Generation of Fuzzy Expert Systems via Machine Learning Methods

Anna Hovakimyan
Yerevan State University
Yerevan, Armenia
e-mail: ahovakimyan@ysu.am

Siranush Sargsyan
Yerevan State University
Yerevan, Armenia
e-mail:siranushs@ysu.am

Elen Soghomonyan
Yerevan State University
Yerevan, Armenia
e-mail:elensoghomonyan@gmail.com

***Abstract*—This work considers one of the automatic construction approaches for fuzzy expert systems. It is based on fuzzy sets, fuzzy logic theory, and machine learning methods.**

The problem is to identify patterns in massive data, build expert knowledge, and use them to create fuzzy logical rules and a fuzzy inference system that combines the special expert system. The effectiveness of the resulting system is evaluated using well-known metrics.

The proposed method is implemented in the Python programming language. The generated system is tested on data taken from the website kaggle.com and medical data about osteoporosis.

***Keywords*—Expert Systems, Fuzzy Logic, Machine Learning, Fuzzy Expert Systems.**

I. Introduction

The concept of an expert system refers to the area of artificial intelligence. The expert system is a computer system that attempts to replicate the decision-making ability of a human expert [1].

There are a variety of formulations of expert systems as well, in particular, according to the “Expert System is a computer program which solves problems that previously required specialists vast experience applying presented industry knowledge and computerized decision process” [1].

Expert systems are composed of the following components: user interface, knowledge base, and inference system.

The user interface is designed to communicate with the user and to collect information from him/her. The information collected later serves as the basis for the system to draw situational conclusions [1].

A knowledge base is a set of facts and rules based on which the conclusion is made. The knowledge base is created from known facts, through expert evaluations and rules.

The rule of expert system mostly stands for

if *cause* ***then*** *effect*

format. The inference engine tool compares user-supplied information with the information in the knowledge base and as a result, provides conclusion(s) based on what is available in the knowledge base of the rules.

The knowledge base and inference tool form expert key components of systems, and therefore their studies are particularly important for building an efficient system.

Since in the reality around us, we often encounter fuzzy concepts, the presence of fuzzy expert systems is in demand. In such systems, both knowledge in databases and inference systems are based on the theory of fuzzy sets and relations, as well as on the theory of fuzzy logic [2].

Here we consider components of expert systems and non-specific one’s relationships between fuzzy sets. Inference in fuzzy expert systems is implemented through the FIS (Fuzzy Inference System) [2]. A fuzzy inference system uses fuzzy set theory to reflect (match) input data to output data [2].

Fuzzy inference systems have the following structure: fuzzy knowledge base, inference engine, fuzzification, and defuzzification mechanisms. As input and output data are crisp values.

II. Building a fuzzy expert system from data

In this paper, we present a machine learning approach to construct a fuzzy expert system based on data [4]. This kind of learning we consider as supervised learning and a problem solved by an expert system as a regression problem. So the observed data are divided into training and testing datasets [4].

Based on the input and output data in the training dataset, we will get a fuzzy rules set that represents the relationship between input and output variables in the dataset. For each of these variables, we introduce a linguistic variable for which it is necessary to construct fuzzy terms and sets [2].

The algorithm for constructing fuzzy rules is as follows [5]. The ranges of values of input/output variables are divided into non fuzzy sets, then to each set of input/output data some fuzzy region with a membership function is assigned. In the paper, we consider data with two input variables x_1 and x_2, and one output variable y(Fig. 1).

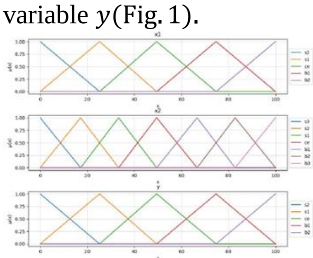

Fig.1. The construction of membership functions for two input and one output variable

Then fuzzy rules from the data are derived.

A fuzzy rule looks like as follows:

if *antecedent clauses* ***then*** *consequent clauses*

We consider a data sample in the training set as a labeled data: the output value is a label in a regression model [4]. For each input and output value in the data sample, we are determining via the membership function the degree of this value belonging to the appropriate section and fixing the membership largest degree.

For example, let's consider the following sample (Fig. 2)

$$\boldsymbol{samlple1}(\boldsymbol{x_1 = 82, x_2 = 25, y = 55})$$

We assign the linguistic variable to the linguistic term with the largest membership function value, namely x_1 to the term b_1 marked red, x_2 to the term s_1 marked green, and y to the term ce marked blue. So we get the rule in the following form

if x1 is b1 and x2 is s1 then y is ce

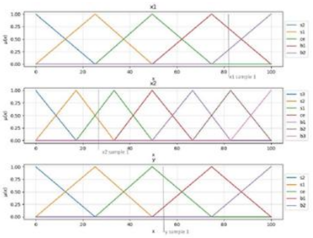

*Fig.2. The construction of a fuzzy rul*e

Then a Wang-Mendel degree to each rule is assigned by multiplying the values of memberships functions according to the terms being in the antecedents as well consequents of the generated rule [5].

