

Prediction of dot gain in flexographic color printing using machine learning

ABSTRACT

This work focuses on using machine learning algorithms in the prediction of dot gain related to flexographic process color printing. The way these advanced aspects of machine learning techniques are applied can revolutionize the various uses of printing technology. The machine learning techniques can be used to a wide range of applications since they adhere to dynamic programming methodology and computational learning theory. The machine learning algorithms can generate a trained input dataset framework, allowing them to make logical and dynamic predictions and judgments based on input data. Two grades of paper substrates with varying surface textures, two levels of anilox screen rulings, and a total of 100 steps of halftone square dot percentages with 4% intervals for each process colors are selected as the experimental process variables. An algorithm for evaluating a flexographic print output response, known as Dot Gain was generated using the Python machine learning technique. For data analysis and performance evaluation, machine learning techniques such as linear regression, decision tree, random forest regression, XG (Extreme Gradient) boost regression, SVM (Support Vector Machine) regression and neural network algorithms were used. The findings of this research work demonstrate that, out of all the machine learning algorithms used in this investigation, neural network methods had the highest accuracy. The accuracy of the neural network algorithm is 96.43, 98.32, 97.01 & 95.30 respectively in the prediction of dot gain for cyan, magenta, yellow and black.

KEY WORDS

flexography, dot gain, machine learning, regression, neural network

Soumen Basak¹ 
Saritha P.C² 
Alenrex Maity³ 
Kanai Chandra Paul¹ 

¹ Jadavpur University,
Department of Printing
Engineering, Kolkata, India

² Institute of Printing Technology
and Government Polytechnic
College, Department of Printing
Technology, Kerala, India

³ Jadavpur University,
Department of Information
Technology, Kolkata, India

Corresponding author:
Soumen Basak
e-mail:
soumen.basak@jadavpuruniversity.in

First received: 2.1.2025.

Revised: 7.4.2025.

Accepted: 5.5.2025.

Introduction

The printing background

A rotary relief image carrier that is appropriately flexible is used in the direct printing process known as flexographic printing to transfer printing ink to a range of substrates, including paper, paperboards, plastic films, foils and more (Abusaq et al., 2023).

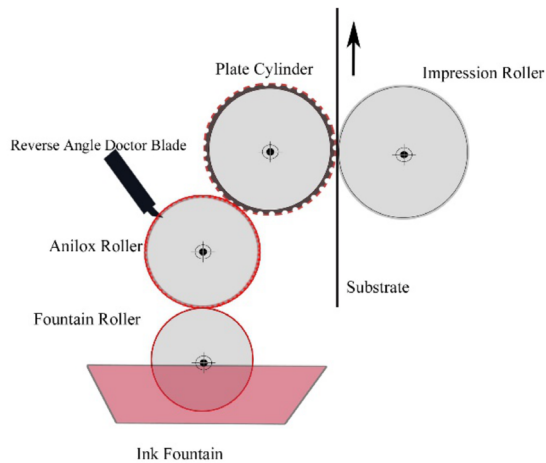
While using the flexible relief printing plates, a very small amount of printing pressure at the printing nip is sufficient to transfer the ink from the plate onto

the substrates which have a variety of surface textures (Joshi, 2022). The anilox roller cell engraving requirements played a significant role in flexographic printing by dictating the amount of ink that is needed to be applied to the substrate. Consequently, this has an impact on the finished print's optical density and halftone reproduction (Bould et al., 2010). Furthermore, the optimum outcome-oriented printing practices in flexographic presses are ensured by the anilox roller cell characteristics (Dendge, 2023).

Smoother paper can increase print density, but an increase in substrate surface roughness can also result in a sudden decrease in print density (Theohari et al., 2014).

The dispersion and absorption of ink, which is influenced by the paper's inherent tone, roughness, and porous surface, affects print quality. The smoothness or roughness of the paper directly affects how much the ink diffuses across the print (Tomasegovic et al., 2021).

The dot gain or Tonal Value Increase (TVI) indicates the increase of dot size in the final print than that of the original dot size given at the input. The Print Contrast indicates the accuracy of halftone reproduction at the shadow areas of a print (Dharavath, Bensen & Gaddam, 2005).



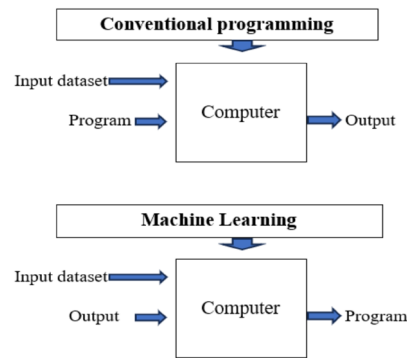
» **Figure 1:** Schematic diagram of flexographic printing unit

The Machine Learning

A branch of artificial intelligence and computer science known as "machine learning" is concerned with simulating human learning processes and gradually improving their accuracy via the use of data and algorithms (Chandel et al., 2022). Alan Turing originally put out the ideas of machine learning in the 1950s.

Machine learning helps computers to draw lessons from their prior experiences and, when needed, form their own opinions. The main objective of machine learning is to create computer programs in advance without requiring human intervention. Any computer can learn by using patterns, predictions, input and prior experience. This allows the machine to make decisions on its own without human intervention and get the right outcomes (Sisodia & Seth, 2022).

Through the use of machine learning techniques, computers are trained to carry out ordinary chores with ease and naturalness. As seen in Figure 2, raw data is fed into a machine learning model, which analyzes it and forecasts the result based on its comprehension of the incoming data (Sodhi, Awasthi & Sharma, 2019).



» **Figure 2:** Conventional Programming vs Machine Learning

As illustrated in Figure 3, there are four common methods for training machine learning algorithms. Machine learning is a collection of algorithms that are learned from data and/or experiences rather than being purposefully coded. Each task necessitates a unique set of algorithms and these algorithms find patterns to complete specific tasks (Çelik & Altunaydin, 2018). They are: Supervised learning, Unsupervised Learning, Reinforcement learning and Semi-supervised Learning.

Some Machine Learning Algorithms

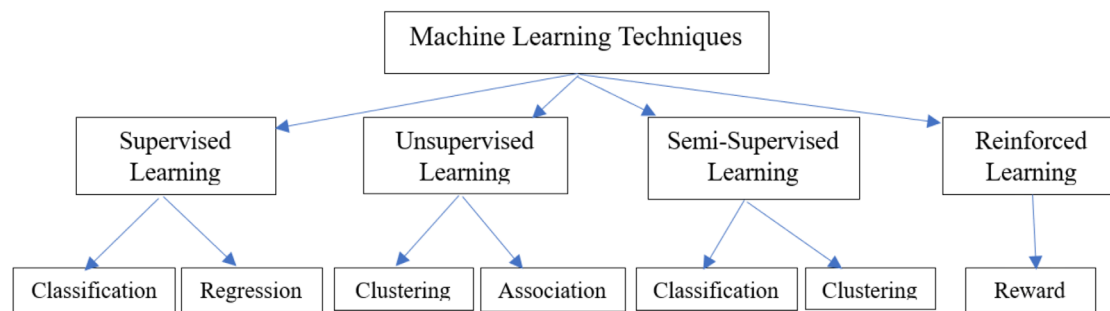
Algorithms are used in the development and instruction of machine learning models. These are the kind of actions that a machine learning model learns via observation and then applies the same technique. Machine learning algorithms have the ability to improve themselves through training (Jin, 2020). Linear regression, decision tree regression, random forest regression, extreme gradient boosting or XG Boost, support vector machine or SVM regression, neural network learning techniques are some of the frequently used algorithms.

