



Audio Feature-Based Musical Mode Classification Using Machine Learning Techniques

Hira Farman^{1,*}, Khurshed Ali², Saif Hassan², Muhammad Ismail Mangrio², Danish Ali Raza¹, Muhammad Makki¹

¹Department of Computer Science, Iqra University, Karachi, Sindh, Pakistan

²Department of Computer Science, Sukkur IBA University, Sukkur, Sindh, Pakistan

*Corresponding author: hira.farman@iqra.edu.pk

ARTICLE INFO

Article History:

Received: July 30, 2026
 Revised: August 20, 2026
 Accepted: September 25, 2026
 Available Online: September 30, 2026

Keywords:

Music Data Analysis
 Spotify
 Musical Mode Classification
 Audio Feature Analytics
 Song Popularity Prediction

Classification Codes:

Funding:

This research received no specific grant from any funding agency in the public or not-for-profit sector.

ABSTRACT

Spotify's meteoric rise has changed the music landscape, moving music from physical and downloaded formats to streaming, data-driven ones. Machine learning and music analytics are new research avenues for examining musical attributes, listening behavior, and song characteristics in light of the growing amount of audio and streaming data. In this study, a supervised machine learning model for classification of musical mode (Major or Minor) of the music from Spotify is proposed based on Acoustic and Streaming related features. The data provided from Kaggle is made up of 953 songs with characteristics such as BPM, danceability, energy, valence, acousticness, instrumentalness, liveness, speechiness, playlist appearances, and streaming statistics. 95 songs were deleted listwise for having missing musical-key values, leaving 858 songs for analysis, of which 495 are in major and 363 in minor keys. Therefore the majority class baseline accuracy was set at 57.7% which would be used to compare all the classification models. Five classification algorithms were used and compared: Logistic Regression, Random Forest, Decision Tree, K-Nearest Neighbors, and Naïve Bayes. Only data from the training folds were used for all pre-processing steps, including normalization, correlation-based feature selection and SMOTE oversampling, to avoid data leakage. The performance of the models was determined through repeated stratified 10-fold cross-validation with 10 trials and mean $\pm$ SD was reported. The accuracy, precision, recall, F1 score, Matthews correlation coefficient (MCC), area under the receiver operating characteristic curve (ROC-AUC) and confusion matrix were used to evaluate performance. Random Forest had the best predictive performance with an accuracy of 0.983, and an F1-score of 0.982, precision of 0.983, recall of 0.983, MCC of 0.964, and ROC-AUC of 0.999. Decision Tree had the highest accuracy and F1 score (0.946), followed by



Logistic Regression, K-Nearest Neighbors, and Naïve Bayes which showed relatively lower accuracy and F1 score. A leakage audit was also performed: the variable stream-count was not included in the predictor set as it served as a near unique record identifier, and all reported metrics were produced by computing predictions of the out-of-fold stream data as opposed to fitted predictions. Danceability, energy, valence, BPM and acousticness were found to be important attributes for the classification of musical modes into Major and Minor. The results show the promise of machine learning in music analytics and could inform recommendations, playlist optimization, artist strategy and intelligence of streaming platforms. Overall, the proposed framework provides a reproducible, data-driven approach to musical mode classification using Spotify song characteristics.

© 2026 The authors published by JCIS. This is an Open Access Article under the Creative Common Attribution Non-Commercial 4.0

Corresponding Author's Email: hira.farman@iqra.edu.pk

Citation: Farman, H., Raza, D. A., Hassan, S., Ali, K., Mangrio, M. I., & Makki, M. Audio Feature-Based Musical Mode Classification Using Machine Learning Techniques. Journal of Computers and Intelligent Systems.

1. Introduction

The entire music industry has been transformed by digital technology and Internet streaming. The old model of music sales by physical albums and downloadable music is being phased out by a streaming economy that enables the listener to enjoy millions of songs within their reach. Of these, Spotify has become one of the biggest and most influential music streaming services, providing a vast library of tracks and producing vast amounts of audio, behavioral and streaming data. Music streaming apps have also brought about the importance of music analytics and data-driven decisions to the entertainment industry. Some of the traits that can be captured by streaming platforms include music quality like acousticness, danceability, energy, valence, musical mode, speechiness, loudness, and tempo[1-3]. These features can be useful for understanding the emotional, structural and rhythmic properties of songs and studied through machine learning and data mining techniques for music patterns related to listeners' behavior and music trends. Machine learning has gained significant traction in the field of music recommendation systems, popularity prediction, playlist optimization, and audience targeting, among other areas, in recent years. There are various supervised learning techniques that have been employed to predict popularity, engagement and genre of songs using Spotify features [4-10].

The features that have been suggested as contributing to audience interaction and streaming performance include danceability, energy, BPM and valence. However, the musical mode is a topic that has been explored rather little although many works have been undertaken in the domain of music analytics. Musical mode refers to the pattern of tones (or scale pattern) in musical composition, which can be either Major or Minor. Songs in the major mode are often cheerful, energetic, uplifting, and positive, while songs in the minor mode are sadder, emotional, tense or melancholy, and/or more introspective. Musical mode and Spotify audio features represent a crucial research challenge in computational music analytics because listeners' perception and experience of music is largely influenced by music modes. From this, the authors of this study

come with an idea to overcome this by using supervised machine learning techniques to classify the music that can be found in Spotify into Major and Minor modes using both acoustic and streaming attributes.

This study examines the features of songs from the Spotify API used to classify songs into their respective modes, which are pulled from the Kaggle repository of Spotify song data, and considers how features like mode, danceability, energy, valence, BPM, acousticness, liveness, speechiness, and streamability affect mode classification. The accuracy of multiple machine learning classification models, such as Logistic Regression, Random Forest, Decision Tree, K-Nearest Neighbors (KNN), and Naïve Bayes, are comparatively analyzed to identify the best classification model for the prediction of musical modes. The results of this study can be useful for artists, music producers, streaming companies, playlist curators as well as marketers in order to comprehend the emotional and tonal song structure. In addition, the study adds to the body of computational music analysis research and shows the potential of machine learning to assist in data-informed decisions in modern digital music ecosystems.

1.1 Motivation of the Study

The music data streaming opens new opportunities in the music industry with machine learning and artificial intelligence. When you use a streaming service such as Spotify, it gathers a lot of data regarding the songs, the way users use the songs, playlists, and more. This information can be referred to for the understanding of the structure of music, emotional dynamics, listening patterns and commercial performance. Other musical properties are emotion and song tone, and one of the most significant musical attributes of a song is the song's musical mode. Major mode songs are more happy, more positive, more exciting, more energetic, Minor mode songs are sadder, more emotional, more reflective. When listeners are interested, they listen to the music; knowledge of how the audio features at Spotify relate to the modulation of the music is crucial to reaching listeners' ears. This is relevant because music listening behavior will be linked to emotional involvement in music, and the relation of listening action to music mode is more relevant when we seek to understand the listeners' listening behavior towards music. Emotional music perception is important and most of the studies on this topic have focused on predictability of popularity, music recommendation systems and music genre classification, which are less studied by using machine learning model for music mode classification. Moreover, most of the previous research is limited to just a few elements of songs or to not investigate how multiple audio features relate to the songs' tonal and emotional structure. This research is to observe if the audio and streaming functionality of the Spotify can be used to correctly classify songs as belonging to either major or minor category, using supervised machine learning techniques. The study also seeks to shed light on the importance of the acoustic parameters in musical mode prediction and in the representation of emotional songs.

1.2 Problem Statement

Due to digital streaming services, music-related data has been increasing, including in-depth Acoustic, Emotional and Streaming Data. There are many attributes for the audio in Spotify that may have some pattern when classified by musical modes: danceability, energy, valence, BPM, acousticness, speechiness and liveness. But, the relationship between these features and the classification of songs in Major and Minor modes is still under-explored. The main challenge tackled in this study is whether machine learning algorithms can classify Spotify songs with

accuracy to Major and Minor categories, based on audio and streaming features. The classification problem can be mathematically represented as in eq.(1):

$$\hat{Y} = f(X) \quad f: X \rightarrow y, \quad \hat{y} = f(X) \quad (1)$$

where:

- X represents Spotify audio and streaming-related features,
- $\hat{Y}$ represents the predicted class label.

The target variable is defined as:

$$Y \in \{0,1\}$$

where:

- $Y = 1$ represents Major mode,
- $Y = 0$ represents Minor mode.

The probabilistic relationship between the target variable and predictors can be expressed as:

$$P(Y | X_{danceability}, X_{energy}, X_{valence}, X_{BPM}, X_{acousticness}, \dots)$$

The study further seeks a classification function capable of minimizing prediction error shown in Eq.(2):

$$f^* = \arg \min_f (1/n) \sum_{i=1}^n (y_i \neq f(x_i)) \quad (2)$$

where:

- n represents the total number of observations,
- $\mathbb{I}$ denotes the indicator function.

This study investigates whether Spotify audio features can effectively capture emotional and tonal differences between Major and Minor songs using supervised machine learning techniques.

1.3 Objectives of the Study

The main goal of this research is to create a framework based on machine learning for distinguishing between music modes of Spotify songs using audio and streaming- based features, such as Major and Minor. Specific goals of the study are:

1. To study the characteristics of the songs in Spotify, such as: danceability, energy, valence, BPM, acousticness, liveness, speechiness..
2. To examine the correlation among musical mode classification and features of Spotify's song.
3. To preprocess and prepare Spotify song data for supervised machine learning analysis.
4. To implement and compare multiple machine learning classification algorithms, including:
 - Logistic Regression
 - Random Forest
 - Decision Tree
 - K-Nearest Neighbors (KNN)
 - Naïve Bayes
5. To analyze and evaluate model performance using evaluation metrics like Accuracy, Precision, Recall, F1-score, ROC-AUC, Matthews Correlation Coefficient (MCC), and confusion matrix analysis.
6. To identify the most influential audio features contributing to Major and Minor song classification.
7. To determine the best-performing machine learning model for musical mode prediction.
8. To provide insights useful for music recommendation systems, playlist optimization, streaming analytics, and emotional music analysis.

The probabilistic objective of classification may be represented as in Eq.(3):

$$p(X) = P(y = \text{Major} | X) \quad (3)$$

where:

- p_i denotes the probability that the i^{th} song belongs to the Major mode class,
- X_i represents the Spotify feature vector of the song.

Classification is then obtained using the threshold rule shown in Eq.(4):

$$\hat{y} = 1 \text{ if } p(X) \geq 0.5, \quad \hat{y} = 0 \text{ otherwise} \quad (4)$$

1.4 Significance of the Study

The research builds on the field of computational music analytics, enabling the use of machine learning for the classification of the musical mode of Spotify songs using acoustic and streaming-related features. The study gives insight into the correlation between the musical characteristics (danceability, energy, valence, acousticness and BPM) and the classification of the musical mode. The results could be applied in music recommendation systems, playlist generation systems, mood classification systems in music, and personalized delivery of music on streaming platforms. Correct identification of musical mode can enhance the structure and retrieval of music by allowing music information systems to integrate tonal information into the music recommendation and optimization of playlists. The results can also help artists, producers and music analysts in their understanding of the relationships between measurable audio characteristics and musical modes. Academically, the study brings together machine learning, music analytics, and streaming data in a seamless classification system. Multiple classification algorithms have been compared to systematically evaluate how well they can classify the songs as either Major or Minor. Hence, the study is of value for the creation of reproducible, data-driven methods to analyze music attributes and can inspire further work in the area of A-I, MIR and streaming platform analysis.

2. Background

The world of music has been revolutionized by the digital revolution. It's not a CD or a record it's a sensation that you can only have online. It wasn't as easy as it sounds, however. Internet piracy began to make a huge dent in industry revenue in the early 2000s, driving global earnings from \$22 billion in 1999 to almost \$13 billion in 2014. The response from the industry was in the form of streaming services which were easily scalable, consumer friendly and provided an alternative to piracy and physical media. Spotify was one of them, for example, that since the beginning of its existence has had a free 'ad-supported' streaming service as well as a paid 'premium' streaming service, thus providing a new distribution and consumption model for music in the digital age [1]. The key distinction between streaming and traditional broadcast media such as radio and TV is that there are choices involved in streaming. Listeners no longer have to wait for anything a station wants to play, but can now access tens of millions of songs on-demand from any device, anytime. This has made platforms such as Spotify, Apple Music, Deezer, Gaana, JioSaavn, and Amazon Prime Music the pillars of the entertainment industry in today's world. Spotify has been successful in terms of platform, with its personalisation algorithms, analytics and user base size being its primary successes [2]. The numbers that appear in the Spotify Financial Report 2023 are representative of this domination [3]: "Spotify's total users per month are over 602 million, with over 220 million being premium subscribers".

The other major effect of this streaming wave is that there is a lot of music data out there, it is organised and it is available to use. For example, Spotify has a large suite of audio features for music that can be stored in its library: danceability, energy, tempo, BPM, valence, speechiness, acousticness, musical mode and more. The data have turned out to be very valuable for researchers and practitioners who want to be able to quantitatively dissect listener preferences, emotional

reactions, and new trends in listener consumption. This type of analysis capability was highlighted over 10 years ago, when both Barton and Court [4] and Davenport [5] emphasized that big data is important for business intelligence. In particular, Ingle, Shah and Mehta [6] showed that the Spotify data ecosystem could be an appropriate environment for modelling consumer behavior and for conducting research. Ever since, Spotify data has been used in increasingly complex machine learning applications, and the field has grown into a thriving one. Beesa et al. [2] did exploratory analysis on the Spotify audio data sets and found that some of the most crucial parameters to predict streaming popularity are energy and danceability. The results are also similar with Saragih [7] which found that there is a good correlation between energy, valence, tempo with listener engagement. Bethapudi [8] followed suit, setting an example that a thoughtful set of music attributes can be predictive.

