

KLASIFIKASI *GENRE* MUSIK MENGGUNAKAN TEKNIK PEMBELAJARAN MESIN

Gst. Ayu Vida Mastrika Giri¹⁾, Made Leo Radhitya²⁾

Program Studi Informatika¹⁾, Program Studi Teknik Informatika²⁾

Universitas Udayana, Badung, Bali¹⁾, Institut Bisnis dan Teknologi Indonesia, Denpasar, Bali²⁾

vida@unud.ac.id¹⁾, leo.radhitya@instiki.ac.id²⁾

ABSTRACT

Music classification plays an important role in music information retrieval because the results of music classification based on genre can be used by music listeners to find new music according to their preferences. Currently, online music streaming providers create playlists and also provide music recommendations based on genres. Therefore, developing a good and accurate machine learning model to classify music automatically based on genre will be very beneficial for online music streaming providers. In this research, automatic classification of music genres was carried out using content-based features that have been extracted from audio signals. The dataset used contains musical features that have been processed from the GTZAN music dataset which consists of 57 content-based features. The machine learning techniques used for music classification were K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), and Multi Layer Perceptron (MLP). The MLP machine learning technique produced the greatest accuracy value when compared to K-NN and SVM in this study with an accuracy value of 71.6%, followed by SVM with an accuracy value of 71%, and K-NN with an accuracy value of 69.6%.

Keywords: *genre, k-nearest neighbor, multi layer perceptron, music classification, support vector machine*

ABSTRAK

Klasifikasi musik memainkan peran penting dalam temu kembali informasi musik karena hasil klasifikasi musik berdasarkan *genre* dapat digunakan oleh pendengar musik untuk menemukan musik baru yang sesuai dengan preferensi. Saat ini, penyedia *streaming* musik *online* membuat *playlist* dan juga memberikan rekomendasi musik berdasarkan *genre*. Oleh karena itu, pengembangan model pembelajaran mesin yang baik dan akurat untuk mengklasifikasikan musik secara otomatis berdasarkan *genre* akan sangat bermanfaat bagi penyedia *streaming* musik *online*. Pada penelitian ini, dilakukan klasifikasi *genre* musik otomatis menggunakan fitur berbasis konten yang telah diekstrak dari sinyal audio. Dataset yang digunakan berupa fitur yang telah diproses dari dataset musik GTZAN yang terdiri dari 57 fitur berbasis konten. Teknik pembelajaran mesin yang digunakan untuk klasifikasi musik adalah K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), dan Multi Layer Perceptron (MLP). Teknik pembelajaran mesin MLP menghasilkan nilai akurasi terbesar jika dibandingkan dengan K-NN dan SVM pada penelitian ini dengan nilai akurasi 71,6%, disusul dengan SVM dengan nilai akurasi 71%, dan K-NN dengan nilai akurasi sebesar 69,6%.

Kata Kunci: *genre, k-nearest neighbor, klasifikasi musik, multi layer perceptron, support vector machine*

PENDAHULUAN

Genre musik adalah kumpulan kata deskriptif yang dapat merepresentasikan informasi tentang musik [1], sehingga dapat digunakan untuk membedakan atau mengelompokkan musik. Musik yang berada pada *genre* yang sama memiliki karakteristik umum seperti instrumentasi, konten harmonik, dan struktur ritmis yang sama [2]. Klasifikasi musik memainkan peran penting dalam temu kembali informasi musik. Hasil klasifikasi musik berdasarkan *genre* dapat digunakan oleh pendengar musik untuk menemukan musik yang sesuai dengan preferensinya.

Penyedia *streaming* musik *online* seperti Spotify dan Apple Music menggunakan *genre* musik untuk membuat *playlist* dan memberikan rekomendasi pada pendengar musik berdasarkan *genre* favorit mereka [3]. Oleh karena itu, pengembangan model pembelajaran mesin yang baik dan akurat untuk mengklasifikasikan musik secara otomatis berdasarkan *genre* akan sangat berharga bagi penyedia *streaming* musik *online*.

