**Aigul Mimenbayeva**
Master of Sciences, Senior Lecturer, Department of Computational and Data Science
aigulka79_79@mail.ru, orcid.org/0000-0003-4652-470X
Astana IT University, Kazakhstan

**Gulmira Bekmagambetova**
PhD, Senior Lecturer of the Department of Information Systems
gulmirabekmagam@gmail.com, orcid.org/0000-0002-8999-793X
S.Seifullin Kazakh Agro Technical Research University, Kazakhstan

**Gulzhan Muratova**
Candidate of Physical and Mathematical Sciences, Acting Associate Professor
mugk1234@gmail.com, orcid.org/0009-0001-1259-0832
S.Seifullin Kazakh Agro Technical Research University, Kazakhstan

**Akgul Naizagarayeva**
Master of Technical Sciences, Senior Teacher of the Department of Information Systems
akgul_1985@mail.ru, orcid.org/0000-0001-6888-1092
S.Seifullin Kazakh Agro Technical Research University,  Kazakhstan

**Tleugaisha Ospanova**
Candidate of Technical Sciences, Acting Professor of the Faculty of Information Technologies
Tleu2009@mail.ru, orcid.org/0000-0002-1729-1321
L.N. Gumilyov Eurasian National University, Kazakhstan

**Assem Konyrkhanova**
PhD, Acting of Associate Professor of the Faculty of Information Security
ErkeshanK@mail.ru, orcid.org/0000-0002-4901-8901
L.N. Gumilyov Eurasian National University, Kazakhstan

# CLASSIFICATION OF KAZAKH MUSIC GENRES USING MACHINE LEARNING TECHNIQUES

**Abstract:** This article analysis a Kazakh Music dataset, which consists of 800 audio tracks equally distributed across 5 different genres. The purpose of this research is to classify music genres by using machine learning algorithms Decision Tree Classifier and Logistic regression. Before the classification, the given data was pre-processed, missing or irrelevant data was removed. The given dataset was analyzed using a correlation matrix and data visualization to identify patterns. To reduce the dimension of the original dataset, the PCA method was used while maintaining variance. Several key studies aimed at analyzing and developing machine learning models applied to the classification of musical genres are reviewed.

Cumulative explained variance was also plotted, which showed the maximum proportion (90%) of discrete values generated from multiple individual samples taken along the Gaussian curve. A comparison of the decision tree model to a logistic regression showed that for f1 Score Logistic regression produced the best result for classical music – 82%, Decision tree classification – 75%. For other genres, the harmonic mean between precision and recall for the logistic regression model is equal to zero, which means that this model completely fails to

classify the genres Zazz, Kazakh Rock, Kazakh hip hop, Kazakh pop music. Using the Decision tree classifier algorithm, the Zazz and Kazakh pop music genres were not recognized, but Kazakh Rock with an accuracy and completeness of 33%. Overall, the proposed model achieves an accuracy of 60% for the Decision Tree Classifier and 70% for the Logistic regression model on the training and validation sets. For uniform classification, the data were balanced and assessed using the cross-validation method.

The approach used in this study may be useful in classifying different music genres based on audio data without relying on human listening.

**Keywords:** Machine learning algorithms; music genre; Decision Tree Classifier; Logistic regression; cross-validation.

### Introduction

The digital era has revolutionized the way we interact with music. With the vast amount of music available online, the ability to accurately categorize songs by genre is crucial for music discovery, recommendation systems, and audience targeting.

Music genre classification is an important sub-discipline of music information retrieval (MIR), a field that intersects various aspects from musicology to AI and ML. The challenges in this field, notably the ambiguous nature of music genres, have led researchers to explore various ML techniques for more effective classification [1]. Historically, music genre classification relied on acoustic and sound characteristics. However, the introduction of ML, especially deep learning and neural networks, has revolutionized this approach. The ability of these models to handle complex, multi-layered data has significantly improved the accuracy and effectiveness of music genre classification [2]. Supervised learning is predominantly used in music classification, involving model training using labeled data to predict genres. The process includes extracting key features from audio and training a model with these features. Evaluation metrics vary depending on the balance of the dataset, with accuracy being a standard measure in balanced cases [3].

Music genre classification can encompass various types of classification problems, ranging from binary to multi-class and multi-label classifications, reflecting the complexity and diversity of musical genres [4]. The standard procedure in music classification involves transforming audio into representative features, then constructing and evaluating a model based on these features. This evaluation is crucial to assess the model's accuracy and effectiveness in real-world applications [5].

Additionally, it is important to note that audio data is initially in analog format. To process it on a computer, it needs to be converted to a digital format using an analog-to-digital converter. This conversion involves sampling the analog signal at regular intervals and quantizing the amplitude values into discrete levels, creating a digital representation of the audio [6].