Linear regression

A relationship between the dependent variable and one or more independent variables—known as the regression line—is established by linear regression using the best-fitting straight line. The value of the target variable can be predicted using the regression equation and the given predictor variable or variables. Multiple linear regression models are an extension of simple linear regression, where a response variable functions as a linear function of two or more predictor variables.

Decision tree regression

The decision tree sorts the tree from the root to a few leaf nodes in order to classify the instances. By analysing the attribute defined by that node, instances are categorized, beginning at the root node of the tree and proceeding along the branch that corresponds to the attribute value.



» **Figure 3:** *The machine learning techniques*

Random forest regression

It is a technique that fits many decision tree classifiers simultaneously on different subsamples of the dataset. The final outcome is decided by averages or majority vote. The outcome is an improvement in control and forecast accuracy while lowering the problem of over-fitting. It creates a series of decision trees with controlled variance by combining random feature selection with bootstrap aggregation or bagging.

Extreme gradient boosting regression

Extreme Gradient Boosting (XG Boost), a kind of gradient boosting, takes into account more accurate approximations. It calculates the loss function's second-order gradients to minimize loss and accomplish advanced regularization. This reduces over-fitting and improves the model's performance and generalization. XG Boost can swiftly evaluate results and process massive datasets effectively.

Support vector machine regression

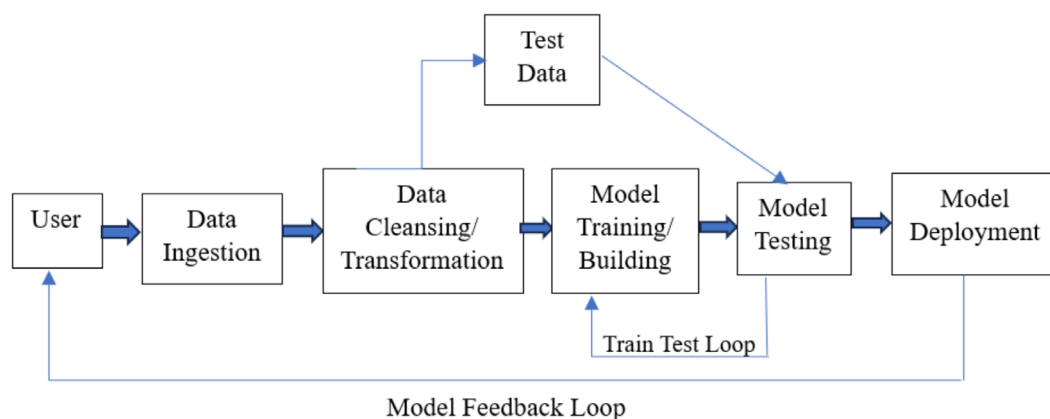
The Support Vector Machine (SVM) algorithm will simultaneously analyze the data to be processed in order to improve the data information. It collects many sets of analysis samples during the

analysis process and determines the sample data of the boundary value in order to improve the scientific quality of the final data analysis results.

Neural network algorithms

The phrase "artificial neural network" refers to a system that simulates human information transmission by classifying different facts into a single neuron and connecting those neurons online to carry out complex memory tasks. However, this dynamic data analysis process forms the basis of the artificial neural network algorithm. A layered structure consisting of multiple interconnected nodes makes up the algorithm. These nodes represent the real neurons in the human brain. Each unique neuron has a high degree of authenticity and the data can complete the process of external output. The human body moves forward, stops and then takes off in a manner similar to this. After the weighting coefficient has been selected, setting the output threshold will facilitate computation and data processing. This makes the process of numerical analysis more systematic overall. This adaptive strategy allows the machines to keep improving by learning from their mistakes.

The machine learning system



» **Figure 4:** *The machine learning workflow: The input dataset is processed and converts into valid format that best fits the needs of the machine learning model being employed*

The fundamental workflow related to machine learning techniques is depicted in Figure 4. The machine learning method that creates a trained model is used in the training phase. The accuracy of the trained model is then repeatedly assessed using the test data. The model is then deployed so that the application can use it after it has reached the required level of accuracy. Establishing a machine learning system to forecast dot gain in the flexographic color printing is the aim study.

Methods

Experiment

Different paper samples having varying paper roughness were printed by flexographic printing varying different input parameters such as anilox screen rulings. Effects of varying dot percentages on different colored dot gain were also considered. Prints were taken in four process colors separately on two sets of different paper samples. Table 1 displays the three input variables that were selected for the research. In the present research, Dot gain was used as the response variable.

Table 1

Parameters in the study

Factors	Unit	Sym- bol	Levels	Response
Anilox Ruling	Lines per inch (lpi)	A	900, 1200	Dot Gain
Paper Roughness	Millilitre per minute (ml/min)	B	76, 109	
Dot Percentage	Percentage (%)	C	4% to 100 % with 4% increments	
Halftone dot shape	-	D	Square Shaped AM dot	

The anilox roller screen rulings used in the work had lines per inch (lpi) of 900 and 1200 respectively. The chosen roughness levels for the paper were 76 and 109 milliliters per minute. The roughness was measured using Roughness tester (Bendtsen type) following IS 1060 (Part 5/ Sec 20) RA2018. The engraved anilox roller cells were 60° hexagonal in form. The OMET LAB230 Iflex flexographic printing machine, which uses process color inks of Cyan (C), Magenta (M), Yellow (Y) and Black (K), was utilized for the printing. The inks used were solvent based liquid inks and have viscosity of 300 cp. A photopolymer plate (Manufacturer – Dupont, Type of the plate – Digital solid photopolymer, Plate production process – Laser engraving process) with a

magnetic backing and of 1.14 mm thickness was used as the printing image carrier. All of the parameters, including the ink characteristics (UV ink), room temperature (23° C), printing speed (35 m/min) and nip pressure (3 mm), were kept constant during the printing process.

