

# ANALISIS DAN STUDI KOMPARATIF ALGORITMA KLASIFIKASI GENRE MUSIK

Eka Angga Laksana<sup>1)</sup>, Feri Sulianta<sup>2)</sup>

<sup>1), 2)</sup> Teknik Informatika Universitas Widyatama  
Jl Cikutra No. 204 A, Bandung – Jawa Barat 40125  
Email : eka.angga@widyatama.ac.id<sup>1)</sup>, feri.sulianta@widyatama.ac.id<sup>2)</sup>

## Abstrak

*Data mining memiliki beragam algoritma klasifikasi yang terus berkembang. Berbagai bidang baik industry maupun akademis telah memanfaatkan teknik klasifikasi dengan berbagai tujuan. Teknik klasifikasi membuat data menjadi berkelompok berdasarkan kategori atau atribut tertentu. Multimedia pun tak luput dari area yang disentuh oleh klasifikasi. Multimedia sangat berkaitan dengan hiburan dan telah menjadi salah satu kebutuhan manusia, salah satunya adalah musik. Dewasa ini penikmat musik semakin dimanjakan oleh perkembangan musik. Ditambah lagi dengan tren teknologi internet yang memungkinkan berbagai perangkat tersambung dengan internet, seseorang dapat mendengarkan musik dari berbagai perangkat kapanpun dan dimanapun. Menikmati musik tentunya tak lepas dari genre yang bersangkutan. Biasanya seseorang dapat dengan mudah mengenali genre berdasarkan pengetahuan yang dimilikinya. Namun dengan teknik data mining, hal tersebut dapat dilakukan dengan beragam teknik klasifikasi. Klasifikasi genre music sebenarnya bukanlah hal yang baru lagi, beberapa peneliti telah mengangkat topik tersebut dengan teknik yang bervariasi. Namun demikian yang membuat penelitian ini berbeda adalah menemukan algoritma dengan performa terbaik melalui proses evaluasi yang sesuai dengan standar data mining. Algoritma yang diujikan adalah: SVM, KNN, Random Forest, Extra Trees, dan Gradient Boosting.*

**Kata kunci:** data mining, klasifikasi, music, genre, algoritma, visualisasi, evaluasi, SVM, KNN, Random Forest, Extra Trees, Gradient Boosting.

## 1. Pendahuluan

Musik telah menjadi bagian dari kehidupan manusia sehari-hari. Keberadaan musik yang dinikmati melalui indra pendengaran manusia, memiliki manfaat utama dalam dari sisi hiburan, pendidikan, budaya dan lain-lain. Jika pada zaman dahulu manusia mendengarkan musik dari pertunjukan langsung, saat ini musik telah dapat disimpan dalam perangkat dan dimainkan dimanapun berada. Hal tersebut dibuktikan dengan perkembangan berbagai perangkat digital seperti *smart phone* dan *digital player*. Saat ini orang-orang cenderung menikmati musik modern. Berbeda dengan musik tradisional yang mengedepankan nuansa etnik dan perangkat tradisional, musik modern banyak dipengaruhi oleh budaya luar dan terutama dari segi perangkat yang

digunakan juga lebih modern. Audio jenis tersebut dapat dikategorikan menjadi beberapa tipe seperti pop, rock, reggae dan lain-lain. Biasanya orang-orang menyukai musik dengan jenis tertentu.

Kategori file audio atau disebut juga sebagai klasifikasi ditentukan oleh beberapa hal. Ritme, tempo, tinggi rendah nada merupakan contoh dari ciri-ciri yang menandai bahwa suatu musik termasuk ke dalam klasifikasi tertentu. Manusia berdasarkan pengalaman yang dimilikinya dapat mengetahui klasifikasi atau genre suatu musik dengan mudah. Kemampuan tersebut berguna dalam menyeleksi jenis musik tertentu untuk dinikmati sebagai hiburan. Namun manusia memiliki kemampuan terbatas dalam mengenali genre suatu audio terutama jika dihadapkan pada file audio digital yang begitu banyak. Meskipun pada dasarnya bisa saja dilakukan, tetap saja waktu menjadi masalah utama.

