Unsupervised Learning Based Rule Generating System with Temporal Features Extractions Tuned for Tinnitus Retraining Therapy

Xin Zhang[1](✉), Xinyan Shi[1], and Pamela Thompson[2]

[1] University of North Carolina at Pembroke, Pembroke, NC 28372, USA
xin.zhang@uncp.edu
[2] Catawba College, Salisbury, NC 28144, USA

Abstract. The paper studies the effect of Tinnitus Retraining Therapy using an experimental approach. Each patient was assigned to either the control group where a treatment involves sound masking tests, or the experimental group in which a treatment involves counseling and/or sound masking tests. One additional feature with these medical records is that the treatment varies from one to another due to different frequencies of the visits. The irregular sampling rate and the sparsity in the treatment records present challenges to data analysis using temporal sound feature extraction design. The authors proposed an unsupervised learning system to uncover the relationship among treatment patterns, symptom improvements, and other factors in the data with association rules based on specially designed temporal features and sound descriptors.

Keywords: Association rules · Sparse data · Sound descriptors · Tinnitus retraining therapy

1 Introduction

Tinnitus, an extremely common condition, is typically defined as a perception of sound that is not related to an external sound source. Even though only a fraction of those who experience tinnitus are significantly disturbed [1–4], it is affecting about 17% of the general population around the world (44 million people in the USA) [5]. Being classified as a symptom and not a disease, tinnitus does require treatment as it can cause significant emotional and somatic distress and can significantly influence patients' quality of life, particularly when it becomes a chronic problem [6]. Tinnitus Retraining Therapy (TRT) [7], developed by Dr. Jastreboff since the mid '80s, is a treatment model with a high rate of success and is based on a novel neurophysical approach. This method uses a combination of sound masking with low level, broad-band noise, and counseling to achieve the habituation of tinnitus, i.e., the patient is no longer aware of their tinnitus, because even though the pitch and loudness of tinnitus might remain the same, tinnitus is no longer so intrusive.

Action rules learning explored the premium treatment operations to improve the efficiency of the therapy and potentially lower the treatment cost, where decision trees were

M. Ceci et al. (Eds.): ISMIS 2022, LNAI 13515, pp. 438–445, 2022.
https://doi.org/10.1007/978-3-031-16564-1_42

used to extract rules [8–12] and the results were promising. Decision trees have been widely used in data mining to efficiently extract rules for the purpose of classification; however, this type of learning algorithm is not flexible to learning potential interesting rules among all factors and features [13]. Association rules based on unsupervised learning have been successfully used in exploring the relationships among attributes without the limitation to preset any class labels to form the right-side part of the rules, the consequent. This characteristic will allow more potentially interesting rules to be uncovered [15]. The authors proposed to use unsupervised learning with association rules to explore more interesting rules to learn essential patterns in the data relative to tinnitus and its related symptoms.

2 Data Preprocessing

Tinnitus Retraining Therapy combines medical evaluation, counseling, and sound therapy to successfully treat a majority of patients. Based on a questionnaire from the patient as well as an audiological test, a preliminary medical evaluation of patients is required before beginning TRT. Sensitive information containing privacy from the medical evaluation were removed before the data analysis stage. However, some medical information, such as a list of medications the patient may take and other conditions that might be present, such as diabetes, was included in the tinnitus dataset since these information is relevant to the research of tinnitus.

Tinnitus is a symptom of many pathologies – even in one patient, several different types of tinnitus might coexist. For example, tinnitus is accompanied by hyperacusis in about 40% of the cases. Hyperacusis is a decreased tolerance of sound and can be a serious problem. Some patients experience hyperacusis without tinnitus. Tinnitus Retraining Therapy can restore totally or partially the normal level of sensitivity to sound. Patient categorization (see the table below) is performed after the completion of the medical evaluation, which is a structured interview guided by special forms and audiological evaluation. Patients were classified into the following categories in the initial visit based on their symptoms as shown in the following table [6].

