

Moisture content prediction model for pharmaceutical granules using machine learning techniques

Haftom Tekie¹, Tibebe Beshah², Fisha Hailelassie³, and Kibrom Gidey³

¹Sun Daero College

²Addis Ababa University College of Natural Sciences

³Debre Tabor University

October 19, 2020

Abstract

The aim of this study was to develop a prediction model and identifying relative important factors in the evaluation of moisture content of pharmaceutical granules using Artificial Neural Network (ANN) and Support Vector Machine (SVM) techniques. Optimal models of ANN and SVM models were developed and compared, utilizing Matlab16.0 as a software tool. The performance of the models is evaluated using a quantitative error metric; mean squared error (MSE), Regression(R) and Confusion Matrix (CM). This study reveals that the ANN model is an optimal model for predicting moisture content of pharmaceutical granules for datasets of APF. The model of ANN, with MSE of 0.016941 and classification accuracy of 98.7% is built and accepted as optimal model for predicting moisture content of pharmaceutical granules. Temperature, Loge-Mixer time, Initial moisture, air flow rate and drying time respectively are the most important factors in determining the moisture content of the granules.

1 Introduction

The most commonly used apparatus in the drying area of pharmaceutical production (manufacturing) is Fluid Bed Dryer (FBD). It is used for drying the granules to get desired moisture content in the tablet formulation granules required for impeccable compression of tablets. The solid particles will be puffed up and become slung in the air stream, if air is allowed to go through a bed of solid powdered material in the upward direction with the speed above the settling rate of the particles. This stage is called fluidized, as the solid bed seems like the boiling liquid. The application of hot air to fluidize the bed speeds up the drying rate of the granules [1]. FBD has a stainless steel chamber containing portable punctured bottom known as the bowl, where material to dry is located. Air, which is introduced from the bottom of the bowl, is heated up to the desired temperature by the heaters and moves through the bed of the material after it is sieved through the filter. The fans, which are found at the top of the equipment, produce this airflow and the control panel is used for regulating the flow rate and the operating temperature. The bed, which is the FBD bag, enlarges in size and particles start motion in a violent way as the flow of air rises. The material loses its moisture because of the frequent contact with air and the air gathers the tiny and dusty particles of the material on leaving the FBD through the filter. The material gets dried quickly as the FBD has a fast drying rate [1]. In the case of conventional moisture prediction strategies for the drying process, the design methods require the development of mathematical models describing the dynamics of the plant to be monitored. Then a control law from this model is to be derived by applying analytical techniques. The system with this control law in relation with the mathematical model of the plant is expected to give the desired performance. But, in practice, it is difficult to obtain the exact representative model of the plant due to the uncertainty, multivariable interaction, complexity and constraints of the plant. These are the main limitations of conventional moisture prediction strategies. Thus, there is a need to develop

sophisticated moisture prediction strategies with the help of the latest mathematical tools such as Artificial Neural Network, Fuzzy Logic, and Genetic Algorithms to deal with the increasingly complex systems and to control them under increased uncertainty. Due to increasing industrial applications of ANN and SVM, the process of fluidized bed drying requires improved modeling techniques incorporating neural and classification concepts for efficient operation and control. They can be used to model complex relationships between inputs and outputs or to find patterns in data [2].

An artificial neural network (ANN), often just called a “neural network” (NN), is a mathematical model or computational model based on biological neural networks, in other words, it is an emulation of the biological neural system. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation [2]. Neural networks are powerful nonlinear data processing techniques, which have proven their reliability in many areas. The neural network is a new approximation method of complex systems, especially useful when these systems are difficult to model by means of conventional statistical methods [3]. In this context, the objective of the research is to develop optimal prediction topology based on ANN and SVM and to figure out the significant factors affecting the moisture content of pharmaceutical granules.

Settled in 1997 E.C. and headquartered at Adigrat, Tigray regional State of Ethiopia, 898 km away from Addis, Addis Pharmaceutical Factory (APF) PLC is the biggest pharmaceutical manufacturing plant in Ethiopia and the town is 2300 meters above sea level. The plant was established and geared with high tech production equipments. In this manufacturing plant above 91 products of different types like Central Nervous system drugs, Cardiovascular drugs, anti diabetic agents, Gastro-intestinal drugs, Antihistamines & Anti allergic, Anti helminthics, analgesics, Antiprotozoals, Antibiotics, Vitamins and Large volume Parenterals has been manufacturing with international standards. There are above 9 manufacturing lines and laboratories which are fitted out with high tech materials and working equipments with the capacity to manufacture different types of syrups and tablets. Products like vials and dry powder for reconstruction are manufactured in a production line which is different from one where Tablets, Capsules, Suspensions, Sterile and Gels are produced. Later on the company starts manufacturing of large volume Parenterals in a branch factory found at Akaki/kaliti sub-city, Addis Ababa which is inaugurated in 2009. Production processes of the firm are fashioned, audited, overseen, guided and supervised according to current Good Manufacturing Practice (cGMP) recommendations. Consequently, the firm acquires GMP Certification by virtue of Certificate of Competence for Drug Manufacturing Plant from the indigenous regulatory authority of Ethiopia [4].

