FUZZY RULE-BASED REGRESSION FOR EXPLAINABLE AI

Martin Dallinger
K12205889
martin.dallinger@outlook.com

15.05.2024

JOHANNES KEPLER
UNIVERSITÄT LINZ

TABLE OF CONTENTS

EXPLAINABLE AI - MOTIVATION

- **Recent successes in AI [1]**
  - Primarily in Machine Learning (ML) and Deep Learning (DL)

[1] W. Ding et al., Explainability of Articial Intelligence Methods, Applications and Challenges: A Comprehensive Survey
[2] A. Adadi and M. Berrada, Peeking Inside the Black-Box: A Survey on Explainable Articial Intelligence (XAI)

EXPLAINABLE AI - MOTIVATION

- **Recent successes in AI [1]**
  - Primarily in Machine Learning (ML) and Deep Learning (DL)
- **But: Complicated “Black Boxes“ [2]**
  - → Lack of trust
  - Critical domains: Healthcare, finance, cybersecurity, …

[1] W. Ding et al., Explainability of Articial Intelligence Methods, Applications and Challenges: A Comprehensive Survey
[2] A. Adadi and M. Berrada, Peeking Inside the Black-Box: A Survey on Explainable Articial Intelligence (XAI)

EXPLAINABLE AI - MOTIVATION

- **Recent successes in AI [1]**
  - Primarily in Machine Learning (ML) and Deep Learning (DL)
- **But: Complicated “Black Boxes“ [2]**
  - → Lack of trust
  - Critical domains: Healthcare, finance, cybersecurity, …

Primary motives [2]

Explain to **justify**	Explain to **control**
Explain to **improve**	Explain to **discover**

[1] W. Ding et al., *Explainability of Articicial Intelligence Methods, Applications and Challenges: A Comprehensive Survey*
[2] A. Adadi and M. Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Articial Intelligence (XAI)*

EXPLAINABLE AI - MOTIVATION

- **Recent successes in AI [1]**
  - Primarily in Machine Learning (ML) and Deep Learning (DL)
- **But: Complicated “Black Boxes“ [2]**
  - → Lack of trust
  - Critical domains: Healthcare, finance, cybersecurity, …
- Not only models explaining themselves but explain existing data

Primary motives [2]

Explain to **justify**	Explain to **control**
Explain to **improve**	Explain to **discover**

[1] W. Ding et al., Explainability of Articial Intelligence Methods, Applications and Challenges: A Comprehensive Survey
[2] A. Adadi and M. Berrada, Peeking Inside the Black-Box: A Survey on Explainable Articial Intelligence (XAI)

EXPLAINABLE AI - MOTIVATION

- **Recent successes in AI [1]**
  - Primarily in Machine Learning (ML) and Deep Learning (DL)
- **But: Complicated “Black Boxes“ [2]**
  - → Lack of trust
  - Critical domains: Healthcare, finance, cybersecurity, …
- Not only models explaining themselves but explain existing data
- Fuzzy rule-based regression is inherently explainable
  - Usable in many areas of XAI (see appendix for a simple taxonomy)

Primary motives [2]

Explain to **justify**	Explain to **control**
Explain to **improve**	Explain to **discover**

[1] W. Ding et al., Explainability of Articial Intelligence Methods, Applications and Challenges: A Comprehensive Survey
[2] A. Adadi and M. Berrada, Peeking Inside the Black-Box: A Survey on Explainable Articial Intelligence (XAI)

SECTION 2

Introduction to Fuzzy Logic

FUNDAMENTALS OF FUZZY SETS

- **Published by Lotfi Zaddeh in 1965 [4]**
  - Traditionall set membership: $\mu_{trad}: \mathbb{R} \rightarrow \{0,1\}$

[4] L. Zadeh, Fuzzy Sets, Information and Control, vol. 8, no. 3, pp. 338-353
[5] E. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis
[6] Q. Liang and J. Mendel, An Introduction to Type-2 TSK Fuzzy Logic Systems

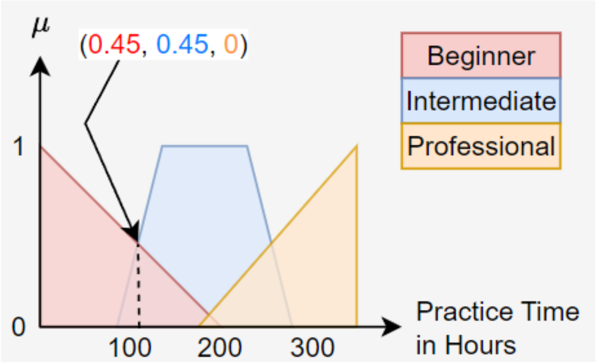
FUNDAMENTALS OF FUZZY SETS

- **Published by Lotfi Zaddeh in 1965 [4]**
  - Traditionall set membership: $\mu_{trad}\colon \mathbb{R} \to \{0,1\}$
  - Fuzzy sets: Use Membership Functions (MFs)
    - $\mu_X\colon \mathbb{R} \to [0,1]$
    - $a \in \mathbb{R}$, fuzzy set $X, \mu_X(a) \in [0,1]$

[4] L. Zadeh, Fuzzy Sets, Information and Control, vol. 8, no. 3, pp. 338-353
[5] E. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis
[6] Q. Liang and J. Mendel, An Introduction to Type-2 TSK Fuzzy Logic Systems

FUNDAMENTALS OF FUZZY SETS

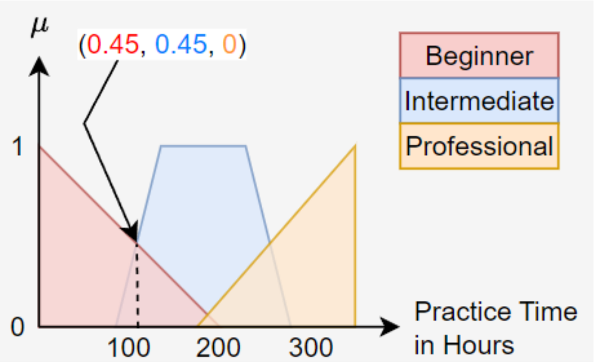
- **Published by Lotfi Zaddeh in 1965 [4]**
  - Traditionall set membership: $\mu_{trad}: \mathbb{R} \rightarrow \{0,1\}$
  - Fuzzy sets: Use Membership Functions (MFs)
    - $\mu_X: \mathbb{R} \rightarrow [0,1]$
    - $a \in \mathbb{R}$, fuzzy set $X, \mu_X(a) \in [0,1]$



[4] L. Zadeh, Fuzzy Sets, Information and Control, vol. 8, no. 3, pp. 338-353
[5] E. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis
[6] Q. Liang and J. Mendel, An Introduction to Type-2 TSK Fuzzy Logic Systems

FUNDAMENTALS OF FUZZY SETS

- **Published by Lotfi Zaddeh in 1965 [4]**
  - Traditionall set membership: $\mu_{trad}: \mathbb{R} \rightarrow \{0,1\}$
  - Fuzzy sets: Use Membership Functions (MFs)
    - $\mu_X: \mathbb{R} \rightarrow [0,1]$
    - $a \in \mathbb{R}$, fuzzy set $X, \mu_X(a) \in [0,1]$

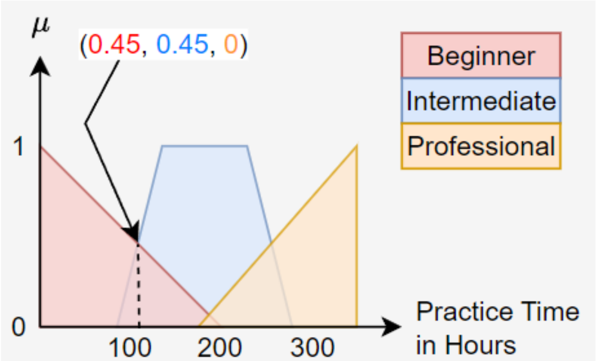


- **Vagueness is important for reasoning [5]**
  - Are bacteria anmials? [4]

[4] L. Zadeh, Fuzzy Sets, Information and Control, vol. 8, no. 3, pp. 338-353
[5] E. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis
[6] Q. Liang and J. Mendel, An Introduction to Type-2 TSK Fuzzy Logic Systems

FUNDAMENTALS OF FUZZY SETS

- **Published by Lotfi Zaddeh in 1965 [4]**
  - Traditionall set membership: $\mu_{trad}: \mathbb{R} \rightarrow \{0,1\}$
  - Fuzzy sets: Use Membership Functions (MFs)
    - $\mu_X: \mathbb{R} \rightarrow [0,1]$
    - $a \in \mathbb{R}$, fuzzy set $X, \mu_X(a) \in [0,1]$



- **Vagueness is important for reasoning [5]**
  - Are bacteria anmials? [4]
- **Types of fuzzy sets**
  - Type 0: Crisp outputs of Takagi-Sugeno-Kang (TSK)-Systems (used in [6])
  - Type 1: Degree of membership (we will constrain ourselves to these)
  - Type 2: Model multiple sources of uncertainties in MFs

[4] L. Zadeh, Fuzzy Sets, Information and Control, vol. 8, no. 3, pp. 338-353
[5] E. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis
[6] Q. Liang and J. Mendel, An Introduction to Type-2 TSK Fuzzy Logic Systems

FUNDAMENTALS OF FUZZY INFERENCE

- **Operations compute memberships:**
  - Let A, B be fuzzy sets, μ_A, μ_B be MFs, $a, b \in \mathbb{R}$.
    Furthermore, $x := \mu_A(a)$ and $y := \mu_B(b)$ are fuzzy variables
  - **AND** $\mu_{A \cap B}$:

[9] J. M. Mendel and P. P. Bonissone, Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems
[10] D. K. Wedding, Fuzzy Sets and Fuzzy Logic: Theory and Applications: George J. Klir and Bo Yuan

FUNDAMENTALS OF FUZZY INFERENCE

- **Operations compute memberships:**
  - Let A, B be fuzzy sets, μ_A, μ_B be MFs, $a, b \in \mathbb{R}$.
    Furthermore, $x := \mu_A(a)$ and $y := \mu_B(b)$ are fuzzy variables
  - **AND** $\mu_{A \cap B}$:
    - $\min\{x, y\}$
    - $x * y$
    - Other t-Norms with $t(x, y)$

[9] J. M. Mendel and P. P. Bonissone, Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems
[10] D. K. Wedding, Fuzzy Sets and Fuzzy Logic: Theory and Applications: George J. Klir and Bo Yuan

