

УДК 519.2:78

9. SPOTIFY-АНАЛИТИКА: КАКИЕ ХАРАКТЕРИСТИКИ ДЕЛАЮТ ПЕСНЮ ХИТОМ

Костромина П.О., Казачина К.А.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Шинкевич Е.А. – кандидат физико-математических наук, доцент

Аннотация. В работе используются методы статистического анализа, позволяющие выявить значимые зависимости между аудио характеристиками треков и их популярностью, а также признаков в формировании «хита».

Ключевые слова. Spotify, стриминговые платформы, популярность треков, музыкальная аналитика, аудиохарактеристики, danceability, energy, acousticness, instrumentalness, loudness, жанровая принадлежность, popularity, машинное обучение, метод случайного леса, взаимная информация, анализ данных, цифровая музыка, предсказание хитов.

Spotify – крупнейшая музыкальная стриминговая платформа в мире: 751 млн ежемесячно активных пользователей, 276 млн платных подписчиков, более 100 млн треков в каталоге и 31,7% мирового рынка стриминга (SQ Magazine, 2026). Ежедневно на платформу загружается около 60 000 новых треков, что формирует беспрецедентную конкуренцию за внимание слушателя.

Глобальная выручка индустрии записанной музыки в 2024 году достигла 29,6 млрд USD, из которых 69% (20,4 млрд USD) приходится на стриминг (IFPI Global Music Report, 2025). Алгоритмическая система Spotify – BaRT (Bandits for Recommendations as Treatments) – генерирует около 40% всех стримов, обрабатывая более 500 млрд точек данных ежедневно и анализируя до 1 500 звуковых характеристик каждого трека (Music Ally, 2026; Productify, 2022). Алгоритм принимает первичное решение о «качестве» трека в течение первых 72 часов после релиза, опираясь на save rate, completion rate и repeat listening (iMusician, 2026).

В условиях большого объема данных и высокой вариативности музыкальных характеристик применение статистического анализа является необходимым инструментом исследования. Статистические методы позволяют выявлять закономерности в данных, оценивать силу и значимость связей между переменными, а также делать обоснованные выводы на основе количественных показателей.

Целью работы является выявление аудиохарактеристик, наиболее значимых для достижения треком статуса «хит» на Spotify, на основе статистического анализа датасета из 87 944 уникальных треков 113 жанров с использованием 15 признаков: 12 аудиохарактеристик Spotify Web API (danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence, tempo, key, mode, time_signature), длительность (duration_min), метка explicit и жанровая принадлежность (genre_encoded).

В качестве целевой переменной использована метрика popularity (0–100). Порог хита установлен на уровне popularity ≥ 50 , что соответствует 75-му перцентилю распределения (верхний квартиль). При данном пороге хиты составляют 24,0% выборки (~21 100 треков), не-хиты – 76,0% (~66 800 треков). Распределение popularity (рис. 1) носит бимодальный характер: значительная доля треков сконцентрирована в области 0–10, второй пик – в диапазоне 30–50. Медиана составляет 34, среднее – 33,4.

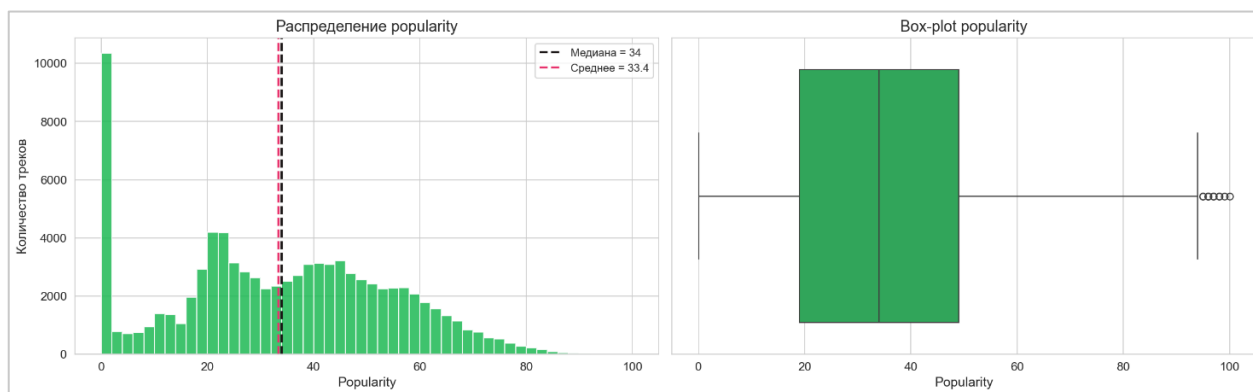


Рисунок 1 – Распределение popularity

Анализ чувствительности к выбору порога (рис. 2) подтверждает оптимальность: при пороге 40 доля хитов возрастает до 40,8% (размывая понятие «хит»), при 60 – сокращается до 11,1%, при 70 – до 3,5%,

при 80 – до 0,7%. Резкое экспоненциальное падение доли хитов при порогах выше 60 делает порог 50 оптимальным компромиссом между разделяющей способностью и объемом обоих классов.