2. Methodology

There are two paradigms characterize much of the research in the Information Technology discipline: behavioral science and design science. The behavioral science paradigm seeks to develop and verify theories that explain or predict human or organizational behavior. The design science paradigm seeks to extend the boundaries of human and organizational capabilities by creating new and innovative artifacts. In the design-science paradigm, knowledge and understanding of a problem domain and its solution are achieved in the building and application of the designed artifact [5]. Design science paradigm has its roots in engineering disciplines it is a relatively young discipline, while the later one has its roots in natural science methods. Design science is fundamentally a problem-solving paradigm. It seeks to create innovations that define the ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of information systems can be effectively and efficiently accomplished. DS create and evaluate artifacts intended to solve identified organizational problems. As the researcher intends to develop a model, design – science paradigm is followed in this research.

2.1. Data Collection

Data collection is a crucial point in developing a model since ML algorithms need historical data in learning patterns of input-output relationships from the historical dataset. Hence collecting and preparing data sets was the first step in developing the models. Accordingly, the real world Datasets from 2006 E.C up to 2011 E.C is collected from APF randomly. These datasets are not fully self-explanatory for the researcher and

hence, experts' advice, discussions, and document analysis is used to identify the factors. Thus, with the help of domain experts at the organization, the datasets are identified and from this dataset, 5 attributes assumed to be determinant factors are nominated. This dataset contains 3570 records, each representing individual of the organization. Out of these datasets, 75 % of the data is used for training datasets and 15% for testing of the model and 15% for validation of the model.

2.2. Data Pre-processing

Since real-world data are usually incomplete, noisy, and inconsistent due to various reasons data preprocessing have been examined for completeness of the raw data available for training and testing models. Data cleaning involves handling missing values, smooth out noisy data, identifying outliers, and correcting inconsistencies and redundancy in the data. In this work, the missing values have been estimated by adding values and dividing to the number of records per batch. Data refining was performed on the raw experimental data by excluding all outliers which were unusual points for the outlier removed data-based experiment.

2.3. Data Normalization

To ensure the statistical distribution of the values for each net input and Output to be roughly uniform, a standard procedure for data preparation called data scaling is carried out. The data sets are scaled/normalized so that they fall within a specified range of 0 to +1.

2.4. Prediction Model

This work uses and investigates the Support Vector Machine (SVM) classifier, with four different kernels and parameter selection investigated, and Artificial Neural Network closely to classify moisture content given historical data sets as inputs. Support Vector Machine (SVM) is a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy while automatically avoiding over-fit to the data [6] and gained popularity due to many promising features such as better empirical performance [7]. Multilayer Perceptron's (MLPs) are the most common form of the feed-forward model architecture and was the architecture used in this work. Feed forward neural networks are widely and successfully used models for classification and problem solving [2], [8].

2.5. Model Evaluation

To determine which network structure is optimal, the performance of a calibrated model is evaluated against one or more criteria. In this work, the ANN and SVM model performance is assessed using a quantitative error metric which includes Mean Squared Error (MSE), Regression (R) and Confusion Matrix (CM).

2.6. Tools Used

In this research Matlab's neural network toolbox 2016a and classification learner app are used as a tool in modeling. Neural Network Toolbox provides flexibility in configuring and developing different models. The classification learner app supports supervised learning especially the toolbox with feed forward, radial basis, and dynamic networks as well as unsupervised learning. With the toolbox and the app, you can design, train and simulate the model. The toolbox gives flexibility (command line and GUI) on built-in functions in developing an ANN model. Besides, Microsoft Excel is used for data preprocessing and further data analysis.

3. Experimentation

This chapter presents the experimentation conducted and discussions on the Experimental results. The significant factors for moisture content determination of the granules are identified according to domain experts' discussions and recommendations and are proved through experimentation. The table below shows these attributes. Average value from a series of six repeated experiments was used. Taking the average results is an approach applied to minimize the instability inherited by the prediction model. In the experiments two procedures were conducted, one for the development of ANN model and the other for the SVM model. These procedures are considered to investigate and compare the classification performance of SVM and ANN, to

identify and assure significant factors and to come up with an optimal predictive model for pharmaceutical granules moisture content of APF.

