

BAB V

HASIL DAN ANALISIS PENELITIAN

5.1 Pendahuluan

Bab ini akan menguraikan hasil penelitian dari perangkat lunak yang digunakan untuk melakukan proses klasifikasi teks berupa pertanyaan berbahasa Indonesia menggunakan algoritma *Support Vector Machine* (SVM) dan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information*. Hasil penelitian akan ditampilkan dalam bentuk *matrix* berupa nilai *accuracy*, *precision*, *recall*, *f-measure*, waktu komputasi dan juga jumlah fitur. Akan dilakukan analisis terkait hasil yang didapatkan.

5.2 Data Hasil Penelitian

5.2.1 Konfigurasi Percobaan

Pengujian perangkat lunak dilakukan dengan menggunakan data masukan berupa data pertanyaan yang sebelumnya telah dijelaskan pada subbab 3.3.3. Proses pengujian yang dilakukan telah disesuaikan dengan perancangan yang dibuat pada bab sebelumnya. Pengujian terbagi menjadi empat konfigurasi diantaranya konfigurasi pertama yaitu pengujian menggunakan algoritma SVM tanpa metode seleksi fitur, konfigurasi kedua yaitu pengujian menggunakan algoritma SVM dan metode seleksi fitur *Information Gain*, konfigurasi ketiga yaitu pengujian menggunakan algoritma SVM dan metode seleksi fitur *Chi Square*, dan konfigurasi keempat yaitu pengujian menggunakan algoritma SVM dan metode seleksi fitur *Mutual Information*. Model yang digunakan akan diuji menggunakan *cross*

validation sebanyak 10-fold dengan parameter kernel, nilai *C* dan *threshold* yang berbeda. Kemudian ditampilkan hasil evaluasi berupa nilai rata-rata yang didapatkan pada hasil pengujian dari 10-fold tersebut. Selanjutnya dilakukan analisa pada perbandingan hasil evaluasi yang telah disajikan dalam bentuk tabel. Hasil evaluasi yang akan ditampilkan pada tabel diantaranya hasil akhir berupa nilai rata-rata dari *accuracy*, *precision*, *recall*, *f-measure*, waktu komputasi dan jumlah fitur.

5.2.1.1 Data Hasil Konfigurasi 1

Pada data konfigurasi 1 dilakukan pengujian menggunakan algoritma SVM tanpa seleksi fitur yang terdiri dari 3 jenis kernel diantaranya kernel Linear, Polynomial dan Rbf. Serta nilai *C* yang terdiri dari nilai *C*: 1 dan nilai *C*: 10. Hasil Konfigurasi ditampilkan pada tabel V-1 sampai V-3.

Tabel V-1. Hasil Evaluasi Metode Klasifikasi SVM Tanpa Seleksi Fitur pada Kernel Linear

Nilai <i>C</i>	Nilai <i>Threshold</i>	Kernel : Linear					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F- Measure</i>	Waktu Komputasi
0.1	-	2004	0.75	0.51	0.59	0.54	0.14
1	-	2004	0.91	0.92	0.86	0.88	0.12
10	-	2004	0.89	0.9	0.86	0.87	0.14

Tabel V-2. Hasil Evaluasi Metode Klasifikasi SVM Tanpa Seleksi Fitur pada Kernel Polynomial

Nilai C	Nilai <i>Threshold</i>	Kernel : Polynomial					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	-	2004	0.43	0.15	0.33	0.2	0.18
1	-	2004	0.75	0.84	0.61	0.59	0.19
10	-	2004	0.76	0.77	0.63	0.62	0.18

Tabel V-3. Hasil Evaluasi Pada Klasifikasi SVM untuk Kernel Rbf

Nilai C	Nilai <i>Threshold</i>	Kernel : Rbf					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	-	2004	0.44	0.35	0.34	0.21	0.16
1	-	2004	0.87	0.9	0.84	0.81	0.18
10	-	2004	0.88	0.9	0.82	0.84	0.19

Bedasarkan tabel V-1, V-2, dan V-3 dapat dilihat hasil evaluasi yang terjadi pada klasifikasi menggunakan algoritma SVM tanpa seleksi fitur mendapatkan hasil yang berbeda pada setiap parameter, yakni penggunaan parameter kernel dan nilai c. Hasil evaluasi yang telah didapat akan dibandingkan dan dianalisa lebih lanjut.

5.2.1.2 Data Hasil Konfigurasi II

Pada data konfigurasi 2 dilakukan pengujian menggunakan algoritma SVM dengan seleksi fitur *Information Gain* yang terdiri dari 3 jenis kernel diantaranya kernel Linear, Polynomial dan Rbf. Serta nilai C yang terdiri dari nilai C: 0.1 ,C: 1

dan nilai C: 10. Kemudian akan dilakukan juga pengujian pada 3 jenis *threshold* diantaranya 1.66, 1.67 dan 1.68. Penggunaan ketiga *threshold* tersebut merupakan *threshold* dengan hasil kinerja terbaik yang didapat dari serangkaian percobaan “*try and error*” yang dilakukan berdasarkan hasil akhir dari bobot pada setiap *term* yang diberikan oleh metode seleksi fitur “*Information Gain*”. Hasil konfigurasi ditampilkan pada tabel V-4 sampai V-6.

Tabel V-4. Hasil Evaluasi Metode Klasifikasi SVM+IG pada Kernel Linear

Nilai C	Nilai <i>Threshold</i>	Kernel : Linear					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	1.66	801	0.76	0.52	0.6	0.55	0.11
	1.67	511	0.44	0.41	0.34	0.22	0.14
	1.68	405	0.72	0.5	0.57	0.52	0.12
1	1.66	801	0.91	0.92	0.85	0.87	0.08
	1.67	511	0.9	0.91	0.85	0.88	0.08
	1.68	405	0.88	0.9	0.83	0.85	0.08
10	1.66	801	0.88	0.87	0.85	0.86	0.08
	1.67	511	0.88	0.88	0.85	0.86	0.07
	1.68	405	0.85	0.83	0.82	0.82	0.07

Tabel V-5. Hasil Evaluasi Metode Klasifikasi SVM+IG pada Kernel Polynomial

Nilai C	Nilai <i>Threshold</i>	Kernel : Polynomial					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	1.66	801	0.43	0.15	0.33	0.2	0.14
	1.67	511	0.45	0.38	0.34	0.22	0.14
	1.68	405	0.5	0.49	0.38	0.29	0.13
1	1.66	801	0.79	0.85	0.69	0.71	0.16
	1.67	511	0.82	0.87	0.74	0.76	0.15

	1.68	405	0.81	0.85	0.73	0.75	0.13
10	1.66	801	0.81	0.84	0.72	0.73	0.15
	1.67	511	0.82	0.87	0.72	0.74	0.13
	1.68	405	0.82	0.85	0.73	0.76	0.13

Tabel V-6. Hasil Evaluasi Pada Klasifikasi SVM+IG untuk Kernel Rbf

Nilai C	Nilai <i>Threshold</i>	Kernel : Rbf					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	1.66	801	0.64	0.47	0.5	0.45	0.14
	1.67	511	0.71	0.49	0.56	0.51	0.12
	1.68	405	0.74	0.51	0.58	0.53	0.12
1	1.66	801	0.88	0.91	0.82	0.84	0.14
	1.67	511	0.89	0.92	0.83	0.86	0.12
	1.68	405	0.88	0.9	0.82	0.84	0.12
10	1.66	801	0.89	0.91	0.84	0.86	0.16
	1.67	511	0.9	0.91	0.86	0.88	0.14
	1.68	405	0.88	0.88	0.84	0.85	0.13

Bedasarkan tabel V-4, V-5, dan V-6 dapat dilihat hasil evaluasi yang terjadi pada klasifikasi menggunakan algoritma SVM menggunakan metode seleksi fitur *Information Gain* mendapatkan hasil yang berbeda pada setiap parameter, yakni penggunaan parameter kernel, nilai C dan *threshold*. Nilai *threshold* yang memiliki hasil evaluasi terbaik dan relatif stabil pada tiap parameternya didapatkan oleh *threshold* 1.67 dengan jumlah fitur sebesar 511. Hal ini dikarenakan *threshold* sebesar 1.67 mampu mengurangi kesalahan model dalam melakukan klasifikasi fitur dengan cara menghilangkan fitur yang kurang relevan. Berdasarkan hasil tersebut, nilai *threshold* yang akan digunakan pada metode seleksi fitur *Information Gain*

ialah 1.67. Hasil evaluasi dengan parameter *threshold* terbaik akan dibandingkan dan dianalisa lebih lanjut.

