Ensemble Enhanced Evidential $k$-NN classifier through random subspaces.

Asma Trabelsi$^{1,2}$, Zied Elouedi$^1$, Eric Lefevre$^2$

$^1$ Université de Tunis, Institut Supérieur de Gestion de Tunis, Larodec
$^2$ Université d’Artois, Laboratoire de Génie Informatique et d’Automatique de l’Artois (LGI2A)
1 Introduction

2 Evidence Theory

3 Enhanced Evidential $k$ Nearest Neighbors classifier

4 Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

5 Experimentation

6 Conclusions & Future works
Introduction

Evidence Theory

Enhanced Evidential $k$ Nearest Neighbors classifier

Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Experimentation

Conclusions & Future works
Classifier fusion is regarded as an effective solution for solving several real world classification problems.
Problematic

Real World Application Data

Uncertain Attribute Values

Certain Class Labels
Problematic

How to deal with uncertain attribute values for solving pattern recognition problems
Problematic

Enhanced Evidential $k$-Nearest Neighbors
Motivation

Ensemble Enhanced Evidential $k$-Nearest Neighbors
Outline

1. Introduction
2. Evidence Theory
3. Enhanced Evidential $k$ Nearest Neighbors classifier
4. Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier
5. Experimentation
6. Conclusions & Future works
Evidence theory (1/2)

Frame of discernment
\[ \Theta = \{ \theta_1, \theta_2, ..., \theta_N \} \]

Basic Belief Assignment (bba)
\[ m : 2^\Theta \rightarrow [0, 1] \]

Combination rule
The Dempster rule allows to combine bbas provided by distinct pieces of evidence. It is set as
\[ m_1 \oplus m_2 (A) = 1 - \frac{1}{K} \sum_{B \cap C = A} m_1 (B)m_2 (C), \quad K = \sum_{B \cap C = \emptyset} m_1 (B)m_2 (C) \]

Decision making
The TBM framework, which consists on two main levels (Credal level, Pignistic level), allows to make decision:
\[ \text{BetP}(A) = \sum_{B \cap A = \emptyset \mid A \cap B} \frac{|A \cap B|}{|B|} m(B), \quad \forall A \in \Theta \]
Evidence theory (1/2)

Frame of discernment

\[ \Theta = \{\theta_1, \theta_2, \ldots, \theta_N\} \]

\[ 2^\Theta = \{A, A \subseteq \Theta\} \]
Evidence theory (1/2)

Frame of discernment

\[ \Theta = \{ \theta_1, \theta_2, \ldots, \theta_N \} \]

\[ 2^\Theta = \{ A, A \subseteq \Theta \} \]

Basic Belief Assignment (bba)

\[ m : 2^\Theta \rightarrow [0, 1] \]

\[ \sum_{A \subseteq \Theta} m(A) = 1 \]
Introduction
Evidence Theory
Enhanced Evidential $k$ Nearest Neighbors classifier
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier
Experimentation
Conclusions & Future works

Evidence theory (1/2)

**Frame of discernment**

$$\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$$

$$2^{\Theta} = \{A, A \subseteq \Theta\}$$

**Basic Belief Assignment (bba)**

$$m : 2^{\Theta} \rightarrow [0, 1]$$

$$\sum_{A \subseteq \Theta} m(A) = 1$$

**Combination rule**

The Dempster rule allows to combine bbas provided by distinct pieces of evidence. It is set as $\forall A \subseteq \Theta$:

$$m_1 \oplus m_2(A) = \frac{1}{1 - K} \sum_{B \cap C = A} m_1(B)m_2(C),$$

$$K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$$
### Evidence theory (1/2)

**Frame of discernment**

\[
\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\} \\
2^\Theta = \{A, A \subseteq \Theta\}
\]

**Basic Belief Assignment (bba)**

\[
m : 2^\Theta \rightarrow [0, 1] \\
\sum_{A \subseteq \Theta} m(A) = 1
\]

