ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250604003051

Original scientific paper
Received: 2025-06-04, Accepted: 2025-07-18

Predictive Modeling with Artificial Neural Networks to Optimize Dosing Accuracy of Galenical Powder Dosing Systems

Jaime Cancho*, Ciro Rodriguez, Ivan Petrlik, Milner Liendo

Abstract: A predictive model is proposed based on artificial neural networks (RNA) to optimize the quality of dosing of galenical powders in bottles, where it is necessary to maintain accuracy and stability, as the current electromechanical control methods have these shortcomings. The experimental development of research fosters new skills, which are key to innovating and facing the challenges of today's knowledge society. The RNA model was applied to the control group, resulting in an experimental group with improved. Six neural network models were trained, achieving the best results with the Recurrent Neural Network (RNN) model. Tests were conducted to optimize process capability indicators, improve process accuracy, and effectively predict the accuracy of dosed weight, considering the system's operating parameters. The RNN model was trained and validated with real data. The findings demonstrate that the application of the proposal will optimize accuracy and weight stability, meeting the quality standards in the industry.

Keywords: dosing; galenic powder; prediction; precision; rigid packaging; RNN

1 INTRODUCTION

In the pharmaceutical industry, the packaging process for galenic powders faces critical challenges of accuracy and consistency [1]. Powders are used in the manufacture of medicines that must comply with strict quality regulations [2]. Galenic powders, due to their physical nature, present properties such as irregular particle size, high cohesion, and sensitivity to environmental factors such as humidity and temperature, which complicate accurate and controlled dosing [3]. The United States Pharmacopeial Convention (USP) expert panel establishes control specifications for weight variation in continuous manufacturing in the pharmaceutical industry [4]. The quality of the packaged product is reflected in its nominal value of the declared weight on the package. [2] In an industrial process for powder packaging, materials, machines, labor, measurements, environment and methods interact, these six elements intervene in the quality of the product and changes inevitably occur over time and consequently variations in the packaging [2], therefore, the world's largest companies place great importance on the constant monitoring of these vital signs, which are known as: critical characteristics or indicators for quality. Continuous manufacturing is a novel process for producing high-quality pharmaceutical products [5]. When a galenic powder is dosed below that indicated on the package, the patient could receive an insufficient amount of the active ingredient, reducing the effectiveness of the treatment [6]. This can prolong the illness or require frequent dosage adjustments. If the weight of the galenic powder is greater than that indicated on the package, the patient could receive an excessive dose, causing serious adverse effects or toxicity, especially in drugs with a narrow therapeutic margin, where the difference between an effective dose and a toxic one is minimal [7]. Inaccuracy in dosing leads to variability in treatment outcomes, complicating the assessment of effectiveness and the ability of healthcare professionals to adjust treatment appropriately. Reliability in the administered dose is essential to ensure a consistent and predictable therapeutic response [8]. Inaccuracy in the weight of the packaged product puts patient safety at risk, especially with medications that require precise dosing [9], such as anticoagulants, cardiovascular drugs, or treatments for chronic diseases. Administering incorrect doses can lead to serious and even fatal complications [10]. Side effects resulting from inadequate doses can demotivate patients, reducing treatment adherence and increasing the risk of premature discontinuation. This could worsen the patient's condition and increase costs for their treatment and the healthcare system due to additional treatments and hospitalizations. Weight variations and inconsistencies during the packaging of galenic products have a significant impact on production costs, generating material waste, as incorrect doses must be discarded or reprocessed, increasing the use of additional raw materials and resources. Weight problems require additional controls and adjustments, increasing operating costs.

For the next 5 years, the global pharmaceutical market is expected to grow by 5% annually, requiring manufacturing companies to improve and maintain the quality of production and distribution of their pharmaceutical products [11].

Peru's economy is characterized by over 70% informal labor [12], which means that many domestic companies are not supervised by health quality regulatory bodies, making it increasingly difficult to identify or control unidentified or poorly controlled companies. In a globalized economy, where profit margins on the sale of pharmaceutical products are very tight, the use of electromechanical technology to control screw-type packaging machines is insufficient.

The research aims to apply a predictive model based on Artificial Neural Network (ANN) to optimize the accuracy of galenic powder packaging, for this purpose a mathematical algorithm will be trained with different ANNs and the results will be measured, choosing the most accurate and stable model.

ANNs work with large amounts of data, their ability to learn and update data allows them to predict errors and correct weight deviations during the packaging process in real time

In the food, pharmaceutical, cosmetics, and related industries, the accurate weight of the products offered is crucial to ensuring their quality and effectiveness. The main challenge in powder packaging is its precise measurement and dosage [13].

2 LITERATURE REVIEW

The proposed research is based on the application of mathematical algorithms that allow the improvement of processes, specifically, in our case, the packaging of sodium bicarbonate x 50 grams. Previous research has proposed methods that apply mathematical algorithms for preprocessing, balancing, data mining and selection, necessary for the development of more efficient predictive models, such as ANNs.

An essential aspect is model validation, which ensures that the predictions meet the accuracy requirements of the packaging MATLAB has been the most widely used tool in related work due to its ability to run and train predictive algorithms [14]. Variations of training algorithms available in MATLAB offer several options that allow improving the model fidelity when applied to galenical powder packaging. This research focuses on weight control, proposing an ANNs model that ensures accuracy and efficiency in galenic powder packaging.

