Outcome Prediction in Cancer Therapy based on Dempster-Shafer Theory and PET Imaging

Chunfeng Lian\textsuperscript{1,2}, Su Ruan\textsuperscript{2}, Thierry Denœux\textsuperscript{1}, Pierre Vera\textsuperscript{3,2}

\textsuperscript{1}Sorbonne Universités, Université de Technologie de Compiègne, CNRS, UMR 7253 Heudiasyc, Compiègne, France

\textsuperscript{2}Université de Rouen, QuantIF – EA 4108 LITIS, Rouen, France

\textsuperscript{3}Centre Henri-Becquerel, Department of Nuclear Medicine, Rouen, France
Contents

- **Background**
  - Outcome Prediction in Cancer Therapy
  - Difficulties in Outcome Prediction
  - Dempster-Shafer Theory

- **Our Approach**
  - Modified EK-NN Classification Rule
  - Evidential Feature Selection

- **Experimental Results**
  - on UCI data sets
  - on Clinical Data Sets

- **Conclusion**
Outcome Prediction in Cancer Therapy

• Outcome prediction prior to or even during the cancer therapy ⇒ tailoring and adapting a treatment planning.

• To this end, there are diverse information sources:
  - **Patient’s Demography**:
    patient gender, patient age, country . . .
  - **Clinical Characteristics**:
    tumor stage, tumor location, histology, genomic data . . .
  - **Medical Imaging**:
    anatomical images and functional images ⇒ tumor volume, intensity, texture features . . .
Difficulties We Have

1. Information sources are imprecise:
   - **Positron Emission Tomography** (PET) is blurring and noisy.
   - Clinical characteristics offered by clinicians are in some sense subjective and inaccuracy.
Difficulties We Have

1. Information sources are imprecise:
   - Positron Emission Tomography (PET) is blurring and noisy.
   - Clinical characteristics offered by clinicians are in some sense subjective and inaccuracy.

2. No consensus to determine the most predictive features:
   - Dozens of PET imaging features, texture features and clinical features. Maybe redundant, irrelevant or even interference.
Difficulties We Have

1. Information sources are imprecise:
   - Positron Emission Tomography (PET) is blurring and noisy.
   - Clinical characteristics offered by clinicians are in some sense subjective and inaccuracy.

2. No consensus to determine the most predictive features:
   - Dozens of PET imaging features, texture features and clinical features. Maybe redundant, irrelevant or even interference.

We need a stable prediction rule and feature selection.
Dempster-Shafer Theory

- Also known as the **Theory of Belief Functions and Evidence Theory**.
- An extension of **Probability theory and Set-Membership Approach**.
- A powerful framework for reasoning and making decision with partial (uncertain, imprecise) knowledge.
Contents

- **Background**
  - Outcome Prediction in Cancer Therapy
  - Difficulties in Outcome Prediction
  - Dempster-Shafer Theory

- **Our Approach**
  - Modified EK-NN Classification Rule
  - Evidential Feature Selection

- **Experimental Results**
  - on UCI data sets
  - on Clinical Data Sets

- **Conclusion**
Modified EK-NN Classification Rule

Given a query instance $X_t$ and training sample $(X_j, Y_j = \omega_q)$, evidence regarding $X_t$’s label can be quantified as [Denoeux, 1995]:

\[
\begin{align*}
    m_{t,j}(\omega_q) &= \alpha \cdot \exp(-\gamma_q \cdot d_{t,j}^2) \\
    m_{t,j}(\Omega) &= 1 - m_{t,j}(\omega_q)
\end{align*}
\]
Modified EK-NN Classification Rule

Given a query instance $X_t$ and training sample $(X_j, Y_j = \omega_q)$, evidence regarding $X_t$’s label can be quantified as [Denoeux, 1995] :

$$
\begin{align*}
    m_{t,j}(\omega_q) &= \alpha \cdot \exp(-\gamma_q \cdot d_{t,j}^2) \\
    m_{t,j}(\Omega) &= 1 - m_{t,j}(\omega_q)
\end{align*}
$$

