

PROPOSAL TUGAS AKHIR

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



I MADE TANGKAS WAHYU KENCANA YUDA

NIM. 1608561031

PROGRAM STUDI INFORMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS UDAYANA

JIMBARAN

2020

LEMBAR PERSETUJUAN PROPOSAL TUGAS AKHIR

Judul : Klasifikasi Musik Berdasarkan Genre pada Layanan Streaming Musik
Spotify Menggunakan algoritma *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor*

Nama : I Made Tangkas Wahyu Kencana Yuda

NIM : 1608561031

Tanggal Seminar : 18 Mei 2020

Disetujui oleh :

Luh Arida Ayu Rahning Putri, S.Kom., M.Cs. Ketua Penguji
NIP. 198209182008122002

I Wayan Supriana, S.Si., M.Cs. Penguji 1
NIP. 1984082920181113001

I Gede Santi Astawa, S.T., M.Cs. Penguji 2
NIP. 198012062006041003

I Komang Ari Mogi, S.Kom., M.Kom. Penguji 3
NIP. 198409242008011007

Dra. Luh Gede Astuti, M.Kom. Penguji 4
NIP. 196401141994022001

Mengetahui,
Komisi Seminar dan Tugas Akhir
Program Studi Informatika
FMIPA UNUD
Ketua,

I Gusti Ngurah Anom Cahyadi Putra, ST., M.Cs
NIP. 198403172019031005

KATA PENGANTAR

Proposal penelitian dengan judul Klasifikasi Musik Berdasarkan Genre pada Layanan Streaming Musik Spotify Menggunakan algoritma *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor* ini disusun dalam rangkaian kegiatan pelaksanaan Tugas Akhir di Jurusan Ilmu Komputer FMIPA UNUD. Proposal ini disusun dengan harapan dapat menjadi pedoman dan arahan dalam melaksanakan penelitian di atas.

Sehubungan dengan telah diselesaikannya proposal ini, maka diucapkan terima kasih dan penghargaan kepada berbagai pihak yang telah membantu pengusul, antara lain :

1. Bapak I Wayan Supriana, S.Si., M.Cs. sebagai calon Pembimbing I yang telah banyak membantu menyempurnakan proposal ini.
2. Bapak I Gede Santi Astawa, S.T., M.Cs. sebagai calon Pembimbing II yang telah bersedia mengkritisi, memeriksa dan menyempurnakan proposal ini.
4. Bapak dan Ibu dosen pengajar di Jurusan Ilmu Komputer Fakultas MIPA Universitas Udayana yang telah meluangkan waktu untuk memberikan saran dan masukan dalam menyempurnakan penelitian ini.
5. Kawan-kawan di Jurusan Ilmu Komputer yang telah memberikan dukungan moral dalam penyelesaian proposal ini.

Disadari pula bahwa sudah tentu proposal ini masih mengandung kelemahan dan kekurangan. Memperhatikan hal ini, maka masukan dan saran-saran penyempurnaan sangat diharapkan.

Jimbaran, 30 April 2020

Penulis

DAFTAR ISI

LEMBAR JUDUL.....	i
LEMBAR PERSETUJUAN PROPOSAL TUGAS AKHIR	ii
KATA PENGANTAR	iii
DAFTAR ISI.....	iv
DAFTAR TABEL.....	vi
DAFTAR GAMBAR	vii
1. Latar Belakang.....	8
2. Rumusan Masalah.....	10
3. Tujuan Penelitian	10
4. Batasan Masalah	11
5. Manfaat Penelitian	11
6. Tinjauan Pustaka.....	12
6.1. Kajian Terkait.....	12
6.2. Kajian Teoritis	13
6.2.1. Lagu dan Musik.....	13
6.2.2. Spotify	20
6.2.3. Identifikasi Fitur Audio Spotify	22
6.2.4. Transformasi Data.....	25
6.2.5. Klasifikasi	25
6.2.6. Confusion Matrix	26
6.2.7. K–Fold Cross–Validation.....	27
6.2.8. K–Nearest Neighbor.....	27

6.2.9. Modified K–Nearest Neighbor.....	28
7. Metodologi Penelitian.....	30
7.1. Sumber Data	30
7.2. Pengumpulan Data.....	30
7.3. Variabel dan Definisi Variabel.....	31
7.4. Alur Penelitian.....	33
7.5. Proses Klasifikasi KNN.....	35
7.6. Proses Klasifikasi MKNN	37
7.7. Skenario Pengujian dan Evaluasi	39
8. Jadwal Pelaksanaan Penelitian.....	41
DAFTAR PUSTAKA	43

DAFTAR TABEL

Tabel 6.1. Confusion Matrix	26
Tabel 7.1. Variabel Fitur Pengguna	31
Tabel 7.2. Variabel Fitur Audio	32
Tabel 8.1. Jadwal Pelaksanaan Kegiatan	41

DAFTAR GAMBAR

Gambar 1.1. Keunggulan Spotify (Iriansyah, 2018)	8
Gambar 7.2. Playlist Genre	31
Gambar 7.3. Alur Penelitian.....	33
Gambar 7.4. Alir Proses klasifikasi KNN	35
Gambar 7.5. Alir Proses Klasifikasi MKNN	37
Gambar 7.6. K–Fold Cross–Validation.....	40

1. Latar Belakang

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 atribut 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 musik 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.

Perbandingan Layanan Streaming Musik				
Nama Layanan	Spotify	Apple Music	JOOX	DEEZER
Asal Negara	 Swedia	 Amerika	 Prancis	 Cina
Jumlah Pengguna	191+ juta pengguna (87+ juta bayar) *Okt 18	50+ juta pengguna berbayar *Mei 18	14+ juta pengguna *Nov 18	Tidak ada data
Jumlah Lagu	35+ juta lagu *Okt 18	50+ juta lagu *Nov 18	53+ juta lagu *Nov 18	Tidak ada data
Kualitas Suara	Up to 320kbps	Up to 256kbps	Up to 1.411kbps (FLAC)	Up to 320kbps
Video	Ada video	Ada video	Tak ada video	Ada video
Mode Offline	Ada	Ada	Ada	Ada
Verat Gratis	Ada	Tidak ada	Ada	Ada
Periode Trial	Ada 1 bulan	Ada 3 bulan	Ada 1 bulan	Ada 1 bulan
Family Plan	Ada	Ada	Tidak Ada	Tidak Ada
Link Lagu	Lirik hanya di beberapa lagu	Lirik hanya di beberapa lagu	Lirik hanya di beberapa lagu	Lirik hanya di beberapa lagu

Gambar 1.1. 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, acoustic–ness, key*, dan *liveliness*). 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 KNN 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 7 dan 11 fitur audio dengan metode KNN serta membahas analisis perbandingan *k-fold cross-validation* pada metode algoritma klasifikasi KNN dengan MKNN menggunakan *confusion matrix*, maka penelitian ini penulis ingin menggunakan 12 fitur audio dengan Spotify serta ingin mengetahui pengaruh atribut lagu yang digunakan terhadap kinerja untuk mengklasifikasi lagu menggunakan KNN dan MKNN serta analisis perbandingan performa *k-fold cross-validation* yang dihasilkan dari algoritma klasifikasi KNN dengan MKNN menggunakan *confusion matrix* pada klasifikasi musik.

