



Correlation between chord guitar and song year era using apriori algorithm

Rifki Fahrial Zainal ¹, Arif Arizal ²

^{1,2} Informatics Engineering, Universitas Bhayangkara Surabaya, Surabaya, Indonesia

Article Info

Article history:

Received Sept 06, 2023

Revised Sept 14, 2023

Accepted Sept 16, 2023

Keywords:

Apriori Algorithm;

Data Mining;

Guitar Chord;

Music;

Song Year Era.

ABSTRACT

There are there are more and more varieties of music, especially if you made the music with guitar. There are also many and variety key combinations for the guitar chord. This is usually taken into consideration by beginners who are just learning to play the guitar to make their own music. Music also must made by feeling for the tone itself, and everyone has a different feel. Beginners usually see references from existing songs to made their own music. They usually make it in any key or chord that they want. But they also need inspiration or suggestions for the next key or chord to use from the key or chord they specified. In this study, we propose a way for beginners to find a combination of chords that can be used to make their own first music. From the results of this study, it was found that of the many songs in the database that were released in the 1990s to 2000s, most of them used three combinations of chords Am, Em and G. These three combinations were the combinations that most often appeared in songs. These three keys can become the user's favourites to be used as a basis for making songs or just to find inspiration from songs from the 1990s to 2000s.

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.



Corresponding Author:

Rifki Fahrial Zainal,

Informatics Engineering,

Universitas Bhayangkara Surabaya,

Jl. Ahmad Yani No.114, Ketintang, Kec. Gayungan, Surabaya, Jawa Timur 60231, Indonesia.

Email: rifki@ubhara.ac.id

Introduction

Music is the result of human behaviour among many other human cultures which is interesting because music plays a very large role in various cultures. Music is a means of fulfilling human needs in the desire for art and creativity. From a social point of view, making music into a song can be said to be a mirror of the social order that existed in society when the song was composed. Songs also can influence listeners in to doing something that proposed inside the lyrics of the music. In general, Indonesian singers, music groups or musicians really like to produce songs with their own style.

Music according to Phetorant (Kharisma, 2021) is a combination or collection of frequencies that can be heard. Referring to the definition above, the scope of music is very broad. The function of music is also very diverse in life. As entertainment, art music, economy, industry (be it tourism, advertising, television), education, film music and others.

Songs have their respective difficulty levels to play or to rearrange. Making a song also has its own level of difficulty. It can depend on the genre, lyrical vocabulary, even the feel of the songwriter. Making songs, of course, must understand what is called tone. Some of the basic notes in the song need to be learned so that it can determine what the song will be made of. To express music, tools or

instruments are needed to support the music itself. There are rhythmic instruments, melodic, and a combination of the two. One of the instruments used to express music is the guitar (Gredy Aprianno Penciptaan Seni Musik Bara & dan Pengkajian, 2020).

The difficulty in making songs is having to memorize guitar chords or existing notes. For some people who have often or know various kinds of tones, it will be easier to make songs. All that is left is to adjust the genre, feel, and how creative that person is. But for beginner who just starting to learn songs, this would be a significant difficulty. Because they do not understand the feel and rotation of the chords on the guitar. There are a lot of notes that are interconnected and it would be a beautiful tune if played in sequence. But again, we must understand what tone is used and must be played in sequence. Being able to memorize some basic chords is also a very valuable capital for making a song.

A very common thing experienced by someone who can or is just learning to play the guitar is wanting to know the chords of the song he wants to play. But having basic chords sometimes is not enough. If the guitar player has been learning and playing guitar for a long time, the feel of finding the tone of a song might be easy to get. So just by hearing his admiration, the guitar player already knows what chords to play. If the guitar player is just learning to play the guitar, he will need more guitar chord arrangements from other people.

To gain access for those guitar chord arrangements, we can learn the pattern from existing song. If we can learn the pattern, we can gain the information pattern combination item set from guitar chord arrangements from existing song. Mining algorithm of association rules is one of the most popular methods of emerge from large databases and get association rules so that new knowledge is discovered (Alfianzah et al., 2020; Irfiani, 2019; Sari et al., 2020).

