KLASIFIKASI MUSIK BERDASARKAN GENRE PADA LAYANAN STREAMING MUSIK SPOTIFY MENGGUNAKAN ALGORITMA *K–NEAREST NEIGHBOR* DAN *MODIFIED K*–*NEAREST NEIGHBOR*

**I Made Tangkas Wahyu Kencana Yuda1) I Wayan Supriana2) I Gede Santi Astawa3)**

Program Studi Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam1) 2) 3)

Universitas Udayana, Badung, Bali1) 2) 3)

[tangkaswahyu98@gmail.com](mailto:tangkaswahyu98@gmail.com) 1) wayansupriana@unud.ac.id2) gedesantiastawa@unud.ac.id3)

# **ABSTRACT**

*In today's technological developments, music can be heard through various media, one of which is streaming music media. The rapid increase in the amount of music in digital form has made manual genre labeling ineffective. Genre labeling can be done automatically by applying artificial intelligence that can classify music based on genre by using the features in the song attributes.*

*This research was conducted to determine the effect of song attributes used on performance to classify songs using the K – Nearest Neighbor (KNN) and Modified K – Nearest Neighbor (MKNN) classification algorithm methods, and the comparison of the performance produced by the KNN and MKNN algorithms in genre classification. This study uses 12 audio features with 14 genres on Spotify, and also selects features using Principal Component Analysis (PCA) by comparing KNN and MKNN using k-fold cross-validation and measuring how precise the classification of classes is using confusion matrix with accuracy for each. genre class.*

*It was found that the effect of the song attributes used on performance to classify songs using KNN and MKNN. In addition, there is also a comparison of the performance produced by the KNN and MKNN algorithms in the classification of 14 genres using 12 feature variables, where accuracy without using PCA feature selection is the highest, while when using Principal Component Analysis, it is concluded that the KNN method without PCA based on the level of accuracy is the best method compared to other methods.*

***Keywords*** *: K–Nearest Neighbor, Modified K–Nearest Neighbor, Genre Classification, Streaming Music Media, Spotify*

# **ABSTRAK**

Dalam perkembangan teknologi era modern saat ini, musik dapat didengarkan melalui berbagai macam media, salah satunya media *streaming music*. Bertambahnya jumlah musik dalam bentuk digital secara pesat menyebabkan pemberian label *genre* secara manual menjadi tidak efektif. Pemberian label *genre* secara otomatis dapat dilakukan dengan menerapkan algoritma kecerdasan buatan yang dapat mengklasifikasi musik berdasarkan *genre* dengan menggunakan fitur–fitur pada atribut lagu.

Penelitian ini dilakukan untuk mengetahui pengaruh atribut lagu yang digunakan terhadap kinerja untuk mengklasifikasi lagu menggunakan metode algoritma klasifikasi *K–Nearest Neighbor* (KNN) dan *Modified K–Nearest Neighbor* (MKNN), dan perbandingan performa yang dihasilkan algoritma KNN dan MKNN pada *klasifikasi genre*. Penelitian ini menggunakan 12 fitur audio dengan 14 *genre* pada Spotify, dan juga melakukan seleksi fitur menggunakan *Principal Component Analysis* (PCA) dengan membandingkan KNN dan MKNN menggunakan *k–fold cross–validation* dan mengukur seberapa tepat klasifikasi terhadap kelas menggunakan *confusion matrix* dengan akurasi pada tiap kelas *genre*.

Dari penelitian ini, didapatkan pengaruh atribut lagu yang digunakan terhadap kinerja untuk mengklasifikasi lagu menggunakan KNN dan MKNN. Selain itu juga didapatkan perbandingan performa yang dihasilkan algoritma KNN dan MKNN pada klasifikasi 14 *genre* dengan menggunakan 12 variabel fitur, yang mana akurasi tanpa menggunakan seleksi fitur PCA adalah yang tertinggi, sedangkan ketika menggunakan seleksi fitur *Principal Component Analysis* (PCA), disimpulkan bahwa metode KNN tanpa PCA berdasarkan pada tingkat akurasi merupakan metode terbaik dibandingkan metode lainnya.

