

## KLASIFIKASI SOAL PILIHAN GANDA BERBAHASA INDONESIA BERDASARKAN LEVEL KOGNITIF PUSPENDIK DENGAN SUPPORT VECTOR MACHINE

Lilis Setyaningsih<sup>1)</sup>, Endang Setyati<sup>2)</sup>, Suhatati Tjandra<sup>3)</sup>  
<sup>1,2,3</sup>Teknologi Informasi, Institut Sains dan Teknologi Terpadu Surabaya  
Jl. Ngagel Jaya Tengah No. 73-77, Baratajaya, Gubeng, Surabaya  
<sup>1</sup>liestyaaq@gmail.com,<sup>2</sup>endang@stts.edu,<sup>3</sup>tati@stts.edu

### ABSTRAK

Bank soal merupakan kebutuhan penting di sekolah, sebagai media evaluasi pembelajaran. Bank soal yang berkualitas, dapat membantu dalam menganalisa kemampuan kognitif peserta didik. Pada akhirnya meningkatkan kualitas pembelajaran di sekolah. Penelitian ini menggunakan dataset berupa pokok soal dari soal PG, yang di dalamnya terdapat stimulus yang berupa text. Soal akan diklasifikasi berdasarkan level kognitif Puspendik, yaitu level pengetahuan dan pemahaman (L1), level aplikasi (L2) dan level penalaran atau HOTS (L3). Tidak semua soal menggunakan kata kerja operasional. Proses penelitian dilakukan dalam 3 tahap, yaitu pengumpulan dan labeling dataset, preprocessing, training model, dan testing data. Penelitian dilakukan dengan *Bag Of Words*, *TF-IDF*, *Latent Semantic Analysis* dan metode *Support Vector Machine*. Hasil uji coba menunjukkan tingkat akurasi dalam mengklasifikasi soal L1 sebesar 64,67%, L2 75,42% dan L3 25,83%. Tingkat akurasi secara keseluruhan dicapai, pada saat ujicoba 30% data testing, yaitu 61,54%.

Kata kunci : Klasifikasi Soal Pilihan Ganda, Level Kognitif Puspendik, Latent Semantic Analysis, Support Vector Machine.

### ABSTRACT

*A question bank is an important need in schools as a medium for evaluating learning. A quality question bank, can help in analyzing the cognitive abilities of students. In the end, it will improve the quality of learning in schools. This study uses a dataset in the form of subject matter from PG questions, in which there is a stimulus in the form of text. Questions will be classified based on Puspendik's cognitive level, namely the level of knowledge and understanding (L1), the level of application (L2) and the level of reasoning or HOTS (L3). Not all questions use operational verbs. The research process was carried out in 3 stages, namely collecting and labeling datasets, preprocessing, training models, and testing data. The research was conducted with Bag of Words, TF-IDF, Latent Semantic Analysis and Support Vector Machine methods. The test results showed that the level of accuracy in classifying L1 questions was 64.67%, L2 75.42% and L3 25.83%. The overall level of accuracy was achieved, at the time of testing 30% of the testing data was 61.54%.*

**Keywords:** *Multiple Choice Question Classification, Puspendik Cognitive Level, Latent Semantic Analysis, Support Vector Machine.*

### PENDAHULUAN

Evaluasi pembelajaran merupakan salah bentuk pengukuran tingkat kognitif peserta didik. Ranah kognitif berkaitan dengan kemampuan berpikir, kemampuan memahami, mengimplementasi dan melakukan penalaran. Evaluasi pembelajaran akhir semester atau ujian sekolah biasanya

dilaksanakan dengan menggunakan soal bertipe pilihan ganda (PG).

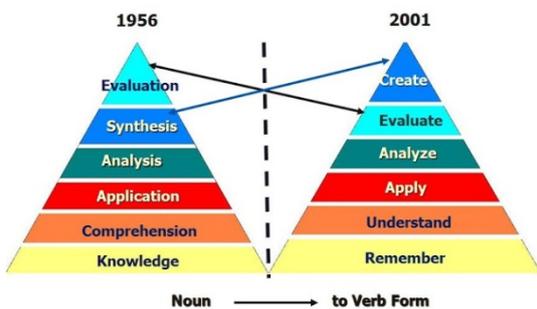
Pada penelitian ini diharapkan mampu mengklasifikasikan soal PG, sehingga sekolah mempunyai bank soal yang berkualitas agar dapat mengukur tingkat kognitif peserta didik.