FUNDAMENTALS OF FUZZY INFERENCE

- **Operations compute memberships:**
  - Let A, B be fuzzy sets, μ_A, μ_B be MFs, $a, b \in \mathbb{R}$.
    Furthermore, $x := \mu_A(a)$ and $y := \mu_B(b)$ are fuzzy variables
  - **AND** $\mu_{A \cap B}$:
    - $\min\{x, y\}$
    - $x * y$
    - Other t-Norms with $t(x, y)$
  - **OR** $\mu_{A \cup B}$: $\max\{x, y\}$

[9] J. M. Mendel and P. P. Bonissone, Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems
[10] D. K. Wedding, Fuzzy Sets and Fuzzy Logic: Theory and Applications: George J. Klir and Bo Yuan

FUNDAMENTALS OF FUZZY INFERENCE

- **Operations compute memberships:**
  - Let A, B be fuzzy sets, μ_A, μ_B be MFs, $a, b \in \mathbb{R}$.
    Furthermore, $x := \mu_A(a)$ and $y := \mu_B(b)$ are fuzzy variables
  - **AND** $\mu_{A \cap B}$:
    - $\min\{x, y\}$
    - $x * y$
    - Other t-Norms with $t(x, y)$
  - **OR** $\mu_{A \cup B}$: $\max\{x, y\}$
  - **NOT** $\mu_{\neg A}$: $1 - x$

[9] J. M. Mendel and P. P. Bonissone, Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems
[10] D. K. Wedding, Fuzzy Sets and Fuzzy Logic: Theory and Applications: George J. Klir and Bo Yuan

FUNDAMENTALS OF FUZZY INFERENCE

- **Operations compute memberships:**
  - Let A, B be fuzzy sets, μ_A, μ_B be MFs, $a, b \in \mathbb{R}$.
    Furthermore, $x := \mu_A(a)$ and $y := \mu_B(b)$ are fuzzy variables
  - **AND** $\mu_{A \cap B}$:
    - $\min\{x, y\}$
    - $x * y$
    - Other t-Norms with $t(x, y)$
  - **OR** $\mu_{A \cup B}$: $\max\{x, y\}$
  - **NOT**$\mu_{\neg A}$: $1 - x$
- **Often correlation instead of implication [9]**
  - 14+ types of fuzzy implications [10]

[9] J. M. Mendel and P. P. Bonissone, Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems
[10] D. K. Wedding, Fuzzy Sets and Fuzzy Logic: Theory and Applications: George J. Klir and Bo Yuan

FUNDAMENTALS OF FUZZY INFERENCE

Name	Symbol	Class	Function $\mathcal{J}(a, b)$	Axioms	Complement $c(a)$	Year
Early Zadeh	$\mathcal{J}_m$	QL	$\max[1-a, \min(a, b)]$	1, 2, 3, 4, 9	$1-a$	1973
Gaines-Rescher	$\mathcal{J}_s$		$\begin{cases} 1 & a \leq b \\ 0 & a > b \end{cases}$	1, 2, 3, 4, 5, 6, 7, 8		1969
Gödel	$\mathcal{J}_g$	R	$\begin{cases} 1 & a \leq b \\ b & a > b \end{cases}$	1, 2, 3, 4, 5, 6, 7		1976
Goguen	$\mathcal{J}_\Delta$	R	$\begin{cases} 1 & a \leq b \\ b/a & a > b \end{cases}$	1, 2, 3, 4, 5, 6, 7, 9		1969
Kleene-Dienes	$\mathcal{J}_b$	S, QL	$\max(1-a, b)$	1, 2, 3, 4, 6, 8, 9	$1-a$	1938, 1949
Lukasiewicz	$\mathcal{J}_a$	R, S	$\min(1, 1-a+b)$	1, 2, 3, 4, 5, 6, 7, 8, 9	$1-a$	1920
Pseudo-Lukasiewicz 1	$\mathcal{J}_\lambda$ $(\lambda > -1)$	R, S	$\min\left[1, \frac{1-a+(1+\lambda)b}{1+\lambda a}\right]$	1, 2, 3, 4, 5, 6, 7, 8, 9	$\frac{1-a}{1+\lambda a}$	1987
Pseudo-Lukasiewicz 2	$\mathcal{J}_w$ $(w > 0)$	R, S	$\min\left[1, (1-a^w+b^w)^{\frac{1}{w}}\right]$	1, 2, 3, 4, 5, 6, 7, 8, 9	$(1-a^w)^{\frac{1}{w}}$	1987
Reichenbach	$\mathcal{J}_r$	S	$1-a+ab$	1, 2, 3, 4, 6, 8, 9	$1-a$	1935
Willmott	$\mathcal{J}_{wl}$		$\min[\max(1-a, b), \max(a, 1-a), \max(b, 1-b)]$	4, 6, 8, 9	$1-a$	1980
Wu	$\mathcal{J}_{wu}$		$\begin{cases} 1 & a \leq b \\ \min(1-a, b) & a > b \end{cases}$	1, 2, 3, 5, 7, 8	$1-a$	1986
Yager	$\mathcal{J}_y$		$\begin{cases} 1 & a = b = 0 \\ b^a & \text{others} \end{cases}$	1, 2, 3, 4, 6		1980
Klir and Yuan 1	$\mathcal{J}_p$	QL	$1-a+a^2b$	2, 3, 4, 9	$1-a$	1994
Klir and Yuan 2	$\mathcal{J}_q$	QL	$\begin{cases} b & a = 1 \\ 1-a & a \neq 1, b \neq 1 \\ 1 & a \neq 1, b = 1 \end{cases}$	2, 4	$1-a$	1994

Table 11.1 from [10]

[10] D. K. Wedding, Fuzzy Sets and Fuzzy Logic: Theory and Applications: George J. Klir and Bo Yuan

FUNDAMENTALS OF FUZZY INFERENCE SYSTEMS (FIS)

- **Computational models built upon fuzzy logic**
  - = Fuzzy Inference Systems (FIS)
  - Popular variants: Mamdani and Takagi-Sugeno-Kang (TSK)

[8] K. Wang, Computational Intelligence in Agile Manufacturing Engineering: The 21st Century Competitive Strategy

FUNDAMENTALS OF FUZZY INFERENCE SYSTEMS (FIS)

- **Computational models built upon fuzzy logic**
  - = Fuzzy Inference Systems (FIS)
  - Popular variants: Mamdani and Takagi-Sugeno-Kang (TSK)

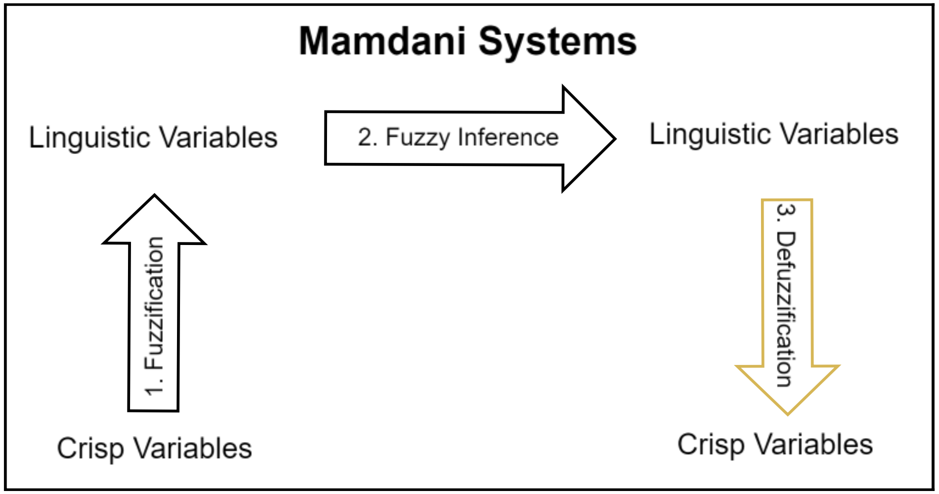


Image adapted from [8]

[8] K. Wang, Computational Intelligence in Agile Manufacturing Engineering: The 21st Century Competitive Strategy

FUNDAMENTALS OF FUZZY INFERENCE SYSTEMS (FIS)

- **Computational models built upon fuzzy logic**
  - = Fuzzy Inference Systems (FIS)
  - Popular variants: Mamdani and Takagi-Sugeno-Kang (TSK)

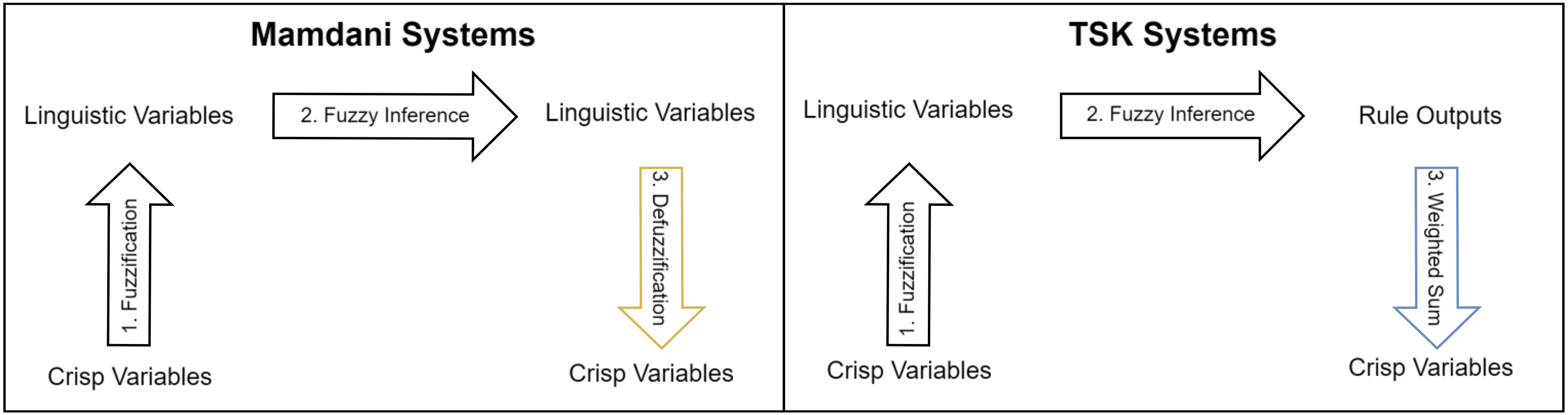


Image adapted from [8]

[8] K. Wang, Computational Intelligence in Agile Manufacturing Engineering: The 21st Century Competitive Strategy

EXAMPLE MAMDANI SYSTEM

Example: Competition success estimation for *Capture The Flag* competitions

EXAMPLE MAMDANI SYSTEM

Example: Competition success estimation for *Capture The Flag* competitions

- **Inputs:**
  - Practice time in hours: *float*
  - Level duration in weeks: *float*