Overall, this research paper revolves around understanding and manipulating song data, whether it involves only musical instruments, acoustic compositions, or a combination of instruments and vocals [7].

Music holds a significant place in Kazakhstan's cultural heritage, reflecting the rich history and traditions of the Kazakh people. The traditional music of Kazakhstan, including both instrumental and vocal genres, has deep roots and has evolved over centuries. Like many other countries, Kazakhstan has a rich musical tapestry, blending traditional Kazakh music with modern genres from around the world. It's important to note that the success of music genre classification using machine learning depends on the quality of the dataset, the selection of relevant features, and the choice of an appropriate model. Additionally, incorporating domain knowledge about Kazakh music and culture can enhance the accuracy and cultural relevance of the classification system

**Literature Review and problem statement**

The literature review underscores the evolving role of ML in music genre classification, highlighting the challenges and advancements in the field. This background sets the stage for the project's objective of developing an ML model to classify songs into Rock or Hip-Hop genres, contributing to the broader field of MIR. The connection of machine learning and music has gained significant attention in recent years, with researchers exploring the development of models capable of automatically classifying music genres. This literature review observes key studies that contribute to the deeper understanding of machine learning models applied to the music genre classification.

In reference [8] authors mainly focused on the practical research using GTZAN dataset and showing the readers different kinds of models that can help to classify audio files into various kinds of music genres. They had approximately 1000 music files and gave us figures and tables with visual information about the overall methodology and formats using spectrogram files and training/test data. Comparisons and testing have been made using Keras framework and Google Colab.

The [9] research introduces an original approach by combining audio and lyrics features for music genre classification. The study explores the complementary nature of audio and information in the text format, employing a "deep neural network architecture" to integrate both modes. This specific work highlights the potential for multimodal approaches to enhance the accuracy and robustness of music genre classification models.

In [10] research paper, authors used the same dataset and techniques (CNN) as the background and framework of their study, to analyze and compare the results and overall information about the whole topic. The main difference that they have in their research is that they did not use visual features and representations with CNN as classifier, the authors represent music segments in the dataset by mel frequency cepstral coefficients (MFCC). Mainly, it was used for reprocessing the music segments in the dataset and only then train the CNN model. To conclude, the results were positive and successful with 94.5 percent of accuracy.

The [11] research is mostly about transfer learning, a technique where a model trained on one task and is adapted for another related task. The authors, for their research, used given technique in the context of music genre classification. The study explores the "transferability" of features learned from a large dataset to improve the performance on a smaller target dataset. Experimental results showed that the transfer learning way can achieve a higher average classification accuracy (95.8 percent) than the same deep Recurrent Neural Network (RNN) which marks the parameters randomly (93.7 percent).

Authors in reference [12] proposed an end-to-end learning approach for music audio analysis, eliminating the need for handcrafted feature engineering. The study introduces a "convolutional neural network architecture" that is capable of directly processing raw audio waveforms. This approach abridged the model pipeline and showcased the overall potential of end-to-end learning for tasks like music genre classification.

This literature review showed the overall evolution of machine learning models for music genre classification, with a range from early comparative studies to the integration of deep learning and multimodal approaches. Possibly, future research in this field may focus on addressing the issues like scalability, interpretability and the integration of additional modalities to advance the possibilities of music genre classification systems.

**Purpose and objectives of the study**

*This project aims* to develop a machine learning model capable of classifying songs based solely on their audio features. The approach involves processing and analyzing audio data, reducing feature dimensions, and applying classification algorithms.

*The scientific novelty* of this research lies in its potential to enhance music recommendation systems and support artists and producers in understanding the categorization of their music in digital platforms.

*The objects of the study* are the sound characteristics of Kazakh music's over the past 8 years.

In the course of the work the following tasks were set:
1. Loading and preparing the dataset.
2. Pairwise relationships between continuous variables
3. Normalize and scale the data
4. Compare decision tree model to a logistic regression
5. Balance data for greater performance
6. Evaluate models.

**Materials and methods**

In order to make a comparative analysis of the considered software tools, the materials of Scopus, Web of Science, ResearchGate international databases on the mentioned topic were analyzed, and machine learning algorithms for increasing the productivity of agricultural crops were considered.

The Kazakh_music dataset is a widely used benchmark dataset for music genre classification, which consists of 1000 audio tracks equally distributed across 5 different genres, dataset was manually collected in 2016-2023. The genres in the dataset are Kazakh Rock, Kazakh pop music, Q-pop, Zazz and Toi. Each audio track is 30 seconds long, sampled at 25,050 Hz, and stored in the WAV format. The dataset is carefully curated and annotated with ground truth labels by music experts.