Table 1. Attributes Selected According Domain Expert Recommendation

#	Attribute Name	Description
1	<i>Loge-Mixer time</i>	<i>The for time Granules in the Mixer</i>
2	<i>Drying time</i>	<i>The time needed to lose its moisture</i>
3	<i>Temperature</i>	<i>The temperature of the bed dryer</i>
4	<i>Initial moisture</i>	<i>Initial water content in granules</i>
5	<i>Air-flow rate</i>	<i>Airflow rate</i>

3.1. Building ANN Models

Experiment 1: Selecting the Optimal Number of Hidden Nodes

The goal of this experiment is to find the number of hidden neurons needed for the network to be able to predict the moisture content of pharmaceutical granules with possibly the highest accuracy given a fixed number of training parameters. To standardize the experiments in finding an optimal number of hidden neurons thumb rule was considered. Accordingly, for n inputs $\geq 2n$ experiment was done. But in this research, the lower bound is set to 1 to see if there is enhanced performance. Starting from 1 hidden neuron, then the number of hidden neurons is increased by one node until $2n$ is reached. For each experiment, six runs of the network are used; each run has different weight initialization. The result of a particular experiment is the average from the runs. A run was completed when the training epochs reached 1000. The table below shows performances obtained via a different number of hidden neurons for ANN models. As we can see from table 2 below, the network had better performance with a hidden neuron of 10. Hence, 10 numbers of hidden neurons are selected for investigating better achievement in performance for subsequent experiments.

Table 2. Performance of # of HN for ANN with Significant Attributes

# of HLN	ANN performance Description	ANN performance Description	ANN performance Description	ANN performance Description
	MSE	MSE	MSE	MSE
1	0.052651	0.050657	0.044153	0.048085
2	0.052688	0.044603	0.044326	0.048285
3	0.049583	0.051434	0.045508	0.04514
4	0.051803	0.04515	0.045111	0.051506
5	0.049792	0.047775	0.044466	0.045659
6	0.051151	0.052351	0.047103	0.051533
7	0.041448	0.041368	0.046321	0.04915
8	0.045464	0.052453	0.047361	0.044403
9	0.046628	0.049691	0.044869	0.05024
10	0.04728	0.054702	0.043113	0.0419

The study found that the network reached its average peak performance in specific ranges of hidden neurons from overall experiments that is when the number of hidden neurons is 10 Hence, 10 hidden neurons are selected for subsequent experiments.

Experiment 2: Selecting Training Function

The goal of this section of experiments is to explore the performance of various training functions and to select the appropriate training functions for the model. The following training algorithms are inspected:

Trainlm, trainscg, trainoss, traingd, traincgp and trainbfg as they are commonly used training functions for classifications. The table below shows experiment performances attained by these training functions for the respective models.

Table 3. Performance of Training Function for ANN

Training Function	# of HLN	ANN performance Description	ANN performance Description	ANN performance Description
		MSE	MSE	MSE
Trainlm	10	0.045851	0.050684	0.044878
Trainscg	10	<i>0.043159</i>	<i>0.043323</i>	<i>0.054939</i>
Trainoss	10	<i>0.060435</i>	<i>0.044584</i>	<i>0.044627</i>
Traingd	10	<i>0.056271</i>	<i>0.054693</i>	<i>0.041859</i>
Traincgp	10	<i>0.054059</i>	<i>0.055978</i>	<i>0.046054</i>
Trainbfg	10	<i>0.046602</i>	<i>0.046923</i>	<i>0.042551</i>

As we can see from the above table, the average result performance of the trainlm algorithm outperforms its counterparts in terms of achieving performance with the selected number of hidden neurons, and it is, therefore, selected as a training algorithm for the next subsequent experiments of our model.

Experiment 3: Selecting a Transfer Function

The aim of this series of experiments in this section is to investigate the performance of the transfer function and select the appropriate transfer functions for the respective model. The most common transfer functions LOGSIG, PURELIN, and TANSIG are used in this experiment. The two hidden layer neurons selected from the above experiment was examined with these transfer function. The table below shows performances attained by these transfer functions.

Table 4. Performance of Different Transfer Functions for ANN Model

Transfer Function in HL	# of HLN	ANN performance Description	ANN performance Description	ANN performance Description
		MSE	MSE	MSE
Logsig	10	<i>0.036594</i>	<i>0.047813</i>	<i>0.054159</i>
Tansig	10	<i>0.051452</i>	<i>0.054591</i>	<i>0.053178</i>
Purelin	10	0.03094	0.043441	0.041213

As it is shown in table 4 above, PURELIN with 10 numbers of hidden neurons performs the lowest average result of MSE, so it can be concluded that PURELIN outperforms TANSIG and LOGSIG. Therefore, PURELIN is the most suitable transfer function for the dataset of our problem domain.

