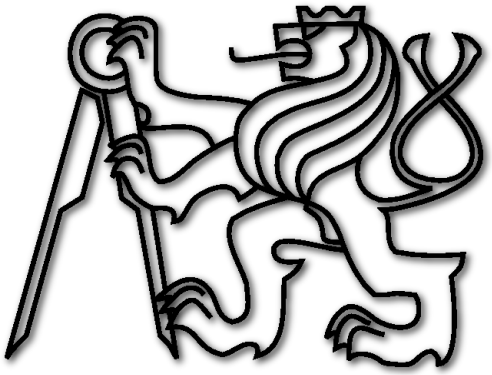
Feature Ranking a Feature Selection založené na induktivních modelech

Aleš Pilný
pilnyale (at) fel.cvut.cz



Computational Intelligence Group
Department of Computer Science and Engineering
Faculty of Electrical Engineering
Czech Technical University in Prague

7.10.2010

Obsah

- Skupina CIG
- Úvod do problematiky
  - Knowledge Discovery, Feature Ranking / Selection
- Stávající metody
  - AMIFS, ReliefF, CFS
- Metody založené na induktivních modelech
  - Úvod do GAME, FAKE-GAME
  - FeRaNGA
  - CBFR, MIFR

CIG (Computational Intelligence Group)

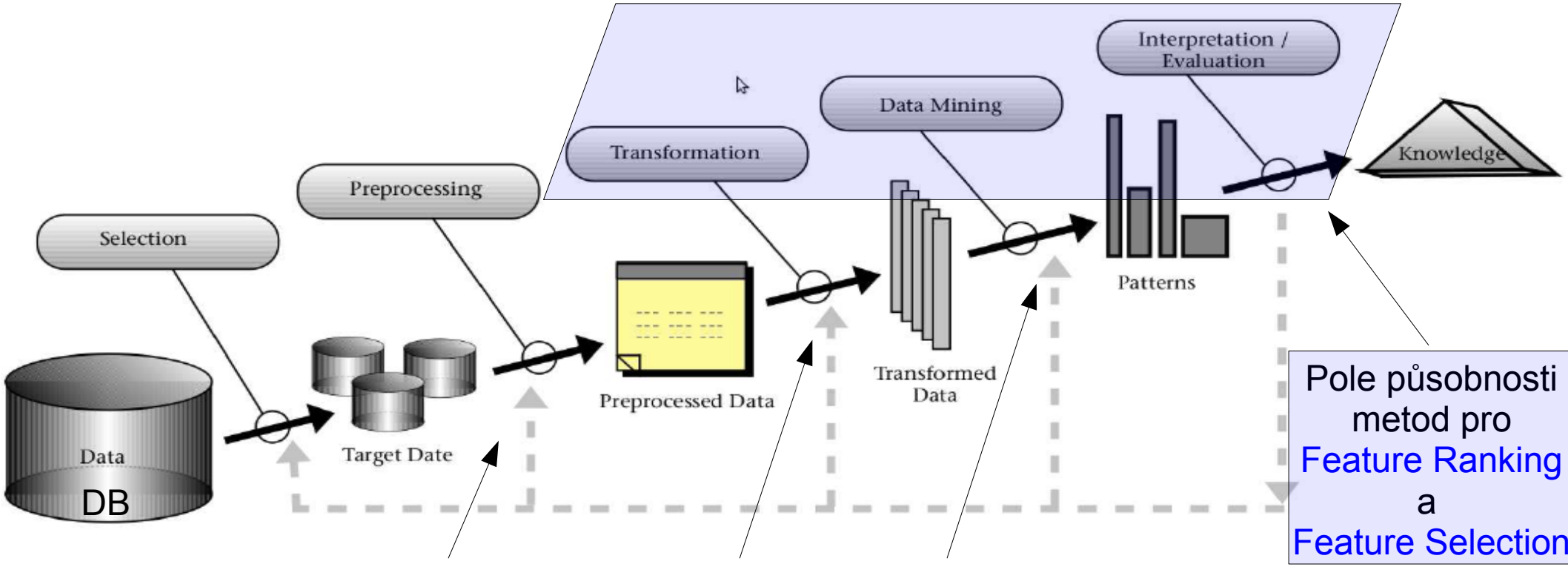
- Umělé neuronové sítě (rekurentní, THSOM, GAME,..)
- Přírodou inspirované algoritmy
  - Genetické algoritmy
  - Mravenčí a jiné optimalizace (JCOOL)
  - hybridní evoluce
- Data Mining & Knowledge Discovery (FAKE-GAME)
- Automatické předzpracování dat (GA)
- Hardwarové akcelerace vědeckých výpočtů (PS)

www.cig.felk.cvut.cz

Projekty spojené s lékařstvím, přírodovědou, HW, ...

Úvod do problematiky

- Proces získávání znalostí – Knowledge Discovery



Pole působnosti metod pro Feature Ranking a Feature Selection

Normalizace, chybějící hodnoty,...

Redukce dat (v počtu instancí i atributů), ...

Hledání vzorů v datech (klasifikace, regrese, clustering...)

Úvod do problematiky II

Redukce dat z pohledu počtu atributů:

Feature Ranking

- ohodnocování atributů

Feature Selection

- vybrání podmnožiny atributů

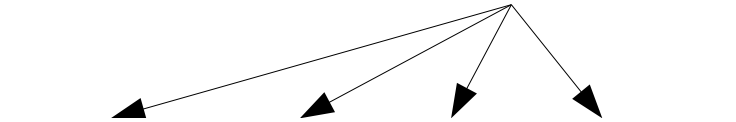
Feature Extraction

Vytváření menší množiny nových atributů z původních

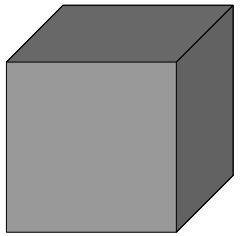
Feature = attribute = factor = … příznak, atribut, faktor, fíčura, vstup, …

Úvod do problematiky III - Feature Ranking / Selection

Jak jsou atributy významné / které?



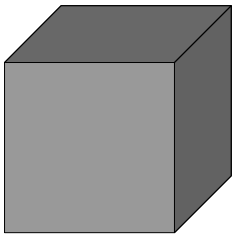
S-length	S-width	P-length	P-width
5,1	3,5	1,4	0,2
4,9	3	1,4	0,2
4,7	3,2	1,3	0,2
4,6	3,1	1,5	0,2
5	3,6	1,4	0,2
5,4	3,9	1,7	0,4
4,6	3,4	1,4	0,3
5	3,4	1,5	0,2
4,4	2,9	1,4	0,2



Feature Ranking

Pořadí (ranks)

1. **P-length (0.9)**
2. **P-width (0.8)**
3. **S-length (0.3)**
4. **S-width (0.1)**



“Feature Selection“

Redukce dimensionality

S-width	P-length	P-width
3,5	1,4	0,2
3	1,4	0,2
3,2	1,3	0,2
3,1	1,5	0,2
3,6	1,4	0,2
3,9	1,7	0,4
3,4	1,4	0,3
3,4	1,5	0,2
2,9	1,4	0,2

