# Applying Fuzzy Inference Systems in the Extraction of Chia Cake Extract: Predicting the Mass Yield

Gracielle Johann, Casi Santos dos Santos,
Paula Fernandes Montanher, Rafael Alves Paes de Oliveira
Federal University of Technology - Paraná
Dois Vizinhos, PR 85660-000, Brazil
grajohann@gmail.com, casi.ufma@gmail.com,
paulamontanher@utfpr.edu.br, raoliveira@utfpr.edu.br

Anderson Chaves Carniel
Department of Computer Science
Federal University of São Carlos
São Carlos, SP 13565-905, Brazil
accarniel@ufscar.br

***Abstract*—*Chia extract* has been increasingly used in the food industry since it is rich in bioactive compounds, such as fatty acids, omega-3 fatty, antioxidants, proteins, vitamins, minerals, and dietary fiber. This extract can be obtained by using conventional extraction techniques (e.g., pressure) on chia seeds. Unfortunately, such techniques are insufficient to access all chemical components present in the seeds matrix, producing a by-product named *chia cake* that is usually discarded. On the other hand, since chia cake contains significant nutraceutical properties, it is still viable and beneficial to perform extractions of chia extract from chia cake. A typical objective of an extraction is to gather a high *mass yield* of chia (cake) extract. Since the extraction process is complex and expensive (e.g., in terms of laboratory resources), there is an increasing interest in determining the mass yield based on variables of the extraction like temperature, extraction time, and solvent. In this paper, we study the viability of applying traditional fuzzy inference systems (e.g., based on Mamdani's method) and adaptive neuro-fuzzy inference systems (ANFIS) for this problem. We propose a fuzzy inference architecture that predicts the mass yield of chia cake extract based on temperature, extraction time, and solvent. Our architecture makes use of fuzzy sets and fuzzy rules in the context of fuzzy inference methods. To design them, we create and use a dataset that contains the mass yield of real extractions conducted in the laboratory under different configurations. Hence, it represents another contribution of this paper and serves as the needed foundation to build the proposed architecture. Further, we conduct a performance evaluation to choose the fuzzy inference system that better fits the architecture. Based on our analysis, ANFIS was the best inference method since it delivered the lesser errors and greater correlations between predicted and observed values. We conclude that fuzzy inference systems are powerful tools for the food industry since they can capture the intrinsic imprecise nature of the extraction process, model the existing non-linear relations of the variables, and represent the expert domain knowledge.**

***Index Terms*—Fuzzy inference systems, adaptive neuro-fuzzy inference systems, chia cake, mass yield, performance evaluation**

## I. Introduction

The extraction of *chia extract* from *chia seeds* (*Salvia hispanica L.*) has attracted the attention of the food industry since chia extract is rich in bioactive compounds, such as fatty acids, omega-3 fatty, antioxidants, proteins, vitamins, minerals, and dietary fiber [1]. Conventional extraction techniques (e.g., pressure extraction) are insufficient to access all chemical components present in the seeds matrix. This inefficient process results in the formation of a by-product named *chia cake* that is usually discarded. However, the extraction of chia extract from chia cake is still viable and benefits the food industry. Hence, this extraction provides economic value to an industrial by-product, extracts the nutraceutical properties of the chia cake, and reduces the disposal of products in the environment.

The extraction of chia extract from chia cake consists of three main steps. The first step is to dry the cake at a specific temperature, which causes changes in the composition and structure of the solid matter. This procedure guarantees the integrity of the cake (i.e., avoids biochemical degradation) by reducing the water activity. The second step is to apply a solvent (e.g., water, methanol, and hexane), which solubilizes and separates the extract from the chia cake. The last step is to perform ultrasound-assisted extraction in a specific duration. This process, due to cavitation, breaks the cell walls and facilitates the release of the extract. Hence, the variables temperature, solvent, and extraction time (i.e., duration) directly affect the *mass yield* obtained in the extraction. The mass yield is calculated by the ratio of the extracted mass (i.e., the mass extracted from the chia cake) and the initial mass (i.e., the mass of the chia cake).

