Evidence-Based Reasoning in Classifier Hierarchies

Rebecca Fay, Friedhelm Schwenker, Günther Palm

Department of Neural Information Processing
University of Ulm

Overview

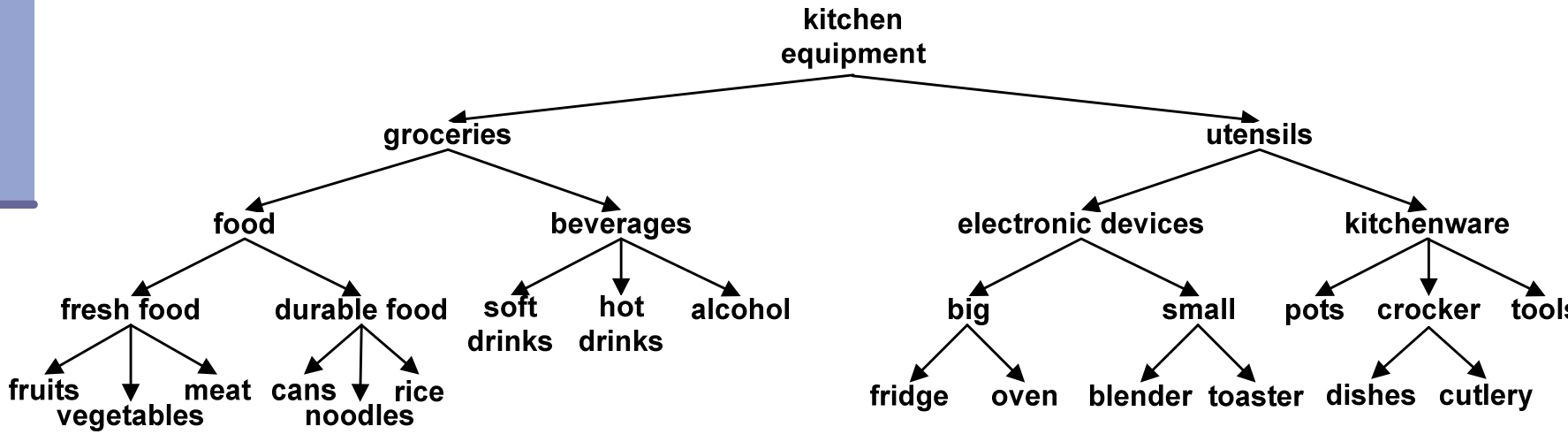
- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

Overview

- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

Motivation

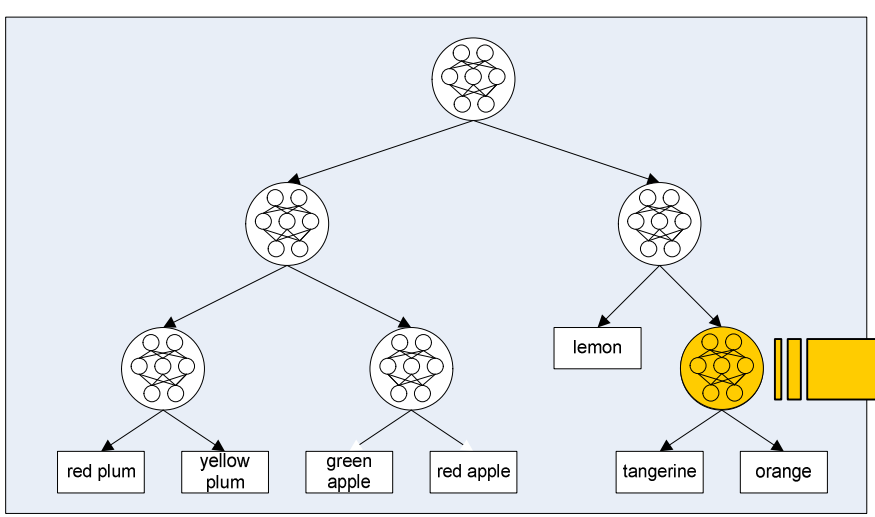
- Object recognition problems are multi-class problems
  - Number of classes greater than 2
- Object recognition problems often show a hierarchical structure