2.1 Artificial Neural Networks (ANN)

ANNs are computer models that mimic the structure and function of the human brain. They are widely used in process prediction and classification, and are notable for their ability to identify trends from large volumes of data. Their biologically inspired design allows ANNs to adapt to different activities and environments, making them valuable tools for data management and promoting important advances in different fields of action where artificial intelligence and machine learning can be applied [15]. ANNs are trained to identify complex patterns and make accurate predictions. This makes them suitable tools for Quality Control and process automation in industry. The learning capacity of ANNs allows for significant improvements in industrial processes and optimization of their results, ensuring precision and efficiency [16].

2.2 Applications in the Pharmaceutical Industry

In the pharmaceutical industry, ensuring accurate packaging of galenic powders is critical to the correct administration of drugs to the patient. ANNs are used to anticipate weight errors and calibrate variables during dosing, which increases production efficiency and reduces waste of resources. This contribution of ANNs significantly improves process quality, ensuring stable and reliable results in drug production [17]. Recent research shows that ANNs stand out significantly in terms of accuracy and flexibility, making them the most advantageous option in various industrial applications compared to conventional methods. [18].

2.3 Predictive Models

Predictive models that use ANN, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), have proven to be highly effective in predicting process control variables. Their ability to handle different types of data highlights their great potential in various analysis and prediction uses [16]. ANNs have the ability to be trained using historical process data, allowing them to accurately predict the weight of galenic powders dosed into rigid containers and self-calibrate the variables that control this process in real time, guaranteeing consistent results. This characteristic makes them valuable tools to optimize industrial processes [15].

2.4 Challenges and Future Trends

Beyond their benefits, ANNs are hampered by the need to process large amounts of data for training and the difficulty in processing the results. These limitations can hinder their application and adaptation in certain environments. [1]. However, looking to the future, we need not only technically efficient models but also more understandable and easier-to-program models. Deep learning techniques are currently being integrated to increase the accuracy and stability of predictions. These advances project a significant improvement in the efficiency and applicability of ANNs [18].

Applying ANN to predict weight accuracy in powder packaging is a revolutionary industrial innovation. ANNs are trained and adapt to various process conditions, optimizing quality and efficiency. This demonstrates a significant technological advance in the industry, demonstrating its potential to optimize industrial processes and ensure stable and reliable results.

3 MATERIALS AND METHODS

The research conducted is of the applied type because it uses existing knowledge. Applied research considers all the norms, regulations, and statutes that govern social behavior, ensuring that its findings and methods are aligned with the legal and ethical context. [1].

As a sample, there is a weight record of 1,430 Sodium Bicarbonate bottles × 50 grams packaged with the METS-01 screw dosing machine from a Pharmaceutical Laboratory in Lima, taken randomly from the production line, at the discretion of the Quality Control Department, these correspond to 11 production batches on consecutive dates since January 2025, the variables that affect the operation of the machine during dosing were also considered.

The sample was collected using a precision digital scale that was initially certified by a competent laboratory specializing in its field. For weight control studies, the use of a precision balance will be necessary. [19].

The input variables involved in controlling the weight of the METS-01 are the rotational speed of the three-phase electric motor. This speed is adjusted by a frequency converter. This rotation drives the worm gear that transports the galenic powder and the electrical pulses per revolution read by the incremental encoder.

Figure 1 METS-01 screw packaging machine

The information obtained is analyzed, which will yield important insights and establish patterns, relationships, and trends for developing the prediction model.

To verify the current accuracy of the galenic powder packaging process, the Process Control Analysis (PCA) was carried out. Applying statistical control through variable control allows for greater control of the final product, guaranteeing its specifications and quality [20].

It was necessary to determine the distribution type of the sample data to clearly define the type of statistic to be used in the inferential statistics; therefore, the Student t-test was used. The effect of the model on the capacity indicators of the galenic powder packaging process was reviewed, and the Cp and Cpk in the control group and their optimization were identified after applying the model to the experimental group.

3.1 Type of Artificial Neural Network

It began with a regression analysis that examines the relationship between process variables, such as the rpm of the motor that drives the auger and the electrical counting pulses per motor revolution, and the weight history. The data was gathered, cleaned, and normalized to ensure consistency. A linear regression model was subsequently implemented as an initial approximation to understand the influence of the variables on the batch weight. This model provided us with a solid foundation for designing and training the neural network, using its results to define the appropriate structure, layers, and activation functions, adjusting the parameters to improve the model's accuracy. There are several types of ANN, each with its own characteristics, advantages, and disadvantages. To decide which type of neural network to use for the project, a mathematical algorithm was developed and trained with six different types of ANN. The results were verified, and it was decided that the research should be done with the RNN model because it offers greater advantages in accordance with our objective. The RNN is a type of ANNs. Once the neural network was implemented and the effect of the verified improvements was measured, the findings were compared with similar research to validate the proposed model. Using neural networks contributes to improved accuracy compared to the previous system [21].

4 ETHICAL CONSIDERATIONS

The owner of the MEPS-01 machine has given his or her consent, and data privacy and confidentiality will be respected. Ethical conduct in research must consider three general dimensions: the human, the political, and civil society [22].

