Frequent Pattern Mining of Risk Factors Predicting Neonatal Seizures Outcomes

Ionela Maniu
Research and Telemedicine Center in Neurological Diseases in Children, Pediatric Clinical Hospital
Department of Mathematics and Computer Science, "Lucian Blaga" University
Bulevardul Victoriei 10, Sibiu 550024, Romania
Phone: 0269 216 062
ionela.maniu@yahoo.ro

George Maniu
Department of Mathematics and Computer Science, "Lucian Blaga" University
Bulevardul Victoriei 10, Sibiu 550024, Romania
Phone: 0269 216 062
george.maniu@ulbsibiu.ro

Gabriela Visa
Research and Telemedicine Center in Neurological Diseases in Children, Pediatric Clinical Hospital
gabiap@yahoo.com

Raluca Costea
Research and Telemedicine Center in Neurological Diseases in Children, Pediatric Clinical Hospital
ralucacostea2004@yahoo.com

Bogdan Neamtu
Research and Telemedicine Center in Neurological Diseases in Children, Pediatric Clinical Hospital
Departament of Computer Science and Electrical Engineering, Faculty of Engineering Lucian Blaga University Sibiu Romania
Preclinical Departament, Faculty of Medicine, Lucian Blaga University, Sibiu, Romania
Bulevardul Victoriei 10, Sibiu 550024, Romania
Phone: 0269 216 062
bogdan.neamtu@ulbsibiu.ro, bogdan.neamtu@pediatriesibiu.ro

Abstract

This study aims to present a possible approach to identify the most common combinations of possible risk factors for the outcomes following neonatal seizures. First, we extract important predictor variables from the published studies featuring aspects regarding neurological outcomes in the context of neonatal seizure. Then, we used association rules to build prediction models to determine associations of risk factors which are frequently identified in real-data / evidence based researches. A total of 15 studies and 14 variables were included to identify frequent patterns and generating association rules processe. We searched for accurate and valuable interrelationships between various risk factors. The FP-Growth algorithm generated different itemsets with the largest including four parameters: electroencephalography (EEG), seizures semiology (SO), aetiology (ET), birthweight (BW), with support 0.200. The insights regarding our results may help in creating evidence-based prevention programmes enhancing existing algorithms for diagnosis and treatment of neonatal seizure outcomes.

Keywords: Neonatal Seizures, Risk Factors, Association Rules, Frequent Patterns.

1. Introduction

Neonatal seizures seem to have paramount importance in forecasting longterm outcomes on neurological maturation. These paroxysmal episodes might manifest as subtle patterns (more than 50%), clonic or myoclonic (20-25%) and tonic (~5%). Consequently, it is crucial for a neonatologist to make an adequate differential diagnosis to exclude other essential pathologies and to intervene in managing this syndrome effectively. Different medical conditions could mimic neonatal seizures namely gastroesophageal reflux in Sandifer syndrome, palid reflex anoxic attacks, cyanotic breath holding spells, hyperplexia, nervous system infections, haemorrhage or mitochondrial disorders, epilepsy syndromes, cardiac arrhythmias (Cazan et. al., 2014; Cazan et. al., 2017).

Among the prognostic factors, aetiology ranks first followed by perinatal conditions, seizures patterns, different biomarkers from the medical investigations, therapy type (Costea et. al. 2017; Azhibekov et. al., 2015).

Current approaches even tackle the possibility to find suggestive biomarkers noninvasively in epileptic toddlers or neonates (Shahar, 2012; Nicolae et. al., 2016; Pitkänen, 2016; Choy et. al., 2014). However, the most advanced angles are used for abnormal outcomes prognosis considering data mining techniques and association algorithms through a list of risk factors.

Many existing models for predicting mortality and or neurodevelopmental outcomes following neonatal seizures have been described. Usually, the goal is to foresee these outcomes based on risk factors. Irrespective of the model design to manage the outcomes the results intend to help the clinicians and researchers but also parents in initiating and planning decisions related to procedures, treatments, follow-ups, assistances, comparisons between different clinical studies or populations, institutions.

However, it is difficult to choose risk factors. As can be seen from the reviewed studies, the initial number of included risk factors (10-20 or more than 50) and different types of factors (perinatal, aetiology, seizures, investigations, therapy) vary considerably between studies. There are also (expected) differences in the factors combinations which are selected by the algorithms as suggestive for high-risk outcomes. In the current study, we propose an approach to identify the most common combinations of possible risk factors published in the literature which could influence an abnormal outcome. Our solution is based on frequent pattern mining (FPM), applied on risk factors identified by different research papers, a data mining technique mainly used for market basket analysis (MBA or Affinity Analysis) to discover relations among vast amounts of business transaction records.