Measurements and data collection

The print quality score was determined using the output response. As a result, the output parameters were quantitatively measured using the X-Rite Spectro Eye Spectro-densitometer. Visual examination of the printed area was done using a digital microscope, the LEICA-S8APO. The machine learning algorithms employed the readings of 100 print trials of input data sets for every color. It is expected that the split ratio of the training to test dataset is 7:3. The collected data is fed into the Python program, which processes it using machine learning techniques.

Results

The Tables 2 and 3 display the results of machine learning techniques for each of the C, M, Y, and K process colors. The evaluation instrument for comparison is the performance accuracy of different algorithms for every color. The capacity of the developed model to forecast the result based on the given input variables is referred to as performance accuracy.

Table 2

Splitting of dataset

Shape of data	(100, 4)
Shape of split data	The shape of X_train : (75, 3)
	The shape of X_test : (25, 3)
	The shape of Y_train : (75, 3)
	The shape of Y_test : (25, 3)

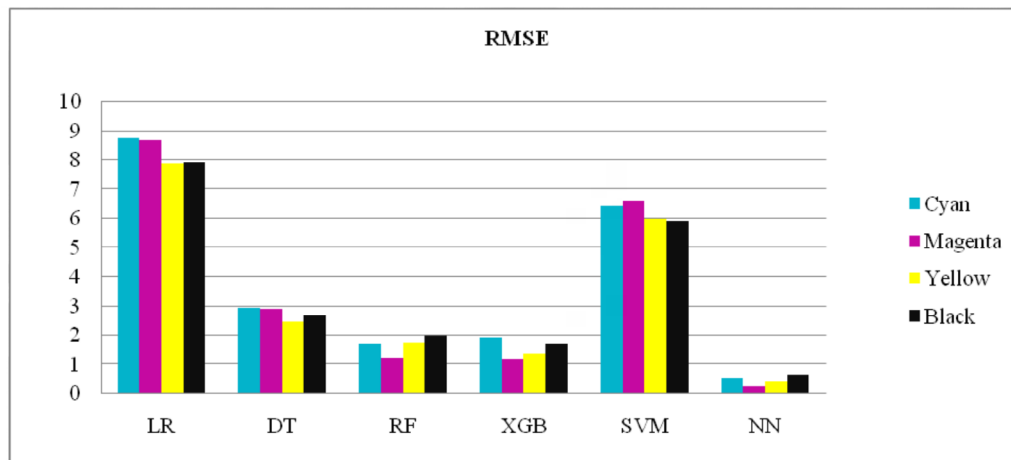
Figure 5 shows the Root Mean Square Error (RMSE) for the different machine learning algorithms. RMSE quantifies the differences between the predicted values and actual values. It has been observed that RMSE is least for Neural Network which implies it is best for this prediction.

Figure 6 represents bar chart of Mean Absolute Error (MAE) for the different machine learning algorithms adopted for prediction MAE measures the average absolute differences between the actual and the predicted values. It has been found that MAE value is the lowest for the Neural Network. This means that Neural Network is the best of all the machine learning algorithms considered for the prediction of dot gain in the present work

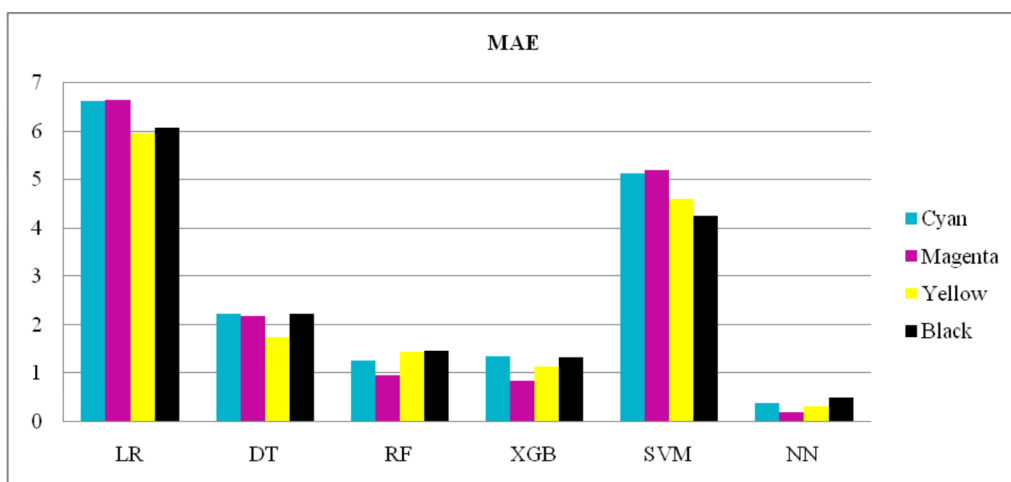
Figure 7 shows the bar chart of R² values for the different machine learning algorithms adopted for the prediction of dot gain.

It indicates the percentage of variation in the target variables that can be obtained by the independent variables. The maximum value of R^2 is 1. The higher the R^2 value, the better is the prediction. It has been found that R^2 value for the Neural Network method

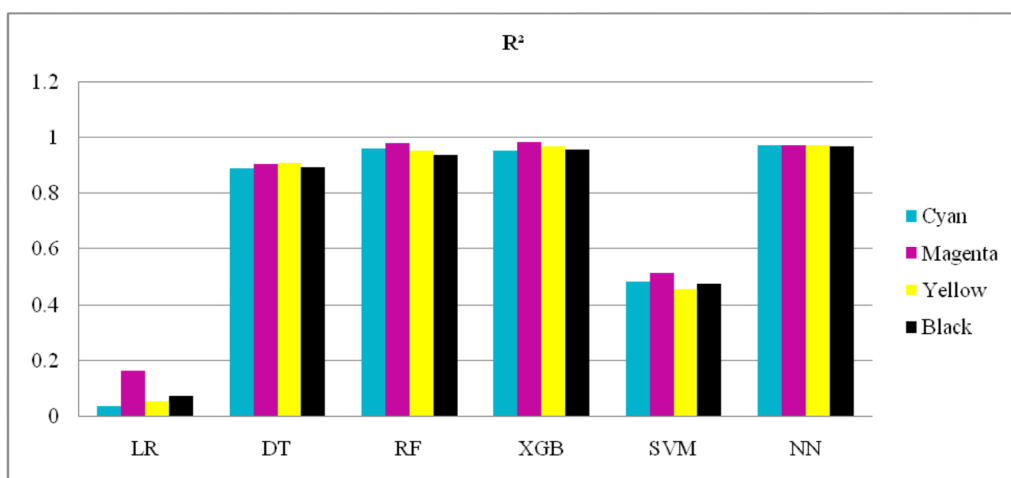
is maximum and near about 1 which implies that this method is the best for the prediction of dot gain in the current investigation. The values of accuracy of different machine learning algorithms given in Table 3 were decided by considering the values given in Figures 5-7.



