



ENHANCED CONFIX STRIPPING STEMMER DAN FUZZY CLUSTERING PADA DATA PELAKU SENI DI MALANG RAYA

Bambang Nurdewanto¹⁾, Elta Sonalitha²⁾, Anis Zubair³⁾, Salnan Ratih Asriningtias⁴⁾

¹⁾Sistem Informasi, Fakultas Teknik, Universitas Merdeka Malang
Email: elta.sonalitha@unmer.ac.id

²⁾Teknik Elektro, Fakultas Teknologi Informasi, Universitas Merdeka Malang
Email: nurdewa@unmer.ac.id

³⁾ Sistem Informasi, Fakultas Teknologi Informasi, Universitas Merdeka Malang
Email: anis_zubair@yahoo.com

⁴⁾Teknologi Informasi, Pendidikan Vokasi, Universitas Brawijaya
Email: salnanratih@gmail.com

Abstrak

Undang-Undang Pemajuan Budaya baru saja dicanangkan dan diberlakukan di Indonesia tahun 2017. Perundang-undangan ini membutuhkan pendataan tentang segala hal yang berkaitan dengan budaya Indonesia, terutama budaya tradisional warisan leluhur bangsa Indonesia. Pendataan meliputi segala fakta di bidang seni dan budaya. Data tentang seni dan budaya Indonesia sangatlah besar. Maka diperlukan perancangan sistem klasifikasi otomatis yang dapat mengkategorisasikan jenis maupun ragam budaya. Data untuk penelitian ini menggunakan data Pelaku Seni dengan Taksonomi bahwa, satu pelaku seni dapat memainkan beberapa seni tradisional misalkan musik, gerak, tari, dan sastra. Klasifikasi yang ditentukan oleh pemerintah adalah metode Taksonomi. Taksonomi data dapat menggunakan beberapa metode. Penelitian ini menggunakan *Enhance Confix Stripping Stemmer* untuk menentukan kluster pada proses *Fuzzy Clustering*. Hasil penggunaan metode *enhanced confix stripping* mampu mencari kata dasar dari kata berimbuhan dengan tingkat akurasi 91,67% dan Hasil stemming menggunakan metode *enhanced confix stripping* mampu memberikan hasil *cluster* yang sangat baik yaitu memiliki rata-rata varian semua *cluster* (v) sebesar 0.0000291216758415.

Kata kunci: *Clustering, Enhance Confix Stripping Stemmer, Fuzzy Clustering, Seniman, Seni*

Abstract

The cultural advancement Law has just been promulgated and implemented in Indonesia in 2017. This law requires the collection of data on all matters relating to Indonesian culture, especially the traditional culture of Indonesian ancestral heritage. The data collection covers all facts in the arts and culture. Data about Indonesian art and culture is huge. Then we need to design an automatic classification system that can categorize the types and diversity of cultures. The data for this research uses the data of Artists with Taxonomy that, one artist can play some traditional arts for example music, motion, dance, and literature. The classification determined by the government is the Taxonomy method. Data taxonomy can use several methods. This research uses Enhance Confix Stripping Stemmer to determine clusters in the Fuzzy Clustering process. The results of using the enhanced confix stripping method are able to find the basic words of the words with an accuracy rate of 91.67% and the results of stemming using the enhanced



confix stripping method are able to provide excellent cluster results that have an average variant of all clusters (v) of 0.0000291216758415.

Keyword: *Clustering, Cultural, Enhance Confix Stripping Stemmer, Fuzzy Clustering, Heritage*

I. PENDAHULUAN

Pelaksanaan Undang-undang Pemajuan Budaya tahun 2017 di Indonesia, membutuhkan langkah-langkah teknis diantaranya adalah Pendataan. Pendataan harus dilakukan oleh semua elemen masyarakat, baik pemerintah maupun pelaku-pelaku seni. Budaya yang dimaksud dalam Undang-undang Pemajuan Budaya adalah segala kegiatan, hasil dan produk seni tradisional terutama yang diturunkan oleh nenek moyang bangsa Indonesia. Hal ini dilakukan untuk menjaga warisan budaya leluhur dari kepunahan.

Pendataan salah satunya adalah pendataan pelaku seni tradisional. Pengolahan data dasar sebagai fakta yang dapat digunakan sebagai acuan untuk langkah-langkah pemajuan budaya. Data Pelaku Seni merupakan salah satu data dasar yang sangat penting digunakan untuk referensi siapa saja pelaku seni dan kemahiran seni apa saja yang dipunyai.

Data seniman ini harus diklasifikasi agar mudah untuk diidentifikasi keahliannya. Pemerintah Indonesia sudah menentukan model klasifikasi untuk pendataan seni dan budaya yaitu menggunakan metode taksonomi. Klasifikasi dalam Pelaku Seni dan keahliannya menggunakan model Taksonomi, dikarenakan seorang pelaku seni dapat menguasai beberapa bidang seni, mungkin seni musik, gerak, dan suara. Pelaku Seni akan di *rating* berdasarkan bobot seberapa ahli di bidangnya.

