An Early Warning System that combines Machine Learning and a Rule-Based Approach for the Prediction of Cancer Patients’ Unplanned Visits

H. F. Witschel*, E. Laurenzi*, S. Jüngling*, Y. Kadvany⁰ and A. Trojan⁰

*FHNW University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, CH-4600 Olten

⁰mobile Health AG, Falkenstrasse 21, CH-8008 Zürich

27 March 2023, Symposium AAAI-MAKE, San Francisco (CA)
The Burden of Cancer Worldwide

- Cancer is among the leading causes of death worldwide.
- Generally, cancer rates are highest in countries whose populations have the highest life expectancy, education level, and standard of living.
- By 2040, the number of new cancer cases per year is expected to rise to 29.5 million and the number of cancer-related deaths to 16.4 million.

Source: International Agency for Research on Cancer

Cancer Tomorrow | IARC - All Rights Reserved 2023 - Data version: 2020
Problem

- Oncologists are under pressure to visit an increasing number of patients!
  - Unplanned visits are problematic (from a Swiss research project).
    - Hard to dedicate adequate time to complex cancer cases
    - Not optimal cancer patient outcomes
  
- A possible approach: an early warning system to predict issues leading to unplanned visits.
- A ML-based approach is proven to be successful for clinical risk prediction (Bull et al. 2020).
  - Problems (Ginestra et al. 2019):
    - reasons for raising alerts are not understood by clinicians AND
    - generic lack of trust of pure ML approach
  - Genuinely human interpretable models can help vs. relying on try to explain black boxes (Rudin 2019).
Our Approach and Underlying Assumptions

- An interpretable Rule-based Machine Learning approach that predicts issues leading to unplanned visits.
  - a learning algorithm that produces directly interpretable rules.
  - rules can be interpreted, formulated and enhanced by experts based solely on an understanding of input-output associations.

- Assumptions:
  - Learned rules express conditions for the doctors to advise a visit
  - Doctors may not be able to externalize all rules
  - Doctors are able to judge the learned rules
    - checking and modifying ML-generated rules
    - extend the rule base with low effort and at the same time use their full medical knowledge to control which cases will lead to warnings or not.
Data Set

- Diary kept by cancer patients in the form of a mobile app where data is entered on a daily basis.
- Total of 16,670 diary entries.
- 266 patients (mostly suffering by breast cancer).
Attributes used to train our EWS

- Ground truth: unplanned visits of patients
  - learn situations that lead to unplanned visits.
  - 1 instance = diary entry = combination patient-day (i.e. training example)
  - Association of class attribute “unplanned visit” to each instance
  - For each actual unplanned visit, we consider a time frame of 3 days prior, where the respective class attribute gets assigned the value “yes”.
    - Time frame estimated by our oncologists.
  - Drugs and symptoms strengths are encoded
    - Patients received a guideline with def.
    - Developed by oncologists and
    - Based on a specific terminology (CTCAE)

- Free-text attributes
  - Diagnosis (entered by the doctor) AND Note (optionally entered by the patient)
  - String attributes were vectorised using TF/IDF weights -> prefixed with «diag_» and «note_», resp.

<table>
<thead>
<tr>
<th>Attribute(s)</th>
<th>number</th>
<th>description</th>
<th>type, value range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth year</td>
<td>1</td>
<td></td>
<td>numeric</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td></td>
<td>[male, female]</td>
</tr>
<tr>
<td>Primary tumor</td>
<td>1</td>
<td></td>
<td>[breast, gut, blood/lymph, lung, prostate]</td>
</tr>
<tr>
<td>Wellbeing</td>
<td>1</td>
<td>subjective wellbeing of patient</td>
<td>numeric [0...100]</td>
</tr>
<tr>
<td>Therapy form</td>
<td>1</td>
<td>frequency of treatment</td>
<td>[daily, weekly, bi-weekly, 3-weekly, 4-weekly]</td>
</tr>
<tr>
<td>Drugs</td>
<td>88</td>
<td>cancer and other drugs</td>
<td>numeric [1, nan]</td>
</tr>
<tr>
<td>Symptom strengths</td>
<td>52</td>
<td>strength of relevant symptoms</td>
<td>numeric [0...1, nan]</td>
</tr>
<tr>
<td>Diagnosis terms</td>
<td>246</td>
<td>Terms occurring in patient’s diagnosis</td>
<td>numeric (TF/IDF)</td>
</tr>
<tr>
<td>Note terms</td>
<td>311</td>
<td>Terms from patient notes</td>
<td>numeric (TF/IDF)</td>
</tr>
<tr>
<td>Unplanned visit</td>
<td>1</td>
<td>Class attribute</td>
<td>[yes, no]</td>
</tr>
</tbody>
</table>
Proposed division of labor between ML and expert (1/3)

1. The human expert states a set of rules \( A \).
   - workshop where the medical expert was invited to first look at a number of example cases where unplanned visits had happened. Next, the expert formulated some rules. E.g.,
   - Rule: diag_nikotine_abuse > 0) and (Chest_Pain > 0) \( \Rightarrow \) UnplannedVisit=yes
   - Problem category: Lung

2. We apply the rule learner to generate a set of rules \( B \). E.g.,
   - Rule: (Wellbeing <= 46) and (Endoxan >= 1) and (diag_lymphangiosis >= 5.296978) \( \Rightarrow \) UnplannedVisit=yes
   - None of the learned rules matched with the rules suggested by the experts.