Apriori algorithm is one of the most influential algorithms in the field of association rule mining. At present, many algorithms are improved or extended based on Apriori algorithm. Apriori algorithm uses the iterative method of layer-by-layer search to find frequent itemsets in transaction database (Henry et al., 2022; Le, 2022).

Association rules or affinity analysis are concerned with the study of "what goes with what". An example could be in the form of a transaction study at a cafe, for example someone who buys coffee also buys toast (Ilmiah & Grafis, n.d.; Samuel et al., 2022). In this case it means coffee along with toast. Because it originally came from a study of customer transaction databases to determine the habits of a product purchased with what product, association rules are also often called market basket analysis.

Association rules want to provide this information in the form of an "if-then" or "if-then" relationship. This rule is calculated from probabilistic data. Association analysis is also known as a data mining method that forms the basis of various other data mining methods. One stage of association analysis called frequent pattern mining has attracted the attention of many researchers to produce efficient algorithms.

The importance of an associative rule can be determined by two parameters, support (support value), which is the percentage of combinations of these items. in the database and confidence (certainty value), namely the strength of the relationship between items in associative rules (Idris et al., 2022; Rumui et al., 2022). Association analysis is defined as a process to find all associative rules that meet the minimum requirements for support (minimum support) and minimum requirements for confidence (minimum confidence) (Fenny Syafariani, 2022; Pan, 2021).

Based on this algorithm, this study shows that this algorithm is the most efficient way to find frequent items, which is the guitar chord arrangements needed based on the era of the existing song. In this paper, an analysis of data associations will be made with the apriori algorithm for decision support systems for guitar chord correlation applications with the era of song releases using PHP programming and data collection with MySQL which can be a solution for all those who are new to the field of music and are eager to create or compile tone by paying attention to the music that existed in the past in order to be inspired and succeed in making a work in the form of a song.

Method

There are several algorithms that have been developed regarding association rules, but there is one classic algorithm that is often used, namely the apriori algorithm. The basic idea of this algorithm is to

develop frequent itemsets. By using one item and recursively expanding the frequent itemset with two items, three items and so on up to a frequent itemset of all sizes.

To develop frequent sets with two items, you can use frequent set items. The reason is that if the set of one item does not exceed the minimum support, then any larger itemset size will not exceed that minimum support. In general, developing sets with fc-items uses the frequent set with $k-1$ items developed in the previous step. Each step requires one check of the entire contents of the database. In the association there are terms antecedent and consequent, antecedent to represent the "if" and consequent to represent the "then". In this analysis, the antecedent and consequent are a group of items that have no relationship together.

Of the large number of rules that may be developed, it is necessary to have rules that are strong enough for the level of dependence between items in the antecedent and consequent. To measure the strength of this association rule, support and confidence measures are used. Support is the ratio between the number of transactions that contain the antecedent and consequent and the number of transactions. Confidence is the ratio between the number of transactions that include all items in the antecedent and consequent to the number of transactions that include all items in the antecedent.

The first step of the a priori algorithm is, the support of each item is calculated by scanning the database (Dewi Sri Mulyani et al., 2021; Yustiana et al., 2021). After support for each item is obtained, items that have support greater than minimum support are selected as high-frequency patterns with length 1 or often abbreviated as 1-itemset. The abbreviation k-itemset means a set consisting of k items.

The second iteration produces a 2-itemset, each of which has two items. First, a 2-itemset candidate is created from a combination of all 1-itemsets. Then for each of these 2-itemset candidates the support is calculated by scanning the database. Support means the number of transactions in the database containing both items in the 2-itemset candidate. After the support of all 2-itemset candidates is obtained, the 2-itemset candidates that meet the minimum support requirements can be designated as 2-itemset which is also a high-frequency pattern with a length of 2.

For the next iteration, the kth iteration can be further divided into several parts:

The k-itemset candidate is formed from the combination of (k-1)-itemset obtained from the previous iteration. One characteristic of the a priori algorithm is the pruning of k-itemset candidates whose subsets containing k-1 items are not included in a high-frequency pattern with length k-1.

Support for each k-itemset candidate is obtained by scanning the database to calculate the number of transactions that contain all items in the k-itemset candidate. This is also a feature of the a priori algorithm, which requires calculations by scanning the entire database as long as the longest k-itemset.

High-frequency patterns that contain k items or k-itemsets are determined from k-itemset candidates whose support is greater than the minimum support. Then calculate the confidence of each item combination.