5.2.1.3 Data Hasil Konfigurasi III

Pada data konfigurasi 3 dilakukan pengujian menggunakan algoritma SVM dengan seleksi fitur *Chi Square* yang terdiri dari 3 jenis kernel diantaranya kernel Linear, Polynomial dan Rbf. Serta nilai C yang terdiri dari nilai C: 1 dan nilai C: 10. Kemudian akan dilakukan juga pengujian pada 3 jenis *threshold* diantaranya *threshold* dengan nilai 2.5, 3.5 dan 4.5. Penggunaan ketiga *threshold* tersebut merupakan *threshold* dengan hasil kinerja terbaik yang didapat dari serangkaian percobaan “*try and error*” yang dilakukan berdasarkan hasil akhir dari bobot pada setiap *term* yang diberikan oleh metode seleksi fitur “*Chi Square*”. Hasil Konfigurasi ditampilkan pada tabel V-7 sampai V-9 :

Tabel V-7. Hasil Evaluasi Metode Klasifikasi SVM+CS pada Kernel Linear

Nilai C	Nilai <i>Threshold</i>	Kernel : Linear					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	2.5	630	0.78	0.53	0.62	0.56	0.17
	3.5	432	0.79	0.53	0.62	0.57	0.13
	4.5	350	0.79	0.53	0.62	0.57	0.06
1	2.5	630	0.92	0.92	0.93	0.87	0.08
	3.5	432	0.92	0.93	0.89	0.91	0.08
	4.5	350	0.91	0.92	0.92	0.9	0.07
10	2.5	630	0.91	0.91	0.88	0.89	0.07
	3.5	432	0.9	0.91	0.88	0.89	0.07
	4.5	350	0.9	0.9	0.87	0.88	0.05

Tabel V-8. Hasil Evaluasi Metode Klasifikasi SVM+CS pada Kernel Polynomial

Nilai C	Nilai Threshold	Kernel : Polynomial					
		Jumlah Fitur	Accuracy	Precision	Recall	F-Measure	Waktu Komputasi
0.1	2.5	630	0.59	0.5	0.47	0.4	0.11
	3.5	432	0.74	0.52	0.58	0.53	0.11
	4.5	350	0.78	0.53	0.62	0.57	0.08
1	2.5	630	0.85	0.89	0.75	0.77	0.12
	3.5	432	0.92	0.9	0.8	0.77	0.12
	4.5	350	0.91	0.91	0.88	0.89	0.16
10	2.5	630	0.85	0.88	0.76	0.78	0.12
	3.5	432	0.87	0.89	0.81	0.83	0.1
	4.5	350	0.9	0.9	0.88	0.89	0.09

Tabel V-9. Hasil Evaluasi Pada Klasifikasi SVM+CS untuk Kernel Rbf

Nilai C	Nilai Threshold	Kernel : Rbf					
		Jumlah Fitur	Accuracy	Precision	Recall	F-Measure	Waktu Komputasi
0.1	2.5	630	0.75	0.51	0.59	0.54	0.11
	3.5	432	0.77	0.52	0.61	0.56	0.09
	4.5	350	0.79	0.53	0.62	0.57	0.08
1	2.5	630	0.91	0.93	0.85	0.87	0.11
	3.5	432	0.91	0.93	0.89	0.9	0.08
	4.5	350	0.92	0.92	0.9	0.91	0.07
10	2.5	630	0.92	0.93	0.88	0.89	0.11
	3.5	432	0.92	0.92	0.9	0.91	0.09
	4.5	350	0.91	0.91	0.9	0.9	0.07

Bedasarkan tabel V-7, V-8, dan V-9 dapat dilihat hasil evaluasi yang terjadi pada klasifikasi menggunakan algoritma SVM menggunakan metode seleksi fitur *Chi-Square* mendapatkan hasil yang berbeda pada setiap parameter, yakni penggunaan parameter kernel, nilai C dan *threshold*. Nilai *threshold* yang memiliki

hasil evaluasi terbaik dan relatif stabil didapatkan oleh *threshold* 3.5 dengan jumlah fitur sebesar 432. Hal ini dikarenakan *threshold* sebesar 3.5 mampu mengurangi kesalahan model dalam melakukan klasifikasi fitur dengan cara menghilangkan fitur yang kurang relevan. Berdasarkan hasil tersebut, nilai *threshold* yang akan digunakan pada metode seleksi fitur *Chi Square* ialah 3.5. Hasil evaluasi dengan parameter *threshold* terbaik akan dibandingkan dan dianalisa lebih lanjut.

5.2.1.4 Data Hasil Konfigurasi IV

Pada data konfigurasi 4 dilakukan pengujian menggunakan algoritma SVM dengan seleksi fitur *Mutual Information* yang terdiri dari 3 jenis kernel diantaranya kernel Linear, Polynomial dan Rbf. Serta nilai C yang terdiri dari nilai C: 1 dan nilai C: 10. Kemudian akan dilakukan juga pengujian pada 3 jenis *threshold* diantaranya *threshold* dengan nilai 0.0002, 0.0003 dan 0.0004. Penggunaan ketiga *threshold* tersebut merupakan *threshold* dengan hasil kinerja terbaik yang didapat dari serangkaian percobaan “*try and error*” yang dilakukan berdasarkan hasil akhir dari bobot pada setiap *term* yang diberikan oleh metode seleksi fitur “*Mutual Information*”. Hasil Konfigurasi ditampilkan pada tabel V-10 sampai V-12.

Tabel V-10. Hasil Evaluasi Metode Klasifikasi SVM+MI pada Kernel Linear

Nilai C	Nilai <i>Threshold</i>	Kernel : Linear					
		Jumlah Fitur	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi
0.1	0.0002	671	0.75	0.51	0.59	0.54	0.11
	0.0003	608	0.77	0.52	0.61	0.56	0.09
	0.0004	240	0.79	0.53	0.62	0.57	0.08
1	0.0002	671	0.91	0.93	0.85	0.87	0.11
	0.0003	608	0.91	0.93	0.89	0.9	0.08

	0.0004	240	0.92	0.92	0.9	0.91	0.07
10	0.0002	671	0.92	0.93	0.88	0.89	0.11
	0.0003	608	0.92	0.92	0.9	0.91	0.09
	0.0004	240	0.91	0.91	0.9	0.9	0.07

Tabel V-11. Hasil Evaluasi Metode Klasifikasi SVM+MI pada Kernel Polynomial

Nilai C	Nilai Threshold	Kernel : Polynomial					
		Jumlah Fitur	Accuracy	Precision	Recall	F-Measure	Waktu Komputasi
0.1	0.0002	671	0.55	0.49	0.44	0.36	0.13
	0.0003	608	0.66	0.5	0.53	0.47	0.11
	0.0004	240	0.76	0.51	0.6	0.55	0.09
1	0.0002	671	0.83	0.87	0.72	0.74	0.12
	0.0003	608	0.86	0.88	0.76	0.78	0.13
	0.0004	240	0.89	0.9	0.84	0.86	0.1
10	0.0002	671	0.84	0.88	0.75	0.77	0.13
	0.0003	608	0.86	0.88	0.78	0.8	0.11
	0.0004	240	0.89	0.88	0.85	0.86	0.11

Tabel V-12. Hasil Evaluasi Pada Klasifikasi SVM+MI untuk Kernel Rbf

Nilai C	Nilai Threshold	Kernel : Polynomial					
		Jumlah Fitur	Accuracy	Precision	Recall	F-Measure	Waktu Komputasi
0.1	0.0002	671	0.74	0.51	0.58	0.53	0.11
	0.0003	608	0.75	0.52	0.6	0.55	0.1
	0.0004	240	0.78	0.52	0.62	0.56	0.08
1	0.0002	671	0.9	0.92	0.85	0.87	0.12
	0.0003	608	0.91	0.93	0.87	0.89	0.12
	0.0004	240	0.91	0.92	0.87	0.89	0.09
10	0.0002	671	0.91	0.92	0.87	0.89	0.14
	0.0003	608	0.91	0.93	0.88	0.89	0.13

	0.0004	240	0.91	0.9	0.88	0.89	0.09
--	--------	-----	------	-----	------	------	------

Berdasarkan tabel V-10, V-11, dan V-12 dapat dilihat hasil evaluasi yang terjadi pada klasifikasi menggunakan algoritma SVM menggunakan metode seleksi fitur *Mutual Information* mendapatkan hasil yang berbeda pada setiap parameter, yakni penggunaan parameter kernel, nilai *C* dan *threshold*. Nilai *threshold* yang memiliki hasil evaluasi terbaik dan relatif stabil pada tiap parameternya didapatkan oleh *threshold* 0.0004 dengan jumlah fitur sebesar 240. Hal ini dikarenakan *threshold* sebesar 0.0004 mampu mengurangi kesalahan model dalam melakukan klasifikasi fitur dengan cara menghilangkan fitur yang kurang relevan. Berdasarkan hasil tersebut, maka nilai *k* yang akan digunakan pada metode seleksi fitur *Mutual Information* ialah 0.0004. Hasil evaluasi dengan parameter *threshold* terbaik akan dibandingkan dan dianalisa lebih lanjut.

5.2.1.5 Perbandingan Data Hasil Konfigurasi

Berikut data perbandingan hasil klasifikasi menggunakan SVM tanpa seleksi fitur, SVM dengan seleksi fitur *Information Gain*, SVM dengan seleksi fitur *Chi Square* dan SVM dengan seleksi fitur *Mutual Information*.

