

# Penerapan Vector Space Model Dalam Klasifikasi Penilaian Thematic Apperception Test

Usman Nurhasan<sup>1</sup>, Rakhmat Arianto<sup>2</sup>, Alwan Ghozi Kurnia<sup>3</sup>

<sup>1,2,3</sup>Jurusan Teknologi Informasi, Politeknik Negeri Malang  
Jalan Soekarno Hatta No. 9 Malang

Email: <sup>1</sup>usmannurhasan@polinema.ac.id, <sup>2</sup>arianto@polinema.ac.id, <sup>3</sup>alwaangk@gmail.com

## Abstract

Psychology is the study of human mental behavior and functions scientifically. In practice, a person's personality can be assessed from psychological tests. One of the psychological tests is the Thematic Apperception Test (TAT). TAT Test is a projective psychological test consisting of various themes presented in the form of an image which is then projected accordingly with the response. The purpose of the TAT is to reveal the dynamics of the subject's personality in the form of encouragement, sediment, complex, and various dominant conflicts. Thematic Apperception Test still uses a card and tape recorder to record Testee's voice. Calculation of results or assessments is still done manually. Errors in the assessment will affect the results, so we need an intelligent information system using the vector space model method based on experience in the field of psychology. A web-based system which can provide personality test results and information about the Thematic Apperception Test. From the test results using 10 stories, it was found that the average precision value was 72.7%, the average recall value was 86.7%, and the average f-measure value was 78.3%.

**Keywords:** Expert System, Text Mining, Vector Space Model, psychology, Thematic Apperception Test.

## Abstrak

Psikologi adalah ilmu yang mempelajari tentang perilaku dan fungsi mental manusia secara ilmiah. Dalam prakteknya, kepribadian seseorang dapat dinilai dari tes psikologi. Salah satu tes psikologi adalah Thematic Apperception Test. Thematic Apperception Test adalah suatu tes psikologi proyektif yang tersusun dari beberapa tema gambar kemudian ditampilkan untuk mendapatkan ulasan dari pengguna. Tujuan Thematic Apperception Test yaitu untuk mengungkap dinamika kepribadian subjek yang berupa dorongan, setimen, kompleks, dan berbagai konflik yang dominan. Thematic Apperception Test masih menggunakan kartu dan tape recorder untuk merekam suara testee. Perhitungan hasil atau penilaian masih dilakukan secara manual. Kesalahan dalam penilaian akan mempengaruhi hasil sehingga diperlukan sistem informasi cerdas yang menerapkan penggunaan metode vector space model dengan dasar pengalaman di bidang psikologi. Sistem berbasis web dimana dapat memberikan hasil tes kepribadian serta informasi mengenai Thematic Apperception Test. Dari hasil pengujian menggunakan 10 cerita didapat rata-rata nilai precision sebesar 72.7%, rata-rata nilai recall sebesar 86.7%, dan rata-rata nilai f-measure sebesar 78.3%.

**Kata kunci:** Sistem Cerdas, Text Mining, Vector Space Model, Psikologi, Thematic Apperception Test.

## 1. PENDAHULUAN

Penilaian psikologi adalah serangkaian kegiatan yang sebagian besar dilakukan oleh psikolog dan biro layanan psikologi di masyarakat untuk berbagai kepentingan [1]. Salah satunya adalah kondisi kesehatan mental



seseorang. Data Riskesdas untuk 2018 menunjukkan bahwa 7 dari 1000 Rumah Tangga memiliki anggota keluarga dengan menyandang Skizofrenia / Psikosis. Lebih dari 19 juta orang berusia di atas 15 tahun terkena gangguan mental emosional atau sering disebut emotional mental disorders, serta lebih dari 12 juta orang berusia pada skala usia di atas 15 tahun diperkirakan telah mengalami depresi. Sementara itu, dari data dipaparkan bahwa kasus bunuh diri di Indonesia menyentuh angka 1,6 hingga 1,8% per 100.000 orang [2].

Dengan adanya permasalahan diatas, penulis terdorong untuk mengangkat permasalahan tersebut dan membuat suatu sistem cerdas yang mampu membantu kinerja psikolog dalam melakukan Thematic Apperception Test (TAT) sebagai penelitian dengan judul Rancang Bangun Sistem Cerdas Thematic Apperception Test Menggunakan Metode Vector Space Model. Dengan menerapkan metode Vector Space Model (VSM) sebagai pemroses ulasan, diharap dapat mengetahui tingkat kemiripan dari cerita yang dibuat oleh testee dengan kategori yang ada di dalam Thematic Apperception Test.