Penggunaan teknik pembelajaran mesin telah digunakan oleh para peneliti untuk mengklasifikasikan musik berdasarkan *genre*. Tiga teknik pembelajaran mesin yaitu Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), dan Naïve Bayes (NB) digunakan untuk mengklasifikasi data musik berdasarkan *metadata* musik yang didapatkan dari Spotify telah dilakukan oleh Setiadi, dkk. [4]. Berdasarkan hasil pengujian yang dilakukan pada penelitian tersebut, SVM memiliki kinerja klasifikasi terbaik dengan akurasi 80%, kemudian diikuti oleh K-NN dengan 77,18% dan NB dengan 76,08%. Penelitian yang dilakukan oleh Lau & Ajoodha [2] membandingkan penggunaan *deep learning* yaitu Convolutional Neural Network (CNN) dan beberapa metode pembelajaran mesin tradisional. Dataset yang digunakan adalah GTZAN yang terdiri dari 1000 musik dan 10 *genre*. Hasil klasifikasi tertinggi pada

penelitian tersebut adalah dengan metode pembelajaran mesin tradisional K-NN dengan rata-rata akurasi sekitar 92%. Dataset GTZAN, CNN, dan teknik pembelajaran mesin tradisional juga digunakan dalam penelitian yang dilakukan oleh Sahil, dkk. [5]. Pada penelitian tersebut, hasil klasifikasi tertinggi didapat dengan menggunakan teknik pembelajaran mesin SVM dengan akurasi 82%, disusul dengan teknik pembelajaran mesin jaringan syaraf tiruan dengan akurasi 80%.

Pada penelitian ini digunakan tiga teknik pembelajaran mesin yang berpotensi berguna dalam mengklasifikasikan *genre* musik berdasarkan nilai akurasi tertinggi yang ada pada penelitian-penelitian sebelumnya, yaitu K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), dan Multi Layer Perceptron (MLP) yang termasuk dalam jaringan syaraf tiruan. Perbedaan penelitian ini dengan penelitian-penelitian sebelumnya [2] [5] adalah perbedaan parameter yang digunakan untuk klasifikasi dengan K-NN dan MLP, serta perbedaan kernel untuk SVM.

TINJAUAN PUSTAKA

Genre Musik

Genre musik adalah label yang dibuat dan digunakan oleh manusia untuk mengkategorikan dan mendeskripsikan musik. *Genre* musik tidak memiliki definisi dan batasan yang ketat, karena muncul melalui interaksi yang kompleks antara faktor publik, pemasaran, sejarah, dan budaya. Anggota *genre* tertentu memiliki karakteristik yang sama dengan anggota lainnya yang terkait dengan instrumentasi, struktur ritmis, dan konten nada musik [6].

Fitur Musik Berdasarkan Konten

Fitur musik berdasarkan konten diekstrak dari sinyal audio musik. Fitur ini dapat direpresentasikan dengan nilai mean, standar deviasi, *variance*, histogram, agregasi MFCC, dan *area moments*. Secara umum, fitur musik berdasarkan konten dapat dibagi menjadi dua

jenis, yaitu fitur domain waktu dan fitur domain frekuensi [2].

Fitur domain waktu diekstrak secara langsung dari sinyal audio musik. Tiga fitur yang termasuk dalam fitur domain waktu adalah [5]

- *Root mean square energy*: nilai rata-rata akar kuadrat energi musik dalam sinyal waktu diskrit
- *Zero crossing rate*: frekuensi dimana sinyal waktu diskrit berubah tanda dari positif ke negatif atau sebaliknya.
- *Tempo* yang mengacu pada kecepatan sebuah musik, dinyatakan dalam ketukan per menit.

Fitur domain frekuensi didapat dengan mengaplikasikan *fourier transform* pada sinyal audio musik. Fitur yang termasuk pada domain frekuensi adalah [2]

- *Mel-Frequency Cepstral Coefficient (MFCC)*: serangkaian spektrum daya jangka pendek dalam *file* audio. MFCC memodelkan karakteristik suara manusia.
- *Chroma*: berhubungan dengan total energi dari sinyal dari setiap 12 *pitch semitone*.
- *Spectral centroid*: frekuensi dimana energi terpusat pada sebuah *frame*.
- *Spectral bandwidth*: lebar pita gelombang atau *range* dari frekuensi dari sebuah gelombang bunyi.
- *Spectral Roll-off*: frekuensi dimana sebuah persentase tertentu dari total energi spektral berada.