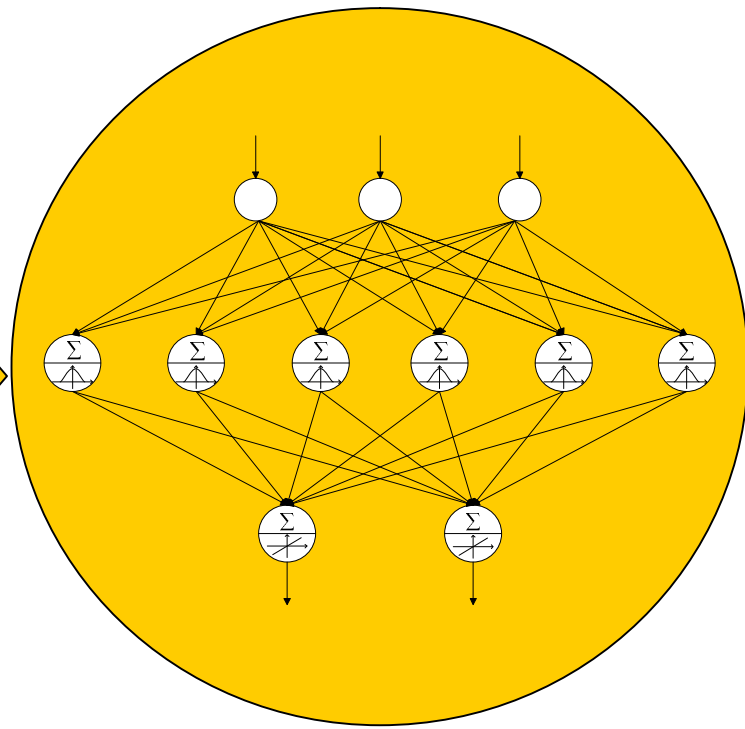


Overview

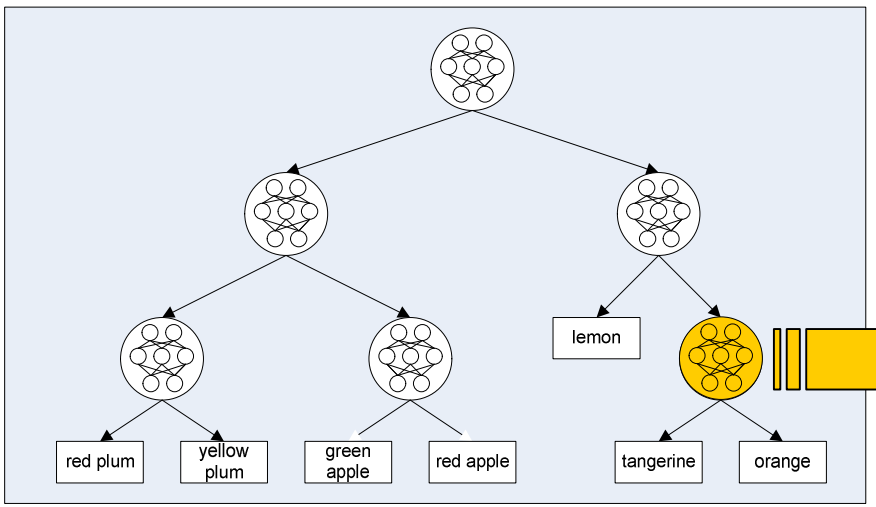
- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

Hierarchical Neural Network Classifiers





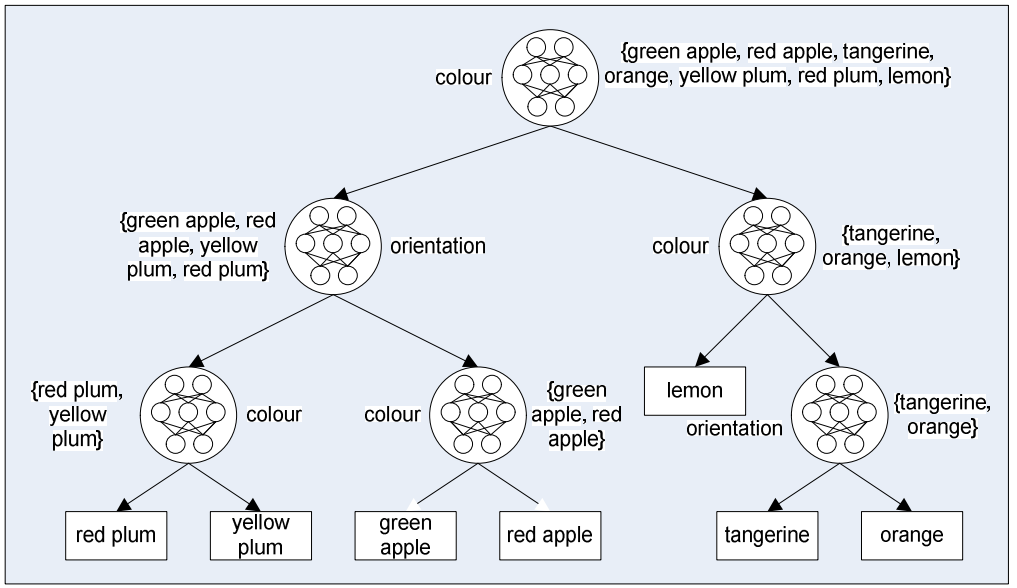
Hierarchical Neural Network Classifiers



Classifiers

- RBF
- k-NN
- Fuzzy k-NN
- LVQ

Hierarchical Neural Network Classifiers



Fay et al. *Learning Object Recognition in a NeuroBotic System.* SOAVE 2004

- Hierarchically organised neural networks
- Coarse to fine output space division
- Hierarchical class grouping
- Different strategies to combine evidences of the individual classifiers to a collective output

Hierarchical Neural Network Classifiers

- Classifiers provide evidence at different levels of abstraction
- Classifiers provide evidence for single classes or sets of classes
- The classifier outputs indicate how likely a given sample belongs to a class or group of classes

➔ Hierarchical neural network classifiers can be interpreted as a group of experts making hierarchically structured statements

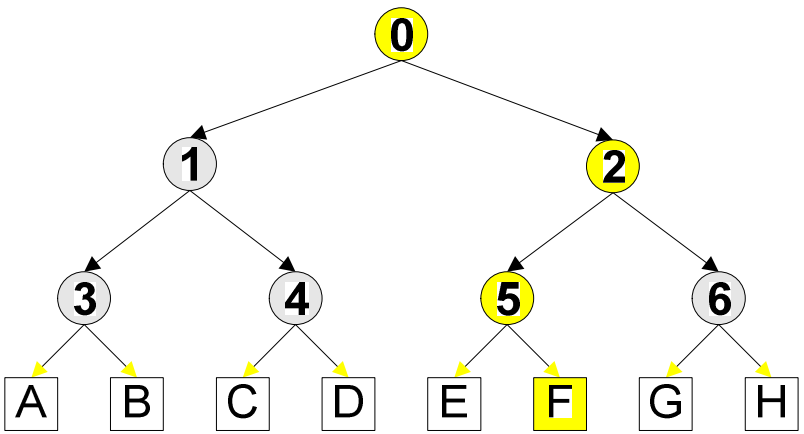
Benefits of Hierarchical Neural Network Classifiers

- Division of complex classification problems into several less complex classification problems
- Links symbolic and sub-symbolic information
- Results are available at different levels of abstraction

