

TEKNIK *VECTOR SPACE MODEL* (VSM) DALAM PENENTUAN PENANGANAN DAMPAK *GAME ONLINE* PADA ANAK

Bania Amburika^{1*}, Yulison Herry Chrisnanto¹, Wisnu Uriawan²

¹ Jurusan Informatika, Fakultas MIPA, Universitas Jenderal Achmad Yani.
Jl. Terusan Jendral Sudirman, Cimahi, Jawa Barat.

² Jurusan Teknik Informatika, Fakultas Sains dan Teknologi,
UN Sunan Gunung Djati.

Jl. A.H. Nasution No.105, Jawa Barat.

*Email: amburikabania@gmail.com

Abstrak

Perkembangan teknologi informasi menjadi lingkungan baru bagi anak, anak dapat dengan mudah bergaul dan mengikuti perkembangan tersebut. Perkembangan teknologi informasi yang semakin pesat salah satu contohnya yaitu banyaknya permainan *game online*. *Game online* merupakan permainan yang dimainkan secara nyata, melawan pihak lain, dengan waktu terpisah dan dibantu oleh PC/handphone, serta dapat dimainkan dengan mudah oleh anak sehingga seorang anak dapat terikat pada *game online* dan mengalami kecanduan. Kecanduan *game online* pada anak dapat menjadikan pengaruh buruk bagi perkembangan anak sehingga orang tua perlu mengetahui dampak buruk serta penanganan dari ketergantungan yang diakibatkan oleh *game online*. Berdasarkan hal tersebut, dapat dibangun sebuah sistem yang dapat menjadi rekomendasi orang tua dalam penentuan penanganan dampak *game online* pada anak yang mengalami kecanduan. Untuk melakukan proses penentuan penanganan, metode yang digunakan adalah *Vector Space Model* (VSM), metode ini merupakan metode yang digunakan untuk mengukur tingkat kemiripan pada setiap gejala yang dimasukkan oleh pengguna dan membandingkan dengan gejala yang sudah tersimpan didalam database. Hasil dari penelitian ini berupa rekomendasi penanganan yang dapat dilakukan berdasarkan gejala yang menjadi masukan.

Kata Kunci : *Game online*, Gejala, VSM (*Vector Space Model*)

1. PENDAHULUAN

Perkembangan anak usia sekolah sangat ketat, karena kemampuan seseorang pada saat dewasa dapat dipengaruhi oleh perkembangan anak pada saat usia sekolah. Anak usia sekolah sangat dengan mudah dipengaruhi oleh perkembangan teknologi informasi terutama oleh *game online*. Kecanduan *game online* pada anak dapat menjadikan pengaruh buruk bagi perkembangan anak sehingga orang tua perlu mengetahui dampak buruk serta penanganan dari ketergantungan yang diakibatkan oleh *game online*. Anak usia sekolah sangat dengan mudah dipengaruhi oleh perkembangan teknologi informasi terutama oleh *game online*. Dibuktikan oleh sebuah penelitian yang menyatakan bahwa 38% responden termasuk dalam kategori tidak adiksi terhadap *game online*, dan 62% responden termasuk dalam kategori adiksi terhadap *game online*. Berdasarkan penelitian tersebut tingkat adiksi terhadap *game online* terjadi pada anak usia sekolah (Sanditaria, et al., 2014).

Kecanduan *game online* pada anak dapat menjadikan pengaruh buruk bagi perkembangan anak sehingga penting untuk mengetahui dampak buruk serta penanganan dari ketergantungan yang diakibatkan oleh *game online*. Banyaknya dampak buruk yang ditimbulkan akibat ketergantungan *game online* tidak dibarengi dengan banyaknya penanganan yang dapat dilakukan. Terbatas dan sulitnya menentukan penanganan itu sendiri karena didalamnya melibatkan atribut gejala yang saling berkaitan.