Dataset menggunakan soal PG yang digunakan di tingkat X SMK. Bagian soal yang digunakan hanya bagian pokok soal (*stem*) tanpa option jawaban. Stimulus soal yang digunakan, hanya stimulus soal yang berupa text.

Klasifikasi soal akan menghasilkan 3 kelas, yaitu L1, L2 dan L3. Klasifikasi dilakukan dengan menggunakan metode Support Vector Machine, dan 3 kernelnya yaitu Linear, Poly dan RBF.

**Taksonomi Bloom Revisi**

Taksonomi Bloom (Revisi) membagi kognitif dalam 6 level, yaitu (C1) *remember* atau mengingat, (C2) *understand* atau memahami, (C3) *apply* atau menerapkan, (C4) *analyze* atau menganalisa, (C5) *evaluate* atau mengevaluasi, dan (C6) *create* atau mencipta (Kusuma, S. F. 2015)



**Gambar 1.**Level Kognitif Taksonomi Bloom Revisi

<https://dreamlikechild.weebly.com/uploads/3/9/0/4/39042447/827490737.jpg>

Pada Taksonomi Bloom Revisi, terdapat pergantian dari penggunaan kata benda menjadi kata kerja yang mampu menggambarkan proses kognitif.

**Level Kognitif Puspendik**

Puspendik (Pusat Penilaian Pendidikan Kemendikbud RI) sejak tahun 2015 telah membagi level kognitif menjadi 3 level, yaitu level pengetahuan dan pemahaman atau L1 (C1,C2), level aplikasi atau L2 (C3) dan level penalaran atau L3 (C4, C5, C6 atau HOTS) (Asmira, O. 2018)

Pada setiap level tersebut mempunyai ciri-ciri katakerja operasional yang sering digunakan seperti pada tabel 1.

**Tabel 1.**Contoh KKO Level Kognitif Puspendik

Level I	KKO
L1	Mengutip Menyebutkan Menjelaskan Membilang Mengidentifikasi Membaca Menamai Menandai Menghafal Meniru Mencatat Mengulang Meninjau Menulis Menyatakan Memperkirakan Menceritakan Mendiskusikan Menerangkan Membedakan
L2	Mengurutkan Menerapkan Mengkalkulasi Memodifikasi Menghitung Mencegah Menentukan Menggambarkan Menggunakan Mengemukakan Mengadaptasi Mempersoalkan Melaksanakan Melakukan Mensimulasikan Mentabulasi Mengklasifikasi Mengoperasikan Meramalkan
L3	Mengaudit Mengatur Memecahkan Menegaskan Menganalisis Merinci Menominasikan Mendiagramkan Mengkorelasikan Menguji Memerintahkan Mengaitkan Menransfer Melatih Mengedit Menemukan Menyeleksi Mendeteksi Menelaah Membangunkan Merancang Membatas Mereparasi Mengkonstruksi Merumuskan Menghubungkan Menciptakan Menampilkan

**Soal Pilihan Ganda**

Soal pilihan ganda terdiri dari bagian pokok soal (*stem*) dan option jawaban. Pada pokok soal terdapat stimulus dan pertanyaan/pernyataan tidak lengkap (Puspendik. 2019). Lihat gambar 2 di bawah ini, bagian-bagian dari pokok soal pilihan ganda.

**Contoh Soal Pilihan Ganda**

1. Tono, Budi, dan Andi membeli alat-alat tulis yang sama di sebuah toko. Tono membeli dua buku tulis, satu pena, dan satu pensil dengan membayar Rp11.000,00, Budi membeli satu buku tulis, satu pena dan satu pensil dengan harga Rp7.000,00, Andi membeli tiga buku tulis dan dua pena, ia membayar seharga Rp16.000,00.

Jika banyak buku tulis, pena, dan pensil yang mereka beli berturut-turut dinyatakan dengan  $x$ ,  $y$ , dan  $z$ , model matematika dari permasalahan tersebut adalah ...