Taksonomi yang dihasilkan akan mengambil kata dasar dari kalimat-kalimat yang diisikan oleh pelaku seni pada instrument yang menggali tentang keahlian pelaku seni. Dalam penelitian ini, Pengambilan kata dasar ini menggunakan

Enhanced Confix Stripping Stemmer yaitu metode mencari kata dasar dengan membuang imbuhan-imbuhan dalam ejaan bahasa Indonesia. Kata dasar yang dihasilkan dari proses ini adalah untuk menentukan klaster dalam *Fuzzy Clustering*. Penelitian ini akan menguji keakuratan dari *Enhanced Confix Stripping Stemmer* dalam menentukan klaster.

II. KAJIAN LITERATUR

A. Text Preprocessing

Text Processing merupakan sebuah fungsi untuk mengubah data text yang tidak terstruktur ke dalam bentuk data terstruktur. Beberapa tahap proses text preprocessing diantaranya sebagai berikut :

- *Case folding*

Menghilangkan karakter selain huruf dan merubah semua huruf menjadi huruf kecil

- *Tokenizing*

Memotong kalimat menjadi kata sejumlah kata penyusunannya

- *Filtering*

Mengambil kata-kata penting dari hasil *tokenizing* menggunakan teknik *stop list* yaitu membuang kata tidak penting

- *Stemming*

Mencari kata dasar dari hasil *filtering* dengan membuang imbuhan

B. Term Frequency Inverse Document Frequency (TFIDF)

Dalam pencarian informasi TF-IDF bertujuan untuk menentukan nilai atau bobot pada sebuah kata (*term*) berdasarkan tingkat kepentingannya dalam dokumen. TF merupakan frekuensi kemunculan sebuah *term* dalam dokumen. IDF merupakan frekuensi kemunculan sebuah *term* pada keseluruhan dokumen. Semakin sering *term*



muncul pada suatu dokumen maka nilai TF semakin besar dan nilai IDF semakin kecil. Persamaan untuk menghitung TF-IDF dapat dilihat dalam (1). Pada persamaan (1) D merupakan jumlah keseluruhan dokumen dan DF merupakan jumlah dokumen yang mengandung *term*.

$$TFIDF = TF \times IDF = TF \times \log_e \frac{|D|}{DF} \quad (1)$$

C. Enhanced Confix Stripping Stemmer

Confix Stripping Stemmer merupakan sebuah metode untuk proses *stemming* pada Bahasa Indonesia yang dibuat oleh Nazief dan Adriani. Metode ini dikembangkan berdasarkan aturan morfologi Bahasa Indonesia yang mengelompokkan imbuhan menjadi awalan (*prefix*), sisipan (*infix*), akhiran (*suffix*) dan gabungan awalan-akhiran (*confixes*). Metode ini menggunakan kamus kata dasar dan mendukung *recoding*, yaitu penyusunan kembali kata-kata yang mengalami proses *stemming* berlebih. Aturan morfologi Bahasa Indonesia mengelompokkan beberapa kategori sebagai berikut :

- *Inflection suffix* yaitu kelompok akhiran yang merubah bentuk kata dasar. Kelompok ini dibagi menjadi dua diantaranya *particle* (P) berupa “lah”, “kah”, “tah” dan “pun” dan *possessive pronoun* (PP) atau kata ganti kepunyaan berupa “-ku”, “-mu” dan “-nya”.
- *Derivation Suffix* (DS) yaitu sekumpulan akhiran asli yang secara langsung ditambahkan ke pada kata dasar berupa “-i”, “-kan” dan “-an”.
- *Derivation Prefixes* (DP) yaitu sekumpulan awalan yang langsung dapat diberikan pada kata dasar murni atau pada kata dasar yang sudah mendapatkan penambahan sampai 2 awalan. Contoh awalan yang dapat bermorfologi adalah “me”, “be”, “te” dan “pe”. Contoh awalan yang tidak

dapat bermorfologi adalah “di”, “ke” dan “se”.

Berdasarkan jenis imbuhan di atas, maka bentuk kata berimbuhan dalam Bahasa Indonesia dapat dimodelkan sebagai berikut :

[DP+[DP+[DP]]]+KataDasar+[[+DS][+PP][+P]]

Beberapa aturan pada proses *stemming* menggunakan metode *Confix Stripping* oleh Nazief dan Adriani adalah sebagai berikut :

- Tidak semua kombinasi awalan dan akhiran diperbolehkan diantaranya ‘be-i’, ‘di-an’, ‘ke-i’, ‘ke-kan’, ‘me-an’, ‘se-i’, ‘se-kan’, dan ‘te-an’
- Penggunaan imbuhan yang sama secara berulang tidak diperkenankan,
- Jika suatu kata hanya terdiri dari satu atau dua huruf tidak diperkenankan
- Penambahan suatu awalan tertentu dapat mengubah bentuk asli kata dasar sebagai contoh awalan “me-” dapat berubah menjadi “meng”.