Pada penelitian mengenai Thematic Apperception Test oleh Bilal Zavanna Sulaiman dengan judul Cognitive Behavior Therapy untuk Meningkatkan Perilaku rutin minum obat pada penderita skizofrenia [3] dapat ditarik kesimpulan bawah TAT menjadi salah satu asesmen untuk mengetahui pribadi objek dalam menggunakan pendekatan cognitive behavior therapy. Hal ini terlihat dari hasil Tes TAT yang menunjukkan bahwa kebutuhan utama subjek sekarang adalah dukungan dari lingkungan sosial agar subjek bisa mencapai kebahagiaannya sesuai yang diharapkan. Penelitian lain yang dilakukan oleh Ranap Sitorus, Harry Soekotjo Dachlan, dan Wijono dengan judul Analisis Pengaruh Frasa Pada Deteksi Emosi Dari Teks Menggunakan Vector Space Model [4]. Penelitian ini dilakukan dengan menggunakan algoritma Term Frequency – Inversed Document Frequency (TF-IDF) dan Vector Space Model. Pada penelitian tersebut Algoritma Term Frequency – Inversed Document Frequency (TF-IDF) dan Vector Space Model mampu mendeteksi dokumen emosi teks Bahasa Indonesia dengan data set 90%, hasil deteksi yang diperoleh dengan frasa dan non frasa berhasil dideteksi 100% untuk emosi senang dan takut. Berbeda dengan emosi kecewa menggunakan frasa 78% sedangkan non frasa 89%, perbandingannya sebesar 11%. Dalam klasifikasi emosi sangat baik akan tetapi hasil emosi dengan pendeteksian menggunakan Frase masih kurang sempurna. Apabila data set yang digunakan kecil maka hasil pendeteksian emosi sangat buruk, yang hasilnya berbanding terbalik dengan non Frase rata – rata 98% keberhasilannya.

Proses analisis data dalam TAT dapat dilakukan dengan menggunakan algoritma VSM. Penggunaan VSM bertujuan untuk mendapatkan pemaknaan yang tepat pada setiap ulasan yang diberikan. Pada penelitian dengan judul Sistem Temu Kembali Informasi Pada Dokumen Dengan Metode Vector Space Model oleh Irmawati [5] diambil kesimpulan bahwa sistem dapat bekerja secara maksimal dan efisien untuk menghasilkan informasi sesuai tujuan.

Dari proses analisis yang dilakukan didapatkan nilai Recall mendekati angka 100%, sedangkan nilai precision mendapatkan kisaran 73,6%. Jika melihat hasil yang didapatkan maka dapat disimpulkan algoritma tersebut relevan untuk melakukan temu-kembali dokumen berdasarkan keyword yang diberikan dengan tingkat keberhasilan 73,6%.

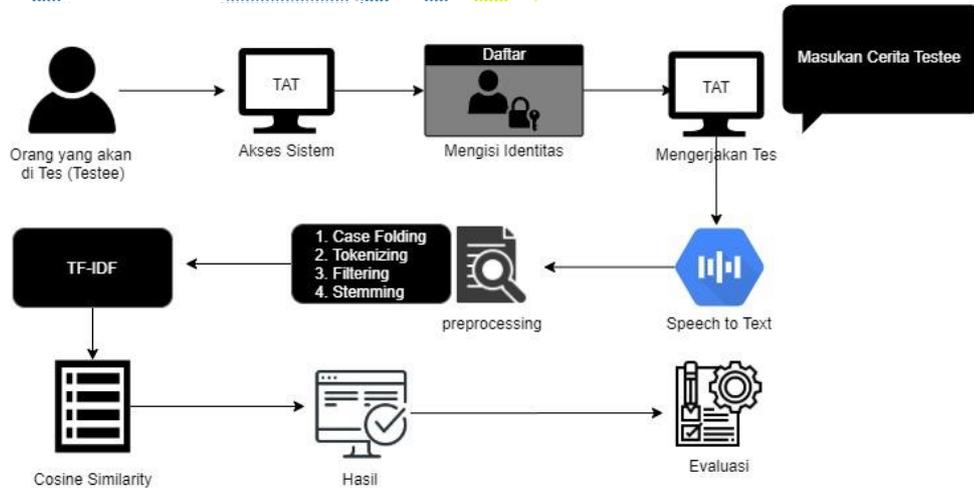
Penelitian lain oleh Anna dan Ade Hendini dengan judul Implementasi Vector Space Model. Pada Sistem Pencarian Mesin Karaoke [6] dapat ditarik kesimpulan bahwa penerapan Vector Space Model pada sistem pencarian lagu berdasarkan judul dapat mempercepat proses pencarian lagu yang relevan sesuai dengan lagu yang ingin dicari pengunjung. Dengan sistem ini pula dapat dilihat nilai bobot dari masing-masing lagu yang relevan menjadi urutan teratas untuk dipilih atau yang sedang populer.