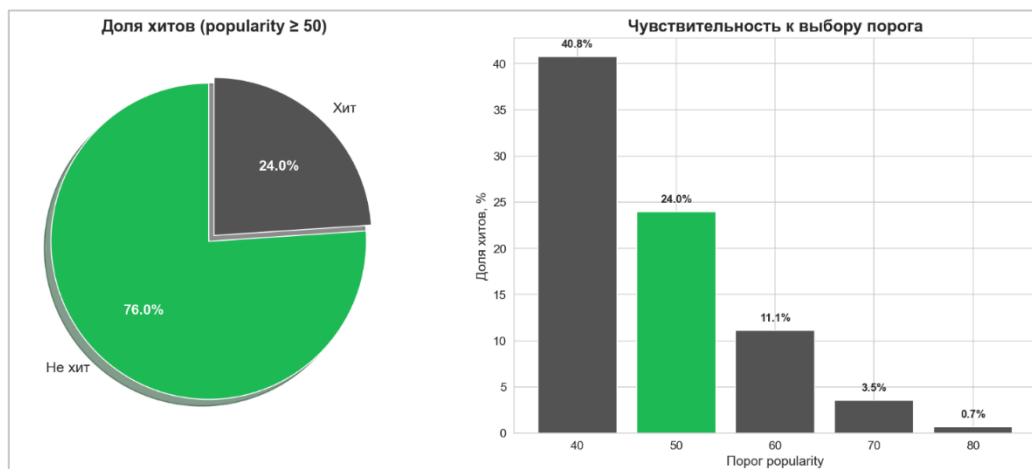


Рисунок 2 – Доля хитов и анализ чувствительности порога

Наложенные гистограммы плотности (рис. 3) и box-plots (рис. 4) для 12 аудиопризнаков выявляют три признака с наиболее выраженными различиями между классами:

1 Danceability – распределение хитов смещено вправо (медиана ~0,60 vs ~0,55), что свидетельствует о более высокой танцевальности популярных треков.

2 Instrumentalness – подавляющее большинство хитов сконцентрировано вблизи нулевых значений, тогда как не-хиты имеют значительно более широкое распределение с длинным «хвостом» до 1,0. Хиты – это преимущественно вокальные композиции.

3 Acousticness – хиты чаще имеют низкие значения, что указывает на преобладание электронно обработанного звука.

Добавленные признаки key, mode и time_signature демонстрируют практически полное перекрытие распределений между классами, что заранее указывает на их слабую дифференцирующую способность.