A mixed combination rule to fuse $K$-NNs’ evidence

1. NNs with the same label : Dempster’s rule $\Rightarrow m_{\Gamma q}(\omega_q) + m_{\Gamma q}(\Omega) = 1$ ;
2. Between different $m_{\Gamma q}$, where $q = 1, \ldots, c$ :
   - Discounting according to each group’s cardinality $|\Gamma_q|$ :
     $$
     \begin{align*}
     dm_{t,q}^{\Gamma q}(\{\omega_q\}) &= (|\Gamma_q|/|\Gamma_{\max}|)^\eta \times m_t^{\Gamma q}(\omega_q) \\
     dm_{t,q}^{\Gamma q}(\Omega) &= 1 - (|\Gamma_q|/|\Gamma_{\max}|)^\eta \times m_t^{\Gamma q}(\omega_q)
     \end{align*}
     $$
   - Global fusion via Yager’s rule :
     $$
     \begin{align*}
     m_t(\{\omega_q\}) &= dm_{t,q}^{\Gamma q}(\{\omega_q\}) \times \prod_{h \notin \{1, \ldots, c\}\setminus q} dm_t^{\Gamma h}(\Omega), \quad \forall q = 1 \ldots c \\
     m_t(\Omega) &= 1 - \sum_{q=1}^{c} m_t(\{\omega_q\})
     \end{align*}
     $$
## Modified EK-NN

### Examples of mixed combination

#### Comparing fusion results between different combination rules

- Assume ♯1 ~ ♯4 are the training samples with the same distance to a query instance:

<table>
<thead>
<tr>
<th>CASE 1: A NORMAL CONFLICTING SITUATION</th>
<th>Dempster’s rule</th>
<th>Yager’s rule</th>
<th>mixed rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>m({ω_1}) 0.8 0.8 0.0</td>
<td>0.8276</td>
<td>0.1920</td>
<td>0.7680</td>
</tr>
<tr>
<td>m({ω_2}) 0.0 0.0 0.8</td>
<td>0.1379</td>
<td>0.0320</td>
<td>0.0080</td>
</tr>
<tr>
<td>m(Ω) 0.2 0.2 0.2</td>
<td>0.0345</td>
<td>0.7760</td>
<td>0.2240</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CASE 2: A HIGH CONFLICTING SITUATION</th>
<th>Dempster’s rule</th>
<th>Yager’s rule</th>
<th>mixed rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>m({ω_1}) 0.8 0.8 0.0 0.0</td>
<td>0.4898</td>
<td>0.0384</td>
<td>0.0384</td>
</tr>
<tr>
<td>m({ω_2}) 0.0 0.0 0.8 0.8</td>
<td>0.4898</td>
<td>0.0384</td>
<td>0.0384</td>
</tr>
<tr>
<td>m(Ω) 0.2 0.2 0.2 0.2</td>
<td>0.0204</td>
<td>0.9232</td>
<td>0.9232</td>
</tr>
</tbody>
</table>
Modified EK-NN
two-step Classification

Assumption
- Using belief functions:
  precise objects \simeq additional evidence for imprecise objects.

Approach
1. Mass functions constructed via the proposed mixed combination rule;
2. Making decision for easy to classify objects
   \Rightarrow bigger group of "training pairs";
3. Calculating prototypes (i.e., class centers)
   \Rightarrow Making decision for imprecise objects.
Modified EK-NN
A example of two-step Classification

**Figure**: (b) and (c) are credal partition results for EK-NN and mEK-NN; (d)-(f) are decision making results. The error rates are respectively 9.80%, 8.80% and 7.80%.
Evidential Feature Selection

Main idea

Requirements

A good feature subset should satisfy three requirements:

- High classification accuracy;
- **Low imprecision and uncertainty** (small overlaps between different classes);
- Sparsity to reduce the risk of over-fitting.
Evidential Feature Selection

Main idea

Requirements

A good feature subset should satisfy three requirements:

- High classification accuracy;
- Low imprecision and uncertainty (small overlaps between different classes);
- Sparsity to reduce the risk of over-fitting.

