Info 411
Machine Learning and Data Mining
Lecture 11: Type-2 Fuzzy Sets and Models Demystified

Brendon J. Woodford

24-Sep, 2019

UNIVERSITY of OTAGO
Te Whare Wānanga o Otāgo
NEW ZEALAND
SAPERE AUDE
BUSINESS SCHOOL

Department of
Information Science

Lecture Outline

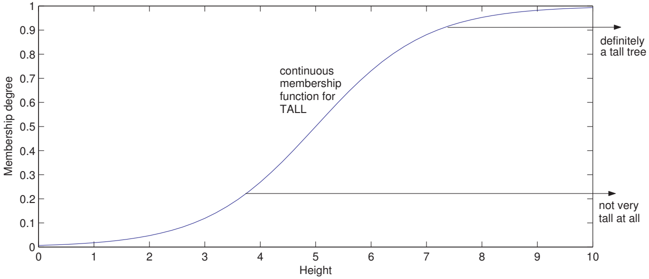
- Background to Fuzzy Sets and Models in general:
  - Crisp Sets
  - Fuzzy Sets
  - Fuzzy Logic
  - Crisp Rules
  - Fuzzy Rules
- The world of Type-2 Fuzzy Sets and Models:
  - A Type-1 Fuzzy Set
  - A Type-2 Fuzzy Set
  - A Worked Example
- Final Words?
- The Evolving Fuzzy Neural Network (EFuNN)

In a Crisp Sets World

- Everything is either true or false.
- No uncertainty is allowed.
- An item either is:
  - Entirely within a set, or
  - Entirely not in a set
- Conforms to the Law of the ***Excluded Middle***:
  - X must be either in set A or in set $not - A$.
  - No middle ground is allowed.
- Opposite sets (A and $not - A$) must between them contain everything.

The World of Fuzzy Sets: Definitions[1]

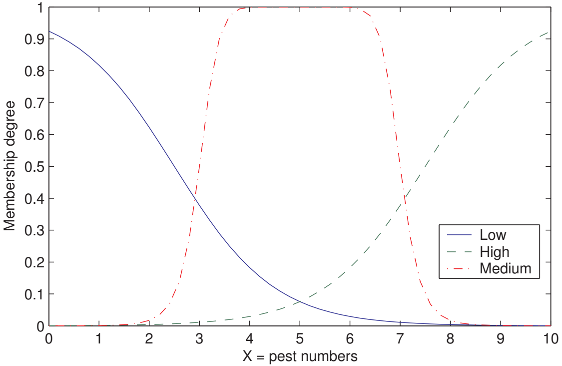
- Items can belong to a fuzzy set to different degrees:
  - Degrees of membership are expressed with membership function.
- Completely within a set is a membership degree of 1.
- Completely outside a set is a membership degree of 0.
- An item can be both A and $not - A$ to different degrees:
  - e.g. $TALL$ to a degree of 0.9, $not - TALL$ to 0.2.



[1] Adapted from [Jang;2016, pp. 2–9].

The World of Fuzzy Sets: Membership Functions[2]

- A Membership Function (MF) describes the degree of membership of a value in a fuzzy set.
- Mathematically written as $\mu_X(x)$.
- There are many different types of MF.
- Which one to use depends on the problem.



[2] Adapted from [Bezdek;1993, p. 4].

The World of Fuzzy Sets: Fuzzification

- The process of determining the degree to which a crisp value belongs in a fuzzy set.
- The value returned by a fuzzy MF.
- Most variables in a fuzzy system have multiple MFs attached to them.
- Fuzzifying that variable involves passing the crisp value through each MF attached to that value.

The World of Fuzzy Sets: Fuzzy Logic

- Same operations and functions as crisp logic.
- Must deal with degrees of truth rather than absolute truths.
- Fuzzy logic is a superset of crisp (Boolean) logic.
- Output of fuzzy logical functions are the same as crisp functions i.e.
  - Just calculated differently.
  - Handle ***degrees*** of truth, rather than ***absolute*** truths.
- The basis of fuzzy rule based systems.