Tabel V-13. Hasil Evaluasi Metode Klasifikasi Model SVM pada Kernel Linear

Algoritma	Nilai <i>C</i>	Nilai <i>Threshold</i>	Kernel : Linear					
			<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	Waktu Komputasi	Jumlah Fitur
SVM	0.1	-	0.75	0.51	0.59	0.54	0.14	2004
	1		0.91	0.92	0.86	0.88	0.12	2004
	10		0.89	0.9	0.86	0.87	0.14	2004
SVM+IG	0.1		0.44	0.41	0.34	0.22	0.14	511

	1	1.67	0.9	0.91	0.85	0.88	0.08	511
	10		0.88	0.88	0.85	0.86	0.07	511
SVM+CS	0.1	3.5	0.79	0.53	0.62	0.57	0.13	432
	1		0.92	0.93	0.89	0.91	0.08	432
	10		0.9	0.91	0.88	0.89	0.07	432
SCM+MI	0.1	0.0004	0.79	0.53	0.62	0.57	0.07	240
	1		0.92	0.92	0.89	0.9	0.07	240
	10		0.9	0.9	0.88	0.89	0.05	240

Tabel V-14. Hasil Evaluasi Metode Klasifikasi Model SVM pada Kernel Polynomial

Algoritma	Nilai C	Nilai Threshold	Kernel : Polynomial					
			Accuracy	Precision	Recall	F-Measure	Waktu Komputasi	Jumlah Fitur
SVM	0.1		0.43	0.15	0.33	0.2	0.18	2004
	1		0.75	0.84	0.61	0.59	0.19	2004
	10		0.76	0.77	0.63	0.62	0.18	2004
SVM+IG	0.1	1.67	0.45	0.38	0.34	0.22	0.14	511
	1		0.82	0.87	0.74	0.76	0.15	511
	10		0.82	0.87	0.72	0.74	0.13	511
SVM+CS	0.1	3.5	0.74	0.52	0.58	0.53	0.11	432
	1		0.92	0.9	0.8	0.77	0.12	432
	10		0.87	0.89	0.81	0.83	0.1	432
SCM+MI	0.1	0.0004	0.76	0.51	0.6	0.55	0.09	240
	1		0.89	0.9	0.84	0.86	0.1	240
	10		0.89	0.88	0.85	0.86	0.11	240

Tabel V-15. Hasil Evaluasi Pada Klasifikasi Model SVM untuk Kernel Rbf

Algoritma	Nilai C	Nilai Threshold	Kernel : Rbf					
			Accuracy	Precision	Recall	F-Measure	Waktu Komputasi	Jumlah Fitur
SVM	0.1	-	0.44	0.35	0.34	0.21	0.16	2004
	1		0.87	0.9	0.84	0.81	0.18	2004
	10		0.88	0.9	0.82	0.84	0.19	2004

SVM+IG	0.1	1.67	0.71	0.49	0.56	0.51	0.12	511
	1		0.89	0.92	0.83	0.86	0.12	511
	10		0.9	0.91	0.86	0.88	0.14	511
SVM+CS	0.1	3.5	0.77	0.52	0.61	0.56	0.09	432
	1		0.91	0.93	0.89	0.9	0.08	432
	10		0.92	0.92	0.9	0.91	0.09	432
SCM+MI	0.1	0.0004	0.78	0.52	0.62	0.56	0.08	240
	1		0.91	0.92	0.87	0.89	0.09	240
	10		0.91	0.9	0.88	0.89	0.09	240

Bedasarkan tabel V-13, V-14, dan V-15 dapat dilihat perbandingan dari hasil evaluasi yang terbagi menjadi 3 kernel diantaranya kernel Linear, Polynomial dan Rbf. Hasil dari tabel perbandingan selanjutnya akan diubah kedalam bentuk grafik. Grafik tersebut ditujukan untuk menjelaskan pengaruh parameter yang digunakan di SVM dan metode seleksi fitur dalam bentuk visualisasi. Selanjutnya, pada analisis hasil penelitian akan dibahas berbagai fenomena yang ditemukan pada hasil klasifikasi. Hasil dari analisis diharapkan dapat melihat pengaruh dari penggunaan metode seleksi fitur dan juga dapat menjawab rumusan masalah yang menjadi landasan pada penelitian ini.

5.2.1.6 Data Konfigurasi Hasil Pengujian Klasifikasi Pertanyaan

Berikut data perbandingan hasil klasifikasi pada pengujian menggunakan berbagai model yang sebelumnya telah dilatih. Masukan berupa pertanyaan berbahasa Indonesia pada pengujian ini terbagi 3, pertanyaan untuk label *factoid* yaitu “Siapa nama presiden rusia”, pertanyaan untuk label *non-factoid* yaitu “Bagaimana cara menggunakan sumpit” dan pertanyaan untuk label *others* yaitu “Apa saja yang menjadi syarat melakukan ibadah puasa”. Status *factoid*, non-

factoid dan others diberi label berhasil jika pertanyaan diprediksi dengan benar berdasarkan data aktual.

Tabel V-13. Data Hasil Pengujian Prediksi Klasifikasi Pertanyaan menggunakan Kernel Linear

Algoritma	Nilai C	Nilai Threshold	Kernel : Linear		
			Status <i>Factoid</i>	Status <i>Non-Factoid</i>	Status <i>Others</i>
SVM	0.1	-	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil
SVM+IG	0.1	1.67	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil
SVM+CS	0.1	3.5	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Tidak Berhasil
SVM+MI	0.1	0.0004	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil

Tabel V-14. Data Hasil Pengujian Prediksi Klasifikasi Pertanyaan menggunakan Kernel Polynomial

Algoritma	Nilai C	Nilai Threshold	Kernel : Polynomial		
			Status <i>Factoid</i>	Status <i>Non-Factoid</i>	Status <i>Others</i>
SVM	0.1	-	Berhasil	Tidak Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Tidak Berhasil
	10		Berhasil	Berhasil	Tidak Berhasil
SVM+IG	0.1	1.67	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Tidak Berhasil
	10		Berhasil	Berhasil	Tidak Berhasil
SVM+CS	0.1		Berhasil	Berhasil	Tidak Berhasil

	1	3.5	Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil
SVM+MI	0.1	0.0004	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil

Tabel V-15. Data Hasil Pengujian Prediksi Klasifikasi Pertanyaan menggunakan Kernel Rbf

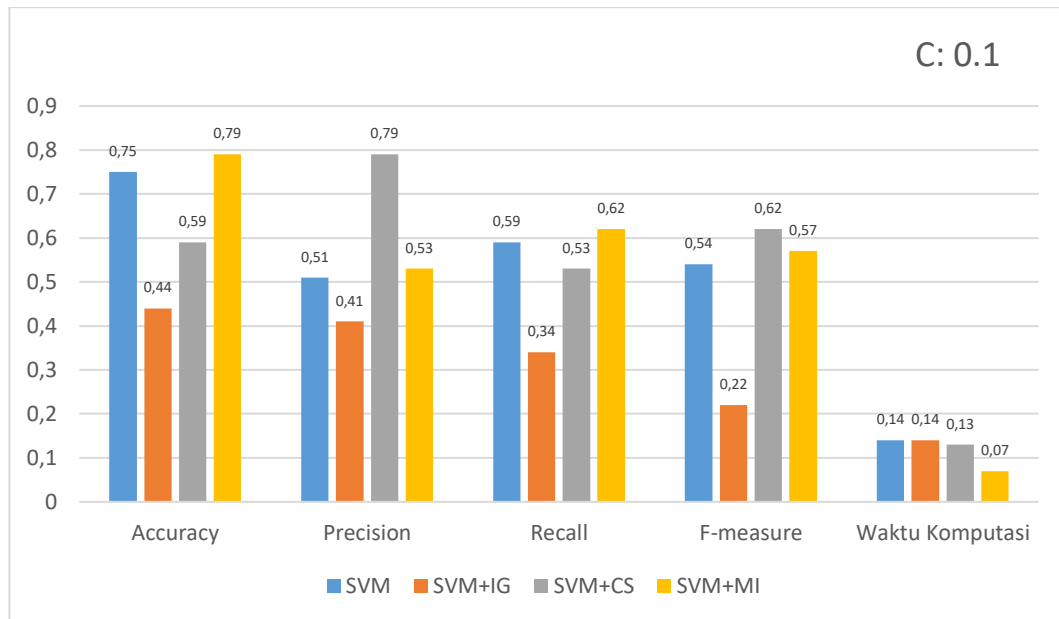
Algoritma	Nilai C	Nilai Threshold	Kernel : Rbf		
			Status <i>Factoid</i>	Status <i>Non-Factoid</i>	Status <i>Others</i>
SVM	0.1	-	Berhasil	Tidak Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Tidak Berhasil
	10		Berhasil	Berhasil	Berhasil
SVM+IG	0.1	1.67	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Tidak Berhasil
	10		Berhasil	Berhasil	Berhasil
SVM+CS	0.1	3.5	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil
SVM+MI	0.1	0.0004	Berhasil	Berhasil	Tidak Berhasil
	1		Berhasil	Berhasil	Berhasil
	10		Berhasil	Berhasil	Berhasil