Aplikasi yang dirancang nantinya akan diterapkan Speech to text, hal ini dilakukan untuk memberikan kemudahan dalam mendokumentasikan ulasan dari testee. Penggunaan teknologi speech to text banyak digunakan di beberapa penelitian. Pada penelitian dengan judul "Speech to Text Patient Complaint for Bahasa Indonesia" oleh Teguh Laksono, A. Fathan Hidayatullah dan Chanifah Indah R [7] dengan metode Mel Frequency Cepstral Coefficient (MFCC) dapat ditarik kesimpulan dari data suara yang digunakan dalam penelitian yaitu kata-kata dalam Bahasa Indonesia dan bahasa Jawa. Peneliti telah mengumpulkan 500 suara yang direkam data dari 10 orang. Dari 50 data uji memperoleh akurasi 64%.

## **2. METODOLOGI PENELITIAN**

### **2.1. Proses Bisnis**

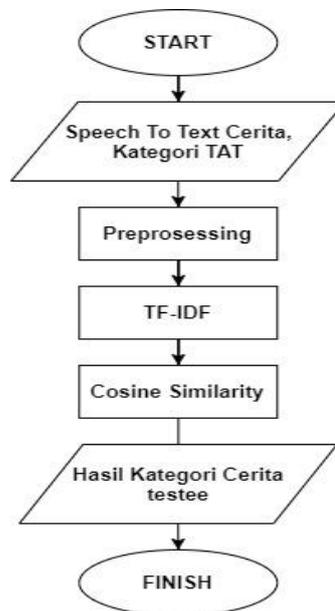
Pada gambar 1 merupakan bisnis proses sistem informasi Thematic Apperception Test dimana wawancara dilakukan untuk mengetahui semua kebutuhan yang akan digunakan oleh sistem. Data diambil dari buku panduan Thematic Apperception Test oleh Murray termasuk kartu serta kategori scoring TAT. Sistem akan memulai tes dengan menampilkan kartu, waktu respon, dan cerita testee (orang yang akan dites). Cerita testee akan diproses dengan menggunakan speech to text). Setelah testee bercerita, cerita tersebut akan dilakukan preprocessing untuk mendapatkan kata dasar dari setiap kata kemudian dihitung TF-IDF untuk mengetahui banyaknya kemunculan kata sehingga dapat dihitung bobot dari tiap kata. Selanjutnya dihitung Cosine Similarity untuk membandingkan kemiripan antara cerita testee dan kategori yang ada dalam TAT. Hasil dari TAT berupa waktu respon testee, heroes, need, press, inner state. Jika sudah didapatkan hasil tes, akan dilakukan analisa terhadap hasil tes oleh sistem dengan hasil pakar untuk mengevaluasi tingkat akurasi dari metode yang digunakan.



**Gambar 1.** Proses Bisnis Aplikasi yang Dirancang

## 2.2. Pengolahan Data

Pengolahan data yang dilakukan pada penelitian ini digunakan algoritma Vector Space Model. Data yang didapat dari input cerita testee merupakan data yang langsung diambil saat Thematic Apperception Test berlangsung. Alur pengolahan data ditampilkan pada gambar 2.



**Gambar 2.** Pengolahan Data

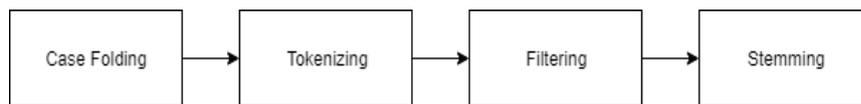
Gambar 2 merupakan alur pengolahan data, data tersebut dilakukan pengolahan agar menjadi data yang mudah digunakan dalam proses interpretasi. Cerita Testee di inputkan dengan menggunakan speech to text kemudian dilakukan preprocessing untuk penyeleksian kata-kata sehingga cerita menjadi lebih ringkas. Setelah melalui preprocessing, data berupa teks akan diubah ke dalam bentuk angka melalui perhitungan TF-IDF berdasarkan

skala yang telah ditentukan. Hasil dari perhitungan TF-IDF ini yang kemudian akan diproses lebih lanjut menggunakan algoritma Vector Space Model. Algoritma Vector Space Model akan menghasilkan kemiripan (cosine similarity) dibandingkan dengan data kategori need, press, inner state dalam Thematic Apperception Test oleh Murray.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Preprocessing

Pre-processing diimplementasikan dengan tujuan untuk mengurangi noise pada dataset, sehingga menghasilkan peningkatan performa klasifikasi [8]. Tahapan preprocessing meliputi Case Folding, Tokenizing, Filtering, Stemming. Tahapan preprocessing ditunjukkan pada gambar 3. Pada gambar tersebut merupakan tahapan dari preprocessing, input cerita Testee akan mengalami preprocessing terlebih dahulu untuk menyeleksi kata-kata. Contoh studi kasus.