The World of Fuzzy Sets: Crisp Rules

- Consists of antecedents and consequents.
- Each part of an antecedent is a logical expression.
  - e.g. $A \geq 0.5$, light is on.
- Consequent will be asserted if antecedent is true.
- Crisp rules:
  - Only one rule at a time allowed to fire.
  - A rule will fire or not fire.
  - Have problems with uncertainty.
  - Have problems with representing concepts like ***small***, ***large***, ***thin***, ***wide***.
  - The ordering of firing is also a problem.

The World of Fuzzy Sets: Fuzzy Rules

- Also have antecedents and consequents.
- Both deal with partial truths.
- Antecedents match fuzzy sets.
- Consequents assign fuzzy sets.
- Fuzzy rules can have weightings:
  - $[0, \ldots, 1]$, the importance of a rule, or commonly set to 1.
- Restaurant tipping example:
  - IF service IS ***poor*** OR food IS ***rancid*** THEN tip IS ***cheap***.
  - IF service IS ***good*** THEN tip IS ***average***.
  - IF service IS ***excellent*** OR food IS ***delicious*** THEN tip IS ***generous***.

The World of Fuzzy Sets: Fuzzy Inference

- Fuzzy inference matches fuzzy facts against fuzzy antecedents.
- Inference process assigns fuzzy output set.
- Common fuzzy inference process is class Zadeh-Mamdani inference.
- Performing fuzzy reasoning is a four step process:
  1. Fuzzify the inputs (fuzzification).
  2. Apply fuzzy logical operators across antecedents (fuzzy inference).
  3. Apply implication method.
  4. Apply defuzzification method.

The World of Fuzzy Sets: Fuzzy Inference (continued)

- Applying fuzzy logical operators across antecedents:
  - Involves calculating the degree of truth for each rules' antecedents.
  - Uses fuzzy logic functions: fuzzy AND, fuzzy OR, etc.
  - This is the degree of support for each rule.
- Each rule is allowed to 'fire' in a fuzzy system.
- Results of fuzzy inference is a fuzzy set for each output variable for each rule.
- The shape of the fuzzy set is based on the degree of support for each rule.

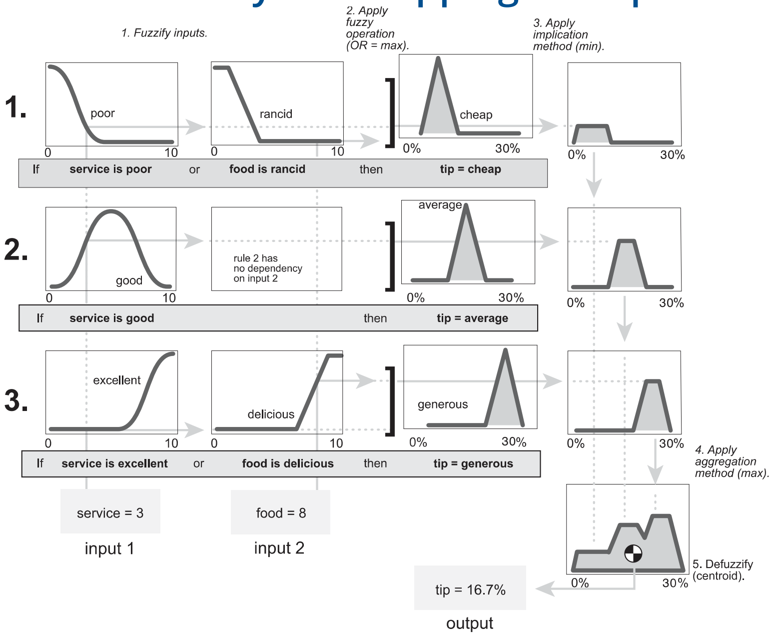
The World of Fuzzy Sets: Fuzzy Inference (continued)

- Applying implication method:
  - Shapes the output fuzzy set using the degree of support.
  - Degree to which the fuzzy output set is supported is determined by the degree of support for each rule.
  - An aggregation of the degree of support of several rules.
- An aggregated MF must be created for each output variable.
- Aggregation is performed once for each output variable.
- The order in which the rules execute doesn't matter as the aggregated MF will still be the same.