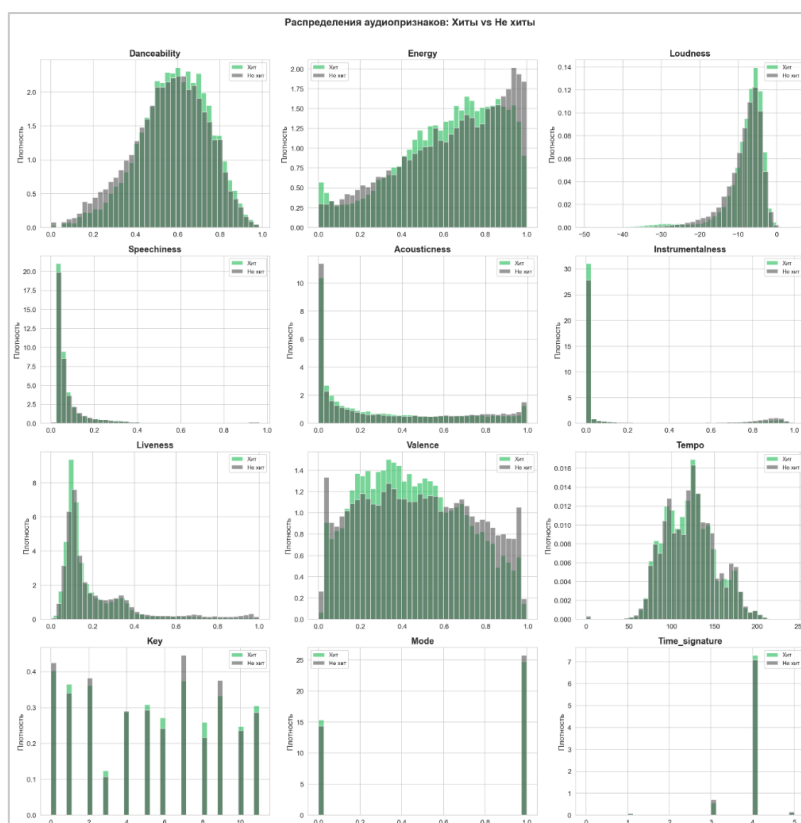


Рисунок 3 – Распределения аудиопризнаков: хиты vs не-хиты

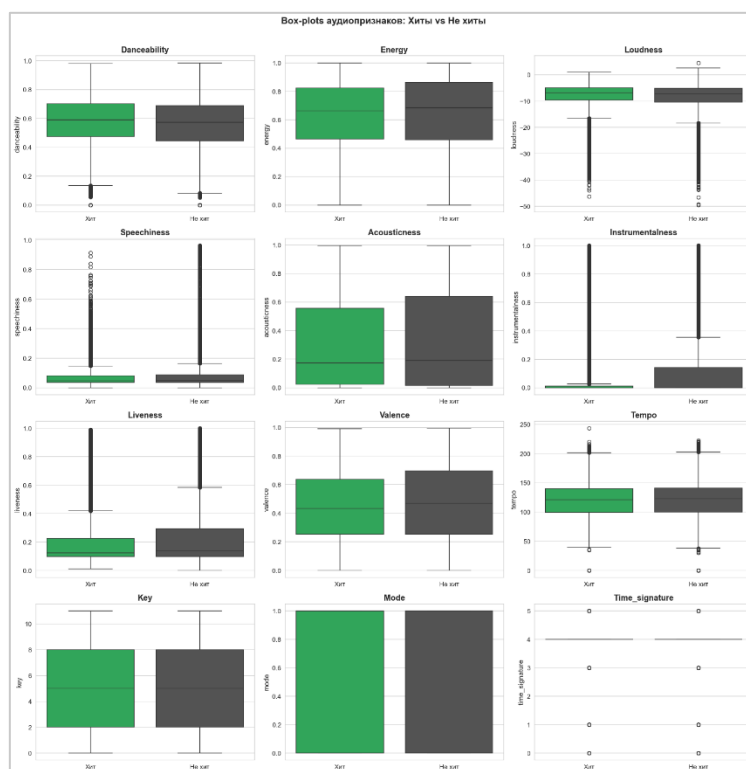


Рисунок 4 – Box-plots аудиопризнаков

Стандартизированные размеры эффекта с использованием pooled standard deviation (рис. 5) рассчитаны для всех 12 аудиопризнаков. Ни одно значение $|d|$ не превышает порога малого эффекта по классификации Коэна ($|d| = 0,2$). Наибольшие по модулю: liveness ($d = -0,206$), instrumentalness ($d = -0,196$), danceability ($d = +0,123$). Добавленные признаки подтвердили свою неинформативность: key ($d = +0,005$), mode ($d = -0,054$), time_signature ($d = +0,059$) – все ниже порога $|d| = 0,1$.

Это ключевой результат: при статистической значимости различий ($p < 0,001$ для всех признаков при $n \approx 88\,000$) практическая значимость каждого отдельного аудиопризнака невелика. Большая выборка делает статистически значимыми даже тривиальные различия – именно поэтому необходимо опираться на размер эффекта, а не на p-value.

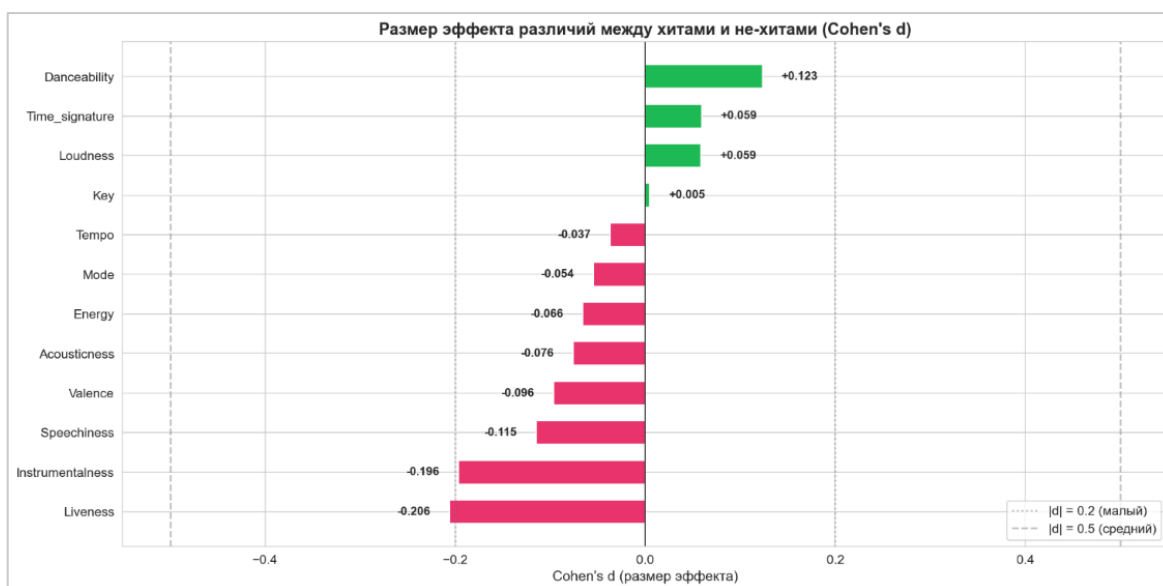


Рисунок 5 – Размер эффекта коэффициента d Коэна

Радарная диаграмма (рис. 6) обобщает «портрет хита»: более высокие danceability ($\sim 0,59$ vs $\sim 0,56$) и energy ($\sim 0,65$ vs $\sim 0,63$), более низкие instrumentalness ($\sim 0,09$ vs $\sim 0,19$) и acousticness ($\sim 0,29$ vs $\sim 0,33$). Области хитов и не-хитов значительно перекрываются, что визуально подтверждает малые значения коэффициента d Коэна.