**Combination rule**

The **Dempster rule** allows to combine bbas provided by distinct pieces of evidence. It is set as \(\forall A \subseteq \Theta:\)

\[
m_1 \oplus m_2(A) = \frac{1}{1 - K} \sum_{B \cap C = A} m_1(B)m_2(C),
\]

\[
K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)
\]

**Decision making**

The TBM framework, which consists on two main levels (Credal level, Pignistic level), allows to make decision:

\[
\text{BetP}(A) = \sum_{B \cap A = \emptyset} \frac{|A \cap B|}{|B|} m(B), \quad \forall A \in \Theta
\]
Evidence theory (2/2)

Dissimilarity between 

Evidence theory

The Jousselme distance between two pieces of evidence \( m_1 \) and \( m_2 \) is found as follows:

\[
d(m_1, m_2) = \sqrt{\frac{1}{2} \left( \overrightarrow{m_1} - \overrightarrow{m_2} \right)^T D(\overrightarrow{m_1} - \overrightarrow{m_2}) \overrightarrow{m_1} - \overrightarrow{m_2}}
\]

\( \overrightarrow{m_1} \) and \( \overrightarrow{m_2} \) are vector representations of \( m_1 \) and \( m_2 \).

\( D \) is the Jaccard similarity measure defined by:

\[
D(A, B) = \begin{cases} 
1 & \text{if } A = B = \emptyset \\
\frac{|A \cap B|}{|A \cup B|} & \text{otherwise}
\end{cases} 
\]

\( A, B \in 2^\Theta \)
Dissimilarity between bbas

The Jousselme distance between two pieces of evidence $m_1$ and $m_2$ is found as follows:

$$d(m_1, m_2) = \sqrt{\frac{1}{2} (\vec{m_1} - \vec{m_2})^T D(\vec{m_1} - \vec{m_2})}$$
Dissimilarity between bbas

The Jousselme distance between two pieces of evidence \( m_1 \) and \( m_2 \) is found as follows:

\[
d(m_1, m_2) = \sqrt{\frac{1}{2} \left( \overrightarrow{m_1} - \overrightarrow{m_2} \right)^T D(\overrightarrow{m_1} - \overrightarrow{m_2})}
\]

- \( \overrightarrow{m_1} \) and \( \overrightarrow{m_2} \) are vector representations of \( m_1 \) and \( m_2 \)
Evidence theory (2/2)

Dissimilarity between bbas

The Jousselme distance between two pieces of evidence $m_1$ and $m_2$ is found as follows:

$$d(m_1, m_2) = \sqrt{\frac{1}{2} (\overrightarrow{m_1} - \overrightarrow{m_2})^T D(\overrightarrow{m_1} - \overrightarrow{m_2})}$$

- $\overrightarrow{m_1}$ and $\overrightarrow{m_2}$ are vector representations of $m_1$ and $m_2$
- $D$ is the Jaccard similarity measure defined by:

$$D(A, B) = \begin{cases} 1 & \text{if } A=B=\emptyset \\ \frac{|A \cap B|}{|A \cup B|} & \forall A, B \in 2^\Theta \end{cases}$$
Let $\Omega = \{w_1, \ldots, w_c\}$ denote the set of classes.

Each instance is described by:
- Uncertain attribute values $x \in R^N$ represented within the belief function framework;
- A certain class label $y \in \Omega$.

Objective: given a learning set $L = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, predict the class label of a new instance described by uncertain attribute values $x$ using the $k$-Nearest Neighbors classifier.
Example

Assume that our data are composed with five instances characterized by three uncertain attributes \( x = \{ \text{Hair, Eye, Height} \} \) and a certain class \( y \) with possible values \( \{ w_1, w_2 \} \). The basic belief assignments, which are affected to the attribute values, will be defined on the frame of discernments \( \Theta_{\text{Hair}} = \{ \text{Blond, Dark} \}, \Theta_{\text{Eye}} = \{ \text{Brown, Blue} \} \) and \( \Theta_{\text{Height}} = \{ \text{Short, Middle, Tall} \} \).