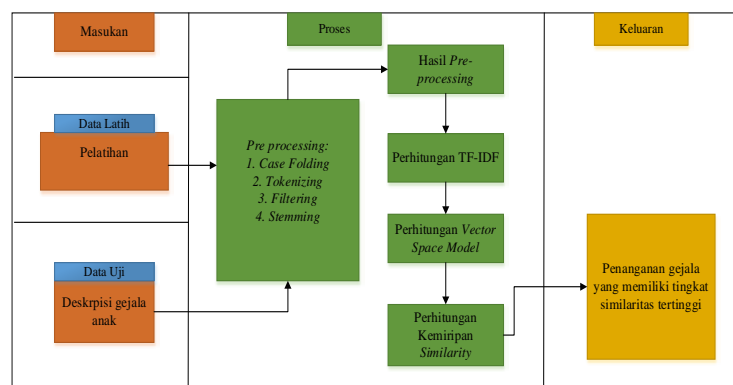
Menurut (Aziz, et al., 2015) tentang pembangkitan *Frequently Asked Questions* (FAQ) otomatis dan solusi yang relevan untuk keluhan pelanggan. Keluhan yang disampaikan pelanggan umumnya memiliki kesamaan dengan keluhan-keluhan sebelumnya, sehingga solusi dari keluhan baru dapat didasarkan pada solusi yang diberikan pada keluhan lama. Penelitian ini menggunakan algoritma TF-IDF dan *Vector Space Model* (VSM) dan kombinasi algoritma TF-IDF untuk penyelesaiannya dan menghasilkan perhitungan accuracy pada masing-masing percobaan, pada threshold 0.5 kombinasi notasi TF-IDF yang memiliki nilai rata-rata accuracy dan precision tertinggi

adalah modifikasi pertama, yaitu masing-masing sebesar 62.09% dan 55.15%. sedangkan untuk threshold 0.65 yang memiliki nilai rata-rata accuracy dan precision tertinggi adalah TF-IDF, yaitu masing-masing sebesar 83.18% dan 68.53%. Pada percobaan dengan menggunakan 171 data, TF-IDF dan threshold 0.65 dapat membangkitkan 27 FAQ, yaitu 70.37% relevan.

Penelitian ini bertujuan untuk membuat sistem dalam menentukan penanganan dampak *game online* dengan menggunakan metode *Vector Space Model* (VSM) pada setiap gejala yang menjadi data latih, sehingga pengguna mudah dalam menentukan penanganan yang dibutuhkan untuk proses penyembuhan.

2. METODOLOGI

Penelitian yang dilakukan yaitu membuat sistem yang mampu menentukan penanganan akibat *game online* pada anak menggunakan metode *Vector Space Model*. Data masukan yaitu deskripsi gejala yang dirasakan oleh anak dan keluarannya berupa penanganan terhadap gejala tersebut serta memiliki tingkat similaritas tertinggi terhadap gejala lain. Gambar 1 menunjukkan gambaran umum sistem yang akan dibangun.



Gambar 1. Gambaran Umum Sistem

2.1 Text Mining

Text mining merupakan salah satu bidang khusus dari data mining. *Text mining* dapat didefinisikan sebagai suatu proses menggali informasi dimana seseorang user berinteraksi dengan sekumpulan dokumen menggunakan tool analisis yang merupakan komponen-komponen dalam data mining (Wisnu & Hetami, 2015). *Text mining* digunakan untuk mengolah dokumen sebelum dilakukan proses similarity. Didalam proses *text mining* terdapat proses *preprocessing* (Asshidiq & M.E, 2013). *Preprocessing text* merupakan tindakan menghilangkan karakter-karakter tertentu yang terkandung dalam dokumen, seperti koma, tanda petik dan lain-lain serta mengubah semua huruf kapital menjadi huruf kecil. Selain itu, dalam tahap *text preprocessing* ini dilakukan *tokenization*. *Tokenization* merupakan proses pengolahan token yang terdapat dalam rangkaian teks, sehingga dokumen akan dipecah-pecah menjadi *term* (Aziz, et al., 2015).