*Kata kunci*— *K–Nearest Neighbor, Modified K–Nearest Neighbor, Klasifikasi Genre, Media Streaming Music, Spotify*

# **PENDAHULUAN**

Menurut KBBI (2020), Musik diartikan sebagai: 1) Ilmu dan seni menyusun nada atau suara dalam urutan, kombinasi, dan hubungan temporal untuk menghasilkan komposisi (suara) yang mempunyai kesatuan dan kesinambungan, 2) Nada atau suara yang disusun sedemikian rupa sehingga mengandung irama, lagu, dan keharmonisan (terutama yang menggunakan alat–alat yang dapat menghasilkan bunyi–bunyi itu). Komposisi menjadi penting karena pada alunan musik yang indah tergantung pada penataan atriut musik. Jenis–jenis atribut pada musik ini dapat menghasilkan musik yang berbeda*–*beda dari yang lain.

Perkembangan teknologi dalam era modern ini, musik dapat didengarkan melalui berbagai macam media, baik melalui media transmisi radio hingga media *streaming*. Adapun jasa layanan *streaming* *music*, diantaranya Spotify, Joox, Apple Music, Deezer dan lain sebagainya. Sejak diluncurkan pada 2008 lalu, Spotify merupakan salah satu jasa layanan *streaming music* digital yang paling banyak digunakan oleh pendengar musik, seperti dijelaskan pada Gambar 1.1 (Iriansyah, 2018). Pengguna Spotify dapat mengakses musik berdasarkan *artist*, album, *genre*, *playlist*, *podcasts*, atau label rekaman.



**Gambar 1.** Keunggulan Spotify (Iriansyah, 2018).

Menurut Giri (2018), *genre* musik adalah salah satu cara pengelompokan musik sesuai dengan kemiripannya satu sama lain yang sangat umum digunakan untuk mengatur *database* musik digital. *Database* atau katalog musik yang diorganisasi berdasarkan *genre* musik memberikan kemudahan pada pendengar musik untuk mencari musik–musik sejenis yang sesuai dengan referensinya. Bertambahnya jumlah musik dalam bentuk digital secara pesat menyebabkan pemberian label *genre* secara manual menjadi tidak efektif. Pemberian label *genre* secara otomatis dapat dilakukan dengan menerapkan algoritma kecerdasan buatan yang dapat mengklasifikasi musik berdasarkan *genre* dengan menggunakan fitur–fitur pada atribut lagu.

Pemberian label *genre* musik yang ada di Spotify cukup banyak, seperti *Jazz, Gospel, Blues, Funk, Rock, EDM, Reggae, Hip–Hop, Pop, Pop Punk, Pop Rock, Slow Pop*, dan lainnya. Dari *genre* musik tersebut kemudian terbagi lagi menjadi beberapa *playlist*, sehingga total *playlist* yang ada di Spotify berjumlah ribuan. Pembagian musik berdasarkan *genre* ini tentunya untuk mempermudah pengguna Spotify dalam memilih atau menemukan lagu kesukaannya sesuai *genre* maupun *playlist*.

Terdapat beberapa penelitian sebelumnya mengenai pengklasifikasian pada *genre* musik. Antara lain, pada penelitian Giri (2018), melakukan penelitian klasifikasi musik berdasarkan 10 *genre* yaitu *classical, EDM, hip–hop, metal, pop, punk, R&B, rap, soul* dan *rock* dengan metode *K–Nearest Neighbor (KNN)* dan menggunakan 11 fitur audio pada atribut lagu (*speechiness, energy, danceability, loudness, tempo, mode, valence, instrumentalness, accousticness, key,* dan *liveness*). Penelitian ini mendapatkan nilai akurasi klasifikasi KNN sebesar 44,8%. Nilai tertinggi ada pada *genre* *Classical* dengan akurasi 100% dan nilai terendah ada pada *genre* *Pop* dengan akurasi 25%.

Selanjutnya Okfalisa, dkk (2017) pada penelitiannya melakukan perbandingan akurasi metode klasifikasi *K–Nearest Neighbor (KNN)* dan *Modified K–Nearest Neighbor* (MKNN). Analisis komparatif didasarkan pada keakuratan kedua algoritma. Sebelum klasifikasi, *k–fold cross validation* dilakukan untuk mencari pemodelan data yang optimal menghasilkan pemodelan data pada *cross* 2 dengan akurasi 93,945%. Hasil pemodelan *k–fold cross–validation* akan menjadi model untuk sampel data pelatihan dan pengujian data untuk menguji KNN dan MKNN untuk klasifikasi. Hasil klasifikasi menghasilkan akurasi berdasarkan aturan *confusion matrix*. Tes menghasilkan akurasi tertinggi KKN sebesar 94,95% dengan akurasi rata–rata selama tes adalah 93,94% dan akurasi tertinggi MKNN adalah 99,51% dengan akurasi rata–rata selama tes adalah 99,20%.