Hierarchy Retrieval

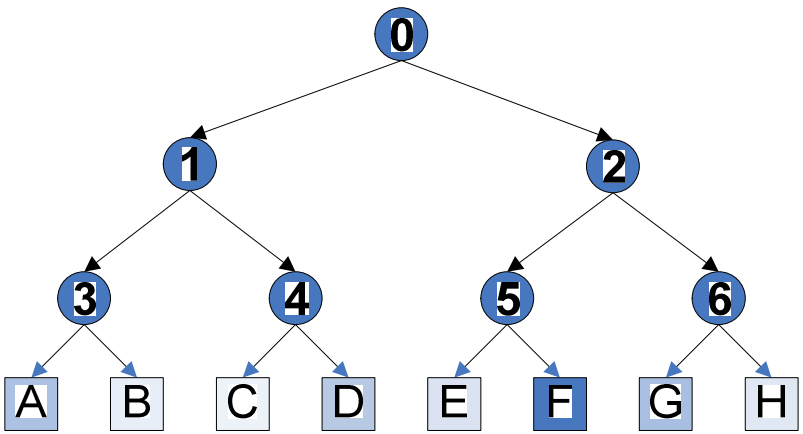
- Retrieval of combined classification result
  - Similar to decision tree approach
  - Utilising Dempster-Shafer theory of evidence

Hierachy Retrieval
Similar to Decision Trees



- Only part of the experts are consulted
- The decision is successively delegated to an expert at the next lower level resulting in a path from the root node to an end node
- Result is a class

Hierarchy Retrieval Utilising Dempster-Shafer Theory of Evidence



- All experts are consulted
- Each expert gives his oppinion
- Result is a class as well as evidences for the individual classes

Overview

- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

Dempster-Shafer Theory of Evidence

- Generalisation of the Bayesian probability theory
- Provides means for dealing with incomplete knowledge, uncertainty and ignorance
- Allows for assigning evidence not only to atomic hypotheses but also to sets of hypotheses
- Provides rule for combining evidence (orthogonal sum)
- Suitable for combining hierarchically structured knowledge incorporating uncertainty

Dempster-Shafer Theory of Evidence

- Basic concepts
  - Frame of discernment $\Omega = \{q_1, q_2, ..., q_q\}$

    Set of mutually exclusive atomic hypotheses
  - Basic probability assignment (or mass) m

    $m : 2^{\Omega} \mapsto [0,1] \quad m(\emptyset) = 0 \quad \sum_{A \subseteq \Omega} m(A) = 1$
  - Dempster's combination rule (orthogonal sum)

    $m_{1,2} = m_1 \oplus m_2$

    Combines evidence from different sources

    $m_{1,2}(C) = K^{-1} \sum_{A,B:A \cap B = C} m_1(A) \cdot m_2(B), \forall C \neq \emptyset$

    $K = 1 - \sum_{A,B:A \cap B = \emptyset} m_1(A) \cdot m_2(B)$

Transferable Belief Model by Smets

- Interpretation of the Dempster-Shafer theory of evidence
- Positive masses can be assigned to the empty set $m(\emptyset) \geq 0$
  - Unnormalised orthogonal sum

$$m_{1,2}(C) = \cancel{K^{-1}} \sum_{A,B:A\cap B=C} m_1(A)\cdot m_2(B), \forall C \subseteq \Omega$$

  - Unnormalised basic probability assignments

$$m : 2^{\Omega} \mapsto [0,1] \quad \cancel{m(\emptyset) = 0} \quad \sum_{A\subseteq\Omega} m(A) = 1$$

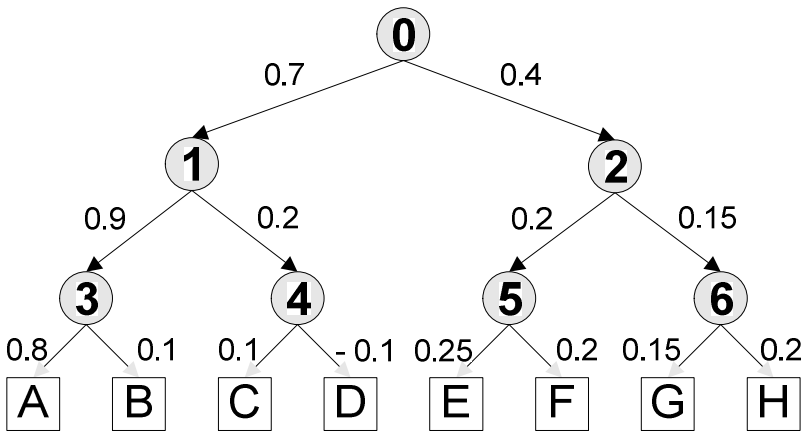
Overview

- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

Evidence-Based Reasoning in Hierarchical Neural Network Classifiers

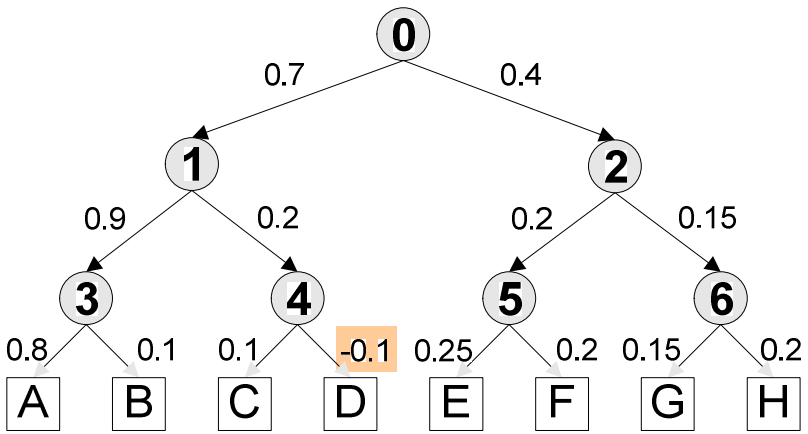
- Mass functions obtained from classifier outputs
  - Enforce positive classifier outputs
  - Normalise classifier outputs with sum greater than one
  - Discount classifier outputs
  - Assign belief to frame of discernment O (Representing ignorance)
- Combine mass functions of all classifiers
- Apply pignistic transformation to obtain final classification result

