Predicting relapse of schizophrenia

Petr NALEVKA

University of Economics, Prague, Czech Republic

Abstract. Schizophrenia is a serious mental disorder which tends to follow a relapsing course leaving a devastating effect on the patients [1]. This article proposes symbolic data-mining methods to predict early stages of schizophrenia relapse for timely intervention.

Keywords. relapse of schizophrenia, symbolic data-mining, Youden

1. Introduction

Relapse of schizophrenia is preceded by various prodromal symptoms [2]. Timely detection of such symptoms may help predict and prevent onset of the relapse [2,3]. This principle is used in the ITAREPS program [4], which collects patients' condition self-evaluations sent by SMS messages and automatically informs the outpatient psychiatrist in case patient's worsening exceeds an expert-defined threshold. Since 2006, in two consequent studies (mirror-design and double-blind) ITAREPS gathered a unique set of demographic and diagnostic data from over 400 patients.

2. Objectives

Employ data-mining methods to improve prediction performance of the ITAREPS program. Especially, lower the amount of false alerts and search for interesting patterns to improve knowledge of schizophrenia development in different patient types.

3. Methods

Symbolic data-mining methods are used to discover patterns which are successful predictors but also interpretable for the psychiatrists. Applicability of such methods has been proven on similar system failure prediction problems [5,6]. Patients' condition history is split into temporal windows either preceding relapse (positive) or not (negative). Different patterns are generated and their ability to distinguish between positive and negative windows is tested using the Youden's index (Sensitivity + Specificity - 1) [7,8] in order to minimize error. Patterns consist of high-level categorical features combined using logical connectives. Features describe

1 Corresponding author
2 In addition, evaluation SMS messages, are also sent by patient's carer – a selected family members.
development of patient's condition within windows such as the level, trend or stability of worsening for individual symptoms or groups of symptoms.

Patterns are evaluated for each patient group which is established based on demographic and diagnostic data. This helps to fight noise in data, lower errors and discover knowledge which is specific to different types of patients.

4. Results

In comparison to ITAREPS's Youden = 0.226 and specificity = 0.769 the described method performed at Youden = 0.324 and specificity = 0.873 in a leave-one-out prediction performance estimate (3 week temporal windows, positive windows in 2-6 weeks before relapse). The overall number of false alerts has been reduced from 1709 to 955. Table 1 shows an example of 3 best patterns discovered.

Table 1. Most successful patterns by training set Youden

<table>
<thead>
<tr>
<th>Youden</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.58</td>
<td>married, no cognitive symptoms, carer sympt.1 (high avg.), all carer symptoms (high avg.)</td>
</tr>
<tr>
<td>0.58</td>
<td>young, non-university, sympt.10 (high avg.), all carer symptoms (low std.deviation)</td>
</tr>
<tr>
<td>0.51</td>
<td>young, no disorganization symptoms, carer sympt.10 (high average)</td>
</tr>
</tbody>
</table>

5. Discussion and Conclusion

The presented results meet the outlined objectives. Prediction performance estimates show an improved Youden with a significant increase in specificity over the current ITAREPS. False alerts are reduced by 44% with 12% false alerts per SMS.

Moreover, table 1 presents an example of discovered background knowledge. Patients with absent cognitive and disorganisational symptoms seem better suited for relapse prediction. In addition evaluation done by carers seem a better predictor.

Acknowledgements. This project was supported by Grant Agency of Czech republic, GACR 201/08/0802.

References