A.  $2x + y + z = 11.000; x + y + 2z = 7.000; 3x + 2y = 16.000$

B.  $2x + y + z = 11.000; x + y + z = 7.000; 3x + y = 16.000$

C.  $2x + y + z = 11.000; x + y + z = 7.000; 3x + 2y = 16.000$

D.  $2x + y + z = 11.000; x + y + z = 7.000; 2x + 3y = 16.000$

E.  $x + 2y + z = 11.000; x + y + z = 7.000; 3x + 2y = 16.000$

Labels in image: Pokok Soal (Stem), Stimulus, Pertanyaan/pernyataan tidak lengkap, Kunci Jawaban, Pengecoh.

**Gambar 2.** Bagian Pokok Soal Pilihan Ganda

**Pembobotan Term dengan TF-IDF**

TF-IDF merupakan suatu algoritma yang digunakan dalam pembobotan terhadap kemunculan suatu kata (token) yang dijadikan sebagai fitur dalam suatu dokumen seleksi (Sudin, S. 2018)

Langkah perhitungan TF-IDF

1. Menghitung Term Frequency, yaitu frekwensi kemunculan term dalam sebuah dokumen

$$tf(t, d) = \left( \frac{\text{jumlah kemunculan term pada satu dokumen}}{\text{jumlah seluruh term dalam satu dokumen}} \right) \dots (1)$$

2. Menghitung Nilai *idf*. *Idf* adalah hubungan antara banyak jumlah kata dengan dokumen

$$idf(t, D) = \log \left( \frac{\text{jumlah seluruh dokumen}}{\text{jumlah dokumen suatu term muncul}} \right) \dots (2)$$

3. Menghitung bobot (w) kata dalam dokumen.

$$W = (tf * idf) \dots (3)$$

**Latent Semantic Analysis**

Latent Semantic Analysis (LSA) merupakan sebuah metode yang memanfaatkan model statistik matematis untuk menganalisa struktur semantik suatu teks (Suhartono, D. 21015). LSA menerapkan matriks SVD (*Singular Value Decomposition*), untuk mereduksi dimensi nilai kompleksitas dalam pemrosesan term dokument matriks.

SVD merupakan teorema aljabar linier yang menyebutkan bahwa persegi panjang dari term-document matrix dapat dipecah/didekomposisikan menjadi tiga matriks, yaitu : matriks ortogonal U, matriks diagonal S, dan transpose dari matriks ortogonal V.

$$A_{mn} = U_{mm} \times S_{mn} \times V_{mn}^T \dots (4)$$

$A_{mn}$  = Matrik awal

$U_{mn}$  = Matrik Ortogonal U

$S_{mn}$  = Matrik Diagonal S

$V_{mn}^T$  = Transpose Matrik Ortogonal V

Cosine similarity digunakan untuk menghitung nilai kosinus sudut antara vektor dokumen dengan vektor kueri.

Semakin kecil sudut yang dihasilkan, maka tingkat kemiripan kata semakin tinggi.

$$\cos \alpha = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \dots (5)$$

A = vektor dokumen

B = vektor kueri

A . B = perkalian dot vektor A dan vektor B

|A| = panjang vektor A

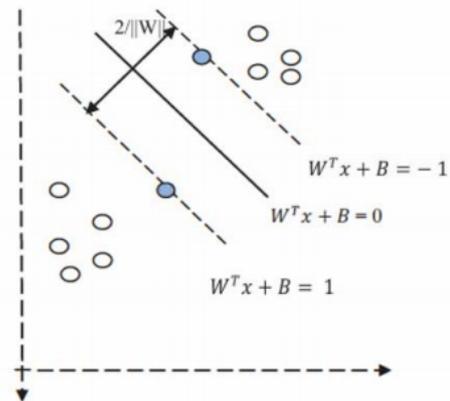
|B| = panjang vektor B

|A| |B| = cross product antara |A| dan |B|

$\alpha$  = sudut yang terbentuk antara vektor A dan B

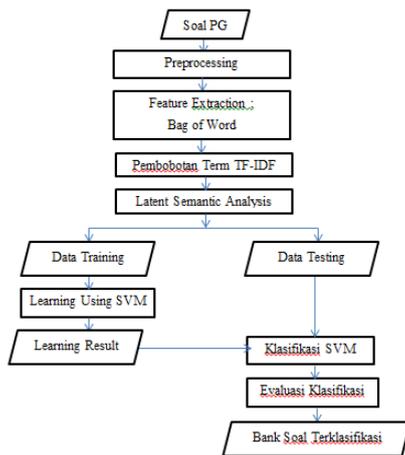
**Support Vector Machine**

SVM (Support Vector Machine) adalah metode klasifikasi untuk data linier dan nonlinier. Pemetaan nonlinear untuk mengubah data pelatihan menjadi dimensi yang lebih tinggi, dan kemudian menemukan hyperplane optimal linier untuk pemisahan kategori (Valtemir, 2018).