The iteration stops when all items have been counted until there are no more item combinations. In summary, the a priori algorithm is as follows:

```

Create L1 = set of supported itemsets of cardinality one
Set k to 2
while (Lk-1 ≠ ∅) {
    Create Ck from Lk-1
    Prune all the items sets in Ck that are not supported, to create Lk
    Increase k by 1
}
The set of all supported itemsets is L1 ∪ L2 ∪ ... ∪ Lk

```

Besides the a priori algorithm, there are also other algorithms such as FP-Growth. The difference between the a priori algorithm and FP-Growth is the number of database scans. The apriori algorithm scans the database every time it iterates, while the FP-Growth algorithm only does it once at the beginning.

A priori is an algorithm to perform frequent itemset searches to get association rules. As the name implies, this algorithm uses prior knowledge regarding previously known frequent itemset properties, to process further information. Apriori uses an iterative approach which is also known as level-wise search where k-itemset is used to search for (k+1)-itemset.

Association analysis is also known as one of the data mining techniques that forms the basis of various other data mining techniques (Saefudin et al., 2022; Xie, 2021). One stage of association analysis called frequent pattern mining has attracted the attention of many researchers to produce efficient algorithms.

The basic methodology of association analysis is divided into two stages (Ariestya et al., 2019; Wang & Gao, 2021):

High-frequency pattern analysis. This stage looks for item combinations that meet the minimum requirements of the support value in the database. The support value of an item is obtained by the following formula:

$$\text{Support } (A) = \frac{(\text{The number of transactions contains } A)}{(\text{Transaction totals})} \quad (1)$$

while the support value of the 2 items is obtained from the following formula:

$$\text{Support } (A \cap B) = \frac{(\text{The number of transactions contains } A \text{ and } B)}{(\text{Transaction totals})} \quad (2)$$

Formation of associative rules. After all high-frequency patterns are found, then the associative rules that meet the minimum requirements for confidence are sought by calculating the confidence of associative rules $A \rightarrow B$. The confidence value of rules $A \rightarrow B$ is obtained from the following formula:

$$\text{Confidence } (B | A) = \frac{(\text{The number of transactions contains } A \text{ and } B)}{(\text{The number of transactions contains } A)} \quad (3)$$

The trimming process is the result of items that have been combined and then trimmed using the minimum support specified by the user.

Results and Discussions

The experiment was conducted on a total of 195 songs released in the 1990s to 2000s. Out of a total of 195 songs, there are 24 chords composing each song. In detail, the basic chord data can be seen in table 1 and the song data and its arrangement can be seen in table 2.

Table 1. Basic Chord

ID	NAMA
1	A
2	Am
3	A#
4	B
5	Bm
6	Bb
7	C
8	Cm
9	C#
10	C#m
11	D
12	Dm
13	D#
14	D#m
15	E
16	Em
17	F
18	Fm