5.3 Analisis Hasil Penelitian

5.3.1 Analisis Kernel, Nilai C dan *Threshold*

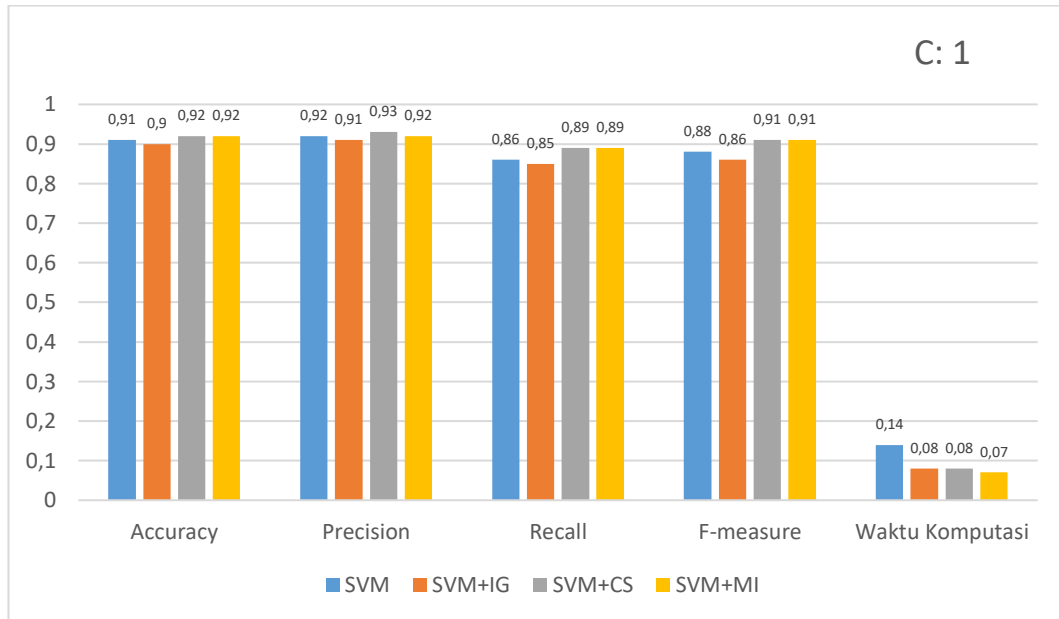
Penggunaan algoritma SVM sebagai model klasifikasi menggunakan masukan parameter yakni parameter kernel dan juga nilai c. Data hasil klasifikasi yang telah ditampilkan pada subbab sebelumnya menunjukkan hasil evaluasi yang berbeda pada tiap penggunaan parameternya. Penggunaan parameter masukan pada model klasifikasi yang digunakan akan dianalisis sebagai berikut.

5.3.1.1 Kernel Linear



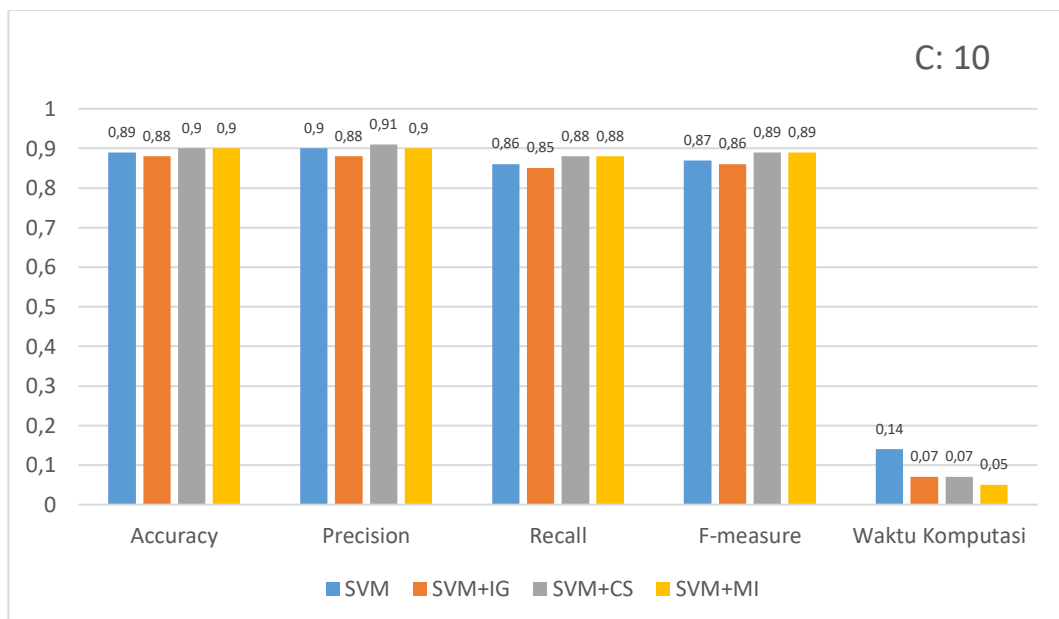
Gambar V-1. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Linear & C: 0.1

Gambar V-1 menunjukkan peningkatan *accuracy* pada metode seleksi fitur *Chi Square* sebesar 4%. Namun, penggunaan parameter ini memberikan penurunan yang cukup signifikan pada metode seleksi fitur *Information Gain* yakni sebesar 31% dan sebesar 16% untuk metode seleksi fitur *Chi Square*.



Gambar V-2. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Linear & C: 1

Gambar V-2 menunjukkan hasil *accuracy* yang cukup stabil untuk setiap skema pengujian pada model klasifikasi.



Gambar V-3. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Linear & C: 10

Gambar V-3 menunjukkan hasil *accuracy* yang cukup stabil untuk setiap skema pengujian pada model klasifikasi.

Grafik yang ditampilkan pada gambar diatas menunjukkan bahwa penggunaan nilai C : 1 dan C : 10 bekerja dengan baik dan cukup stabil pada kernel linear. Hasil akurasi terbaik sebesar 0.92 diperoleh oleh penggunaan metode seleksi fitur yakni *Chi Square* dan *Mutual Information* dengan penggunaan nilai C:1. Sedangkan, pada penggunaan nilai C : 0,1 memberikan hasil kinerja yang buruk pada setiap model klasifikasi di penggunaan kernel Linear. Hal ini memberikan fenomena bahwa pemilihan nilai C yang tepat diperlukan untuk mempengaruhi hasil akhir model klasifikasi.

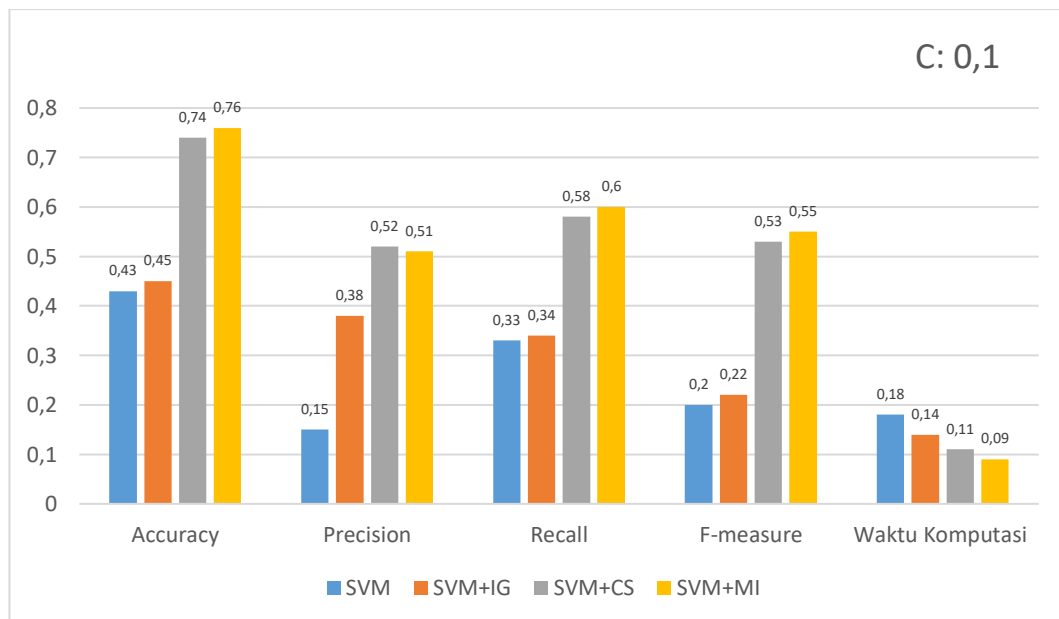
Nilai C sendiri digunakan untuk mengontrol trade off antara margin dan error klasifikasi. Hal ini bertujuan untuk menginformasikan optimasi SVM seberapa banyak kesalahan klasifikasi yang ingin dihindari pada proses pelatihan. Untuk penggunaan nilai C yang sangat kecil menyebabkan pengoptimal mencari hyperlane dengan fungsi pemisah dengan margin yang lebih besar. Hal ini menyebabkan kemungkinan adanya dot.product atau fitur yang diabaikan sehingga rasio kesalahan dalam prediksi menjadi lebih besar.

