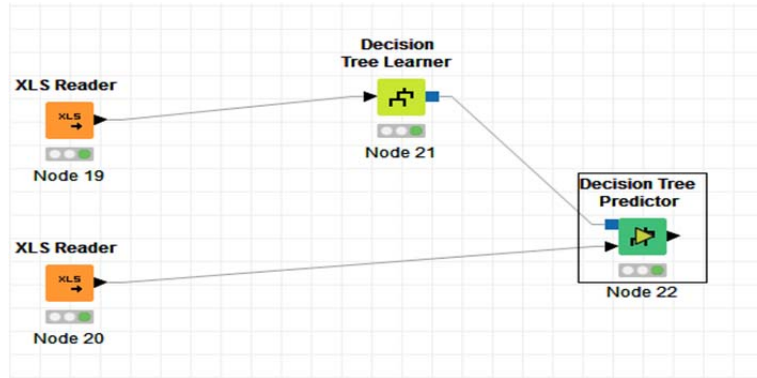
E.V.Ramana / International Journal of Engineering and Technology (IJET)
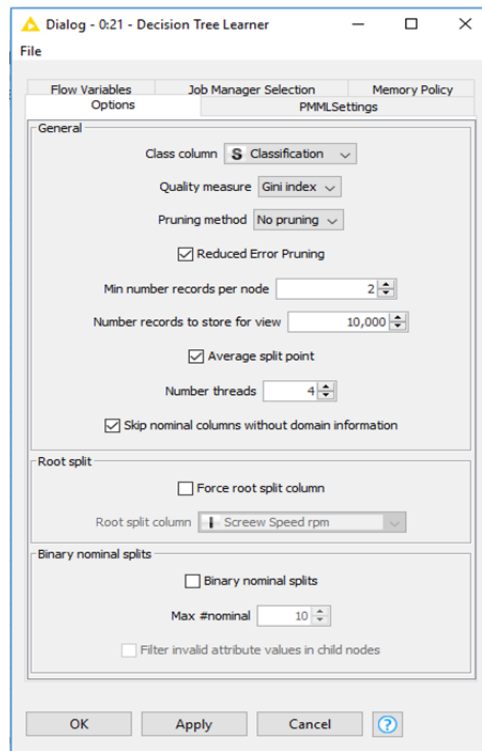

Fig.1. Work Flow of Decision Tree

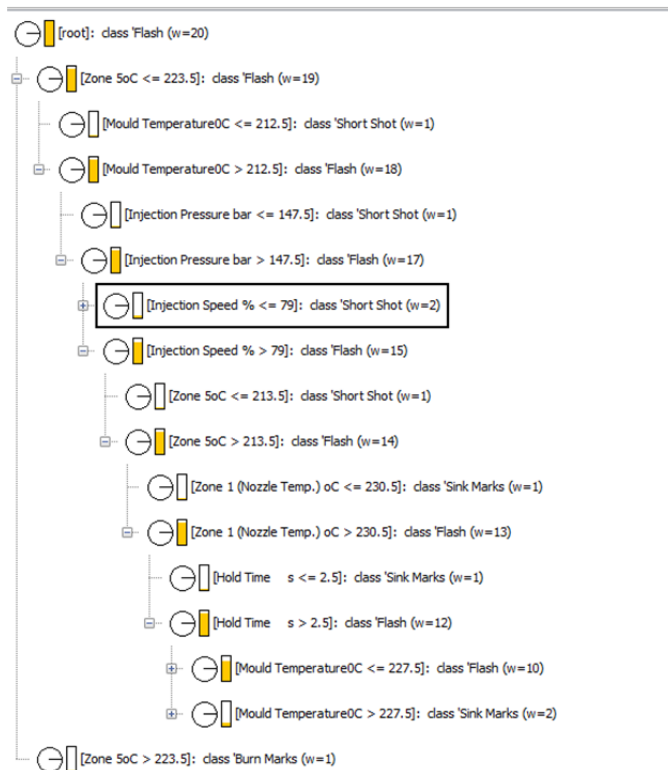Fig.2. Options for Decision Tree Learner Node

Fig.3. Simple Tree View

3.1 Classification by Decision Tree Model

The classified data by Decision Tree Predictor of *Knime* software and observed values are shown in Fig.4. The model predicted class labels of test cases correctly except in two cases, wherein short shot is wrongly predicted as sink marks and burn marks as flash.