Jika penelitian sebelumnya, hanya membahas penggunaan Spotify dengan 11 fitur audio dengan metode KNN serta membahas metode *k–fold cross–validation*, maka pada penelitian ini penulis ingin menggunakan 12 fitur audio dengan 14 *genre* pada Spotify dengan membandingkan kedua metode algoritma klasifikasi *K–Nearest Neighbor* (KNN) dengan metode algoritma klasifikasi *Modified K–Nearest Neighbor* (MKNN) menggunakan *k–fold cross–validation* untuk membagi sampel data secara acak dan mengelompokkan data tersebut sebanyak nilai 10*–fold* dan mengukur seberapa tepat klasifikasi terhadap kelas menggunakan *confusion matrix* untuk akurasi pada tiap kelas *genre*. Kemudian pada penelitian ini juga akan menggunakan seleksi fitur menggunakan *Principal Component Analysis* (PCA) dengan membandingkan kedua metode algoritma klasifikasi *K–Nearest Neighbor* (KNN) dengan metode algoritma klasifikassi *Modified K–Nearest Neighbor* (MKNN) menggunakan *k–fold cross–validation* dan mengukur seberapa tepat klasifikasi terhadap kelas menggunakan *confusion matrix* dengan akurasi pada tiap kelas *genre*.

# **TINJAUAN PUSTAKA**

## **Lagu dan Musik**

Menurut KBBI (2020), Musik diartikan sebagai: 1) Ilmu dan seni menyusun nada atau suara dalam urutan, kombinasi, dan hubungan temporal untuk menghasilkan komposisi (suara) yang mempunyai kesatuan dan kesinambungan, 2) Nada atau suara yang disusun sedemikian rupa sehingga mengandung irama, lagu, dan keharmonisan (terutama yang menggunakan alat–alat yang dapat menghasilkan bunyi–bunyi itu).

Lagu merupakan ragam suara yang berirama (dalam bercakap, bernyanyi, membaca, dan sebagainya). Selain itu lagu juga sering dikenal dengan suatu syair atau lirik yang mempunyai irama, sebuah lagu biasanya selalu diiringi dengan alat musik untuk menghasilkan musik yang mengandung suara yang berirama. Secara umum lagu memang tidak lepas dengan musik. Pengertian seni musik adalah suatu yang membuahkan hasil karya seni, berupa bunyi berbentuk lagu atau komposisi yang mengungkapkan pikiran serta perasaan penciptanya lewat unsur–unsur pokok musik, yakni irama, melodi, harmoni, tempo, birama, timbre, tangga nada, dinamika serta ekspresi sebagai satu kesatuan susunan lagu (Jamalus, 1988).

Dari sekian banyak lagu yang ada saat ini terbagi lagi menjadi beberapa *genre* atau jenis musik yang sesuai dengan lagu tersebut. *Genre* merupakan pengelompokan musik yang sesuai dengan kemiripan antara satu sama lainnya. Pada umumnya sebuah *genre* musik dapat didefinisikan berdasarkan teknik musik, gaya, konteks, maupun temanya. *Genre* secara umum berarti tipe atau kelas dari musik yang kita dengar. *Genre* musik memberikan ekspektasi bagaimana bunyi musik, berapa lama musik tersebut, dan bagaimana pendengar harus berperilaku. Pada era *Mozart*, ada lima *genre* utama, yaitu *symphony, string quartet, sonata, concerto,* dan opera. Pada era musik modern, musik dapat dibagi menjadi berbagai *genre* seperti *Blues, Classical, Country, Gospel, Hip Hop, Jazz, Metal, Pop, Reggae, Rock, EDM, Funk, R&B,* dan *Soul*. Deskripsi dari *genre*–*genre* tersebut adalah sebagai berikut:

1. Musik *Blues*

Musik *Blues* merupakan sebuah aliran musik yang berasal dari Amerika Serikat. *Blues* berkembang dari musik–musik spiritual dan puji–pujian yang muncul dari komunitas budak–budak Afrika di AS silam. Sebelum abad ke–20, musik *blues* hanya populer di kalangan orang Amerika. Ciri – ciri musik ini yaitu pola irama yang sering digunakan, terkesan sedih dan permainan gitarnya selalu penuh dengan improvisasi.