Normalisasi Min-Max

Perbedaan rentang data pada setiap fitur dapat menyebabkan data tidak terdistribusi dengan baik dan mempengaruhi hasil klasifikasi. Normalisasi data dilakukan untuk menyamakan semua rentang data fitur yang didapat. Persamaan 1 digunakan untuk menormalisasi data [7].

$$\text{norm}(x) = \frac{x - \text{minValue}}{\text{maxValue} - \text{min value}} \quad (1)$$

Dimana $\text{norm}(x)$ adalah nilai dari sebuah data dalam sebuah fitur yang telah dinormalisasi, x

adalah data dalam sebuah fitur yang belum dinormalisasi, minValue adalah nilai terendah dari fitur tertentu, dan maxValue adalah nilai tertinggi dari fitur tertentu.

K-Nearest Neighbor (K-NN)

Sebuah teknik klasifikasi K-NN merepresentasikan setiap data sebagai titik data dalam ruang d -dimensi, dimana d adalah jumlah atribut. Jika diberikan sebuah data pengujian, nilai kedekatan/jarak dari data pengujian tersebut dihitung dengan setiap data pelatihan dengan pengukuran jarak tertentu. K tetangga terdekat dari data pengujian diambil dari data pelatihan dengan jarak terkecil [8].

Pemilihan nilai K menjadi penting dalam klasifikasi dengan K-NN. Jika nilai K terlalu kecil, maka hasil klasifikasi rentan *overfitting*. Jika nilai K terlalu besar, pengklasifikasian tetangga terdekat bisa salah karena menyertakan data pelatihan yang sebenarnya terletak jauh dari data pengujian.

Support Vector Machine (SVM)

Teknik pembelajaran mesin SVM memakai subset dari data pelatihan yang paling sulit untuk diklasifikasi, yang disebut dengan *support vector*, untuk mengklasifikasi data. Sebuah *hyperplane* untuk memisahkan data dibuat dengan menggunakan *support vector* sebagai acuan.

Dengan menggunakan fungsi kernel, teknik pembelajaran mesin SVM bisa langsung bekerja dengan merepresentasikan *hyperplane* dalam klasifikasi non linear tanpa mengalami *curse of dimensionality*. Ini telah menjadi alasan penggunaan SVM dalam permasalahan klasifikasi non linear [8].

Multi Layer Perceptron (MLP)

Sebuah jaringan syaraf multi layer atau bisa juga disebut MLP menggeneralisasi konsep dasar perceptron ke arsitektur node yang lebih kompleks yang mampu mempelajari batas keputusan nonlinier. Arsitektur umum dari MLP adalah terdiri dari beberapa *layer/lapisan* yaitu *input layer*, *hidden layer*, dan *output*

layer. Lapisan-lapisan tersebut saling berhubungan sehingga setiap lapisan beroperasi berdasarkan keluaran dari lapisan sebelumnya. Keluaran akhir dihasilkan pada lapisan terakhir/*output layer*, yang digunakan untuk membuat prediksi [8].

Learning rate dalam ANN secara umum dan MLP adalah *hyperparameter* yang mengontrol seberapa banyak perubahan model sebagai respons terhadap perkiraan kesalahan setiap kali bobot model diperbarui. Memilih *learning rate* sangat penting karena nilai yang terlalu kecil dapat mengakibatkan proses pelatihan yang lama, sedangkan nilai yang terlalu besar dapat menghasilkan bobot yang kurang optimal atau proses pelatihan yang tidak stabil.

K-Fold Cross validation

Metode *K-Fold Cross validation* adalah sebuah metode statistik yang dapat digunakan untuk memilih model dan memprediksi kesalahan uji model prediktif dengan lebih baik.

Dalam validasi silang, dataset dibagi dengan sejumlah *K-Folds*. Dimana pada setiap iterasi setiap *fold* digunakan satu kali sebagai data uji dan *fold* yang tersisa digunakan sebagai data training, proses tersebut diulang sampai semua data dievaluasi [9].