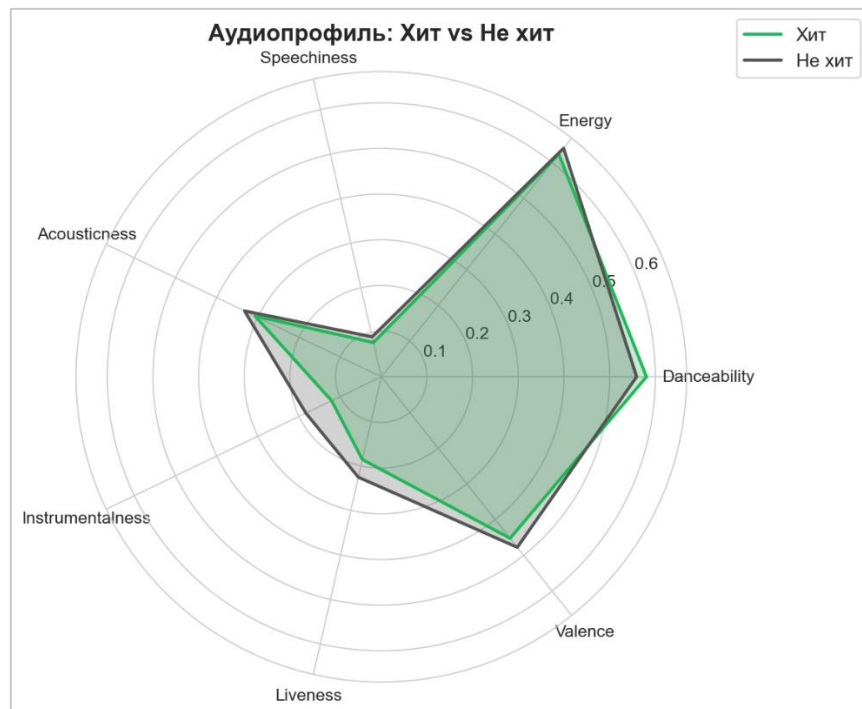


Рисунок 6 – Аудиофиль: хит vs не-хит

Матрица корреляций Пирсона (рис. 7), рассчитанная для 15 переменных, выявляет сильные межпризнаковые связи (мультиколлинеарность): energy–loudness ($r = 0,76$), energy–acousticness ($r = -0,73$), danceability–valence ($r = 0,49$). Однако корреляции всех аудиопризнаков с целевой переменной popularity крайне слабые: максимальное $|r| = 0,12$ (instrumentalness). Добавленные признаки key ($r = 0,00$), mode ($r = -0,02$) и time_signature ($r = 0,03$) подтвердили отсутствие линейной связи с popularity.

Scatter-plots (рис. 8) четырёх наиболее коррелирующих признаков наглядно демонстрируют облачный характер зависимости: даже для instrumentalness ($r = -0,119$) коэффициент детерминации составляет лишь $R^2 \approx 0,014$, что означает объяснение $\sim 1,4\%$ дисперсии popularity. Корреляция не является причиной – наблюдаемые связи могут объясняться конфаундерами (жанр, известность артиста и др.).

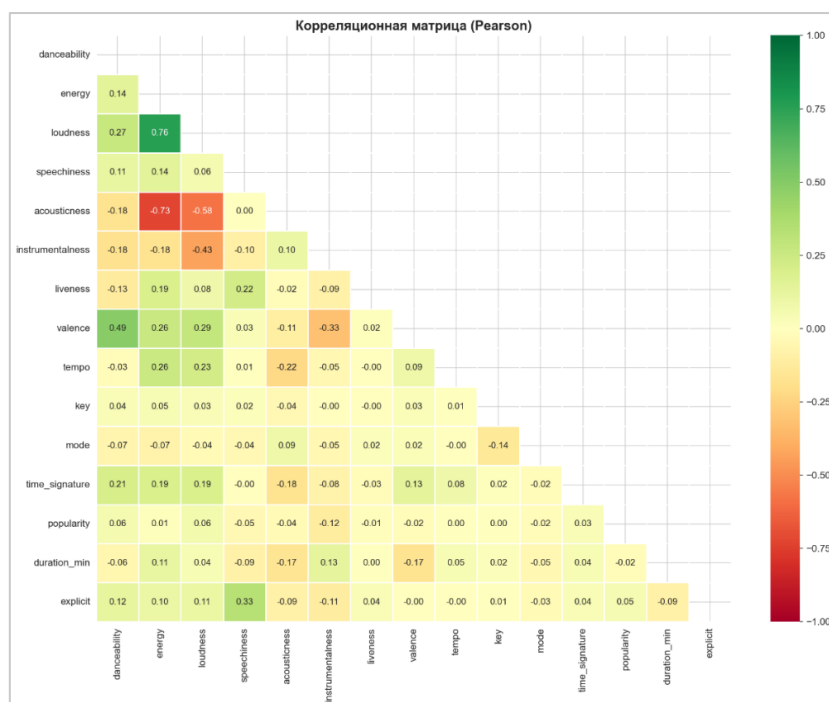


Рисунок 7 – Корреляционная матрица (Пирсон)

