Evolving Fuzzy Systems (EFS)

Fundamentals, Reliability, Interpretability, Useability and Applications

Dr. Edwin Lughofer
Department of Knowledge-Based Mathematical Systems
Johannes Kepler University, Linz, Austria
http://www.flll.jku.at/staff/edwin
Our Department - Location in Linz AND Hagenberg (Upper Austria)

Main Campus at the University of Linz

Softwarepark Hagenberg

Fuzzy Logic Labor

Dr. Edwin Lughofer

http://www.flll.jku.at/staff/edwin
Surveys about EFS

Edwin Lughofer

Evolving Fuzzy Systems
Methodologies
Advanced Concepts
Applications
Springer, Heidelberg, 2011
(approx. 150 figures, 250 images, 460 pages)

+ book chapter (60 pages) in
Handbook on Computational Intelligence
Editor: Plamen Angelov
Title: EFS – Fundamentals, Reliability, Interpretability, Useability and Apps
Publisher: World Scientific, January 2016
Properties of Fuzzy Systems

• **Rule-Based Models** which build upon the concept of **Fuzzy Logic** (going back to Zadeh: truth values are fuzzified in [0,1] by **Fuzzy Sets**)

![Temperature Fuzzy Sets Diagram](image)

20 25 28 => to 0.2 degree “Warm” to 0.8 degree “Hot”

• **Formation of Linguistic Rules** by Conjunction (t-norm) of Fuzzy Sets over Input Vars, e.g. (example from premise price modeling):
  
  IF  Area is LOW AND Age is OLD  Then Price is LOW

• **Tradeoff** between
  
  » **Universal Approximation**: Models with arbitrary degree of non-linearity by piecewise local approx.
  
  » **Interpretability** (Readable Linguistic Terms and Rules)

• **Data-Driven Case** => **Gray-Box Models** (ranging from light to dark grey)

• Similarities Between Certain Architectures with Certain Types of **Neural Networks** (e.g. TS with Gaussian sets ⇔ RBF networks)
Research Directions in Fuzzy Sys Design

• **Old-School (70ties, 80ties):** Knowledge-Based Design (White Box)

• **New-School (90ties, 00s):** Data-Driven Learning with Machine Learning/Optimization Techniques (genfis2+3, LOLIMOT (tree-like structures), FMCLUST, ANFIS, Genetic FS …) --- batch design!

• **Emerging Topic 1 (in infants):** Hybrid Design
  » Input: Expert-based Design + Measurement Data
  » Output: Refined Fuzzy Systems meeting Interpretability Constraints
  » Variants:
    – Movement of Expert-based Partitioning
    – Optimization of Fuzzy Set Combinations (=> New Rules)
    – Model Transfer (Mamdani=>Takagi-Sugeno, consequent re-learning)

• **Emerging Topic 2: Adaptive Evolving from Streams**
The Idea of Evolving Fuzzy Systems
(emerged approx. 2004/05)

Evolving ≠ Evolutionary (=>Genetic FS)

Learning Fuzzy Systems in (Single-Pass) Incremental and Evolving Manner from Streaming (Block/Sample-wise Loaded) Data

Characterization of a Data Stream (Gama, 2010):
• The data samples or data blocks are continuously arriving on-line over time
• The data samples are arriving in a specific order, over which the system has no control.
• Data streams are usually not bounded in a size
• Once a data sample/block is processed, it is usually discarded immediately, afterwards
Concepts in Evolving Fuzzy Systems

- **Incrementality** accounts for a step-wise (sample or block) processing of data and model building, omitting time-intensive re-training (*on-line capability*)

- **Adaptivity** accounts for (*recursively*) adapting parameters with newly loaded samples

- **Evolving** means that **structural components** (rules, neurons) are *added on demand* due to new system states, operating conditions etc.

- **Single-Pass Capability** = Sample is loaded, sent into the incremental learning engine and discarded immediately, afterwards (*achieving low computation time and virtual memory demand*)
Structural Evolution vs. Parameter Adaptation

Target Y

…. Old Data Samples

Original Model (Quadr., Dashed Line)

Input Feature X
Structural Evolution vs. Parameter Adaptation

Input Feature X

Target Y

- .... Old Data Samples
- .... New Data Samples

Original Model (Quadr., Dashed Line)
Structural Evolution vs. Parameter Adaptation

Target Y

- .... Old Data Samples
- .... New Data Samples

Original Model (Quadr., Dashed Line)

Adapted Model, only parameter updates (Dotted Line)

Input Feature X
Structural Evolution vs. Parameter Adaptation

Target Y

- .... Old Data Samples
- .... New Data Samples

Adapted Model (Dotted Line)

Original Model (Quadr., Dashed Line)

Evolved Model (Cubic, Solid Line)

Input Feature X
Failure Case (Parameter Adaptation)

Significant Range Extension

Dr. Edwin Lughofer

http://www.flll.jku.at/staff/edwin
Success of Update with EFS (evolving rule base)

- Update Params only
- Update Structure + Params

Dr. Edwin Lughhofer

http://www.fll.jku.at/staff/edwin
Failure Case (Parameter Adaptation)

Bad Approx.

(a)

(b)
Success of Update with EFS (evolving rule base)

Update Params only

Update Structure + Params
Some (Industrial) Requirements

- **Fast Online Identification** of Models/Classifiers (from scratch), usage in On-line Production/Decision/Plaus.
- Updating and extending models to dynamic changes, new operating conditions, environmental influences etc. => improvement of generalization capability of models

- **Hybrid Modelling**: Refining Knowledge-based Models (Rule-Base Systems) with Data

- **Huge Data Bases [Big Data]**: data which cannot be loaded into virtual memory at once => has to be processed block-wise

- **Enhanced Human-Machine Interaction Scenarios** (Operators give feedback and provide their expertise)
Evolving Modelling Framework

Inclusion of target values from measurement/feedback; Never include estimated/predicted values from models (risk of self-error propagation)

Incremental Feedback Loop (only new data is processed)

Refine Expand Evolve Models

Pool of Evolved Models

Response (predictions, classifications, ...) of models for new data

Feedback on Quality of Response

Operator

Internal Algorithm

Usually required for classification problems => Problem: high effort/costs to give feedback => Active learning req.