*Data Sources*

CSV File: Contains basic track information and genre labels.

JSON File: Includes musical features like danceability and acousticness.

*Tools and Technologies*

Programming Language: Python

Libraries: Pandas for data manipulation, Matplotlib and Seaborn for data visualization, scikit-learn for machine learning.

*Methodology*

Data Cleaning: Handling missing or irrelevant data in the datasets. Exploratory Data Analysis: Visualizing data to uncover patterns and insights.

Feature Reduction: Using techniques like PCA to reduce feature dimensions while retaining variance.

Model Training: Implementing Decision Tree and Logistic Regression models and evaluating their performance.

*Heatmap*

A heatmap is a graphical representation of data in which values are depicted using a color scale. Mathematically, a heatmap can be represented as a matrix $M$, where each element $Mij$ corresponds to the value of a variable at position $(i,j)$ in the dataset. Once you have this matrix representation, you can apply various mathematical operations or visualization techniques to analyze or display the data. For example, it is uses to calculate statistics like the mean or standard deviation of the values in the matrix, or for to visualize the matrix as an image where different colors represent different values [13].

*Logistic regression*

Logistic regression predicts "probability value" through a linear combination of the given features plugged inside a logistic function given as below (1):

$$p(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1)}} \tag{1}$$

where $\beta_0 = \frac{-\vartheta}{k}$ –intercept, $\vartheta = \frac{\beta_0}{\beta_1}$, $k = \frac{1}{\beta_1}$, $x$ as the input feature vector (a column vector of predictors or independent variables), $y$ as the binary output variable (0 or 1, representing the two classes).

The hypothesis function in logistic regression is defined as (2):

$$h_\theta(x) = \sigma(\theta^T x) = \frac{1}{1 + e^{\theta^T x}} \tag{2}$$

Here, $\theta^T x$ is the dot product of the parameter vector and the feature vector [14].

*Decision Tree*

A decision tree is a popular supervised machine learning algorithm used for classification and regression tasks. It works by partitioning the data into subsets based on certain attributes/features, with the goal of creating a tree-like structure of decisions. Each internal node represents a "decision" based on the value of a feature, leading to different branches, and each leaf node represents the outcome or predicted value [15].
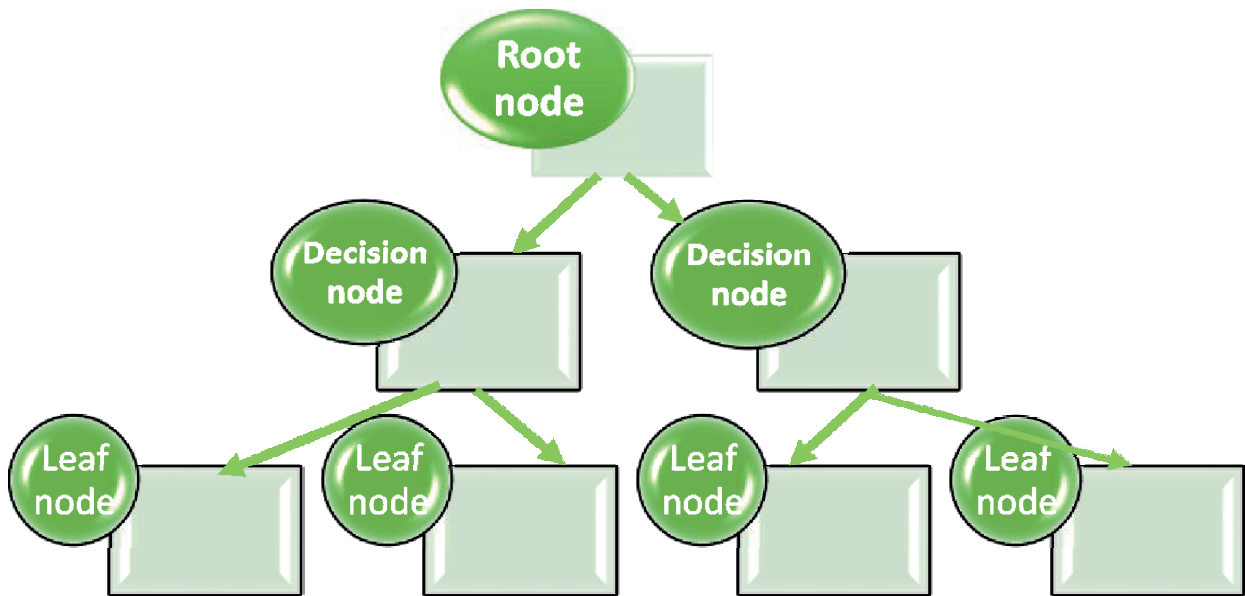


Figure 1. Decision Tree structure

*Gini Impurity in Decision Tree*

The algorithm selects the partition that minimizes impurities and maximizes the purity of the split. Informally, the impurity is a gauge of the similarity of the labels at the node.