5 RESULTS

5.1 Data Analysis and Visualization

The dataset to be analyzed by the model is created from the sample. The upper and lower limits, range, class number, class size, standard deviation, variance, mean, median, and mode are calculated. The histogram is created from this data.

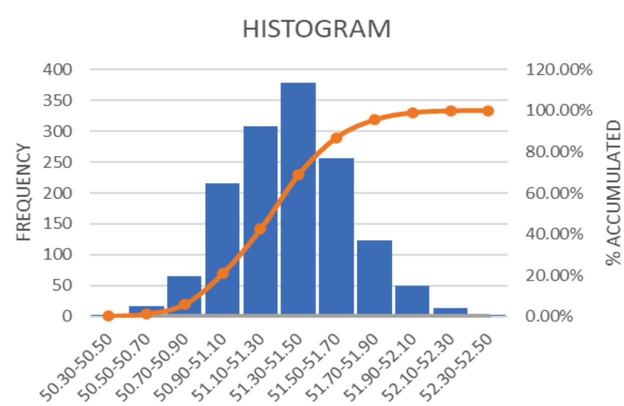

Figure 2 Visual representation of data

It is observed which frequency is almost symmetrical and bell-shaped (Fig. 2), suggesting that the data have a nearly normal distribution since the highest peak is in the center, specifically in frequency bar 379; as we move away to the left or right, the frequencies progressively decrease, indicating that there is less data at the extremes.

The frequencies do not show skew to the left or right, that is, they are not tilted, which reinforces the idea of a normal distribution.

To give a better visualization of the statistical diagram (Fig. 2) we will increase the interval to a smaller scale and propose an interval of 0.1, where the frequency and its distribution are found.

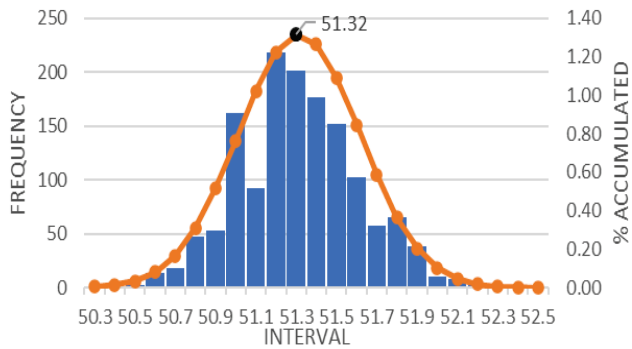

Figure 3 Data frequency distribution

In the Frequency Distribution (Fig. 3), it is verified that the distribution of the data is centered according to the mean; we can affirm that the sample weight data are within a normal distribution.

For the scatter plot, we group the data that were taken at the same time, 10 records per group, obtaining 143 groups, and the scatter diagram is developed.

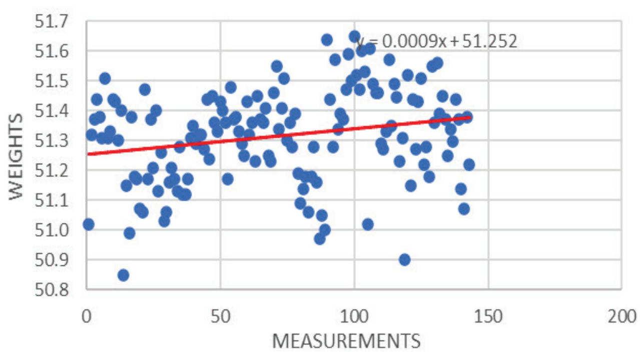

Figure 4 A wide dispersion is seen, that is, there are many outliers that do not follow the general trend line pattern

5.2 Analysis and Efficiency of the METS-01 Galenic Powder Packaging Machine

X-R charts are used for statistical monitoring of part quality control in industry. They allow for the detection of variability, consistency, control, and improvement of a production process. They are composed of the following elements:

- Upper control limit (*UCL*), or maximum tolerance.
- Lower control limit (*LCI*), or minimum tolerance.
- Control Line (*CL*), or average of the minimum and maximum tolerances.

Measurement variables, which in our research are the weights of the bottles packed with Sodium Bicarbonate × 50 grams in the Pharmaceutical Laboratory.

The Tab. 1 must be completed to calculate the control limits for means and ranges.

Table 1 Formulas to obtain the limits by means and ranges

Graphics for	*LCS*	*LC*	*LCI*
Averages	$\bar{\bar{x}} + A_2 \cdot \bar{R}$	$\bar{\bar{x}}$	$\bar{\bar{x}} - A_2 \cdot \bar{R}$
Ranks	$D_4 \cdot \bar{R}$	$\bar{R}$	$D_3 \cdot \bar{R}$

To perform the calculations, the values of A_2, D_3, and D_4 are required, which will be extracted from the XR Constants Table. n = 10 must be considered, then it is obtained.

Table 2 Summary of data obtained

Graphics for	*LCS*	*LC*	*LCI*
Averages	51.56	51.3	51.07
Ranks	1.417	0.8	0.178

Integrating the sample and Tab. 2, we have Fig. 5.