**Gambar 3.** Preprocessing

Cerita : *“Ini Deni, anak usia 11 tahun yang ingin mencoba bermain biola tetapi dia tidak bisa untuk memainkannya. Dia sedang berpikir bagaimana caranya dia mulai bermain biola ini. Dia sedang berpikir bagaimana untuk memulai main biola, mulai main darimana, perkenalannya dengan alat bantuannya, dan dengan senar-senarnya. Tetapi dia belum tau nama alat bantuannya apa, dia hanya tau biola saja. Lalu akhirnya dia mencoba untuk mencari guru untuk mengajarkan dia bermain biola.”* Dan contoh tabel kategori dapat dilihat pada tabel 1.

**Tabel 1.** Kategori

Kategori	Daftar Kata
N-Achievement	Coba, Giat, Gigih, Usaha, Ambisi, Ulet, Tekun, Upaya, Teguh, Tekat,
N-Understanding	ajar sekolah, baca, ilmu, pikir, solusi
N-Acquisition Sosial	Serakah, Rakus, Tamak, Barter, Dagang, Judi, kerja
Konflik	coba, cara, pikir, himbau, kontrol, atur, ikut, benci, usaha, wujud, bangkang,

*Case Folding* yaitu tahap mengubah semua karakter masukan dokumen menjadi huruf kecil (*lowercase*).

“ini deni, anak usia 11 tahun yang ingin mencoba bermain biola tetapi dia tidak bisa untuk memainkannya. dia sedang berpikir bagaimana caranya dia mulai bermain biola ini. dia sedang berpikir bagaimana untuk memulai main biola, mulai main darimana, perkenalannya dengan alat bantuannya, dan dengan senar-senarnya. tetapi dia belum tau nama alat bantuannya apa, dia

hanya tau biola saja. lalu akhirnya dia mencoba untuk mencari guru untuk mengajarkan dia bermain biola.”

*Tokenizing* yaitu pemotongan string input. Hasil dari *Tokenizing* dapat dilihat di tabel 2.

**Tabel 2. Tokenizing**

deni	anak	usia	11
tahun	yang	ingin	mencoba
bermain	biola	tetapi	dia
tidak	bisa	untuk	memainkanya
Dia	sedang	berfikir	bagaimana
caranya	dia	mulai	bermain
biola	ini	dia	sedang
berfikir	bagaimana	untuk	memulai
main	biola	mulai	main
darimana	perkenalannya	dengan	alat
bantunya	dan	dengan	senarnya
Tetapi	dia	belum	tau
nama	alat	bantunya	apa
dia	hanya	tau	bantunya
apa	dia	hanya	tau
bantunya	apa	dia	hanya
tau	bantunya	apa	dia
hanya	tau	biola	saja
Lalu	akhirnya	dia	mencoba
untuk	mencari	guru	untuk
mengajarkan	dia	bermain	biola

*Filtering* yaitu tahap pengambilan kata kata penting dari *tokenizing*. Hasil *filtering* dapat dilihat pada tabel 3.

**Tabel 3. Filtering**

deni	anak	bermain	tahun
memainkanya	memulai	berfikir	biola
caranya	usia	bermain	biola
main	mencoba	berfikir	mulai
bantunya	main	biola	alat
bantunya	perkenalannya	senarnya	alat
biola	mencari	nama	
mencoba	bermain	guru	
mengajarkan	11	biola.	

*Stemming* yaitu proses transformasi kata ke kata-kata dasar. Hasil *stemming* dapat dilihat pada tabel 4.

**Tabel 4. Stemming**

Deni	anak	main	biola
main	mulai	pikir	biola
cara	usia	main	mulai
main	coba	pikir	alat

bantu	main	biola	alat
bantu	kenal	senar	
biola	cari	nama	
coba	main	guru	
ajar	11	biola	

### 3.2. Implementasi TF-IDF

*Term Frequency – Inverse Document Frequency* (TF-IDF) adalah teknik pembobotan berbasis statistik yang sering diterapkan di berbagai permasalahan penggalian informasi. Namun, secara umum TF-IDF tidak banyak dikenal sebagai algoritma untuk peringkasan teks otomatis. Pada peringkasan teks otomatis menggunakan TF-IDF, ide dasarnya adalah memberikan bobot pada setiap kalimat dalam sebuah dokumen. Setelah masing-masing kalimat diberikan bobot, tahap selanjutnya adalah kalimat akan diurutkan berdasarkan bobot dimana kalimat k teratas dengan bobot paling besar akan dipilih sebagai hasil akhir ringkasan. Bobot kalimat diperoleh dari penjumlahan bobot term dari kosa kata pembentuk kalimat, dimana term dapat berupa kata, frasa atau tipe sintatik lainnya [9]. Proses perhitungan Vector Space Model melalui tahapan perhitungan term frequency (tf). TF mempresentasikan frekuensi kata kunci tertentu. TF digunakan untuk menunjukkan tingkat prioritas kata dalam kalimat [10]. TF dapat menggunakan persamaan (1)