<table>
<thead>
<tr>
<th></th>
<th>( \Theta_{\text{Hair}} )</th>
<th>( \Theta_{\text{Eye}} )</th>
<th>( \Theta_{\text{Height}} )</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_1 )</td>
<td>( m_1^{\Theta_{\text{Hair}}} { \text{Dark} } ) = 0.5</td>
<td>( m_1^{\Theta_{\text{Eye}}} { \text{Brown} } ) = 1</td>
<td>( m_1^{\Theta_{\text{Height}}} { \text{Middle} } ) = 0.95</td>
<td>( \Omega_1 )</td>
</tr>
<tr>
<td></td>
<td>( m_1^{\Theta_{\text{Hair}}} \Theta_{\text{Hair}} ) = 0.5</td>
<td>( m_1^{\Theta_{\text{Eye}}} \Theta_{\text{Eye}} ) = 0</td>
<td>( m_1^{\Theta_{\text{Height}}} \Theta_{\text{Height}} ) = 0.05</td>
<td></td>
</tr>
<tr>
<td>( O_2 )</td>
<td>( m_2^{\Theta_{\text{Hair}}} { \text{Blond} } ) = 0.1</td>
<td>( m_2^{\Theta_{\text{Eye}}} { \text{Blue} } ) = 0.82</td>
<td>( m_2^{\Theta_{\text{Height}}} { \text{Middle} } ) = 1</td>
<td>( \Omega_1 )</td>
</tr>
<tr>
<td></td>
<td>( m_2^{\Theta_{\text{Hair}}} \Theta_{\text{Hair}} ) = 0.9</td>
<td>( m_2^{\Theta_{\text{Eye}}} \Theta_{\text{Eye}} ) = 0.18</td>
<td>( m_2^{\Theta_{\text{Height}}} \Theta_{\text{Height}} ) = 0</td>
<td></td>
</tr>
<tr>
<td>( O_3 )</td>
<td>( m_3^{\Theta_{\text{Hair}}} { \text{Blond} } ) = 0.6</td>
<td>( m_3^{\Theta_{\text{Eye}}} { \text{Brown} } ) = 0.2</td>
<td>( m_3^{\Theta_{\text{Height}}} { \text{Tall} } ) = 0.55</td>
<td>( \Omega_2 )</td>
</tr>
<tr>
<td></td>
<td>( m_3^{\Theta_{\text{Hair}}} \Theta_{\text{Hair}} ) = 0.4</td>
<td>( m_3^{\Theta_{\text{Eye}}} \Theta_{\text{Eye}} ) = 0.8</td>
<td>( m_3^{\Theta_{\text{Height}}} \Theta_{\text{Height}} ) = 0.45</td>
<td></td>
</tr>
<tr>
<td>( O_4 )</td>
<td>( m_4^{\Theta_{\text{Hair}}} { \text{Dark} } ) = 0.7</td>
<td>( m_4^{\Theta_{\text{Eye}}} { \text{Brown} } ) = 0</td>
<td>( m_4^{\Theta_{\text{Height}}} { \text{Short} } ) = 1</td>
<td>( \Omega_1 )</td>
</tr>
<tr>
<td></td>
<td>( m_4^{\Theta_{\text{Hair}}} \Theta_{\text{Hair}} ) = 0.3</td>
<td>( m_4^{\Theta_{\text{Eye}}} \Theta_{\text{Eye}} ) = 1</td>
<td>( m_4^{\Theta_{\text{Height}}} \Theta_{\text{Height}} ) = 0</td>
<td></td>
</tr>
<tr>
<td>( O_5 )</td>
<td>( m_5^{\Theta_{\text{Hair}}} { \text{Blond} } ) = 1</td>
<td>( m_5^{\Theta_{\text{Eye}}} { \text{Blue} } ) = 0.18</td>
<td>( m_5^{\Theta_{\text{Height}}} { \text{Middle} } ) = 0.15</td>
<td>( \Omega_2 )</td>
</tr>
<tr>
<td></td>
<td>( m_5^{\Theta_{\text{Hair}}} \Theta_{\text{Hair}} ) = 0</td>
<td>( m_5^{\Theta_{\text{Eye}}} \Theta_{\text{Eye}} ) = 0.82</td>
<td>( m_5^{\Theta_{\text{Height}}} \Theta_{\text{Height}} ) = 0.85</td>
<td></td>
</tr>
</tbody>
</table>
Let $N_k(x) \subseteq L$ denotes the set of the $k$ nearest neighbors of $x$ in $L$, based on the Jousselme distance measure.