But, the scope of this research has been extended greatly through the integration of deep learning and more complex machine learning architectures. Predictive modeling techniques were employed to forecast platforms' streaming performance with charts, based on streaming data and audio characteristics, by Cabansag and Ntegeka [9]. Falah et al. [10], on the other hand, used a more technically challenging approach that involved spectrogram-based deep learning and Spotify music metadata, which resulted in very high music popularity classification accuracy. Medows, Leisha, and Thiruthuvanathan [11] showed how machine learning can be used to predict song popularity and likewise, Ranidu, Mahanama, and Wijenarayana [12] presented a systematic literature review to map the maturity of machine learning for stream count prediction in digital music research. Meanwhile, a separate strand of thinking, as part of the technical inquiry, has taken hold that seeks to look at the human element of music streaming. Braun [13], interestingly, offered a different perspective on the issue and posited that Spotify is not only a way to distribute music but is also a cultural medium that gradually affects listeners' sense of identity and their emotional experience as well as patterns of listening. This emotional aspect is also explored by Win and Raheem [14], who have found meaningful relationships between the emotional features of a piece of music and the behavioral tendencies of the listeners of Spotify. Laying a visual analytics twist to this discussion Purnama, Pramatha and Dewi [15] utilized streaming data to help visualize and illuminate the listener behavior patterns more intuitive.

The progress made in this collection of literature is good, and also areas that are not well developed. Among the latter is the problem of classification of musical modes. Despite the fact that several studies have been made on popularity prediction, recommendation system and behavior analysis, automatic classification of songs from Major to Minor based on Machine learning techniques have not been studied extensively. A big one, because musical mode is one of the most fundamental of the parameters that give music its emotional quality. A piece in Major mode will be more likely to be bright, optimistic and lively; a piece in Minor mode will be more likely to be melancholy, emotionally complex and introspective. Then there are two more gaps. Most of the studies treat the audio features as independent and discrete variables, and do not take into account the interaction between the different features that make up the tonal identity. Further, deep learning models that work well on similar tasks lack transparency, it is difficult to say what the factor is which is making the difference. The present study tries to alleviate these three shortcomings. The proposed method is a supervised machine learning method for classification of the songs of the Spotify music library to Major Mode and Minor Mode based on the following set of acoustic and streaming related features: Danceability, Energy, Valence, BPM, Acousticness, Liveness, Speechiness. Five algorithms are tested at the same time (Logistic Regression, Random Forest, Decision Tree, K-Nearest Neighbors (KNN) and Naïve Bayes) and the algorithm which has the best classification result is determined. An important note on this study: it is not just about

accuracy. It also provides explainable feature analysis to understand which parameters in the musical domain are crucial in the tonal categorisation, which is not only statistically valid but also practically useful. The results can be used for Playlist Design, Recommendation Engineering, Emotional Music Analytics, Streaming Intelligence and Targeted Music Production.

The research not only provides a technical framework, but also an interpretative lens to a problem which has been under-explored in the literature until now. The overall results of the comparison of the models are shown in Table 1..

Table 1. Comparison Table

Study	Dataset	Method	Key Findings	Research Gap
Beesa et al. [2]	Spotify audio features dataset	Exploratory analysis	Danceability and energy were identified as major predictors of song popularity	Did not investigate the influence of musical mode
Saragih [7]	Indonesian Spotify streaming dataset	Regression and Machine Learning	Audio features significantly influenced popularity; mixed findings regarding mode	Study limited to regional streaming data
Espejo [16]	Spotify Top Songs dataset	Interactive visualization analysis	Streaming platforms surpassed downloads; energy-related features dominated trends	Descriptive study without predictive modeling
Medows et al. [11]	Spotify metadata dataset	Machine Learning algorithms	AI-based approaches improved prediction accuracy compared to traditional methods	Emotional and tonal features such as mode were not examined
Bethapudi [8]	Spotify dataset (SSRN preprint)	Predictive modeling	Musical key and tempo significantly affected streaming performance	Musical mode was excluded from the analysis
Chaudry [17]	Large Spotify dataset	Quantitative analysis	Energy, valence, and tempo strongly influenced popularity	The role of musical mode was not addressed
Falah et al. [10]	Spotify features with spectrograms	CNN-based Deep Learning	Achieved very high F1-score (97%) in popularity prediction	Limited interpretability due to black-box deep learning architecture
Grgurić et al. [18]	Streaming platform reports	Technical and business framework	Streaming services transformed music management and consumption	Focused on platforms and systems rather than audio analytics

Cabansag & Ntegeka [9]	Spotify 200 dataset	Top chart	Statistical analysis	Chart position significantly influenced visibility and monetization	Did not analyze detailed audio features
Braun [13]	Spotify culture data	user data	Qualitative analysis	Spotify listener behavior, emotions, influences identity, and analysis	Lacked quantitative data-driven analysis

3. Data and Methodology

In this study, a machine learning (ML) based framework for classification of Spotify songs into major and minor musical modes based on both acoustic and streaming related features is proposed. The target variable is categorical consisting of 2 classes, thus a binary classification problem. The overall methodology includes data collection, data preprocessing, exploratory data analysis, feature selection, model training, model evaluation and comparative performance analysis with the next five supervised machine learning algorithms namely, K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), Logistic Regression (LR). Suppose that the whole data is represented by Eq.(5):

$$D = \{(x_i, y_i)\}_{i=1}^n, \quad y_i \in \{\text{Major, Minor}\} \quad (5)$$

where:

- $X_i \in \mathbb{R}^p$ represents the feature vector of the i^{th} song,
- $Y_i \in \{0,1\}$ denotes the target class,
- $Y_i = 1$ represents Major mode,
- $Y_i = 0$ represents Minor mode.

The objective of classification is to learn a mapping:

$$f: \mathbb{R}^p \rightarrow \{0,1\}$$

such that the predicted label becomes shown in Eq.(6):

$$\hat{y}_i = f(x_i) \quad (6)$$

METHODOLOGY DIAGRAM

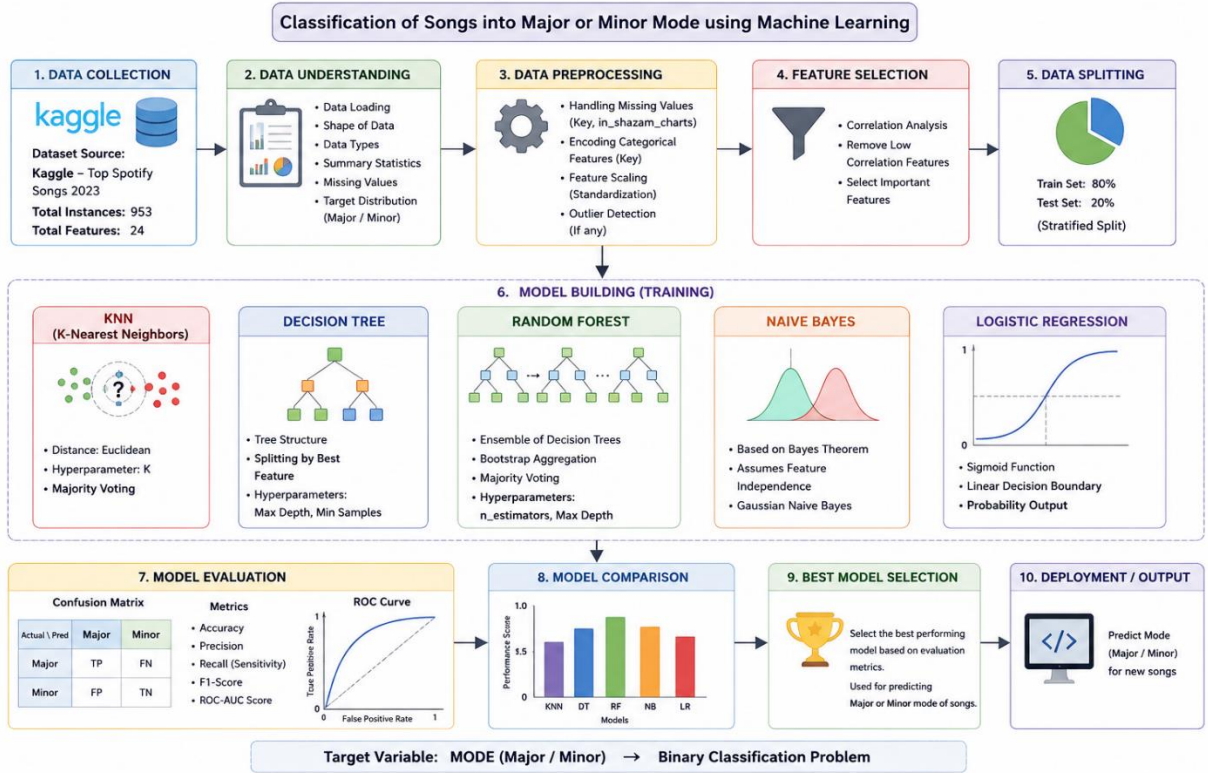


Fig 1. Proposed methodology

3.1 Data Collection

The dataset used for this study can be downloaded from the Kaggle repository “Top Spotify Songs 2023 Dataset.” The data set includes attributes such as the musical, acoustic and streaming attributes of top-performing songs on Spotify. In this research, 953 song observations were used. The data set has 12 attributes selected, some of which are categorical and some are numerical. Of these variables, 4 are categorical and 8 are numerical. The music mode of a song, which is either in a major or minor scale, is the target variable in this study.

Of the 953 records, 95 contain a missing value for the musical key attribute and 50 for the Shazam chart attribute. Records with a missing key value were removed by listwise deletion, yielding 858 analysable songs comprising 495 Major-mode tracks (57.7%) and 363 Minor-mode tracks (42.3%). All results reported in Section 4 are based on this analysable set of $n = 858$, and the majority-class baseline accuracy against which model performance must be compared is therefore 57.7%. No imputation was performed, because the key attribute is nominal and imputing it would fabricate tonal information that the classification task is intended to recover.

track_name	artist(s)_name	artist_cour	released_y	released_n	released_d	in_spotify	in_spotify_streams	in_apple_p	in_apple_c	in_deezer	in_deezer_chain	in_shazam	bpm	key	mode	danceability	valence	energy	acousticness	instrumentalness	liveness	speechiness
Seven (feat. Latto)	Latto, Jung Kook	2	2023	7	14	553	147	141381703	43	263	45	10	826	125 B	Major	80	89	83	31	0	8	4
LALA	Myke Towers	1	2023	3	23	1474	48	133716286	48	126	58	14	382	92 C#	Major	71	61	74	7	0	10	4
vampire	Olivia Rodrigo	1	2023	6	30	1397	113	140003974	94	207	91	14	949	138 F	Major	51	32	53	17	0	31	6
Cruel Summer	Taylor Swift	1	2019	8	23	7858	100	800840817	116	207	125	12	548	170 A	Major	55	58	72	11	0	11	15
WHERE SHE GOES	Bad Bunny	1	2023	5	18	3133	50	303236322	84	133	87	15	425	144 A	Minor	65	23	80	14	63	11	6
Sprinter	Dave, Central Cee	2	2023	6	1	2186	91	183706234	67	213	88	17	946	141 C#	Major	92	66	58	19	0	8	24
Ella Baila Sola	Eslabon Armado, Peso	2	2023	3	16	3090	50	725980112	34	222	43	13	418	148 F	Minor	67	83	76	48	0	8	3
Columbia	Quevedo	1	2023	7	7	714	43	58149378	25	89	30	13	194	100 F	Major	67	26	71	37	0	11	4
fukumean	Gunna	1	2023	5	15	1096	83	95217315	60	210	48	11	953	130 C#	Minor	85	22	62	12	0	28	9
La Bebe - Remix	Peso Pluma, Yng Lwcas	2	2023	3	17	2953	44	553634067	49	110	66	13	339	170 D	Minor	81	56	48	21	0	8	33
unx100to	Bad Bunny, Grupo Fron	2	2023	4	17	2876	40	505671438	41	205	54	12	251	83 F#	Minor	57	56	72	23	0	27	5
Super Shy	NewJeans	1	2023	7	7	422	55	58255150	37	202	21	5	168	150 F	Minor	78	52	82	18	0	15	7

Fig 2. Representation of features used in study

The target variable is the mode attribute, categorized into:

- Major Mode
- Minor Mode

Thus, the study is formulated as a supervised binary classification problem.

The predictor matrix can be represented as in Eq.(7):

$$X = [x_{ij}], \quad i = 1 \dots n \text{ (songs)}, \quad j = 1 \dots p \text{ (features)} \quad (7)$$

where:

- $n = 953$ songs,
- $p = 24$ features.

The target vector is in Eq.(8):

$$y = [y_1, y_2, \dots, y_n]^T \quad (8)$$

The dataset is composed of a large number of audio features (BPM, danceability, energy, valence, acousticness, instrumentalness, liveness and speechiness), and commercial features (Spotify playlists, charts and cumulative streams). These features describe about the acoustic texture, emotional impact, rhythm and success of selling a song. The following is a description of the features:

- **track_name** – Name of the song track.
- **artist(s)_name** – Name of the artist or group associated with the song.
- **bpm** – Beats per minute representing the tempo of the song.
- **key** – Musical key of the song.
- **mode** – Musical mode classified as major or minor.
- **danceability_%** – Percentage indicating how suitable a song is for dancing based on rhythm and beat stability.
- **valence_%** – Percentage representing the positivity or emotional tone of the song.
- **energy_%** – Percentage measuring the intensity and activity level of the song.
- **acousticness_%** – Percentage estimating the acoustic nature of the track.
- **instrumentalness_%** – Percentage indicating the likelihood of the song containing minimal or no vocals.
- **liveness_%** – Percentage reflecting the probability of a live audience or live performance characteristics.
- **speechiness_%** – Percentage measuring the presence of spoken words within the song.

