Theses of the PhD thesis

The Applicability of Fuzzy Theory in Machine Learning

Abrar Hussain

Supervisor:

József Dombi, PhD, Professor

Doctoral School of Computer Science
University of Szeged

Department of Computer Algorithms and Artificial Intelligence

2022

Introduction

In this dissertation, we examined the use of fuzzy membership functions and fuzzy operators in the domain of machine learning. We first developed an arithmetic-based Fuzzy Inference System (FIS) which is computationally efficient than the existing FISs. The design was then extended to produce arithmetic-based type2 FIS. In the next step, data-driven algorithms were developed to build these type1 and type2 arithmetic-based FISs directly from the system data. This eliminated the dependency on the expert knowledge for identifying the important rules for FIS. In the final step, we replaced the neuron in the neural network with the Rule-based Neuron (RBN). The RBN is a neuron with built in arithmetic-based FIS. This led to the development of the Rule-based Neural Network (RBNN). Compared to the Deep Neural Network (DNN), the RBNN is more interpretable, robust to feature noise and produce better results (higher F1 score) on skewed datasets. Next, we describe each of these contribution in detail:

Arithmetic-based Type1 Fuzzy System

Based on the Fuzzy Rule Inference (FRI), various types of Fuzzy Inference Systems (FISs) have been developed. The two best known are the Mamdani and model-based Takagi Sugeno (TS) type FISs. There are however some drawbacks with these two inference techniques. These are:

1. The input space is not completely covered by the membership functions i.e. they cover only a limited subspace. For example if we have two inputs, each with 7 categories, then 49 rules

are required to cover the whole input space. If the number of input variables increases from two (working in a higher dimensional space), then the problem grows exponentially.

2. Most of the membership functions are not analytical i.e. the derivative is not defined at every point and higher derivatives do not exist. This is a drawback because the gradient-based optimization techniques cannot be used to tune the parameters of these membership functions.

3. It is not clear how to choose a fuzzy operator system for the antecedent part of a rule. The choice of membership function and operator system is completely arbitrary and we cannot get a proper efficient design.

4. Different implication operators are introduced. Surprisingly, the product operator is mainly used. The product operator is a strict t-norm and it is not an implication operator.

5. The result of an evaluation of the consequent part of the rule is not a membership function (it is an α-cut of the membership function).

6. Centre of gravity (COG) defuzzification usually involves the integral evaluation of the aggregated function and it is computationally expensive.

Using these techniques, designing an adaptive FIS is a challenging task. As we mentioned above, both of these techniques have some advantages and disadvantages. There is a need to combine these advantages into a single design approach. We attempted to solve these issues using a new approach, which has the following good features:

1. A new type of parametric membership function called the Distending Function (DF) is introduced. With a few rules, it can cover the whole input space. It has three parameters and each has a semantic meaning. It has two types, namely the symmetric and asymmetric DF and both can be utilized for developing an FIS.

2. The DF is analytical i.e. higher derivatives exist at each point. This property is used in optimization procedures to tune the parameters of the DF.

3. The Generalized Dombi Operator (GDO) is used for evaluating the antecedent part of the rule. The GDO and the DF are consistent with each other.

4. Our approach does not involve the implication step. Instead the activation strength of each rule is multiplied by the consequent DF to get the fuzzy output of each rule.

5. The consequent of each rule is a DF. Aggregation is carried out using the weighted arithmetic mean of these consequent DFs of all the rules. A linear combination is closed for DFs and so the result of an aggregation is also a DF.

6. Defuzzification in this case is only a single-step calculation (finding the point that has the highest value of the aggregated DF).

7. Using our proposed approach, we designed an adaptive FIS. It consists of tuning the DF parameters using gradient descent optimization. The adaptive FIS can handle the changing process dynamics.

We combined the advantages of Mamdani and TS methods and developed the arithmetic-based FIS. In our study, the symmetric and asymmetric DFs are used. The asymmetric DF provides more flexibility in adaptive controller design. The efficiency of this new approach was shown by designing an adaptive control system for a water tank level and vehicle lateral dynamics.

