

SISTEM REKOMENDASI MUSIK DENGAN MULTIMODAL FUSION ATTENTIVE NETWORKS

Jacky Chan¹, Neveline², Sherina³, Frans Mikael Sinaga⁴, Arwin Halim^{5*}

^{1,2,3,4,5}Teknik Informatika, Universitas Mikroskil; Jl. Thamrin no 124 Medan, Sumatera Utara

Riwayat artikel:

Received: 3 Oktober 2022

Accepted: 29 Desember 2023

Published: 1 Januari 2024

Keywords:

BERT, deep learning, MFAT

Correspondent Email:

arwin@mikroskil.ac.id

Abstrak. Musik merupakan salah satu media hiburan yang sering dimanfaatkan oleh masyarakat. Platform musik terus meningkatkan organisasi musik dengan menyediakan berbagai lagu untuk didengarkan oleh penggunanya. Secara umum sistem rekomendasi musik menggunakan metode Collaborative Filtering yang memiliki kinerja yang baik, tetapi metode ini memiliki masalah terhadap informasi yang tidak memiliki interaksi umpan balik antara item dan pengguna (*cold-start*). Sistem rekomendasi Content-Based diusulkan untuk mengatasi masalah tersebut yang akan difokuskan pada pendekatan *multimodal fusion* agar dapat mengoptimalkan hasil rekomendasi. Sistem rekomendasi ini mempertimbangkan dua pendekatan dengan fokus utama mengekstraksi fitur dari modalitas individu seperti judul lagu menggunakan model Word2Vec dan lirik menggunakan model BERT yang kemudian digabungkan melalui arsitektur *deep learning* untuk memprediksi rekomendasi lagu berikutnya. Hasil pengujian menggunakan *black-box testing* menunjukkan bahwa perangkat lunak secara fungsional mengeluarkan hasil yang sesuai dengan yang diharapkan dan sudah mencapai tujuan yang sesuai dengan kebutuhan pengguna. Penerapan algoritma rekomendasi akan diuji menggunakan *confusion matrix* dan memperoleh nilai akurasi yang cukup akurat yaitu sebesar 84%.

Abstract. Music is one of the entertainment media that is often used by the public lately. The music platform continues to improve the organization of music by providing a variety of songs for its users to listen to. In general, the music recommendation system uses the Collaborative Filtering method which has good performance, but this method has problems with information that does not have feedback interaction between items and users (*cold-start*). A Content-Based recommendation system is proposed to overcome this problem which will focus on a multimodal fusion approach in order to optimize the recommendation results. This recommendation system considers two approaches with the main focus on extracting features from individual modalities such as song titles using the Word2Vec model and lyrics using the BERT model which are then combined through deep learning architecture to predict the next song recommendation. The test results using black-box testing show that the software functionally produces results that are as expected and have achieved the goals according to user needs. Meanwhile, in the recommendation test, the confusion matrix was used and obtained a fairly accurate accuracy value of 84%.

1. PENDAHULUAN

Musik merupakan salah satu media hiburan yang dibutuhkan oleh sebagian orang untuk menenangkan pikiran dan suasana hatinya.

Dalam beberapa tahun terakhir, industri musik semakin bergeser ke arah digital yang membuat masyarakat lebih memilih mendengarkan musik secara *online*. Hal ini dapat ditunjukkan dari

survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), dengan 35,5% dari populasi pengguna internet di Indonesia atau sekitar 46,9 juta orang Indonesia yang mendengarkan musik secara *online* [1]. Spotify adalah layanan penyedia *streaming* musik yang memiliki lebih dari 70 juta data musik, dimana penggunaannya hanya melakukan interaksi dengan sebagian dari data tersebut. Musik dengan interaksi yang sedikit seperti inputan data musik baru (*cold-start*) akan sulit untuk direkomendasikan, sehingga diperlukan sistem rekomendasi yang dapat mengatasi masalah ini. Musik memiliki representasi dari beberapa modalitas seperti fitur akustik, tag (contoh: artis, *genres*, instrumen), maupun lirik. Sistem rekomendasi musik yang dikembangkan pada bidang penelitian *Music Information Retrieval* (MIR) pada umumnya hanya memanfaatkan informasi dari salah satu modalitas [2], tetapi dalam penelitian ini digunakan sistem multimodal dengan memanfaatkan beberapa informasi dari modalitas musik.

