



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB III

METODOLOGI PENELITIAN

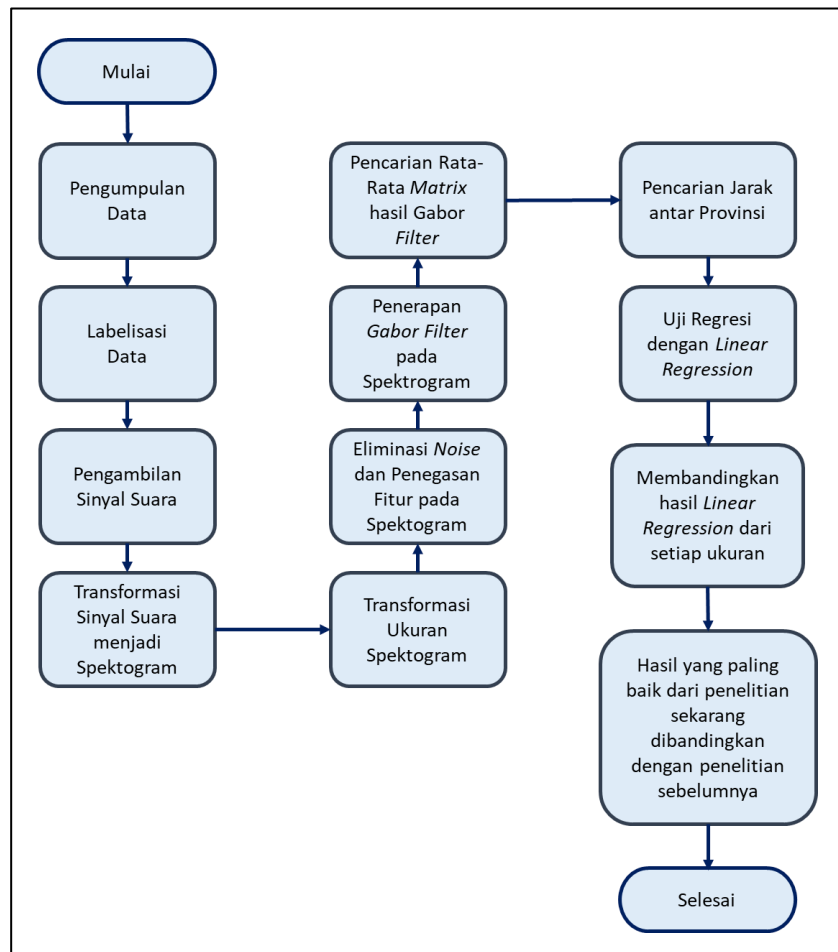
3. 1. Objek Penelitian

Objek penelitian berupa lagu-lagu daerah di seluruh provinsi Pulau Jawa yang digunakan untuk bahan penelitian diperoleh dari buku Rangkuman Pengetahuan Umum Lengkap (RPUL) Indonesia dan Dunia untuk SD, SMP, SMA, dan Umum dari Penerbit Erlangga dengan ISBN: 978-602-486-064-6. Lagu dari setiap provinsi yang diuji akan dicek ketersediaannya pada Spotify dan diunduh melalui *platform* tersebut. Adapun lagu yang dipakai hanya fokus pada Pulau Jawa saja karena adanya keterbatasan waktu pada penelitian. Spotify dipilih karena seluruh musik yang ada di dalamnya dikirim dari distributor ataupun label rekaman yang bekerja sama dengan distributor sehingga tidak sembarang orang dapat mengunggah data (Spotify, 2019). Selain itu, dipilihnya Spotify daripada *platform streaming* lain contohnya seperti JOOX karena perbedaan jumlah lagu. JOOX memiliki total 5 juta lagu (JOOX, 2019), sedangkan Spotify memiliki total 20 juta lagu.

3. 2. Alur Penelitian

Penelitian ini dilakukan menggunakan metode yang pernah dipakai oleh penelitian sebelumnya, namun beberapa hal disesuaikan dengan keadaan data penelitian sekarang. Adapun metode yang dipakai sama bertujuan untuk menjaga perbandingan hasil yang lebih akurat. Proses setelah pengumpulan data

berdasarkan penelitian sebelumnya (Kuanca, 2019) setelah penyesuaian terbagi menjadi 9 tahap seperti pada gambar 3.2 yaitu: (1) pengumpulan data, (2) labelisasi data, (3) pengambilan sinyal suara dari file *Waveform Audio Format* (WAV), (4) transformasi sinyal suara menjadi spektrogram, (5) transformasi ukuran spektrogram, (6) eliminasi noise dan penegasan fitur pada spektrogram, (7) penerapan Gabor Filter pada spektrogram, (8) pencarian rata-rata Gabor Matrix untuk angka penunjuk fitur, (9) pencarian jarak antar provinsi, (10) pencarian pengaruh fitur dengan jarak geografi antar provinsi, (11) Membandingkan hasil Linear Regression dari setiap ukuran, dan (12) Hasil yang paling baik dari penelitian sekarang dibandingkan dengan penelitian sebelumnya.



Gambar 3.1. Alur metode penelitian dari awal hingga akhir

3. 2. 1 Pengumpulan Data

Proses pengumpulan data dimulai dengan mengambil lagu daerah yang hanya berasal dari seluruh provinsi Pulau Jawa dari buku acuan. Setelah itu, data yang sudah dikumpulkan dicek ketersediaanya pada Spotify. Pengecekan ini dapat menggunakan fitur *search* pada kolom yang telah tersedia dan memasukkan judul lagu ke dalamnya. Jika tersedia, maka lagu akan diunduh untuk kemudian dilanjutkan ke proses selanjutnya, jika tidak ada maka lagu akan dilewati dan lanjut pengecekan lagu selanjutnya.

Jika terdapat lebih dari satu judul lagu yang sama dalam pencarian, maka lagu yang akan dipakai adalah lagu yang memiliki jumlah *streaming* tertinggi pada Spotify.

3. 2. 2 Labelisasi Data

Sebelum labelisasi, data dilakukan *cleansing* terlebih dahulu dengan memotong bagian lagu yang tidak memiliki nada. Labelisasi selanjutnya digunakan sebagai pemberi identitas dari masing-masing data. Nama provinsi beserta jarak *latitude* dan *longitude* daerah provinsi dan pulau (letak geografis) dipakai pada tahap labelisasi ini.

3. 2. 3 Pengambilan Sinyal Suara dari File WAV

Pengambilan sinyal suara dilakukan menggunakan *library* Scipy.io. Sinyal suara yang menjadi input adalah sinyal suara WAV 1 *channel* (Wolfe & Godsill, 2003). Hasil dari tahap pengambilan suara ini berupa audio sinyal *numpy array* dengan dimensi satu dan frekuensi (Kuanca, 2019).

3. 2. 4 Transformasi Sinyal Suara menjadi Spektogram

Agar audio sinyal bisa tergambarkan secara visual, maka sinyal tersebut diubah menjadi bentuk spektogram. Spektogram yang dihasilkan menggambarkan intensitas frekuensi suatu sinyal lagu. Spektogram dihasilkan dari proses *Short Time Fourier Transform* (STFT) yang dikomputasi juga menggunakan *library* Scipy.io.

3. 2. 5 Transformasi Ukuran Spektogram

Spektogram yang dihasilkan dari tahapan sebelumnya lalu ditransformasi guna menjaga akurasi dan mengurangi *noise* yang ada pada gambar. Selain itu, proses ini juga berpengaruh pada efisiensi kerja algoritma. Penelitian sebelumnya memakai dimensi 256 x 256 piksel dengan *anti-aliasing* (Kuanca, 2019). Penelitian kali ini menggunakan tiga dimensi ukuran yaitu 256 x 256 piksel itu sendiri ditambah kenaikan ukuran dua kali yaitu 512 x 512 piksel dan 1024 x 1024 piksel sebagai bagian dari pembaharuan penelitian.

3. 2. 6 Eliminasi Noise dan Penegasan Fitur pada Spektogram

Spektogram lagu daerah dari proses sebelumnya memiliki fitur yang belum tegas, maka dari itu dilakukan eliminasi *noise* dan penegasan dengan menggunakan pembulatan biner menjadi tipe data uint8, perubahan warna menggunakan *colormap*, dan pengaplikasian Gaussian Filter (Kuanca, 2019). Perubahan warna dilakukan dengan warna paling kontras atau memperlihatkan perbedaan yang nyata apabila dibandingkan. Warna yang dipilih adalah warna hitam dan putih sebagai yang pertama dan terakhir dalam spektrum warna (Frery, Dos Santos Melo, & Fernandes, 2000).

3. 2. 7 Penerapan Gabor Filter pada Spektogram

Sama seperti penelitian sebelumnya, penggunaan *Gabor Filter* dilakukan dengan beberapa parameter, yaitu frekuensi spasial dalam bentuk frekuensi piksel, *theta* sebagai orientasi *filter* dalam satuan *radians*, dan *offset* sebagai fase *offset* harmonik suatu pengaplikasian *Gabor Filter*

dengan satuan *radians* itu sendiri. Untuk itu, digunakan 19 orientasi *theta* berbeda untuk mencari fitur naik atau turun dari spektrogram lagu daerah tersebut. Prosesnya akan dilakukan dengan perhitungan *Gabor* dari 0 derajat, 10 derajat, 20 derajat, dan begitu seterusnya sampai 19 fitur orientasi sesuai yang dilakukan pada penelitian sebelumnya dan berguna mencari seberapa cepat atau lambat nada yang naik dan turun (Kuanca, 2019).

3. 2. 8 Pencarian Rata-rata Gabor Matrix untuk Angka Penunjuk Fitur

Hasil gambar yang sudah melalui filtrasi *Gabor* nilainya akan dirata-rata sehingga menjadi suatu angka yang menunjukkan fitur tertentu. Selain itu, hasil dari proses ini juga menunjukkan rata-rata dari semua dimensi yang ada, dapat dikatakan hasil ini dapat mewakili kecenderungan setiap nilai daerah uji. Proses ini dilakukan dengan menggunakan *np.mean()*, yaitu rumus dalam *library* Numpy yang digunakan untuk menghitung rata-rata.

3. 2. 9 Pencarian Jarak antar Provinsi

Pencarian jarak berdasar pada *longitude* dan *latitude* setiap daerah. Perhitungan dilakukan dengan menggunakan rumus hukum *Haversine*. Rumus *Haversine* ini dipakai untuk mencari jarak antara dua titik yang ada di permukaan bola, jadi rumus ini tidak mengabaikan kelengkungan yang dimiliki oleh bumi saat menghitung jarak tersebut (Yulianto et al., 2018).

3. 2. 10 Pencarian Pengaruh Fitur dengan Jarak Geografi Antar Provinsi

Pencarian pengaruh fitur selanjutnya dilakukan menggunakan metode *linear regression* dari hasil-hasil yang telah didapatkan sebelumnya. Dengan bantuan *library* Statsmodel, hasil dari regresi linear ini adalah *p-value* untuk menentukan signifikansi antara fitur *Gabor* dengan jarak antar provinsi. Jika *p-value* pada fitur tersebut di bawah 0.05 ($p < 0.05$), maka fitur bernilai signifikan entah itu berbanding lurus maupun berbanding terbalik. Hubungan perbandingan dapat dilihat dari nilai koefisien hasil, jika nilainya positif maka pada fitur tersebut nilai perhitungan Gabor dan jarak geografis yang sangat berbanding lurus, jika nilainya negatif maka sangat berbanding terbalik.

3. 2. 11 Membandingkan Hasil *Linear Regression* dari Setiap Ukuran

Hasil dari pengujian *linear regression* pada setiap ukuran spektogram selanjutnya diambil yang paling baik. Baik tidaknya hasil tersebut dilihat berdasarkan jumlah *p-value* yang signifikan (nilai di bawah 0.05) pada setiap ukuran, yaitu semakin banyak jumlahnya maka pengujian menggunakan ukuran tersebut akan semakin baik. Ukuran dengan jumlah *p-value* terbanyak tersebut akan dipakai pada tahap selanjutnya.

3. 2. 12 Hasil Penelitian yang Paling Baik dari Penelitian Sekarang dibandingkan dengan Penelitian Sebelumnya

Linear regression dari ukuran terpilih pada penelitian sekarang selanjutnya dibandingkan dengan penelitian terdahulu. Variabel yang

dibandingkan adalah jumlah *p-value* signifikan dan koefisien (*coef*) yang bernilai positif. Hasil yang lebih baik dinilai dari jumlah *p-value* yang lebih banyak.