2. Methods

Study selection

We reviewed the literature searching for evidence based studies with a focus on the existing scoring systems to predict outcomes after neonatal seizures. Five scoring systems emerged (Ellison, 1981, 1986; Pisani et. al., 2009; Garfinkle, & Shevell, 2011; Salamon et. al., 2014; Hur & Chung, 2016). A detailed analysis of these research papers was already presented in a previous review (Maniu et. al., 2017). However, we also considered other articles presenting case studies on computational models developed to identify risk factors (Miller et. al., 2005; Ambalavanan et. al., 2006; Nunes, 2008; Pisani, 2012; Yildiz, 2012; Lai, 2013; Vargas, 2013; Anand, 2014; Shah, 2014; Pisani, 2016). In our previuos work (Maniu et. al., 2018) factor analysis was used to determine groups of main risk factors considering the articles mentioned above. A total of 15 studies were included in the current study analysis to identify the most common combinations of possible risk factors using this time frequent pattern mining and association rules techniques.

Variables

First, we extracted the neurological risk factors from the prediction models presented in each study in Excel format. Then, we coded each outcome variable binary with 1 if the

factor was considered as risk factor by the prediction model and 0 otherwise. A total of 14 variables were investigated. Eventually, we used RapidMiner implementation to build a model for frequent items/variables set mining and association rules generation. The process involved the following steps: (1) the Excel table with 14 binomial type attributes and 15 records (for 15 reviewed articles) was imported in RapidMiner, (2) FP-Growth operator was added to the model, (3) Create Association Rules operator was connected with the FP-Growth operator. In the case of FP-Growth operator, we considered 0.1 as the min_support value (to include even outliers in the analysis) while for Association rules operator we considered 0.5 as the min_confidence parameter.

Frequent pattern mining

Association rules mining (ARM) is a branch of data mining methods, including various algorithms that identify similarities between items/variables. It relies on techniques such as frequent itemset mining (horizontal, vertical, projected layout), sequential pattern(pattern-growth, apriori approaches), or structured pattern. Agrawal and Srikant proposed for the first time in 1994 a FPM algorithm for MBA named AIS, in the form of association rule mining. Many scalable algorithms followed, having different approaches regarding the form, passes, reduced data set, candidate itemsets problems used data structures or storage space. The goal was a rule set simplification and/or performance improvements (Liu et. al., 1999, Hu & Chen, 2006, Hunyadi, 2009, Hunyadi, 2011, Kiran 2011, & Inan 2013). Beside MBA, ARM techniques have been used in many areas such as medical data (Downs, 2000, Aswani Kumar, 2010, Chaves, 2013, & Inan 2013), Customer Relationship Management (CRM) (Chen, 2005), spatial and spatiotemporal data (Koperski, 1995; Shiwei, 2018).

FP-Growth algorithm(frequent pattern growth)

The table no.1 presents basic concepts of FPM technique.

Table 1. Frequent pattern mining: Basic concepts

Set of all items (ex: in a market basket data)	$I = \{i_1, i_2, \dots, i_m\}$
itemset / k-itemset	a collection of zero or more items / itemset containing k items
transaction	a subset of items chosen from *I*
set of all transactions	$T = \{t_1, t_2, \dots, t_n\}$
support count	the number of transactions that contain a particular itemset $\sigma(X) = \lvert \{t_i \lvert t_i \in T, X \subseteq t_i\} \rvert$
support	fraction of transactions in which an itemset occurs support (X) = $\frac{\sigma(X)}{n}$
frequent itemset	an itemset (ex. X) is called frequent if: support(X) > *minsup* where *minsup* – user-defined threshold

An association rule is an implication of the form $X \rightarrow Y$ where X and Y are sets of predictors ($X \subset I, X \neq \emptyset, Y \subset I, Y \neq \emptyset, X \cap Y = \emptyset$) representing the antecedent and the consequent part of the rule.

The most common measures to quantify the strength for generated rules are presented in table 2.

