Long Song Type Classification based on Lyrics

Bayarsaikhan Namjil[1], Nandinbilig Ganbaatar[1], Suvdaa Batsuuri[1*]

Abstract

Mongolian folk songs are inspired by Mongolian labor songs and are classified into long and short songs. Mongolian long songs have ancient origins, are rich in legends, and are a great source of folklore. So it was inscribed by UNESCO in 2008. Mongolian written literature is formed under the direct influence of oral literature. Mongolian long song has 3 classes: ayzam, suman, and besreg by their lyrics and structure. In ayzam long song, the world perfectly embodies the philosophical nature of world phenomena and the nature of human life. Suman long song has a wide range of topics such as the common way of life, respect for ancestors, respect for fathers, respect for mountains and water, livestock and animal husbandry, as well as the history of Mongolia. Besreg long songs are dominated by commanded and trained characters. In this paper, we proposed a method to classify their 3 types of long songs using machine learning, based on their lyrics structures without semantic information. We collected lyrics of over 80 long songs and extracted 11 features from every single song. The features are the name of a song, number of the verse, number of lines, number of words, general value, double value, elapsed time of verse, elapsed time of 5 words, and the longest elapsed time of 1 word, full text, and type label. In experimental results, our proposed features show on average 78% recognition rates in function type machine learning methods, to classify the ayzam, suman, and besreg classes.

Key Words: Mongolian Long Song, Song Genre Classification by Lyrics, Machine Learning, Feature Extraction from Song Lyrics.

I. INTRODUCTION

The well-being, history, and culture of Mongolians can be seen in any genre of folklore only by reading the poetry of long songs [1]. The lyrics of long songs were first composed by someone, developed, and spread through word of mouth, becoming a form of folklore. Looking at the historical roots of Mongolian literature and culture, there are very ancient sources, and the poems of long songs are of ancient origin, and they are examples of elegant poems and poems composed by our sages. But now, because the long song is not sung in all its tones, the full meaning of the long song is not heard or known. In this study, we will not only express the poetic meaning of the long song in terms of literature, but also identify and evaluate long song data using machine learning methods in 3 categories: ayzam, suman, and besreg. There is a lot of research in this field in many fields around the world, and the Mongolian linguistics sector has developed this interdisciplinary research and made great achievements. Our work is new, as it is the first of its kind in the field of folklore and literature. In particular, it is very important to start with long poems.

Our mission is to study and promote artificial intelligence, which is widely used in multidisciplinary research around the world, in combination with its unique folklore heritage.

II. RELATED WORKS

The first person to study Mongolian long songs on a scientific basis was the Russian scientist A. Pozdneev. In 1880, he included Buryat, Khalkh, and Ould songs in his Mongol long song, and the Russian scholar AD Rudneev also studied the melodies of Mongolian long songs. В. Я. Vladimertsov carefully studied and recorded Oirat songs [2].

In addition to the oral source of long songs, the written source has become a major research tool in this field. Mongolian scholars such as P. Khorloo, H. Sampildendev, Sh. Dorj, J. Badraa, and S. Tsoodol have studied long songs, and some senior scholars of the SCC of the Mongolian Academy of Sciences have collected written scriptures and books that were widely distributed in Mongolia. The famous long song singer J. Dorjdagva is not only a great singer but also a researcher who has a great place in the history of long song studies. J. Badraa published a book about his story called "The Great Singer's Speech" which is a val-

uable work among scholars and researchers in this field. Dr. A. Alimaa, Head of the Institute of Linguistics and Oral Studies of the Mongolian Academy of Sciences, has studied, discovered, and put into circulation more than 3,000 long songs sung in Mongolia [3]. The study of Mongolian poetry, it has been studied by Western Mongol scholars since the middle of the 19th century. They have been observing and emphasizing the uniqueness of Mongolian poetry [4-5] and [6-7]. Also the Long song was inscribed by UNESCO in 2008 [8].