Sistem rekomendasi musik secara umum menggunakan metode *Collaborative Filtering* (CF). Metode CF cukup efisien dan memiliki kinerja yang baik, tetapi metode ini memiliki masalah terhadap informasi yang tidak memiliki interaksi umpan balik antara *item* dan pengguna (*cold-start*). Metode *Content-Based* disarankan untuk mengatasi masalah yang terjadi pada CF dengan menggunakan konten berbasis *track* untuk mendapatkan hasil rekomendasi [3].

Penelitian ini akan menyelesaikan permasalahan yang ada dengan menerapkan algoritma *deep learning* yaitu *Attentive Neural Networks* dan *Recurrent Neural Networks* pada sistem rekomendasi musik. Sistem rekomendasi ini dioptimalkan kepada pengguna dengan menggunakan pendekatan *Multimodal Fusion*, dengan fokus utama mengekstraksi fitur dari modalitas individu seperti judul lagu dan lirik yang kemudian digabungkan melalui arsitektur *deep learning* untuk memprediksi preferensi lagu. Selain dapat memprediksi preferensi lagu, dalam sistem ini dapat melakukan pencarian lagu, album dan artis, dapat mendengarkan lagu secara *streaming online* dan dapat melihat lirik lagu.

2. TINJAUAN PUSTAKA

2.1. Musik

Musik berasal dari bahasa Yunani yaitu *mousike* yang berarti kesenian dalam menyusun suara –nada sedemikian rupa sehingga mengandung irama, lagu dan keharmonisan. Johansson[4] berpendapat bahwa “Musik merupakan suatu keunikan istimewa yang diciptakan oleh manusia yang mempunyai kapasitas yang sangat kuat untuk menyampaikan emosi dan mengatur emosi”. Menurut Kamus Besar Bahasa Indonesia [5] musik adalah ilmu atau seni menyusun nada atau suara diutarakan, kombinasi dan hubungan temporal untuk menghasilkan komposisi suara yang mempunyai keseimbangan dan kesatuan nada atau suara yang disusun sedemikian rupa sehingga mengandung irama, lagu dan keharmonisan.

2.2. Sistem Rekomendasi

Sistem rekomendasi merupakan suatu perangkat lunak yang mendukung pengguna dalam menyarankan informasi yang berguna atau memperkirakan apa yang akan dilakukan pengguna untuk mencapai tujuannya. Rekomendasi yang diberikan mempunyai hubungan dengan beragam proses dalam melakukan pengambilan keputusan seperti objek apa yang layak untuk dibeli, musik apa yang sesuai untuk didengarkan, ataupun informasi online apa yang sebaiknya untuk dibaca. Sistem rekomendasi merupakan sebuah perangkat lunak yang bertujuan untuk membantu user dengan cara memberikan rekomendasi kepada user ketika pengguna dihadapkan dengan jumlah informasi yang besar. Rekomendasi yang diberikan diharapkan dapat membantu pengguna dalam proses pengambilan keputusan, seperti barang apa yang akan dibeli, buku apa yang akan dibaca, atau musik apa yang akan didengar, dan lainnya [6].

2.3. Multimodal Fusion

Multimodal Fusion adalah salah satu penelitian multimodal yang populer dan dianggap sebagai penelitian yang penting meskipun merupakan penelitian yang baru [7], multimodal fusion dapat membantu dalam membedakan setiap kata secara dinamis. Dalam multimodal fusion terdapat jaringan yang

dimanfaatkan untuk mendeteksi gangguan [8], dan terdapat sebuah algoritma yang digunakan untuk menggabungkan informasi dari berbagai jenis data dengan tujuan untuk mencapai kinerja yang lebih baik [9]. Dalam mencapai kinerja yang lebih baik, multimodal fusion menetapkan bahwa dalam sebuah kombinasi terdapat dua maupun lebih dari dua modalitas. Dalam multimodal fusion terdapat teknik utama seperti tahapan awal, tahapan menengah dan tahapan akhir. Pada tahapan awal, data yang diperoleh dari modalitas yang berbeda akan digabungkan dalam satu representasi sebelum masuk pada tahapan menengah. Pada tahapan menengah, tahapan ini akan menangani masalah data yang tidak sempurna, masalah keandalan dan asinkron diantara modalitas yang berbeda. Dan pada tahapan terakhir, data akan diproses secara terpisah kemudian pada akhirnya akan digabungkan. [10].