1. Musik *Classical*

Musik *classical*atau klasik adalah kumpulan dari *genre* musik pada era Mozart dengan bentuk seperti *symphony, string quartet, sonata, concerto,* dan opera. Musik *classical* biasa dimainkan dengan alat–alat musik tertentu seperti biola, beranjak pada abad ke–9 hingga abad ke–21 dan biasanya merujuk pada musik klasik Eropa. Ciri–ciri musik klasik yakni penggunaan dinamika *crescendo* dan *decrescendo*. *Crescendo* yaitu perubahan dinamika lagu dari lembut menjadi keras dan nyaring. Sedangkan *decrescendo* sebaliknya, perubahan dinamikanya dari keras menjadi lembut. Ciri lainnya yaitu perubahan *tempo* nya menggunakan *accelerando* (semakin cepat) dan *ritardando* (semakin lembut).

1. Musik *Country*

Musik *Country* adalah campuran dari unsur–unsur musik Amerika yang berasal dari Amerika Serikat Bagian Selatan dan Pegunungan Appalachia. Musik ini berakar dari lagu rakyat Amerika Utara, musik kelt, musik *gospel*, dan berkembang sejak tahun 1920–an di Amerika bagian selatan, saat itu musik *country* dimainkan memakai gitar, bass, dan steel guitar. Istilah musik *country* mulai dipakai sekitar tahun 1940–an untuk menggantikan istilah musik *hillbilly* yang berkesan merendahkan. Pada tahun 1970–an, istilah musik *country* telah menjadi istilah populer. Istilah lain untuk *genre* musik ini adalah *country* and *western*, namun sudah semakin jarang dipakai kecuali di Britania Raya dan Irlandia. Contoh orang yang menggunakan *Genre* *Country*: Taylor Swift.

1. Musik *Gospel*

Musik *Gospel* merupakan salah satu jenis musik religius yang pertama kali dikenalkan gereja–gereja Afro–Amerika pada tahun 1930–an dan dibawakan oleh kalangan kulit hitam. Dalam perkembangannya jenis musik ini terpecah menjadi dua aliran, *Gospel* kulit hitam dan *Gospel* kulit putih. Hal ini terjadi karena adanya diskriminasi ras pada saat itu di Amerika.

Ciri khas dari *genre* *Gospel* adalah vokal yang dominan dengan harmoni yang kuat dan liriknya yang religius, khususnya Kristen. *Gospel* juga identik dengan paduan suara, koor dan choir. Lambat laun musisi *gospel* mencoba untuk memperkenalkan jenis musik ini ke kancah yang lebih luas. Salah satunya Mahalia Jackson, membawakan *Gospel* yang sudah dikombinasikan dengan *genre* musik lain. Kemudian diikuti oleh Golden Gate Quartet dan Clara War yang lebih berani membawakan tembang–tembang gereja ini di club malam dalam performnya. Hingga kini Musik *Gospel* sudah lagi tidak ada perbedaan yang kentara dengan musik kontemporer.

1. Musik *Hip Hop*

Musik *Hip Hop* merupakan salah satu *genre* musik yang terdiri dari perpaduan *rapping, DJing, Breakdance* dan *Graffiti*. *Hip Hop* mulai ada sekitar tahun 1970–an yang dipelopori oleh kalangan Afro–Amerika dan Amerika Latin. Awalnya *Hip Hop* hanyalah musik dari *Disk Jockey* yang membuat variasi dengan memutar disk sehingga menciptakan bunyi yang unik. Kemudian *Rapping* masuk untuk menghiasi vokal dari bunyi–bunyi tersebut. Sedangkan untuk koreografinya diisi dengan tarian patah–patah yang biasa disebut *breakdance*. Seperti *genre* musik yang lain, *Hip Hop* juga memiliki era dan melahirkan musisi dalam perkembangannya. Diantara musisi *Hip Hop* yang terkenal ialah Nate Dogg, Snoop Dogg and Ludacris.

1. Musik *Jazz*

Musik *jazz* merupakan jenis musik yang tumbuh dari penggabungan *blues, ragtime,* dan musik eropa, terutama musik band. Aliran musik ini berasal dari masyarakat Afro–Amerika Selatan pada akhir abad ke–19 dan awal abad ke–20. Kata *Jazz* berasal dari bahasa slang (bahasa daerah pinggiran pantai barat Amerika Serikat) dan pertama kali dipakai istilah *jazz* pada tahun 1915 di Chicago. Ciri–ciri dari *genre* ini yaitu banyak menggunakan instrumen gitar, trombone, terompet, saksofon dan vokal liriknya cenderung dianggap sebagai bagian dari bunyi instrumen, ritme dan melodinya memiliki cenderung mengimprovisasi nada.