From a productive point of view, there is an interest in gathering a high mass yield of chia cake extract. However, the determination of mass yield is challenging due to the following facts. First, the extraction process demands expensive laboratory resources, such as specialized equipment and supplies. Second, to calculate the mass yield at a specific time, the extraction is stopped and then the extracted mass is weighed up. This means that the chia cake cannot be reused in other extractions. For instance, the evaluation of mass yield at three different extraction times (e.g., 15, 45, and 60 min) requires three independent extractions (one from 0 to 15 min, another from 0 to 45 min, and lastly another from 0 to 60 min). Third, the temperature applied to the extraction is not precise and may vary during the drying process. Due to this, it is difficult to vary temperature values over a wide range to estimate mass yield. Finally, the relationship of temperature, extraction time, and mass yield is highly complex and non-linear, which makes it difficult to relate them physically in the laboratory.

The use of fuzzy set theory [2] and fuzzy inference systems [3] can mitigate the aforementioned problems and provide predictions of mass yield. Such systems can capture the intrinsic imprecise nature of the extraction process and model the existing non-linear relations by representing domain knowledge with fuzzy if-then rules. In this paper, we study the viability of applying fuzzy inference systems to determine the mass yield of chia cake extract when performing extractions from chia cake. For this, we pursue three main goals. The first goal is to adequately model and represent the uncertainty of the extraction process. We model fuzzy sets as the scope of each continuous variable of the extraction process (i.e., temperature, extraction time, and mass yield). By using such fuzzy sets that are in turn represented by linguistic values, we are able to design fuzzy if-then rules sets to express domain knowledge.

The second goal is to propose a fuzzy inference architecture that predicts the mass yield of the extraction process of chia cake extract. Given the temperature, extraction time, and solvent of an extraction process, the proposed architecture makes use of our fuzzy sets and fuzzy rules in a fuzzy inference method to return the predicted mass yield. The use of fuzzy methods has been a promising approach when applied to other organic matrices, such as seeds of almond [4], sandbox [5], pomegranate [6], and dragon fruit peel [7]. In this paper, we expand the applicability of fuzzy inference methods to another organic matrix that is important for the food industry.

To choose the fuzzy inference method that better fits into our architecture, the last goal of this paper consists of evaluating different fuzzy inference systems and adaptive neuro-fuzzy inference systems (ANFIS) modeled for the problem. For this, we evaluate the Mamdani and Larsen inference methods [3], and the ANFIS architecture [8] with the Takagi-Sugeno-Kang (TSK) method [9], [10]. This evaluation employs a real dataset, produced in laboratory, that contains mass yield for different input values. This dataset is also utilized to optimize ANFIS architecture.

This paper is organized as follows. Section II surveys related work. Section III provides an overview of our architecture. Section IV presents the real dataset and explains the knowledge base. Section V discusses the use of fuzzy inference systems in the architecture. Finally, Section VI concludes the paper and presents future work.

## II. Related work

To the best of our knowledge, there is a lack of studies that aim at predicting the mass yield of chia cake extract by employing fuzzy set theory and fuzzy logic. However, a few studies in the literature have applied fuzzy approaches to improving the extraction process of other organic matrices. In this section, we discuss these studies and sketch that our work provides additional insight into extraction processes in the food industry by using fuzzy approaches.

The work in [6] compares simulation systems for predicting the mass yield of pomegranate oil from super-critical extraction. The pomegranate seed contains oil with lipids on a dry basis ranging from 66g to 193g per kg of fruit. By using temperature and pressure as inputs, the authors compared the performance of back-propagation neural networks, radial basis function neural networks, and ANFIS for their context.

In [7], the authors employ a fuzzy assessment method to determine better suitable conditions for microwave-assisted extraction of pectin from dragon fruit peels. This extraction considers power and heating time as inputs and measures yield and viscosity of pectin as outputs. The output values are converted into fuzzy performance grades that are later used to create a fuzzy performance grade matrix. This matrix, together with weight values, is employed to indicate the overall performance index of the extraction. The authors reveal that a microwave power of 450 W and an extraction time of 5 min was the best combination to their goal.

In [5], the authors compare ANFIS and the response surface methodology (RSM) to aid the solvent-based extraction of oil from *Hura crepitans*. *Hura crepitans* seeds are present in sandbox tree and it is known to be rich in non-edible oil with oil content that ranges from 36.4 to 72.2 wt%. In this comparison, ANFIS presented better results than RSM.