2.4. Deep Learning

Deep Learning merupakan salah satu metode implementasi dari machine learning yang bertujuan untuk meniru cara kerja otak manusia dengan menggunakan artificial neural network atau jaringan nalar buatan. Deep Learning adalah teknik dalam Jaringan Saraf Tiruan yang menggunakan teknik tertentu seperti Restricted Boltzmann Machine (RBM) untuk mempercepat proses pembelajaran dalam jaringan saraf tiruan yang menggunakan lapisan yang banyak atau lebih dari 7 lapis. Dengan adanya Deep Learning, waktu yang dibutuhkan untuk training akan semakin sedikit karena masalah hilangnya gradien pada propagasi balik akan semakin rendah. Beberapa jenis Deep Learning antara lain Deep Auto Encoder, Deep Belief Nets, CNN, dan lain lain [11].

Pada penelitian ini, beberapa jaringan saraf tiruan yang digunakan antara lain:

1. Attentive Neural Networks. Attentive Neural Networks digunakan untuk memanfaatkan urutan item yang dipilih oleh pengguna untuk memodelkan kepentingan jangka pendek pengguna pada setiap item akan diberikan bobot yang berbeda [12]
2. Recurrent Neural Networks
Recurrent Neural Network (RNN) adalah salah satu model deep learning yang biasa digunakan untuk mengelola data yang bersifat sekuensial. RNN

disebut berulang karena melakukan sebuah tugas yang sama untuk setiap elemen yang berurutan, dengan output tergantung pada perhitungan sebelumnya [13]. Peningkatan popularitas ini dikarenakan hasil yang dicapai dalam pemrosesan data terstruktur sekuensial yang kompleks. Tidak seperti teknik lain dalam deep learning, RNN memiliki memori untuk mengingat perhitungan sebelumnya dan juga memiliki mekanisme perulangan internal yang memberikan kemampuan untuk belajar dari keadaan sebelumnya [14-17].

2.5. BERT

BERT atau Bidirectional Encoder Representations dari Transformer, meningkatkan transformer standar dengan menghilangkan kendala searah dengan menggunakan pra-pelatihan model Masked Language (MLM). Model masked language secara acak menutupi beberapa token dari input dan tujuannya adalah untuk memprediksi id kosakata asli dari masked word hanya berdasarkan konteksnya. Tidak seperti pra-pelatihan model left-to-right language, tujuan MLM memungkinkan representasi untuk memadukan konteks kiri dan kanan, yang memungkinkan kita untuk melatih transformer dua arah yang dalam. Selain model masked language, BERT menggunakan tugas prediksi kalimat berikutnya yang secara bersama-sama melatih representasi pasangan teks, terdapat dua langkah dalam BERT yaitu pra-pelatihan dan fine-tuning. Selama pra-pelatihan, model dilatih pada data yang tidak berlabel selama tugas pra-pelatihan yang berbeda. Untuk fine-tuning, model BERT pertama-tama diinisialisasi dengan parameter yang telah dilatih sebelumnya, dan semua parameter disesuaikan dengan menggunakan data berlabel dari tugas-tugas downstream, setiap tugas downstream memiliki model fine-tuned terpisah, meskipun mereka diinisialisasi dengan parameter pra-pelatihan yang sama [18].

2.6. Confusion Matrix

Confusion Matrix merupakan metode yang digunakan untuk mengukur kinerja suatu metode klasifikasi. Confusion matrix juga dapat memperoleh nilai accuracy, recall dan

precision. Nilai accuracy menggambarkan seberapa akurat model dengan mengklasifikasikan dengan benar. Nilai accuracy dapat dihitung menggunakan persamaan (1)

$$Accuracy = \frac{TN+TP}{TN+FP+FN+TP} \quad (1)$$

Nilai precision menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model. Nilai precision dapat dihitung menggunakan persamaan (2)

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Nilai recall yang menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. Nilai recall dapat dihitung menggunakan persamaan (3).

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Dalam confusion matrix terdapat beberapa istilah yang dapat dilihat pada Tabel 1 seperti TP (True Positive) merupakan jumlah data positive yang benar diprediksi, TN (True Negative) merupakan jumlah data negative yang benar diprediksi, FP (False Positive) merupakan jumlah data positive yang salah diprediksi dan FN (False Negative) merupakan jumlah data negative yang salah diprediksi [19-20]

Tabel 1. Tabel Confusion Matrix

Confusion Matrix		Prediction Value	
		Positive	Negative
Actual Values	Positive	TP	FP
	Negative	FN	TN

3. METODE PENELITIAN

Metodologi yang digunakan dalam pengembangan sistem ini mengacu pada metodologi waterfall. Waterfall [5] merupakan metodologi terstruktur yang pada pengembangan perangkat lunak dilakukan secara sistematis. Pada prinsipnya waterfall dijalankan secara konsisten berdasarkan perencanaan ataupun penjadwalan terhadap semua aktivitas yang telah ditentukan. Adapun rincian tahapannya antara lain:

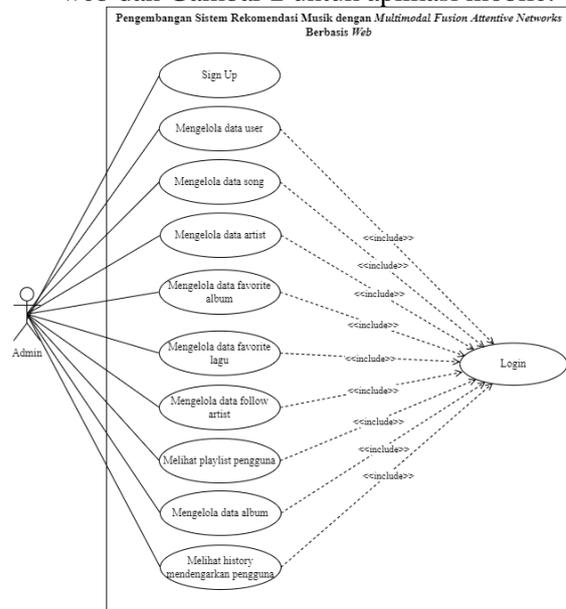
1. Perencanaan

Pada tahapan ini, dilakukan pengumpulan dataset musik yang berasal dari dataset

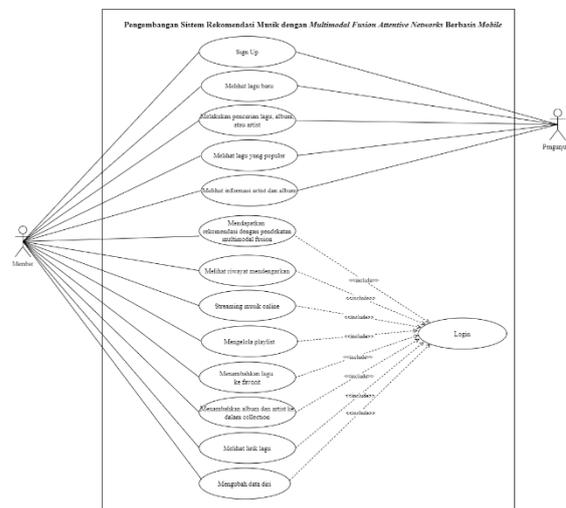
Spotify Genius Track Dataset yang diperoleh dari Kaggle.

2. Analisis Kebutuhan.

Pada tahap ini, analisis kebutuhan fungsional akan dirancang dengan menggunakan use case diagram yang ditunjukkan pada Gambar 1 untuk aplikasi web dan Gambar 2 untuk aplikasi mobile.



Gambar 1. Use Case Diagram Aplikasi Web

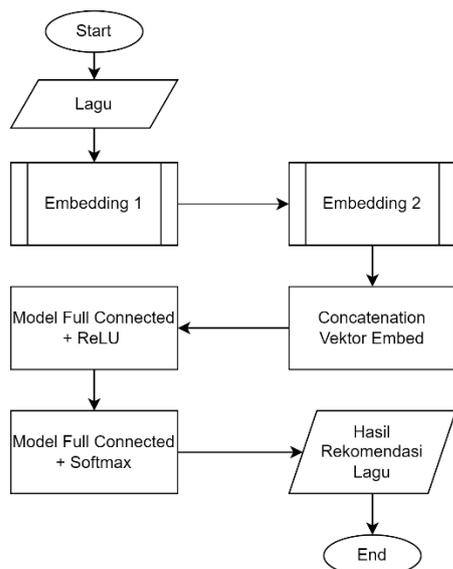


Gambar 2. Use Case Diagram Aplikasi Mobile

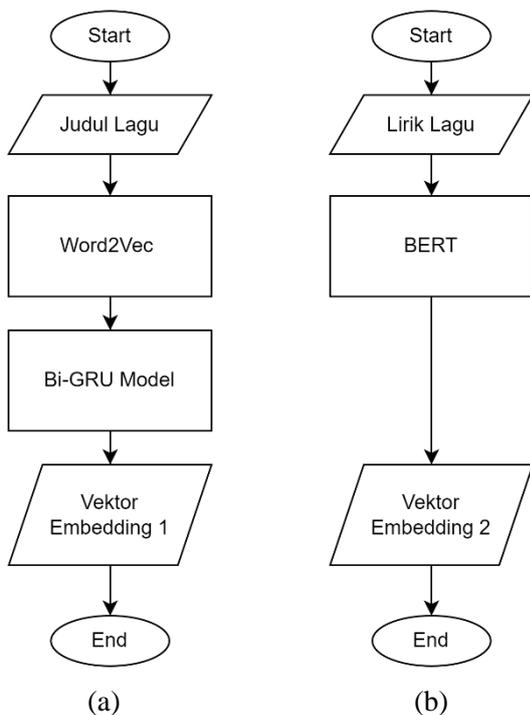
3. Analisis Proses.