Pada kernel Linear ditampilkan juga hasil dari data evaluasi berupa waktu komputasi yang diperlukan untuk kinerja SVM menunjukkan hasil yang cukup baik untuk setiap model klasifikasi yang digunakan. Kernel linear bekerja dengan membagi data secara *linear* yang mana cara ini lebih sederhana dibandingkan

dengan penggunaan kernel non-linear. Sehingga pada penggunaan kernel ini membutuhkan waktu pemrosesan yang lebih sedikit dibanding penggunaan kernel lain. Dan juga penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information* menunjukkan fenomena bahwa penggunaan metode seleksi fitur terbukti efektif mampu mengurangi waktu komputasi pada kinerja algoritma SVM. Metode seleksi fitur *Mutual Information* menunjukkan kinerja terbaik dalam mengurangi waktu komputasi pada kinerja SVM dengan memberikan waktu komputasi sebesar 4 s. Hal ini disebabkan penggunaan treshold pada metode seleksi fitur mampu mengurangi jumlah fitur secara signifikan yaitu menjadi sebesar 240 fitur. Semakin sedikit jumlah fitur maka waktu komputasi yang diperlukan pada kinerja SVM juga semakin berkurang.

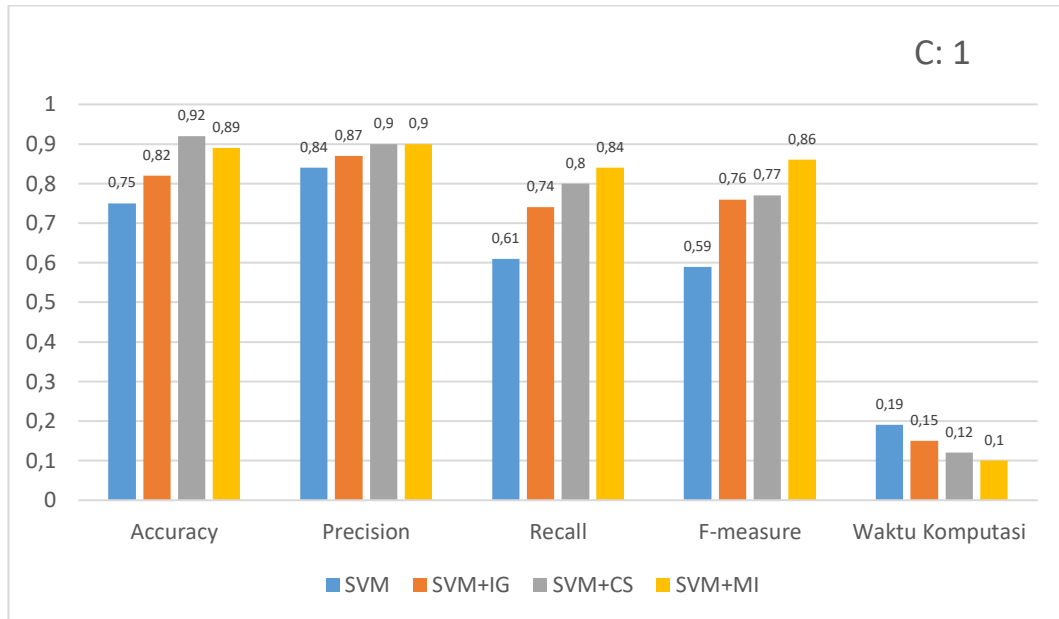
Secara keseluruhan penggunaan model SVM baik tanpa seleksi fitur maupun dengan seleksi fitur bekerja cukup efektif dan stabil pada kernel linear terutama dengan nilai C:1 dan C:10. Penggunaan metode seleksi fitur *Mutual Information* dan *Chi Square* mampu memberikan peningkatan pada hasil akhir data evaluasi namun peningkatan hasil yang diberikan tidak terlalu signifikan. Sedangkan, pada penggunaan seleksi fitur *Information Gain* memberikan hasil evaluasi yang sedikit menurun jika dibandingkan dengan model SVM tanpa seleksi fitur.

5.3.1.2 Kernel Polynomial



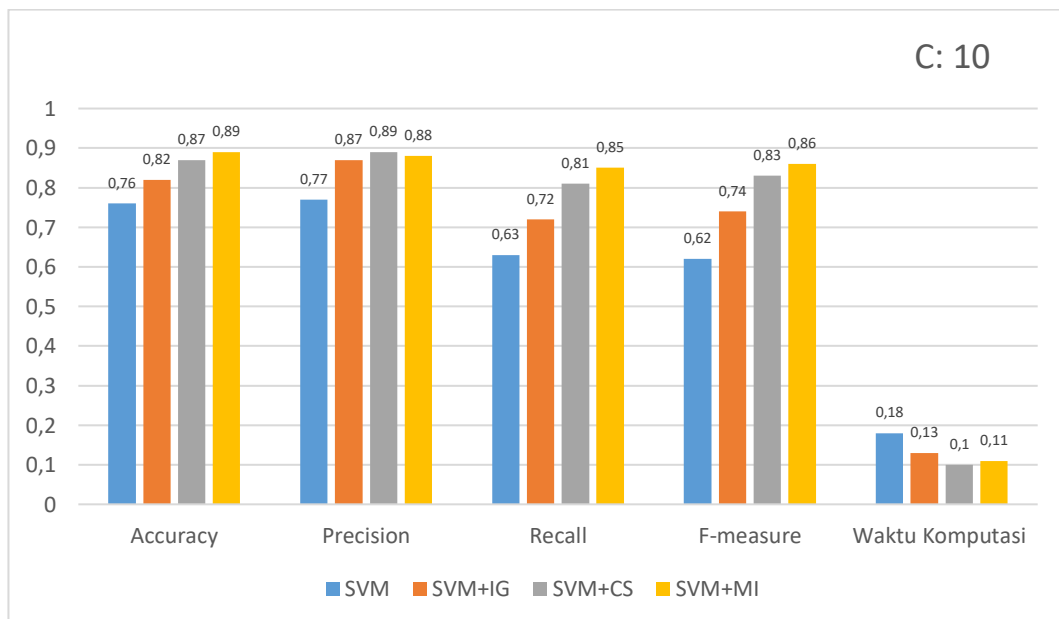
Gambar V-4. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Polynomial & C: 0.1

Gambar V-4 menunjukkan peningkatan *accuracy* yang sangat signifikan pada metode seleksi fitur *Chi Square* dan *Mutual Information* yakni sebesar 29% untuk *Chi Square* dan 31% untuk *Mutual Information*. Sedangkan, pada *Information Gain* hanya memberikan peningkatan sebesar 2%.



Gambar V-5. Grafik Data Perbandingan Hasil Model Klasifikasi
Kernel Polynomial & C: 1

Gambar V-5 menunjukkan peningkatan *accuracy* yang sangat signifikan pada metode seleksi fitur *Chi Square* sebesar 17% dan *Mutual Information* 14%.



Gambar V-6. Grafik Data Perbandingan Hasil Model Klasifikasi

Kernel Polynomial & C: 10

Gambar V-6 menunjukkan peningkatan *accuracy* yang sangat signifikan pada metode seleksi fitur *Chi Square* dan *Mutual Information* yakni sebesar 11% untuk *Chi Square* dan 23% untuk *Mutual Information*. Sedangkan pada metode seleksi fitur *Information Gain* juga terjadi peningkatan yakni sebesar 6%.

Grafik yang ditampilkan pada gambar diatas menunjukkan bahwa penggunaan kernel polynomial dengan nilai C : 0.1 , C : 1 dan C : 10 menunjukkan hasil evaluasi yang kurang baik dan cenderung tidak stabil pada tiap model klasifikasi yang digunakan. Berdasarkan grafik tersebut didapatkan fenomena bahwa penggunaan kernel polynomial menghasilkan hasil evaluasi yang kurang baik jika dibandingkan dengan penggunaan kernel lain. Sehingga didapatkan fenomena bahwa pemilihan kernel yang tepat diperlukan untuk melakukan optimasi pada algoritma SVM.