$$tf = tf_{ij} \tag{1}$$

Dengan tf adalah term frequency, dan  $tf_{ij}$  adalah banyaknya kemunculan atau frekuensi kemunculan kata terpilih dalam dokumen. Term frequency (tf) dihitung dengan cara jumlah kemunculan kata terpilih dibagi dengan jumlah kata. Proses selanjutnya adalah perhitungan idf. Pada persamaan (2), ditulis rumus Perhitungan Inverse Document Frequency (idf)

$$idf_i = \log \frac{N}{df_i} \tag{2}$$

Dengan idfi adalah inverse document frequency untuk mengukur seberapa sering sebuah kata muncul dalam teks kalimat yang berbeda [10], N adalah jumlah dokumen yang terambil oleh sistem, dan dfi adalah banyaknya dokumen dalam koleksi dimana term  $t_i$  muncul di dalamnya, maka Perhitungan idfi digunakan untuk mengetahui banyaknya term yang dicari (dfi) yang muncul dalam dokumen lain yang ada pada database. Perhitungan term frequency Inverse Document Frequency (tfidf), menggunakan persamaan (3)

$$W_{ij} = tf_i \cdot idf_i \tag{3}$$

Dengan  $W_{ij}$  adalah bobot dokumen, N adalah Jumlah dokumen yang terambil oleh sistem,  $tf_{ij}$  adalah banyaknya kemunculan term  $t_i$  pada dokumen  $d_j$ , dan  $df_i$  adalah banyaknya dokumen dalam koleksi dimana term  $t_i$  muncul di dalamnya. Bobot dokumen ( $W_{ij}$ ) dihitung untuk didapatkannya suatu bobot hasil perkalian atau kombinasi antara term frequency ( $tf_{ij}$ ) dan Inverse Document Frequency (idf). Hasil TF-IDF dapat dilihat pada tabel 5.

**Tabel 5. TF-IDF**

Token	term frequency (tf)					document frequency (df)	d/df	Idf = log10 (D/df)
	Q	D1	D2	D3	D4			
coba	2	1	0	0	1	2	2	0.301029
giat	0	1	0	0	0	1	4	0.602059
gigih	0	1	0	0	0	1	4	0.602059
usaha	0	1	0	0	0	1	4	0.602059
ambisi	0	1	0	0	0	1	4	0.602059
ulet	0	1	0	0	0	1	4	0.602059
teguh	0	1	0	0	0	1	4	0.602059
tekat	0	1	0	0	0	1	4	0.602059
ajar	1	0	1	0	0	1	4	0.602059
sekolah	0	0	1	0	0	1	4	0.602059
baca	0	0	1	0	0	1	4	0.602059
ilmu	0	0	1	0	0	1	4	0.602059
pikir	2	0	1	0	1	2	2	0.301029
solusi	0	0	1	0	0	1	4	0.602059
serakah	0	0	0	1	0	1	4	0.602059
rakus	0	0	0	1	0	1	4	0.602059
tamak	0	0	0	1	0	1	4	0.602059
barter	0	0	0	1	0	1	4	0.602059
dagang	0	0	0	1	0	1	4	0.602059
judi	0	0	0	1	0	1	4	0.602059
kerja	0	0	0	1	0	1	4	0.602059
cara	1	0	0	0	1	1	4	0.6020
himbau	0	0	0	0	1	1	4	0.6020
kontrol	0	0	0	0	1	1	4	0.6020
atur	0	0	0	0	1	1	4	0.6020
ikut	0	0	0	0	1	1	4	0.6020
benci	0	0	0	0	1	1	4	0.6020
wujud	0	0	0	0	1	1	4	0.6020

Dari TF-IDF diatas didapatkan hasil pembobotan (w) pada query dan tiap dokument. Hasil pembobotan dapat dilihat pada tabel 6.

**Tabel 6. Hasil Pembobotan**

bobot (W)				
Q	D1	D2	D3	D4
0.6020	0.301030	0.000000	0.000000	0.301030
0	0.602059	0	0	0
0	0.602059	0	0	0
0	0.602059	0	0	0
0	0.602059	0	0	0
0	0.602059	0	0	0
0	0.602059	0	0	0
0.6020	0	0.602059	0	0
0	0	0.602059	0	0
0	0	0.602059	0	0
0	0	0.602059	0	0
0.6020	0	0.301029	0	0.301029

bobot (W)				
Q	D1	D2	D3	D4
0	0	0.602059	0	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0	0	0	0.602059	0
0.6020	0	0	0	0.602059
0	0	0	0	0.602059
0	0	0	0	0.602059
0	0	0	0	0.602059
0	0	0	0	0.602059
0	0	0	0	0.602059
0	0	0	0	0.602059
0	0	0	0	0.602059