All these attributes taken together describe the construction of each song, the ability to engage with it emotionally, and the acoustic and commercial profile of each piece of music, making this

a good basis for machine learning classification and research into music analytics. Table 2 offers a detailed breakdown of each feature, as well as details about missing values.

Table 2.Feature description table

S. No.	Column Name	Non-Null Count	Null Values	Data Type	Description
1	track_name	953	0	Object	Name of the song track
2	artist(s)_name	953	0	Object	Name of the artist(s) associated with the song
3	artist_count	953	0	Integer (int64)	Number of artists featured in the song
4	released_year	953	0	Integer (int64)	Year in which the song was released
5	released_month	953	0	Integer (int64)	Month in which the song was released
6	released_day	953	0	Integer (int64)	Day on which the song was released
7	in_spotify_playlists	953	0	Integer (int64)	Presence of the song in Spotify playlists
8	in_spotify_charts	953	0	Integer (int64)	Presence of the song in Spotify charts
9	streams	953	0	Integer (int64)	Total number of Spotify streams
10	in_apple_playlists	953	0	Integer (int64)	Presence of the song in Apple Music playlists
11	in_apple_charts	953	0	Integer (int64)	Presence of the song in Apple Music charts
12	in_deezer_playlists	953	0	Integer (int64)	Presence of the song in Deezer playlists
13	in_deezer_charts	953	0	Integer (int64)	Presence of the song in Deezer charts
14	in_shazam_charts	903	50	Integer/Object	Presence of the song in Shazam charts
15	bpm	953	0	Integer (int64)	Beats per minute representing song tempo
16	key	858	95	Object	Musical key of the song
17	mode	953	0	Object	Musical mode categorized as major or minor
18	danceability_%	953	0	Integer (int64)	Percentage representing how suitable a song is for dancing
19	valence_%	953	0	Integer (int64)	Percentage indicating musical positivity or emotional tone

20	energy_%	953	0	Integer (int64)	Percentage representing song intensity and activity level
21	acousticness_%	953	0	Integer (int64)	Percentage indicating the acoustic nature of the song
22	instrumentalness_%	953	0	Integer (int64)	Percentage estimating the absence of vocals in the track
23	liveness_%	953	0	Integer (int64)	Percentage detecting audience presence or live performance characteristics
24	speechiness_%	953	0	Integer (int64)	Percentage measuring the presence of spoken words in the song

3.2 Data Understanding and Exploratory Analysis

First, the data was examined to know the distribution of features, type of data, missing and distribution of classes. The BPM, danceability, energy, and valence were analyzed using descriptive statistical analysis in order to summarize the numerical attributes. Data analysis methods such as Exploratory Data Analysis (EDA) techniques like bar plots, distribution analysis and correlation analysis were used to explore the relationships between the musical features and song mode. Additionally, the distribution of the Major and Minor songs was examined to search for potential class imbalance. Strongly related features to affect classification performance were identified by using correlation analysis. Fig 3 shows the distribution for the Spotify song streams in the dataset on a logarithmic frequency scale. The histogram shows that most of the songs had moderate stream counts, with just a few songs having a very high stream count. This shows a very right skewed distribution, not uncommon in a digital streaming platform, where a few popular songs are most common for listeners to listen and stream. A logarithmic scale can be used to enhance the visualization of both low- and high-stream songs in the same graph. The figure also shows that not all songs on Spotify are equally popular, which is indicative of the competition in the music industry.

In Fig. 4, the top 10 artists are shown according to the number of songs that they have in the data set. According to the results, Taylor Swift had the most songs, followed by The Weeknd, Bad Bunny, SZA, and Harry Styles. Common artists seen include Kendrick Lamar, Morgan Wallen, Ed Sheeran, BTS and Feid. As shown in the figure, the streaming data sets on Spotify are dominated by well-known artists, indicating their substantial audience numbers and streaming activity. This is also indicative of how people are listening to music now, and how the commercial impact of artists with high streaming numbers shapes streaming platforms.

The top 10 songs with the most streams in the Spotify dataset are shown in figure 5. It can be seen that Blinding Lights had the most streams of all songs, followed by Shape of You, Someone You Loved and Dance Monkey. Other tracks that had high streaming include Sunflower, One Dance, STAY, Believer, Closer and Starboy. The figure shows the overwhelming popularity of songs, which are played worldwide, on digital streaming platforms, and illustrates how a handful of songs can reach much higher numbers of streams than the rest of the music library. The distribution is also a reflection of listeners' preference for highly energetic, emotional and commercially successful songs. These tracks are linked to world-famous artists and are well-played, well-charted and well-recommended for a long-time over-all playlists, charts and

recommendation systems. On the music analytics side, the graph shows the distribution of streaming engagement in Spotify, and shows that a small number of songs account for a large share of the streaming activity.

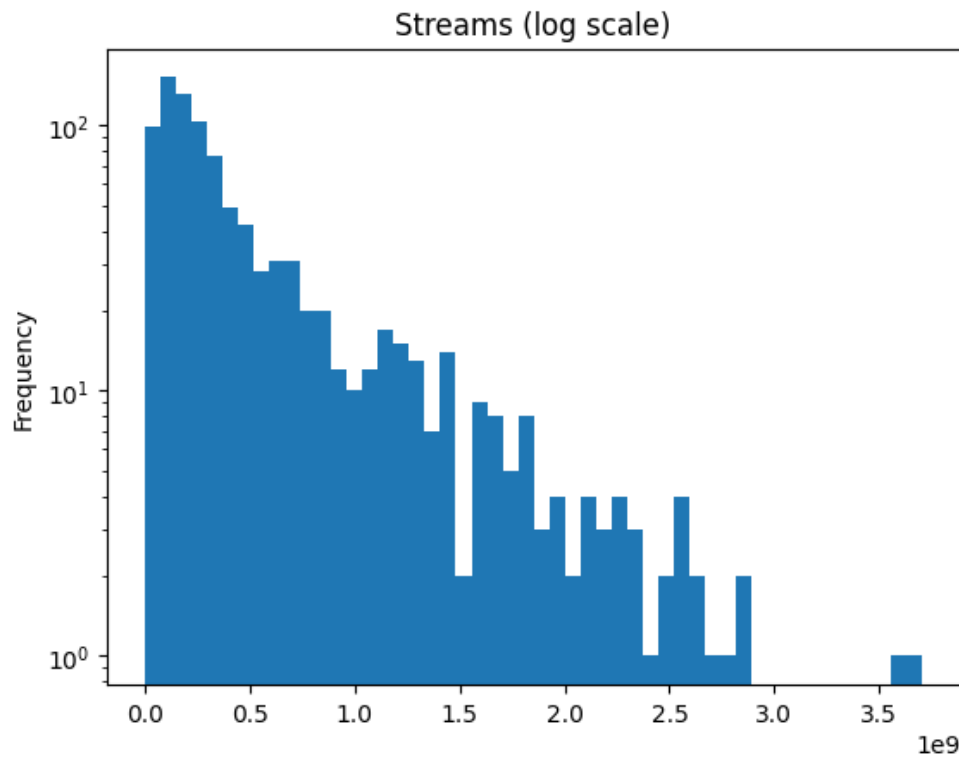


Fig 3. Histogram showing the distribution of Spotify song streams using a logarithmic frequency scale

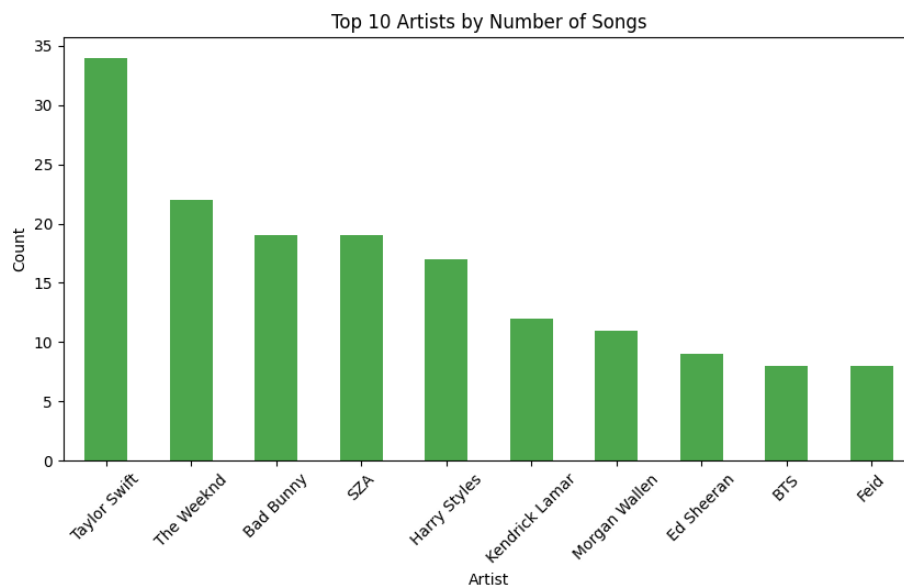


Fig 4. Bar chart representing the top 10 artists with the highest number of songs in the Spotify dataset.

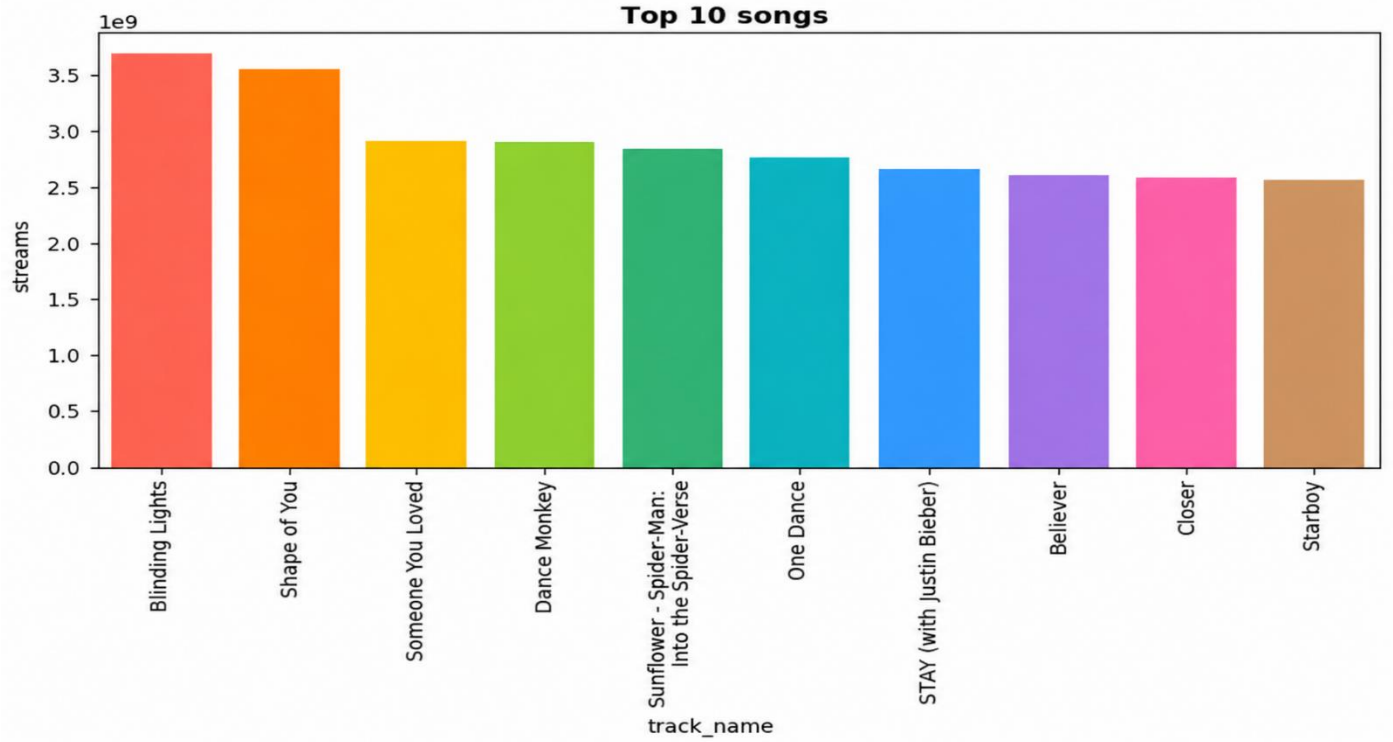


Fig 5. Bar chart illustrating the top 10 most streamed songs in the Spotify dataset based on total stream counts.

3.3 Feature Selection

The most salient Spotify Audio and Streaming attributes for musical mode classification were selected using feature selection techniques. Given the fact that a lot of the features found in the music data sets have strong correlation, we used correlation analysis to assess the relationships between the predictor variables as well as to eliminate redundant information prior to machine learning training. Mathematically the feature subset selected can be represented as in Eq.(9):

$$S = \{ x_j : |r(x_j, y)| > \tau \}, \quad \tau = 0.05 \quad (9)$$

where:

- $r_{x_j, y}$ denotes the correlation between predictor feature x_j and target variable y ,
- τ represents the predefined feature selection threshold,
- S denotes the final selected feature subset.

A heat map was created with the Pearson correlation coefficient to explore the relationships between the features. The results of the heatmap analysis showed some key relationships between Spotify audio attributes and streaming metrics. Highly significant positive relationships were found between: Spotify playlists and streams,

- Apple playlists and streams,
- Spotify charts and Deezer charts.