» **Figure 5:** Bar chart of RMSE values for different machine learning algorithms



» **Figure 6:** Bar chart of MAE values for different machine learning algorithms



» **Figure 7:** Bar chart of R^2 score for different machine learning algorithms

Table 3

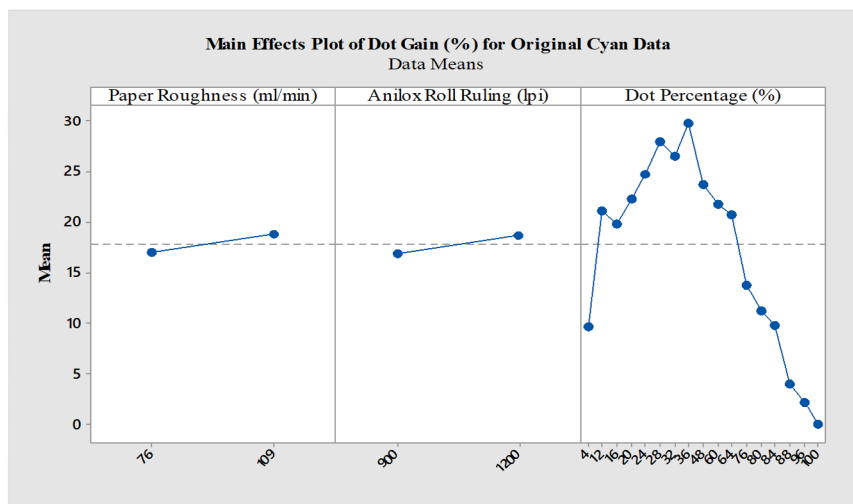
The prediction of Dot Gain with different algorithms

Sl. No.	Process Colors	Machine Learning Algorithms					
		Linear Regression (LR)	Decision Tree (DT)	Random Forest Regression (RF)	XG Boost Regression (XGB)	Support Vector Machine (SVM)	Neural Network (NN)
1.	Cyan	66.43	78.29	87.76	84.78	71.47	96.43
2.	Magenta	66.07	79.76	88.84	89.10	71.35	98.32
3.	Yellow	67.54	88.09	85.05	88.24	73.10	97.01
4.	Black	67.72	85.18	85.22	86.38	75.67	95.30

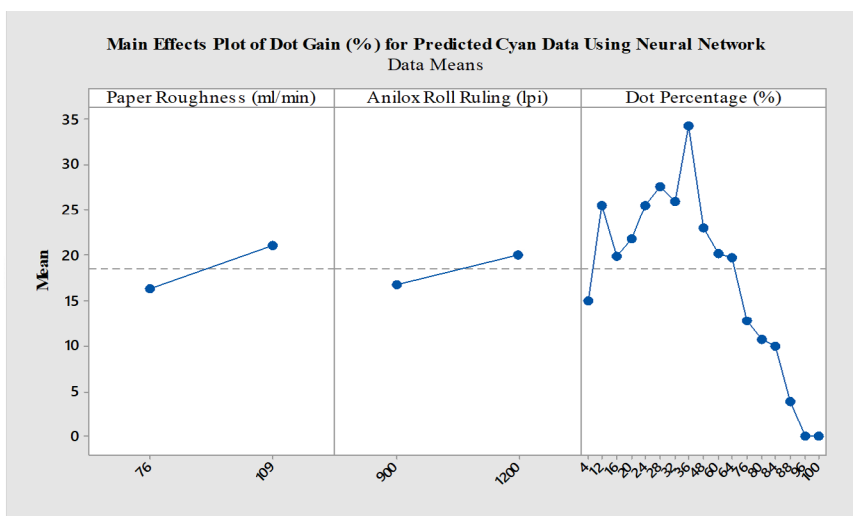
Graphical analysis

The graphical depictions that display the interplay between the various experimental factors that was identified during the analysis stage. The Figures 8-15 display the main effects plot of each experimental factors over

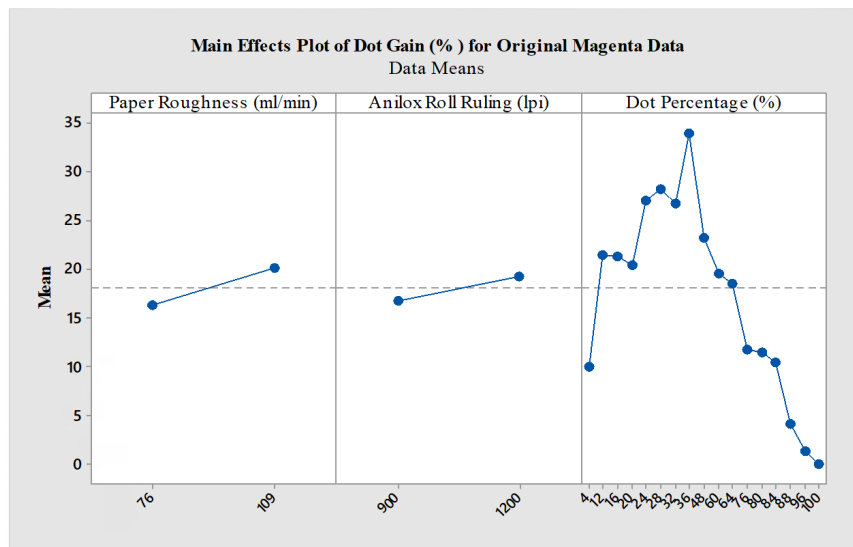
dot gain of original data and predicted data obtained by neural network and the Figures 16-23 displays the interaction effects of original experimental data as well as predicted data. The Figures 8-23 have been plotted using Minitab 17.



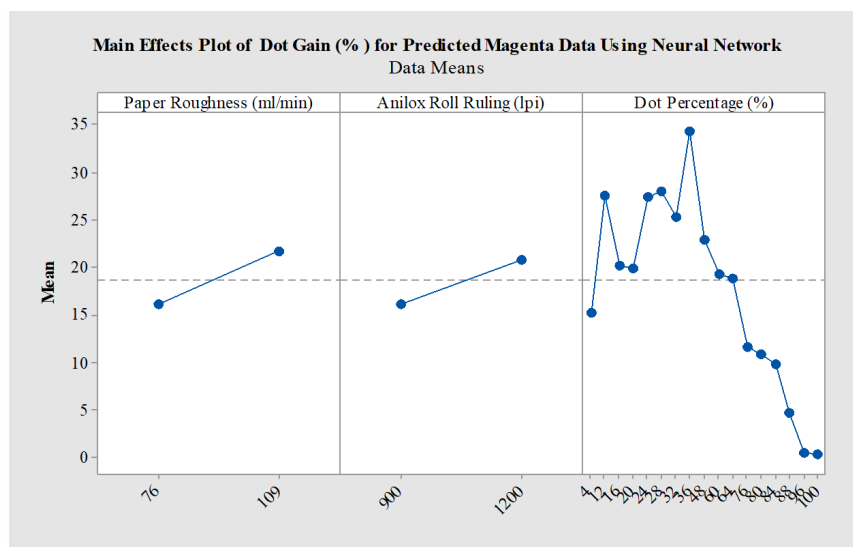
» **Figure 8:** Main effects plot of dot gain for original cyan data