3.2. Building SVM Models

The following experiments are performed to build an optimal SVM model.

A. Kernel Selection

B. Parameter Selection and Optimization

SVM training is carried out using Matlab's 2016 Statistics and Machine Learning Toolbox classification learner application. The classification learner application gives the same result as the experiment is done repeatedly, hence the experiment is done one time, i.e. the average mechanism is not applied. For example in kernel selection, only one experiment is performed for 5 kernels.

3.2.1. Kernel Selection

In this experiment, we have compared the performance of SVM by changing the kernel function and select the best kernel function of SVM for our problem domain.

Table 5. SVM Kernel Functions Performance

Kernel Function	Accuracy (%)
Linear	80.2%
Fine Gaussian	80.2%
Medium Gaussian	80.2%
Coarse Gaussian	80.2%
Quadratic	80.2%
Cubic	80.2%

As you can see from the table above all the SVM kernels performs good results (accuracy). As far as the performance of kernel functions is the same we can't able to select one with better accuracy among them to exceed the experimentation with SVM. Hence All SVMs are selected as they perform the same accuracy when compared with each other; hence other subsequent experiments are performed with All SVMs.

3.2.1.1. Parameter Selection and Optimization

3.2.1.1.1. Feature Selection

In Classification Learner, you can specify different features (or predictors) to include in the model to see if you can improve models by removing features with low predictive power [9]. Hence, in this experiment we have examined the performance as the effect of each attribute/parameter is missing. Principal component analysis (PCA) is used to reduce the dimensionality of the predictor space, even the explained variance is set to default. Reducing the dimensionality can create classification models in Classification Learner that help prevent over fitting. PCA linearly transforms predictors to remove redundant dimensions and generates a new set of variables called principal components [9].

Table 6. Accuracy as Loge-Mixer Time Factor not Selected

Kernel Function	Accuracy	Accuracy
	With PCA	Without PCA
All SVMs	80.2	80.2

Table 7. Accuracy as Drying Time Factor not Selected

Kernel Function	Accuracy	Accuracy
	With PCA	Without PCA
All SVMs	80.2	80.2

Table 8. Accuracy as Temperature Factor not Selected

Kernel Function	Accuracy	Accuracy
	With PCA	Without PCA
All SVMs	80.2	80.2

Table 9. Accuracy as Initial Moisture Factor not Selected

Kernel Function	Accuracy	Accuracy
	With PCA	Without PCA
All SVMs	80.2	80.2

Table 10. Accuracy as Flow Rate Factor not Selected

Kernel Function	Accuracy	Accuracy
	With PCA	Without PCA
All SVMs	80.2	80.2

Table 11. Accuracy as All the 5 Factors are Selected

Kernel Function	Accuracy	Accuracy
	With PCA	Without PCA
All SVMs	80.2	80.2

As can be observed from the above tables the best performance is achieved by All SVMs, which means when all attributes/features are selected, hence we can deduce that the attribute selection with the discussion of domain expert during data collection is assured.

3.2.1.1.2. Performance Optimization

Table 12. Accuracy of SVM as Box Constraint Level Change

Kernel Function	Accuracy	Accuracy	Accuracy	Accuracy
	Box Constraint Level 1	Box Constraint Level 2	Box Constraint Level 3	Box Constraint Level 4
All SVMs	80.2%	80.2%	80.2%	80.4%

Table 13. Accuracy of SVM as Kernel Scale Mode Change

Kernel Function	Accuracy	Accuracy	Accuracy	Accuracy
	Kernel Scale Mode Automatic	Kernel Scale Mode 0.56	Kernel Scale Mode 2.2	Kernel Scale Mode 8.9
All SVMs	80.2%	80.2%	80.2%	80.4%

As a result of the experiments on SVM in our domain problem, the optimal model developed is using All SVMs, with all features when PCA is on (i.e. when the explained variance is 95%) and Kernel scale mode is manual (i.e. specifically when it is set to 0.56, 1.2 and 8.9 for fine Gaussian, Medium Gaussian and coarse Gaussian respectively) and box constraint level is 4 with prediction Accuracy 80.4 %. Normc function is used to normalize/scale-out the dataset values and cross-validation is applied to overcome overfitting.

4. Result Analysis and Discussion

In this section, analysis and discussions on experimental results are made which includes model performance comparison based on performance measures between the models developed by the two techniques.