Confusion Matrix

Performa model pengklasifikasi dapat dievaluasi dengan membandingkan label yang diprediksi dengan label data yang sebenarnya. Informasi ini dapat diringkas dalam sebuah tabel yang disebut *confusion matrix*.

Confusion matrix merangkum jumlah data yang diprediksi dengan benar atau salah oleh pengklasifikasi menggunakan empat hitungan berikut [8]:

- *True positive* (TP) adalah jumlah data positif yang diprediksi dengan benar sebagai positif oleh classifier.
- *False positive* (FP) adalah jumlah data negatif yang diprediksi salah sebagai positif oleh classifier.

- *False negative* (FN) adalah jumlah data positif yang diprediksi salah sebagai negatif oleh classifier.
- *True negative* (TN) adalah jumlah data negatif yang diprediksi dengan benar sebagai negatif oleh classifier.

Akurasi

Meskipun *confusion matrix* telah memberikan informasi yang diperlukan untuk menentukan seberapa baik performa model klasifikasi, meringkas informasi ini menjadi satu angka berupa nilai akurasi membuatnya lebih mudah untuk membandingkan performa relatif dari model yang berbeda. Nilai akurasi menghitung secara keseluruhan perbandingan dari jumlah data yang diprediksi kelasnya dengan benar dengan jumlah total prediksi. Persamaan 2 digunakan untuk menghitung nilai akurasi [8].

$$Accuracy = \frac{\text{Number of correct prediction}}{\text{Total number of prediction}} \quad (2)$$

Precision, Recall, dan F1-Score

Precision adalah ukuran seberapa banyak prediksi positif yang benar (*true positive*). Nilai *precision* adalah perbandingan antara *true positive* dengan banyaknya data yang diprediksi positif. Persamaan 3 [8] dapat digunakan untuk menghitung nilai *precision*, dimana TP adalah *true positive*, dan FP adalah *false positive*.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Nilai *recall* adalah ukuran berapa banyak kasus positif yang diprediksi oleh pengklasifikasi dengan benar, atas semua kasus positif dalam data. Nilai *recall* dapat dihitung dengan membandingkan nilai *true positive* dengan banyaknya data yang sebenarnya positif. Persamaan 4 [8] dapat digunakan untuk menghitung nilai *recall*, dimana TP adalah *true positive*, dan FN adalah *false negative*.

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F1-Score atau *F1-Measure* adalah ukuran yang menggabungkan *precision* dan *recall*. *F1-Score* dapat menggambarkan rata-rata harmonik dari *precision* dan *recall*. Nilai *F1-Score* yang tinggi

merepresentasikan nilai *precision* dan *recall* yang tinggi juga. Persamaan 5 digunakan untuk menghitung F1 Score [8]

$$F1\ Score = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (5)$$

METODOLOGI PENELITIAN

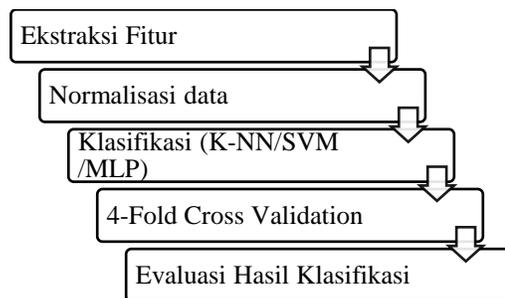
Dataset

Dataset yang digunakan pada penelitian ini adalah dataset GTZAN [6] yang terdiri dari 1000 musik dengan durasi 30 detik yang telah diklasifikasi ke dalam 10 *genre* dengan 100 musik pada setiap *genre*. Sepuluh *genre* yang ada pada dataset GTZAN adalah *blues*, *classical*, *country*, *disco*, *hip-hop*, *jazz*, *metal*, *pop*, *reggae*, dan *rock*.

Dataset GTZAN terdiri dari 57 fitur, yaitu *chroma STFT*, *RMSE*, *spectral-centroid*, *spectral-bandwidth*, *spectral-rolloff*, MFCC 1-20, *tempo*, *energy*, *zero crossing rate*, dan *harmony*. Setiap fitur, kecuali *tempo*, terdiri dari dua dimensi yaitu nilai mean dan standar deviasi dari rekaman musik selama 30 detik.