On the other side, computer science researchers are re-searching to classify sound types. It is common to process signals from audio data and classify them into rock, pop, rap, and classical [9]. Although there are fewer classifica-tions based on verse alone than audio, there are also works using natural language processing and machine learning. In recent years, the use of deep learning has increased, and as a result, deep learning methods such as RNN have been used to classify verse data. For example, the work of Ale-xandros Tsaptsinos [10]. Anna Boonyanit's work [11] cate-gorizes hip-hop, rock, and pop with about 60 % recognition rate. However, no research has been conducted in Mongolia to classify the types of long songs and the meaning of the poems automatically using machine learning. Therefore, in this study, we purposed to Mongolian long song type clas-sification using machine learning methods.

III. LONG SONG TYPE CLASSIFICATION

3.1. Mongolian Poetry, Poetic Tradition, Regularity, In-terpretation of the Meaning of the Verse

Mongolian folk songs are one of the major genres of Mongolian folklore. Oral literature is a work of art that orig-inated from the life of the people and spread through word of mouth as an expression of Mongolian customs, history, culture, and wisdom. The main types of folklore include fairy tales, epics, legends, riddles, proverbs, blessings, praises, the three worlds, and old sayings. Many of these genres are poetic. It seeks to study song poetry, including long song poetry, which includes verse patterns, word inter-pretations, the meaning of lyrics, and the ability to classify lyrics by machine learning methods.

The poetry of long songs is mostly composed of written words. The noble composition of the Mongolian script and the choice of rare words in the Mongolian language fund show that the Mongolian long song is not only a genre of oral literature but also written poetry. It seems that most of the poems in long songs were written by highly educated people. This is especially evident in the long songs of state-related reverence. The expression of the above is that the verses of the long song are read to gain a wide range of knowledge and teachings, such as the phenomena of the

universe, nature, customs, and respect and love.

Long songs are divided into three types according to their melody and size: ayzam, suman, and besreg, and these types are also reflected in the meaning and content of the poem. Many works of Mongolian poetry are thematically categorized only in terms of verse content. Ayzam (large-scale song) is a song with a wide range of melodies, and a large number of retro folds, which are larger than the other two categories. Suman (medium-sized song) long song has a wide range of melodies, is fast and has a lot of ornaments, and is widely sung in Mongolia. In addition to popular top-ics such as farming, there is a wide range of topics that can be explored to understand the history of Mongolia. Besreg (short or small songs) long songs have a wider melody than short folk songs, but they are not short, they have short per-cussion and ornamentation, and the meaning and content of the words are dominated by syllables and teachings. There is a tradition of using this type of song as a learning tool for beginners.

3.2. Ability to Classify Long Song Types by Lyrics

Mongolian folk song is innovative and important to study long songs around the world, combine them with computer science research methods, and expand it into in-terdisciplinary research. The most important thing to do in this area is to collect a large amount of data. However, there is a lot of data collected for the written sources of long songs, and we decided to experiment with the example of Central Khalkh songs in this study. Khalkh long song is widespread in the heart of Mongolia, so most of the songs commonly sung today fall into this category. Every song has verses (badag/turleg in Mongolian), every verse con-sists of lines, and each line has several words. In this work, we studied the possibility of classifying three types of ayzam, suman, and besreg based on the data of the long song verses by machine learning method. The following figure shows the general scheme of work.

Data preparation and features: This time, we collected and experimented with 14 ayzam types, 45 suman types, 21
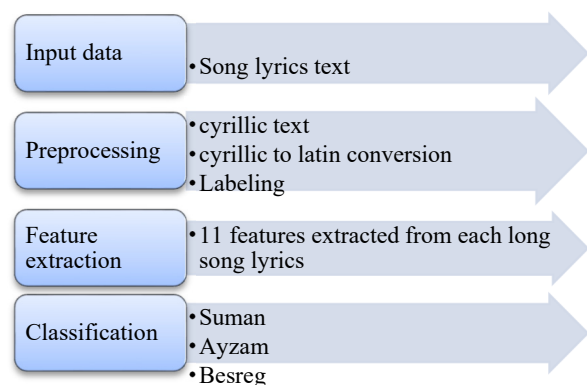

Fig. 1. Long song type classification general scheme.

besreg types, and a total of 80 lyrics.