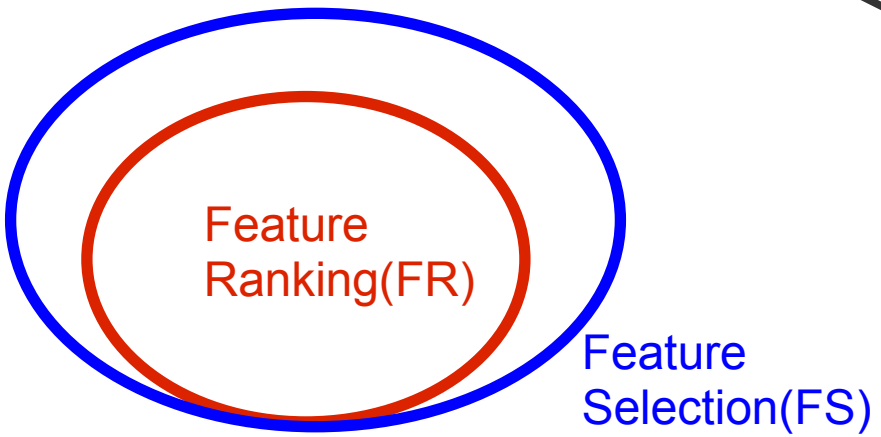
Úvod do problematiky IIIb

FS metody

Wrappers – vyhodnocení každé podmnožiny atributů napříč stav. prostorem atributů na daném modelu
- náročné, náchylné na přeučení

Filters – obdoba Wrapper metod
+ místo modelu jednodušší filtr – zohledňuje jen vlastnosti dat

Embedded – metody zabudované v alg. pro tvorbu modelu
+ selekce dle informací z vytváření modelu



Stávající metody - Feature Ranking

ReliefF metoda

- Varianty pro klasifikaci i regresi (RReliefF)
- Průběžné ohodnocování dle vzdálenosti mezi blízkými instancemi (weighting – near Hit & Miss)

AMIFS metoda (klasifikace)

- Výběr k-nejlepších (= selekce. If k = N ranking)
- Mutual Information kritérium výběru (rozdíl entropií)
  - „Vybírej postupně atributy, které mají nejvíce společné informace s výstupní třídou, ale nejméně se zbývajícími nevybranými atributy“.

Stávající metody - Feature Selection

CFS metoda - klasifikace

- Ohodnocuje podmnožiny atributů dle hodnoty <u>korelačně-heuristické funkce</u>

$$M_S = \frac{k\,\overline{r_{cf}}}{\sqrt{k + k(k-1)\,\overline{r_{ff}}}}$$

S	Podmnožina atributů
k	Počet atributů podmnožiny S
$\overline{r_{cf}}$	Průměrná korelace mezi atributem a výstupní třídou
$\overline{r_{ff}}$	Průměrná korelace mezi atributy

Čitatel – predikovatelnost třídy podmnožinou atributů

Jmenovatel – hodnota redundance mezi atributy

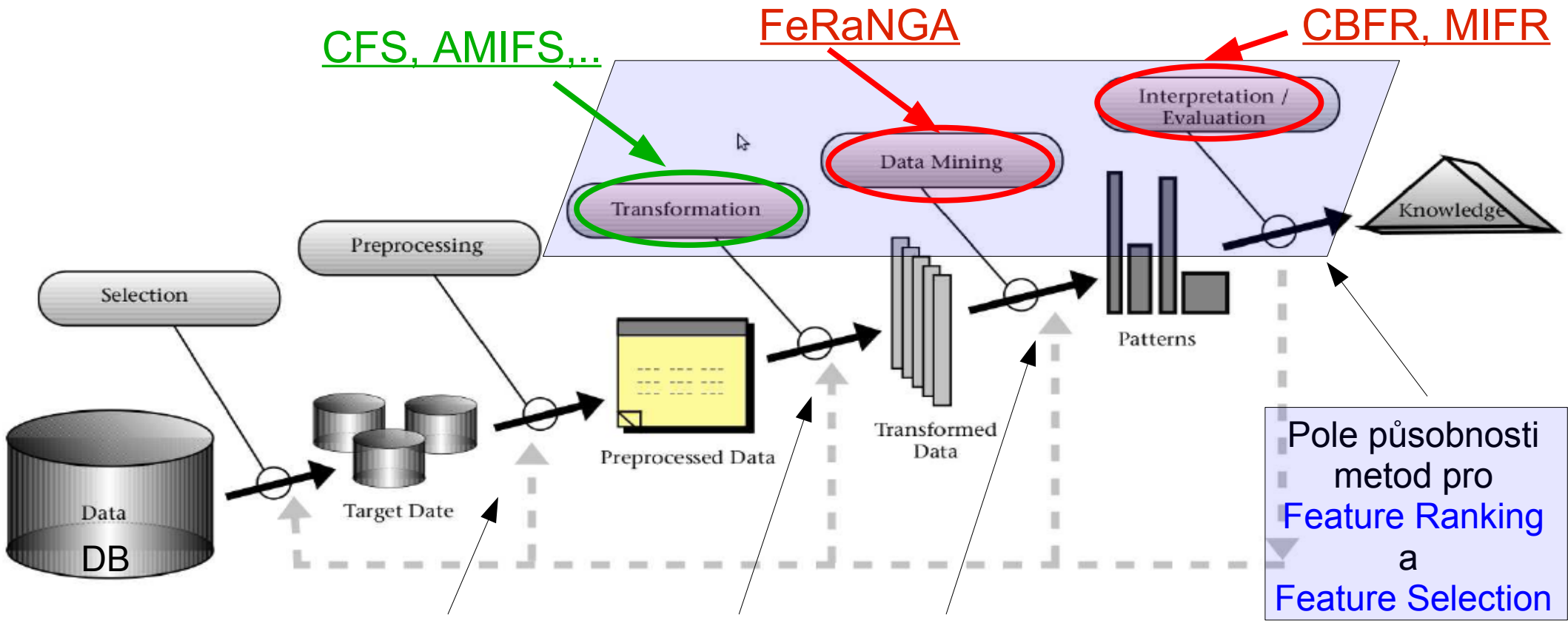
Výsledkem je S maximalizující M_S

Metody založené na induktivních modelech I

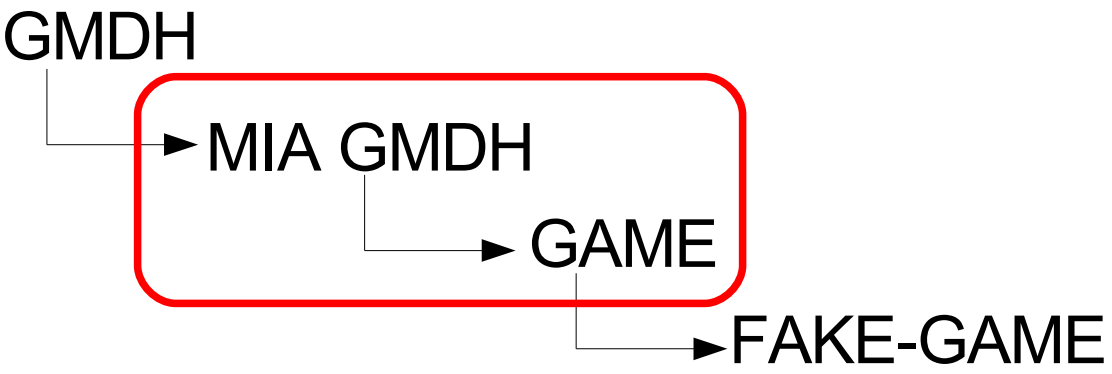
- FeRaNGA
  - Ohodnocování atributů během stavby sítě GAME
- CBFR (Correlation Based Feature Ranking methods)
- MIFR (Mutual Information Based Feature Ranking methods)

CBFR a MIFR - ohodnocování atributů na hotové síti GAME dle zpracování vzájemné korelace či Mutual Information sousedních neuronů

Přeheld


CFS, AMIFS,..
FeRaNGA
CBFR, MIFR
Interpretation / Evaluation
Data Mining
Transformation
Knowledge
Preprocessing
Selection
Patterns
Transformed Data
Preprocessed Data
Target Date
Data
DB
Pole působnosti metod pro Feature Ranking a Feature Selection
Normalizace, chybějící hodnoty,...
Redukce dat (v počtu instancí i atributů), ...
Hledání vzorů v datech (klasifikace, regrese, clustering...)