Proses *stemming* metode *Confix Stripping* oleh Nazief dan Adriani adalah sebagai berikut :

1. Di awal proses *stemming* dan setiap langkah yang selanjutnya melakukan pengecekan kata di kamus kata dasar. Jika kata ditemukan maka proses *stemming* dihentikan,
2. Melakukan pengecekan *rule precedence* yaitu larangan kombinasi awalan dan akhiran. Jika kata mengandung *prefix-suffix* dari “be-lah”, “be-an”, “me-i”, “di-i”, “pe-i” atau “te-1” maka langkah berikutnya adalah (5,3,4,6) selain itu maka langkah berikutnya (3,4,5,6).
3. Menghilangkan *Inflection Particle* P (“lah”, “kah”, “tah”, “-pun”) jika terdapat dalam kata dan dilanjutkan dengan menghilangkan *possessive pronoun* PP (“-ku”, “-mu”, “-nya”)
4. Menghilangkan *Derivation Suffix* DS (“-I”, “-kan”, “-an”)



5. Menghilangkan *Derivation Prefixes* DP (“di-”, “ke-”, “se-”, “me-”, “be-”, “pe-”, “te-”) dengan maksimum iterasi sebanyak tiga kali :

a. Berhenti jika :

- Jika *prefix* merupakan bentuk larangan kombinasi awalan dan akhiran yang sudah dihapus pada tahap sebelumnya
- Jika *prefix* sama persis dengan *prefix* yang sudah dihapus pada tahap sebelumnya
- Tiga *prefix* sudah dihapus

b. Identifikasi jenis *prefixed an* menghilangkannya. Terdapat dua jenis *prefixes* diantaranya :

- *Plain*: *prefix* yang berupa “di”, “ke”, dan “se” yang dapat dihilangkan secara langsung
- *Complex*: *prefix* yang berupa “me”, “be”, “pe” dan “te” yang mungkin merubah kata dasar dapat menggunakan aturan pada Tabel 1,2,3 atau 4.

c. Langkah 5 diulangi jika pencarian di kamus dasar gagal untuk kata saat ini.

6. Jika semua langkah telah selesai tetapi tidak juga berhasil, maka kata awal diasumsikan sebagai kata dasar atau kata dikembalikan seperti semula yaitu sebelum dilakukan proses *stemming* dan proses diakhiri.

Algoritma ini masih memiliki kesalahan diantaranya :

- Kurangnya aturan pemenggalan awalan untuk kata-kata dengan format “mem+p...”, “mem+s...”, dan “peng+k...”
- Kurang relevannya pemenggalan awalan pada kata-kata dengan format “menge+kata dasar” dan “penge+kata dasar”
- Adanya elemen pada beberapa kata dasar yang menyerupai suatu imbuhan.

Untuk memperbaiki kesalahan tersebut, maka metode *Enhanced Confix Stripping* melakukan perbaikan kesalahan diantaranya sebagai berikut :

- Memodifikasi aturan pada Tabel 3 dan Tabel 4. Perubahan aturan dapat dilihat pada Tabel 5 dan Tabel 6.
- Menambahkan langkah untuk mengatasi kesalahan pemenggalan akhiran yang seharusnya tidak dilakukan.

Tabel 1. Aturan Menghilangkan *Prefix* “me-“

Aturan	Bentuk Kata	Pemenggalan
1	me{llr w y}V...	me-{ r w y}V...
2	mem{b f v}..	mem-{ b f v}..
3	mempe...	mem-pe...
4	mem{rV V}...	me-m{ rV V}...
5	men{c d j z}...	men-{ c d j z}...
6	menV...	me-nV... me-tV
7	meng{g h q k}	meng-{ g h q k}
8	mengV...	meng-V... meng-kV...
9	menyV...	meny-sV...
10	mempV...	mem-pV... dimana V!=‘e’

Tabel 2. Aturan Menghilangkan *Prefix* “pe-“

Aturan	Bentuk Kata	Pemenggalan
1	pe{w y}V...	pe-{ w y}V...
2	perV...	per-V... pe-rV...
3	perCAP...	per-CAP... dimana C!=‘r’
4	perCAerV...	per-CAerV... dimana C!=‘r’
5	pem{b f V}...	pem-{ b f V}...
6	pem{rV V}...	pe-m{ rV V}... pe-p{ rV V}...
7	pen{c d j z}...	pen-{ c d j z}...
8	penV...	pen-V... pe-tV...
9	peng{g h q}...	peng-{ g h q}...
10	pengV...	peng-V... peng-kV...
11	penyV...	peny-sV...
12	pelV...	pe-IV kecuali “pelajar” menjadi “ajar”
13	peCerV...	per-erV... dimana C!={ r w y l m n}
14	peCP...	pe-CP... dimana C!={ r w y l m n} dan P!=‘er’
15	peC ₁ erC ₂	pe-C ₁ erC ₂ ... dimana C ₁ !={ r w y l m n}

Tabel 3. Aturan Menghilangkan *Prefix* “be-“

Aturan	Bentuk Kata	Pemenggalan
1	berV...	ber-V... be-rV
2	berCAP...	ber-CAP... dimana C!=‘r’ dan P!=‘er’
3	berCAerV...	ber-CAerV... dimana C!=‘r’
4	belajar	bel-ajar
5	beC ₁ erC ₂ ...	Be-C ₁ erC ₂ ... dimana C ₁ !={ r l}