Features: From the long song data, 11 features such as song name (string), number of verses (numeric), number of lines (numeric), number of words (numeric), generalvalue (string), doublevalue(string), elapsed time of verse(numeric), elapsed time of 5 words(numeric), the longest elapsed time of 1 word (numeric), full text(string), category name (string -suman, ayzam or besreg). Fig. 2 shows the average values of verse/line/word numbers in 3 types.

Fig. 3 shows the average continued time of verse/5 words/ the longest 1 word in 3 types.

Here, it is clear to the long song means because the longest continued time for 1 word is 27 seconds in the Suman type case.

Because machine learning methods are relatively effective depending on the data distribution and characteristics,
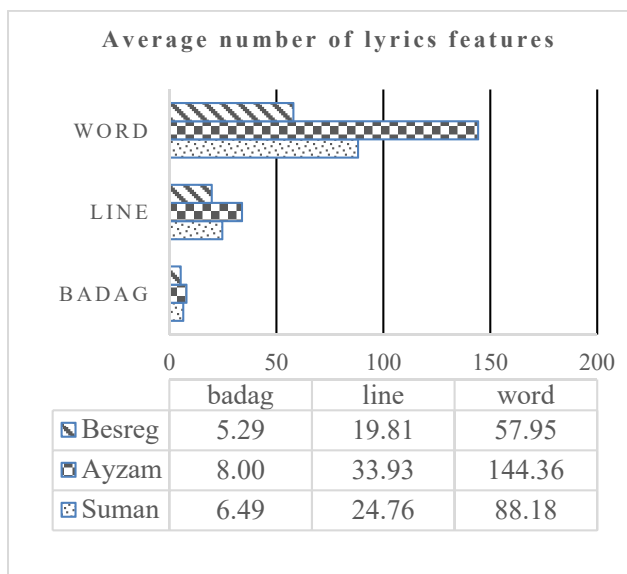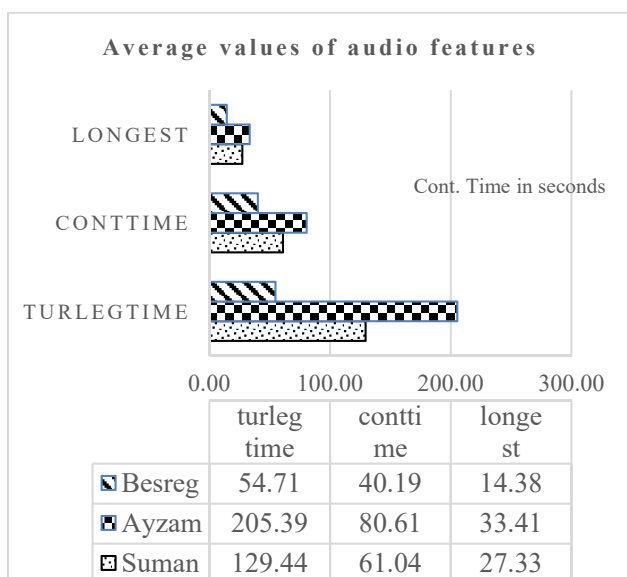
possible methods have been tested using the weka program. The best method for our data was Multilayer perceptron algorithm.

3.3. MLP Neural Network

A Multilayer Perceptron has input and output layers, and one or more hidden layers with many neurons stacked together. And while in the Perceptron the neuron must have an activation function that imposes a threshold, like ReLU or sigmoid, neurons in a Multilayer Perceptron can use any arbitrary activation function. Multilayer Perceptron falls under the category of feedforward algorithms because inputs are combined with the initial weights in a weighted sum and subjected to the activation function, just like in the Perceptron. But the difference is that each linear combination is propagated to the next layer. Each layer is feeding the next one with the result of their computation, their internal representation of the data. This goes all the way through the hidden layers to the output layer. If the algorithm only computed the weighted sums in each neuron,

Average number of lyrics features

	badag	line	word
▨ Besreg	5.29	19.81	57.95
▥ Ayzam	8.00	33.93	144.36
⊞ Suman	6.49	24.76	88.18

Fig. 3. The average values of audio features for the 3 classes.