IV. k-NEAREST NEIGHBOR

k- Nearest Neighbor algorithm has been used to classify the test dataset. Numeric columns along with Euclidean distance are used in the implementation of algorithm.   The work flow incorporating K-Nearest Neighbor and input nodes has been presented in Fig.4. Number of nearest neighbors is selected as 2 to classify new instance. Weight by distance option has been chosen so that closer neighbors have more influence on the class [8].
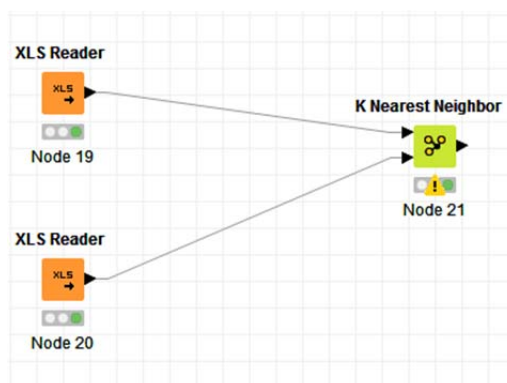

Fig.4. Work Flow of K-Nearest Neighbor Node

TABLE I. Classified Data – Decision Tree Predictor

Row ID	Injection Pressure (bar)	Injection Speed (%)	Zone 1 (Nozzle Temp.) °C	Zone 2 (°C)	Zone 3 (°C)	Zone 4 (°C)	Zone 5(°C)	Mould temp(°C)	Injection time (s)	Hold pressure (bar)	Hold Time (s)	Refill timing (s)	Cooling time (s)	Total cycle time (s)	Machine clamping force (ton)	Material Grade	Screw speed (rpm)	Classification	Prediction
1	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
2	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
3	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
4	148	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
5	150	80	233	233	228	225	222	225	6	58	1	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
6	150	80	200	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
7	150	70	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
8	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
9	145	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
10	150	80	233	220	215	205	195	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
11	150	76	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Sink Mark
12	150	80	233	233	228	225	222	200	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
13	150	80	245	245	235	230	225	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Burn Mark
14	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Flash
15	156	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
16	150	80	250	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
17	150	80	233	233	228	225	222	225	6	68	5	3.2	6	15	120	E7	70	Flash	Flash
18	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
19	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
20	150	80	233	233	228	225	222	250	6	58	3	3.2	6	15	120	E7	70	Flash	Flash

4.1 Classification by K-Nearest Neighbor

Actual and predicted values by KNN model for class label have been displayed in TABLE II. The model could predict correctly except for one test case wherein flash is wrongly predicted as burn marks.

V.   SUPPORT VECTOR MACHINE

The work flow involving support vector leaner and predictor nodes, and XLS reader nodes supplying training and test datasets has been shown in Fig.5. Support Vector Machine is trained by SVM Learner node on input data.RBF (Radial Basis Function) kernel has been chosen with γ value as 0.2. Overlapping penalty is selected as 2 through SVM dialog box as shown in Fig.6.

Fig.5. Work Flow of Support Vector Machine
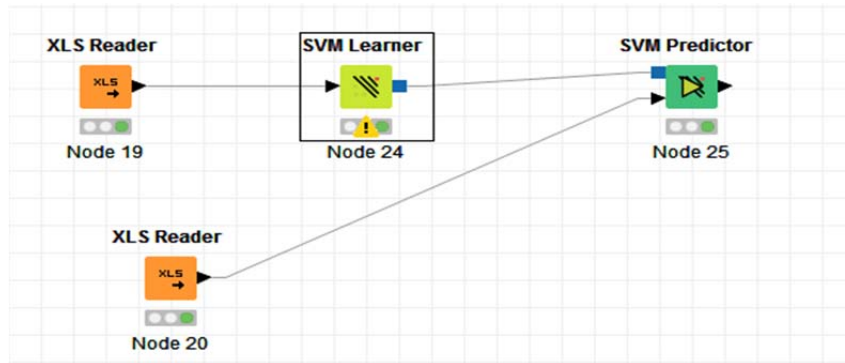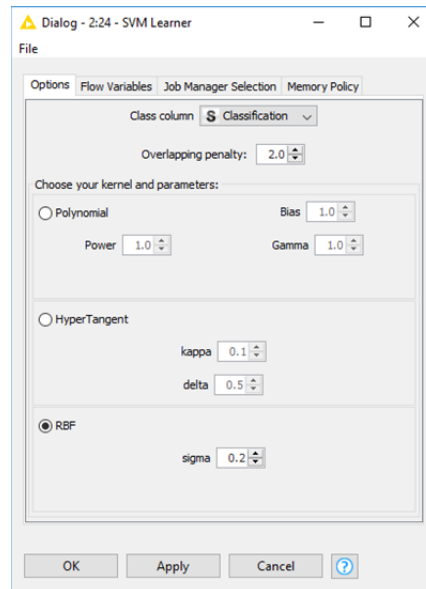

