

BAB 4

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Implementasi Sistem

Implementasi sistem merupakan tahap untuk menerapkan hasil dari perancangan yang telah dilakukan terhadap sistem yang akan dibangun. Sehingga sistem tersebut dapat digunakan. Implementasi dalam penelitian ini mencakup implementasi perangkat keras, implementasi perangkat lunak, implementasi basis data serta implementasi antarmuka.

4.1.1 Implementasi Perangkat Keras

Untuk dapat menjalankan sistem yang dibangun, membutuhkan perangkat keras yang dapat mendukung berjalannya sistem. Adapun perangkat keras yang digunakan untuk mengimplementasikan sistem ekstraksi informasi sebagai berikut.

Tabel 4.1 Implementasi Perangkat Keras

Nama Perangkat Keras	Keterangan
Prosesor	Intel Core i3 2.60 Ghz
RAM	6 GB
Hardisk	500 GB

4.1.2 Implementasi Perangkat Lunak

Selain membutuhkan perangkat keras, untuk dapat menjalankan sistem yang dibangun juga membutuhkan perangkat lunak. berikut perangkat lunak yang digunakan untuk mengimplementasikan sistem ekstraksi informasi sebagai berikut.

Tabel 4.2 Implementasi Perangkat Lunak

Nama Perangkat Lunak	Keterangan
Sistem Operasi	Windows 8.1
Web Browser	Mozilla Firefox, Google Chrome
IDE Python	PyCharm Community Edition 2018
DBMS	PostgreSQL 9.3

4.1.3 Implementasi Basis Data

Implementasi basis data merupakan implementasi dari struktur tabel yang telah dirancang sebelumnya. Dalam penelitian ini basis data atau *database* diterapkan dengan menggunakan PostgreSQL. Adapun sintaks dari pembuatan *database* ekstraksi informasi riwayat hidup adalah sebagai berikut.

Tabel 4.3 Implementasi Basis Data

No	Nama Tabel	Perintah SQL
1	datalatih_datalatih	<pre>CREATE TABLE datalatih_datalatih (id_data_latih bigserial NOT NULL, hasil_tokenization text NOT NULL, x1 integer NOT NULL, x2 integer NOT NULL, x3 integer NOT NULL, x4 integer NOT NULL, x5 integer NOT NULL, x6 integer NOT NULL, x7 integer NOT NULL, x8 integer NOT NULL, x9 integer NOT NULL, x10 integer NOT NULL, x11 integer NOT NULL, x12 integer NOT NULL, x13 integer NOT NULL, x14 integer NOT NULL, x15 integer NOT NULL, x16 integer NOT NULL, x17 integer NOT NULL, x18 integer NOT NULL, x19 integer NOT NULL, x20 integer NOT NULL, x21 integer NOT NULL, x22 integer NOT NULL, x23 integer NOT NULL, kelas integer NOT NULL,</pre>

		<pre> nama_file character varying(100) NOT NULL, CONSTRAINT datalatih_datalatih_pkey PRIMARY KEY (id_data_latih)) WITH (OIDS=FALSE); </pre>
2	datauji_datauji	<pre> CREATE TABLE datauji_datauji (id_data_uji bigserial NOT NULL, hasil_tokenization text NOT NULL, x1 integer NOT NULL, x2 integer NOT NULL, x3 integer NOT NULL, x4 integer NOT NULL, x5 integer NOT NULL, x6 integer NOT NULL, x7 integer NOT NULL, x8 integer NOT NULL, x9 integer NOT NULL, x10 integer NOT NULL, x11 integer NOT NULL, x12 integer NOT NULL, x13 integer NOT NULL, x14 integer NOT NULL, x15 integer NOT NULL, x16 integer NOT NULL, x17 integer NOT NULL, x18 integer NOT NULL, x19 integer NOT NULL, x20 integer NOT NULL, x21 integer NOT NULL, x22 integer NOT NULL, x23 integer NOT NULL, kelas integer NOT NULL, </pre>