Text mining dalam prakteknya mencari pola-pola tertentu, mengasosiasikan suatu bagian teks dengan yang lain berdasarkan aturan-aturan tertentu, kata-kata yang dapat mewakili sehingga dapat dilakukan analisa keterhubungan antar satu dengan yang lain (Wisnu & Hetami, 2015). Berikut tahapan-tahapan proses didalam *text mining*:

- Casefolding*: mengubah semua huruf dalam dokumen menjadi huruf kecil (*lowercase*).
- Tokenizing*: memotong tiap kata dalam kalimat atau parsing dengan menggunakan spasi sebagai delimiter yang akan menghasilkan token berupa kata.
- Filtering*: menyaring kata yang didapat dari *tokenizing* yang dianggap tidak penting atau tidak memiliki makna dalam proses *text mining* yang disebut *stopword*. *Stopword* berisi kata-kata umum yang sering muncul dalam sebuah dokumen dalam jumlah banyak namun tidak memiliki kaitan dengan tema tertentu.
- Stemming*: mengembalikan kata-kata yang diperoleh dari hasil *filtering* ke bentuk dasarnya, menghilangkan imbuhan awal (*prefix*) dan imbuhan akhir (*suffix*) sehingga didapat kata dasar.

2.2 Term Frequency-Inversed Document Frequency (TF/IDF)

Basis pembobotan TF-IDF merupakan jenis pembobotan yang melibatkan pengukuran statistik untuk mengukur seberapa penting sebuah kata dalam kumpulan dokumen. Tingkat kepentingan meningkat ketika sebuah kata muncul beberapa kali dalam sebuah dokumen tetapi diimbangi dengan frekuensi kemunculan kata tersebut dalam kumpulan dokumen (Wisnu & Hetami, 2015). TF merupakan pembobotan yang sederhana dimana penting tidaknya sebuah kata diasumsikan sebanding dengan jumlah kemunculan kata tersebut dalam dokumen, sementara IDF merupakan pembobotan yang mengukur seberapa penting sebuah kata dalam dokumen apabila dilihat secara global pada seluruh dokumen (M. Isa & Abidin, 2013). Perhitungan IDF menggunakan persamaan 1.

$$IDF_{(t)} = \log (D/df_{(t)}) \quad (1)$$

Dimana:

$df_{(t)}$ = Jumlah dokumen yang mengandung kata ke-t dari kata kunci

D = Jumlah semua dokumen yang ada di dalam database

IDF = Rasio frekuensi dokumen pada kata ke-t dari kata kunci

Perhitungan TF-IDF menggunakan persamaan 2

$$TF-IDF_{(d,t)} = TF_{(d,t)} * IDF_{(t)} \quad (2)$$

Dimana:

d = dokumen ke-d

t = kata ke-t dari kata kunci

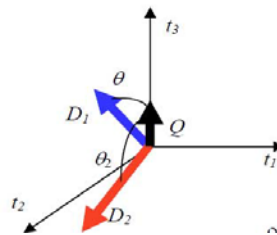
tf = frekuensi banyaknya kata ke-t dari kata kunci pada dokumen ke-d

TF-IDF = bobot dokumen ke-d terhadap kata kunci ke-t

IDF = rasio frekuensi dokumen pada kata ke-t dari kata kunci

2.3 Vector Space Model (VSM)

Vector Space Model (VSM) sering digunakan untuk mempresentasikan sebuah dokumen dalam ruang vector (P.D & Patrick, 2010). VSM merupakan model Information Retrieval yang mempresentasikan dokumen dan query sebagai vektor pada ruang multidimensi. Kesamaan suatu dokumen dengan query dapat diukur dengan vektor dokumen dan vektor query (Aziz, et al., 2015).