19	F#
20	F#m
21	G
22	Gm
23	G#
24	G#m

Table 2. Some of the Song List

ID LAGU	JUDUL LAGU	KUNCI LAGU	TAHUN
17	/Rif - Loe To Ye	D A Bm G D A Bm G	1990an
18	/Rif - Raja	G Em D C G Em D G	1990an
19	Agnes Monica - Cinta Diujung Jalan	C Am G D C C Am C D Bm Am C D C D Em	1990an
20	Akademi Fantasi - Menuju Puncak	G C G D Bm C D	1990an
21	Aldi Bragi - Janji	Am F D E Am Em F C E A G D A F E	1990an
22	Andre Hehanusa - Kutahu Engkau Begitu	G D Em C D Bm Em Am Bm C D C Bm Am D	1990an
23	Anggun - Mimpi	C G Am Em F D G Am B C C Em F Am G B C C Em Am G F Em D G C Am Em F C C F Am G Em F D G	1990an
24	Asbak band - Ternyata Salah Mengenalmu	G D C G Am D G D Am Em C G Am Em CD	1990an
25	Base Jam - Bukan Pujangga	C Em F G Am F Am G Am Em F G	1990an
26	Bebi Romeo - Bunga Terakhir	Am G F Em Dm G C Bm E C Bm E Am Gm C F Em Dm F G	1990an
27	Boomerang - Pelangi	Em G Bm C G D C D G D Bm C Em C Em CD	1990an
28	Broery marantika - Jangan Ada Dusta Diantara Kita	C G Am F Em Dm G Am Em F C Am Em FD G F C G C	1990an
29	Caffeine - Hidupku Kan Damaikan Hatimu	C Am F C Am Em F C Bm F#m G D	1990an
30	Chrisye - Pergilah Kasih	C F G C Am Dm F G C G Am F C G Dm C F G C	1990an
31	Desy Ratnasari - Tenda Biru	Am F G Am Am E Am A Dm Am B E Dm Am E Am	1990an
32	Dewa 19 - Kamulah Satu - Satunya	F C Am G C Am G F C Am G F Am	1990an
33	Dewa 19 - Satu Hati	G Em C D Em C G D D Em C D Em G Am C	1990an
34	Dewa 19 - Tak Kan Ada Cinta Yang Lain	Am F G Em F C E Am Dm G Em F C G C	1990an
35	Dewa 19 - Cemburu	G D Em Em D C Em D B Em C D G A# Gm A# C Am Cm	2000
36	Dewa 19 - Pupus	G D Em C G Am C D G C C D Am G D C G A D G D Em D C Cm	2002
37	Dewa 19 - Kirana	Am F G Dm Am F G Bb Dm F C Am G Em D# Bb F C G Em G D Bm A F#m Bm G A Em Bm G A C C F Am C	1990an
38	Dewa 19 - Risahkan Hati	Am Em F G F Em Dm G F G Em A Dm Em F G	2000
39	Dr PM - Damai Mimpi	C G Em D Am D G C G D Em C C D Em F#m Am	1990an
40	Dr PM - Pernah Mencoba	C B F G Dm C G Dm F C G F F C	1990an
41	Dygtta - Karena Ku Sayang Kamu	C Am Dm G F G Em Am Dm G D F G	1990an

Based on these chord data, an apriori process is then carried out in order to obtain the rules with the highest support and confidence values. From the experimental results using the entire song, the most dominant chords are G, Am and Em. The value of support from the chord is:

$$\text{Support (G)} = (\text{The number of songs contain G}) / (\text{Song totals}) = 168/195 = 0,86$$

$$\text{Support (Am)} = (\text{The number of songs contain Am}) / (\text{Song totals}) = 128/195 = 0,65$$

$$\text{Support (Em)} = (\text{The number of songs contain Em}) / (\text{Song totals}) = 126/195 = 0,64$$

From the results of the calculation above, if a minimum support value of 0.4 is used, the three chords can be processed in the calculation of itemset 2. The next process is the calculation of the value of itemset 2. Because the results of itemset 1 are obtained, the three chords have a higher support value than the

value minimum support, then the three chords are combined. The results of the calculation of the combination of the three chords are as follows:

$$\text{Support (G and Am)} = \frac{\text{(The number of songs contain G and Am)}}{\text{(Song totals)}} = \frac{122}{195} = 0,62$$

$$\text{Support (G and Em)} = \frac{\text{(The number of songs contain G and Em)}}{\text{(Song totals)}} = \frac{121}{195} = 0,62$$

$$\text{Support (Am and Em)} = \frac{\text{(The number of songs contain Am and Em)}}{\text{(Song totals)}} = \frac{96}{195} = 0,49$$

Because the support value of the combination of the three chords meets the minimum support value, the itemset can be continued into the next itemset. It is just that because itemset 2 consists of only 1 combination, namely the combination Am, G and Em, the itemset calculation process is stopped. The next process is to form a priori rules from the available itemsets. Each rule then calculates its confidence value as follows:

$$\text{Confidence (Em | G and Am)} = \frac{\text{(The number of transactions contains G, Am and Em)}}{\text{(The number of transactions contains G and Am)}} = \frac{93}{122} = 0,7622950819672131$$

$$\text{Confidence (Am | G and Em)} = \frac{\text{(The number of transactions contains G, Am and Em)}}{\text{(The number of transactions contains G and Em)}} = \frac{93}{121} = 0,768595041322314$$

$$\text{Confidence (G | Em and Am)} = \frac{\text{(The number of transactions contains G, Am and Em)}}{\text{(The number of transactions contains Em and Am)}} = \frac{93}{96} = 0,96875$$