The World of Fuzzy Sets: Defuzzification

- The inferred MF can then be turned into a single, crisp, value.
- This process is called defuzzification.
- Converts an inferred, aggregated MF into a crisp number.
- Many types in existence.
- Two common ones:
  - Centre of Gravity.
  - Mean of Maxima.
- Each combines the fuzzy and corresponding crisp values in a different way to create a single number.

The World of Fuzzy Sets: Tipping Example[3]



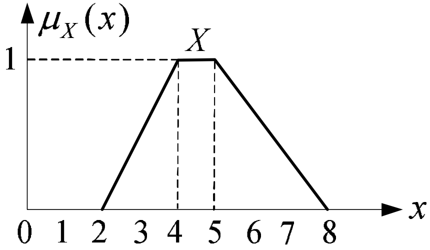
[3] Reproduced from [Jang;2016, p. 2–29].

Advantages of Fuzzy Systems[4]

- Comprehensibility.
- Parsimony.
- Modularity.
- Consistency.
- Explainability.
- Uncertainty.
- Parallelism.
- Robust.

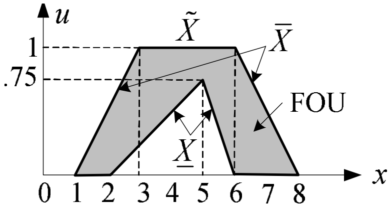
[4] Sourced and modified from E. Lughofer, Evolving Fuzzy Systems – Methodologies, Advanced Concepts and Applications (2011): Springer Publishing Company, Incorporated.

A Type-1 Fuzzy Set – Revisited



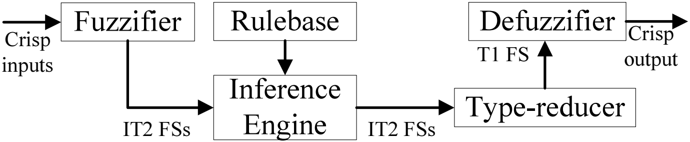
- This is a standard Type-1 Fuzzy Set (T1 FS), X.
- Integer numbers are considered in the x domain so it can be represented as $\{$`2/0,3/0.5,4/1,5/1,6/0.67,7/0.33,8/0`$\}$.
- The MF, $\mu_X(x)$ of a T1 FS can be chosen based on users' opinions or be designed using optimisation procedures.
- The limitation of T1 FS is that the membership grades are crisp values but these values have inherent uncertainty i.e. how do we model "***about 0.5***"?

A Type-2 Fuzzy Set



- This is a standard Interval Type-2 Fuzzy Set (IT2 FS), $\tilde{X}$.
- The MFs grades in an IT2 FS are ***themselves fuzzy***.
- The membership for each x becomes an interval.
- An IT2 FS is bounded from the above and below by two T1 FSs $\overline{X}$ and $\underline{X}$ which is effectively the uncertainty associated with each MF grade.
- $\overline{X}$ is the ***upper MF (UMF)*** and $\underline{X}$ is the ***lower MF (LMF)***.
- The area between $\overline{X}$ and $\underline{X}$ is the ***Footprint Of Uncertainty*** (FOU).