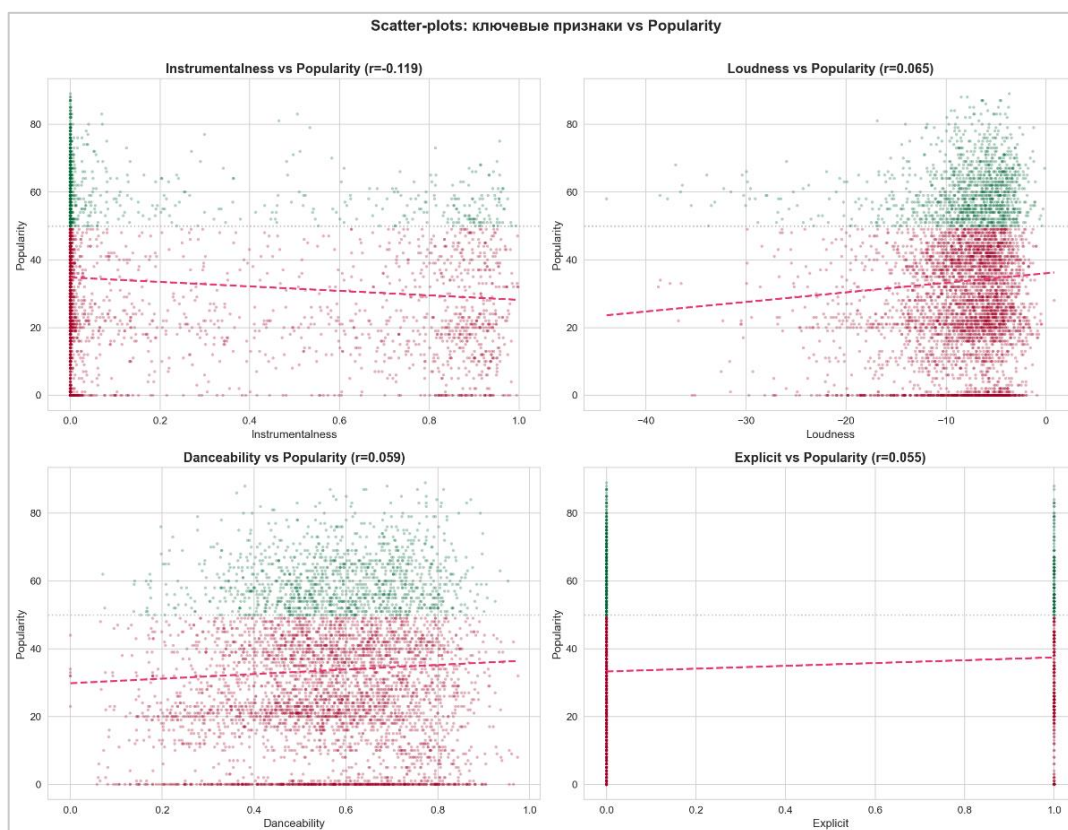


Рисунок 8 – Scatter-plots ключевых признаков

Анализ топ-20 жанров (рис. 9) показывает выраженную жанровую дифференциацию: k-pop (средний popularity ≈ 59 , доля хитов $\approx 89\%$), pop-film (≈ 56 / $\approx 92\%$), metal (≈ 53 / $\approx 73\%$). Тепловая карта нормализованных аудиопрофилей (рис. 10) демонстрирует, что каждый жанр обладает уникальной комбинацией аудиопризнаков: reggaeton – максимальные danceability (0,96) и valence (1,00); metal – максимальные energy (1,00) и loudness (1,00) при нулевой acousticness; piano – максимальная acousticness (1,00) и instrumentalness (1,00) при минимальных energy и loudness. Таким образом, жанр является фактором, определяющим комбинацию аудиопризнаков, – и именно этот фактор, как будет показано далее, несёт наибольшую предсказательную информацию.