Perhitungan Jarak antar query, menggunakan persamaan (4)

$$|q| = \sqrt{\sum_{j=1}^t (W_{iq})^2} \tag{4}$$

Dengan |q| adalah Jarak antar query, dan  $W_{iq}$  adalah bobot query dokumen ke-i, maka Jarak query (|q|) dihitung untuk didapatkan jarak antar query dari bobot query dokumen ( $W_{iq}$ ) yang terambil oleh sistem. Jarak antar query bisa dihitung dengan persamaan akar jumlah kuadrat dari query. Perhitungan Jarak antar dokumen, menggunakan persamaan (5)

$$|d_j| = \sqrt{\sum_{i=1}^t (W_{ij})^2} \tag{5}$$

Dengan |d<sub>j</sub>| adalah jarak antar dokumen, dan  $W_{ij}$  adalah bobot dokumen ke-i, maka Jarak dokumen (|d<sub>j</sub>|) dihitung untuk didapatkan jarak antar dokumen dari bobot dokumen dokumen ( $W_{ij}$ ) yang terambil oleh sistem. Jarak antar dokumen bisa dihitung dengan persamaan akar jumlah kuadrat dari dokumen. Setelah didapatkan hasil pembobotan untuk masing masing query dan data dari TF-IDF dapat dihitung jarak query dan dokumen serta index term untuk mendapatkan cosine similarity atau jarak antar dokumen. Berikut Tabel 7 Merupakan hasil perhitungan jarak query dan jarak dokumen.

**Tabel 7.** Hasil Jarak Query dan Jarak Dokumen

Jarak Query	Jarak Dokumen			
	D1	D2	D3	D4
1.20412	1.621096	1.379493	1.592901	1.648809

### 3.3. Implementasi VSM

*Vector Space Model* (VSM) adalah salah satu metode atau algoritma yang sering digunakan untuk sebuah sistem temu kembali informasi.

Algoritma ini merupakan sebuah model yang digunakan untuk mengukur kemiripan atau kesamaan (similarity term) antar suatu dokumen dengan suatu query dengan cara pembobotan term. Pada tahap Information Retrieval System, kemiripan antar dokumen didefinisikan berdasarkan representasi bag-of-words dan diubah ke suatu model ruang vector. Relevansi sebuah dokumen ke sebuah query didasarkan pada similaritas diantara vektor dokumen dan vektor query [6]. Pengukuran *Cosine Similarity* menggunakan persamaan (6).

$$sim d_j = \frac{q \cdot d_j}{|q| * |d_j|} \tag{6}$$

Similaritas antara query dan dokumen atau Sim(q,dj) berbanding lurus terhadap jumlah bobot query (q) dikali bobot dokumen (dj) dan berbanding terbalik terhadap akar jumlah kuadrat q (|q|) dikali akar jumlah kuadrat dokumen (|dj|). Perhitungan similaritas menghasilkan bobot dokumen yang mendekati nilai 1 atau menghasilkan bobot dokumen yang lebih besar dibandingkan dengan nilai yang dihasilkan dari perhitungan inner product. Hasil pengukuran cosine similarity dapat dilihat pada tabel 8.

**Tabel 8. Cosine Similarity**

Cosine Similarity	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	0.0928477	0.3273268	0	0.3651484

Setelah pengujian, perbandingan antara hasil manual dan hasil keluaran sistem diperoleh dengan menggunakan metode Vector Space Model. Hasil perbandingan pakar dan sistem dapat dilihat pada tabel 9.

Table 9. Hasil Perbandingan Pakar dan Sistem

No	KARTU	HASIL	
		PAKAR	SISTEM
1	1	N-Achievement	N-Achievement
		N-Understanding	N-Understanding
		Konflik	Konflik
2	2	N-Achievement	N-Achievement
		N-Abbasement (submission)	N-Understanding
		P-Dominance Paksaan	P-Dominance Paksaan
3	3GF	N-Acquisition Sosial	N-Achievement
		P-Lack Kekurangan	N-Acquisition Sosial
		Dejection	P-Lack Kekurangan
			Konflik
4	4	N-Affiliation Emosional	N-Achievement
		N-Nurturance	N-Affiliation Emosional
		P-Affiliation Emosional	N-Nurturance
		Konflik	P-Affiliation Emosional
			Konflik
5	5	N-Nurturance	N-Nurturance
		P-Deference Patuh	P-Deference Patuh
		Kecemasan	Kecemasan