Let's assume that a dataset $D_{Tree}$ contains examples from n classes. Its Gini Index, $Gini(D_{Tree})$, is defined as (3):

$$Gini(D_{Tree}) = 1 - \sum_{j=1}^{n} p_j^{2} \tag{3}$$

where $p_j$ is the relative frequency of class $j$ in $n$. The lowest value is zero, which signifies a node that contains the elements of the same class [16].

**Research Results and Discussion**

*1. Load and prepare the dataset*
The heat map between the musical characteristics of a given dataset was computed.

| | acousticness | danceability | energy | instrumentalness | liveness | speechiness | tempo | valence |
|---|---|---|---|---|---|---|---|---|
| acousticness | 1.000000 | 0.099400 | -0.316881 | 0.010640 | -0.004815 | 0.100149 | -0.088235 | 0.006583 |
| danceability | 0.099400 | 1.000000 | -0.257554 | -0.180838 | -0.186983 | 0.116911 | -0.183284 | 0.496265 |
| energy | -0.316881 | -0.257554 | 1.000000 | 0.173595 | 0.160971 | -0.014125 | 0.238089 | 0.004980 |
| instrumentalness | 0.010640 | -0.180838 | 0.173595 | 1.000000 | -0.047121 | -0.129386 | 0.003878 | -0.122700 |
| liveness | -0.004815 | -0.186983 | 0.160971 | -0.047121 | 1.000000 | 0.160832 | 0.000469 | -0.083941 |
| speechiness | 0.100149 | 0.116911 | -0.014125 | -0.129386 | 0.160832 | 1.000000 | -0.057482 | -0.060068 |
| tempo | -0.088235 | -0.183284 | 0.238089 | 0.003878 | 0.000469 | -0.057482 | 1.000000 | 0.080164 |
| valence | 0.006583 | 0.496265 | 0.004980 | -0.122700 | -0.083941 | -0.060068 | 0.080164 | 1.000000 |

Figure 2. Heat map of musical features of tracks

From the above graph, we can see that none of the features have a strong correlation. Therefore, it is not necessary to remove any attributes from the given data.

*2. Normalize the feature data*
The importance of simplifying models and using as few features as possible without sacrificing performance was mentioned above. Since we did not find any strong correlations between our features, we can use a common dimensionality reduction technique called principal component analysis (PCA). PCA works by rotating the data along the axis of highest variance, which allows us to identify the features that contribute the most to the differences between classes.

However, since PCA uses the absolute variance of a feature to rotate the data, features with a wider range of values can overshadow and bias the algorithm. To avoid this, we must first normalize the data, which means scaling all features so that they have a mean of 0 and a standard deviation of 1 [17].

*3. Principal component analysis on scaled data*
After preprocessing our data, we can use Principal Component Analysis (PCA) to reduce its dimensionality. Scree plots and cumulative explained ratio plots can help us determine the number of components to retain for further analysis. Scree plots show the number of components on the x-axis and the variance explained by each component on the y-axis, sorted by decreasing variance. This helps us visualize which components explain the most variance in our data. To determine an appropriate cutoff, we look for an "elbow" in the plot, where there is a steep drop in variance from one component to the next (Fig. 3).
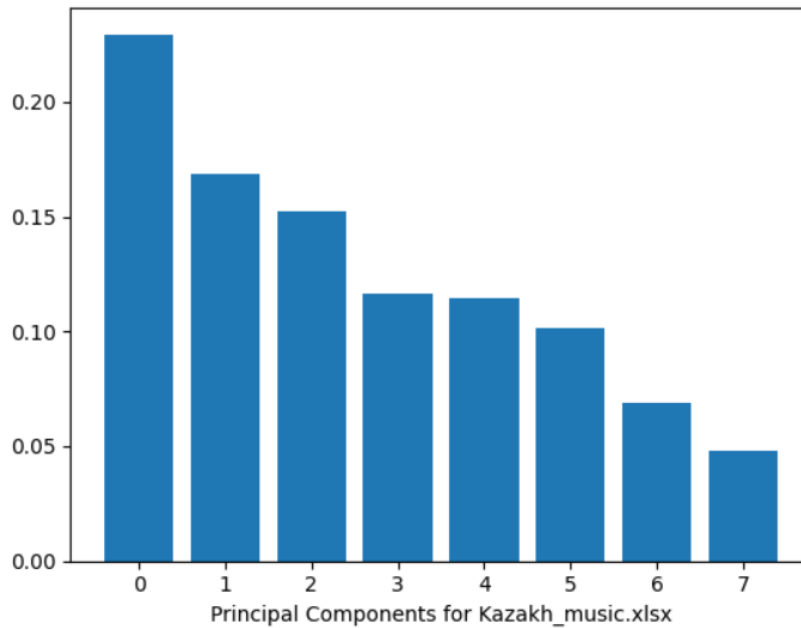
Figure 3. Principal components for Kazakh_music dataset

Unfortunately, there does not appear to be a clear elbow in this scree plot, which means it is not straightforward to find the number of intrinsic dimensions using this method.