Average values of audio features

Cont. Time in seconds

	turleg time	contti me	longe st
▨ Besreg	54.71	40.19	14.38
▥ Ayzam	205.39	80.61	33.41
⊞ Suman	129.44	61.04	27.33

Fig. 3. The average values of audio features for the 3 classes.
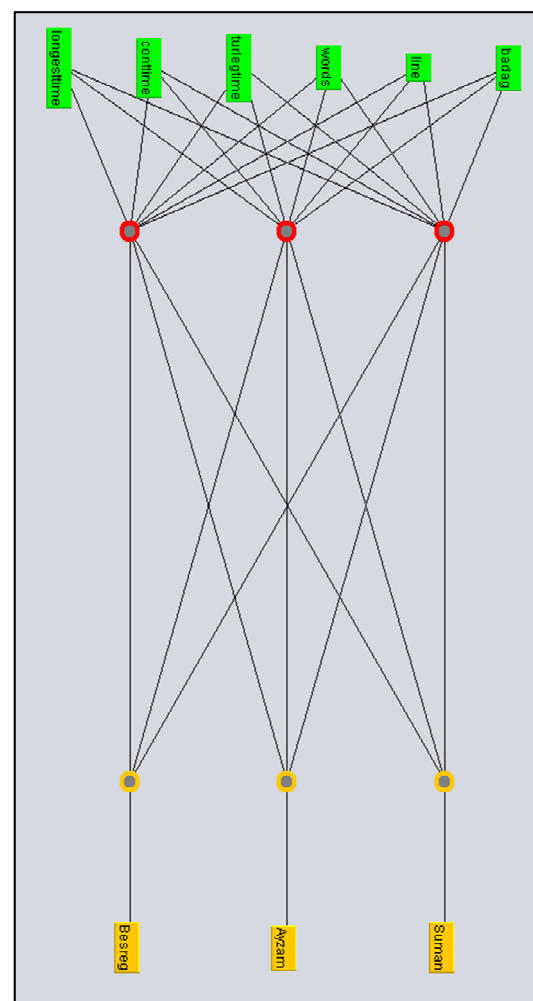

Fig. 4. Multilayer Perceptron example of Long song numeric values.

propagated results to the output layer, and stopped there, it wouldn't be able to learn the weights that minimize the cost function. If the algorithm only computed one iteration, there would be no actual learning. Backpropagation is the learning mechanism that allows the Multilayer Perceptron to iteratively adjust the weights in the network, to minimize the cost function. In each iteration, after the weighted sums are forwarded through all layers, the gradient of the Mean Squared Error is computed across all input and output pairs. Then, to propagate it back, the weights of the first hidden layer are updated with the value of the gradient. This process keeps going until the gradient for each input-output pair has converged, meaning the newly computed gradient hasn't changed more than a specified convergence threshold, compared to the previous iteration.

The results are described in detail in the experimental results section. The following figure shows an example of how song data is prepared in a * .arff file to be read in a weka program [12].

```
@relation 'longsong'
@attribute badag numeric
@attribute line numeric
@attribute words numeric
@attribute turlegtime numeric
@attribute conttime numeric
@attribute longesttime numeric
@attribute class {Suman,Ayzam,Besreg}
@data
12,48,144,150,51,20,Suman
5,20,61,194,67,35,Ayzam
2,8,27,52,52,15,Besreg
4,20,81,173,56,19,Suman
3,12,65,126,76,40,Ayzam
```

```
@relation 'longsong'
@attribute name string
@attribute badag numeric
@attribute line numeric
@attribute words numeric
@attribute generalvalue string
@attribute doublevalue string
@attribute turlegtime numeric
@attribute conttime numeric
@attribute longesttime numeric
@attribute Text string
@attribute class {Suman,Ayzam,Besreg}
@data
'Urkhan khongor salkhi',12,48,144,'KH
uzegdeltey khosuulsan','KHuleg morio,
khosuulsan',150,51,20,'Urkhan khongor
gankhdag bilee \n Uyakhan gol shig kh
dasdag bilee \n KHaychin chikhtey sha
tuulna daa \n KHayrtay amrag chamtayg
daa \n unegen chikhtay shargaaraa \n
unekheer khayrtay khuuteygee \n Udam
deltey shargaaraa \n SHargyn goviyg t
amragtaygaa \n Sardaa neg uulzana bil
```

Fig. 5. Data format for machine learning algorithms input (Weka tool).