Kemampuan manusia dalam belajar, dapat ditiru oleh mesin komputer atau disebut juga dengan istilah kecerdasan buatan. Pada umumnya manusia manusia menebak genre suatu lagu berdasarkan beberapa hal, yaitu: tempo dan ritme lagu. Hal inilah yang mendasari bahwa pada setiap file audio pun memiliki pola. Dengan mempelajari pola tersebut komputer juga mampu belajar dan merekam keterkaitan pola dengan genre music tertentu. Proses tersebut disebut juga sebagai *machine learning* atau dengan kata lain disebut sebagai teknik *Data Mining*. Terdapat dua tipe proses *learning* yaitu supervised dan unsupervised learning[1].

Agar komputer dapat melakukan klasifikasi, diperlukan metode bagaimana mesin tersebut mampu untuk 'belajar'. Metode pada *machine learning* digolong menjadi dua macam yaitu supervised dan unsupervised learning[2]. Istilah supervised didefinisikan sebagai proses dalam mempelajari suatu pola yang seolah-olah didampingi dan berkaitan erat dengan prediktif modelnya. Hal ini dibuktikan dengan dataset yang memiliki label pada kolom terakhir, berperan sebagai kesimpulan atas paparan fakta-fakta yang tercatat. Berbeda halnya dengan *unsupervised learning* yang mana data tidak memiliki label, pengelompokan atau cluster adalah teknik yang umum dilakukan pada teknik jenis ini. Pada dasarnya, data akan membentuk suatu cluster berdasarkan persamaan pola yang dimilikinya. Pada penelitian ini, merupakan klasifikasi berdasarkan pada teknik supervised learning, dimana label ditentukan

prediktif model sebagai urutan fakta penyusun label tersebut akan dijelaskan pada poin preprocessing.

Terdapat beberapa algoritma yang umum dipakai dalam supervised learning. Penelitian ini ditujukan untuk manakah di antara algoritma tersebut yang memiliki performa paling baik dalam klasifikasi genre. Uji performa dilakukan dengan kriteria tertentu yang umum dipakai sebagai teknik evaluasi dalam data mining. Penelitian ini diharapkan dapat memberikan kontribusi terhadap ilmu pengetahuan dengan mengetahui teknik manakah yang memiliki performa paling baik sehingga dapat menginspirasi baik bagi kami maupun peneliti lain untuk mengembangkannya lebih lanjut lagi.

## 2. LANDASAN PENELITIAN

Pada poin ini menjelaskan tentang hal apa saja yang mendasari teknik klasifikasi audio.

### A. Ekstraksi fitur

Komponen paling utama dalam audio analisis adalah *feature extraction*. *Feature extraction* merupakan proses mengubah segmen audio menjadi bentuk numerik yang mana merepresentasikan karakteristik dari audio yang bersangkutan. Perjalanan audio analisis dari *feature extraction* hingga klasifikasi dapat dibagi menjadi dua tahap. Pertama, deskripsi fitur yaitu membangun jenis *feature* apa saja sebagai komponen utama dari pengenalan pola. Tentu saja hal ini sangat menantang sebab menentukan hasil akhir proses analisis jenis apapun yang dilakukan setelahnya. Kedua, setelah fitur dikenali maka metode analisis klasifikasi dengan berbagai algoritma *machine learning* dapat dilakukan. Pada dasarnya *feature extraction* bukan hal yang baru lagi sehingga penelitian ini menggunakan teknik-teknik yang sebenarnya sudah dibuat oleh peneliti lain sebelumnya. Oleh karena penelitian ini lebih difokuskan pada beberapa algoritma klasifikasi data mining untuk menguji jenis metode klasifikasi mana yang terbaik.

