

DEVELOPMENT OF A SOFTWARE MODULE FOR COLLECTING AND PROCESSING MEDIA PLAYBACK STATISTICS USING APIS

Vinnitsia National Technical University

Анотація:

В даній роботі представлено результати розробки програмного модуля для збору та обробки статистичних даних із відкритих API музичних платформ Spotify та YouTube. Основною метою дослідження було створення інструменту для автоматизованого збору метаданих про треки та відео, їхню обробку та аналіз популярності музичного контенту. Модуль дозволяє отримувати актуальну інформацію про кількість прослуховувань, вподобань, жанри, тривалість композицій тощо. Проведений статистичний аналіз виявив характерні тенденції в динаміці популярності контенту, а також кореляцію між активністю користувачів на різних платформах. Результати тестування підтвердили стабільність роботи системи та її придатність для обробки великих обсягів даних. Отримані результати можуть бути використані для побудови аналітичних систем, прогнозування трендів та вивчення поведінки користувачів у цифровому музичному середовищі.

Ключові слова: *Spotify, YouTube, API, музична аналітика, обробка даних, статистика, програмний модуль, Python, популярність контенту, машинна обробка.*

Abstract:

This paper presents the results of the development of a software module for collecting and processing statistical data from open APIs of music platforms Spotify and YouTube. The main goal of the research was to create a tool for automated collection of metadata about tracks and videos, their processing and analysis of the popularity of musical content. The module allows you to obtain up-to-date information about the number of listenings, likes, genres, duration of compositions, etc. The statistical analysis revealed characteristic trends in the dynamics of content popularity, as well as correlation between user activity on different platforms. The test results confirmed the stability of the system and its suitability for processing large amounts of data. The results obtained can be used to build analytical systems, predict trends and study user behavior in the digital music environment.

Keywords: *Spotify, YouTube, API, music analytics, data processing, statistics, software module, Python, content popularity, automated analysis.*

Introduction

In today's digital environment, platforms like Spotify and YouTube play a significant role in distributing music content and shaping user preferences. Listening and viewing statistics are valuable sources for analyzing track popularity and identifying media trends. The aim of this study is to develop a software module for the automated collection and processing of statistical data from the public APIs of these services. The solution enables preliminary data analysis, helps detect current trends, and lays the foundation for further research in the field of digital analytics.

Research results

As a result of the conducted study, a software module was developed to automate the collection and preliminary processing of statistical data from the Spotify and YouTube APIs [1]. The goal of the project was to obtain up-to-date information on music content, its popularity, listening/viewing dynamics, and user interaction with specific artists or tracks.

The software module establishes full interaction with the public APIs of both platforms, enabling real-time access to data on tracks and videos, including titles, artist names, number of streams or views, likes, genres, duration, publication dates, and other attributes [2]. The data is retrieved in a structured

format (JSON) and further processed using Python with libraries such as pandas, requests, matplotlib, and seaborn.

Special attention was paid to building an efficient mechanism for filtering and aggregating data. This made it possible to structure information by time periods, genres, regional features, and content popularity levels. Based on the collected and processed data, statistical analysis was performed, revealing general trends in music consumption – such as the rising popularity of short video formats, increasing interest in independent artists, and noticeable seasonality in user preferences [3].

The study demonstrated a degree of correlation between user activity on Spotify and YouTube: tracks gaining popularity on one platform often went viral on the other [4]. However, some discrepancies were identified, attributed to differences in target audiences, recommendation algorithms, and content types consumed on each platform.

The module was tested on a sample of over 1,000 tracks and videos, with data collected over a two-week period. The system showed stable performance regardless of data volume or request frequency. The average execution time per request and data processing was approximately 150–200 milliseconds, which is suitable for real-time use or integration into larger analytics platforms [5].

Looking ahead, the module can be expanded by integrating with databases for storing historical data, applying machine learning methods to predict content popularity shifts, and creating a visual interface for dynamic analytics exploration.

Conclusion

Thus, the results of the study confirmed the effectiveness of the developed software module for collecting and processing statistical data using the Spotify and YouTube APIs. The implemented software solution allows for the automation of the process of obtaining information about music content, structuring it, and performing preliminary analysis with subsequent construction of analytical conclusions. Practical tests confirmed the stability of the module's operation, its performance, and its adaptability to changes in the data formats of external APIs. The conducted statistical analysis revealed general trends in content consumption and allowed for the assessment of correlations between track popularity on different platforms. The results obtained can serve as a foundation for further research in the field of music analytics, the development of recommendation systems, or marketing strategies in the digital media space.

LIST OF REFERENCES

1. Karypis, G. Evaluation of item-based collaborative filtering algorithms / Karypis G. // Proceedings of the 2001 international conference on Knowledge discovery and data mining, ACM, 2001. - pp. 143-148.
2. Dastin, J. Spotify API: Developer's Guide / Dastin J. // O'Reilly Media, 2015. - 350 p.
3. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & King, I. Neural collaborative filtering / He X., Liao L., Zhang H., Nie L., Hu X., King I. // Proceedings of the 26th International Conference on World Wide Web, 2017. - pp. 173-182.
4. Zhang, Y., & Chen, X. A Survey of Hybrid Recommender Systems / Zhang Y., Chen X. // International Journal of Computer Science Issues (IJCSI), 2020. - 17(2), pp. 48-57.
5. Shardanand, U., & Maes, P. Social information filtering: algorithms for automating "word of mouth" / Shardanand U., Maes P. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1995. - pp. 210-217.

Карась Олександр Володимирович – доктор філософії, старший викладач, Вінницький національний технічний університет, м. Вінниця, e-mail: karas@vntu.edu.ua

Щерба Антон Олегович – студент групи 2ПІ-21б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, Україна, email: anton.shcerba04@gmail.com

Karas Oleksandr Volodymyrovych – PhD, Senior Lecturer, Vinnytsia National Technical University, Vinnytsia, e-mail: karas@vntu.edu.ua

Shcherba Anton Olegovich – student of group 2PI-21b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, Ukraine, email: anton.shcerba04@gmail.com