Metody založené na induktivních modelech II



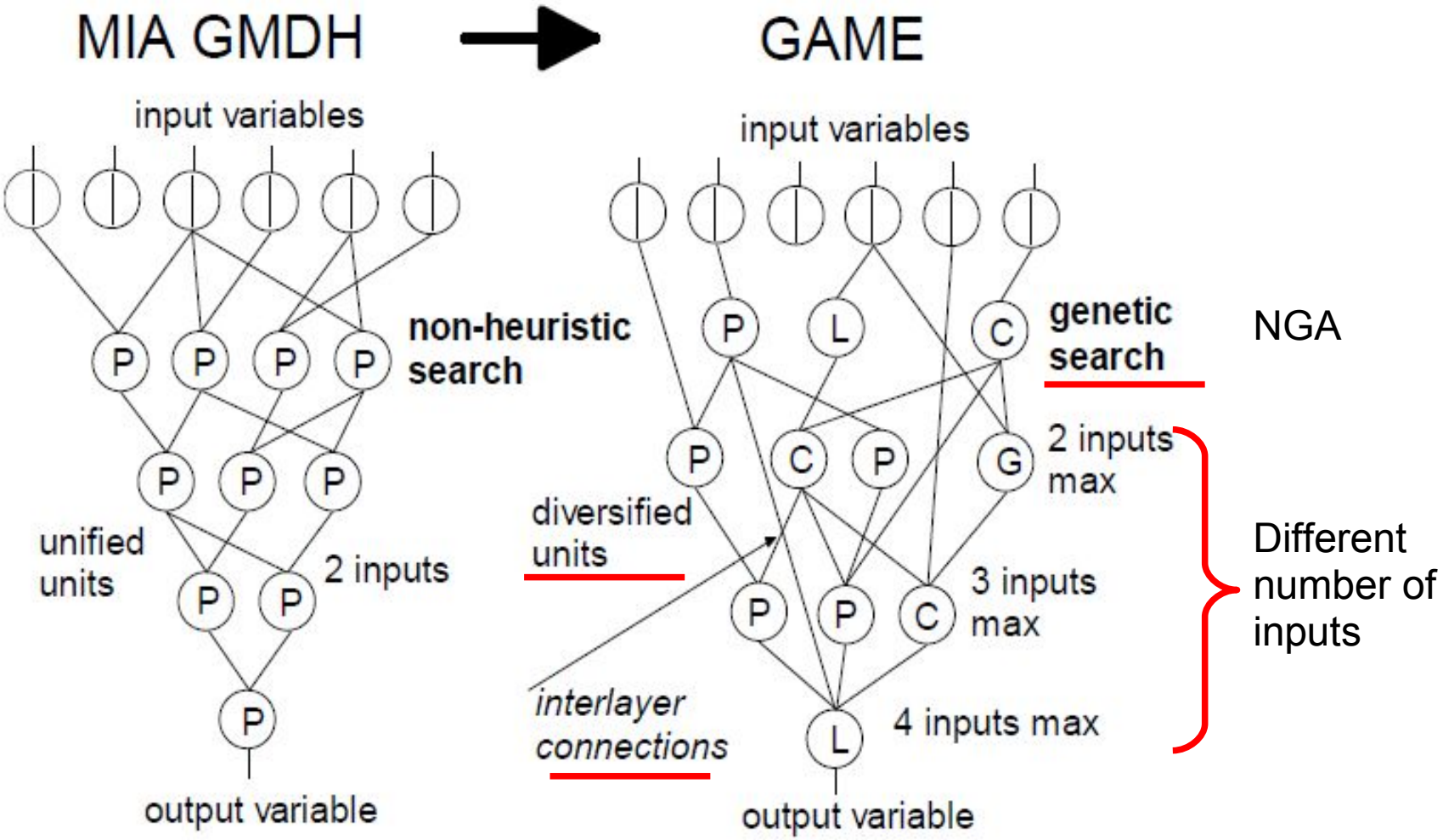
GMDH … Group Method of Data Handling

MIA GMDH … Multilayered Iterative Algortithm GMDH

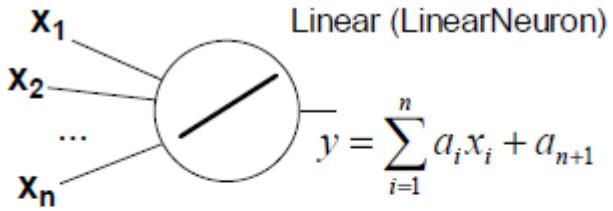
GAME … Group of Adaptive Methods Evolution

FAKE-GAME … Fully Automated Knowledge Extraction - GAME

Metody založené na induktivních modelech III

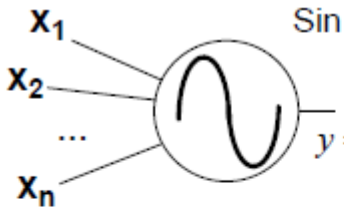


Metody založené na induktivních modelech III



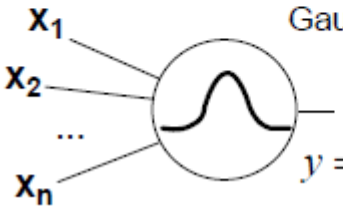
Linear (LinearNeuron)

$$y = \sum_{i=1}^{n} a_i x_i + a_{n+1}$$



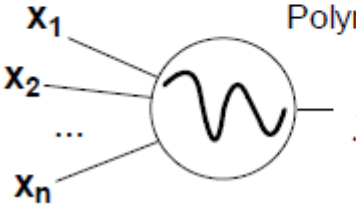
Sin (SinusNeuron)

$$y = a_{n+1} * \sin\left[a_{n+2} * \left(\sum_{i=1}^{n} a_i x_i + a_{n+3} \right) \right] + a_0$$



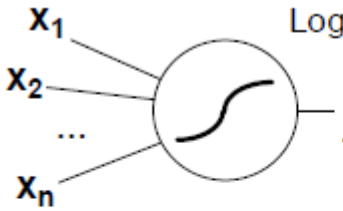
Gaussian (GaussianNeuron)

$$y = (1 + a_{n+1}) * e^{-\frac{\sum_{i=1}^{n} (x_i - a_i)^2}{(1+a_{n+2})^2}} + a_0$$