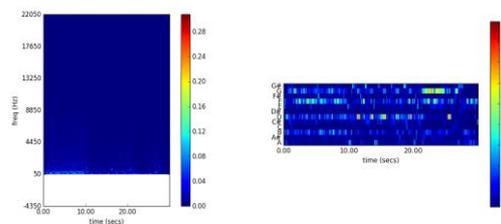
*Feature extraction* pada umumnya berupa beberapa bentuk representasi dari waktu dan frekuensi. Salah satu bentuk frekuensi yang dipakai dalam klasifikasi audio adalah MFCC (Mel-Frequency cepstral coefficient)[3]. MFCC berupa representasi spectral yang mencerminkan sinyal energi dan berpusat pada koefisien pertama. Jika dijumlahkan, maka total terdapat 34 jenis *feature extraction* yang dipakai dalam audio analisis[4]. Terbagi menjadi dua jenis fitur yaitu time-domain *feature* (jenis ke 1 hingga 3) secara langsung diekstrak dari sampel raw signal. Jenis kedua yaitu fitur frekuensi domain (4 hingga 34, terlepas dari MFCC) berdasarkan tingkat Discrete Fourier Transform (DFT). Beberapa contoh dari jenis pertama yaitu: *Zero crossing rate*, *Energy*, *Entropy of energy*. Sedangkan pada jenis fitur kedua yaitu: *spectral spread*, *spectral centroid*, *MFCC* dan *chroma Vector*.

Cara paling mudah untuk menggambarkan *feature extraction* adalah melalui visualisasi Spectrogram dan Chromagram. Spectrogram menunjukkan frekuensi tiap satuan waktu dalam sebuah file audio digital. Sedangkan chromagram menunjukkan nada yang dicapai tiap satu satuan waktu sebuah file audio. Untuk lebih jelasnya, dapat dilihat pada gambar 1 dan 2.

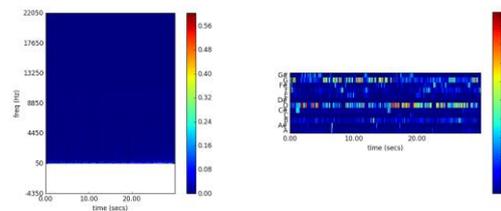
### B. Dataset

Penelitian ini menggunakan musik genre dataset. Dataset yang sama pernah pula digunakan oleh G. Tzanetakis dan P.Cook pada paper yang berjudul “*Musical genre classification of audio signals*”[5]. Dataset audio digunakan untuk kepentingan akademis dan penelitian dan tidak ada unsur komersial yang terlibat. Dilihat dari sejarah terkumpulnya dataset tersebut, bermula dari tahun 2000 – 2001 diperoleh dari berbagai sumber termasuk CD, radio, rekaman microphone dan berbagai media rekaman lainnya. Dataset berjumlah 1000 audio track dimana masing-masing berdurasi 30 detik. Atribut track lainnya berupa frekuensi sebesar 22050Hz mono 16 bit dalam format .wav, dan jika ditotal ukuran dataset mencapai lebih dari 1 Gb. Genre audio yang diolah meliputi: blues, classic, country, disco, hiphop, jazz, metal, pop, reggae dan rock.

Implementasi ekstraksi fitur pada dataset dapat dilihat pada gambar 1. Gambar berikut diambil dari dua file dengan genre yang berbeda. Pada masing-masing gambar terbagi menjadi dua jenis yaitu: spectrogram dan chromagram. Spectrogram merupakan perbandingan frekuensi dengan panjang waktu tertentu sedangkan chromagram merupakan pencapaian nada dalam durasi tertentu.



Gambar 1. Spectrogram & Chromagram sample music rock



Gambar 2. Spectrogram dan Chromagram sample music Pop

Jika diperhatikan, kedua gambar tersebut memiliki perbedaan. Ekstraksi fitur mampu mengenali karakteristik dari setiap genre audio yang diujikan.

Gambar di atas merupakan contoh ekstraksi fitur pada atribut frekuensi dan nada. Namun dalam implementasinya terdapat 37 jenis atribut sehingga sangat mencukupi jumlah fitur set yang dibutuhkan dalam proses *machine learning*.

### C. Penunjang Eksperimen

Eksperimen penelitian ini dilakukan dengan bahasa pemrograman Python dan didukung oleh library Scikit Learn[6] dan segala dependensinya untuk melakukan proses klasifikasi. Riset di bidang multimedia sangatlah kompleks sehingga peneliti ini memerlukan library khusus terutama pada audio processing. Library yang digunakan adalah pyAudioAnalysis yang mana merupakan salah satu Python Library yang mendukung audio analisis, meliputi: feature extraction, klasifikasi sinyal audio, segmentasi dan visualisasi. Ditambah lagi dengan penerapan metode evaluasi yang diambil dari Scikit Learn menambah lengkap proses pengujian pada penelitian. Dengan informasi ini diharapkan dapat memberi inspirasi bagi peneliti lain untuk mengembangkan penelitian selanjutnya yang lebih baik lagi.