2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan permasalahan dari penelitian ini, yaitu sebagai berikut.

- a. Bagaimana pengaruh atribut lagu yang digunakan terhadap kinerja untuk mengklasifikasi lagu menggunakan KNN dan MKNN?
- b. Bagaimana perbandingan performa yang dihasilkan algoritma KNN dan MKNN pada klasifikasi musik?

3. Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, tujuan dari penelitian ini adalah sebagai berikut.

- a. Untuk mengetahui pengaruh atribut lagu yang digunakan terhadap kinerja untuk mengklasifikasi lagu menggunakan KNN dan MKNN.

- b. Untuk mengetahui perbandingan performa yang dihasilkan algoritma KNN dan MKNN pada klasifikasi musik.

4. Batasan Masalah

Beberapa batasan masalah yang akan dijadikan acuan dalam penelitian ini adalah sebagai berikut.

- a. Data yang digunakan adalah lagu dengan 14 genre musik yang ada di layanan streaming musik Spotify yaitu *Blues, Classical, Country, Gospel, Hip Hop, Jazz, Metal, Pop, Reggae, Rock, EDM, Funk, R&B, dan Soul*.
- b. Variabel fitur pada atribut lagu yang digunakan adalah 12 fitur, antara lain *danceability, valence, energy, tempo, loudness, speechiness, instrumentalness, liveness, acousticness, key, mode, dan duration*.
- c. Perbandingan metode yang digunakan adalah metode algoritma *K-Nearest Neighbor* dan *Modified K-Nearest Neighbor*.

5. Manfaat Penelitian

Beberapa manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut.

- a. Bagi Masyarakat
Penelitian ini dapat membantu untuk mendapatkan informasi tentang pengaruh atribut lagu yang digunakan terhadap kinerja untuk mengklasifikasi lagu menggunakan KNN.
- b. Bagi Penulis
Penelitian ini diharapkan dapat menambah pengetahuan bagi penulis mengenai perbandingan performa algoritma KNN dan MKNN pada klasifikasi musik diajukan.
- c. Bagi Keilmuan
Penelitian ini dapat menjadi referensi perbandingan performa algoritma KNN dan MKNN pada penelitian lain yang memiliki karakteristik yang sama.

6. Tinjauan Pustaka

6.1. Kajian Terkait

Penelitian tugas akhir oleh Supriyadi (2018) menggunakan lagu dengan genre musik pop dan klasik pada layanan streaming musik Spotify dengan mengimplementasikan Artificial Neural Network (ANN) metode *Backpropagation* guna untuk melakukan pengklasifikasian genre musik pop dan klasik. Data input 7 fitur audio yang digunakan adalah *acousticness*, *danceability*, *energy*, *loudness*, *speechiness*, *tempo* dan *valence* kemudian outputnya adalah genre musik pop dan klasik. Dari hasil implementasi ANN dengan metode backpropagation digunakan arsitektur jaringan saraf tiruan 7 neuron input, 1 hidden layer dengan 4 neuron dan 1 output. Hasil akurasi dari pengujian diperoleh sebesar 99,5%. Baik pada penelitian Supriyadi (2018) maupun penelitian penulis sama-sama melakukan klasifikasi musik pada layanan *streaming* Spotify. Perbedaanannya, jika pada penelitian Supriyadi (2018) menggunakan algoritma Backpropagation, maka pada penelitian ini penulis menggunakan algoritma KNN dan MKNN.

Pada jurnal oleh Ravi, dkk (2019), dilakukan identifikasi jenis penyakit gigi dan mulut yang ditentukan dari gejala yang dialami dengan menggunakan metode klasifikasi *Modified K-Nearest Neighbor* (MKNN) menggunakan 6 kelas data yang meliputi Pulpitis, Gingivitis, Karies Gigi, Periodontitis, Deposits, dan Nekrosis Pulpa. Pada penelitian ini membuktikan bahwa pada data latih sebanyak 70 dan data uji 30 serta nilai $K=60$, metode MKNN dapat melakukan identifikasi jenis penyakit gigi dan mulut dengan mencapai 86,6%. Pada penelitian ini juga membuktikan bahwa metode MKNN cenderung lebih tinggi akurasinya dibandingkan dengan metode KNN dimana metode MKNN memiliki tingkat akurasi 76,66% sedangkan KNN 43,33%. Baik pada penelitian Ravi, dkk (2018) maupun penelitian penulis sama-sama menggunakan metode klasifikasi *Modified K-Nearest*

Neighbor (MKNN). Perbedaannya, jika pada penelitian Supriyadi (2018) melakukan klasifikasi identifikasi jenis penyakit gigi dan mulut, maka pada penelitian ini penulis melakukan klasifikasi musik pada layanan *streaming* Spotify.

Pada jurnal Parvin, dkk (2010), dilakukan pengujian untuk membandingkan *K-Nearest Neighbor* (KNN) dengan *Modified K-Nearest Neighbor* (MKNN) dengan empat dataset yaitu (a) Iris (b) Balance scale (c) Bupa (d) SAHeart dan nilai K yang berbeda. Hasil dari penelitian adalah akurasi K = 3, 5 dan 7. Metode *Modified K-Nearest Neighbor* (MKNN) lebih baik dari *K-Nearest Neighbor* (KNN). Penelitian Parvin, dkk (2010) menjadi latar belakang penulis dalam memilih metode *Modified K-Nearest Neighbor* (MKNN) sebagai metode dalam melakukan klasifikasi musik pada layanan *streaming* Spotify.

6.2. Kajian Teoritis

6.2.1. 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, serta bentuk atau susunan lagu dan ekspresi sebagai satu kesatuan (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 music 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

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.

2. Musik Classical

Musik classical/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 temponya menggunakan accelerando (semakin cepat) dan ritardando (semakin lembut).

3. Musik Country

Genre 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.

4. Musik Gospel

Gospel merupakan salah satu jenis musik religius yang pertama kali dikenalkan gereja-gereja Afro-Amerika pada tahun 1930-an dan dibawa 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 Ward 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.

5. Musik Hip Hop

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.

6. Musik Jazz

Genre 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.

7. 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 Nu Metal, Glam Metal, Death Metal, Doom Metal, Black Metal, Trash Metal, Folk Metal dan Power Metal.

8. 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, temponya dapat divariasikan.

9. 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 rocksteady. 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.

10. Musik Rock

Lahirnya musik Rock berawal dari kombinasi dari genre musik rhythm and 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 soft rock, glam rock, hard rock, progressive rock, punk rock dan alternative 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.

11. Musik EDM

Musik Electronic/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 di setiap club malam dunia. Dan beberapa musisi yang terkenal dalam genre musik ini adalah Carl Cox dan Richie Hawtin.

12. Musik Funk

Funk merupakan jenis musik yang berasal dari kolaborasi genre rhythm & blues, jazz dan soul music 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.

13. 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.

14. 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 Rhythm 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 RnB 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.

6.2.2. Spotify

Spotify merupakan layanan streaming musik 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 musik, 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 musik 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.

2. *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.

3. Radio

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

4. 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.

5. 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.