Arithmetic-based Type2 Fuzzy System

Compared to type1, the type2 fuzzy systems are better at handling uncertainties, produce smoother output response, more adaptive and use a smaller rule base. For practical and computational reasons, interval type2 fuzzy systems were introduced. The design of interval type2 fuzzy system consists of five steps: 1) Fuzzification of the inputs using type2 membership functions; 2) Calculation of rules firing strengths; 3) Implication and aggregation is used to produce the outputs. These operations also produce also a type2 fuzzy set; 4) Type reduction is applied to convert type2 fuzzy set into type1 fuzzy sets; 5) Defuzzification is performed to get the crisp output value. This process is similar to design of type1 fuzzy system but here we have type reduction as an additional step. This step converts type2 fuzzy sets to type1 fuzzy sets. The type reduction is achieved using the so called Karnik Mendel (KM) iterative algorithm. There are some drawbacks in the above-mentioned approach. These are:

1. The choice of type2 membership function and its systematic connection with the uncertainty are not clear. Different typeI membership functions can be combined to generate type2 membership function. And it is not clear which type of membership

functions should be used for particular uncertainty case.

2. The type reduction step is based on the KM algorithm, which is computational expensive. Due to its iterative nature, it is not suited for on-line applications.

3. Although type2 Fuzzy Logic System (FLS) require fewer rules compared to type1 fuzzy systems, but the number of parameters is comparatively large. So the optimization is not easy in this case.

4. The implication and aggregation steps also increases the computation complexity of the type2 FLS.

Here, we solved some of these issues by proposing a new type of interval type2 FLS. It overcame these issues using the following unique features:

1. A type2 extension of the DF called the Type2 Distending Function (T2DF) is proposed. Different types of uncertainties can be expressed by associating it with the parameters of T2DF.

2. Fuzzy arithmetic approach is utilized here for designing type2 fuzzy logic controller. So it has no type reduction step and it does not require the iterative algorithms. It is simple and suitable for on-line implementations.

3. Most of the parameters of the T2DF are fixed. Usually just the parameter associated with the uncertainty is varied. Therefore the optimization process is easy to perform.

4. There are no implication, aggregation and type reduction steps. Therefore it is computationally fast.

Because of these features, the proposed approach provides a complete framework for handling the uncertainty and noise using type2 fuzzy systems.

Data-Driven Arithmetic-Based Fuzzy Type1 System

A rule base system represents the human expertise in the form of linguistic rules. These rules describe the dependencies between the input and output variables in the form of IF-THEN statements. As the number of input variables increases, the required number of rules and the system complexity increases exponentially. In most cases, expert knowledge is not available or it is poorly described. So the exact description of fuzzy rules is not an easy task. If the working data of the process is available then a data-driven based design is an attractive option. In this case, the problem reduces to identifying a suitable fuzzy model which fits the given data.
The data-based identification of a fuzzy model can be divided into two parts namely; qualitative and quantitative identification. Qualitative identification focuses on the number and description of fuzzy rules. Quantitative identification is concerned with the identification of parameter values. In the case of qualitative identification, soft computing methodologies are used like evolutionary algorithms, genetic algorithms and swarm optimization etc. Neural networks are mostly used for learning parameters (the quantitative part). Combining these two leads to the development of a Adaptive Neuro-Fuzzy Inference System (ANFIS).
The above mentioned approaches however have some drawbacks:

1. Qualitative identification:

The identified rule base has a so-called flat structure (curse of dimensionality) problem. To cover the input space completely, a huge number of rules are required. Each rule is applicable only within a specific area and its strength is zero outside. If the training data of the system does not fully span the input and output space, then this will cause serious problems when modeling the system. If the input falls in these uncovered areas then the identified rule bases do not generate any action.

2. Quantitative identification:
   The computation complexity of the quantitative part of the identified fuzzy model also increases with the number of rules. As the number of rules increases, the number of parameters of the membership function and operators also grow exponentially.

3. Interpretability:
   In most cases, the interpretability of the identified fuzzy rule base is not clear. It is easier to interpret a few rules and get an insight into the working model. However, if the number of rules grows exponentially, then for a given set of input values, it is not possible to predict the response of the model and analyze its performance.