4.1. Model Comparison

After performing the experiments the next step is comparing the optimal models developed for our problem domain and selecting the 'best' model. The experiments are conducted to build an optimal (high performance) ANN and SVM models, hence performance measures MSE, R and CM are used to compare the models. The optimal model topology for the ANN model is shown in fig 1. For a direct comparison of the two models, the optimal performances of each model are compared with the respective measure. The accuracy of the models is computed by taking the percentage of the total number of correct classifications in both categories from the total number in both categories.

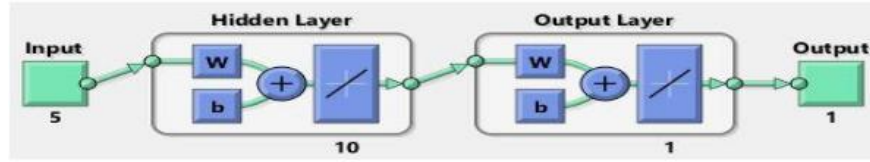


Figure 1 Optimal Model Topology for ANN Model

In fig 1, the topology shows 5 inputs which are the factors which were identified with the help of domain experts during data collection. 1 hidden layer with 10 hidden neurons Purelin as transfer function of hidden neurons and pureline as transfer function of output layer and with 1 output node which shows the moisture level of the granules.

4.1.1. Comparing Validation Performance (MSE)

The mean squared error (MSE) is the average squared difference between outputs and targets. Lower values of MSE are better while zero means no error which is ideal [9]. Hence, experiments are done to achieve minimum MSE close to zero.

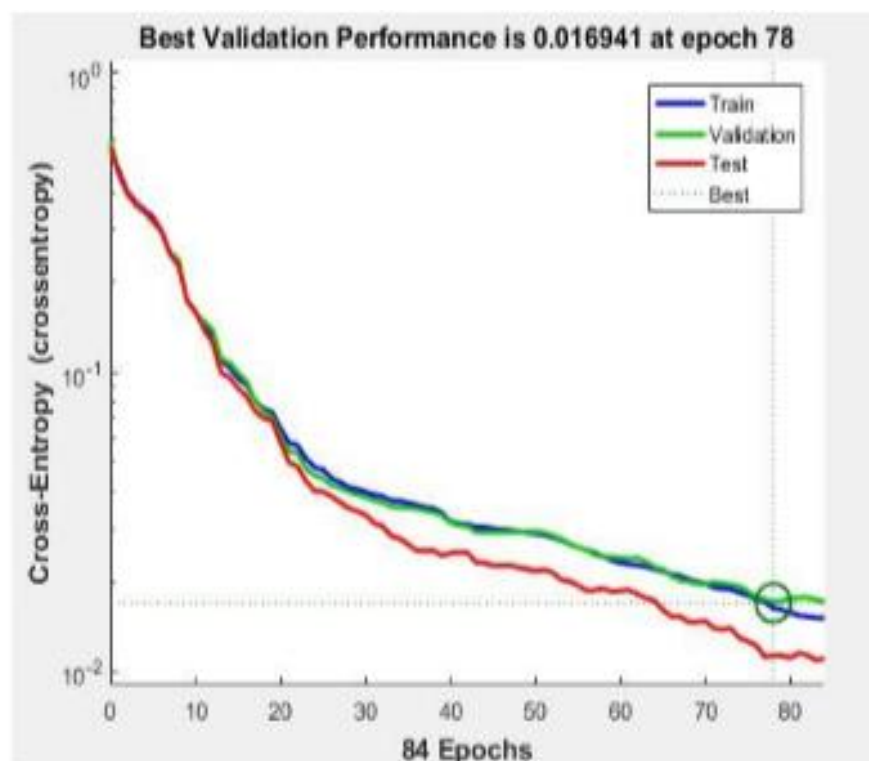


Figure 2 Validation Performance (MSE) of ANN Model

Best performance with 0.016941 at cycle 78 was attained by ANN. The performance achieved is good as it was close to zero. This figure does not indicate any major problems with the training because the validation and test curves are very similar. If the test curve had increased significantly before the validation curve increased, then it is possible that some overfitting might have occurred. Mean Squared error found in epsilon SVR was 0.1427 and it is observed that ANN outperforms SVM in validation performance.

4.1.2. Comparing Regression Performance (R)

The next step in validating the network is to investigate a regression plot, which shows the relationship between the outputs of the network and the targets. If the training were perfect, the network outputs and the targets would be exactly equal, but the relationship is rarely perfect in practice i.e. R-value of 1 means a close relationship, 0 a random relationship [9]. The figures below show the overall regression plot for the two models.