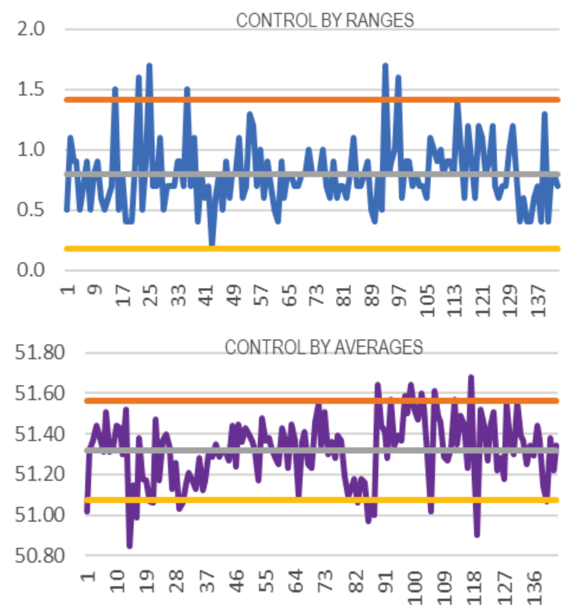

Figure 5 Graph of Averages and Ranges, the process is outperforming the LCS and LCI

Both graphs indicate that the process is statistically out of control. Because the operator's procedure was observed during packaging, we can confirm that the machine is out of control in terms of weight accuracy and consistency during the dosing of galenic powders into rigid packaging.

5.3 Analysis of Packaging Process Capacity

Packaging process capacity can be expressed as the packaging machine's ability to perform its intended purpose within the permitted tolerances. There are two main indicators: one C_p shows whether the data distribution can fit within the required production specifications, and the other C_{pk} shows whether the overall data average is located at the center of the limits.

The legal tolerable weight requirements for prepackaged products are specified, indicating predetermined nominal values [24]. From the Peruvian Metrological Standard [23], the tolerance for 50-gram presentations is –0.0 and +1 grams. In our case, the Control Line or production average LC = 51.3 grams, the Pharmaceutical Laboratory has seen fit to consider ±1 gram of tolerance for the analysis of the packaging process capacity. In this sense, USL = 52.3 grams and LSL = 50.3, with these limits, the control chart is drawn by specific averages.

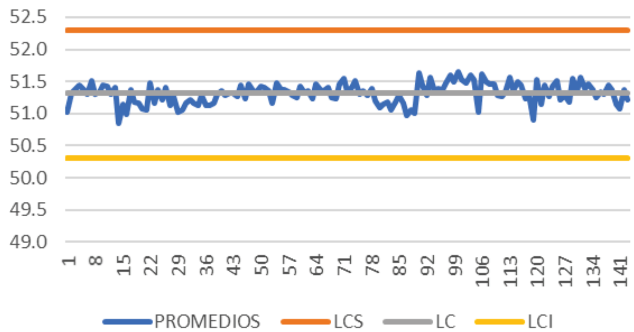

Figure 6 Specific means are considered as tolerances specifically given by the manufacturer

It can be seen that all means of the data subgroups are fully controlled, with slight alterations inherent to the process.

The process capability analysis C_p and the actual process capability index C_{pk} of the packaging will be carried out to determine if the process is capable of producing Sodium Bicarbonate × 50 grams according to the tolerance decided by the Pharmaceutical Laboratory in ±1 gram.

For C_p:

$$C_p = \frac{USL - LSL}{6 \cdot \sigma} = \frac{52.3 - 50.3}{6 \cdot 0.3031} = 1.0997 \tag{1}$$

For C_{pk}:

$$\begin{aligned} C_{pk} &= \text{minimum}\left(\frac{USL - \overline{\overline{x}}}{3 \cdot \sigma}, \frac{\overline{\overline{x}} - LSL}{3 \cdot \sigma} \right) = \\ &= \text{minimum}\left(\frac{52.3 - 51.3}{3 \cdot 0.3031}, \frac{51.3 - 50.3}{3 \cdot 0.3031} \right) = 1.0997 \end{aligned} \tag{2}$$

$C_p > 1$, the process is moderately capable, and requires observation.

If $C_{pk} > 1$, the process is suitable for producing within the given specifications.

5.4 Analysis of Prediction Models

The mathematical algorithm was designed, and six Artificial Neural Network models were trained. Optimized Artificial Neural Network – TUNED_ANN, Deep Neural Network – DNN, Long Short-Term Memory Network – LSTM, One-Dimensional Convolutional Neural Network – CNN_1D, Multilayer Perceptron – MLP, Recurrent Neural Network – RNN.

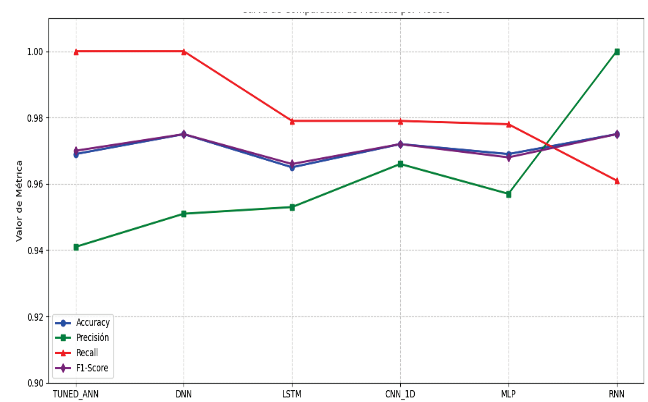

Figure 7 Metrics of the 6 models tested

The term added as LSTM means Long Short-Term Memory. The comparison of metrics by model allows us to compare between the reviewed models and conclude that the RNN model outperforms all the ANN models trained for the prediction of weight in the packaging of galenic powders in rigid packaging, therefore, it is stated that the RNN model is the best model to predict the accuracy of the weight of the packaging of Sodium Bicarbonate × 50 grams in rigid packaging with the METS-01 packaging machine.