Pada tahap ini, proses kerja aplikasi dalam memberikan sistem rekomendasi dengan mengoptimalkan pendekatan Multimodal Fusion untuk pemanfaatan berbagai representasi lagu dengan menggunakan Attentive Neural Networks dan Recurrent Neural Networks dalam mengumpulkan

informasi *sequential*. Analisis proses akan dimodelkan dengan menggunakan diagram *flowchart*. Gambar 3 menunjukkan *flowchart* keseluruhan proses sistem rekomendasi musik. Gambar 4(a) menunjukkan proses Embedding 1 untuk modal judul lagu dan Gambar 4(b) menunjukkan proses Embedding 2 untuk modal lirik lagu.



Gambar 3. Flowchart Rekomendasi Multimodal Fusion Attentive Network



Gambar 4. (a) Flowchart Proses Embedding 1 (b) Flowchart Proses Embedding 2

4. Perancangan Sistem.
Pada tahap ini, rancangan mock-up prototipe aplikasi akan dibuat dengan menggunakan figma.
5. Implementasi.
Pada tahap ini, rancangan akan diimplementasi dengan menggunakan bahasa pemrograman PHP dengan framework Laravel untuk aplikasi web dan kotlin untuk aplikasi mobile, serta menggunakan database MySQL.
6. Pengujian Sistem
Pada tahap ini, pengujian akan dilakukan untuk aplikasi web dan mobile dengan pendekatan black-box testing dan pengujian algoritma rekomendasi dengan *confusion matrix*.

4. HASIL DAN PEMBAHASAN

4.1. Hasil

Hasil dari pengembangan sistem rekomendasi berupa aplikasi berbasis web dan mobile. Aplikasi berbasis web digunakan oleh admin untuk mengelola data master dan user. Tampilan awal dari aplikasi adalah halaman login yang ditunjukkan pada Gambar 5



Gambar 5. Halaman Login pada Web

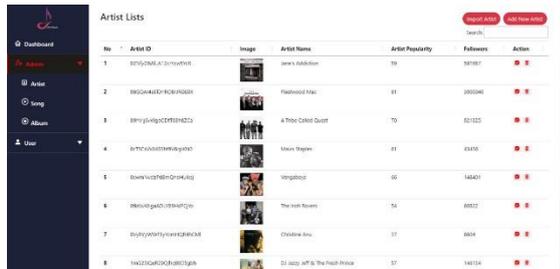
Setelah login, website akan menampilkan halaman dashboard dari admin yang ditunjukkan pada Gambar 6.



Gambar 6. Halaman Dashboard Admin pada Web

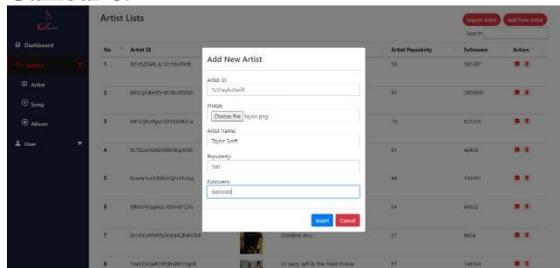
Setelah itu, admin dapat mengelola data administrator dan data user sesuai dengan menu

yang ada di sebelah kiri. Pada halaman awal dashboard juga tersedia rekapitulasi data seperti jumlah judul lagu, album, artis dan pengguna. Gambar 7 menunjukkan contoh halaman pengelolaan daftar artis. Pada halaman ini, admin dapat melakukan aksi penambahan, edit, atau penghapusan pada data artis.



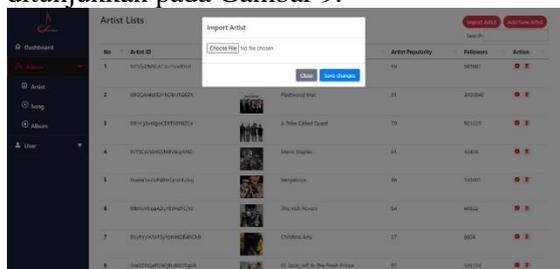
Gambar 7. Halaman Daftar Artis untuk Pengelolaan Artis pada Web

Jika admin ingin menambahkan data baru, maka menu Add New Artis dapat digunakan untuk menampilkan halaman dialog untuk penambahan data baru yang ditunjukkan pada Gambar 8.