1. Musik *Metal*

Musik *Metal* mengutamakan permainan gitar dan biasanya dimainkan dengan *tempo* lagu yang sangat cepat. Musik *Metal* berawal dari kata Heavy *Metal* yang terdapat dari lagu band Hard *Rock* tahun 1960–an, Steppenwolf. Seiring berkembangnya musik *rock* pada tahun 1970–an, nama *Heavy* *Metal* kemudian dijadikan aliran tersendiri, dengan merujuk pada *blues* *rock* dan *psychedelic* *rock*. *Genre* baru ini bercirikan solo gitar yang panjang, distorsi gitar yang kuat dan ketukan yang cepat di semua instrumen alat musiknya.

*Heavy metal* pada era ini melahirkan band–band besar seperti Led Zeppelin, Black Sabbath dan Deep Purple. Namun elemen *blues* masih terlihat kental mempengaruhi di era ini. Selanjutnya Judas Priest mengimprovisasi *genre* ini dengan meniadakan unsur *blues* nya, sehingga hanya mengandalkan distorsi, beat yang lebih cepat dan harmoni.

Perkembangan demi perkembangan terus terjadi dalam sejarah salah satu jenis musik ini sampai akhirnya menjadikannya terbagi menjadi beberapa sub–*genre*, diantaranya *Deathcore, Metalcore, Progressive Metal, Mathcore, Nu Metal* atau *Groove Metal, Glam Metal, Death Metal, Doom Metal* atau *Ghotic Metal, Black Metal, Trash Metal, Folk Metal, Grindcore, Alternative Metal* dan *Power Metal*.

1. Musik *Pop*

Musik *pop* atau musik populer merupakan jenis–jenis musik yang saat ini digemari oleh masyarakat awam. *Genre* ini mulai berkembang dan populer pada tahun 1950–an. *Genre* dinamakan populer karena memiliki daya tarik masa yang lebih besar dibandingkan dengan *genre* lainnya. Ciri– ciri musik *pop* yaitu melodi yang digunakan mudah diterapkan dengan berbagai karakter lirik, fleksibel dan mudah dipadukan dengan jenis lain, lagu mudah disenandungkan dan mudah dipahami, harmoni tidak rumit, *tempo* nya dapat divariasikan.

1. Musik *Reggae*

Musik *Reggae* mempunyai ciri khas dengan ritme backbeat dan progresif kord sederhana. *Genre* musik ini berasal dari Jamaika dan mirip dengan gaya musik ska dan *rock*steady. Walaupun unsur utama *Reggae* dari *ska* dan *rocksteady*, salah satu jenis musik ini juga mengadopsi *jazz*, *R&B* dan Jamaican mento. Amerika Serikat dan Jamaika mengakui *Reggae* sebagai aliran musik tersendiri pada tahun 1960–an. Dahulu musik reggae sering dianggap dengan jiwa muda, *rude boys* atau anak laki–laki kasar dan pemberontakan. Dan kita semua tahu, *genre* musik ini melahirkan beberapa musisi hebat seperti Bob Marley, Prince Buster, Desmond Dekker dan Jackie Mittoo sebagai legendanya meskipun sebetulnya masih banyak lagi yang lain.

1. Musik *Rock*

Lahirnya musik *Rock* berawal dari kombinasi dari *genre* musik *rhythm* dan *blues*, *country*, *jazz*, musik klasik termasuk musik rakyat (*folk music*). Jenis musik ini mulai dikenal secara umum pada pertengahan tahun 1950–an. *Rock* mempunyai ciri khas yang didominasi oleh vokal, gitar, drum, keyboard maupun *synthesizer*. Musik *Rock* mirip dengan metal, mengutamakan permainan gitar tetapi memiliki *tempo* lagu yang lebih lambat dari *genre* metal. Dalam beberapa dekade saja *genre* musik ini sudah menelurkan sub–*genre* baru seperti *Alternative Rock, Art Rock, Blues rock, Britpop, Country rock, Emo, Folk rock, Garage rock, Glam rock, Grunge, Hard rock, Heavy Metal, Indie rock,* [*Jazz*](https://www.onoini.com/pengertian-musik-jazz/)*rock, New wave, Pop rock, Post britpop, Post grunge, Post punk, Progressive rock, Psychedelic rock, Punk rock, Rock and roll, Rockabily, Soft rock, Southern rock* dan *Surf rock*. Dalam perkembangannya hingga tersebar ke seluruh dunia, musik *Rock* juga melahirkan beberapa band yang bisa dibilang legendaris seperti The Beatles, Dream Theater dan lainnya.