According to these requirements, we developed an Evidential Feature Selection (EFS) method based on DST and mEK-NN.
Evidential Feature Selection

- The dissimilarity between two training instances $X_i$ and $X_j$ is measured by a weighted euclidian distance:

$$d_{j,i} = \sqrt{\sum_{p=1}^{m} \lambda_p \cdot (d_{j,i}^p)^2}$$

$\Rightarrow \lambda_p$ is the binary coefficient for feature selection.
The dissimilarity between two training instances $X_i$ and $X_j$ is measured by a weighted euclidian distance:

$$d_{j,i} = \sqrt{\sum_{p=1}^{m} \lambda_p \cdot (d_{j,i}^p)^2}$$

⇒ $\lambda_p$ is the binary coefficient for feature selection.

Based on mEK-NN, features are selected through:

$$\arg \min_{\lambda} \frac{1}{n} \sum_{i=1}^{n} \sum_{q=1}^{c} (Pl_i(\{\omega_q\}) - t_{i,q})^2 + \frac{\rho}{n} \sum_{i=1}^{n} m_i(\Omega) + \delta \times l_0$$

- $m_i$ and $Pl_i$ are the mass and plausibility function;
- label indicator $t_{i,q} = 1$ iff $Y_i = \omega_q$;
- $\rho$ and $\delta$ are two hyper-parameters.
Evidential Feature Selection
Specified loss function

\[
\arg \min_{\lambda_1, \ldots, \lambda_m} \frac{1}{n} \sum_{i=1}^{n} \sum_{q=1}^{c} (1 - t_{i,q} - \sum_{h \neq q} B^i_h)^2 + \rho \times \frac{1}{n} \sum_{i=1}^{n} (1 - \sum_{q=1}^{c} B^i_q) + \delta \sum_{p=1}^{m} [1 - \exp(-5 \lambda_p)] \quad (1)
\]

with

\[B^i_q = A^i_q \prod_{s \in \{1, \ldots, c\} \setminus q} (1 - A^i_s)\]

and

\[A^i_q = \left( \frac{|\Gamma^i_q|}{|\Gamma^i_{max}|} \right)^\eta \left( 1 - \prod_{j \in \Gamma^i_q} [1 - \alpha \exp(-\gamma_q \cdot d^2_{i,j})] \right).\]

⇒ Solved via integer genetic algorithm [Deep et al., 2009].
Evidential Feature Selection

A test on synthetic data

- Data generation [Perkins et al., 2003]:
  There were $n_r$ relevant, $n_c$ redundant and $n_i$ irrelevant features uniformly distributed between $[-1,1]$. Class label was determined only by relevant features:

  \[
  y = \begin{cases} 
  \omega_1 & \text{if } \max_i(x_i) > 2^{1-\frac{1}{n_r}} - 1, \\
  \omega_2 & \text{otherwise.}
  \end{cases}
  \]

  where $x_i$ ($1 \leq i \leq n_r$) is the $i$th relevant feature.
Evidential Feature Selection
A test on synthetic data

- Data generation [Perkins et al., 2003]:

There were \( n_r \) relevant, \( n_c \) redundant and \( n_i \) irrelevant features uniformly distributed between \([-1,1]\). Class label was determined only by relevant features:

\[
y = \begin{cases} 
\omega_1 & \text{if } \max_i(x_i) > 2^{1 - \frac{1}{n_r}} - 1, \\
\omega_2 & \text{otherwise.}
\end{cases}
\]

where \( x_i \) (\( 1 \leq i \leq n_r \)) is the \( i \)th relevant feature.

- Obtained results:

<table>
<thead>
<tr>
<th>( n_r )</th>
<th>( n_c )</th>
<th>( n_i )</th>
<th>subset size</th>
<th>EK-NN</th>
<th>mEK-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>no fs</td>
<td>with fs</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>14.67</td>
<td>12.67</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>16</td>
<td>2</td>
<td>17.33</td>
<td>12.00</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>26</td>
<td>2</td>
<td>23.33</td>
<td>18.67</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>36</td>
<td>2</td>
<td>28.67</td>
<td>26.67</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>46</td>
<td>2</td>
<td>29.33</td>
<td>23.33</td>
</tr>
</tbody>
</table>
Contents

- Background
  - Outcome Prediction in Cancer Therapy
  - Difficulties in Outcome Prediction
  - Dempster-Shafer Theory

- Our Approach
  - Modified EK-NN Classification Rule
  - Evidential Feature Selection

- Experimental Results
  - on UCI data sets
  - on Clinical Data Sets

- Conclusion
TABLE: comparing EFS with classical wrapper methods using 10-fold cross validation. The robustness is evaluated via [Somol and Novovicova, 2010].