### 4. Further visualization of PCA

If the scree plot doesn't help us decide how many features to keep, we can use the cumulative explained variance plot. This plot shows how much variance each component explains, and we can use it to find the point where 90% of the variance is explained. This point represents the number of features we need to keep (Fig. 4).
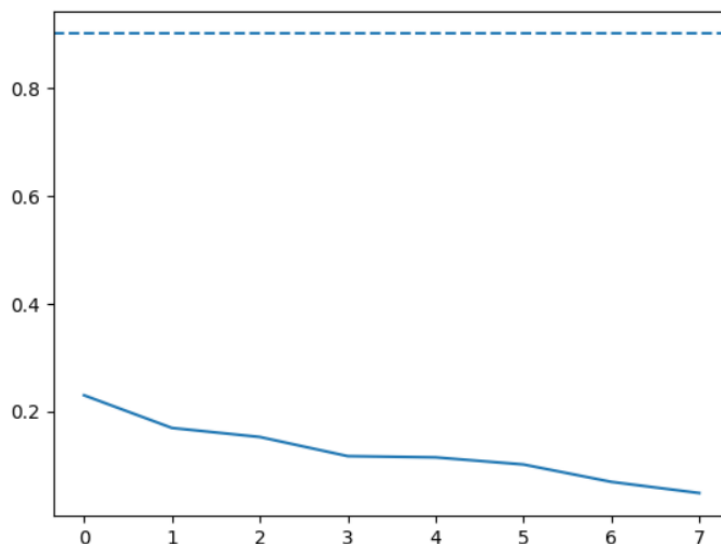


Figure 4. Variance plot PCA visualization

Furthermore, its used to classify songs into genres the lower dimensional PCA projection of the data. Before dataset was spitted into 'train' and 'test' subsets, where the 'train' subset will be used to train our model while the 'test' dataset allows for model performance validation.

*5. Train a decision tree and a logistic regression model to classify the genre*

Decision trees are rule-based classifiers that work by asking a series of binary questions about the data until they reach a conclusion. Decision Tree model was built to classify music genres, Gini index was used for splitting criteria [18].
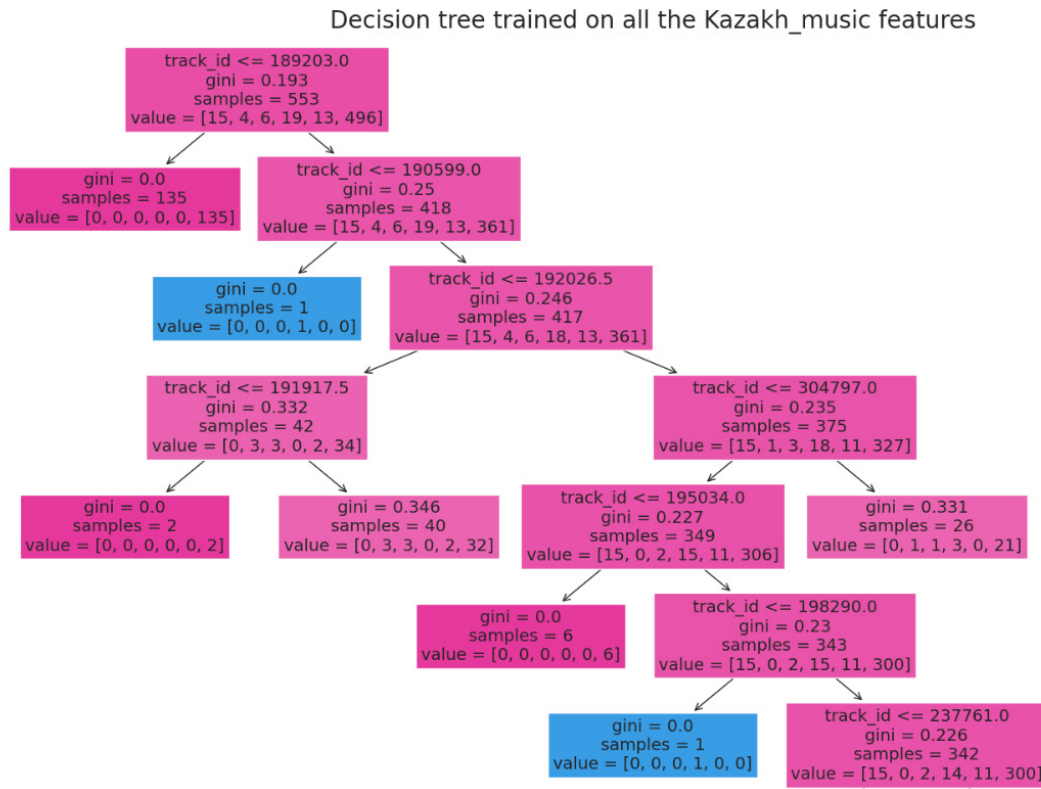