1. Musik *Electronic Dance Music* (EDM)

Musik *Electronic* atau *Techno* atau dikenal dengan *Electronic Dance Music* (EDM) mulai ada sekitar tahun 1980 di Detroit, awal mula musik ini dibentuk sebagai bentuk tari dan musik pesta. Seperti namanya, *genre* musik ini mengedepankan ragam suara dari teknologi dalam musiknya. Sebagian besar musik *Techno* merupakan kombinasi *synthesizer*, hentakan drum dan sequencer. Biasanya musik *techno* menghiasi disetiap club malam dunia. Dan beberapa musisi yang terkenal dalam *genre* musik ini adalah Carl Cox dan Richie Hawtin. Dalam beberapa dekade saja *genre* musik ini sudah menelurkan sub–*genre* baru seperti *Electro House*, *Progressive House*, *Big Room House*, *Melbourne Bounce*, *Deep House*, *Future House*, *Trance*, *Tropical House*, *Grime*, *Dubstep*, *Brostep* dan lain–lainnya.

1. Musik *Funk*

Musik *Funk* merupakan jenis musik yang berasal dari kolaborasi *genre* *rhythm* & *blues*, *jazz* dan soul musik yang dijadikan sebuah ritmik. Jenis seni musik ini terlahir pada tahun 1960–an ketika para musisi Afro–Amerika mencoba berimprovisasi. Musik *Funk* dapat dikenali dari bunyi gitarnya yang kuat, drum yang dominan, alunan nada yang terpotong singkat dan memberikan kesan dansa dan gembira ketika mendengarnya.

1. Musik *Rhythm* and *Blues* (R&B)

Musik *R&B* atau *Rhythm* dan *Blues* merupakan gabungan dari *genre* *jazz*, *blues* dan *gospel*. Pada awalnya musik ini hanya ditujukan bagi pendengar kulit hitam. Nama *Rhythm* dan *Blues* sendiri dibuat pada akhir tahun 1940–an di Amerika Serikat sebagai istilah pemasaran agar tidak dianggap rasis.

Periode awal musik *R&B* berfokus pada ritme boogie dengan musisi ternama seperti Bo Diddley dan Chuck Berry. Sementara jenis musik ini berkembang selama bertahun–tahun sampai akhirnya menumbuhkan beberapa *genre* baru. Musik *R&B* kontemporer mempunyai susunan tersendiri yang berbeda dengan pendahulunya, dengan beralih fokus pada unsur *pop* (walaupun elemen *jazz*, *gospel* dan *blues* terkadang masih tetap muncul). Dari waktu ke waktu, kesan *genre* *R&B* sebagai musik kalangan kulit hitam semakin memudar dengan menyebarnya musik ini ke seluruh dunia.

1. Musik *Soul*

Musik *Soul* adalah salah satu *genre* musik populer yang berasal dari Amerika Serikat. Aliran musik ini muncul dan berkembang pada akhir 1950 dan awal 1960. Jenis musik ini menggabungkan dua aliran musik yang telah lahir sebelumnya, yakni *Rythm* and *blues* serta musik *gospel* Afrika Amerika. Menurut *Rock* *and Roll Hall of Fame*, musik *Soul* adalah jenis musik yang muncul melalui transformasi musik *gospel* dan *R&B* menjadi satu bentuk yang lebih *funky* serta irama yang *catchy*. Sehingga irama yang dihasilkan mampu membuat tubuh bergerak, mengalir mengikuti alunan musik. Hal ini yang kemudian menjadi satu hal penting dalam musik *soul*.

## **Spotify**

Spotify merupakan layanan *streaming music* asal Swedia yang menyediakan hak digital manajemen yang dilindungi konten dari label rekaman dan perusahaan media. Musik yang ada pada Spotify dapat diakses atau dicari berdasarkan *artist*, album, *genre*, *playlist*, *podcasts*, atau label rekaman. Selain itu, Spotify juga berperan sebagai penyedia hak cipta digital atas musik yang diunggah, sehingga lagu–lagu yang disajikan merupakan lagu resmi dan pengguna tentunya dapat menikmati koleksi yang beragam tersebut secara resmi pula. Hal yang cukup menarik dari Spotify adalah ia menawarkan dua model berlangganan, yaitu gratis dan *premium* (Spotify, 2020).

Sebagai layanan *streaming music*, Spotify memiliki banyak fitur yang beragam guna memberikan pelayanan yang terbaik bagi penggunanya. Berikut beberapa fitur–fitur pengguna pada atribut lagu tersebut.