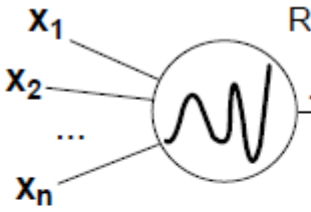
Polynomial (CombiNeuron)

$$y = \sum_{i=1}^{n} \left(a_i \prod_{j=1}^{m} x_j^r \right) + a_0$$



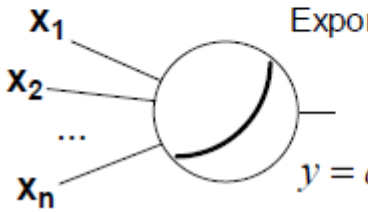
Logistic (SigmNeuron)

$$y = \frac{1}{1 + e^{-\sum_{i=1}^{n} a_i x_i}} + a_0$$



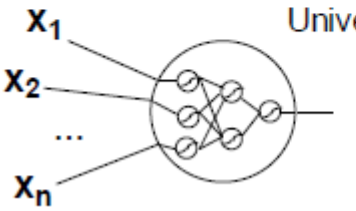
Rational (PolyFractNeuron)

$$y = \frac{a_{n^2+2}}{\sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{n*i+j} x_i x_j + a_{n^2+1}} + a_0$$



Exponential (ExpNeuron)

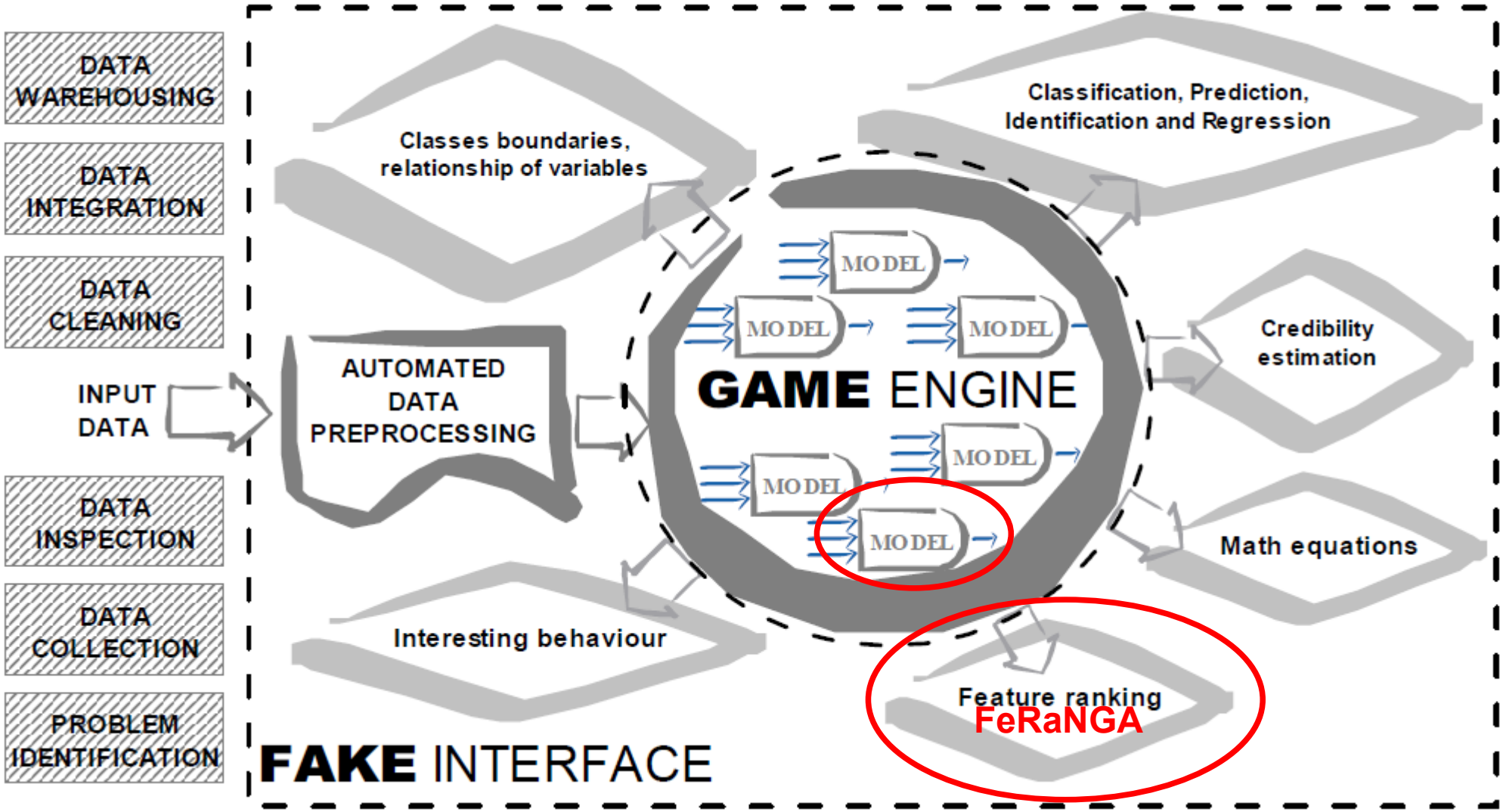
$$y = a_{n+2} * e^{a_{n+1} * \sum_{i=1}^{n} a_i x_i} + a_0$$



Universal (BPNetwork)

$$y = \sum_{q=1}^{2n+1} \psi_q \left(\sum_{p=1}^{n} \phi_{pq}(x_p) \right)$$

Metody založené na induktivních modelech IV



FeRaNGA, FeRaNGA-n I

- FeRaNGA - Klasifikace i regrese
  - Během stavby sítě GAME (pomocí NGA) aktualizujeme váhy (významnosti) atributů dle jejich procentuálního využití
    - využití - z monitorování genů v populaci
  - Při stavbě se model snaží nevyužít redundantní a irelevantní atributy – učení s učitelem (trénovací, testovací, validační množiny dat)

    => Feature Selection atributů jako vedlejší efekt stavby modelu (sítě GAME)

Nejlepší výsledky z první vrstvy sítě GAME - paradoxně kvůli NGA

FeRaNGA, FeRaNGA-n II

Experiment s umělými daty:

Uniform Hypercube Data Set

50 atributů, 2 třídy bodů ze dvou různých 10 D normálních gaussovských rozložení

Od 1 do 10 klesající relevance, 11-20 irrelevantní, 21-50 redundantní

FeRaNGA, FeRaNGA-n III

FR na Hypercube Data setu z Weky a FeRaNGA metody

Method	1	2	3	4	5	6	7	8	9	10	Redundant(21-50) and Irrelevant(11-20)																						
ChiSquare	1	2	3	4	5	6	7	8	9	10	22	37	48	35	39	43	47	26	45	27	30	49	29	32	28	33	46	31	42	17	18	14	
GainRatio	1	2	3	4	5	6	7	8	9	10	23	38	27	35	37	30	34	39	28	25	22	26	50	32	36	21	31	46	33	17	18	14	
InfoGain	1	2	3	4	5	6	7	8	9	10	22	37	48	43	35	39	26	47	44	30	27	49	36	29	33	28	23	46	42	17	18	14	
OneR	1	2	3	4	5	6	7	8	9	10	48	37	43	50	40	30	25	44	39	45	22	49	38	34	36	33	23	28	24	18	17	14	
ReliefF	1	2	3	4	8	7	9	10	5	6	29	48	47	25	31	30	49	26	27	28	21	50	24	43	45	42	23	20	44	15	41	17	
SVM	1	2	48	3	29	43	4	9	35	7	36	45	41	44	34	25	8	11	12	30	37	19	15	16	6	38	24	20	5	13	40	28	
SU	1	2	3	4	5	6	7	8	9	10	23	38	35	37	27	48	39	34	26	25	50	45	49	29	32	21	31	46	24	17	18	14	
GAME 1	1	2	4	3	5	27	49	45	6	47	32	7	8	33	35	38	24	43	13	16	21	23	37	14	28	40	9	12	18	20	29	50	
GAME 2	1	2	3	4	44	38	5	48	25	7	22	40	32	19	47	50	20	46	31	35	43	18	37	8	9	10	12	14	16	21	30	49	
GAME 3	1	2	3	5	7	47	37	4	48	6	8	26	30	43	45	13	14	18	42	9	10	11	12	17	19	20	23	25	29	32	38	50	
GAME 4	1	2	3	33	4	30	41	7	44	29	46	47	8	5	9	25	34	36	48	10	35	40	6	13	14	16	18	20	22	24	31	50	
GAME 5	1	6	3	2	27	8	50	4	5	48	7	22	25	28	39	47	9	16	29	30	31	11	13	40	46	49	10	15	20	26	35	45	

Defaultní konfigurace GAME i Weka

Nevybrané atributy

- FeRaNGA má horší výsleky
- Nestabilní

FeRaNGA, FeRaNGA-n IV

Značný vliv konfigurace NGA na výsledek lze ovlivnit nastavením parametrů GA (epochy, individuálové,..)

⇒ nestabilita stále přetrvá

Nastavení NGA konfigurace můžeme doplnit výpočtem ranků/významností z ensemble n GAME modelů jako MEDIÁNU z hodnot významností jednotlivých atributů z těchto modelů.

= FeFaNGA-n

FeRaNGA-n

- Uniform Hypercube data set, FeRaNGA-5
  - NGA: # epoch = 150 a počáteční populace = 150 jedinců
- UNC = # unikátních chromozomů, od nejlepších po všechny

UNC	1	2	3	4	5	6	7	8	9	10
2	1	2	3	4	5	6	7	8	9	10
3	1	2	3	4	5	6	7	8	9	10
1 / 4	1	2	3	4	5	6	7	8	9	10
1 / 3	1	2	3	4	5	6	7	8	9	10
1 / 2	1	2	3	4	5	6	7	8	9	10
2 / 3	1	2	3	4	5	6	7	8	9	10
All	1	2	3	4	5	6	7	8	9	10

- S rostoucím UNC klesá trend restrikce metody FeRaNGA v počtu selektovaných atributů
- Všechny ranky jsou korektní (díky FeRaNGA-5)

Influence of NGA configuration on FeRaNGA-7 results

Gaussian multivariate data set

- two clusters of points generated from two different 10th-dimensional normal Gaussian distributions
- 1-10 are equally relevant, 11-20 are irrelevant, 21-50 are highly redundant with the first ten features

Influence of NGA configuration on FeRaNGA-7 results (on Gaussian Data Set)

Default configuration of NGA: 30 individuals and 15 epochs

Layer	Important part of ranks of features																				
0	9	1	2	3	4	5	6	7	10	0	0	0	0	0	0	0	0	0	0	0	0
1	9	10	5	6	7	4	1	2	3	0	0	0	0	0	0	0	0	0	0	0	0
2	9	7	3	5	10	4	6	2	1	31	8	28	38	46	20	37	26	29	41	12	39
Overall	9	5	6	7	10	3	4	1	2	0	0	0	0	0	0	0	0	0	0	0	0

Ranks computed as a medians over all layers of medians.

Influence of NGA configuration on FeRaNGA-7 results (on Gaussian Data Set)

Default configuration of NGA: 30 individuals and 15 epochs

Layer	Important part of ranks of features																			
0	9	1	2	3	4	5	6	7	10	0										
1	9	10	5	6	7	4	1	2	3	0										
2	9	7	3	5	10	4	6	2	1	31	8	28	38	46	20	37	26	29	41	12	39
Overall	9	5	6	7	10	3	4	1	2	0	0	0	0	0	0	0	0	0	0	0	0

9 correct ranks in first two layers

Redundant features

Incorrect features!

Ranks computed as a medians over all layers of medians

Influence of NGA configuration on FeRaNGA-7 results (on Gaussian Data Set)

Default configuration of NGA: 30 individuals and 15 epochs

Layer	Important part of ranks of features																				
0	9	1	2	3	4	5	6	7	10	0											
1	9	10	5	6	7	4	1	2	3	0											
2	9	7	3	5	10	4	6	2	1	31	8	28	38	46	20	37	26	29	41	12	39
Overall	9	5	6	7	10	3	4	1	2	0	0	0	0	0	0	0	0	0	0	0	0

9 correct ranks in first two layers

Redundant features

Incorrect features!

Configuration of NGA: 75 individuals and 75 epochs

Layer	Important part of ranks of features																				
0	9	6	1	10	3	7	2	5	8	4	0	0	0	0							
1	9	1	2	6	10	7	5	3	8	4	31	45	24	25	29	30	33	40	41	49	0
2	3	2	6	9	5	10	8	45	28	33	1	7	47	26	49	50	13	4	37	41	25
Overall	9	6	2	1	10	3	7	5	8	4	45	25	33	41	49	24	31	30	0	0	0

10 correct ranks

Influence of NGA configuration on FeRaNGA-7 results (on Gaussian Data Set)

Default configuration of NGA: 30 individuals and 15 epochs

Layer	Important part of ranks of features																				
0	9	1	2	3	4	5	6	7	10	0											
1	9	10	5	6	7	4	1	2	3	0											
2	9	7	3	5	10	4	6	2	1	31	8	28	38	46	20	37	26	29	41	12	39
Overall	9	5	6	7	10	3	4	1	2	0	0	0	0	0	0	0	0	0	0	0	0

9 correct ranks in first two layers

Redundant features

Incorrect features!

Configuration of NGA: 75 individuals and 75 epochs

Layer	Important part of ranks of features																				
0	9	6	1	10	3	7	2	5	8	4	0	0	0	0							
1	9	1	2	6	10	7	5	3	8	4	31	45	24	25	29	30	33	40	41	49	0
2	3	2	6	9	5	10	8	45	28	33	1	7	47	26	49	50	13	4	37	41	25
Overall	9	6	2	1	10	3	7	5	8	4	45	25	33	41	49	24	31	30	0	0	0

10 correct ranks

Configuration of NGA: 150 individuals and 150 epochs

Layer	Important part of ranks of features																
0	7	10	2	9	6	3	8	5	1	4	0	0	0	0	0	0	0
1	2	7	10	8	6	3	9	5	1	4	25	22	28	37	44	45	0
Overall	7	2	10	9	6	3	8	5	1	4	25	22	28	37	44	45	0

All features have correct ranks.