Initial Models (from batch offline or former on-line training cycles)

New Data

Dr. Edwin Lughofer

http://www.flll.jku.at/staff/edwin
Definition of Fuzzy Rule Types in EFS (Regression)

Rule_i : IF \( x_1 \) IS \( \mu_{i1} \) AND...AND \( x_p \) IS \( \mu_{ip} \) THEN \( l_i \) IS \( \Phi_i \)

- **Mamdani** (Mamdani, Assilian, FSS, 1977)
  > \( \Phi_i \) : fuzzy set

- **Sugeno**
  > \( \Phi_i \) : singleton numerical (real) value

- **Takagi-Sugeno** (Takagi and Sugeno, IEEE SMC, 1985)
  > \( \Phi_i \) : linear function (hyperplane)

- **Takagi-Sugeno-Kang** (Sugeno and Kang, FSS, 1988)
  > \( \Phi_i \) : polynomial functions, Gamma, Kernels (local SVM)

- **Generalized Takagi-Sugeno – non-axis parallel contours!** New developm. in Lughofe, Cernuda et al., Evolving Sys, 2015
  > Antecedent part is a multidimensional kernel, e.g. \( \exp(-(X - C_i)^T \Sigma_i^{-1}(X - C_i)) \)
Basic Geometric Interpretation of Rule Contours

Contour of classical Rule:
IF x1 is A1 AND x2 is B1 THEN

Contour of classical Rule (t-norm):
IF x1 is A2 AND x2 is B2 Then...

Shape Depends on types of fuzzy sets

- Noisy Sample Data
- Regression Trendcurve based on 3 rules

Characteristic Spreads of Fuzzy Sets define Contour of Rules
Basic Geometric Interpretation of Rule Contours

Contour of classical Rule:
IF $x_1$ is A1 AND $x_2$ is B1 THEN ...

Contour of classical Rule (t-norm):
IF $x_1$ is A2 AND $x_2$ is B2 Then...

Contour of Generalized Rule:
IF $(x_1,x_2)$ is G1 Then...

More exact representation, less interpretability

Characteristic Spreads of Fuzzy Sets define Contour of Rules
Extended Projection Concept to Assure Interpretability (Lughofer, Cernuda, Pratama, ES, 2015)

Case: Large Span Rules

$$\sigma_i = \max_{k=1,...,u} \left( \frac{r}{\sqrt{\lambda_k}} \cos(\phi(e_k, a_i)) \right)$$

$$a_i = (0,0,...,1,...0)$$

Dr. Edwin Lughofer

http://www.fll.jku.at/staff/edwin
Evolving Fuzzy Classifiers

• **Single Model Architecture**
  (classical: Kruse, Nauck, Ishibuchi), two options
  » \( \Phi_i \): Singleton consequent labels
  » \( \Phi_i \): Confidence Vector for each Class
  (able to represent class overlaps in rules)

• **Multi-Model Architecture for Class Decomposition:**
  » **Regression-based on Indicator = one-versus-rest** (Lughofer, Angelov, Zhou, FSS, 2008) resolving masking effect of linear version
  » **All-Pairs** (Lughofer and Buchta, IEEE TFS, vol. 21 (4), 2013) - binary classifier between each class pair!
    - less complex decision boundaries,
    - new upcoming classes can be integrated quicker
      (lower class imbalance effect)
    - enhanced output interpretation based on preference relation matrix

\[
R = \begin{bmatrix}
0 & conf_{1,2} & conf_{1,3} & \ldots & conf_{1,K} \\
conf_{2,1} & 0 & conf_{2,3} & \ldots & conf_{2,K} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
conf_{K,1} & conf_{K,2} & conf_{K,3} & \ldots & 0
\end{bmatrix}
\]

Preference of Class #1 over all others
Preference degree of Class #2 over #1
Resolving Masking Effect in Indicator-Based Multi-Class Classification

Linear Regr. By Indicator matrix, 3 classes

Fuzzy regression by indicator matrix, 3 classes

from all $K$ models:

$$\text{conf} = \frac{\max_{m=1,\ldots,K} \hat{g}_m(\vec{x})}{\sum_{m=1}^{K} \hat{g}_m(\vec{x})}$$

where $\hat{g}_m(\vec{x}) = \hat{f}_m(\vec{x}) + |\min(0, \min_{m=1,\ldots,K} \hat{f}_m(\vec{x}))|$.
Geometric Interpretation of Classification in EFC

Winner-Takes-it-all

\[ L = k^* \text{ with } k^* = \arg \max_{1 \leq k \leq K} \left( \frac{\mu_1(x)h_{1,k} + \mu_2(x)h_{2,k}}{\mu_1(x) + \mu_2(x)} \right) \]

\[ \text{conf}_L = \frac{\mu_1(x)h_{1,L} + \mu_2(x)h_{2,L}}{\mu_1(x) + \mu_2(x)} \]

Membership of nearest rule

New Weighted Average Voting (including class confidences)

Lughofer, Evolving Systems, 2012
Lughofer, Weigl et al. ASOC, 2015

Movement of decision boundary closer to the uncertain rule (due to gravitation concept) => output class: rectangle

Membership Second nearest rule with different majority class label L_2 than nearest rule

http://www.fll.jku.at/staff/edwin
Behavior Decision Boundary

Example: two nearest rules having different majority classes => decision boundary in between, query point with high conflict

Example: two nearest rules having same majority class => decision boundary between these two and a third rule with different majority class => query point with low conflict