Gambaran Umum Penelitian

Penelitian ini melakukan klasifikasi *genre* musik otomatis menggunakan fitur berbasis konten yang telah diekstrak dari sinyal audio. Dataset yang digunakan berupa fitur yang telah diproses dari dataset musik GTZAN yang terdiri dari 57 fitur berbasis konten. Alur pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Data fitur yang didapat tidak pada *range* data yang sama, sehingga diperlukan normalisasi min-max untuk menjadikan semua fitur dalam *range* 0-1. Dataset yang telah dinormalisasi

dibagi menjadi 2 bagian, yaitu data training dan data testing. Pada penelitian ini, 75% data digunakan menjadi data training dan 25% data menjadi data testing.

Klasifikasi menggunakan K-Nearest Neighbor (K-NN) pada penelitian ini menggunakan library dari Scikit Learn [10] dengan menggunakan jumlah tetangga (*neighbor*) ganjil dari 3 hingga 50. Pengukuran jarak yang digunakan adalah *euclidean distance*. Nilai akurasi klasifikasi dengan jumlah tetangga yang memiliki *cross validation* tertinggi akan dibandingkan dengan hasil klasifikasi menggunakan metode SVM dan MLP.

Klasifikasi menggunakan SVM pada penelitian ini menggunakan library dari Scikit Learn [10] dengan kernel berbeda pada tiap percobaan. Kernel yang digunakan yaitu *linear*, *polynomial*, dan *RBF*. Nilai akurasi klasifikasi SVM dengan kernel yang memiliki nilai *cross validation* tertinggi akan dibandingkan dengan hasil klasifikasi metode K-NN dan MLP.

Klasifikasi menggunakan Multi Layer Perceptron (MLP) pada penelitian ini menggunakan library dari Scikit Learn [10] dengan menggunakan *learning rate* 0,3, 0,5, 0,7, dan 0,9 dengan pengurangan *learning rate* 0,1, 0,3, 0,5, 0,7, dan 0,9. Dalam setiap variasi *learning rate* dan pengurangan *learning rate*, akan digunakan satu buah *hidden layer* dengan jumlah neuron pada *hidden layer* sejumlah angka genap yang ada pada *range* 20 sampai 50. Nilai akurasi klasifikasi dari kombinasi *learning rate* dan pengurangan *learning rate* dengan nilai *cross validation* tertinggi akan dibandingkan dengan hasil teknik pembelajaran mesin K-NN dan SVM.

Model pembelajaran mesin yang digunakan akan dievaluasi dengan menggunakan nilai mean dari *K-Fold Cross validation* dengan jumlah $K=4$. Evaluasi hasil klasifikasi pada penelitian ini menggunakan nilai *precision*, *recall*, *F1-Score*, dan nilai akurasi. Hasil

klasifikasi musik berdasarkan *genre* tertentu akan ditampilkan dalam *confusion matrix*.

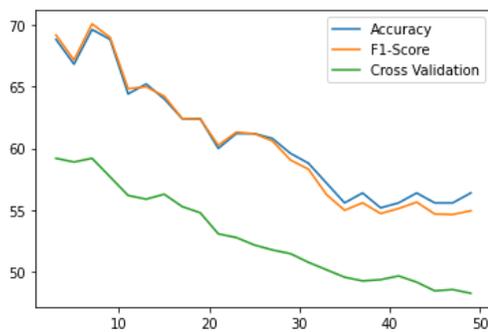
HASIL DAN PEMBAHASAN

Pada klasifikasi dengan teknik pembelajaran mesin K-NN, nilai K diambil dari angka ganjil yang berada di antara 3 dan 50. Tabel 1 menunjukkan lima data dengan nilai tertinggi yang terdiri dari nilai K yang digunakan, nilai *cross validation*, dan akurasi. Nilai *cross validation* tertinggi ada pada nilai K 7 dan 3, yaitu 59.2 dan nilai akurasi tertinggi ada pada nilai K=7, yaitu 68,8%.