Fig.6 Options for SVM Learner Node

5.1 Classification by Support Vector Machine

The classification by SVM Predictor showing predicted class labels of test cases is shown in TABLE III. This model is able to classify the labels correctly of all test cases representing acceptable products, but misrepresented in some cases of products with sink marks, flash and burn marks.

TABLE II. Classified Data – K-Nearest Neighbor

Row ID	Injection Pressure (bar)	Injection Speed (%)	Zone 1 (Nozzle Temp.) °C	Zone 2 (°C)	Zone 3 (°C)	Zone 4 (°C)	Zone 5(°C)	Mould temp(°C)	Injection time (s)	Hold pressure (bar)	Hold Time (s)	Refill timing (s)	Cooling time (s)	Total cycle time (s)	Machine clamping force (ton)	Material Grade	Screw speed (rpm)	Classification	Prediction
1	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
2	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
3	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
4	148	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
5	150	80	233	233	228	225	222	225	6	58	1	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
6	150	80	200	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
7	150	70	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
8	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
9	145	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
10	150	80	233	220	215	205	195	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
11	150	76	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
12	150	80	233	233	228	225	222	200	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
13	150	80	245	245	235	230	225	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Burn Mark
14	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Burn Mark
15	156	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
16	150	80	250	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
17	150	80	233	233	228	225	222	225	6	68	5	3.2	6	15	120	E7	70	Flash	Flash
18	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
19	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Burn Mark
20	150	80	233	233	228	225	222	250	6	58	3	3.2	6	15	120	E7	70	Flash	Flash

TABLE III. Classified Data – SVM Predictor

Row ID	Injection Pressure (bar)	Injection Speed (%)	Zone 1 (Nozzle Temp.) °C	Zone 2 (°C)	Zone 3 (°C)	Zone 4 (°C)	Zone 5(°C)	Mould temp(°C)	Injection time (s)	Hold pressure (bar)	Hold Time (s)	Refill timing (s)	Cooling time (s)	Total cycle time (s)	Machine clamping force (ton)	Material Grade	Screw speed (rpm)	Classification	Prediction
1	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
2	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
3	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
4	148	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
5	150	80	233	233	228	225	222	225	6	58	1	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
6	150	80	200	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Flash
7	150	70	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
8	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
9	145	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
10	150	80	233	220	215	205	195	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Flash
11	150	76	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Sink Mark
12	150	80	233	233	228	225	222	200	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
13	150	80	245	245	235	230	225	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Burn Mark
14	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Burn Mark
15	156	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
16	150	80	250	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
17	150	80	233	233	228	225	222	225	6	68	5	3.2	6	15	120	E7	70	Flash	Flash
18	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
19	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Burn Mark
20	150	80	233	233	228	225	222	250	6	58	3	3.2	6	15	120	E7	70	Flash	Flash

VI. NEURAL NETWORK

PNN Learner trains Probabilistic Neural Network (PNN) using DDA (Dynamic Decay Adjustment) method on training dataset. The underlying Constructive Training of Probabilistic Algorithm generates rules based on numeric data. Each rule generated by PNN Node is defined by Gaussian function and is adjusted to avoid conflicts with rules of different classes. PNN model and test dataset are given as input to PNN Predictor Node to predict the class label. The workflow showing PNN Learner and Predictor nodes along with XLS Reader nodes are shown in Fig.7.
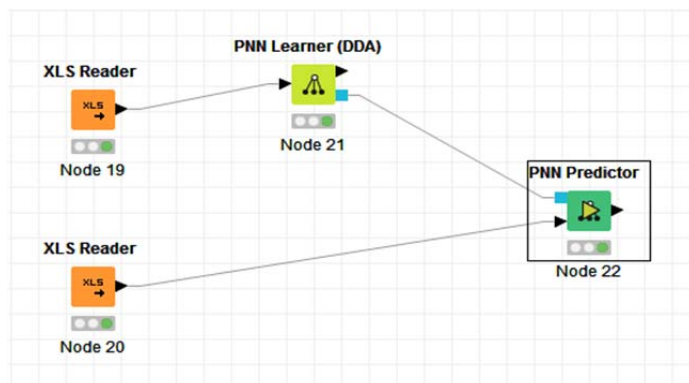