Dependency of accuracy on Nr. of models for FeRaNGA-n method (on Hypercube Data set)

First ten ranks from first layers of FeRaNGA-7 on the Hypercube Data Set.

Model	NGA configuration																			
	Default										75									
1	1	2	4	34	6	8	26	39	44	0	1	4	2	5	3	6	7	10	8	9
2	1	4	2	26	34	3	5	6	8	10	1	2	4	3	5	6	7	8	10	9
3	1	4	3	5	6	26	44	0	0	0	1	2	4	3	5	6	7	8	0	0
4	1	2	4	3	5	6	44	8	26	0	1	2	4	3	5	6	7	8	9	0
5	1	2	3	4	5	6	8	44	0	0	1	2	3	4	5	6	7	8	0	0
6	1	4	2	3	5	6	44	8	37	0	1	2	3	4	5	6	7	8	9	0
7	1	4	2	3	5	6	8	44	0	0	1	2	3	4	5	6	7	8	9	0

- Ranks computed from a higher Nr. of models depend on significance of features from previous models.

Dependency of accuracy on Nr. of models for FeRaNGA-n method (on Hypercube Data set)

First ten ranks from first layers of FeRaNGA-7 on the Hypercube Data Set.

Model	NGA configuration																			
	Default										75									
1	1	2	4	34	6	8	26	39	44	0	1	4	2	5	3	6	7	10	8	9
2	1	4	2	26	34	3	5	6	8	10	1	2	4	3	5	6	7	8	10	9
3	1	4	3	5	6	26	44	0	0	0	1	2	4	3	5	6	7	8	0	0
4	1	2	4	3	5	6	44	8	26	0	1	2	4	3	5	6	7	8	9	0
5	1	2	3	4	5	6	8	44	0	0	1	2	3	4	5	6	7	8	0	0
6	1	4	2	3	5	6	44	8	37	0	1	2	3	4	5	6	7	8	9	0
7	1	4	2	3	5	6	8	44	0	0	1	2	3	4	5	6	7	8	9	0

Dependency of accuracy on Nr. of models for FeRaNGA-n method (on Hypercube Data set)

First ten ranks from first layers of FeRaNGA-7 on the Hypercube Data Set.

Model	NGA configuration																			
	Default										75									
1	1	2	4	34	6	8	26	39	44	0	1	4	2	5	3	6	7	10	8	9
2	1	4	2	26	34	3	5	6	8	10	1	2	4	3	5	6	7	8	10	9
3	1	4	3	5	6	26	44	0	0	0	1	2	4	3	5	6	7	8	0	0
4	1	2	4	3	5	6	44	8	26	0	1	2	4	3	5	6	7	8	9	0
5	1	2	3	4	5	6	8	44	0	0	1	2	3	4	5	6	7	8	0	0
6	1	4	2	3	5	6	44	8	37	0	1	2	3	4	5	6	7	8	9	0
7	1	4	2	3	5	6	8	44	0	0	1	2	3	4	5	6	7	8	9	0

- For NGA configuration 75 are correct ranks from 5, 6 and 7 models.

Dependency of accuracy on Nr. of models for FeRaNGA-n method (on Hypercube Data set)

First ten ranks from first layers of FeRaNGA-7 on the Hypercube Data Set.

Model	NGA configuration																			
	Default										75									
1	1	2	4	34	6	8	26	39	44	0	1	4	2	5	3	6	7	10	8	9
2	1	4	2	26	34	3	5	6	8	10	1	2	4	3	5	6	7	8	10	9
3	1	4	3	5	6	26	44	0	0	0	1	2	4	3	5	6	7	8	0	0
4	1	2	4	3	5	6	44	8	26	0	1	2	4	3	5	6	7	8	9	0
5	1	2	3	4	5	6	8	44	0	0	1	2	3	4	5	6	7	8	0	0
6	1	4	2	3	5	6	44	8	37	0	1	2	3	4	5	6	7	8	9	0
7	1	4	2	3	5	6	8	44	0	0	1	2	3	4	5	6	7	8	9	0

- For NGA configuration 75 are correct ranks from 5, 6 and 7 models.
- Growing Nr. of models and stronger NGA config. cause improving of accuracy.

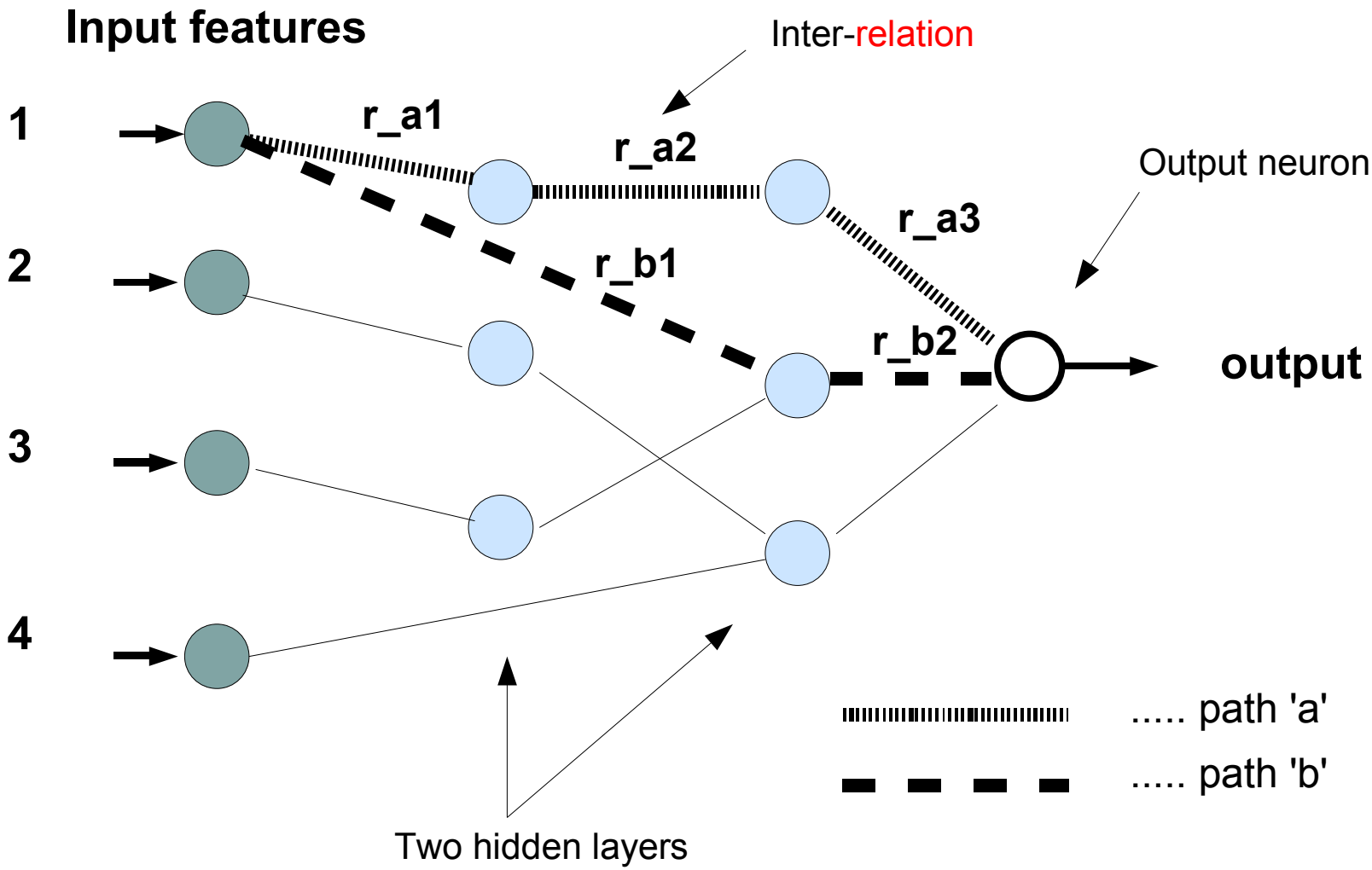
Dependency of accuracy on Nr. of models for FeRaNGA-n method (on Hypercube Data set)

First ten ranks from first layers of FeRaNGA-7 on the Hypercube Data Set.