The authors in [4] make use of a fuzzy assessment method to support the solid-liquid-based Soxhlet extraction and drying pretreatment for the extraction of oil from almond seed powder. The method receives extraction time, temperature, moisture content, and solvent-to-sample ratio as inputs and infers oil recovery and stability index as outputs. The authors show that the inputs of the extraction can be incorporated into fuzzy inference systems for improving the accuracy of the predicted output values.

Based on the previous studies, we can note (i) the importance of adequately representing the uncertainty of the extraction process and its input parameters, and (ii) the essential use of previous training samples collected in laboratory to improve such inference systems. We not only apply such items but also extend them in our work by modeling a fuzzy inference architecture (Section III), and evaluating different fuzzy inference methods designed for our problem (Section V).

## III. An Overview of the Architecture

Figure 1 depicts our proposed architecture for the problem. Its goal is to return the predicted mass yield that can be obtained for a given combination of solvent, extraction time, and temperature informed by the expert user of the domain, such as a chemical engineer. This is also useful for discovering the best combination of solvent, extraction time, and temperature to maximize mass yield. For this, our architecture leverages four interacting components: (i) experimental base, (ii) knowledge base, (iii) fuzzy inference system, and (iv) evaluation process.

The experimental base consists of a real dataset that contains the mass yield of chia cake extract for some specific configurations of input values that were evaluated in laboratory (Section IV). We have built this dataset aiming to comprehend the extraction process and discover correlations that serve as a basis for designing our knowledge base and optimizing fuzzy inference systems. Further, this dataset is also used to validate the predictions performed by our architecture (Section V-A).

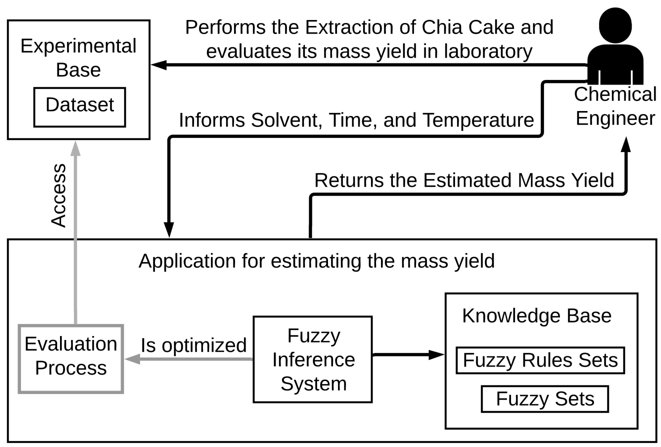


Fig. 1. Our proposed architecture to estimate the mass yield of chia cake extract based on solvent, extraction time, and temperature.

The knowledge base models the intrinsic fuzziness of the problem through fuzzy sets and fuzzy rules set. We define them according to experiences learned from the creation of our real dataset (Section IV). To design the fuzzy sets, we define a set of linguistic values for each variable that expresses continuous values (i.e., mass yield, extraction time, and temperature). Since our extraction is based on solvents with different characteristics, we deal with solvents separately. That is, for each solvent, we design a specific set of fuzzy if-then rules to be used by fuzzy inference systems.

We conduct an experimental evaluation of different fuzzy inference systems to choose the best inference method for the problem, and consequently, for our architecture (Section V). The evaluation process allows us to optimize ANFIS according to samples of our real dataset. This can be performed every time that the experimental base is changed (e.g., new values are inserted into the dataset).

## IV. Data Acquisition and Knowledge Base

In this section, we present our real dataset and discuss how it exerts influence in the definition of the knowledge base of the proposed architecture (Figure 1). The dataset contains the mass yield from a series of extractions conducted according to the following methodology. For the first step, we have dried the chia cake in a forced circulation oven (Model SL-102/64) at a specific *temperature*. Both fresh and dried samples were put in falcon tubes, with a ratio *solute/solvent* of 1:10, and into a Cristofoli ultrasound bath (frequency of 42 KHz) maintained at 30 °C. Finally, the ultrasound-assisted extraction was performed during a specific *period* (i.e., *extraction time*). We have varied the following parameters: (i) temperature values: *in natura* (∼30), 40, 50, 60, 70, and 80 °C, (ii) three solvents: hexane, methanol, and water, and (iii) extraction times: 15, 45, and 60 min. Each extraction was performed twice and the resulting mass yield was measured accordingly. Our dataset has 216 samples that are very representative to describe the different nuances for the problem since we employed a set of input values commonly used by the food industry.