Gambar 2. Representasi Dokumen dan Query pada VSM (Mandala & Setiawan, 2002)

Perhitungan kemiripan antara vektor dokumen dan vektor query dilihat dari sudut yang paling kecil. Sudut yang dibentuk oleh dua buah vektor dapat dihitung dengan melakukan *inner product* (ER, et al., 2012). Kemiripan antara vektor dokumen dan vektor query akan dihitung dengan pendekatan *cosine similarity* (Aziz, et al., 2015). Pengukuran *Cosine Similarity* menggunakan persamaan 3.

$$Sim (d_i, d_j) = \frac{D_i D_j}{\|D_1\| \|D_2\|} = \frac{W_{iq} \cdot W_{ij}}{\sqrt{W_{iq}^2} * \sqrt{W_{ij}^2}} \quad (3)$$

Dimana:

$Sim (d_i, d_j)$ = similaritas antara query dan dokumen

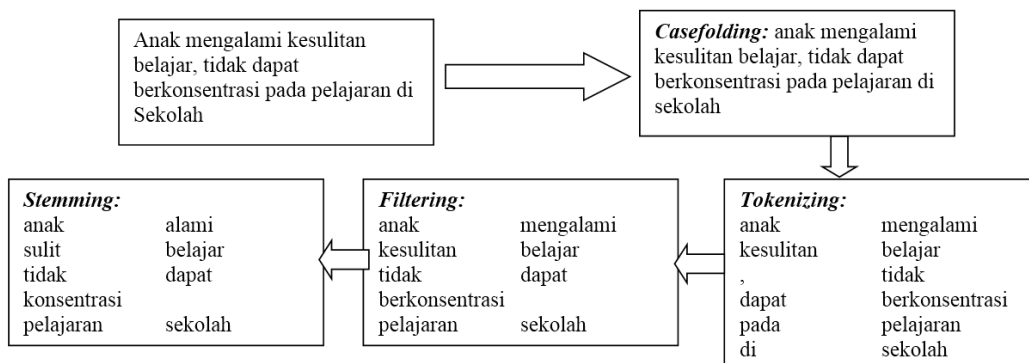
$\|D_1\|$ = panjang vektor dokumen 1

$\|D_2\|$ = panjang vektor dokumen 2
 W_{ij} = bobot term dalam dokumen
 W_{iq} = bobot query dalam dokumen

3. HASIL DAN PEMBAHASAN

3.1 Implementasi

Pada penelitian ini pembobotan terhadap data uji dilakukan dengan menggunakan metode *Term-Frequency – Inverse Document Frequency* (TF-IDF) dan dilanjutkan dengan metode *Vector Space Model* (VSM). Sebelum melakukan proses pembobotan, setiap data uji harus terlebih dahulu melalui tahap *preprocessing*, yang terdiri dari *casefolding*, *tokenizing*, *filtering*, dan *stemming*. Adapun proses *preprocessing* dapat dilihat pada gambar 3, perhitungan TF-IDF dilihat pada tabel 1, dan perhitungan VSM dilihat pada tabel 2.



Gambar 3. Tahap *Preprocessing*

Tabel 1. Perhitungan TF-IDF

Kata-kata Penting	TF								DF	D/DF	IDF	W									
	D1	D2	D3	D4	D5	D6	D7	D8				D1	D2	D3	D4	D5	D6	D7	D8		
sekolah	1	1	0	0	0	0	0	0	1	8	0.9031	0.9031	0	0	0	0	0	0			
sulit	1	0	0	1	1	0	0	1	0	3	2.6667	0.426	0	0	0.426	0.426	0	0	0.426	0	
tidak	1	0	1	0	0	0	0	0	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0	
alami	1	0	0	1	0	0	0	1	0	2	4	0.6021	0	0	0.6021	0	0	0	0	0.602	0
konsentrasi	1	0	0	0	0	0	0	1	0	1	8	0.9031	0	0	0	0	0	0	0	0.903	0
jika	0	0	1	0	0	0	0	0	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0	0
interaksi	0	0	0	1	0	0	0	0	0	1	8	0.9031	0	0	0.9031	0	0	0	0	0	0
main	0	0	1	0	0	0	0	0	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0	0
game	0	0	1	0	0	0	0	0	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0	0
dalam	0	0	1	0	0	0	0	1	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0.903	0
satu	0	0	1	0	0	0	0	0	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0	0
hari	0	0	1	0	0	0	0	0	0	1	8	0.9031	0	0.9031	0	0	0	0	0	0	0