Each $x_i \in N_k(x)$ can be considered as a piece of evidence regarding the class of $x$.

The strength of this evidence decreases with the distance $d_i$ between $x$ and $x_i$. 

Asma Trabelsi
Enhanced Evidential $k$ Nearest Neighbors classifier

If $y_i = w_k$, the evidence of $(x_i, y_i)$ can be represented by the simple mass function:

$$m_i(\{w_k\}) = \phi_k(d_i)$$

$$m_i(\{w_l\}) = 0 \quad \forall l \neq k$$

$$m_i(\Omega) = 1 - \phi_k(d_i)$$

$d_i$ is calculated as the sum of Jousselme distances between the uncertain attribute values.

$\phi_k$ is a decreasing function from $[0, +\infty)$ to $[0, 1]$ such that

$$\lim_{d \to +\infty} \phi_k(d) = 0.$$
If \( y_i = w_k \), the evidence of \((x_i, y_i)\) can be represented by the simple mass function:

\[
\begin{align*}
    m_i(\{w_k\}) &= \varphi_k(d_i) \\
    m_i(\{w_l\}) &= 0 \quad \forall \ l \neq k \\
    m_i(\Omega) &= 1 - \varphi_k(d_i)
\end{align*}
\]
If \( y_i = w_k \), the evidence of \((x_i, y_i)\) can be represented by the simple mass function:

\[
\begin{align*}
    m_i(\{w_k\}) &= \varphi_k(d_i) \\
    m_i(\{w_l\}) &= 0 \quad \forall \ l \neq k \\
    m_i(\Omega) &= 1 - \varphi_k(d_i)
\end{align*}
\]

\(d_i\) is calculated as the sum of Jousselme distances between the uncertain attribute values.
If $y_i = w_k$, the evidence of $(x_i, y_i)$ can be represented by the simple mass function:

$$m_i(\{w_k\}) = \varphi_k(d_i)$$
$$m_i(\{w_l\}) = 0 \quad \forall \ l \neq k$$
$$m_i(\Omega) = 1 - \varphi_k(d_i)$$

$d_i$ is calculated as the sum of Jousselme distances between the uncertain attribute values.

$\varphi_k$ is a decreasing function from $[0, +\infty)$ to $[0, 1]$ such that $\lim_{d \to +\infty} \varphi_k(d) = 0$. 

If \( y_i = w_k \), the evidence of \((x_i, y_i)\) can be represented by the simple mass function:

\[
\begin{align*}
m_i(\{w_k\}) &= \varphi_k(d_i) \\
m_i(\{w_l\}) &= 0 \quad \forall \ l \neq k \\
m_i(\Omega) &= 1 - \varphi_k(d_i)
\end{align*}
\]

\( d_i \) is calculated as the sum of Jousselme distances between the uncertain attribute values.

\( \varphi_k \) is a decreasing function from \([0, +\infty)\) to \([0, 1]\) such that \( \lim_{d \to +\infty} \varphi_k(d) = 0 \).
Enhanced Evidential $k$ Nearest Neighbors classifier

Choice of the function $\phi_k$:

$$\phi_k(d) = \alpha \exp(-\gamma_k d^2).$$

Parameters $\gamma_1, \ldots, \gamma_c$ can be optimized using:

- Exact method relying on a gradient search procedure for medium-sized databases.
- A linearization method for large training sets.