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AHLMN[119]</td>
<td>TS fuzzy</td>
<td>Incremental</td>
<td>RFWLS</td>
<td>Yes</td>
<td>Yes, cons. only</td>
<td>No</td>
<td>1-2</td>
<td></td>
</tr>
<tr>
<td>DENSFI[12]</td>
<td>neuro-fuzzy</td>
<td>Evolving</td>
<td>RFWLS with</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>1-2</td>
<td>one of the first approaches</td>
</tr>
<tr>
<td>EFC-AP[90]</td>
<td>all-pairs</td>
<td>Evolving</td>
<td>RFWLS or</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>1-2</td>
<td>Very high accuracy</td>
</tr>
<tr>
<td>EFP[116]</td>
<td>Mamdani + TS</td>
<td>Evolutionary</td>
<td>RFWLS + RLS</td>
<td>Yes</td>
<td>Yes, cons. only</td>
<td>No</td>
<td>1-2</td>
<td>first attempt of recursive learning of non-linear params</td>
</tr>
<tr>
<td>eFPT[70]</td>
<td>Tree-based</td>
<td>Dynamic</td>
<td>None</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>3</td>
<td>good interpretability, exploits architecture from[130]</td>
</tr>
<tr>
<td>eFT[71]</td>
<td>Tree-based</td>
<td>Replacement</td>
<td>RFWLS double-weight.</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>4</td>
<td>first approach of evolving fuzzy decision tree</td>
</tr>
<tr>
<td>eFuMo[115]</td>
<td>dynamic TS</td>
<td>Evolving</td>
<td>RFWLS double-weight.</td>
<td>Yes</td>
<td>Yes, cons. only</td>
<td>No</td>
<td>3-5</td>
<td>capable of splitting of rules</td>
</tr>
<tr>
<td>eMG[35]</td>
<td>generalized</td>
<td>Participatory</td>
<td>RFWLS double-weight.</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4</td>
<td>first approach using generalized rules</td>
</tr>
<tr>
<td>ENFM[117]</td>
<td>generalized</td>
<td>Recursive</td>
<td>RFWLS</td>
<td>Yes</td>
<td>Yes, cons. only</td>
<td>No</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>eClass[50]</td>
<td>single model</td>
<td>Evolving</td>
<td>RFWLS + or</td>
<td>No</td>
<td>Yes, cons. only</td>
<td>No</td>
<td>1-2</td>
<td>first one-vs-rest in EFC</td>
</tr>
<tr>
<td>ePL[24]</td>
<td>TS fuzzy</td>
<td>Participatory</td>
<td>RLS (global)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>5</td>
<td>first usage of participatory learning</td>
</tr>
<tr>
<td>eT2FIS[74]</td>
<td>Type-2 Mamdani</td>
<td>Native</td>
<td>same as antecedent learning</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>3</td>
<td>first evolving Type-2 Mamdani system</td>
</tr>
<tr>
<td>eTS(+)[114,115]</td>
<td>TS fuzzy system</td>
<td>Evolving</td>
<td>RFWLS double-weight.</td>
<td>Yes</td>
<td>Yes, later in[131]</td>
<td>Yes</td>
<td>1-2</td>
<td>one of the pioneering methods[204]</td>
</tr>
<tr>
<td>eTS-IS-SVM[55]</td>
<td>extended TS fuzzy system</td>
<td>Recursive clustering, suitability</td>
<td>Recursive Gauss-Newton</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>6 (2+4)</td>
<td>first approach for evolving extended TS fuzzy system</td>
</tr>
</tbody>
</table>

Dr. Edwin Lughoffer

Huge Diversity

http://www.fll.jku.at/staff/edwin

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FLEXFIS(4+4)</td>
<td>TS fuzzy system</td>
<td>Evolving clustering (eVQ)</td>
<td>RFWLS</td>
<td>Yes</td>
<td>Yes, conseq.</td>
<td>1-2</td>
<td>first approach for getting in ante, enhanced robustness</td>
</tr>
<tr>
<td>FLEXFIS-Class</td>
<td>single model, one vs-rest classifiers</td>
<td>Evolving clustering (eVQ)</td>
<td>RFWLS</td>
<td>No</td>
<td>Yes, conseq.</td>
<td>1-2</td>
<td>first one-vs-rest in EFC</td>
</tr>
<tr>
<td>FPA</td>
<td>single model classifiers</td>
<td>incremental constraint-based optimization</td>
<td>weight vector update</td>
<td>No</td>
<td>No</td>
<td>2</td>
<td>first joint concept on-line dim. red for gen., rules pruning</td>
</tr>
<tr>
<td>Gen-FLEXFIS</td>
<td>generalized TS fuzzy system</td>
<td>Evolving clustering (extended eVQ)</td>
<td>RFWLS</td>
<td>Yes</td>
<td>Yes, conseq.</td>
<td>1-2</td>
<td>gen. rules + on-line dim. red</td>
</tr>
<tr>
<td>GENEFIS</td>
<td>generalized TS fuzzy system</td>
<td>Gen. adaptive resonance theory + stat. influence</td>
<td>FWGRLS</td>
<td>Yes</td>
<td>No</td>
<td>3-4</td>
<td>gen. rules + on-line dim. red</td>
</tr>
<tr>
<td>HAW-NEPS</td>
<td>neuro-fuzzy system</td>
<td>inc. opt. (non-linear mod. of Kaczmarz)</td>
<td>RLS (global)</td>
<td>No</td>
<td>No</td>
<td>3</td>
<td>first wavelet-based neuro-fuzzy</td>
</tr>
<tr>
<td>LOLIMOT Inc.</td>
<td>Tree-based structure</td>
<td>Node replacement (recursive split)</td>
<td>Recursive non-linear least squares</td>
<td>No</td>
<td>No</td>
<td>1-2</td>
<td>first approach for an incremental LOLIMOT</td>
</tr>
<tr>
<td>PANFIS</td>
<td>generalized TS fuzzy system</td>
<td>Extended Self-Organizing Map, statistical influence</td>
<td>FWGRLS</td>
<td>Yes</td>
<td>No</td>
<td>2</td>
<td>first proj. concept of gen. rules, stability proofs</td>
</tr>
<tr>
<td>tGR</td>
<td>TS fuzzy system</td>
<td>Recursive clustering (GK)</td>
<td>RLS (global), RFWLS (local)</td>
<td>Yes</td>
<td>No</td>
<td>3-5</td>
<td>no rule evolution</td>
</tr>
<tr>
<td>(E)SAFIS</td>
<td>TS fuzzy system (MIMO)</td>
<td>statistical influence, distance criterion</td>
<td>RLS (global), inc. with Lyapunov</td>
<td>Yes</td>
<td>No</td>
<td>3</td>
<td>one of the pioneering methods (2008, 2006)</td>
</tr>
<tr>
<td>SAFIS-Class</td>
<td>single model TS fuzzy system</td>
<td>statistical influence, distance criterion</td>
<td>RLS (global)</td>
<td>Yes</td>
<td>No</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>SEIT2FNN</td>
<td>Type-2 TS fuzzy system</td>
<td>coverage criterion, incremental steepest descent</td>
<td>RLS (global)</td>
<td>No</td>
<td>Yes, conseq.</td>
<td>3</td>
<td>first approach for an evolving Type-2 fuzzy system</td>
</tr>
<tr>
<td>SOFMLS</td>
<td>Mamdani</td>
<td>Evolving clustering (nearest neighbor)</td>
<td>modified RLS with increased stability</td>
<td>Yes</td>
<td>No</td>
<td>4</td>
<td>first approach of an evolving Mamdani fuzzy system</td>
</tr>
<tr>
<td>SofNN (imp)</td>
<td>neuro-fuzzy system</td>
<td>coverage and system error criteria, rule enlargement</td>
<td>modified RLS (global) with weight parameter</td>
<td>Yes</td>
<td>No</td>
<td>2</td>
<td>one of the pioneering methods (2005)</td>
</tr>
</tbody>
</table>
Algorithmic Key Steps in EFS (Common Denominator)