EXAMPLE MAMDANI SYSTEM

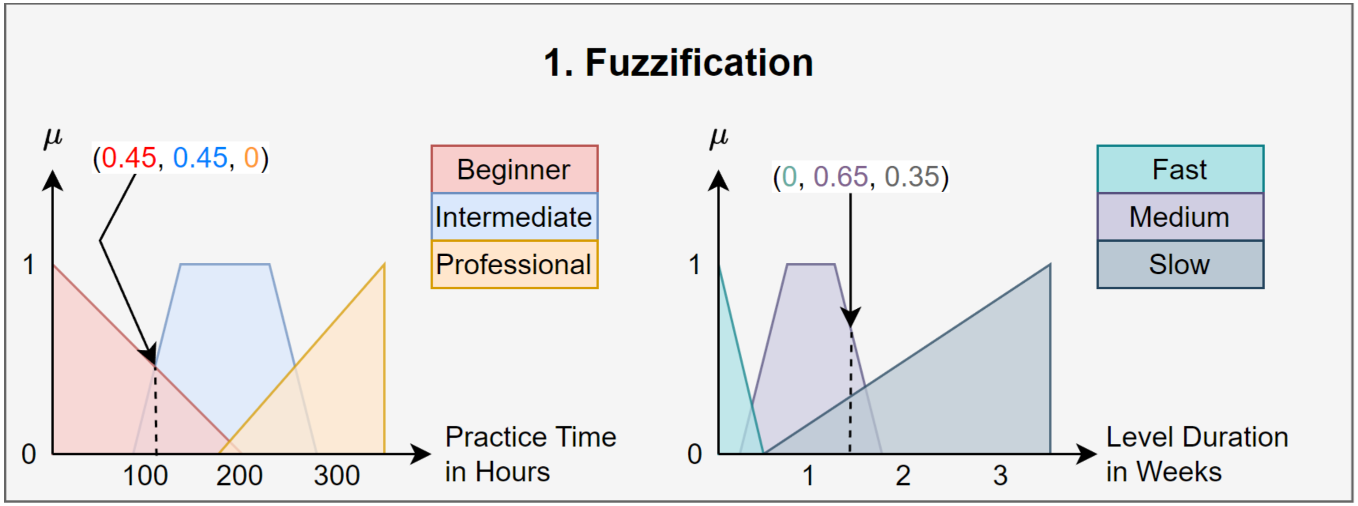
Example: Competition success estimation for *Capture The Flag* competitions

- **Inputs:**
  - Practice time in hours: *float*
  - Level duration in weeks: *float*
- **Output:**
  - Expected placement in competition: *float*

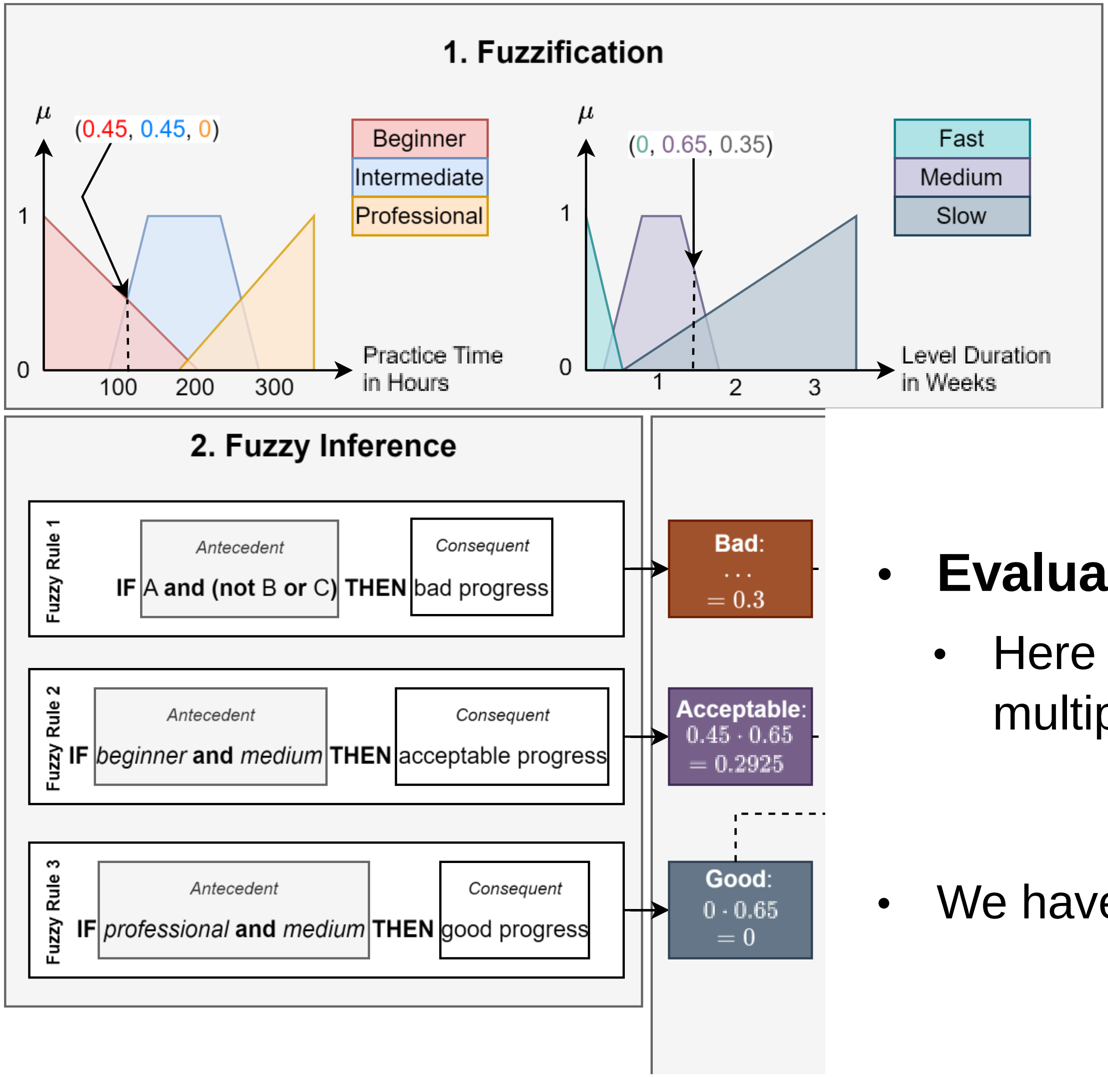
EXAMPLE MAMDANI SYSTEM

Example: Competition success estimation for *Capture The Flag* competitions

- **Inputs:**
  - Practice time in hours: *float*
  - Level duration in weeks: *float*
- **Output:**
  - Expected placement in competition: *float*



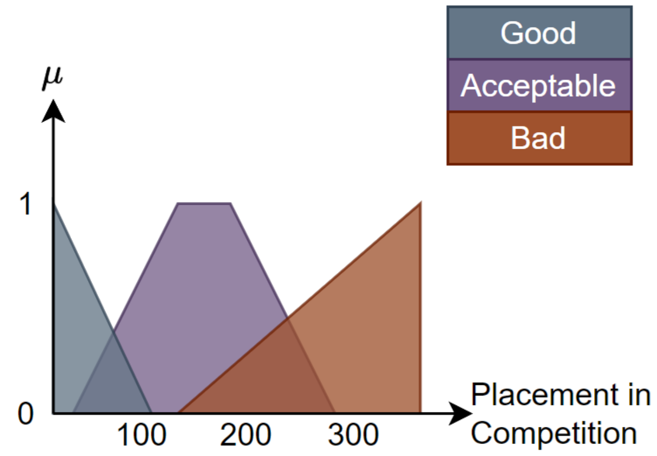
EXAMPLE MAMDANI SYSTEM: FUZZY INFERENCE



- **Evaluate Antecedents**
  - Here **AND** is modeled as multiplication
- We have natural language results!

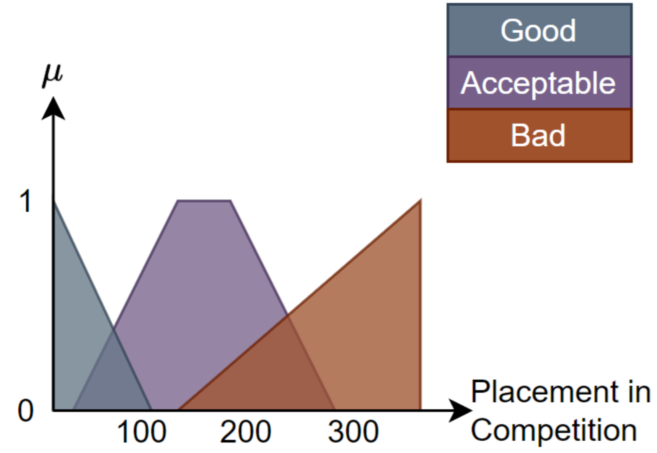
MEMBERSHIP FUNCTION FOR DEFUZZIFICATION

- **Consequent will cap defuzzification**
  - $\min(\mu_X, c_X)$ for all sets X and consequent values c
  - Good = 0
  - Acceptable ≈ 0.3
  - Bad = 0.3