Gambar 3. Proses Klasifikasi SVM

**METODE PENELITIAN**



**Gambar 4.** Alur Sistem

Pada penelitian ini dilakukan tahapan-tahapan proses seperti terlihat pada gambar 4 di bawah ini :

**Dataset Penelitian**

Soal PG dikumpulkan dari bank soal sekolah dan buku pelajaran. Soal-soal tersebut diklasifikasi secara manual berdasarkan KKO (L1, L2 dan L3) oleh Musyawarah Guru Mata Pelajaran (MGMP) Sekolah yang terdiri dari 2 guru yang mempunyai pengalaman mengajar lebih dari 10 tahun serta divalidasi oleh Tim Kurikulum.

**Tabel 2.** Contoh Dataset

Kelas	Mapel	Soal
L1	INDO	Asap kendaraan bermotor yang berwarna putih disebabkan adanya kebocoran dibagian ring seker. Jika asapnya hitam berarti terlalu banyak campuran bensin daripada udara di dalam karburator. Laporan di atas menjelaskan tentang jenis laporan ....
L2	SJRH	Adakalanya sejarah menggunakan ilmu sosial dan sebaliknya ilmu sosial menggunakan sejarah. Artinya, terjadi persilangan antara konsep diakronik dan sinkronik. Penerapan persilangan Konsep tersebut dapat dilakukan pada pilihan ...
L3	PKN	Perhatikan karakteristik berikut! 1) Daerah yang berbatasan dengan laut bebas. 2) Wilayah perairan ini diukur sejauh 200

Kelas	Mapel	Soal
		mil dari garis pantai yang paling jauh menjorok ke laut (garis dasar). 3) Pada daerah ini Negara pantai memiliki kedaulatan dan prioritas utama dalam mengelola sumber daya alam yang terkandung di dalamnya. Berdasarkan karakteristik tersebut, dapat disimpulkan bahwa wilayah yang dimaksud adalah ...

Soal PG terdiri dari 3 mata pelajaran yaitu Bahasa Indonesia (INDO), Pendidikan Pancasila dan Kewarganegaraan (PKN), dan Sejarah Indonesia (SJRH). Jumlah dataset 135 soal, yang terdiri dari 45 soal level pengetahuan dan pemahaman (L1), 45 soal level aplikasi (L2) dan 45 soal level penalaran (L3). Jumlah soal per mapel adalah 45 soal.

**Preprocessing**

Pada Text Mining, preprocessing data digunakan untuk mengekstraksi fitur-fitur menarik dan pengetahuan dari data teks yang tidak terstruktur. Text data biasanya mengandung beberapa format seperti angka, tanggal, kata-kata umum yang tidak digunakan pada Text Mining seperti preposisi, pronominal dapat dihilangkan. (Kannan, S. 2014)

Berikut adalah penjelasan dari alur preprocessing yang terdapat pada gambar 4:

- a. *Cleaning data*  
Pada tahap ini dilakukan penghapusan URL, spasi ganda, angka dan tanda baca (*Punctuation*) kecuali tanda hubung (-). Kata berulang seperti manfaat-manfaat, akan diperlakukan sebagai satu kata.
- b. *Casefolding*  
Tahapan mengubah huruf besar menjadi huruf kecil. Proses casefolding akan mempermudah pencarian kata yang sama.
- c. *Tokenisasi*  
Tokenisasi adalah langkah merubah kalimat menjadi dalam bentuk kata (*token*). Setiap kata dipisahkan dengan spasi (*white space*).
- d. *Pos tagging*