5.5 The Indicators were obtained from the RNN Model

The Accuracy compares the overall performance of the models in terms of global precision. The RNN model obtains the highest Accuracy (0.975), which suggests high performance in correct predictions. The Precision shows how accurate each model's predictions are, avoiding false positives. The RNN model has a perfect precision (1.0000), ensuring that all its predictions are correct. The F1-score evaluates the balance between Precision and Recall. The RNN model has the best F1 Score (0.9822), which indicates that it correctly balances both metrics. Recall is the ability to effectively detect correct values. The RNN model has the best Recall (0.973), balancing both metrics.

We will work with an RNN with an LSTM architecture, which is trained to predict the binary classification of the weight dispensed by the METS-01 machine, based on the input variables. This classification is defined based on whether the predicted weight value is below or above the median, transforming a regression problem into a classification problem.

This RNN-LSTM architecture represents a systemic and integrated solution for addressing prediction problems in industrial processes, aligning with the fundamental principles of Systems Engineering, such as:

- Dynamic modeling of complex systems.
- Decision-making under uncertainty
- Performance optimization in multidimensional systems.

Interoperability between mathematical models, real data, and physical systems.

We evaluate whether our data belongs to a normal distribution. For this, we will work with the help of IBM SPSS STATISTICS Viewer software and Tab. 2. The

Kolmogorov-Smirnov normality test was performed, and a p-value of 0.2 was obtained. Since we have considered a significance level $a = 0.05$, it is necessary that $p > a$; therefore, it is accepted that "the data follows a normal distribution".

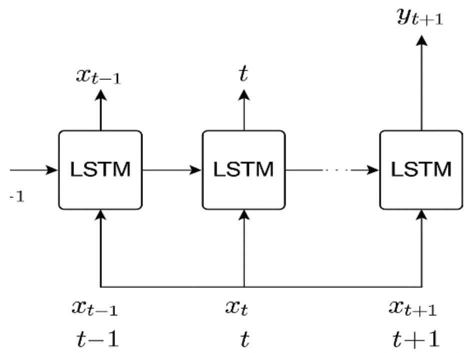

Figure 8 Sequential flow of information through time steps and the propagation of hidden states, which clearly illustrates the dynamic and adaptive nature of the model.

Groups of 5 data points were formed, obtaining 28 subgroups as the Control Group (CG), the data for USL and LSL are considered, and formulas 1 and 2 are applied to find Cp and Cpk for each subgroup. The defined RNN-LSTM model is applied to this control group and Cp and Cpk are obtained with values higher than 5, mathematically it is possible due to the high capacity of the model, but in industrial application this value is not realistic, therefore we will consider a statistically acceptable and technically viable value of C_p and $C_{pk} = 1.333$.

The groups and variables are reordered, and with the help of SPSS, $p < 0.001$ is obtained for the C_p Group of the CG and EG, and $p < 0.001$ for the C_{pk} Group of the CG and EG. Then it is true that: $p < 0.001$ a. Therefore, we can affirm that neural networks optimize the capacity indicators of the dosing process of galenic powders in rigid packaging.

From the Control Group, 30 samples were randomly selected as Initial Weight to perform the Student's t-test. The defined RNN-LSTM model was applied, and the Modeled Weight was obtained, with an average of 50.6 grams. Analyzed with SPSS software.

Table 3 From the Student t-test for sample correlation, $p < 0.001$

Sample correlations				
	N	Correlation	Significance	
			p of a factor	Two-factor p
Initial weight & model weight	30	–.762	< .001	< .001

Maintaining the same value of the significance level to $a = 0.05$, we have that $p < a$, therefore, it is stated that the prediction model based on neural networks determines the best accuracy of packaging galenic powders in rigid packaging.

Powder packaging between 50.2 and 50.7 grams is acceptable. We will consider precision as a dichotomous variable: 1 indicates weights that are within tolerance (precise) and 0 indicates weights that are outside the tolerance (imprecise). Binary language favors the application of prediction models such as RNN and the evaluation of performance metrics such as *Precision*, *Accuracy*, *F1-Score,* and *Recall*. The 1,430 data points were grouped into 286 groups and converted into binary language. The Student t-test for related samples is performed to determine if there is a statistically significant difference between the actual values and the predicted values, as well as the key result *p*-value (two-tailed sig.). With SPSS software, we have

Table 4 Paired sample test, $p < 0.083$ was obtained, then: $p > a$

Paired sample test				
	Standard deviation	Mean standard error	p of a factor	Two-factor p
Actual value & predicted value	0.102	0.006	< 0.042	< 0.083

It also performed simple linear regression.