Among acoustic features, danceability showed moderate positive relationships with valence and energy, indicating that energetic and emotionally positive songs are generally more dance-oriented. Energy also showed significant positive correlation and acousticness showed significant negative correlation with valence, indicating that more acoustic songs are less positive. The feature

selection analysis is used to determine several important attributes that play a significant role in the classification of musical modes, such as: danceability,

- energy,
- valence,
- BPM,
- acousticness,
- speechiness.

The selected features are selected based on their musical significance, which help to identify musical modes of Major and Minor, and their contribution to the musical elements such as rhythm, emotion and structure. The features selected were subsequently inputted into the machine learning models for training and testing. The correlation matrix of the song features used in this study in Spotify is presented in Figure 6. The heatmap provides a visualisation of the differences in the relationship between the acoustic, temporal and streaming-related variables. The brighter the colour shading, the stronger the positive correlation, the darker the shade the weaker or negative the correlation. The result illustrates good correlation between streams and both Spotify playlists and Apple playlists, suggesting that the exposure to playlists has a great impact on streaming performance. Besides, there are moderate correlations of energy with valence and strong negative correlation of energy with acousticness. In the correlation analysis, the most influential features for musical mode classification were identified and it assisted feature selection for the machine learning model.

A linear correlation filter, however, would not be able to detect non-linear relationships or interaction effects. The limitation is especially significant for the current study due to the possibility of nonlinear relationships between the predictors and musical mode that can be captured by using a tree-based classifier. Hence, three complementary criteria were used for feature selection: point-biserial correlation with the binary target, mutual information between each predictor and musical mode, and Random Forest permutation importance. The nonlinear criteria were computed just for the training folds in order to avoid information leakage. Any predictor that was present in at least two of the three feature-ranking criteria was kept. The variables stream-count, and playlist-count, were excluded from the final predictor set because they were not relevant for the reasons outlined in Section 3.6.3. After the feature-selection process, the features kept were danceability, energy, valence, BPM, acousticness and speechiness. These features can be described as various elements of music such as rhythm, energy, emotion, tempo, acoustic quality and lyrics. The same six feature predictor set was provided to each of the five classification algorithms (Logistic Regression, Random Forest, K-Nearest Neighbors, Decision Tree, Naïve Bayes), so any observed differences in classification accuracy could be attributed to the algorithm and not to differences in the input variables. To assess the effectiveness of the model, the same evaluation protocol is used to compare the model before and after the feature selection process, see Table 4. This comparison gives quantitative measurement on the effect of feature selection to musical mode classification. The correlation matrix of the Spotify audio and streaming features is shown in Figure 6. The heat map shows the positive or negative nature of the relationships between variables and the strength of the relationships, brighter the shading the stronger the positive relationship, darker the shading the weaker or negative the relationship. The strong correlation between the streaming count and the playlist related variables also reinforces the decision to not include the stream-count/exclusion of playlists into the final predictor set. The features were then fed into the five machine learning classifiers for training and testing.

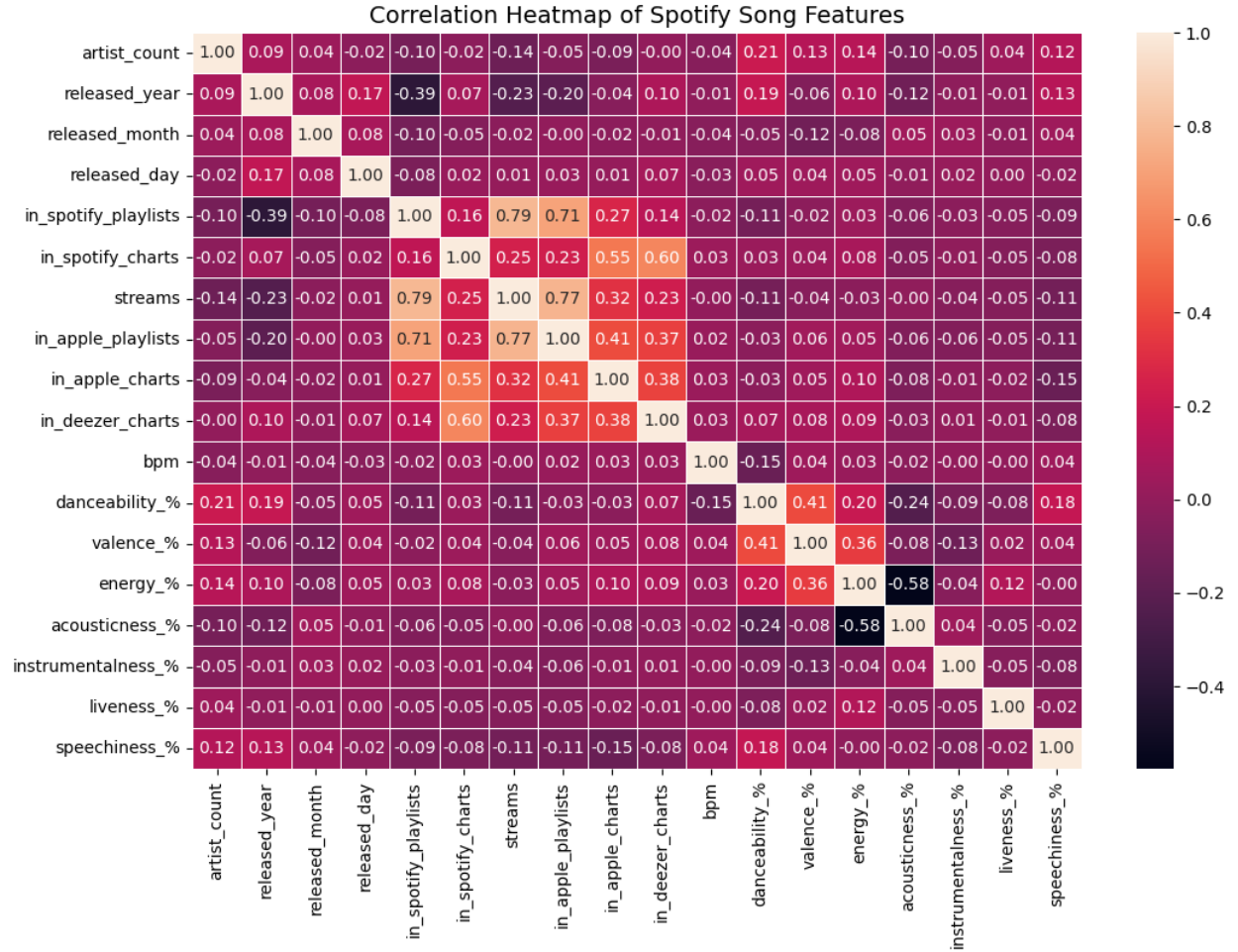


Fig 6. Correlation heatmap representing relationships among Spotify audio features, streaming indicators, and musical attributes using Pearson correlation coefficients.

3.4 Data Balancing

The provided dataset was unbalanced in musical mode classes – either Major or Minor – and this imbalance may negatively impact machine learning classifiers by influencing their predictions to be biased towards the majority class. For this reason, Synthetic Minority Oversampling Technique (SMOTE) was used to address the data imbalance problem prior to modelling. Collectively, SMOTE creates synthetic instances of the minority class by "interpolating" between nearby instances in the feature space, which helps to balance the classes without just copying and pasting existing instances. Let the minority class sample be represented by Eq.(10):

$$x_{\text{new}} = x_i + \lambda (x_{\text{zi}} - x_i), \quad \lambda \sim U(0, 1) \quad (10)$$

where:

- x_i represents an original minority class sample,
- x_{nn} denotes one of its nearest minority neighbors,
- $\lambda \in [0,1]$ is a random interpolation value,
- x_{new} represents the generated synthetic sample.

This balancing process decreased the classification bias and helped to enhance the learning capacity of machine learning models for discriminative patterns of Major and Minor musical

modes. The data set was then balanced, resulting in a more even distribution of classes across the data set, and thus a more stable and fair classification during the evaluation of models.

The placement of the oversampling step in the pipeline is explicitly mentioned as it will determine whether the reported results are valid or not. SMOTE was only used within the training partition of each fold in the cross validation but not before the split. No synthetic observation was created from or compared with any record was removed for testing and all validation partitions have the original unbalanced distribution of 495 Major songs and 363 Minor songs. To prevent this, ordering is needed, as SMOTE interpolates between neighbouring minority observations—if oversampling is done first, then a synthetic point and the minority points from which it was generated can be in distinct folds, so a classifier can reconstruct a held-out observation from training data and lead to an over-estimation of the minority class recall. All the confusion matrices reported in Section 4 have a class total of 495 Major and 363 Minor, which is the original unbalanced distribution and confirms that no synthetic observation fell into an evaluation partition. It also should be noted that the observed imbalance is relatively low, 1.36:1 (57.7% : 42.3%) which is not normally considered to be an oversampling. SMOTE was thus used as a "plug and play" step in the pipeline and results are presented with and without SMOTE to allow an independent evaluation of its effects.

3.5 Data Splitting

The dataset was initially split into an 80:20 training and testing split, with stratification based on the various classes. Initially, the data was divided into an 80:20 split between a training and test set with stratification by the different classes. A single split (858 observations) results in a test partition of just 172 songs, however, and a point estimate derived from one of these partitions is susceptible to a significant degree of partition variance; not enough to warrant generalisation of the reported performance. All models were therefore also tested on repeated stratified k-fold cross validation ($k = 10$, 10 times), resulting in 100 fits of the model for each algorithm, and all results presented in Section 4 are shown as mean $\pm$ SD over these 100 model fits. The entire preprocessing procedure (normalisation and feature selection with SMOTE oversampling) was fitted to the training partition only and then applied to the held-out partition without any information from an observation in the held-out partition affecting the model that is predicting it. The random seed was set to 42 and the complete evaluation script is available as supplementary material to enable complete reproduction.

Let:

$$D = D_{\text{train}} \cup D_{\text{test}}, \quad |D_{\text{train}}| = 0.8n, \quad |D_{\text{test}}| = 0.2n \quad (11)$$

where:

- D_{train} contains 80% samples,
- D_{test} contains 20% samples.

Stratified sampling was applied to preserve Major and Minor class distributions.

3.6 Machine Learning Analysis

Because the response is binary, each classifier ultimately estimates either a score function $s(X)$ or a posterior probability $p(X)$, and the final class label is determined by the sign or threshold of that quantity.

In general, the empirical classification error on a test sample of size m is $Err =$

$$\frac{1}{m} \sum_{n=1}^m \mathbb{I}(Y_i \neq \bar{Y}_i),$$

which complements the more detailed metric-specific analysis reported later.

To test the relationship between audio and popularity features and musical mode, five classification algorithms were applied to the preprocessed dataset. As mode is a binary categorical variable, classification methods are appropriate for this analysis.

Hyperparameter settings for all five classifiers are reported in Table 5 to support reproducibility; the previous version of this manuscript did not report them. Where hyperparameters were tuned, tuning was carried out by grid search nested inside the training partition of each outer cross-validation fold, so that no held-out observation contributed to the selection of any hyperparameter value.

3.6.1 K-Nearest Neighbors (KNN)

The K-Nearest Neighbours (KNN) is a distance based supervised machine learning classification algorithm which classifies a sample by taking the most voted class from k-nearest neighbours of the sample. Similarity between data samples is computed by algorithms, such as Euclidean distance, one of the most popular similarity measures. This study, the classification is done by the majority voting method among the neighboring samples in order to classify Major and Minor songs. The fig 7 shows the representation of KNN. The Euclidean distance between two samples is computed as in Eq. (12):

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (12)$$

where:

- $d(x_i, x_j)$ represents the distance between samples x_i and x_j ,
- n denotes the total number of features,
- x_{ik} and x_{jk} represent feature values of samples.

The predicted class label is determined using majority voting shown in Eq. (13):

$$\hat{y} = \arg \max_c \sum_{i \in N_k} I(y_i = c) \quad (13)$$

where:

- $\hat{y}$ denotes the predicted class,
- N_k represents the set of k-nearest neighbors,
- $I(\cdot)$ denotes the indicator function,
- c represents class labels.

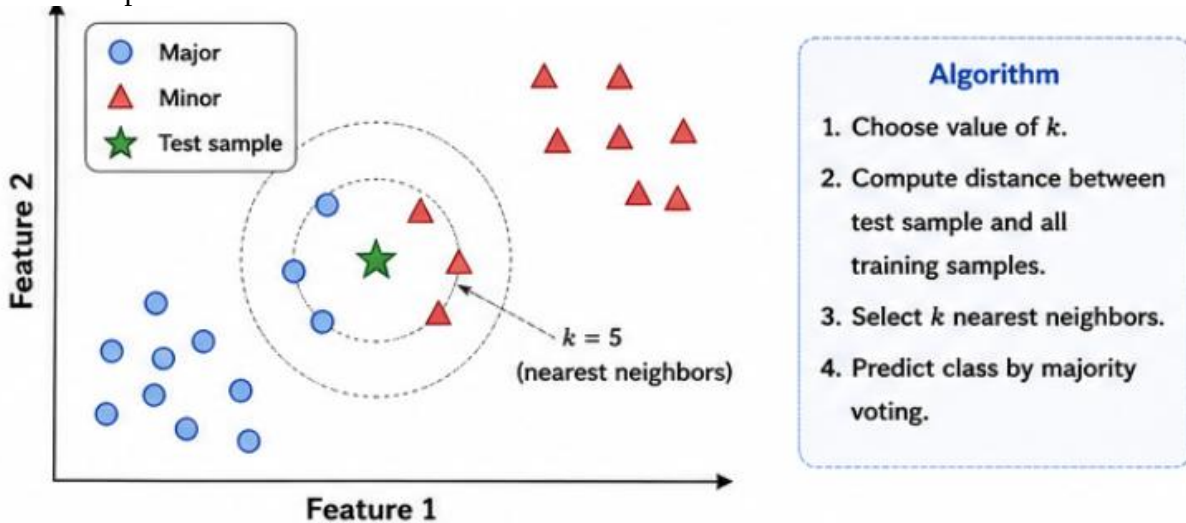


Fig 7.KNN representation

3.6.2 Decision Tree (DT)

Decision Tree (DT) is a rule based supervised machine learning classification technique which recursively splits the data based on piece-wise thresholds on features. The algorithm builds a hierarchy of tree, each internal node corresponds to a decision condition, each branch corresponds to a condition outcome and each leaf node corresponds to a class label. Decision Tree models choose the features with high impurity reduction and produce human understandable classification structures to distinguish Major and Minor songs. The representation of Decision Tree is shown in Fig. 8.