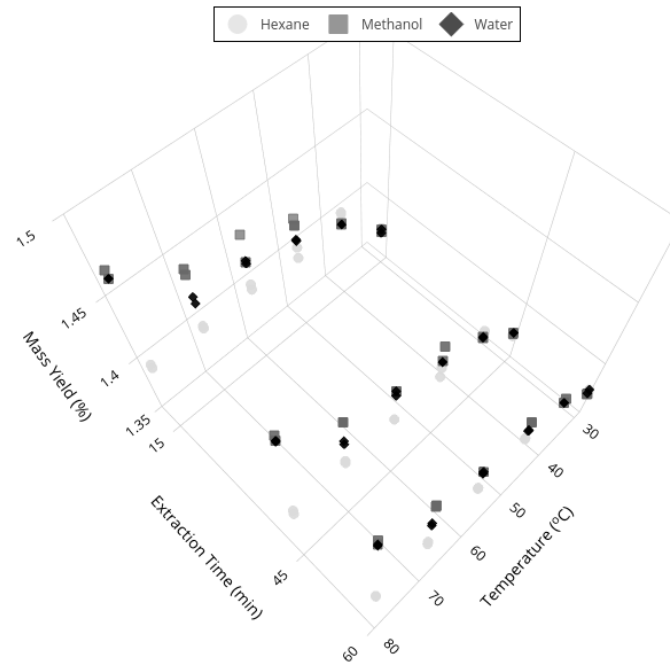


Fig. 2. 3D scatter plot that shows the relation between the variables of the problem: solvent, temperature, extraction time, and mass yield.

Figure 2 depicts, for each solvent, how the variables temperature and extraction time are related to the obtained mass yield. For the three solvents, mass yield stays around 1.5. Although the low ratio, it has enough content for the processes in industrial applications since chia cake extract has high added value and, without the extraction process, would be thrown away, as a worthless waste. For all three kinetics, the increase in mass yield is highly non-linear and more pronounced in the initial steps. The main reason is that there is more extract inside chia cake that can be extracted by the solvent. As the extraction time increases, the extract is reduced and becomes harder to be accessed by the solvent.

The three solvents applied in our extractions have distinct costs, environmental impact, and different capacities to access extract inside the chia cake solid matrix (due to each solvent polarity). We observe that the use of hexane, a non-polar solvent, gathered a lower mass yield than methanol and water, which are polar solvents. Based on that, the final decision about what experimental condition would be used (i.e., solvent, extraction time, and temperature) must be taken by considering the following combination: (i) mass yield, (ii) economical study, and (iii) environmental analysis. Hence, our dataset is another important contribution of this paper and advances the development of expert systems for the food industry.

Based on our dataset, we design fuzzy sets and fuzzy rule sets. Figure 3 shows the fuzzy sets and their linguistic values for the linguistic variables temperature and extraction time. We define them by using common ranges applied in the food industry. We specify six linguistic terms for characterizing

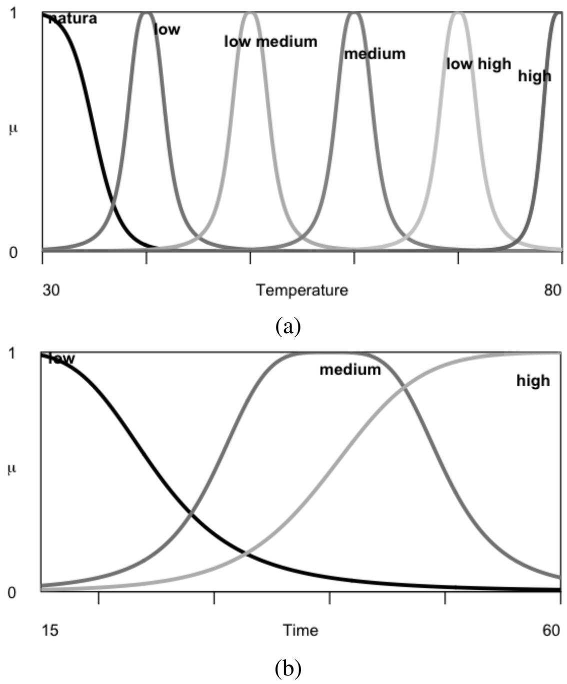


Fig. 3. Fuzzy sets for the knowledge base: temperature (a), and extraction time (b).

mass yield: *very low*, *low*, *low medium*, *medium*, *low high*, and *medium high*. We did not show their fuzzy sets since we evaluate different fuzzy inference systems that use distinct methods for defining the output values (see Section V-A).