Table 2. Association rule evaluation metrics

Confidence Predictability $X \rightarrow Y$ $(c\%)$	A rule has confidence c in the dataset if c% of the dataset transactions that contain X also contain Y	confidence $(X \rightarrow Y)$ = $P(Y\|X) = \frac{\sigma(X \cup Y)}{\sigma(X)}$
	how frequently items in Y appear in transactions that contain X	
Support Prevalence $X \rightarrow Y$ $(c\%, s\%)$	A rule has support s in the dataset if s% of the dataset transactions contain X also contain Y	support $(X \rightarrow Y)$ = $P(X \cup Y) = \frac{\sigma(X \cup Y)}{n}$ where n = number of transactions
	how often a rule is applicable to a given data set	

ARM process identifies all the rules with the support and confidence equal or grather than coresponding support (minsup) and confidence thresholds (minconf). It aims to predict an item occurrence based on other items occurrences. This process is divided in two subtasks: frequent itemset generation and rule genaration. The first task objective is to generate all itemsets having the support >= minsup treshold. The second task objective is to generate high confidence rules using the frequent itemsets previously identified. There are many algorithms presented in literature for performing these tasks (as efficiently as possible).

FP-growth method is a frequent itemset mining, projected layout based method, proposed by Han et. al in 2000. The algorithm strategy involves a two steps approach: (1) build the FP-tree (a compressed form of the initial dataset), (2) extract frequent item sets (frequent patterns and corresponding support) from the FP-tree, using the recursively build conditional pattern base and conditional FP-tree (for each frequent item from the FP-tree) (Han 2000, 2004, 2006; Borgelt 2013). The FP-tree algorithm implementation is the following:

procedure FP growth(Tree, α)
(1) if Tree contains a single path P then
(2) for each combination (denoted as β) of the nodes in the path P
(3) generate pattern β ∪α with support_count = minimum support count of nodes in β;
(4) else for each a_i in the header of Tree {
(5) generate pattern $\beta = a_i \cup \alpha$ with support_count = a_i .support count;
(6) construct β's conditional pattern base and then β's conditional FP_tree $Tree_\beta$;
(7) if $Tree_\beta \neq \emptyset$ then
(8) call FP_growth($Tree_\beta$, β); }

3. Results

The 15 studies analysis covered a period from 1975 to 2015, representing neonatal seizure real-data based studies from different states/countries (California, Maryland, New York, Brazil, Italy, Taiwan, Colombia, India, UK, Canada, Slovenia, South Korea). The samples size varied from 55 to 403 cases. The type and number of risk factors included in the models differed in the analysed studies. Most of them considered a range between 10 and 20 factors, but there were papers with over 50 analysed factors (Ambalavanan, 2006). Multiple logistic regression model was the most widely used technique for risk factor identification. On average, a study has identified 4 risk factors (with mode values: 4 and 5) while the largest results had 6 risk factors.

In all this studies the generated model had presence/absence of post neonatal outcomes (epilepsy, developmental delay, mortality) as target variable and different combination of