### D. Penelitian Terkait

Poin ini menjelaskan mengenai beberapa penelitian terdahulu terkait dengan klasifikasi genre musik. Dataset yang sama pernah digunakan pula oleh Baniya et. al[7] dengan menggunakan metode ELM (Extreme Learning Machine) dan mencapai tingkat akurasi sebesar 85,15%. Ajoodha et. al menggunakan dataset GTZAN pun, dan melakukan perbandingan akurasi terhadap beberapa algoritma seperti: Naïve Bayes, SVM, Multilayer Perceptron, Linier Logistic regression models, K-NN dan Random Forest dan akurasi terbaik dicapai oleh Linier Logistic regression models sebesar 81%[8]. Namun dari kedua penelitian tersebut tidak disebutkan tool dan Bahasa pemrograman apa yang digunakan. Selain itu, penelitian lain terkait dengan klasifikasi musik dilakukan juga oleh [9] dengan membandingkan antara Naïve Bayes, BayesNet, J48, SMO dan Logistic Classifier menggunakan Java-Weka Library hingga didapatkan akurasi mencapai 100% oleh SMO. Weka library juga digunakan oleh [10][11], namun lebih menekankan pada tahap pre processing yaitu pada space time decomposition dan feature selection menggunakan algoritma genetika.

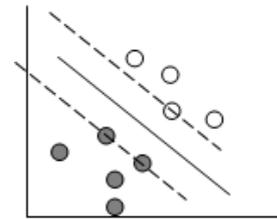
### 3. Algoritma Klasifikasi

Berikut adalah penjelasan algoritma yang dipakai dalam eksperimen.

#### a. SVM

SVM (Support Vector Machine) membangun hyper-plane atau beberapa set hyper-plane di dalam ruang dimensi yang tinggi bahkan tak terbatas[12]. Algoritma SVM dapat digunakan pada klasifikasi, regresi dan lain lain. Pada klasifikasi, algoritma ini memerlukan output berupa optimal hyper-plane yang mana mengkategorikan sample set yang baru. Pada gambar di bawah garis lurus

merupakan batasan optimal dari hyperplane dan diapit oleh dua garis putus-putus yang berperan sebagai margin maksimal dari training data.



Gambar 3.SVM Hyperplane

Diketahui training vector  $x_i \in \mathbb{R}^p, i = 1, \dots, n$  dalam dua kelas, dan vector  $y \in \{1, -1\}^n$ , svm menyelesaikan permasalahan berikut

$$\min_{w,b,c} \frac{1}{2} w^T w + C \sum_{i=1}^n C_i \quad (1)$$

$$\text{Subject to } y_i (w^T \phi(x_i) + b) \geq 1 - C_i, \\ C_i \geq 0, i = 1, \dots, n$$

Dengan elemen keduanya adalah

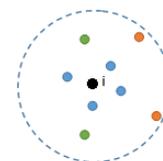
$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \quad (2) \\ \text{Subject to } y^T \alpha = 0 \\ 0 \leq \alpha_i \leq C, i = 1, \dots, n$$

Dimana  $e$  adalah vector dari keseluruhan,  $C > 0$  adalah batas atas,  $\phi$  adalah  $n$  dari  $n$  positif semidefinite matrix.

#### b. KNN

K-Nearest neighbour adalah salah satu algoritma klasifikasi berdasarkan pengenalan pola kedekatan object dengan sekelompok object tertentu sejumlah k [2]. Kedekatan pola dihitung berdasarkan jarak kemiripan, ada beberapa metode untuk menghitung jarak antara training set dengan suatu object dan salah satunya yang populer yaitu Euclidian Distance. Berikut adalah langkah penyusunan algoritma nearest neighbour:

- Bentuk dataset training
- Hitung jarak (metric) untuk menghitung tingkat kemiripan antar object
- Nilai k, sejumlah k training object untuk nantinya dipakai sebagai klasifikasi objek baru



Gambar 4.K-Nearest neighbour

Setelah tahap di atas dilakukan, suatu object baru dapat diklasifikasikan dengan mengikuti tahapan berikut:

Hitung jarak antara keseluruhan training dengan object baru. Gambar 4 menunjukkan kedekatan item i sebagai pusat dari lingkaran dengan item item lainnya. Jumlah parameter k dapat diubah menurut hasil evaluasi yang paling optimal.