$\alpha$ is a parameter such that $0 < \alpha < 1$.

The evidence of the $k$ nearest neighbors of $x$ is pooled using Dempster's rule of combination:

$$m = \bigoplus_{i \in N_k(x)} m_i.$$
Choice of the function $\varphi_k$: $\varphi_k(d) = \alpha \exp(-\gamma_k d^2)$. 

Parameters $\gamma_1, ..., \gamma_c$ can be optimized using:

- Exact method relying on a gradient search procedure for medium sized databases.
- A linearization method for large training sets.

$\alpha$ is a parameter such that $0 < \alpha < 1$. 

The evidence of the $k$ nearest neighbors of $x$ is pooled using Dempster's rule of combination:

$m = \oplus_{i \in N_k(x)} m_i$.
Enhanced Evidential $k$ Nearest Neighbors classifier

Choice of the function $\varphi_k$: $\varphi_k(d) = \alpha \exp(-\gamma_k d^2)$.

Parameters $\gamma_1, \ldots, \gamma_c$ can be optimized using:
- Exact method relying on a gradient search procedure for medium sized databases.
- A linearization method for large training sets.
Enhanced Evidential $k$ Nearest Neighbors classifier

Choice of the function $\varphi_k$: $\varphi_k(d) = \alpha \exp(-\gamma_k d^2)$.

Parameters $\gamma_1, \ldots, \gamma_c$ can be optimized using:
- Exact method relying on a gradient search procedure for medium sized databases.
- A linearization method for large training sets.

$\alpha$ is a parameter such that $0 < \alpha < 1$. 
Enhanced Evidential $k$ Nearest Neighbors classifier

Choice of the function $\varphi_k$: $\varphi_k(d) = \alpha \exp(-\gamma_k d^2)$.

Parameters $\gamma_1, \ldots, \gamma_c$ can be optimized using:
- Exact method relying on a gradient search procedure for medium sized databases.
- A linearization method for large training sets.

$\alpha$ is a parameter such that $0 < \alpha < 1$.

The evidence of the $k$ nearest neighbors of $x$ is pooled using Dempster’s rule of combination:

$$m = \bigoplus_{x_i \in N_k(x)} m_i$$
Outline

1. Introduction
2. Evidence Theory
3. Enhanced Evidential $k$ Nearest Neighbors classifier
4. Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier
5. Experimentation
6. Conclusions & Future works
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Diversity between classifiers is a substantial factor for achieving a good ensemble. Diversity may be achieved by diversifying the input features using the Random Subspace Method.
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Diversity between classifiers is a substantial factor for achieving a good ensemble.
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Diversity between classifiers is a substantial factor for achieving a good ensemble.

Diversity may be achieved by diversifying the input features.
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Diversity between classifiers is a substantial factor for achieving a good ensemble.

Diversity may be achieved by diversifying the input features

Ensemble Enhanced Evidential $k$-NN classifier through feature subspaces.
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Diversity between classifiers is a substantial factor for achieving a good ensemble.

Diversity may be achieved by diversifying the input features.

Ensemble Enhanced Evidential $k$-NN classifier through feature subspaces.

Generate feature subspaces using the Random Subspace Method.
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Asma Trabelsi

Dependent combination rules
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Number of created classifiers

25 EE $k$-NNs classifiers are sufficient for reducing the error rate and for improving performance.

Size of feature subsets

Randomly select the subspace size, relative to each individual EE $k$-NN classifier, in the range $[n/3; 2n/3]$.
Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

Number of created classifiers

25 EE$k$-NNs classifiers are sufficient for reducing the error rate and for improving performance.
### Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier

<table>
<thead>
<tr>
<th>Number of created classifiers</th>
<th>Size of feature subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 EE$k$-NNs classifiers are sufficient for reducing the error rate and for improving performance.</td>
<td>Randomly select the subspace size, relative to each individual EE$k$-NN classifier, in the range $[n/3;2n/3]$</td>
</tr>
</tbody>
</table>
Outline

