

**TESIS**

**ANALISIS AUDIO : KLASIFIKASI ALAT MUSIK GAMELAN JAWA  
MENGUNAKAN DEEP LEARNING DENGAN PYTHON**



Disusun oleh:

**Nama : Nadea Cipta Laksmi**  
**NIM : 19.52.1195**  
**Konsentrasi : Informatics Technopreneurship**

**PROGRAM STUDI S2 TEKNIK INFORMATIKA  
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2021**

**TESIS**

**ANALISIS AUDIO : KLASIFIKASI ALAT MUSIK GAMELAN JAWA  
MENGUNAKAN DEEP LEARNING DENGAN PYTHON**

**AUDIO ANALYSIS : CLASSIFICATION OF JAVA GAMELAN  
INSTRUMENT USING DEEP LEARNING WITH PYTHON**

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Magister



Disusun oleh:

**Nama** : Nadea Cipta Laksmita  
**NIM** : 19.52.1195  
**Konsentrasi** : Informatics Technopreneurship

**PROGRAM STUDI S2 TEKNIK INFORMATIKA  
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

**HALAMAN PENGESAHAN**

**ANALISIS AUDIO : KLASIFIKASI ALAT MUSIK GAMELAN JAWA  
MENGUNAKAN DEEP LEARNING DENGAN PYTHON**

**AUDIO ANALYSIS : CLASSIFICATION OF JAVA GAMELAN INSTRUMENT  
USING DEEP LEARNING WITH PYTHON**

Dipersiapkan dan Disusun oleh

**Nadea Cipta Laksmita**

**19.52.1195**

Telah Ditujikan dan Dipertahankan dalam Sidang Ujian Tesis  
Program Studi S2 Teknik Informatika  
Program Pascasarjana Universitas AMIKOM Yogyakarta  
pada hari Rabu, 7 April 2021

Tesis ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Magister Komputer

Yogyakarta, 7 April 2021

**Rektor**

**Prof. Dr. M. Suyanto, M.M.**  
**NIK. 190302001**

## HALAMAN PERSETUJUAN

### ANALISIS AUDIO : KLASIFIKASI ALAT MUSIK GAMELAN JAWA MENGUNAKAN DEEP LEARNING DENGAN PYTHON

### AUDIO ANALYSIS : CLASSIFICATION OF JAVA GAMELAN INSTRUMENT USING DEEP LEARNING WITH PYTHON

Dipersiapkan dan Disusun oleh

**Nadea Cipta Laksmita**

**19.52.1195**

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis  
Program Studi S2 Teknik Informatika  
Program Pascasarjana Universitas AMIKOM Yogyakarta  
pada hari Rabu, 7 April 2021

**Pembimbing Utama**

**Dr. Arief Setyanto, S.Si., M.T.**  
**NIK. 190302106**

**Anggota Tim Penguji**

**Dr. Andi Sunyoto, M.Kom.**  
**NIK. 190302052**

**Pembimbing Pendamping**

**Ferry Wahyu Wibowo, S.Si., M.Cs**  
**NIK. 190302235**

**Alva Hendi M., S.T., M.Eng., Ph.D.**  
**NIK. 190302493**

Tesis ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Magister Komputer

Yogyakarta, 7 April 2021  
**Direktur Program Pascasarjana**

**Dr. Kusrini, M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nadea Cipta Laksmita  
NIM : 19.S2.1195  
Konsentrasi : Informatics Technopreneurship

Menyatakan bahwa Tesis dengan judul berikut:

**ANALISIS AUDIO : KLASIFIKASI ALAT MUSIK GAMELAN JAWA  
MENGUNAKAN DEEP LEARNING DENGAN PYTHON**

Dosen Pembimbing Utama : Dr. Anief Setyanto, S.Si., M.T.

Dosen Pembimbing Pendamping : Ferry Wahyu Wibowo, S.Si., M.Cs.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 15 Februari 2021

Yang Menyatakan,



Nadea Cipta Laksmita

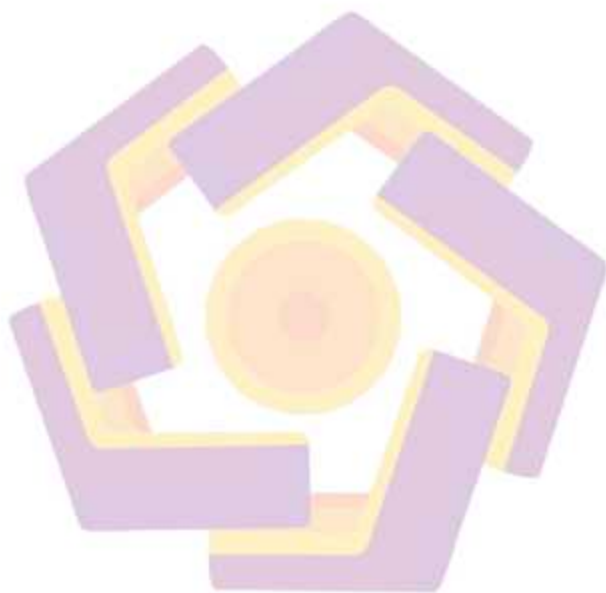
## HALAMAN PERSEMBAHAN

Saya juga sangat berterima kasih kepada semua pihak yang telah secara langsung maupun tidak langsung membantu saya dalam menyelesaikan tesis ini. Tesis ini saya persembahkan kepada:

1. Ibu dan Ayah yang tidak pernah lelah mendoakan, selalu support baik finansial maupun dalam bentuk dukungan lainnya. Serta Kakak Ninda sayang dan Keluarga dekat saya yang selalu mendukung.
2. Bapak Dr. Arief Setyanto, S.Si., M.T. dan Bapak Ferry Wahyu Wibowo, S.Si., M.Cs. selaku dosen pembimbing yang selalu memberikan masukan serta bimbingan positif dalam menyelesaikan tesis ini.
3. R. Anggitan Dewa Citra, Anisa Hardiyanti, Yannasa Farah yang selalu menjadi orang terbaik saya untuk penyemangat.
4. Paradise, Khoironi, Firman Tri Anggara, Hengky Garbo Setiawan yang menjadi panutan dan membantu saya dalam belajar, dan selalu menjadi inspirasi saya dalam belajar.
5. Teman seperjuangan di Magister Teknik Informatika Universitas Amikom Yogyakarta.

## HALAMAN MOTTO

Kurang Cerdas Bisa Diperbaiki, Namun Tidak Jujur Sulit Diperbaiki



## KATA PENGANTAR

Segala puji bagi Allah yang telah memberikan rahmat, kekuatan, kesempatan, sehingga penulis dapat menyelesaikan tesis dengan judul **"ANALISIS AUDIO : KLASIFIKASI ALAT MUSIK GAMELAN JAWA MENGGUNAKAN DEEP LEARNING DENGAN PYTHON"**. Penyusunan tesis ini untuk memenuhi persyaratan kelulusan Program Magister di Universitas AMIKOM Yogyakarta pada Jurusan Teknik Informatika.

Dalam penyusunan tesis ini tidak lepas dari dukungan berbagai pihak, oleh karena itu pada kesempatan ini penulis mengucapkan terimakasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas "AMIKOM" Yogyakarta.
2. Ibu Dr. Kusriani, M.kom selaku direktur program pascasarjana Universitas "AMIKOM" Yogyakarta.
3. Bapak Dr. Arief Setyanto, S.Si., M.T selaku dosen pembimbing pertama yang telah meluangkan waktu untuk membimbing, mengarahkan, dan memberikan ilmu saat penyusunan tesis.
4. Bapak Ferry Wahyu Wibowo, S.Si., M.Cs. selaku dosen pembimbing kedua yang telah meluangkan waktu untuk membimbing, mengarahkan, dan memberikan ilmu saat penyusunan tesis.
5. Kedua orang tua penulis, untuk doa dan dukungan yang terus mengalir.
6. Bapak Ibu dosen dan pegawai Universitas "AMIKOM" Yogyakarta.
7. Semua pihak telah membantu dalam proses penyusunan tesis yang terlibat langsung maupun tidak yang tidak dapat penulis sebutkan satu persatu.

Yogyakarta, 15 Februari 2021

Penulis



## DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PERNYATAAN KEASLIAN TESIS .....	v
HALAMAN PERSEMBAHAN .....	vi
HALAMAN MOTTO .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
INTISARI.....	xiii
<i>ABSTRACT</i> .....	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	8
1.3. Batasan Masalah.....	9
1.4. Tujuan Penelitian.....	9
1.5. Manfaat Penelitian .....	10
BAB II TINJAUAN PUSTAKA.....	11
2.1. Tinjauan Pustaka .....	11
2.2. Keaslian Penelitian.....	15
2.3. Landasan Teori.....	21
2.3.1. Pengertian Klasifikasi .....	21
2.3.2. Gamelan Jawa .....	21
2.3.2.1. Unsur-unsur dalam Gamelan .....	22
2.3.2.2. Alat- Alat Gamelan .....	26
2.3.3. <i>Deep Learning</i> .....	34
2.3.4. <i>Machine Learning</i> .....	36
2.3.5. MFCC ( <i>Mel-Frequency Cepstral Coefficients</i> ) .....	39

2.3.6. <i>Python</i> .....	42
2.3.7. <i>Convolutional Neural Network</i> .....	44
2.3.8. Arsitektur Jaringan CNN.....	49
2.3.9. Musik .....	50
2.3.10. Audio.....	51
2.3.11. Wav/ Audio Waveform .....	52
2.3.12. Music Information Retrieval (Informasi Pencarian Musik)....	54
<b>BAB III METODE PENELITIAN</b> .....	<b>54</b>
3.1. Jenis, Sifat dan Pendekatan Penelitian.....	54
3.1.1. Jenis Penelitian.....	54
3.1.2. Sifat Penelitian .....	55
3.1.3. Pendekatan Penelitian .....	55
3.2. Metode Pengumpulan.....	56
3.3. Metode Analisa Data.....	57
3.4. Alur Penelitian .....	60
3.5. Alur Klasifikasi Alat Musik Gamelan Jawa .....	60
3.5.1. Data Eksplorasi dan Visualisasi .....	62
3.5.2. Preprocessing Data.....	64
3.5.3. Model Training dan Evaluasi .....	67
<b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN</b> .....	<b>67</b>
4.1. Implementasi Pengumpulan Data .....	70
4.2. Eksplorasi dan Visualisasi Data .....	70
4.2.1. Dataset Suara Alat Musik Gamelan Jawa .....	71
4.2.2. Data File Sampel Audio .....	72
4.2.3. Menganalisis Data Audio.....	72
4.2.4. Pemeriksaan Auditori.....	73
4.2.5. Inspeksi Visual .....	74
4.2.6. Pengamatan .....	75
4.2.7. Metada.....	76
4.2.8. Distribusi Kelas.....	76
4.2.9. Properti File Sampel Audio.....	77

4.2.10. Channel Audio .....	78
4.2.11. Sample Rate .....	78
4.2.12. Bit Depth .....	79
4.3. Algoritma dan Teknik .....	82
4.4. Metodologi Pemrosesan .....	82
4.4.1. Data dan Pemisahan Data .....	82
4.4.1.1. Properti Audio Yang Memerlukan Normalisasi .....	83
4.4.1.2. Tahapan Preprocessing .....	84
4.4.2. Fitur Ekstraksi .....	85
4.4.3. Mengonversi Data dan Label .....	87
4.4.4. Memisahkan Dataset .....	87
4.5. Implementasi .....	87
4.5.1. Penerapan Model Awal MLP .....	89
4.5.2. Mengkompilasi Model .....	91
4.5.3. Training .....	92
4.5.4. Pengujian Model .....	92
4.5.5. Prediksi .....	93
4.5.6. Uji Validasi .....	93
4.6. Hasil Evaluasi .....	106
4.6.1. Menghitung Akurasi .....	107
4.6.2. Menghitung Precision .....	108
4.6.3. Menghitung Recall .....	109
BAB V PENUTUP .....	111
5.1. Kesimpulan .....	113
5.2. Saran .....	114
DAFTAR PUSTAKA .....	114

## DAFTAR TABEL

Tabel 2. 1. Matriks Literatur Review dan Posisi Penelitian.....	15
Tabel 4.1.1. Pengambilan Data Audio alat musik gamelan .....	16
Tabel 4.2.1. Dataset Suara Alat Musik Gamelan Jawa .....	17
Tabel 4.5.7.1 Data Pengujian Angklung .....	18
Tabel 4.5.7.2 Data Pengujian Bonang.....	19
Tabel 4.5.7.3 Data Pengujian Demung .....	20
Tabel 4.5.7.4 Data Pengujian Gong .....	69
Tabel 4.5.7.5 Data Pengujian Kendang.....	71
Tabel 4.5.7.6 Data Pengujian Rebab .....	95
Tabel 4.5.7.7 Data Pengujian Saron.....	96
Tabel 4.5.7.8 Data Pengujian Suling.....	97
Tabel 4.5.6. Hasil perbandingan data testing pada Angklung.....	97
Tabel 4.5.6.1. Hasil perbandingan data testing pada alat musik Bonang..	98
Tabel 4.5.6.2. Hasil perbandingan data testing pada alat musik Demung	98
Tabel 4.5.6.3. Hasil perbandingan data testing pada alat musik Gong.....	99
Tabel 4.5.6.4. Hasil perbandingan data testing pada alat musik Kendang	101
Tabel 4.5.6.5. Hasil perbandingan data testing pada alat musik Rebab...	101
Tabel 4.5.6.6. Hasil perbandingan data testing pada alat musik Saron...	102
Tabel 4.5.6.7. Hasil perbandingan data testing pada alat musik Suling..	103
Tabel 4.6. Hasil Confusion Matrix.....	103
Tabel 4.6.1. Hasil Akurasi Data .....	103
Tabel 4.6.2. Hasil Klasifikasi Precision .....	104
Tabel 4.6.2.1. Hasil dari pencarian FP dan Perhitungan Precision .....	105
Tabel 4.6.3.1 Perhitungan Recall dan All Recall .....	110

## DAFTAR GAMBAR

Gambar 2.3.2.2.1. Alat Musik Kedang .....	27
Gambar 2.3.2.2.2 Alat Musik Gong.....	28
Gambar 2.3.2.2.3 Alat Musik Rebab.....	29
Gambar 2.3.2.2.4 Alat Musik Saron .....	30
Gambar 2.3.2.2.5 Alat Musik Bonang .....	31
Gambar 2.3.2.2.6 Alat Musik Angklung.....	31
Gambar 2.3.2.2.7 Alat Musik Demung .....	31
Gambar 2.3.2.2.8 Alat Musik Demung.....	34
Gambar 3.1 Ilustrasi cakupan data mining.....	35
Gambar 2.2.5 Alur Metode MFCC.....	39
Gambar 2.2.5.1. Alur <i>frame</i> bloking pada sinyal audio.....	40
Gambar 2.2.5.2. Proses fitur MFCC pada audio.....	42
Gambar 2.2.6. proses <i>Convolutional Neural Network (CNN)</i> .....	44
Gambar 2.2.8. <i>MLP</i> .....	44
Gambar 2.2.8.1 <i>Convolutional Layer</i> .....	45
Gambar 2.2.8.2 <i>Feature Map</i> .....	46
Gambar 2.2.8.3. <i>Pooling Layer</i> .....	48
Gambar 2.2.11. Gelombang <i>Waveform</i> .....	52
Gambar 3.4. Alur Penelitian.....	59
Gambar 3.5 Alur Klasifikasi .....	60
Gambar 4.1. Perekaman Suara alat musik gamelan.....	68
Gambar 4.2.4. Perintah dan Hasil Pemeriksaan Audiotori .....	73
Gambar 4.2.5. Hasil Inspeksi Visual “Angklung” .....	74
Gambar 4.2.6. Perbedaan Hasil Visual “Bonang” .....	75
Gambar 4.2.7. Memuat File Metadata .....	75
Gambar 4.2.8. Tampilan Hasil Distribusi Kelas .....	76
Gambar 4.2.9. Properti File Sampel Audio.....	77
Gambar 4.2.10. Memunculkan Channel Audio.....	77

Gambar 4.2.11. Hasil Tampil <i>Sample Rate</i> .....	78
Gambar 4.2.12. Hasil Tampil <i>Bit Depth</i> .....	79
Gambar 4.3. Proses Ekstrasi MFCC dengan CNN .....	81
Gambar 4.4.1.2. Tahapan <i>Preprocessing Sample Rate</i> .....	83
Gambar 4.4.1.2.1. Tahapan <i>Preprocessing Bit Depth</i> .....	84
Gambar 4.4.2 Fitur Ekstrasi .....	85
Gambar 4.4.3. Mengonversi Data dan Label .....	85
Gambar 4.4.4. Memisahkan Dataset .....	86
Gambar 4.5.1. Penerapan Model MLP dengan <i>Python</i> .....	87
Gambar 4.5.2.1. Penerapan arsitektur CNN pada <i>Python</i> .....	88
Gambar 4.5.2.2. Hasil perubahan layer arsitektur CNN .....	89
Gambar 4.5.3. Training .....	90
Gambar 4.5.4 Pengujian Model .....	90
Gambar 4.5.5. Penerapan Prediksi audio dengan <i>python</i> .....	91
Gambar 4.5.6. Hasil Validasi .....	92
Gambar 4.5.6. Tes Audio Lain.....	93

## INTISARI

Perkembangan musik akhir-akhir ini sangat pesat. Sisi produksi musik yang mengalami perkembangan, juga alat-alat musik yang semakin banyak macamnya. Di Indonesia banyak sekali jenis alat musik yang menjadi ciri khas di setiap daerahnya, terlebih utama musik khas Jawa yang sekarang banyak menjadi pilihan pelaku produksi musik untuk mengolah musik mencampur dengan instrumen gamelan. *Deep Learning* memungkinkan inovasi dan perubahan di semua aspek kehidupan modern kita. Sebagian besar terobosan kecerdasan buatan yang dapat didengar di media didasarkan pada *deep learning*. Peneliti menggunakan dataset yang dibuat sendiri dengan memainkan alat musik gamelan yang direkam. Implementasi dari instrumen alat musik gamelan yaitu bonang, gambang dan kendang dianalisis menggunakan metode *convolutional neural network* (CNN) dengan *library keras dan librosa*. Dalam proses analisis menggunakan CNN melalui proses *preprocessing audio* dan selanjutnya dilakukan permodelan hingga didapatkan mode yang baik dalam data audio yang uji. Berdasarkan model yang diterapkan, hasil validasi data latih memiliki nilai *loss* 0.0749 dan nilai akurasi sebesar 0.9832 yang berasal dari hasil *training* dengan *epoch* sebanyak 100. Pada tabel prediksi dapat dilihat bahwa suara alat musik gamelan dari data asli ke audio lain adalah mendekati angka yang sama. Sehingga *confidence level* sebesar 0.925 dapat dipastikan valid dan sesuai dengan klasifikasi. Lebih Rinci dari hasil testing sebanyak 20 kali data masukan audio yang tidak berasal dari dataset adalah Angklung sebanyak 0.730142499; Bonang sebesar 0.983872325; Demung sebesar 0.572570784; Sedangkan Gong, Kendang, Rebab, Saron dan Suling masing-masing sebesar 0.999924088; 0.970878; 0.996331604; 0.52538331; 0.999755581. Dengan demikian, perhitungan akurasi berdasarkan *confusion matrix* yang telah dilaksanakan adalah sebesar 0.88125; dengan presisi sebesar 0,98388499 dan seluruh *recall* berjumlah 10,575

Kata kunci: *deep learning*, MFCC, analisis audio

## ***ABSTRACT***

The development of music has been very fast lately. The music production technique is experiencing development, as well as musical instruments of which there are more and more varieties. In Indonesia, there are many types of unique traditional musical instruments in each region. Javanese music instrument is a popular the choice of many music producers mixed. *Deep Learning* enables innovation and change in all aspects of our modern life. Much of the artificial intelligence breakthrough are based on *deep learning*. This research proposed using dataset by playing a recorded directly with gamelan instrument. The implementation of the gamelan instruments, namely Angklung, Bonang, Demung, Gong, Kendang, Rebab, Saron, Suling was analyzed using the method *convolutional neural network (CNN)* with *loud and librosa libraries*. In the analysis process using CNN through the process *audio preprocessing* and modeling the training and testing audio data. Based on the model obtained the *loss value* achieve 0.0749 and the accuracy at 0.9832 from an *epoch* of 100. In the prediction table it can be seen that the sound of gamelan instruments from original data to other audio is close to the same number. So that an *confidence level* of 0.925 can be ascertained valid and in accordance with the classification. More detailed from the test results as much as 20 times the audio input data that did not come from the dataset was Angklung as much as 0.730142499; Bonang of 0.983872325; Demung of 0.572570784; Meanwhile, Gong, Kendang, Rebab, Saron and Suling each amounted to 0.999924088; 0.970878; 0.996331604; 0.52538331; 0.999755581. So, the accuracy calculation based on the confusion matrix that has been implemented is 0.88125; with a precision of 0.98388499 and all recall amounts to 10.575

*Keyword: deep learning, MFCC, audio analysis*



# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Perkembangan musik akhir-akhir ini sangat pesat. Sisi produksi musik yang mengalami perkembangan, juga alat-alat musik yang semakin banyak macamnya. Di Indonesia banyak sekali jenis alat musik yang menjadi ciri khas di setiap daerahnya, terlebih utama musik khas Jawa yang sekarang banyak menjadi pilihan pelaku produksi musik untuk mengolah musik mencampur dengan instrumen khas Jawa, yaitu gamelan. Dunia audio pada era digital memang sangat erat kaitannya dengan musik, karena audio dapat diubah menjadi sinyal digital sedemikian rupa bagi para pelakunya menjadi sebuah musik yang terpadu. Musik yang terpadu akan dinikmati banyak pendengar, akan sangat baik jika pendengar mengetahui jenis-jenis alat musik khas Indonesia ini. Sebagai pondasi, istilah "audio" mengacu pada segala sesuatu yang berhubungan dengan suara, yang nantinya dapat diputar, diproses, dicampur, terutama akan didengar. Banyak sekali perkembangan format audio yang mengalami perkembangan yang sangat drastis dari waktu ke waktu.

Gamelan Jawa merupakan alat musik yang muncul dari sejarah kebudayaan Jawa yang di dalam perkembangannya selalu dipakai mengiringi pagelaran wayang maupun pengisi suatu pagelaran adat istiadat orang Jawa (Prasetyo, 2012). Seni gamelan Jawa memiliki nilai-nilai historis dan filosofis Bangsa Indonesia khususnya bagi masyarakat Jawa dan gamelan Jawa juga mempunyai fungsi estetika yang berkaitan dengan nilai-nilai sosial, moral dan spiritual (Purwadi, 2006). Mengutip

pernyataan Fatmawati (2017) gamelan jawa adalah ensambel musik yang biasanya menonjolkan metalofon, gambang, gendang, dan gong. Musik yang tercipta pada gamelan berasal dari perpaduan alat musik yang ada di dalamnya yang kemudian disebut musik gamelan jawa atau dalam Bahasa Jawa disebut karawitan.

Penyebutan suara komputer yang paling pertama dating dari penciptaan game yang primitive pada saat itu, di mana suara dimainkan melalui dinamika system. Tetapi tidak peduli seberapa keras pengembang perangkat lunak tersebut, dengan kualitas yang diperlukan, kompatibel dengan reel-to-reel atau kaset atau perekam, berusaha untuk mencapai, itu tidak mungkin. Itulah sebabnya banyak pelaku audio mencari solusi untuk mengubah format audio sehingga suaranya alami. Ini mengarah pada kompetisi lebih lanjut yang kita miliki sekarang. Ini tidak hanya berlaku untuk bahan yang direproduksi, tetapi juga untuk suara studio, pertunjukan langsung, kualitas atau pengaturan parameter dasar dari sudut pandang pengetahuan fisika, akustik, dll. Dipercayai bahwa kualitas format audio penuh pertama dikaitkan dengan munculnya standar dan ekstensi file .wav (singkatan ini berasal dari kata bahasa Inggris "gelombang" atau gelombang). Sama saja, ia menjadi anak sulung, yang dapat diproses dalam program komputer pada tingkat profesional. Dengan digital musik yang berkembang pada tingkat yang sangat potensial, kebutuhan untuk yang inovatif dan klasifikasi alat musik jawa menjadi lebih jelas. Alat musik khas Indonesia ini sangatlah baik ketika dapat dikenalkan ke masyarakat di seluruh dunia, yang akan membawa Indonesia semakin dikenal dunia.

Pertumbuhan database musik yang berkembang secara sangat pesat menyebabkan sukarnya proses pengelompokkan musik dalam kategori tertentu, sehingga dapat berakibat sulitnya pencarian suatu kategori musik dalam jumlah banyak dan skala yang besar. Selama ini pengklasifikasian jenis musik dapat dilakukan dengan dua cara, secara manual dan secara otomatis. Secara manual, manusia dapat mengklasifikasikan suatu jenis musik dengan cara mendengar langsung suatu musik baik hanya sepele bagian maupun secara keseluruhan. Kemampuan ini didapat dari pola pendengaran manusia selama jangka waktu tertentu. Kekurangan dari pengklasifikasian jenis musik secara manual ini adalah rawannya terjadi kesalahan pada saat pengklasifikasian yang dilakukan oleh orang awam. Hal ini disebabkan oleh perbedaan persepsi tiap individu dalam menentukan jenis suatu musik. Herry Sujaini (2019) pernah melakukan penelitian analisis citra alat musik tradisional dengan metode KNN, Random Forest, dan SVM yang mana dari hasil eksperimennya sistem pengklasifikasi 10 jenis citra alat musik tradisional Indonesia dengan metode tersebut yang memiliki akurasi hingga 92,1%. Ini yang menjadi tolok ukur penting penulis untuk melakukan analisis klasifikasi alat musik Gamelan Jawa. Pada penelitian yang dilakukan Muhammad Arif Maula (2018) tentang karakteristik suara peking pada gamelan yang mana terjuntuk data klasifikasi dengan frekuensi. Pada penelitian ini didapatkan hasil frekuensi dominan peking slendro yang ada di Yogyakarta yaitu pada bilah 1, 1082 – 1103 Hz, bilah 2, 1237 – 1267 Hz, bilah 3, 1432 – 1469 Hz, bilah 5, 1642 – 1686 Hz, bilah 6, 1902 – 1925 Hz, bilah i, 2155 – 2238 Hz. Untuk Peking Pelog yaitu pada bilah 1, 1102 – 1178 Hz, bilah 2, 1177 – 1282 Hz, bilah 3, 1276 – 1400 Hz, bilah

4, 1515 – 1640 Hz, bilah 5, 1603 – 1778 Hz, bilah 6, 1714 – 1865 Hz, bilah 7, 1868 – 2118 Hz. Untuk range kerja frekuensi peking slendro yaitu antara, 1082 Hz – 2238 Hz dan peking pelog memiliki range frekuensi antara, 1102 Hz – 2118 Hz. Hasil data klasifikasi ini menunjukkan frekuensi saja dimana frekuensi itu bisa terdengar dan terbentuk untuk alat musik lainnya. Penelitian Diah P. Wulandari (2013) juga melakukan analisis dengan fitur-fitur spectral *phase slope* (PS), *weighted phase deviation* (WPD), *spectral flux* (SF), dan *Short-time Fourier Transform* (STFT) dengan data hasil Bonang dengan metode PS dengan *presisi, recall, f-measure* sebesar 0.86, 0.85, 0.86. metode WPD sebesar 0.62, 0.62, 0.62, SF sebesar 0.76, 0.75, 0.75 dan RCD sebesar 0.74, 0.75, 0.75. Demung dengan metode PS sebesar 0.91, 0.93, 0.92, WPD 0.43, 0.43, 0.43. Dengan performansi keseluruhan dari metode adalah PS 0.91, WPD 0.65, SF 0.80, RCD 0.78 yang tertunjuk dari saron peking. Dari hasil tersebut kami dapat mengambil tiga kesimpulan. Pertama, PS menghasilkan lebih baik kinerja menggunakan jendela panjang sementara yang lain (WPD, SF, dan RCD) bekerja lebih baik menggunakan jendela pendek. Penggunaan metode ini sangat kurang untuk mengumpulkan dataset alat musik gamelan, tidak semua metode dapat mencapai titik akurasi terbesarnya. Maka dari itu kami menggunakan MFCC dengan CNN untuk melakukan klasifikasi alat musik gamelan jawa.

Menilik dari perbandingan dengan beberapa metode dan fitur penelitian terhadap audio guna melakukan klasifikasi terhadap audio, yang pertama adalah dari penelitian Sita Purnama Dewi (2019) yang menganalisis ekstrasi fitur LFCC (Linear Frequency Cepstral Coefficient) pada klasifikasi suara bayi menangis

menggunakan KNN (K-Nearest Neighbour) dilakukan analisis perubahan pra-penekanan, jumlah bank filter dan jumlah cepstral. Pemilihan nilai bank filter yang diterapkan harus lebih besar dari nilai cepstral yang diterapkan, metode LFCC tidak memberikan pengaruh yang signifikan terhadap proses klasifikasi tetapi merekomendasikan untuk menggunakan pra-penekanan untuk menghindari data suara dengan kualitas rendah. Dengan ini penelitian tentang suara dengan MFCC dan CNN akan lebih efektif, karena batas frekuensi dan resolusi suara pada data yang akan diteliti bias lebih detail, sedangkan dengan LFCC dan KNN kurang bisa detail, LFCC dan KNN ini hanya hingga batas 400-500 Hz. Sedangkan penelitian kami bisa mencapai kualitas 22.000 Hz. Berikutnya adalah dari penelitian tahun 2007 oleh Giannakopoulos yang mengklasifikasikan audio dengan Bayes menggunakan KNN. Hasilnya klasifikasi audio dengan cara ini hanya mencakup sample 16KHz. Penelitian kami bisa mencakup 22Khz yang dapat menghasilkan klasifikasi dengan audio kualitas lebih tinggi dari pada cara *bayes knn*. Lalu yang ketiga adalah analisis klasifikasi audio menggunakan *fuzzy c-means* (FCM) yang hanya menyediakan pembacaan sampel audio SP dan MS yang mana akan sangat cocok dihitung klasifikasinya menggunakan cara *fuzzy c-means*, sedangkan penelitian kami menggunakan format *.wav* yang dalam pengambilan audionya tidak banyak diterapkan filtering, sehingga tidak banyak juga kompresi untuk sampel yang kita ambil. Jadi inilah alasan yang jadi penentu kami dalam melakukan penelitian klasifikasi audio menggunakan MFCC dengan CNN.

*Deep Learning* memungkinkan inovasi dan perubahan di semua aspek kehidupan modern kita. Sebagian besar terobosan kecerdasan buatan yang dapat

didengar di media didasarkan pada *deep learning*. Hasilnya, apakah kita seorang pebisnis yang tertarik untuk meningkatkan efisiensi organisasi kita, pembuat kebijakan yang peduli dengan etika dan privasi di dunia Big Data, peneliti yang bekerja dengan data kompleks, atau warga negara yang ingin tahu yang menginginkan pemahaman yang lebih baik tentang potensi kecerdasan buatan dan bagaimana hal itu akan mengubah hidup kita, penting bagi Anda untuk memiliki pemahaman tentang deep learning. Fakta bahwa deep learning adalah sekumpulan algoritme dan model berarti bahwa memahami pembelajaran dalam memerlukan pemahaman tentang bagaimana algoritme dan model ini memproses data. Deep learning akan digunakan dalam klasifikasi alat musik Jawa ini karena data yang kompleks akan diolah secara mendalam. Dengan perkembangan jaman, sekarang telah mulai dikembangkan metode agar suatu audio file dapat dikenali secara otomatis dari fitur-fitur yang telah diekstrak sebelumnya dengan bantuan MFCC (*Mel Frequency Cepstral Coefficients*), MFCC (*Mel Frequency Cepstral Coefficients*) adalah koefisien yang mewakili audio karena frekuensi, yang ekstraksi fitur dalam proses ini ditandai dengan konversi data suara menjadi gambar spektrum gelombang. File audio dilakukan oleh MFCC sehingga objek suara dapat dikonversi menjadi matriks, suara akan menjadi vector yang akan diproses sebagai output untuk mempermudah mesin dalam membaca. Salah satu metode yang penting dalam deep learning adalah Convolutional Neural Network (CNN). CNN dirancang khusus untuk pengenalan dan klasifikasi gambar. CNN memiliki beberapa lapisan (*layer*) yang mengekstrak informasi dari gambar dan menentukan klasifikasi dari gambar tersebut. MFCC dan CNN digunakan menjadi rekomendasi

penulis untuk melakukan penelitian. Danny Lionel (2019) juga melakukan penelitian klasifikasi pada audio yang menggunakan CNN dan MFCC. Sangatlah penting bagi penulis untuk melakukan analisis klasifikasi alat musik gamelan Jawa dengan CNN dan MFCC karena dalam implementasinya tingkat akurasi saat testing pada pembuatan convolutional neural network model sangat baik dan dalam pengujiannya banyaknya dataset, iterasi training, dan spesifikasi komputer sangat mempengaruhi tingkat akurasi dan lama pembuatan neural network model yang optimal. Tujuan dari digunakan CNN dengan MFCC adalah mengekstraksi fitur untuk memampatkan sinyal audio ke vector yang mewakili informasi yang terdapat dalam audio yang akan diekstraksi. Mungkin bagi manusia ini terlihat sederhana karena kita dapat mengenali suatu suara yang dicerna oleh otak kita yang memungkinkan kita dengan mudah mengaitkan jenis jenis suara. Dalam input CNN diproses oleh unit komputasi kecil atau disebut neuron yang diatur dalam struktur berlapis. Sebuah neuron akan beroperasi dengan menerapkan pemfilteran non-linier ke sejumlah inputnya dengan parameter jaringan yang dipelajari selama pengujian audio dapat diinginkan bahwa pola energi dalam ekstraksi dikenali. Mengapa dalam penelitian ini menggunakan CNN untuk proses klasifikasi audio, karena CNN memiliki fitur yang efektif untuk memperkirakan Fundamental Frequency yang mana akan membaca dengan baik proses dari raw waveform yang diinputkan. Karena CNN haus akan gambar, kami ingin mengubah suara menjadi gambar. Sinyal audio juga dapat direpresentasikan dengan cara lain. Alih-alih memplot amplitudo sinyal audio terhadap waktu, kita juga dapat memplotnya sehubungan dengan frekuensi. Plot yang akan kita buat disebut spektrogram.

Beberapa penelitian yang menggunakan citra sebagai data dan metode *convolutional neural network* seperti penelitian klasifikasi citra menggunakan *convolutional neural network (cnn)* pada *caltech 101* milik Suartika, implementasi *deep learning* menggunakan *convolutional neural network* untuk klasifikasi citra candi berbasis GPU milik Kefin Pudi D, implementasi *convolutional neural network* untuk sortasi mutu salak ekspor berbasis citra digital milik Rismiyati dan penelitian lain tentang *convolutional neural network* menggunakan data citra menjadi pertimbangan bagi peneliti untuk melakukan penelitian tentang implementasi *convolutional neural network* terhadap instrument alat musik gamelan menggunakan *keras*, akan tetapi peneliti menggunakan dataset yang dibuat sendiri dengan memainkan alat musik gamelan yang direkam. Tetapi belum ada pengklasifikasian jenis alat musik khas Indonesia. Biasanya, jenis alat musik khas Indonesia ini hanya dapat dikenali jika memang mengenal alat musik ini. Maka dari itu, penelitian ini diharapkan menjadikan terobosan baru dalam mengklasifikasikan jenis alat musik Jawa di Indonesia.

## 1.2. Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut :

- a. Bagaimana klasifikasi alat musik gamelan Jawa berbasis *wav* dengan ekstraksi MFCC dengan metode CNN?
- b. Bagaimana hasil akurasi implementasi klasifikasi *deep learning* terhadap audio pada alat musik gamelan Jawa yang berupa angklung, bonang, demung, gong, kendang, rebab, saron, suling?



### 1.3. Batasan Masalah

Untuk dapat mencapai tujuan dari penelitian ini secara terarah maka ditentukan batasan masalah variabel sebagai berikut :

- a. Penelitian ini hanya mengimplementasikan audio berbasis *.wav*.
- b. Penelitian ini menggunakan musik nuansa Jawa atau gamelan jawa.
- c. Dalam pelaksanaan implementasi klasifikasi, penelitian ini menggunakan Bahasa *Python*.
- d. Dataset instrumen gamelan jawa diambil dan dibuat dari sampling musik yang dibuat atau dimainkan oleh penulis itu sendiri.
- e. Instrumen musik gamelan jawa yang akan diklasifikasikan adalah angklung, bonang, demung, gong, kendang, rebab, saron, suling.
- f. Klasifikasi alat musik gamelan jawa berbasis *wav* dengan ekstraksi MFCC dengan metode CNN.
- g. Audio yang akan diekstrasi adalah berupa *wav* dengan resolusi 16 bit dan *sample rate* 11-22 khz.

### 1.4. Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

- a. Mengetahui apa saja alat musik gamelan dengan audio berformat *.wav*.
- b. Mengenalkan budaya Indonesia ke masyarakat dari musik gamelan jawa berbasis teknologi.
- c. Menciptakan dataset audio alat musik gamelan jawa yang akan bermanfaat untuk penelitian-penelitian selanjutnya.

- d. Mengembangkan sebuah penelitian berbasis CNN dengan MFCC yang digunakan dalam mengidentifikasi suara alat musik gamelan Jawa, sehingga menunjukkan performansi dari klasifikasi yang diterapkan menggunakan metode tersebut.
- e. Sebagai syarat untuk menyelesaikan pendidikan Magister Teknik Informatika di Universitas AMIKOM Yogyakarta.

### 1.5. Manfaat Penelitian

Manfaat dari penelitian yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

#### a. Bagi Penulis

Sebagai syarat kelulusan S2 di Magister Teknik Informatika Universitas AMIKOM Yogyakarta dan menerapkan teori yang telah diperoleh selama mengikuti perkuliahan di Universitas AMIKOM Yogyakarta dan sejauh mana kemampuan menulis memberikan solusi pada suatu masalah yang dihadapi.

#### b. Bagi Masyarakat

Dapat mengenalkan kepada masyarakat luas dan tentunya budaya Indonesia semakin dikenal luas dengan adanya klasifikasi instrumen musik gamelan Jawa yang berbasis teknologi serta Menciptakan dataset audio alat musik gamelan Jawa yang akan bermanfaat untuk penelitian-penelitian selanjutnya.

## BAB II

### TINJAUAN PUSTAKA

#### 2.1. Tinjauan Pustaka

Beberapa penelitian yang menjadi referensi dalam penelitian yang penulis lakukan merujuk dari beberapa dari penelitian terdahulu yang sedikit banyak ada keterkaitan metode dan topik untuk peneliti selanjutnya baik secara langsung maupun tidak langsung.

Farrokhmanesh (2018) dalam penelitiannya tentang klasifikasi musik untuk deteksi *malware* menurut kami telah berhasil melakukan klasifikasi jenis *malware* pada aplikasi virus dengan musik yang digunakan untuk mendeteksi *malware* yang didasarkan pada fitur tingkat tinggi opcode, pemanggilan fungsi, atau grafik aliran kontrol program (CFG). Dalam metode yang diusulkan peneliti sebelumnya ini, *byte* program diubah menjadi sinyal audio yang berarti, kemudian teknik *Music Information Retrieval* (MIR) digunakan untuk membuat model klasifikasi musik pembelajaran mesin dari sinyal audio untuk mendeteksi instance baru dan yang tidak terlihat. Eksperimen mengevaluasi pengaruh berbagai strategi yang mengubah *byte* menjadi sinyal audio dan keefektifan metode. Ini sangat berbeda dengan yang dilakukan pada penelitian kami. Jika Farrokhmanesh (2018) membuat musik untuk menandai adanya *malware*, yang kami lakukan adalah mengklasifikasikan alat musik gamelan Jawa sebagai pengenalan. Lalu peneliti sebelumnya ini menggunakan teknik MIR sebagai model klasifikasinya, jika kami adalah menggunakan MFCC (*Mel Frequency Cepstral Coefficients*) dan CNN

*Convolutional Neural Network* sebagai metode dan model untuk mengklasifikasikan suara alat musik gamelan jawa.

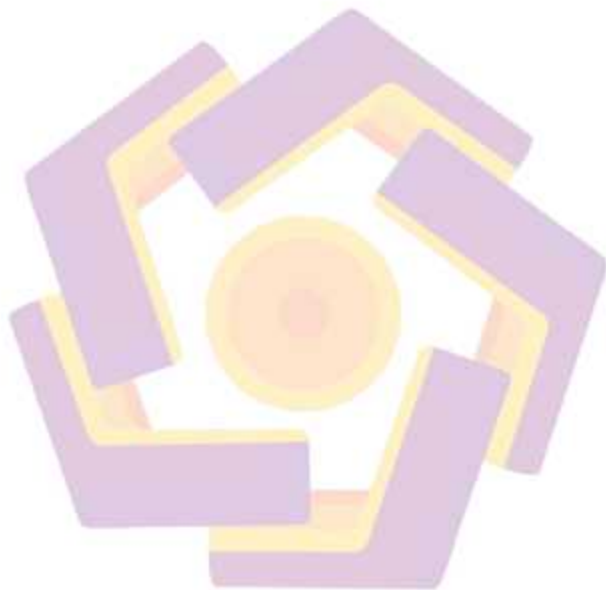
Cory McKay (2010) pada penelitiannya tentang klasifikasi musik otomatis dengan jMIR telah berhasil melakukan penerapan klasifikasi musik otomatis yang mana bidang penyelidikan yang luas dan multidisiplin yang menawarkan manfaat signifikan dari perspektif akademis dan komersial. Disertasi ini berfokus pada pengembangan jMIR, seperangkat alat perangkat lunak yang kuat, fleksibel, dapat diakses, dan asli yang dapat digunakan untuk merancang, berbagi, dan menerapkan berbagai teknologi klasifikasi musik otomatis. Menurut kami itu sangat bagus tetapi hanya untuk pengembangan metode yang telah digunakan. Untuk menggolongkan jenis musik pada dasarnya sudah banyak digunakan, bahkan tidak perlu menggunakan mesin, file musik yang musisi buat pun sudah mencantumkan data tentang jenis musiknya pada metadatanya. Penelitian kami akan mengklasifikasikan suara alat musik gamelan jawa yang pastinya sangat berbeda dari peneliti sebelumnya, karena kami mengambil data dari alat musik gamelan yang sudah peneliti rekam sebelumnya dengan menggunakan *software* perekam yang dijadikan dataset untuk penelitian terbaru yang kami lakukan.

Peneliti lainnya Jiaxing Ye (2016) dan Mike Smales (2018) adalah tentang klasifikasi suara perkotaan. Ini bertujuan untuk mengembangkan skema berbasis pembelajaran mesin yang efisien untuk klasifikasi suara perkotaan dalam kondisi kebisingan kehidupan nyata. Tidak seperti metode klasifikasi peristiwa suara konvensional yang utamanya membahas pola spektral temporal lokal, Jiaxing Ye (2016) mengusulkan skema agregasi untuk menggabungkan fitur akustik lokal dan

global, Mike Smales (2018) mengklasifikasikan suara urban dengan *Deep Learning*. Menurut kami ini akan membutuhkan waktu yang sangat lama untuk melakukan penelitian, dikarenakan dataset suara haruslah sangat banyak, sehingga untuk keefektifan penelitiannya kurang. Penelitian kami akan mengklasifikasikan suara alat musik gamelan jawa yang pastinya sangat berbeda dari peneliti sebelumnya, karena kami mengambil data dari alat musik gamelan yang sudah peneliti rekam sebelumnya dengan menggunakan *software* perekam yang dijadikan dataset untuk penelitian terbaru yang kami lakukan. Sehingga penelitian kami akan sangat bermanfaat dan menambah citra pengenalan budaya Indonesia semakin dikenal baik oleh dunia. Penelitian kami juga menggunakan *Deep Learning* tetapi kami menggunakan CNN yang kami pastikan akan lebih efektif untuk kinerja penelitian ini.

Penelitian lain oleh Nilesh (2019) tentang Klasifikasi Audio Berbasis Konten dan Pengambilan Menggunakan Segmentasi, Ekstraksi Fitur dan Pendekatan Jaringan Saraf Tiruan. Penelitian ini dilakukan karena Volume data audio meningkat pesat setiap hari di jaringan publik seperti Internet. Menurut kami hal ini meningkatkan kesulitan dalam mengakses data audio tersebut. Oleh karena itu, diperlukan mekanisme pengindeksan dan anotasi yang efisien. Non-stasioneritas dan diskontinuitas yang ada dalam sinyal audio meningkatkan kesulitan dalam segmentasi dan klasifikasi sinyal audio. Sangat menyita waktu. Penelitian kami akan mengklasifikasikan suara alat musik gamelan jawa yang pastinya sangat berbeda dari peneliti sebelumnya, karena kami mengambil data dari alat musik gamelan yang sudah peneliti rekam sebelumnya dengan menggunakan *software*

perekam yang dijadikan dataset untuk penelitian terbaru yang kami lakukan. Sehingga penelitian kami akan sangat bermanfaat dan menambah citra pengenalan budaya Indonesia semakin dikenal baik oleh dunia. Penelitian kami juga menggunakan *Deep Learning* tetapi kami menggunakan CNN yang kami pastikan akan lebih efektif untuk kinerja penelitian ini.



## 2.2. Keaslian Penelitian

Tabel 2. 1. Matriks Literatur Review dan Posisi Penelitian

### Analisis Audio : Klasifikasi Alat Musik Gamelan Jawa Menggunakan Deep Learning Dengan Python

No	Judul	Peneliti	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1.	Klasifikasi Musik Sebagai Pendekatan Baru Untuk Deteksi Malware	Farrokhmanesh, Ali Hamzeh, <i>Journal of Computer Virology and Hacking Techniques</i> , 2018	Mendeteksi malware dengan teknik pemrosesan sinyal audio yang disajikan.	Metode tingkat byte baru untuk mendeteksi malware dengan teknik pemrosesan sinyal audio disajikan. Dalam metode yang kami usulkan, byte program diubah menjadi sinyal audio yang berarti, kemudian teknik Music Information Retrieval (MIR) digunakan untuk membuat model klasifikasi musik pembelajaran mesin dari sinyal audio untuk mendeteksi instance baru dan yang tidak terlihat.	Tantangan paling penting adalah memicu algoritme yang mengesankan untuk mengekstrak fitur dari file yang dapat dieksekusi yang dapat menjadi representasi yang baik dari konten file. Banyak cara tersedia untuk meningkatkan metode ini. Pertama, strategi baru untuk memetakan file biner ke sinyal audio dapat dikembangkan yang dapat meningkatkan akurasi akhir.	Pada penelitian kami, menggunakan <i>deep learning</i> yang penerapannya digunakan untuk mendeteksi klasifikasi alat musik jawa gamelan. Penelitian yang kami lakukan adalah mendeteksi file audio berformat .wav, jika pada penelitian sebelumnya adalah membuat deteksi <i>malware</i> dengan musik yang dibuat dengan MIDI <i>Instrument</i> .

Tabel 2.1. Matriks Literatur Review Dan Posisi Penelitian

Analisis Audio : Klasifikasi Alat Musik Gamelan Jawa Menggunakan Deep Learning Dengan Python (Lanjutan)

No.	Judul	Peneliti, Publikasi, dan Tahun	Media dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
2.	Klasifikasi Peristiwa Suara Perkotaan Berdasarkan Agregasi Fitur Lokal Dan Global	Jiaxing Ye, Takumi Kobayashi, Masahiro Murakawa, Elsevier Journal, 2016		Makalah ini bertujuan untuk mengembangkan skema berbasis pembelajaran mesin yang efisien untuk klasifikasi suara perkotaan dalam kondisi kebisingan kehidupan nyata. Tidak seperti metode klasifikasi peristiwa suara konvensional yang utamanya membahas pola spektral temporal local.	Pemahaman yang lebih baik tentang suara perkotaan dapat sangat berkontribusi pada kelayakan hidup perkotaan, seperti kontrol kebisingan yang lebih efisien, pengawasan keselamatan yang andal, dan perencanaan lingkungan akustik yang lebih baik.	Untuk mengumpulkan kekuatan diskriminatif di kedua ruang fitur untuk klasifikasi, kami mengadopsi skema fusi bersyarat kelas yang memperkirakan kontribusi relatif dari dua fitur tingkat sehubungan dengan setiap jenis suara perkotaan dengan cara yang diawasi.	Pada penelitian yang akan dilakukan nanti akan diimplementasikan menggunakan <i>deep learning</i> yang penerapannya digunakan untuk mendeteksi klasifikasi alat musik jawa gamelan. Penelitian yang kami lakukan adalah mendeteksi file audio berformat .wav. Audio yang akan jadi bahan klasifikasi adalah berasal dari alat musik gamelan



Tabel 2.1. Matriks Literatur Review Dan Posisi Penelitian

Analisis Audio : Klasifikasi Alat Musik Gamelan Jawa Menggunakan Deep Learning Dengan Python (Lanjutan)

No.	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan.
3.	Deteksi Nada Tunggal Alat Musik Kecapi Bugis Makassar Menggunakan Metode Mel Frequency Cepstral Coefficient (MFCC) Dan Klasifikasi K-Nearest Neighbour (KNN)	Alif Fajri Ryamizard, e-Proceeding of Engineering, 2018	Kecapi sering digunakan dalam festival musik Sulawesi di berbagai daerah, namun sering terdapat permasalahan pada saat penyetalan alat musik kecapi karena membutuhkan waktu cukup lama. Pada Tugas Akhir ini telah dirancang sebuah sistem yang dapat mengidentifikasi nada yang terdapat pada alat musik kecapi melalui pengolahan suara.	<ol style="list-style-type: none"> <li>1. Sistem ini mampu mengidentifikasi asi nada alat musik kecapi secara real time menggunakan metode MFCC dan klasifikasi K-Nearest Neighbor.</li> <li>2. Akurasi sistem yang paling terbaik didapatkan ketika nilai windowing 120 dengan akurasi 81.42% distance den</li> </ol>	<ol style="list-style-type: none"> <li>1. Dilakukan penelitian dengan menggunakan ekstraksi ciri yang lain sebagai pembanding untuk melihat akurasi terbaik.</li> <li>2. Digunakan objek penelitian berupa alat musik tradisional yang lain seperti angklung, seruling, gendang, dan lain-lain, kemajuan dan kelestarian budaya asli bangsa Indonesia</li> </ol>	Perbandingannya adalah pada penelitian yang akan dilakukan nanti akan diimplementasikan menggunakan <i>deep learning</i> yang penerapannya digunakan untuk mendeteksi klasifikasi alat musik jawa gamelan. Penelitian yang kami lakukan adalah mendeteksi file audio berformat .wav. dari alat musik gamelan yang dimainkan peneliti yang dibuat menjadi data set.

Tabel 2.1. Matriks Literatur Review Dan Posisi Penelitian

Analisis Audio : Klasifikasi Alat Musik Gamelan Jawa Menggunakan Deep Learning Dengan Python (Lanjutan)

No.	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
4.	Klasifikasi Genre Musik Menggunakan Metode Deep Learning Convolutional Neural Network dan Mel Spektrogram	Danny Lionel, Rudy Adipranata, Endang Setyati, Jurnal Infra, 2019	Dalam penelitian ini akan digunakan metode Mel-spektrogram. Dimana Mel spektrogram merupakan hasil pemetaan fitur yang telah diambil oleh metode MFCC, yang akan diklasifikasikan dan dimasukkan kedalam Convolutional Neural Network. Yang akan dibedakan activation function nya yaitu ReLU dan ELU gar suatu audio file dapat dikenali secara otomatis	Tingkat akurasi saat <i>testing</i> pada pembuatan <i>convolutional neural network model</i> sangat bergantung pada lama <i>training</i> dan banyak <i>dataset</i> yang dimasukkan kedalam <i>neural network</i> . Tingkat akurasi tertinggi dicapai pada epoch ke 60, dengan tingkat akurasi sebesar 99% untuk RELU dan 95% untuk ELU.	Dapat dilakukan penambahan dataset sehingga akan memperluas jumlah variasi yang dapat dikenali neural network model. Lama training dapat diperpanjang untuk mendapatkan model yang optimal. Menggunakan mesin komputer yang lebih baik sehingga waktu training dapat dipersingkat. Dapat dilakukan modifikasi pada struktur neural network	Perbandingannya adalah pada penelitian yang akan dilakukan nanti sama-sama akan diimplementasikan proses <i>deep learning</i> yang penerapannya digunakan untuk mendeteksi klasifikasi alat musik jawa gamelan. Penelitian yang kami lakukan adalah mendeteksi file audio berformat .wav. Audio yang akan jadi bahan klasifikasi adalah berasal dari alat musik gamelan yang dimainkan peneliti yang dibuat menjadi data set.

Tabel 2.1. Matriks Literatur Review Dan Posisi Penelitian

Analisis Audio : Klasifikasi Alat Musik Gamelan Jawa Menggunakan Deep Learning Dengan Python (Lanjutan)

No.	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
5.	Arsitektur CNN Untuk Klasifikasi Audio Skala Besar	Shawn Hershey, IEEE, 2017	Tujuan penelitian ini adalah untuk menyelidiki berbagai ukuran dari set pelatihan dan kosa kata label, menemukan bahwa analog dari CNN yang digunakan dalam klasifikasi gambar bekerja dengan baik pada tugas klasifikasi audio kami, dan pelatihan yang lebih besar serta set label membantu sampai titik tertentu.	Jaringan gambar yang canggih mampu memberikan hasil yang sangat baik pada klasifikasi audio bila dibandingkan dengan jaringan sederhana yang terhubung sepenuhnya atau arsitektur klasifikasi gambar sebelumnya.	Melakukan penelitian harus dengan objektif dan dengan data yang valid, sehingga tidak menimbulkan rasa subjektif yang ada dalam penelitian ini walaupun dalam penelitian ini terdapat banyak aspek seni yang dinilai secara subjektif	Perbandingannya adalah pada penelitian yang akan dilakukan nanti diimplementasikan proses <i>deep learning</i> yang pencrapannya digunakan untuk mendeteksi klasifikasi alat musik jawa gamelan. Penelitian yang kami lakukan adalah mendeteksi file audio berformat .wav. Audio yang akan jadi bahan klasifikasi adalah berasal dari alat musik gamelan yang dimainkan peneliti yang dibuat menjadi data set. Data set yang dibuat akan dimasukan ke dalam program dengan Bahasa <i>Python</i> yang mana akan terlihat akurasi klasifikasi. Peneliti sebelumnya menggunakan CNN untuk klasifikasi, sedangkan peneliti selanjutnya akan menggunakan <i>deep learning</i> .

Tabel 2.1. Matriks Literatur Review Dan Posisi Penelitian

Analisis Audio : Klasifikasi Alat Musik Gamelan Jawa Menggunakan Deep Learning Dengan Python (Lanjutan)

No.	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
6.	Klasifikasi Dan Retrieval Musik Berdasarkan Genre.	Gst. Ayu Vida Mastrika Giri. Jurnal Ilmiah Ilmu Komputer Universitas Udayana, 2017	Untuk mengetahui fitur-fitur akustik musik yang dapat digunakan untuk klasifikasi berdasarkan genre. Teknik-teknik klasifikasi yang dapat digunakan untuk mengimplementasi klasifikasi genre otomatis.	Fitur-fitur akustik musik yang dapat digunakan untuk klasifikasi berdasarkan genre adalah timbre, rhythm, melody/harmony, dan pitch. Fitur akustik musik terbaik yang dapat digunakan adalah Mel Frequency Cepstrum Coefficients (MFCC). Teknik-teknik klasifikasi yang dapat digunakan untuk mengimplementasi klasifikasi genre otomatis adalah sistem pakar, klasifikasi supervised, dan klasifikasi unsupervised.	Tidak adanya system yang akan membedakan atau mengklasifikasi genre musik. Tidak ada objek yang akan diklasifikasi kan contoh dalam audio berformat seperti apa.	Perbandingannya adalah pada penelitian yang akan dilakukan nanti sama-sama akan diimplementasikan proses <i>deep learning</i> yang penerapannya digunakan untuk mendeteksi klasifikasi alat musik jawa gamelan. Penelitian yang kami lakukan adalah mendeteksi file audio berformat .wav. Audio yang akan jadi bahan klasifikasi adalah berasal dari alat musik gamelan yang dimainkan peneliti yang dibuat menjadi data set. Data set yang dibuat akan dimasukkan ke dalam program dengan Bahasa <i>Python</i> yang mana akan terlihat akurasi klasifikasi.

## **2.3. Landasan Teori**

### **2.3.1. Pengertian Klasifikasi**

Menurut Ibrahim Bafadal (2009), klasifikasi adalah suatu proses memilih dan mengelompokkan buku-buku perpustakaan sekolah atau bahan pustaka lainnya atas dasar tertentu serta diletakkannya secara bersama-sama di suatu tempat. Menurut Sulisty Basuki (1991) klasifikasi berasal dari kata Latin “classis” atau proses pengelompokan, artinya mengumpulkan benda/entitas yang sama serta memisahkan benda/entitas yang tidak sama. Sulisty-Basuki (1999:298) mendefinisikan klasifikasi yang diterapkan di pusat informasi dan perpustakaan adalah penyusunan sistematik terhadap buku atau bahan pustaka lain atau katalog atau entri indeks berdasarkan subjek, dalam cara yang paling berguna bagi mereka yang membaca atau mencari informasi.

### **2.3.2. Gamelan Jawa**

Indonesia merupakan negara yang besar dan kaya dengan budaya berbagai tradisi musik, salah satu musik khas Indonesia yang telah diakui oleh dunia, yakni gamelan Jawa. Gamelan berasal dari kata *nggamel* (dalam bahasa Jawa)/*gamel* yang berarti memukul/menabuh, diikuti akhiran “an” yang menjadikannya sebagai kata benda (Ferdiansah, 2010). Gamelan Jawa merupakan alat musik yang muncul dari sejarah kebudayaan Jawa yang di dalam perkembangannya selalu dipakai mengiringi pagelaran wayang maupun pengisi suatu pagelaran adat istiadat orang Jawa (Prasetyo, 2012). Seni gamelan Jawa memiliki nilai-nilai historis dan filosofis Bangsa Indonesia khususnya bagi masyarakat Jawa dan gamelan Jawa juga

mempunyai fungsi estetika yang berkaitan dengan nilai-nilai sosial, moral dan spiritual (Purwadi, 2006). Mengutip pernyataan Fatmawati (2017) gamelan Jawa adalah ensambel musik yang biasanya menonjolkan metalofon, gambang, gendang, dan gong. Musik yang tercipta pada gamelan berasal dari perpaduan alat musik yang ada di dalamnya yang kemudian disebut musik gamelan Jawa atau dalam Bahasa Jawa disebut karawitan. Mengacu pada hal-hal yang telah dijelaskan di atas, dapat ditarik kesimpulan bahwa pengertian mendengarkan musik gamelan Jawa adalah suatu aktivitas menangkap rangsangan berupa suara yang dihasilkan oleh harmonisasi alat musik gamelan Jawa berupa metalofon, gambang, gendang dan gong dengan menggunakan indera pendengaran untuk tujuan tertentu.

#### **2.3.2.1. Unsur-unsur dalam Gamelan**

Dalam musik gamelan terdapat enam unsur pokok, yaitu:

a. Laras (tangga nada atas scale)

Secara umum laras atau tangga nada dibedakan menjadi dua yaitu pentatonis dan diatonis. Hampir setiap jenis musik menggunakan laras tertentu, hanya saja dibedakan lagi dengan adanya larasan (daerah wilayah/ambitus suara) maupun embat (pergeseran/ variasi interval).

Larasan berasal dari kata laras yang mempunyai arti, pertama, indah menarik hati (Jawa: *nengsemake*), suara dan lagu dalam gamelan. Istilah laras dalam karawitan Jawa mempunyai makna sesuatu yang bersifat enak atau nikmat untuk dihayati. Kedua, laras bermakna sebagai nama nada dalam bilah gamelan yang telah ditentukan jumlah frekuensinya. Ketiga, adalah tangga nada yaitu

susunan nada yang jumlah, urutan dan pola interval nada-nadanya telah ditentukan. Laras dalam karawitan Jawa ada dua macam yaitu laras slendro dan laras pelog.

Laras slendro adalah tangga nada yang mempunyai pola interval relatif sama panjang (sama rata) antara nada satu dengan nada berikutnya. Laras slendro memiliki urutan nada-nada yang terdiri dari lima nada dalam satu gembyang (istilah gembyang adalah untuk menyebut nada yang sama tetapi berbeda frekuensinya) dengan pola jarak yang hampir sama rata. Susunan dan pola interval itu diatur sebagai berikut:

1 2 3 5 6 1

Ada pun nama nada yang digunakan dalam laras slendro adalah:

1. Penunggal (barang) yang diberi simbol angka satu (1) dan dibaca siji disingkat dengan ji.
2. Gulu (jangga), diberi simbol angka dua (2) dibaca loro atau ro.
3. Dhadha (jaja), diberi simbol angka tiga (3) dibaca telu atau lu.
4. Lima, diberi simbol angka lima (5) dibaca lima atau ma.
5. Enem, diberi simbol angka enam (6) dibaca nem.

Perlu diketahui bahwa di dalam laras slendro terdapat sub laras yaitu slendro pathet nem yang mempunyai teba nada rendah, slendro pathet sanga yang mempunyai teba nada sedang, dan slendro manyura yang mempunyai teba nada tinggi. Sub-sub laras ini akan sangat berperan bila dikaitkan dengan permainan instrument gamelan.

Laras pelog dalam karawitan Jawa adalah tangga nada yang mempunyai pola interval jauh dan dekat yang hampir menyerupai tangga nada diatonic. Ada

dua pendapat tentang laras pelog, pertama: sistem urutan nada-nada yang terdiri dari tujuh nada dalam satu gembyang yang menggunakan pola jarak yang tidak sama rata dengan susunan pola interval sebagai berikut:

1    2    3            4    5    6    7            1

Kedua: sistem urutan nada yang terdiri dari lima nada dalam satu gembyang menggunakan pola jarak sebagai berikut:

1    2    3            5    6            1

Urutan nada dalam laras pelog dapat digeser dengan menggunakan pola interval yang sama. Hal ini hampir sama dengan tangga nada diatonik yang dapat digeser tonikanya. Laras pelog juga terdapat sub-sub laras yaitu pelog pathet lima, dan pelog pathet barang. Mengacu pada penjelasan tentang laras di atas, laras yang akan digunakan pada penelitian ini adalah laras slendro.

b. Ritme

Unsur terpenting dari musik selain nada atau tangga nada adalah ritme. Ritme atau irama adalah gerak nada yang teratur mengalir karena munculnya aksentuasi secara tetap. Keindahan irama akan lebih terasa karena adanya jalinan perbedaan nilai dari satuan bunyi. Pola irama pada musik dapat membedakan perasaan tertentu karena pada hakikatnya irama adalah gerak yang menggerakkan perasaan dan erat hubungannya dengan gerak fisik. Ritme juga berhubungan dengan pola jarak waktu yang diperlukan oleh suatu nada atau bunyi ke nada atau bunyi berikutnya. Dalam sebuah pagelaran Gamelan Jawa, unsur ritme dimiliki oleh kendang. Dinamika dan level yang dihasilkan oleh kendang menentukan ritme dari suatu pagelaran.

c. Ricikan atau instrumen atau alat musik



Ricikan adalah sarana fisik utama (untuk instrumental, disamping vokal) bagi seniman untuk menyampaikan gagasan atau perasaannya. Ricikan atau instrumen yang ada pada gamelan Jawa antara lain kendang, bonang, bonang penerus, demung, saron, kenong, slenthem, gender, gong, gambang, rebab, siter, suling dan kempul.

d. Organisasi musikal/struktur

Bermusik, terutama dalam kebiasaan musik tradisi, hampir selalu melibatkan lebih dari satu orang atau alat musik dan vokal, maka dikenal adanya istilah organisasi musikal, yang menyangkut beberapa hal, diantaranya adalah distribusi kedudukan dari masing-masing ricikan atau vokal. Organisasi musikal dalam pagelaran Gamelan Jawa, terbangun oleh nada-nada yang dihasilkan oleh instrument-instrumen gamelan dengan frekuensi rendah sampai menengah dengan reverb yang cukup besar (sound envelope yang lama). Nada-nada yang dihasilkan oleh instrumen ini rata-rata menghasilkan nada dengan kesan kuat dan megah seperti gong, kenong slentem dan kempul

e. Melodi

Unsur melodi dalam sebuah pagelaran Gamelan membuat warna dari bunyi yang dihasilkan. Unsur ini biasanya dihasilkan oleh instrumen yang memiliki komponen nada dengan frekuensi tinggi dan sound envelope yang rendah (reverb cenderung kecil). Instrumen yang tergolong dalam struktur ini antara lain adalah demung, seruling, gender, dan bonang

(Gamelan Jawa : Awal Mula, Makna, Masa Depan. Bambang Yudoyono, 1984)

### 2.3.2.2. Alat- Alat Musik Gamelan Jawa

Dalam gamelan jawa sebenarnya terdiri kumpulan dari beberapa alat musik yang dimainkan secara bersamaan dengan menghasilkan suara atau irama yang bagus. Komposisi dari musik gamelan ini diciptakan dengan beberapa aturan yang terdiri dari beberapa utaran, pathet dan dibatasi oleh satu gongan serta melodinya. Ada beberapa alat musik yang terdapat dalam gamelan tersebut, diantaranya sebagai berikut:

#### 1. Kendang

Fungsi kendang dalam suatu lagu yang dimainkan memiliki peran yang penting, yaitu sebagai pengatur tempo dan irama disetiap gendhing, baik tempo pokok maupun irama cepat ataupun lambat dengan tangkap, yang mana diatur dengan bunyi kendang termasuk didalamnya untuk mengawali dan mengakhiri gendhing.

Selain itu ada juga fungsi dari kendang ini, diantaranya sebagai berikut:

- Anggeran Wiletan yaitu sebagai penjaga irama agar bisa terdengar bagus.
- Anceran Wiletan yakni pemberi irama yang baik pada awal maupun tengah lagu yang sesuai dengan kebutuhan.
- Amardawalagu adalah sebagai melodi pada setiap lagu.
- Angkruh lagu yaitu sebagai kerangka dalam suatu lagu.
- Adumanis lagu yakni untuk pendukung ritmis pada *wadrita-wadrita* lain dan biasanya *sinden* yang memberi variasi.

Berikut adalah gambar dari alat musik kendang



Gambar.2.3.2.2.1. Alat Musik Kendang

## 2. Gong

Alat ini memiliki fungsi untuk memberi tanda berakhirnya sebuah *gatra* dan juga untuk memberi tanda dimulainya serta berakhirnya *gendhing*. Selain itu berikut fungsi gong di berbagai daerah di Indonesia, seperti:

- Di Jawa alat musik yang satu ini juga pernah dimainkan saat kematian seseorang dari keluarga kerajaan, akan tetapi anggota kerajaan tidak diperbolehkan memainkannya.
- Di daerah Bali biasanya alat musik ini dimainkan pada suatu acara yang berkaitan dengan ritual keagamaan.
- Untuk di daerah Minangkabau, gong digunakan pada pesta pernikahan atau selamatan.
- Di Jawa Barat digunakan pada acara sekuler membawa acara kenningratan dan juga tidak dimiliki oleh semua orang

- Sedangkan di daerah Kalimantan, Sulawesi dan NTT, alat musik gong ini dimainkan pada acara pengobatan, kematian, dan pernikahan namun tidak dimiliki oleh banyak orang. Berikut adalah gambar dari alat musik Gong



Gambar 2.3.2.2.2 Alat Musik Gong

### 3. Rebab

Cara memainkan alat musik ini memang hampir sama dengan biola yaitu dengan cara digesek. Ukuran dari alat ini biasanya lebih kecil dari biola. Pada bagian badan berbentuk bulat dan dibagian yang lainnya memiliki ujung yang panjang. Fungsi dari rebab adalah sebagai salah satu instrumen pemuka, selain itu rebab juga dijuluki sebagai pemimpin lagu dalam ansambel terutama dalam gaya tabuhan yang lirih. Berikut adalah gambar dari alat musik rebab.



Gambar 2.3.2.2.3 Alat Musik Rebab

#### 4. Saron

Menurut ukuran dan fungsinya, alat musik yang satu ini dibedakan menjadi 3 jenis, yaitu sebagai berikut:

- Demung Saron, alat ini memiliki ukuran yang besar dan *berokaf* tengah. Pada umumnya dalam satu perangkat gamelan biasanya terdapat satu atau dua demung. Akan tetapi ada juga gamelan di keraton yang mempunyai lebih dari dua demung.
- Saron Barung, memiliki ukuran yang sedang dan *berokaf* tinggi. Sama halnya dengan demung saron, pada satu paket gamelan biasanya terdapat satu atau dua saron barung saja. Suatu paket gamelan juga mempunyai saron wayangan yang berbilang sembilan. Sebagaimana namanya yang menunjukkan bahwa saron ini dimainkan untuk mengiringi pertunjukan wayang saja.
- Saron Penerus (Peking), jenis saron ini memiliki ukuran yang paling kecil akan tetapi *berokaf* paling tinggi. Saron penerus atau peking ini memainkan tabuhan

rangkap dua atau rangkap empat lagu *balungan*. Lagu pada peking juga berusaha menguraikan lagu *balungan* dalam konteks lagu *gendhing*.

Berikut adalah gambar dari alat musik saron.



Gambar 2.3.2.2.4 Alat Musik Saron

#### 5. Bonang

Alat musik bonang ini dibedakan menjadi 2 berdasarkan fungsinya, yaitu:

- Bonang Barung, alat musik ini berfungsi sebagai *pemurba* lagu, yang bertugas untuk memulai jalannya sajian *gendhing*. Bonang Barung biasanya menjadi pembuka pada setiap permainan gamelan.
- Bonang Penerus, fungsi dari alat musik ini adalah sebagai pengisi harmoni bunyi bonang barung. Bonang penerus ini memiliki suara satu *okaf* yang lebih tinggi dibandingkan dengan bonang barung. Dalam memainkan alat musik bonang penerus ini lebih cepat jika dibandingkan dengan bonang berung.

Berikut adalah gambar dari alat musik bonang.



Gambar 2.3.2.2.5 Alat Musik Bonang

#### 6. Angklung

Angklung merupakan alat musik tradisional Jawa Barat yang terbuat dari potongan bambu. Alat musik ini terdiri dari 2 sampai 4 tabung bambu yang dirangkai menjadi satu dengan tali rotan. Tabung bambu diukir detail dan dipotong sedemikian rupa untuk menghasilkan nada tertentu ketika bingkai bambu digoyang. Pada pagelaran musik gamelan sekarang banyak dimasukan alat musik angklung guna memberi nuansa yang bervariasi. Berikut adalah gambar dari alat musik angklung.



Gambar 2.3.2.2.6 Alat Musik Angklung

## 7. Demung

Demung adalah salah satu instrumen gamelan yang termasuk keluarga balungan. Dalam satu set gamelan biasanya terdapat 2 demung, keduanya memiliki versi pelog dan slendro. Demung menghasilkan nada dengan oktaf terendah dalam keluarga balungan, dengan ukuran fisik yang lebih besar. Demung memiliki wilahan yang relatif lebih tipis namun lebih lebar daripada wilahan saron, sehingga nada yang dihasilkannya lebih rendah. Tabuh demung biasanya terbuat dari kayu, dengan bentuk seperti palu, lebih besar dan lebih berat daripada tabuh saron.

Cara menabuhnya ada yang biasa sesuai nada, nada yang imbal, atau menabuh bergantian antara demung 1 dan demung 2, menghasilkan jalinan nada yang bervariasi namun mengikuti pola tertentu. Cepat lambatnya dan keras lemahnya penabuhan tergantung pada komando dari kendang dan jenis gendhingnya. Pada gendhing Gangsaran yang menggambarkan kondisi peperangan misalnya, demung ditabuh dengan keras dan cepat. Pada gendhing Gati yang bernuansa militer, demung ditabuh lambat namun keras. Ketika mengiringi lagu ditabuh pelan. Ketika sedang dalam kondisi imbal, maka ditabuh cepat dan keras. Berikut adalah gambar dari alat musik demung.





Gambar 2.3.2.2.7 Alat Musik Demung

## 8. Suling

Suling adalah alat musik dari keluarga alat musik tiup kayu atau terbuat dari bambu. Suara suling berciri lembut dan dapat dipadukan dengan alat musik lainnya dengan baik. Suling modern untuk para ahli umumnya terbuat dari perak, emas atau campuran keduanya. Sedangkan suling untuk pelajar umumnya terbuat dari nikel-perak, atau logam yang dilapisi perak. Suling konser standar ditalakan di C dan mempunyai jangkauan nada 3 oktaf dimulai dari middle C. Akan tetapi, pada beberapa suling untuk para ahli ada kunci tambahan untuk mencapai nada B di bawah middle C. Ini berarti suling merupakan salah satu alat musik orkes yang tinggi, hanya piccolo yang lebih tinggi lagi dari suling. Piccolo adalah suling kecil yang ditalakan satu oktaf lebih tinggi dari suling konser standar. Piccolo juga umumnya digunakan dalam orkes. Berikut adalah gambar dari alat musik suling.



Gambar 2.3.2.2.8 Alat Musik Demung

### 2.3.3. Deep Learning

Deep Learning adalah salah satu bagian dari machine learning yang menggunakan *Deep Neural Network* untuk menyelesaikan permasalahan pada domain machine learning. Deep neural network merupakan network yang mempunyai jumlah hidden layer lebih dari satu. Salah satu tantangan dalam mengimplementasikan deep learning adalah untuk mengatur nilai untuk berbagai hiper-parameter, salah satunya adalah topologi jaringannya, yang sangat berkaitan erat dengan jumlah hidden layer dan jumlah neuron.

*Deep Learning* merupakan salah satu bidang dari *Machine Learning* yang memanfaatkan jaringan syaraf tiruan untuk implementasi permasalahan dengan dataset yang besar. Teknik *Deep Learning* memberikan arsitektur yang sangat kuat untuk *Supervised Learning*. Dengan menambahkan lebih banyak lapisan maka model pembelajaran tersebut bisa mewakili data citra berlabel dengan lebih baik.

Pada *Machine Learning* terdapat teknik untuk menggunakan ekstraksi fitur dari data pelatihan dan algoritma pembelajaran khusus untuk mengklasifikasi citra maupun untuk mengenali suara. Namun, metode ini masih memiliki beberapa

kekurangan baik dalam hal kecepatan dan akurasi. Aplikasi konsep jaringan syaraf tiruan yang dalam (banyak lapisan) dapat ditanggihkan pada algoritma *Machine Learning* yang sudah ada sehingga komputer sekarang bisa belajar dengan kecepatan, akurasi, dan skala yang besar. Prinsip ini terus berkembang hingga *Deep Learning* semakin sering digunakan pada komunitas riset dan industri untuk membantu memecahkan banyak masalah data besar seperti *Computer vision*, *Speech recognition*, dan *Natural Language Processing*.

*Feature Engineering* adalah salah satu fitur utama dari *Deep Learning* untuk mengekstrak pola yang berguna dari data yang akan memudahkan model untuk membedakan kelas. *Feature Engineering* juga merupakan teknik yang paling penting untuk mencapai hasil yang baik pada tugas prediksi. Namun, sulit untuk dipelajari dan dikuasai karena kumpulan data dan jenis data yang berbeda memerlukan pendekatan teknik yang berbeda juga. Algoritma yang digunakan pada *Feature Engineering* dapat menemukan pola umum yang penting untuk membedakan antara kelas.

Menurut Goodfellow, dkk. (2016) *deep learning* adalah sebuah pendekatan dalam penyelesaian masalah pada sistem pembelajaran komputer yang menggunakan konsep hierarki. Konsep hierarki membuat komputer mampu mempelajari konsep yang kompleks dengan menggabungkan dari konsep-konsep yang lebih sederhana. Jika digambarkan sebuah graf bagaimana konsep tersebut dibangun di atas konsep yang lain, graf ini akan dalam dengan banyak layer, hal tersebut menjadi alasan disebut sebagai *deep learning* (pembelajaran mendalam)

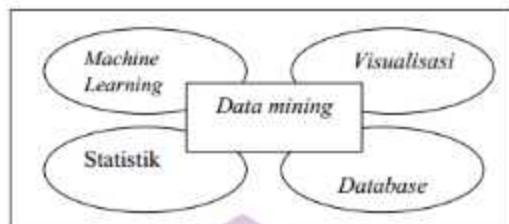
Dalam *Deep Learning*, metode CNN atau *Convolutional Neural Network* sangatlah bagus dalam menemukan fitur yang baik pada citra ke lapisan berikutnya untuk membentuk *hipotesis nonlinier* yang dapat meningkatkan kekompleksitasan sebuah model. Model yang kompleks tentunya akan membutuhkan waktu pelatihan yang lama sehingga di dunia *Deep Learning* penggunaan GPU sudah sangatlah umum.

#### 2.3.4. Machine Learning

Istilah *machine learning* pertama kali didefinisikan oleh Arthur Samuel ditahun 1959. *Machine learning* merupakan salah satu bidang ilmu komputer yang dapat memberikan pembelajaran kepada perangkat computer untuk mengetahui sesuatu tanpa pemrogram yang jelas berdasarkan apa yang disampaikan oleh Arthur Samuel. Menurut Mohri et al. (2012) *machine learning* dapat didefinisikan sebagai metode komputasi berdasarkan pengalaman untuk meningkatkan performa atau membuat prediksi yang akurat. Definisi pengalaman disini ialah informasi sebelumnya yang telah tersedia dan bisa dijadikan data pembelajar. *Machine learning* juga merupakan cabang dari ilmu kecerdasan buatan yang mempelajari pembangunan sistem berdasarkan data.

*Machine learning* adalah suatu area dalam *artificial intelligence* atau kecerdasan buatan yang berhubungan dengan pengembangan teknik-teknik yang bisa diprogramkan dan belajar dari data masa lalu. Pengenalan pola, *data mining* dan *machine learning* sering dipakai untuk menyebut sesuatu yang sama. Bidang ini bersinggungan dengan ilmu probabilitas dan statistik kadang juga

optimasi. *Machine learning* menjadi alat analisis dalam *data mining*, bagaimana bidang-bidang ini berhubungan bisa dilihat dalam **Gambar 3.1**. (Santoso, 2007).



(Sumber : Santoso, 2007)

Gambar 3.1 Ilustrasi cakupan data mining

Penerapan metode *machine learning* ke dalam *database* dalam jumlah besar disebut *data mining*. Hal ini dapat di analogikan seperti ibarat tanah yang luas di permukaan bumi yang mengandung bahan mentah dari alam dapat dilakukan penambangan, sehingga mampu menghasilkan sedikit bahan yang sangat berharga. Demikian pula, dalam *data mining*, sejumlah data besar diolah untuk membangun model sederhana untuk mendapatkan informasi yang berharga (Alpaydin, 2010).

Dalam pembelajaran *machine learning*, terdapat beberapa skenario-skenario. Seperti:

1. *Supervised Learning*

Penggunaan skenario *supervised learning*, pembelajaran menggunakan masukan data pembelajaran yang telah diberi label. Setelah itu membuat prediksi dari data yang telah diberi label.

## 2. *Unsupervised Learning*

Penggunaan skenario *unsupervised learning*, pembelajaran menggunakan masukan data pembelajaran yang tidak diberi label. Setelah itu mencoba untuk mengelompokan data berdasarkan karakteristik-karakteristik yang ditemui.

## 3. *Reinforcement Learning*

Pada skenario *reinforcement learning* fase pembelajaran dan tes saling dicampur. Untuk mengumpulkan informasi pembelajar secara aktif dengan berinteraksi ke lingkungan sehingga untuk mendapatkan balasan untuk setiap aksi dari pembelajar.

Saat ini telah banyak pendekatan *machine learning* yang digunakan untuk deteksi *spam*, *Optical character recognition* (OCR), pengenalan wajah, deteksi penipuan *online*, NER (*Named Entity Recognition*), *Part-of-Speech Tagger*.

Contoh penerapan machine learning dalam kehidupan adalah sebagai berikut:

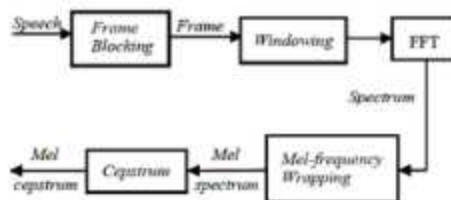
1. Dalam bidang kedokteran contohnya adalah mendeteksi penyakit seseorang yang berasal dari gejala sebelumnya. Misal mendeteksi penyakit jantung dari rekaman *elektrokardiogram*.
2. Dalam bidang *computer vision* contohnya adalah proses pengenalan wajah yang terdapat pada *social media Facebook*. Permisalan lainnya adalah proses penerjemahan tulisan.

3. Pada bidang *information retrieval* contohnya adalah penterjemahan bahasa dengan menggunakan komputer, mengubah suara menjadi teks, dan *filter email spam*. (Codepolitan.com,2016)

### 2.3.5. MFCC (Mel-Frequency Cepstral Coefficients)

Spektrum suara menampilkan frekuensi berbeda yang ada dalam suara (Wolfe, 2014). *Mel Frequency Cepstrum*(MFC) adalah representasi transformasi linear kosinus dari spektrum daya log waktu singkat dari sinyal ucapan pada skala frekuensi Mel non-linear. MFCC (*Mel-Frequency Cepstral Coefficient*) telah menjadi teknik yang sangat awam dan efisien untuk *signal processing* (Lalitha, 2015). MFCC mewakili model telinga, dan dapat menghasilkan pengenalan suara terutama ketika menggunakan jumlah koefisien yang tinggi (Wanli,2013). MFCC merupakan cara yang paling sering digunakan pada berbagai bidang area pemrosesan suara, karena dianggap cukup baik dalam merepresentasikan sinyal. Cara kerja MFCC didasarkan pada perbedaan frekuensi yang dapat ditangkap oleh telinga manusia sehingga mampu merepresentasikan sinyal suara sebagaimana manusia merepresentasikannya.

Proses MFCC akan tergambar di bawah ini:



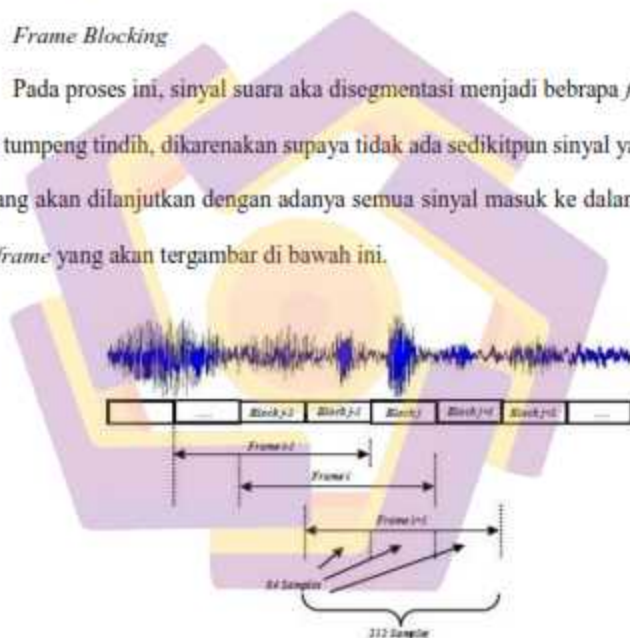
Gambar 2.2.5 Alur Metode MFCC

- *Preemphasis*

Dalam proses ini pengolahan sinyal suara filter akan digunakan setelah proses sampling. Dengan maksud adalah agar mendapatkan bentuk pectral frekuensi sinyal wicara yang lebih halus. Dimana bentuk spectral yang relatif bernilai tinggi untuk daerah rendah dan cenderung turun secara tajam untuk daerah fekuensi diatas 2000 Hz.

- *Frame Blocking*

Pada proses ini, sinyal suara aka disegmentasi menjadi bebrapa *frame* yang saling tumpang tindih, dikarenakan supaya tidak ada sedikitpun sinyal yang hilang, dan yang akan dilanjutkan dengan adanya semua sinyal masuk ke dalam satu atau lebih *frame* yang akan tergambar di bawah ini.



Gambar 2.2.5.1. Alur *frame* bloking pada sinyal audio

- *Windowing*

Kemudian sesudah itu, sinyal analog yang sudah menjadi sinyal digital dibaca *frame* ke *frame*, dan pada setiap *frame* nya dilakukan *windowing* dengan



fungsi *window* tertentu. Ini dimaksudkan meminimalisir ketidakberlanjutan sinyal pada awal dan akhir setiap frame.

Sama halnya dengan proses fitur MFCC pada audio gamelan yang berformat *.wav* ini. Ketika audio sudah dimasukkan, selanjutnya akan ke tahap preprosesing di mana sebelum diekstrasi dia akan digolongkan sesuai blok- blok pada *frame* nya. Setelah *preprocessing*, barulah metode MFCC bekerja untuk menentukan apakah blok yang sudah dikategorikan berdasarkan *frame by frame* itu sesuai atau tidak. Berikut di bawah adalah alur proses fitur MFCC pada audio.

- FFT

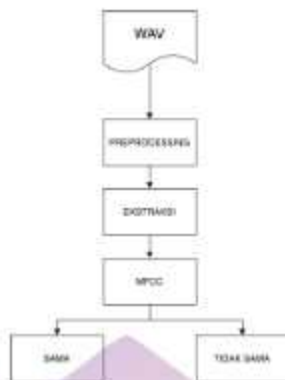
Sesudah itu FFT digunakan untuk mengubah sunyal waveform menjadi satuan frekuensi yang mana akan diubah lagi menjadi satuan energi dari frekuensi tersebut

- Mel Frequency Wrapping

Mengetahui ukuran energi dari frekuensi tertentu dalam sinyal audio. Sehingga akan mendapatkan mel's spektrumnya

- Capstrum

Menghasilkan nilai yang muncul pada mel spektrum dengan menggunakan cepstral filtering



Gambar 2.2.5.2. Proses fitur MFCC (Mel-Frequency Cepstral Coefficients) pada audio

### 2.3.6. Python

Menurut pengertian dari Python Software Foundation (2016), Python adalah bahasa pemrograman interpretatif, berorientasi objek dan semantik yang dinamis. Python memiliki high-level struktur data, dynamic typing dan dynamic binding. Python memiliki sintaks sederhana dan mudah dipelajari untuk penekanan pada kemudahan membaca dan mengurangi biaya perbaikan program. Python mendukung modul dan paket untuk mendorong kemandirian program dan code reuse. Interpreter Python dan standard library-nya tersedia secara gratis untuk semua platform dan dapat secara bebas disebarakan.

### 2.3.7. Convolutional Neural Network (CNN)

Salah satu metode yang penting dalam deep learning adalah *Convolutional Neural Network* (CNN). CNN dirancang khusus untuk pengenalan dan klasifikasi

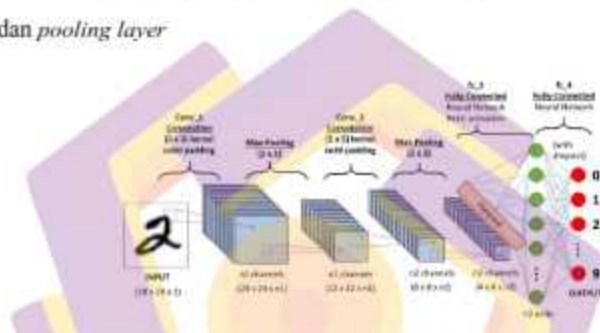
gambar. CNN memiliki beberapa lapisan (layer) yang mengekstrak informasi dari gambar dan menentukan klasifikasi dari gambar tersebut

Terdapat tiga layer yang membentuk CNN yaitu *Convolutional Layer*, *Pooling Layer* dan *Fully Connected Layer*. Secara umum, layer tersebut diatur dengan cara menumpuk beberapa convolutional layer diikuti dengan pooling layer. Pola ini diulangi hingga input gambar semakin mengecil. Layer terakhir adalah fully connected layer yang memberikan hasil output akhir, berupa skor klasifikasi.

*Convolutional Neural Network* (CNN) merupakan pengembangan dari multilayer perceptron (MLP) yang didesain untuk mengolah data dua dimensi dalam bentuk citra. CNN ini termasuk kedalam jenis Deep Neural Network karena kedalaman jaringan yang tinggi dan banyak diaplikasikan pada data citra. Pada dasarnya klasifikasi citra dapat digunakan dengan MLP, akan tetapi dengan metode MLP kurang sesuai untuk digunakan karena tidak menyimpan informasi spasial dari data citra dan menganggap setiap piksel adalah fitur yang independen sehingga menghasilkan hasil yang kurang baik. Penelitian awal yang mendasari penemuan CNN ini pertama kali dilakukan oleh Hubel dan Wiesel (Hubel & Wiesel, T, 1968) mengenai visual cortex pada indera penglihatan kucing. Secara teknis, CNN adalah sebuah arsitektur yang dapat dilatih dan terdiri dari beberapa tahap. Masukan (input) dan keluaran (output) dari setiap tahap adalah terdiri dari beberapa array yang biasa disebut feature map. Setiap tahap terdiri dari tiga layer yaitu konvolusi, fungsi aktivasi layer dan pooling layer.

Metode CNN terbukti berhasil mengungguli metode *Machine Learning* lainnya seperti SVM pada kasus klasifikasi objek pada citra. (Rully Soelaiman dkk).

Secara teknis, *convolutional network* adalah arsitektur yang bisa ditraining dan terdiri dari beberapa tahap. Masukan dan keluaran dari masing-masing tahap adalah beberapa *array* yang disebut *feature map* atau peta fitur. Contoh untuk citra *grayscale*, masukan adalah berupa matriks dua dimensi. Keluaran dari masing-masing tahap adalah *feature map* hasil pengolahan dari semua lokasi pada citra masukan. Masing-masing tahap terdiri dari tiga layer yaitu konvolusi, *activation layer* dan *pooling layer*



Gambar 2.3.7. proses Convolutional Neural Network (CNN)

### 2.3.8. Arsitektur Jaringan CNN

Arsitektur dari CNN dibagi menjadi 2 bagian besar, *Feature Extraction Layer* dan *Fully-Connected Layer (MLP)*.



Gambar 2.3.8. MLP

### a. Feature Extraction Layer

Proses yang terjadi pada bagian ini adalah melakukan "encoding" dari sebuah *image* menjadi *features* yang berupa angka-angka yang merepresentasikan *image* tersebut (*Feature Extraction*). *Feature extraction layer* terdiri dari dua bagian. *Convolutional Layer* dan *Pooling Layer*. Namun kadang ada beberapa riset/paper yang tidak menggunakan *pooling*.

### b. Convolutional Layer (Conv. Layer)

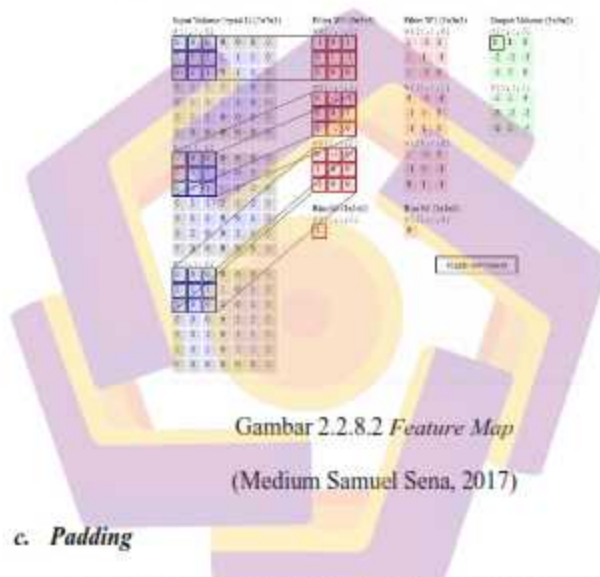


Gambar 2.2.8.1 Convolutional Layer

Gambar diatas adalah RGB (*Red, Green, Blue*) *image* berukuran  $32 \times 32$  *pixels* yang sebenarnya adalah multidimensional *array* dengan ukuran  $32 \times 32 \times 3$  (3 adalah jumlah *channel*). *Convolutional layer* terdiri dari *neuron* yang tersusun sedemikian rupa sehingga membentuk sebuah filter dengan panjang dan tinggi (*pixels*). Sebagai contoh, *layer* pertama pada *feature extraction layer* biasanya adalah *conv. layer* dengan ukuran  $5 \times 5 \times 3$ . Panjang

5 pixels, tinggi 5 pixels dan tebal/jumlah 3 buah sesuai dengan *channel* dari *image* tersebut.

Ketiga filter ini akan digeser keseluruh bagian dari gambar. Setiap pergeseran akan dilakukan operasi “dot” antara *input* dan nilai dari filter tersebut sehingga menghasilkan sebuah *output* atau biasa disebut sebagai *activation map* atau *feature map*.



Gambar 2.2.8.2 *Feature Map*  
(Medium Samuel Sena, 2017)

### c. *Padding*

Sedangkan *Padding* atau *Zero Padding* adalah parameter yang menentukan jumlah *pixels* (berisi nilai 0) yang akan ditambahkan di setiap sisi dari *input*. Hal ini digunakan dengan tujuan untuk memanipulasi dimensi *output* dari *conv. layer* (*Feature Map*).

Tujuan dari penggunaan *padding* adalah :

Dimensi *output* dari *conv. layer* selalu lebih kecil dari *inputnya* (kecuali penggunaan 1x1 filter dengan *stride* 1). *Output* ini akan digunakan

kembali sebagai *input* dari *conv. layer* selanjutnya, sehingga makin banyak informasi yang terbuang.

Dengan menggunakan *padding*, kita dapat mengatur dimensi *output* agar tetap sama seperti dimensi *input* atau setidaknya tidak berkurang secara drastis. Sehingga kita bisa menggunakan *conv. layer* yang lebih dalam/deep sehingga lebih banyak *features* yang berhasil di-*extract*.

Meningkatkan performa dari model karena *conv. filter* akan fokus pada informasi yang sebenarnya yaitu yang berada diantara *zero padding* tersebut. Pada ilustrasi diatas, dimensi dari input sebenarnya adalah 5x5, jika dilakukan *convolution* dengan *filter* 3x3 dan *stride* sebesar 2, maka akan didapatkan *feature map* dengan ukuran 2x2. Namun jika kita tambahkan *zero padding* sebanyak 1, maka *feature map* yang dihasilkan berukuran 3x3 (lebih banyak informasi yang dihasilkan). Untuk menghitung dimensi dari *feature map* kita bisa gunakan rumus seperti dibawah ini:

$$\text{output} = \frac{W+N+2P}{S} + 1 \quad (2.1)$$

W = Panjang/Tinggi Input

N = Panjang/Tinggi Filter

P = Zero Padding

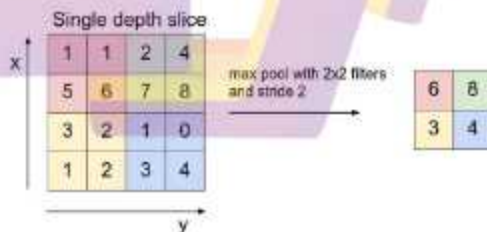
S = Stride

(Le Cun,1998)

#### d. Pooling Layer

*Pooling* atau *subsampling* adalah pengurangan ukuran matriks dengan menggunakan operasi *pooling*. Terdapat dua macam *pooling* yang sering dipakai yaitu *average pooling* dan *max-pooling*. Dalam *average pooling*, nilai yang diambil adalah nilai rata-rata, sementara pada *max pooling*, nilai yang diambil adalah nilai maksimal.

*Pooling layer* biasanya berada setelah *conv. layer*. Pada prinsipnya *pooling layer* terdiri dari sebuah filter dengan ukuran dan stride tertentu yang akan bergeser pada seluruh area feature map. *Pooling* yang biasa digunakan adalah *Max Pooling* dan *Average Pooling*. Sebagai contoh jika kita menggunakan *Max Pooling 2x2* dengan *stride 2*, maka pada setiap pergeseran *filter*, nilai maximum pada area *2x2 pixel* tersebut yang akan dipilih, sedangkan *Average Pooling* akan memilih nilai rata-ratanya.



Gambar 2.2.8.3. *Pooling Layer*

(Medium Samuel Sena, 2017)



Tujuan dari penggunaan *pooling layer* adalah mengurangi dimensi dari *feature map (downsampling)*, sehingga mempercepat komputasi karena parameter yang harus *diupdate* semakin sedikit dan mengatasi *overfitting*(Samuel Sena/medium.com).

### 2.3.9. Musik

Musik pada umumnya dapat dengan mudah ditemukan di kehidupan sehari-hari. Bahkan dapat dikatakan musik akan selalu bersinggungan dengan kehidupan manusia. Musik merupakan hasil olahan dari suara atau bunyi yang diberikan irama, sehingga memiliki nilai keselarasan. Para ahli mendefinisikan musik antara lain sebagai berikut :

1. Menurut Koentjaraningrat musik merupakan bagian dari kesenian. Kesenian merupakan salah satu unsur kebudayaan manusia.
2. Menurut Jamalus : Musik adalah suatu hasil karya seni berupa bunyi dalam bentuk lagu atau komposisi yang mengungkapkan pikiran dan perasaan penciptanya melalui unsur-unsur pokok musik yaitu irama, melodi, harmoni, dan bentuk atau struktur lagu serta ekspresi sebagai suatu kesatuan.
3. Menurut pendapat Soeharto. M dalam buku "Kamus Musik"(1992 : 86) Pengertian musik adalah pengungkapan melalui gagasan melalui bunyi, yang unsur dasarnya berupa melodi, irama, dan harmoni dengan unsur pendukung berupa gagasan, sifat dan warna bunyi. Dari pengertian musik menurut para ahli dapat ditarik kesimpulan bahwa musik merupakan seni yang timbul dari perasaan atau pikiran manusia sebagai pengungkapan ekspresi diri, yang diolah

dari suatu nada-nada atau suara-suara yang diatur dengan irama atau hitungan, dan mengandung unsur harmonis atau keselarasan.

### **2.3.10. Audio**

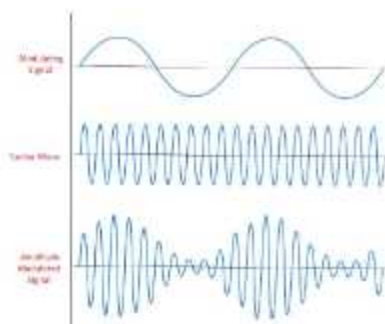
Menurut Heinich, 2002; Ibrahim, 1997; Ibrahim, 2001 (Daryanto, 2010: 4), kata media merupakan bentuk jamak dari kata medium. Medium dapat didefinisikan sebagai perantara atau pengantar terjadinya komunikasi dari pengirim menuju penerima. Menurut Gagne (Arief S. Sadiman, dkk., 2009: 6), media adalah berbagai jenis komponen dalam lingkungan peserta didik yang dapat merangsangnya untuk belajar. Sementara itu, Briggs (Arief S. Sadiman, dkk., 2009: 6), berpendapat bahwa media adalah segala alat fisik yang dapat menyajikan pesan serta merangsang peserta didik untuk belajar. Buku, film, kaset, film bingkai adalah contoh-contohnya. Berdasarkan definisi di atas, dapat disimpulkan bahwa media pembelajaran adalah segala bentuk perantara atau pengantar yang dapat digunakan untuk menyampaikan pesan dari pengirim (pendidik) menuju penerima (peserta didik) dalam kegiatan pembelajaran sehingga dapat merangsang pikiran, perasaan, perhatian, dan minat serta perhatian peserta didik agar proses belajar mengajar dapat terjadi. Berdasarkan Kamus Besar Bahasa Indonesia edisi ketiga (Tim Penyusun, 2007: 76), audio merupakan alat peraga yang bersifat dapat didengar. Daryanto (2010: 37), audio berasal dari kata audible, yang artinya 8 suaranya dapat diperdengarkan secara wajar oleh telinga manusia. Bahan ajar audio merupakan salah satu jenis bahan ajar noncetak yang di dalamnya mengandung suatu sistem yang menggunakan sinyal audio secara langsung, yang dapat dimainkan atau

diperdengarkan oleh pendidik kepada peserta didiknya guna membantu mereka dalam menguasai kompetensi tertentu (Andi Prastowo, 2011: 264). Menurut Arief S. Sadiman, dkk. (2009: 49), media audio adalah media untuk menyampaikan pesan yang akan disampaikan dalam bentuk lambang-lambang auditif, baik verbal (ke dalam kata-kata atau bahasa lisan) maupun non verbal. Dari uraian tersebut, dapat disimpulkan bahwa media audio adalah salah satu bentuk perantara atau pengantar noncetak yang dapat digunakan untuk menyampaikan pesan dari pendidik kepada peserta didik dengan cara dimainkan atau diperdengarkan secara langsung sehingga peserta didik mampu menguasai kompetensi tertentu dari kegiatan pembelajaran yang dilakukan.

### **2.3.11. Wav/ Audio Waveform**

Merupakan bentuk audio digital yang disimpan secara digital dan tidak bisa mengkompresi audio. Namun bentuk lossless menjamin untuk audio yang berkualitas. Bentuk WAV mudah untuk diubah dan dikompresi ke bentuk MP3 atau lainnya. File ini disimpan dalam bentuk extension \*.WAV.

Audio Waveform adalah daya temporal yang diberikan oleh nilai hopSize, yang terdiri dalam menggambarkan setiap frame garis vertikal dari min range ke max range. Tujuannya untuk menampilkan ringkasan dari file audio. Dengan demikian, fitur ini dapat digunakan untuk perbandingan kecepatan gelombang.



Gambar 2.2.11. Gelombang Waveform

### 2.3.12. Musik Information Retrieval (Informasi Pencarian Musik)

“Music Information Retrieval (MIR) is looking into describing the bits of the digital music in ways that facilitate searching through this abundant world without structure.” ( Jehan, 2005 ) Beberapa topik yang paling terpopuler di bidang Music Information Retrieval adalah :

- A. Fingerprinting : bertujuan menggambarkan permukaan audio lagu dengan representasi yang rapi secara simetris dari beberapa lagu yang disebut musical signature. Menurut Cano didalam penelitian Jehan “ The technology enables, for example, cell-phone carriers or copyright management services to automatically identify audio by comparing unique “fingerprints” extracted live from the audio with fingerprints in aspecially compiled music database running on a central server.” (Jehan, 2005)
- B. Query by description : Terdiri dari sebuah MIDI yang besar atau database suara yang disediakan dalam bentuk kualitatif deskripsi teks dari musik atau “irama” yang di mainkan dari sebuah lagu yang masuk didalam microphone. Menurut B. Whitman and R. Rifkin didalam penelitian Jehan “The system typically

compares the entry with a pre-analyzed database metric, and usually ranks the results by similarity.” (Jehan, 2005)

- C. Music similarity adalah percobaan mengkalkulasikan sinyal dari musik. Menurut M. Welsh didalam penelitian Jehan, “There are many criteria with which we may estimate similarities, including editorial (title, artist, country), cultural (genre, subjective qualifiers), symbolic (melody, harmony, structure), perceptual (energy, texture, beat), and even cognitive (experience, reference).” (Jehan, 2005)
- D. Classification Menurut L. Lu, H. Jiang, dan H.-J. Zhang didalam penelitian Jehan mengatakan, “tasks integrate similarity technologies as a way to cluster music into a finite set of classes, including genre, artist, rhythm, instrument, etc.” (Jehan, 2005) Kemiripan dan klasifikasi seringkali tampak paling utama dari cara II-5 mendefenisikan “kebenaran” untuk memberi sebagai fakta tentang perubahan hasil tanpa ada kesalahan.
- E. Thumbnailing bertujuan untuk membangun “representative” kesimpulan suara dari beberapa musik, seperti halnya memindahkan yang paling banyak dan menonjolkan yang paling sedikit dari file tersebut. Menurut Peeters didalam penelitian Jehan, “The task is to detect the boundaries and similarities of large musical structures, such as verses and choruses, and finally assemble them together.” (Jehan, 2005)

## BAB III

### METODE PENELITIAN

#### 3.1. Jenis, Sifat, dan Pendekatan Penelitian

Adapun jenis, sifat dan pendekatan penelitian adalah sebagai berikut:

##### 3.1.1. Jenis Penelitian

Adapun jenis penelitian yang digunakan oleh penulis adalah menggunakan metode eksperimen. Metode eksperimen adalah metode atau cara penelitian kuantitatif yang digunakan untuk mengetahui pengaruh *variable independent* atau yang disebut perlakuan/treatment terhadap *variable dependen* / hasil dalam kondisi yang terkendalikan. Kondisi yang dikendalikan dimaksudkan agar tidak ada *variable* lain (selain *variable treatment*) yang mempengaruhi *variable hasil*. Dibutuhkan sebuah eksperimen di dalam penelitian agar kondisi tersebut dapat dikendalikan, juga diperlukan sebuah kontrol. Dikatakan jenis penelitian eksperimen, karena peneliti akan melakukan pengambilan data audio yang akan dijadikan penelitian ini berdasarkan perekaman yang dilakukan secara pribadi suara alat musik gamelan. Sehingga tercipta dataset yang menjadi bahan penelitian klasifikasi suara alat musik gamelan jawa.

##### 3.1.2. Sifat Penelitian

Pada penelitian ini adalah dengan menggunakan klasifikasi dengan deteksi. Maksud dan tujuan deteksi yang nantinya akan dilakukan pada audio berbasis wav yang akan di persiapkan ekstrasi deteksi dengan menggunakan *MFCC* dan *deep learning*. Dengan menganalisis audio pada musik gamelan nantinya akan terlihat klasifikasi

beragam alat musik gamelan yang telah digunakan. Metode ini sudah tidak diragukan lagi merupakan metode yang sangat tepat untuk digunakan dalam menganalisis.

### 3.1.3. Pendekatan Penelitian

Dalam penelitian ini pendekatan yang dilakukan yaitu kuantitatif. Metode penelitian kuantitatif adalah sebagai metode penelitian yang berlandaskan pada filsafat positivisme digunakan untuk meneliti pada populasi atau sampel tertentu, teknik pengambilan sampel pada umumnya dilakukan secara random, pengumpulan data menggunakan instrument penelitian analisis data bersifat kuantitatif atau statistik (Hayati, 2014). Penelitian ini tentu terdapat proses perhitungan dengan data yang dapat dihitung jumlahnya, sehingga dapat keluar informasi berdasarkan data yang ada. Dataset suara alat musik gamelan yang jumlahnya sangat banyak akan membantu proses ekstraksi yang menghasilkan suatu klasifikasi sesudah dilakukan deteksi. Dengan MFCC dan *deep learning*, pemrosesan pada mesin akan berjalan dan menunjukkan proses data secara terukur.

### 3.2. Metode Pengumpulan Data

Dalam penelitian, metode pengumpulan data yang dilakukan peneliti dibagi kedalam 3 bentuk pokok yakni sebagai berikut :

#### a. Studi Pustaka

Dalam metode ini, peneliti melakukan pengumpulan data dengan cara menelaah berbagai buku, jurnal, makalah, serta laporan hasil penelitian

sebelumnya yang sejenis dan berguna untuk mendapatkan landasan teori mengenai masalah yang diteliti.

b. Sumber Data

Dalam penelitian ini, yang dimaksud dengan sumber data adalah di mana data itu diperoleh. Sumber data bias diperoleh dari data primer dan data sekunder. Sumber data primer pada penelitian ini adalah dengan peneliti membuat sendiri data set pada musik gamelan yang penulis mainkan dan rekam sendiri.

c. Dokumenter

Yang dilakukan penulis adalah dengan metode documenter dikarenakan sumber data yang telah dibuat, akan diuji cobakan pada proses eksperimen. Dalam eksperimen dilakukan pencatatan dan pengamatan yang nantinya akan menghasilkan sebuah dokumen yang selanjutnya akan menjadi bahan Analisa sebuah klasifikasi yang diterapkan dalam musik gamelan jawa.

### 3.3. Metode Analisis Data

Setelah data-data dari peneliti itu sudah cukup terkumpul, maka yang harus dilakukan selanjutnya adalah dengan menganalisa data. Analisa data yang penulis lakukan adalah bersifat kuantitatif. Teknik analisis data yang dilakukan adalah dengan menggunakan statistik akurasi. Akurasi ini akan ditunjukkan pada eksperimen yang dilakukan peneliti.



### 3.4. Alur Penelitian

Pada sub-sub bab ini akan dijelaskan alur penelitian yang akan digunakan dalam penelitian ini antara lainya:

1. Studi Pustaka

Mengetahui penelitian-penelitian terdahulu yang mempunyai kaitan dengan penelitian yang dilakukan, yaitu dengan cara membaca buku, jurnal, makalah, dan laporan jurnal terkait.

2. Identifikasi Masalah

Merumuskan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, dan manfaat penelitian.

3. Pengumpulan Data

Pengumpulan data yang dibuat oleh penulis sebagai dataset pengujian pada audio wav.

4. Melakukan Eksperimen

Setelah data set telah siap digunakan, maka eksperimen dilakukan untuk menguji apakah data yang telah dibuat mampu bekerja sebagai klasifikasi masalah.

5. Melakukan Dokumentasi

Dokumentasi dilakukan untuk mengamati eksperimen yang dilakukan yang nantinya akan menjadi bahan bukti eksperimen guna dilakukan sebuah analisis pada klasifikasi ini.

6. Melakukan Identifikasi dengan Python

Identifikasi apakah klasifikasi ini dilakukan dengan Bahasa Python yang mana akan mengolah dataset yang digunakan untuk klasifikasi alat musik gamelan jawa.

#### 7. Mengolah Hasil Akurasi Klasifikasi

Di dalam tahap ini akan muncul hasil dari klasifikasi berupa sebuah akurasi yang menunjukkan dataset itu bekerja dalam melakukan sebuah identifikasi alat musik gamelan jawa yang diteliti.

#### 8. Mengolah Analisis

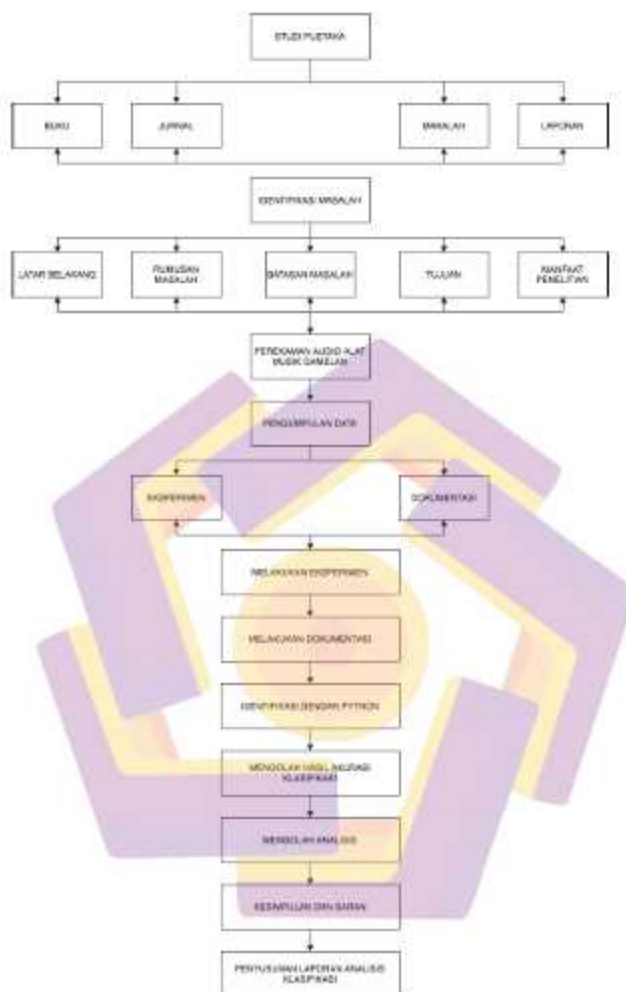
Tahap selanjutnya adalah mengolah analisis, dimana dari hasil akurasi pada klasifikasi alat musik gamelan ini akan bekerja dengan baik. Sehingga ketika akan digunakan untuk pengembangan selanjutnya oleh peneliti dengan hasil yang baik dan berkualitas.

#### 9. Kesimpulan dan Saran

Pada tahap ini adalah bagaimana penulis dapat menemukan konklusi dalam penelitiannya. Kelemahan dalam penelitian yang dilakukan juga perlu dicetuskan guna menjadi perbaikan pada penelitian-penelitian selanjutnya.

#### 10. Penyusunan Laporan Analisis Klasifikasi

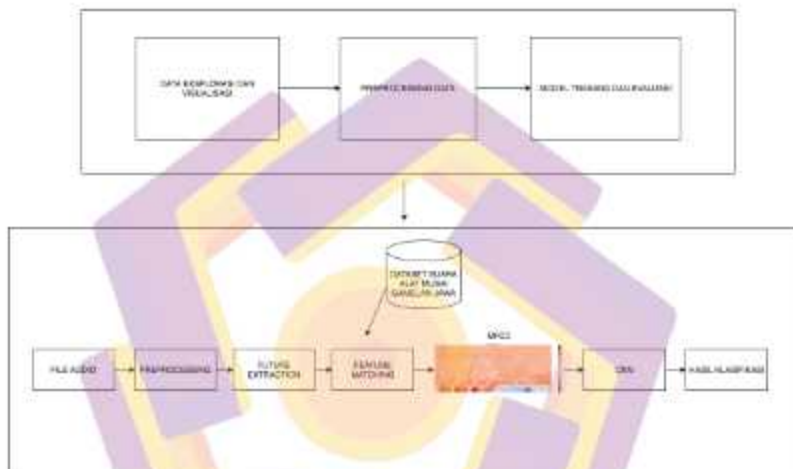
Pada tahap ini adalah penyusunan laporan dari penelitian yang telah dilakukan. Sehingga menjadikan sebuah dokumen dari analisis klasifikasi yang nantinya dapat menjadi manfaat untuk peneliti lainnya.



Gambar 3.4. Alur Penelitian

### 3.5. Alur Klasifikasi Alat Musik Gamelan Jawa

Pada penelitian ini, klasifikasi alat musik gamelan menggunakan ekstraksi MFCC (*Mel Frequency Cepstral Coefficients*) dengan metode CNN (*Convolutional Neural Network*). Pada klasifikasi ini, akan digunakan audio diuji sesuai hasilnya. Alurnya akan ditunjukkan pada gambar di bawah ini:



Gambar 3.5 Alur Klasifikasi

#### 3.5.1. Data Eksplorasi dan Visualisasi

Langkah klasifikasi pertama yang dilakukan adalah data eksplorasi dan visualisasi. Untuk itu, diperlukan sebuah dataset yang berisi 374 kutipan suara kurang lebih sama dengan 34 detik suara alat musik gamelan berupa angklung, bonang, demung, gong, kendang, rebab, saron, dan suling. Ringkasan data file sampel audio digital dalam format .wav. Gelombang suara didigitalkan dengan pengambilan sampel pada interval diskrit yang dikenal sebagai laju pengambilan sampel. Kami mengambil 11kHz untuk kualitas audio

berarti sample diambil 11.000 kali per detik. Setiap sampel adalah amplitude gelombang pada interval waktu tertentu, di mana kedalaman bit menentukan seberapa detail sampel akan juga dikenal sebagai rentang dinamis sinyal, kami mengambil 16 bit yang berarti sampel dapat berkisar dari nilai amplitude 65.536. Oleh karena itu, data yang akan kita analisis untuk setiap kutipan suara pada dasarnya adalah *array* satu dimensi atau vector nilai amplitude. Kemudian adalah dengan menganalisis audio data. Di sini akan digunakan libraris dari *iPython.display.Audio* yang mana ini akan membuat audio dapat kita putar dengan menggunakan *Jupyter Notebook* (suatu *workspace* untuk mengolah data). Lalu kita menggunakan *Librosa* yang merupakan paket dari *Python* untuk musik dan pemrosesan audio oleh Brian McFee dan akan memungkinkan kita untuk membuat audio di *notebook* kita sebagai *array numpy* untuk analisis. Kemudian kita lakukan obserasi di mana dari inspeksi visual kita dapat melihat bahwa sulit untuk memvisualisasikan perbedaan antara beberapa kelas. Bentuk gelombang dari alat musik gamelan memiliki bentuk yang serupa. Telinga manusia secara alami dapat mendeteksi perbedaan antara harmonis akan menarik untuk melihat seberapa baik model pembelajaran yang dalam akan dapat mengekstrak fitur yang diperlukan untuk membedakan antara kelas-kelas ini. Lalu dataset akan kita *load* ke *Panda dataframe* yang akan dikelompokkan ke kelas distribusi, *channel* audionya, *sample rate* nya, *bitdepth* nya. Pengamatan di sini kita bisa melihat label kelas tidak seimbang. Bisa jadi suara kenong sama dengan demung. Ini akan menjadi perhatian dan sesuatu yang mungkin perlu kita tangani. Properti *file* sampel audio adalah hal

yang selanjutnya kita lakukan yaitu iterasi melalui setiap file sampel audio dan ekstrak, jumlah saluran audio, kecepatan sampel dan kedalaman bit. Saluran sebagian besar sampel memiliki dua saluran audio (artinya stereo) dengan beberapa hanya dengan satu saluran (mono). Pilihan termudah di sini untuk membuatnya seragam adalah dengan menggabungkan dua saluran dalam sampel stereoid menjadi satu dengan rata-rata nilai dari dua saluran. Ada berbagai tingkat Sampel yang telah digunakan di semua sampel yang menjadi perhatian. Ini berarti bahwa kita harus menerapkan teknik konversi laju sampel (baik konversi naik atau konversi turun) sehingga kita dapat melihat representasi agnostik dari bentuk gelombangnya yang memungkinkan kita melakukan perbandingan yang adil. Ada juga berbagai kedalaman bit. Kemungkinan kita perlu menormalkannya dengan mengambil nilai amplitudo maksimum dan minimum untuk kedalaman bit tertentu. Properti audio lain yang perlu dipertimbangkan. Kita mungkin juga perlu mempertimbangkan untuk menormalkan level volume (nilai amplitude gelombang) jika ini terlihat sangat bervariasi, dengan melihat volume puncak atau volume RMS.

### 3.5.2. Preprocessing Data

Langkah kedua dari klasifikasi alat musik gamelan Jawa ini adalah *data preprocessing* dan *data splitting*. Properti audio yang perlu dinormalisasi. Mengikuti dari buku catatan sebelumnya, kami mengidentifikasi properti audio berikut yang memerlukan pemrosesan awal untuk memastikan konsistensi di seluruh kumpulan data nya adalah saluran audio, tingkat sampel, kedalaman bit.

Penulis akan menggunakan Librosa yang akan berguna untuk pra-pemrosesan dan ekstraksi fitur. Lalu tahap preprocessing. Untuk sebagian besar preprocessing kita akan dapat menggunakan *fungsi load () Librosa*. Kami akan membandingkan output dari *Librosa* dengan output default dari *library wavfile scipy* menggunakan file yang dipilih dari dataset. Kemudian adalah konversi tingkat sampel. Secara default, fungsi beban Librosa mengubah kecepatan pengambilan sampel menjadi 22,05 KHz yang dapat kita gunakan sebagai tingkat perbandingan kita. Lalu kedalaman bit. Fungsi muat Librosa juga akan menormalkan data sehingga nilainya berkisar antara -1 dan 1. Ini menghilangkan kerumitan kumpulan data yang memiliki kedalaman bit yang sangat beragam. Setelah itu gabungkan saluran audio. Librosa juga akan mengubah sinyal menjadi mono, artinya jumlah saluran akan selalu 1. Properti audio lain yang perlu dipertimbangkan juga. Pada tahap ini, belum jelas apakah faktor lain juga perlu diperhitungkan, seperti durasi sampel dan tingkat volume. Penulis akan melanjutkan apa adanya untuk sementara waktu dan kembali membahasnya nanti jika dianggap memengaruhi validitas metrik target kami. Fitur Ekstrak. Seperti yang dijelaskan dalam proposal, kami akan mengekstrak Koefisien *Cepstral Frekuensi Mel* (MFCC) dari sampel audio. MFCC merangkum distribusi frekuensi di seluruh ukuran jendela, sehingga memungkinkan untuk menganalisis karakteristik frekuensi dan waktu suara. Representasi audio ini akan memungkinkan kami mengidentifikasi fitur untuk klasifikasi. Mengekstrak MFCC. Untuk ini kita akan menggunakan fungsi *mfcc () Librosa* yang menghasilkan MFCC dari data audio deret waktu. Lalu

mengekstrak MFCC untuk setiap file. Kami sekarang akan mengekstrak MFCC untuk setiap file audio dalam dataset dan menyimpannya dalam Panda Dataframe bersama dengan label klasifikasinya. Ubah data dan labelnya.

Penulis akan menggunakan *sklearn.preprocessing.LabelEncoder* untuk menyandikan data teks kategorikal ke dalam data numerik yang dapat dipahami model. Lalu dalam tahap kedua ini pisahkan kumpulan data. Di sini kita akan menggunakan *sklearn.model\_selection.train\_test\_split* untuk membagi dataset menjadi set pelatihan dan pengujian. Ukuran set pengujian akan menjadi 20% dan kami akan menetapkan status acak.

### 3.5.3. Model Training dan Evaluasi

Langkah ketiga adalah model training dan evaluasi. Arsitektur model awal - MLP. Kita akan mulai dengan membangun Jaringan *Neural Multilayer Perceptron* (MLP) menggunakan *Keras* dan backend *Tensorflow*. Dimulai dengan model sekuensial sehingga kita dapat membangun model lapis demi lapis. Kita akan mulai dengan arsitektur model sederhana, yang terdiri dari tiga lapisan, lapisan masukan, lapisan tersembunyi, dan lapisan keluaran. Ketiga lapisan tersebut akan menjadi jenis lapisan padat yang merupakan jenis lapisan standar yang digunakan dalam banyak kasus untuk jaringan saraf. Lapisan pertama akan menerima bentuk masukan. Karena setiap sampel berisi 40 MFCC (atau kolom), kami memiliki bentuk (1x40) ini berarti kami akan mulai dengan bentuk input 40. Dua lapisan pertama akan memiliki 256 node. Fungsi aktivasi yang akan kita gunakan untuk 2 lapisan pertama kita adalah ULT, atau Aktivasi Linear Rektifikasi. Fungsi aktivasi ini telah terbukti berfungsi dengan baik di



jaringan saraf. Kami juga akan menerapkan nilai Dropout 50% pada dua lapisan pertama kami. Ini akan mengecualikan node secara acak dari setiap siklus pembaruan yang pada gilirannya menghasilkan jaringan yang mampu melakukan generalisasi yang lebih baik dan kecil kemungkinannya untuk menyesuaikan data pelatihan. Lapisan keluaran kita akan memiliki 10 node (`num_labels`) yang cocok dengan jumlah klasifikasi yang mungkin. Aktivasi untuk lapisan keluaran kami adalah softmax. Softmax membuat jumlah output menjadi 1 sehingga output dapat diartikan sebagai probabilitas. Model kemudian akan membuat prediksi berdasarkan opsi mana yang memiliki probabilitas tertinggi. Kemudian kita mengompilasi model. Untuk mengkompilasi model kita, kita akan menggunakan tiga parameter yaitu yang pertama fungsi kerugian. Kita akan menggunakan `categorical_crossentropy`. Ini adalah pilihan paling umum untuk klasifikasi. Skor yang lebih rendah menunjukkan bahwa model tersebut berperforma lebih baik. Kami akan menggunakan metrik akurasi yang memungkinkan kami melihat skor akurasi pada data validasi saat kami melatih model. Lalu akan dilakukan pengoptimalan, di sini kita akan menggunakan `adam` yang secara umum merupakan pengoptimal yang baik untuk banyak kasus penggunaan. Lalu kita akan mulai dengan 100 epoch yang merupakan frekuensi model akan bergilir data. Model tersebut akan diperbaiki pada setiap siklus hingga mencapai titik tertentu. Kami juga akan mulai dengan ukuran tumpukan yang rendah, karena memiliki ukuran tumpukan yang besar dapat mengurangi kemampuan generalisasi model. Lalu kita menguji model Di sini kita akan meninjau

keakuratan model pada kumpulan data pelatihan dan pengujian. Lalu kita harus memprediksi. Di sini kita akan membuat metode yang akan memungkinkan kita untuk menguji prediksi model pada file audio .wav yang ditentukan. Kemudian dilakukan sebuah validasi uji dengan data sampel. Pemeriksaan kesucian awal untuk memverifikasi prediksi menggunakan sub-bagian dari sampel file audio yang kami jelajahi di buku catatan pertama. Kami berharap sebagian besar ini diklasifikasikan dengan benar. Lalu kita lakukan pengamatan. Dari pemeriksaan kesehatan singkat ini model tampaknya memprediksi dengan baik. Audio lainnya juga akan kita jadikan sampel uji fitur. Di sini kita akan menggunakan sampel berbagai suara bebas hak cipta yang bukan merupakan bagian dari data pengujian atau pelatihan untuk lebih memvalidasi model kita. Terakhir kita lakukan pengamatan dari performa model awal kami memuaskan dan telah digeneralisasi dengan baik, tampaknya memprediksi dengan baik saat diuji terhadap data audio baru.

## BAB IV

### HASIL PENELITIAN DAN PEMBAHASAN

#### 4.1. Implementasi Pengumpulan Data

Proses pengolahan data dilakukan dengan menggunakan *software* *Jupyter Notebook* dan *Python*. Sebelum melakukan proses pengambilan data, beberapa atribut yang harus diperlukan adalah sebagai berikut:

1. *Software* *Jupyter Notebook*, *Python* dan *web browser google chrome* yang telah terinstal di perangkat komputer;
2. Koneksi *internet*.

Pada penelitian ini dilakukan pembuatan data / informasi dari penulis penelitian yang sudah memiliki sampel suara alat musik gamelan jawa dengan menggunakan *Digital Audio Workstation (DAW)* sebagai media perekaman alat musik gamelan jawa yang kemudian di olah sedemikian rupa hingga menjadi sampe audio yang dijadikan sebuah data penelitian. DAW yang digunakan adalah *LogicProX* yang dirender dengan kualitas sesuai dengan Batasan masalah pada penelitian ini, yaitu 11-22 kHz dan 16 Bit. 16 Bit ditetapkan karena mengikuti library dari penelitian ini hanya mencapai 16 Bit saja. Data yang akan digunakan adalah data suara alat musik angklung, bonang, demung, gong, kendang, rebab, saron, suling. Cara memainkan alat musik gamelan yang akan direkam dan dijaidkan data memiliki waktu yang berbagai macam, dan tentunya memainkannya sesuai dengan kaidah permainan alat musik gamelan, hanya dengan menggunakan DAW *LogicProX* akan dibuat rentang waktu 1-34 detik secara acak, agar data semakin banyak untuk dapat terbaca dengan metode MFCC dengan CNN ini.

Berikut adalah gambar saat pengambilan suara alat musik gamelan pada DAW software LogicProX.



Gambar 4.1. Perekaman Suara alat musik gamelan

Saat diberi sampel audio dalam format yang dapat dibaca komputer (seperti file .wav) dengan durasi beberapa detik, penulis ingin dapat menentukan apakah sampel tersebut berisi salah satu suara alat musik gamelan jawa target dengan skor kemungkinan yang sesuai. Sebaliknya, jika tidak ada suara target yang terdeteksi, kami akan disajikan dengan skor yang tidak diketahui. Berikut adalah data pengambilan suara masing- masing alat musik gamelan dengan rentang frekuensi 11-22 kHz. Total dari data yang akan dijadikan sumber penelitian adalah masing-

masing 150 data pada 8 alat musik gamelan. Jadi total keseluruhan adalah 1200 data.

Tabel 4.1.1. Pengambilan Data Audio alat musik gamelan

No	Class ID	IDName	NameClass	Pengambilan Data			Total
				11 kHz	12 kHz	22 kHz	
1	0	10001	Angklung	60	45	45	1200
2	1	10002	Bonang	79	55	36	
3	2	10003	Demung	100	25	25	
4	3	10004	Gong	100	25	25	
5	4	10005	Kendang	51	49	50	
6	5	10006	Rebab	64	43	43	
7	6	10007	Saron	100	25	25	
8	7	10008	Suling	50	50	50	

Total dari data yang akan dijadikan sumber penelitian adalah masing-masing 150 data pada 8 alat musik gamelan. Jadi total keseluruhan adalah 1200 data. Pada pengambilan data menggunakan DAW LogicproX, terdapat Angklung dengan rentang kualitas frekuensi atau yang disebut *sample rate* sebesar 11 Khz sebanyak 60, 12 Khz adalah 45 dan 22 kHz sebanyak 45. Selanjutnya Bonang dengan 11 kHz nya sebanyak 79, 12 kHz sebanyak 35 dan 22 kHz sebanyak 36. Untuk alat musik Demung adalah sebanyak 100, 25, 25 untuk rentang 11-22 kHz. Gong sebanyak 100, 25,25. Kendang sebanyak 51, 49, 50. Rebab sebanyak 64, 43, 43. Saron sejumlah 100,25,25. Dan suling sebanyak 50, 50, 50. Jumlah pengambilan diambil secara acak, tetapi untuk jumlah masing-masing alat musik sejumlah 150. Ini digunakan sebagai pembeda dalam system melakukan klasifikasi nanti. Tentunya cara memainkan alat musik gamelan ini telah sesuai nada dan pattern pada masing tiap-tiap alat musik

## 4.2. Eksplorasi dan Visualisasi Data

### 4.2.1. Dataset Suara Alat Musik Gamelan Jawa

Dalam penelitian yang kami lakukan, menggunakan dataset yang kami bernama 'gamel'. Kumpulan data berisi 1200 suara alat musik gamelan jawa dari 8 kelas yaitu antara lain:

1. Angklung
2. Bonang
3. Demung
4. Gong
5. Kendang
6. Rebab
7. Saron
8. Suling

Metadada suara alat musik gamelan masing-masing berisi ID unik untuk setiap kutipan suara alat musik gamelan jawa bersama dengan nama kelas yang diberikan. Berikut di bawah ini adalah tabel tampilan metadada alat musik gamelan yang berjumlah 1200 data yang terdiri dari masing-masing alat musik gamelan sebanyak 150.

Tabel 4.2.1. Dataset Suara Alat Musik Gamelan Jawa

slice_file_name	fsID	start	end	sallence	fold	classID	class_name
10001-0-0-0.wav	10001	0	34	0	30	0	Angklung
10001-0-0-1.wav	10001	0	34	0	30	0	Angklung
10001-0-0-2.wav	10001	0	34	0	30	0	Angklung
10001-0-0-3.wav	10001	0	34	0	30	0	Angklung
10001-0-0-4.wav	10001	0	34	0	30	0	Angklung
10001-0-0-5.wav	10001	0	34	0	30	0	Angklung
10001-0-0-6.wav	10001	0	34	0	30	0	Angklung
10001-0-0-7.wav	10001	0	34	0	30	0	Angklung
10001-0-0-8.wav	10001	0	1	0	30	0	Angklung
10001-0-0-9.wav	10001	0	3	0	30	0	Angklung
10001-0-0-10.wav	10001	0	5	0	30	0	Angklung
10001-0-0-11.wav	10001	0	7	0	30	0	Angklung
10001-0-0-12.wav	10001	0	9	0	30	0	Angklung
10001-0-0-13.wav	10001	0	11	0	30	0	Angklung
10001-0-0-14.wav	10001	0	14	0	30	0	Angklung
10001-0-0-15.wav	10001	0	15	0	30	0	Angklung
10001-0-0-16.wav	10001	0	16	0	30	0	Angklung

#### 4.2.2. Data File Sampel Audio

Kutipan suara alat musik gamelan jawa ini adalah file audio digital dalam format *.wav*. Gelombang suara didigitalkan dengan pengambilan sampel pada interval diskrit yang dikenal sebagai laju pengambilan sampel. Data adalah 11-22kHz untuk kualitas audio berarti sample diambil 11.000-22.000 kali per detik. Setiap sampel adalah amplitude gelombang pada interval waktu tertentu, di mana kedalaman bit menentukan seberapa detail sampel akan juga dikenal sebagai rentang dinamis sinyal, kami mengambil 16 bit yang berarti sampel dapat berkisar dari nilai amplitude 65.536. Oleh karena itu, data yang akan kita analisis untuk setiap kutipan suara pada dasarnya adalah *array* satu dimensi atau vector nilai amplitude. Konversi pertama yang diterapkan pada sinyal audio adalah memplot

amplitude sampelnya dari waktu ke waktu, sesuai dengan data yang diambil peneliti.

#### 4.2.3. Menganalisis Data Audio

Kemudian adalah dengan menganalisis audio data. Di sini akan digunakan libris dari *iPython.display.Audio* yang mana ini akan membuat audio yang dapat diputar dengan menggunakan *Jupyter Notebook* (suatu *workspace* untuk mengolah data). Lalu kita menggunakan *Librosa* yang merupakan paket dari *Python* untuk musik dan pemrosesan audio oleh Brian McFee dan akan memungkinkan kita untuk membuat audio di *notebook* kita sebagai array *numpy* untuk analisis. Peneliti harus menginstal *librosa* menggunakan *pip install librosa*.

Dengan menganalisis audio data dengan menggunakan *librosa* dan keras yang merupakan library dari Bahasa pemrograman *python*. Dalam penggunaan tersebut memiliki batasan rentang *sample rate* dan *bit depth* yang dapat dibaca oleh library tersebut adalah 11-22 kHz dan 16 Bit saja. Oleh karena itu, kami mengambil data audio alat musik gamelan ini dengan rentang yang telah ditetapkan.

#### 4.2.4. Pemeriksaan Audiotori

Kita akan menggunakan *IPython.display.Audio* untuk memutar file audio sehingga dapat memeriksa secara aural. Ini akan memeriksa, apakah benar file audio yang menjadi data penelitian benar adanya. Berikut adalah gambar perintah yang dijalankan pada *Jupyter Notebook*.



```

In [14]: import IPython.display as ipd
         ipd.Audio('F:/Macity-ML-Capstone-master/EthnicSound_Dataset/sample/audio/10001-0-0-10.wav')
Out[14]:
▶ 0:01 / 0:20 ————— 🔊 ⓘ

In [15]: import IPython.display as ipd
         ipd.Audio('F:/Macity-ML-Capstone-master/EthnicSound_Dataset/sample/audio/10002-1-0-4.wav')
Out[15]:
▶ 0:00 / 0:34 ————— 🔊 ⓘ

In [17]: import IPython.display as ipd
         ipd.Audio('F:/Macity-ML-Capstone-master/EthnicSound_Dataset/sample/audio/10003-2-0-12.wav')
Out[17]:
▶ 0:00 / 0:19 ————— 🔊 ⓘ

```

Gambar 4.2.4. Perintah dan Hasil Pemeriksaan Auditori

#### 4.2.5. Inspeksi Visual

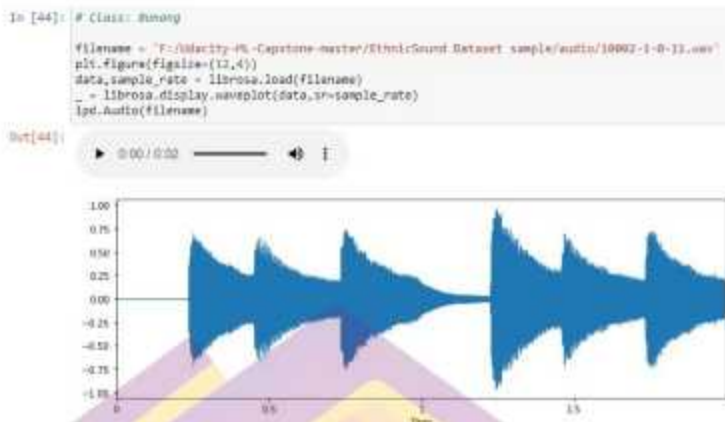
Kami akan memuat sampel dari setiap kelas dan secara visual memeriksa data untuk pola apa pun. Kami akan menggunakan *librosa* untuk memuat file audio ke dalam *array* kemudian *librosa.display* dan *matplotlib* untuk menampilkan bentuk gelombang. Berikut tampilan hasilnya dalam bentuk gambar dibawah ini.



Gambar 4.2.5. Hasil Inspeksi Visual Suara Alat Musik Gamelan  
"Angklung"

#### 4.2.6. Pengamatan

Kemudian dilakukan observasi di mana dari inspeksi visual kita dapat melihat bahwa sulit untuk memvisualisasikan perbedaan antara beberapa kelas. Bentuk gelombang dari alat musik gamelan memiliki bentuk yang serupa. Telinga manusia secara alami dapat mendeteksi perbedaan antara harmonis akan menarik untuk melihat seberapa baik model pembelajaran yang dalam akan dapat mengekstrak fitur yang diperlukan untuk membedakan antara kelas-kelas ini.



Gambar 4.2.6. Perbedaan Hasil Visual Alat Musik Gamelan “Bonang”

#### 4.2.7. Metadata

Dalam penelitian ini akan dilakukan pemuatan file metadata suara alat musik gamelan Jawa dengan nama ‘gamel.csv’ ke dalam *dataframe* Panda. Berikut di bawah ini adalah tampilan hasil memuat file metadatanya.

```
In [51]: import pandas as pd
metadata = pd.read_csv('E:/idacity-PL-Capstone-master/EthnicSound-Dataset-sample/metadata/gamel.csv')
metadata.head()
```

```
Out[51]:
```

	file_name	fid	start	end	silence	fid	classID	class_name
0	10001-0-0-0.wav	10001	0	34	0	30	0	Angklung
1	10001-0-0-0.wav	10001	0	34	0	30	0	Angklung
2	10001-0-0-2.wav	10001	0	34	0	30	0	Angklung
3	10001-0-0-3.wav	10001	0	34	0	30	0	Angklung
4	10001-0-0-4.wav	10001	0	34	0	30	0	Angklung

Gambar 4.2.7. Memuat File Metadata

Metadata yang muncul pada hasil perintah dengan import panda akan sama dengan data yang sebenarnya.

#### 4.2.8. Distribusi Kelas

Dalam penelitian ini akan ditampilkan cetak distribusi kelas di mana, setiap kelas akan muncul berapa banyak data yang tersedia atau dibuat. Berikut di bawah ini adalah hasil dalam bentuk gambar.

```
In [47]: print(metadata.class_name.value_counts())
```

Suling	150
Kendang	150
Bonang	150
Angklung	150
Saron	150
Demung	150
Rebab	150
Gong	150
Name: class_name, dtype: int64	

Gambar 4.2.8. Tampilan Hasil Distribusi Kelas

Hasil dari distribusi data menunjukkan 150 data masing masing alat musik gamelan yang artinya data terdistribusi secara merata dengan jumlah total 1200 dataset.

#### 4.2.9. Properti File Sampel Audio

Pada tahap ini akan diulangi setiap file sampel audio dan mengekstrak nya, jumlah *channel* nya, *sample rate*, dan *bit depth* nya. Berikut dibawah ini adalah hasil mengimpor data nya.

```

In [53]: # load various imports
import pandas as pd
import os
import librosa
import librosa.display

from helpers.wavfilehelper import wavfilehelper
wavfilehelper = wavfilehelper()

audiodata = []
for index, row in metadata.iterrows():
    file_name = os.path.join(os.path.abspath('../libcity-10-Capstone-master/EthnicSound Dataset samples'),
                             row['file_name'])
    data = wavfilehelper.read_wav_properties(file_name)
    print(data)
    audiodata.append(data)

# Convert into a Pandas dataframe
audiodef = pd.DataFrame(audiodata, columns=['num_channels', 'sample_rate', 'bit_depth'])

```

(22, 0, 0)

Gambar 4.2.9. Properti File Sampel Audio

#### 4.2.10. Channel Audio

Sebagian besar sampel memiliki dua saluran audio (artinya stereo) dengan beberapa hanya dengan satu saluran (mono). Opsi termudah di sini untuk membuatnya Uniform akan menggabungkan dua saluran dalam sampel stereo menjadi satu dengan rata-rata nilai dari dua saluran. Berikut dibawah ini adalah hasil memunculkan channel audio pada sampel suara.

```

# num of channels
print(audiodef.num_channels.value_counts(normalize=True))

```

1	0.045576
2	0.077748

Gambar 4.2.10. Memunculkan Channel Audio

#### 4.2.11. Sample Rate

Rentang luas rasio sampel yang telah digunakan di semua sampel suara harus kita perhatikan. Ini kemungkinan berarti bahwa kami harus menerapkan teknik konversi laju sampel (baik konversi naik atau konversi turun) sehingga kita dapat melihat representasi agnostik dari bentuk gelombang mereka yang memungkinkan kami melakukan perbandingan yang adil. Data yang kami gunakan adalah 11-22 kHz. Kami mengambil 11-22kHz untuk kualitas audio berarti sampel diambil 11.000-22.000 kali per detik. Berikut adalah hasil tampilan *sample rate* pada audio alat musik gamelan.

```
In [55]: # sample rates
print(audioidf.sample_rate.value_counts(normalize=True))
0    1.0
Name: sample_rate, dtype: float64
```

Gambar 4.2.11. Hasil Tampil *Sample Rate*

Di sini dapat kita lihat, tertunjuk 1.0, dikarenakan data audio yang kita ambil yaitu 11-22 kHz.

#### 4.2.12. Bit Depth

Ada juga lebar kisaran kedalaman bit. Kemungkinan perlu menormalkannya dengan mengambil nilai amplitudo maksimum dan minimum untuk kedalaman bit tertentu. Setiap sampel adalah amplitudo gelombang pada

interval waktu tertentu, di mana kedalaman bit menentukan seberapa detail sampel akan juga dikenal sebagai rentang dinamis sinyal, kami mengambil 16 bit yang berarti sampel dapat berkisar dari nilai amplitude 65.536. Oleh karena itu, data yang akan dianalisis untuk setiap kutipan suara pada dasarnya adalah *array* satu dimensi atau vector nilai amplitude. Berikut hasil tampil *bit depth* dengan hasil gambar di bawah ini.

```
In [56]: # bit_depth
print(audiodef.bit_depth.value_counts(normalize=True))
0    1.0
Name: bit_depth, dtype: float64
```

Gambar 4.2.12. Hasil Tampil Bit Depth

Di sini dapat kita lihat, tertunjuk 1.0, dikarenakan data audio yang kita ambil sama semua yaitu 16bit.

#### 4.3. Algoritma dan Teknik

Langkah pertama yang dilakukan adalah data eksplorasi dan visualisasi. Solusi yang diusulkan untuk penelitian ini adalah menerapkan *teknik Deep Learning* yang telah terbukti sangat berhasil di bidang klasifikasi citra. Pertama kita akan mengekstrak Koefisien *Cepstral Frekuensi Mel* (MFCC) dari sampel audio pada basis per bingkai dengan ukuran jendela beberapa milidetik. MFCC merangkum distribusi frekuensi di seluruh ukuran jendela, sehingga

memungkinkan untuk menganalisis karakteristik frekuensi dan waktu suara. Representasi audio ini akan memungkinkan untuk mengidentifikasi fitur untuk klasifikasi.

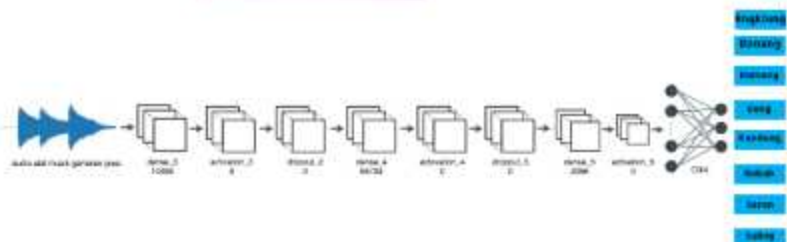
Langkah kedua dari klasifikasi alat musik gamelan Jawa ini adalah *data preprocessing* dan *data splitting*. Properti audio yang perlu dinormalisasi. Mengikuti dari buku catatan sebelumnya, dapat dilakukan identifikasi properti audio berikut yang memerlukan pemrosesan awal untuk memastikan konsistensi di seluruh kumpulan data nya adalah saluran audio, tingkat sampel, kedalaman bit. Penulis akan menggunakan Librosa yang akan berguna untuk pra-pemrosesan dan ekstraksi fitur. Multi-layer perceptron (MLP) digolongkan sebagai jenis Deep Neural Network karena mereka terdiri dari lebih dari satu lapisan perceptron dan menggunakan aktivasi non-linear yang membedakannya dari perceptron linier. Arsitekturnya terdiri dari lapisan masukan, lapisan keluaran yang pada akhirnya membuat prediksi tentang masukan, dan di antara dua lapisan tersebut terdapat sejumlah lapisan tersembunyi yang berubah-ubah.

Langkah ketiga adalah model training dan evaluasi. Arsitektur model awal - MLP. Kita akan mulai dengan membangun Jaringan *Neural Multilayer Perceptron* (MLP) menggunakan *Keras* dan backend *Tensorflow*. Algoritme untuk melatih MLP dikenal sebagai Backpropagation. Dimulai dengan semua bobot dalam jaringan yang ditetapkan secara acak, masukan melakukan penerusan melalui jaringan dan keputusan lapisan keluaran diukur berdasarkan kebenaran dasar dari label yang ingin Anda prediksi. Kemudian bobot dan bias dipropagasi



kembali melalui jaringan tempat metode pengoptimalan, biasanya Stochastic Gradient descent digunakan untuk menyesuaikan bobot sehingga mereka akan bergerak selangkah lebih dekat ke kesalahan minimum pada lintasan berikutnya. Fase pelatihan akan terus melakukan siklus ini di jaringan sampai kesalahan tidak bisa lebih rendah yang dikenal sebagai konvergensi.

Convolutional Neural Networks (CNNs) dibangun di atas arsitektur MLP tetapi dengan sejumlah perubahan penting. Pertama, lapisan diatur menjadi tiga dimensi, lebar, tinggi dan kedalaman. Kedua, node dalam satu lapisan tidak selalu terhubung ke semua node di lapisan berikutnya, tetapi sering kali hanya sub-wilayahnya. Untuk alasan yang dibahas, baik MLP dan CNN biasanya membuat pengklasifikasi yang baik, di mana CNN khususnya bekerja sangat baik dengan tugas klasifikasi gambar karena ekstraksi fitur dan bagian klasifikasinya. Saya percaya bahwa ini akan sangat efektif dalam menemukan pola dalam MFCC seperti efektif dalam menemukan pola dalam gambar. Ilustrasi proses ekstraksi audi ke MFCC dengan CNN tergambar dalam gambar di bawah ini.



Gambar 4.3. Proses Ekstraksi MFCC dengan CNN

Pada penelitian yang telah kami lakukan yang tergambar dalam proses ekstraksi MFCC dengan CNN di atas bahwa kinerja untuk memisahkan masing-masing audio yang menjadi beberapa layer. Ada 8 layer yang berhasil mengolah audio untuk diklasifikasikan ke dengan CNN, dense\_3, activation\_3, dropout\_2, dense\_4, activation\_4, dropout\_3, dense\_5, dan activation\_5. Jika dilihat dari hasilnya, terlihat pada layer dense\_3, dense\_4, dan dense\_5 yang mengalami perubahan jumlah param nya.

#### **4.4. Metodologi Pemrosesan**

##### **4.4.1. Data dan Pemisahan Data**

###### **4.4.1.1. Properti Audio Yang Memerlukan Normallsasi**

Mengikuti bagian sebelumnya, penelitian ini mengidentifikasi properti audio berikut yang perlu preprocessing untuk memastikan konsistensi di seluruh dataset:

- Audio Channels
- Sample rate
- Bit-depth

Kami akan terus menggunakan Librosa yang akan berguna untuk preprocessing dan ekstraksi fitur.

#### 4.4.1.2. Tahapan Preprocessing

Untuk sebagian besar *preprocessing* akan digunakan fungsi *load ()* *Librosa*. Kami akan membandingkan output dari *Librosa* dengan output default dari *library wavfile scipy* menggunakan file yang dipilih dari dataset. Konversi laju sampel Secara default, fungsi beban *Librosa* mengubah laju pengambilan sampel menjadi 22,05 KHz yang dapat digunakan sebagai tingkat perbandingan. Berikut adalah hasil *preprocessing* dari data audio dengan gambar.

```
In [20]: import librosa
         from scipy.io import wavfile as wav
         import numpy as np

         filename = 'F:/Ukharity ML Capstone-easter/EtherealSound Dataset sample/audio/10000-1-9-20.wav'

         librosa_audio, librosa_sample_rate = librosa.load(filename)
         scipy_sample_rate, scipy_audio = wav.read(filename)

         print('Original sample rate:', scipy_sample_rate)
         print('librosa sample rate:', librosa_sample_rate)

Original sample rate: 11025
librosa sample rate: 22050
```

Gambar 4.4.1.2. Tahapan *Preprocessing Sample Rate*

Hasil menunjukkan original sample rate sebesar 11025, dan sampel dari *Librosa* sebesar 22050. Bit-depth Fungsi muat *Librosa* juga akan menormalkan data sehingga rentang nilai antara -1 dan 1. Ini menghilangkan kerumitan kumpulan data yang memiliki rentang kedalaman bit yang luas. Hasil bit depth akan ditampilkan dalam gambar dibawah ini.

```
In [20]: import librosa
from scipy.io import wavfile as wav
import numpy as np

filename = 'F:/Macity-PL-Capstone-kaster/EthnicSound Dataset sample/audio/10004-1-0-20.wav'

librosa_audio, librosa_sample_rate = librosa.load(filename)
scipy_sample_rate, scipy_audio = wav.read(filename)

print('Original sample rate:', scipy_sample_rate)
print('Librosa sample rate:', librosa_sample_rate)

Original sample rate: 11025
Librosa sample rate: 22050
```

Gambar 4.4.1.2.1. Tahapan *Preprocessing Bit Depth*

Hasilnya adalah original audio file berkisar minimal hingga maksimalnya adalah -27378 hingga 32312. Pada file audio *Librosa* berkisar minimal hingga maksimalnya adalah -0.61050653 hingga 0.7709209. Gabungan saluran audio *Librosa* juga akan mengubah sinyal menjadi mono, yang berarti jumlah saluran akan selalu 1. Berikut adalah hasil impor gambar audio asli dengan 2 *channel* dengan audio pada *Librosa* dengan *channel* gabungan. Pada tahap ini, belum jelas apakah faktor lain juga perlu dipertimbangkan, seperti durasi sampel dan level volume.

Akan dilanjutkan apa adanya untuk sementara waktu dan kembali membahasnya nanti jika dianggap memengaruhi validitas metrik target kami.

#### 4.4.2. Fitur Ekstraksi

Seperti yang dijelaskan dalam proposal, akan dilakukan ekstraksi Koefisien *Cepstral Frekuensi Mel* (MFCC) dari sampel audio. MFCC merangkum distribusi frekuensi di seluruh ukuran jendela, sehingga memungkinkan untuk menganalisis

karakteristik frekuensi dan waktu suara. Representasi audio ini akan memungkinkan kami mengidentifikasi fitur untuk klasifikasi. Mengekstrak MFCC, kami akan menggunakan *fungsi mfcc () Librosa* yang menghasilkan MFCC dari data audio deret waktu. Berikut dibawah adalah gambar hasilnya.

```
In [32]: mfccs = librosa.feature.mfcc(y=librosa_audio, sr=librosa_sample_rate, n_mfcc=40)
print(mfccs.shape)
(40, 175)
```

Gambar 4.4.2 Fitur Ekstraksi

Ini menunjukkan *librosa* menghitung serangkaian 40 MFCC pada 175 frame. Selanjutnya akan diimpor hasil *librosa* tadi dalam bentuk grafik gambar dibawah ini. Mengekstrak MFCC untuk setiap file akan diekstrak MFCC untuk setiap file audio dalam dataset dan menyimpannya dalam *Panda Dataframe* bersama dengan label klasifikasinya. Jika sudah selesai, maka hasilnya akan ada 1200 file.

#### 4.4.3. Mengonversi Data dan Label

Akan digunakan *sklearn.preprocessing.LabelEncoder* untuk menyandikan data teks kategorikal menjadi data numerik yang dapat dimengerti model. Berikut adalah gambarnya.

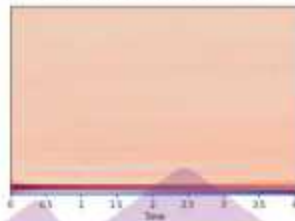
```
In [37]: from sklearn.preprocessing import LabelEncoder
from keras.utils import to_categorical

# Convert features and corresponding classification labels into numpy arrays
X = np.array(features).astype('float32')
y = np.array(class_label.tolist())

# Encode the classification labels
le = LabelEncoder()
yy = to_categorical(le.fit_transform(y))
```

```
In [12]: import librosa.display
librosa.display.spectrogram(wfc12, sr=librosa.sample_rate, x_axis='time')
Out[12]: 

```



Gambar 4.4.3. Mengonversi Data dan Label.

Karena CNN haus akan gambar, kami ingin mengubah suara menjadi gambar. Sinyal audio juga dapat direpresentasikan dengan cara lain. Selain memplot amplitudo sinyal audio terhadap waktu, kita juga dapat memplotnya sehubungan dengan frekuensi. Plot yang akan kita buat disebut spektrogram. Untuk memplot spektrogram kami memecah sinyal audio menjadi potongan milidetik dan menghitung *Short-Time Fourier Transform (STFT)* untuk setiap potongan. Kemudian memplot potongan waktu ini sebagai garis vertikal berwarna dalam spektrogram. Spektrogram mewakili konten frekuensi dalam audio sebagai warna dalam gambar. Isi frekuensi potongan milidetik dirangkai bersama sebagai batang vertikal berwarna. Spektrogram pada dasarnya adalah grafik dua dimensi, dengan dimensi ketiga diwakili oleh warna. Spektrogram adalah representasi gambar dari digit suara yang dihasilkan, setiap digit audio adalah sesuai dengan spektrogramnya.

#### 4.4.4. Memisahkan Dataset

Di sini akan digunakan `sklearn.model_selection.train_test_split` untuk membagi set data menjadi set pelatihan dan pengujian. Berikut adalah hasil gambarnya.

```
In [37]: from sklearn.preprocessing import LabelEncoder
from keras.utils import to_categorical

# Convert features and corresponding classification labels into numpy arrays
X = np.array(featuresdf.feature_values())
y = np.array(featuresdf.class_label.tolist())

# Encode the classification labels
le = LabelEncoder()
yy = to_categorical(le.fit_transform(y))

In [38]: # split the dataset
from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(X, yy, test_size=0.1, random_state=0)

Store the preprocessed data

In [39]: # save the preprocessed data for use by the next notebook
import pickle

pickle.dump(x_train, open('x_train.pkl', 'wb'))
pickle.dump(x_test, open('x_test.pkl', 'wb'))
pickle.dump(y_train, open('y_train.pkl', 'wb'))
pickle.dump(y_test, open('y_test.pkl', 'wb'))
pickle.dump(yy, open('yy.pkl', 'wb'))
pickle.dump(le, open('le.pkl', 'wb'))

Stored "x_train" (numpy)
Stored "x_test" (numpy)
Stored "y_train" (numpy)
Stored "y_test" (numpy)
Stored "yy" (numpy)
Stored "le" (LabelEncoder)
```

Gambar 4.4.4. Memisahkan Dataset

### 4.5. Implementasi

#### 4.5.1. Penerapan Model Awal -MLP

Akan dimulai dengan membangun Multilayer Jaringan Neural Perceptron (MLP) menggunakan Keras dan backend Tensorflow. Dimulai dengan model sekuensial sehingga kita dapat membangun model lapis demi lapis. Lalu dengan arsitektur model sederhana, yang terdiri dari tiga lapisan, lapisan masukan, lapisan

tersembunyi, dan lapisan keluaran. Ketiga lapisan tersebut akan menjadi jenis lapisan padat yang merupakan jenis lapisan standar yang digunakan dalam banyak kasus untuk jaringan saraf. Lapisan pertama akan menerima bentuk masukan. Karena setiap sampel berisi 40 MFCC (atau kolom) kita memiliki bentuk (1x40) ini berarti akan dimulai dengan bentuk masukan 40. Dua lapisan pertama akan memiliki 256 node. Fungsi aktivasi yang akan kita gunakan untuk 2 lapisan pertama adalah ULT, atau Aktivasi Linear Terserah. Fungsi aktivasi ini telah terbukti berfungsi dengan baik di jaringan saraf. Kami juga akan menerapkan nilai Dropout 50% pada dua lapisan pertama. Ini secara acak akan mengecualikan node dari setiap siklus pembaruan yang pada gilirannya menghasilkan jaringan yang mampu menghasilkan generasi yang lebih baik dan kecil kemungkinannya untuk menyesuaikan data pelatihan.

Lapisan keluaran akan memiliki 10 node (`num_labels`) yang cocok dengan jumlah klasifikasi yang mungkin. Aktivasi untuk lapisan keluaran adalah softmax. Softmax membuat jumlah keluaran menjadi 1 sehingga keluaran dapat diartikan sebagai probabilitas. Model kemudian akan membuat prediksi berdasarkan opsi mana yang memiliki probabilitas tertinggi. Berikut di bawah ini adalah hasilnya

```
In [21]: # retrieve the preprocessed data from previous notebook
%store -r x_train
%store -r x_test
%store -r y_train
%store -r y_test
%store -r yy
%store -r le
```



```

In [22]: import numpy as np
         from keras.models import Sequential
         from keras.layers import Dense, Dropout, Activation, Flatten
         from keras.layers import Convolution2D, MaxPooling2D
         from keras.optimizers import Adam
         from keras.utils import np_utils
         from sklearn import metrics

         num_labels = yy.shape[1]
         filter_size = 2

         # Construct model
         model = Sequential()

         model.add(Dense(256, input_shape=(48,)))
         model.add(Activation('relu'))
         model.add(Dropout(0.5))

         model.add(Dense(256))
         model.add(Activation('relu'))
         model.add(Dropout(0.5))

         model.add(Dense(num_labels))
         model.add(Activation('softmax'))

```

Gambar 4.5.1. Penerapan Model MLP dengan *Python*

#### 4.5.2. Mengkompilasi Model

Untuk mengkompilasi model kita, akan digunakan tiga parameter berikut:

- Fungsi kerugian - akan menggunakan `kategorikal_crossentropy`.

Ini adalah pilihan paling umum untuk klasifikasi. Skor yang lebih rendah menunjukkan bahwa model tersebut berperforma lebih baik.

- Metrik - akan menggunakan metrik akurasi yang memungkinkan kami melihat skor akurasi pada data validasi saat kami melatih model.

- Pengoptimal - akan menggunakan `adam` yang secara umum merupakan pengoptimal yang baik untuk banyak kasus penggunaan.

Berikut di bawah ini adalah hasil kompilasi modelnya

```
import numpy as np
from keras.models import Sequential
from keras.layers import Dense, Dropout, Activation, Flatten
from keras.layers import Convolution2D, MaxPooling2D
from keras.optimizers import Adam
from keras.utils import np_utils
from sklearn import metrics

num_labels = yy.shape[1]
filter_size = 2

# Construct model
model = Sequential()

model.add(Dense(256, input_shape=(40,)))
model.add(Activation('relu'))
model.add(Dropout(0.5))

model.add(Dense(256))
model.add(Activation('relu'))
model.add(Dropout(0.5))

model.add(Dense(num_labels))
model.add(Activation('softmax'))
```

Gambar 4.5.2.1. Penerapan arsitektur CNN pada Python

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 256)	10496
activation (Activation)	(None, 256)	0
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 256)	65792
activation_1 (Activation)	(None, 256)	0
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 0)	2056
activation_2 (Activation)	(None, 0)	0
-----		
Total params:	78,344	
Trainable params:	78,344	
Non-trainable params:	0	
-----		
Pre-training accuracy:	16.2500%	

Gambar 4.5.2.2. Hasil perubahan layer dengan penerapan arsitektur CNN

Dengan Layer dense 3 terdapat 256 bentuk yang menghasilkan 10496 param. Ketika layer aktivasi 3, dropout 2 dan selanjutnya dense 4 yang dibentuk sekitar 50%, akan menghasilkan 65792 param. Setelah itu layer aktivasi 4 sampai dense 5 bekerja hingga menjadikan 8 bentuk. Jadi total params ada 78.344, training param 78.344 dan menghasilkan pre training akurasi sebesar 16.2500%

### 4.5.3. Training

Akan dimulai dengan 100 epoch yang merupakan frekuensi model akan mengilir data. Model tersebut akan diperbaiki pada setiap siklus hingga mencapai titik tertentu. Lalu akan dimulai dengan ukuran tumpukan yang rendah, karena memiliki ukuran tumpukan yang besar dapat mengurangi kemampuan generalisasi model. Berikut adalah hasilnya berupa gambar.

```
in [30]: from keras.callbacks import ModelCheckpoint
from keras.callbacks import EarlyStopping
from keras import optimizers

# Compile model
model.compile(optimizer=optimizer, loss=loss, metrics=['accuracy'])

# Train model
model.fit(train_data_loader_generator, validation_data_loader_generator,
        epochs=100, verbose=1, validation_freq=10)

# Evaluate model
test_loss, test_acc = model.evaluate(test_data_loader_generator)
print("Test Loss: %s, Test Accuracy: %s" % (test_loss, test_acc))
```

```
Epoch 001/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 002/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 003/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 004/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 005/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 006/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 007/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 008/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 009/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 010/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 011/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 012/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 013/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 014/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 015/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 016/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 017/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 018/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 019/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 020/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 021/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 022/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 023/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 024/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 025/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 026/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 027/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 028/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 029/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 030/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 031/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 032/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 033/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 034/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 035/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 036/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 037/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 038/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 039/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 040/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 041/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 042/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 043/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 044/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 045/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 046/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 047/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 048/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 049/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
Epoch 050/100: Train Loss: 0.8257, Validation Loss: 0.8257, Validation Accuracy: 0.0000
```

Gambar 4.5.3. Training

Dalam penelitian ini menggunakan Keras sebagai perintah dalam Bahasa pemrogramannya, kami memberitahu keras untuk menerima pengujian yang

ditetapkan sebagai set dari validasinya. Model melakukan iterasi lebih dari 10 epoch dan meningkatkan parameternya sampai 100 epoch hingga mendapatkan nilai akurasi tertinggi. Dalam hal ini adalah kualitas keakuratan pengujian saja. Idealnya, peneliti harus memiliki jumlah epoch yang lebih besar dan harus menghentikan jaringan saat akurasi pengujian berhenti meningkat.

#### 4.5.4. Pengujian Model

Akan ditinjau keakuratan model pada set data pelatihan dan pengujian.

Berikut adalah hasil nya.

```
In [7]: # Evaluating the model on the training and testing set
score = model.evaluate(x_train, y_train, verbose=0)
print("Training Accuracy: ", score[1])

score = model.evaluate(x_test, y_test, verbose=0)
print("Testing Accuracy: ", score[1])

Training Accuracy:  0.9770833253860474
Testing Accuracy:  0.925000011920929
```

Gambar 4.5.4 Pengujian Model

Dengan pengujian model suara alat musik gamelan Jawa yang datanya dibuat dengan 16 bit dan 11-22 kHz, mendapatkan training akurasi sebesar 0,9770 yang artinya adalah 97%, testing akurasi yang didapat sebesar 0,9250; yang artinya 92%.

#### 4.5.5. Prediksi

Di sini akan dibuat metode yang memungkinkan untuk menguji prediksi model pada file audio .wav tertentu. Berikut adalah hasil nya.

```

In [27]: import librosa
import numpy as np

def extract_feature(file_name):

    try:
        audio_data, sample_rate = librosa.load(file_name, res_type='kaiser_fast')
        mfccs = librosa.feature.mfcc(y=audio_data, sr=sample_rate, n_mfcc=40)
        mfccscaled = np.mean(mfccs.T,axis=0)

    except Exception as e:
        print("Error encountered while parsing file: ", file)
        return None, None

    return np.array([mfccscaled])

In [28]: def print_prediction(file_name):
prediction_feature = extract_feature(file_name)

predicted_vector = model.predict_classes(prediction_feature)
predicted_class = le.inverse_transform(predicted_vector)
print("The predicted class is: ", predicted_class[0], "\n")

predicted_proba_vector = model.predict_proba(prediction_feature)
predicted_proba = predicted_proba_vector[0]
for i in range(len(predicted_proba)):
    category = le.inverse_transform(np.array([i]))
    print(category[0], " : ", format(predicted_proba[i], '.32f') )

```

Gambar 4.5.5. Penerapan Prediksi audio dengan *python*

#### 4.5.6. Uji Validasi

Uji dengan data sampel. Pemeriksaan kewarasan awal untuk memverifikasi prediksi menggunakan sub-bagian dari file audio sampel yang kami jelajahi di metadata pertama. Diharapkan sebagian besar ini diklasifikasikan dengan benar.

```

In [29]: # Class: suling

filename = 'F:\Macity-PC-Capture-master\EthicSound_Dataset_sample\audio\10007-0-0-2.wav'
print_prediction(filename)

The predicted class is: Suling

Angklung      : 0.00000000000045070463460592758187
Bonang        : 0.000000000000840433451998467972995
Demung        : 0.0000000000000000000000002755161129
Gong          : 0.000000000001091206322709391471903
Kendang       : 0.00000000000000000000004835146627024794
Rebab        : 0.000000000000796749000052557715054
Saron         : 0.00000000022740023752900621634099
Suling        : 1.0000000000000000000000000000000000

In [30]: # Class: saron

filename = 'F:\Macity-PC-Capture-master\EthicSound_Dataset_sample\audio\10007-0-0-1.wav'
print_prediction(filename)

The predicted class is: Saron

Angklung      : 0.0000000000000000000000002687354229144
Bonang        : 0.0000000000000000000000001338942
Demung        : 0.0000000000000000000000002973813794281
Gong          : 0.00000000000000000000000013377630193021
Kendang       : 0.0000000000000000000000000000000000
Rebab        : 0.0000000000000000000000000000000000
Saron         : 0.0000000000000000000000000000000000
Suling        : 0.0000000000000000000000000000000000

```

Gambar 4.5.6. Hasil Validasi

Dari ini Model pemeriksaan singkat tampaknya memprediksi dengan baik. Dalam hasil validasi yang akan muncul adalah *confidence level* dari masing-masing alat musik gamelan yang sudah diinputkan. Dari *confidence level* yang paling tertinggi adalah hasil dari validasi, apakah alat musik itu.

Di sini akan digunakan sampel berbagai suara bebas hak cipta yang bukan merupakan bagian dari data pengujian atau pelatihan kami untuk memvalidasi lebih lanjut model kami. Data yang kami gunakan untuk melakukan validasi dari masing-masing alat musik gamelan adalah berasal dari selain audio yang ada di metadata kami. Tentunya pengambilan audio ini sudah memenuhi syarat permainan standar gamelan serta nada dan pattern nya. Berikut adalah detail data pengujian dan validasi alat musik gamelan yang tersaji dalam tabel di bawah ini.

Kendang merupakan satu dari banyak alat musik gamelan yang kami jadikan penelitian ini tidak memiliki nada, dikarenakan alat musik ini menggunakan pola ritmis atau ketukan saja, tetapi kendang dapat direkam dan memiliki frekuensi tersendiri sehingga tetap bisa dijadikan sebagai validasi klasifikasi alat musik gamelan yang kami buat pada penelitian ini. Kami menggunakan 20 pola nada atau pattern terhadap masing-masing file audio yang di uji.

Tabel 4.5.7.1 Data Pengujian Angklung

No	Data Pengujian			
	IDName	Sample Rate	Nada/Pattern	Waktu
1	tes_angklung_1	11 kHz	5	1 detik
2	tes_angklung_2	12 kHz	5	1 detik
3	tes_angklung_3	22 kHz	5	1 detik
4	tes_angklung_4	11 kHz	5	1 detik
5	tes_angklung_5	12 kHz	7	2 detik
6	tes_angklung_6	22 kHz	8	3 detik
7	tes_angklung_7	11 kHz	8	4 detik
8	tes_angklung_8	22 kHz	1,2,3,4,5,6,7,8	1menit 30 detik
9	tes_angklung_9	11 kHz	5	1 detik
10	tes_angklung_10	11 kHz	1	7 detik
11	tes_angklung_11	11 kHz	1	34 detik
12	tes_angklung_12	11 kHz	8	34 detik
13	tes_angklung_13	11 kHz	4	34 detik
14	tes_angklung_14	11 kHz	6	34 detik
15	tes_angklung_15	11 kHz	3	34 detik
16	tes_angklung_16	11 kHz	2	34 detik
17	tes_angklung_17	11 kHz	7	34 detik
18	tes_angklung_18	11 kHz	5	34 detik
19	tes_angklung_19	11 kHz	1	1 detik
20	tes_angklung_20	11 kHz	1	3 detik

Tabel 4.5.7.2 Data Pengujian Bonang

21	tes_bonang_1	11 kHz	4,1,4,1	3 detik
22	tes_bonang_2	12 kHz	4,4,1	3 detik
23	tes_bonang_3	22 kHz	4	2 detik
24	tes_bonang_4	22 kHz	1	2 detik
25	tes_bonang_5	11 kHz	4,1	2 detik
26	tes_bonang_6	12 kHz	3,3,1	2 detik
27	tes_bonang_7	12 kHz	4,4,1,1,4	3 detik
28	tes_bonang_8	22 kHz	1,4,4	3 detik
29	tes_bonang_9	11 kHz	3,3,1	2 detik
30	tes_bonang_10	12 kHz	4,4,1,1,4	3 detik
31	tes_bonang_11	22 kHz	4,1,4	2 detik
32	tes_bonang_12	22 kHz	4,3,1,4,3	11 detik
33	tes_bonang_13	12 kHz	5,5,4,5,5,4	5 detik
34	tes_bonang_14	12 kHz	3	3 detik
35	tes_bonang_15	12 kHz	4,3,1,4,3	3 detik
36	tes_bonang_16	22 kHz	1,3	1 detik
37	tes_bonang_17	22 kHz	3,1,1,3	2 detik
38	tes_bonang_18	11 kHz	5,1,4,4,1,1	0 detik
39	tes_bonang_19	12 kHz	5,1,4,4,1,1	0 detik
40	tes_bonang_20	22 kHz	3,4,3,5,4,3,1	7 detik

Tabel 4.5.7.3 Data Pengujian Demung

41	tes_demung_1	11 kHz	7,4	1 detik
42	tes_demung_2	12 kHz	3,1	1 detik
43	tes_demung_3	22 kHz	3,4	1 detik
44	tes_demung_4	11 kHz	5,7	1 detik
45	tes_demung_5	11 kHz	4,5	1 detik
46	tes_demung_6	12 kHz	4,3	1 detik
47	tes_demung_7	22 kHz	7,5	1 detik
48	tes_demung_8	11 kHz	4,5	1 detik
49	tes_demung_9	12 kHz	3,4	1 detik
50	tes_demung_10	22 kHz	5,7	1 detik
51	tes_demung_11	11 kHz	8,7	1 detik
52	tes_demung_12	12 kHz	5,7	1 detik
53	tes_demung_13	22 kHz	5,2	1 detik
54	tes_demung_14	11 kHz	5,7	1 detik
55	tes_demung_15	12 kHz	5,1	1 detik
56	tes_demung_16	12 kHz	3,1	1 detik
57	tes_demung_17	12 kHz	1,3	1 detik
58	tes_demung_18	22 kHz	4,1	1 detik
59	tes_demung_19	22 kHz	7,3	1 detik
60	tes_demung_20	11 kHz	1,3	1 detik



Tabel 4.5.7.4 Data Pengujian Gong

01 tes_gong_1	11 kHz	1	24 detik
02 tes_gong_2	11 kHz	1	32 detik
03 tes_gong_3	11 kHz	1	20 detik
04 tes_gong_4	11 kHz	1	30 detik
05 tes_gong_5	11 kHz	1	24 detik
06 tes_gong_6	11 kHz	1	20 detik
07 tes_gong_7	11 kHz	1	24 detik
08 tes_gong_8	22 kHz	1	2 detik
09 tes_gong_9	22 kHz	1	3 detik
70 tes_gong_10	11 kHz	1	1 detik
71 tes_gong_11	12 kHz	1	2 detik
72 tes_gong_12	22 kHz	1	3 detik
73 tes_gong_13	11 kHz	5	1 detik
74 tes_gong_14	22 kHz	5	1 detik
75 tes_gong_15	22 kHz	5	2 detik
76 tes_gong_16	11 kHz	4	1 detik
77 tes_gong_17	12 kHz	4	2 detik
78 tes_gong_18	11 kHz	1	4 detik
79 tes_gong_19	12 kHz	1	2 detik
80 tes_gong_20	22 kHz	1	2 detik

Tabel 4.5.7.5 Data Pengujian Kendang

81 tes_kendang_1	11 kHz	Pattern 1	3 detik
82 tes_kendang_2	12 kHz	Pattern 2	3 detik
83 tes_kendang_3	22 kHz	Pattern 3	3 detik
84 tes_kendang_4	11 kHz	Pattern 4	3 detik
85 tes_kendang_5	12 kHz	Pattern 5	3 detik
86 tes_kendang_6	22 kHz	Pattern 6	3 detik
87 tes_kendang_7	11 kHz	Pattern 7	3 detik
88 tes_kendang_8	12 kHz	Pattern 8	3 detik
89 tes_kendang_9	22 kHz	Pattern 9	3 detik
90 tes_kendang_10	11 kHz	Pattern 10	3 detik
91 tes_kendang_11	12 kHz	Pattern 11	3 detik
92 tes_kendang_12	22 kHz	Pattern 12	3 detik
93 tes_kendang_13	11 kHz	Pattern 13	3 detik
94 tes_kendang_14	12 kHz	Pattern 14	3 detik
95 tes_kendang_15	22 kHz	Pattern 15	3 detik
96 tes_kendang_16	11 kHz	Pattern 16	3 detik
97 tes_kendang_17	12 kHz	Pattern 17	3 detik
98 tes_kendang_18	22 kHz	Pattern 18	3 detik
99 tes_kendang_19	11 kHz	Pattern 19	3 detik
100 tes_kendang_20	22 kHz	Pattern 20	3 detik

Tabel 4.5.7.6 Data Pengujian Rebab

101 tes_rebab_1	22 kHz	0	2 detik
102 tes_rebab_2	22 kHz	4	2 detik
103 tes_rebab_3	22 kHz	0#	1 detik
104 tes_rebab_4	11 kHz	0	3 detik
105 tes_rebab_5	12 kHz	7	1 detik
106 tes_rebab_6	22 kHz	7,3,7	2 detik
107 tes_rebab_7	22 kHz	2#	2 detik
108 tes_rebab_8	11 kHz	2,0,5,4,	4 detik
109 tes_rebab_9	12 kHz	1	1 detik
110 tes_rebab_10	12 kHz	2	1 detik
111 tes_rebab_11	22 kHz	3	1 detik
112 tes_rebab_12	22 kHz	3,2,3	2 detik
113 tes_rebab_13	12 kHz	3#	2 detik
114 tes_rebab_14	22 kHz	4	1 detik
115 tes_rebab_15	22 kHz	0	5 detik
116 tes_rebab_16	11 kHz	3,0	2 detik
117 tes_rebab_17	12 kHz	1#	1 detik
118 tes_rebab_18	22 kHz	2	1 detik
119 tes_rebab_19	11 kHz	1,2	1 detik
120 tes_rebab_20	22 kHz	1,2	1 detik

Tabel 4.5.7.7 Data Pengujian Saron

121 tes_saron_1	11 kHz	0	1 detik
122 tes_saron_2	12 kHz	0#	1 detik
123 tes_saron_3	22 kHz	5#	1 detik
124 tes_saron_4	11 kHz	2#	1 detik
125 tes_saron_5	12 kHz	4#	1 detik
126 tes_saron_6	22 kHz	5#	1 detik
127 tes_saron_7	11 kHz	7	1 detik
128 tes_saron_8	12 kHz	4	1 detik
129 tes_saron_9	22 kHz	5	1 detik
130 tes_saron_10	11 kHz	4	1 detik
131 tes_saron_11	12 kHz	5,7,5	1 detik
132 tes_saron_12	22 kHz	4,5,2	2 detik
133 tes_saron_13	11 kHz	5,4,3,5,4,3,1	3 detik
134 tes_saron_14	12 kHz	3,5,3,7,3,4	3 detik
135 tes_saron_15	22 kHz	7,3,7,4,3	4 detik
136 tes_saron_16	11 kHz	7,5,4,3,1,3,4	5 detik
137 tes_saron_17	12 kHz	3,5,3,7,3,4,7,5,4,3,1,3,4	7 detik
138 tes_saron_18	22 kHz	1,5,1,7,1,4,7,5,4,1,1,1,4	7 detik
139 tes_saron_19	11 kHz	7,5,4,3,1,3,4,5,4,3,5,4,3,1	7 detik
140 tes_saron_20	12 kHz	4,5,2,5,4,3,5,4,3,1,5,7,5	7 detik

Tabel 4.5.7.8 Data Pengujian Suling

141 tes_suling_1	11 kHz	0,7,0,6,4,3	1 detik
142 tes_suling_2	12 kHz	0,4	1 detik
143 tes_suling_3	22 kHz	0,4	1 detik
144 tes_suling_4	11 kHz	5,4,5,4	1 detik
145 tes_suling_5	12 kHz	4	1 detik
140 tes_suling_6	22 kHz	4	1 detik
147 tes_suling_7	11 kHz	4	1 detik
148 tes_suling_8	12 kHz	3	1 detik
149 tes_suling_9	22 kHz	3,2,1	2 detik
150 tes_suling_10	11 kHz	4,3,2,1,4,3,2	2 detik
151 tes_suling_11	12 kHz	0	1 detik
152 tes_suling_12	22 kHz	2	1 detik
153 tes_suling_13	11 kHz	0,7,6	2 detik
154 tes_suling_14	12 kHz	0,7	1 detik
155 tes_suling_15	22 kHz	1,2,7	2 detik
156 tes_suling_16	11 kHz	3,3,3	2 detik
157 tes_suling_17	12 kHz	3,3,2	2 detik
158 tes_suling_18	22 kHz	0,3	1 detik
159 tes_suling_19	11 kHz	0,7,3,1	2 detik
160 tes_suling_20	12 kHz	0,7,0,7,3,1	4 detik

Berikut adalah hasilnya yang akan ditunjukkan dengan gambar sebagai berikut.

```
In [36]: filename = 'F:/udacity-ML-capstone-master/Evaluation audio/tes_suling_1.wav'
print_prediction(filename)
The predicted class is: Suling

Angklung      | 0.00550214092836117744445800781250
Bonang        | 0.00159959669690579175640096679688
Demung        | 0.00000000021615588485798482309312
Gong          | 0.00149287190288305282592773437500
Kendang       | 0.000019245400039564352244138717651
Rebab         | 0.00054759002523406747016906738281
Saron         | 0.00010058253974420949616703796387
Suling        | 0.09073797464370727539062580000000
```

```
In [42]: filename = 'F:/udacity-ML-capstone-master/Evaluation audio/tes_rebab_1.wav'
print_prediction(filename)
A sample data weighted towards gun shot - peak in the dog barking sample is similar in shape to the
The predicted class is: Rebab

Angklung      | 0.00000000099491334141500767297298
Bonang        | 0.00000110279147104109062702083588
Demung        | 0.0000000000000001701096350991174
Gong          | 0.00000000218968748842625174991554
Kendang       | 0.00000000000121348461252021536702
Rebab         | 0.09999600558471579687100000000000
Saron         | 0.00000000000022292093842597029218
Suling        | 0.00000000318395443497762149490082
```



Berikut adalah tabel hasil perbandingannya. Perbandingan data testing masing masing alat musik gamelan akan menunjukkan *confidence level* yang nantinya jumlahnya akan terhitung akurasinya.

Tabel 4.5.6. Hasil perbandingan data testing pada alat musik Angklung

No	Nama File Testing	Class	Hasil Prediksi Class	Hasil Confidence Level Testing Class	Rata-Rata Confidence Level
1	tes_angklung_1		Angklung	0.837981304	
2	tes_angklung_2		Rebab	0.039010034	
3	tes_angklung_3		Rebab	0.478294738	
4	tes_angklung_4		Rebab	0.179050891	
5	tes_angklung_5		Angklung	0.37022113	
6	tes_angklung_6		Angklung	0.38311038	
7	tes_angklung_7		Gong	0.389579147	
8	tes_angklung_8		Gong	0.006823433	
9	tes_angklung_9		Suling	0.00007252	
10	tes_angklung_10		Angklung	1	
11	tes_angklung_11	1. Angklung	Angklung	1	0.730142496
12	tes_angklung_12		Angklung	1	
13	tes_angklung_13		Angklung	0.889999915	
14	tes_angklung_14		Angklung	0.933999959	
15	tes_angklung_15		Angklung	0.962805095	
16	tes_angklung_16		Angklung	1	
17	tes_angklung_17		Angklung	0.93687483	
18	tes_angklung_18		Angklung	1	
19	tes_angklung_19		Angklung	0.888779321	
20	tes_angklung_20		Angklung	1	
Jumlah Confidence Level Angklung				14.00284896	

Dari hasil penelitian dengan audio Angklung yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul rebab sebanyak 3 kali, gong sebanyak 2 kali dan suling sebanyak 1 kali. Sehingga menghasilkan *confidence level* rata-rata 0,7301. Ini artinya klasifikasi terhadap angklung dapat digunakan dengan baik.

Tabel 4.5.6.1. Hasil perbandingan data testing pada alat musik Bonang

22 tes_bonang_1		Bonang	1	
23 tes_bonang_2		Bonang	1	
24 tes_bonang_3		Bonang	1	
25 tes_bonang_4		Bonang	0.889999915	
26 tes_bonang_5		Bonang	1	
27 tes_bonang_6		Bonang	1	
28 tes_bonang_7		Bonang	1	
29 tes_bonang_8		Bonang	1	
30 tes_bonang_9		Bonang	1	
31 tes_bonang_10		Bonang	1	
32 tes_bonang_11		Bonang	1	
33 tes_bonang_12		Bonang	0.889999915	
34 tes_bonang_13		Bonang	0.93687483	
35 tes_bonang_14		Bonang	0.933999959	
36 tes_bonang_15		Bonang	0.888779321	
37 tes_bonang_16		Bonang	0.962805095	
38 tes_bonang_17		Bonang	0.933999959	
39 tes_bonang_18		Bonang	0.933999959	
40 tes_bonang_19		Bonang	0.888779321	
41 tes_bonang_20		Bonang	1	
Jumlah Confidence Level Bonang				33.6734891

Dari hasil penelitian dengan audio Bonang yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul bonang 20 kali. Sehingga menghasilkan *confidence level* rata-rata 0,9838. Ini artinya klasifikasi terhadap bonang dapat digunakan dengan baik.

Tabel 4.5.6.2. Hasil perbandingan data testing pada alat musik Demung

41 tes_demung_1	Demung	0.654100838
42 tes_demung_2	Demung	0.530050834
43 tes_demung_3	Saron	0.494300593
44 tes_demung_4	Demung	0.500320457
45 tes_demung_5	Demung	0.649260638
46 tes_demung_6	Demung	0.731225331
47 tes_demung_7	Demung	0.552844226
48 tes_demung_8	Demung	0.640658049
49 tes_demung_9	Demung	0.409155641
50 tes_demung_10	Saron	0.400515788
51 tes_demung_11	Demung	0.720178986
52 tes_demung_12	Demung	0.637282394
53 tes_demung_13	Demung	0.527392785
54 tes_demung_14	Demung	0.601236418
55 tes_demung_15	Demung	0.555110275
56 tes_demung_16	Saron	0.510182837
57 tes_demung_17	Demung	0.518634798
58 tes_demung_18	Demung	0.558331173
59 tes_demung_19	Demung	0.408881594
60 tes_demung_20	Demung	0.562905390
Jumlah Confidence Level Demung :		11.45142567
		0.572570784

Dari hasil pengujian penelitian ini terhadap alat musik Demung muncul prediksi klasifikasi saron sebanyak 3 kali dalam jumlah 20 kali pengujian total. Sehingga menghasilkan *confidence level* rata-rata sebanyak 0,5725. Diketahui dari semua pengujian muncul rentang angka 0.4-0.7.

Tabel 4.5.6.3. Hasil perbandingan data testing pada alat musik Gong

61 tes_gong_1		Gong	0.00000000	
62 tes_gong_2		Gong	0.00000000	
63 tes_gong_3		Gong	1	
64 tes_gong_4		Gong	1	
65 tes_gong_5		Gong	1	
66 tes_gong_6		Gong	1	
67 tes_gong_7		Gong	0.00000000	
68 tes_gong_8		Gong	1	
69 tes_gong_9		Gong	1	
70 tes_gong_10		Gong	0.00000001	
71 tes_gong_11	4. Gong	Gong	0.00000007	0.00992408
72 tes_gong_12		Gong	0.00011250	
73 tes_gong_13		Gong	0.00000045	
74 tes_gong_14		Gong	0.00004227	
75 tes_gong_15		Gong	0.00007055	
76 tes_gong_16		Gong	0.00043565	
77 tes_gong_17		Gong	0.00043556	
78 tes_gong_18		Gong	1	
79 tes_gong_19		Gong	1	
80 tes_gong_20		Gong	1	
Jumlah Confidence Level Gong :			19.99948178	

Dari hasil penelitian dengan audio Gong yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul Gong 20 kali. Sehingga menghasilkan *confidence level* rata-rata 0,9999. Ini artinya klasifikasi terhadap bonang dapat digunakan dengan baik.

Tabel 4.5.6.4. Hasil perbandingan data testing pada alat musik Kendang

81 tes_kendang_1		Kendang	0.00540070	
82 tes_kendang_2		Kendang	0.00040342	
83 tes_kendang_3		Kendang	0.00038924	
84 tes_kendang_4		Kendang	0.00743265	
85 tes_kendang_5		Kendang	0.00088779	
86 tes_kendang_6		Kendang	0.00706727	
87 tes_kendang_7		Kendang	0.00874015	
88 tes_kendang_8		Kendang	0.00770282	
89 tes_kendang_9		Kendang	0.00874556	
90 tes_kendang_10	5. Kendang	Kendang	0.00024713	0.0070878
91 tes_kendang_11		Kendang	0.00423067	
92 tes_kendang_12		Kendang	0.00011636	
93 tes_kendang_13		Kendang	0.00306509	
94 tes_kendang_14		Kendang	0.00210517	
95 tes_kendang_15		Kendang	0.00145500	
96 tes_kendang_16		Kendang	0.00553430	
97 tes_kendang_17		Kendang	0.00148348	
98 tes_kendang_18		Kendang	0.53000846	
99 tes_kendang_19		Kendang	0.00124301	
100 tes_kendang_20	Kendang	0.00048722		
Jumlah Confidence Level Kendang :			20.34175590	

Dari hasil penelitian dengan audio Kendang yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul Kendang 20 kali. Sehingga menghasilkan *confidence level* rata-rata 0,9608. Ini artinya klasifikasi terhadap bonang dapat digunakan dengan baik.

Tabel 4.5.6.5. Hasil perbandingan data testing pada alat musik Rebab

101 tes_rebab_1	Rebab	1
102 tes_rebab_2	Rebab	1
103 tes_rebab_3	Rebab	1
104 tes_rebab_4	Rebab	1
105 tes_rebab_5	Rebab	0,99955484
106 tes_rebab_6	Rebab	1
107 tes_rebab_7	Rebab	1
108 tes_rebab_8	Rebab	1
109 tes_rebab_9	Rebab	1
110 tes_rebab_10	Rebab	1
111 tes_rebab_11	Rebab	1
112 tes_rebab_12	Rebab	1
113 tes_rebab_13	Rebab	0,99999998
114 tes_rebab_14	Rebab	0,99999999
115 tes_rebab_15	Rebab	1
116 tes_rebab_16	Rebab	1
117 tes_rebab_17	Rebab	1
118 tes_rebab_18	Rebab	1
119 tes_rebab_19	Rebab	0,999999765
120 tes_rebab_20	Rebab	1
Jumlah Confidence Level Rebab		18,93863207

ii. Rebab

0,96081604

Dari hasil penelitian dengan audio Rebab yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul Rebab 20 kali. Sehingga menghasilkan *confidence level* rata-rata 0,9708. Ini artinya klasifikasi terhadap bonang dapat digunakan dengan baik.



Tabel 4.5.6.6. Hasil perbandingan data testing pada alat musik Saron

221 tes_saron_1	Demung	0.39931495
222 tes_saron_2	Demung	0.47136493
223 tes_saron_3	Saron	0.984217087
224 tes_saron_4	Saron	1.158223488
225 tes_saron_5	Saron	0.851842682
226 tes_saron_6	Demung	0.47128875
227 tes_saron_7	Saron	0.530942889
228 tes_saron_8	Demung	0.45070191
229 tes_saron_9	Saron	0.89171770
230 tes_saron_10	Saron	0.538813520
231 tes_saron_11	Demung	0.324668736
232 tes_saron_12	Demung	0.438421388
233 tes_saron_13	Saron	0.312380174
234 tes_saron_14	Demung	0.486211811
235 tes_saron_15	Saron	0.812612137
236 tes_saron_16	Demung	0.418899221
237 tes_saron_17	Saron	0.525368994
238 tes_saron_18	Demung	0.451468812
239 tes_saron_19	Demung	0.429282941
240 tes_saron_20	Saron	0.330144219
Jumlah Confidence Level Saron		0.52338331
		0.5253

Dari hasil penelitian dengan audio Saron yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul Saron sebanyak 10, Demung sebanyak 10. Sehingga menghasilkan *confidence level* rata-rata 0,5253. Ini artinya klasifikasi terhadap Saron banyak memerlukan tambahan data, dikarenakan secara subjektif, Suara saron dan Demung memiliki has yang sama.

Tabel 4.5.6.7. Hasil perbandingan data testing pada alat musik Suling

141 tes_suling_1	Suling	1
142 tes_suling_2	Suling	0.999999981
143 tes_suling_3	Suling	1
144 tes_suling_4	Suling	0.995117825
145 tes_suling_5	Suling	0.999999945
146 tes_suling_6	Suling	1
147 tes_suling_7	Suling	0.999999981
148 tes_suling_8	Suling	1
149 tes_suling_9	Suling	1
150 tes_suling_10	Suling	1
151 tes_suling_11	Suling	0.999999907
152 tes_suling_12	Suling	0.999999376
153 tes_suling_13	Suling	0.999999908
154 tes_suling_14	Suling	1
155 tes_suling_15	Suling	1
156 tes_suling_16	Suling	1
157 tes_suling_17	Suling	1
158 tes_suling_18	Suling	1
159 tes_suling_19	Suling	1
160 tes_suling_20	Suling	0.999997917
Jumlah Confidence Level Suling		18.98511382
		0.98755381

Dari hasil penelitian dengan audio Suling yang diuji sebanyak 20 kali menunjukkan hasil prediksi *class* nya muncul Suling 20 kali. Sehingga

menghasilkan *confidence level* rata-rata 0,9997. Ini artinya klasifikasi terhadap bonang dapat digunakan dengan baik.

#### 4.6. Hasil Evaluasi

Evaluasi untuk penelitian ini adalah “Akurasi Klasifikasi” yang akan didefinisikan sebagai presentase prediksi yang benar. *Akurasi sama dengan klasifikasi yang benar dibagi dengan jumlah klasifikasi*. Klasifikasi Akurasi dianggap sebagai metrik pilihan yang optimal karena diasumsikan bahwa dataset akan relatif simetris dengan ini menjadi pengklasifikasi multi-kelas dimana target kelas data umumnya akan memiliki ukuran yang seragam. Sebelumnya, ketika kita dihadapkan dengan permasalahan *machine learning*, akan banyak sekali pertanyaan bagaimana cara mendapatkan model yang baik, oleh karena itu kita harus mengukur kinerja suatu model yang telah kita buat, dan ini adalah langkah yang terpenting. Salah satu Teknik yang dapat kita gunakan untuk mengukur kinerja suatu model terutama tentang klasifikasi dengan *machine learning* adalah menggunakan *confusion matrix*. *Confusion matrix* juga sering disebut *error matrix*. Pada dasarnya *confusion matrix* memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem (model) dengan hasil klasifikasi sebenarnya.

Berikut adalah sajian data pada *confusion matrix* klasifikasi alat musik gamelan jawa.

Tabel 4.6. Hasil Confusion Matrix

Class/ Sample	Angklung	Bonang	Demung	Gong	Kendang	Rebab	Saron	Suling	Jumlah
Angklung	14	0	0	2	0	3	0	1	20
Bonang	0	20	0	0	0	0	0	0	20
Demung	0	0	17	0	0	0	3	0	20
Gong	0	0	0	20	0	0	0	0	20
Kendang	0	0	0	0	20	0	0	0	20
Rebab	0	0	0	0	0	20	0	0	20
Saron	0	0	10	0	0	0	10	0	20
Suling	0	0	0	0	0	0	0	20	20
Jumlah	14	20	27	22	20	23	13	21	160

Pada tabel ini, menghasilkan beberapa angka yang menunjukkan 4 istilah yang akan kita gunakan sebagai representasi hasil proses klasifikasi pada alat musik gamelan Jawa. Ke empat itu adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). True Positive (TP) merupakan data positif yang diprediksi benar. True Negative (TN) merupakan data negatif yang diprediksi benar. False Positive (FP) merupakan data negatif namun diprediksi sebagai data positif. False Negative (FN) merupakan data positif namun diprediksi sebagai data negatif. Matrix di atas merupakan hasil dari evaluasi dari 8 kelas alat musik gamelan Jawa yang telah berhasil diidentifikasi kebenarannya.

#### 4.6.1. Menghitung Akurasi

Akurasi adalah presentase dari total data yang diidentifikasi dan dinilai benar. Akurasi juga salah satu metrik untuk mengevaluasi model klasifikasi. Secara deskripsi, akurasi merupakan fraksi prediksi model yang kita terapkan yang sudah benar. Menghitung akurasi sama dengan TP (True Positive) dibagi dengan total data. Berikut adalah tabel akurasi.

Tabel 4.6.1. Hasil Akurasi Data

Class/ Sample	Angklung	Bonang	Demung	Gong	Kendang	Rebab	Saron	Suling	Jumlah
Angklung	14	0	0	0	2	0	3	0	20
Bonang	0	20	0	0	0	0	0	0	20
Demung	0	0	17	0	0	0	3	0	20
Gong	0	0	0	20	0	0	0	0	20
Kendang	0	0	0	0	20	0	0	0	20
Rebab	0	0	0	0	0	20	0	0	20
Saron	0	0	10	0	0	0	10	0	20
Suling	0	0	0	0	0	0	0	20	20
Jumlah	14	20	17	20	20	20	13	21	160

Ditunjukkan hasil kebenaran dari angklung, bonang, demung, gong, kendang, rebab, saron, dan suling sebanyak masing-masing 14,20,17,20,20,20,10,20. Jika dijumlahkan adalah 141. Dengan total data yang diuji adalah sebanyak 160. Jadi Akurasi adalah 141 dibagi dengan 160 menjadi 0,8812.

#### 4.6.2. Menghitung Precision

Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Berikut adalah tabel hasilnya.

Tabel 4.6.2. Hasil Klasifikasi Precision

Class/ Sample	Angklung	Bonang	Demung	Gong	Kendang	Rebab	Saron	Suling	Jumlah
A. Angklung	14	0	0	0	2	0	3	0	20
B. Bonang	0	20	0	0	0	0	0	0	20
C. Demung	0	0	17	0	0	0	3	0	20
D. Gong	0	0	0	20	0	0	0	0	20
E. Kendang	0	0	0	0	20	0	0	0	20
F. Rebab	0	0	0	0	0	20	0	0	20
G. Saron	0	0	10	0	0	0	10	0	20
H. Suling	0	0	0	0	0	0	0	20	20
Jumlah	14	20	17	20	20	20	13	21	160

Terdapat warna-warna yang menunjukkan presisi dari setiap kelas yang ditunjukkan. Dalam kelasnya kami menyebut sesuai dengan abjad. A adalah Angklung, B adalah Bonang, C adalah Demung dan seterusnya hingga H adalah Suling. Precision sama dengan TP dibagi TP ditambah FP. Poin nomor satu kita sudah mendapatkan nilai TP Nah selanjutnya adalah kita harus mencari nilai FP (false positive). Pada tahap kali ini, kita akan mencari nilai FP dari masing-masing kelas. perhatikan warna tulisan dibawah ini dan kolom ditabel atas sehingga menghasilkan tabel di bawah ini.

Tabel 4.6.2.1. Hasil dari pencarian FP (False Positive) dan Perhitungan

	Pressicion
FP(A)	0 $P(A) = 141/(141+0) = 1$
FP(B)	0 $P(B) = 141/(141+0) = 1$
FP [C]	10 $P[C] = 141/(141+10) = 0,93377483$
FP (D)	2 $P(D) = 141/(141+2) = 0,98601399$
FP (E)	0 $P[E] = 141/(141+0) = 1$
FP (F)	3 $P[F] = 141/(141+3) = 0,97916667$
FP (G)	3 $P[G] = 141/(141+3) = 0,97916667$
FP (H)	1 $P(H) = 141/(141+1) = 0,99295775$

Total precision adalah jumlah dari hasil precision masing-masing kelas di bagi jumlah kelas. Maka hasilnya adalah 7,871 dibagi dengan 8 adalah 0,9838.

#### 4.6.3. Menghitung Recall

Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Berikut adalah hasil tabel nya untuk menunjukkan hitungan hasil recall

Tabel 4.6.3 Klasifikasi False Negative

Class/ Sample	Angklung	Bonang	Demung	Gong	Kendang	Rebab	Saron	Suling	Jumlah
A Angklung	14	0	0	2	0	0	0	1	20
B Bonang	0	20	0	0	0	0	0	0	20
C Demung	0	0	17	0	0	0	0	0	20
D Gong	0	0	0	20	0	0	0	0	20
E Kendang	0	0	0	0	20	0	0	0	20
F Rebab	0	0	0	0	0	20	0	0	20
G Saron	0	0	10	0	0	0	10	0	20
H Suling	0	0	0	0	0	0	0	20	20
Jumlah	14	20	27	22	20	20	10	21	160

Dengan warna yang sesuai dengan masing-masing kelas klasifikasi alat musik gamelan, kita mendapatkan hasil FN (False negative), yang mana akan kita gunakan untuk menghitung recall. Karena Recall adalah TP (True Positive) dibagi dengan FN. Hasilnya seperti tabel di bawah ini. Hasil TP adalah jumlah dari akurasi sebesar 141.

Tabel 4.6.3.1 Perhitungan Recall dan All Recall

FN[A] =	6	R[A]	23.5
FN[B] =	0	R[B]	0.00
FN[C] =	3	R[C]	47
FN[D] =	0	R[D]	0.00
FN[E] =	0	R[E]	0.00
FN[F] =	0	R[F]	0.00
FN[G] =	10	R[G]	14.1
FN[H] =	0	R[H]	0
All Recall =			10.575

Hasil dari perhitungan menunjukkan all recall sebesar 10,575. Recall A menunjukkan 23,5, Recall B 0, Recall C sebesar 47, Recall D,E,F sebesar 0. Recall G sebesar 14,1 dan recall H sebesar 0.

## BAB V

### PENUTUP

#### 5.1. Kesimpulan

Berdasarkan hasil analisis dan pembahasan pada bab sebelumnya, penulis dapat menarik beberapa kesimpulan sebagai berikut :

1. Implementasi dari instrument alat musik gamelan yaitu bonang, gambang dan kendang dianalisis menggunakan metode *convolutional neural network* (CNN) dengan *library keras dan librosa*. Dalam proses analisis menggunakan CNN melalui proses *preprocessing audio* dan selanjutnya dilakukan permodel hingga didapatkan permodelan yang baik dalam data citra latih dan uji. Fitur Mel Frequency Cepstral Coeffisients (MFCC) berhasil mengeksplorasi untuk pengenalan suara alat musik gamelan jawa dengan pelatihan data menggunakan audio berformat .wav yang diklasifikasikan dengan menggunakan CNN (*convolutional neural network*) di mana sinyal suara dianalisis dan diekstrasi sehingga akan menentukan hasil klasifikasinya dengan akurasi yang diperoleh.
2. Berdasarkan model yang diterapkan, hasil validasi data latih memiliki nilai *loss* 0,0749 dan nilai sebesar 0,9832 yang berasal dari hasil training model dari *epoch* sebanyak 100. Pada tabel prediksi dapat dilihat bahwa suara alat musik gamelan dari data asli ke audio lain adalah mendekati angka yang sama. Nilai akurasi sebesar 0,88125 dapat dipastikan sesuai dengan klasifikasi berdasarkan hitungan akurasi adalah True Positive dibagi dengan total data yang diuji.

Hasil dari pengujian data testing yang diambil bukan dari dataset sebanyak 20 kali pengujian audio lain menunjukkan rata-rata *confidence level* pengujian yaitu pada Angklung sebanyak 0,7301; Bonang sebesar 0,9838; Demung sebesar 0,5725; Sedangkan Gong, Kendang, Rebab, Saron dan Suling masing-masing sebesar 0,9999; 0,9608; 0,9963; 0,5253; 0,9997. Jumlah *confidence level* dari masing-masing alat musik gamelan adalah pada Angklung sebesar 14,6028; Bonang sebesar 19,6774; Demung sebesar 11,4514; Gong diperoleh sebesar 19,9984, Kendang, Rebab, Saron, dan Suling masing masing berjumlah 19,3417; 19,9266; 10,5076; 19,9951.

3. Hasil presisi dari masing masing klasifikasi alat musik gamelan adalah 1;1; 0,9337; 0,9860; 1; 0,9791; 0,9791; 0,9929. Sehingga jumlah presisinya 7,871. Sedangkan total presisi dibagi dengan jumlah kelas adalah 0,9838. Hasil dari perhitungan menunjukkan all recall sebesar 10,575. Recall A menunjukkan 23,5, Recall B 0, Recall C sebesar 47, Recall D,E,F sebesar 0. Recall G sebesar 14,1 dan recall H sebesar 0.
4. *Convolutional neural network* dapat diimplementasikan pada instrument alat musik gamelan yaitu angklung, bonang, demung, gong, kendang, rebab, saron, suling dengan uji coba data baru memiliki akurasi yang seluruhnya benar dan tepat. Pada hasil kelas prediksi menunjukan akurasi dari hasil data *test* seluruhnya benar. Dengan hasil tersebut maka dapat dikatakan bahwa penggunaan metode *convolutional neural network* relevan diimplementasi terhadap instrument alat musik Gamelan jawa yaitu angklung, bonang, demung, gong, kendang, rebab, saron, suling.



## 5.2. Saran

Berdasarkan hasil analisis dan kesimpulan, dapat diberikan saran sebagai berikut:

1. Perlu diketahui bahwa secara subjektif oleh pendengaran manusia, suara alat musik gamelan Jawa demung, saron, bonang memiliki ciri khas yang sama. Maka akan lebih baik lagi jika dataset dikumpulkan lebih banyak lagi dengan berbagai macam *bit* dan *sample rate*, sehingga suara yang akan dijadikan validasi tidak rancu. Audio yang baik juga akan menghasilkan data yang baik juga. Karena dari penelitian audio *bit* nya sama dan *sample rate* nya yaitu 16 bit dan 11-22kHz, ini kurang memuaskan walaupun hasilnya akurat. Dikarenakan library yang terbaca hanya dengan 16 bit. Sebaiknya bias sampai 48 bit dan 48Khz.
2. Pengujian audio yang bukan berasal dari dataset akan lebih baik jika isinya tidak hanya satu suara, melainkan dalam satu audio terdapat berbagai alat musik gamelan, sehingga akan terlihat dominan mana alat musiknya guna mengecek kebenaran dari alat musik gamelan Jawa.
3. Hasil dari klasifikasi suara alat musik ini sebaiknya dipublikasikan ke masyarakat luas agar budaya Indonesia semakin dikenal dengan baik oleh masyarakat dan akan mengenalkan budaya Indonesia ke mancanegara.
4. Membangun arsitektur *convolutional neural network* dengan perhitungan yang baik sehingga menghasilkan akurasi yang tinggi.

## DAFTAR PUSTAKA

### PUSTAKA BUKU

- Teguh, W., 2018, Dasar- Dasar Pemrograman Python untuk Machine learning dan Kecerdasan Buatan, Gava Media, Yogyakarta
- Kelleher, J., 2019, Deep Learning, The MIT Press, London

### PUSTAKA JURNAL

- Sujaini, H., (2019). Klasifikasi Citra Alat Musik Tradisional dengan Metode k-Nearest Neighbor, Random Forest, dan Support Vector Machine. JSINBIS (Jurnal Sistem Informasi Bisnis), 9(2), 00. 185-191.
- Ryamizard, A.F., Hidayat, B. and Saidah, S., (2018). Deteksi Nada Tunggal Alat Musik Kecapi Bugis Makassar Menggunakan Metode Mel Frequency Cepstral Coefficient (mfcc) Dan Klasifikasi K-nearest Neighbour (knn). eProceedings of Engineering, 5(3).
- Lionel, D., Adipranata, R. and Setyati, E., 2019. Klasifikasi Genre Musik Menggunakan Metode Deep Learning Convolutional Neural Network dan Mel-Spektrogram. Jurnal Infra, 7(1), pp.51-55.
- Afida, A.M., (2020). Klasifikasi Jenis Burung Berdasarkan Suara Menggunakan Algoritme Support Vector Machine. Mathunesa: Jurnal Ilmiah Matematika, 8(1).
- Nam, J., Choi, K., Lee, J., Chou, S.Y. and Yang, Y.H., (2018). Deep learning for audio-based music classification and tagging: Teaching computers to

- distinguish rock from bach. *IEEE signal processing magazine*, 36(1), pp.41-51.
- Mahardika, M.N., (2018). Implementasi MIR (Music Information Retrieval) pada Modul Genre Recognition dan Deep Learning Classification untuk Aplikasi Musicmoo (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- Lee, J., Park, J., Kim, K.L. and Nam, J., (2018). Samplecnn: End-to-end deep convolutional neural networks using very small filters for music classification. *Applied Sciences*, 8(1), p.150.
- Bayle, Y., Robine, M. and Hanna, P., (2019). SATIN: a persistent musical database for music information retrieval and a supporting deep learning experiment on song instrumental classification. *Multimedia Tools and Applications*, 78(3), pp.2703-2718.
- Farrokhmanesh, M. and Hamzeh, A., (2019). Music classification as a new approach for malware detection. *Journal of Computer Virology and Hacking Techniques*, 15(2), pp.77-96.
- Hershey, S., Chaudhuri, S., Ellis, D.P., Gemmeke, J.F., Jansen, A., Moore, R.C., Plakal, M., Platt, D., Saurous, R.A., Seybold, B. and Slaney, M., (2017), March. CNN architectures for large-scale audio classification. In 2017 IEEE international conference on acoustics, speech and signal processing (icassp) (pp. 131-135). IEEE.
- Ye, J., Kobayashi, T. and Murakawa, M., 2017. Urban sound event classification based on local and global features aggregation. *Applied Acoustics*, 117, pp.246-256.

Arafah, M.A., Efendi, R. and Sazaki, Y., (2018). *Klasifikasi Jenis Musik Menggunakan Support Vector Machine (SVM)* (Doctoral dissertation, Sriwijaya University).