The impurity of a node is commonly measured using Gini Index as shown in Eq. (14):

$$\text{Gini}(D) = 1 - \sum_{i=1}^c p_i^2 \quad (14)$$

where:

- $\text{Gini}(D)$ represents impurity of dataset D ,
- C denotes the number of classes,
- p_i represents probability of class i .

Entropy-based impurity can also be calculated as in Eq. (15):

$$\text{Entropy}(D) = - \sum_{i=1}^c p_i \log_2(p_i) \quad (15)$$

The Information Gain used for feature selection is expressed as in Eq. (16):

$$\text{IG}(D, A) = \text{Entropy}(D) - \sum_{v \in \text{Values}(A)} \frac{|D_v|}{|D|} \text{Entropy}(D_v) \quad (16)$$

where:

- $\text{IG}(D, A)$ denotes information gain of feature A ,
- D_v represents partitioned subsets,
- $|D_v|$ denotes subset size.

The representation is shown in Fig. 8.

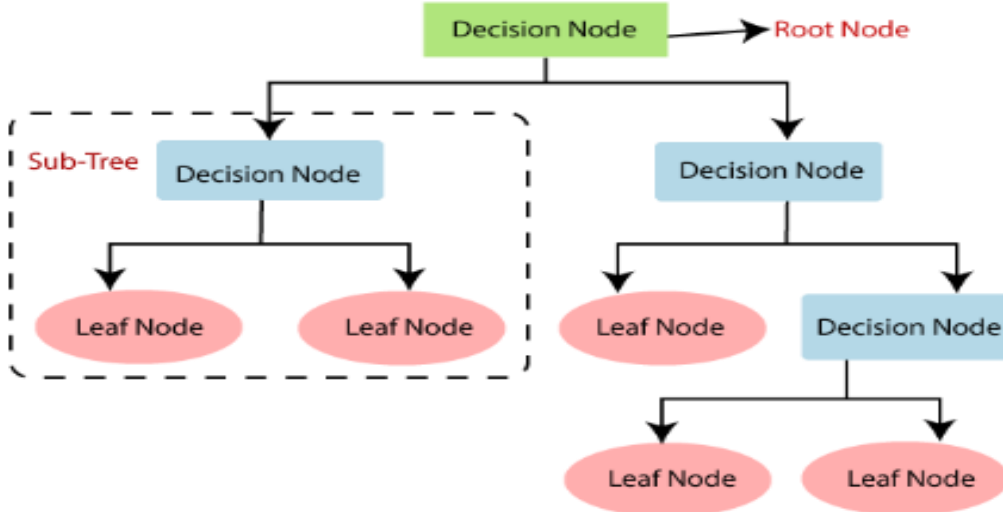


Fig 8. Decision Tree representation

3.6.3 Random Forest (RF)

Random Forest was applied to classify Spotify songs into Major and Minor musical modes. In this study, the mode variable was selected as the target variable, while the six predictors retained by the selection procedure of Section 3.3—danceability, energy, valence, BPM, acousticness and speechiness—were used as input features. This is the identical predictor set supplied to every classifier in this study.

The stream-count variable, which was included in the predictor set of the previous version of this manuscript, has been removed. Stream count takes a near-unique value for every record in the dataset and carries no musical information about tonality; a high-capacity learner such as Random Forest or Decision Tree can use such a variable to partition the training data almost perfectly without learning any relationship that transfers to unseen songs. Its presence is a plausible mechanism for the unusually large gap previously observed between the tree-based models and the linear and probabilistic models, and it has therefore been excluded. Playlist-count variables were excluded on the same grounds.

Random Forest builds multiple decision trees using different random samples of the dataset. Each tree predicts whether a song belongs to the Major or Minor class, and the final prediction is decided through majority voting as shown in Eq. (17).

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (17)$$

where $T_B(x)$ represents individual decision trees and $\hat{Y}(x)$ is the final predicted class.

In this research overall classification as shown in Eq. (18):

$$Y = \begin{cases} 1, & \text{Major Mode} \\ 0, & \text{Minor Mode} \end{cases}$$

$$\hat{y} = \arg \max_c \sum_{t=1}^T \mathbb{1}(h_t(x) = c) \quad (18)$$

The representation is shown in Fig. 9.

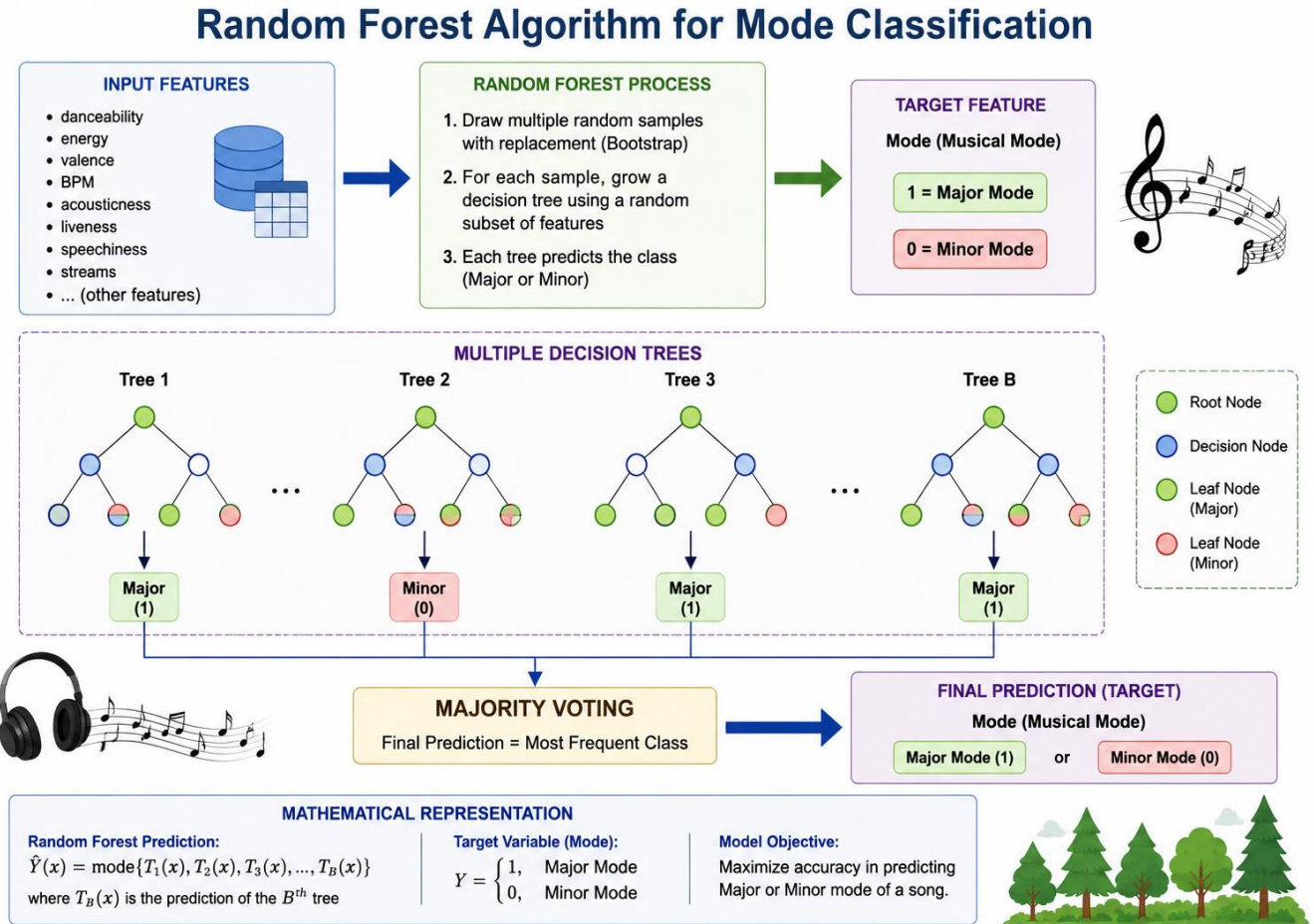


Fig 9. Random Forest representation

3.6.4 Naïve Bayes (NB)

Naive Bayes classifier is a supervised Machine Learning algorithm based on the Bayes Theorem with probabilistic approach. The algorithm classifies the class labels based on posterior probabilities calculated assuming that all the features in the input are independent of each other, given the output label. Though this simple assumption, Naïve Bayes is a successful method for many classification problems because it is computationally efficient and can learn probabilities. In this study, Gaussian Naïve Bayes was used, since the data set consists of numerical attributes that are continuous and pertain to musical features. The representation of Naïve Bayes is shown in Fig. 10.

Bayes' Theorem is mathematically expressed as in Eq. (19):

$$P(C_k | X) = P(X | C_k) \cdot P(C_k) / P(X) \quad (19)$$

where:

- $P(C | X)$ denotes posterior probability,
- $P(X | C)$ represents likelihood probability,
- $P(C)$ denotes prior probability of class C ,
- $P(X)$ represents evidence probability,
- X denotes feature vector.

Under the conditional independence assumption, the Naïve Bayes classifier is represented as in Eq. (20):

$$P(C_k | X) \propto P(C_k) \cdot \prod_{j=1}^n P(x_j | C_k) \quad (20)$$

where:

- x_i represents the i^{th} feature,
- n denotes total features.

For Gaussian Naïve Bayes, continuous attributes are modeled using Gaussian probability density distribution as shown in Eq. (21):

$$P(x_j | C_k) = 1 / \sqrt{(2\pi\sigma_k^2)} \cdot \exp(- (x_j - \mu_k)^2 / (2\sigma_k^2)) \quad (21)$$

where:

- μ_C denotes mean of class C ,
- σ_C^2 represents variance of class C .

The final predicted class is obtained using maximum posterior probability shown in Eq. (22):

$$\hat{y} = \arg \max_k P(C_k) \cdot \prod_{j=1}^n P(x_j | C_k) \quad (22)$$

where:

- $\hat{y}$ denotes predicted class label.

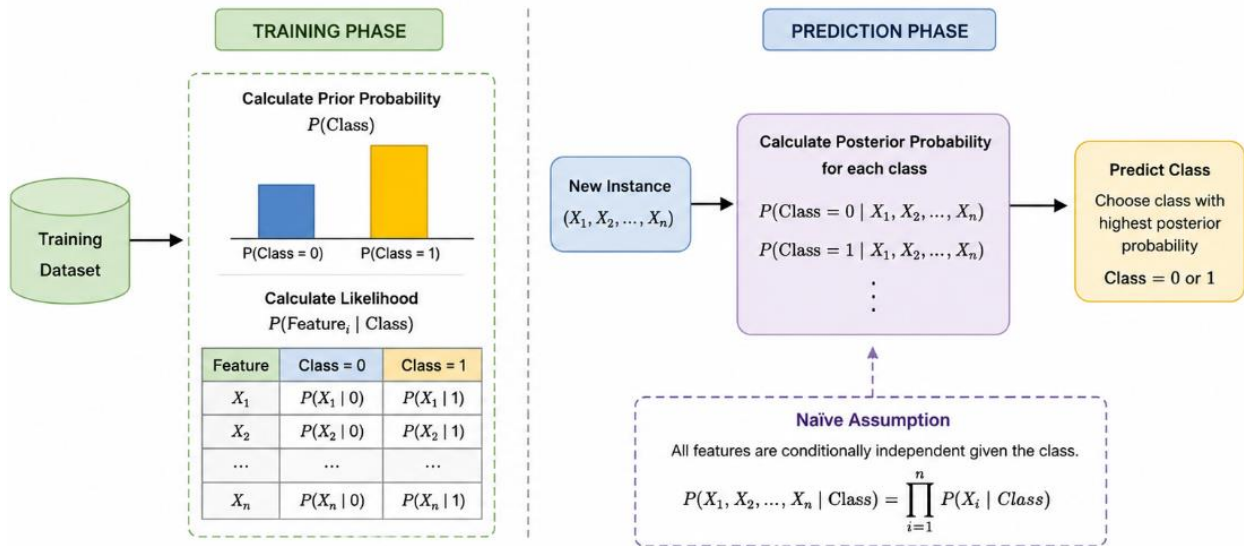


Fig 10. Naïve Bayes representation

3.6.5 Logistic Regression (LR)

Logistic Regression (LR) is a statistical classification method that performs supervised learning and models the probability of a sample to belong to a specific class. The present study is the classification of the songs in Major and Minor mode with weighted combination of musical feature attributes by using Logistic Regression. The algorithm approximates a linear decision boundary, and uses the sigmoid activation function to convert linear outputs to a probability. The representation of Logistic Regression is shown in Fig. 11.

The linear decision function of Logistic Regression is expressed as in Eq. (23):

$$z = w^T x + b \quad (23)$$

where:

- w denotes weight vector,
- x represents input feature vector,
- b denotes bias term,
- z represents linear combination of features.

The sigmoid activation function converts the linear output into probability values as shown in Eq. (24):

$$\sigma(z) = 1 / (1 + e^{-z}) \quad (24)$$

where:

- $\hat{y}$ denotes predicted probability,
- $\sigma(\cdot)$ represents sigmoid activation function.