Berdasarkan grafik diatas didapatkan informasi bahwa penggunaan Nilai C : 0,1 menghasilkan hasil evaluasi yang sangat buruk bila dibandingkan dengan penggunaan nilai C : 1 dan nilai C : 10. Bahkan, penggunaan nilai C : 0,1 pada model SVM tanpa seleksi fitur hanya mendapatkan *accuracy* sebesar 0.44. Hal ini menunjukkan bahwa diperlukan pemilihan nilai C yang tepat untuk mendapatkan hasil evaluasi yang baik pada algoritma SVM. Parameter nilai C bertujuan untuk memberikan informasi pada optimasi SVM seberapa banyak kesalahan klasifikasi yang ingin dihindari pada proses pelatihan. Untuk penggunaan nilai C yang sangat kecil menyebabkan fungsi pengoptimal mencari hyperlane pada fungsi pemisah

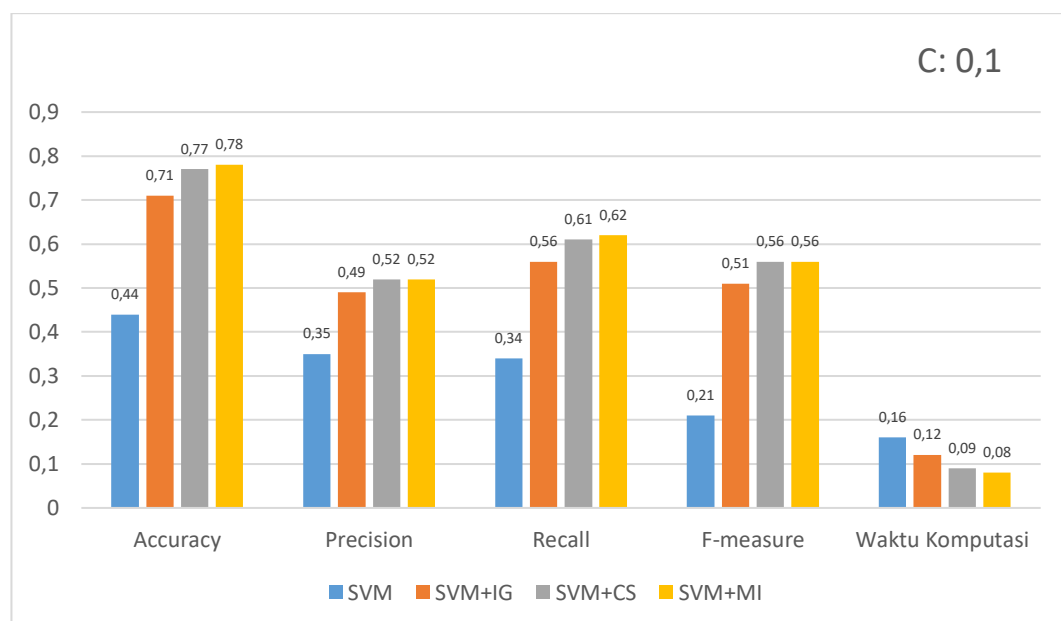
dengan margin yang lebih besar. Hal ini menyebabkan kemungkinan adanya dot.product atau fitur yang diabaikan sehingga rasio kesalahan dalam prediksi menjadi lebih besar.

Pada kernel polynomial juga ditampilkan hasil dari data evaluasi berupa waktu komputasi yang diperlukan untuk kinerja SVM pada tiap model menunjukkan hasil yang lebih lama dibandingkan dengan waktu komputasi pada penggunaan kernel linear. Kernel Polynomial termasuk kedalam *non-linear* kernel model yang bekerja dengan cara mengubah data ke dalam bentuk ruang fitur (*feature space*) berdimensi tinggi. Oleh karena itu, data dapat dipisah secara linear pada ruang fitur. Ruang fitur memiliki dimensi yang lebih tinggi dari vector *input*. Sehingga komputasi pada ruang fitur menjadi sangat besar yang mengakibatkan ruang fitur akan memiliki jumlah fitur yang tidak terhingga. Penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information* menunjukkan fenomena bahwa penggunaan metode seleksi fitur tersebut mampu mengurangi waktu komputasi pada kinerja algoritma klasifikasi SVM. Metode seleksi fitur *Mutual Information* menunjukkan kinerja terbaik dalam mengurangi waktu komputasi pada kinerja SVM, dengan memberikan waktu komputasi sebesar 9 s pada penggunaan parameter nilai $C : 0.1$. Penggunaan *threshold* 0.0004 pada metode seleksi fitur *Mutual Information* mampu mengurangi jumlah fitur secara signifikan yaitu menjadi sebesar 240 fitur. Semakin sedikit jumlah fitur maka waktu komputasi yang diperlukan pada kinerja SVM juga semakin berkurang.

Penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information* menunjukkan kinerja yang baik pada kernel polynomial.

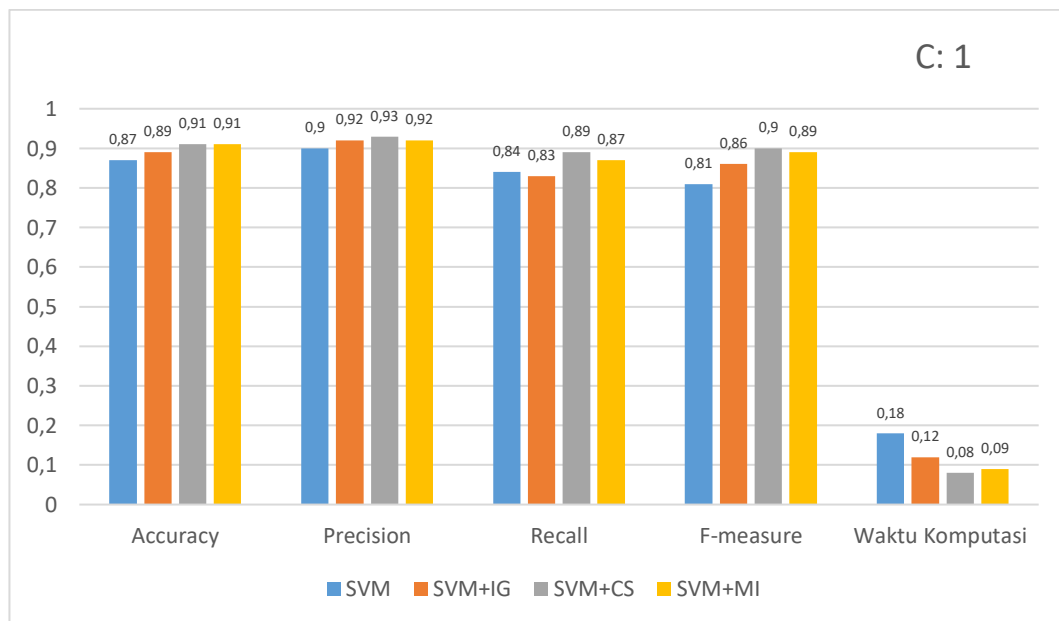
Walaupun secara keseluruhan hasil evaluasi yang didapat pada kernel ini kurang baik dan cenderung tidak stabil, penggunaan metode seleksi fitur tersebut mampu memberikan hasil yang cukup signifikan dalam meningkatkan hasil evaluasi pada proses klasifikasi SVM menggunakan kernel polynomial, terutama pada metode seleksi fitur *Chi Square* dan *Mutual Information*. Pada penggunaan parameter nilai $C : 1$ SVM tanpa seleksi fitur hanya mendapatkan hasil akurasi sebesar 0.75. Dengan menggunakan metode *Information Gain* hasil akurasi mampu meningkat sebesar 0.82, metode *Chi Square* mampu meningkatkan hasil akurasi menjadi 0.92 dan Metode *Mutual Information* mampu meningkatkan hasil akurasi sebesar 0.89. Hal ini memberikan fenomena bahwa metode *Chi Square* mampu memberikan fenomena terbaik dalam meningkatkan hasil evaluasi pada penggunaan kernel polynomial.

5.3.1.3 Kernel Rbf



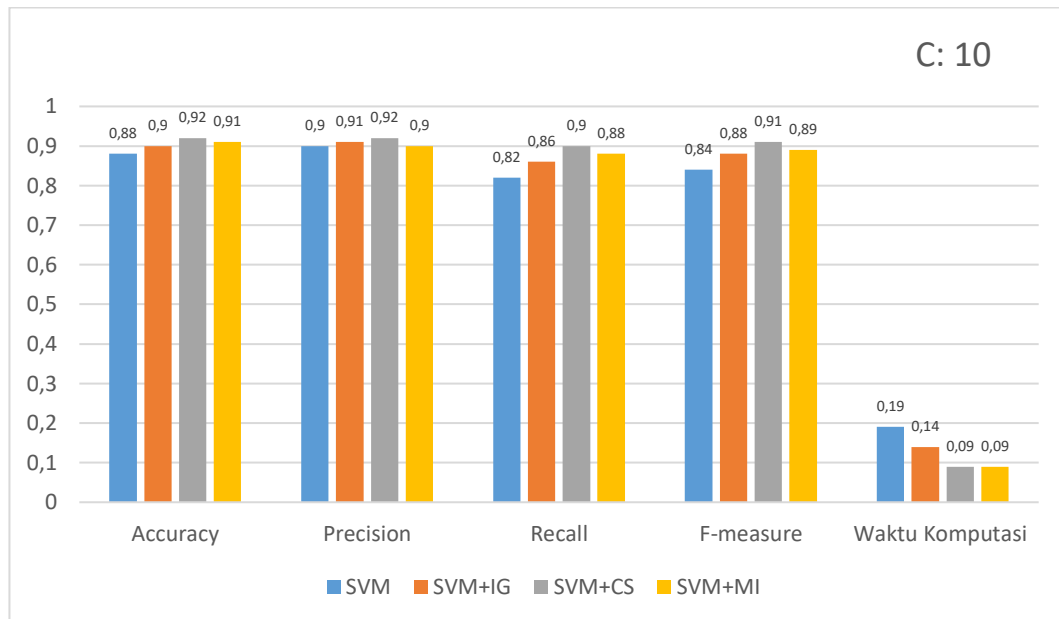
Gambar V-7. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Rbf & $C: 0.1$

Gambar V-7 menunjukkan peningkatan *accuracy* yang sangat signifikan pada metode seleksi fitur *Chi Square* dan *Mutual Information* yakni sebesar 34% untuk *Chi Square* dan 33% untuk *Mutual Information*. Sedangkan pada metode seleksi fitur *Information Gain* juga terjadi peningkatan yakni sebesar 27%.