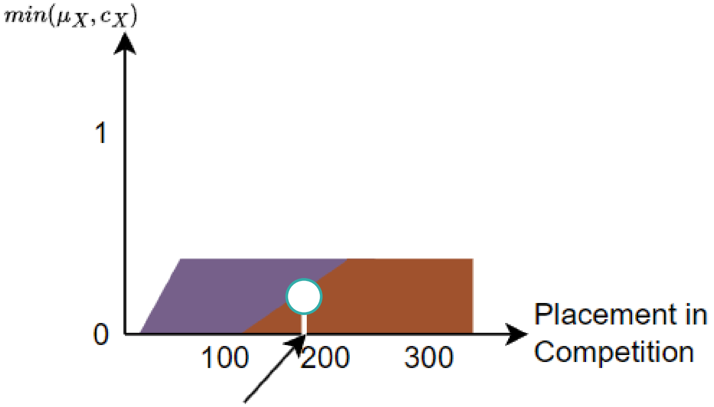


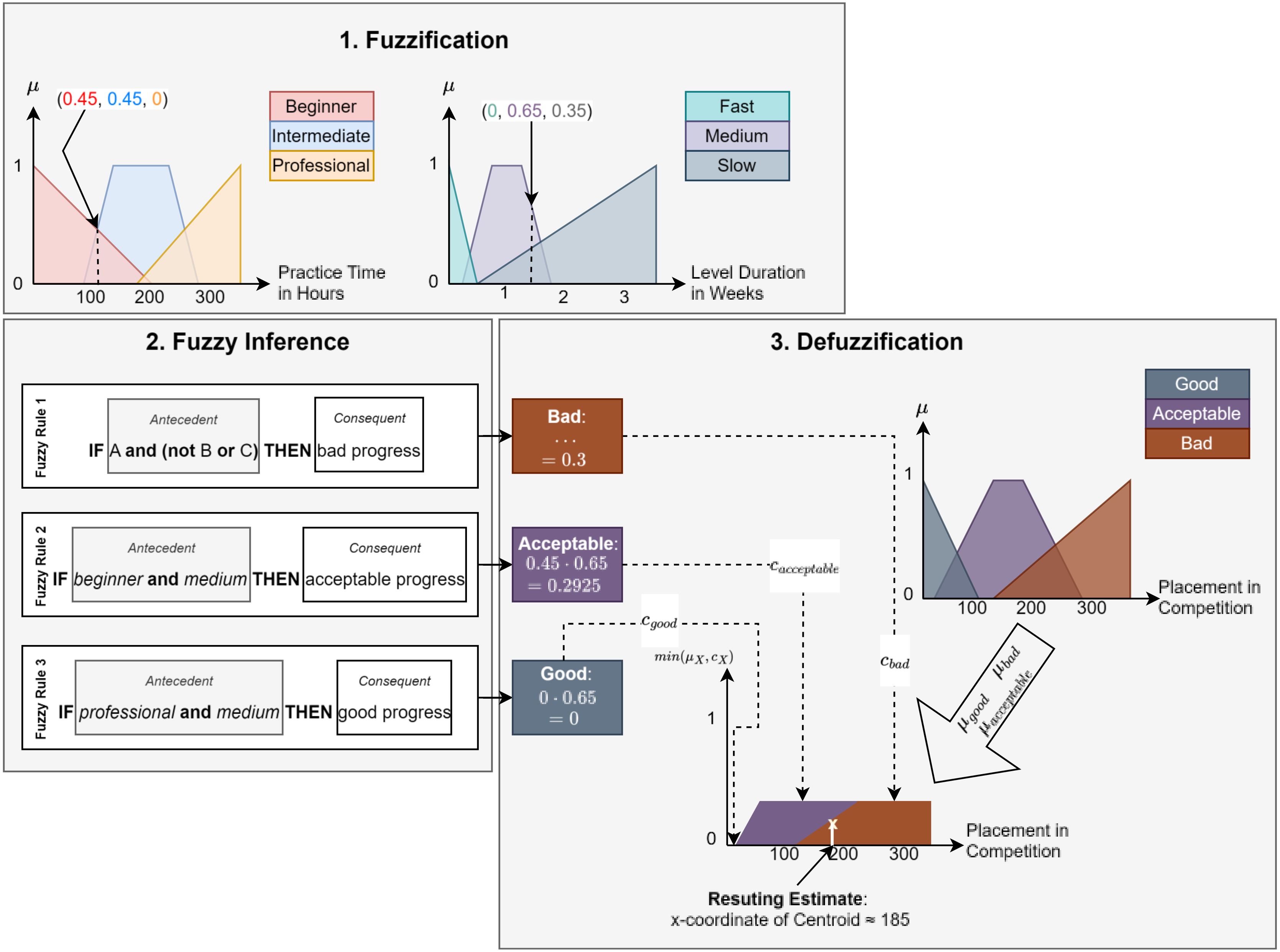
MEMBERSHIP FUNCTION FOR DEFUZZIFICATION

- **Consequent will cap defuzzification**
  - $\min(\mu_X, c_X)$ for all sets X and consequent values c
  - Good = 0
  - Acceptable ≈ 0.3
  - Bad = 0.3



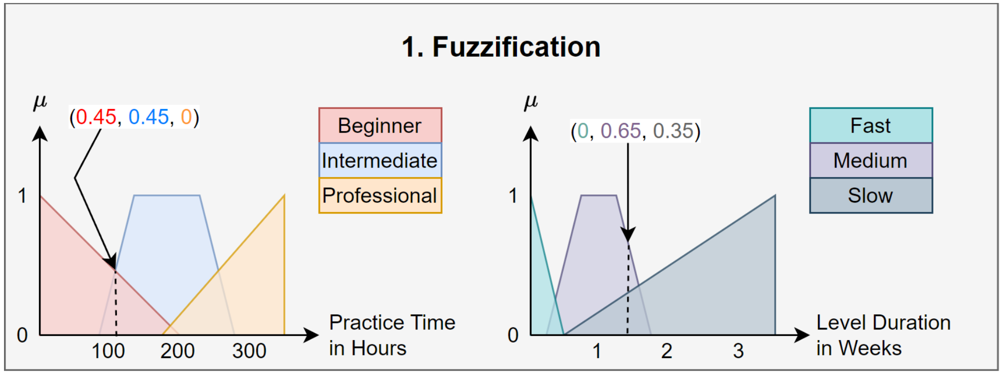
- **Domain-Value of MF-Centroid will be the output**
  - Only on the remaining shape




1. Fuzzification
μ
(0.45, 0.45, 0)
Beginner
Intermediate
Professional
1
0
100
200
300
Practice Time in Hours
μ
(0, 0.65, 0.35)
Fast
Medium
Slow
1
0
1
2
3
Level Duration in Weeks
2. Fuzzy Inference
Fuzzy Rule 1
Antecedent
Consequent
IF A and (not B or C) THEN bad progress
Fuzzy Rule 2
Antecedent
Consequent
IF beginner and medium THEN acceptable progress
Fuzzy Rule 3
Antecedent
Consequent
IF professional and medium THEN good progress
3. Defuzzification
Bad: ... = 0.3
Acceptable: 0.45 · 0.65 = 0.2925
Good: 0 · 0.65 = 0
Good
Acceptable
Bad
μ
1
0
100
200
300
Placement in Competition
c_acceptable
c_good
c_bad
min(μ_X, c_X)
μ_good μ_bad μ_acceptable
1
0
100
200
300
Placement in Competition
Resuting Estimate:
x-coordinate of Centroid ≈ 185

TAKAGI-SUGENO-KANG DEFUZZIFICATION

- **Inference outputs are “Type-0“ (=crisp) polynomials [11]**
  - Use antecedent values as weight for conclusions
  - Weighted sum as output



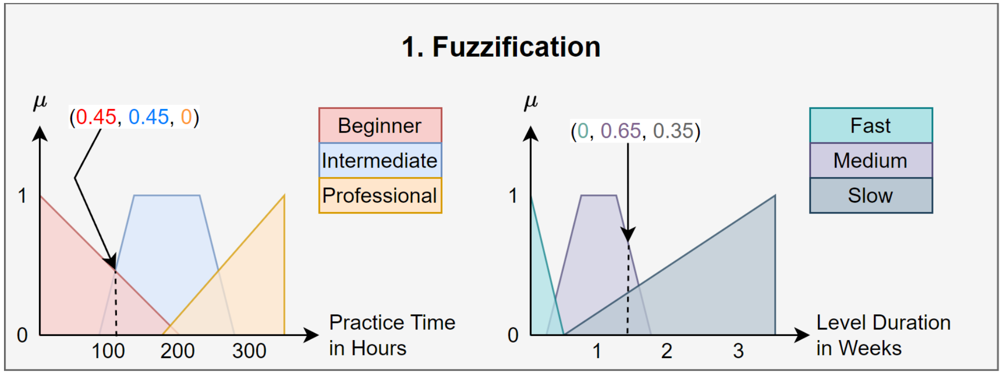
Example

h: Practice time in hours = 110
d: Level duration in weeks = 1.5

[11] T. Takagi and M. Sugeno, Fuzzy Identification of Systems and Its Applications to Modeling and Control

TAKAGI-SUGENO-KANG DEFUZZIFICATION

- **Inference outputs are "Type-0" (=crisp) polynomials [11]**
  - Use antecedent values as weight for conclusions
  - Weighted sum as output



Example

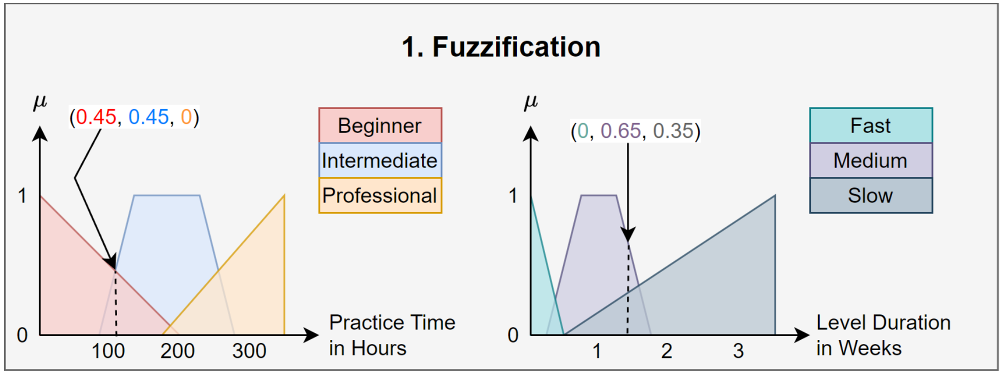
h: Practice time in hours = 110
d: Level duration in weeks = 1.5

```
IF beginner AND fast THEN placement = 150 – h + 20*d
IF professional OR medium THEN placement = 200-h
```

[11] *T. Takagi and M. Sugeno, Fuzzy Identification of Systems and Its Applications to Modeling and Control*

TAKAGI-SUGENO-KANG DEFUZZIFICATION

- **Inference outputs are "Type-0" (=crisp) polynomials [11]**
  - Use antecedent values as weight for conclusions
  - Weighted sum as output



Example

h: Practice time in hours = 110
d: Level duration in weeks = 1.5

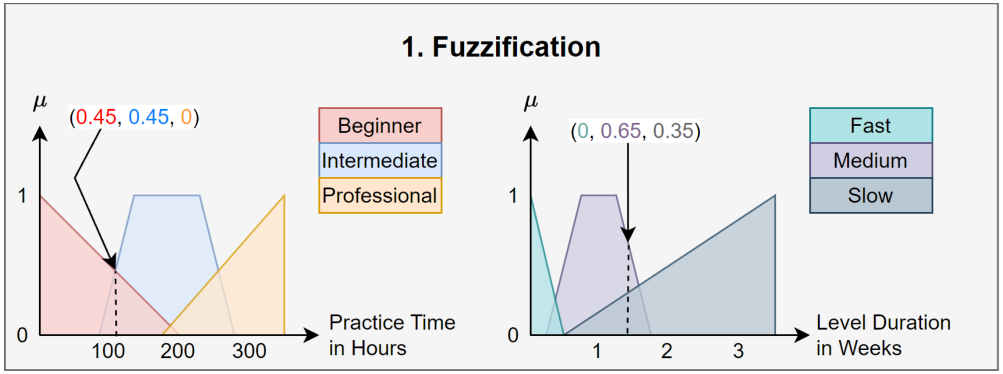
```
IF beginner AND fast THEN placement = 150 – h + 20*d
IF professional OR medium THEN placement = 200-h
```