By using the linguistic values, we design three sets of fuzzy rules, one set for each solvent. We define 12 rules for water, 19 for methanol, and 14 for hexane. Each solvent has its own set of fuzzy rules due to the behavior presented during the data acquisition. The antecedent of each rule is composed of the linguistic values defined for the input variables extraction time and temperature, whereas the consequent consists of the linguistic values specified for the output variable (mass yield). Some examples of fuzzy rules for water are:

IF **extraction time** is *high* AND **temperature** is *high*
THEN **mass yield** is *low high*
IF **temperature** is *medium*
THEN **mass yield** is *low medium*
IF **extraction time** is *high* AND **temperature** is *low*
THEN **mass yield** is *low*

## V. Evaluating Fuzzy Inference Systems for Predicting the Mass Yield

In this section, we conduct an evaluation of different fuzzy inference systems to choose the inference method for our architecture (Figure 1). Section V-A details the employed methodology and Section V-B discusses the obtained results.

### A. Experimental Setup

As discussed in Section IV, our knowledge base is defined by combining human expertise of the extraction process and exploratory data analysis of our dataset. Hence, the knowledge base provides the needed foundation to implement fuzzy inference systems for our architecture. As a first approach, we employ traditional fuzzy inference methods like the Mamdani's inference and Larsen's inference. For each fuzzy rule, these methods calculate the membership degrees of each input based on its corresponding linguistic value (i.e., fuzzy set). These degrees are combined, according to the logical operator of the antecedent part, to determine the degree of fulfillment of the rule (i.e, firing strength of the rule). This resulting degree is used by the *implication* operator, which reshapes the fuzzy set of the consequent. The resulting implications are combined by the *composition* operator, which yields the reasoning conclusion. While the Mamdani's inference uses the minimum operator for the implication, the Larsen's inference deploys the product operator. Both inference methods use the maximum operator for the composition and require a defuzzification technique.

We further make use of ANFIS with the TSK method in our evaluation. ANFIS combines fuzzy inference systems with adaptive-neural networks to build a feed-forward network with five layers. The first layer consists of adaptive nodes that adjust the parameters of the membership functions of the antecedent part of the rules. The second layer, which is formed by fixed nodes, computes the degree of fulfillment of the rules. The third layer normalizes the resulting degrees and is also composed of fixed nodes. The fourth layer represents the consequent part of the rules as adaptive nodes. In the TSK, it adjusts the parameters of the linear combination of the inputs and a constant value. The last layer consists of a fixed node that aggregates the outputs by computing the weighted sum. The training phase of ANFIS is responsible for tuning the parameters of the membership functions in the antecedent and the parameters of the consequent part. For this, a hybrid learning algorithm of the gradient method and the least-squares estimate is often deployed.

We employ FuzzyR [11], an R package that extends the work in [12] and contains a set of functions for handling fuzzy inference systems. For ANFIS, FuzzyR offers a function that optimizes a fuzzy inference system based on the TSK method [13], [14]. We use FuzzyR to implement four fuzzy inference systems by using the knowledge base described in Section IV. The first two systems are based on the Mamdani and Larsen methods. Thus, we vary the implication operator. We deploy the centroid as defuzzification technique for these systems. The remaining systems are based on ANFIS; we call them *ANFIS 1* and *ANFIS 2*. *ANFIS 1* uses the product operator for evaluating the antecedent part of the rules, whereas *ANFIS 2* employs the minimum operator. Since the output of a fuzzy rule in the TSK method can be a polynomial function of inputs, FuzzyR allows us to specify a linguistic value for a given linear combination of the inputs and their coefficients (e.g., two parameters in our case) with a constant value (see *linearmf* for more details[1]). Hence, we can assign, for each

[1]https://cran.r-project.org/web/packages/FuzzyR/index.html

TABLE I
EMPLOYED ACCURACY MEASURES BY OUR ANALYSIS.