Gambar V-8. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Rbf & C: 1

Gambar V-8 menunjukkan peningkatan *accuracy* untuk setiap model klasifikasi yakni untuk *Information Gain* sebesar 2%, kemudian untuk *Chi Square* dan *Mutual Information* sebesar 4%.



Gambar V-9. Grafik Data Perbandingan Hasil Model Klasifikasi Kernel Rbf & C: 10

Gambar V-8 menunjukkan peningkatan *accuracy* untuk setiap model klasifikasi yakni untuk *Information Gain* sebesar 2%, kemudian untuk *Chi Square* sebesar 4% dan *Mutual Information* 3%.

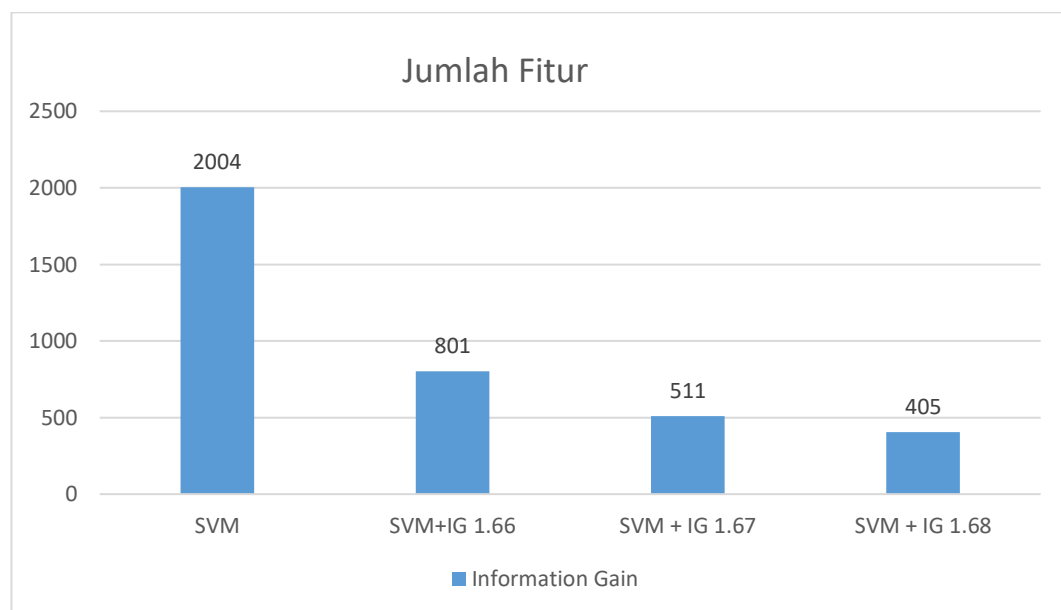
Grafik yang ditampilkan pada gambar diatas menunjukkan bahwa penggunaan kernel Rbf terutama dengan parameter nilai C : 1 dan C : 10 menunjukkan hasil evaluasi yang sangat baik dan cenderung stabil pada tiap model klasifikasi. Sedangkan kinerja Rbf cenderung kurang baik dan tidak stabil pada penggunaan parameter C : 0,1. Hal ini menunjukkan fenomena bahwa pemilihan nilai C yang tepat dapat memberikan hasil evaluasi terbaik. Penggunaan nilai C yang sangat kecil menyebabkan pengoptimal mencari hyperlane dengan fungsi pemisah dengan margin yang lebih besar. Hal ini menyebabkan kemungkinan adanya dot.product atau fitur yang diabaikan sehingga rasio kesalahan dalam prediksi menjadi lebih besar.

Pada kernel Rbf juga ditampilkan hasil dari data evaluasi berupa waktu komputasi yang diperlukan untuk kinerja SVM pada tiap model menunjukkan hasil yang lebih lama dibandingkan penggunaan kernel linear. Kernel Rbf termasuk kedalam *non-linear* kernel model yang bekerja dengan cara mengubah data ke dalam bentuk ruang fitur (*feature space*) berdimensi tinggi. Oleh karena itu, data dapat dipisah secara linear pada ruang fitur. Ruang fitur memiliki dimensi yang lebih tinggi dari vector input. Sehingga komputasi pada ruang fitur menjadi sangat besar yang mengakibatkan ruang fitur akan memiliki jumlah fitur yang tidak terhingga. Hal ini menyebabkan waktu komputasi pada kernel Rbf menjadi lebih besar. Penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information* menunjukkan fenomena bahwa penggunaan metode seleksi fitur terbukti efektif mampu mengurangi waktu komputasi pada kinerja algoritma SVM. Metode seleksi fitur *Mutual Information* menunjukkan kinerja terbaik dalam mengurangi waktu komputasi pada kinerja SVM dengan memberikan waktu komputasi sebesar 8 s pada parameter nilai $C : 0.1$. Hal ini disebabkan penggunaan treshold pada metode seleksi fitur mampu mengurangi jumlah fitur secara signifikan yaitu menjadi sebesar 240 fitur. Semakin sedikit jumlah fitur maka waktu komputasi yang diperlukan pada kinerja SVM juga semakin berkurang.

Penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information* menunjukkan kinerja yang baik pada kernel Rbf. Penggunaan metode seleksi fitur tersebut mampu meningkatkan hasil evaluasi pada proses klasifikasi SVM menggunakan kernel Rbf pada metode seleksi fitur *Information Gain*, *Chi Square* dan *Mutual Information*. Pada penggunaan parameter nilai $C : 10$

SVM tanpa seleksi fitur mendapatkan hasil akurasi yaitu 0.88. Dengan menggunakan metode *Information Gain* mampu meningkatkan hasil akurasi menjadi 0.9, metode *Chi Square* mampu meningkatkan hasil akurasi menjadi 0.92 dan metode *Mutual Information* mampu meningkatkan hasil akurasi sebesar 0.91. Hal ini memberikan fenomena bahwa metode *Chi Square* mampu memberikan kinerja terbaik dalam meningkatkan hasil evaluasi pada penggunaan kernel Rbf.