Gambar 8. Halaman Dialog Penambahan Artis Baru untuk Pengelolaan Artis pada Web

Selain memasukkan data satu per satu, aplikasi web juga menyediakan fasilitas untuk melakukan import data sekaligus yang ditunjukkan pada Gambar 9.



Gambar 9. Halaman Import Daftar Artis dari File untuk Pengelolaan Artis pada Web

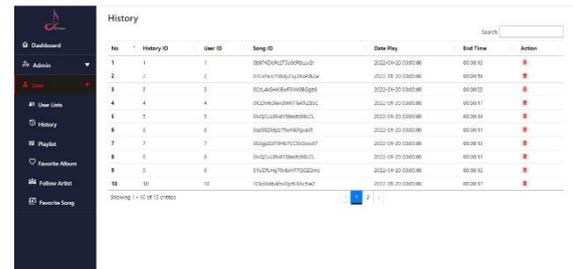
Seluruh pengelolaan data artis untuk penambahan, edit dan penghapusan data juga diterapkan untuk data master lainnya seperti

data lagu dan data album. Selain penambahan data master, admin juga dapat mengakses informasi terkait pengguna, seperti daftar pengguna, history pemutaran musik pengguna, playlist lagu pengguna, favorit album, favorit lagu dan follow artis. Gambar 10 menunjukkan halaman web untuk daftar pengguna.



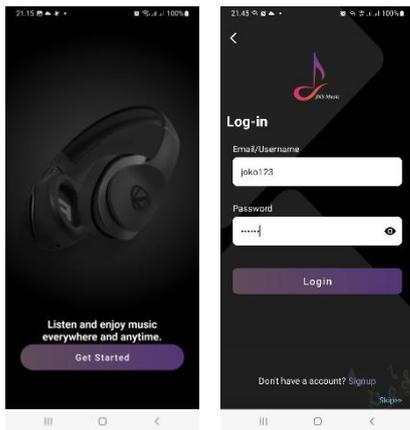
Gambar 10. Halaman Daftar Pengguna untuk Pengelolaan User pada Web

Gambar 11 menunjukkan contoh history dari pemutaran musik seluruh pengguna aplikasi mobile.



Gambar 11. Halaman History Pemutaran Musik oleh Pengguna pada Web

Aplikasi web digunakan oleh admin dan tidak dapat diakses oleh pengguna biasa. Pengguna akhir dapat menggunakan aplikasi berbasis mobile dan menggunakan fitur sesuai dengan fungsionalitas pada use case diagram pada Gambar 2. Aplikasi berbasis mobile digunakan oleh pengguna akhir untuk mendengarkan musik dan mendapatkan rekomendasi. Gambar 12 menunjukkan tampilan awal dari aplikasi mobile.

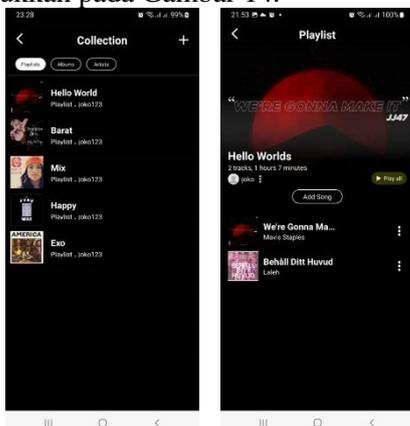


Gambar 12. Tampilan Login Aplikasi *Mobile* Setelah pengguna login, pengguna dapat mengakses halaman awal yang berisi lagu yang terakhir dimainkan, lagu baru, follower terbanyak dari artis, dan lagu yang paling banyak difavoritkan. Pengguna juga dapat memilih menu yang dapat dipilih pada sebelah kiri atas, yang ditunjukkan pada Gambar 13.



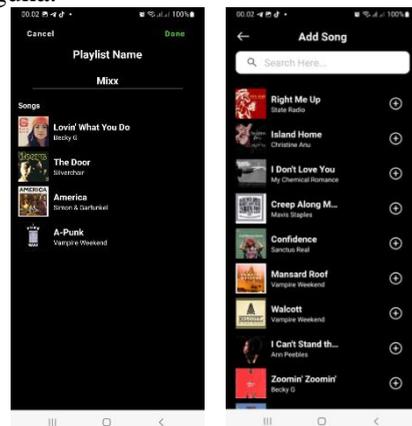
Gambar 13. Tampilan Awal Pengguna yang Terdaftar pada Aplikasi *Mobile*

Pengguna dapat mengakses collection yang ada dan melihat playlist yang telah dibuat, yang ditunjukkan pada Gambar 14.