		<pre> kelas_klasifikasi integer NOT NULL, nama_file character varying(100) NOT NULL, CONSTRAINT datauji_datauji_pkey PRIMARY KEY (id_data_uji)) WITH (OIDS=FALSE); </pre>
3	hasilekstraksi_dataekstraksi	<pre> CREATE TABLE hasilekstraksi_dataekstraksi (id_ekstraksi bigserial NOT NULL, sub_judul text NOT NULL, data_diri text NOT NULL, pendidikan_formal text NOT NULL, pendidikan_non_formal text NOT NULL, pengalaman_kerja text NOT NULL, kemampuan text NOT NULL, lain_lain text NOT NULL, nama_file character varying(100) NOT NULL, CONSTRAINT hasilekstraksi_dataekstraksi_pkey PRIMARY KEY (id_ekstraksi)) WITH (OIDS=FALSE); </pre>

4.1.4 Implementasi Antarmuka

Implementasi antarmuka dalam penelitian ini mendeskripsikan setiap tampilan dari sistem yang akan dibangun. Adapun implementasi antarmuka terdiri dari nama antarmuka dan *file HTML* yang mewakilinya sebagai berikut.

Tabel 4.4 Implementasi Antarmuka

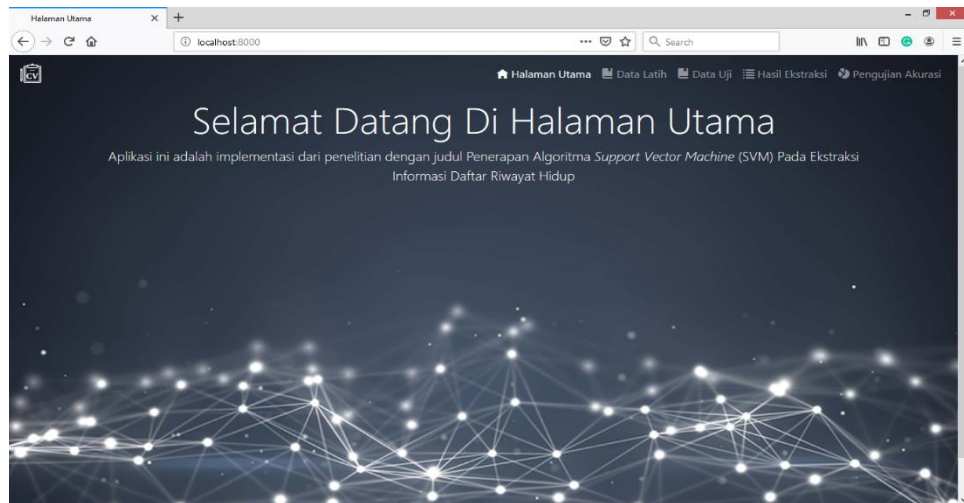
Nama Halaman	Deskripsi	File <i>HTML</i>
Halaman Utama	Merupakan halaman utama yang menampilkan judul penelitian saja	index.html
Halaman Data Latih	Merupakan halaman untuk menampilkan tabel data latih serta merekap data latih sesuai kelas dan <i>file CV</i>	datalatih.html
Halaman Tambah Data Latih	Merupakan halaman untuk mengupload <i>file CV</i> untuk data latih	tambah_data_latih.html
Halaman Pelatihan <i>SVM</i>	Merupakan halaman untuk proses pelabelan kelas serta menampilkan hasil pembobotan	pelatihansvm.html
Halaman Data Uji	Merupakan halaman untuk menampilkan tabel data uji serta merekap data uji sesuai kelas dan <i>file CV</i>	datauji.html