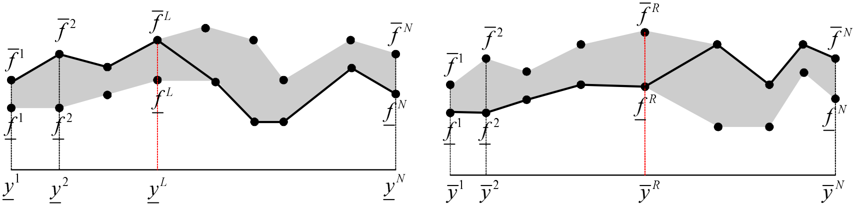
An IT2 Fuzzy Logic System (IT2 FLS)[5]



- At least one of the FS in the rule base is an IT2 FS.
  i.e. R^n: IF x_1 is $\tilde{X}_1^n$ and x_I is $\tilde{X}_I^n$ THEN y is Y^n,
  $n = 1, 2, \ldots, N$ where
  $\tilde{X}_i^n$ $(i = 1, \ldots, I)$ are IT2 FS and $Y^n = [\underline{y}^n, \overline{y}^n]$ is an interval.
- A problem is that we need to convert the IT2 FSs into a T1 FS before defuzzification can be carried out.
- For each R^n a ***type-reducer*** determines which value of its firing interval, $[\underline{f}^n, \overline{f}^n]$, is combined to compute y_l and y_r i.e. the lower and upper bounds of the output T1 FS.

[5] Sourced and reproduced from [Wu;2012, p. 2]

Which membership grade to use?[6]



(a) Computing y_l: switch from the upper firing level to the lower firing level.

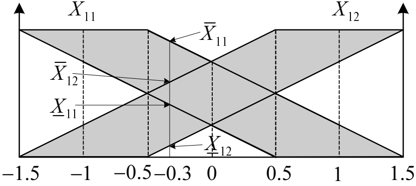
(b) Computing y_r: switch from the lower firing level to the upper firing level.

- The Karnik-Mendel (KM) algorithm computes two values: L for $\overline{f}^L$ and R for $\underline{f}^R$.
- These two values become ***switch points*** for $\underline{y}^L$ and $\overline{y}^R$ respectively and indicate which membership grade from $\underline{f}^n$ or $\overline{f}^n$, is used for defuzzification where $n = 1, 2, \dots, N$.

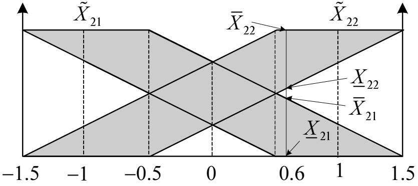
[6] Sourced and reproduced from [Wu;2012, p. 4]

Worked example: The MFs for inputs x_1 and x_2[7]

- Consider an IT2 FLS that has two inputs (x_1 and x_2) and one output y.
- Each domain consists of two IT2 FSs.



(a) Input MFs of x_1.



(b) Input MFs of x_2.

7 Sourced and reproduced from [Wu;2012, p. 5]

Worked example: The rulebase and consequents of the IT2 FLS[8]

R^1: IF x_1 is $\tilde{X}_{11}$ and x_2 is $\tilde{X}_{21}$ THEN y is Y^1
R^2: IF x_1 is $\tilde{X}_{11}$ and x_2 is $\tilde{X}_{22}$ THEN y is Y^2
R^3: IF x_1 is $\tilde{X}_{12}$ and x_2 is $\tilde{X}_{21}$ THEN y is Y^3
R^4: IF x_1 is $\tilde{X}_{12}$ and x_2 is $\tilde{X}_{22}$ THEN y is Y^4

x_1 \ x_2	$\tilde{X}_{21}$	$\tilde{X}_{22}$
$\tilde{X}_{11}$	$Y^1 = [\underline{y}^1, \overline{y}^1] = [-1, -0.9]$	$Y^2 = [\underline{y}^2, \overline{y}^2] = [-0.6, -0.4]$
$\tilde{X}_{12}$	$Y^3 = [\underline{y}^3, \overline{y}^3] = [0.4, 0.6]$	$Y^4 = [\underline{y}^4, \overline{y}^4] = [0.9, 1]$

[8] Sourced and reproduced from [Wu;2012, p. 5]