116

IV. EXPERIMENTAL RESULTS

We classified the collected data using machine learning methods. This time, we have collected and experimented with 14-ayzam types, 45-suman types, 21-besreg types, and a total of 80 lyrics. We tested the data with only text features, and only numeric values and combined text and numeric values in a 10-fold cross-validation (Table 1).

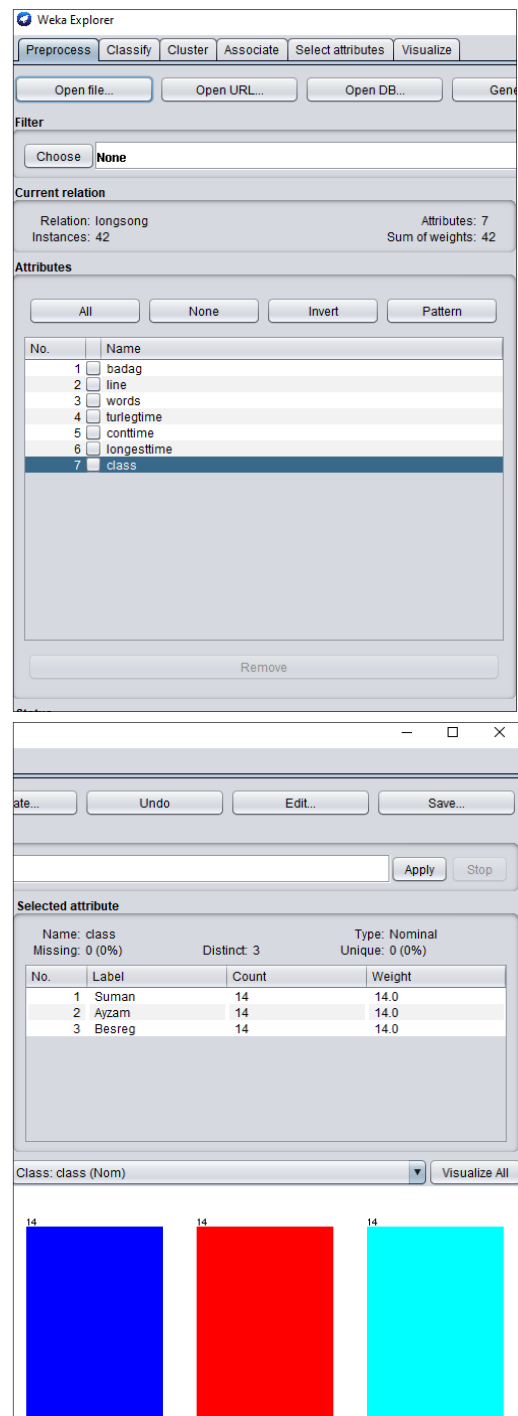Fig. 6 shows the example of balanced and numeric valued data in Weka.

Fig. 6. Data inputs(balanced numeric values) in Weka.

Table 1. The number of experimental data samples.

	Suman	Ayzam	Besreg	Total
Unbalanced	45	14	21	80
Balanced	14	14	14	42

Fig. 7 shows an example of the best result of classification by the MLP method in Weka.