| Measure | Definition | Description |
|---|---|---|
| MAE | $\frac{1}{n}\sum_{i=1}^{n}\lvert\check{y}_i - y_i\rvert$ | It is the mean absolute difference between the predicted and observed values. |
| RMSE | $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \check{y}_i)^2}$ | It is the square root of the average squared difference between the predicted and observed values. |
| MAPE | $\frac{1}{n}\sum_{i=1}^{n}\lvert\frac{y_i-\check{y}_i}{y_i}\rvert$ | It is the simple average of absolute (percentage) errors. |
| $R^2$ | $1 - \frac{\sum_{i=1}^{n}(y_i-\check{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})}$ | It is the coefficient of determination that measures the proportion of the variance in the dependent variable that is predictable from the independent variable(s). |

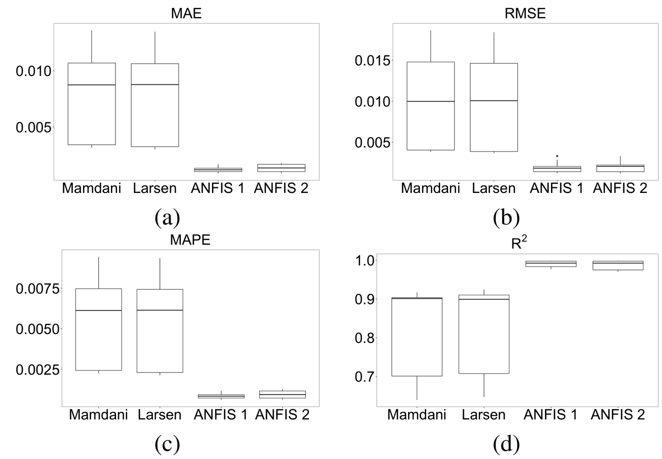


Fig. 4. Box plots that show the variability of accuracy measures after executing the repeated holdout validation for each fuzzy inference system: MAE (a), RMSE (b), MAPE (c), and $R^2$ (d).

solvent, specific linear combinations to the six linguistic terms that characterize mass yield. Similarly to [11], we use the hybrid learning algorithm for training *ANFIS 1* and *ANFIS 2*, set the number of epochs to 100, and assign 0.01 to step size.

We compare the fuzzy inference systems by using the holdout validation executed 5 times (i.e., repeated holdout validation). Our training set had 80% of the entries of our dataset, whereas the test set contained 20%. The entries were randomly selected for each set. Further, each execution of the validation led to different partitions of the dataset. For evaluating the Mamdani and Larsen methods, we have used only the test set since they do not have a training phase. Hence, we guarantee that all fuzzy inference systems were evaluated by using the same test set in a particular execution.

To conduct our evaluation, we analyze the performance of the fuzzy inference systems by computing the accuracy measures shown in Table I. In this table, $y$ means the observed value, $\check{y}$ refers to the predicted value, and $n$ is the number of observations. These measures are commonly employed in the food industry, such as the studies summarized in Section II. We use them to compare the predicted values and the measured values of mass yield (i.e., observed values) in the test set.

### *B. Performance Results*

Figure 4 shows the obtained accuracy measures of each fuzzy inference system after executing our repeated holdout validation. Since this kind of validation provides information about the stability and variability, in this figure, we employ box plots to report the results of the computed accuracy measures. That is, the box plots allow us to analyze how each fuzzy inference system modeled for the problem changes its performance with different training sets.

We observe a standard behavior of the fuzzy inference systems when graphically analyzing their quartiles in each box plot of the accuracy measures. First, according to the meaning of the accuracy measures, *ANFIS 1* and *ANFIS 2* presented the best results. Considering the obtained results for MAE, RMSE, and MAPE, these promising systems showed that they

TABLE II
AVERAGE VALUES OF THE ACCURACY MEASURES FOR EACH SOLVENT AND FUZZY INFERENCE SYSTEM.