Algorithm 1.1. Key Steps in an Evolving Fuzzy Systems Learning Engine

1. Load new data sample $\bar{x}$
2. Pre-process data sample (e.g. normalization)
3. If rule-base is empty, initialize the first rule with its center to the data sample $\bar{c} = \bar{x}$ and its spread (range of influence) to some small value; goto Step (1).
4. Else, perform the following steps (5-10):
5. Check if rule evolution criteria are fulfilled
   (a) If yes, evolve a new rule (Section 3.4) and perform body of Step (3) (without if-condition).
   (b) If no, proceed with next step.
6. Update antecedents parts of (some or all) rules (Sections 3.4 and 3.2)
7. Update consequent parameters (of some or all) rules (Sections 3.3 and 3.1)
8. Check if the rule pruning/merging criteria are fulfilled
   (a) If yes, prune or merge rules (Section 3.4); goto Step (1).
   (b) If no, proceed with next step.
10. Goto Step (1).
Algorithmic Key Steps in EFS
(Common Denominator)

Algorithm 1.1. Key Steps in an Evolving Fuzzy Systems Learning Engine

1. Load new data sample $\bar{x}$
2. Pre-process data sample (e.g., normalization)
3. If rule-base is empty, initialize the first rule with its center to the data sample $\bar{c} = \bar{x}$ and its spread (range of influence) to some small value; goto Step (1).
4. Else, perform the following steps (5-10):
5. Check if rule evolution criteria are fulfilled
   a. If yes, evolve a new rule (Section 3.4) and perform body of Step (3) (without if-condition).
   b. If no, proceed with next step.
6. Update antecedents parts of (some or all) rules (Sections 3.4 and 3.2)
7. Update consequent parameters (of some or all) rules (Sections 3.3 and 3.1)
8. Check if the rule pruning/merging criteria are fulfilled
   a. If yes, prune or merge rules (Section 3.4); goto Step (1).
   b. If no, proceed with next step.
10. Goto Step (1).

Very approach dependent

Common Baseline: RFWLS

Very approach dependent:
Generalization Concept in Lughofer et al., EUSFLAT, ES, 2011
Consequent Learning (More or less Standardized)

- **Updating (Linear) Consequent Parameters** in TS(K)-fuzzy Systems by *recursive fuzzily weighted least squares*
  
  » => induces a **local learning** effect
  » **Converges in one iteration step** (=> global optimality) due to parabola of objective function!
  » => Fast and robust convergence to hypothetical batch case

  » **Recent extensions introduce:**
  1.) weight decay regularization (GFWLS)
  2.) adaptation regulation factor (based on the current approximation error),
  3.) forgetting of older samples
Incremental Learning of Consequent Parameters (in TS fuzzy systems)

Local learning by a recursive fuzzily weighted least squares (consequ. func. for each rule separately)

Objective Function

\[ J_i = \sum_{k=1}^{N} \Psi_i(\vec{x}(k)) \epsilon_i^2(k) \rightarrow \min_{w_i} \]

\[ e_i(k) = y(k) - \hat{y}_i(k) \]

\[ \hat{w}_i(k+1) = \hat{w}_i(k) + \gamma(k)(y(k+1) - r^T(k+1)\hat{w}_i(k)) \]

\[ \gamma(k) = \frac{P_i(k+1)r(k+1)}{1 + r^T(k+1)P_i(k)r(k+1)} \]

\[ P_i(k+1) = (I - \gamma(k)r^T(k+1))P_i(k) \]

Membership degree to Rule \#i

Weighted Inverse Hessian Matrix:

\[ X^T Q_i X \text{ with } Q_i \text{ weighting matrix} \]
Local Verus Global Learning
(All Params in All Rules at Once)

- **More Flexibility during learning**: Rules can be added, pruned or merged by not 'disturbing' the recursive parameter update of the others (Global learning requires a spec. Handling of the inverse Hessian in such cases as size changes)

- **Comp. Complexity**: $O(C(p+1)^2)$ versus $O((C(p+1))^2)$
- **More stability** as dealing with smaller Hessian matrices
- **Local feature weighting capability** (per local region)
- **Better Interpretability** (piecewise linear approx. along real trend of curve, Yen, Wang and Gillespie, IEEE TFS, 1998)
Antecedent Learning - Principles

• **Incremental Partitioning of the Feature Space:**
  » Concept 1: **Rule Evolution** (expansion of the knowledge base to explore new regions) – *distance and density-based criteria*
  » Concept 2: **Rule Pruning** (contraction of the knowledge base to remove redundancies, obsolete rules etc.) => more compact, less time-intensive for on-line – *distance and geometric criteria*

• **Incremental Adaptation of Non-Linear Parameters:**
  1.) **analytical** (incremental optimization, *e.g.* RLM), - concept by Wang, Vrbanek, IEEE TFS, 2008!
  2.) **heuristics-based** (movement of rules, fuzzy sets), often done with inc. Clustering (many possibilities: eVQ, incremental fuzzy c-means, recursive Gustafsson.Kessel, recursive subtractive clustering, potentials as density estimators, evolving participatory learning, …)
Incremental Learning of Non-Linear Antecedent Parameters (for LS Problem)