Figure 5. Decision tree model visualization of given dataset

Table 1. Impurities of decision tree model of Kazakh_music dataset

| Impurities of Decision tree music genre classification model | |
|---|---|
| $Gini(D_{Tree})$ | array [0.1928851, 0., 0.24951352, 0., 0.24612253, 0.33219955, 0., 0.34625, 0.23478044, 0.22651702, 0.  0.2301252, 0., 0.22586437, 0., 0.23016262, 0.23397788, 0. 0.33136095]) |

Thus, the Decision tree music genre classification model (Fig. 5) was constructed, using the Gini index to calculate the probability that at random selection a certain feature will be classified incorrectly. Next, we use Logistic regression to classify genres and then compare the classification accuracy of both methods.

*6. Logistic Regression Model*

Lets assume that $y$ is the desired probability of class membership, $x_1, x_2, x_3, ...$ are the features of the audio tracks that are used for prediction, and $\beta_0, \beta_1, \beta_2, \beta_3, ...$ are the coefficients describing the effect of each characteristic of the given Kazakh_music.csv dataset on the probability.

The logistic regression model of genre classification of given dataset is (4):

$$p(x_{genre}) = \frac{1}{1 + e^{-1.0765}}, \tag{4}$$

Sigmoid function is used for optimization, which basically compresses the entire function in the range of 0 to 1 [19].

© Aigul Mimenbayeva, Gulmira Bekmagambetova,
   Gulzhan Muratova, Akgul Naizagarayeva,
   Tleugaisha Ospanova, Assem Konyrkhanova

Table 1. Logistic regression model attributes for Kazakh_music dataset

| intercept | ([ 0.96465051, -2.04110411, -1.82763437, -1.69836001, -0.55197223, 5.15442021]) |
|---|---|
| coefficients | [[-0.290511, 0.481960, -0.60660100, -0.23564811, 0.19516111, 0.469844, 0.45450513, 0.04646012], [-0.38772, -0.93905, 0.524279, 0.24939019, -0.53690511, -0.93844211, 0.405309821 0.0446847], [0.1859421, -0.38303, 0.81202414, -0.86606412, -0.25127821, -0.81656126, 0.090108, -0.2274476], [0.5275934, -1.48951, 0.338776, -1.39109234, -0.1614646, -1.20359806, 0.51258125, 0.217572], [-0.248506, 1.281675, -1.253564, 1.51284903, 0.84236487, 1.61254275, -0.901272, -0.255725], [ 0.213209, 1.047963, 0.185094, 0.7305663, -0.0878782, 0.87621433, -0.56123212, 0.17445588]]) |

Table 1 shows the coefficients of Logistic regression model for Kazakh_music dataset.

### 7. Compare decision tree model to a logistic regression

Once we have trained two models, we can compare how well they predict by looking at how many data points they get wrong.

```
Decision Tree:
                      precision    recall   f1-score    support

             Jazz        0.00       0.00      0.00          5
      Kazakh Rock        1.00       0.50      0.67          2
   Kazakh hip hop        0.33       0.33      0.33          3
 Kazakh pop music        0.00       0.00      0.00          4
  Classical music        0.72       0.79      0.75         33

         accuracy                             0.60         47
        macro avg        0.41       0.32      0.35         47
     weighted avg        0.57       0.60      0.58         47

Logistic Regression:
                      precision    recall   f1-score    support

             Jazz        0.00       0.00      0.00          5
      Kazakh Rock        0.00       0.00      0.00          2
   Kazakh hip hop        0.00       0.00      0.00          3
 Kazakh pop music        0.00       0.00      0.00          4
  Classical music        0.70       1.00      0.82         33

         accuracy                             0.70         47
        macro avg        0.14       0.20      0.16         47
     weighted avg        0.49       0.70      0.58         47
```

Figure 5. Comparative table of results of Decision Tree Classifier and Logistic Regression models

Both models do similarly well, boasting an average precision of 60% and 70%. However, looking at our classification report, we can see that rock song are fairly well classified, but hip-hop songs are disproportionately misclassified as rock songs.