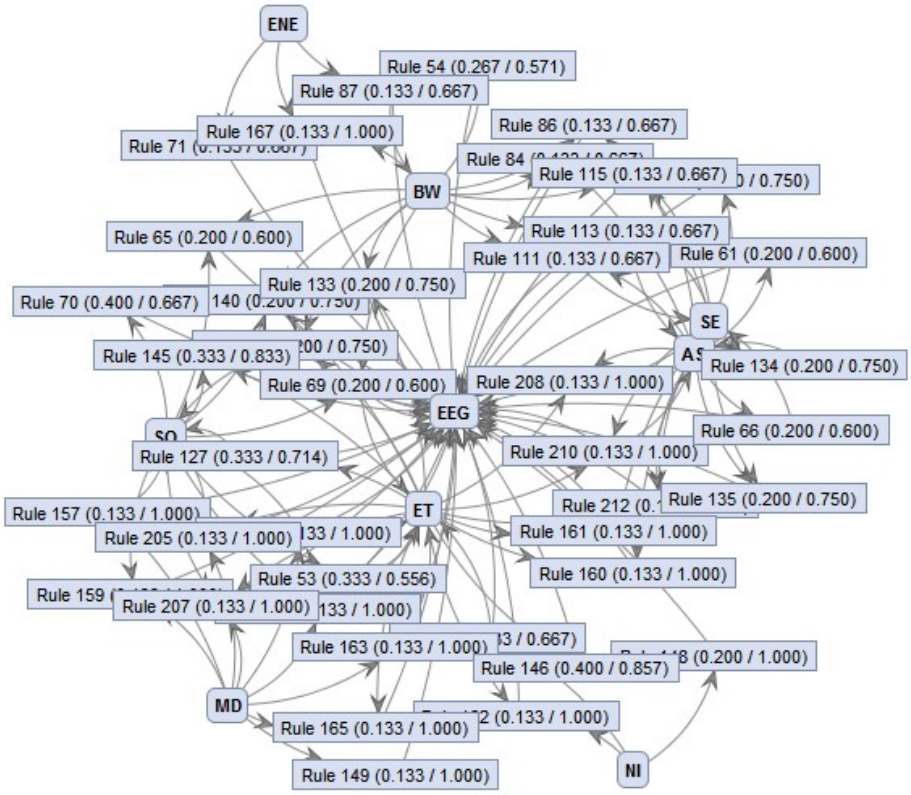
clinical variables (perinatal factors, etiology factors, seizures characteristics factors, investigations findings factors, therapy related factors) as predictors. Analyzing retrospectively the 15 studies, from the point of view of variables identified as risk factors, for each study we recorded the information for the following variables: birth weight (BW), gestational age (GA), mode of delivery (MD), Apgar score (AS), resuscitation (RE), seizure characteristics: type, onset, duration, semiology (SO), status epilepticus (SE), etiology (ET), early neurologic examination (ENE), EEG findings (EEG), cerebral ultrasound scan findings (UBS), neuroimaging (NI), treatment duration (TD), response to treatment (TR).
In the context of our data set, the set of all items I consisted of the 14 risk factors while the 15 reviewed studies represented the set of all transactions T.

The implementation in RapidMiner using FP-Growth operator and Association rules operator effected different patterns. The process results are presented in the table no 3 with a perspective including frequent pattern generation for 2-itemset, 3-itemset and 4-itemset (with support >=0.2 for 2 and 3 itemset). The 2-itemset was counting 38 patterns, the 3-itemset 29 while the 4-itemset 8. The largest itemset consisted of 4 items, the combination with the largest support being EEG, SO, ET, BW items(s0.200). The combination of EEG, SO, ET was also the most common combination of 3 items (s0.333).

Table 3. Frequent itemset generation

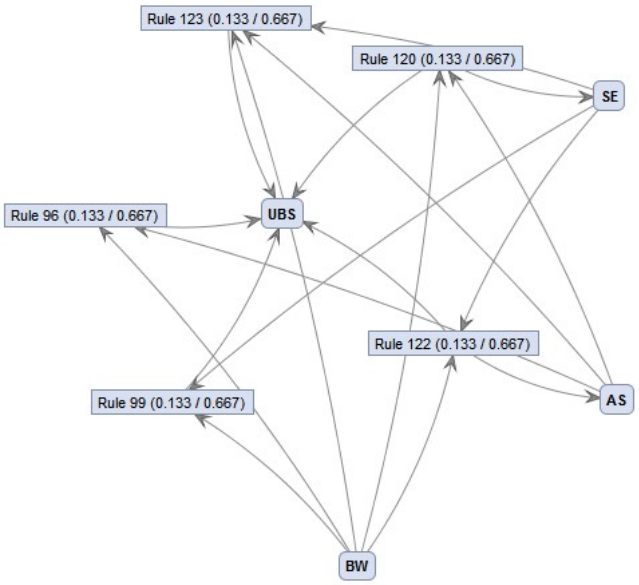
2-itemset			3-itemset				4-itemset				
0.400	EEG	SO	0.333	EEG	SO	ET	0.200	EEG	SO	ET	BW
0.400	EEG	ET	0.267	SO	ET	BW	0.133	EEG	SO	ET	MD
0.400	SO	ET	0.200	EEG	SO	BW	0.133	EEG	ET	AS	SE
0.333	SO	BW	0.200	EEG	ET	BW	0.133	EEG	BW	AS	SE
0.267	EEG	BW	0.200	EEG	AS	SE	0.133	SO	BW	AS	SE
0.267	ET	BW	0.200	BW	AS	SE	0.133	BW	AS	UBS	SE
0.267	AS	SE	0.133	EEG	SO	MD	0.133	AS	RE	TR	TD
0.200	EEG	AS	0.133	EEG	ET	AS					
0.200	EEG	SE	0.133	EEG	ET	SE					
0.200	EEG	NI	0.133	EEG	ET	NI					
0.200	SO	AS	0.133	EEG	ET	MD					
0.200	BW	AS	0.133	EEG	BW	AS					
0.200	BW	UBS	0.133	EEG	BW	SE					
0.200	BW	SE	0.133	EEG	BW	ENE					
0.200	AS	TR	0.133	SO	ET	MD					