1. Introduction
2. Evidence Theory
3. Enhanced Evidential $k$ Nearest Neighbors classifier
4. Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier
5. Experimentation
6. Conclusions & Future works
Experimentation setups

Databases

<table>
<thead>
<tr>
<th>Database</th>
<th>Instances</th>
<th>Attributes</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart</td>
<td>267</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Monks</td>
<td>195</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Lymphography</td>
<td>148</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Audiology</td>
<td>226</td>
<td>69</td>
<td>24</td>
</tr>
</tbody>
</table>

All these databases do not contain uncertain condition attributes represented within the belief function framework.

Generate synthetic databases by taking into account the original databases and a degree of uncertainty $P$ to transform actual condition attribute value $v_A^k$ of each object $u_i$, where $A^k \in A$, into a basic belief assignment:

$$m_{\Theta}^k(i)({v_A^k}) = 1 - P$$

The degree of uncertainty $P$ takes value in the interval $[0,1]$:

- Certain Case ($P=0$)
- Low Uncertainty ($0 \leq P < 0.4$)
- Middle Uncertainty ($0.4 \leq P < 0.7$)
- High Uncertainty ($0.7 \leq P \leq 1$)
## Experimentation setups

<table>
<thead>
<tr>
<th>Databases</th>
<th>Instances</th>
<th>Attributes</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voting records</td>
<td>435</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Heart</td>
<td>267</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Monks</td>
<td>195</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Lymphography</td>
<td>148</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Audiology</td>
<td>226</td>
<td>69</td>
<td>24</td>
</tr>
</tbody>
</table>
All these databases do not contain uncertain condition attributes represented within the belief function framework.

<table>
<thead>
<tr>
<th>Databases</th>
<th>Instances</th>
<th>Attributes</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voting records</td>
<td>435</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Heart</td>
<td>267</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Monks</td>
<td>195</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Lymphography</td>
<td>148</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Audiology</td>
<td>226</td>
<td>69</td>
<td>24</td>
</tr>
</tbody>
</table>
### Experimentation setups

<table>
<thead>
<tr>
<th>Databases</th>
<th>Instances</th>
<th>Attributes</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voting records</td>
<td>435</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Heart</td>
<td>267</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Monks</td>
<td>195</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Lymphography</td>
<td>148</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Audiology</td>
<td>226</td>
<td>69</td>
<td>24</td>
</tr>
</tbody>
</table>

All these databases do not contain uncertain condition attributes represented within the belief function framework.

### Generate synthetic databases

Generate synthetic databases by taking into account the original databases and a degree of uncertainty $P$ to transform actual condition attribute value $v_{A^k}$ of each object $u_i$, where $A^k \in A$, into a basic belief assignment:

$$m_i^{\Theta_k}(\{v_{A^k}\}) = 1 - P$$

$$m_i^{\Theta_k}(\Theta_k) = P$$
Experimentation setups

<table>
<thead>
<tr>
<th>Databases</th>
<th>Instances</th>
<th>Attributes</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voting records</td>
<td>435</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Heart</td>
<td>267</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Monks</td>
<td>195</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Lymphography</td>
<td>148</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>Audiology</td>
<td>226</td>
<td>69</td>
<td>24</td>
</tr>
</tbody>
</table>

All these databases do not contain uncertain condition attributes represented within the belief function framework.