Recursive LM (used in EFP by Wang and Vrbanek, 2008)

\[ \Phi_{\text{nonlin}}(k + 1) = \Phi_{\text{nonlin}}(k) + P(k)Jac^T(k)e(k) \]

(2.87)

with

\[ P(k) = \frac{1}{\lambda_k} (P(k - 1) - P(k - 1)US^{-1}U^T P(k - 1)) \]

(2.88)

\( \lambda_m \) a forgetting factor (set to 1 in case of no forgetting) and

\[ S = \lambda_k V + U^T P(k - 1)U \]

(2.89)

The matrix \( S \) is a \( 2 \times 2 \) matrix, therefore easily and quickly calculated in each update step; this is because

\[ U^T = \begin{bmatrix} \text{Jac}^T(k) \\ 0 \ldots 010 \ldots 0 \end{bmatrix} \]

and

\[ V^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & D\alpha_k \end{bmatrix} \]
Principal Cases of Data Stream Influence on Rules (over time)

Case 1: rule expansion (parameter updates)
- Old rule contour
- Updated rule contour

Case 2: knowledge expansion (new rule evolved)
- Old rule contour
- New single sample may induce new rule
- The rule latter becomes overlapping
- Case 3: too early knowledge expansion => back-merge desired
- + untypical rule growing => split desired: in infants

... Old Samples
... New Sample(s)

... New Samples next steps

Contours of original Rules
Contours of updated/new Rules

http://www.flll.jku.at/staff/edwin
Concrete Example: Antecedent Learning in FLEXFIS (Lughofer, IEEE TFS, 2008)

- is performed in the Clustering Space => each cluster is associated with one rule (local region)
- Based on evolving version of Vector Quantization (eVQ, Lughofer, PR, 2008)

- **Three central steps** (Start: first center = first sample):
  - Checking whether a data sample x fits into cluster partition
    - Main Pos (Euclid): distance to nearest cluster higher than vigilance
    - Arb. Pos (Mahal): Statistical Tolerance Region
  - If yes, Movement of winning = nearest center towards sample
    \[ c_{win}^{(new)} = c_{win}^{(old)} + \eta (x - c_{win}^{(old)}) \]
    Update ranges of influence of win. Cluster

    Ellipsoids in Main Position (Euclid):
    \[ (k_i + 1) \sigma_{ij}^2 = k_i \sigma_{ij}^2 + (k_i + 1) \Delta c_{ij}^2 + (c_{ij} - x_{kj})^2 \quad \forall j = 1, ..., p + 1 \]

    Ellipsoids in Arbitrary Position (Mahal):
    \[ \Sigma^{-1}(k + 1) = \frac{\Sigma^{-1}(k)}{1 - \alpha} - \frac{\alpha}{1 - \alpha} \left( \frac{\Sigma^{-1}(k)(x - c)}{\Sigma^{-1}(k)(x - c)^T} \right) \Sigma^{-1}(k)(x - c) \]

- If no, evolve new Rule
  \[ c_{C+1} = x \quad \Sigma_{C+1}^{-1} = \frac{\sum_{i=1}^{C} \Sigma_i^{-1} C}{\sum_{i=1}^{C} \Sigma_i^{-1}} \quad \Sigma_1^{-1} = \text{diag}(\frac{\text{frac range}}{\text{range}^2}) \]
Concepts for Increased Stability

- Automatic Drift handling (detection + reaction)
  » Older learned relations become obsolete

  => More flexibility in model updates required => Forgetting

  » Global and Local Drift Handling (different intensity! per Rule):
    Shaker, Lughofer, Evolving Systems, 2014
Drift Reaction - Example

Conventional Update => cluster joins both distribution

Update with forgetting => Cluster covers new distribution (old samples completely forgotten)
Concepts for Increased Stability

• Dynamic Curse of Dimensionality Reduction
  » Next slide

• Incremental Smoothing (Regularization)
  » Rosemann, Brockmann, INS, 2012

• Convergence Analysis / Ensurance (FLEXFIS, PANFIS)
  » Bounded Sub-optimality in the least squares sense (correction terms integration), Proofed bounds on the system error
Idea of Dynamic Smooth Dim. Reduction

- **Introduction of the Concept of Feature Weighting**
  (EFS Approach Independent, Lughofer, Pratama, ES, 2015)
  » Criterion: expected statistical contribution
  » Important Features => High Weights (in \([0,1]\))
  » Update Feature Weights over Time => smooth change in the importance of features – No crisp selection/deletion (as features may become important later again!)

- **Integration of Feature Weights in Rule Evolution Criterion**
  (Approach Dependent)
  » In Gen-Smart-EFS: weights into Mahalnobis distance
  » => re-scaled version of inverse covariance matrix:

\[
\Sigma_i^{-1}(\text{rescaled}) = V \ast \text{diag}(\lambda \ast) \ast D \ast \text{diag}(\lambda \ast) \ast V^T \\
\text{diag}(\lambda \ast) = \text{diag}(A \ast) = \text{diag}(V^T \ast \text{diag}(\lambda \ast) \ast V)
\]
Concepts for Increased Useability

- **Interpretability**: understanding the model components and behavior
  - **Criteria examined**: Distinguishability, Simplicity, Consistency, Coverage, Input/Output Behavior, Feature Weights, Rule Lengths, Rule Weights, Local Property, Interpretation of Consequents (red = essential for transparency)

- **Reliability**: interpretation of model predictions (useable for AL)

---

Dr. Edwin Lughofer

http://www.fll.jku.at/staff/edwin
Merge Criteria for Assuring Distinguishability and Compactness

Final Criterion: AND Connection of 1, 2 and 3:

1. Euclidean distance between centers
   \[ d(c_A, c_B) \leq \frac{\sum_{j=1}^{p} |c_{A;j} - c_{B;j}|(f_{ac} \ast \sigma_{A;j} + f_{ac} \ast \sigma_{B;j})}{\sum_{j=1}^{p} |c_{A;j} - c_{B;j}|} + \epsilon \]
   Touching (\( \epsilon = 0 \))
   Slightly Overlapping (\( \epsilon < 0 \) big)
   Close (\( \epsilon > 0 \) small)