Worked example: The rulebase and consequents of the IT2 FLS[9] (continued)

- Consider an input vector $\mathbf{x}^{'} = (x_1^{'}, x_2^{'}) = (-0.3, 0.6)$
- The firing intervals of the four IT2 FSs are:

$$
\begin{aligned}
[\mu_{\underline{X}_{11}}(x_1^{'}), \mu_{\overline{X}_{11}}(x_1^{'})] &= [0.4, 0.9] \\
[\mu_{\underline{X}_{12}}(x_1^{'}), \mu_{\overline{X}_{12}}(x_1^{'})] &= [0.1, 0.6] \\
[\mu_{\underline{X}_{21}}(x_2^{'}), \mu_{\overline{X}_{21}}(x_2^{'})] &= [0, 0.45] \\
[\mu_{\underline{X}_{22}}(x_2^{'}), \mu_{\overline{X}_{22}}(x_2^{'})] &= [0.55, 1]
\end{aligned}
$$

[9] Sourced and reproduced from [Wu;2012, p. 5]

Worked example: Firing the rules[10]

The firing intervals, $[\underline{f}^N, \overline{f}^N]$, of the four rules are:

Rule No.:	Firing Interval	$\rightarrow$	Consequent
R^1:	$[\underline{f}^1, \overline{f}^1] = [\mu_{\underline{X}_{11}}(x_1') \cdot \mu_{\underline{X}_{21}}(x_2'), \mu_{\overline{X}_{11}}(x_1') \cdot \mu_{\overline{X}_{21}}(x_2')]$ $= [0.4 \times 0, 0.9 \times 0.45] = [0, 0.405]$	$\rightarrow$	$[\underline{y}^1, \overline{y}^1] = [-1, -0.9]$
R^2:	$[\underline{f}^2, \overline{f}^2] = [\mu_{\underline{X}_{11}}(x_1') \cdot \mu_{\underline{X}_{22}}(x_2'), \mu_{\overline{X}_{11}}(x_1') \cdot \mu_{\overline{X}_{22}}(x_2')]$ $= [0.4 \times 0.55, 0.9 \times 1] = [0.22, 0.9]$	$\rightarrow$	$[\underline{y}^2, \overline{y}^2] = [-0.6, -0.4]$
R^3:	$[\underline{f}^3, \overline{f}^3] = [\mu_{\underline{X}_{12}}(x_1') \cdot \mu_{\underline{X}_{21}}(x_2'), \mu_{\overline{X}_{12}}(x_1') \cdot \mu_{\overline{X}_{21}}(x_2')]$ $= [0.1 \times 0, 0.6 \times 0.45] = [0, 0.27]$	$\rightarrow$	$[\underline{y}^3, \overline{y}^3] = [-0.4, -0.6]$
R^4:	$[\underline{f}^4, \overline{f}^4] = [\mu_{\underline{X}_{12}}(x_1') \cdot \mu_{\underline{X}_{22}}(x_2'), \mu_{\overline{X}_{12}}(x_1') \cdot \mu_{\overline{X}_{22}}(x_2')]$ $= [0.1 \times 0.55, 0.6 \times 1] = [0.055, 0.6]$	$\rightarrow$	$[\underline{y}^4, \overline{y}^4] = [0.9, 1]$

[10] Sourced and reproduced from [Wu;2012, p. 6]