Tabel 4. Aturan Menghilangkan *Prefix* “te-“

Aturan	Bentuk Kata	Pemenggalan
1	terV...	ter-V... te-rV...
2	terCerV...	ter-CerV...dimana C!=‘r’



3	terCP...	ter-CP... dimana C!='r' dan P!='er'
4	terC ₁ erC ₂	ter-C ₁ erC ₂ ... dimana C ₁ !='r'

Tabel 5. Modifikasi Aturan Tabel 1

Aturan	Bentuk Kata	Pemenggalan
5	men{c d j s z}...	men-{c d j s z}...
8	mengV...	meng-V... meng-kV (mengV-... jika V='e')
10	mempA	mem-pA... dimana A!='e'

Tabel 6. Modifikasi Aturan Tabel 2

Aturan	Bentuk Kata	Pemenggalan
9	pengC...	peng-C...
10	pengV...	peng-V... peng-kV... (pengV-... jika V='e')

Keterangan Simbol Huruf :

C : huruf konsonan

V : huruf Vokal

A : huruf konsonan atau vocal

P : partikel atau fragmen dari suatu kata, misalnya "er"

D. Fuzzy Clustering

Fuzzy clustering yang merupakan salah satu teknik untuk menentukan kluster data yang mana keberadaan tiap-tiap titik data dalam suatu kluster ditentukan oleh derajat keanggotaan. Penentuan kluster data di dasarkan pada bentuk *Euclidean Distance* untuk mengukur jarak kedekatan antar data. Konsep dari *fuzzy clustering* adalah menentukan pusat kluster terlebih dahulu yang merupakan lokasi rata-rata untuk tiap pusat kluster. Pada kondisi awal nilai pusat cluster masih belum akurat, maka dilakukan perbaikan secara berulang pada pusat cluster dan derajat keanggotaan di tiap-tiap titik data sehingga menghasilkan pusat kluster yang bergerak ke posisi yang tepat. Perulangan ini didasarkan pada minimisasi fungsi objektif yang menggambarkan jarak dari titik data ke pusat cluster yang terbobot oleh derajat keanggotaan titik data tersebut. Hasil *output* berupa deretan pusat kluster dan beberapa derajat keanggotaan untuk tiap-tiap data. Algoritma dari *fuzzy clustering* adalah sebagai berikut :

1. Menentukan parameter dari *fuzzy clustering* diantaranya :

- Jumlah *cluster* (c)
- Pangkat (w)
- Maximum iterasi (*MaxIter*)
- Error* terkecil yang diharapkan (ϵ)

e. Fungsi objektif awal ($P_0=0$)

f. Iterasi awal (t=1)

- Membangkitkan bilangan acak (μ_{ik}) 0-1 (matrik partisi U) sejumlah n x m dimana n adalah jumlah data dan m adalah jumlah kluster
- Menentukan derajat keanggotaan untuk tiap-tiap data dan atribut (matrik partisi U)
 - Menghitung jumlah setiap kolom sesuai persamaan (2)

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (2)$$

- Menghitung derajat keanggotaan tiap-tiap kluster sesuai persamaan (3)

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (3)$$

dimana,

i = data ke-i

c= jumlah kluster.

- Menghitung pusat kluster sesuai persamaan (4)

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^2 * x_{ij})}{\sum_{i=1}^n (\mu_{ik})^2} \quad (4)$$

dimana,

n = jumlah data,

i = data ke-i,

k = kluster ke-k,

j = atribut ke-j

- Menghitung fungsi objektif pada iterasi ke-t, sesuai persamaan (5)

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (x_{ij} - V_{jk})^2 \right] (\mu_{ik})^w \right) \quad (5)$$

dimana,

n = jumlah data

c = jumlah kluster



m = jumlah atribut,
 X_{ij} = data ke-i atribut ke-j
 V_{kj} = pusat kluster ke k atribut ke j
 μ_{ik} = bobot data ke-i kluster ke-k
 w = pangkat

6. Menghitung perubahan matrik partisi sesuai persamaan (6)

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{w-1}}} \quad (6)$$

7. Cek kondisi berhenti, jika (|Pt - P(t-1)| < ϵ) atau I > *MaxIter* maka berhenti. Jika tidak maka t=t+1 ulangi langkah ke 4.