6.2.3. Identifikasi Fitur Audio Spotify

Untuk setiap lagu di platform mereka, Spotify menyediakan data untuk 13 fitur audio pada atribut lagu. Panduan pengembang API Web Spotify mendefinisikannya sebagai berikut (Ashrith, 2018) :

1. *Danceability*

Menjelaskan seberapa kecocokan sebuah lagu atau trek musik untuk menari berdasarkan kombinasi elemen musik termasuk tempo, stabilitas ritme, kekuatan beat, dan keteraturan keseluruhan. Nilai danceability antara 0,0 sampai dengan 1,0, semakin mendekati 1.0 maka trek pada musik tersebut semakin cocok untuk menari begitu pula sebaliknya.

2. *Valence*

Mengindikasikan jenis emosi. Lagu dengan valence tinggi terdengar lebih positif (mis. Bahagia, ceria, gembira), sedangkan trek dengan valence rendah terdengar lebih negatif (misal sedih, tertekan, marah). Hal ini juga diukur pada skala 0.0 sampai 1.0, dengan nilai yang lebih rendah yang mewakili emosi negatif dan nilai yang lebih tinggi mewakili emosi positif.

3. *Energy*

Merupakan ukuran intensitas dan aktivitas suatu tingkat dari kemampuan suatu musik untuk meningkatkan emosi dari pendengarnya. Maksudnya ada suatu musik yang memiliki energi meningkat dan ada juga yang menurun. Biasanya kontrol pada energi musik dapat dilakukan dengan meningkatkan jumlah instrument, mengatur dinamika, maupun

irama pada musik tersebut. Semakin energik suatu musik, maka nilai energi akan semakin besar, begitu pula sebaliknya, semakin lembut suatu musik, maka semakin kecil nilai energinya. Sebagai contoh, death metal memiliki energi tinggi, sedangkan prelude Bach mendapat skor rendah pada skala.

4. *Tempo*

Tempo merupakan cepat atau lambatnya sebuah lagu dan ukuran dari sebuah tempo adalah *beat*. *Beat* sendiri dapat diartikan sebagai ketukan dasar yang menunjukkan banyaknya ketukan dalam satu menit *Beats Per Minute* (BPM).

5. *Loudness*

Keseluruhan kenyaringan trek dalam desibel (dB). *Loudness* dihitung dalam desibel dengan mengukur intensitas gelombang audio selama durasi lagu. Lagu yang lebih keras cenderung lebih energik atau agresif, sedangkan lagu yang lembut cenderung menggunakan instrumen yang lembut dan menggambarkan emosi yang lebih tenang. Nilai tipikal berkisar antara -60 dan 0 db.

6. *Speechiness*

Mendeteksi keberadaan kata-kata yang diucapkan di dalam sebuah trek lagu, seperti eksklusif pidato atau rekaman (misal acara bincang-bincang, buku audio, puisi). Nilai di atas 0,66 menggambarkan trek yang mungkin seluruhnya terbuat dari kata-kata yang diucapkan. Nilai antara 0,33 dan 0,66 menggambarkan trek yang mungkin berisi musik dan ucapan baik dalam bagian atau berlapis, termasuk kasus seperti musik rap. Nilai di bawah 0,33 kemungkinan besar mewakili musik dan trek *nonspeech*.

7. *Instrumentalness*

Memprediksi apakah suatu lagu mengandung *vocal* atau tidak. Suara “Ooh” dan “Aah” diperlakukan sebagai instrumen dalam konteks ini. Rap atau trek kata yang diucapkan jelas

"vokal". Nilai instrumentalness mempunyai range dari 0,0 hingga 1,0. Semakin dekat ke nilai 1,0, semakin tinggi kecenderungan musik tersebut tidak disertai dengan vokal.

8. *Liveness*

Mendeteksi keberadaan *audiens* dalam rekaman. Nilai *liveness* yang lebih tinggi mewakili probabilitas yang meningkat bahwa trek dilakukan secara langsung.

9. *Acousticness*

Suatu ukuran yang menunjukkan apakah trek tersebut akustik atau tidak. Ukuran kepercayaan dari 0,0 hingga 1,0 dari apakah trek akustik. Semakin mendekati nilai 1.0 pada trek tersebut menunjukkan semakin tinggi kepercayaan lintasannya akustik, begitu pula sebaliknya apabila semakin mendekati 0.0 maka semakin rendah kepercayaan lintasnya akustik.

10. *Key*

Taksiran keseluruhan kunci trek. *Integer* memetakan ke pitch menggunakan notasi Pitch Class standar. Misalnya 0 = C, 1 = C \sharp /D \flat , 2 = D, dan seterusnya.

11. *Mode*

Mengindikasikan modalitas (besar atau kecil) dari suatu trek, jenis skala dari mana konten melodinya diturunkan. Mayor diwakili oleh 1 dan minor adalah 0.

12. *Duration*

Durasi trek musik dalam milidetik. Nilai *duration* mempunyai *range* dari 0,0 hingga 1,0. Semakin dekat ke nilai 1,0, semakin panjang durasi trek musik.

13. *Time Signature*

Perkiraan *Time Signature* keseluruhan dari sebuah lagu. *Time Signature* (meter) adalah konvensi notasi untuk menentukan berapa banyak ketukan di setiap bar (atau ukuran).

Penelitian ini akan menggunakan variabel fitur audio pada atribut lagu spotify yang digunakan adalah 12 fitur antara lain

danceability, valence, energy, tempo, loudness, speechiness, instrumentalness, liveness, acousticness, key, mode dan *duration* yang diambil dengan menggunakan Spotify API.

6.2.4. 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 (2.1).

$$V' = \frac{V - \text{Min}_A}{\text{Max}_A - \text{Min}_A} (\text{New_max}_A - \text{New_min}_A) + \text{New_min}_A \quad \text{.. (2.1)}$$

Keterangan :

V' = Nilai yang dicari

V = Nilai pada variabel A

Min_A = Nilai minimal variabel A

Max_A = Nilai maksimal variabel A

New_min_A = Rentang nilai minimal pada variabel A

New_max_A = Rentang nilai maksimal pada variabel A

6.2.5. 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.

1. Pembangunan model sebagai *prototype* untuk disimpan sebagai memori.
2. 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*.

6.2.6. Confusion Matrix

Confusion Matrix adalah tabel yang sering digunakan untuk menggambarkan kinerja model pada klasifikasi (atau "*classifier*") pada serangkaian data uji yang nilai sebenarnya diketahui (Nomleni, 2015). Tabel 6.1 menunjukkan *confusion matrix* untuk *classifier*.

Tabel 6.1. *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.

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

- FP adalah *false positive*, merupakan jumlah data prediksi yang salah diklasifikasikan positif.

Pada penelitian ini ukuran kinerja dari sistem klasifikasi yaitu akurasi yang merupakan persentase dari total data yang diidentifikasi dan dinilai. Berdasarkan *confusion matrix*, kemudian dihitung nilai akurasi seperti pada (2.2).

$$Akurasi = \frac{(TP + TN)}{(TP + TN + FN + FP)} \times 100 \quad (2.2)$$

6.2.7. K-Fold Cross-Validation

Cross-validation (CV) adalah metode statistik yang dapat digunakan untuk mengevaluasi kinerja model atau algoritma dimana data dipisahkan menjadi dua *subset* yaitu data proses pembelajaran dan data validasi atau evaluasi. 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* pelatihan untuk evaluasi. *Ten-fold cross-validation* adalah salah satu *k-fold cross-validation* yang direkomendasikan untuk pemilihan model terbaik (Wibowo, 2017).

6.2.8. K-Nearest Neighbor

Pada algoritma *K-Nearest Neighbor* (KNN), sebagian atau seluruh data latih disimpan, kemudian digunakan dalam proses prediksi. Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Teknik dari algoritma *K-Nearest Neighbor* (KNN) yakni mengelompokkan suatu data baru berdasarkan jarak data baru tersebut ke beberapa data dengan jarak ketetanggaan terdekat (jarak *euclidean distance*) (Widiantara, 2018).

Dalam menghitung *euclidean distance*, langkah pertama yang harus dilakukan adalah menentukan data latih. Setelah mendapat data latih, kemudian melakukan proses perhitungan untuk mencari jarak dengan tetangga terdekatnya. Selanjutnya dilakukan proses perhitungan jarak menggunakan *euclidean distance*. Dalam metode KNN, yang dituliskan dalam (2.3) mendefinisikan rumus jarak *euclidean distance* (Parvin dkk, 2010).

$$d_{(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (2.3)$$

Keterangan :

$d(x,y)$: Jarak antara data uji dengan data latih

x : Data latih

y : Data uji

i : Mempresentasikan nilai atribut ke x dan y

n : Jumlah dimensi atribut

Keunggulan dari algoritma *K-Nearest Neighbor* ini adalah tidak memerlukan waktu yang lama untuk proses trainingnya, jika dibandingkan dengan algoritma *Backpropagation* (Redjeki, 2013).

6.2.9. **Modified K-Nearest Neighbor**

Modified K-Nearest Neighbor (MKNN) merupakan algoritma pengembangan dari kelemahan algoritma *K-Nearest Neighbor* (KNN). Kelemahan *K-Nearest Neighbor* (KNN) terdapat pada penanganan data latih yang hanya berbasis jarak untuk menentukan ketetanggan pada setiap data latih tanpa memperhitungkan nilai bobot pada setiap data latih (Parvin dkk, 2010).

Pada algoritma *Modified K-Nearest Neighbor* (MKNN) dilakukan proses perhitungan data latih dengan tetangga terdekatnya. Kemudian hasil dari perhitungan data latih tersebut

diklasifikasikan dengan data uji yang telah ditentukan. Dalam proses algoritma *Modified K-Nearest Neighbor* (MKNN), terdapat beberapa tambahan proses dibandingkan dengan *K-Nearest Neighbor* (KNN) yaitu, menghitung validitas dan *weight voting*. Sedangkan proses KNN tidak menggunakan perhitungan validitas dan *weight voting* (Parvin dkk, 2010).

6.2.9.1. Menghitung Nilai Validitas

Dalam algoritma MKNN, perhitungan validitas dihitung dari hasil *euclidean distance* setiap data berdasarkan tetangganya. Perhitungan validitas dilakukan setelah perhitungan jarak *euclidean* dilakukan. Dalam perhitungan validitas, dilakukan pencarian nilai tertinggi dari hasil perhitungan jarak. Dalam metode MKNN yang dituliskan dalam (2.4) mendefinisikan rumus untuk menghitung nilai validitas pada data *training* (Parvin dkk, 2010).

$$Validitas(x) = \frac{1}{H} \sum_{i=1}^H S(lbl(x), lbl(N_i(x))) \quad \dots\dots (2.4)$$

Keterangan :

H : Jumlah titik terdekat

$lbl(x)$: Kelas x

$lbl(N_i(x))$: Label kelas titik terdekat x

Fungsi S digunakan untuk menghitung kesamaan antara titik x dan data ke- i dari tetangga terdekat. Persamaan yang dituliskan dalam (2.5) mendefinisikan fungsi S .

$$S(a,b) = \begin{cases} 1 & \text{jika } a = b \\ 0 & \text{jika } a \neq b \end{cases} \quad \dots\dots\dots (2.5)$$

Keterangan :

a : Kelas a pada data *training*

b : Kelas lain selain a pada data *training*

6.2.9.2. Menghitung Weighted Voting

Dalam metode MKNN, perhitungan *weight voting* menggunakan hasil dari nilai validitas dan perhitungan jarak data uji dan data latih. Masing-masing data tetangga dihitung dengan menggunakan rumus bobot $1/(d_e + \alpha)$, validitas dari tiap data pada data *training* dikalikan dengan *weighted* berdasarkan pada jarak *Euclidian*. Dalam metode MKNN yang dituliskan dalam (2.6) mendefinisikan rumus *weight voting* tiap tetangga (Parvin dkk, 2010).

$$w_{(i)} = Validitas(i) \times \frac{1}{d_e + \alpha} \dots\dots\dots (2.6)$$

Keterangan :

$w_{(i)}$: Perhitungan *weight* bobot

$Validitas(i)$: Nilai validitas

d_e : Jarak *Euclidean Distance*

α : Nilai α bernilai 0,5 *smoothing* (pemulusan)

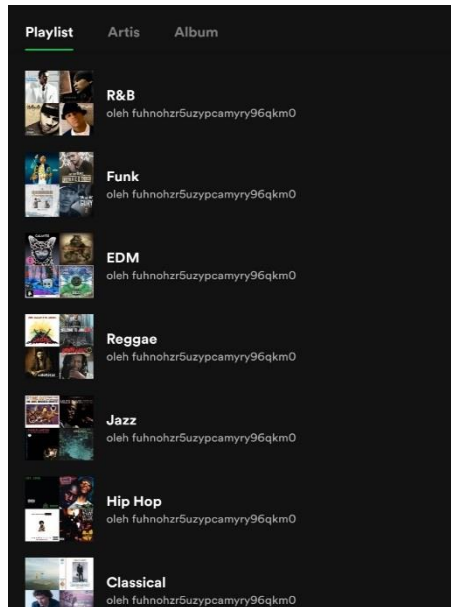
7. Metodologi Penelitian

7.1. Sumber Data

Dalam penelitian ini, data yang digunakan adalah data sekunder yang diperoleh dari website <https://www.spotify.com> dengan akses pengambilan data API Spotify menggunakan Python.

7.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini berjumlah 50 musik untuk setiap *genre*. *Genre* yang digunakan berjumlah 14 *genre*, yaitu *Blues*, *Classical*, *Country*, *Gospel*, *Hip Hop*, *Jazz*, *Metal*, *Pop*, *Reggae*, *Rock*, *EDM*, *Funk*, *R&B*, dan *Soul*. Jadi, total musik yang digunakan dalam penelitian ini adalah 700 musik di layanan streaming musik Spotify, seperti pada Gambar 7.2.



Gambar 7.2. Playlist Genre

7.3. Variabel dan Definisi Variabel

Secara keseluruhan, terdapat 18 variabel atribut lagu dalam API Spotify. Namun, hanya 13 variabel yang digunakan pada penelitian ini, yaitu 1 fitur pengguna (*playlist*) dan 12 fitur audio (*danceability*, *valence*, *energy*, *tempo*, *loudness*, *speechiness*, *instrumentalness*, *liveness*, *acousticness*, *key*, *mode*, dan *duration*). Penjelasan dan definisi dari variabel yang digunakan dalam penelitian ini ditampilkan dalam Tabel 7.1 dan Tabel 7.2.