Worked example: Computing the final crisp result

- From the KM algorithms, we find that $L = 1$ and $R = 3$:

$$y_l = \frac{\overline{f}^1 \underline{y}^1 + \underline{f}^2 \underline{y}^2 + \underline{f}^3 \underline{y}^3 + \underline{f}^4 \underline{y}^4}{\overline{f}^1 + \underline{f}^2 + \underline{f}^3 + \underline{f}^4}$$

$$= \frac{0.405 \times (-1) + 0.22 \times (-0.6) + 0 \times 0.4 + 0.55 \times 0.9}{0.405 + 0.22 + 0 + 0.055}$$

$$= -0.7169$$

$$y_r = \frac{\underline{f}^1 \overline{y}^1 + \underline{f}^2 \overline{y}^2 + \underline{f}^3 \overline{y}^3 + \overline{f}^4 \overline{y}^4}{\underline{f}^1 + \underline{f}^2 + \underline{f}^3 + \overline{f}^4}$$

$$= \frac{0 \times (-0.9) + 0.22 \times (-0.4) + 0 \times 0.6 + 0.6 \times 1}{0 + 0.22 + 0 + 0.6}$$

$$= 0.6244$$

- Finally, the crisp output of the IT2 FLS is:

$$y = \frac{y_l + y_r}{2} = \frac{-0.7169 + 0.6244}{2} = -0.0463$$

Applications of Type-2 Fuzzy Logic

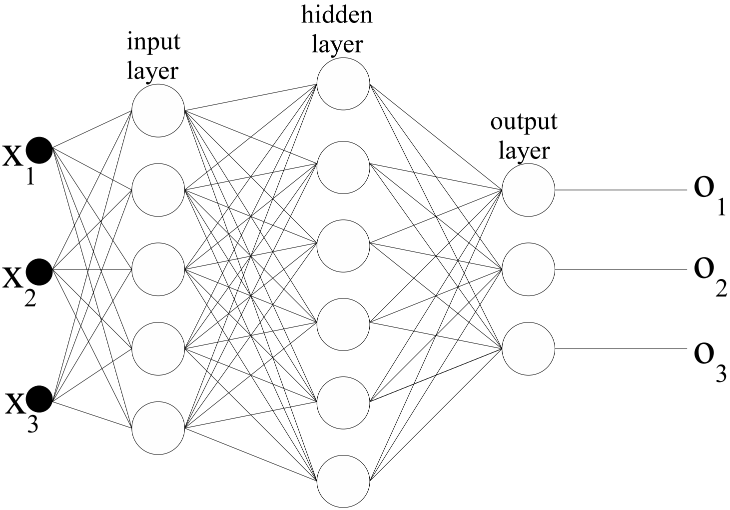
- Type-2 fuzzy logic filtering to reduce noise in colour images and contributions to perceptual computer applications i.e. Computing With Words (CWW)[11].
- A comprehensive review on IT2 FL in Intelligent Control [Castillo;2014].
- Stock market prediction using an IT2 FLS [Bernado;2012].

[11] IEEE Computational Intelligence Magazine, August **(7)**:3, 2012.

So why haven't we heard of Type-2 Fuzzy Logic beforehand?

- Learning about and applying T1 FS seems natural to do, notice their shortcomings, then address them using IT2 FS.
- Although IT2 FS were introduced by [Zadeh;1975], there were initially few people who actually published anything about them.
- As we try to more frequently model uncertain environments, IT2 FS should be able to address them so their popularity should rise.
- It is possible to extend the principle of T2 FS to Tn FS to account for situations when there is even more uncertainty to be modelled.

Neural networks (NN) in general



Neural Networks (NN) in general (continued)

- Supervised learning algorithm:
  - Finds a mapping between the input layer and output layer via the hidden layer(s).
  - Reduces the error between “state” the NN is currently in and the “state” the NN is “required” to be in.
  - Training data must be presented many times for the neural network to learn.
  - The addition of a learning rate and momentum constant reduces the problem of the NN being trapped in local minima.
- The result of learning:
  - **For prediction:** a non-linear function that maps input values to output values.
  - **For classification:** a mapping that associates input values with a particular class.
- Ultimately a NN that can generalise well to unseen or new data.