3. Rules from both \( A \) and \( B \) are evaluated based on a cost matrix where false negatives (FNs) have higher cost than false positives (FPs, false alarms). Rules are ranked by cost.

<table>
<thead>
<tr>
<th>Rules</th>
<th>FN: situations that should lead to unplanned visits are not recognized!</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>diag_nikotine_abuse &gt; 0 and Chest_Pain &gt; 0 ( \Rightarrow ) UnplannedVisit=yes</td>
<td>214.39</td>
<td>37.63</td>
</tr>
<tr>
<td>Wellbeing &lt;= 46 and Endoxan &gt;= 1 and diag_lymphangiosis &gt;= 5.296978 ( \Rightarrow ) UnplannedVisit=yes</td>
<td>71.54</td>
<td>3.14</td>
</tr>
<tr>
<td>Wellbeing &lt;= 46 and Endoxan &gt;= 1 and diag_lymphangiosis &gt;= 5.296978 ( \Rightarrow ) UnplannedVisit=yes</td>
<td>71.54</td>
<td>3.14</td>
</tr>
<tr>
<td>Wellbeing &lt;= 46 and Endoxan &gt;= 1 and diag_lymphangiosis &gt;= 5.296978 ( \Rightarrow ) UnplannedVisit=yes</td>
<td>63.61</td>
<td>3.01</td>
</tr>
<tr>
<td>Wellbeing &lt;= 46 and Endoxan &gt;= 1 and diag_lymphangiosis &gt;= 5.296978 ( \Rightarrow ) UnplannedVisit=yes</td>
<td>63.61</td>
<td>3.01</td>
</tr>
<tr>
<td>Wellbeing &lt;= 46 and Endoxan &gt;= 1 and diag_lymphangiosis &gt;= 5.296978 ( \Rightarrow ) UnplannedVisit=yes</td>
<td>63.61</td>
<td>3.01</td>
</tr>
</tbody>
</table>
Performance of machine-learned rule set

- Our machine-learned rule set (cost ratio 1:10):
  - discovered 47 out of the 166 critical situations (28.3% recall)
  - generated 263 false alarms (15.2% precision).
- When we use a ratio of 1:20 for the cost-sensitive rule learner:
  - the model recognised 54 critical situations, i.e. 7 more than with the 1:10 model,
  - but at a cost of an extra 104 false alarms, i.e. overall 367 false positives

Confusion matrix (1:10)
Proposed division of labor between ML and expert (2/3)

4. ML rules are inspected by the human expert in the ranked order. The expert can suggest to drop a rule, but also to modify it, e.g., dropping or adding a condition. Modified rules will be evaluated and accepted if their cost on the test set is acceptable.

- Out of the 12 rules, 3 were accepted in their original version.
- Another 3 rules were rejected - symptoms were not critical or unclear interpretation.
- The remaining 6 rules were accepted with modifications
  - In 2 cases, these modifications were additional conditions (e.g., additional symptoms that would make a situation truly critical). The additional conditions involved new attributes related to the trend, i.e., it will require to construct new features that take the historical development of symptom strengths into account.
  - In the remaining 4 cases, we discovered new insights:
    - predict a situation that requires action (a different kind of alert), but not necessarily a visit. Not possible when working with black-box machine learning models!!!
Example of two learned and interpreted rules

- **Rule 1**: (diag_clipmarker >= 4.661854) and (Wellbeing <= 74) => UnplannedVisit=yes
- **Medical interpretation**: that breast cancer patients receive a neoadjuvant chemotherapy before a surgery, during which the tumor shrinks (which is why its position is marked with a “clip marker” – i.e., that term correlates with a specific treatment). The chemotherapy impacts the wellbeing negatively.
- **Action**: suggest a visit if two additional conditions apply:
  - a) the trend of wellbeing is negative over several days and,
  - b) when nausea or fatigue appear as accompanying symptoms.

*example of human recommended additional conditions in ML-discovered rules*

- **Rule 2**: (Wellbeing <= 46) and (Endoxan >= 1) and (diag_lymphangiosis >= 5.296978) => UnplannedVisit=yes
- **Medical interpretation**: situation of palliative care
- **Suggested actions**: instead of a visit, advise the home care service to intensify measures to alleviate the suffering and ensure a higher wellbeing. E.g., increase the dose of painkiller medication or reduce that of Endoxan

*example of human recommended new type of alert in ML-discovered rules*
Conclusion

- We have proven that a machine learning approach to the discovery of rules may provide value and that a corresponding model will be able to discover several critical situations.

- The utility of a machine-learned rule set will be limited because increasing its coverage (recall) is possible, but comes at the cost of lower precision, i.e. more false alarms.

- The analysis of rules by medical experts not only results in rule modifications and suggestions for feature engineering, but also in specific types of actions entailed by predictions that allow for more fine-grained recommendations to be made by the Early-Warning System.
Outlook

- Advanced feature engineering
  - Consider the history of well-being, symptom development over time, missing patient entries of previous days
- Knowledge engineering
  - Grouping symptoms, drugs and their connection meaningfully
    - including the resulting symptom categories as new features.
Thank you.
emanuele.laurenzi@fhnw.ch