Model	NGA configuration																			
	Default										75									
1	1	2	4	34	6	8	26	39	44	0	1	4	2	5	3	6	7	10	8	9
2	1	4	2	26	34	3	5	6	8	10	1	2	4	3	5	6	7	8	10	9
3	1	4	3	5	6	26	44	0	0	0	1	2	4	3	5	6	7	8	0	0
4	1	2	4	3	5	6	44	8	26	0	1	2	4	3	5	6	7	8	9	0
5	1	2	3	4	5	6	8	44	0	0	1	2	3	4	5	6	7	8	0	0
6	1	4	2	3	5	6	44	8	37	0	1	2	3	4	5	6	7	8	9	0
7	1	4	2	3	5	6	8	44	0	0	1	2	3	4	5	6	7	8	9	0

- For NGA configuration 75 are correct ranks from 5, 6 and 7 models.
- Growing Nr. of models and stronger NGA config. cause improving of accuracy.
- With NGA config. 150 are all ranks of features correct.

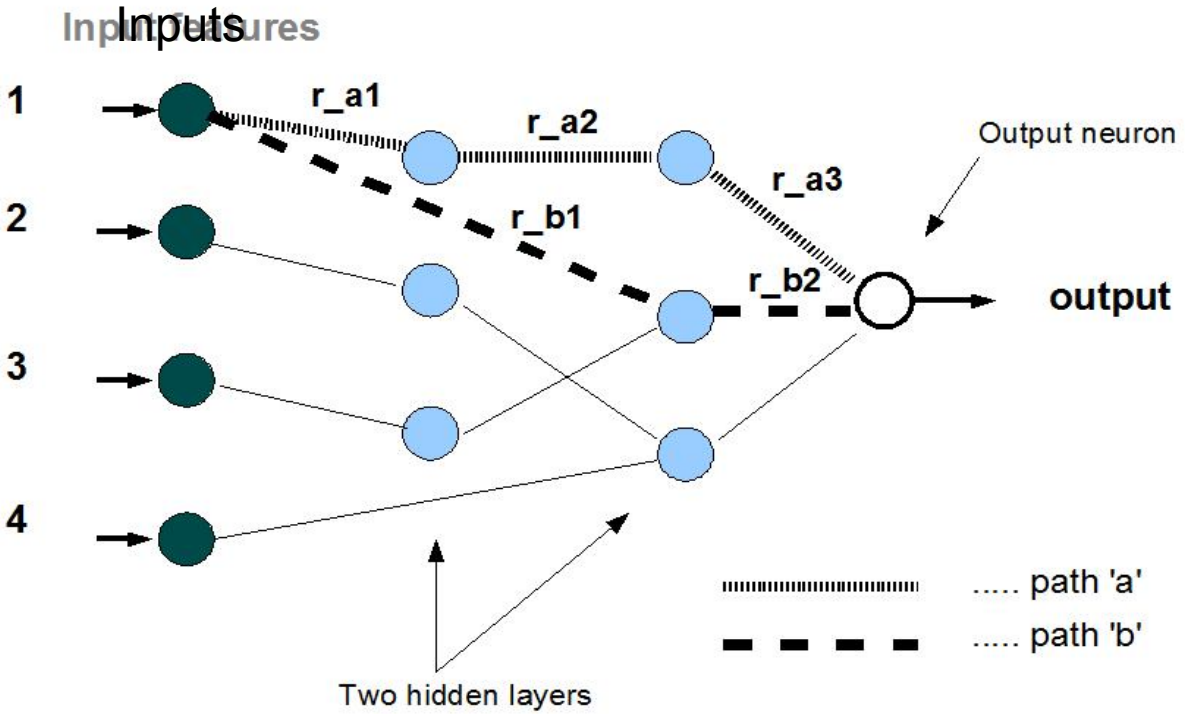
CBFR a MIFR


Input features
Inter-relation
1
2
3
4
r_a1
r_a2
r_a3
r_b1
r_b2
Output neuron
output
..... path 'a'
..... path 'b'
Two hidden layers

CBFR a MIFR metody

- Kalkul
- Fuzzy logika
- Certainty Factors

- Klasifikace i regrese
- Ohodnocení atributů dle zpracování vzájemného vztahu výstupů sousedních jednotek



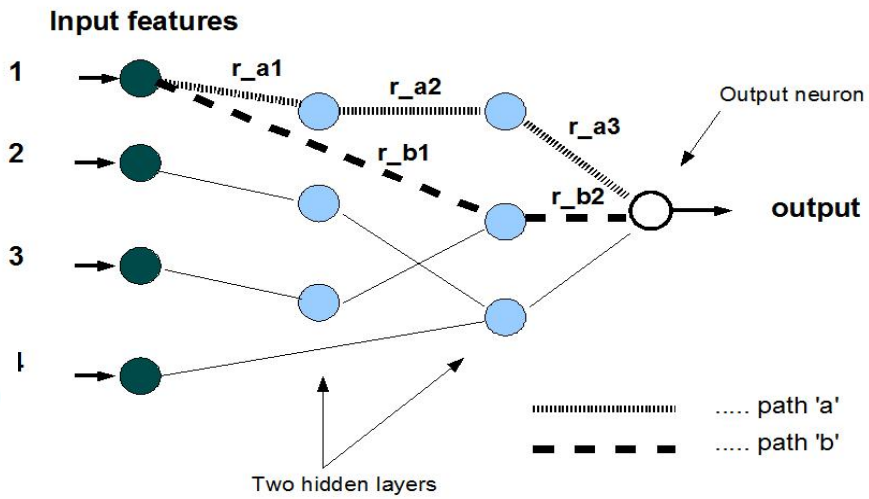
Nejůspěšnější přístupy:
FL-FR metoda (fuzzy logic)
a
CCF-FR metoda
(combine certainty factors)

CBFR metody

Zpracovává KORELACI:

- FL-FR ~ Fuzzy Logic (sets)
  - std. sjednocení / průnik

$$S_i = \max(\min(r_{11}, ..., r_{1K_1}), ..., \min(r_{N1}, ..., r_{NK_N}))$$



- CCF-FR ~ Combine Certainty Factors
  - $CF_{cobmi}(r_{Nj}, r_{Nj+1}) = r_{Nj} + (1 - r_{Nj}) * r_{Nj+1}$