Besides the above patient categories, a clustering based on clinic visit pattern had been performed. We postulate that the visit frequency and the number of total visits are important factors to the success of the therapy according to the clinic doctors. Therefore, we used a clustering method, where records were grouped by similar visiting patterns, while the visiting history of each patient is discretized into durations, starting from the date of his/her initial visit date in terms of weeks, and transformed as vectors of the same length for grouping. Assuming two patients denoted by p, q, patient p visits are represented by a vector $v_p = [v_1, v_2, \ldots, v_n]$ whereas vector $w_q = [w_1, w_2, \ldots, w_m]$ represents visits of patient q where the number of total visits of patient p is n and the number of total visits of patient q is m. The distance $\rho\,(p, q)$ between p, q and the distance $\rho\,(q, p)$ between q, p is defined as

$$\rho(q,p) = \frac{\sum_{i=1}^{k} \min(v_i - w_i, v_i - w_{i+1}, v_i - w_{i+2})}{k}, k = \min(m,n), \quad (1)$$

where w_{i+1} will be set to 0 when $i + 1 > m$ so on and so forth. Also, $v_i - 0$ would be bigger than any other distance differences and therefore will be ignored. This algorithm quickly

Table 1. Categories of patients with Tinnitus and Hyperacusis.

Type	Impact on life	Tinnitus	Subjective hearing loss	Hyperacusis	Prolonged sound induced exacerbation	Treatment
0	Low	Present	–	–	–	Abbreviated counseling
1	High	Present	–	–	–	Sound generators set at mixing point
2	High	Present	Present	–	–	Hearing aid with stress on enrichment of the auditory background
3	High	Irrelevant	Irrelevant	Present	–	Sound generators set above threshold of hearing
4	High	Irrelevant	Irrelevant	Present	Present	Sound generators set at the threshold

drops the less matching dates of a visit and selected pairs of visits even of different index number.

A threshold of an empirical value was applied to remove those records which have large distance values to form a tolerance class, so that all members of a patient cluster were ensured to have similar visiting patterns. Then visit-based temporal features were computed for all group members. Constructing tolerance classes helps to identify the right groups of patients for which useful temporal features can be built. By increasing the threshold value, larger classes would be produced for the process of knowledge extraction, but the information included in temporal features would be less accurate. On the other hand, if the threshold value is too small, the size of tolerance classes might be also too small to get any useful information through the knowledge extraction process. Since the original database was small, the authors only produced a small number of classes so that a good number of the patient records were kept in several classes. Additionally, visits except for those producing the minimum distances were kept. This means, after the data preprocessing, all the patients have the same total number of visits with similar visiting frequency in the dataset.

Important to the equation is the constraint that recovery of any patient with only one visit cannot be evaluated. Due to the intuition of the process, it is reasonable to assume that one visit would not be enough to fix the problem and that it was due to an interruption with some unknown cause. Therefore, such records have been removed from dataset for the rest of the experiments.

During a period of medical treatment, a doctor may change the treatment from one category to another based on the recovery specifics of the patients and the symptoms occurred during treatment. Additionally, the category of patient may change over time (e.g., the patient category may be changed from 3 to 1 if hyperacusis is disappearing). Other typical categorical features which may change over time in the dataset include sound-instrument types as well as visiting frequencies. At this stage, there were two types of data: one is numerical, such as scores for emotions, functions, and catastrophes related to the tinnitus problems; the other is categorical, such as instruments used in the therapy and patient categories. In terms of stability, we categorize the data into two types: Stable and Flexible [14]. Stable, compared with Flexible, is defined as an attribute that should maintain the same value over time throughout most of the records.