Issues with learning in a traditional NN

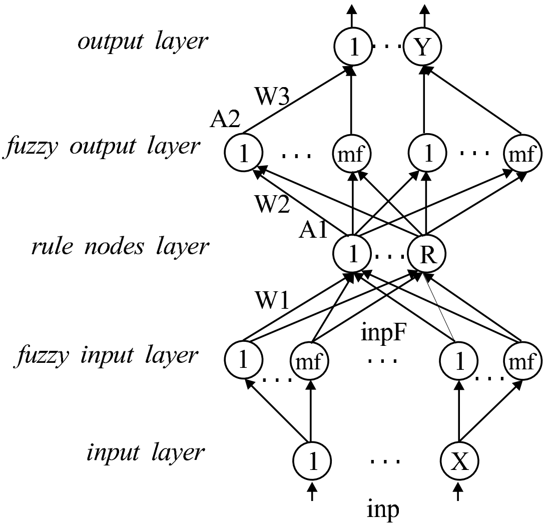
- How to select the “right” number of hidden nodes ***a priori***?
- How many times (epochs) should we train the NN for?
- What values for the learning rate and momentum constant should we use?
- What does the hidden layer and the connections within the neural network actually represent?
- Incremental learning systems have addressed some of these problems but other challenges remain.

The Evolving Fuzzy Neural Network (EFuNN)[12]

- On-line/real-time neurocomputing model.
- Incrementally learns/creates its structure from its input.
- Adapts itself to a changing environment by adding or pruning its hidden nodes (rule nodes).
- Fast learning $\rightarrow$ normally just single pass of the training data set is required.
- Can accurately learn and generalise using a small set of examples.
- Can extract rules to explain its processing.

[12] "Adaptive learning system and method", ***Patent No. 503882–Kasabov*** in [Watts;2009].

The Evolving Fuzzy Neural Network (EFuNN)[13]



[13] "Adaptive learning system and method", ***Patent No. 503882–Kasabov*** in [Watts;2009].

Learning in the EFuNN

- Starts with no rule nodes.
- Each input/output data vector (inp, out) is fuzzified using mf membership functions to form a fuzzified data vector $(inpF, outF)$.
- Normalised Euclidean distance used to measure similarity between $(inpF, outF)$ and each existing rule node, R_n.
- Parameters $sThr$ and $errThr$ are used as thresholds to determine when a new rule node should be added:
  - **If** any distance between $inpF < sThr$ or $outpF > errThr$
    then add a new rule node

    else

    existing rule node centre(s) are updated to accommodate the new data instance $\rightarrow$ (similar to on-line k-means clustering)

    end
- Parameter lr is used to control the learning rate much like traditional Neural Network (NN) backpropagation learning.

References

- D. Bernado, H. Hargras & E. Tsang
  "An Interval Type-2 Fuzzy Logic System for the Modeling and Prediction of Financial Applications"
  Autonomous and Intelligent Systems 2012, Lecture Notes in Computer Science 7326, 95–105, 2012
- J. E. Bezdek
  "Editorial: Fuzzy Models-What Are They, and Why?"
  IEEE Transactions on Fuzzy Systems. **1**(1): 1–6, 1993
- O. Castillo & P. Melin
  "A Review on Interval Type-2 Fuzzy Logic Applications in Intelligent Control"
  Information Sciences. **279**: 615–631, 2014
- R. Jang
  "Fuzzy Logic Toolbox, V2.2.24"
  The Mathworks, Inc.: Natick, MA, 2016

References (continued)

- J. M. Mendel
  "Advances in Type-2 Fuzzy Sets and Systems"
  Information Sciences, **177**: 84–110, 2007
- D. Wu
  "A Brief Tutorial on Interval Type-2 Fuzzy Sets and Systems"
  Technical Report – University of Southern California, 2012
- M. J. Watts
  "A Decade of Kasabov's Evolving Connectionist Systems: A Review"
  IEEE Transactions on Systems, Man, and Cybernetics: Part C, **39**(3):253–269, 2015
- L. A. Zadeh
  "The Concept of a Linguistic Variable and its Application to Approximate Reasoning – I"
  Information Sciences, **8**(3): 199–249, 1975