Example
Evidence-Based Reasoning



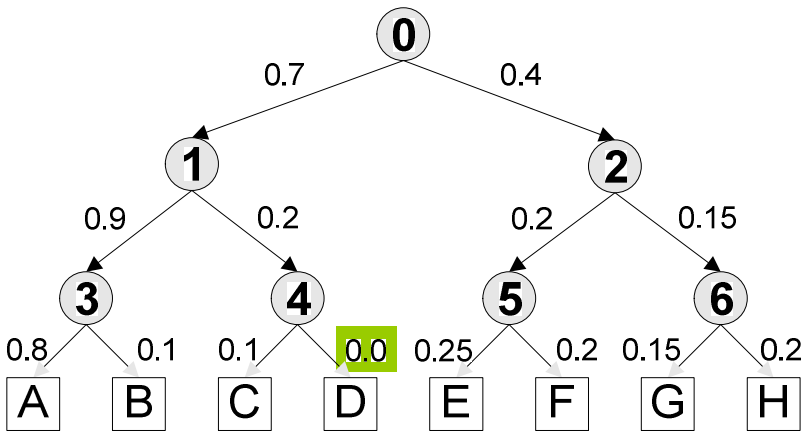
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



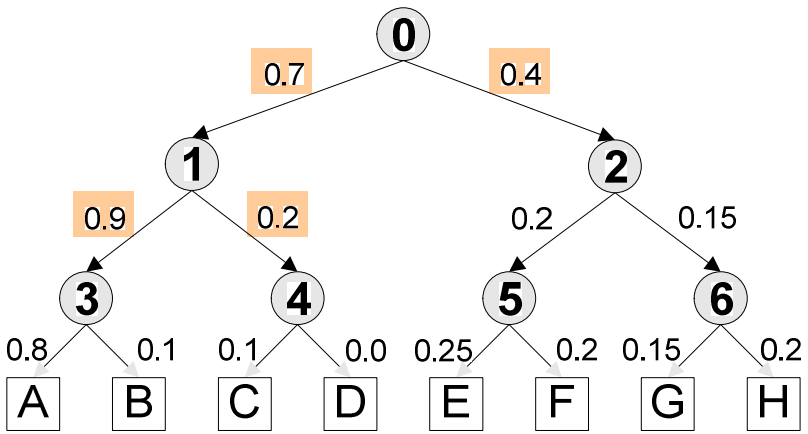
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



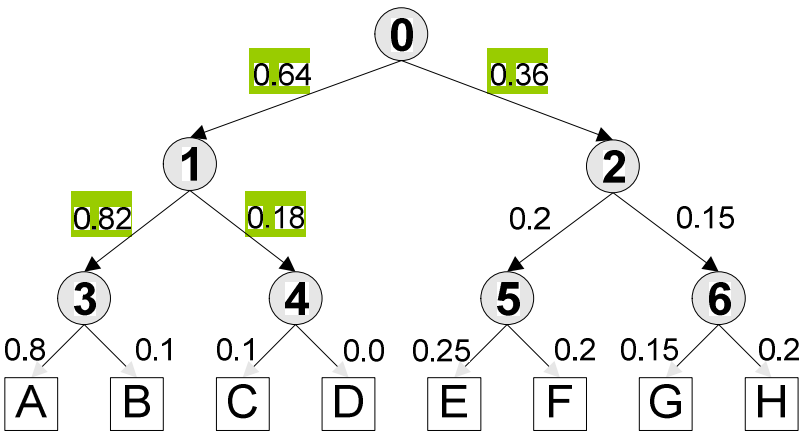
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



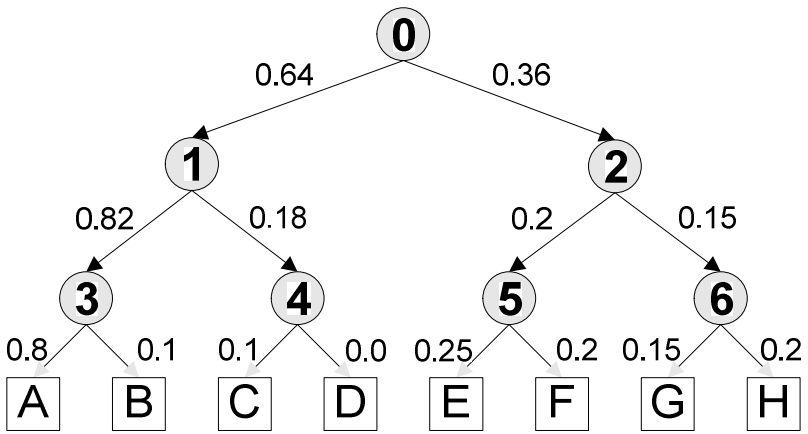
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



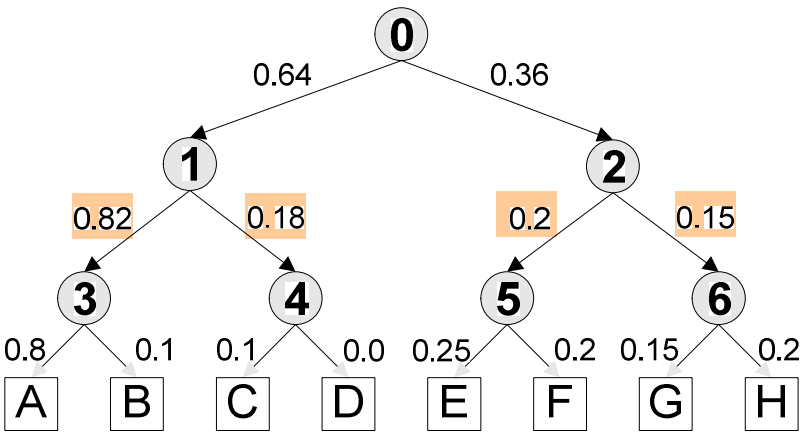
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