Halaman Tambah Data Uji	Merupakan halaman untuk mengupload <i>file CV</i> untuk data uji	tambah_data_uji.html
Halaman Hasil Ekstraksi	Merupakan halaman untuk menampilkan data hasil ekstraksi dari data uji yang telah terklasifikasi	dataekstraksi.html
Halaman Mulai Pengujian Akurasi	Merupakan halaman untuk memberikan label kelas sebenarnya pada data uji	mulaipengujian.html
Halaman Hasil Pengujian Akurasi	Merupakan halaman untuk menampilkan hasil akurasi dari data uji yang telah melalui proses pemberian label kelas sebenarnya	pengujianakurasi.html

Dari tabel implementasi diatas, akan ditampilkan implementasi antarmuka dari sistem ekstraksi informasi daftar riwayat hidup sebagai berikut.

1. Halaman Utama

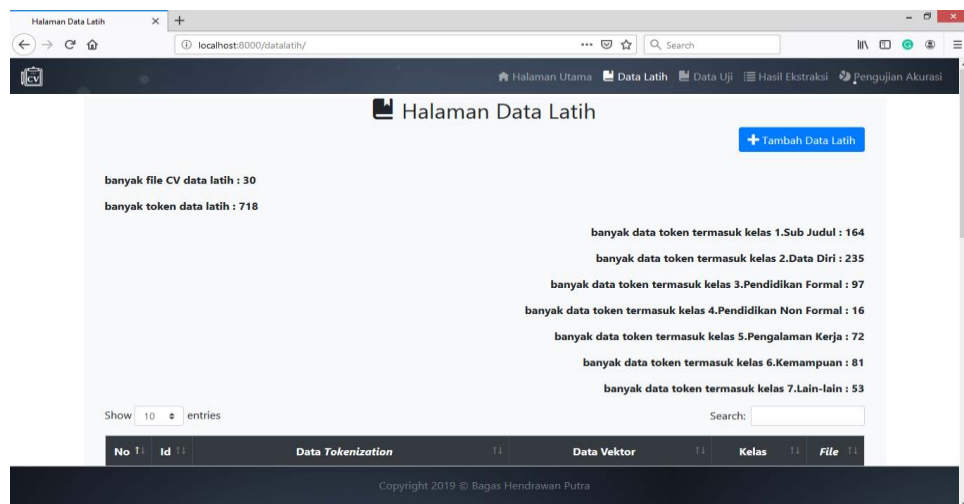
Berikut ini tampilan antarmuka halaman utama.



Gambar 4.1 Hasil Implementasi Antarmuka Halaman Utama

2. Halaman Data Latih

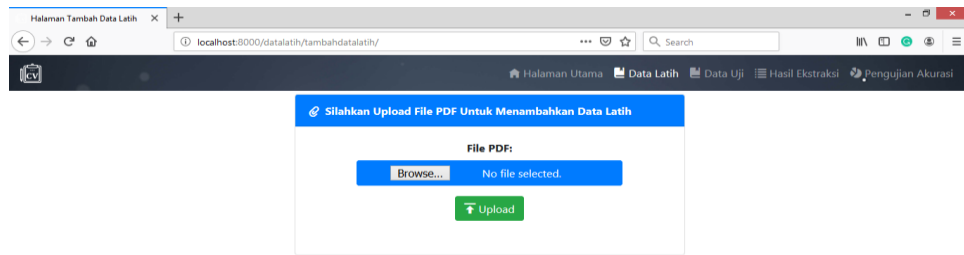
Berikut ini tampilan antarmuka halaman data latih.



Gambar 4.2 Hasil Implementasi Antarmuka Data Latih

3. Halaman Tambah Data Latih

Berikut ini tampilan antarmuka halaman tambah data latih.