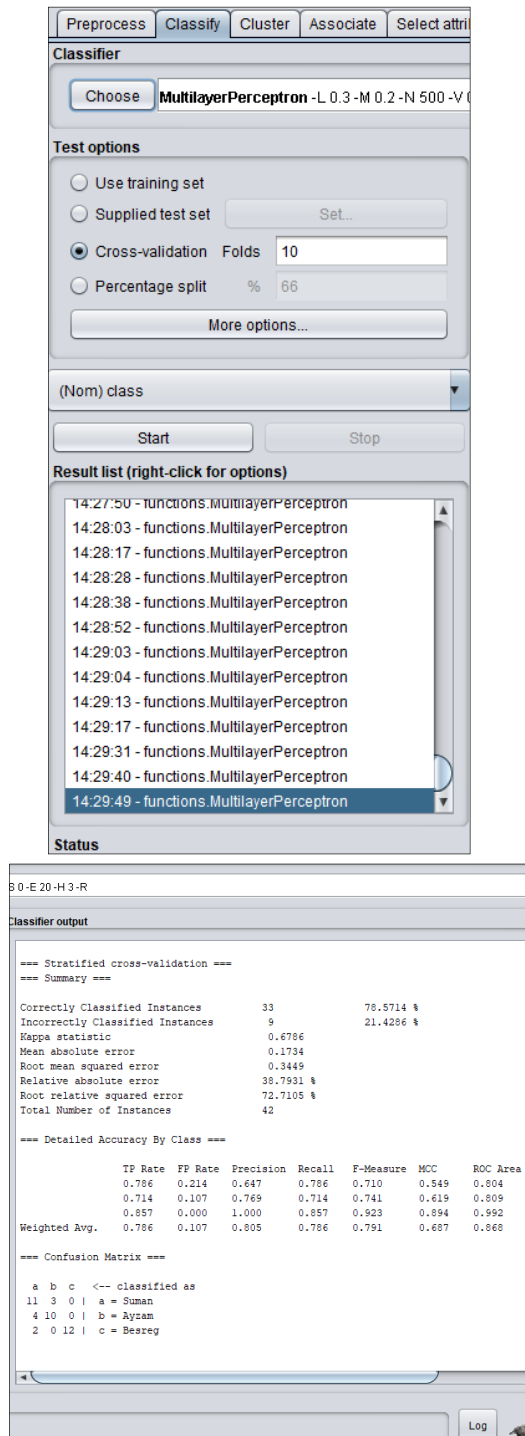Unbalanced data in the three categories in terms of data

Fig. 7. An example of classification results in Weka.

may affect the classification results. Most of these three categories are suman type songs.

Table 2 shows comparison results of song type/genre classification in balanced and unbalanced data.

In our case, the best result is function methods show 78% accuracy in balanced data with only numeric values. There are 6 features that have numeric values and 3 of them are about lyrics structure, then 3 of them are about audio information. The worst case is text acceptable methods show 20-56% in accuracy balanced and unbalanced data with texts. Because we did not use any natural language processing methods. We converted Cyrillic text to Latin text character by character. Therefore it is meaningless about semantics. They compared 2 methods are used language models such as BERT, LSTM, etc. So their work is meaningful and shows better results.

On the other side, song genres are big differences compared with the 3 types of one genre. Our goal is to classify 3 types of only Long song genres. Table 4 shows the results of 3 types each and weighted average scores.

According to the definition of the 3 types of long songs, Suman and Ayzam songs are similar, and Ayzam and Besreg songs are similar too. Besger is the shortest one. Therefore, the Besreg songs have the highest values in Table 4.

Table 2. Experimental results comparison.

	Unbalanced	Balanced
NB	61.25	69.04
SMO	63.75	78.57
MLP	56.25	76.19
IBK	63.75	66.66
J48	68.75	59.52
RandomForest	67.5	69.04

Table 3. Experimental results comparison with related works. Only text values of Long song.

# of classes	Long song	Sam et al. [13]	Akalp et al. [14]	
			6 genres	
	3	10 genres	BERT model	Distil BERT model
NB	56.25% with text value acceptable methods	0.603	77.63% in one-label	74.38% in one-label
SMO		0.574		
MLP		-		
IBK		-		
J48		0.358		
RF		-		

117

Table 4. Classification results for 3 types.

	Precision	Recall	F1 score
Suman	0.647	0.786	0.710
Ayzam	0.769	0.714	0.741
Besreg	1	0.857	0.923
Weighted average	0.805	0.786	0.791

Fig. 8. A confusion matrix of the models with balanced data.

Fig. 8 shows the confusion matrix of 3 types of classification results. As mentioned above, Suman and Ayzam songs are misclassified compared with Besreg songs.

In the other viewpoint, we may mislabel the 3 types of long songs, because there is no exact correct answer which is suman type, which is ayzam type, etc.