Evaluasi untuk pengujian *fuzzy clustering* dilakukan dengan menggunakan *cluster variance* sesuai persamaan (7)

$$v_c^2 = \frac{1}{n_c - 1} \sum_{i=1}^{n_c} (d_i - \bar{d}_i)^2 \quad (7)$$

Dimana :

v_c^2 = variance pada *cluster c*
 $c = 1 \dots k$, k adalah jumlah *cluster*
 n_c = jumlah data pada *cluster c*
 d_i = data ke i pada suatu *cluster*
 \bar{d}_i = rata-rata dari data pada suatu *cluster*

Terdapat dua jenis *cluster variance* yaitu :

- *Variance within cluster* (V_w)
 Varian ini digunakan untuk melihat hasil variansi penyebaran data yang ada dalam sebuah *cluster*. Semakin kecil nilai V_w maka semakin bagus *clusternya*. Persamaan untuk menghitung V_w dapat dilihat pada persamaan (8).

$$v_w = \frac{1}{N - k} \sum_{i=1}^k (n_i - 1)v_i^2 \quad (8)$$

Dimana,

N = jumlah semua data
 k = jumlah cluster
 n_i = jumlah data anggota dalam cluster i
 • *Variance between cluster* (V_b)

Varian ini digunakan untuk melihat hasil variansi penyebaran data antar *cluster*. Semakin besar nilai V_b maka semakin bagus

clusternya. Persamaan untuk menghitung V_b dapat dilihat pada persamaan (9)

$$v_b = \frac{1}{k - 1} \sum_{i=1}^k n_i (d_{ij} - \bar{d})^2 \quad (9)$$

Dimana,

k = jumlah *cluster*
 d_{ij} = data ke j pada suatu *cluster* ke i
 \bar{d} = rata-rata dari \bar{d}_i

Untuk melihat varian semua *cluster* maka dapat diukur dengan persamaan (10)

$$v = \frac{V_w}{V_b} \quad (10)$$

Jika nilai V semakin kecil maka nilai *cluster* semakin baik

III. METODE PENELITIAN

Pada penelitian ini, dilakukan proses klasifikasi pelaku seni berdasarkan data deskripsi pelaku seni yang diperoleh dari kuisioner. Terdapat 29 data yang dapat dilihat pada Tabel 7.

Tabel 7. Data Pelaku Seni

No	Nama	Deskripsi
1	Katijan	Dinas Pendidikan Prov. Jawa Timur#1980 Duta Seni Kota Batu#2014 Sapto Darmo#pelatih#1987 Karawitan Sapto Darmo, Margi Laras Jati, Larasati, SMP Arjuno, Wahyu Budaya Pelatih Duta Seni Kota Batu Aktif sebagai seniman dan pelatih di Kota Batu sampai di Jombok, Ngantang, Kab. Malang Wayang Wong;Ketoprak;Ludruk;Campursari Jaranan;Tayub Iringan Jaranan/Bantengan;Karawitan Gagrag Metaraman;Karawitan Gagrag Surakarta Pelatih;Niyaga;Pelaku Telaten dalam melatih,
2	Mirza Baihaqie	2016 2016 Cempoko Wahyu Budoyo#sekretaris#2017 Rogo wilis#sekretaris#2017 Rogo wilis, cempoko wahyu budaya APEKSI Kota Batu Pemuda pegiat seni tradisi Bantengan;Jaranan Kidal Gambuh/Pawang/Dukun Pelaku Gambuh dalam pertunjukan
3	Karin Mirda Yuwana	Kografer PSP dan FLS2N terbaik #Kota Batu#2015 Penyaji terbaik FKT#Jawa Timur#2011 Duta Tari Kota Batu ke Korea Selatan Duta Tari Kota Batu ke Malaysia Duta seni Indonesia ke Taipei, Australia, Beijing Sanggar Tari Maheswari Sanggar Taru Maheswari Duta Tari Kota Batu ke Korea Selatan Duta Tari Kota Batu ke Malaysia Duta seni Indonesia ke Taipei, Australia, Beijing Tari Bali;Tari Topeng Kedungmonggo;Tari Remo



		Surabayan/Gaya Jombangan/Gaya Sawunggaling.;Tari Beskalan;Tari tradisi Malangan, Mataraman, kreasi Tata Rias Koreografer;Penari;Pelaku Ahli bidang seni tradisi
4	Alvianitaa	Matta fair malaysia # 2016 Festival karya tari 2019 Tidak ada Sanggr tari maheswari , malang dance company Penari Bisa ke luar negeri hanya dari tari Tari Tari Bali;Tari Jaipong;Tari Topeng Kedungmonggo;Tari Remo Surabayan/Gaya Jombangan/Gaya Sawunggaling.;Tari Gandrung;Tari Beskalan Jaranan;Reog;Singo Ulung Baju dan aksesoris seni;Tata Rias Pencipta;Pelatih;Koreografer;Tata busana;Tata Panggung;Pengriasi;Penari;Pelaku Tidak punya
5	Yusak Santoso	sutradara terbaik teater pelajar se-jawa timur PentasTunggal "Kunjungan Nyonya Tua" Tahun 2002 10 Penyaji Terbaik Non Peringkat PSP Banyuwangi "Los Bagados De Los Pencos Tahun 2003 5 Penyaji Terbaik Non Peringkat PSP Madiun "Monumen Keadilan" Tahun 2005 5 Sutradara Terbaik Non Peringkat PSP Madiun "Monumen Keadilan" Tahun 2005 Juara II Festival Teater Cinta Tanah Air tingkat nasional "Semar Gugat Semar Ruwat" Tahun 2007 5 Penyaji Terbaik Non Peringkat fragmen Sejarah Se- Jawa Timur "Kudeta Para Brahmana" 2008 Sutradara Drama Tari "Sandya Nuswantara" Anjungan Jawa timur TMII Tahun 2016 Juara 1 PSP Tingkat Kota Batu "Srikandi Edian" 2016 Sutradara "Joko Tarub Single Parents" Pentas Bareng Hari Teater Sedunia di Gedung Kesenian MbatuAJi 2017 Sutradara Drama Tari "Semar Mendem" Duta Seni Kota Batu Anjungan jawa timur TMII 2018 Teater Pandu#Pembina#1997 - Sekarang Lembaga Budaya Akur#Sekretaris#2018 - Sekarang Teater Pandu Pembina Pantang Melacurkan diri pada tepuk tangan penonton Teater/Drama Musikal/Parodi;Opera;Film atau Sinetron;Wayang Wong;Ketoprak;Ludruk;Kentrung Bantengan;Jaranan;Tayub Irianan Jaranan/Bantengan Sastra Sutradara;Penulis Naskah;Pelatih;Tata Panggung;Pelaku Videography, Photography, Editor Multimedia
...
29	Suwandi	belum ada pentas dari event ke event turonggo seto anggota sanduk besta mulyo anggota sanduk besta mulyo anggota ikut terlibat dalam event festival sanduk kota batu tari sanduk jaranan reog iringan jaranan bantengan iringan reog karawitan gagrag surakarta penari pelaku menguasai beberapa seni tradisi lainnya yang pernah digeluti seperti seni reog dan jaranan