Tabel 2. Perhitungan VSM

Kata-kata Penting	TF-IDF								VECTOR SPACE MODEL									
	D1	D2	D3	D4	D5	D6	D7	D8	D1	D2	D3	D4	D5	D6	D7	D8		
sekolah	1	0.816	0	0	0	0	0	0	0.8156	0	0	0	0	0	0	0		
sulit	1	0	0	0.181	0.181	0	0	0.181	0	0	0	0.1814	0.1814	0	0	0.1814	0	
tidak	1	0	0.816	0	0	0	0	0	0	0.8156	0	0	0	0	0	0	0	
alami	1	0	0	0.362	0	0	0	0.362	0	0	0	0.3625	0	0	0	0.3625	0	
konsentrasi	1	0	0	0	0	0	0	0.816	0	0	0	0	0	0	0	0.8156	0	
jika	0	0	0.816	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
interaksi	0	0	0	0.816	0	0	0	0	0	0	0	0	0	0	0	0	0	
main	0	0	0.816	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
game	0	0	0.816	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
dalam	0	0	0.816	0	0	0	0	0.816	0	0	0	0	0	0	0	0	0	
satu	0	0	0.816	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
hari	0	0	0.816	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Sqrt	2.24	1.41	2.70	1.59	1.95	1.56	1.53	1.95	1.64	0.903	0.903	0.736	0.425	0	0	1.165	0

Setelah melalui proses perhitungan TF-IDF dan VSM, data uji tersebut kemudian dihitung untuk menentukan bobot similaritas kedekatan antara data uji dengan data latih yang ada didalam sistem. Perhitungan *Cosine Similarity* dapat dilihat pada tabel 3.

Tabel 3. Perhitungan *Cosine Similarity*

Data Latih	<i>Cosine Similarity</i>	Rangking
Ke-1	$0,816 / 3,15 = 0,259 * 100\% = 25,9\%$	Rank 3
Ke-2	$0,816 / 6,04 = 0,135 * 100\% = 13,5\%$	Rank 4
Ke-3	$0,543 / 1,40 = 0,387 * 100\% = 38,7\%$	Rank 1
Ke-4	$0,181 / 4,36 = 0,041 * 100\% = 4,1\%$	Rank 5
Ke-5	$0 / 3,49 = 0 * 100\% = 0$	Rank 6
Ke-6	$0 / 3,42 = 0 * 100\% = 0$	Rank 7
Ke-7	$1,395 / 4,36 = 0,319 * 100\% = 31,9\%$	Rank 2
Ke-8	$0 / 3,67 = 0 * 100\% = 0$	Rank 8

Dari hasil akhir tersebut dapat diketahui bahwa D3 memiliki tingkat similaritas tertinggi terhadap data uji (DU) yaitu 38,7%.

3.2 Hasil Pengujian Akurasi

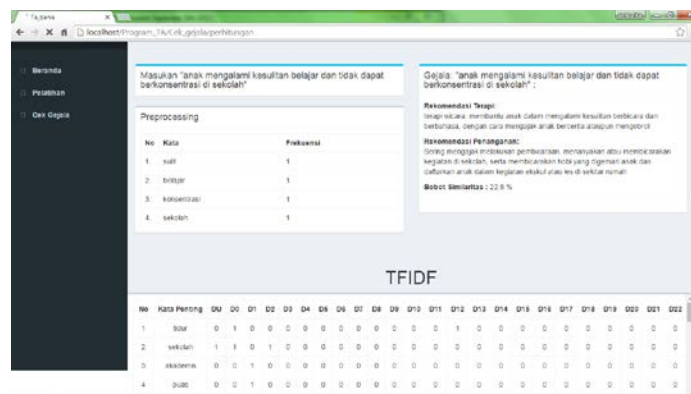
Dalam tahap ini dilakukan pengujian akurasi sistem, dimana pengujian sistem yang dihasilkan akan disamakan dengan masukan gejala. Pengujian dilakukan dengan mengambil 10 sampel masukan berupa teks yang berbeda tetapi memiliki makna yang sama, dan akan mengeluarkan hasil yang serupa. Hasil pengujian akurasi dapat dilihat pada tabel 4.