Table 5 A linear correlation $R = 0.974$ indicates a positive and very strong correlation between predicted and actual values

Model summary					
Model	R	R^2	Adjusted R^2	Standard error of the estimate	Regression coefficient B
1	0.974	0.948	0.948	0.101	0.962

The coefficient of determination $R^2 = 0.948$, meaning that the model explains 94.8% of the variability in the accuracy of the actual weight. The adjusted $R^2 = 0.948$ confirms the stability of the model and that there is no overfitting or loss due to complexity. The standard error of the estimate indicates the average of the errors between the predicted and actual values. A low value, such as 0.101, suggests that the model is fairly accurate in its estimates.

Statistical indicators obtained confirm that the applied RNN-LSTM model effectively predicts the weight accuracy in the packaging of galenic powders, since, due to the three conclusive findings, it is stated that the RNN-LSTM model optimizes the accuracy of dosing galenic powders in rigid packaging.

6 DISCUSSION

Initially, data from the production line or control group demonstrate that the process is unable to meet the limits specified for the galenic powder packaging process, with an average C_p and C_{pk} of 0.295 and an average standard deviation of 0.13 grams. When the galenic powder packaging process capacity indicators show an indicator lower than 1, it indicates that the process is at considerable risk, which is a level below the permitted quality standards.

Levene's test shows a significant result ($F = 37.187$; $p < 0.001$), which confirms that the variability of the EG is much lower than that of CG. The research carried out confirms that by applying the RNN model to the control group, an experimental group is obtained with a better centered process and with a reduction in its dispersion, which directly affects the improvement of the process capacity indicators C_p and C_{pk} for the packaging of galenic powders. The results obtained in this research agree significantly with the findings described by [25] using the quality indicators C_p and C_{pk}, they also use RNN-LSTM to predict the process quality performance. In addition, the implementation of RNN allows it to reduce the standard deviation of the analysis process, that is, it has less variability, so the values of C_p and C_{pk} increase significantly. They report that the RNN model achieved process improvement through a marked reduction in the standard deviation, which increased the process capability indices, considering those obtained with models such as Random Forest or ARIMA. Similarly, in our research, RNN was applied to the control group, and an experimental group with a standard deviation of 0.025 grams was obtained. This considerable improvement reflects an increase in the process capability indicators, obtaining a C_p and C_{pk} of 1.333. This value shows that the process is stable and capable, according to standards in the galenic industry, and also indicates that the C_{pk} has a proportion of less than 0.01% of non-conforming products. Then (1.333/0.295 = 4.53), it is verified that the application of the model offers a 4.53-fold improvement in the capacity of the galenic powder packaging process. The Student t-test shows statistically significant results, $t = –50.254$ and $p < 0.001$.

The verification of the performance of the predictive model based on recurrent neural networks (RNNs) developed to design a prediction model for the accuracy of galenic powder packaging is based on the technical analysis of its performance indicators and their comparison with documented studies of a high methodological level. The model was designed to predict weight accuracy in the galenic powder packaging process, using sequential variables such as the rotation speed of the auger and the electrical pulses per revolution of the auger. In this context, key metrics such as *Accuracy*, *Precision*, *Recall* (sensitivity), and *F1-score* (harmonic mean of precision and sensitivity) were applied, obtaining values above 80% in all cases, demonstrating a high predictive capacity.

This performance is supported on a study [25], where a hybrid ensemble model is proposed that integrates multiple deep architectures to classify genetic mutations in cancer patients. In this work, the authors applied their model to the Kaggle clinical dataset "MSK-Personalized Medicine", obtaining high-level validation indicators: *Accuracy* = 80.6%, *Precision* = 81.6%, *Recall* = 80.6%, and *F1-score* = 83.1%. These results are lower than those obtained in our research "Predictive model based on neural networks to optimize the accuracy of dosing galenic powders in rigid containers", where the following were obtained: *Accuracy* = 97.5%, *Precision* = 100%, *Recall* = 97.3%, and *F1-score* = 97.5%. The findings reinforce the idea that when RNNs exceed 80% in their validation indicators, they are highly effective in modeling sequential data and can be used in an automated production line.

Furthermore, the work [26] establishes that predictive models with F1 and AUC metrics above 80% are suitable for implementation in industrial environments. The authors emphasize that the *F1-Score* provides a balanced metric, ideal for processes where false positives and false negatives have an economic impact, such as in weight control in the packaging of galenic powders. This assertion reinforces the reliability of the proposed model for plant application by avoiding losses due to overfilling or underfilling, which entails raw material losses, or deficient filling, also known as underfilling, which affects product quality.

Regarding statistical validation, the model underwent normality testing (Kolmogorov-Smirnov) and t-tests comparing means to evaluate the difference in performance between the control group with data from the production line and the experimental group obtained after applying the RNN model. A statistically significant difference was found ($p < 0.05$), indicating that the RNN model substantially improves process accuracy.

7 CONCLUSIONS

Regarding the variables that affect the capacity of the galenic powder packaging process, the C_p and C_{pk} indicators of the control group were optimized 4.53 times with the application of recurrent neural networks, using historical data as inputs.