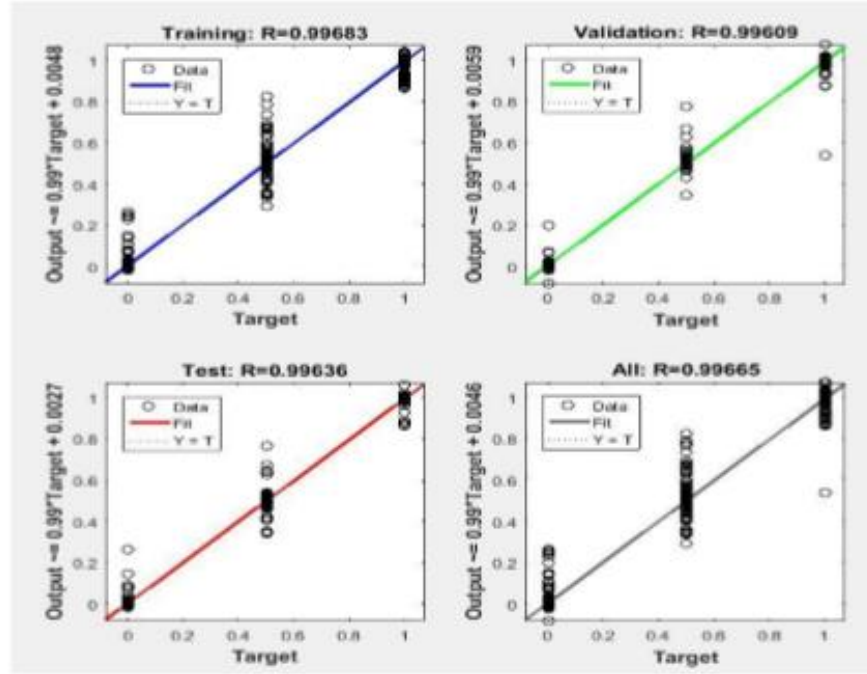


Figure 3 Regression Plot of ANN

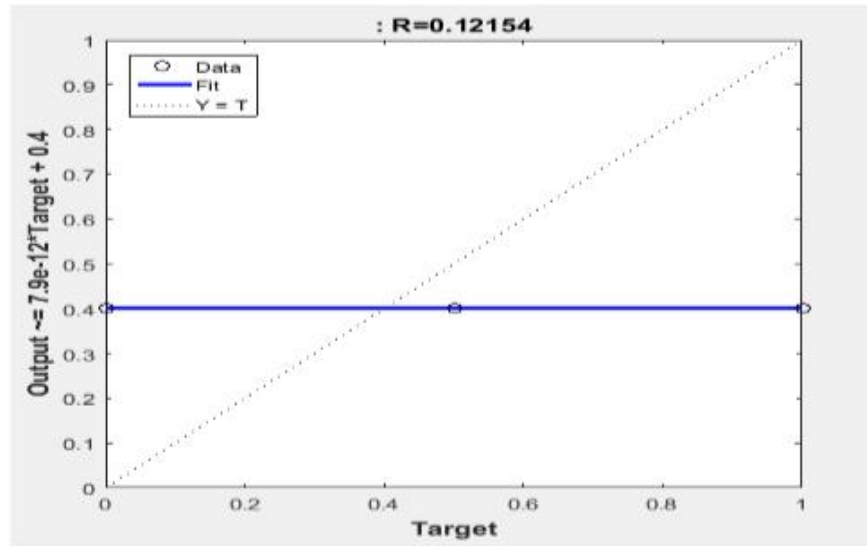


Figure 4 Regression Plot of SVM

The dashed line in each plot of the above two figures indicates the direct linear relationship between input and output which is perfect/ideal result where outputs = targets. The solid line represents the best fit linear regression line between outputs and targets. The R-value is an indication of the relationship between the outputs and targets. If $R = 1$, this indicates that there is an exact linear relationship between outputs and targets. If R is close to zero, then there is no linear relationship between outputs and targets [8]. From our experiment, the R-value attained is 0.99665 for ANN and 0.12154 for epsilon SVR. Result attained for

the regression performance by ANN is good as it is close to 1 and the Result attained for the regression performance by epsilon SVR is poor. The tiny circles in the plot represent datasets used in building the model. For our dataset in ANN, the training data indicates a good fit because it tends to fit the linear line with some scattered circles.

4.1.3. Comparing Classification Performance (CM)

The final step is comparing the classification performance of developed models hence confusion matrix is used to evaluate the performance as it is a widely used classification metric. The below figure shows the training, validation, and testing confusion matrix which compares/counts the predicted classes by the models against the actual classes during training, validation, and testing. Generally, a confusion matrix displays the number of correct and incorrect predictions made by the classification model compared to the actual outcomes (target values) in the data. The diagonal cells in each confusion matrix of the below figures show the number of cases that are correctly classified, and the off-diagonal cells show the misclassified cases. The blue cell in the lower right illustrates the overall percent of correctly classified cases in green and the total percent of misclassified cases in red. The classification performance can be interpreted using the following measures; Accuracy, Sensitivity, Specificity and Positive predictive value or Precision. The figure below shows the confusion matrix attained by the two models.