Tabel 1. Hasil Teknik pembelajaran mesin K-NN

Nilai K	Cross validation	Accuracy
7	59,2	69,6
3	59,2	68,8
5	58,9	66,8
9	57,7	68,8
15	56,3	64,0

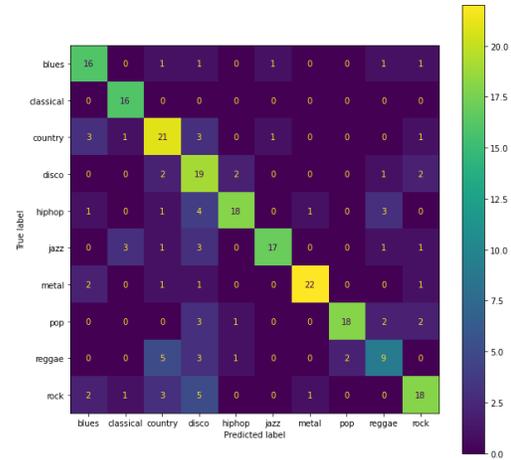
Hasil teknik pembelajaran mesin K-NN dengan nilai K lainnya tidak menghasilkan nilai *cross validation* dan akurasi yang lebih baik dari yang tercantum pada Tabel 1. Setelah nilai K=15 nilai akurasi dan *cross validation* cenderung menurun seperti yang ditunjukkan pada Gambar 2.



Gambar 2. Nilai Akurasi, *F1-Score*, dan *Cross validation* dari seluruh percobaan dengan teknik pembelajaran mesin K-NN

Hasil klasifikasi dari nilai K=7 akan digunakan untuk perbandingan dengan teknik pembelajaran mesin SVM dan MLP. Detail dari

hasil klasifikasi KNN dengan nilai K=7 dapat dilihat pada Gambar 3. *Genre* yang paling banyak diprediksi benar adalah *metal*, dengan jumlah prediksi benar sebanyak 22 musik.



Gambar 3. *Confusion matrix* hasil klasifikasi menggunakan teknik pembelajaran mesin K-NN dengan K=7

Nilai *precision*, *recall*, dan *F1-Score* untuk setiap *genre* hasil klasifikasi dengan K-NN dapat dilihat pada Tabel 2. Nilai *F1-Score* tertinggi ada pada *genre classical* dan *metal* dengan nilai 0,86, hal ini menunjukkan bahwa musik-musik pada *genre* tersebut dapat terklasifikasi dengan baik. Nilai *F1-Score* terendah ada pada *genre reggae* dengan nilai 0,49, hal ini menunjukkan bahwa musik pada *genre reggae* masih banyak salah diklasifikasi ke *genre* lainnya.

Tabel 2. Nilai *precision*, *recall*, dan F-1 Score hasil klasifikasi menggunakan teknik pembelajaran mesin K-NN dengan K=7

	Precision	Recall	F1-Score
Blues	0,67	0,76	0,71
Classical	0,76	1,00	0,86
Country	0,60	0,70	0,86
Disco	0,45	0,64	0,72
Hip-hop	0,82	0,64	0,72
Jazz	0,89	0,65	0,76
Metal	0,92	0,81	0,86
Pop	0,90	0,69	0,78

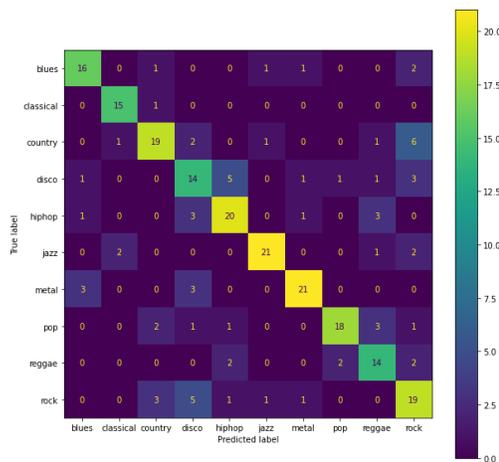
Reggae	0,53	0,45	0,49
Rock	0,69	0,60	0,64

Klasifikasi dengan teknik pembelajaran mesin SVM memakai tiga jenis kernel berbeda yaitu linear, *polynomial*, dan RBF. Tabel 3 menunjukkan nilai *cross validation* dan *accuracy* dari tiga jenis kernel tersebut. Nilai *cross validation* dan akurasi tertinggi dihasilkan dengan memakai kernel *polynomial*. Hasil ini akan dibandingkan dengan hasil teknik pembelajaran mesin K-NN dan MLP.