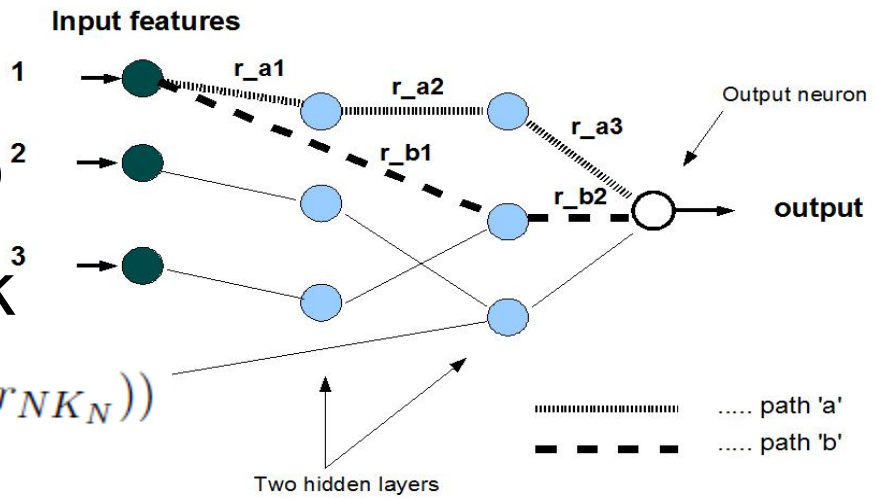
$$S_i = \max(CF_{combi1}, ..., CF_{combiN})$$

S_i … významnost i-tého atributu
r_{NK_N} … korelace výstupu neuronů na N-té cestě

MIFR metody

Zpracovává MI:

- MI-FL-FR ~ Fuzzy Logic (sets)
  - std. sjednocení / průnik



$$S_i = \max(\min(r_{11}, ..., r_{1K_1}), ..., \min(r_{N1}, ..., r_{NK_N}))$$

- MI-CCF-FR ~ Combine Certainty Factors
  - $CF_{cobmi}(r_{Nj}, r_{Nj+1}) = r_{Nj} + (1 - r_{Nj}) * r_{Nj+1}$

$$S_i = \max(CF_{combi1}, ..., CF_{combiN})$$

S_i … významnost i-tého atributu

r_{NK_N} … MI výstupu neuronů na N-té cestě

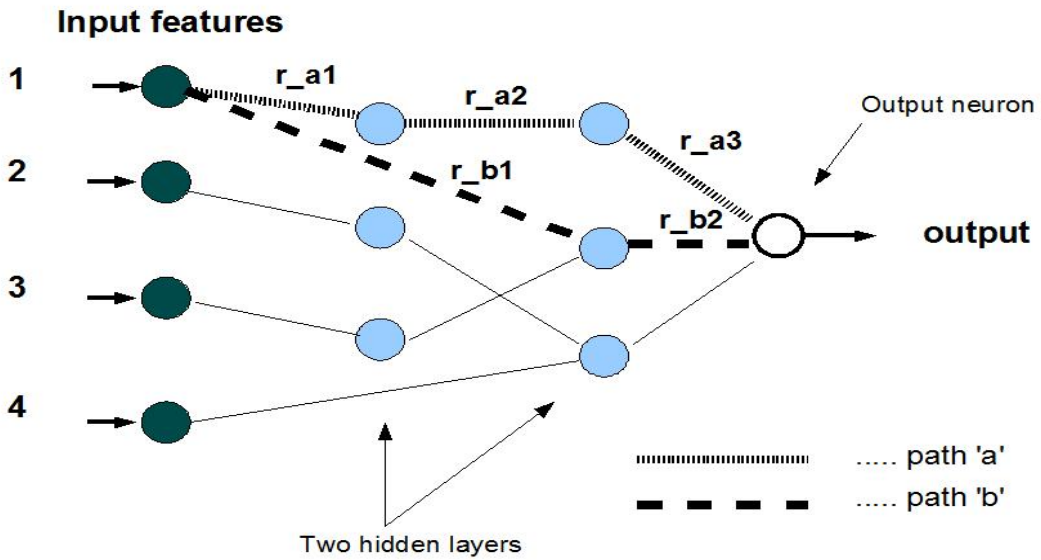
MI-CCF-FR method

IF E is observed THEN H is true (with cetainty factor, CF = n)
E … evidence, H … hypothesis

CF is a number from 0.0 to 1.0

CCF… Combine CF

CF's combined along the paths



$$CF_{cobmi}(r_{Nj}, r_{Nj+1}) = r_{Nj} + (1 - r_{Nj}) * r_{Nj+1}$$

Rank assigned to the maximal conclusion

$$S_i = \max(CF_{combi1}, \ldots, CF_{combiN})$$

CB-FR and MI-FR Experiments

Housing real-word data set

- **Regression problem / data from ML UCI repository (*http://www.ics.uci.edu/*)**

Comparison of proposed on RMS error among our proposed methods and ICA-FX method

- ICA-FX - N. Kwak, C. Kim, and H. Kim. Dimensionality reduction based on ica for regression problems. Neurocomputing, 71:2596 2603, 2008

Comparison on Housing real-word data set

FL-FR and CCF-FR – correlation based methods

MI-FL-FR and MI-CCF-FR – MI based methods

ICA-FX ... averages of five regresion methods (MLP,SVM, 1-NN, 3-NN and 5-NN).

Better results than the rest of methods

Average RMS error from ten runs

Average of standard deviations from ten experiments

The best preformance of five regression methods or of ten runs of our proposed methods

method\ # of att.	2	3	5	7	8	9	11
FL-FR	3.78 (0.08) 3.65	3.93 (0.41) 3.64	3.15 (0.23) 2.91	3.9 (0.32) 3.55	3.75 (0.2) 3.45	-	
MI-FL-FR	2.87 (0.10) 2.75	3.54 (0.12) 3.44	3.15 (0.09) 3.03	3.83 (0.16) 3.64	3.55 (0.11) 3.35	3.96 (0.25) 3.66	3.24 (0.17) 3.04
CCF-FR	5.79 (0.05) 5.71	3.98 (0.08) 3.9	4.08 (0.41) 3.79	3.48 (0.28) 3.2	4.51 (0.52) 3.761	-	
MI-CCF-FR	3.84 (0,08) 3.71	3.63 (0.12) 3.52	3.48 (0.08) 3.35	3.01 (0.09) 2.85	3.25 (0.1) 3.11	3.22 (0.09) 3.08	3.24 (0.17) 3.04
ICA-FX	-	4.09 (0.53) 3.35 (MLP)	3.74 (0.51) 3.43 (5-NN)	3.37 (0.55) 3.25 (3-NN)	-	3.48 (0.63) 3.20 (MLP)	3.61 (0.72) 3.27 (SVM)

Závěr

FR a FS metody založené na

- realtime využití stavby induktivního modelu
  - FeRaNGA-n
- post-processingu struktury hotového modelu
  - CB-FR metody využívající korelaci výstupu neuronů
    - FL-FR
    - CCF-FR
  - MI-FR metody využívající MI výstupu neuronů
    - MI-FL-FR
    - MI-CCF-FR

Dotazy

Děkuji za pozornost

pilnyale@fel.cvut.cz

Dotazy?