$$\text{Confidence (Am and Em | G)} = \frac{\text{(The number of transactions contains G, Am and Em)}}{\text{(The number of transactions contains G)}} = \frac{93}{168} = 0,5535714285714286$$

$$\text{Confidence (G and Em | Am)} = \frac{\text{(The number of transactions contains G, Am and Em)}}{\text{(The number of transactions contains Am)}} = \frac{93}{128} = 0,75$$

$$\text{Confidence (G and Am | Em)} = \frac{\text{(The number of transactions contains G, Am and Em)}}{\text{(The number of transactions contains Em)}} = \frac{93}{126} = 0,7619047619047619$$

If using a Minimum Confident value of 0.7, the rules that can be used are as shown in table 3. Meanwhile, other rules are not used because they are below the minimum confident value used.

Table 3. Overall Song Rule Results

Rules	Confidence
{ Am, Em } → { G }	0,96875
{ G, Em } → { Am }	0,768595041322314
{ G, Am } → { Em }	0,7622950819672131
{ Em } → { G, Am }	0,7619047619047619
{ Am } → { G, Em }	0,75

Based on the results shown in table 3, the songs collected in the database from the 1990s to 2000s, mostly used three chord combinations Am, Em and G. From table 3, the rule with the highest confidence value is the rule:

$$\{Am, Em\} \rightarrow \{G\}$$

Shows that for most chord combinations that start with Am and Em, it is always followed by a G key throughout the entire chord of the song.

Conclusions

From the experiments that have been carried out, several conclusions can be drawn. The first conclusion is that of all the songs in the database, the G chord is the chord that appears the most, with 168 followed by the Am chord with 128 and the Em chord with 126, which shows that these chords were the most popularly used in Indonesian pop songs in the era of 1990 to 2000s. The second conclusion about the association rules that can be drawn from Indonesian pop songs chord is $\{Am, Em\} \rightarrow \{G\}$, which show that in Indonesian pop song in the era of 1990 to 2000s, the chord that start with Am and Em, will be followed by a G key. Then main finding from the research show that only one association rules that can be drawn from Indonesian pop song in the era of 1990 to 2000s. This finding need more research in the future to broaden the area of Indonesian song, by genre or the era, to show another possibilities association rules. The more song that included in the research can give more insight and another perspectives that may give us a new association rules. Another way is not only included the first chord, but the chord that included in the chorus might be able to show us another great results.