| Solvent | Measure | Fuzzy Inference System | | | |
|---|---|---|---|---|---|
| | | Mamdani | Larsen | ANFIS 1 | ANFIS 2 |
| Water | MAE | 0.0118 | 0.0117 | 0.0010 | 0.0010 |
| | RMSE | 0.0163 | 0.0161 | 0.0013 | 0.0013 |
| | MAPE | 0.0082 | 0.0081 | 0.0007 | 0.0007 |
| | $R^2$ | 0.6699 | 0.6772 | 0.9978 | 0.9978 |
| Methanol | MAE | 0.0089 | 0.0089 | 0.0014 | 0.0014 |
| | RMSE | 0.0101 | 0.0102 | 0.0026 | 0.0026 |
| | MAPE | 0.0062 | 0.0063 | 0.0010 | 0.0010 |
| | $R^2$ | 0.8997 | 0.8979 | 0.9931 | 0.9931 |
| Hexane | MAE | 0.0033 | 0.0031 | 0.0013 | 0.0017 |
| | RMSE | 0.0039 | 0.0038 | 0.0018 | 0.0021 |
| | MAPE | 0.0024 | 0.0022 | 0.0009 | 0.0012 |
| | $R^2$ | 0.9082 | 0.9161 | 0.9807 | 0.9807 |

can capture the nuances of the extraction process of chia cake extract by predicting values of mass yield with a low error (i.e., very near to 0). While MAE and RMSE values express the average errors of predicted values, MAPE measures the accuracy in terms of relative errors (i.e., relative deviation). The lower (greater) value of MAE, RMSE, and MAPE is, the better (worst) the accuracy of the system is. On the other hand, $R^2$ is a measure of the proportion of the variance explained by the system. An $R^2$ value of 1.0 implies a perfect fit. *ANFIS 1* and *ANFIS 2* stand out by delivering values near to 1.0 for this accuracy measure, which means that the predicted values are highly correlated to the observed values. Second, *ANFIS 1* and *ANFIS 2* distinguished themselves by showing that their accuracy measures' values have lower variability than the values obtained by *Mamdani* and *Larsen*. This means that *ANFIS 1* and *ANFIS 2* led to results that do not highly vary if compared to their competitors. As a consequence, we can consider that *ANFIS 1* and *ANFIS 2* showed more stability in the accuracy results.

To comprehend the performance of the fuzzy inference systems, Table II depicts the average value of the five executions of the holdout for each accuracy measure. In this table, we split the results for each solvent since the extraction process is oriented by solvent and we formulated specific fuzzy rules for each type of solvent. This table allows us to better visualize the gains of *ANFIS 1* and *ANFIS 2* compared to *Mamdani* and *Larsen* for each solvent. These systems provided very similar results for water and methanol, whereas *ANFIS 1* was slightly better than *ANFIS 2* if we consider the values for MAE, RMSE, and MAPE. *Mamdani* and *Larsen* showed better results when performing predictions for hexane. We believe that this is related to the behavior of kinetics observed when employing this solvent in the extraction process (as shown in Figure 2). Considering the measures calculated based on errors, we can note that all fuzzy inference systems delivered very low deviation. We highlight *ANFIS 1* and *ANFIS 2* since they presented the lowest values. Another behavior is related to $R^2$. Studies in the literature (e.g., [15]) recommend a value greater than 0.9 for $R^2$ for expert systems since this value shows a strong correlation between predict values and observed values.

The final decision of which fuzzy inference system should be used in real scenarios is often performed by taking into account the combination of four accuracy measures. This means that we should deploy the system that shows the lesser errors (MAE, RMSE, and MAPE) and greater correlations ($R^2$). In our architecture (Figure 1), therefore, the experiments suggest the use of *ANFIS 1*, independently of the solvent employed in the extraction process.

## VI. Conclusions and Future Work

In this paper, we have studied the applicability of fuzzy inference systems to predict the mass yield of the extraction process of chia extract from chia cake. For this, we proposed an architecture that leverages four components. The first component consists of a dataset containing the mass yield of real extractions conducted in laboratory under different configurations. In this dataset, we varied the following parameters: (i) temperature, (ii) extraction time, and (iii) solvent. The second component is composed of fuzzy sets and fuzzy rules sets that model the intrinsic fuzziness of the problem based on the experiences learned from the creation of our dataset. The third component is the fuzzy inference system, which predicts the mass yield of extractions based on the solvent, temperature, and extraction time informed by the user (e.g., chemical engineer). The last component refers to possible optimizations performed on the fuzzy inference system.

To choose the best fuzzy inference system for our architecture, we have evaluated two traditional fuzzy systems (*Mamdani* and *Larsen*) and two ANFIS methods (*ANFIS 1* and *ANFIS 2*). In our evaluation, *ANFIS 1* and *ANFIS 2* presented the best overall results by showing the lowest values for MAE, RMSE, and MAPE, and the biggest values for $R^2$. If we analyze the results of each solvent, *ANFIS 1* distinguished itself since it delivered slightly better accuracy results compared to *ANFIS 2*. Hence, based on our analysis, the best candidate for our architecture is *ANFIS 1*.