We get Placement = $\frac{(0.45*0)*(150-110+20*1.5)+\max\{0,0.65\}*(200-h)}{(0.45*0)+\max\{0,0.65\}} = 90$

[11] T. Takagi and M. Sugeno, Fuzzy Identification of Systems and Its Applications to Modeling and Control

TAKAGI-SUGENO-KANG DEFUZZIFICATION

- **Inference outputs are “Type-0“ (=crisp) polynomials [11]**
  - Use antecedent values as weight for conclusions
  - Weighted sum as output



Example

h: Practice time in hours = 110
d: Level duration in weeks = 1.5

```
IF beginner AND fast THEN placement = 150 – h + 20*d
IF professional OR medium THEN placement = 200-h
```

We get Placement $= \dfrac{\cancel{(0.45*0)*(150-110+20*1.5)}+\max\{0,0.65\}*(200-h)}{\cancel{(0.45*0)}+\max\{0,0.65\}} = 90$

[11] T. Takagi and M. Sugeno, Fuzzy Identification of Systems and Its Applications to Modeling and Control

VISUALIZING FUZZY INFERENCE SYSTEMS

Simple methods highlighted by [18]

[18] J. Cao, et al., Fuzzy Inference System With Interpretable Fuzzy Rules: Advancing Explainable Artifcial Intelligence for Disease Diagnosis

[12] https://github.com/tenaci-hand-grip/fuzzy-visualization, 05.04.2024

VISUALIZING FUZZY INFERENCE SYSTEMS

Simple methods highlighted by [18]

Explicit Rule List

IF (A **AND** B) THEN C=2
IF (A **OR** D) THEN E=4, C=1
IF (G **OR** (**NOT** F)) THEN C=0

[18] J. Cao, et al., Fuzzy Inference System With Interpretable Fuzzy Rules: Advancing Explainable Artifcial Intelligence for Disease Diagnosis
[12] https://github.com/tenaci-hand-grip/fuzzy-visualization, 05.04.2024

VISUALIZING FUZZY INFERENCE SYSTEMS

Simple methods highlighted by [18]

Explicit Rule List

IF (A **AND** B) THEN C=2
IF (A **OR** D) THEN E=4, C=1
IF (G **OR** (**NOT** F)) THEN C=0

Explicit Rule List with MF

	Antecedent parameter 1	Consequent parameters
Rule 1	$\mu = 0.61; \sigma^2 = 0.5$	$p^0 = 0.589$
Rule 2	$\mu = 0; \sigma^2 = 1$	$p^0 = 0.195$
...	...	...
Rule k	$\mu = 12; \sigma^2 = 4$	$p^3 = 0.613$

[18] J. Cao, et al., Fuzzy Inference System With Interpretable Fuzzy Rules: Advancing Explainable Artifcial Intelligence for Disease Diagnosis
[12] https://github.com/tenaci-hand-grip/fuzzy-visualization, 05.04.2024

VISUALIZING FUZZY INFERENCE SYSTEMS

Simple methods highlighted by [18]

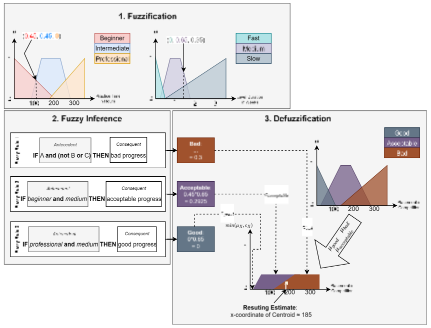
Explicit Rule List

IF (A **AND** B) THEN C=2
IF (A **OR** D) THEN E=4, C=1
IF (G **OR** (**NOT** F)) THEN C=0

Explicit Rule List with MF

	Antecedent parameter 1	Consequent parameters
Rule 1	$\mu = 0.61; \sigma^2 = 0.5$	$p^0 = 0.589$
Rule 2	$\mu = 0; \sigma^2 = 1$	$p^0 = 0.195$
...	...	...
Rule k	$\mu = 12; \sigma^2 = 4$	$p^3 = 0.613$

Showing Inference for One Sample



[18] J. Cao, et al., Fuzzy Inference System With Interpretable Fuzzy Rules: Advancing Explainable Artifcial Intelligence for Disease Diagnosis
[12] https://github.com/tenaci-hand-grip/fuzzy-visualization, 05.04.2024

VISUALIZING FUZZY INFERENCE SYSTEMS

Simple methods highlighted by [18]

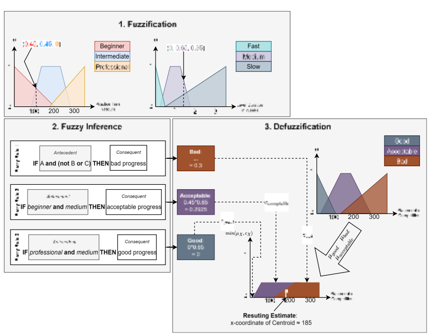
Explicit Rule List

IF (A **AND** B) THEN C=2
IF (A **OR** D) THEN E=4, C=1
IF (G **OR** (**NOT** F)) THEN C=0

Explicit Rule List with MF

	Antecedent parameter 1	Consequent parameters
Rule 1	$\mu = 0.61; \sigma^2 = 0.5$	$p^0 = 0.589$
Rule 2	$\mu = 0; \sigma^2 = 1$	$p^0 = 0.195$
...	...	...
Rule k	$\mu = 12; \sigma^2 = 4$	$p^3 = 0.613$

Showing Inference for One Sample



3D Surface-Plots

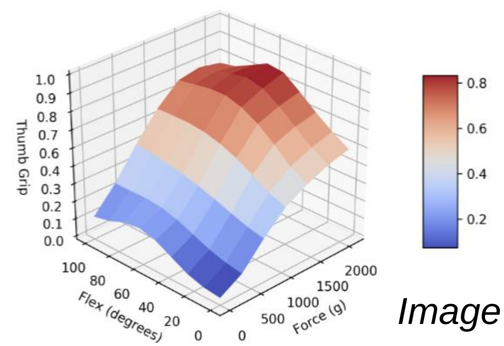


Image taken from [12]

[18] J. Cao, et al., Fuzzy Inference System With Interpretable Fuzzy Rules: Advancing Explainable Artifcial Intelligence for Disease Diagnosis
[12] https://github.com/tenaci-hand-grip/fuzzy-visualization, 05.04.2024

SECTION 3

Fuzzy Rule-Based Regression

EXAMPLE: ASSET RISK VALIDATION

- **Risk estimation of IT-Assets (BAKK)[1]**

1: Recommended to me by Univ.-Prof. Priv.-Doz. DDI Dr. Rass

EXAMPLE: ASSET RISK VALIDATION

- **Risk estimation of IT-Assets (BAKK)[1]**
  - You are the Chief Information Security Officer (CISO) of a big company
  - Get asset data and risk scores from different departments
  - Trust it blindly?

1: Recommended to me by Univ.-Prof. Priv.-Doz. DDI Dr. Rass

EXAMPLE: ASSET RISK VALIDATION

- **Risk estimation of IT-Assets (BAKK)[1]**
  - You are the Chief Information Security Officer (CISO) of a big company
  - Get asset data and risk scores from different departments
  - Trust it blindly?

  - **Automatically create many rules and FIS**
  - Linear regression

1: Recommended to me by Univ.-Prof. Priv.-Doz. DDI Dr. Rass

EXAMPLE: ASSET RISK VALIDATION

- **Risk estimation of IT-Assets (BAKK)[1]**
  - You are the Chief Information Security Officer (CISO) of a big company
  - Get asset data and risk scores from different departments
  - Trust it blindly?

  - **Automatically create many rules and FIS**
  - Linear regression

    → Very high or low (negative coefficients) indicate important rules

1: Recommended to me by Univ.-Prof. Priv.-Doz. DDI Dr. Rass

EXAMPLE: ASSET RISK VALIDATION

- **Risk estimation of IT-Assets (BAKK)[1]**
  - You are the Chief Information Security Officer (CISO) of a big company
  - Get asset data and risk scores from different departments
  - Trust it blindly?

  - **Automatically create many rules and FIS**
  - Linear regression

    → Very high or low (negative coefficients) indicate important rules
  - Statistical testing for rule (ir-)relevance

1: Recommended to me by Univ.-Prof. Priv.-Doz. DDI Dr. Rass

EXAMPLE: ASSET RISK VALIDATION

- **Risk estimation of IT-Assets (BAKK)[1]**
  - You are the Chief Information Security Officer (CISO) of a big company
  - Get asset data and risk scores from different departments
  - Trust it blindly?

  - **Automatically create many rules and FIS**
  - Linear regression

    → Very high or low (negative coefficients) indicate important rules
  - Statistical testing for rule (ir-)relevance
  - **Explain the data generation process**
    - Deviations from best practices/agreed frameworks?
    - Conflicts or opposing actions?

1: Recommended to me by Univ.-Prof. Priv.-Doz. DDI Dr. Rass

LINEAR REGRESSION OVER FIS

- **Goal: Find rules that explain the data**
- **Combine** $k \in \mathbb{N}$ **such FIS** $g_1, \ldots, g_k: \mathbb{R}^d \to \mathbb{R}$ **linearly** *similar to [13,14]*

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[14] C. Cichy and S. Rass, Fuzzy Expert Systems for Automated Data Quality Assessment and Improvement Processes

LINEAR REGRESSION OVER FIS

- **Goal: Find rules that explain the data**
- **Combine** $k \in \mathbb{N}$ **such FIS** $g_1, \ldots, g_k: \mathbb{R}^d \to \mathbb{R}$ **linearly** *similar to [13,14]*
  - Let $\boldsymbol{x_i} \in \mathbb{R}^d$ be the i-th input vector with d dimensions
  - Let $\alpha_0, \alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be the **rule-weights** (parameters) we will tune later

$$f(x_i, \alpha_0, \alpha_1, \ldots, \alpha_k) = \alpha_0 + \alpha_1 g_1(\boldsymbol{x_i}) + \cdots + \alpha_k g_k(\boldsymbol{x_i})$$

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[14] C. Cichy and S. Rass, Fuzzy Expert Systems for Automated Data Quality Assessment and Improvement Processes

LINEAR REGRESSION OVER FIS

- **Goal: Find rules that explain the data**
- **Combine** $k \in \mathbb{N}$ **such FIS** $g_1, \dots, g_k: \mathbb{R}^d \to \mathbb{R}$ **linearly** *similar to [13,14]*
  - Let $\boldsymbol{x_i} \in \mathbb{R}^d$ be the i-th input vector with d dimensions
  - Let $\alpha_0, \alpha_1, \dots, \alpha_k \in \mathbb{R}$ be the **rule-weights** (parameters) we will tune later

$$f(x_i, \alpha_0, \alpha_1, \dots, \alpha_k) = \alpha_0 + \alpha_1 g_1(\boldsymbol{x_i}) + \cdots + \alpha_k g_k(\boldsymbol{x_i})$$

- **Minimize Ordinary Least Squares**
  - Let $y_i \in \mathbb{R}$ be the i-th output label (could also be defined multidimensional)

$$SSE = \sum_{i=1}^{n} \left(y_i - f(\boldsymbol{x_i}, \alpha_0, \alpha_1, \dots, \alpha_n)\right)^2$$

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[14] C. Cichy and S. Rass, Fuzzy Expert Systems for Automated Data Quality Assessment and Improvement Processes

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Minimize Ordinary Least Squares**
  - Let $y_i \in \mathbb{R}$ be the i-th output label (could also be defined multidimensional)

$$SSE = \sum_{i=1}^{n} \left(y_i - \alpha_0 - \alpha_1 g_1(\boldsymbol{x_i}) - \cdots - \alpha_k g_k(\boldsymbol{x_i})\right)^2$$

[15] W. H. Press et al., Numerical Recipes in C,2nd ed.