References

- Alfianzah, R., Handayani, R. I., & Murniyati, M. (2020). Implementation of Apriori Algorithm Data Mining for Increase Sales. *Sinkron*, 5(1), 17. <https://doi.org/10.33395/sinkron.v5i1.10587>
- Ariestya, W. W., Supriyatin, W., & Astuti, I. (2019). MARKETING STRATEGY FOR THE DETERMINATION OF STAPLE CONSUMER PRODUCTS USING FP-GROWTH AND APRIORI ALGORITHM. *Jurnal Ilmiah Ekonomi Bisnis*, 24(3), 225–235. <https://doi.org/10.35760/eb.2019.v24i3.2229>
- Dewi Sri Mulyani, E., Purnama Putra, Y., Badar Sambani, E., Siti Sundari, S., Mufizar, T., & Satrio Nugraha, M. (2021). Student Competency Association Analysis For Learning Evaluation Using Apriori Algorithm. *Elinvo (Electronics, Informatics, and Vocational Education)*, 6(2), 120–130. <https://doi.org/10.21831/elinvo.v6i2.42264>
- Fenny Syafariani, R. (2022). Association Analysis with Apriori Algorithm for Electronic Sales Decision Support System. *International Journal of Informatics Information System and Computer Engineering*, 3(1), 61–72. <https://doi.org/10.34010/injiiscom.v3i1.8089>
- Gredy Aprianno Penciptaan Seni Musik Bara, R., & dan Pengkajian, P. (2020). PENGARUH ETUDE TERHADAP TEKNIK PERMAINAN GITAR KLASIK DAN GITAR ELEKTRIK. 2, 103–108.
- Henry, M., Chandra, W., & Zahra, A. (2022). Implementation of Apriori Algorithm for Music Genre Recommendation. *Jurnal Online Informatika*, 7(1), 110. <https://doi.org/10.15575/join.v7i1.819>
- Idris, A. I., Sampetoding, E. A. M., Ardhana, V. Y. P., Maritsa, I., Sakri, A., Ruslan, H., & Manapa, E. S. (2022). Comparison of Apriori, Apriori-TID and FP-Growth Algorithms in Market Basket Analysis at Grocery Stores. *The IJICS (International Journal of Informatics and Computer Science)*, 6(2), 107. <https://doi.org/10.30865/ijics.v6i2.4535>
- Ilmiah, J., & Grafis, K. (n.d.). Implementasi Algoritma Apriori Pada Sistem Rekomendasi Produk Rameiki Mart. 15(2), 330–347. <http://journal.stekom.ac.id/index.php/pixel/page330>
- Irfiani, E. (2019). Application of Apriori Algorithms to Determine Associations in Outdoor Sports Equipment Stores. *Sinkron*, 3(2), 218. <https://doi.org/10.33395/sinkron.v3i2.10089>
- Kharisma, B. (2021). Performing Arts Education ANALISIS KOMPOSISI SOUNDTRACK DALAM VIDEO GAME “GENSHIN IMPACT.” <https://doi.org/10.24821/ijopaed>
- Le, Y. (2022). Research on data resource management of biomass energy engineering based on data mining. *Energy Reports*, 8, 1482–1492. <https://doi.org/10.1016/J.EGYR.2022.02.048>
- Pan, T. (2021). An Improved Apriori Algorithm for Association Mining Between Physical Fitness Indices of College Students. *International Journal of Emerging Technologies in Learning*, 16(9), 235–246. <https://doi.org/10.3991/ijet.v16i09.22747>

- Rumui, N., Andi Roy, rd, & Pujo Hari Saputro Informatics Manajemen Politeknik Negeri FakFak FakFak, th. (2022). Comparison of Apriori Algorithm and FP-Growth in Managing Store Transaction Data 1 st Syukron Anas, 2 nd. In International Journal of Computer and Information System (IJCIS) Peer Reviewed-International Journal (Vol. 03). <https://ijcis.net/index.php/ijcis/index>
- Saefudin, M., Diana, D., Hendrato, H., Informasi, S., Komputer, S., & Jakarta STI, S. (2022). DEVELOPMENT OF DATA MINING SOFTWARE USING ASSOCIATION TECHNIQUES BASED ON APRIORI ALGORITHM METHOD (Pengembangan Perangkat Lunak Data Mining Menggunakan Teknik Asosiasi Berdasarkan Metode Algoritma Apriori). Journal of Information System, Informatics and Computing Issue Period, 6(1), 125-136. <https://doi.org/10.52362/jisicom.v6i1.800>
- Samuel, S., Sani, A., Budiyantra, A., Ivone, M., & Frieyadie, F. (2022). SALES LEVEL ANALYSIS USING THE ASSOCIATION METHOD WITH THE APRIORI ALGORITHM. Jurnal Riset Informatika, 4(4), 331-340. <https://doi.org/10.34288/jri.v4i4.422>
- Sari, N. D., W A, B. S., & Nasiri, A. (2020). Implementation of Apriori Algorithm for Determining Product Purchase Patterns. SISFOTENIKA, 11(1), 1. <https://doi.org/10.30700/jst.v11i1.1033>
- Wang, H. Bin, & Gao, Y. J. (2021). Research on parallelization of Apriori algorithm in association rule mining. Procedia Computer Science, 183, 641-647. <https://doi.org/10.1016/j.PROCS.2021.02.109>
- Xie, H. (2021). Research and Case Analysis of Apriori Algorithm Based on Mining Frequent Item-Sets. Open Journal of Social Sciences, 09(04), 458-468. <https://doi.org/10.4236/jss.2021.94034>
- Yustiana, I., Sutarha, I., Maulawati, N., & Yusuf, I. M. (2021). BASKET MARKET ANALYSIS USING R-BASED APRIORI ALGORITHM TO FIND INFORMATION FROM SALES DATA. International Journal of Engineering and Applied Technology (IJEAT, 4(2), 79-92. <https://doi.org/10.52005/ijeat.v3i1.xxx>