No.	KARTU	HASIL	
		PAKAR	SISTEM
6	6GF	n. Abasement (submission)	N-Construction
		n. Deference (patuh)	N-Abasement submission
		p. Dominance (paksaan)	N-Deference Patuh
		p. Exposition	P-Dominance Paksaan
		Konflik	P-Exposition
			Konflik
7	7GF	n. Rejection	N-Nurturance
		p. Dominance	N-Rejection
		Dejection	Dejection
8	8GF	N-Achievement	N-Achievement
		N-Recognition	N-Understanding
		P-Deference Hormat	N-Recognition
		Konflik	P-Deference Hormat
			Konflik
9	9GF	N-Rejection	N-Nurturance
		P-Rejection	N-Autonomy Kebebasan
		Konflik	P-Rejection
			Konflik
10	10	N-Affiliation Emosional	N-Affiliation Emosional
		P-Sex	N-Nurturance
		Eksaltasi	N-Succorance
			P-Affiliation Emosional
			P-Sex
		Eksaltasi	

Tabel 9 adalah hasil output manual yang diperoleh dari analisis psikolog yang telah melakukan TAT dan output sistem menggunakan metode Vector Space Model.

### 3.4. Hasil Analisis

Diskusi ini bertujuan untuk menyimpulkan hasil dari pengujian sistem yang telah dibuat berdasarkan tes yang dilakukan oleh testee. Dari hasil pada tabel 9 dapat dihitung nilai precision, recall, dan f-measure. Precision didefinisikan sebagai pengukuran ketepatan. Jika data diprediksi positif, seberapa seringkah data prediksi tersebut benar. Sedangkan recall didefinisikan sebagai pengukuran kelengkapan. Dari jumlah data sebenarnya yang bernilai positif, sebanyak apakah data yang diprediksi positif [11]. Sedangkan, f-measure dapat digunakan untuk mempertimbangkan nilai precision dan recall menjadi satu kesatuan yang saling berhubungan satu sama lain [12].

Rumus masing masing precision, recall, dan f-measure dapat dilihat pada persamaan (7), (8), dan (9) sebagai berikut :

$$precision = \frac{TP}{TP + FP} \tag{7}$$

$$recall = \frac{TP}{TP + FN} \tag{8}$$

$$f - \text{measure} = 2x \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} \quad (9)$$

Dengan TP adalah True Positif dimana TP adalah jumlah kategori yang berhasil di ekstrak pada sistem yang sesuai dengan hasil manual (pakar). FP adalah False Positif dimana FP jumlah kategori yang di ekstrak sistem namun tidak terdapat pada hasil manual (pakar) dan FN adalah False Negatif dimana jumlah kategori dari hasil manual (pakar) tetapi tidak terdapat pada hasil kategori yang diekstrak sistem. Misalkan pada kartu 1, kartu 1 memiliki output sistem sebanyak 3 kategori yang tepat dari 3 jumlah kategori yang muncul pada sistem, maka cerita pada kartu 1 memiliki :

$$\text{nilai precision} = TP/(TP+FP) = 3/(3+0) = 1,$$

$$\text{nilai recall} = TP/(TP+FN) = 3/(3+0) = 1, \text{ dan}$$

$$\text{nilai f-measure} = 2x ((\text{precision} \times \text{recall})) / ((\text{precision} + \text{recall})) = 2x ((1 \times 1)) / ((1+1)) = 1.$$

**Table 10.** Perhitungan Precision, Recall, dan F-measure

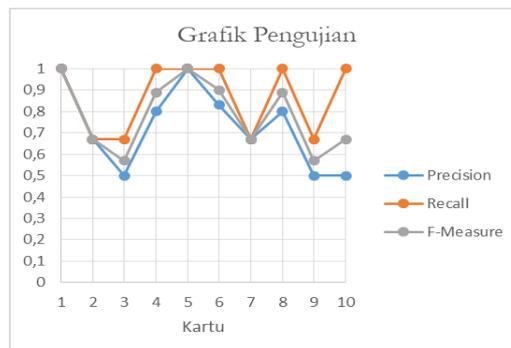
No	Kartu	Precision	Recall	F-Measure
1	1	1	1	1
2	2	0.67	0.67	0.67
3	3GF	0.5	0.67	0.57
4	4	0.8	1	0.89
5	5	1	1	1
6	6GF	0.83	1	0.9
7	7GF	0.67	0.67	0.67
8	8GF	0.8	1	0.89
9	9GF	0.5	0.67	0.57
10	10	0.5	1	0.67
Rata-rata		0.727	0.868	0.783

Contoh lain pada kartu 4, kartu 4 memiliki output sistem sebanyak 4 kategori yang tepat dari jumlah 5 kategori yang muncul pada sistem, maka cerita pada kartu 4 memiliki :

$$\text{nilai precision} = TP/(TP+FP) = 4/(4+1) = 0.8,$$

$$\text{nilai recall} = TP/(TP+FN) = 4/(4+0) = 1, \text{ dan}$$

$$\text{nilai f-measure} = 2x ((\text{precision} \times \text{recall})) / ((\text{precision} + \text{recall})) = 2x ((0.8 \times 1)) / ((0.8+1)) = 0.89.$$



**Gambar 4.** Grafik Pengujian

Hasil perhitungan precision, recall, dan f-measure dapat dilihat pada tabel 10. Tabel tersebut adalah hasil pengujian dari 10 kartu Tes TAT dengan nilai rata-rata precision sebesar 0.727 (73%), rata-rata recall 0.868 (87%), dan rata-rata f-measure 0.783 (78%). Pada gambar 4 merupakan grafik pengujian dari tiap kartu dengan menghitung nilai nilai precision, recall, dan f-measure.