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Minimize Ordinary Least Squares**
  - Let $y_i \in \mathbb{R}$ be the i-th output label (could also be defined multidimensional)

$$SSE = \sum_{i=1}^{n} \left(y_i - \alpha_0 - \alpha_1 g_1(\boldsymbol{x_i}) - \cdots - \alpha_k g_k(\boldsymbol{x_i})\right)^2$$

  - General least squares problem with g_i as basis functions [15]

[15] W. H. Press et al., Numerical Recipes in C,2nd ed.

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Minimize Ordinary Least Squares**
  - Let $y_i \in \mathbb{R}$ be the i-th output label (could also be defined multidimensional)

$$SSE = \sum_{i=1}^{n} \left(y_i - \alpha_0 - \alpha_1 g_1(\boldsymbol{x_i}) - \cdots - \alpha_k g_k(\boldsymbol{x_i})\right)^2$$

  - General least squares problem with g_i as basis functions [15]

  - Solution: Set first derivative to zero → normal equations [15]

[15] W. H. Press et al., Numerical Recipes in C,2nd ed.

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Minimize Ordinary Least Squares**
  - Let $y_i \in \mathbb{R}$ be the i-th output label (could also be defined multidimensional)

$$SSE = \sum_{i=1}^{n} \left(y_i - \alpha_0 - \alpha_1 g_1(\boldsymbol{x_i}) - \cdots - \alpha_k g_k(\boldsymbol{x_i}) \right)^2$$

  - General least squares problem with g_i as basis functions [15]

  - Solution: Set first derivative to zero → normal equations [15]
    - No guarantee for global optimality → more examination/tries
    - Numerically more stable with QR-Decomposition

[15] W. H. Press et al., Numerical Recipes in C,2nd ed.

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Testing Significance Levels of Coefficients**
  - Can some g_i be dropped with same performance? [13]

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[17] M. Chvosteková and V. Witkovský, Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model With Normal Errors

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Testing Significance Levels of Coefficients**
  - Can some g_i be dropped with same performance? [13]
    - Statsitical test: Likelihood Ratio Test (let $\mathcal{L}$ denote likelihood)
    - H_0: Basis function g_i can be dropped
    - Choose:

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[17] M. Chvosteková and V. Witkovský, Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model With Normal Errors

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Testing Significance Levels of Coefficients**
  - Can some g_i be dropped with same performance? [13]
    - Statsitical test: Likelihood Ratio Test (let $\mathcal{L}$ denote likelihood)
    - H_0: Basis function g_i can be dropped
    - Choose:
      - Significance level (0.05 common) and
      - Degree of freedom: difference in parameters

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[17] M. Chvosteková and V. Witkovský, Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model With Normal Errors

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Testing Significance Levels of Coefficients**
  - Can some g_i be dropped with same performance? [13]
    - Statsitical test: Likelihood Ratio Test (let $\mathcal{L}$ denote likelihood)
    - H_0: Basis function g_i can be dropped
    - Choose:
      - Significance level (0.05 common) and
      - Degree of freedom: difference in parameters
  - **Likelihood Ratio Test**: $\Lambda = 2(\log(\mathcal{L}_{full}) - \log(\mathcal{L}_{reduced}))$
    - Assume Gaussian error distribution: $\Lambda \sim \mathrm{X}^2$ [17]
      - Support: “Central Limit Theorem“ and “Principle of Maximum Entropy“
    - If Λ > $\mathrm{X}^2_{0.05,1}$(crit.) → reject H_0 → g_i is important, keep!

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[17] M. Chvosteková and V. Witkovský, Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model With Normal Errors

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Testing Significance Levels of Coefficients [13]**
  - **Likelihood Ratio Test**: $\Lambda = 2(\log(\mathcal{L}_{full}) - \log(\mathcal{L}_{reduced}))$~ X^2
    - If Λ > $\mathrm{X}^2_{0.05,1}$(crit.) → reject H_0 → Variable is important, keep!

[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[16] E. L. Lehmann and J. P. Romano, Testing Statistical Hypotheses, 3rd ed

REGRESSION OVER FUZZY INFERENCE SYSTEMS

- **Testing Significance Levels of Coefficients [13]**
  - **Likelihood Ratio Test**: $\Lambda = 2(\log(\mathcal{L}_{full}) - \log(\mathcal{L}_{reduced}))$~ X^2
    - If Λ > $\mathrm{X}^2_{0.05,1}$(crit.) → reject H_0 → Variable is important, keep!

  - Under Gaussian distribution assumptions for the data and Maximum Likelihood Estimation (MLE) for variance $\sigma^2 = \frac{SSE}{n}$ [16]:

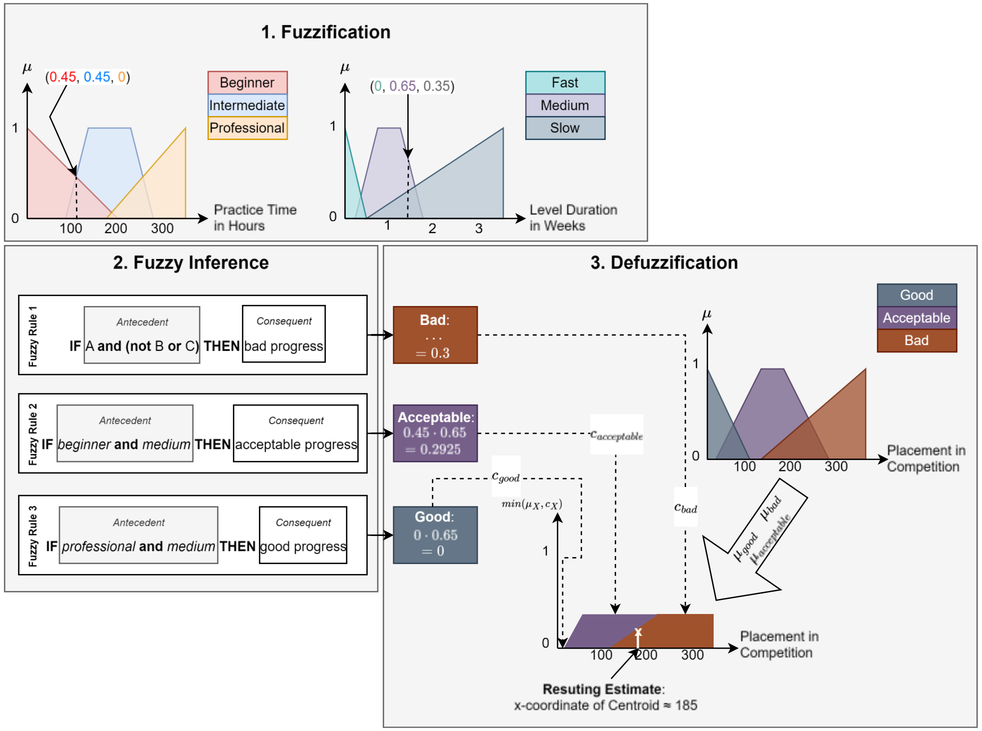
$$\log \mathcal{L}(SSE, n) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\frac{SSE}{n}\right) - \frac{n}{2}$$

n: number of samples
SSE: Sum of Squared Errors

  - n is equal so we only test the SSE of both models

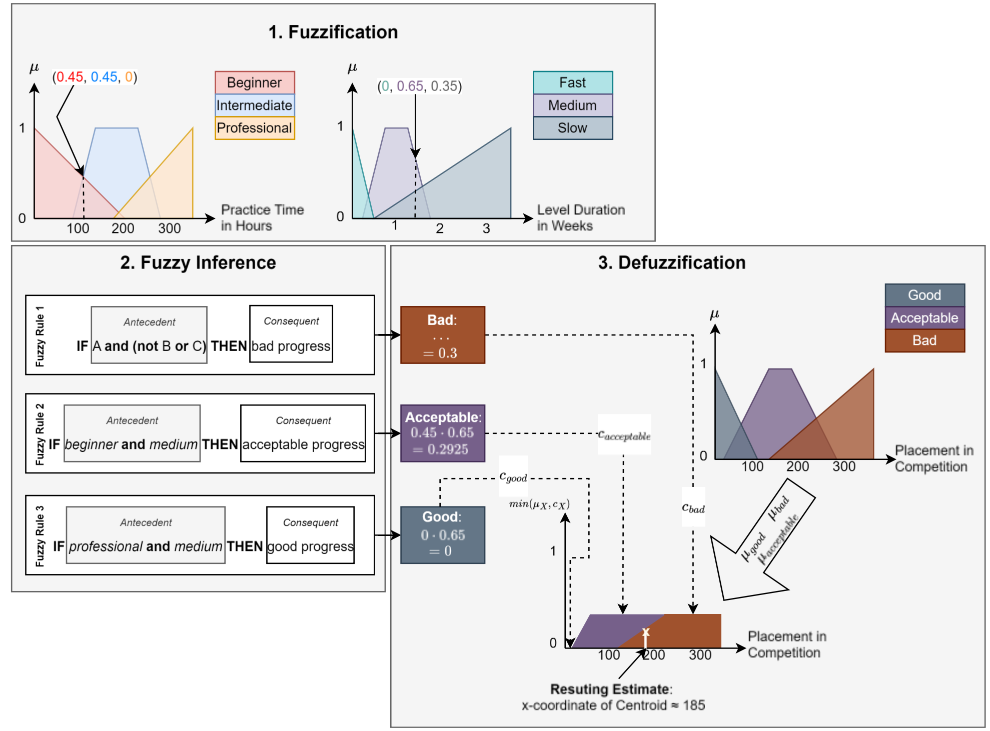
[13] C. Cichy and S. Rass, A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment
[16] E. L. Lehmann and J. P. Romano, Testing Statistical Hypotheses, 3rd ed