2. Homogeneity Criterion (restricting the volume of the merged ellipsoid)
   \[ V_{merged} \leq \rho(V_A + V_B) \]
   Dimensionality

3. For Regression: Similarity of Consequents high (dihedral angle)
   \[ S_{cons}(\vec{w}_A, \vec{w}_B) \geq \text{thresh} \]

3. For Classification: majority class in both rules are the same
Homogeneity Criterion: Impact

» Basic Idea: **Merging of rules which are touching**, slightly overlapping, very close to each other and **fulfilling homogeneity criterion**

Candidate for Merging

Not a Candidate for Merging
Concepts for Increased Useability

• **Single-Pass Active Learning**
  (Lughofer, PR / ES, 2012 for classification; Cernuda, Lughofer et al. 2014 for regression)
  » **Reducing Annotation / Measurement Effort**
  » Based on Reliability Concepts and Beyond (certainty-based sampling)

• **(Towards) Plug-and-Play Functionality:** Compensating „unlucky“ initial settings of learning parameter(s)
  » **Possibility 1:** adaptive learning parameters (e.g. steering rule evolution vs. Rule update) according to the current data stream characteristics
  » **Possibility 2:** dynamic split-and-merge of clusters (rules), resolving intra-heterogeneity (=> split) and inter-homogeneity (=> merge)

=> Histogram along Pri-Comp 1 of Cluster 4

Clustering and feature analysis diagrams illustrating the concepts and methods.
Applications of EFS and Some Results (from own Projects)
## Applications of EFS - Overview

<table>
<thead>
<tr>
<th>Application Type/Class (alphabetically)</th>
<th>EFS approaches (+ refs)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active learning / human-machine interaction</td>
<td>FLEXFIS-Class, FLEXFIS-AP, FLEXFIS-PLS</td>
<td>reducing the annotation effort and measurement costs in industrial processes</td>
</tr>
<tr>
<td>Adaptive on-line control</td>
<td>evolving PID and MRC controllers in, eFuMo, rGK, self-evolving NFC, adaptive controller in</td>
<td>design of fuzzy controllers which can be updated and evolved on-the-fly</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>EFuNN</td>
<td>specific applications such as ribosome binding site (RBS) identification, gene profiling</td>
</tr>
<tr>
<td>Chemometric Modeling and Process Control</td>
<td>FLEXFIS++, the approach in</td>
<td>The application of EFS onto processes in chemical industry (high-dim. NIR spectra)</td>
</tr>
<tr>
<td>EEG signals classification and processing</td>
<td>eTS, eSNNr</td>
<td>time-series modeling with the inclusion of time delays</td>
</tr>
<tr>
<td>Evolving Smart Sensors (e-Sensors)</td>
<td>eTS+, (gas industry), (chemical process industry), FLEXFIS and PANFIS (NOx emissions)</td>
<td>evolving predictive and forecasting models in order to substitute cost-intensive hardware sensors</td>
</tr>
<tr>
<td>Forecasting and prediction (general)</td>
<td>AHLTNM (daily temp.), eT2FIS (traffic flow), eFPT (Statlog from UCI), eFT and eMG (short-term electricity load), FLEXFIS and GENEFIS (house prices), LOLIMOT inc. (maximum cylinder pressure), rGK (sales prediction) and others</td>
<td>various successful implementations of EFS</td>
</tr>
</tbody>
</table>
## Applications of EFS - Overview

<table>
<thead>
<tr>
<th>Financial domains</th>
<th>eT2FIS, eT2FIS,07 evolving granular systems, ePL, PAN-FIS, SOFNN</th>
<th>time-series modeling with the inclusion of time delays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of dynamic benchmark problems</td>
<td>DENFIS, eT2FIS, eTS+, FLEXFIS, SAFIS, SEIT2ENN, SOFNN</td>
<td>Mackey-Glass, Box-Jenkins, etc.</td>
</tr>
<tr>
<td>On-line fault detection and condition monitoring</td>
<td>eMG for classification, FLEXFIS++</td>
<td>EFS applied as SysID models for extracting residuals</td>
</tr>
<tr>
<td>On-line monitoring</td>
<td>eTS+ (gas industry)</td>
<td>supervision of system behaviors</td>
</tr>
<tr>
<td>Robotics</td>
<td>eTS+</td>
<td>in the area of self-localization</td>
</tr>
<tr>
<td>Time-series modeling</td>
<td>DENFIS, ENFM, and eTS-LS-SVM (sun spot)</td>
<td>local modeling of multiple time-series versus instance-based learning</td>
</tr>
<tr>
<td>User behavior identification</td>
<td>eClass and eTS, eTS+, FPA</td>
<td>analysis of the user’s behaviors in multi-agent systems, on computers, indoor environments etc.</td>
</tr>
<tr>
<td>Video processing</td>
<td>eTS, eTS+</td>
<td>including real-time object id., obstacles tracking and novelty detection</td>
</tr>
<tr>
<td>Visual quality control</td>
<td>EFC-AP, FLEXFIS-Class, pClass</td>
<td>image classification tasks based on feature vectors</td>
</tr>
</tbody>
</table>

Dr. Edwin Lughofer

http://www.fll.jku.at/staff/edwin
Evolvable Chemometric Models in Chemical Production Systems

Feedback Loop

FTNIR-Spectrometer

PLC (ARM-processor) + Chemometrics-Unit

Evolving Chemometric Models

Immersion probe

eChemo paradigm (recent development 2012-2015)

Monitoring

Trigger analytical reference value

process control system

CANopen

Remote access (WEB)

Storage

Dr. Edwin Lughofe
Some Results (eChemo paradigm)
(Cernuda, Lughofer et al, ACA 2012 - EFS + PLS for Chemo models)

• Dynamic Adaptation of Chemometric Models in Viscose Production

Prediction On-line Data with Static PLS Models (State of the Art)

Prediction On-line Data with Dynamic Evolving Fuzzy Models (eChemo paradigm)

• Coupling with Active Learning => Reduction of Measurements + Model Updates by approx. 98% => Significant Cost Reduction (Cernuda, Lughoffer et. al, CILS, 2014)
Active Learning Results in On-line Melamin Resine Production (Sliding Window Based Approach)

No Active Sample Selection, blind equidistant adaptation

Active Sample Selection

Savings: ~13% (~87% Samples used for Selection)
**Automatic On-line Surface Inspection with Adaptivity of Image Classifiers**

(binary: good/ bad reward)

Eitzinger, Lughofer et. al, MVA, 2010
Lughofer, Smith et al., IEEE SMC-A, 2009

Getting Rid of Manual Checks!