<table>
<thead>
<tr>
<th></th>
<th>Error(%)</th>
<th>Robustness(%)</th>
<th>Subset Size</th>
<th>Error(%)</th>
<th>Robustness(%)</th>
<th>Subset Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Iris</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>2.67</td>
<td>n/a</td>
<td>4</td>
<td>7.62</td>
<td>n/a</td>
<td>7</td>
</tr>
<tr>
<td>SFS</td>
<td>4.67</td>
<td>54.55</td>
<td>1</td>
<td>11.90</td>
<td>57.97</td>
<td>2</td>
</tr>
<tr>
<td>SBS</td>
<td>5.33</td>
<td>21.05</td>
<td>2</td>
<td>10.95</td>
<td>23.88</td>
<td>3</td>
</tr>
<tr>
<td>SFFS</td>
<td>5.33</td>
<td>21.62</td>
<td>3</td>
<td>5.24</td>
<td>54.93</td>
<td>2</td>
</tr>
<tr>
<td>EFS*</td>
<td>2.00</td>
<td>100</td>
<td>3</td>
<td>4.76</td>
<td>81.18</td>
<td>3</td>
</tr>
<tr>
<td><strong>Seeds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>13.04</td>
<td>n/a</td>
<td>13</td>
<td>38.87</td>
<td>n/a</td>
<td>8</td>
</tr>
<tr>
<td>SFS</td>
<td>30.50</td>
<td>75</td>
<td>1</td>
<td>61.99</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>SBS</td>
<td>6.24</td>
<td>42.47</td>
<td>5</td>
<td>48.35</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>SFFS</td>
<td>7.29</td>
<td>57.58</td>
<td>4</td>
<td>36.21</td>
<td>40</td>
<td>5</td>
</tr>
<tr>
<td>EFS*</td>
<td>5.13</td>
<td>91.89</td>
<td>3</td>
<td>32.51</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td><strong>Wine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>13.04</td>
<td>n/a</td>
<td>13</td>
<td>38.87</td>
<td>n/a</td>
<td>8</td>
</tr>
<tr>
<td>SFS</td>
<td>30.50</td>
<td>75</td>
<td>1</td>
<td>61.99</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>SBS</td>
<td>6.24</td>
<td>42.47</td>
<td>5</td>
<td>48.35</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>SFFS</td>
<td>7.29</td>
<td>57.58</td>
<td>4</td>
<td>36.21</td>
<td>40</td>
<td>5</td>
</tr>
<tr>
<td>EFS*</td>
<td>5.13</td>
<td>91.89</td>
<td>3</td>
<td>32.51</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td><strong>Yeast</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>7.20</td>
<td>n/a</td>
<td>30</td>
<td>13.37</td>
<td>n/a</td>
<td>22</td>
</tr>
<tr>
<td>SFS</td>
<td>14.44</td>
<td>80</td>
<td>1</td>
<td>15.82</td>
<td>33.33</td>
<td>1</td>
</tr>
<tr>
<td>SBS</td>
<td>19.67</td>
<td>22.22</td>
<td>2</td>
<td>19.03</td>
<td>23.91</td>
<td>2</td>
</tr>
<tr>
<td>SFFS</td>
<td>9.87</td>
<td>25</td>
<td>4</td>
<td>13.79</td>
<td>43.65</td>
<td>3</td>
</tr>
<tr>
<td>EFS*</td>
<td>5.80</td>
<td>92.37</td>
<td>3</td>
<td>8.63</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td><strong>WDBC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>7.20</td>
<td>n/a</td>
<td>30</td>
<td>13.37</td>
<td>n/a</td>
<td>22</td>
</tr>
<tr>
<td>SFS</td>
<td>14.44</td>
<td>80</td>
<td>1</td>
<td>15.82</td>
<td>33.33</td>
<td>1</td>
</tr>
<tr>
<td>SBS</td>
<td>19.67</td>
<td>22.22</td>
<td>2</td>
<td>19.03</td>
<td>23.91</td>
<td>2</td>
</tr>
<tr>
<td>SFFS</td>
<td>9.87</td>
<td>25</td>
<td>4</td>
<td>13.79</td>
<td>43.65</td>
<td>3</td>
</tr>
<tr>
<td>EFS*</td>
<td>5.80</td>
<td>92.37</td>
<td>3</td>
<td>8.63</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td><strong>Parkinsons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>7.20</td>
<td>n/a</td>
<td>30</td>
<td>13.37</td>
<td>n/a</td>
<td>22</td>
</tr>
<tr>
<td>SFS</td>
<td>14.44</td>
<td>80</td>
<td>1</td>
<td>15.82</td>
<td>33.33</td>
<td>1</td>
</tr>
<tr>
<td>SBS</td>
<td>19.67</td>
<td>22.22</td>
<td>2</td>
<td>19.03</td>
<td>23.91</td>
<td>2</td>
</tr>
<tr>
<td>SFFS</td>
<td>9.87</td>
<td>25</td>
<td>4</td>
<td>13.79</td>
<td>43.65</td>
<td>3</td>
</tr>
<tr>
<td>EFS*</td>
<td>5.80</td>
<td>92.37</td>
<td>3</td>
<td>8.63</td>
<td>100</td>
<td>3</td>
</tr>
</tbody>
</table>
UCI Data Sets