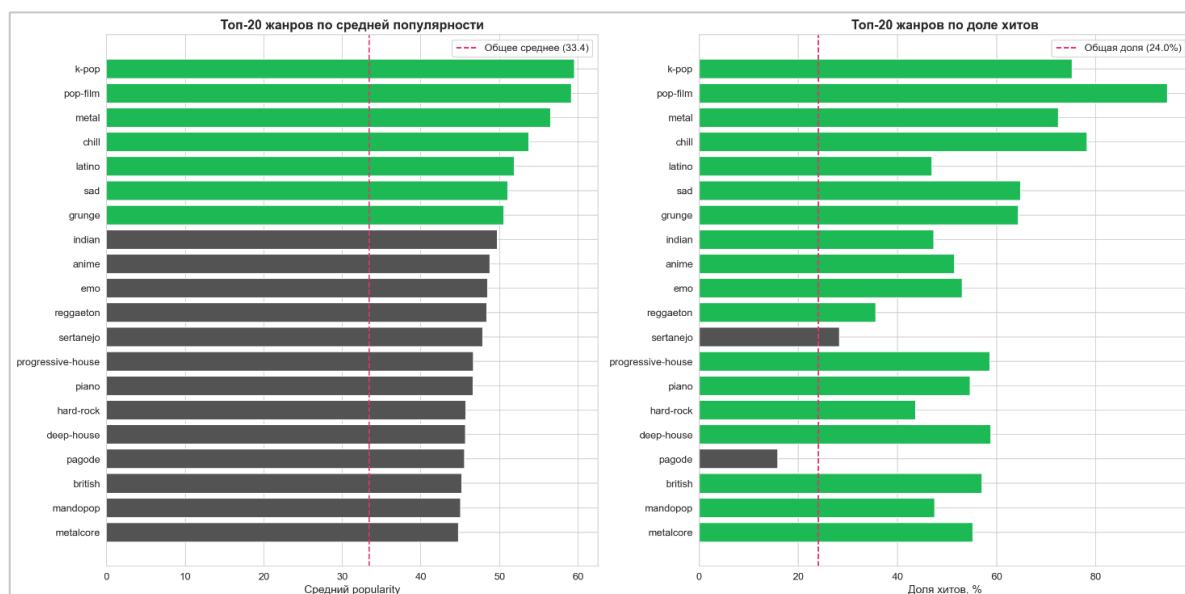


Рисунок 9 – Топ-20 жанров

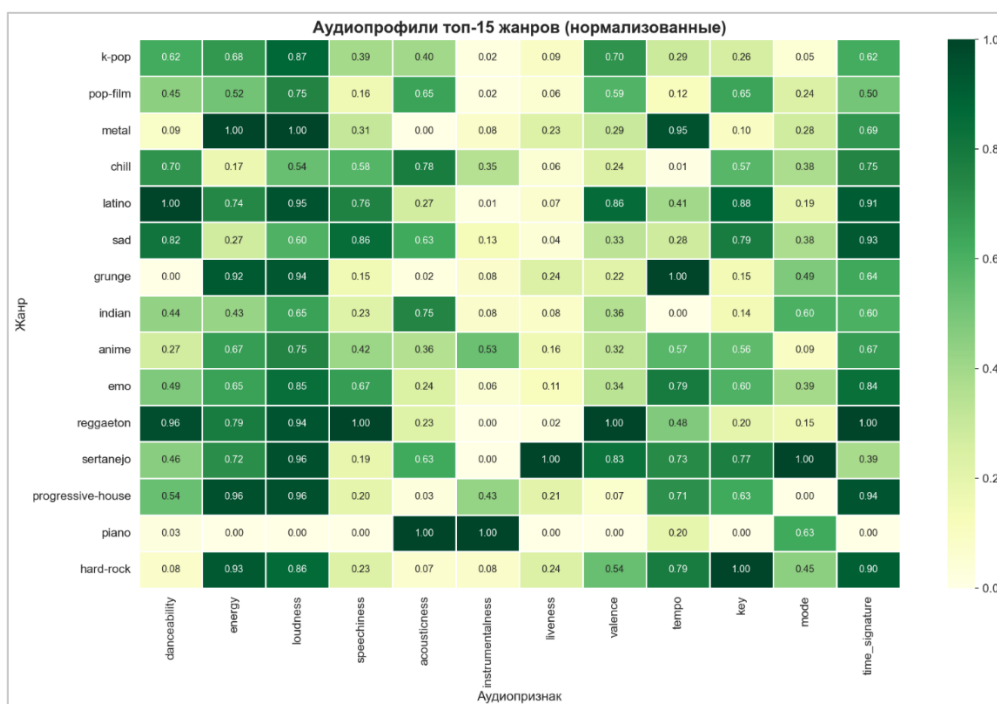


Рисунок 10 – Аудиопрофили топ-15 жанров

Треки с explicit-контентом имеют среднюю popularity 37,1 против 33,0 (рис. 11), доля хитов – 32,6% vs 23,1%. Данная ассоциация не свидетельствует о причинности: explicit-контент сконцентрирован в жанрах hip-hop и gar, чья аудитория (18–34 лет) наиболее активна на Spotify.

Оптимальная длительность хита – 3,0-4,5 минуты (рис. 12): треки этого диапазона демонстрируют наивысшую среднюю popularity (~35). При длительности свыше 5 минут средняя popularity стабилизируется на уровне ~27–28, что согласуется с индустриальной практикой.