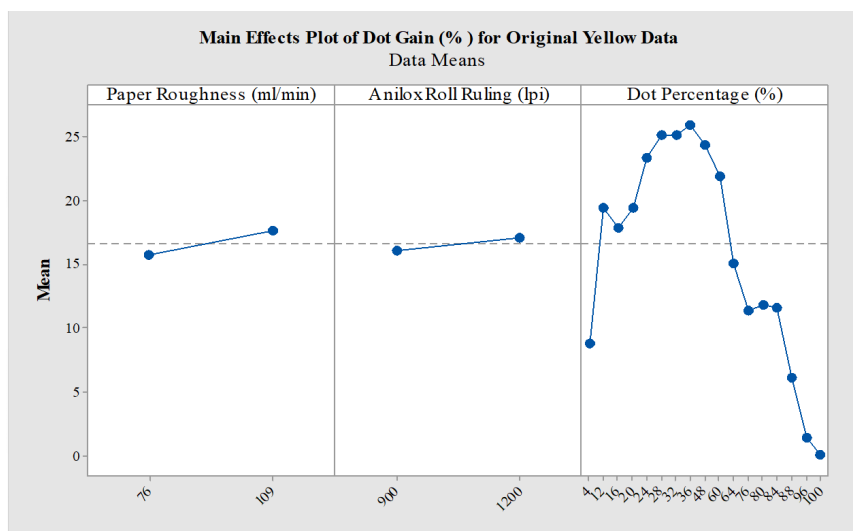
» **Figure 9:** Main effects plot of dot gain for predicted cyan data using neural network



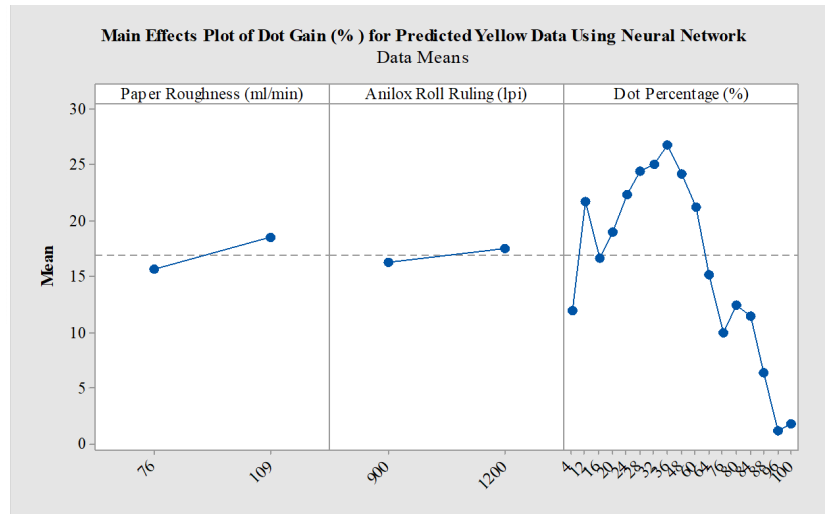
» **Figure 10:** Main effects plot of dot gain for original magenta data



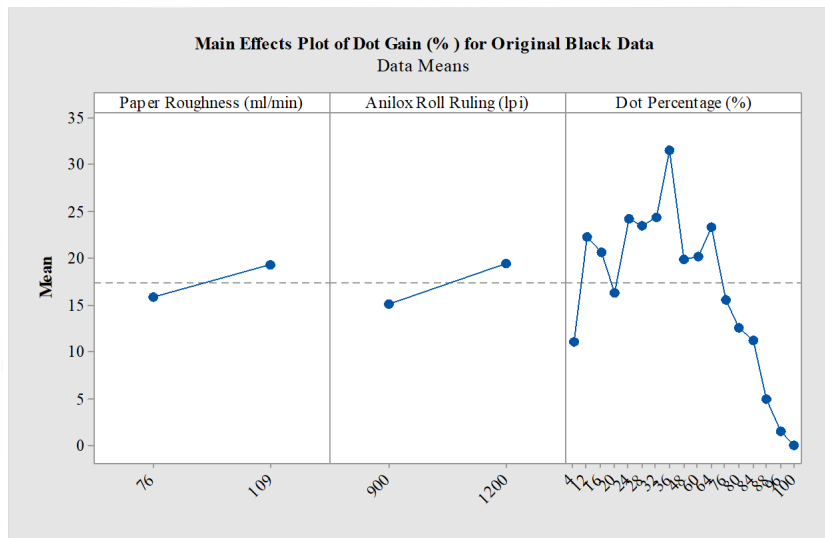
» **Figure 11:** Main effects plot of dot gain for predicted magenta data using neural network



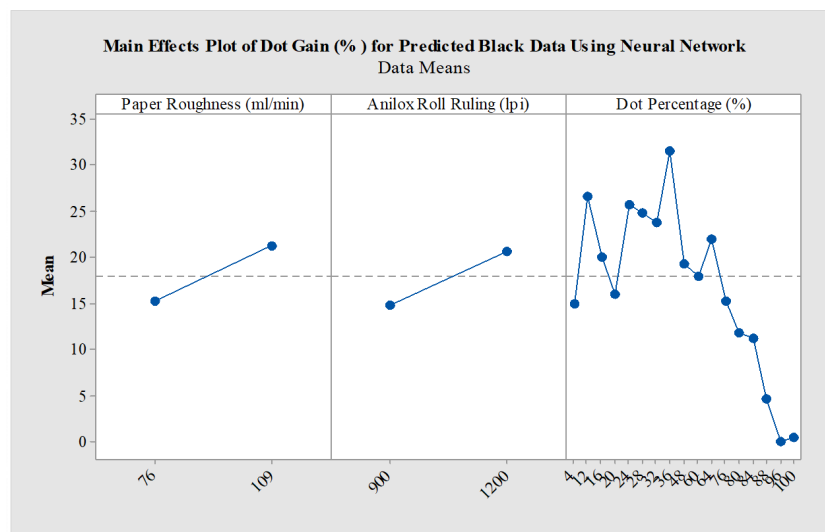
» **Figure 12:** Main effects plot of dot gain for original yellow data