3 Temporal Feature Extraction

We employ temporal feature extraction method in order to study the effect of treatment duration on patients' recovery, which are evaluated based on three areas: functional, emotional, and catastrophic. We examine both the individual score and the total scores of them to evaluate whether the improvement has been made in a short or a long duration of treatment. Specifically, temporal feature extraction involves segmenting the visits into vectors based on the duration of patients' visit and use the recovery speed for each duration vector to evaluate the performance of different durations.

$$A_i(p) = \frac{A_k^i - A_{k-1}^i}{W_k - W_{k-1}}, k = 2, \ldots n, p \in m_j, j = 1, ..t \tag{2}$$

$$B_i(p) = \frac{A_n^i - A_1^i}{W_n - W_1}, p \in m_j, j = 1, ..t \tag{3}$$

$$C_i(p) = A_k^i - A_{k-1}^i, k = 2, \ldots n, p \in m_j, j = 1, ..t \tag{4}$$

$$D_i(p) = A_n^i - A_1^i, p \in m_j, j = 1, \ ..t, \tag{5}$$

where $A_i\ (p)$, representing the derivative of a segment of a visit vector, is the recovery speed of the i_{th} type for patient p of a group m_j, W is the number of weeks from the initial visit till the k_{th} visit, and n is the total visit; $B_i\ (p)$, representing the derivative of the whole visit vector, is the overall recovery speed of a patient; $C_i\ (p)$ represents the difference of the scores of a segment of a visit vector; $D_i\ (p)$ represents the difference of the scores of the whole vector.

The sounds pieces used in the treatment are not recorded completely as a full piece of acoustic recording; instead, only their loudness level, instrument type, and the main fundamental frequencies are documented in each visit. Therefore, complex sound descriptors for large sampling rate about sound vibrations are not applicable in this case. Moreover, to evaluate individual sound descriptors of all the visits for a patient, the median values were included in a larger dataset into different experiments in this research. The next section describes the details of the experiment design.

4 Experiment Design and Results

Two types of dataset were used in the experiments: the control group in which patients had similar visiting patterns, and the experimental group in which patients received sound masking treatment. The original dataset included 744 patients, whose total visits varies from 1 to over 15. We removed 203 patient records since they only have one visit. In the control group, each sub-dataset had 37 features, while the experiment group had 62 features. Altogether, nineteen clusters were generated, among which five clusters contained patients with four transformed visits, while fourteen clusters contained patients with three transformed visits. The temporal features showed that the majority of the patients had the best recovery rate in their first week.

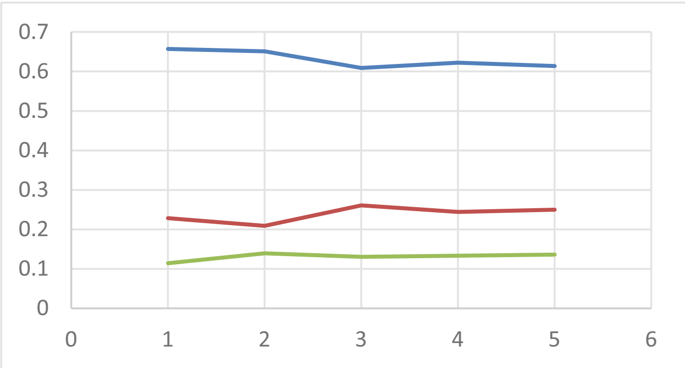

Fig. 1. Recovery rate comparison in the three-week durations of the five clusters with four transformed visits (blue – first week, red – second week, green – third week).

The above figure shows that, among those patients who took four visits in the three week-durations of TRT, about 60%–65% of them had the best recovery results in the first week (blue line). The red line shows the percentage of patients (about 20%–27%), whose best recovery occurred in the second week. The green line shows the percentage of those (about 11%–15%), whose best recovery occurred in the third week. Nevertheless, the results did not indicate that a patient should stop after the first week of treatment.

Figure 2 shows that among patients, who took three visits in a two-week duration of TRT, about 58%–65% of them have the best recovery in the first week for each of the fourteen clusters as shown in the blue line in the figure below.

We use five clinical categories as listed in Table 1. Among 212 patients, 33 were in category 0, 75 were in category 1, 50 were in category 2, 45 were in category 3, 9 were in category 4. The figure below shows that patients of the clinic category 0 had the best average recovery speed, while patients of category 4 had the least. The original dataset is very unbalanced, where patients of category 4 were less than 4% (26 out of 735 patients); therefore, patterns found related to this category may need more data to support. The blue marks represent individual patient recovery speed of each category, while the red line denotes the average speed of each category.