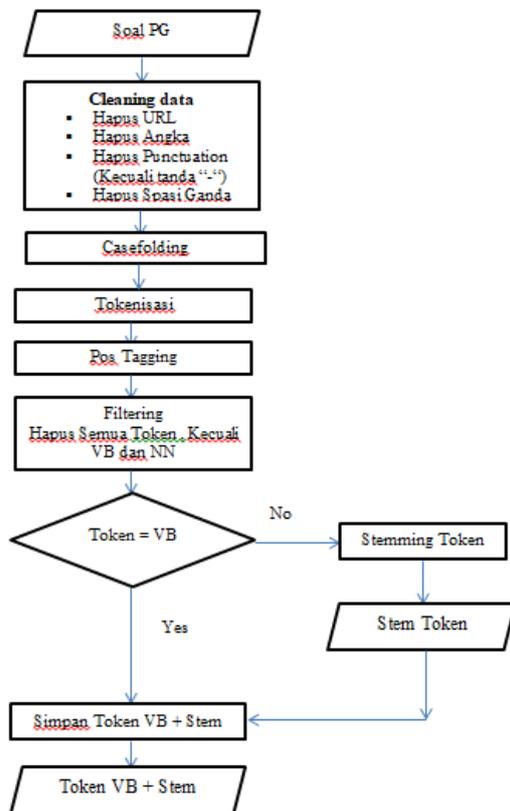
Proses pos tagging merupakan proses pemberian label jenis kata. Contoh : “VB” = kata kerja, “NN” = kata benda, “ADV” = Adverb, IN = preposisi, dan lain-lain. Proses ini diperlukan untuk proses filtering.

e. *Filtering/ Stopword*

Filtering dilakukan untuk menghapus kata-kata yang tidak mempengaruhi klasifikasi. Klasifikasi ini berkaitan dengan penggunaan KKO dan kata yang terkait dengannya, maka seluruh jenis kata kecuali VB dan NN akan dihapus.

f. *Stemming*

Proses stemming akan merubah kata ke dalam bentuk kata dasar (Silva, 2018). Contoh : “penjelasan” setelah di stemming akan menjadi kata “jelas”. Kata hasil proses stemming akan di simpan dan digabung dengan kata kerja.



Gambar 5. Alur Preprocessing

Pada tabel 3, dapat dilihat hasil dari proses preprocessing.

Tabel 3. Hasil Preprocessing

Soal	Hasil Preprocessing
Asap kendaraan bermotor yang berwarna putih disebabkan adanya kebocoran dibagian ring seker. Jika asapnya hitam berarti terlalu banyak campuran bensin daripada udara di dalam karburator. Laporan di atas menjelaskan tentang jenis laporan ....	[berwarna, disebabkan, berarti, menjelaskan, asap, kendara, motor, ada, bocor, bagi, campur, bensin, udara, dalam, karburator, lapor, atas, jenis, lapor]
Adakalanya sejarah menggunakan ilmu sosial dan sebaliknya ilmu sosial menggunakan sejarah. Artinya, terjadi persilangan antara konsep diakronik dan sinkronik. Penerapan persilangan Konsep tersebut dapat dilakukan pada pilihan ...	[menggunakan, ilmu, menggunakan, terjadi, dilakukan, sejarah, ilmu, sejarah, silang, konsep, sinkronik, terap, silang, konsep, pilih]
Perhatikan karakteristik berikut! 1) Daerah yang berbatasan dengan laut bebas. 2) Wilayah perairan ini diukur sejauh 200 mil dari garis pantai yang paling jauh menjorok ke laut (garis dasar). 3) Pada daerah ini Negara pantai memiliki kedaulatan dan prioritas utama dalam mengelola sumber daya alam yang terkandung di dalamnya. Berdasarkan karakteristik tersebut, dapat disimpulkan bahwa wilayah yang dimaksud adalah...	[perhatikan, berbatasan, diukur, mil, menjorok, memiliki, terkandung, berdasarkan, disimpulkan, dimaksud, adalah, karakteristik, ikut, daerah, laut, wilayah, air, garis, pantai, laut, garis, dasar, daerah, negara, pantai, daulat, prioritas, kelola, sumber, daya, alam, karakteristik, wilayah]

**Bag of Words**

Bag of word merupakan kumpulan kantung (*bag*) kata yang mewakili dokumen tanpa memandang urutan kata dan tata bahasa tapi tetap mempertahankan keberagamannya (Putri, W.T. H.

2018). Hasil Bag of Word merupakan gabungan dari seluruh hasil preprocessing.