The following figures describe the graphical rule visualisation (using KKLayout format) and associated rules evaluation metrics, in the case of filtering after variables EEG (figure no. 1) and USB (figure no. 2). Apgar score, birth weight and status epilepticus are factors that appear in both EEG findings and cerebral ultrasound scan findings (with same support (0.133) and confidence(0.667)). It is important to remind that the extracted rules are useful to understand correlations / associations but not causality. For example, the rule X → Y indicates that Y has a high probability to happen when X is present, but this does not mean that X is the cause of Y. Aditional information / domain experience is needed to make this claim.



No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Conviction
208	ET, AS	EEG, SE	0.133	1	1	-0.133	0.107	5	∞
210	ET, SE	EEG, AS	0.133	1	1	-0.133	0.107	5	∞
134	SE	EEG, AS	0.200	0.750	0.947	-0.333	0.147	3.750	3.200
111	BW, AS	EEG, SE	0.133	0.667	0.944	-0.267	0.093	3.333	2.400
113	BW, SE	EEG, AS	0.133	0.667	0.944	-0.267	0.093	3.333	2.400
66	AS	EEG, SE	0.200	0.600	0.900	-0.467	0.133	3	2
201	MD	EEG, SO, ET	0.133	1	1	-0.133	0.089	3	∞
87	ENE	EEG, BW	0.133	0.667	0.944	-0.267	0.080	2.500	2.200
157	MD	EEG, SO	0.133	1	1	-0.133	0.080	2.500	∞
163	MD	EEG, ET	0.133	1	1	-0.133	0.080	2.500	∞
203	SO, MD	EEG, ET	0.133	1	1	-0.133	0.080	2.500	∞
205	ET, MD	EEG, SO	0.133	1	1	-0.133	0.080	2.500	∞
139	ET, BW	EEG, SO	0.200	0.750	0.947	-0.333	0.093	1.875	2.400
127	ET	EEG, SO	0.333	0.714	0.909	-0.600	0.147	1.786	2.100
81	NI	EEG, ET	0.133	0.667	0.944	-0.267	0.053	1.667	1.800
69	SO, BW	EEG, ET	0.200	0.600	0.900	-0.467	0.067	1.500	1.500
148	NI	EEG	0.200	1	1	-0.200	0.067	1.500	∞
149	MD	EEG	0.133	1	1	-0.133	0.044	1.500	∞
159	SO, MD	EEG	0.133	1	1	-0.133	0.044	1.500	∞
160	ET, AS	EEG	0.133	1	1	-0.133	0.044	1.500	∞
161	ET, SE	EEG	0.133	1	1	-0.133	0.044	1.500	∞
162	ET, NI	EEG	0.133	1	1	-0.133	0.044	1.500	∞
165	ET, MD	EEG	0.133	1	1	-0.133	0.044	1.500	∞
167	BW, ENE	EEG	0.133	1	1	-0.133	0.044	1.500	∞
207	SO, ET, MD	EEG	0.133	1	1	-0.133	0.044	1.500	∞
212	ET, AS, SE	EEG	0.133	1	1	-0.133	0.044	1.500	∞
53	SO	EEG, ET	0.333	0.556	0.833	-0.867	0.093	1.389	1.350
146	ET	EEG	0.400	0.857	0.955	-0.533	0.089	1.286	2.333
145	SO, ET	EEG	0.333	0.833	0.952	-0.467	0.067	1.250	2
128	SE	EEG	0.200	0.750	0.947	-0.333	0.022	1.125	1.333

Figure 1. Graphical and tabelar rule vizualization, filtering after EEG factor



No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Conviction
120	BW, AS	UBS, SE	0.133	0.667	0.944	-0.267	0.107	5	2.600
122	BW, SE	AS, UBS	0.133	0.667	0.944	-0.267	0.107	5	2.600
96	BW, AS	UBS	0.133	0.667	0.944	-0.267	0.080	2.500	2.200
99	BW, SE	UBS	0.133	0.667	0.944	-0.267	0.080	2.500	2.200
123	BW, AS, SE	UBS	0.133	0.667	0.944	-0.267	0.080	2.500	2.200