Proses klasifikasi pelaku seni dilakukan dengan beberapa tahap diantaranya sebagai berikut :

- *Text preprocessing* yang meliputi proses *case folding*, *tokenizing*, *filtering* dan *stemming* yang dilakukan dengan menggunakan metode *enhanced confix stripping*
- Pembobotan *Term* yang dilakukan dengan melakukan pembobotan terhadap *term* yang merupakan hasil *stemming* menggunakan TF-IDF
- Proses klasifikasi data pelaku seni dengan *Fuzzy Clustering* yang mana data yang diklasifikasi merupakan hasil dari pembobotan *term*.
- Proses pengujian yang dilakukan dengan cara menghitung prosentasi kesalahan pencarian kata dasar pada proses *stemming* dan perhitungan *cluster variance*

IV. HASIL DAN PEMBAHASAN

Proses *stemming* dilakukan setelah proses *filtering*. Hasil dari proses *filtering* dapat dilihat pada Tabel 8.

Tabel 8. Hasil *Filtering*

No	Nama	Deskripsi
1	Katijan	dinas pendidikan prov jawa timur duta seni kota batu sapto darmo pelatih karawitan sapto darmo margi laras jati larasati smp arjuno wahyu budaya pelatih duta seni kota batu aktif seniman pelatih kota batu jombok ngantang kab malang wayang ketoprak ludruk campursari jaranan tayub iringan jaranan bantengan karawitan gagrag metaraman karawitan gagrag surakarta pelatih niyaga pelaku telaten melatih
2	Mirza Baihaqie	cempoko wahyu budoyo sekretaris rogo wilis sekretaris rogo wilis cempoko wahyu budaya apeksi kota batu pemuda pegiat seni tradisi bantengan jaranan kidal gambuh pawang dukun pelaku gambuh pertunjukan
3	Karin Mirda Yuwana	kografer psp flsn terbaik kota batu penyaji terbaik fkt jawa timur duta tari kota batu korea selatan duta tari kota batu malaysia duta seni indonesia taipei australia beijing sanggar tari maheswari sanggar taru maheswari duta tari kota batu korea selatan duta tari kota batu malaysia duta seni indonesia taipei australia beijing tari bali tari topeng kedungmonggo tari remo surabayan gaya jombangan gaya



		sawunggaling tari beskalan tari tradisi malangan mataraman kreasi tata rias koreografer penari pelaku ahli bidang seni tradisi
4	Alvianitaa	matta fair malaysia festival karya tari sanggr tari maheswari malang dance company penari negeri tari tari tari bali tari jaipong tari topeng kedungmonggo tari remo surabayan gaya jombangan gaya sawunggaling tari gandrung tari beskalan jaranan reog singo ulung baju aksesor seni tata rias pencipta pelatih koreografer tata busana tata panggung pengriasi penari pelaku
5	Yusak Santoso	sutradara terbaik teater pelajar jawa timur pentastunggal kunjungan nyonya tua penyaji terbaik non peringkat psp banyuwangi los bagados de los pencos penyaji terbaik non peringkat psp madiun monumen keadilan sutradara terbaik non peringkat psp madiun monumen keadilan juara ii festival teater cinta tanah air tingkat nasional semar gugat semar ruwat penyaji terbaik non peringkat fragmen sejarah jawa timur kudeta brahmana sutradara drama tari sandya nuswantara anjungan jawa timur tmii juara psp tingkat kota batu srikandi edian sutradara joko tarub single parents pentas bareng teater sedunia gedung kesenian mbatuaji sutradara drama tari semar mendem duta seni kota batu anjungan jawa timur tmii teater pandu pembina sekarang lembaga budaya akur sekretaris teater pandu pembina pantang melacurkan tepuk tangan penonton teater drama musikal parodi opera film sinetron wayang ketoprak ludruk kentrung bantengan jaranan tayub iringan jaranan bantengan sastra sutradara penulis naskah pelatih tata panggung pelaku videography photography editor multimedia
...
29	Suwandi	pentas event event turonggo seto anggota sanduk besta mulyo anggota sanduk besta mulyo anggota terlibat event festival sanduk kota batu tari sanduk jaranan reog iringan jaranan bantengan iringan reog karawitan gagrag surakarta penari pelaku menguasai seni tradisi digeluti seni reog jaranan