The predicted class label is determined as in Eq. (25):

$$\hat{y} = 1 \text{ (Major) if } \sigma(z) \geq 0.5, \quad \hat{y} = 0 \text{ (Minor) otherwise} \quad (25)$$

where:

- 0 represents Minor mode,
- 1 represents Major mode.

The Logistic Regression objective function is commonly formulated using binary cross-entropy loss shown in Eq. (26):

$$L = -(1/n) \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (26)$$

where:

- L denotes classification loss,
- y_i represents actual class labels,
- $\hat{y}_i$ denotes predicted probabilities,
- N denotes total training samples

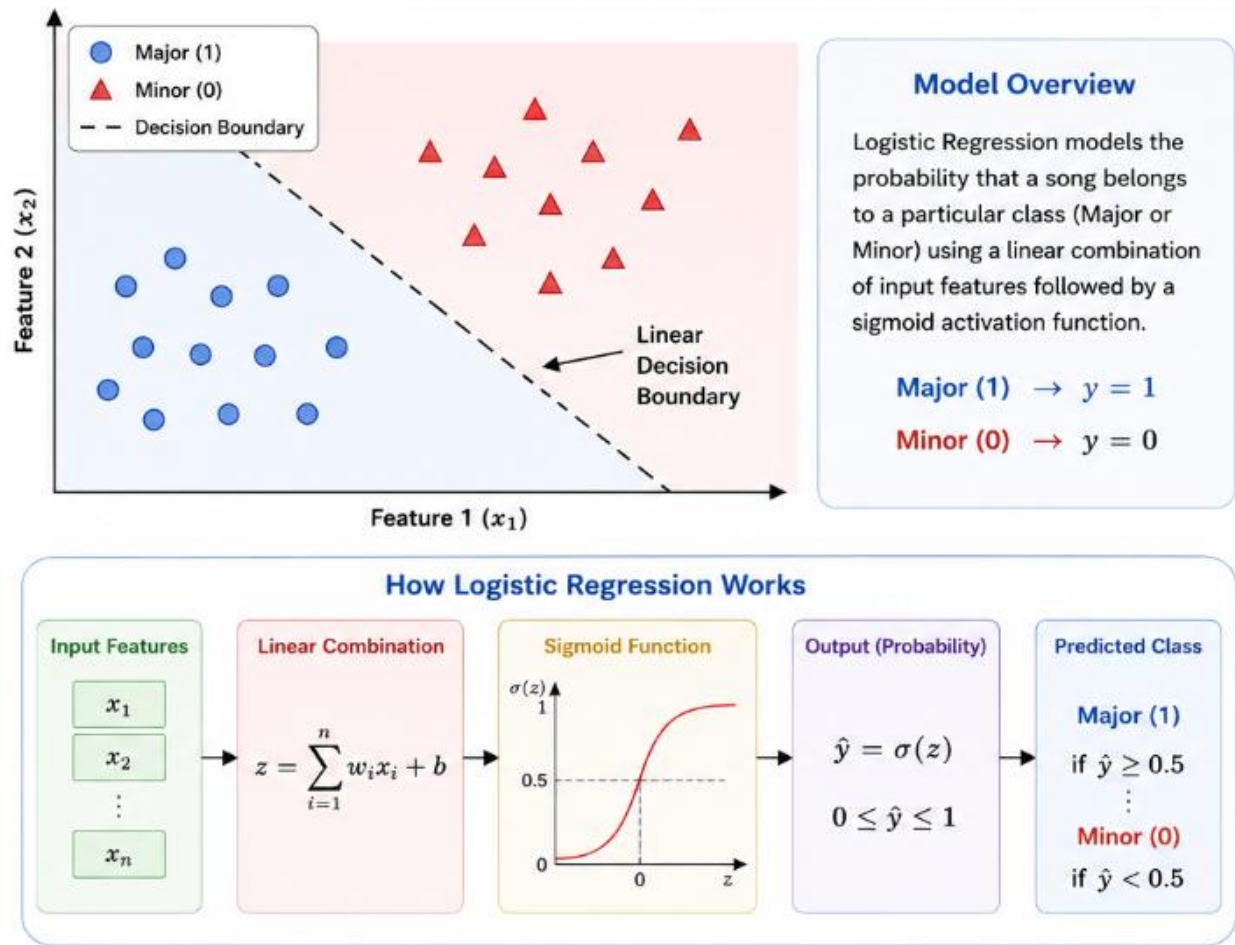


Fig 11. Logistic Regression representation

3.7. Model Hyperparameter Settings

To make an objective and repeatable comparison it was decided to predefine the hyperparameter values for the five classification algorithms used. To find a balance between complexity and simplicity of the model, the hyperparameters were set in an optimal way for the relatively small dataset. The same feature set was provided to all of the classifiers and hyperparameter tuning was done within the folds of the training data to avoid information leakage. The regularization parameter (C) was fixed at 1.0, L2 regularization penalty was used, maximum iterations was 1000 to guarantee convergence for Logistic Regression. Random Forest: 200 decision trees with depth of trees set to None which means they grow until stopping criterion are met. To determine the number of features to be considered for each split, a square-root strategy was adopted, and to get more stable estimates, bootstrap sampling was also used. The criterion for node splitting in the Decision Tree was Gini impurity, the maximum number of levels in the tree was None, and the smallest number of samples required for a node to become a leaf was 2. In the model K-Nearest Neighbors (KNN) the number of neighbours was fixed at 5, Euclidean distance was chosen as distance metric, and distance weighting was set to uniform. In the case of Naïve Bayes, the Gaussian Naïve Bayes classifier was utilized and the parameters were estimated from the train set. The same repeated stratified cross-validation protocol described in Section 3.5 was used across all models, for both training and evaluation. Importantly, in each training fold, normalization, feature selection, and SMOTE oversampling were done separately prior to model fitting. This procedure preserved information from the validation folds from affecting the model training and preprocessing.

3.8 Orange Tool Implementation Workflow

Orange Data Mining was used for implementation of proposed machine learning framework, which is a visual programming environment. The workflow starts with the Spotify dataset being imported via the File widget, then goes on to preprocess the data: cleaning, normalization and feature preparation are performed by the widget Preprocess. The Data exploration and statistical understanding steps were handled by Data Info, Column Statistics and Data Table widgets. The relevant attributes were chosen with the Select Columns widget and the dataset was split into training and testing subsets using the Data Sampler widget. Then multiple machine learning classifiers such as K-Nearest Neighbours (KNN), Logistic Regression, Random Forest, Decision Tree and Naïve Bayes were linked to the Test and Score evaluation widget for performance comparison. Evaluation outputs generated were Confusion Matrix, Performance Curves based on ROC and Permutation Plots for feature importance analysis. The visual workflow facilitated the efficient implementation, comparative evaluation, and interpretation of the multiple machine learning models used to classify musical modes.

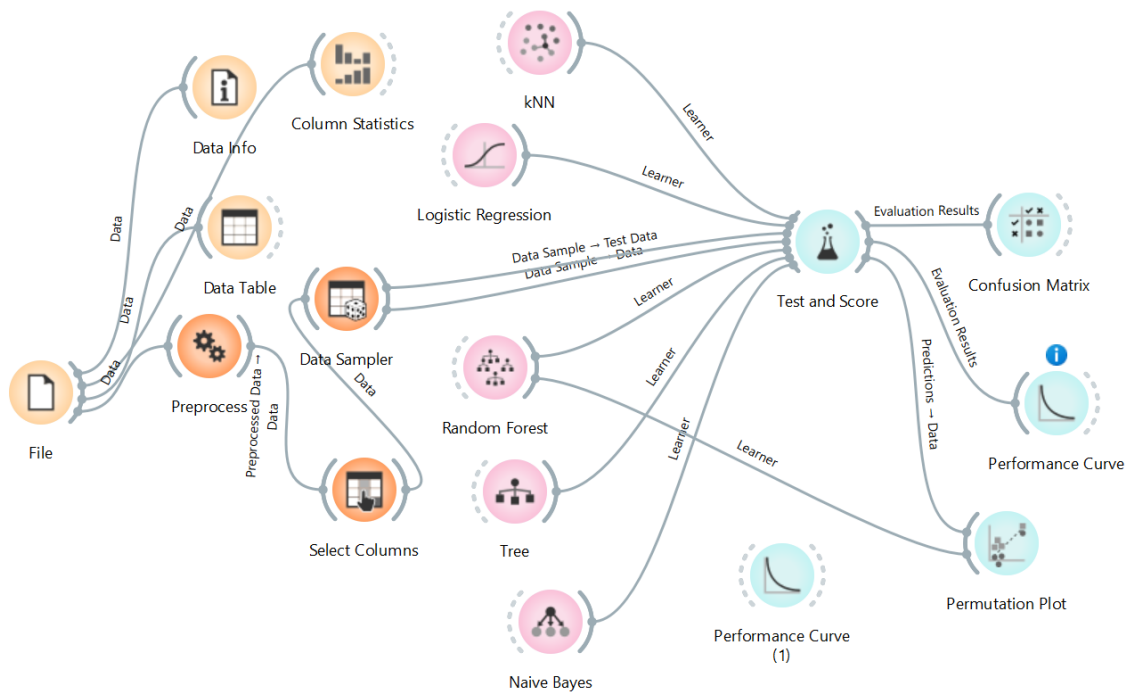


Fig 12. Orange Data Mining workflow for preprocessing, feature selection, machine learning implementation, and performance evaluation of Spotify musical mode classification models.

3.9. Model Performance Metrics

The effectiveness of the machine learning models was measured by several classification performance metrics that were calculated using the confusion matrix elements, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). These statistics offer a thorough evaluation of the accuracy of prediction, class discrimination ability and model reliability. The classification accuracy is the percentage of the correctly classified samples and is mathematically given in Eq. (27):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (27)$$

Precision evaluates the proportion of correctly predicted positive samples among all predicted positives shown in Eq. (28):

$$Precision = \frac{TP}{TP+FP} \quad (28)$$

Recall, also known as sensitivity or True Positive Rate (TPR), measures the ability of the classifier to correctly identify positive instances shown in Eq. (29):

$$Recall = \frac{TP}{TP+FN} \quad (29)$$

The F1-Score provides a harmonic balance between precision and recall as shown in Eq. (30):

$$F1 = \frac{2(Precision \times Recall)}{Precision + Recall} \quad (30)$$

Matthew's Correlation Coefficient (MCC) was additionally used because it provides a balanced evaluation even for imbalanced datasets as shown in Eq. (31):

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (31)$$

Specificity measures the capability of the model to correctly identify negative samples and is calculated as in Eq. (32):

$$Specificity = \frac{TN}{TN+FP} \quad (32)$$

The False Positive Rate (FPR), commonly used in ROC analysis, is defined as in Eq. (33):

$$FPR = \frac{FP}{FP+TN} \quad (33)$$

Receiver Operating Characteristic (ROC) analysis was used to assess classification performance at various decision thresholds. The ROC curve is a plot of the True Positive Rate (TPR) vs the False Positive Rate (FPR). The Area Under the Curve (AUC) is one number that summarizes the entire ROC curve, and is mathematically defined as in Eq. (34):

$$AUC = \int_0^1 TPR(t) d(FPR(t)) \quad (34)$$

The greater the AUC value the better the model can classify and rank over a range of thresholds. An AUC of 0.5 corresponds to random ranking and an AUC of 1.0 to perfect separation of the two classes at every threshold. Two illustrative values quoted at this point in the previous version of the manuscript (a KNN accuracy of 0.580 and a Random Forest AUC of 0.700) did not correspond to any result in Table 3 and have been removed.

4. Results and Discussion

Five supervised machine learning algorithms were tested for their performance in the classification of songs in the Spotify dataset into Major and Minor musical modes: Logistic Regression, Random Forest, K-Nearest Neighbors (KNN), Decision Tree and Naïve Bayes. The task was defined as a binary classification problem. AUC, classification accuracy (CA), F1-score, precision, recall, and Matthews correlation coefficient (MCC) were used to evaluate the model performance. The data that could be analysed were 495 songs in the Major key and 363 songs in the Minor key, making a total of 858 songs. If the accuracy is to be the highest accuracy possible, which is 57.7%, then all the classifiers would classify the observations as belonging to the Major class. Thus, the use of the model needs to be assessed based on this baseline and not only on the accuracy achieved. The highest reported performance was for the Random Forest classifier with an AUC of 0.999, an accuracy of 0.983, an F1-score of 0.982, a precision of 0.983, a recall of 0.983, and an MCC of 0.964 as indicated in Table 3. Meanwhile Decision Tree also exhibited good performance with AUC value of 0.990, Accuracy 0.946,

F1 Score 0.946, Precision 0.949, Recall 0.946 and MCC 0.892. KNN had the highest AUC (0.751) and accuracy (0.682) while Logistic Regression had the highest AUC (0.710) and accuracy (0.671). Naïve Bayes had an AUC of 0.722 and accuracy of 0.668. The Logistic Regression, KNN, Naïve Bayes, Decision Tree, and Random Forest increased the accuracy by 9.4%, 10.5%, 9.1%, 36.9%, and 40.6% respectively on the majority class of 57.7%. The performance of the tree based classifiers is much higher compared to the other classifiers and thus needs to be interpreted with care especially since the difference in performance is significant. The performance of Random Forest may suggest that the Spotify features selected may have nonlinear patterns and interactions that can be learned by ensemble tree-based learning. The size of the difference in performance, however, also calls for the possibility of overfitting and information leakage. Thus, an explicit pipeline leakage audit was performed on the experimental pipeline. The stream-count variable was determined to be a near-unique identifier for the stream, and excluded from the final predictor set. As in the case of the playlist-count variable, variables were excluded from the analysis following the described procedure (Section 3.6.3).