1. Katalog musik

Beragamnya musik yang disediakan oleh suatu layanan *streaming music* tentu menjadi daya tarik tersendiri bagi penggunanya. Pada tahun 2012 saja, Spotify mengklaim mempunyai 20 juta koleksi lagu dari berbagai *genre* musik di seluruh dunia.

1. *Playlist*

*Playlist* yang disediakan oleh aplikasi resmi Spotify sangat banyak, selain itu pengguna juga bisa membuat *playlist* musik sendiri sesuai dengan musik yang disukai, sehingga lagu–lagu yang disukai dapat diputar dengan mudah tanpa harus mencari–cari lagi.

1. Radio

Spotify juga menyediakan akses ke beberapa stasiun radio populer dunia, disini pengguna dapat memilih sendiri *genre* musik yang digemari.

1. Mode *Offline*

Bagi pengguna *premium,* fitur ini dapat digunakan tetapi tidak untuk pengguna gratis. Fitur ini dapat memutar musik yang sudah diunduh sebelumnya tanpa harus terhubung ke internet terlebih dahulu.

1. Integrasi Media Sosial

Spotify memungkinkan akun penggunanya untuk terintegrasi dengan akun media sosial pengguna tersebut yaitu Facebook dan Twitter. Setelah terhubung mereka dapat mengakses musik favorit dan juga dapat mengikut *playlist* dari teman–temannya.

## **Transformasi Data**

Menurut Supriyadi (2018), proses transformasi dilakukan untuk menskalakan atribut numerik dalam *range* yang lebih kecil, seperti –1.0 sampai 1.0 atau 0.0 sampai 1.0. Ada banyak metode transformasi, salah satunya *min–max normalization*. *Min–max normalization* merupakan metode normalisasi dengan melakukan transformasi linier terhadap data asli, yang dituliskan dalam pada Persamaan (2.1).

....(2.1)

Keterangan:

*V’* = Nilai yang dicari.

*V* = Nilai pada variabel A.

𝑀𝑖𝑛𝐴 = Nilai minimal variabel A.

𝑀𝑎𝑥𝐴 = Nilai maksimal variabel A.

𝑁𝑒𝑤\_𝑚𝑖𝑛𝐴 = Rentang nilai minimal pada variabel A.

𝑁𝑒𝑤\_𝑚𝑎𝑥𝐴 = Rentang nilai maksimal pada variabel A.

## **Klasifikasi**

Klasifikasi adalah salah satu tugas dari *data* *mining* yang bertujuan untuk memprediksi label kategori benda yang tidak diketahui sebelumnya, dalam membedakan antara objek yang satu dengan yang lainnya berdasarkan atribut atau fitur (Mutrofin, dkk 2014).

Di dalam klasifikasi terdapat dua pekerjaan yang dilakukan, (Prasetya, 2012) yaitu sebagai berikut:

* Pembangunan model sebagai *prototype* untuk disimpan sebagai memori.
* Penggunaan model tersebut untuk melakukan pengenalan klasifikasi atau prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya.

Terdapat beberapa banyak algoritma klasifikasi yang sudah dikembangkan oleh para peneliti diantaranya adalah *K–Nearest Neighbor* (KNN), *Modified K–Nearest Neighbor* (MKNN), *Decision Tree Classifier*, *Rule–Based Classifier*, *Neural–Network, Artificial Neural Network* (ANN), *Support Vector Machine*, dan *Naive Bayes Classifier*.

## **Principal Component Analysis**

Menurut Jolliffe (2002), *Principal Component Analysis* (PCA) merupakan metode untuk mengurangi dimensi kumpulan data yang terdiri dari sejumlah besar variabel yang saling terkait, sambil mempertahankan sebanyak mungkin variasi yang ada dalam kumpulan data. Tujuan dari pca adalah mentransformasikan ke satu set variabel baru, mengecek korelasi antar variabel data dan mempertahankan sebagian besar variasi yang ada di semua variable asli. Dengan melihat korelasi antar variabel tersebut akan diperoleh faktor*–*faktor yang mempengaruhi pada penelitian ini. Langkah*–*langkah PCA digunakan sebagai berikut (Dananjaya, 2019).

1. Menghitung *variance*

Menghitung *variance* menggunakan rumus pada Persamaan (2.2).

................. (2.2)

1. Menghitung *covariance*

Menghitung *covariance* menggunakan rumus pada Persamaan (2.3). Setelah itu dihasilkan *covariance matrix*.