Tabel 4. Pengujian akurasi

No	Data Uji	Tepat Dikenali	Bobot Similaritas
1	Data Uji ke-1	Ya	31,6%
2	Data Uji ke-2	Ya	57,7%
3	Data Uji ke-3	Ya	27,7%
4	Data Uji ke-4	Ya	27,4%
5	Data Uji ke-5	Ya	57,7%
6	Data Uji ke-6	Ya	44,7%
7	Data Uji ke-7	Tidak	50%
8	Data Uji ke-8	Ya	28,9%
9	Data Uji ke-9	Ya	27,7%
10	Data Uji ke-10	Tidak	70,7%

Dari tabel 4 memperlihatkan hasil pengujian akurasi bahwa dengan data uji gejala berupa teks yang berbeda tapi memiliki makna yang sama, sistem dapat menentukan dengan ketepatan 80%, dimana ditemukan 2 masukan gejala yang tidak sesuai dari 10 masukan gejala yang di masukan. Berdasarkan 8 masukan gejala yang menghasilkan rata-rata bobot similaritas sebesar 37,925%.

Perangkat lunak telah diimplementasikan dan diuji kemampuannya untuk menentukan penanganan terhadap gejala yang dirasakan oleh anak akibat *game online*. Gambar 4 memperlihatkan hasil perangkat lunak.



Gambar 4. Hasil Perangkat Lunak

4. KESIMPULAN

Dari hasil penelitian ini, maka diambil kesimpulan bahwa metode *Vector Space Model* (VSM) dapat membantu pengguna dalam menentukan penanganan yang dijadikan sebagai tahap awal dalam proses penyembuhan pada anak akibat *game online*.

Hasil pengujian akurasi dengan masukan gejala berupa teks yang berbeda tapi memiliki makna yang sama, sistem ini dapat menentukan dengan ketepatan 80%, dimana ditemukan 2 masukan gejala yang tidak sesuai dari 10 masukan gejala yang di masukan. Dari 10 gejala yang menjadi contoh masukan oleh sistem, 8 masukan gejala termasuk menghasilkan rekomendasi sesuai sehingga diperoleh rata-rata bobot similaritas sebesar 37,925%.

DAFTAR PUSTAKA

- Asshidiq, A., Saptono, R. & Sulisty, M., 2013. Penilaian Ujian Bertipe Essay Menggunakan Metode Text Similarity. *Jurnal Informatika*.
- Aziz, A., Saptono, R. & P.S, K., 2015. Implementasi Vector Space Model dalam Pembangkitan Frequently Asked Questions dan Solusi yang Relevan Keluhan Pelanggan. *Scientific Journal Informatics*, 2(2).
- ER, N. A. S., Muliantara, A. & Widiartha, I. M., 2012. Peningkatan Relevansi Hasil Pencarian Kata Kunci dengan Penerapan Model Ruang Vektor pada Sistem Informasi Ruang Baca di Jurusan Ilmu Komputer Universitas Udayana. *Ilmu Komputer*, 5(2), pp. 36-42.
- M.Isa, T. & Abidin, T. F., 2013. Mengukur Tingkat Kesamaan Paragraf menggunakan Vector Space Model untuk Mendeteksi Plagiarisme. *Seminar Nasional dan Expo Teknik Elektro*.
- Sanditaria, W., Fitri Rahayu, S. Y. & Mardiyah, A., 2014. Adiksi Bermain Game Online pada Anak Usia Sekolah di Warung Internet Penyedia Game Online Jatinangor Sumedang. *Jurnal Kesehatan*.
- Wisnu, D. & Hetami, A., 2015. Perancangan Informtaion Retrieval (IR) untuk Pencarian Ide Pokok Teks Artikel Berbahasa Inggris dengan Pembobotan Vector Space Model. *Jurnal Ilmiah Teknologi dan Informasi*, Volume 9.