Hitung nilai kedekatan sejumlah k object, urutkan berdasarkan object terdekat Tentukan label, jika ada lebih dari 1 k maka ditentukan berdasarkan frekuensi tertinggi.

$$Euclidean = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (3)$$

X dan y adalah atribut dari masing masing item i. K nearest neighbor tepat dipakai untuk variable berkelanjutan dimana nilai atribut berupa angka.

**c. Random Forest**

Random forest merupakan algoritma yang dapat dipakai pada klasifikasi dan regresi. Diperkenalkan oleh Leo Breinan[13] dimana teknik ini dapat menghasilkan banyak pohon klasifikasi. Bagaimana Random Forest dapat menghasilkan klasifikasi berawal mula dari input vector yang bergerak menuruni masing masing pohon. Masing-masing pohon merupakan klasifikasi berdasarkan mekanisme suara terbanyak atau vote untuk menandai class tersebut. Sehingga pepohonan (forest) dapat menentukan klasifikasi berdasarkan hasil voting tersebut.

Sesuai dengan penelitian [13] forest error rate bergantung pada dua hal berikut:

- Korelasi antara dua pohon di forest. Meningkatnya korelasi meningkat pula error rate.
- Ketahanan tiap individu pohon di forest. Tree dengan error rate rendah memiliki klasifikasi yang akurat.

Berdasarkan penelitian Leo Breinan, diperoleh karakteristik Random Forest sebagai berikut:

- Akurasi sebaik Adaboost, terkadang lebih baik.
- Robust terhadap noise dan outliers.
- Lebih cepat daripada metode bagging atau boosting.
- Dapat menunjukkan informasi berguna mengenai estimasi error, strength, korelasi dan variable.
- Sederhana dan dapat dijalankan parallel.

**d. Extra trees**

Klasifikasi dengan metode extra trees atau yang disebut juga sebagai “Extremly randomized Trees” merupakan varian pengembangan dari decision tree acak pada berbagai sub bagian dataset dan menghitung rata-ratanya untuk meningkatkan akurasi prediksi dan pengendalian over vitting. Berbeda dengan Random Forest dimana pada setiap tahapannya, sample dan keputusan diambil secara acak dan bukan diambil dari yang terbaik[14]. Estra trees membangun grup decision tree sesuai dengan prosedur top-down.

**e. Gradient boosting**

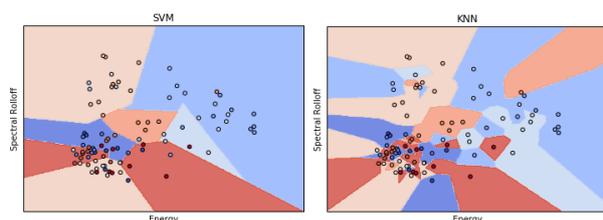
Gradient bossting, seperti halnya keluarga algoritma Boosted lainnya memiliki kemampuan untuk meningkatkan akurasi prediktif model. Beberapa algoritma boosting lainnya seperti: XGBoost, AdaBoost dan GentleBoost memiliki formula matematika tersendiri dan bervariasi. Konsep Gradient Boosting terletak pada pengembangannya yang mana memiliki ekspansi tambahan terhadap fitting criterion [15]. Berawal dari metode Bagging yaitu mengambil sampel data secara acak, bangun algoritma dan hitung rata-rata segala kemungkinan yang terjadi termasuk error dan akurasi. Daripada mengambil secara acak, pemilihan sample dapat dilakukan secara lebih cerdas dengan menggunakan fungsi Boost. Misalkan diketahui sebuah model sebagai M dengan akurasi sebesar 82%. Meningkatkan akurasi dapat dilakukan dengan cara membangun ulang model secara keseluruhan menggunakan input variable yang baru dan bangun ulang model. Gradient Boost melakukan ekspansi terhadap model matematika berikut [15]:

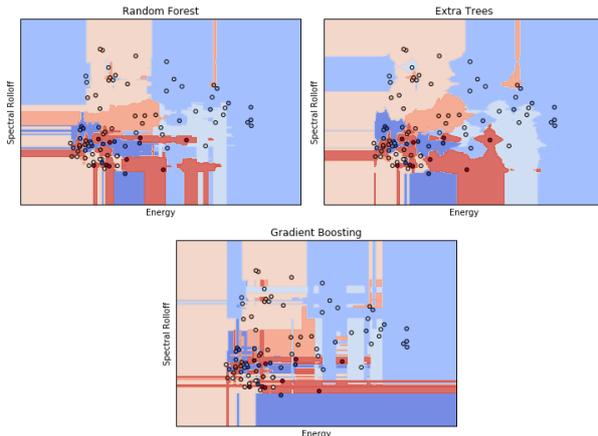
$$F(x; \beta_m, a_m) = \sum_{m=1}^M \beta_m h(x; a_m) \quad (4)$$

Fungsi  $h(x; a)$  biasanya merupakan fungsi parameter untuk input variable x, dikaitkan dengan parameter  $a = \{a_1, a_2, \dots\}$ . Kondisi berbeda pada nilai  $a_m$  yang dipilih untuk parameter ini. Formula tersebut merupakan inti dari beberapa metode seperti Neural Network, Radial Vector Machine dan Support Vector Machine[15].

**3.1 Visualisasi klasifikasi**

Poin ini menjelaskan mengenai apa yang terjadi pada data setelah selesai dilakukan klasifikasi. Ada beberapa persiapan sebelum sebelum membentuk klasifikasi, seperti tercantum pada gambar 5. Seperti halnya visualisasi audio yang telah dijelaskan pada poin sebelumnya, maka dipilihlah dua dari total 37 atribut. Dikarenakan visualisasi berupa scatter plot merupakan bentuk dua dimensi, maka dipilihlah Energy dan Spectral Rolloff. Kemudian karena keterbatasan komputasi maka dipilih 9 file audio per genre sehingga total terdapat 90 dataset. Area di dalam plot melambangkan dua hal yaitu titik-titik yang merepresentasikan tiap audio dan area latar belakang dengan batas daerahnya masing-masing yang mana merupakan hasil prediksi. Bagaimana hal tersebut dapat dilakukan? Pada dasarnya, membentuk plot merupakan visualisasi perbandingan hasil prediksi dengan kondisi sesungguhnya. Jika dilihat lebih detail, ada beberapa titik yang ‘tersesat’, menempel pada area yang bukan domainnya.





Gambar 5. Plot klasifikasi untuk masing masing algoritma

#### 4. Teknik pengujian

Pembuktian suatu model belumlah lengkap jika belum dilakukan proses evaluasi. Sesuai dengan tujuan utama dari penelitian ini yaitu untuk menemukan algoritma dengan performa terbaik, maka teknik evaluasi menjadi hal yang wajib dilakukan. Proses evaluasi menggunakan teknik yang umum dipakai dalam data mining. Berikut adalah teknik evaluasi [6] yang telah diimplementasikan dalam penelitian ini dan penjelasan singkatnya:

##### a. Akurasi

Seberapa akurat suatu model menghasilkan klasifikasi ditunjukkan oleh akurasi. Model akurasi bermula dari  $y'$  dimana merupakan nilai prediksi dari sample ke  $i$  dan  $y_i$  adalah nilai prediksi benarnya. Formula pembagian prediksi benar untuk  $n_{samples}$  adalah sebagai berikut:

$$accuracy(y, y') = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(y'_i = y_i) \quad (5)$$

##### b. Presisi dan recall

Presisi adalah kemampuan suatu model klasifikasi untuk tidak melabeli positif pada sample yang sebenarnya adalah negatif. Sedangkan recall adalah kemampuan model untuk menemukan semua sample positif, dengan kata lain menguji seberapa jauh benar atau salahnya hasil prediksi bila dibandingkan dengan kondisi sebenarnya. Berbeda dengan klasifikasi binari yang memiliki keanggotaan label true dan false, penelitian ini melibatkan klasifikasi multi class dan multi label. Hal ini ditandai dengan jumlah genre music berjumlah 10 jenis yang mana menjadi keanggotaan label atau target prediksi:

$$P(A, B) := \frac{|A \cap B|}{|A|} \quad (6)$$

$$R(A, B) := \frac{|A \cap B|}{|B|} \quad (7)$$

##### c. F1 score

Presisi dan recall merupakan properti yang bertolak belakang, sebab presisi yang tinggi berdampak pada recall yang rendah, demikian sebaliknya. F Measure atau disebut juga F1 Score menunjukkan keterkaitan antara presisi dan recall atau sering didefinisikan dengan istilah rata-rata harmonic dari presisi dan recall. Berikut adalah formula untuk menghitung F1 score antara A dan B:

$$F_{\beta}(A, B) := (1 + \beta^2) \frac{P(A, B) \times R(A, B)}{\beta^2 P(A, B) + R(A, B)} \quad (8)$$

#### 5. Hasil dan Pembahasan

Poin utama dari penelitian ini dapat dilihat dari hasil eksperimen yang telah dilakukan, sebab hal tersebut akan menjawab objektif yang dipaparkan pada bagian pendahuluan. Berdasarkan eksperimen yang telah dilakukan menunjukkan bahwa algoritma Extratree memiliki akurasi paling baik, diikuti Random Forest dan algoritma lainnya sesuai dengan Gambar 6.



Gambar 6. Perbandingan akurasi

Seperti yang telah dipaparkan sebelumnya bahwa skor F1 merupakan rata-rata harmoni presisi dan recall. Maka untuk menggambarkan keterkaitan antara presisi dan recall cukup dengan menampilkan perbandingan skor F1 antar algoritma seperti nampak pada gambar 7. Algoritma Extra Tree masih menunjukkan performa terbaik dengan 72.3% terpaut tipis dengan Random Forest.



Gambar 7. Perbandingan skor F1

Terkait dengan berapa jumlah prediksi benar dan salah, confusion matrix merupakan teknik yang tidak boleh dilewatkan. Tabel confusion matrix menunjukkan detail misklasifikasi per genre. Tabel bagian atas mepresentasikan hasil prediksi dan kolom paling kiri

merupakan atribut yang diobservasi[2]. Setiap cell menampilkan jumlah atribut asli yang termasuk ke dalam hasil prediksinya. Setiap algoritma memiliki tabel confusion matrix masing-masing, namun untuk penelitian ini hanya menampilkan khusus pada Extra Trees sebagai algoritma dengan performa terbaik, seperti dilihat pada tabel 1.

**Tabel 1.** Confusion Matrix hasil klasifikasi Extra Trees

```
Confusion Matrix:
      blu  cla  cou  dis  hip  jaz  met  pop  reg  roc,
blu  7.8  0.0  0.7  0.5  0.1  0.2  0.3  0.0  0.1  0.3,
cla  0.0  9.3  0.0  0.0  0.0  0.2  0.0  0.0  0.0  0.4,
cou  0.4  0.3  7.0  0.7  0.0  0.2  0.2  0.2  0.2  0.8,
dis  0.1  0.0  0.8  6.9  0.5  0.1  0.2  0.5  0.5  0.4,
hip  0.1  0.0  0.0  0.6  7.1  0.1  0.4  0.5  1.1  0.1,
jaz  0.6  1.1  0.4  0.4  0.0  6.9  0.3  0.0  0.0  0.3,
met  0.1  0.0  0.1  0.4  0.1  0.0  8.7  0.0  0.0  0.6,
pop  0.0  0.1  0.8  0.6  0.5  0.3  0.0  7.2  0.1  0.4,
reg  0.4  0.0  0.4  0.4  1.0  0.1  0.1  0.6  6.8  0.1,
roc  0.4  0.5  1.3  0.8  0.1  0.1  1.2  0.3  0.6  4.7,
Selected params: 200.00000],
dtype=|S200)
```