So we will get the weight of the first rule

$$Degr(Rule1) = m_{b_1}(x_1) \times m_{s_1}(x_2) \times m_{ce}(y) =$$
$$= 0.8 \times 0.7 \times 0.9 = 0.504$$

Note that a fuzzy rule is built for each sample in the dataset, and the algorithm may build a large number of rules. Furthermore, it is possible to generate conflicting rules, that is, rules that have the same antecedent clauses but different consequent clauses. In these cases, we have to make some optimization. We compare the degree of the rules and keep the rule that has the highest degree.

It is also possible to add human knowledge to the knowledge base built from data. That can be made by multiplying the weight of the rules by some factor.

III. THE INFERENCE IN FUZZY EXPERT SYSTEM

The inference process in the fuzzy system is the determination of the crisp value of the output linguistic variable (y in our example) for given crisp values of input linguistic variables via the Wang and Mendel method [5].

For a given combination of values of input linguistic variables, the antecedents of a given rule are combined to determine the degree of output linguistic variable using the product operator.

So in our example, for each input (x_1, x_2) to determine the weight m_o^i of the output variable o in the i^{th} rule we use the multiplication operator using antecedent segments.

Let the i^{th} rule be

$$\boldsymbol{if\ x_1\ is\ l_1\ and\ x_2\ is\ l_2\ then\ o\ is\ l,}$$

then

$$\boldsymbol{m_o^i = m_{l_1}(x_1) \times m_{l_2}(x_2),}$$

where o is the output linguistic variable, l_1 and l_2 are linguistic terms, and (x_1, x_2) is some combination of input values of linguistic variables.

Then the centre $\underline{y^i}$ of the fuzzy region for the i^{th} rule as the point that has the smallest value of all points for which the membership function value is equal to 1 is defined.

The crisp value of the output variable y is calculated via the formula

$$y = \frac{\sum_{i=1}^{K} m_o^i\, \underline{y^i}}{\sum_{i=1}^{K} m_o^i}$$

where K is the number of rules contained o in the consequent.

IV. THE TESTING AND MEASURING SYSTEM PERFORMANCE

To evaluate the generated system, we use the training as well as the testing data [4]. To be able to evaluate the resulting system, we will use the coefficient of determination R^2 [6].

$$sum\_of\_square_{\_total} = \sum_{i=1}^{N} (y_i - \hat{y})^2$$

$$sum\ \_of\_square_{residual} = \sum_{i=1}^{N} (y_i - \underline{y^i})^2$$

where N is the number of data samples in the training or testing set, y_i is the output in the i^{th} data sample, $\hat{y}$ is the arithmetic mean of the dependent variable, and $\underline{y^i}$ is the calculated value of the output variable for the i^{th} sample.

The R^2 we define as follows:

$$R^2 = 1 - \frac{sum\_of\_square_{residual}}{sum\_of\_square_{total}}$$

Note that R^2 will take a value in [0, 1] interval, and the bigger the better. If $R^2 = 1$, this means that the calculated values will be equal to the real ones. The value of R^2 calculated on training and testing data points to overfitting or underfitting of the machine learning system [4].

VI. CONCLUSION AND FURTHER WORK

The conducted research shows that the approach to the development of expert systems, based on fuzzy methods of data analysis, is acceptable for solving practical problems. At the same time, such expert systems are more flexible decision support systems compared to clear expert systems.

In the future, it is planned to collect data online based on the results of laboratory examinations of patients to enrich the training data bank and create an expert system as an additional tool for a doctor to diagnose a disease.

REFERENCES

[1] M. Gondran, *An introduction to expert systems*, McGraw-Hill, New York, 1986.

[2] E. Trillas, L. Eciolaz,, *Fuzzy Logic: An Introductory Course for Engineering Students*. Springer, New York, 2015.

[3] W. Siler, *Fuzzy expert systems and fuzzy reasoning*, Wiley, New York 2005.

[4] Aurélien Géro, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*, Second Edition, O'Reilly, Sebastopol, 2019.

[5] L. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples", *IEEE Transactions on Systems, Math, and Cybernetics*, vol. 22, no. 6, pp. 1414-1427, 1992, doi: 10.1109/21.199466,

[6] Orozco-Arias, Johan S. Piña, R. Tabares-Soto., "Measuring Performance Metrics of Machine Learning.Algorithms for Detecting and Classifying Transposable Elements*", Processes*,vol.8, no.6, pp.638-656, 2020, doi:10.3390/pr8060638.