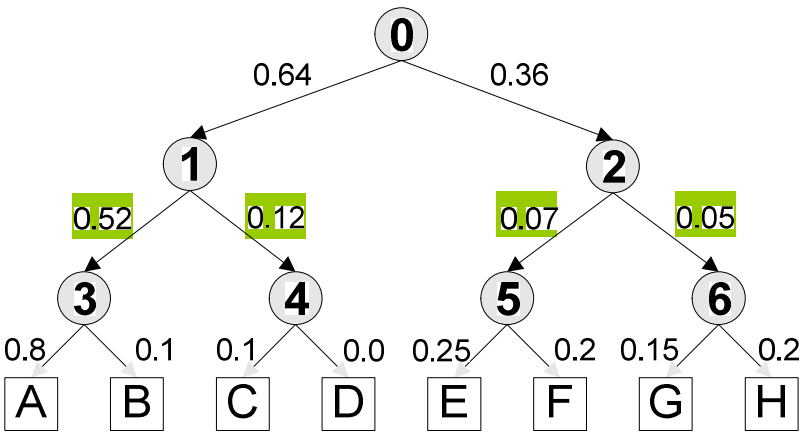
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



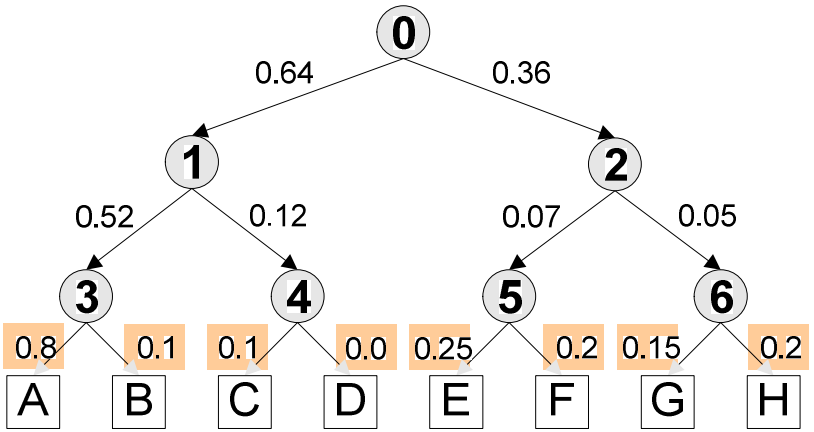
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



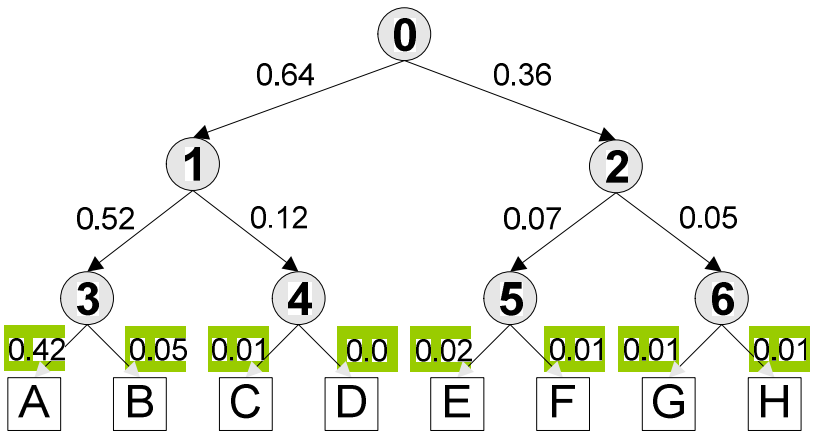
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



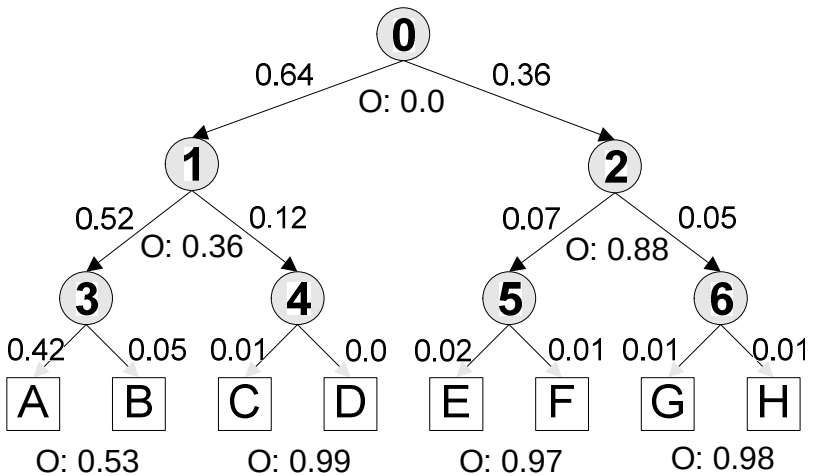
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