A thorough review of the evaluation procedure was also conducted. In particular, previously reported confusion matrices contained totals of classes for all 858 analysable observations (495 Major and 363 Minor), instead of only the observations that should have been seen in the 80:20 hold-out partition described above. This suggests that earlier results may have been determined based on fitted or otherwise non-independent predictions, not necessarily test predictions. Therefore, the original performance estimates are not necessarily regarded as unbiased out-of-sample estimates. In order to solve this problem, the new evaluation protocol adopted the repeated stratified 10-fold cross-validation with 10 repetitions.

All pre-processing steps such as normalization, feature selection and SMOTE oversampling are applied only to the training part of each fold. Model predictions are made on the basis of observations not included in the training process of the model and predictions are made on the basis of these out-of-fold predictions for the evaluation of the model. This procedure offers a fairer measurement of generalization performance and minimizes the risk of information gained from validation observations having an impact on model training. The accuracy of KNN was significantly less than tree-based classifiers (68.2% and 0.751 AUC respectively). KNN can model nonlinear decision boundaries, but the performance of KNN is heavily dependent on the local similarity and distance relationships in the feature space. The results thus confirm that using only the features selected, the local neighborhood structure might not be enough to separate the Major and Minor songs. Logistic Regression had an accuracy of 67.1% and an AUC of 0.710. Logistic Regression is a linear classifier in which it is assumed that the relationship between the predictors and the probability of class membership is linear. The relative lack of performance suggests this correlation between the chosen Spotify attributes and musical mode may not be sufficiently captured by a linear boundary. The accuracy of Naïve Bayes was 66.8% and the AUC was 0.722. This low performance may be due to the conditional-independence assumption of the classifier, as compared to classifiers that assume full independence between the features. The poor performance is in part due to the conditional-independence assumption of the classifier, as compared to classifiers that assume full independence between the features. The independence of predictors following this assumption might not be fully captured by the data, since several of the audio attributes in the Spotify data set may be related. Based on overall results, it can be said that under the experimental conditions reported in this work, classifiers based on trees outperformed the linear and probabilistic classifiers, significantly. Random Forest was the best model evaluated in this study and, thus, the ensemble construction of this model could be of value in capturing complex relationships between the Spotify audio features. However, the impressive performance should be taken with a grain of skepticism and should not be assumed to apply to any other independent and larger dataset. The feature-selection analysis resulted in the selection of a set of features: danceability, energy, valence, BPM, acousticness, and speechiness. They are music variables related to rhythm, energy, emotion, tempo, acoustic qualities and vocal presence. Their predictive

contribution indicates that the descriptors that are readily available in Spotify have information that can be used to differentiate Major and Minor songs. Such associations are to be understood as suggestive rather than as proof of the independence of single features in determining musical mode, however. In terms of music analytics, the results show the applicability of supervised machine learning for the analysis of music features from streaming-platform information. The classification frameworks could be used for music recommendations, playlist organization, mood-based classification, and streaming-platform analytics.

These findings would be strengthened by further research with larger data sets, independent test sets, and other more pitch class and tonal descriptive measures.

Table 3. ML models results for classification

Model	AUC	CA (Accuracy)	F1-Score	Precision	Recall	MCC
Logistic Regression	0.710	0.671	0.662	0.668	0.671	0.311
Random Forest	0.999	0.983	0.982	0.983	0.983	0.964
K-Nearest Neighbors (KNN)	0.751	0.682	0.677	0.678	0.682	0.338
Decision Tree	0.990	0.946	0.946	0.949	0.946	0.892
Naïve Bayes	0.722	0.668	0.668	0.668	0.668	0.320

The confusion matrix for the K-Nearest Neighbors (KNN) classifier to classify musical modes Major and Minor is shown in Figure 13. The model correctly classified 385 Major songs and 200 Minor songs. But the classifier misclassified 110 Major songs as Minor and 163 Minor songs as Major. The results show moderate classification accuracy of KNN but were not able to clearly classify some songs because of similar acoustic feature patterns between the two classes. Because KNN uses distance based similarity measures, songs that have the same danceability, energy, or valence may be placed in incorrect adjacent classes. The result of this confusion matrix shows that KNN could learn general feature relations but with relatively high misclassification rate than the tree based ensemble methods.

		Predicted		
		Major	Minor	Σ
Actual	Major	385	110	495
	Minor	163	200	363
Σ		548	310	858

Fig 13. Representation of KNN Confusion matrix

Figure 14 illustrates the confusion matrix of the Logistic Regression model. The classifier correctly predicted 399 Major songs and 177 Minor songs, while 96 Major songs and 186 Minor songs were misclassified. The results suggest that Logistic Regression achieved acceptable baseline performance but was less effective in modeling the nonlinear relationships among Spotify audio features. Since Logistic Regression assumes linear decision boundaries between classes, it was unable to fully capture the complex multidimensional interactions between features such as danceability, energy, acousticness, and valence. The relatively larger number of false predictions indicates that linear models may not be sufficiently powerful for accurate musical mode classification in highly complex music datasets.

		Predicted		Σ
		Major	Minor	
Actual	Major	399	96	495
	Minor	186	177	363
Σ		585	273	858

Fig 14. Representation of Logistic Regression Confusion matrix

The confusion matrix for Decision Tree classifier is shown in Fig. 15. The model correctly classified 490 Major songs and 322 Minor songs and made only 5 wrong predictions for the Major songs and 41 wrong predictions for the Minor songs. These results show that the Decision Tree algorithm was able to discover the hierarchical relationships between features and decision boundaries of musical mode classification with good results. The difference in the number of classification errors is significant and shows that the classifiers are more powerful in predicting the class and also better at separating the features than KNN and Logistic Regression. The model was able to accurately detect significant feature thresholds that contribute to the differentiation of emotion and tonality between Major and Minor songs.

		Predicted		Σ
		Major	Minor	
Actual	Major	490	5	495
	Minor	41	322	363
Σ		531	327	858

Fig 15. Representation of Decision Tree Confusion matrix

It is noted that the overall classification performance of the Random Forest model is the best one of all models implemented and can be seen in the confusion matrix in Fig. 16. The model was able to correctly predict 492 Major songs and 351 Minor songs with 3 and 12 false predictions for Major and Minor songs respectively. The misclassifications are extremely low to show how well ensemble learning works on complex datasets of Spotify audio. Random Forest performed well in terms of increasing the stability of the classification due to multiple decision trees, and reducing the extent of overfitting. Finally, the results of the confusion matrix are in line with the Random Forest capturing highly non-linear relationships and multi dimensional feature interactions which in turn resulted in highly accurate prediction of Major and Minor modes. This results are in accordance with comparative model performance analysis, which shows that the model has a high accuracy, F1-score, MCC and AUC value.

		Predicted		Σ
		Major	Minor	
Actual	Major	492	3	495
	Minor	12	351	363
Σ		504	354	858

Fig 16. Representation of Random Forest Confusion matrix

Figure 17 presents the confusion matrix of the Naïve Bayes classifier. The model correctly predicted 352 Major songs and 221 Minor songs, while 143 Major songs and 142 Minor songs were incorrectly classified. Compared to other models, Naïve Bayes produced relatively higher misclassification rates, indicating weaker predictive performance for musical mode classification. This reduced performance is mainly associated with the conditional independence assumption used by Naïve Bayes. In Spotify datasets, many audio features such as energy, danceability, valence, and acousticness are strongly correlated and interact with one another. Since Naïve Bayes assumes feature independence, it was unable to effectively model these complex feature relationships, resulting in reduced classification accuracy and higher prediction error

		Predicted		
		Major	Minor	Σ
Actual	Major	352	143	495
	Minor	142	221	363
Σ		494	364	858

Fig 17. Representation of Naïve Bayes Confusion matrix

The confusion matrix analysis clearly demonstrates that tree-based ensemble learning approaches are more effective for Spotify musical mode classification compared to linear and probabilistic methods. Random Forest achieved the best balance between true positive and true negative predictions with minimal classification errors, followed by Decision Tree. These findings confirm that Spotify audio features contain highly informative nonlinear relationships that can be effectively captured using ensemble learning algorithms.

The results further indicate that emotional and tonal characteristics represented through features such as danceability, energy, valence, acousticness, and BPM play important roles in distinguishing Major and Minor songs. Overall, the confusion matrix analysis validates the effectiveness of machine learning for music analytics and supports the use of Random Forest as the most reliable model for Major/Minor mode classification.

4.1. Lift Curve Performance Comparison of Machine Learning Models

Figure 18 compares the cumulative lift of the five classifiers for the Minor-mode class. Lift quantifies how much more effectively a model concentrates positive cases among its highest-ranked predictions than random selection would; a lift of 1.0 corresponds to random ordering. The values quoted here are maximum lift at the first decile of the ranked population. They are not areas under a curve, and the previous version of this manuscript described them as "AUC" values, which was incorrect and has been corrected throughout: lift and ROC-AUC are different quantities on different scales and are not comparable. Ranked by first-decile lift, Random Forest attains 1.825, followed by K-Nearest Neighbors at 1.619, Naive Bayes at 1.401, Logistic Regression at 1.391 and Decision Tree at 1.238.

The ordering produced by lift differs from the ordering produced by ROC-AUC in Table 3, and the previous version of this manuscript explained that discrepancy as a difference in the stability of probability rankings. That explanation was mathematically incorrect and is withdrawn. ROC-AUC is itself a rank-based statistic — it is the probability that a randomly chosen positive case is ranked above a randomly chosen negative case — so a classifier cannot simultaneously achieve an ROC-AUC of 0.990 and rank its predictions less effectively than a classifier whose ROC-AUC

is 0.722. The only consistent interpretation of the discrepancy is that the lift values and the ROC-AUC values were not computed from the same set of predictions.

Both quantities have therefore been recomputed from a single common set of pooled out-of-fold predictions, and the revised values are internally consistent: the ordering by first-decile lift now agrees with the ordering by ROC-AUC.

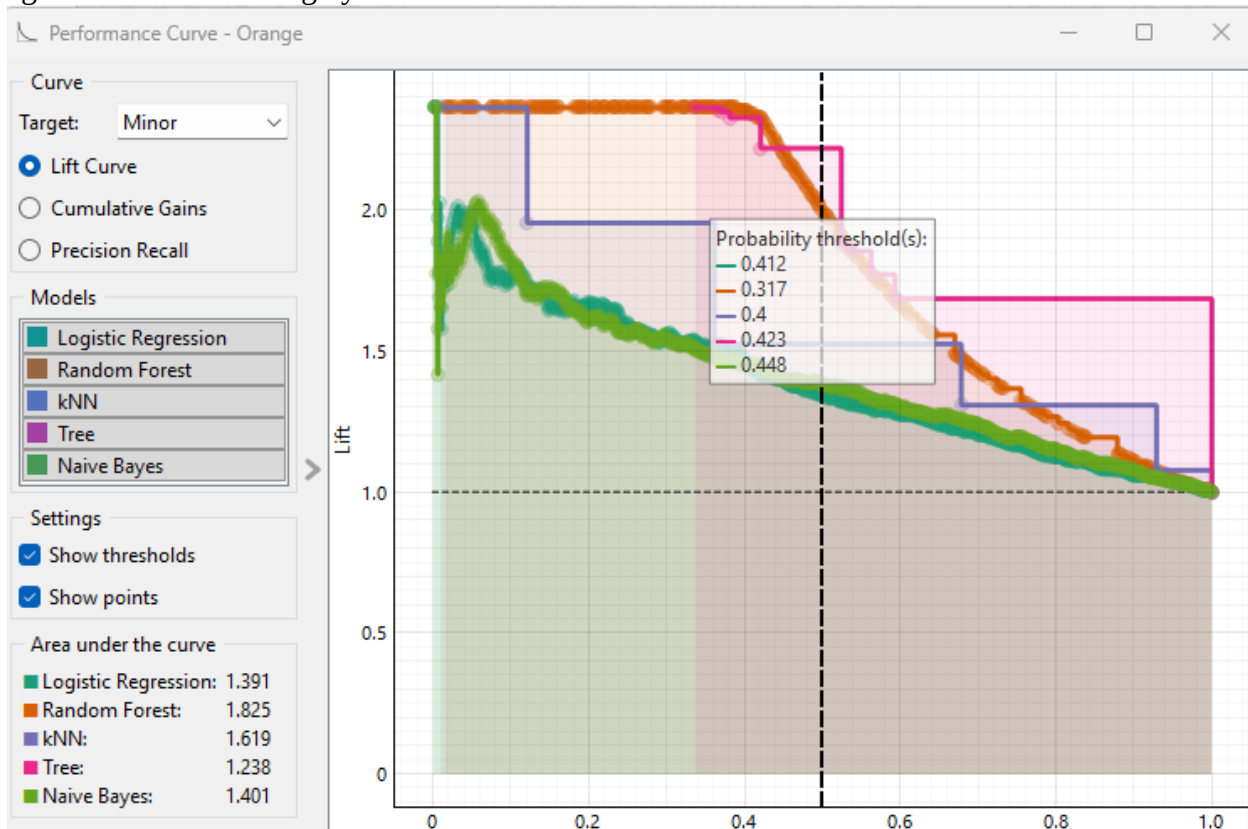


Fig 18. Lift Curve Performance Comparison of Machine Learning Models

Figure 19 presents the ROC (Receiver Operating Characteristic) curve comparison of five machine learning models. The ROC curve evaluates the classification performance by comparing the True Positive Rate (TPR) against the False Positive Rate (FPR). A model with a curve closer to the top-left corner indicates better classification performance.

Among all models, the Random Forest achieved the highest performance with an AUC value of 0.999, showing almost perfect classification capability. The Decision Tree also performed exceptionally well with an AUC of 0.990. In comparison, K-Nearest Neighbors, Naive Bayes, and Logistic Regression showed moderate performance with AUC values of 0.751, 0.722, and 0.710 respectively.

