Predicting Malignancy from Mammography Findings and Image Guided Core Biopsies

Breast Cancer Workshop 2013 – June 19\textsuperscript{th} 2013
Porto, Portugal

Pedro Miguel Ferreira
Nuno A. Fonseca
Inês Dutra
Ryan Woods
Elizabeth Burnside
Outline

• Breast Cancer
• Objectives
• Dataset
• Methodology
• Results and Analysis
• *MammoClass* (online application)
• Conclusions and Future Work
Outline

• Breast Cancer
• Objectives
• Dataset
• Methodology
• Results and Analysis
• *MammoClass* (online application)
• Conclusions and Future Work
Breast Cancer

- **USA:**
  - 1 woman dies of breast cancer every 13 minutes
  - **In 2011:**
    - 230,480 invasive cancers
    - 39,520 (~17%) expected to die


- **Portugal:**
  - **Per year:**
    - 4,500 new cases
    - 1,500 deaths (33%)

Source: *Liga Portuguesa Contra o Cancro* – accessed June 2013
Breast Screening Programs

- Reduction of death rate in 30%

- **Mammography:**
  
The cheapest and most efficient method to detect cancer in a preclinical stage
Mammography

**Nodule/Mass:**

Solid lesion with more than 1 cm of width and usually well defined.

Also known as tumour.
Mammography
Outline

• Breast Cancer
• Objectives
• Dataset
• Methodology
• Results and Analysis
• MammoClass (online application)
• Conclusions and Future Work
Objectives

- Build classifiers capable of predicting mass density and malignancy from a reduced set of mammography findings

- Reduce the number of unnecessary biopsies

Outline

• Breast Cancer
• Objectives
• Dataset
• Methodology
• Results and Analysis
• *MammoClass* (online application)
• Conclusions and Future Work
Dataset

- Source:
  - Ryan Woods (M.D.)
  - Elizabeth Burnside (M.D.)

- 348 cases

- Each case refers to a breast nodule retrospectively classified according to BI-RADS® system

- From mammographies results

- Collected between October 2005 and December 2007
Attributes

13 attributes

118 (33.9%) malignant (+)
230 (66.1%) benign (-)
Masses classification

**Prospective**

- **Classification** of feature **mass density just by one radiologist**:
  - low density;
  - iso-dense;
  - high density;

- **Brief** and superficial medical **report** (at the time of imaging);

- **Classification under stress**.

**Retrospective**

- **Classification** by a **group of experienced physicians** that **re-assess** all **exams**;

- **Review of mass density** **classification** made by radiologist (prospective study);

- **Classification without stress**;

- **Reference standard** for **mass density**.
Masses classification

348 cases (retrospectively classified)

180
(≈ 52%)

168
(≈ 48%)

(prospectively classified)
Outline

• Breast Cancer
• Objectives
• Dataset
• Methodology
• Results and Analysis
• *MammoClass* (online application)
• Conclusions and Future Work
Methodology

- **WEKA**
- **Paired Corrected T-Tester**
  - Significance level: 0.05

348

180

168

10 x strat. c. v.

test
Methodology - Experiments

10 x stratified. c. v.

- $E_1$ – Predicting malignancy with $\text{retro\_density}$
- $E_2$ – Predicting malignancy with $\text{density\_num}$
- $E_3$ – Predicting malignancy without mass density
- $E_4$ – Predicting $\text{retro\_density}$*
- $E_5$ – Predicting $\text{density\_num}$*

* in all experiments the low and iso densities were merged into a single class
Methodology - Algorithms applied