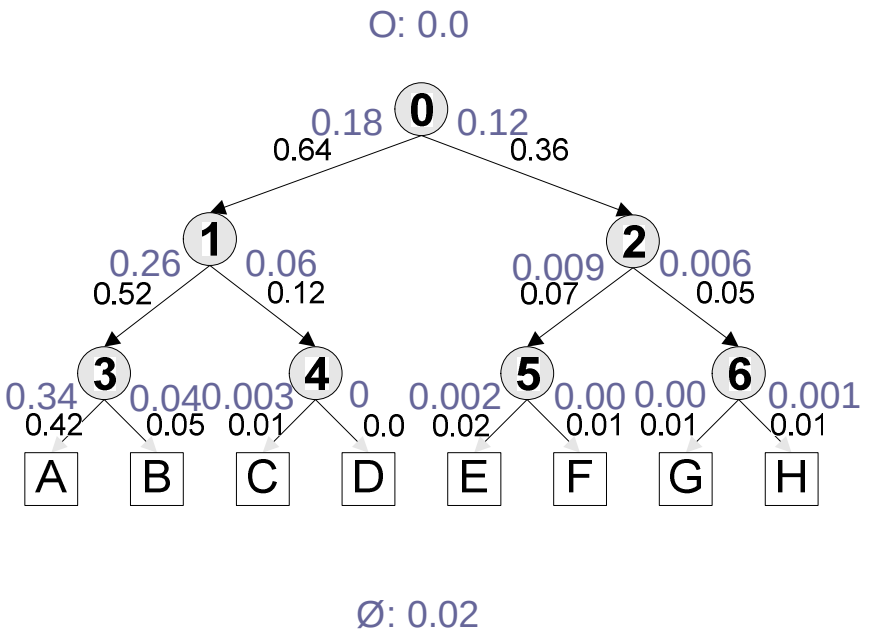
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



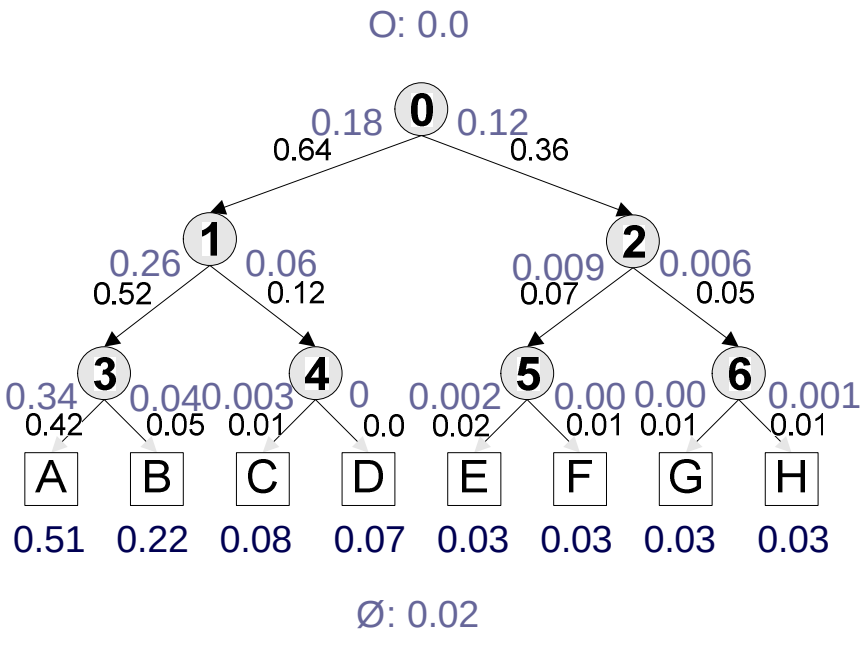
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



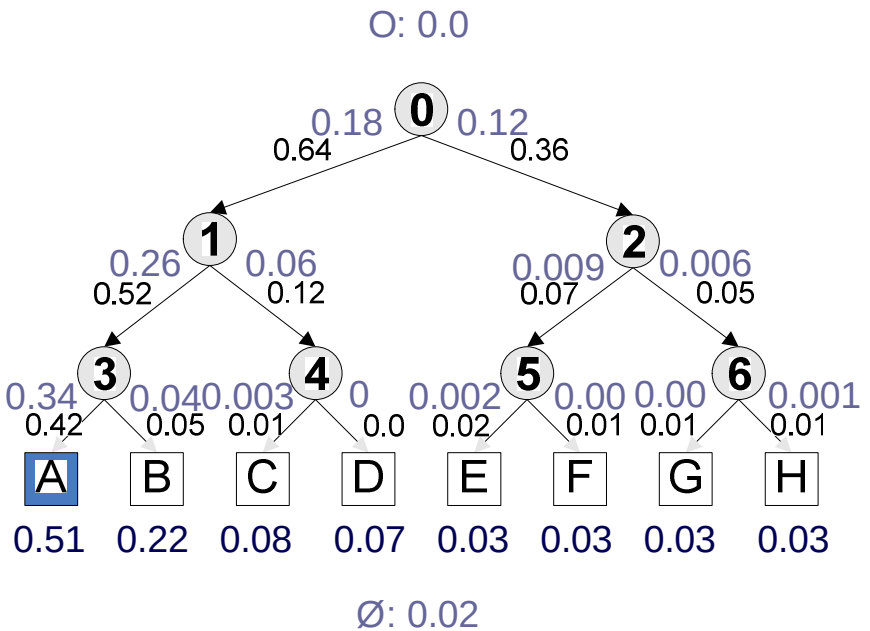
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Example
Evidence-Based Reasoning



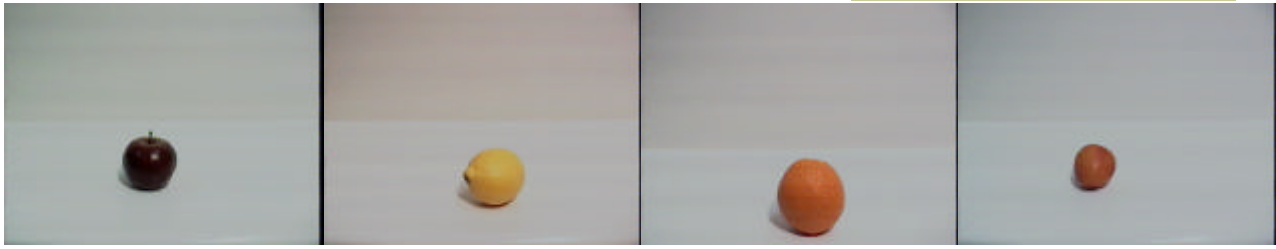
- Enforce positive outputs
- Normalise outputs > 1
- Discount outputs
- Assign belief to frame of discernment O
- Combine evidence of all classifiers
- Calculate pignistig transformation
- Obtain combined classification result