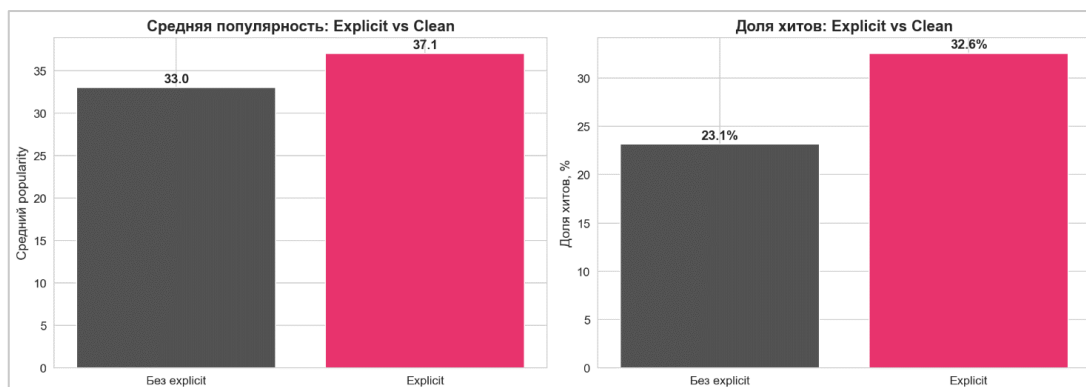


Рисунок 11 – Влияние explicit-метки

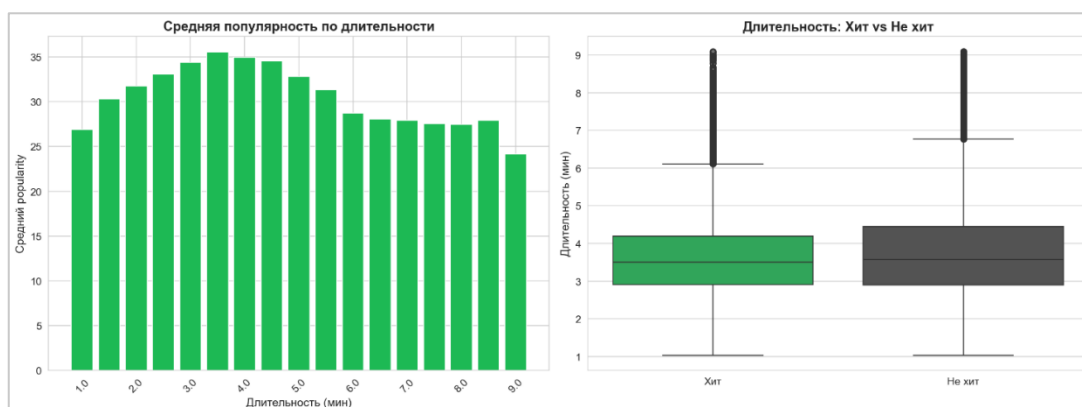


Рисунок 12 – Анализ длительности

Для ранжирования 15 признаков использованы два независимых метода:

1 Взаимная информация (рис. 13) оценивает нелинейную зависимость между признаком и бинарной целевой переменной `is_hit`. Результат: `genre_encoded` является абсолютным лидером с $MI \approx 0,135$ – это в 9 раз больше, чем у ближайшего аудиопризнака ($time_signature \approx 0,015$, $duration_min \approx 0,015$). Все прочие признаки группируются в узком диапазоне 0,001–0,015.

2 Важность признаков (метод случайного леса) (рис. 14) оценивает вклад признака через снижение Gini impurity. Результат: `genre_encoded` лидирует с $importance \approx 0,24$ – в 3 раза больше ближайшего ($loudness \approx 0,085$). Точность ООВ (вневыборочная оценка) 76,5% означает, что модель на 15 признаках корректно классифицирует более трёх четвертей треков без выделения отдельной тестовой выборки – значительно лучше случайного угадывания (50%).

Итоговый комбинированный рейтинг (рис. 15) усредняет нормализованные оценки обоих методов. `Genre_encoded` занимает абсолютное первое место (1,000), далее с колоссальным отрывом: `duration_min` (0,201), `loudness` (0,196), `instrumentalness` (0,174), `energy` (0,172), `valence` (0,170), `acousticness` (0,166), `liveness` (0,160). Добавленные признаки `key` (0,013), `mode` (0,029) и `time_signature` (0,053) заняли последние позиции, что подтверждает их неинформативность – сам по себе научный результат, демонстрирующий, что тональность и музыкальный размер не влияют на популярность.