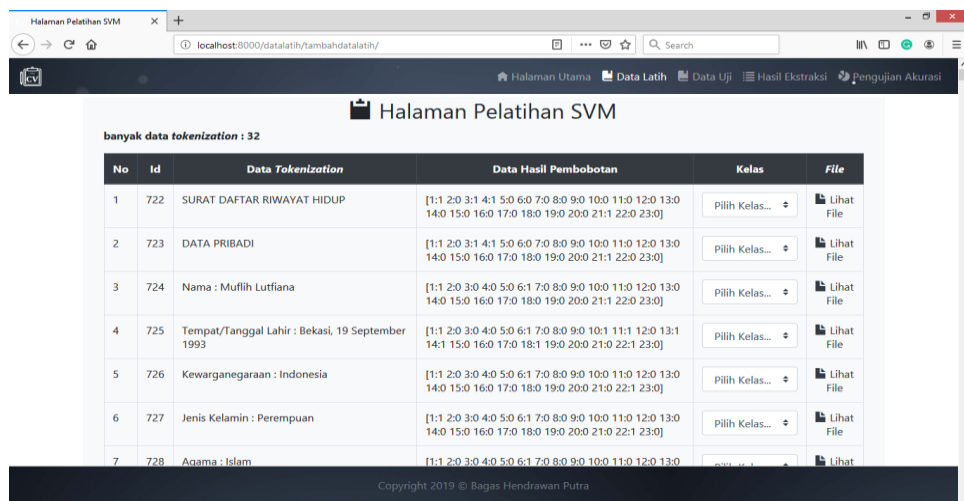


Copyright 2019 © Bagas Hendrawan Putra

Gambar 4.3 Hasil Implementasi Antarmuka Tambah Data Latih

4. Halaman Pelatihan SVM

Berikut ini tampilan antarmuka halaman pelatihan SVM.