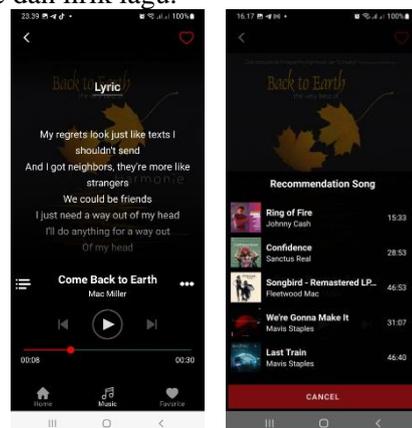
Gambar 14. Tampilan Collection dan Playlist Pengguna yang Terdaftar pada Aplikasi *Mobile*

Pengguna dapat menambahkan playlist baru dengan mengakses tampilan collection dan menekan tombol '+' di sebelah kanan atas. Nama Playlist dapat dimasukkan pengguna dan hasilnya, playlist dapat diakses untuk ditambahkan lagu yang baru. Gambar 15 menunjukkan tampilan input nama playlist baru dan penambahan lagu baru pada playlist pengguna.



Gambar 15. Tampilan Playlist Baru dan Penambahan Lagu pada Playlist Pengguna yang Terdaftar pada Aplikasi *Mobile*

Aplikasi mobile juga menyediakan fasilitas untuk menampilkan lirik lagu yang sedang dimainkan. Lirik akan otomatis muncul setelah pengguna memutar lagu yang diinginkan. Gambar 16 menunjukkan tampilan pemutaran music dan lirik lagu.



Gambar 16. Tampilan Pemutaran Musik dan Lirik Lagu Musik pada Aplikasi *Mobile*

4.2. Pembahasan

Setelah aplikasi web dan mobile dikembangkan, aplikasi tersebut perlu untuk dievaluasi. Evaluasi untuk fungsionalitas dari aplikasi menggunakan pendekatan *black-box testing* untuk memastikan fitur dapat digunakan dan menghasilkan luaran yang benar. Pada pendekatan *black-box testing*, *test scenario* akan dibuat untuk semua fitur yang disediakan aplikasi dan memastikan *actual result* dari aplikasi sesuai dengan *expected result* yang ditetapkan. Salah satu *test scenario* yang telah dibuat untuk menguji penambahan *playlist* baru pada aplikasi *mobile*:

1. Test scenario name: menambah playlist baru
2. Test scenario description: verifikasi fungsi add new playlist
3. Test case description: menambahkan playlist baru dengan mengklik tombol +.
4. Test step:
 - a. Pengguna login ke aplikasi
 - b. Pengguna membuka halaman collection
 - c. Pengguna memilih tombol + disebelah kanan atas
 - d. Pengguna mengisi nama playlist baru
 - e. Pengguna menyimpan dengan klik tombol 'Done'
5. Expected result: aplikasi akan menampilkan playlist baru pada halaman collection
6. Actual result: aplikasi telah menampilkan playlist baru pada halaman collection
7. Test Status: passed

Evaluasi untuk algoritma rekomendasi akan menggunakan *confusion matrix* dengan menghitung nilai akurasi, presisi dan *recall* dari hasil rekomendasi musik yang diterima oleh pengguna. Pengujian akan dilakukan dengan membagi dataset pelatihan dengan perbandingan 70:30. Hasil pengujian dari sistem rekomendasi menghasilkan nilai akurasi sebesar 84.29%, nilai presisi sebesar 76.76%, dan nilai *recall* sebesar 83.1%.

KESIMPULAN

Berdasarkan pengujian yang dilakukan terhadap sistem rekomendasi musik menggunakan pendekatan *multimodal fusion* berbasis *mobile* dan *web*, maka dapat ditarik kesimpulan sebagai berikut:

- a. Berdasarkan pengujian yang dilakukan dengan *confusion matrix*, akurasi dari model *multimodal fusion* yang diterapkan dalam penelitian ini adalah sebesar 84% sehingga dapat disimpulkan bahwa model yang dibangun telah berhasil untuk merekomendasikan lagu yang memiliki preferensi yang sesuai dengan lagu yang sedang didengarkan oleh pengguna.
- b. Rekomendasi musik dengan pendekatan *content-based* dengan mencari kesamaan lagu berdasarkan judul dan lirik lagu yang digunakan dalam penelitian ini berhasil mengatasi masalah *cold-start*.
- c. Aplikasi *web* dan *mobile* berjalan dengan baik tanpa ada masalah dan melewati *black-box testing* dalam pengujian.
- d. Penambahan lagu baru akan memperbaharui hasil rekomendasi lagu untuk pengguna sehingga lagu baru dapat muncul dan dipilih oleh pengguna.

