

Klasifikasi Alat Musik Tradisional dengan Metode Machine Learning dengan Librosa dan Tensorflow pada Python

Puja Anggeli¹, Suroso², M.Zakuan Agung³

Jurusan Teknik Elektro Program Studi Teknik Telekomunikasi,
Politeknik Negeri Sriwijaya Palembang, Indonesia
Jl. Sriwijaya Negara, Bukit Besar, Kec. Ilir Barat I, Kota Palembang,
Sumatera Selatan, (0711)353414

pujaaanggeli19@gmail.com, osorus11@gmail.com, mzakuanagung30115@gmail.com

Abstract

The development of artificial intelligence technology AI (Artificial Intelligence) has been widely applied in various fields of daily life. AI (Artificial Intelligence) is divided into several branches, one of which is Machine Learning. Machine Learning is developed based on statistics, mathematics and data mining so that machines can learn by analyzing data without needing to be reprogrammed. With the development of the music world, not many people and the current generation know about traditional music from their respective regions. Traditional musical instruments produce sound art that has its own characteristics and uniqueness which is passed down from generation to generation. Therefore, to simplify the process of recognizing each musical instrument, a system was created that can classify traditional musical instruments using machine learning. The methods used in this research are librosa and tensorflow, where tensorflow is used for numerical computing and large-scale machine learning projects that have the best performance in classifying. In this study using Python 3.6 as a programming language and using PyCharm as a Integrated Development Environment (IDE). From the results, the system accuracy as expected after being tested several times, namely 91%.

Keywords: machine learning, classification, librosa, tensorflow

Abstrak

Perkembangan teknologi kecerdasan buatan AI (Artificial Intelligence) sudah banyak diterapkan di berbagai bidang kehidupan sehari-hari. AI (Artificial Intelligence) terbagi menjadi beberapa cabang, salah satunya adalah Machine Learning. Machine Learning dikembangkan berdasarkan ilmu statistika, matematika dan data mining sehingga mesin dapat belajar dengan menganalisa data tanpa perlu di program ulang. Dengan berkembangnya dunia musik, tidak banyak masyarakat dan generasi sekarang yang mengetahui tentang musik tradisional dari daerahnya masing-masing. Alat musik tradisional mengeluarkan seni suara yang memiliki khas dan keunikannya tersendiri yang di wariskan secara turun temurun. Oleh karena itu, untuk mempermudah proses dalam pengenalan setiap alat musik, dibuatlah sebuah model yang dapat mengklasifikasi alat musik tradisional menggunakan machine learning. Metode yang digunakan pada penelitian ini adalah librosa dan tensorflow, dimana tensorflow digunakan untuk komputasi numerik dan project machine learning berskala besar yang memiliki kinerja terbaik dalam mengklasifikasi. Pada penelitian ini menggunakan Python 3.6 sebagai bahasa pemrograman dan menggunakan PyCharm sebagai Integrated Development Environment (IDE). Dari hasil penelitian, di dapatkan hasil tingkat akurasi sesuai yang diharapkan setelah di lakukan nya uji coba beberapa kali yaitu 91%.

Kata kunci: machine learning, klasifikasi, librosa, tensorflow

1. PENDAHULUAN

Indonesia memiliki beragam jenis alat musik tradisional, *The United Nations Educational, Scientific, and Cultural Organization* (UNESCO) menyebutkan bahwa Indonesia sebagai negara superpower dalam bidang budaya. Indonesia adalah negara yang penuh akan kebudayaannya yang beragam dan setiap alat musik tradisional memiliki ciri khas dan keunikannya tersendiri. Seiring berkembangnya dunia musik, tidak sedikit masyarakat yang melupakan ciri khas kebudayaan dari daerahnya masing-masing.

Artificial Intelligence (AI) sudah banyak diterapkan di berbagai bidang kehidupan sehari-hari. *Artificial Intelligence* (AI) terbagi menjadi beberapa cabang, salah satunya adalah *Machine Learning*. *Machine Learning* dikembangkan berdasarkan ilmu statistika, matematika dan data mining sehingga mesin dapat belajar dengan menganalisa data tanpa perlu di program ulang. *Machine learning* merupakan cabang dari *Artificial Intelligence* (AI) dan ilmu komputer yang berfokus pada penggunaan data dan algoritma untuk meniru cara manusia belajar, secara bertahap meningkatkan akurasi. [1]

Pada penelitian ini menggunakan metode *Librosa* yang merupakan sebuah fitur pada *python* untuk menganalisis musik dan audio, sedangkan *Tensorflow* adalah platform open source end-to-end untuk *Machine Learning*. Dimana *Tensorflow* memiliki ekosistem yang luas, fleksibel dari alat, perpustakaan dan sumber daya komunitas yang mendorong state-of-the-art dalam *machine learning* dan pengembang mudah membangun dan menyebarkan aplikasi *machine learning* [2]. *Tensorflow* merupakan pustaka matematika simbolis berdasarkan dataflow dan pemrograman yang diciptakan oleh tim Google Brain yang digunakan untuk komputasi numerik dan project *machine learning* berskala besar, sehingga *Tensorflow* dapat memudahkan dalam proses memperoleh data, melatih model, prediksi dan menyempurnakan hasil yang akan di peroleh [3]. Penelitian ini menggunakan *Python* sebagai bahasa pemrograman dan *PyCharm* sebagai *Integrated Development Environment* (IDE). Penggunaan *PyCharm* sebagai IDE (*Integrated Development Environment*) karena *Pycharm* memiliki kemampuan untuk bekerja dengan beberapa database secara langsung tanpa mengintegrasikannya dengan alat lain, dapat memfasilitasi pengembangan kode yang lebih cepat, dan proses instalasi yang mudah.

Penelitian ini bertujuan untuk membantu seseorang yang kemungkinan merasa kesulitan dalam membedakan alat musik tradisional melalui suara, sehingga dibuat sebuah model atau sistem yang dapat mengklasifikasi alat musik tradisional dengan *machine learning* yang menggunakan metode *Librosa* dan *Tensorflow* pada bahasa pemrograman *Python*, agar dapat mempermudah penggunaannya untuk mempelajari atau mengetahui pustaka dari sebuah alat musik tradisional yang telah dipelajari oleh *machine learning*. Penggunaan metode *Librosa* dan *Tensorflow* dipilih dengan pemikiran bahwa metode yang akan digunakan dapat menghasilkan tingkat

akurasi atau pengenalan dalam proses mengklasifikasi dan memiliki kualitas yang lebih baik dalam proses keakuratan yang lebih tinggi.

2. METODOLOGI PENELITIAN

Tahapan penelitian diawali dengan pengumpulan data set berupa file audio berbentuk .wav hingga hasil berupa teks hasil prediksi.

2.1. Pengumpulan Data

Pada tahap pengumpulan data, data berupa audio alat musik tradisional yang akan di uji coba terbagi menjadi tiga, yaitu Angklung, Gamelan, dan Sasando. Jumlah masing-masing data yang akan di uji coba sebanyak 50 data untuk Angklung, 50 data untuk Gamelan, dan 50 data untuk Sasando. Sehingga total data yang akan di training sebanyak 150 data, dimana format file berupa .wav agar saat di uji coba dapat di proses oleh sistem dengan lebih baik.

2.2. Implementasi Sistem

Bahasa pemrograman yang digunakan adalah bahasa pemrograman *Python 3.6*, menggunakan *Pycharm* sebagai Integrated Development Environment (IDE), dan *Anaconda* sebagai *packet manager* yang dapat mengelola semua paket atau *module library* yang akan digunakan. Adapun beberapa library yang mendukung sistem ini, yaitu *glob*, *os*, *librosa*, *numpy*, *tensorflow*, *sklearn*, *matplotlib*, *scipy*, *StandardScaler*.

2.3. Ekstrasi Fitur pada Audio

Tahapan ekstrasi fitur pada audio menggunakan *librosa* untuk mendapatkan amplitudo gelombang pada grafik. Data dari setiap jenis alat musik yang sudah berbentuk file .wav akan di ekstrasi fiturnya dengan mencari *mfccs*, *chroma*, *melspectrogram*, *spectral contrast*, dan *tonnetz*.

Mel Frequency Cepstral Coefficients (MFCCs) adalah fitur yang digunakan untuk pengenalan suara dan pembicara secara otomatis. MFCCs diperkenalkan oleh Davis dan Mermelstein pada 1980-an. Skala Mel menghubungkan frekuensi yang nada dari nada murni dengan frekuensi terukur yang sebenarnya [4]. MFCC telah menjadi teknik yang sangat awam dan efisien untuk signal processing [8]. Kode untuk mfcc, $M : np.ndarray[\text{shape}=(n_mfcc,t)]$ MFCC *sequence*. Rumus untuk mengubah frekuensi ke skala Mel :

$$M(f) = 1125 \ln(1+f/700) \quad (1)$$

Chroma pada *librosa*, menghitung *chromagram* dari bentuk gelombang atau *power spectrogram*. Chromagram : $np.ndarray[\text{shape}=(n_chroma,t)]$, energi yang dinormalisasi untuk setiap *chroma* di setiap frame [5]. Spektrogram dapat dibentuk dengan menggunakan *Fourier Transform* Spektrogram didefinisikan sebagai besar kuadrat dari STFT, memberikan kekuatan suara untuk frekuensi dan waktu tertentu dalam dimensi ketiga [9].

Melspectrogram adalah spektrogram yang memvisualisasikan suara pada skala Mel sebagai lawan dari domain frekuensi. Skala Mel adalah transformasi logaritmik dari frekuensi sinyal. Skala Mel mendasar dalam aplikasi Machine Learning ke audio, karena memicu persepsi suara. Transformasi dari skala Hertz ke Skala Mel adalah sebagai berikut:

$$m = 1127 \cdot \log \left(1 + \frac{f}{700} \right) \quad (2)$$

Jika input spektrogram S diberikan, maka input tersebut dipetakan langsung ke basis Mel oleh $\text{mel_f.dot}(S)$. Jika input deret waktu y, sr diberikan, maka spektrogram besarnya S dihitung terlebih dahulu, dan kemudian dipetakan ke skala mel dengan $\text{mel_f.dot}(S^{**\text{power}})$. Secara default, $\text{power}=2$ beroperasi pada spektrum daya. Kode untuk *Melspectrogram* S : `np.ndarray [shape=(n_mels, t)]`.

Pada *Spectral Contrast*, setiap frame dari spektrogram dibagi menjadi sub-band. Untuk setiap sub-pita, kontras energi diperkirakan dengan membandingkan energi rata-rata di kuantil atas (*peak energy*) dengan energi kuantil bawah (*valley energy*). Nilai kontras yang tinggi umumnya sesuai dengan sinyal pita sempit yang jelas, sedangkan nilai kontras rendah sesuai dengan noise pita lebar. `contrast : np.ndarray [shape=(n_bands + 1, t)]`.

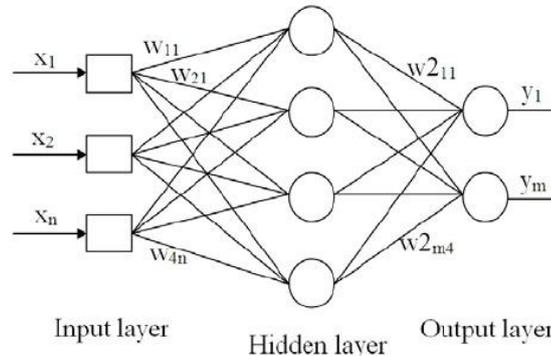
Tonnetz, representasi ini menggunakan metode 1 untuk memproyeksi fitur *chroma* ke basin 6 dimensi yang mewakili kelima sempurna, sepertiga minor, dan sepertiga mayor masing-masing sebagai koordinat dua dimensi. Kode untuk *tonnetz* adalah `tonnetz : np.ndarray [shape(6, t)]`. *Tonnetz dimensions* : 0: Fifth x-axis, 1: Fifth y-axis, 2: Minor x-axis, 3: Minor y-axis, 4: Major x-axis, 5: Major y-axis [5].

2.4. Klasifikasi Data

Tahapan klasifikasi, klasifikasi adalah pengelompokkan yang sistematis daripada sejumlah obyek, gagasan, buku atau benda-benda lain ke dalam kelas atau golongan tertentu berdasarkan ciri-ciri yang sama [7]. Klasifikasi ini menggunakan teknik *Supervised*, dimana *Supervised* merupakan salah satu teknik *machine learning* yang menggunakan dataset (data training) yang sudah berlabel (labeled data) untuk melakukan pembelajaran pada mesin, sehingga mesin mampu mengidentifikasi label input dengan menggunakan fitur yang dimiliki selanjutnya melakukan klasifikasi [10].

Pada tahap ini akan dilakukan training data yang telah di ekstrasi fitur oleh *librosa* kemudian *tensorflow* akan bekerja untuk melatih dan menjalankan neural network untuk mengklasifikasi data yang telah di ekstrasi fitur nya. Klasifikasi pada *Tensorflow* penelitian ini menggunakan *hidden layer* sebanyak n , dimana *hidden layer* adalah jaringan syaraf tiruan yang merupakan lapisan di antara *input layers* dan *output layers*. Neuron buatan mengambil satu set input berbobot dan menghasilkan output melalui

fungsi aktivasi. Ini adalah bagian dari hampir dan saraf di mana peneliti mensimulasikan jenis aktivitas yang terjadi di otak manusia [6].



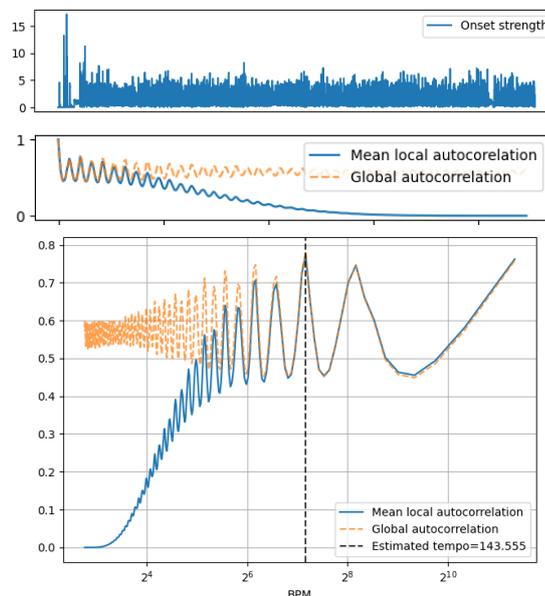
Gambar 1. Visualisasi Hidden Layer pada Tensorflow⁶

3. HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai proses implementasi sistem yang telah di uji, untuk mengetahui tingkat keberhasilan dari keseluruhan program yang telah di buat.

3.1. Implementasi Sistem

Data yang akan di uji coba terdiri atas, Angklung, Gamelan, dan Sasando dalam bentuk folder. Kemudian, setiap folder masing-masing mempunyai 50 data audio digital yang telah diubah dari .mp3 menjadi .wav . Maka, jumlah data secara keseluruhan sebanyak 150 data audio digital yang akan di uji coba.



Gambar 2. Visualisasi Tempogram dari 1 audio

Tahap selanjutnya ialah membuat project pada *Pycharm* dengan menggunakan *Python 3.6* dan *Conda Environment* sebagai *module library*.

Pada penelitian ini menggunakan *PyCharm Community Edition 2021.1.3* dan *Anaconda3 2021.05*.

Setelah melakukan tahap peng-kodean pada *Pycharm*, maka tahap selanjutnya meng-install *module library*. *Module library* yang diperlukan seperti *librosa*, *tensorflow*, *numpy*, *sklearn*, dan yang lainnya. Saat mengunduh pustaka ada beberapa *module* yang tidak dapat di unduh atau gagal di baca oleh sistem seperti *librosa* dan *tensorflow* dikarenakan versi dari *module* atau saat mengunduh pustaka program nya tidak mendukung. Sehingga, pada penelitian ini pengujian menggunakan *Tensorflow* versi 1.15 dan *Librosa* versi 0.8.1.

3.2. Ekstrasi Fitur dan Klasifikasi Data

Tahap ekstrasi fitur menggunakan *librosa* untuk menganalisis audio. Pada tahap ekstrasi fitur untuk mengetahui *mfcc*, *chroma*, *melspectrogram*, *spectral contrast*, dan *tonnetz* pengujian menggunakan *Array* agar dapat mengumpulkan variabel yang memiliki tipe data yang sama dan menyimpan beberapa nilai dalam satu variabel.

Setelah melakukan ekstrasi fitur pada setiap data, sistem akan mengeluarkan output berupa file *features.npy* dan *labels.npy* yang artinya data telah berhasil di ekstrasi fitur. Selanjutnya menggunakan *StandardScaler* untuk menormalisasikan fitur-fitur dari setiap variabel yang telah di ekstrasi menjadi *fit_params.npy* yaitu data parameter, data tersebut dapat membantu dalam test data atau klasifikasi data tersebut pada *Tensorflow*.

Pada klasifikasi dengan *Tensorflow*, pengujian melakukan *training epochs* sebanyak 205 *epoch* dan learning rate 0.01. Setelah dilakukan training untuk klasifikasi data dengan *Tensorflow* di dapat hasil terlihat dari tabel sebagai berikut.

Tabel 1. Hasil Klasifikasi Data

Dataset	Jumlah Testing	F-Score	Precision	Recall
Angklung	50	0.92856673	0.91494931	0.94259562
Gamelan	50	0.90442637	0.90885458	0.90004111
Sasando	50	0.91095362	0.92455625	0.89774545

Tabel 2. Hasil Training Epoch

Jumlah Training	Epoch	Cost
150	0	0.91715527
150	100	0.34633723
150	200	0.24289356
150	205	0.2328656

Dari hasil pengujian yang tertera pada tabel dapat dilihat bahwa hasil klasifikasi yang telah di uji coba dengan melakukan *epochs* sebanyak 205 menghasilkan tingkat akurasi yang optimal dengan *cost* atau error yang persentasenya sedikit, dari hasil tersebut dapat dilihat bahwa pada data Angklung memiliki tingkat akurasi yang cukup tinggi.

Tabel 3. Hasil Tingkat Akurasi

Dataset	Jumlah Testing	Akurasi
Angklung	50	91%
Gamelan	50	90%
Sasando	50	92%
Jumlah Testing	150	
Rata-rata Akurasi	91%	

Pada Tabel 6 dapat dilihat bahwa rata-rata dari hasil tingkat akurasi pada klasifikasi data cukup tinggi, dengan Angklung memiliki akurasi 91% , Gamelan memiliki akurasi 90% , dan Sasando memiliki tingkat akurasi yang sangat tinggi yaitu 92%. Sehingga, rata-rata akurasi pada pengujian ini dengan klasifikasi data sebanyak 150 data adalah 91%.

4. SIMPULAN

Berdasarkan hasil dari penelitian serta evaluasi dari klasifikasi menggunakan metode *machine learning* dengan *Librosa* dan *Tensorflow* dapat disimpulkan bahwa :

- Pada pembuatan model klasifikasi alat musik tradisional menggunakan metode *machine learning* dengan *Librosa* dan *Tensorflow* pada pemrograman *Python* dapat berjalan dengan sesuai yang diharapkan.
- Tingkat akurasi dari hasil testing klasifikasi data menggunakan *tensorflow* sangat tinggi pada 205 *epoch* dengan 0.23 *cost*, yaitu dengan rata-rata tingkat akurasi sebesar 91%
- Untuk mendapatkan hasil yang lebih baik, pada saat melakukan training pada data sangat bergantung pada lama waktu training dan banyak data yang di uji coba. Semakin banyak *epoch* yang dilakukan, maka semakin baik tingkat akurasinya.
- Pada saat melakukan pengujian pada program tersebut membutuhkan spesifikasi laptop atau komputer yang cukup dan baik seperti *random access memory* yang cukup besar agar program atau saat melakukan training data dapat dijalankan dengan baik.

DAFTAR PUSTAKA

- [1] IBM Cloud Education. 2020. Machine Learning, <https://www.ibm.com/cloud/learn/machine-learning>.
- [2] Tensorflow, <https://www.tensorflow.org/>.
- [3] GitHub, Inc. 2021. Tensorflow, <https://github.com/tensorflow/tensorflow>.
- [4] Mel Frequency Cepstral Coefficient (MFCC) tutorial, <http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/>.
- [5] Librosa Development Team. Librosa. 2013-2021, <https://librosa.org/doc/main/index.html>.

- [6] Hidden Layer Perceptron In Tensorflow, <https://www.javapoint.com/hidden-layer-perceptron-in-tensorflow>.
- [7] Arik Dian Pratama, Hozairi, & Tony Yulianto. Klasifikasi Musik Berdasarkan Genre Menggunakan Jaringan Syaraf Tiruan. Vol 2, No 1. Oktober 2016.
- [8] Danny Lionel, Rudy Adipranata, & Endang Setyati. Klasifikasi Genre Musik Menggunakan Metode Deep Learning Convolutional Neural Network dan Mel-Spektrogram. Vol 7, No 1. 2019.
- [9] Huzaifah, M. 2017. Comparison of Time-Frequency Resperesentations for Environmental Sound Classification using Convolutional Neural Network. arXiv: 1706.07156.
- [10] Endang Retnoningsih, Rully Pramudita. Mengenal *Machine Learning* Dengan Teknik *Supervised dan Unsupervised Learning* Menggunakan *Python*. Vol 7, No.2, Desember 2020. 156-165