Tabel 3. Hasil Teknik pembelajaran mesin SVM

Kernel	Cross validation	Accuracy
<i>Polynomial</i>	0,65	0,71
Linear	0,64	0,65
RBF	0,62	0,66

Detail dari hasil klasifikasi dengan teknik pembelajaran mesin SVM dan kernel *polynomial* dapat dilihat pada *confusion matrix* di Gambar 4.



Gambar 4. *Confusion matrix* hasil klasifikasi menggunakan teknik pembelajaran mesin SVM dengan kernel *polynomial*

Nilai *precision*, *recall*, dan *F1-Score* untuk setiap *genre* hasil klasifikasi dengan SVM dapat dilihat pada Tabel 3. Nilai *F1-Score* tertinggi ada pada *genre classical* dengan nilai

0,88, hal ini menunjukkan bahwa musik-musik pada *genre classical* dapat terklasifikasi dengan baik. Nilai *F1-Score* terendah ada pada *genre disco* dengan nilai 0,52.

Tabel 4. Nilai *precision*, *recall*, dan F-1 Score hasil klasifikasi menggunakan teknik pembelajaran mesin SVM dengan kernel *polynomial*

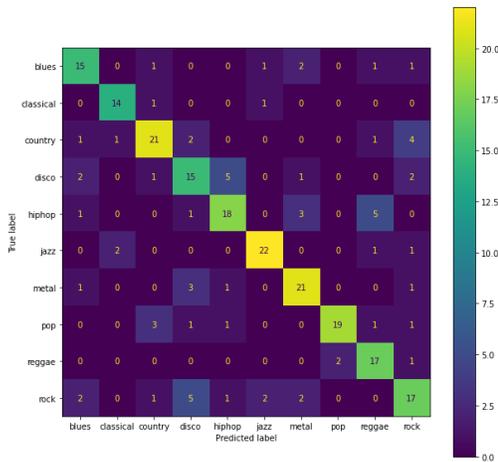
	Precision	Recall	F1-Score
Blues	0,76	0,76	0,76
Classical	0,83	0,94	0,88
Country	0,73	0,63	0,68
Disco	0,50	0,54	0,52
Hip-hop	0,69	0,71	0,70
Jazz	0,88	0,81	0,84
Metal	0,84	0,78	0,81
Pop	0,86	0,69	0,77
Reggae	0,61	0,70	0,65
Rock	0,54	0,63	0,58

Klasifikasi dengan menggunakan MLP dilakukan dengan menggunakan variasi pada *learning rate*, pengurangan *learning rate*, dan jumlah neuron pada *hidden layer*. Tabel 5 menunjukkan lima nilai *cross validation* dan *accuracy* dari semua klasifikasi dengan MLP. Nilai *cross validation* dan akurasi tertinggi dihasilkan dengan memakai kombinasi *learning rate* = 0,3, pengurangan *learning rate* = 0,3, dan jumlah neuron pada *hidden layer* = 48. Hasil ini akan dibandingkan dengan hasil teknik pembelajaran mesin K-NN dan SVM.

Tabel 5. Hasil Teknik pembelajaran mesin MLP

Alpha	Power	Neuron	CV	Accuracy
0,3	0,3	48	58,4	71,6
0,3	0,9	28	58,3	72,4
0,3	0,1	46	58,2	70,0
0,3	0,7	36	58,2	70,8
0,5	0,9	28	58,1	67,6

Hasil klasifikasi dari nilai *learning rate* (α) dan pengurangan *learning rate* ($power$) akan digunakan untuk perbandingan dengan teknik pembelajaran mesin SVM dan MLP. Detail dari hasil klasifikasi SVM dengan nilai $\alpha=0,3$, $power=0,3$, dan jumlah *hidden neuron*=48 dapat dilihat pada Gambar 5.