Tabel 7.1. Variabel Fitur Pengguna

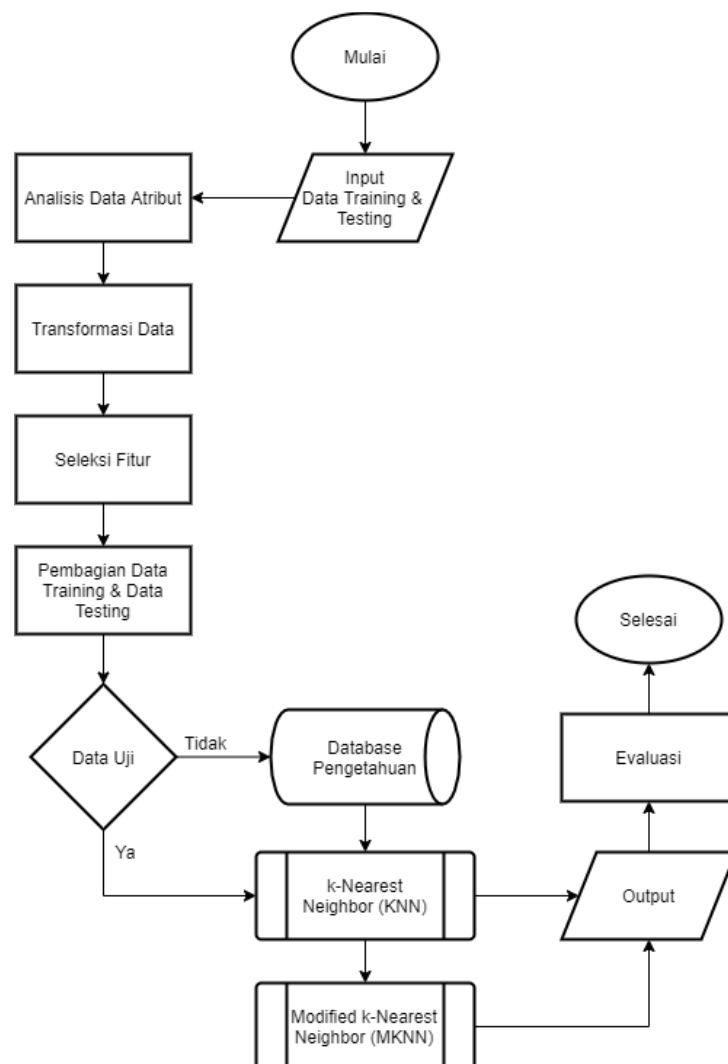
No	Variabel	Definisi Variabel	Ukuran
1.	<i>Playlist</i>	<i>Playlist</i> adalah atribut yang didalamnya berisikan data lagu. <i>Playlist</i> akan dibuat untuk setiap genre. Misalkan daftar <i>playlist</i> 0 untuk <i>genre</i> musik <i>Blues</i> , 1 untuk <i>genre</i> musik <i>Classical</i> , 2 untuk <i>genre</i> musik <i>Country</i> , dan seterusnya.	—

Tabel 7.2. Variabel Fitur Audio

1.	<i>Danceability</i>	Menjelaskan seberapa kecocokan sebuah lagu atau trek musik untuk menari	0,0 – 1,0
2.	<i>Valence</i>	Mengindikasi jenis emosi (seperti bahagia atau sedih)	0,0 – 1,0
3.	<i>Energy</i>	Ukuran intensitas dan aktivitas suatu tingkat dari kemampuan suatu musik untuk meningkatkan emosi dari pendengarnya	0,0 – 1,0
4.	<i>Tempo</i>	Tempo merupakan cepat atau lambatnya sebuah lagu. Ukuran dari sebuah tempo adalah <i>beats per minute</i> (BPM).	<i>Beats Per Minute</i> (BPM)
5.	<i>Loudness</i>	Kenyaringan suatu lagu dalam desibel (dB)	–60 – 0 Db
6.	<i>Speechiness</i>	Mendeteksi keberadaan kata-kata yang diucapkan di dalam sebuah trek lagu	> 0,66 = Trek Kata-kata 0,66 – 0,33 = Musik & Kata-kata < 0,33 = Trek NonSpeech
7.	<i>Instrumentalness</i>	Memprediksi apakah suatu lagu mengandung <i>vocal</i> atau tidak, seperti suara “Ooh” dan “Aah”	0,0 – 1,0
8.	<i>Liveness</i>	Mendeteksi keberadaan <i>audiens</i> dalam rekaman.	0,0 – 1,0
9.	<i>Acousticness</i>	Suatu ukuran yang menunjukkan apakah trek tersebut akustik atau tidak	0,0 – 1,0
10.	<i>Key</i>	Taksiran keseluruhan kunci trek. <i>Integer</i> memetakan ke <i>pitch</i> menggunakan	0 = C, 1 = C#/D b,

		notasi <i>Pitch Class</i> standar.	2 = D
11.	<i>Mode</i>	Mengindikasikan modalitas (besar atau kecil) dari suatu trek, jenis skala dari mana konten melodinya diturunkan	Mayor diwakili oleh 1 dan minor adalah 0
12.	<i>Duration</i>	Durasi trek musik dalam milidetik	Milidetik

7.4. Alur Penelitian



Gambar 7.3. Alur Penelitian

Pada alur penelitian seperti pada Gambar 7.3, dijelaskan diagram alur tahapan-tahapan dalam menyelesaikan keseluruhan penelitian ini. Penelitian ini akan dimulai dari melakukan *input data training* dan data

testing, lalu dilakukan analisis data atribut, transformasi data, seleksi fitur, pembagian data *training* dan data *testing*, klasifikasi dengan metode KNN dan MKNN, *output* yang diinginkan, dan evaluasi. Berikut definisi dari alur penelitian pada Gambar 7.2.

1. Data Training & Data Testing

Pada tahap ini, akan dilakukan pembuatan *playlist* lagu yang mana bekerja sama dengan mahasiswa lulusan Institut Seni Indonesia Denpasar jurusan Seni Karawitan. *Playlist* dibuat untuk masing-masing 14 *genre* pada layanan streaming Spotify, yang kemudian diunduh menggunakan API Web dan disimpan ke dalam *spreadsheet*.

2. Analisis Data Atribut

Pada tahap ini, penulis membuat *plot* sebaran data atribut lagu berdasarkan variable pada Tabel 7.1 dan 7.2 pada *genre* musik yang digunakan.

3. Transformasi Data

Transformasi data dilakukan guna mendapatkan *range* data yang sama dari 2 variabel data atau lebih yang memiliki *range* berbeda jauh dari variabel yang lainnya.

4. Seleksi Fitur

Menurut Putri (2017), seleksi fitur adalah metode yang penting untuk menghasilkan klasifikasi yang baik. Pada penelitian ini, proses seleksi fitur akan menggunakan Principal Component Analysis (PCA) dengan *tools* SPSS atau Matlab.

5. Pembagian Data *Training* dan Data *Testing*

Pembagian data *training* dan data *testing* dilakukan dengan menggunakan metode *k-fold cross-validation*. Metode *k-fold cross-validation* membagi data secara acak sejumlah *k subset* yang saling bebas dengan ukuran yang kira-kira sama atau berbeda. Setelah *subset* terbentuk, dilakukan pengulangan sebanyak *k*-kali untuk data *training* dan data *testing*.

6. Database Pengetahuan

Pada tahap ini, data *training* akan disimpan pada *database* pengetahuan yang nantinya akan digunakan untuk referensi dalam pengujian.

7. Proses klasifikasi KNN dan MKNN

Pada tahap ini, proses pengklasifikasian dilakukan menggunakan metode *K-Nearest Neighbor* (KNN) dan *Modified K-Nearest Neighbor* (MKNN).

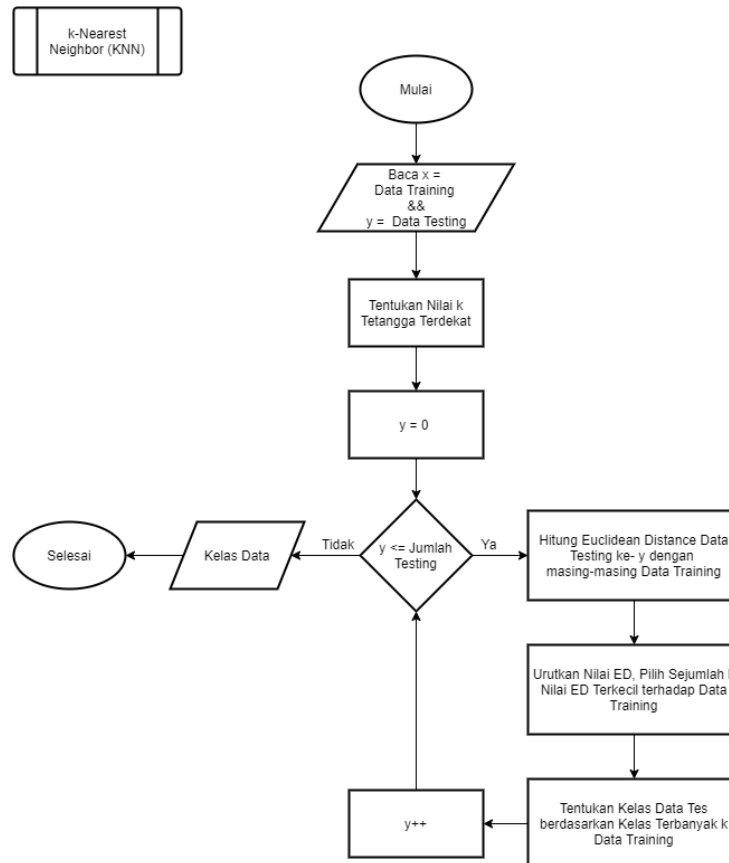
8. Output

Tahap ini berfungsi untuk mendapatkan hasil klasifikasi metode KNN dan MKNN.

9. Evaluasi

Tahap ini berfungsi untuk membandingkan performa dari klasifikasi metode KNN dan MKNN.

7.5. Proses Klasifikasi KNN



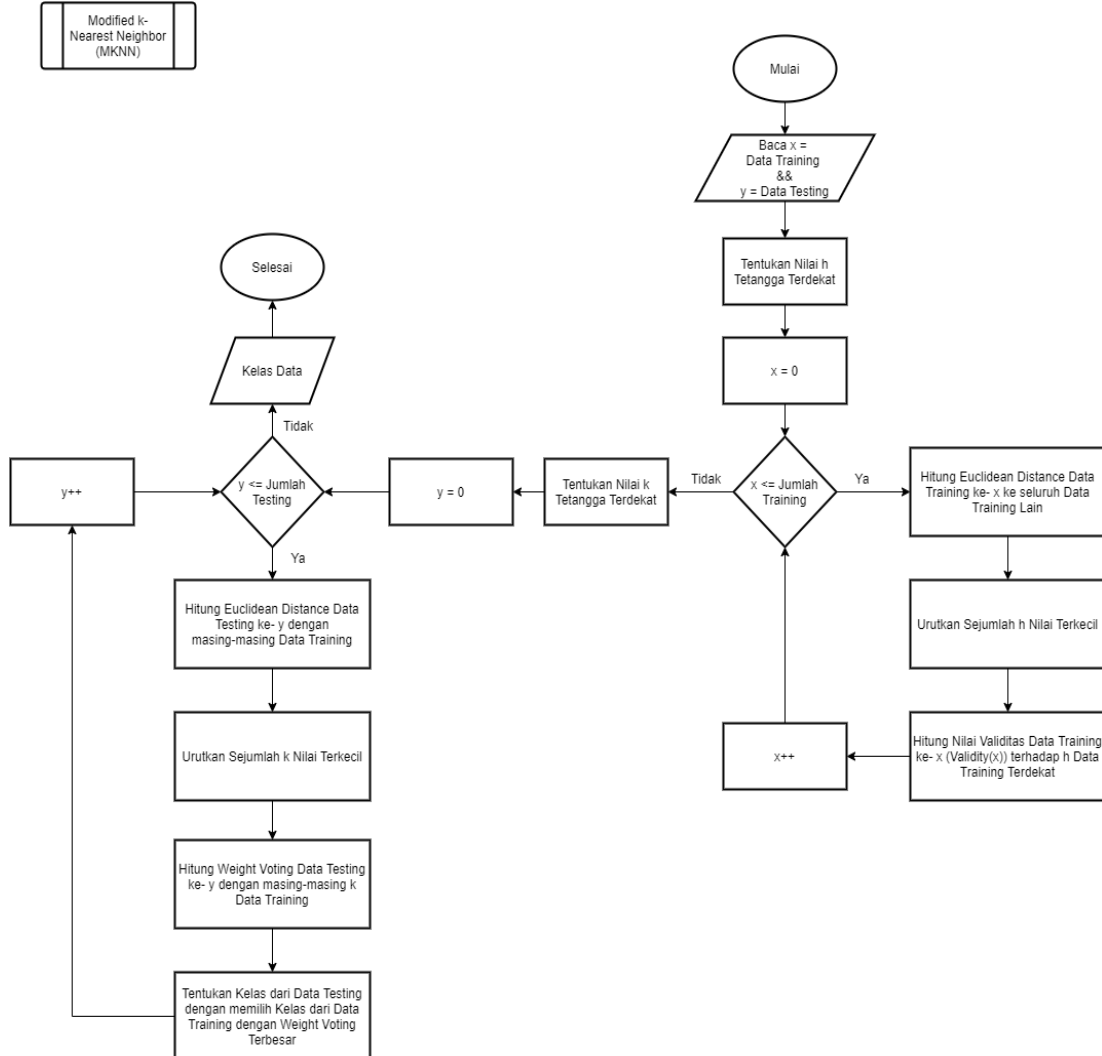
Gambar 7.4. Alir Proses klasifikasi KNN

Pada alir proses klasifikasi KNN pada Gambar 7.4, dijelaskan diagram alir tahapan-tahapan proses klasifikasi metode algoritma K-*Nearest Neighbor* (KNN) dalam menyelesaikan keseluruhan penelitian ini. Proses klasifikasi KNN dimulai dari membaca nilai x yang menyatakan data *training* dan nilai y yang menyatakan data *testing*. Setelah itu, dilakukan penentuan nilai k jumlah tetangga terdekat, penghitungan *Euclidean Distance*, pengurutan nilai *Euclidean Distance*, penentuan kelas data, hingga menghasilkan *output* untuk mendapatkan kelas data. Berikut definisi dari alir proses klasifikasi KNN pada Gambar 7.4.

1. Pada proses KNN menggunakan iterasi yaitu nilai x yang menyatakan data *training*, nilai y yang menyatakan data *testing*, dan nilai k yang menyatakan jumlah tetangga terdekat.
2. Variabelnya yaitu *Euclidean Distance* (ED).
3. Pada tahap ini, memisahkan data lagu menggunakan *k-fold cross-validation*.
4. Pada tahap ini, dilakukan penentuan nilai k . Penentuan nilai k dipertimbangkan berdasarkan banyaknya data yang ada dan ukuran dimensi yang dibentuk oleh data.
5. Pada tahap ini, iterasi y bernilai 0.
6. Pada tahap ini, jika nilai y kurang dari sama dengan data *testing* maka akan lanjut ke tahap berikutnya. Jika nilai y sama dengan data *testing* maka dihasilkan *output* berupa kelas data.
7. Pada tahap ini, data akan dihitung jarak ketetanggaannya (*Euclidean Distance*) pada data *testing* ke- y dengan masing-masing data *training*.
8. Pada tahap ini, dilakukan pengurutan nilai jarak *Euclidean Distance*, lalu memilih sejumlah k pada nilai *Euclidean Distance* terkecil terhadap data *training*.
9. Pada tahap ini, dilakukan penentuan kelas data *testing* berdasarkan kelas terbanyak dari sejumlah k data *training*.
10. Nilai y akan melalui perulangan hingga nilai y sama dengan jumlah *testing*.

11. Jika nilai y sama dengan data *testing*, maka dihasilkan *output* berupa kelas data. Selesai.

7.6. Proses Klasifikasi MKNN



Gambar 7.5. Alir Proses Klasifikasi MKNN

Pada alir proses klasifikasi MKNN pada Gambar 7.5, dijelaskan diagram alir tahapan–tahapan proses klasifikasi metode algoritma *Modified K-Nearest Neighbor* (MKNN) dalam menyelesaikan keseluruhan penelitian ini. Proses klasifikasi MKNN dimulai dari membaca x (yang menyatakan data *training*) dan y (yang menyatakan data *testing*). Variabel h dan k menyatakan jumlah tetangga terdekat. Selanjutnya adalah menghitung *Euclidean Distance*, yang dilanjutkan

dengan melakukan perhitungan validitas yang dihitung dari hasil *Euclidean Distance* setiap data berdasarkan tetangganya. Setelah mendapatkan validitas, dilakukan perhitungan *weight voting* menggunakan hasil dari nilai validitas dan perhitungan jarak data *training* dan data *testing* hingga menghasilkan *ouput* berupa kelas data. Berikut definisi dari alir proses klasifikasi MKNN pada Gambar 7.5.

1. Pada proses MKNN menggunakan iterasi yaitu variable x , yang menyatakan data *training*, dan variabel y , yang menyatakan data *testing*. Variabel h dan k menyatakan jumlah tetangga terdekat.
2. Variabelnya yaitu *Euclidean Distance* (ED), *validity* (validitas), dan *weight voting*.
3. Pada tahap ini, iterasi nilai x bernilai 0.
4. Pada tahap ini, jika nilai x kurang dari atau sama dengan data *training*, maka akan lanjut ke tahap *validity* atau validitas. Namun, jika nilai x sama dengan data *training*, maka lanjut ke tahap menghitung *weight voting* saja tanpa melakukan *validity* atau validitas.
5. Pada tahap *validity* atau validitas, data akan dihitung jarak ketetanggaannya (*Euclidean Distance*) pada data *training* ke- x ke seluruh data *training* lain.
6. Pada tahap ini, diurutkan sejumlah iterasi h nilai terkecil.
7. Pada tahap ini, menghitung nilai validitas data *training* ke- x atau *validity(x)* terhadap sejumlah h data *training* terdekat, lalu melakukan *increment* nilai x .
8. Nilai x akan melalui perulangan hingga nilai x sama dengan jumlah *training*. Selanjutnya menuju tahap perhitungan dengan *weight voting*.
9. Pada tahap *weight voting*, iterasi nilai y bernilai 0.
10. Pada tahap ini, jika nilai y kurang dari atau sama dengan data *testing*, maka akan lanjut ke tahap *weight voting*. Namun, jika kelas nilai x sama dengan data *testing*, maka proses selesai.
11. Pada tahap ini, menghitung *Euclidean Distance* data *testing* ke- y dengan masing-masing data *training*.
12. Pada tahap ini, dilakukan pengurutan sejumlah iterasi k nilai terkecil.

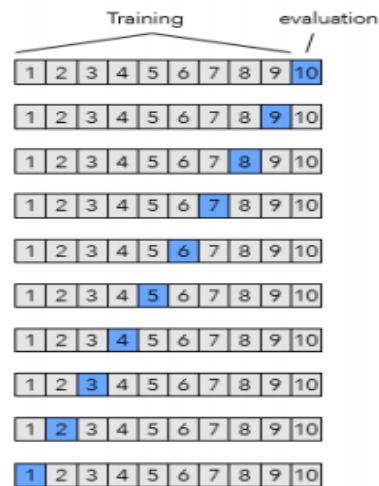
13. Pada tahap ini, dilakukan perhitungan *weight voting* data *testing* ke- y dengan masing-masing k data *training*.
14. Jika y sama dengan k , maka dilanjutkan ke proses menentukan kelas data *testing* berdasarkan *weight voting* data *training* terbesar.
15. Pada tahap ini yaitu proses menentukan kelas dari data *testing* dengan memilih kelas dari data *training* dengan *weight voting* terbesar.
16. Nilai y akan melakukan perulangan hingga nilai y sama dengan jumlah *testing*.
17. Jika nilai y sama dengan data *testing*, lakukan *output* berupa kelas data. Selesai.

7.7. Skenario Pengujian dan Evaluasi

7.7.1. Skenario Pengujian

Pengujian ini berfungsi untuk mengetahui kinerja dari sistem itu sendiri dalam melakukan tugas yaitu melakukan klasifikasi. Pada tahap ini terdapat beberapa skenario pengujian yang akan dilakukan pada penelitian ini, yaitu (Subramanian, 2019) :

- Penentuan nilai k pada algoritma KNN dan MKNN. Terdapat 3 cara, yaitu (Potdar, & Kinnerkar, 2016) :
 1. Penentuan k menggunakan langkah *trial and error*. Dengan cara ini mencari nilai k dengan cara coba-coba, misalnya $k=1$ sampai dengan $k=10$.
 2. Penentuan k melalui *k-fold cross-validation*. Salah satu cara untuk memilih dataset validasi silang dari *dataset training* dan *dataset testing*. Misalnya data total 700 data, maka data akan dibagi sebanyak 10 iterasi yang berisi 9 bagian data *training* dan 1 bagian data *testing*, seperti pada Gambar 7.6.



Gambar 7.6. K-Fold Cross-Validation

3. Penentuan k dengan rumus $k = \sqrt{N}$, di mana N adalah jumlah sampel dalam set data training.

Dari penjelasan di atas, pada penelitian ini penulis menggunakan penentuan nilai k pada algoritma KNN dan MKNN menggunakan penentuan *k-fold cross-validation*.

- Dilakukan analisis atribut dengan menampilkan *plot* penyebaran jarak variabel fitur audio pada data.
- Dilakukan transformasi data untuk mendapatkan *range* data yang sama dari 2 variabel fitur atau lebih yang memiliki range berbeda jauh, seperti fitur *instrumentalness*, *tempo*, *loudness*, dan *duration* yang datanya berbeda dari yang lainnya.
- Dilakukan proses seleksi fitur menggunakan Principal Component Analysis (PCA) dengan *tools* SPSS atau Matlab. Principal Component Analysis (PCA) merupakan salah satu teknik untuk mereduksi dimensi dari suatu data. Tidak seperti metode seleksi fitur yang membandingkan masing-masing variabel fitur kelas datanya. Jika nilainya diluar batas fitur, maka membuang jumlah atribut fitur yang tidak relevan yang dianggap tidak penting tanpa membentuk variabel fitur baru. Sedangkan PCA, mengurangi dimensi data dengan cara ‘mengkombinasikan’ intisari dari atribut dengan membentuk

alternatif subset fitur yang lebih kecil menjadi terbentuk variabel fitur baru (Wibawa, 2017).

- Dilakukan klasifikasi. Pada proses ini, data seleksi fitur yang sudah dilakukan akan dilakukan klasifikasi dengan metode KNN dan MKNN menggunakan *k-fold cross-validation* dalam menentukan *k* terbaik.

7.7.2. Evaluasi

Pada tahap ini terdapat beberapa evaluasi yang ingin didapatkan pada penelitian ini yaitu sebagai berikut.

- Analisis performa penelitian ini tidak hanya menggunakan pengaruh seleksi fitur saja, melainkan bisa digunakan terhadap perbandingan performa mencari nilai *k* terbaik yang dihasilkan dari algoritma klasifikasi KNN dan MKNN dengan *k-fold cross-validation* menggunakan akurasi *confusion matrix* pada persamaan 2.2.

8. Jadwal Pelaksanaan Penelitian

Pelaksanaan dari kegiatan penelitian yang penulis lakukan direncanakan menghabiskan waktu selama lima bulan. Rincian dari kegiatan yang dilakukan dapat dilihat pada Tabel 8.1 di bawah ini.

Tabel 8.1. Jadwal Pelaksanaan Kegiatan

	Minggu Ke–														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Studi Literatur															
Pengumpulan Data Lagu Genre															
Analisis Fitur Atribut Lagu															

Pengujian Penelitian															
Evaluasi Penelitian															

DAFTAR PUSTAKA

- Ashrith. (2018, December 04). *Analyzing Spotify's Top Tracks Of 2017 Using Data Visualization*. Retrieved from Toward Data Science:
<https://towardsdatascience.com/what-makes-a-song-likeable-dbfdb7abe404>
- Giri, G.A.V.M.G. (2018). Klasifikasi Musik Berdasarkan Genre dengan Metode K-Nearest Neighbor. *Jurnal Ilmu Komputer*, VOL. XI No. 2, 103-108.
- Iriansyah, F.Y. (2018, November 13). *[Update] Apa Beda Spotify dengan Apple Music, Joox, dan Deezer?* Retrieved from Techinasia:
<https://id.techinasia.com/komparasi-layanan-streaming-spotify>
- Jamalus. (1998). *Panduan Pengajaran buku Pengajaran musik melalui pengalaman musik*. Jakarta: Proyek Pengembangan Lembaga Pendidikan.
- Mutrofin,S., Izzah,A., Kurniawardhani,A. & Masrur,Mukhamad. (2014). Optimasi Teknik Klasifikasi Modified K-Nearest Neighbor Menggunakan Algoritma Genetika. *JURNAL GAMMA*, 130-134.
- Nomleni, P. (2015). Sentiment Analysis Menggunakan Support Vector Machine (SVM). Surabaya: Program Pasca Sarjana Bidang Keahlian Telematika (CIO) Jurusan Teknik Elektro Fakultas Teknologi Industri Institut Teknologi Sepuluh Nopember.
- Okfalisa., Mustakim., Gazalba, Ikbali. & Reza, N.G.I. (2017). Comparative Analysis of K-Nearest Neighbor and Modified K-Nearest Neighbor Algorithm for Data Classification. *International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 294-298.
- Parvin, H., Alizadeh, H. & Minati, B. (2010). A Modification on K-Nearest Neighbor Classifier. *Global Journal of Computer Science and Technology*, 37-41.

- Potdar, K. & Kinnerkar, R. (2016). A Comparative Study of Machine Learning Algorithms applied to Predictive Breast Cancer Data. *International Journal of Science and Research (IJSR)*, Vol 5 Issue 9, 1550-1553.
- Prasetya, E. (2012). *Data Mining-Konsep dan Aplikasi menggunakan MATLAB*. Yogyakarta: Andi.
- Ravi, M.R., Indriati, & Adinugroho, S. (2019). Implementasi Algoritme Modified K-Nearest Neighbor (MKNN) Untuk Mengidentifikasi Penyakit Gigi Dan Mulut. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol 3, 2596-2602.
- Redjeki, S. (2013). Perbandingan Algoritma Backpropagation dan K-Nearest Neighbor (K-NN) untuk Identifikasi Penyakit. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI) Yogyakarta*, 1-5.
- Spotify. (2020, June 12). <https://www.spotify.com/us/about-us/contact/>. Retrieved from <https://newsroom.spotify.com/company-info/>: <https://www.spotify.com/>
- Subramanian, D. (2019, June 08). *A Simple Introduction to K-Nearest Neighbors Algorithm*. Retrieved from Toward Data Science: <https://towardsdatascience.com/a-simple-introduction-to-k-nearest-neighbors-algorithm-b3519ed98e>
- Supriyadi. (2018). *Analisis Klasifikasi Genre Musik Pop dan Klasik pada Layanan Streaming Musik Spotify Menggunakan Artificial Neural Network (ANN) (Studi Kasus: Lagu dengan Genre Musik Pop dan Klasik di Layanan Streaming Musik Spotify)*. Yogyakarta: Universitas Islam Indonesia.
- Wibawa, M.S. & Novianti, K.D.V. (2017). Reduksi Fitur Untuk Optimalisasi Klasifikasi Tumor Payudara Berdasarkan Data Citra FNA. *Konferensi Nasional Sistem & Informatika 2017, STMIK STIKOM Bali*, 71-78.
- Wibowo, A. (2017). *Binus University Graduate Program*. Retrieved February 8, 2020, from <https://mti.binus.ac.id/2017/11/24/10-fold-cross-validation/>

Widiantara, W.D. (2018). *Klasifikasi Genre Musik Menggunakan Metode Modified K-Nearest Neighbor*. Riau: Universitas Islam Negeri Sultan Syarif Kasim Riau.