- **ZeroR** (baseline classifier)
- **OneR**
- **DTNB**
- **PART**

- **NaiveBayes**
- **BayesNet (TAN)**

**rules**

- **J48**
- **DecisionStump**
- **RandomForest**
- **SimpleCart**
- **NBTTree**

**trees**

- **SMO**

**bayes**

**functions**

**internal parameter variation**
Results

348

180 168

10 x strat. c.v. test
## Results - Experiments

### 10 x stratified. c. v.

<table>
<thead>
<tr>
<th>Exp</th>
<th>Algorithm</th>
<th>CCI</th>
<th>K</th>
<th>F</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>SMO</td>
<td>85.6±7.3</td>
<td>0.69±0.16</td>
<td>0.80±0.11</td>
<td>0.84±0.08</td>
</tr>
<tr>
<td>E1</td>
<td>DTNB</td>
<td>81.6±8.2</td>
<td>0.60±0.18</td>
<td>0.74±0.13</td>
<td>0.88±0.07</td>
</tr>
<tr>
<td>E1</td>
<td>NaiveBayes</td>
<td>81.3±9.5</td>
<td>0.61±0.20</td>
<td>0.76±0.12</td>
<td>0.88±0.08</td>
</tr>
<tr>
<td>E1</td>
<td>J48</td>
<td>80.7±9.3</td>
<td>0.59±0.20</td>
<td>0.75±0.13</td>
<td>0.79±0.11</td>
</tr>
<tr>
<td>E2</td>
<td>SMO</td>
<td>83.9±7.7</td>
<td>0.66±0.17</td>
<td>0.78±0.11</td>
<td>0.82±0.08</td>
</tr>
<tr>
<td>E2</td>
<td>NaiveBayes</td>
<td>80.3±9.3</td>
<td>0.59±0.19</td>
<td>0.75±0.12</td>
<td>0.87±0.09</td>
</tr>
<tr>
<td>E2</td>
<td>DTNB</td>
<td>79.8±9.5</td>
<td>0.56±0.21</td>
<td>0.72±0.15</td>
<td>0.86±0.09</td>
</tr>
<tr>
<td>E2</td>
<td>J48</td>
<td>75.4±9.5</td>
<td>0.47±0.21</td>
<td>0.65±0.15</td>
<td>0.73±0.12</td>
</tr>
<tr>
<td>E3</td>
<td>SMO</td>
<td>83.8±7.7</td>
<td>0.65±0.17</td>
<td>0.78±0.11</td>
<td>0.82±0.09</td>
</tr>
<tr>
<td>E3</td>
<td>J48</td>
<td>76.3±9.9</td>
<td>0.49±0.22</td>
<td>0.67±0.15</td>
<td>0.76±0.13</td>
</tr>
<tr>
<td>E3</td>
<td>NaiveBayes</td>
<td>76.2±9.9</td>
<td>0.51±0.20</td>
<td>0.71±0.13</td>
<td>0.85±0.09</td>
</tr>
<tr>
<td>E3</td>
<td>DTNB</td>
<td>75.7±9.0</td>
<td>0.48±0.19</td>
<td>0.67±0.13</td>
<td>0.81±0.10</td>
</tr>
<tr>
<td>E4</td>
<td>SMO</td>
<td>81.3±8.2</td>
<td>0.52±0.21</td>
<td>0.64±0.17</td>
<td>0.75±0.11</td>
</tr>
<tr>
<td>E4</td>
<td>J48</td>
<td>74.4±8.8</td>
<td>0.32±0.24</td>
<td>0.47±0.21</td>
<td>0.67±0.15</td>
</tr>
<tr>
<td>E4</td>
<td>DTNB</td>
<td>73.5±10.0</td>
<td>0.34±0.24</td>
<td>0.51±0.19</td>
<td>0.76±0.12</td>
</tr>
<tr>
<td>E4</td>
<td>NaiveBayes</td>
<td>72.8±9.9</td>
<td>0.37±0.23</td>
<td>0.56±0.18</td>
<td>0.77±0.11</td>
</tr>
<tr>
<td>E5</td>
<td>NaiveBayes</td>
<td>67.2±12.1</td>
<td>0.33±0.25</td>
<td>0.62±0.15</td>
<td>0.72±0.14</td>
</tr>
<tr>
<td>E5</td>
<td>SMO</td>
<td>66.8±10.7</td>
<td>0.31±0.22</td>
<td>0.55±0.16</td>
<td>0.65±0.11</td>
</tr>
<tr>
<td>E5</td>
<td>J48</td>
<td>63.6±10.1</td>
<td>0.26±0.21</td>
<td>0.56±0.15</td>
<td>0.62±0.13</td>
</tr>
<tr>
<td>E5</td>
<td>DTNB</td>
<td>62.1±11.9</td>
<td>0.22±0.24</td>
<td>0.54±0.16</td>
<td>0.64±0.14</td>
</tr>
</tbody>
</table>

*Predicting malignancy with retro_density*
Results - Experiments

Predicting density
Results - Experiments

10 x stratified. c. v.

- $E_4$ – Predicting \textit{retro\_density}

SVM’s

- CCI: 81.3\% (+/- 8.2)
- Sens: 0.57 (+/- 0.20)
- Spec: 0.92 (+/- 0.07)
- F: 0.64 (+/- 0.17)

Radiologist’s accuracy = 70 \%
Classifier \approx 81 \%
Results - Experiments

TEST

• $\textbf{E}_6$ – Predicting \textit{retro\_density}
  (model $\textbf{E}_4$ applied)

  \begin{itemize}
    \item \textbf{CCI}: 84.5%
    \item Sens: 0.57
    \item Spec: 0.90
    \item F: 0.55
  \end{itemize}
Results - Experiments

Predicting malignancy
Results - Experiments

**10 x stratified. c. v.**

- $E_1 \rightarrow$ Predicting malignancy with *retro_density*

SVM’s

- CCI: 85.6% (+/- 7.3)
- Sens: 0.78 (+/- 0.15)
- Spec: 0.91 (+/- 0.07)
- F: 0.80 (+/- 0.11)
Results - Experiments

TEST

• E₈ – Predicting malignancy with \textit{retro\_density}
  (model E₁ applied)

\begin{align*}
\text{CCI: } & 81.0\% \\
\text{Sens: } & 0.57 \\
\text{Spec: } & 0.90 \\
\text{F: } & 0.63
\end{align*}