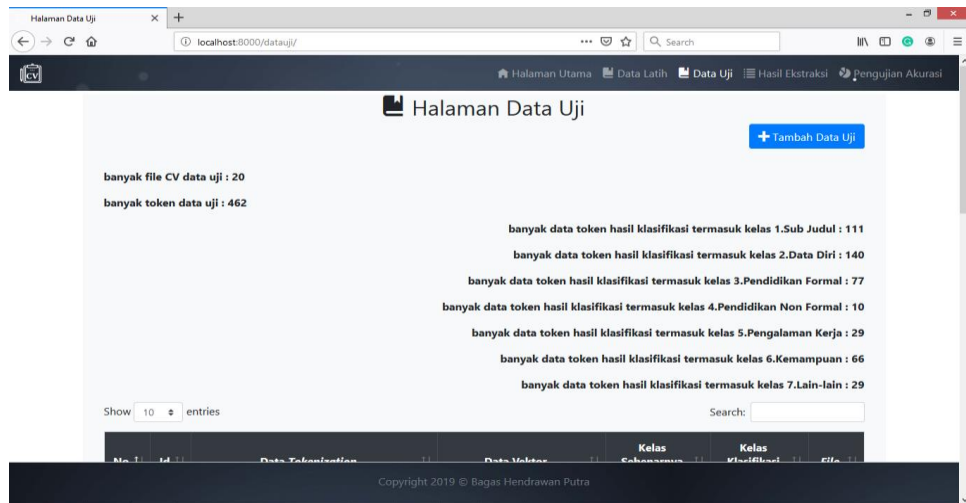


Copyright 2019 © Bagas Hendrawan Putra

Gambar 4.4 Hasil Implementasi Antarmuka Pelatihan SVM

5. Halaman Data Uji

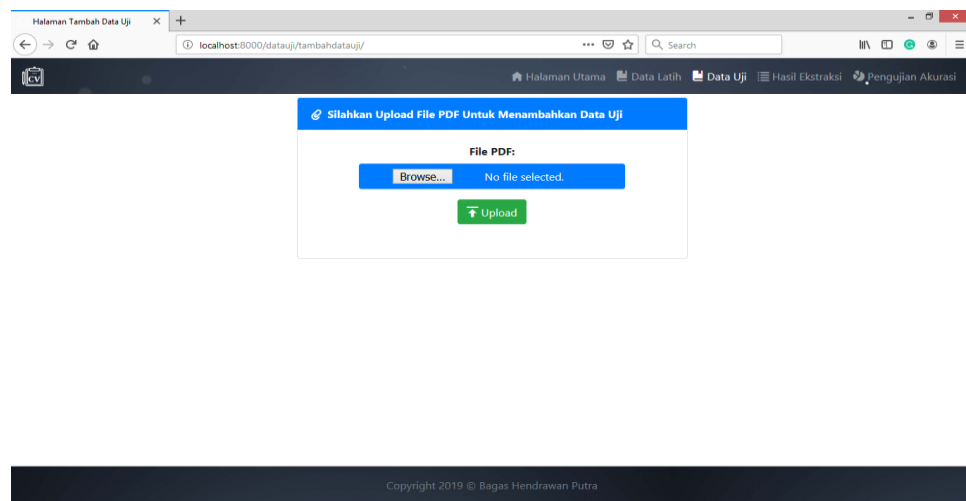
Berikut ini tampilan antarmuka halaman data uji.



Gambar 4.5 Hasil Implementasi Antarmuka Data Uji

6. Halaman Tambah Data Uji

Berikut ini tampilan antarmuka halaman tambah data uji.



Gambar 4.6 Hasil Implementasi Antarmuka Tambah Data Uji

7. Halaman Hasil Ekstraksi

Berikut ini tampilan antarmuka halaman hasil ekstraksi.

Halaman Hasil Ekstraksi

Show 10 entries

No	Id	Sub Judul	Data Diri	Pendidikan Formal	Pendidikan Non Formal	Pengalaman Kerja
1	20	DATA PRIBADI PENDIDIKAN FORMAL PENDIDIKAN NON FORMAL PENGALAMAN KERJA KEMAMPUAN	Nama: Devis Ayu Candra TTL: Tuban, 02 Juli 1992 Alamat: Jln. Dr. Wahidin SH 873, Tuban Jenis Kelamin: Perempuan Status: Belum menikah Agama : Islam HP: 082225559900 Email: ayu_devisdra@gmail.com	2016 - Universitas Yos Soedarmo Surabaya 51 Ilmu Hukum 2010 - SMK PGRI 2 Tuban 2007 - SMP Negeri 1 Jenu 2004 - SD Negeri Temaji 1 Jenu	2010 LULUS Pelatihan Program Pengembangan Komputer 2010 LULUS Pelatihan Program Secretary And Accounting	1. PT. SATYAMITRA SURYA PERKASA Sebagai PLTU Tahun 2017 2. PT. KERETA API INDONESIA Sebagai Pramugari Tahun 2018
2	19	CURRICULUM VITAE PROFIL PENDIDIKAN KEMAMPUAN	Nama: Dimass Prambudi Tempat, Tanggal Lahir: Bandung, 15 Juli 1996 Jenis Kelamin: Laki-Laki Status : Single Agama: Islam Kewarganegaraan : Indonesia Alamat: Galangan Rt 01/Rw 12	1. SDN 6 Batujajar, BDG - IDN Tahun 2002 - 2008 2. SMPN 1 Batujajar, BDG - IDN Tahun 2008 - 2011 3. SMK TIP Cimahi Tahun		

Copyright 2019 © Bagas Hendrawan Putra

Gambar 4.7 Hasil Implementasi Antarmuka Hasil Ekstraksi

8. Halaman Mulai Pengujian

Berikut ini tampilan antarmuka halaman mulai pengujian.