Fig.7. Workflow of Neural Network

6.1 Classification by Neural Network Model

Shrink after commit option has been selected to avoid conflicts with other rules of different classes. Use class with coverage option is selected to ensure the maximum degree of coverage of target columns during training [8-9]. The options dialog box for PNN Learner node is shown in Fig.8.
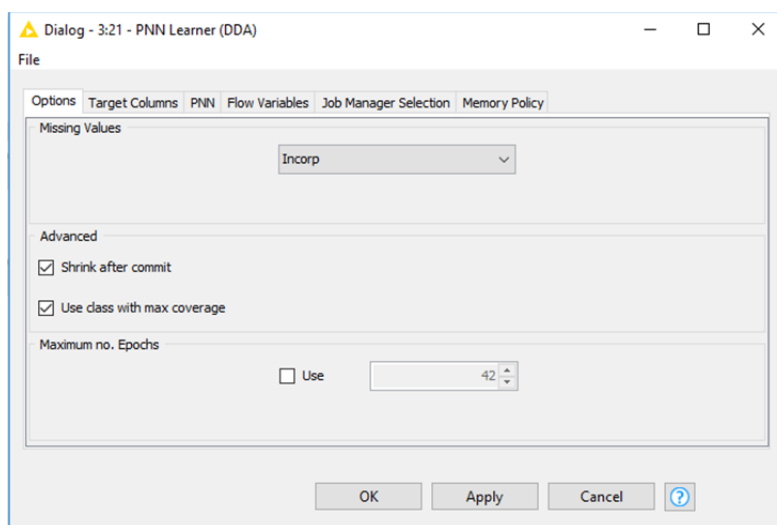

Fig.8. Options Dialog Box of PNN Learner

The predicted data by PNN Predictor for class labels of test cases has been presented in TABLE IV. The learner statistics produced by PNN learner is given in Fig.9. This model is not able to classify the labels of most of the test cases representing acceptable products and products with sink marks, flash, and burn marks.

TABLE IV. Predicted Data – PNN Predictor

| Row ID	Injection Pressure (bar)	Injection Speed (%)	Zone 1 (Nozzle Temp.) °C	Zone 2 (°C)	Zone 3 (°C)	Zone 4 (°C)	Zone 5(°C)	Mould temp(°C)	Injection time (s)	Hold pressure (bar)	Hold Time (s)	Refill timing (s)	Cooling time (s)	Total cycle time (s)	Machine clamping force (ton)	Material Grade	Screw speed (rpm)	Classification	Prediction
1	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
2	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
3	152	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
4	148	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Acceptable	Acceptable
5	150	80	233	233	228	225	222	225	6	58	1	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
6	150	80	200	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Short Shot
7	150	70	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Short Shot
8	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Sink Mark	Sink Mark
9	145	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
10	150	80	233	220	215	205	195	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
11	150	76	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
12	150	80	233	233	228	225	222	200	6	58	3	3.2	6	15	120	E7	70	Short Shot	Short Shot
13	150	80	245	245	235	230	225	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Burn Mark
14	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Burn Mark	Short Shot
15	156	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
16	150	80	250	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
17	150	80	233	233	228	225	222	225	6	68	5	3.2	6	15	120	E7	70	Flash	Flash
18	150	80	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Short Shot
19	150	95	233	233	228	225	222	225	6	58	3	3.2	6	15	120	E7	70	Flash	Short Shot
20	150	80	233	233	228	225	222	250	6	58	3	3.2	6	15	120	E7	70	Flash	Flash
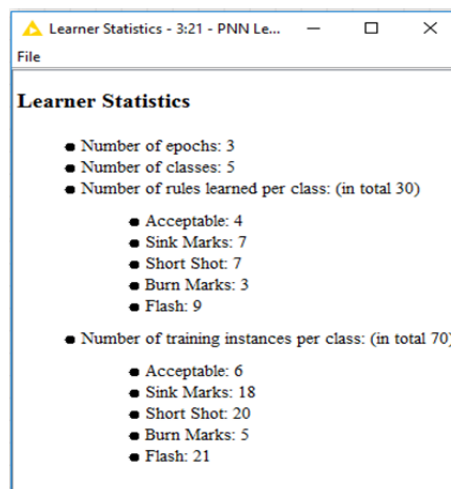
Fig.9. Learner Statistics– PNN Learner

VII. CONCLUSION

The data mining models that are built on the injection molding dataset for analysis by *Knime* software shall be used in predicting acceptable products and products rejected due to short shot, flash, sink and burn marks. Data mining models are built on dataset by applying Decision Tree, K-Nearest Neighbor, Support Vector Machine and Probabilistic Neural Network (PNN) algorithms.