\textbf{SVM’s} with \textit{real values of \textit{retro\_density}}

\begin{align*}
\text{CCI: } & 80.4\% \\
\text{Sens: } & 0.57 \\
\text{Spec: } & 0.89 \\
\text{F: } & 0.62
\end{align*}

\textbf{SVM’s} with \textit{predicted values of \textit{retro\_density}} by classifier E₆

\begin{align*}
\text{CCI: } & 85.6\% (+/- 7.2) \\
\text{Sens: } & 0.78 (+/- 0.15) \\
\text{Spec: } & 0.91 (+/- 0.07) \\
\text{F: } & 0.80 (+/- 0.11)
\end{align*}
MammoClass

- Online application freely available at:
  - http://cracs.fc.up.pt/mammoclass/
Conclusions and Future Work

a) We built models that predict malignancy and mass density based on mammography findings;

b) Machine learning classifiers to predict mass density may aid radiologists during the prospective mass classification

c) One of our classifiers can predict malignancy even in the absence of mass density, since we can fill up this attribute using our mass density predictor.
Conclusions and Future Work

a) Apply other machine learning techniques based on statistical relational learning;

b) Investigate how other features can affect malignancy or are related to the other attributes.
Future Work - Challenges

- Correct classification of BIRADS categories:

  - BIRADS 5: 39 instances
  - BIRADS 4: 131 instances
  - BIRADS 0: 178 instances

- Problems:
  - multi-class problem
  - classes not balanced

348 cases
Future Work - Challenges

- Correct **classification of BIRADS** categories:

  - BIRADS 5 → 39 instances
  - BIRADS 4 → 131 instances
  - BIRADS 0 → 178 instances

  \[ \text{348 cases} \]

- **Approaches**:
  - oversampling
  - undersampling
  - **nested cross-validation** on 348 cases (best results so far)
  - cost-sensitive learning (to be applied)
Future Work - Challenges

- Correct **classification of BIRADS** categories:

  - **BIRADS 5** → 39 instances
  - **BIRADS 4** → 131 instances
  - **BIRADS 0** → 178 instances

  - 348 cases

- **nested cross-validation** on 348 cases (best results so far)

  - \( \text{PPV} = 0.57 \) (B5)
  - \( \text{PPV} = 0.42 \) (B4)
  - \( \text{PPV} = 0.57 \) (B0)

  - \( \text{PPV} = 0.67 \) (B5)
  - \( \text{PPV} = 0.06 \) (B4)
  - \( \text{PPV} = 0.09 \) (B3)

Thank you!

http://cracs.fc.up.pt/mammoclass

pedroferreira@dcc.fc.up.pt
nunofonseca@acm.org
ines@dcc.fc.up.pt
rwoods@gmail.com
eburnside@uwhealth.org
Appendices
Data distribution

- 348

<table>
<thead>
<tr>
<th></th>
<th>retro_density</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
<td>iso</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>malignant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>59 (70.2%)</td>
<td>59 (22.3%)</td>
<td>118</td>
<td>(33.9%)</td>
</tr>
<tr>
<td></td>
<td>benign</td>
<td>25 (29.8%)</td>
<td>205</td>
<td>(77.7%)</td>
</tr>
<tr>
<td>Total</td>
<td>84 (24.1%)</td>
<td>264 (75.9%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Data distribution

- **180**

<table>
<thead>
<tr>
<th>outcome_num</th>
<th>retro_density</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
<td>iso</td>
</tr>
<tr>
<td>malignant</td>
<td>42 (75.0%)</td>
<td>29 (23.4%)</td>
</tr>
<tr>
<td>benign</td>
<td>14 (25.0%)</td>
<td>95 (76.6%)</td>
</tr>
<tr>
<td>Total</td>
<td>56 (31.1%)</td>
<td>124 (68.9%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>outcome_num</th>
<th>density_num</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
<td>iso</td>
</tr>
<tr>
<td>malignant</td>
<td>51 (63.0%)</td>
<td>20 (20.2%)</td>
</tr>
<tr>
<td>benign</td>
<td>30 (37.0%)</td>
<td>79 (79.8%)</td>
</tr>
<tr>
<td>Total</td>
<td>81 (45.0%)</td>
<td>99 (55.0%)</td>
</tr>
</tbody>
</table>
Data distribution

- 168

<table>
<thead>
<tr>
<th>outcome_num</th>
<th>retro_density</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
<td>iso</td>
</tr>
<tr>
<td>malignant</td>
<td>17 (60.7%)</td>
<td>30 (21.4%)</td>
</tr>
<tr>
<td>benign</td>
<td>11 (39.3%)</td>
<td>110 (78.6%)</td>
</tr>
<tr>
<td>Total</td>
<td>28 (16.7%)</td>
<td>140 (83.3%)</td>
</tr>
</tbody>
</table>