classification

**TABLE:** comparing the classification performance. SFFS was used to select features for other methods. For BK-NN [Liu et al., 2013] and CCR [Liu et al., 2014], $R_e$ and $R_i$ represent, respectively, the error rate and imprecision rate.

<table>
<thead>
<tr>
<th></th>
<th>Iris</th>
<th>Seeds</th>
<th>Wine</th>
<th>Yeast</th>
<th>WDBC</th>
<th>Parkinson</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>8.00</td>
<td>7.62</td>
<td>9.64</td>
<td>32.57</td>
<td>9.15</td>
<td>9.63</td>
</tr>
<tr>
<td>CART</td>
<td>8.00</td>
<td>7.14</td>
<td>9.09</td>
<td>37.55</td>
<td>10.04</td>
<td>11.21</td>
</tr>
<tr>
<td>SVM</td>
<td>6.00</td>
<td>7.14</td>
<td>6.83</td>
<td>36.14</td>
<td>8.28</td>
<td>13.26</td>
</tr>
<tr>
<td>EK-NN</td>
<td>5.33</td>
<td>6.67</td>
<td>6.18</td>
<td>35.07</td>
<td>9.70</td>
<td>16.39</td>
</tr>
<tr>
<td>BK-NN ($R_e$, $R_i$)</td>
<td>4.00</td>
<td>2.38</td>
<td>6.74</td>
<td>16.31</td>
<td>7.22</td>
<td>9.18</td>
</tr>
<tr>
<td></td>
<td>4.67</td>
<td>11.90</td>
<td>5.13</td>
<td>40.84</td>
<td>8.44</td>
<td>13.37</td>
</tr>
<tr>
<td>CCR ($R_e$, $R_i$)</td>
<td>4.00</td>
<td>3.81</td>
<td>3.99</td>
<td>19.53</td>
<td>5.99</td>
<td>16.42</td>
</tr>
<tr>
<td></td>
<td>4.67</td>
<td>18.57</td>
<td>15.33</td>
<td>36.11</td>
<td>15.83</td>
<td>12.26</td>
</tr>
<tr>
<td><strong>our method</strong></td>
<td><strong>2.00</strong></td>
<td><strong>4.76</strong></td>
<td><strong>5.13</strong></td>
<td><strong>32.51</strong></td>
<td><strong>5.80</strong></td>
<td><strong>8.63</strong></td>
</tr>
</tbody>
</table>
Clinical Data Sets
Lung Tumor Data

- **twenty-five patients** with stage II-III non small cell lung cancer were treated with curative intent chemo-radiotherapy.
- FDG-PETs: before treatment, after chemotherapy and during radiotherapy.
- definition of recurrence after one year: local/distant recurrence (19 patients) and no recurrence (6 patients).
Clinical Data Sets
Esophageal Tumor Data

- thirty-six patients with esophageal squamous cell carcinomas were treated with chemo-radiotherapy, and followed up in a long term up to five years.

- FDG-PETs: only before treatment is available.