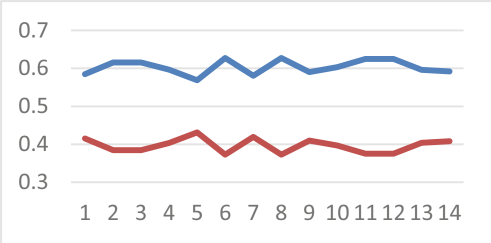

Fig. 2. Recovery rate comparison in two-week durations of the 14 clusters with three transformed visits (blue – first week, red – second week)

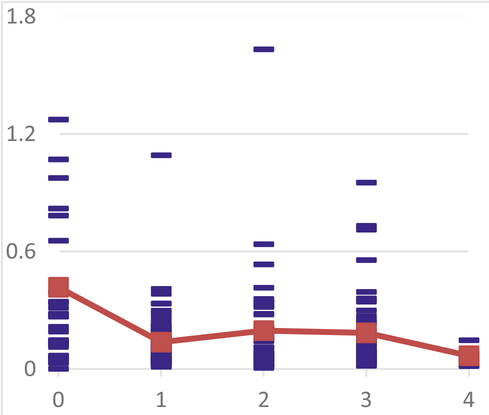

Fig. 3. Overall recovery speeds among the five categories of patients.

Weka 3.8.6 was used to generate association rules by the Apriori algorithm. Similar rules from multiple clusters were observed, while only representative rules were listed and compared in this paper, since the total number of association rules were very large.

LR12=120 ˆ LL12=120== > class=y (confidence 88%, lift 1.01, support 31%) (1)

LL8=120== > class=recovery (confidence 88%, lift 1.01, support 35%) (2)

LR8=120== > class=recovery (confidence 88%, lift 1.01, support 34%) (3)

R1=(3–16] ˆ R2=(7.5–20]== > class=y (confidence 91%, lift 1.05, support 22%) (4)

p2=T ˆ fu1=A== > class=y (confidence 91%, lift 1.06, support 62%) (5)

p1=T ˆ fu1=A== > class=y (confidence 97%, lift 1.13, support 63%) (6)

fu1=A== > class=y (confidence 89%, lift 1.04, support 67%) (7)

fu2=T== > class=y (confidence 90%, lift 1.04, support 37%) (8)

C1=0== > C2=0 (confidence 29%, lift 1.02, support 43%) (9)

C1=1== > C2=1 (confidence 94%, lift 3.40, support 34%) (10)

C1=2== > C2=2 (confidence 90%, lift 4.56, support 28%) (11)

C1=3== > C2=3 (confidence 86%, lift 4.24, support 17%) (12)

C1=4== > C2=4 (confidence 81%, lift 20.52, support 5%) (13)

Analyzing the audiology dataset, we changed the data type of the sound loudness values of different levels into nominal data type and found that recovery was usually associated with the Loudness Discomfort Level (LDL) of levels 4, 6, 8, and 12 of the value of 120 dB by thousands of rules. Therefore, from the resultant rules, we observed that LDL 4, 6, 8, and 12 may be set to the value of 120 dB to achieve positive results. Rules (1)–(3) as shown above are three examples of such rules. The authors also observed that the rules were consistent with the clinic operations where the lower the LDL level, the lower its value range when the dataset about the audiology were discretized into 10 bins per sound loudness column. Because 85% of the records in this dataset were related to recovery, while only 15% were related to no recovery, rules of confidence over 85% are useful to improve the recovery rate. For example, rule (4) indicates that when R1 is in the range of 3–16 and R2 is in the range of 7.5–20, they can be used together in the same clinic visit to improve recovery rate, where R1 and R2 are the loudness values of two types of clinically used noises.

Another example, rules (5)–(8) came from a cluster of 52 patients, where the total recovery rate was 79%. It implies that a patient with tinnitus needs at least two following-up treatments with both audiology and counseling, where in between such treatments a telephone visit may be used. Sometimes, a patient's category may vary after a period of treatment. The rules (9)–(13) about the stability of the patient categories indicate that, among the 551 patients who have more than 1 visit, the patient categories 1 and 2 are relatively stable, while category 0 is the least stable and tends to change to other categories.