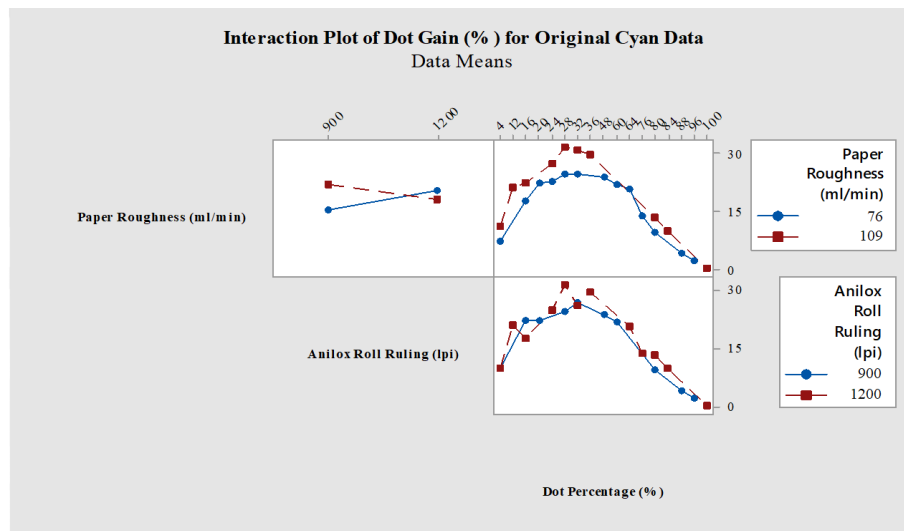
» **Figure 13:** Main effects plot of dot gain for predicted yellow data using neural network



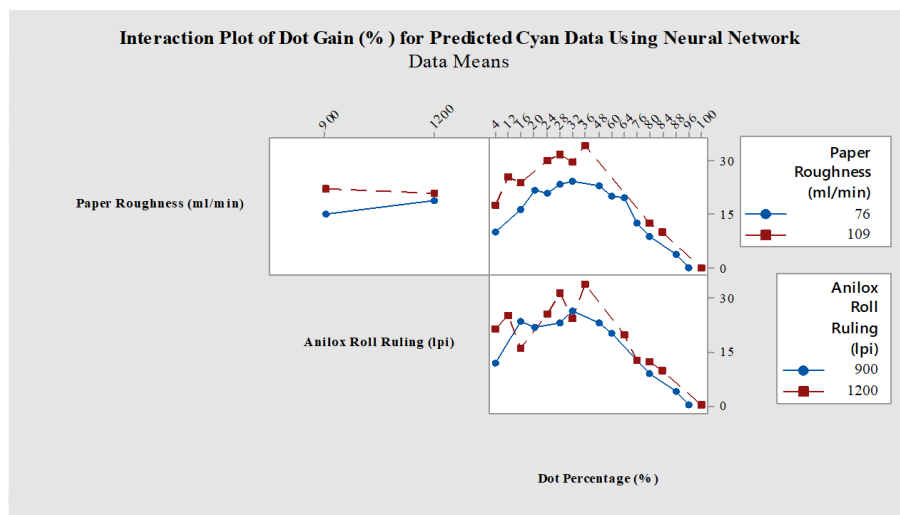
» **Figure 14:** Main effects plot of dot gain for original black data



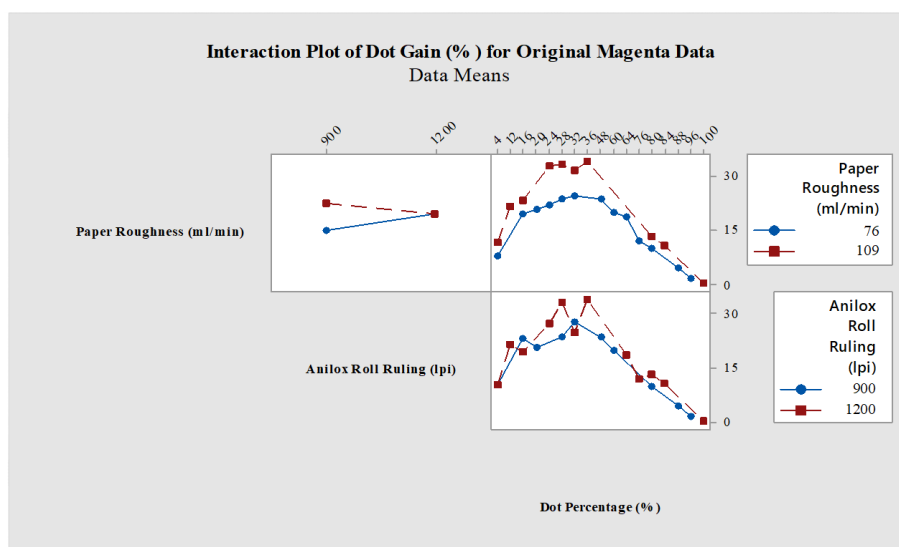
» **Figure 15:** Main effects plot of dot gain for predicted black data using neural network



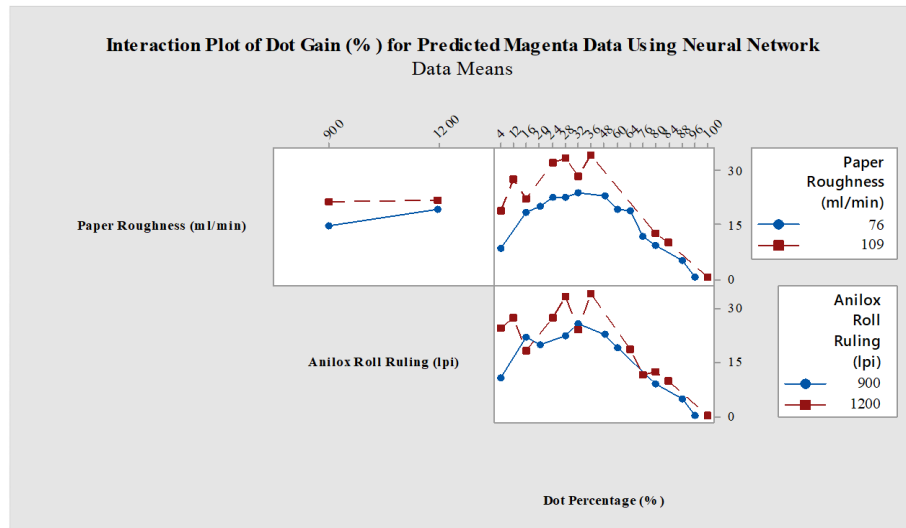
» **Figure 16:** Interaction plot of dot gain for original cyan data