#### 4. SIMPULAN

Berdasarkan dari hasil analisis, perancangan, implementasi serta pengujian yang telah dilakukan maka dapat diambil kesimpulan bahwa implementasi speech to text bahasa indonesia dengan API mozilla.org dapat digunakan untuk merubah cerita testee dari suara menjadi teks. Saat dilakukan pengujian terhadap user, terdapat beberapa kata yang tidak sesuai dengan apa yang diucapkan testee namun secara keseluruhan aplikasi ini dapat menjadi alternatif dalam pengerjaan interpretasi test TAT.

Metode *Vector Space Model* dapat digunakan untuk menentukan kategori pada tiap cerita. Metode tersebut menggunakan TF-IDF untuk mendapatkan bobot dari masing masing dokumen yang didapatkan dari hasil perkalian Term Frequency denan Index Document Frequency. Dari hasil pembobotan tesebut dapat dihitung jarak antar query dan jarak antar dokumen untuk dapat menghitung nilai dari cosine similarity. Dari perhitugnan dengan metode *Vector Space Model* didapatkan hasil dengan nilai rata-rata precision sebesar 0.727 (73%), rata-rata recall 0.868 (87%), dan rata-rata f-measure 0.783 (78%).

Untuk pengembangan lebih lanjut, penulis menyarankan agar dilakukan penggunaan algoritma lain dalam penentuan kategori. Selain itu dapat pula dilakukan pengembangan dengan API Speech to Text lain yang mungkin memberikan hasil yang lebih akurat.

#### DAFTAR PUSTAKA

- [1] C. Suwartono, "Alat Tes Psikologi Konteks Indonesia: Tantangan Psikologi di Era MEA," *J. Psikol. Ulayat*, vol. 3, no. 1, p. 1, 2017, doi: 10.24854/jpu12016-51.
- [2] Kemkes, "Pentingnya Peran Keluarga, Institusi dan Masyarakat Kendalikan Gangguan Kesehatan Jiwa," 2019.
- [3] B. Z. Sulaiman, "Cognitive Behavior Therapy untuk Meningkatkan Perilaku Rutin Minum Obat pada Penderita Skizofrenia," *J. Univ. Muhammadiyah Malang*, no. 2002, pp. 19–20, 2016.
- [4] R. Sitorus and H. S. Dachlan, "Analisis Pengaruh Frasa Pada Deteksi Emosi Dari Teks Menggunakan Vector Space Model," vol. 11, no. 1, pp. 41–47, 2017.
- [5] Irmawati, "Sistem Temu Kembali Informasi Pada Dokumen Dengan Metode Vector Space Model," *J. Ilm. FIFO*, vol. IX, no. 1, pp. 74–80, 2017.
- [6] A. Hendini, "Implementasi Vector Space Model Pada Sistem Pencarian Mesin Karaoke," vol. 6, no. 1, pp. 1–6, 2018.

- [7] T. P. Laksono, A. F. Hidayatullah, and C. I. Ratnasari, "Speech to Text of Patient Complaints for Bahasa Indonesia," *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, pp. 79–84, 2019, doi: 10.1109/IALP.2018.8629161.
- [8] F. F. Zain and Y. Sibaroni, "Effectiveness of SVM Method by Naïve Bayes Weighting in Movie Review Classification," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 5, no. 2, pp. 108–114, 2019, doi: 10.23917/khif.v5i2.7770.
- [9] D. H. Wahid and A. SN, "Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 10, no. 2, p. 207, 2016, doi: 10.22146/ijccs.16625.
- [10] H. Sujaini, "Performance of Methods in Identifying Similar Languages Based on String to Word Vector," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 6, no. 1, pp. 9–14, 2020, doi: 10.23917/khif.v6i1.8199.
- [11] R. P. S. Putri and I. Waspada, "Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 4, no. 1, p. 1, 2018, doi: 10.23917/khif.v4i1.5975.
- [12] V. J. L. Engel, E. Joshua, and M. M. Engel, "Detection of Cyber Malware Attack Based on Network Traffic Features Using Neural Network," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 6, no. 1, pp. 26–32, 2020, doi: 10.23917/khif.v6i1.8869.