Figure 2. Graphical and tabelar rule vizualization, filtering after UBS factor

4. Discussions and conclusions

This study underlines the applicability and suitability of association rules mining algorithms to determine the most common combinations of possible risk factors for outcomes following neonatal seizures. The insights regarding our approach may be useful along with clinical judgment and existing protocols for the epileptologists, neurophysiologists, pediatricians, neonatologists throughout the complex decision-making process in managing these cases. From a medical point of view more complex ET, lower AS or BW values, moderate or severe abnormal EEG (SO) or UBS, increase the odds for an abnormal outcome. In the proposed design we intended to provide a simulation of the association rules to identify combinations of risk factors based on the identified risk factors from 15 literature studies. Our study does not claim to have considered all possible risk factors for outcomes in the context of neonatal seizure, but it successfully bundled key factors such as ET, AS, with EEG, SO and UBS.

The analyzed articles presented models based mostly on the regression method, in particular logistic regression, but also other methods such as decision trees. To achieve the dataset of the present paper, the choice of variables did not take into account the prediction model, but only the fact that certain factors were considered/identified by the model (used in the study) as risk factors.

This retrospective study has not considered a hierarchy of the included studies in terms of quality assessment criteria, models reliability and/or validity. We were focused on proposed predicted models results, with an emphasis on identified risk factors. Also, some studies considered distinct prediction models for different types of outcome while other studies considered all outcome variants as one group using the prediction model for all types of outcomes as a group.

The considered risk factors are very different from each other, in the sense that they don’t have the same frequency. Some are common type of factors (BW, GA, AS) and they

have naturally more chance of being frequent than items that are much less likely to be considered (known as rare items problem, such as UBS).

In this context, our future work will focus on the following aspects: (1) including other studies in the analysis, (2) classifying studies according to quality assessments, (3) classifying studies according to prediction model, (4) classifying different individual outcome variants, (4) including, in the generation of association rules, of variables considered in the initial phase of the model, beside the risk factors variables identified by the model (5) considering other algorithms (MSApriori (Liu 1999), CFPGrowth (Hu 2006), CFPGrowth++ (Kiran 2011)), to deal with issues such as items skewed support distribution (rare items)).

Acknowledgment

This work has been conducted in the Pediatric Clinic Hospital Sibiu, within Research and Telemedicine Center in Neurological Diseases in Children - CEFORATEN project (ID 928 SMIS-CSNR 13605) financed by ANCSI with the grant number 432 / 21. 12. 2012 thru the Sectoral Operational Programme „Increase of Economic Competitiveness".

References

Agrawal R. and R. Srikant (1994) Fast algorithms for mining association rules. In VLDB'94, pp. 487-499

Ambalavanan N, Carlo WA, Shankaran S, Bann CM, Emrich SL, Higgins RD, Tyson JE, OShea TM, Laptook AR, Ehrenkranz RA, Donovan EF, Walsh MC, Goldberg RN, Das A; (2006) Predicting outcomes of neonates diagnosed with hypoxemic-ischemic encephalopathy. Pediatrics;118(5):2084-2093.

Anand V, Nair P. (2014) Neonatal seizures: Predictors of adverse outcome. J Pediatr Neurosci;9:97-9

Azhibekov T, Soleymani S, Lee BH, Noori S, Seri I. Hemodynamic monitoring of the critically ill neonate: An eye on the future. Seminars in Fetal & Neonatal Medicine. 2015;20(4):246–254.

Borgelt, C. (2010) An Implementation of the FP-growth Algorithm. Proceedings of the 1st International Workshop on Open Source Data Mining: Frequent Pattern Mining Implementations. 10.1145/1133905.1133907

Cazan C, Dobrota L, Neamtu M PO-0813 Sandifer Syndrome – A Challenge For The Paediatrician Archives of Disease in Childhood 2014; 99:A520. http://dx.doi.org/10.1136/archdischild-2014-307384.1447

Cazan C, Dobrota L, Neamtu M, Bogdan Neamtu, Gabriela Visa, Nicolae Bodrug, Viorel Istrate Elevated aminotransferase activity as a clue to muscular dystrophy Archives of Disease in Childhood 2017;102:A148. http://dx.doi.org/10.1136/archdischild-2017-313273.385

Chaves, R., Ramrez, J., Grriz, J.M. (2013) Integrating discretization and association rule-based classification for alzheimers disease diagnosis. Expert Syst. Appl. 40(5), 1571–1578