**Kesimpulan**

Musik merupakan jenis multimedia audio yang telah banyak dinikmati oleh berbagai kalangan. Ditambah lagi dengan perkembangan ‘internet of everything’ dimana setiap orang dapat menikmati file audio berkat sambungan internet, semakin menambah kemudahan dalam mengakses hiburan berupa audio. Dalam rangka menikmati hiburan musik setiap orang memiliki selera yang berbeda-beda, terutama dengan genre musik tertentu. Genre suatu file audio berkaitan dengan erat dengan karakteristik yang dimilikinya. Sebuah perangkat atau sistem harus mampu meniru kemampuan manusia dalam mengenali genre. Klasifikasi merupakan teknik yang dapat dipakai dalam memprediksi genre suatu audio. Klasifikasi pun terus berkembang dan tekniknya pun beragam. Setelah dilakukan uji coba dengan membandingkan antara SVM, KNN, Random Forest, Extra Trees dan Gradient Boosting ditemukan bahwa algoritma Extra Trees memiliki performa yang paling baik. Performa yang diujikan yaitu: akurasi, presisi & recall dan F1 score. Manfaat klasifikasi tentu sangat banyak, baik di dunia industri maupun akademis. Di industri misalkan dapat dipakai oleh penyedia layanan komunikasi audio streaming untuk menyusun daftar koleksi musik agar lebih pelanggan dalam memilih suatu lagu.

**Daftar Pustaka**

[1] I. H. Witten, *Data mining : practical machine learning tools and techniques*, 3rd ed. Burlington: Morgan Kaufmann, 2011.  
 [2] Florin Gorunescu, *Data Mining: Concepts, Models and Techniques*. Springer, 2011.  
 [3] N. J. Hunt, N. Lennig, P. Mermelestein, and B. N. Reece, “Syllable-based recognition,” no. 3, pp. 880–883, 1980.  
 [4] T. Giannakopoulos, “PyAudioAnalysis: An open-source python library for audio signal analysis,” *PLoS One*, vol. 10, no. 12, pp. 1–17, 2015.  
 [5] G. Tzanetakis, S. Member, and P. Cook, “Musical Genre Classification of Audio Signals,” vol. 10, no. 5, pp. 293–302, 2002.  
 [6] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2012.  
 [7] B. K. Baniya, D. Ghimire, and J. Lee, “Automatic music genre

classification using timbral texture and rhythmic content features,” *Int. Conf. Adv. Commun. Technol. ICACT*, vol. 2015–August, no. 3, pp. 434–443, 2015.  
 [8] R. Ajoodha, R. Klein, and B. Rosman, “Single-labelled Music Genre Classification Using Content-Based Features,” 2015.  
 [9] M. M. Panchwagh, “Music Genre Classification Using Data Mining Algorithm,” pp. 49–53, 2016.  
 [10] C. N. Silla, A. L. Koerich, and C. A. A. Kaestner, “A Machine Learning Approach to Automatic Music Genre Classification,” *J. Brazilian Comput. Soc.*, vol. 14, no. 3, pp. 7–18, 2008.  
 [11] C. A. A. Jr, Carlos N Silla;Koerich, Alessandro L., Kaestner, “A Feature Selection Approach for Automatic Music Genre Classification,” *Int. J. Semant. Comput.*, vol. 3, no. 2, pp. 183–208, 2009.  
 [12] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.  
 [13] L. Breiman, “RANDOM FORESTS,” pp. 1–35, 1999.  
 [14] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely randomized trees,” no. October 2005, 2006.  
 [15] J. H. Friedman, “Greedy Function Approximation: A Gradient Boosting Machine,” *Ann. Stat.*, vol. 29, p. 5, 2001.

**Biodata Penulis**

**Eka Angga Laksana**, Saat ini menjadi dosen di Universitas Widyatama. Pada tahun 2010 meraih beasiswa S2 BPKLN DIKTI untuk melakukan studi di Universitas Dian Nuswantoro dan mengikuti program dual degree dengan Universiti Teknikal Malaysia Melaka (UTeM). Hingga menyelesaikan studi Master pada tahun 2012 dengan gelar Magister Komputer dan Master of Computer Science. Fokus penelitian di bidang Recommender system, collaborative filtering dan pemanfaatan Data mining di industri hingga multimedia mining seperti audio dan citra. Seorang web developer dan mendalami bahasa pemrograman Python terutama dalam melakukan teknik machine learning.

**Feri Sulianta**, berprofesi sebagai dosen sejak tahun 2005 dan pada tahun 2011 bergabung sebagai dosen tetap di Universitas Widyatama. Ruang lingkup penelitian: data mining, basis data, sosial masyarakat.