

Title: Evaluation of risk factors for falls in the elderly based on a real data set using Bayesian networks

Abstract:

Falls are a significant problem for the elderly, with identification and evaluation of risk factors being essential to reduce fall rates. However, fall prevention requires a pedagogical and repeated approach, time, and expertise to target actionable risk factors accurately. This thesis aims to evaluate the risk factors for falls using a real data set and Bayesian networks. Using real data poses challenges, particularly in data preprocessing, which is time-consuming and requires expertise. Additionally, an application based on AI raises new challenges such as trustability, which depends on the interpretability and explainability of the results.

To address these challenges, this thesis proposes a knowledge model (Bayesian networks) that automatically evaluates the main actionable risk factors. The model is trained on a real data set combined with expert knowledge. Two iterations of the data preprocessing steps are presented and explained, including missing value imputation, variable selection, and the use of balancing techniques for imbalanced data. The first iteration included only the main variables to validate the process's feasibility, and the second iteration included as many variables as possible to improve the prediction and the process.

The model is compared with other well-known classifiers through different measures, including all or partial observation, and using or not balancing methods to manage the delicate question of imbalanced data. A Bayesian network is presented as a good solution, combining the quality of the results to evaluate the risk factors and the interpretability/explainability of the model from the expert's point of view.

The results show that predicting the presence or absence of risk factors for falls is a challenging task. While Bayesian Networks and other classifiers perform equivalently in terms of measures such as balanced accuracy and f1-score, the interest of Bayesian networks lies in their interpretability and the ability to use partial observations.

In summary, this thesis presents a contribution toward an application for fall prevention that facilitates automatic risk factor evaluation from partial observations of the patient, using a real data set and Bayesian networks. The proposed knowledge model (Bayesian networks) addresses the challenges of using real data and AI-based applications, respectively.

Keywords: Classification, Machine Learning, Problem of Falls, Fall Prevention, Bayesian Networks