- neither loco regional nor distant recurrence (13 patients) and disease-positive (23 patients).
Clinical Data Sets
Feature Extraction

- Three types of PET imaging features.
  - **SUV-based features**: $\text{SUV}_{\text{max}}$, $\text{SUV}_{\text{mean}}$, $\text{SUV}_{\text{peak}}$, metabolic tumor volume (MTV) and total lesion glycolysis (TLG);
  - **texture features**: gray level size zone matrix [Tixier et al., 2012];
  - **longitudinal change**: relative difference between baseline features and follow-up features.
Clinical Data Sets
Feature Extraction

- Three types of PET imaging features.
  - SUV-based features: $SUV_{\text{max}}$, $SUV_{\text{mean}}$, $SUV_{\text{peak}}$, metabolic tumor volume (MTV) and total lesion glycolysis (TLG);
  - texture features: gray level size zone matrix [Tixier et al., 2012];
  - longitudinal change: relative difference between baseline features and follow-up features.

- Patients’ clinical characteristics for esophageal tumor data.
  - gender, tumor stage, dysphagia grade, WHO performance status, weight loss, tumor location.
Clinical Data Sets
results

TABLE: Comparing feature selection performance using leave-one-out cross-validation. EFS* denotes the proposed method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Lung Tumor Data</th>
<th></th>
<th>Esophageal Tumor Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>Robustness</td>
<td>subset size</td>
<td>Accuracy</td>
</tr>
<tr>
<td>All features</td>
<td>76±44</td>
<td>n/a</td>
<td>52</td>
<td>64±49</td>
</tr>
<tr>
<td>SFS</td>
<td>84±37</td>
<td>60</td>
<td>3</td>
<td>53±44</td>
</tr>
<tr>
<td>SFFS</td>
<td>72±46</td>
<td>54</td>
<td>4</td>
<td>81±40</td>
</tr>
<tr>
<td>SVMRFE</td>
<td>92±28</td>
<td>57</td>
<td>5</td>
<td>75±44</td>
</tr>
<tr>
<td>EFS*</td>
<td>100±0</td>
<td>94</td>
<td>4</td>
<td>81±40</td>
</tr>
</tbody>
</table>

TABLE: Comparing classification performance. mEK-NN* denotes the proposed classification method.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Lung Tumor Data</th>
<th></th>
<th>Esophageal Tumor Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>without EFS</td>
<td>with EFS</td>
<td>without EFS</td>
<td>with EFS</td>
</tr>
<tr>
<td>ANN</td>
<td>68±48</td>
<td>92±28</td>
<td>67±48</td>
<td>83±38</td>
</tr>
<tr>
<td>SVM</td>
<td>76±44</td>
<td>100±0</td>
<td>64±49</td>
<td>81±40</td>
</tr>
<tr>
<td>EK-NN</td>
<td>68±48</td>
<td>96±20</td>
<td>64±49</td>
<td>83±38</td>
</tr>
<tr>
<td>mEK-NN*</td>
<td>56±51</td>
<td>100±0</td>
<td>53±44</td>
<td>89±32</td>
</tr>
</tbody>
</table>
Contents

- Background
  - Outcome Prediction in Cancer Therapy
  - Difficulties in Outcome Prediction
  - Dempster-Shafer Theory

- Our Approach
  - Modified EK-NN Classification Rule
  - Evidential Feature Selection

- Experimental Results
  - on UCI data sets
  - on Clinical Data Sets

- Conclusion
Conclusions

- More information about this presentation:
Conclusions

More information about this presentation:


Future work:

- Tackling *imbalanced learning problem* and *small sample size effect*, so as to improve the performance;
- Evaluating the proposed method on larger clinical data sets.
Thanks for Your Attention.

chunfeng.lian@utc.fr
A real coded genetic algorithm for solving integer and mixed integer optimization problems. 
*Applied Mathematics and Computation*, 212(2) :505–518.

A k-nearest neighbor classification rule based on dempster-shafer theory. 

A new belief-based k-nearest neighbor classification method. 
*Pattern Recognition*, 46(3) :834–844.

Credal classification rule for uncertain data based on belief functions. 
*Pattern Recognition*, 47(7) :2532–2541.

Grafting : Fast, incremental feature selection by gradient descent in function space. 
References II
