Improved Boosting Algorithms by Pre-Pruning and Associative Rule Mining on Decision Trees for Predicting Obstructive Sleep Apnea

Article *in* Journal of Computational and Theoretical Nanoscience · November 2017

DOI: 10.1166/asl.2017.10335

CITATIONS
9

READS
147

3 authors, including:

Doreen Ying Ying Sim
University of Nottingham
16 PUBLICATIONS **122** CITATIONS

SEE PROFILE

Chee Siong Teh
University Malaysia Sarawak
53 PUBLICATIONS **297** CITATIONS

SEE PROFILE

Advanced Science Letters
Vol.23, No. 11, 11593–11598, 2017

Improved Boosting Algorithms by Pre-Pruning and Associative Rule Mining on Decision Trees for predicting Obstructive Sleep Apnea

Doreen Ying Ying Sim[1*], Chee Siong Teh[1], Ahmad Izuanuddin Ismail[2]

[1]Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak, Kuching, Sarawak, Malaysia
[2]Respiratory Medicine Unit, Department of Respiratory Medicine, UiTM Medical Specialist Centre, Faculty of Medicine, Universiti Teknologi MARA, Selangor, Malaysia

An improved Boosting algorithm, named as Boosted PARM-DT, was developed by pre-pruning techniques and Associative Rule Mining (ARM) on decision trees built from the clinical datasets** collected for Obstructive Sleep Apnea (OSA). The Pruned-Associative-Rule-Mined Decision Trees (PARM-DT) developed by adopting pre-pruning techniques on tree depth, minimum leaf and/or parent node size observations and maximum number of tree splits, based on Apriori and/or Adaptive Apriori (AA) frameworks, is boosted to achieve better predictive accuracies. The improved algorithms were implemented in OSA dataset and UCI online databases for comparisons. Better predictive accuracies were achieved in all the applied datasets/databases when comparing the classical algorithm, i.e. Boosted DT, with the improved one, i.e. Boosted PARM-DT.

Keywords: pre-pruning techniques, Associative Rule Mining, Apriori, Adaptive Apriori (AA), Boosted PARM-DT

1. INTRODUCTION

Obstructive Sleep Apnea (OSA), like some other diseases and medical illnesses, usually has an attribute or a set of attributes which can perfectly or almost perfectly confirm the medical diagnosis[2-4]. This attribute or set of attributes, however, usually has low support threshold(s). Since boosting algorithms, such as AdaBoost, are white-box methods, and this research has raw data** collection on OSA patients' records** with the characteristics of it which is fully known and understood, using pre-pruning techniques, Associative Rule Mining (ARM) and Apriori /Adaptive Apriori framework is a great advantage.

Sleep apnea affects both adults and children which can result in as many as around 30 breathless episodes per night. Untreated sleep apnea can cause death during sleep or can incur serious health problems such as diabetes, hypertension, stroke, and other cardiovascular diseases[2-4]. If a person has Obstructive Sleep Apnea (OSA), his or her tongue and throat muscles may become so relaxed and floppy during sleep that those muscles can cause a narrowing or even complete blockage of the airway(s)[2-5]. Narrowing or complete blockage of the airway(s) can be caused by the cephalometric anatomical abnormalities or morphological defects (in this case, we concentrated just on retrognathia, micrognathia and posterior pillar webbing) and/or other anatomical defects such as throat and/or tongue muscles flow back due to poor blood circulation incurring muscle floppiness and relaxation[2,4,5].

Table 1 shows the *minimum support* and *minimum confidence thresholds* as per stated i.e. (1) bilateral Tonsils' Size or **TS** (size ranges from 0 to 4, i.e. normal case to the worst case); (2) crowding of oropharynx, i.e. **MP** (Mallampati score ranges from 1 to 4); (3) Neck Circumference or **NC** (greater than or equal to 40cm); (4) Epworth Sleepiness Scale or **ESS** (i.e. ESS ranges from 0 to 24); (5) Morbid Obesity or **MO** (BMI greater than or equal to 40); (6) Posterior Pillar Webbing or **PPW**; (7) Retrognathia / Retro-positioned maxilla or **RN** (over-slung or jutting lower jaw); and (8) Micrognathia / receding lower jaw or receding chin or **MN** (short mentohyoid distance or inferiorly displaced hyoid bone).

This paper is organized as follows: Section 2 deploys ARM and pre-pruning techniques on decision trees of OSA dataset; in Section 3, improved algorithms, i.e. Boosted PARM-DT, are implemented. Experimental results of comparing the algorithms proposed with classical approaches are shown and analyzed in Section 4. Conclusions and discussions are summarized in Section 5.

*Email: dsdoreenyy@gmail.com **see Acknowledgments

 doi:10.1166/asl.2017.10335