Hasil dari proses *stemming* dapat dilihat pada Tabel 9.

Tabel 9. Hasil Stemming

No	Nama	Deskripsi
1	Katijan	dinas didik prov jawa timur duta seni kota batu sapto darmo latih karawitan sapto darmo margi laras jati larasati smp arjuno wahyu budaya latih duta seni kota batu aktif seniman latih kota batu jombok ngantang kab malang wayang ketoprak ludruk campursari jaranan tayub iring jaranan bantengan karawitan gagrag metaraman karawitan gagrag surakarta latih niyaga laku telaten latih
2	Mirza Baihaqie	cempoko wahyu budoyo sekretaris rogo wilis sekretaris rogo wilis cempoko wahyu budaya apeksi kota batu pemuda giat seni

		tradisi bantengan jaranan kidal gambuh pawang dukun laku gambuh pertunjukan
3	Karin Mirda Yuwana	kografer psp flsn baik kota batu saji baik fkt jawa timur duta tari kota batu korea selatan duta tari kota batu malaysia duta seni indonesia taipei australia beijing sanggar tari maheswari sanggar taru maheswari duta tari kota batu korea selatan duta tari kota batu malaysia duta seni indonesia taipei australia beijing tari bal tari topeng kedungmonggo tari remo surabayan gaya jombangan gaya sawunggaling tari tari tradisi malangan mataraman kreasi tata rias koreografer tari laku ahli bidang seni tradisi
4	Alvianitaa	matta fair malaysia festival karya tari sanggr tari maheswari malang dance company tari neger tari tari tari bal tari jaipong tari topeng kedungmonggo tari remo surabayan gaya jombangan gaya sawunggaling tari gandrung tari jaranan reog singo ulung baju aksesor seni tata rias cipta latih koreografer tata busana tata panggung rias tari laku
5	Yusak Santoso	sutradara baik teater ajar jawa timur pentastunggal kunjung nyonya tua saji baik non peringkat psp banyuwangi los bagados de los pencos saji baik non peringkat psp madiun monumen keadilan sutradara baik non peringkat psp madiun monumen keadilan juara ii festival teater cinta tanah air tingkat nasional mar gugat mar ruwat saji baik non peringkat fragmen sejarah jawa timur kudeta brahmana sutradara drama tari sandya nuswantara jung jawa timur tmii juara psp tingkat kota batu srikandi edian sutradara joko tarub single parents pentas bareng teater dunia gedung nian mbatuaji sutradara drama tari mar mendem duta seni kota batu jung jawa timur tmii teater pandu bina sekarang lembaga budaya akur sekretaris teater pandu bina pantang melacurkan tepuk tangan tonton teater drama musikal parodi opera film sinetron wayang ketoprak ludruk kentrung bantengan jaranan tayub iring jaranan bantengan sastra sutradara tulis naskah latih tata panggung laku videography photography editor multimedia
...
29	Suwandi	pentas event event turonggo seto anggota sanduk besta mulyo anggota sanduk besta mulyo anggota libat event festival sanduk kota batu tari sanduk jaranan reog iring jaranan bantengan iring reog karawitan gagrag surakarta tari laku seni tradisi seni reog jaranan

Berdasarkan Tabel 8 dan Tabel 9 Pengujian akurasi hasil *stemming* dapat dilihat pada Tabel 10.

Tabel 10. Pengujian Akurasi Hasil Stemming



No	Nama	Jumlah Kata Imbuhan	Hasil Stemming Tidak Akurat	% Akurasi
1	Katijan	8	0	100
2	Mirza Baihaqie	4	1	75
3	Karin Mirda Yuwana	5	0	100
4	Alvianitaa	6	0	100
5	Yusak Santoso	21	4	80.95
6	Maylian Giofany	5	1	80
7	Syamsu S. Soeid	19	2	89.47
8	Astika Putri Devi	1	0	100
9	Theresia Olivia Carolina Angga Setyorini	10	3	70
10	Safiudin Tristiandika Saputra	16	2	87.5
11	Safiudin Tristiandika Saputra	16	1	93.75
12	Kiki Devi Novita Yanti	7	1	85.71
13	Rezsa Bagas Styo Anggoro	10	2	80
14	Bagus wigno priyango	8	1	87.5
15	Ainun nabila azhar	3	0	100
16	RM Ananda Iman Akbar Zulkarnain	8	2	75
17	Kayla niken pradita	5	1	80
18	Mochamad Haidar Athallah	3	0	100
19	Dimas Al Hafid	11	1	90.90
20	Bagas Maulana	2	0	100
21	Muhammad Attar Marzuq	12	1	91.66667
22	Aisyah Devka Safitri	5	0	100
23	Ramadhan Arya Ismoyo	1	0	100
24	Warsito	7	0	100
25	Suyati	5	0	100
26	Afrian Saputra	14	0	100
27	Agus Triwahyudi	10	0	100
28	NURHANIEF	11	1	90.90
29	Suwandi	7	0	100
Rata-rata akurasi				91.67