Generate synthetic databases

Generate synthetic databases by taking into account the original databases and a degree of uncertainty $P$ to transform actual condition attribute value $v_{A^k}$ of each object $u_i$, where $A^k \in A$, into a basic belief assignment:

$$m_i^{\Theta_k} (\{v_{A^k}\}) = 1 - P$$

$$m_i^{\Theta_k} (\Theta_k) = P$$

The degree of uncertainty $P$ takes value in the interval [0,1]:

- Certain Case ($P=0$)
- Low Uncertainty ($0 \leq P < 0.4$)
- Middle Uncertainty ($0.4 \leq P < 0.7$)
- High Uncertainty ($0.7 \leq P \leq 1$)
## Experimentation results

### Results for Heart database (%)

<table>
<thead>
<tr>
<th></th>
<th>(k = 1)</th>
<th>(k = 3)</th>
<th>(k = 5)</th>
<th>(k = 7)</th>
<th>(k = 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EE(k)-NN</td>
<td>Ensemble EE(k)-NN</td>
<td>EE(k)-NN</td>
<td>Ensemble EE(k)-NN</td>
<td>EE(k)-NN</td>
</tr>
<tr>
<td>No</td>
<td>61.15</td>
<td>67.30</td>
<td>63.84</td>
<td>70.38</td>
<td>67.30</td>
</tr>
<tr>
<td>Low</td>
<td>58.46</td>
<td>68.84</td>
<td>64.23</td>
<td>66.15</td>
<td>66.92</td>
</tr>
<tr>
<td>Middle</td>
<td>60</td>
<td>69.23</td>
<td>63.07</td>
<td>65.38</td>
<td>66.15</td>
</tr>
<tr>
<td>High</td>
<td>63.84</td>
<td>68.46</td>
<td>63.07</td>
<td>65.76</td>
<td>66.36</td>
</tr>
</tbody>
</table>

### Results for Vote Records database (%)

<table>
<thead>
<tr>
<th></th>
<th>(k = 1)</th>
<th>(k = 3)</th>
<th>(k = 5)</th>
<th>(k = 7)</th>
<th>(k = 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EE(k)-NN</td>
<td>Ensemble EE(k)-NN</td>
<td>EE(k)-NN</td>
<td>Ensemble EE(k)-NN</td>
<td>EE(k)-NN</td>
</tr>
<tr>
<td>No</td>
<td>92.79</td>
<td>92.05</td>
<td>92.32</td>
<td>92.65</td>
<td>93.02</td>
</tr>
<tr>
<td>Low</td>
<td>92.09</td>
<td>93.14</td>
<td>93.02</td>
<td>93.65</td>
<td>92.55</td>
</tr>
<tr>
<td>Middle</td>
<td>91.62</td>
<td>92.79</td>
<td>91.39</td>
<td>92.56</td>
<td>91.39</td>
</tr>
<tr>
<td>High</td>
<td>84.18</td>
<td>87.20</td>
<td>87.67</td>
<td>88.60</td>
<td>88.60</td>
</tr>
</tbody>
</table>
Outline

1. Introduction
2. Evidence Theory
3. Enhanced Evidential $k$ Nearest Neighbors classifier
4. Ensemble Enhanced Evidential $k$ Nearest Neighbors classifier
5. Experimentation
6. Conclusions & Future works
Conclusions

An ensemble EE $k$-NN classifier through random subspaces. An ensemble EE $k$-NN classifier has outperformed the $E_k$-NN that is learned in the full feature space.

Asma Trabelsi
Conclusions

- An ensemble EEk-NN classifier through random subspaces.
Conclusions

- An ensemble EE$k$-NN classifier through random subspaces.

- An ensemble EE$k$-NN classifier has outperformed the $E_k$-NN that is learned in the full feature space.
Future works

Solutions allowing to pick out the best feature subsets. Compare an ensemble EE $k$-NN classifier through random subspaces with ensemble EE $k$-NN classifier learned through other feature subpace methods.

Asma Trabelsi

Dependent combination rules

21/22
Future works

- Solutions allowing to pick out the best feature subsets.
Future works

- Solutions allowing to pick out the best feature subsets.

- Compare an ensemble EE\(k\)-NN classifier through random subspaces with ensemble EE\(k\)-NN classifier learned through other feature subspace methods.
THANK YOU FOR your ATTENTION! ANY QUESTIONS?