CONCLUSION



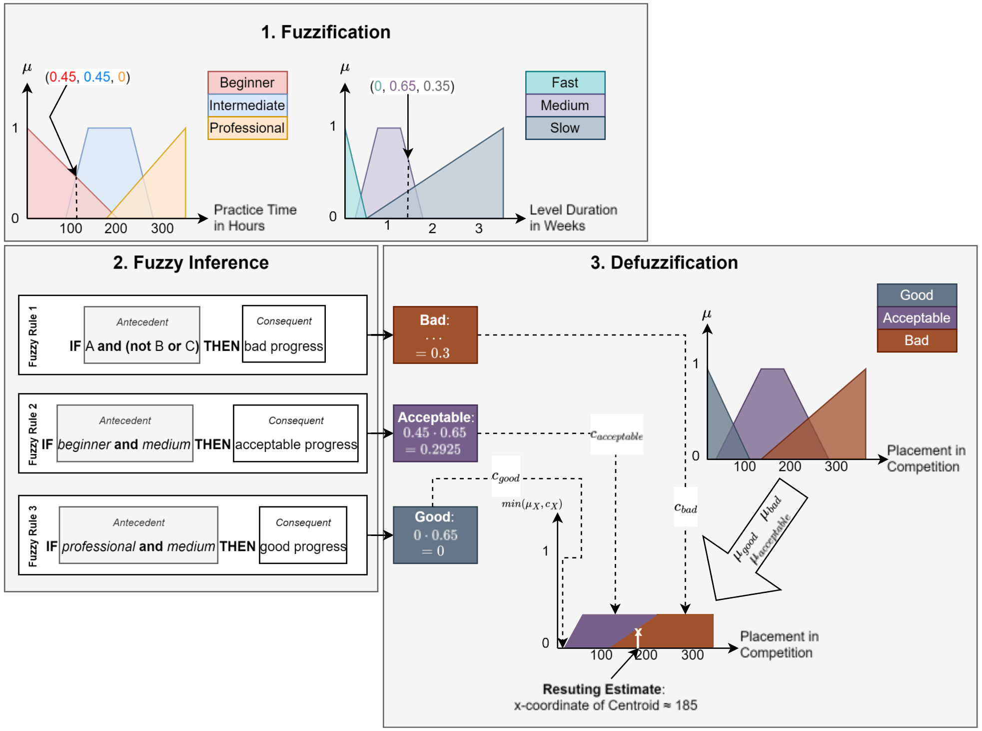
- XAI has many facets and requirements
- Fuzzy logic is inherently explainable
  - Natural language
  - Many possible visualizations
- Determining rule importance by regressing over fuzzy systems

CONCLUSION



- XAI has many facets and requirements
- Fuzzy logic is inherently explainable
  - Natural language
  - Many possible visualizations
- Determining rule importance by regressing over fuzzy systems
- Only scratched the surface – also interesting: Adaptive Neuro-Fuzzy-Inference-Systems (ANFIS)

CONCLUSION



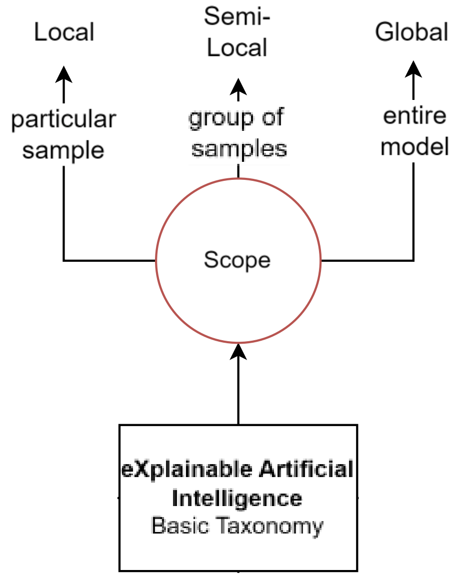
- XAI has many facets and requirements
- Fuzzy logic is inherently explainable
  - Natural language
  - Many possible visualizations
- Determining rule importance by regressing over fuzzy systems
- Only scratched the surface – also interesting: Adaptive Neuro-Fuzzy-Inference-Systems (ANFIS)
- Time for Questions!

APPENDIX

Martin Dallinger; 15.05.2024

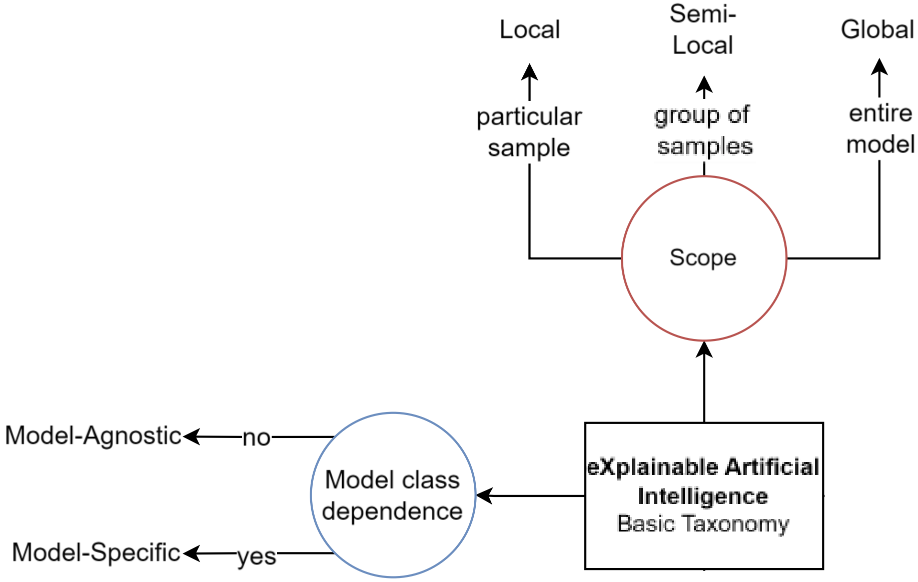
JOHANNES KEPLER
UNIVERSITÄT LINZ
Altenberger Straße 69
4040 Linz, Österreich
www.jku.at

EXPLAINABLE AI – TAXONOMY[1,3]



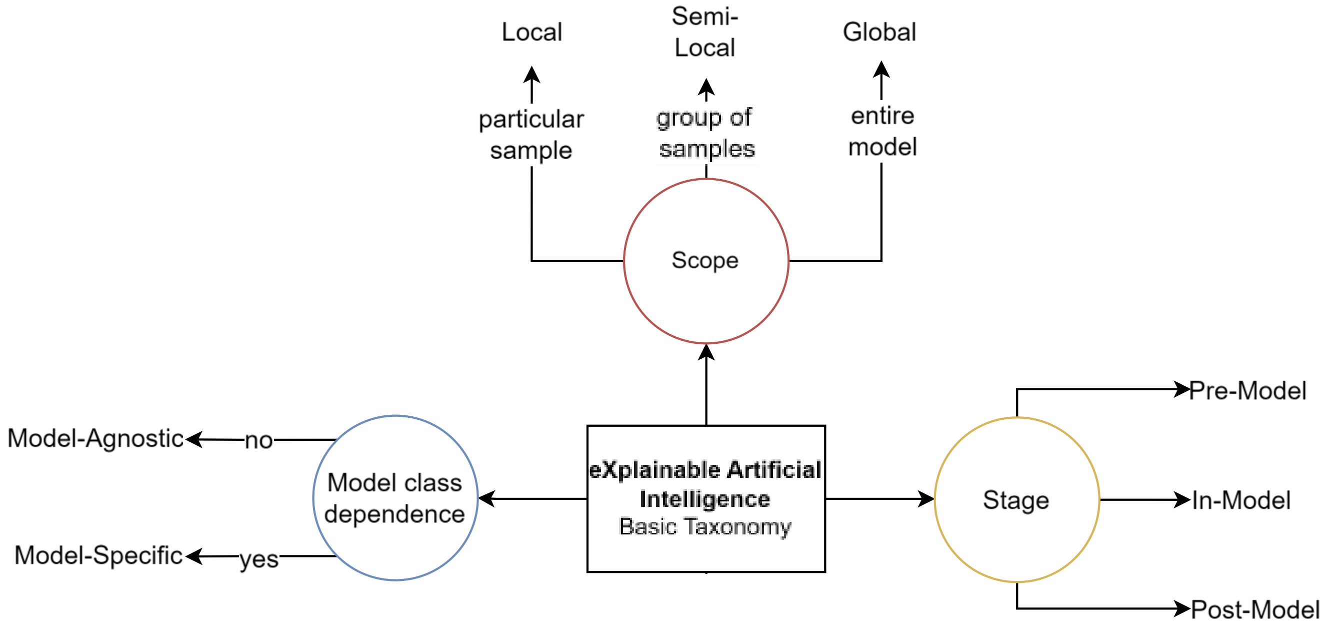
[1] *W. Ding et al., Explainability of Artificial Intelligence Methods, Applications and Challenges: A Comprehensive Survey*
[3] *A. Singh et al., Explainable Deep Learning Models in Medical Image Analysis*

EXPLAINABLE AI – TAXONOMY[1,3]