Prediction accuracies of these Decision Tree and KNN models are found to be satisfactory with misclassification in two cases and one case out of 20 test cases respectively by each model. But prediction accuracies of SVM and PNN models are not satisfactory with misclassification instances of four and eight respectively out of 20 by each model. Sink marks are caused by low hold time ($<= 2.5$ s), nozzle temperature ($<= 230.5^oC$), and high molding temperature ($> 227.5^oC$ and $<=242.5^oC$). High barrel temperature at Zone 5 ($> 223.5^oC$) is resulting in burn marks. Short shot is caused by low injection pressure ($<=147.5$ bar), injection speed ($<= 73\%$), mould temperature ($<= 212.5^oC$) and barrel temperature at Zone 5 temperature ($<= 213.5^oC$).

High injection pressure (>154 bar), nozzle temperature ($> 230.5^oC$), hold time (> 5 s), injection speed ($< 93.5\%$) and mould temperature ($<= 227.5^oC$ and $> 242.5^oC$) and low clamping force ($<=117.5$ ton) are responsible for flash to occur. These are few causes mainly responsible for the above mentioned defects to occur in the products. These models are used in avoiding the production of defective products by eliminating causes of defects and thereby improving the quality of products.

REFERENCES

[1] Park, Hong Seok, and Trung Thanh Nguyen, "Optimization of injection molding process for car fender in consideration of energy efficiency and product quality." Journal of Computational Design and Engineering, 1(4), pp256-265, 2014.
[2] Seaman, Christopher M., Alan A. Desrochers, and George F. List. "Multiobjective optimization of a plastic injection molding process." IEEE Transactions on Control Systems Technology, Vol.2, pp 157-168, 1994.
[3] Lin, Tony, and Belli Chananda. "Quality improvement of an injection-molded product using design of experiments: a case study." Quality engineering, Vol.16, pp99-104, 2003.
[4] Costa, N., and B. Ribeiro. "Artificial neural networks for data modelling of a plastic injection moulding process." in Neural Information Processing, 1999. Proceedings. ICONIP'99. 6th International Conference, Vol. 3. IEEE, pp1081-1087.
[5] Bai, L., L. Gong, and S. Chen, "IMPOS: A Method and System for Injection Molding Optimization." in Industrial Electronics and Applications, 2006 1ST IEEE Conference, IEEE, pp1-5.
[6] Dr. E V Ramana, S Sapthagiri and P Srinivas, 2016, " Data Mining Approach For Quality Prediction Of Injection Molding Process Through Statistica SVM, KNN and GC & RT Techniques", International Journal Of Mechanical Engineering & Technology,Vol.7,pp 22-30,2016.
[7] Dr. E V Ramana, S Sapthagiri and P Srinivas, "Data Mining Approach for Quality Prediction and Improvement of Injection Molding Process through SANN, GCHAID and Association Rules", International Journal of Mechanical Engineering and Technology, Vol.7, pp31–40, 2016.
[8] KNIME Quickstart Guide, KNIME GmbH
[9] Jiawei Han, Micheline Kamber, Han J,, Jian Pei, 3rd ed., Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, 2011.

AUTHOR PROFILE

Dr. E. V. Ramana received his Ph.D degree in Mechanical Engineering in 2015 from JNTUH, Hyderabad. He received his M.Tech degree in Energy Systems and second M.Tech degree in CAD/CAM in the years 1992 and 1997 from JNT University, Hyderabad and P.G.D.C.A from Institute of Public Enterprise, Hyderabad. He possesses 20 years of experience in teaching and 6 years in Industry and Research and published many research papers in international journals, national and international conferences. He is currently working as a Professor of Mechanical Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad and his active areas of research include Data Mining in Manufacturing, CAD/CAM, and automation.