*8. Balance data for greater performance*

To compensate for the fact that there are more data points in some classes than others, we can give more weight to the correct classification of data points in the smaller classes. This will help us to train a model that is more accurate for all classes (Fig. 6).

```python
hop_only = echo_tracks.loc[echo_tracks["genre_top"] == "Hip-Hop"]

rock_only = echo_tracks.loc[echo_tracks["genre_top"] == "Rock"].sample(len(hop_only), random_state=10)

rock_hop_bal = pd.concat([rock_only, hop_only])

features = rock_hop_bal.drop(['genre_top', 'track_id'], axis=1)
labels = rock_hop_bal['genre_top']
pca_projection = pca.fit_transform(scaler.fit_transform(features))

train_features, test_features, train_labels, test_labels = train_test_split(pca_projection, labels, random_state=10)
```

Figure 6. Code fragment for balancing the original dataset

We've now balanced our dataset, but in doing so, we've removed a lot of data points that might have been crucial to training our models.

*9. Does balancing our dataset improve model bias?*

We have already made our dataset smaller, so we will not use any more dimensionality reduction techniques. We would normally use dimensionality reduction more carefully when working with very large datasets or when it takes too long to run our model. Balancing our data has fixed the problem where our model was biased towards the more common class. To get a better idea of how well our models are actually working, we can use a technique called cross-validation (CV). CV lets us compare models more fairly.

*10. Using cross-validation to evaluate models*

Since the way our data is split into train and test sets can impact model performance, CV attempts to split the data multiple ways and test the model on each of the splits. Although there are many different CV methods, all with their own advantages and disadvantages, we will use what's known as *K-fold cross-validation* here. K-fold first splits the data into K different, equally sized subsets. Then, it iteratively uses each subset as a test set while using the remainder of the data as train sets [20]. Finally, we can then aggregate the results from each fold for a final model performance score (Fig. 7).

```python
from sklearn.model_selection import KFold, cross_val_score

kf = KFold(n_splits=10, random_state=10, shuffle=True)

tree = DecisionTreeClassifier(random_state=10)
logreg = LogisticRegression(random_state=10)

tree_score = cross_val_score(tree, pca_projection, labels, cv=kf)
logit_score = cross_val_score(logreg, pca_projection, labels, cv=kf)

print("Decision Tree:", np.mean(tree_score), "Logistic Regression:", np.mean(logit_score))

Decision Tree: 0.7719780219780219 Logistic Regression: 0.823076923076923
```

Figure 7. Code snippet for cross validation

Now, that we have performed k fold cross-validation on our dataset, we can be pretty sure that our model will generalize 77% of the times on the future unseen data points.

**Conclusion**

This report presents a comprehensive study on the classification of music genres using machine learning techniques, with a specific focus on distinguishing between 5 types of genres. The analysis demonstrates the potential of machine learning algorithms to effectively categorize music based on audio data, without relying on human listening.

Key takeaways from the research include:

Effectiveness of Machine Learning in Music Classification: The study successfully employed machine learning algorithms, namely decision trees and logistic regression, to classify music genres. This underscores the capability of these algorithms in handling complex pattern recognition tasks like music genre classification.

Critical Role of Data Preparation and Feature Analysis: The study highlighted the importance of meticulous data preparation, including loading, cleaning, and standardizing data. The absence of strong correlations among features and the application of principal component analysis (PCA) for dimensionality reduction were crucial steps that enhanced the efficiency and effectiveness of the models.

Importance of Data Balance: Addressing the imbalance in the dataset was a pivotal aspect of the study. Balancing the number of data points for each class was essential in reducing model bias and ensuring that the model's accuracy was not skewed towards the more prevalent class.

Model Validation and Reliability: The use of cross-validation, specifically the K-fold method, provided a thorough evaluation of the model's performance. This step was instrumental in assessing the generalization capability of the models, ensuring that they could reliably predict genre classifications on new, unseen data.

Implications for Music Streaming Services: The findings of this study have significant implications for music streaming services. music genres using machine learning can lead to more precise and user-friendly music recommendation systems. This advancement can enhance user experience on streaming platforms, helping users discover music that aligns more closely with their tastes.

In conclusion, this research not only demonstrates the practical application of machine learning in the field of music genre classification but also sets a precedent for future research. It opens avenues for exploring more sophisticated machine learning models and incorporating a broader range of musical features to further refine genre classification systems. The methodologies and insights gained from this study can be pivotal in advancing digital music platforms, making them more responsive and tailored to individual user preferences.