5 Conclusion and Future Work

Tinnitus is not a well-defined disease. The authors developed an unsupervised learning system to extract interesting rules among the medical records of the TRT process to uncover the essentials of tinnitus and its related symptoms. The authors discovered that the 1st week treatment achieves the best recovery rate for the majority of the patients in the dataset though the recovery continued through the rest of the treatment process. The system successfully identified treatment patterns to improve the recovery rate. Action rules may be uncovered among the vast amounts of the rules from the association rules learning to identify how changes in treatment affects the recovery rate.

References

1. McFadden, D.: Tinnitus: facts, theories, and treatments. National Academy Press, Washington (DC) (1982)
2. Coles, R.R.A.: Epidemiology, aetiology and classification. In: Reich, G.E, Vernon, J.A. (eds.) Proceedings of the Fifth International Tinnitus Seminar, 1995, Portland, Or., U.S.A. Portland, Or. :American Tinnitus Association, pp. 25–30 (1996)
3. Pilgramm, M., et al.: Tinnitus in the federal Republic of Germany: a representative epidemiological study. In: Hazell, J.W.P. (ed.) Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK, London. Tinnitus and Hyperacusis Centre, pp. 7–64 (1999)
4. Davis, A., El Refaie, A.: Epidemiology of tinnitus. In: Tyler, R. (ed.) Tinnitus handbook. San Diego, Singular, Thomson Learning, pp. 1–23 (2000)

5. Tinnitus and Hyperacusis Center Homepage. https://www.tinnitus-pjj.com/. Accessed 01 May 2021
6. Snow, J., Ballenger, J.: Ballenger's Otorhinolaryngology Head and Neck Surgery, 16th edn. BC Decker Inc, Hamilton, Ontario, Canada (2003)
7. Jastreboff, P.J.: Phantom auditory perception (tinnitus): mechanisms of generation and perception. Neurosci. Res. **8**(4), 221–254 (1990)
8. Zhang, X., Ras, Z.W., Jastreboff, P.J., Thompson, P.L.: From tinnitus data to action rules and tinnitus treatment. In: 2010 IEEE International Conference on Granular Computing (GrC), pp. 620–625. IEEE (2010)
9. Thompson, P.: Mining for knowledge to build decision support system for diagnosis and treatment of tinnitus. In: Unc Charlotte Electronic Theses And Dissertations (2011)
10. Tarnowska, K.A., Ras, Z.W., Jastreboff, P.J.: Decision Support System for Diagnosis and Treatment of Hearing Disorders. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-51463-5
11. Tarnowska, K.A., Ras, Z.W., Jastreboff, P.J.: Mining for actionable knowledge in Tinnitus datasets. In: Wang, G., Skowron, A., Yao, Y., Ślęzak, D., Polkowski, L. (eds.) Thriving Rough Sets. Studies in Computational Intelligence, vol. 708, pp. 367–395. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-54966-8_18
12. Tarnowska, K.A.: Emotion-based music recommender system for Tinnitus patients (EMOTIN). In: Ras, Z.W., Wieczorkowska, A., Tsumoto, S. (eds.) Recommender Systems for Medicine and Music. Studies in Computational Intelligence, vol. 946, pp. 197–221. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66450-3_13
13. Tan, P., Steinbach, M., Karpatne, A., Kumar, V.: Introduction to Data Mining, 2nd edition, Pearson Education (2019)
14. Ras, Z.W., Wieczorkowska, A.: Action-rules: how to increase profit of a company. In: Zighed, D.A., Komorowski, J., Żytkow, J. (eds.) Principles of Data Mining and Knowledge Discovery. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), vol. 1910, pp. 587–592. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-45372-5_70
15. Martin, D., Rosete, A., Alcalá-Fdez, J., Herrera, F.: A new multi-objective evolutionary algorithm for mining a reduced set of interesting positive and negative quantitative association rules. IEEE Trans. Evol. Comput. **18** (1), 54–69 (2014)