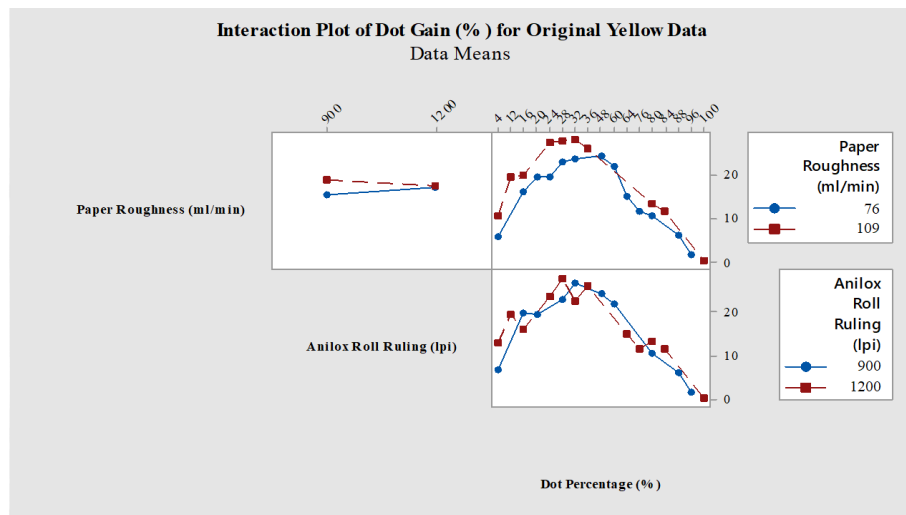
» **Figure 17:** Interaction plot of dot gain for predicted cyan data using neural network



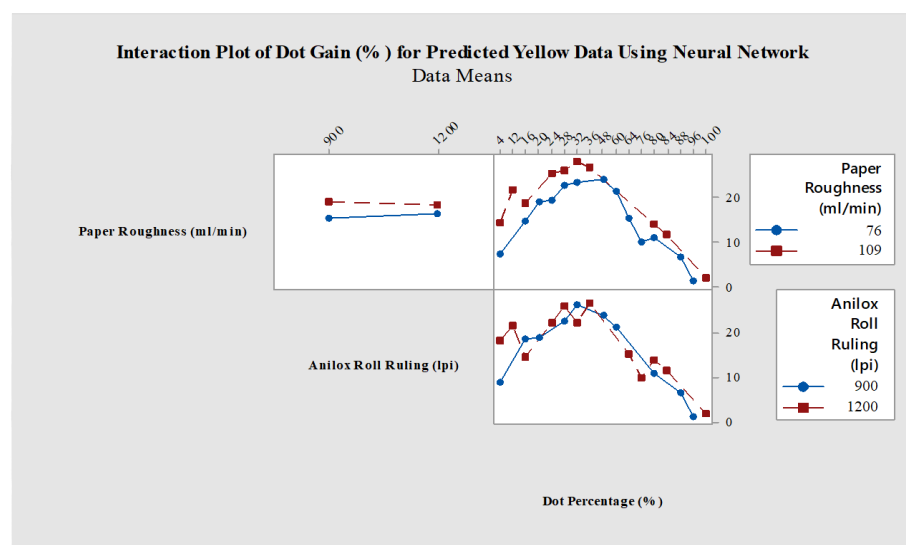
» **Figure 18:** Interaction plot of dot gain for original magenta data



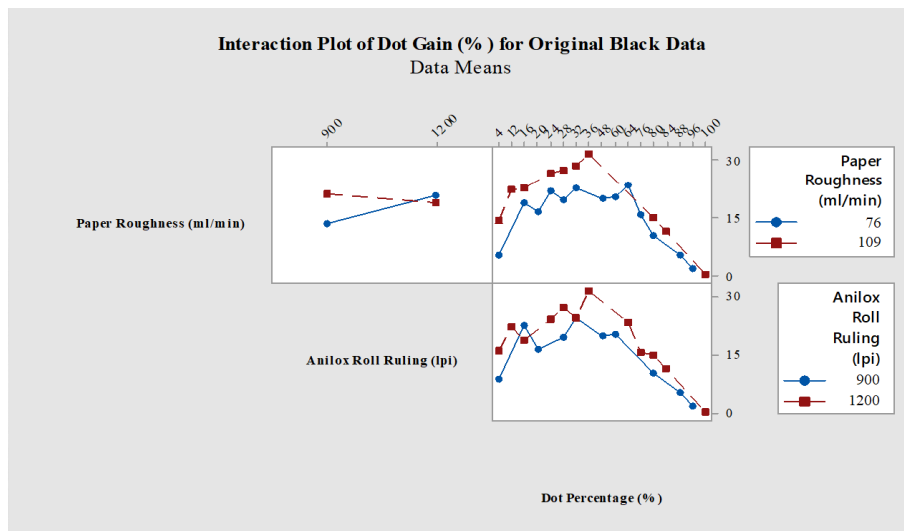
» **Figure 19:** Interaction plot of dot gain for predicted magenta data using neural network



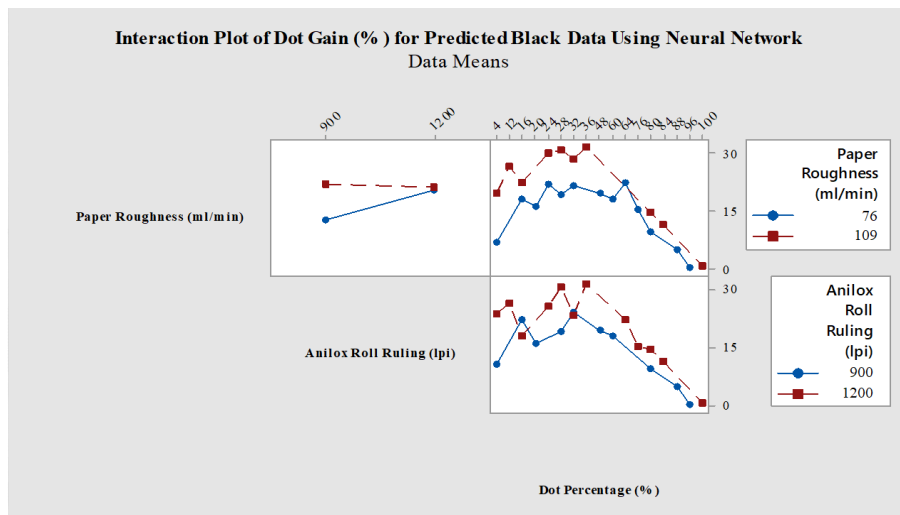
» **Figure 20:** Interaction plot of dot gain for original yellow data