DAFTAR PUSTAKA

- [1] S. Y. M. Netti and I. Irwansyah, "Spotify: Aplikasi Music Streaming untuk Generasi Milenial," J. Komun., vol. 10, no. 1, p. 1, 2018, doi: 10.24912/jk.v10i1.1102.
- [2] K. Vaswani, Y. Agrawal, and V. Alluri, "Multimodal Fusion Based Attentive Networks for Sequential Music Recommendation," 2021, [Online]. Available: <http://arxiv.org/abs/2110.01001>.
- [3] H. Cherifi, B. Gonçalves, R. Menezes, and R. Sinatra, "Preface," Stud. Comput. Intell., vol. 644, pp. v–vi, 2016, doi: 10.1007/978-3-319-30569-1.
- [4] R. Supradewi, "Otak , Musik, Dan Proses Belajar," Bul. Psikol., vol. 18, no. 2, pp. 58–68, 2010.
- [5] "KAMUS BAHASA INDONESIA."
- [6] F. Ricci, L. Rokach, and B. Shapira, Recommender Systems Handbook. 2011.
- [7] H. Zhu, Z. Wang, Y. Shi, Y. Hua, G. Xu, and L. Deng, "Multimodal Fusion Method Based on Self-Attention Mechanism," Wirel. Commun. Mob. Comput., vol. 2020, 2020, doi: 10.1155/2020/8843186.
- [8] C. Hori et al., "Attention-Based Multimodal Fusion for Video Description," Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-Octob, pp. 4203–4212, 2017, doi: 10.1109/ICCV.2017.450.
- [9] R. G. Prasetya, "Multimodal Fusion: A Theory And Applications," pp. 6–18, 2017.

- [10] G. K. Verma and U. S. Tiwary, "Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals," *Neuroimage*, vol. 102, no. P1, pp. 162–172, 2014, doi: 10.1016/j.neuroimage.2013.11.007.
- [11] L. Sharma and A. Gera, "A Survey of Recommendation System: Research Challenges Related papers Survey on Recommendation System Using Semantic Web Mining A Survey of Recommendation System: Research Challenges," *Int. J. Eng. Trends Technol.*, 2013, [Online]. Available: <http://www.ijettjournal.org>.
- [12] N. Sachdeva, K. Gupta, and V. Pudi, "Attentive neural architecture incorporating song features for music recommendation," in *RecSys 2018 - 12th ACM Conference on Recommender Systems*, Sep. 2018, pp. 417–421, doi: 10.1145/3240323.3240397.
- [13] A. Yanuar, "Recurrent Neural Network (RNN)," 1 July 2018, 2018. <https://machinelearning.mipa.ugm.ac.id/2018/07/01/recurrent-neural-network-rnn/> (accessed Nov. 07, 2021).
- [14] N. Nelaturi and G. L. Devi, "A product recommendation model based on recurrent neural network," *J. Eur. des Syst. Autom.*, vol. 52, no. 5, pp. 501–507, 2019, doi: 10.18280/jesa.520509.
- [15] F. Rahutomo, D. S. E. Ikawati, and O. A. Rohman, "Evaluasi Fitur Word2Vec Pada Sistem Ujian Esai Online," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 4, no. 1, pp. 36–45, 2019.
- [16] D. Qiu, H. Jiang, and S. Chen, "Fuzzy information retrieval based on continuous bag-of-words model," *Symmetry (Basel)*, vol. 12, no. 2, 2020, doi: 10.3390/sym12020225.
- [17] M. K. Dahouda and I. Joe, "A Deep-Learned Embedding Technique for Categorical Features Encoding," *IEEE Access*, vol. 9, no. August, pp. 114381–114391, 2021, doi: 10.1109/ACCESS.2021.3104357.
- [18] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, no. Mlm, pp. 4171–4186, 2019.
- [19] M. D. Nguyen and Y. S. Cho, "A Variational Autoencoder Mixture Model for Online Behavior Recommendation," *IEEE Access*, vol. 8, pp. 132736–132747, 2020, doi: 10.1109/ACCESS.2020.3010508.
- [20] A. Kulkarni, D. Chong, and F. A. Batarseh, *Foundations of data imbalance and solutions for a data democracy*. Elsevier Inc., 2020.