Overview

- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

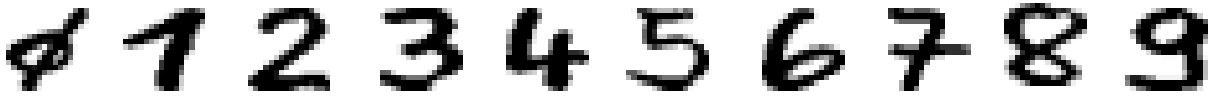
Data Sets

- Fruits

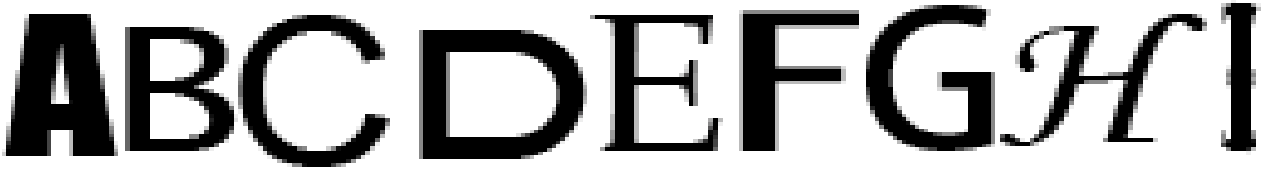


- Objects (COIL)

- Digits



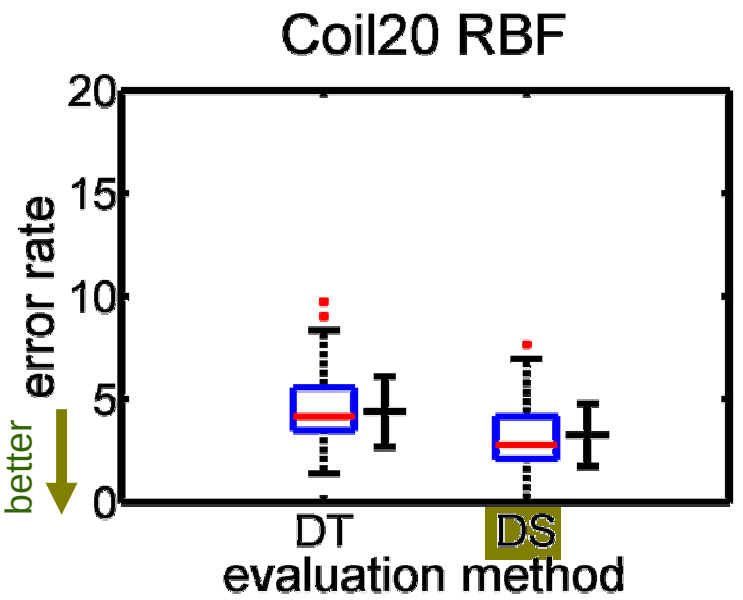
- Letters



Evaluation
Hierarchy Retrieval

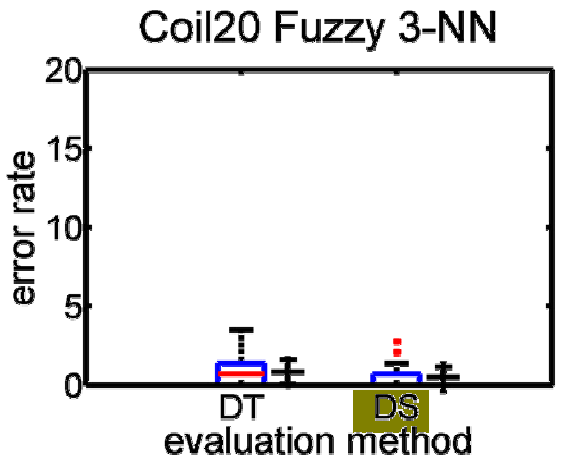
- Compare evidence theoretic retrieval strategy and decision tree retrieval strategy
- Experiments on different data sets:
  - COIL-20
  - Digits
  - Letters
- Different neural classifiers:
  - RBF
  - Fuzzy k-NN classifier
- 10-times 10-fold cross-validation
- Both strategies are evaluated on the same hierarchy