Feedback Loop for Classifier Adaptation
(based on native good/bad reward)

• **Achievements:**
  » Common Interfaces between Images and Classifiers
  » **Extracting Arbitrary Regions of Interests with Clustering Methods**
  » **Machine-learning Classifiers with Accuracy > 99%** (on three real-world surface inspection scenarios)
  » **Resolving Contradicting Input** from Several Operators (Ensembles)
  » **On-Line Adaptation and Evolution** of Image Classifiers (with EFC)

---

Dr. Edwin Lughofer

http://www.fll.jku.at/staff/edwin
# On-line Surface Inspection - Results


<table>
<thead>
<tr>
<th>Static Image Classifiers</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(trained in off-line mode)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eVQ-Class variant A</td>
<td>75.69</td>
<td>91.55</td>
<td>66.67</td>
<td>63.75</td>
</tr>
<tr>
<td>eVQ-Class variant B</td>
<td>88.82</td>
<td>90.11</td>
<td>66.67</td>
<td>64.65</td>
</tr>
<tr>
<td>EFC SM</td>
<td>78.82</td>
<td>95.20</td>
<td>66.67</td>
<td>60.73</td>
</tr>
<tr>
<td>EFC MM</td>
<td>73.53</td>
<td>95.89</td>
<td>54.67</td>
<td>55.59</td>
</tr>
<tr>
<td>k-NN</td>
<td>79.61</td>
<td>91.51</td>
<td>53.33</td>
<td>58.30</td>
</tr>
<tr>
<td>CART</td>
<td>78.82</td>
<td>91.78</td>
<td>52.00</td>
<td>65.26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evolved Image Classifiers</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(updated in on-line mode)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eVQ-Class variant A</td>
<td>89.22 (+13.5)</td>
<td>91.12 (+0.4)</td>
<td>86.67 (+20)</td>
<td>67.67 (+3.9)</td>
</tr>
<tr>
<td>eVQ-Class variant B</td>
<td>90.39 (+1.6)</td>
<td>93.33 (+3.2)</td>
<td>86.67 (+20)</td>
<td>67.98 (+3.3)</td>
</tr>
<tr>
<td>EFC SM</td>
<td>78.82 (+0.0)</td>
<td>96.21 (+1.0)</td>
<td>64.00 (-2.6)</td>
<td>63.14 (+2.4)</td>
</tr>
<tr>
<td>EFC MM</td>
<td>87.65 (+14.1)</td>
<td>97.19 (+1.3)</td>
<td>78.67 (+24)</td>
<td>65.56 (+10.0)</td>
</tr>
<tr>
<td>k-NN (re-trained)</td>
<td>90.98 (+11.4)</td>
<td>96.06 (+4.6)</td>
<td>74.67 (+21.3)</td>
<td>59.52 (+1.2)</td>
</tr>
<tr>
<td>CART (re-trained)</td>
<td>90.59 (+11.8)</td>
<td>97.02 (+5.2)</td>
<td>52.00 (+0.0)</td>
<td>69.18 (+3.9)</td>
</tr>
</tbody>
</table>

Very slow – retraining on each sample, just for benchmark purposes

## Multi-class problem

<table>
<thead>
<tr>
<th>Data Set</th>
<th>EFC SM</th>
<th>EFC MM</th>
<th>EFC AF</th>
<th>eVQ-Class A</th>
<th>eVQ-Class B</th>
<th>CART</th>
<th>k-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD imprint</td>
<td>62.0 ± 2.5</td>
<td>73.1 ± 1.1</td>
<td><strong>82.6 ± 1.5</strong></td>
<td>64.1 ± 2.4</td>
<td>74.9 ± 1.6</td>
<td>75.81 ± 2.1</td>
<td>73.85 ± 2.9</td>
</tr>
</tbody>
</table>

Dr. Edwin Lughofer

http://www.fll.jku.at/staff/edwin
Accumulated Accuracies

(a) CD-Imprint with merging and feat weights

(b) Eggs with merging and feat weights

(c) Micro Fluid Chips with Active Learning

(d) Rotor with merging and feat weights

http://www.fll.jku.at/staff/edwin

Dr. Edwin Lughofer
On-line Human-Machine Interaction @ VI Systems with Evolving Fuzzy Classifiers, AL and Drift Detection embedded (Weigl, Lughofer et al., MVA, 2015)
Operator labels from time to time new samples within active learning cycles => new classes may be defined

=> Extension in EFCs, class decomposition in EFC-AP favorable (K new binary classifiers introduced)
Explaining Decisions Linguistically (Enhanced Model Output Interpretation)
Lughofer, Richter et al., ESWA, submitted, 2016

Labeling by Dropping (still Native Feedback)

Classifier Output

Linguistic Explanation for Reason

Das Beispiel ist mit sehr hoher Sicherheit (90 %) Dirty.
Das Beispiel kann mit sehr niedriger Sicherheit (10 %) auch Inclusion sein.
Der Bereich ist durch das Training nicht (23 %) abgedeckt.
Grund: Aspect Ratio ist klein (208.7 +/- 74.6), Min Max Len ist klein (8.3 +/- 4.4), Dipol angle ist klein (0.0 +/- 0.0) und Sqrt Area ist klein (5.2 +/- 3.4)
Condition Monitoring in Rolling Mills
Serdio, Lughofer et al, INS, 2014/15 (FD + FI)
Serdio, Lughofer et al, INF, 2014 (dynamic models with lags)
Condition Monitoring in Rolling Mills

Update, Evolve Model

Sparse Fuzzy Sys Training with GA Tuning
Results from Condition Monitoring

Effect of Adaptive Filters

Absolute residuals of a model with R2: 0.9928

Mu + Sigma band around the absolute residuals

Mu + Sigma band around the absolute residuals, after an average filter (n=10)