From 1430 data, 143 groups were formed, three groups were omitted and 28 subgroups of 5 data each were formed, each of the 28 subgroups had their C_p and C_{pk} calculated for the control group and the experimental group, the data are integrated and the average C_p and C_{pk} for the CG = 1.333, likewise the average C_p and C_{pk} of the SG = 1.333. Then: (1.333 – 0.2954)/(0.2954) = 3.5125. Then, it is stated that the indicators of the capacity of the galenic powder packaging process C_p and C_{pk} have been optimized by 351.25% by the application of the RNN model.

To verify that the design of a prediction model based on neural networks determines the best accuracy of the packaging of galenic powders, 30 random data were taken as a control group to which the RNN model was applied and the experimental group was obtained, the averages were calculated for GC = 52.28 g and GE = 50.64 g, it is known that the tolerance allowed for packaging applying the RNN model is 50.2 to 50.7 grams, the average of the allowed tolerance being 50.45 grams, then the CG error: 51.28-50.45 = 0.83 grams and the *EG error*: 50.64 – 50.45 = 0.19 grams, from these data it is obtained that (0.83 – 0.19)/(0.83) = 0.7711, then it is stated that the RNN model improved by a 77.11% weight accuracy in packaging of Sodium Bicarbonate × 50 grams

To verify that a model based on neural networks can effectively predict the weight accuracy in the packaging of galenic powders, from the operating parameters of the system, it was necessary to experiment with the data taken from the production line (control group), the already trained RNN model was applied and results were obtained (experimental group) within the established tolerance of 50.2 to 50.7 grams, the metrics were observed, being the Accuracy 97.5% the Precision 100%, the Recall 97.3% and the F1-Score 97.5% these values are very favorable which ensure an effectiveness of the model over 97%.

Of the 1430 data taken from 11 production batches of Sodium Bicarbonate × 50 grams in rigid packaging, from the

Pharmaceutical Laboratory, the highest weight is 52.5 grams and the lowest is 50.3 grams, with an average of 51.3 grams of samples, applying the RNN model, the average is 50.6 grams, which allows us to affirm that there will be a saving of 0.7 grams per bottle of Sodium Bicarbonate × 50 grams, which is equivalent to 1.36% of raw material per bottle.

From the study carried out, it is stated that the proposed RNN model significantly improves weight accuracy in the packaging of galenic powders, considering historical variables.

8 REFERENCES

[1] Castro Maldonado, J. J., Gómez Macho, L. K., & Camargo Casallas, E. (2023). La investigación aplicada y el desarrollo experimental en el fortalecimiento de las competencias de la sociedad del siglo XXI. *Tecnura, 27*(75), 140–174. (in Spanish) https://doi.org/10.14483/22487638.19171

[2] Díaz, E. E., Díaz, C., Flores, L. C., & Heyser, S. (2009). Estudio de la Variabilidad de Proceso en el Área de Envasado de un Producto en Polvo. *Información tecnológica, 20*(6). (in Spanish) https://doi.org/10.4067/S0718-07642009000600013

[3] Gallo, L., Chanampa, L., Gonzales, A., Calcagno, A., Razuc, M., & Natalini, P. (2019). *Tecnología farmacéutica: Diseño y preparación de formulaciones*. Editorial de la Universidad Nacional del Sur. Ediuns. (in Spanish)

[4] United States Pharmacopeia. (2023). Perspectiva (Farmacopeica) de la USP sobre Fabricación Continua de Medicamentosa. *Pharmacopoeial Forum, 44*(6), 1–30. (in Spanish)

[5] Santamaría Alvarez, F. N., Kleinebudde, P., & Thies, J. (2024). Influence of refilling on dosing accuracy of loss-in-weight powder feeders in continuous manufacturing. *Powder Technology, 436*, 119501. https://doi.org/10.1016/j.powtec.2024.119501

[6] Mauro, L. (2024, February). Non-Sterile Clinical GALENIC LABORATORY: A Scientific Discipline between Laboratory Practice Clinical Pharmacy and Personalized Pharmacological Therapy. *Archives of Pharmacy & Pharmacology Research, 4*(1). https://doi.org/10.33552/APPR.2024.04.000579

[7] Luisetto, M. et al. (2024). Galenic Laboratory: State of the Art—A Scientific and Technological Discipline, Innovation and Management. *Open Access Journal of Pharmaceutical Research, 8*(2). https://doi.org/10.23880/oajpr-16000316

[8] Juárez-Hernández, J. E., & Carleton, B. C. (2022). Pediatric oral formulations: Why don't our kids have the medicines they need? *British Journal of Clinical Pharmacology, 88*(10), 4337–4348. https://doi.org/10.1111/bcp.15456

[9] Desideri, I., Martinelli, C., Ciuti, S., Uccello Barretta, G., & Balzano, F. (2022). Lopinavir/ritonavir, a new galenic oral formulation from commercial solid form, fine-tuned by nuclear magnetic resonance spectroscopy. *European Journal of Hospital Pharmacy, 29*(5), 259–263. https://doi.org/10.1136/ejhpharm-2020-002389

[10] da Silva, M. R. M., Dysars, L. P., dos Santos, E. P., & Ricci Júnior, E. (2020). Preparation of extemporaneous oral liquid in the hospital pharmacy. *Brazilian Journal of Pharmaceutical Sciences, 56*. https://doi.org/10.1590/s2175-97902019000418358