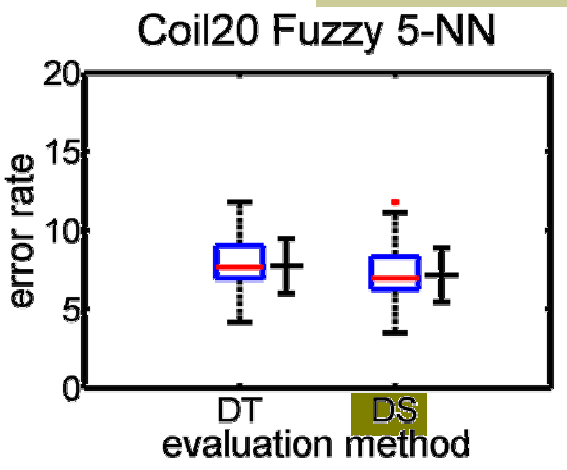
Results
Hierarchy Retrieval

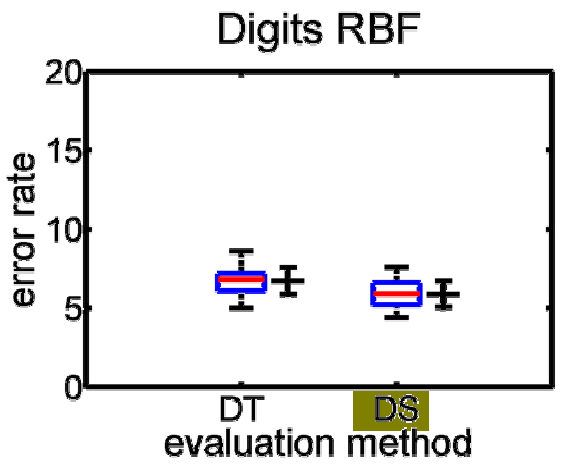


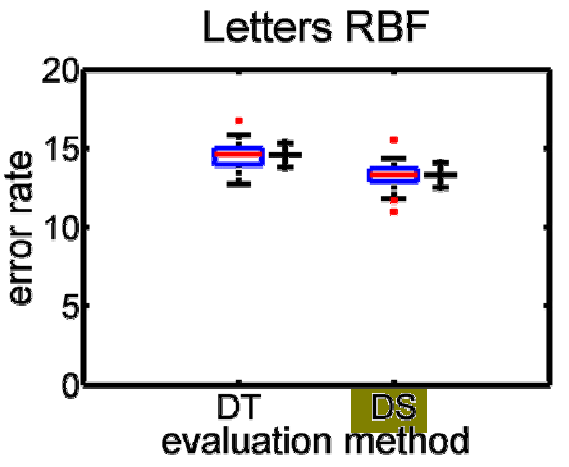
- Data set:
  - COIL-20
- Classifier:
  - RBF
- Dempster-Shafer (DS):
  - Mean error: **3.24±1.52%**
- Decision tree (DT):
  - Mean error: **4.38±1.71%**
- p-value: 0.025
- t-value: 2.26
- Significant at 5%

Results
Hierarchy Retrieval

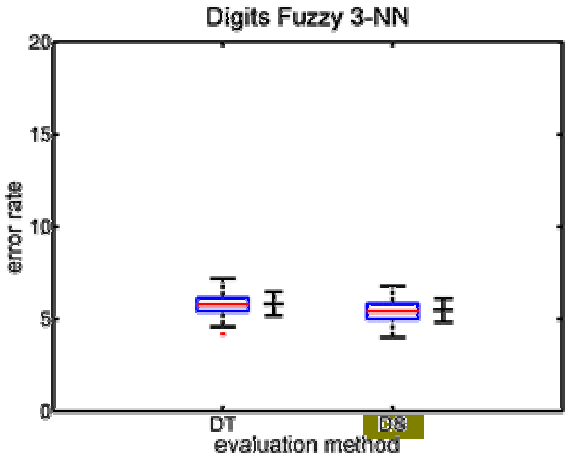


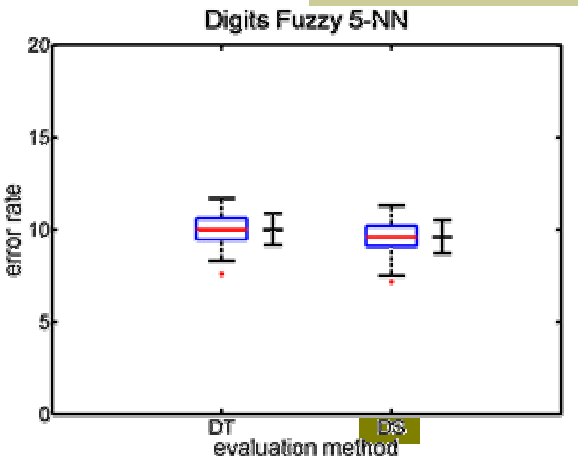


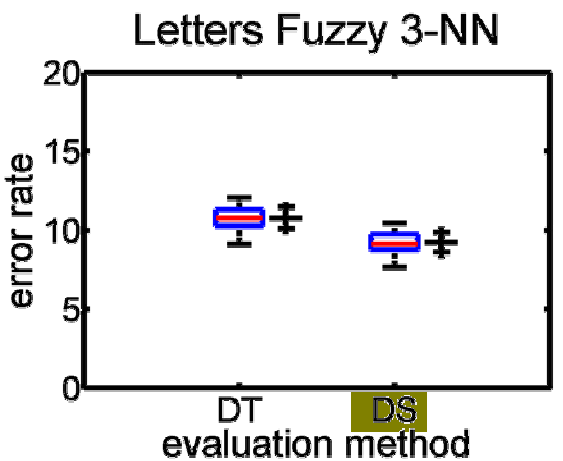


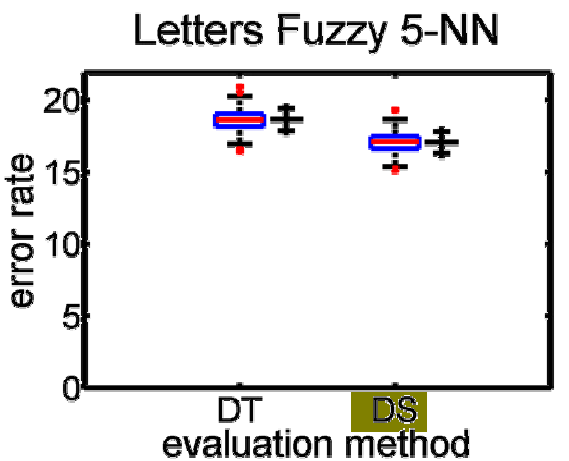


Results
Hierarchy Retrieval









Overview

- Motivation
- Hierarchical Neural Network Classifiers
- Dempster-Shafer Theory of Evidence / Transferable Belief Model
- Evidence-Based Reasoning
- Results
- Conclusions

Conclusions

- Dempster-Shafer theory of evidence is suitable for the application to hierarchical neural network classifiers
- Evidence theoretic hierarchy retrieval outperformed the decision tree retrieval in all tested cases

Thank you for your attention!

This work was partially supported by the European Union award #IST-2001-35282 of the Mirrorbot project.