**Tabel 4.** Hasil Bag of Words

Bag of Words	
berwarna, disebabkan, berarti, menjelaskan, asap, kendaraan, motor, ada, bocor, bagi, campur, bensin, udara, dalam, karburator, lapor, atas, jenis, lapor, menggunakan, ilmu, menggunakan, terjadi, dilakukan, sejarah, ilmu, sejarah, silang, konsep, sinkronik, terap, silang, konsep, pilih, perhatikan, berbatasan, diukur, mil, menjorok, memiliki, terkandung, berdasarkan, disimpulkan, dimaksud, adalah, karakteristik, ikut, daerah, laut, wilayah, air, garis, pantai, laut, garis, dasar, daerah, negara, pantai, daulat, prioritas, kelola, sumber, daya, alam, karakteristik, wilayah	

**Pembobotan Term TF-IDF**

Pembobotan term menghasilkan tabel bobot masing-masing term, seperti pada tabel 5.

**Tabel 5.** Pembobotan Term

	abad	abang	abdul	abris	abstrak	abstraksi	acara	ada	adalah
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.051495
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.354163	0.205586
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.186389
...	...	...	...	...	...	...	...	...	...
130	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000
131	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.040033
132	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.076183
133	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.068373
134	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000

**Latent Semantic Anlysis**

Pada tahap ini dilakukan pengelempokan term yang mempunyai frekwensi tertinggi di masing-masing level soal. Pada tabel 6, merupakan contoh hasil pemrosesan LSA pada soal L1.

**Tabel 6.** Hasil LSA pada L1

Hasil LSA pada L1
sejarah, peristiwa, lembaga, perintah, adalah, kata, nonkementrian, disebut, teks, erti, bagi, manusia, melaksanakan, negara, tugas, berasal, hasil, contoh, kelompok, hidup, memiliki, merupakan, presiden, penggunaanya, ditunjukkan, berbeda, dibentuk, observasi, masa, berarti, syajaratun, pembabakan, klasifikasi, pohon, berdasarkan, eksposisi, selenggara, waktu, dilakukan, definisi, sebu, gambar, lapor, orang, struktur, dikategorikan, bidang, nilai, paragraf, indonesia, tuju, isi, dicontohkan, berulang, pola, aspek, jalan, utuh, konsep, memberikan, arti, pilih, pemerintah,

nonkemetrian, arsip, awas, sifat, purba, mengubah, kepala,

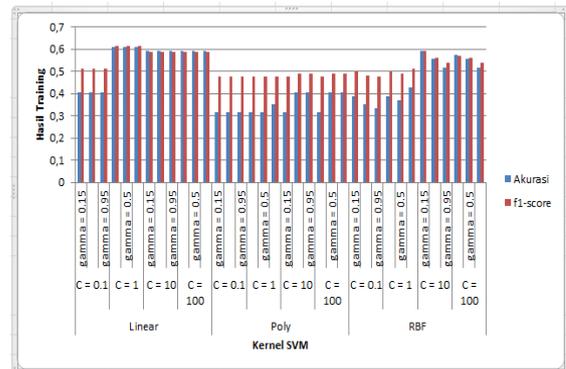
Perhitungan matriks SVD dan Cosine Similarity, seperti pada tabel 7.

**Tabel 7.** Matriks Cosine Similarity

	0	1	2	3	4
0	1.000000	0.000000	0.000000	0.000000	0.000000
1	0.000000	1.000000	0.065798	0.020031	0.036194
2	0.000000	0.065798	1.000000	0.079971	0.144498
3	0.000000	0.020031	0.079971	1.000000	0.479051
4	0.000000	0.036194	0.144498	0.479051	1.000000

**Training Model SVM**

Data training yang digunakan dalam ujicoba sebesar 60% dari dataset. Klasifikasi dilakukan dengan menggunakan metode Support Vector Machine dengan 3 kernel, yaitu Linier, Poly dan RBF. Nilai parameter C yang digunakan dalam ujicoba dalam range 0.1; 1; 10; 100, sedangkan nilai  $\gamma$  dalam range 0.15; 0.5 ; 0.95. Hasil training dapat dilihat pada gambar 6.



**Gambar 6.** Grafik Hasil Training Model SVM

Hasil training dengan menggunakan 60% data training, menghasilkan nilai akurasi paling tinggi di kernel linier dengan nilai C = 1, yaitu dengan akurasi 61,11% dan f1-score 61,52%.

**HASIL DAN PEMBAHASAN**

Model terbaik yang dihasilkan dari data training, akan diterapkan dalam pengujian data testing. Uji coba dilakukan dengan berulang menggunakan data test 20%, 30% dan 40 %, dengan nilai C =1 pada kernel linear.