4. Complexity of control design:
   The identified rule base is used to generate a crisp output using the Mamdani or TS inference engines. Both of these are computationally expensive due to implication, aggregation requirements and defuzzification step.

In this study, we propose a new methodology for a data-driven fuzzy system design. This new method has the following unique features:

1. DFs have been used to cover the input space entirely. A DF has a long tail and it is defined on [-∞, ∞]. The grade of membership always has a non-zero value and the whole input space can be covered. If we apply the Dombi operator on the DFs in the input space, the results is also a DF in the output (higher dimensional) space. So the whole area of the output space can be covered just using a few rules.

2. The DF has three parameters. Usually two parameters can be kept constant and one parameter is used for tuning. The latter increases/decreases the influence area of the DF function. Also, a single step calculation is required to calculate this parameter. Because of this, due to the smaller number of identified rules and the deterministic nature of single parameter, the computation complexity of the quantitative part is negligible.

3. The interpretability of the model increases due to the significant decrease in the number of fuzzy rules.

4. We used an arithmetic-based type1 FIS to generate the output signal. There is no implication operator involved and aggregation is based on arithmetic operations. The aggregation of fuzzy rules results in a single value that is used as a output signal. This single value arises from the defuzzification of consequent DFs, which is a single step calculation. So the overall complexity of the data-driven fuzzy controller design decreases significantly.

Data-Driven Arithmetic-Based Interval Type2 Fuzzy System

Fuzzy type2 modeling techniques are increasingly being used to model uncertain dynamical systems. However, some challenges arise when applying the existing techniques. A large number of rules are required to completely cover the whole input space. A large of parameters associated with type2 membership functions have to be determined and this leads to increased computation time and resources. The identified fuzzy model is usually difficult to interpret due to the large number of rules. Designing a fuzzy type2 controller using these models is also a computationally expensive task.

We presented the solutions to some of the limitations associated with the existing fuzzy type2 modeling and control techniques. A procedure was proposed to identify the type2 model directly from the data, which we called the Distending Function-based Fuzzy Inference System (DFIS) model. This model consists of rules and T2DFs. The whole input space is covered using a few rules. T2DFs can model various types of uncertainties using its parameters. A rule reduction procedure is also proposed. It combines the T2DFs in the close vicinity and it significantly reduces the number of rules. The DFIS model was compared with ANFIS type1, ANFIS type2 and various other fuzzy models. The DFIS model produced a smooth surface with comparatively fewer number of rules and tunable parameters. Furthermore, a procedure was proposed to design an arithmetic-based interval type2 fuzzy controller using the rules. As the controller design does not include the implication and type reduction steps, This greatly reduces the computational complexity and paves the way for the real-time implementation of the proposed design.

The effectiveness of the whole procedure was demonstrated by designing an altitude controller for the Parrot Mambo quadcopter. The proposed controller performed better than the ANFIS-based controllers and regulated the altitude of the quadcopter even in the presence of noisy (uncertain) sensor measurements. This robustness to noisy data is due to the use of T2DFs. The designed controller was then deployed and tested in the flight control system on the quadcopter hardware. Real-time hardware implementations produced the same results as those obtained in the simulations.

Rule-Based Neural Networks

DNNs and FISs seem to complement each other. It is the natural to expect that a combination of these two approaches might have the advantages of both. On the one hand, FISs can benefit from the computational learning procedures of the DNN. The parameters of the rule (sometimes the whole rule) can be learnt directly from the data. On the other hand, the DNN can take advantages of interpretability offered by an FIS. The incorporation of fuzzy logic with a DNN leads to the development of deep fuzzy neural networks (DFNNs). In our research, we generalized the concept of DFNNs by presenting a Rule-Based Neural Network (RBNN). It has the following special features:

1. RBNN can be trained to solve various real world regression and classification tasks. RBNN has a similar architecture to a DNN but it has relatively few trainable parameters.