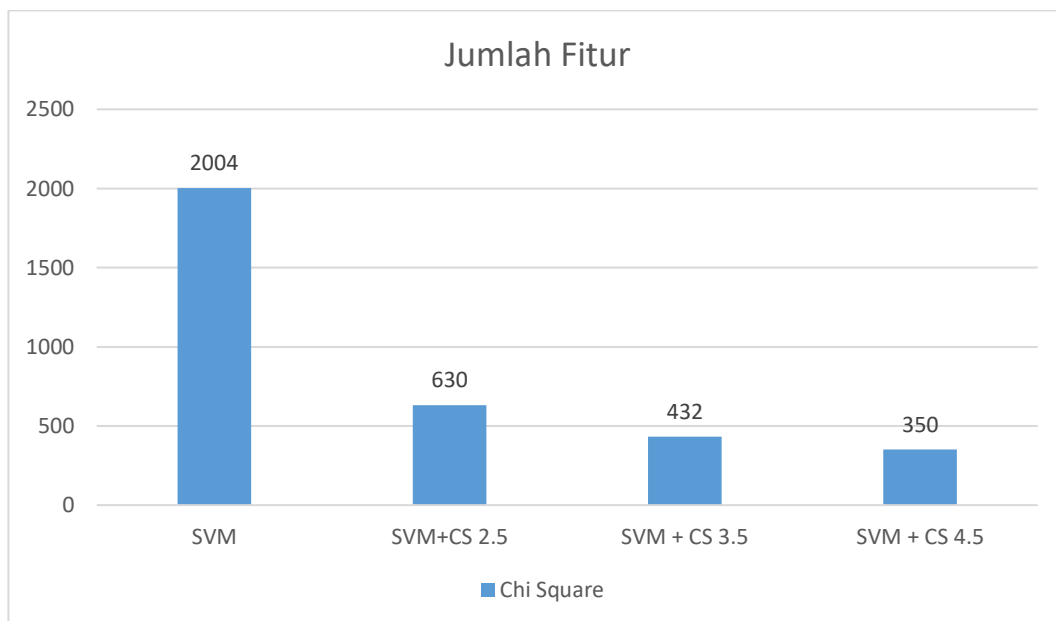
5.3.2 Analisis Jumlah Fitur



Gambar V-10. Grafik Data Perbandingan Jumlah Fitur Metode Information Gain

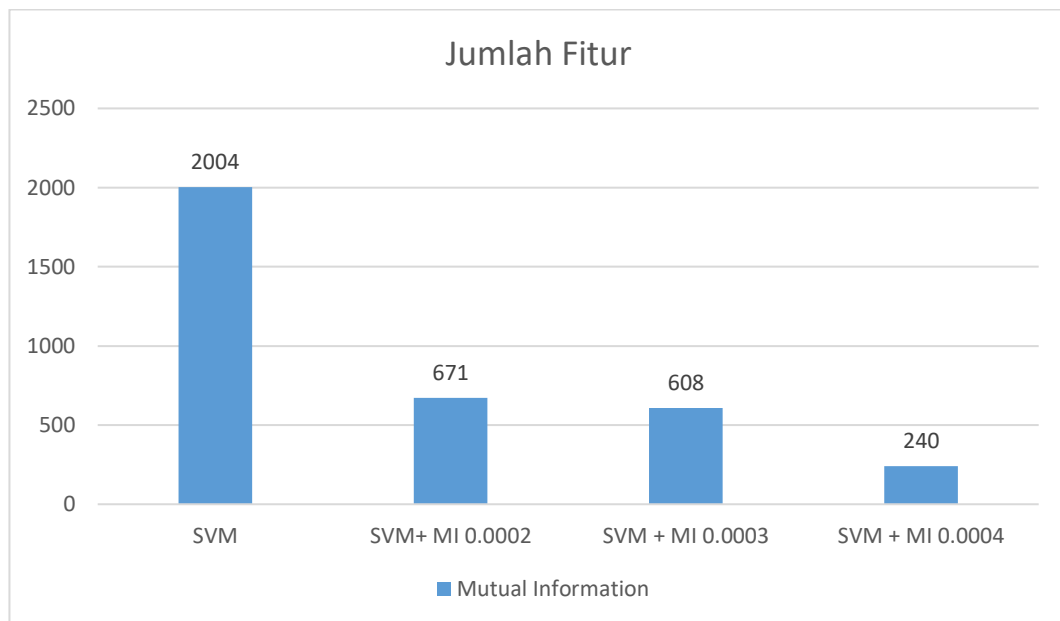
Berdasarkan grafik yang ditampilkan pada gambar V-10 penggunaan metode seleksi fitur *Information Gain* dengan *threshold* 1.66, 1.67 dan 1.68 memberikan hasil berupa pengurangan pada jumlah fitur. Terlihat pada grafik bahwa semakin tinggi *threshold* yang digunakan maka jumlah fitur akan semakin berkurang, dan berlaku pula sebaliknya. Hal ini dikarenakan semakin tinggi nilai

threshold yang diberikan maka nilai ambang batas untuk fitur dapat terseleksi juga semakin besar.



Gambar V-11. Grafik Data Perbandingan Jumlah Fitur Metode Chi Square

Berdasarkan grafik yang ditampilkan pada gambar V-11 penggunaan metode seleksi fitur *Chi Square* dengan *threshold* 2.5, 3.5 dan 4.5 memberikan hasil berupa pengurangan pada jumlah fitur. Terlihat pada grafik bahwa semakin tinggi *threshold* yang digunakan maka jumlah fitur akan semakin berkurang, dan berlaku pula sebaliknya. Hal ini dikarenakan semakin tinggi nilai *threshold* yang diberikan maka nilai ambang batas untuk fitur dapat terseleksi juga semakin besar.



Gambar V-12. Grafik Data Perbandingan Jumlah Fitur Metode Mutual Information

Berdasarkan grafik yang ditampilkan pada gambar V-12 penggunaan metode seleksi fitur *Mutual Information* dengan *threshold* 0.0002, 0.0003 dan 0.0004 memberikan hasil berupa pengurangan pada jumlah fitur. Terlihat pada grafik bahwa semakin tinggi *threshold* yang digunakan maka jumlah fitur akan semakin berkurang, dan berlaku pula sebaliknya. Hal ini dikarenakan semakin tinggi nilai *threshold* yang diberikan maka nilai ambang batas untuk fitur dapat terseleksi juga semakin besar.

Berdasarkan analisis pada uraian diatas maka didapatkan fenomena bahwa penggunaan *threshold* yang semakin tinggi maka semakin banyak fitur yang akan tereliminasi. Pengurangan jumlah fitur dapat mengurangi waktu komputasi pada proses klasifikasi namun belum tentu memberikan kinerja yang baik dikarenakan adanya kemungkinan fitur dengan bobot tinggi yang hilang.

5.3.3 Analisis Hasil Kinerja Metode Seleksi Fitur

Penggunaan metode seleksi fitur berpengaruh dalam mereduksi data yang kurang relevan sehingga dapat mempercepat waktu komputasi dan meningkatkan hasil evaluasi. Berdasarkan data hasil kinerja berupa grafik yang ditampilkan pada subbab 5.3.1 ditunjukkan bahwasanya penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square*, dan *Mutual Information* mampu mempercepat waktu komputasi dan memberikan peningkatan hasil kinerja pada sebagian besar parameter yang ada pada Algoritma *Support Vector Machine*. Metode seleksi fitur *Chi Square* dan *Mutual Information* memberikan hasil yang sangat baik dan relatif stabil untuk setiap parameter di SVM. Perbedaan hasil kinerja yang ditunjukkan oleh dua metode seleksi fitur tersebut tidak terlalu signifikan, hal ini disebabkan karna cara kerja perhitungan bobot pada metode seleksi fitur *Chi Square* dan *Mutual Information* memiliki kesamaan yaitu dengan cara mencari informasi pada setiap term untuk menghitung ketergantungan kelas pada suatu fitur. Metode seleksi fitur *Chi Square* mendapatkan nilai akurasi terbaik sebesar 0.92 dengan jumlah fitur sebesar 432 dan waktu komputasi sebesar 9 s pada kernel linear dengan nilai $C : 1$ dan *threshold* : 3.5. Kemudian, Metode seleksi fitur *Mutual Information* mendapatkan nilai akurasi terbaik sebesar 0.92 dengan jumlah fitur sebesar 240 dan waktu komputasi sebesar 7 s pada kernel linear dengan nilai $C : 1$ dan *threshold* : 0.0004.

Sedangkan, penggunaan metode seleksi fitur *Information Gain* menunjukkan hasil kinerja yang cukup baik namun relatif kurang stabil dikarenakan terjadi penurunan hasil kinerja pada beberapa parameter yang ada di algoritma

SVM. Metode seleksi fitur *Information Gain* bekerja dengan menghitung informasi pada setiap term kemudian melakukan teknik *scoring* untuk menentukan bobot menggunakan maksimal entropy. Metode seleksi fitur *Information Gain* mendapatkan nilai akurasi terbaik sebesar 0.9 dengan jumlah fitur sebesar 511 dan waktu komputasi sebesar 8 s pada kernel linear dengan nilai $C : 1$ dan *threshold* : 1.67. Hal ini memberikan fenomena bahwa perhitungan bobot pada setiap metode seleksi fitur sangat berpengaruh dalam mereduksi term yang kurang relevan sehingga dapat memberikan peningkatan pada hasil kinerja.

5.3.4 Analisis Hasil Prediksi Pengujian Klasifikasi

Berdasarkan data hasil prediksi pengujian klasifikasi yang ditampilkan pada tabel V-13, V-14 dan V-15 dengan data masukan berupa pertanyaan berbahasa Indonesia menunjukkan hasil yang cukup baik terutama saat melakukan prediksi untuk data masukan berupa pertanyaan dengan label *factoid* dan *non-factoid*. Sedangkan untuk data masukan berupa pertanyaan dengan label *other*, model cukup banyak melakukan kesalahan saat melakukan prediksi pada data masukan dengan label tersebut. Hal ini dikarenakan dataset yang digunakan pada penelitian ini belum dapat dikatakan seimbang dimana data untuk pertanyaan dengan label *factoid* berjumlah 519, data dengan label *non-factoid* berjumlah 491, dan data dengan label *others* berjumlah 185. Jumlah data yang terlalu sedikit pada data pertanyaan dengan label *others* menyebabkan terjadinya permasalahan *imbalance data*. Sehingga model cukup banyak melakukan kesalahan pada saat melakukan prediksi pengujian klasifikasi pada data uji dengan label *other*.

5.4 Kesimpulan

Pada bab ini diuraikan analisa yang dilakukan berdasarkan hasil pengujian yang didapatkan pada penelitian perbandingan metode seleksi fitur pada klasifikasi pertanyaan berbahasa Indonesia menggunakan algoritma *Support Vector Machine*. Berdasarkan analisa yang telah diuraikan sebelumnya, dapat disimpulkan bahwa penggunaan metode seleksi fitur berupa *Information Gain*, *Chi Square* dan *Mutual Information* memberikan pengaruh dalam meningkatkan hasil akurasi, mengurangi jumlah fitur, dan mengurangi waktu komputasi pada kinerja algoritma SVM. Penggunaan metode seleksi fitur *Chi Square* pada algoritma SVM dengan kernel linear dan parameter C: 1 menghasilkan kinerja terbaik dengan rata-rata *accuracy* 0.92, *precision* 0.93, *recall* 0.89, *f-measure* 0.91 dan waktu komputasi 8 detik. Pemilihan parameter serta nilai *threshold* yang tepat diperlukan untuk mendapatkan hasil evaluasi terbaik pada setiap model klasifikasi yang digunakan pada penelitian ini.