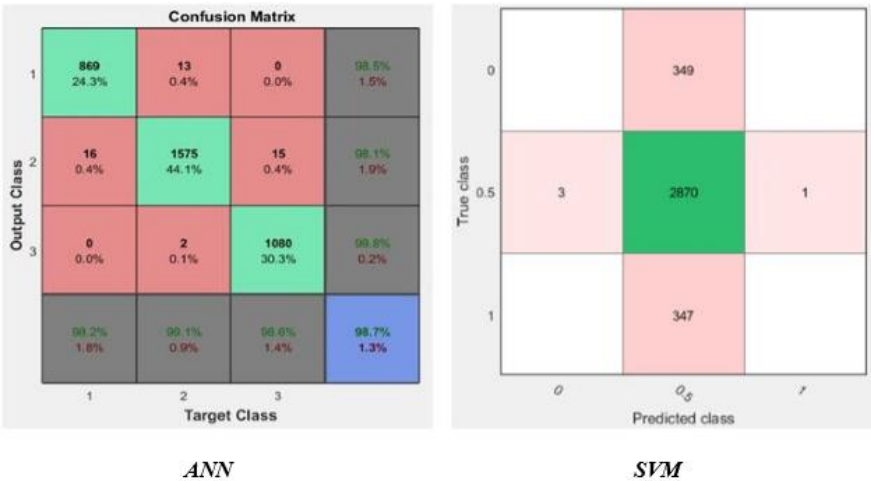


Figure 5 Confusion Matrix Obtained by the Two Models

In the figure above ANN, we can see that 869, 1575, 1080 (total 3524) cases are correctly (Green) classified by the model to their respective class. 869 cases are correctly classified to class 1 i.e. 869 samples were correctly classified as low-level moisture content. However, 13 cases are incorrectly classified to class 2 but they are class 1. 1575 cases are correctly classified to class 2. But 16 and 15 cases are incorrectly classified to class 1 and class 3 respectively but they are supposed to be class 2. 1080 cases are correctly classified to class 3. However, 2 cases are incorrectly classified to class 2 but they are supposed to be class 3. The experimental results presented in this work show that 98.7 % and 80.4 % classification accuracy is obtained by ANN and SVM models respectively, as depicted in the figure below.

Table 14 Classification Performance Measures for the Two Models

Model	Sensitivity (%)	Sensitivity (%)	Sensitivity (%)	Specificity (%)	Specificity (%)	Specificity (%)	Precision (%)	Precision (%)	Precision (%)
	Low	Low	Low	Low	Low	Low	Low	Low	Low
	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
	High	High	High	High	High	High	High	High	High

Model	Sensitivity (%)	Sensitivity (%)	Sensitivity (%)	Specificity (%)	Specificity (%)	Specificity (%)	Precision (%)	Precision (%)	Precision (%)
ANN	98.5	98.1	99.8	99.4	99.2	98.9	98.2	99.1	98.6
SVM	-	>99	-	100	<1	100	-	81	-

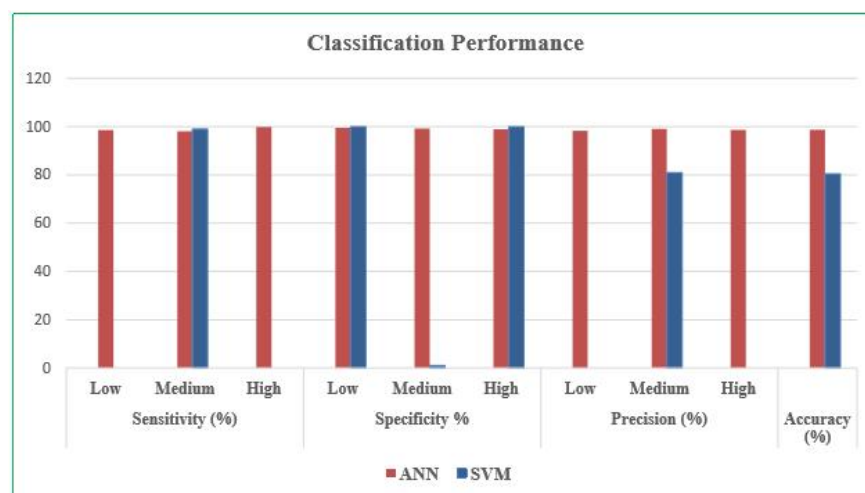


Figure 6 Classification Performances of the Two Models

Figure 6 above shows the classification performance achieved by the models. As you can see from the figure, the sensitivity, specificity, positive predicted value (precision) and accuracy were significantly increased in the ANN model when you compare it with SVM. In addition to this Temperature, Loge-Mixer time, Initial moisture, flow rate and drying time respectively are found the most important factors in determining the moisture content of the granules.