2. The input layer in RBNN has the normalization functionality. We have proposed a new type of normalization technique and

it is called the Ordered Normalization (ON). ON is specially useful when the training data has an asymmetric distribution.

3. The training of RBNN is similar to that of DNN. Stochastic gradient (SG), batch gradient descent (BGD) and Levenberg-Marquadt (LM) optimization methods are used in the parameter update of back propagation.

4. The results of RBNN are interpretable and hence it is not a black box model.

5. It is robust to (input) feature noise and compared to DNN, it produce a higher prediction accuracy even in the presence of large feature noise.

6. The performance of the RBNN on a skewed dataset is comparatively better than that of the DNN and it produces higher F1 scores for the minority (smaller) classes.

To conclude, we first developed an arithmetic-based FIS which is computationally efficient than the existing FISs. The design was extended to produce arithmetic-based type2 FIS. Then data-driven algorithms were developed to build the type1 and type2 FISs directly from the system data. This eliminated the dependency on the expert knowledge for identifying important rules for FIS. In the final step, we replaced the neuron in the neural network with RBN. The RBN is a neuron with built in arithmetic-based FIS. This led to the development of the RBNN. The RBNN was used to solve various regression and classification tasks. Compared to the DNN, the RBNN is more interpretable, robust to feature noise and produce better results (higher F1 score) on skewed datasets.

Összefoglalás

Egyes típusú aritmetikai fuzzy következtetési rendszer fejlesztésének megvalósítása új módon történt. Az eljárás az ún. paraméteres felfújó (distending) halmazhoztartozási függvény alkalmazásával került kifejlesztésre. Megmutattuk, hogy az aritmetikai alapú egyes típusú (Type-1) fuzzy következtetési rendszer (FIS) és a gradiens alapú optimalizálási módszer alkalmazásával egy kettes tipusú (Type-2) következtetési rendszert is kezelni tudunk. Az aritmetikai alapú egyes tipusú (Type-1) következtetési rendszer 20-50-szer gyorsabb, mint a hagyományos eljárások.
A szabályokat meg lehet alkotni a rendszer tanító adatainak segítségével és ezek alapján lehet az adatvezérelt fuzzy következtetési rendszert létrehozni. A klasszikus adatvezérelt technikák hátránya, hogy a fuzzy szabályok száma exponenciálisan nő az input dimenzóinak számával, így a szabályok nem interpretálhatók. Ezeket a hátrányokat a dolgozatban kiküszöböltük, megadtuk a különböző dimenziókban konstruált felfújó (distending) függvények aggregációját, aminek az eredménye egyetlen felfújó függvény. A zaj és a bizonytalanság kiküszöbölését a kettes típusú (Type-2) halmazhoztartozási függvények alkalmazásával lehet kezelni. Ennek elérése céljából a dolgozatban megalkottuk a kettes típusú (Type-2) felfújó (distending) függvényeket, amelyeket úgy generáltunk, hogy a paraméterek értékei helyett intervallumokat választottunk. Az utóbbi időben egyre fontosabbá vált a bizonytalan dinamikus rendszerek leírása. Ennek egyik eszköze a fuzzy kettes típusú (Type-2) modellek alkalmazása. Sok alkalmazás esetében a jelenleg létező technikák nem megfelelőek, ugyanis a teljes input tér szabályokkal való lefedésére volna szükség. A felmerülő nehézségek leküzdéséhez egy új megoldást határoztunk meg, ahol a

kettes típusú (Type-2) modell direkt módon származtatható az adatokból.
A mesterséges intelligencia legfontosabb kutatási területéhez tartoznak a mély neurális hálózatok. Ezeknek nagyon sok fajtája létezik, amik mind fekete doboz jellegű megoldások, ezért az eredmények nem interpretálhatók. A mesterséges intelligencia másik fontos kutatási területe a szakértői rendszerek. Itt is sok megoldás létezik, melyek interpretálhatók, de a megoldásuk nem hatékony. A két terület összekapcsolásának segítségével létrehoztuk a szabályalapú neurális hálózatokat. Az így kialakult struktúra tanítható mind regressziós, mind osztályozás alapú feladatokra.