V. CONCLUSION

This study tested the features of long song lyrics, such as long song verses, poetry, its structure, and the symbolic meaning of long songs, as well as the possibility of combining traditional long song research with modern technological advances. The main result of this study was that long song researchers showed how it is possible to classify data by machine learning when preparing data and classifying it by the human mind. In this time, we tested 80 songs, future it is possible to experiment with the lyrics of more than 300 popular songs.

ACKNOWLEDGEMENT

REFERENCES

[1] Long song definition, https://en.wikipedia.org/wiki/Long_song

[2] B. Y. Vladimirtsov and J. R. Krueger, "The oirat-mongolian heroic epic,"*Mongolian Studies*, vol. 8, pp. 5-58, 1983.

[3] A. Alimaa, "Features, distribution and release characteristics of long songs," Ulaanbaatar, 2013.

[4] G. Galbayar, "The method of connecting the head of Mongolian poetry and its regularity," Ulaanbaatar, 2014.

[5] Mend-oyoo, http://www.mend-ooyo.mn/, 2014.

[6] K. Sampildendev and K. N. Yatsovskoi, Mongolian folk long song, Ulaanbaatar, 1984.

[7] S. Yoon, "Remains and renewals: The process of preserving Urtyn duu in contemporary Mongolia," *Mongolian Studies*, vol. 35, pp. 119-131, 2013.

[8] Unesco report, https://ich.unesco.org/en/RL/urtiin-duu-traditional-folk-long-song-00115.

[9] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293-302, Jul. 2002.

[10] A. Tsaptsinos, Lyric-Based Music Genre Classification using a Hierarchical Attention Network, Jul. 2017. https://arxiv.org/abs/1707.04678

[11] A. Boonyanit, A. Dahl, and M. Leszczynski, Music Genre Classification using Song Lyrics, Stanford CS224N Custom Project, https://web.stanford.edu/class/cs224n/reports/final_reports/report003.pdf

[12] F. Eibe, A. H. Mark, and H. W. Ian, The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Fourth Edition, 2016. https://doc1.bibliothek. li/acb/FLMF040119.pdf

[13] H., Sam, N. Carlos, Jr. Silla, and C. G. Johnson. Automatic Lyrics based Music Genre Classification in a Multilingual Setting, in *Thirteenth Brazilian Symposium on Computer Music*. https://kar.kent.ac.uk/33266/. 2011.

[14] H. Akalp, E. F. Cigdem, S. Yilmaz, N. Bölücü, and B. Can, "Language representation models for music genre classification using lyrics," *International Symposium on Electrical, Electronics and Information Engineering*, pp. 408-414, Feb. 2021.

AUTHORS

Bayarsaikhan Namjil received her BS and MS degrees in the Department of Mongolian language and literature from the National University of Mongolia, Mongolia in 2003 and 2006, respectively.

She started studying for a doctorate in 2018 at the Department of Literature and Arts at the National University of Mongolia. Her research interests include Mongolian folklore, long songs, and artificial intelligence.

Nandinbilig Ganbaatar received her BA and MA degrees in the Department of Literature from National University of Mongolia, in 1996 and 1998 respectively. In 2001, she received a PhD degree in the Department of Literature from National University of Mongolia. In the September of 1998, she joined in the Nomadic Civilization research team at the National University of Mongolia, where she was a lecturer.   From the September of 2002 to the September of 2012, she was an associate professor in the Department of Literature, since the September of 2012, she is currently a professor in the Department of Literature and Arts, at National University of Mongolia. She's published two monographies and over fifty international journals and conference papers. Her research interests include The Idea of Beauty in Heroic epic and Symbolic coding National identity in Mongol Folklore.
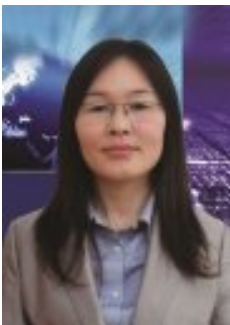
Suvdaa Batsuuri received her BS and MS degrees in Electronics from the National University of Mongolia, in 2002 and 2004, respectively. In 2011, she received a Ph.D. degree in Computer Science, in the Department of Computer Engineering from Kumoh National Institute of Technology, Korea. Her research interests include computer vision, human-computer interaction, and artificial intelligence.