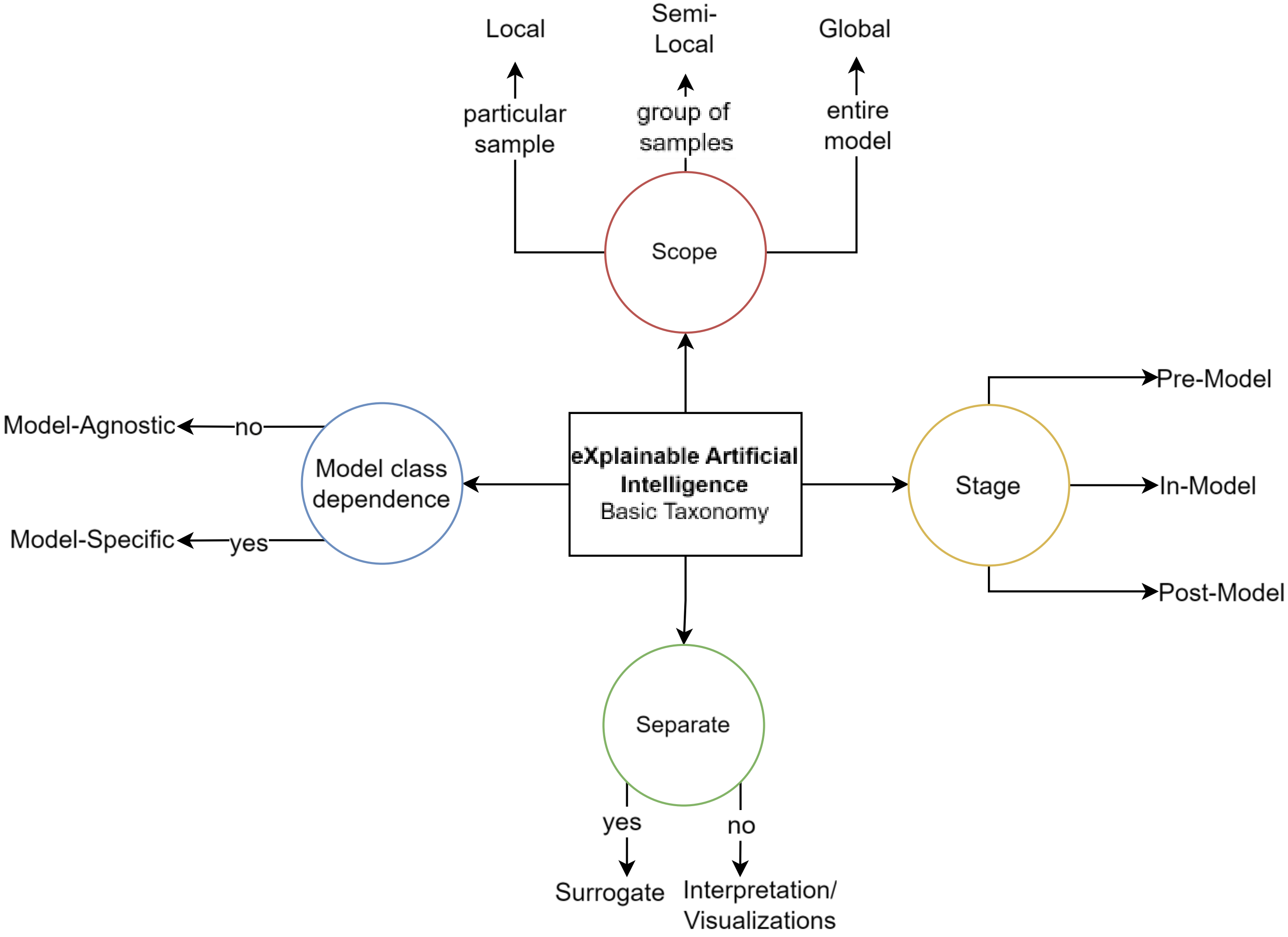
[1] W. Ding et al., Explainability of Artificial Intelligence Methods, Applications and Challenges: A Comprehensive Survey
[3] A. Singh et al., Explainable Deep Learning Models in Medical Image Analysis

EXPLAINABLE AI – TAXONOMY[1,3]



[1] *W. Ding et al., Explainability of Artificial Intelligence Methods, Applications and Challenges: A Comprehensive Survey*
[3] *A. Singh et al., Explainable Deep Learning Models in Medical Image Analysis*

EXPLAINABLE AI – TAXONOMY[1,3]



[1] W. Ding et al., Explainability of Artificial Intelligence Methods, Applications and Challenges: A Comprehensive Survey
[3] A. Singh et al., Explainable Deep Learning Models in Medical Image Analysis

FUNDAMENTALS OF FUZZY SETS CONT.

- **Type 1 fuzzy hedges [7]**
  - Linguistic operators
  - Enhance descriptive capabilities
  - “slightly, highly, very, more or less, much“

[7] L. Zadeh, A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges, Journal of Cybernetics

FUNDAMENTALS OF FUZZY SETS CONT.

- **Type 1 fuzzy hedges [7]**
  - Linguistic operators
  - Enhance descriptive capabilities
  - “slightly, highly, very, more or less, much“

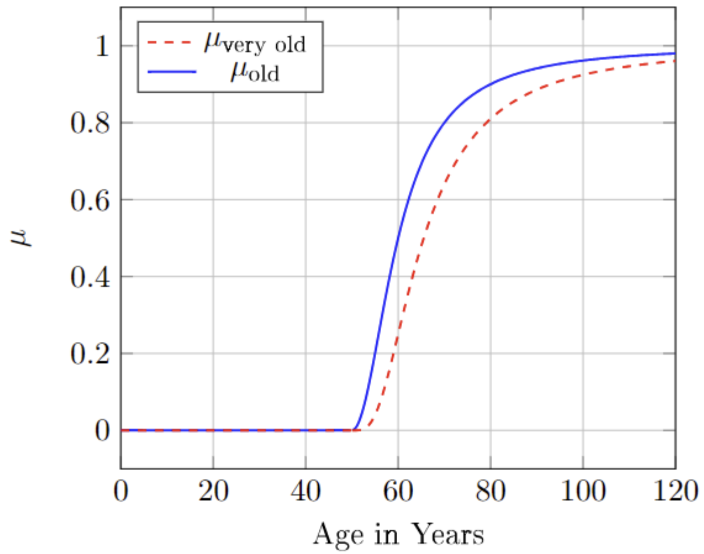


Image adapted from [7]

[7] L. Zadeh, A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges, Journal of Cybernetics

FUNDAMENTALS OF FUZZY SETS CONT.

- **Type 1 fuzzy hedges [7]**
  - Linguistic operators
  - Enhance descriptive capabilities
  - “slightly, highly, very, more or less, much“

- **Type 2 fuzzy hedges [7]**
  - Concept-related beyond MFs (out of scope)
  - “safe“ vs. “essentially safe“ – basic criteria

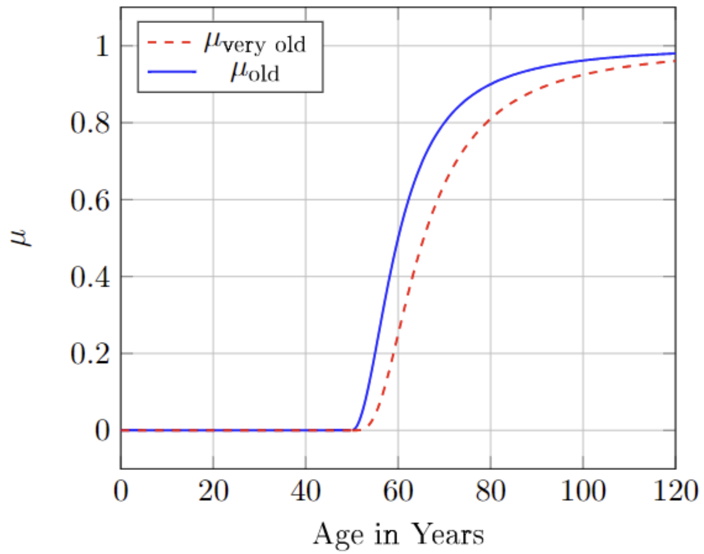


Image adapted from [7]

[7] L. Zadeh, A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges, Journal of Cybernetics

REFERENCES

[1] Ding, W., Abdel-Basset, M., Hawash, H., & Ali, A. (2022). Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. Inf. Sci., 615, 238-292. DOI: 10.1016/j.ins.2022.10.013

[2] Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). IEEE Access, 6, 52138-52160. DOI: 10.1109/ACCESS.2018.2870052

[3] Singh, A., Sengupta, S., & Lakshminarayanan, V. (2020). Explainable Deep Learning Models in Medical Image Analysis. Journal of Imaging, 6. DOI: 10.3390/jimaging6060052

[4] Goguen, J.A. (1973). Zadeh L. A.. Fuzzy sets. Information and control , vol. 8 (1965), pp. 338–353. Zadeh L. A.. Similarity relations and fuzzy orderings. Information sciences , vol. 3 (1971), pp. 177–200. Journal of Symbolic Logic, 38, 656 - 657. DOI: 10.1016/S0019-9958(65)90241-X

[5] Mamdani, E.H. (1976). Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis. IEEE Transactions on Computers, C-26, 1182-1191. DOI: 10.1109/TC.1977.1674779

REFERENCES

[6] Liang, Q., & Mendel, J. (1999). An introduction to type-2 TSK fuzzy logic systems. FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No.99CH36315), 3, 1534-1539 vol.3. DOI: 10.1109/FUZZY.1999.790132

[7] Zadeh, L.A. (1972). A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges. DOI: 10.1080/01969727208542910

[8] Wang, K. (2001). Computational Intelligence in Agile Manufacturing Engineering. In B. H. R. Bettina Buth (Ed.), Engineering and Food for the 21st Century (pp. 297-315). Elsevier. DOI: 10.1016/B978-008043567-1/50016-4

[9] Mendel, J.M., & Bonissone, P.P. (2021). Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems. IEEE Transactions on Fuzzy Systems, 29, 3579-3593. DOI: 10.1109/TFUZZ.2021.3079503

[10] Simon, D. (1996). Fuzzy sets and fuzzy logic: Theory and applications: by George J. KLIR and Bo YUAN. Prentice Hall; Upper Saddle River, NJ; 1995; xv + 574 pp. Control Engineering Practice, 4, 1332-1333. DOI: 10.1016/0967-0661(96)81492-4

REFERENCES

[11] Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics, SMC-15, 116-132. DOI: 10.1109/TSMC.1985.6313399

[12] Image: https://github.com/tenaci-hand-grip/fuzzy-visualization, 05.04.2024

[13] Cichy, C., & Rass, S. (2019), A Fuzzy-Approximation-Approach to Explainable Information Quality Assessment, International Business Information Management Association (IBIMA), pp.3919-3931

[14] Cichy, C., & Rass, S. (2020). Fuzzy Expert Systems for Automated Data Quality Assessment and Improvement Processes. In Proceedings of the CEUR Workshop (Vol. 2751, pp. 7-11)

[15] Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P. (2002). Numerical recipes in C., ISBN:978-0-521-43108-8

[16] Lehmann, E. L., & Romano, J. P. (2005). Testing statistical hypotheses (3rd ed.). Springer texts in statistics. New York: Springer. ISBN: 978-0-387-98864-1

REFERENCES

[17] Chvosteková, M., & Witkovský, V. (2009). Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model with Normal Errors. DOI: 10.2478/v10048-009-0003-9

[18] Cao, J., Zhou, T., Zhi, S., Lam, S., Ren, G., Zhang, Y., Wang, Y., Dong, Y., & Cai, J. (2024). Fuzzy Inference System with Interpretable Fuzzy Rules: Advancing Explainable Artificial Intelligence for Disease Diagnosis—A Comprehensive Review. Information Sciences. DOI: 10.1016/j.ins.2024.120212