Chen R. S.,R. C. Wu and J. Y. Chen, (2005) "Data Mining Application in Customer Relationship Management Of Credit Card Business", In Proceedings of 29th Annual International Computer Software and Applications Conference (COMPSAC'05), Volume 2, pages 39-40

Choy M, Dube CM, Patterson K, Barnes SR, Maras P, Blood AB, Hasso AN, Obenaus A, Baram TZ. A novel, noninvasive, predictive epilepsy biomarker with clinical potential. Journal of Neuroscience 2014; 34: 8672-8684

Costea R., Neamtu M, Neamtu B. (2016) Febrile seizures – EEG study and risks factors for recurrence and for epilepsy. Eur J Pediatr (2016) 175: 1393. https://doi.org/10.1007/s00431-016-2785-8

Downs, S., Wallace, M. (2000) "Mining Association Rules from a Pediatric Primary Care Decision Support System." Proceeding of the 2000 Annual Symposium of American Medical Informatics Association, Los Angels, CA, USA, November 4-8, pp.200-204

Ellison PH, Largent JA, Bahr JP. (1981) A scoring system to predict outcome following neonatal seizures. J Pediatr [Internet].;99(3):455–9

Ellison PH, Horn JL, Franklin S, Jones MG. (1986) The results of checking a scoring system for neonatal seizures. Neuropediatrics [Internet].;17(3):152–7

Garfinkle J, Shevell MI. Prognostic factors and development of a scoring system for outcome of neonatal seizures in term infants. (2011) Eur J Paediatr Neurol [Internet]. Elsevier Ltd;15(3):222–9

Han J, Pei J, Yin Y (2000) Mining Frequent Patterns without Candidate Generation, In Proc. ACM SIGMOD Intl. Conference on Management of Data, 29(2), pages 1-12

Han J, Pei J, Yin Y, Mao R (2004) Mining frequent patterns without candidate generation: A frequent-pattern tree approach. In Data mining and knowledge discovery. 8(1): 53–87

Han J, Kamber M, Pei J (2006) Data mining: concepts and techniques, Morgan Kaufmann

Hu, YH, Chen, YL. (2006) Mining association rules with multiple minimum supports: a new mining algorithm and a support tuning mechanism. Decision Support Systems, 42(1):1–24.

Hunyadi D (2009) Improvements of Apriori Algorithms, First International Conference on Modelling and Development of Intelligent Systems – MDIS, October 22-25, Sibiu, Romania, pg. 106-114, "Lucian Blaga" University Publishing House, ISSN 2067-3965

Hunyadi D (2011) Performance Comparison of Apriori and FP-Growth Algorithms in Generating Association Rules, The 5th European Computing Conference (ECC' 11), Paris, France, April 28-30 2011, pg. 376-381,ISBN 978-960-474-297-4

Hunyadi D (2011) Performance algorithms in generating association rules, International Journal of Mathematics and Computers in Simulation, Issue 3, Volume 5, July 2011, pg. 282-289, ISSN 1998-0159

Hur J.Y, Chung ML. (2016) Risk Factor Analysis and Scoring System for Neurodevelopmental Outcomes after Neonatal Seizures. J Neonatal Biol, 5(4).

Inan, O., Uzer, M.S., Yılmaz, N. (2013) A new hybrid feature selection method based on association rules and PCA for detection of breast cancer. Int. J. Innov. Comput. Inf. Control 9(2), 727–729

Kiran, R. U, Reddy, PK. (2011) Novel techniques to reduce search space in multiple minimum supports-based frequent pattern mining algorithms. In: Proc. 14th Intern. Conf. Extending Database Technology, Uppsala, Sweden, 21-24 March, 11–20

Koperski K, Han J (1995) Discovery of spatial association rules in geographic information databases. In: Proceeding of the 1995 international symposium on large spatial databases (SSD'95), Portland, ME, pp 47–66

Kumar Aswani, Ch., Srinivas, S. (2010) Mining associations in health care data using formal concept analysis and singular value decomposition. J. Biol. Syst. 18(04), 787–807

Lai YH, Ho CS, Chiu NC, Tseng CF, Huang YL. (2013) Prognostic factors of developmental outcome in neonatal seizures in term infants. Pediatr Neonatol; 54: 166-72