Tabel 8 menjelaskan tentang jumlah soal berdasarkan mata pelajaran yang digunakan sebagai data test. Penggunaan data test ini dilakukan secara acak.

**Tabel 8.** Jumlah Data Testing Berdasarkan Mata Pelajaran

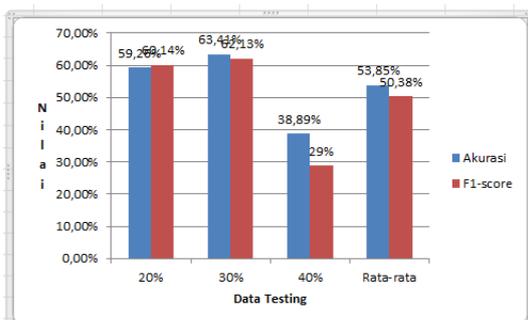
Data Uji	Mata Pelajaran		
	Indo	Sejarah	PKN
20%	8	9	10
30%	13	13	15
40%	13	22	19

Jumlah Data testing berdasarkan Level soal, terlihat pada Tabel 9.

**Tabel 9.** Jumlah Data Testing Berdasarkan Level Soal

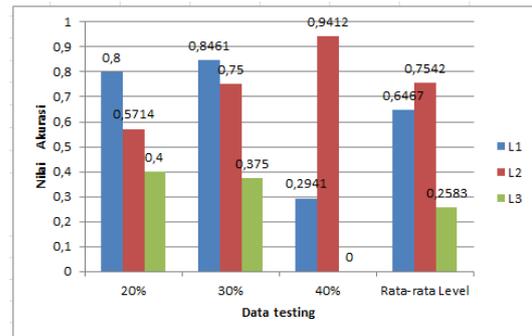
Data Uji	Level		
	L1	L2	L3
20%	10	7	10
30%	13	12	16
40%	17	17	20

Setelah diuji coba dengan kernel linear dan nilai C=1, menghasilkan nilai akurasi dan f1-score seperti pada grafik di bawah ini.



**Gambar 7.** Grafik Hasil Testing dengan Kernel Linear

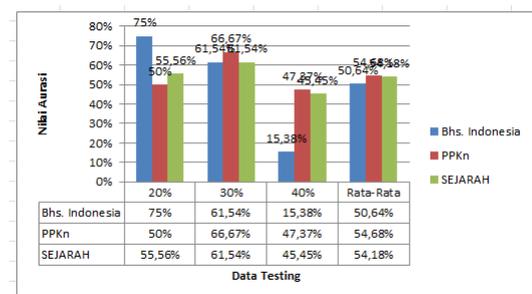
Nilai akurasi dan *f1-score* tertinggi, diperoleh saat ujicoba data testing 30%, yaitu sebesar 63,41% dan 62,13%. Nilai rata-rata dari ujicoba data testing diperoleh nilai akurasi 53,85% dan *f1-score* 50,38%.



**Gambar 8.** Grafik Hasil Testing Berdasarkan Level Soal

Pada gambar 8, nilai akurasi tertinggi berdasarkan level soal pada L1, bernilai 84,61% saat uji coba data testing 30%. L2 diperoleh nilai akurasi 94,12% saat ujicoba data testing 40%. L3 mempunyai nilai akurasi paling rendah, saat data uji 40% karena tidak ada satupun soal yang bisa terklasifikasi. Akhirnya dari 3 kali ujicoba, rata-rata akurasi L1 adalah 64,67%, L2 75,42% dan L3 25,83%.

Analisa lebih lanjut akurasi hasil ujicoba data testing berdasarkan mata pelajaran, seperti data pada grafik di bawah ini.

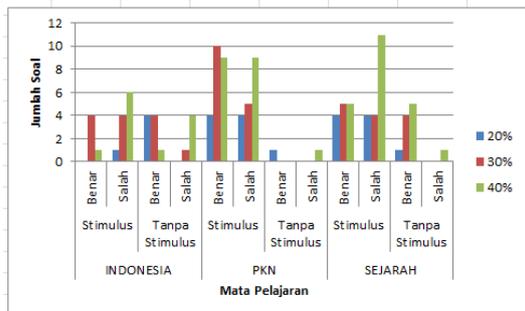


**Gambar 9.** Grafik Hasil Testing berdasarkan Mata Pelajaran

Soal Bahasa Indonesia pada saat data testing 20%, terdapat 5 soal yang diuji, dimana 4 soal tanpa ada stimulus terklasifikasi dan 1 soal dengan stimulus yang tidak terklasifikasi dengan benar.