We conclude that this work presented a successful combination of fuzzy inference systems and domain knowledge to develop an expert system for solving a problem in the food industry. Further, recall that chia cake is a by-product of the extraction of chia extract from chia seeds. Hence, this paper also contributes to the development of expert systems that aim to (i) provide economic value to an industrial by-product, (ii) obtain nutraceutical properties of the chia cake, and (iii) reduce the disposal of products in the environment. With our architecture, it is possible to suggest an experimental condition (i.e., solvent, temperature, extraction time) based on the necessity of the user (e.g., mass yield, the economic cost of the extraction based on the employed supplies).

Future work will deal with two main topics. First, we aim to evaluate type-2 fuzzy inference systems [14]. Second, we will analyze the impact of automatically generating the fuzzy rules of the fuzzy inference systems.

## References


[1] R. Ullah, M. Nadeem, A. Khalique, M. Imran, S. Mehmood, A. Javid, and J. Hussain, "Nutritional and therapeutic perspectives of Chia (*Salvia hispanica* l.): a review," *Journal of Food Science and Technology*, vol. 53, no. 4, pp. 1750–1758, 2016.

[2] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[3] C.-C. Lee, "Fuzzy logic in control systems: Fuzzy logic controller, part ii," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 419–435, 1990.

[4] D. J. A. Mukherjee, K. Roy and S. Hossain, "Qualitative model optimization of almond (*Terminalia catappa*) oil using soxhlet extraction in type-2 fuzzy environment," *Soft Computing*, vol. 24, pp. 41–51, 2020.

[5] B. S. R.O. Omilakin, A.P. Ibrahim and B. E, "Process modeling of solvent extraction of oil from *Hura crepitans* seeds: adaptive neuro-fuzzy inference system versus response surface methodology," *Biomass Conversion and Biorefinery*, pp. 1–14, 2020.

[6] J. Sargolzaei and A. Moghaddam, "Predicting the yield of pomegranate oil from supercritical extraction using artificial neural networks and an adaptive-network-based fuzzy inference system," *Frontiers of Chemical Science and Engineering*, vol. 7, no. 3, pp. 357–365, 2013.

[7] P. L. N. Tongkham, B. Juntasalay and N. Sengkhamparn, "Dragon fruit peel pectin: Microwave-assisted extraction and fuzzy assessment," *Agriculture and Natural Resources*, vol. 51, pp. 262–267, 2017.

[8] J. S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.

[9] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.

[10] M. Sugeno and G. T. Kang, "Structure identification of fuzzy model," *Fuzzy Sets and Systems*, vol. 28, no. 1, pp. 1–33, 1983.

[11] C. Chen, T. R. Razak, and J. M. Garibaldi, "FuzzyR: An extended fuzzy logic toolbox for the R programming language," in *IEEE Int. Conf. on Fuzzy Systems*, 2020, pp. 1–8.

[12] C. Wagner, S. Miller, and J. M. Garibaldi, "A fuzzy toolbox for the R programming language," in *IEEE Int. Conf. on Fuzzy Systems*, 2011, pp. 1185–1192.

[13] C. Chen, R. John, J. Twycross, and J. M. Garibaldi, "An extended ANFIS architecture and its learning properties for type-1 and interval type-2 models," in *IEEE Int. Conf. on Fuzzy Systems*, 2016, pp. 602–609.

[14] ——, "Type-1 and interval type-2 ANFIS: a comparison," in *IEEE Int. Conf. on Fuzzy Systems*, 2017, pp. 1–6.

[15] Y. Zhang, L. Tang, P. L. Croteau, O. Favez, Y. Sun, M. R. Canagaratna, Z. Wang, F. Couvidat, A. Albinet, H. Zhang, J. Sciare, A. S. H. Prévôt, J. T. Jayne, and D. R. Worsnop, "Field characterization of the $PM_{2.5}$ aerosol chemical speciation monitor: insights into the composition, sources, and processes of fine particles in eastern china," *Atmospheric Chemistry and Physics*, vol. 17, no. 23, pp. 14 501–14 517, 2017.