Liu, B., Hsu, W., Ma, Y. (1999) Mining Association Rules with Multiple Minimum Supports. In: Proc. ACM SIGKDD Intern. Conf. Knowledge Discovery and Data Mining, San Diego, USA, 15-18 August, 337-341

Mahoney, L. G. S. (2016) *Assessment of the transitional circulation in late preterm and term neonates using non-invasive biomarkers: a longitudinal analysis and evaluation of repeatability* (Doctoral dissertation, University of Brighton)

Maniu I, Maniu G C, Bălan C, Neamțu B (2017) SCORING SYSTEMS FOR PREDICTING OUTCOMES OF NEONATAL SEIZURE, Acta Medica Transilvanica, vol. 22, ISSUE No 4, pp 48-50

Maniu I, Maniu G, Dospinescu C, Visa G. (2018) A Factor Analysis Model for Dimension Reduction of Outcome Factors in Neonatal Seizure Context, BRAIN. Broad Research in Artificial Intelligence and Neuroscience , Volume 9, Issue 2, ISSN 2067-8957, 95-103

Miller, S. P., Ramaswamy, V., Michelson, D., Barkovich, A. J., Holshouser, B., Wycliffe, N., … Ferriero, D. M. (2005). Patterns of brain injury in term neonatal encephalopathy. *J Pediatr*, *146*(4), 453–460. https://doi.org/10.1016/j.jpeds.2004.12.026

Neamtu B, Picu O, Martu Stefanache MA, Ibric Cioranu VS (2016) The Comparative Evaluation of Salivary Biomarkers (Calcium, Phosphate, Salivary pH) in Cario-resistance Versus Cario-activity Vasile Nicolae, REV.CHIM.(Bucharest), 67, No. 4 , 2016 http://www.revistadechimie.ro

Nunes ML, Martins MP, Barea BM, Wainberg RC, da Costa JC. (2008) Neurological outcome of newborns with neonatal seizures. Arq Neuropsiquiatr.;66:168–174

Pisani F, Sisti L, Seri S. (2009) A Scoring System for Early Prognostic Assessment After Neonatal Seizures. Pediatrics;124(4):e580–7.

Pisani F, Piccolo B, Cantalupo G, et al. (2012) Neonatal seizures and post-neonatal epilepsy: a seven-year follow-up study. Pediatr Res;72:186e93

Pisani F, Facini C, Pelosi A, Mazzotta S, Spagnoli C, Pavlidis E (2016) Neonatal seizures in preterm newborns: A predictive model of outcome. Eur J Paediatr Neurol 20:243-251

Pitkänen A, Löscher W, Vezzani A, et al. Biomarkers in epilepsy. *Lancet Neurol* 2016;15:843–856

Salamon S.A, Neubauer D, Petrovcic A, Paro-Panjan D. (2014) Risk factors and scoring system as a prognostic tool for epilepsy after neonatal seizures. Pediatr Neurol [Internet]. Elsevier Ltd;50(1):77–84

Shahar, E., Nagler R. (2012) Oxidative Stress in Children with Intractable Epilepsy Reflected by Salivary Glands Anti-Oxidant Profile , Neurology , 78 (1 Supplement) P01.073; ISSN:1526-632X

Shah, D.K., Wusthoff C.J., Clarke, P., (2014). Electrographic seizures are associated with brain injury in newborns undergoing therapeutic hypothermia, *Archives of Disease in Childhood - Fetal and Neonatal Edition* 2014;99:F219-F224

Shiwei Lian, Jinning Gao, Hongwei Li (2018) A Method of Mining Association Rules for Geographical Points of Interest, International Journal of Geo-Information

Vargas N.P, Sánchez E.F, Segura O.E (2014) Risk factors associated with a poor response to first-line treatment in neonatal seizures a case-control study done at Mercy Foundation Hospital, Revista Academia Nacional de Medicina, Vol 103

Yıldız EP, Tatlı B, Ekici B, et al. (2012) Evaluation of etiologic and prognostic factors in neonatal convulsions. Pediatr Neurol; 47: 186-92

Ionela Maniu is a lecturer professor Ph.D at Lucian Blaga University of Sibiu, Faculty of Sciences, Department of Mathematics and Computer Science and researcher in the Neurology Research Department within the Research and Telemedicine Center in Neurological Diseases in Children affiliated to the Pediatric Clinical Hospital from Sibiu Romania. Current research is focused on data mining, machine learning and artificial intelligence techniques, with special interest in medical data sets.