................. (2.3)

1. Menghitung *eigenvalue* dan *eigenvectors*

Menghitung *eigenvalue* dan *eigenvector* pada Persamaan (2.4). Sebelum menghitung *eigenvalue* dan *eigenvector*, nilai lamda harus dihitung terlebih dahulu pada Persamaan (2.5).

................. (2.4)

................. (2.5)

1. Mengurutkan *eigenvalue* dan *eigenvectors*, setelah itu tentukan *principical component* (komponen utama).
2. Menghitung korelasi antara variabel utama dengan *principal component* (komponen utama)

Setelah menentukan komponen utama, selanjutnya mengkorelasikan komponen utama dengan variabel utama akan dihitung pada Persamaan (2.6). Mengurangi variabel dengan menghilangkan komponen rendah.

................. (2.6)

Keterangan:

: *Matrix* n x n.

: *Eigenvalue*.

: *Identity Matrix*.

: Korelasi antara variabel utama dengan komponen utama.

: *eigenvectors*.

: *Covariance Matrix*.

## **K–Fold Cross Validation**

*Cross–validation* (CV) adalah metode statistik yang dapat digunakan untuk mengevaluasi kinerja model atau algoritma dengan membagi sampel data secara acak dan mengelompokkan data menjadi dua *subset* yaitu data proses pembelajaran dan data validasiatauevaluasi. Model atau algoritma dilatih oleh *subset* pembelajaran dan divalidasi oleh *subset* validasi. *K–Fold Cross–Validation* adalah kasus khusus dari *cross–validation* di mana mengulangi set *dataset* *k* kali. Dalam setiap putaran, *dataset* akan dibagi menjadi bagian *k* satu bagian digunakan untuk validasi, dan bagian yang tersisa digabung ke dalam *subset* *training* (data latih) untuk evaluasi. *Ten–fold cross****–****validation* adalah salah satu *k–fold cross–validation* yang direkomendasikan untuk pemilihan model terbaik (Wibowo, 2017).

## **Confusion Matrix**

*Confusion Matrix* adalah tabel yang sering digunakan untuk menggambarkan kinerja model pada klasifikasi (atau "*classifier*") pada serangkaian data *testing* (data uji) yang nilai sebenarnya diketahui (Nomleni, 2015). Tabel 2.1 menunjukkan *confusion matrix* untuk *classifier*.

Tabel 2.1. Confusion Matrix

|  |  |  |
| --- | --- | --- |
| Data Prediksi | Data Aktual | |
| Ya | Tidak |
| Ya | TP | FN |
| Tidak | FP | TN |
| Total | P | N |

Pada penelitian ini, entri dalam *confusion matrix* memiliki arti seperti berikut ini:

* + TP adalah *true positive*, merupakan jumlah data prediksi yang benar diklasifikasikan *positif*.
  + TN adalah *true negatif*, merupakan jumlah data prediksi yang benar diklasifikasikan *negatif*.
  + FN adalah *false negatif*, merupakan jumlah data prediksi yang salah diklasifikasikan *negatif*.
  + FP adalah *false positif*, merupakan jumlah data prediksi yang salah diklasifikasikan *positif*.

Pada penelitian ini ukuran kinerja dari sistem klasifikasi yaitu akurasi pada Persamaan (2.7) untuk mengukur seberapa tepat klasifikasi terhadap kelas. Akurasi adalah rasio untuk mengukur kedetakan antara prediksi benar (*positif*) dengan aktual (sebenarnya) dari keseluruhan data.

............................ (2.7)

## **K–Nearest Neighbor**

Metode Algoritma *K–Nearest Neighbor* (KNN) yaitu algoritma yang bertujuan untuk klasifikasi data. Secara umum, metode ini digunakan untuk menentukan kelas data dari suatu data dengan mengukur jarak antar tetangganya (Widiantara, 2018). Jarak yang dipilih adalah jarak terpendek dari data *testing* (data uji). Perbandingan jarak antara data *testing* (data uji) dan data *training* (data latih) dilakukan sebanyak *k* buah. Rumus perhitungan jarak tersebut dapat dilakukan dengan *euclidean distance* dituliskan pada Persamaan (2.8).

..... (2.8)

Keterangan:

*Euclidean Distance()* : Jarak antara data latih dengan data uji.

: Atribut fitur data latih.

: Atribut fitur data uji.

*i* : Mempresentasikan index nilai atribut fitur.

*n*  : Jumlah panjang dimensi atribut fitur.