Pada testing 30%, terdapat 8 soal dengan stimulus, dan 5 soal tanpa stimulus. Soal Bahasa Indonesia yang berstimulus mempunyai akurasi yang sama antara yang terklasifikasi dan yang tidak terklasifikasi dengan benar. Sedangkan yang tanpa stimulus, 4 soal diantaranya terklasifikasi dengan benar.

Pada saat testing 40%, soal bahasa indonesia yang diuji lebih banyak dari kedua level sebelumnya. Jumlah soal yang berstimulus 7 soal, 6 soal tidak terklasifikasi dengan benar. Sedangkan soal yang tanpa stimulus, 4 soal tidak terklasifikasi dengan benar. Lihat grafik di gambar 10, analisa soal mata pelajaran terkait ada atau tidaknya stimulus, sehingga bisa terklasifikasi dengan benar atau tidak.



Gambar 10. Analisa Soal Mata Pelajaran Berdasarkan Stimulus

Pada grafik di atas terlihat pada soal-soal yang berstimulus, terdapat banyak kesalahan prediksi, dan soal yang tanpa stimulus tingkat kesalahannya jauh lebih kecil.

## PENUTUP

Pada penelitian ini klasifikasi soal PG berdasarkan Level Puspendik mempunyai tingkat akurasi L1 64,67%, L2 75,42% dan L3 25,83% dengan menggunakan Support Vector Machine. Tingkat Akurasi klasifikasi paling tinggi pada saat uji coba data testing 30%, yaitu 65,70%. Keberadaan stimulus sangat mempengaruhi hasil prediksi soal.

Berdasarkan analisa hasil testing dan analisa soal, perlu dilakukan penelitian tentang pengaruh stimulus pada klasifikasi soal PG dan penambahan feature extraction lainnya sehingga hasil klasifikasi lebih maksimal.

## DAFTAR PUSTAKA

Asmira, O., Ariyana, Y., Bestary, R., Pudjiastuti, A., Setiawati, W. 2018.

*Buku Pegangan Pembelajaran Berorientasi pada Keterampilan Berpikir Tingkat Tinggi*. Direktorat Jendral Guru dan Tenaga Kependidikan dan Kebudayaan. Jakarta.

Kannan, S., Gurusamy, V. 2014. *Preprocessing Techniques for Text Mining*. [https://www.researchgate.net/publication/273127322\\_Preprocessing\\_Techniques\\_for\\_Text\\_Mining](https://www.researchgate.net/publication/273127322_Preprocessing_Techniques_for_Text_Mining). Diakses tanggal 14 September 2020

Kusuma, S. F., Siahaan, D., Yuhana, U. L. 2015. *Automatic Indonesia's Questions Classification Based On Bloom's Taxonomy Using Natural Language Processing*. Bandung :International Conference on Information Technology Systems and Innovation (ICITSI). Hal 1-6.

Puspendik. 2019. *Penulisan Soal Pilihan Ganda*. <https://www.slideshare.net/MuslihatunSyarifah/ppt-penulisan-soal-pilihan-ganda>. Diakses tanggal 16 September 2020

Putri, W. T. H., Hendrowati, R. 2018. *Penggalan Teks Dengan Model Bag Of Words terhadap Data Twitter*. Jurnal Muara Sains, Teknologi, Kedokteran, dan Ilmu Kesehatan. Vol 2 No 1. Hal .129-138.

Suhartono, D. 2015. *Penggunaan Latent Semantic Analysis dalam Pemrosesan Text*. <https://socs.binus.ac.id/2015/08/03/penggunaan-latent-semantic-analysis-lsa-dalam-pemrosesan-teks/>. Diakses tanggal 20 September 2020.

Sudin, Sakina. 2018. Analisis Jenis Pertanyaan Berbahasa Indonesia pada Question and Answering System Menggunakan Metode Support Vector Machine (SVM). Jurnal Ilmiah Teknologi Sistem Informasi

Valtemir A. Silva, Ig I. Bittencourt, and José C. Maldonado. 2018. *Automatic Question Classifiers: a Systematic Review*, IEEE Transactions on Learning Technologies. Vol. 12. Hal 485-502.