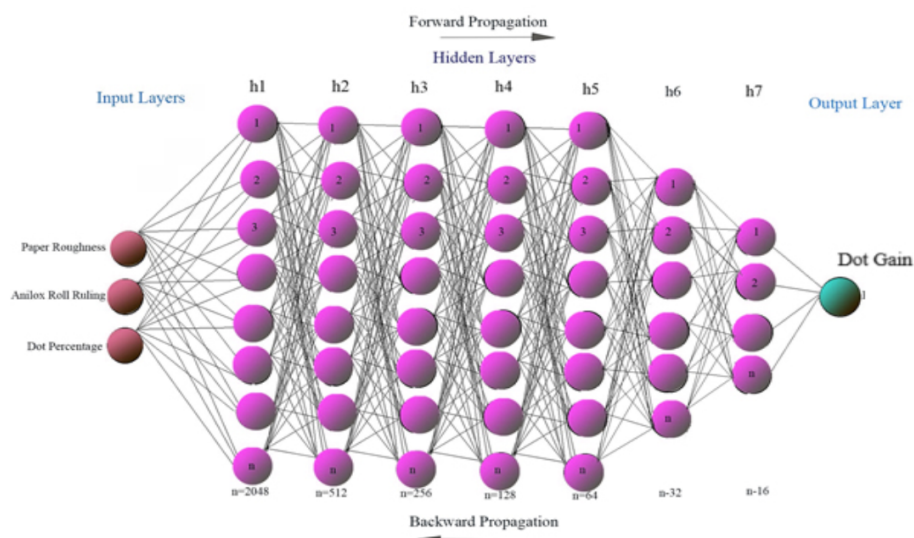
» **Figure 21:** Interaction plot of dot gain for predicted yellow data using neural network



» **Figure 22:** Interaction plot of dot gain for original black data



» **Figure 23:** Interaction plot of dot gain for predicted black data using neural network



» **Figure 24:** The Neural Network Architecture in the study

Discussions

In recent years, machine learning has rapidly grown inside the computing and data analysis framework, which typically allows programs to function intelligently. A few machine learning techniques have been applied in this work to evaluate the dot gain of four-color flexographic printing.

The summary of performance accuracy attained for several machine learning models in dot gain prediction is shown in Table 3. It is discovered through the execution of a Python program that the Neural Network algorithm provides a very notable method for estimating optical dot gain for cyan, magenta, yellow, and black colors. This demonstrates that neural networks outperform tested algorithms like SVM, XG boost regression, random forest, decision tree, and linear regression in the prediction process. With its unique properties, such as a somewhat higher number of hyperparameters, epochs, layers, etc., than other algorithms, the neural network performs exceptionally well.

The main effects plot of original data as well as predicted data for Cyan, Magenta, Yellow and Black reproductions with Dot Gain are shown from Figures 8-15 and the interaction plots of original data and predicted data are illustrated by the Figures 16-23. The results of these plots indicate that the dot gain is comparatively less with coated paper than uncoated paper. As the anilox line frequency increases the dot gain tends to decrease. The lower anilox screen frequency together with the uneven texture of uncoated paper accelerates the dot gain effects considerably.

Figures 8-15 shows the nature of the output response (dot gain) of main effects plot with respect to paper roughness and anilox roll ruling (lpi) is more or less similar in both the original data and predicted data. Interaction plots represented by Figures 16-23 of original data and predicted data obtained by neural networking machine learning shows the nature of the curve is almost similar. Figure 24 shows that the suggested design of the Neural Network algorithm in this study has three input layers, one output layer, seven hidden layers, and 5000 epochs. This method of operation enhanced its level of automation and general effectiveness.

Conclusion

The goal of this research is to develop a machine learning model for predicting dot gain effect associated with flexographic four-process color printing. This paper's primary sections are as follows: a theoretical review of the history of printing and machine learning, Python programming using some key machine learning algorithms, performance assessment and outcome analysis.

The different machine learning techniques are implemented and assessed using neural networks, XG boost regressors, SVM regressors, random forest regressors, decision trees and linear regression. The result shows that when compared to all other evaluated algorithms, Neural Network algorithm performs noticeably better due to the superior properties of neural networks.

This research may be extended by varying the different parameters such as dot shapes, anilox roller cell shapes, varying ink viscosity and also the printing speed.

Funding

The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Abusaq, Z., Zahoor, S., Habib, S. M., Rehman, M., Mahmood, J., Kanan, M. & Mustaq, T. R. (2023) Improving energy performance in flexographic printing process through lean and AI techniques: A Case Study. *Energies*. 16 (4), 1972. Available from: doi: 10.3390/en16041972
- Bould, C. D., Hamblyn, S. M., Gethin, T. D. & Claypole, C.T. (2010) Effect of impression pressure and anilox specification on solid and halftone density. *Journal of Engineering Manufacture*. 225 (5). Available from: doi: 10.1177/2041297510394072
- Çelik, Ö. & Altunaydin, S. S. (2018) A research on machine learning methods and its applications. *Journal of Educational Technology and Online Learning*. 1 (3), 25-40. Available from: doi: 10.31681/jetol.457046
- Chandel, M., Silakari, S., Pandey, R. & Sharma, S. (2022) A study on machine learning and Python's framework. *International Journal of Computer Sciences and Engineering*. 10 (5). Available from: doi: 10.26438/ijcse/v10i5.5864
- Dendge, R. R. (2023) Analysis of dot gain produced by interactions of flexographic plate and anilox roll screen frequencies. *Acta Graphica*. 31 (1), 35-44
- Dharavath, N. H., Bensen, M. T. & Gaddam, B. (2005) Analysis of print attributes of Amplitude Modulated (AM) vs. Frequency Modulated (FM) screening of multicolour offset printing. *Journal of Industrial Technology*. 21 (3).
- Jin, W. (2020) Research on machine learning and its algorithms and development. In: *5th International Conference on Intelligent Computing and Signal Processing: Journal of Physics Conference Series, ICSP, 20-22 March 2020, Suzhou, China*. Purpose-LED Publishing. Available from: doi: 10.1088/1742-6596/1544/1/012003
- Joshi, V. A. (2022) Optimization of flexo process parameters to reduce the overall manufactur-

- ing cost. *An International Journal of Optimization and Control: Theories & Applications*. 12 (1). Available from: doi: 10.11121/ijocta.2022.1137
- Sisodia, P. & Seth, B. (2022) An implementation on Python for data science and machine learning. *International Journal of Creative Research Thoughts*. 10 (3).
- Sodhi, P., Awasthi, N. & Sharma, V. (2019) Introduction to machine learning and its basic application in Python. In: *Proceedings of 10th International Conference on Digital Strategies for Organizational Success, 5-7 January 2019*. pp. 1354-1375. Available from: doi: 10.2139/ssrn.3323796
- Theohari, S., Fraggedakis, E., Tsimis, D., Tsigonias, M. & Mandis, D. (2014) Effect of paper properties on print quality by flexographic method. In: *46th Annual International Conference on Graphic Arts and Media Technology, Management and Education, 25-29 May 2014, Athens, Greece*. Athens, Hellenic Union of Graphic Arts and Media Technology Engineers.
- Tomašegović, T., Pibernik, J., Mahović Poljaček, S. & Madžar, A. (2021) Optimization of flexographic print properties on ecologically favorable paper substrates. *Journal of Graphic Engineering and Design*. 12 (1), 37-44. Available from: doi: 10.24867/JGED-2021-1-037



© 2025 Authors. Published by the University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license 4.0 Serbia (<https://creativecommons.org/licenses/by/4.0/deed.en>).