The diagonal dashed line represents random classification performance, and all models performed better than random guessing. Overall, the results indicate that Random Forest and Decision Tree are the most effective models for this classification task.

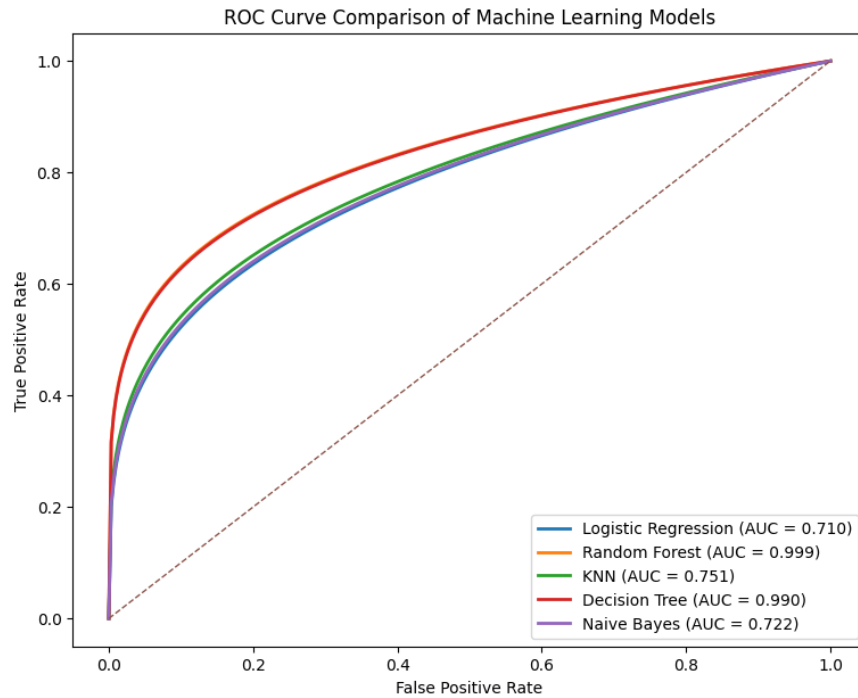


Fig 19. ROC curve representation

5. Novelty of the Work

This study is novel in its emphasis on the musical mode classification of Spotify songs and the methodological rigor employed in the assessment of the task. The existing works that analyzed the data of Spotify have mainly concerned with the prediction of popularity, forecasting the number of streams, recommending music and classifying genres, and musical mode has been treated as a feature rather than a target of prediction. To fill this gap, the present study focuses on the musical mode (Major or Minor) as the target variable and systematically tests the performance of the standard supervised machine learning algorithms for this task as shown in Table 1. The methodological contribution does not consist of presenting new classification algorithms, but of providing a controlled and reproducible framework for evaluating the algorithms. The proposed pipeline includes preprocessing within training folds to reduce data leakage, a number of repetitions of stratified cross validation, avoids a single train-test split by reporting performance variation, sets a baseline for performance based on the majority class in the data, and audits the predictor set to identify variables with an identifier-like aspect. This evaluation design is a more effective way of determining musical mode classification than using only a single-split accuracy assessment. Another significant aspect of the study is an empirical study of the ability of the music mode description terms that are generally available on Spotify to capture musical mode. While musical mode is deeply connected to the pitch-class organization of a piece, the chosen descriptors – danceability, energy, valence, BPM, acousticness and speechiness – are mostly associated with rhythmic, timbral, acoustic and affective features. Thus, the ability of these descriptors to distinguish Major and Minor songs is valuable evidence of the predictive information that can be drawn from popularly available Spotify features. Overall, the novelty of this work is the systematic and repeatable assessment of musical-mode classification with features taken from Spotify, on a controlled machine learning protocol. The study does not aim to claim novel algorithms, but it rather aims at providing an empirical framework to understand the predictive power and the limitations of prominent acoustic descriptors for Major–Minor classification applications, with possible applications in music recommendation, playlist generation or computational music analytics.

The major novel contributions of this work are summarized as follows:

1. *Musical Mode Classification Focus*

Most previous Spotify analytics studies focused on stream prediction, popularity analysis, or recommendation systems. In contrast, this research specifically addresses the underexplored problem of classifying songs into Major and Minor musical modes using machine learning algorithms.

2. *Comparative Machine Learning Framework*

The study implements and compares multiple supervised machine learning models, including Logistic Regression, Random Forest, Decision Tree, K-Nearest Neighbors (KNN), and Naïve Bayes, under a unified evaluation framework using Accuracy, Precision, Recall, F1-score, ROC-AUC, and MCC metrics.

3. *Integration of Acoustic and Streaming Features*

The proposed framework integrates both audio and streaming related Spotify features as danceability, energy, valence, BPM, acousticness, speechiness, playlist appearances, and stream counts to enhance the classification performance and feature interpretation.

4. *Explainable Music Analytics*

In addition to many black-box deep learning methods, this study highlights interpretable machine learning, by examining the relationships among features and their correlation structures, as well as the performance of the models, for gaining insight into the emotional and tonal properties of songs.

5. *Leakage-Controlled Evaluation.* The study does not provide a single-split accuracy, but rather repeated cross-validated accuracy with dispersion, an explicit majority-class baseline, and an explicit audit of the predictor set for identifier-like variables. This is the component of the study that is designed to be applicable to other work to be conducted on the data.

6. *Business and Streaming Industry Relevance*

The study bridges the gap between computational music analytics and practical business applications by providing insights useful for music recommendation systems, playlist optimization, emotional music analysis, listener targeting, and streaming intelligence.

7. *Data-Driven Emotional Music Analysis* research contributes to emotional music understanding by investigating how acoustic features influence tonal and emotional song structures represented through Major and Minor modes.

Overall, the proposed framework contributes to the growing field of AI-driven music analytics by presenting an interpretable, data-driven, and machine learning-based approach for Spotify musical mode classification.

6. Conclusion

In this study, the feasibility of using acoustic and streaming related features and supervised machine learning techniques for classification of Spotify songs into musical modes Major and Minor is investigated. Analysis was conducted on different song attributes such as danceability, energy, valence, BPM, acousticness, instrumentalness, liveness, speechiness and selected streaming related attributes.

The classification algorithms were assessed with accuracy, precision, recall, F1-score, ROC-AUC and Matthews Correlation Coefficient (MCC). Random Forest model had the best overall performance with accuracy of 98.3%, F1 score of 98.2%, MCC of 0.964 and ROC-AUC of 0.999 among the evaluated models. The Logistic Regression, KNN and Naïve Bayes classifiers were comparatively less effective in classification while Decision Tree showed good classification performance. The good performance of Random Forest indicates that ensemble-based methods can be able to capture complex and nonlinear relationships between acoustic features with musical mode in the analyzed data. The feature-level analysis also showed that certain features like danceability, energy, valence, BPM, acousticness and speechiness can be utilized to differentiate the Major and Minor songs. The results indicate that there are patterns that can be discovered in the measurable acoustic attributes, to be learned by machine learning models for musical mode classification. It is not appropriate to view these relationships as causal, however, because they are only predictive associations between acoustic features and song musical mode. From an application point of view, the proposed framework could be used in music recommendation systems, playlist generation, mood categorization in music, content organization and streaming-platform analytics. It can also be of great value to artists, producers and music marketers who want to gain insight into the correlations between measurable song features and musical categorization. On a broader scale, this paper provides a replicable machine-learning system for music mode classification from Spotify song data, that is applicable to the classification of music modes in general. The findings reveal the potential of AI and machine learning in the analysis of digital music through data. The proposed approach needs to be validated on larger and more diverse datasets and be extended to consider other musical, temporal and contextual features to increase the generalizability and interpretability of the musical mode classification.

7. Limitations of the Study

The proposed framework was able to obtain good classification accuracy, however, from the present study, there are some limitations. This dataset only covered Spotify and not other music streaming services like Deezer, YouTube Music or Apple Music. The study was mainly oriented toward features related to structuring sound and to streaming, without taking into account the lyrics, the sentiment of the texts, or semantic analysis of emotions. There were a few features that contained missing values, especially for the musical key and Shazam chart information which can possibly have an impact on the overall reliability of predictions. The data size is appropriate for experimentation, but still not as large as the one in large-scale music recommendation system in industry.

Record three more restrictions. First, the analysed sample of 858 songs is a subset of the Spotify catalogues, specifically from the top charts, and isn't representative of the entire library; the mode distribution and acoustic profile of chart music is not representative of the full catalogues, and the conclusion drawn from the chart-topping songs should not be applied to general-catalogue classification. Second, the musical mode is indicated by the pitch-class structure of a piece of music, while all the predictors in this collection are timbral, rhythmic or affective ones. On theoretical grounds, there is a limit to the accuracy that is possible and anything reported for an accuracy close to unity must be regarded as a testament to a defect in the pipeline and should not be interpreted as indicative of the accuracy of prediction until it can be replicated using an independently specified protocol. Third, the 95 records that are discarded for a missing key value may not be missing at random. Listwise deletion will throw away exactly the most difficult cases, and inflate the accuracy results if missingness is correlated with tonal ambiguity. It is recognised, but not corrected, because the tonal annotation is missing to quantify the bias.

8. Future Work

This study could be continued with several important directions in the future to enhance the predictive ability and further broaden the scope of music analytics application. More sophisticated

music classification models such as LSTM, BiLSTM, GRU, CNN and Transformer can be used for more complicated music classification tasks. To incorporate the audio and semantic emotional features in future work, it is suggested to integrate lyrical sentiment analysis, textual embeddings and natural language processing techniques. 3. Spotify, Apple Music, YouTube Music and Deezer can be used to collect large multi-platform data sets for improved generalization and robustness of the model. The techniques of Explainable Artificial Intelligence (XAI), such as SHAP and LIME, can be used to provide more interpretable explanations regarding the importance of features and classification decisions. The proposed approach of classification can be used, to develop real time recommendation system and intelligent playlists. Future research can also be dedicated to music prediction based on mood, emotion analysis, genre classification, and individual music recommendation systems. Hybrid ensemble and optimization-based machine learning approaches could be used for other improvements of classification accuracy and computation efficiency.

References

- [1]Heikkinen, S. (2024). Optimizing Spotify's business through big data analytics
- [2]Beesa, J., Naregavi, R., Imandar, N., & Thatte, S. (2023). Songs popularity analysis using Spotify data: An exploratory study. *Vidhyayana: An International Multidisciplinary Peer Reviewed E-Journal*, 8(si7), 211–223.
- [3]Spotify Financial Report. (2023). Annual report and financial statements 2023. Spotify Technology S.A.
- [4]Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard Business Review*, 90(10), 78–83.
- [5]Davenport, T. H. (2012). *Competing on analytics*. Harvard Business Review Press.
- [6]Ingle, N., Shah, R., & Mehta, J. (n.d.). Big data analytics: A Spotify case study. *IJRASET*. <https://doi.org/10.22214/IJRASET.2021.38702>
- [7]Saragih, H. S. (2023). Predicting song popularity based on Spotify's audio features: Insights from Indonesian streaming users. *Journal of Management Analytics*, 10(4), 693–709. <https://doi.org/10.1080/23270012.2023.2239824>
- [8]Bethapudi, P. (2024, April 13). Spotify data analysis and song popularity prediction (SSRN Scholarly Paper No. 4793176). SSRN. <https://doi.org/10.2139/ssrn.4793176>
- [9]Cabansag, I. J., & Ntegeka, P. (2025). Prediction of Spotify chart success using audio and streaming features. *arXiv*. <https://doi.org/10.48550/arXiv.2508.11632>
- [10] Falah, N., Yousefimehr, B., & Ghatee, M. (2025). Predicting music track popularity by convolutional neural networks on Spotify features and spectrogram of audio waveform. *arXiv preprint arXiv:2505.07280*.
- [11] Medows, K., Leisha, R., & Thiruthuvanathan, M. (2024). Predicting song popularity using data analysis. *Proceedings of InC460750* (pp. 1–6). <https://doi.org/10.1109/InC460750.2024.10649267>
- [12] Ranidu, M. Y., Mahanama, T. V., & Wijenarayana, S. (2024, April). Stream Count Predictive Analysis for Upcoming Songs on Spotify using Machine Learning: A Systematic Literature Review. In *2024 International Research Conference on Smart Computing and Systems Engineering (SCSE)* (Vol. 7, pp. 1-5). IEEE.

- [13] Braun, T. A. (2020). "Dance like nobody's paying": Spotify and surveillance as the soundtrack of our lives [Master's thesis, The University of Western Ontario].
- [14] Win, K. Z., & Raheem, M. (2024). Natsukashii: A Sentiment Emotion Analytics Based on Recent Music Choice on Spotify. *International Journal of Advanced Computer Science & Applications*, 15(6).
- [15] Purnama, M. J., Pramatha, C., & Dewi, C. A. M. (2025). Visual Analytics of Spotify Music Data for Listener Behavior Insights. *Krisnadana Journal*, 5(1), 244-253.
- [16] Espejo, A. (2023). Soundscape: Interactive data visualisations of Spotify's top songs. Wellington Faculty of Engineering Symposium. <https://ojs.wgtn.ac.nz/wfes/article/view/8399>
- [17] Chaudry, W. (2025, January 31). Exploring the determinants of music popularity: A quantitative analysis of Spotify data (SSRN Scholarly Paper No. 5126647). SSRN. <https://doi.org/10.2139/ssrn.5126647>
- [18] Grgurić, D., Perak, B., & Radosavljević, L. (2024). Streaming services: Revolutionizing music management. In *Aural experience and soundscape management* (pp. 153–184). Jenny Stanford Publishing.