[11] Toktonalieva, N., & Toktonaliev, I. (2020). History and Background of the Implementation of Good Manufacturing Practice Standards in the Pharmaceutical Industry (Review). *Bulletin of Science and Practice, 6*(9), 182–191. https://doi.org/10.33619/2414-2948/58/17

[12] La Rosa Basurco, J. (2023, April). *El reto de formalizar*. Ministerio de Economía y Finanzas del Perú. (in Spanish)

[13] Techflow Pack. (2024, May 27). *Envasado eficiente de polvo; agilización de los procesos de envasado con máquinas avanzadas*. TECHFLOW PACKAGING SOLUTIONS ENGINEERING CORP. (in Spanish)

[14] Ishfaque, M., Dai, Q., ul Haq, N., Jadoon, K., Shahzad, S. M., & Janjuhah, H. T. (2022). Use of Recurrent Neural Network with Long Short-Term Memory for Seepage Prediction at Tarbela Dam, KP, Pakistan. *Energies, 15*(9), 3123. https://doi.org/10.3390/en15093123

[15] Varela-Arregoces, E., & Campells-Sanchez, E. (2011, June). Redes Neuronales Artificiales: una revisión del estado del arte, aplicaciones y tendencias futuras. *Universidad Simon Bolivar*, 18–27. (in Spanish)

[16] Perez Ortiz, J. A. (2002). *Modelos predictivos basados en redes neuronales recurrentes de tiempo discreto* [Doctoral thesis, Universidad de Alicante]. España. (in Spanish)

[17] George Reyes, C. E. (2019). Methodological strategy to develop the state of the art as a product of educative research. *Praxis Educativa, 23*(3), 1–14. (in Spanish) https://doi.org/10.19137/praxiseducativa-2019-230307

[18] Cruz Martinez, A. G. (2018). El estado del arte en la construcción de la tesis doctoral y la gestión de nuevos conocimientos. *Revista Caribeña de Ciencias Sociales*, Universidad Pedagógica Nacional, México. Retrieved October 26, 2024, from https://www.eumed.net/rev/caribe/2018/04/gestion-nuevos-conocimientos (in Spanish)

[19] Arcos, L., & Armas, A. (2019, June). Estabilidad en cuanto a color y peso, de resinas compuestas tipo flow tras contacto con bebidas gaseosa: estudio in vitro. *Revista Odontología Vital*, 59–64. (in Spanish) https://doi.org/10.59334/ROV.v1i30.138

[20] Atencia, C., Reyes, C., & Bautista, A. (2022, August). Aplicación de control estadístico de calidad en bloques bio-climáticos de arcilla mediante gráficos de control por variables. *Universidad Nacional Experimental de Táchira*, 202–2108. (in Spanish)

[21] Xavier, O. C., Pires, S. R., Marques, T. C., & Soares, A. da S. (2022, June). Identificação de evasão fiscal utilizando dados abertos e inteligência artificial. *Revista de Administração Pública, 56*(3), 426–440. (in Spanish) https://doi.org/10.1590/0034-761220210256

[22] Herrera Medina, N., Rivera Gutiérrez, S., & Espinoza-Navarro, O. (2022). Marco Ético y Jurídico en la Investigación Científica en Chile: Acreditación de Comités Ético Científicos. *International Journal of Morphology, 40*(4), 953–958. (in Spanish) https://doi.org/10.4067/S0717-95022022000400953

[23] INDECOPI. (2008). *Norma metrológica peruana* (pp. 1–27). Perú: Lima. (in Spanish)

[24] Pheng, T., Chuluunsaikhan, T., Ryu, G.-A., Kim, S.-H., Nasridinov, A., & Yoo, K.-H. (2022). Prediction of Process Quality Performance Using Statistical Analysis and Long Short-Term Memory. *Applied Sciences, 12*(2), 735. https://doi.org/10.3390/app12020735

[25] Aburass, S., Dorgham, O., & Al Shaqsi, J. (2024, December). A hybrid machine learning model for classifying gene mutations in cancer using LSTM, BiLSTM, CNN, GRU, and GloVe. *Systems and Soft Computing, 6*, 200110. https://doi.org/10.1016/j.sasc.2024.200110

[26] Esraa, R. A. (2024, September). Improving Industrial Quality Control by Machine Learning Techniques. *Journal La Multiapp, 5*(5), 692–711.

Authors' contacts:

Jaime Antonio Cancho Guisado, Mg.
(Corresponding author)
Universidad Nacional Federico Villarreal, Graduate School,
Jr. Prolongations Camaná N°1014, Lima 01, Perú
(511) 998945990,
jcancho@unfv.edu.pe

Ciro Rodriguez Rodriguez, Dr.
Universidad Nacional Mayor de San Marcos,
Lima, Perú
(511) 990427136
crodriguezro@unmsm.edu.pe

Ivan Carlo Petrlik Azabache, Dr.
Universidad Nacional Federico Villarreal,
Lima, Perú
(511) 992502130
ipetrlik@unfv.edu.pe

Milner David Liendo Arévalo, Mg.
Newman Escuela de Posgrado,
Tacna, Perú
(511) 841983001
Mdliendoa1051@gmail.com