Mu + Sigma band around the absolute residuals, after a modified average filter (n=10)

Mu + Sigma band around the absolute residuals, after a median filter (n=10)

Mu + Sigma band around the absolute residuals, after a gaussian filter (2 bells)

Effect of Genetic Tuning

5% Fault Intensity

10% Fault Intensity

20% Fault Intensity

Dr. Edwin Lugofer

http://www.fll.jku.at/staff/edwin
Results from Condition Monitoring

Effect of Lags

Fault intensity 5 %  Fault intensity 10 %  Fault intensity 20 %  Fault intensity 50 %  Fault intensity 100 %

% detection

% overdetection

TS fuzzy Systems

Dr. Edwin Lughofer

http://www.fll.jku.at/staff/edwin
NOx Emission Modeling (eSensor)
Lughofer, Guardiola et al., ASOC, 2011

Evolving Fuzzy Systems Modeling with Compact Rulebases

Analytical Models requiring long development and setup phases for new types of engines!

Static Points of Engine Map (off-line test case)
Dynamic Driving Trajectory (real on-line case)
Results eSensor for NOx
(separate test set)

Error Analysis
Evolving Fuzzy Systems

Error Analysis
Static Fuzzy Systems

Outperformance of analytical model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy</td>
<td>1.32</td>
<td>2.04</td>
<td>1.61</td>
<td>2.46</td>
</tr>
<tr>
<td>Physical</td>
<td>1.57</td>
<td>2.23</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Ridge Repr.</td>
<td>2.91</td>
<td>3.04</td>
<td>5.76</td>
<td>2.76</td>
</tr>
<tr>
<td>SVR</td>
<td>3.61</td>
<td>3.44</td>
<td>4.94</td>
<td>4.61</td>
</tr>
<tr>
<td>ANFIS</td>
<td>2.37</td>
<td>3.26</td>
<td>4.04</td>
<td>4.74</td>
</tr>
<tr>
<td>NN</td>
<td>1.49</td>
<td>2.65</td>
<td>7.06</td>
<td>3.49</td>
</tr>
</tbody>
</table>

http://www.fll.jku.at/staff/edwin

Dr. Edwin Lugohefer
Adaptive Dynamic House Pricing
(Lughofer, Trawinski, Trawinski, Information Sciences, 2011)

Adaptivity for new Recordings

Quality Zones

State-of-the-Art before:
hard coded Expert System,
No (flexible) data-driven model at all
Results from Prediction of House Prices

<table>
<thead>
<tr>
<th>Tr. / Test</th>
<th>FLEXFIS</th>
<th>FLEXFIS+pr.</th>
<th>Expert</th>
<th># Rules</th>
<th># R. pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-2002 / 2003</td>
<td>0.049</td>
<td>0.049</td>
<td>0.078</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>1999-2003 / 2004</td>
<td>0.072</td>
<td>0.074</td>
<td>0.106</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>2000-2004 / 2005</td>
<td>0.072</td>
<td>0.068</td>
<td>0.121</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>2001-2005 / 2006</td>
<td>0.130</td>
<td>0.125</td>
<td>0.134</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>2002-2006 / 2007</td>
<td>0.089</td>
<td>0.089</td>
<td>0.138</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>2003-2007 / 2008</td>
<td>0.110</td>
<td>0.115</td>
<td>0.145</td>
<td>13</td>
<td>7</td>
</tr>
</tbody>
</table>

Default vigilance of 0.3 optimal in 3 out of 4 cases!

<table>
<thead>
<tr>
<th>Tr. / Test</th>
<th>evolved TS</th>
<th>equiv. Mamdani</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-2002 / 2003</td>
<td>0.049</td>
<td>0.059</td>
</tr>
<tr>
<td>1999-2003 / 2004</td>
<td>0.072</td>
<td>0.113</td>
</tr>
<tr>
<td>2000-2004 / 2005</td>
<td>0.072</td>
<td>0.093</td>
</tr>
<tr>
<td>2001-2005 / 2006</td>
<td>0.130</td>
<td>0.158</td>
</tr>
<tr>
<td>2002-2006 / 2007</td>
<td>0.089</td>
<td>0.126</td>
</tr>
<tr>
<td>2003-2007 / 2008</td>
<td>0.110</td>
<td>0.118</td>
</tr>
</tbody>
</table>
Model Insights House Price Prediction

Figure 1. Fuzzy partitions for the five input variables (a) to (c): Area, Age, Storey, Rooms, Centre and the output variable Price (d)

Figure 2. Fuzzy partitions for the two input variables Storeys and Rooms when not using any merging/pruning option in FLEXFIS – compare with those in Figure 1.

Dr. Edwin Lughofer

http://www.flll.jku.at/staff/edwin
Rules House Price Prediction

Rule 1: If Area is LOW and Age is NEW and Storeys is AV and Rooms is FEW and Centre is CLOSE
    Then Price is LOW
Rule 2: If Area is LOW and Age is VeryNEW and Storeys is MANY and Rooms is FEW and Centre is CLOSE
    Then Price is LOW
Rule 3: If Area is LOW and Age is NEW and Storeys is AV and Rooms is USUAL and Centre is CLOSE
    Then Price is LOW
Rule 4: If Area is LOW and Age is OLD and Storeys is AV and Rooms is FEW and Centre is CLOSE
    Then Price is LOW
Rule 5: If Area is LOW and Age is NEW and Storeys is FEW and Rooms is FEW and Centre is FAR
    Then Price is LOW
Rule 6: If Area is LOW and Age is MEDIUM and Storeys is MANY and Rooms is FEW and Centre is VeryCLOSE
    Then Price is MEDIUM
Rule 7: If Area is MEDIUM and Age is NEW and Storeys is FEW and Rooms is USUAL and Centre is MEDIUMDist
    Then Price is MEDIUM
Rule 8: If Area is HIGH and Age is MEDIUM and Storeys is AV and Rooms is MANY and Centre is CLOSE
    Then Price is HIGH
Rule 9: If Area is MEDIUM and Age is MEDIUM and Storeys is AV and Rooms is USUAL and Centre is CLOSE
    Then Price is HIGH
Thank you for your attention!