Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 113 (2017) 571–576

Procedia Computer Science

www.elsevier.com/locate/procedia

International Workshop on Healthcare Interoperability and Pervasive Intelligent Systems (HiPIS 2017)

Predicting the need of Neonatal Resuscitation using Data Mining

Ana Morais[a], Hugo Peixoto[b], Cecília Coimbra[b], António Abelha[b] and José Machado[b]*

[a]*University of Minho, Campus Gualtar, Braga 4710, Portugal*
[b]*Algoritmi Research Center, University of Minho, Campus Gualtar, Braga 4710, Portugal*

Abstract

It is estimated that approximately 10% of newborns require some kind of assistance for breathing at birth. Aiming to prevent neonatal mortality, the goal behind this paper is to predict the need for neonatal resuscitation given some health conditions of both the newborn and the mother, and also the characteristics of the pregnancy and the delivery using Data Mining (DM) models induced with classification techniques. During the DM process, the CRISP-DM Methodology was followed and the WEKA software tool was used to induce the DM models. For some models, it was possible to achieve sensitivity results higher than 90% and specificity and accuracy results superior to 98%, which were considered to be satisfactory.

© 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the Conference Program Chairs.

Keywords: Data Mining; Decision Support Systems; CRISP-DM; Classification; Neonatal Resucitation

1. Introduction

Nowadays, information systems are essential for organizations since they provide useful information for decision-making processes by storing, processing and analysing large amounts of data[1]. DM is considered as the set of methods and techniques for exploring and analyzing large datasets in an automatic form with the aim of finding unknown or hidden rules, associations or patterns. The DM techniques can be classified as descriptive, which includes clustering techniques, or predictive, which includes classification and regression techniques. This paper is focused on classification techniques, in which the model or classifier predicts categorical labels[2].

* Corresponding author. Tel.: +351 253 604 430; fax: +351 253 604 471.
E-mail address: jmac@di.uminho.pt

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the Conference Program Chairs.
10.1016/j.procs.2017.08.287

Due to the large amount and the complex nature of data generated by transactions in healthcare environments, the interest of healthcare organizations on using DM has been increasing since it can greatly benefit all parties involved. Namely, the analysis provided by DM allows healthcare institutions to improve operating efficiency while maintaining a high level of quality of care. However, DM's applications in healthcare have limitations such as the often difficult accessibility to data, which can be due to the existence of raw inputs in different settings and systems. Thus, the data needs to be collected and integrated before it can be used in DM[3].

Decision Support Systems (DSS) can be defined as the class of computer-based information systems that support decision making activities[4]. In order to provide quick and reliable decision support, DSS require the automated analysis of data to find tendencies and extract knowledge, which can be performed by DM techniques[5]. When applied to healthcare, DSS are called Clinical DSS and offer support to clinicians at the various stages of the care process. They can also take over some routine tasks by warning clinicians of potential problems or providing suggestions for them to consider and provide clinicians or even patients with knowledge and person-specific information that can be intelligently filtered and presented at appropriate times. In addition, other advantages of the usage of Clinical DSS in medical practice include ensuring accurate and timely diagnoses for preventing diseases, lowering operating costs, improving efficiency and reducing patient inconvenience[6].

2. Background and Related Work

Pereira et al. (2015) proved the viability of using DM models to predict the most appropriate type of delivery considering the pregnancy characteristics of patients[7]. Namely, real data from the perinatal and maternal care unit of Centro Hospitalar of Oporto (CHP) was used and four different classification techniques were implemented: Decision Trees (DT), Generalized Linear Models (GLM), Support Vector Machines (SVM) and Naïve Bayes (NB). The best induced model acquired satisfactory results by achieving sensitivity values around 90% and was afterwards included in the Business Intelligence platform already employed in CHP[7].

Portela et. al (2015) tested the usage of DM classification techniques, namely GLM, SVM, DT and NB, to predict the probability of a patient to have a blood pressure critical event in the following hours by combining a set of patient data extracted in real-time from CHP. The achieved results demonstrated to be quite promising, with sensitivity values around 95% and the best induced model was afterward included in the INTCare ensemble engine, representing an important step in helping the prevention of possible cases of hypertension or hypotension[8].

Neonatal resuscitation is defined as the resuscitation of newborns with birth asphyxia and its goal is to re-establish adequate spontaneous respiration and cardiac output and to prevent neonatal mortality[9].It is estimated that approximately 10% of newborns require some degree of assistance to start breathing at birth. The risk factors for these occurrences include malpresentation of the fetus, gestational age lower than 35 weeks, low birth weight infants, increased or decreased maternal age, multiple birth pregnancy, among others[10]. An accurate evaluation of the risk factors can lead to the anticipation of resuscitation need which, in turn, allows adequate preparation of the necessary equipment and staff to perform neonatal resuscitation. In fact, how quickly and successfully the resuscitation is performed can be truly decisive for the infant's health, namely for avoiding hypoxic damage on the organs or even brain damage[11]. Thus, having a decision-support system that can accurately predict the need for resuscitation would be of great interest for both patients and health professionals. Indeed, obstetricians could know beforehand whether the newborn needs resuscitation and could perform all the necessary procedures right after the newborn's birth, improving the efficiency of the provided care and reducing medical errors.

3. Methodologies, Material and Methods

The data presented in this work was extracted from Electronic Health Records (EHR) and admission records from the obstetrics service of a Portuguese hospital. This data comprises the year of 2016 and has information about 3163 newborns, along with information about their mothers and the respective delivery episodes. During the Data Mining Process, the Cross Industry Standard Process for Data Mining (CRISP-DM) Methodology was followed, which is a hierarchical process model that divides the process of data mining into six phases: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation* and *Deployment*[12]. This

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات