**KLASIFIKASI PADA MIXED TYPE AUDIO MENGGUNAKAN DECISION TREE**

**Disusun untuk Sidang Proposal Skripsi**

**Semester Genap Tahun Akademik 2015/2016**

**OLEH:**

HILMI ABDUL AZIZ

10112810



**Program Studi Teknik Informatika**

**Fakultas Teknik Dan Ilmu Komputer**

**Universitas Komputer Indonesia**

**2016**

**HALAMAN PENGAJUAN PEMBIMBING**

**Klasifikasi Pada Mixed Type Audio Menggunakan**

**Decision Tree**

NIM: 10112810

Nama: Hilmi Abdul Aziz



Pembimbing Usulan:

1. Galih Hermawan, S.Kom., M.T.
2. Nelly Indriani W. , S.Si., M.T

# DAFTAR ISI

[DAFTAR ISI ii](#_Toc443246675)

[DAFTAR TABEL iii](#_Toc443246676)

[DAFTAR GAMBAR iv](#_Toc443246677)

[1. Latar Belakang Masalah 1](#_Toc443246678)

[2. Identifikasi Masalah 2](#_Toc443246679)

[3. Maksud dan Tujuan 2](#_Toc443246680)

[4. Batasan Masalah 2](#_Toc443246681)

[5. Metodologi Penelitian 3](#_Toc443246682)

[5.1 Metode Pengumpulan Data 3](#_Toc443246683)

[5.2 Metode Pembangunan Perangkat Lunak 3](#_Toc443246684)

[6. Deskripsi Umum Sistem 5](#_Toc443246685)

[7. Review Literatur 5](#_Toc443246686)

[8. Jadwal Penelitian 6](#_Toc443246687)

[9. Sistematika Penulisan 7](#_Toc443246688)

[10. Daftar Pustaka 9](#_Toc443246689)

# DAFTAR TABEL

[Tabel 8.1 Tabel Jadwal Kegiatan 6](#_Toc443246690)

# DAFTAR GAMBAR

[Gambar 5‑1 Skema Waterfall 4](file:///D:\Kuliah\SMT%208\Proposal%20Skripsi.docx#_Toc443246691)

[Gambar 6.1 Langkah Pengelompokan jenis Audio 5](#_Toc443246692)

# Latar Belakang Masalah

Audio merupakan media yang banyak digunakan pada saat ini dalam berbagai aspek teknologi, dengan tujuan dan cara penyampaian yang berbeda-beda. Dengan tujuan yang berbeda pula maka munculah berbagai tipe audio yang digunakan, misalnya musik sebagai sarana hiburan, *Audio Book*, rekaman pembicaraan/ceramah dengan ataupun tanpa tambahan suara latar, hingga audio yang berisi suara lingkungan/alam yang disebut pula sebagai *Mixed Type Audio*.

Banyakanya jenis audio pun melahirkan suatu pendekatan untuk dapat mengklasifikasikan jenis audio berdasarkan isi dari audio tersebut. Untuk melakukan pengklasifikasian maka dibutuhkan berkas audio yang diolah sehingga dapat diambil kesimpulan apakah jenis audio tersebut. Salah satu cara yang dilakukan adalah *pre-processing* yaitu pemrosesan awal terhadap *Mixed Type Audio* yang akan diklasifikasi jenisnya.

Tahap pre-processing pada setiap pemrosesan speech/musik termasuk didalamnya pemisahan suara vokal/latar, pengenalan pembicara tidak bisa dihindari. Pentingnya tahap ini adalah berdasar pada kebutuhan untuk menentukan apakah frame dari sinyal yang diterima merupakan bagian dari suatu kelas tertentu, misalnya kelas *speech* saja, musik saja, atau gabungan kelas tesebut [[1](#PEJ13)]. Dengan tahap ini akan mengurangi beban perhitungan pada pengklasifikasian jenis audio. Oleh karena itu dibutuhkan sebuah metode untuk mengenali jenis dari audio pada tahap *pre-processing* ini. Teknik yang dapat digunakan ialah teknik klasifikasi.

Salah satu metode yang digunakan untuk klasifikasi adalah metode klasifikasi *Decision Tree*. *Decision Tree* adalah sebuah pohon dimana setiap node cabang merupakan pilihan antara sejumlah alternatif, dan setiap simpul daun merupakan keputusan. *Decision Tree* biasanya digunakan untuk mengklasifikasikan dokumen teks dengan akurasi yang lebih baik dibandingkan dengan metode *naïve Bayes* dan *KNN* [[2](#SMe13)].

Berdasarkan latar belakang diatas, maka penulis tertarik untuk melakukan penelitian dengan judul “Klasifikasi jenis *Mixed Type Audio* Menggunakan *Decision Tree*”.

# Identifikasi Masalah

Dalam tugas akhir ini, penulis akan merancang dan membangun sistem yang akan melakukan proses klasifikasi pada berkas *Mixed Type Audio*.

Rumusan masalah yang akan dibuat oleh sistem secara garis besar adalah sebagai berikut:

1. Bagaimana menerapkan metode *Decision Tree* untuk klasifikasi jenis audio pada *Mixed Type Audio*
2. Berapa persen tingkat akurasi dari penerapan *Decision Tree* dari klasifikasi jenis audio pada *Mixed Type Audio*?

# Maksud dan Tujuan

Berdasarkan permasalahan yang diteliti, maka maksud dari penulisan tugas akhir ini adalah mengklasifikasikan jenis dari *Mixed Type Audio* dengan menggunakan metode *Decision Tree*.

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menunjukkan apakah *Decision Tree* dapat digunakan dalam mengklasifikasikan audio kedalam jenis *Mixed Type Audio*.
2. Mengukur tingkat akurasi klasifikasi audio yang dihasilkan oleh *Decision Tree*.

# Batasan Masalah

Dalam penulisan penelitian ini terdapat batasan-batasan, agar permasalahan tidak terlalu melebar adapun batasan-batasan yang telah di tentukan adalah sebagai berikut :

1. Metode Klasifikasi untuk membuat *Decision Tree* yang digunakan dalam penelitian ini adalah C5.0/See5.
2. Jenis Klasifikasi *Mixed Type Audio* terbagi 5 yaitu : Musik, *Speech*, *Environtment*, *Speech* dengan background Musik, atau *Speech* dengan background *Environtment*.

# Metodologi Penelitian

Metode yang digunakan dalam penulisan laporan ini menggunakan dua metode, yaitu metode pengumpulan data dan metode pembangunan perangkat lunak.

## Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam membantu penelitian ini menggunakan dua cara, yaitu dengan cara studi literatur dan dokumen.

1. Studi Literatur

Mengumpulkan literatur, jurnal, dan bacaan-bacaan yang berhubungan dengan judul penelitian.

1. Pengumpulan Data

(Obersevasi) Mengumpulkan training dan testing data yang diambil dari berkas audio berjenis *Mixed Type Audio*.

## Metode Pembangunan Perangkat Lunak

Pembangunan perangkat lunak ini menggunakan model klasik yang bersifat sistematis, berurutan dalam membangun software. Berikut ini tahapan dari model *Waterfall* (air terjun) menurut Pressman [[3](#Pre10)]:

1. *Communication*

Langkah ini merupakan analisis terhadap kebutuhan software, dan tahap untuk mengadakan pengumpulan data dengan melakukan pertemuan dengan ahli, maupun mengumpulkan data-data tambahan baik yang ada di jurnal, artikel, maupun dari internet.

1. *Planning*

Proses planning merupakan lanjutan dari proses communication (analysis requirement). Tahapan ini akan menghasilkan dokumen user requirement atau bisa dikatakan sebagai data yang berhubungan dengan keinginan user dalam pembuatan software, termasuk rencana yang akan dilakukan.

1. *Modeling*

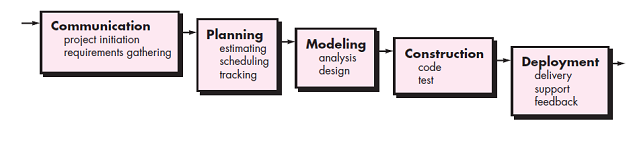
Proses modeling ini akan menerjemahkan syarat kebutuhan ke sebuah perancangan software yang dapat diperkirakan sebelum dibuat coding. Proses ini berfokus pada rancangan struktur data, arsitektur *software*, representasi *interface*, dan detail (algoritma) prosedural. Tahapan ini akan menghasilkan dokumen yang disebut *software requirement*.

1. *Construction*

Construction merupakan proses membuat kode. *Coding* atau pengkodean merupakan penerjemahan desain dalam bahasa yang bisa dikenali oleh komputer. *Programmer* akan menerjemahkan transaksi yang diminta oleh user. Tahapan inilah yang merupakan tahapan secara nyata dalam mengerjakan suatu *software*, artinya penggunaan komputer akan dimaksimalkan dalam tahapan ini. Setelah pengkodean selesai maka akan dilakukan testing terhadap sistem yang telah dibuat tadi. Tujuan testing adalah menemukan kesalahan-kesalahan terhadap sistem tersebut untuk kemudian bisa diperbaiki.

1. *Deployment*

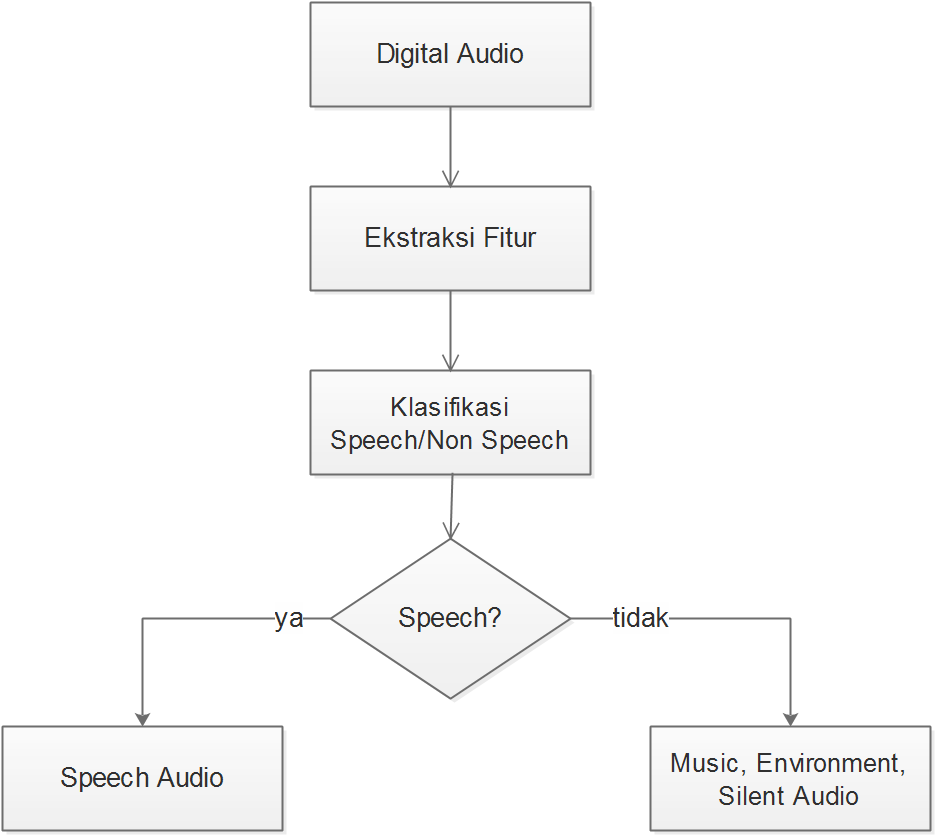
Tahapan ini bisa dikatakan final dalam pembuatan sebuah *software* atau sistem. Setelah melakukan analisis, desain dan pengkodean maka sistem yang sudah jadi akan digunakan oleh user. Kemudian *software* yang telah dibuat harus dilakukan pemeliharaan secara berkala.



Gambar 5‑1 Skema Waterfall

# Deskripsi Umum Sistem

Dalam pengklasifikasian dari *Mixed Type Audio* ini terdiri dari beberapa langkah yang dijelaskan dalam gambar berikut [[4](#Lie02)]:



Gambar 6.1 Langkah Pengelompokan jenis Audio

Dari gambar diatas dapat dilihat proses dalam klasifikasi *Mixed Audio Type. Audio* yang diproses akan di ekstraksi terlebih dahulu fiturnya agar dapat diketahui ciri-ciri dan karakteristiknya.

Setelah itu dengan metode klasifikasi *Decision Tree*, hasil ekstraksi akan dibandingkan dengan kriteria setiap kelas pada aplikasi. Kriteria setiap kelas adalah hasil dari data latih dari jenis audio yang bersangkutan. Perbandingan/Pencocokan antara kriteria kelas dan kriteria data masukan akan menghasilkan pengklasifikasian berdasarkan tingkat kemiripan tertinggi antara data audio masukan dan kriteria kelas.

# Review Literatur

1. Mixed Type Audio

Mixed Type Audio adalah sebuah berkas audio yang memiliki lebih dari satu jeis suara, misalnya ialah speech yang digabung dengan musik, ataupun digabung dengan suara lingkungan sekitar [[5](#Lei06)]. Contoh jenis mixed speech adalah audio dari sebuah film.

1. Decision Tree dan Algoritma C5.0/See5

Decision Tree merupakan sebuah teknik klasifikasi data yang bekerja dengan membagi data manjadi bagian-bagian yang lebih kecil, dengan ujung daun/*leaf* adalah sebuah keputusan, setiap node adalah fitur/ciri yang diklasifikasikan dan setiap *branch* yang memiliki nilai.

Algoritma C5.0/See5 adalah salah satu algoritma dalam pembuatan *Decision Tree*, Algoritma ini merupakan pengembangan dari algoritma C4.5 dan ID3. Algoritma ini memiliki kelebihan dari pendahulunya, diantaranya penggunaan memori lebih sedikit, pohon keputusan yang lebih kecil, dan mendukung penghapusan atribut yang tidak terpakai [[6](#Rut15)].

Beberapa algoritma yang pernah digunakan dalam bidang klasifikasi audio diantaranya *Support Vector Machine* (SVM) dan *Decision Tree*. [[7](#Zho11)] Oleh karena itu maka diharapkan metode *Decision Tree* dapat menyelesaikan masalah pada penelitian ini

# Jadwal Penelitian

Jadwal kegiatan dibuat untuk mengetahui kegiatan yang dilakukan dengan merinci setiap tahap penyelesaian penelitian yang akan dilakukan pada waktu yang ditentukan serta secara tidak langsung memberikan kedisiplinan untuk menyelesaikan target yang telah disusun. Jadwal kegiatan penelitian sebagai berikut :

Tabel 8.1 Tabel Jadwal Kegiatan

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No | Jenis Kegiatan | Bulan | | | | |
| 1 | 2 | 3 | 4 | 5 |
| 1 | Studi Literatur |  |  |  |  |  |
| 2 | Pengumpulan Data |  |  |  |  |  |
| 3 | Communication |  |  |  |  |  |
| 4 | Planning |  |  |  |  |  |
| 5 | Modeling |  |  |  |  |  |
| 6 | Construction |  |  |  |  |  |
| 7 | Deployment |  |  |  |  |  |

# Sistematika Penulisan

Sistematika penulisan proposal penelitian ini untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini menguraikan latar belakang permasalahan, merumuskan inti permasalahan, mencari solusi atas masalah tersebut, menentukan maksud dan tujuan pembatasan masalah, dan sistematika penulisan dari penelitian mengenai pengklasifikasian jenis audio ini.

BAB 2 LANDASAN TEORI

Bab ini mengkaji dan menggunakan berbagai konsep dari teori-teori dari para ahli yang berkaitan dengan topik penelitian yaitu klasifikasi jenis audio. Antara lain *Mixed Type Audio*, Metode klasifikasi, Perangkat Lunak yang digunakan untuk membangun sistem.

BAB 3 ANALISIS DAN PERCANGAN SISTEM

Bab ini menganalisis masalah sentimen acara televisi di indoensia untuk mengetahui hal atau masalah apa yang timbul kemudian memecahkan permasalahan tersebut dengan menggunakan metode klasifikasi yang ada pada bab sebelumnya. Sedangkan perancangan adalah suatu tahap pada penelitian dimana perangkat yang digunakan ditentukan, mengidentifikasi data yang digunakan terhadap objek yang dirumuskan.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menjelaskan proses penerapan pada bab sebelumnya yang telah dipersiapkan untuk selanjutnya diimplementasikan menjadi perangkat lunak sehingga dapat mengklasifikasikan jenis audio.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi mengenai kesimpulan tentang keseluruhan dari penelitian dan saran untuk kemungkinan penelitian selanjutnya.

# Daftar Pustaka

x

|  |  |
| --- | --- |
| [1] | Pejman Mowlaee and Abolghasem Sayadiyan, "Audio Classification Of Music/Speech Mixed Signals Using Sinusoidal Modeling With Svm And Neural Network Approach," *Journal of Circuits, Systems and Computers*, vol. 22, no. 02, February 2013. |
| [2] | Menaka S. and Radha R., *Text Classification using Keyword Extraction*. Coimbatore, India: Peelamedu, Desember 2013. |
| [3] | Pressman, *Software Engineering : A Practitioner’s Approach*, 7th ed.: McGrow Hill, 2010. |
| [4] | Lie Lu and Hong-Jiang Zhang, "Content Analysis for Audio Classification and Segmentation," *IEEE Transactions On Speech And Audio Processing*, vol. 10, no. 7, Oktober 2002. |
| [5] | Lei Chen, Sule Gunduz, and M. Tamer Ozsu, "Mixed Type Audio Classification With Support Vector Machine," *2006 IEEE International Conference on Multimedia and Expo*, 2006. |
| [6] | Rutvija Pandya and Jayati Pandya, "C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning," *International Journal of Computer Applications* , vol. 117, no. 16, 2015. |
| [7] | Zhong-Hua Fu, Wei Feng, and Yong Luo, "Pitch-density-based features and an SVM binary tree approach for multi-class audio classification in broadcast news," *Multimedia Systems*, vol. 17, no. 2, Maret 2011. |

x