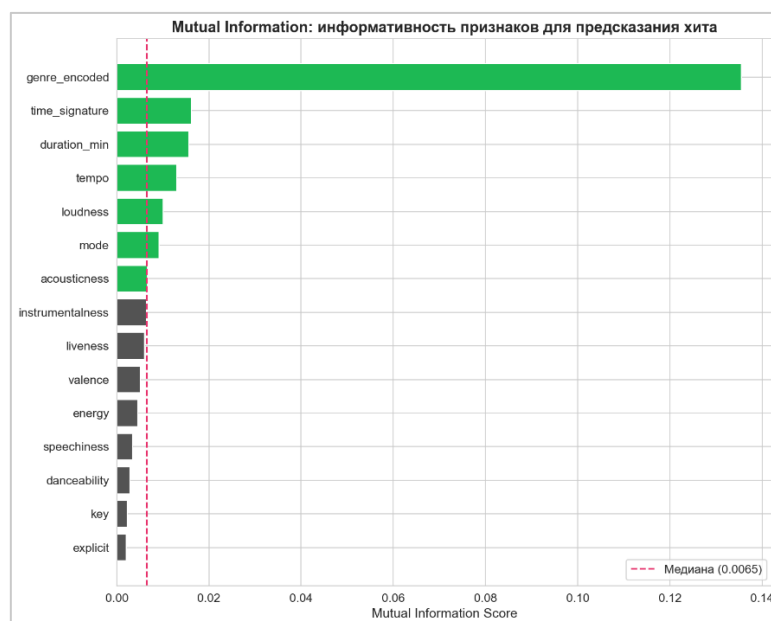


Рисунок 13 – Mutual Information

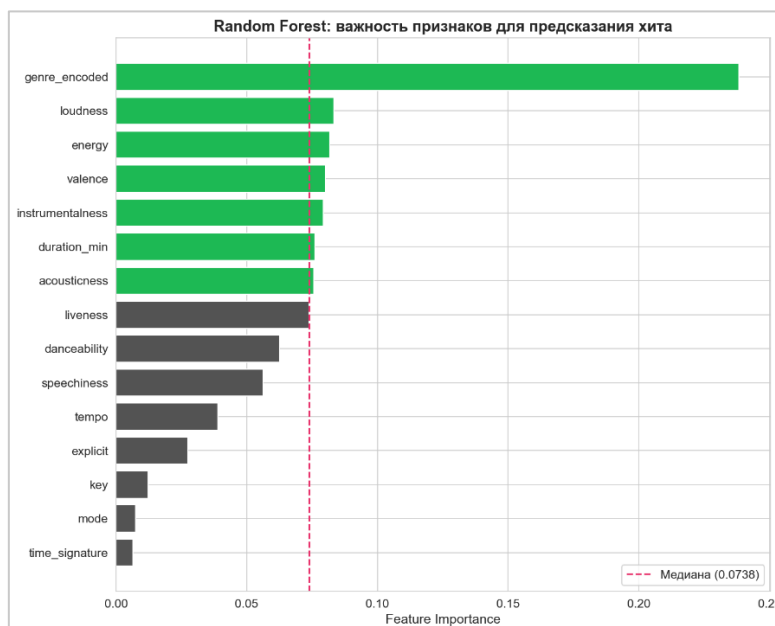


Рисунок 14 – Random Forest Feature Importance

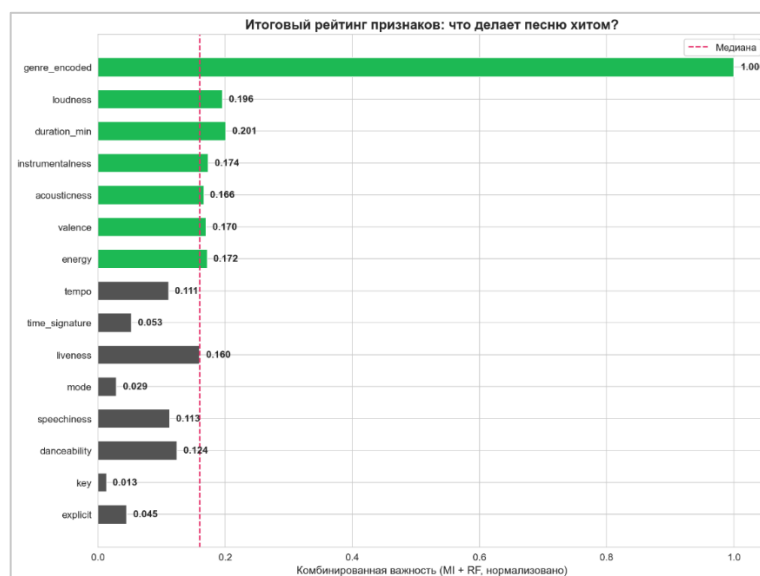


Рисунок 15 – Итоговый рейтинг признаков

Проведённый статистический анализ 87 944 треков по 15 признакам позволяет сформулировать следующие выводы:

Жанр – главный предиктор хита. Жанровая принадлежность (`genre_encoded`) несёт в 5–9 раз больше информации о принадлежности трека к классу «хит», чем любой отдельный аудиопризнак ($MI = 0,135$ vs $\leq 0,015$; RF importance = $0,24$ vs $\leq 0,085$). Жанр выступает латентным фактором, определяющим комбинацию аудиохарактеристик.

Аудиопризнаки по отдельности – слабые предикторы. Все корреляции с popularity не превышают $|r| = 0,12$, все значения коэффициента d Коэна $< 0,21$. Тональность (`key`), лад (`mode`) и музыкальный размер (`time_signature`) оказались неинформативными ($|d| < 0,06$, $|r| < 0,03$).

В совокупности признаки работают. Модель Random Forest на 15 признаках достигает Точность ООВ (вневыборочная оценка) 76,5% – значительно лучше случайного угадывания (50%). «Формула хита» существует, но определяется не одним-двумя параметрами, а комбинацией жанра и набора аудиохарактеристик.

Обобщённый «портрет хита»: вокальная композиция (низкая instrumentalness), с высокой танцевальностью ($danceability \approx 0,59$), умеренно громкая ($loudness \approx -7$ дБ), с низкой акустичностью, длительностью 3–4,5 минуты.

За пределами аудиоанализа остаются доминирующие факторы: известность артиста, маркетинг, плейлисты, поведенческие сигналы – их включение в модель предположительно существенно повысит точность предсказания.

Список использованных источников:

1. Burtsev, M. Spotify Dataset – 114,000 songs with audio features and genre labels [Electronic resource] / M. Burtsev // Kaggle, 2023. – URL: <https://www.kaggle.com/datasets/maharshipandya/spotify-tracks-dataset>. – Date of access: 13.03.2026.
2. Google Colaboratory [Electronic resource] // Google LLC. – URL: <https://colab.research.google.com>. – Date of access: 13.03.2026.
3. McKinney, W. Data Structures for Statistical Computing in Python / W. McKinney // Proceedings of the 9th Python in Science Conference, 2010. – P. 51–56.

UDC 519.2:78

SPOTIFY ANALYTICS: WHICH CHARACTERISTICS MAKE A SONG A HIT

Kostromina P.O., Kazachina K.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Shinkevich E.A. – PhD in Physics and Mathematics, Associate Professor

Annotation. The paper applies statistical analysis methods to identify significant relationships between audio features of tracks and their popularity, as well as key factors influencing the formation of a “hit” on the Spotify platform. The study is based on a dataset of music tracks and focuses on evaluating the importance of various audio characteristics using data analysis and machine learning techniques.

Keywords. Spotify, music streaming, track popularity, audio features, danceability, energy, acousticness, instrumentalness, genre classification, machine learning, Random Forest, mutual information, data analysis, hit prediction.