Berdasarkan Tabel 10 diperoleh hasil *stemming* dengan akurasi 91.67%. Hal ini menunjukkan bahwa metode *enhanced confix stripping* mampu mencari kata dasar

dari kata berimbuhan dengan akurat. Hasil *stemming* yang digunakan untuk proses klasifikasi memberikan hasil klasifikasi yang akurat hal ini dapat dilihat pada hasil pengujian yang dilakukan sebanyak lima kali pada Tabel 11.

Tabel 10. Pengujian Akurasi Hasil *Stemming*

No	Vw	Vb	V
1	0.814814815	10683.14587	0.0000762710557756
2	0.185185185	10683.14587	0.0000173343308580
3	0.185185185	10683.14587	0.0000173343308580
4	0.185185185	10683.14587	0.0000173343308580
5	0.185185185	10683.14587	0.0000173343308580
Rata-rata V			0.0000291216758415

Berdasarkan Tabel 11. Menunjukkan bahwa rata-rata nilai varian semua *cluster* (v) untuk lima kali pengujian sangat kecil yaitu sebesar 0.0000291216758415 sehingga hasil *cluster* menunjukkan sangat baik

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan proses pengujian yang dilakukan maka dapat ditarik kesimpulan bahwa :

1. Metode *enhanced confix stripping* mampu mencari kata dasar dari kata berimbuhan dengan tingkat akurasi 91,67%
2. Hasil *stemming* menggunakan metode *enhanced confix stripping* mampu memberikan hasil *cluster* yang sangat baik yaitu memiliki rata-rata varian semua *cluster* (v) sebesar 0.0000291216758415

B. Saran

Perlu menambahkan data pada database kamus dasar yang lebih lengkap, karena kegagalan proses pencarian kata dasar dari kata berimbuhan disebabkan oleh kamus dasar yang dimiliki masih kurang lengkap.

REFERENSI

- Asriningtias, Salnan Ratih; Sonalitha, Elta.
Aplikasi K-Means Clustering untuk Pengelompokan Siswa ke Dalam Kelas Berdasarkan Nilai Akademis, Jenis Kelamin, Perilaku dan Nama Siswa di SMA Negeri 1 SRENGAT.



- Seminar Nasional Sistem Informasi (SENASIF). 2018. Vol 1. Issue 1. Halaman 1179-1187
- Ayeelden H., Hassanien, A.E., Fahmy, A.A., *Fuzzy Clustering and Categorization of Text Document*, IEEE, 2013, pp 262-266.
- Divya, P., Kumar, G.S.N., *Study of Feature Selection Method for Text Mining*, IJARTET, 2015, Vol. 2, Issue 1, pp 11-19.
- Kumar, S., Varma, B., *A Different Type of Feature Selection Method for Text Categorization on Imbalanced Data*, IJARCCCE, 2016, Vol.5, Issue 9, pp 297-303
- Langgeni, D.P., Baizal, Z.A., Firdaus, Y., *Clustering Artikel Berita Berbahasa Indonesia Menggunakan Unsupervised Feature Selection*, Seminar Nasional Informatika UPN Veteran Yogyakarta, 2010, ISSN : 1979-2328 pp 1-10
- Liu, T., Liu, S., Chen, Z., *An Evaluation on Feature Selection for Text Clustering*, 2003, ICML.
- Purbo, O.W., 2019, *Text Mining*, Yogyakarta : ANDI
- Ramya, M., Pinakas, J.A., *Different Type of Feature Selection for Text Classification*, IJCTT, 2014, ISSN: 2231-28-3, pp 102-107
- Sonalitha, Elta; Nurdewanto, Bambang; Ratih, Salnan; Sari, Nadia Roosmalita; Setiawan, Aries Boedi; Tutuko, Pindo. *Comparative Analysis of Tsukamoto and Mamdani Fuzzy Inference System on Market Matching to Determine the Number of Exports for MSMEs*. 2018 Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS). Page 440-445.
- Sruthi, S., Shalini, S., *Sentence Clustering in Text Document Using Fuzzy Clustering Algorithm*, 2014, ICCICCT.
- Sruthi, S., Shalini, S., *Sentence Clustering in Text Document Using Fuzzy Clustering Algorithm*, 2014, ICCICCT.
- Suresh, H., Raj., G., *An Unsupervised Fuzzy Clustering Method for Twitter Sentiment Analysis*, 2016, Internasional Conference on Computational System and Information Systems for Sustainable Solution.