5. Conclusion

In this study, the aim was to develop a prediction model and identifying relative important factors in the evaluation of moisture content of pharmaceutical granules using ANN and SVM techniques. Datasets of 6 years were collected from APF containing 3570 instances and were preprocessed. In this study, optimal models of ANN and SVM models were developed and compared. The models were built on the preprocessed dataset with ANN and SVM classifiers through a supervised learning approach using Matlab's 16.0 as a software tool (Neural Network toolbox for ANN and Classification Learner and LIBSVM for SVM). The performance of the models was evaluated using a quantitative error metric; mean squared error (MSE), Regression(R) and Confusion Matrix (CM). The model classification performance was interpreted using the following measures: Accuracy, sensitivity, specificity, and positive predictive value or precision.

The results of this study revealed that the ANN model was the optimal model for predicting moisture content of pharmaceutical granules for datasets of APF. The model of ANN, with neural network topology 5-1-10-1 i.e. 5 inputs, one hidden layers, 10 neurons per each layer and one output node, trainlm training algorithm, PURELIN activation function in the hidden layers and PURELIN in the output layer, dividerand as divide function and MSE of 0.016941 with classification accuracy of 98.7% was built and accepted as optimal model for predicting moisture content of pharmaceutical granules of APF. With regards to the result of the experiments on SVM, the optimal model developed was using All SVMs with all features when PCA was enabled, Kernel scale was manual and the box constraint level was 4 with prediction Accuracy 80.4 %. Normc function was used to normalize/scale-out the dataset values and cross-validation with 5 folds was applied

to overcome over fitting. From this work, it is concluded that classification using ANN results in better classification accuracy than using SVM and is also best when compared to the results of related works. The result from the study also discloses *Temperature, Loge-Mixer time, Initial moisture, flow rate and drying time* respectively as the most important factors in determining the moisture content of the granules. Accuracy attained by the ANN model of this research was compared with the accuracy achieved in the related work. Our model shows better classification accuracy besides their dataset size difference. Finally, the research also confirms an Artificial Neural Network as a promising method for moisture classification and prediction.

Acknowledgments

First of all, we would like to acknowledge to the almighty god for his blessings. Additionally, we want to give great thanks to our families for their moral support. Finally, we would like to express our sincere gratitude to our advisor Dr. Tibebe Beshah for providing his invaluable guidance, comments, and suggestions, throughout the course of the thesis. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] H. Chen, X. Liu, C. Bishop, and B. J. Glasser, "Fluidized bed drying of a pharmaceutical powder: A parametric investigation of drying of dibasic calcium phosphate," *Dry. Technol.*, vol. 35, no. 13, pp. 1602–1618, 2017, doi: 10.1080/07373937.2016.1265552.
- [2] Y. Singh and A. Chauhan, "Neural Networks in Data Mining," *Journnal Theor. Appl. Inf. Technol.*, vol. 4, no. 3, pp. 01–06, 2014, doi: 10.9790/3021-04360106.
- [3] P. Maindi W, "Research Paper on Basic of Artificial Neural Network," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 2, no. 1, pp. 96–100, 2014.
- [4] Addis Pharmaceutical Factory, "Addis Pharmaceutical Factory." <http://www.apfethio.com/> (accessed Jun. 03, 2020).
- [5] S. Y. R. Esearch, B. A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research," *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.
- [6] V. Jakkula, "Tutorial on Support Vector Machine (SVM)," *Sch. EECS, Washingt. State Univ.*, pp. 1–13, 2011, [Online]. Available: <http://www.ccs.neu.edu/course/cs5100f11/resources/jakkula.pdf>.
- [7] Y. Karimi, S. O. Prasher, A. Madani, and S. Kim, "Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations," *Can. Biosyst. Eng. / Le Genie des Biosyst. au Canada*, vol. 50, 2008.
- [8] A. Krenker, J. Bester, and A. Kos, "Introduction to the Artificial Neural Networks," in *Artificial Neural Networks - Methodological Advances and Biomedical Applications*, 2011.
- [9] Mathworks Inc., "Matlab Statistics and Machine Learning Toolbox," 2015, 2019. <https://www.mathworks.com/products/statistics.html>.