## References

[1]    Patil, S.A., Pradeepini, G., & Komati, T.R. (2023). Novel mathematical model for the classification of music and rhythmic genre using deep neural network. *J Big Data*, 10, 108. https://doi.org/10.1186/s40537-023-00789-2

[2]    Liu, C., Feng, L., Liu, G., & Shenglan, L. (2021). Bottom-up broadcast neural network for music genre classification. *Multimedia Tools Application*, 80, 7313-7331. https://doi.org/ 10.1007/s11042-020-09643-6

[3]    Ghildiyal, A., Singh, K., & Sharma, S. (2020). Music genre classification using machine learning. *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 1368–1372. https://doi.org/10.1109/ICECA49313.2020.9297444.

[4]   Cheng, Y., Chang, P., Nguyen, D., & Che-Nan, K. (2021). Automatic music genre classification based on CRNN. *Eng Lett.*, 29(1), 36-48.

[5]   Foleis, J.H., & Tavares, T.F. (2020). Texture selection for automatic music genre classification. *Appl Soft Comput.*, 89, 106127.

[6]   Elbir, A., & Aydin, N. (2020). Music genre classification and music recommendation by using deep learning. *Electron Lett.*, 56(12), 627–9.

[7]   Ramírez, J., & Flores, M.J. (2020). Machine learning for music genre: multifaceted review and experimentation with audio set. *J Intell INF Syst.,* 55, 469–499. https://doi.org/10.1007/s10844-019-00582-9

[8]   Mounika, K., Deyaradevi, S., Swetha, K., & Vanitha, V. (2021). Music Genre Classification Using Deep Learning. *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA),* 1-7. https://doi.org/10.1109/ICAECA52838.2021.9675685

[9]   Pelchat, N., & Gelowitz, C.M. (2020). Neural network music genre classification. *Can J Electr Comput Eng.*, 43(3), 170–173. https://doi.org/10.1109/CJECE.2020.2970144.

[10]   Suero, M., Gassen, C.P., Mitic, D., Xiong, N., & Leon, M. (2020). A Deep Neural Network Model for Music Genre Recognition. *Advances in Intelligent Systems and Computing*, 1074. https://doi.org/10.1007/978-3-030-32456-8_41

[11]   Liang, B., & Gu, M. (2020). Music genre classification using transfer learning. *2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 392-393. https://doi.org/10.1109/MIPR49039.2020.00085.

[12]   Ghosal, S., & Sarkar, I. (2020). Novel approach to music genre classification using clustering augmented learning method (CALM). *AAAI Spring Symposium Combining Machine Learning with Knowledge Engineering,* 29 (1).

[13]   Parmezan, A., Silva, D., & Batista, G. (2020). A combination of local approaches for hierarchical music genre classification. *Proceedings of the 21st ISMIR conference, Montreal, Canada, October 11–16.* http://dx.doi.org/10.5281/zenodo.4245540

[14]   Yuan, C., Ma, Q., Chen G., Jhou W., Zhang X., Tang X., Han J., & Hu S. (2020). Exploiting heterogeneous artist and listener preference graph for music genre classification. *Poster Session A3: Multimedia search and recommendation & multimedia system and middleware.*

[15]   Zhuang, Y., Chen, Y., & Zheng, J. (2020). Music genre classification with transformer classifier. *Proceedings of the 2020 4th international conference on digital signal processing*, 155-159. https://doi.org/10.1145/3408127.3408137

[16]   De Araújo, L.R., de Sousa, R., & Barbosa, S.D. (2020). Brazilian lyrics-based music genre classification using a BLSTM network. *Lecture notes in computer science*, 415, 522-534. https://doi.org/10.1007/978-3-030-61401-0_49

[17]   Budhrani, A., Patel, A., & Ribadiya, S. (2021). Music2Vec: Music Genre Classification and Recommendation System. *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India*, 1406-1411 http://dx.doi.org/10.1109/ ICECA49313.2020.9297559

[18]   Chang, S., Abdul A., Chen J., &, Liao, H. (2018). A personalized music recommendation system using convolutional neural networks approach. *2018 IEEE International Conference on Applied System Invention (ICASI)*, 47-49. https://doi.org/10.1109/ICASI.2018.8394293

[19]   Hyung, Z., Park, J.S., & Lee, K. Utilizing context-relevant keywords extracted from a large collection of user-generated documents for music discovery. *Info. Processing and Management*, 53 (5), 1185-1200. https://doi.org/10.1016/j.ipm.2017.04.006

[20]   Jandaghian, M., Setayeshi, S., Razzazi, F., & Sharifi, A. (2023). Music emotion recognition based on a modified brain emotional learning model. *Multimed Tools Appl.*, 82, 26037–26061.