Gambar 5. *Confusion matrix* hasil klasifikasi menggunakan teknik pembelajaran mesin MLP dengan nilai $\alpha=0,3$, $power=0,3$, dan jumlah *hidden neuron*=48

Nilai *precision*, *recall*, dan *F1-Score* untuk setiap *genre* hasil klasifikasi dengan MLP dapat dilihat pada Tabel 6. Nilai *F1-Score* tertinggi ada pada *genre classical* dengan nilai 0,85, hal ini menunjukkan bahwa musik-musik pada *genre classical* dapat terklasifikasi dengan baik. Nilai *F1-Score* terendah ada pada *genre disco* dengan nilai 0,57.

Tabel 6. Nilai *precision*, *recall*, dan *F-1 Score* hasil klasifikasi menggunakan teknik pembelajaran mesin MLP dengan nilai $\alpha=0,3$, $power=0,3$, dan jumlah *hidden neuron*=48

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Blues	0,68	0,71	0,70
Classical	0,82	0,88	0,85
Country	0,75	0,70	0,72
Disco	0,56	0,58	0,57

Hip-hop	0,69	0,64	0,67
Jazz	0,85	0,85	0,85
Metal	0,72	0,78	0,75
Pop	0,90	0,73	0,81
Reggae	0,65	0,85	0,74
Rock	0,61	0,57	0,59

SIMPULAN

Tiga teknik pembelajaran mesin yaitu K-NN, SVM, dan MLP telah digunakan pada penelitian ini untuk mengklasifikasi musik berdasarkan *genre* dengan menggunakan dataset GTZAN yang terdiri dari 57 fitur berbasis konten.

Model terbaik dipilih dari setiap metode yang digunakan. Model K-NN terbaik adalah dengan $K=7$, model SVM terbaik adalah yang menggunakan kernel *polynomial*, dan model MLP terbaik adalah dengan nilai α 0,3, $power$ 0,3, dan jumlah *hidden neuron* 48.

Teknik pembelajaran mesin MLP menghasilkan nilai akurasi terbesar jika dibandingkan dengan K-NN dan SVM pada penelitian ini dengan akurasi 71,6%, disusul dengan SVM dengan akurasi 71%, dan K-NN dengan akurasi sebesar 69,6%.

DAFTAR PUSTAKA

- [1] J. K. Bhatia, R. D. Singh and S. Kumar, "Music Genre Classification," in *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, Mathura, 2021.
- [2] D. Lau and R. Ajooda, "Music Genre Classification: A Comparative Study Between Deep-Learning And Traditional Machine," in *Proceedings of Sixth International Congress on Information and Communication Technology*, Singapore, 2021.
- [3] M. S. Islam, M. M. Hasan, M. A. Rahim, A. M. Hasan, M. Mynuddin, I. Khandokar and M. J. Islam, "Machine

- Learning-Based Music Genre Classification with PreProcessed Feature Analysis," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 7, no. 3, pp. 491-502, 2021.
- [4] D. R. I. M. Setiadi, D. S. Rahardwika, E. H. Rachmawanto, C. A. Sari, C. Irawan, D. P. Kusumaningrum, Nuri and S. L. Trusthi, "Comparison of SVM, KNN, and NB Classifier for Genre Music Classification based on Metadata," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, 2020.
- [5] P. Sahil, V. Chetan and M. Nikita, "10.1007/978-981-19-2177-3_83," *10.1007/978-981-19-2177-3_83*, vol. 13, no. 1, pp. 15-21, 2022.
- [6] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293-302, 2002.
- [7] I. N. Y. T. Giri, L. A. A. R. Putri, G. A. V. M. Giri, G. N. A. C. Putra, I. M. Widiartha and I. W. Supriyana, "Music Genre Classification Using Modified K-Nearest Neighbor (MK-NN)," *Jurnal Elektronik Ilmu Komputer Udayana*, vol. 10, no. 3, pp. 261-270, 2022.
- [8] P.-N. Tan, M. Steinbach, A. Karpatne and V. Kumar, *Introduction to Data Mining*, Harlow: Pearson Education Limited, 2019.
- [9] Z. R. Tembusai, H. Mawengkang and M. Zarlis, "K-Nearest Neighbor with K-Fold CrossValidation and Analytic Hierarchy Processon Data Classification," *International Journal of Advances in Data and Information Systems*, vol. 2, no. 1, pp. 1-8, 2021.
- [10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg, "Scikit-learn: Machine Learning in Python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.