Halaman Mulai Pengujian

banyak data : 30

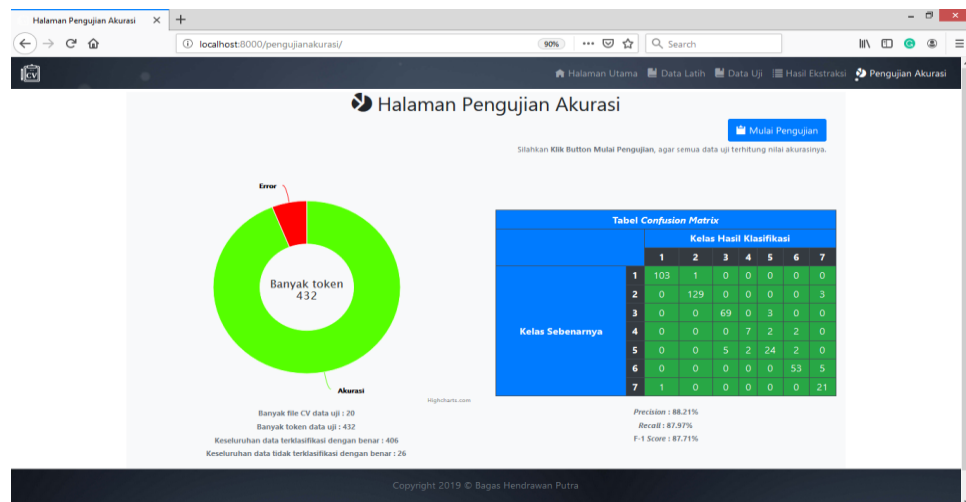
No	Id	Data Tokenization	Kelas
1	433	CURRICULUM VITAE	Pilih Kelas...
2	434	DATA PRIBADI	Pilih Kelas...
3	435	Nama :SYAMSUDAR	Pilih Kelas...
4	436	Tempat,Tanggal Lahir : Pinrang,22 Agustus 1991	Pilih Kelas...
5	437	Jenis Kelamin : Laki-laki	Pilih Kelas...
6	438	Agama : Islam	Pilih Kelas...
7	439	Tinggi Badan :170cm	Pilih Kelas...

Copyright 2019 © Bagas Hendrawan Putra

Gambar 4.8 Hasil Implementasi Antarmuka Mulai Pengujian

9. Halaman Hasil Pengujian Akurasi

Berikut ini tampilan antarmuka halaman hasil pengujian akurasi.



Gambar 4.9 Hasil Implementasi Antarmuka Hasil Pengujian Akurasi

4.2 Pengujian Sistem

Tahap pengujian ini dilakukan untuk menemukan kesalahan atau kekurangan dari sistem yang telah dibangun. Pengujian ini dimaksudkan untuk mengetahui apakah sistem yang dibangun sudah memenuhi kriteria yang dibutuhkan atau tidak. Adapun pengujian sistem yang akan dilakukan memiliki tahapan rencana pengujian yaitu rencana pengujian fungsionalitas dan pengujian *confusion matrix* sebagai berikut.

4.2.1 Pengujian Fungsionalitas

Pada penelitian ini, pengujian sistem yang dilakukan menggunakan metode *black box*. Pengujian *black box* ini dilakukan untuk menguji fungsi-fungsi yang terdapat pada sistem yang dibangun. Adapun skenario pengujian pada sistem yang dibangun adalah sebagai berikut.

Tabel 4.5 Skenario Pengujian Fungsionalitas

No	Komponen Yang Diuji	Skenario	Jenis Pengujian
1	Data Latih	Menampilkan data latih dan informasi hasil filter data latih	<i>Black Box</i>
		Lihat <i>file CV</i> data latih	<i>Black Box</i>

2	Tambah data latih	Masukan <i>file CV</i> data latih dengan benar	<i>Black Box</i>
		Masukan <i>file CV</i> data latih kosong	<i>Black Box</i>
		Menekan tombol <i>upload</i>	<i>Black Box</i>
3	Pelatihan <i>SVM</i>	Masukan label kelas data latih dengan benar	<i>Black Box</i>
		Masukan label kelas data latih kosong	<i>Black Box</i>
		Lihat <i>file CV</i> data latih	<i>Black Box</i>
		Menekan tombol simpan	<i>Black Box</i>
4	Data Uji	Menampilkan data uji dan informasi hasil filter data uji	<i>Black Box</i>
		Lihat <i>file CV</i> data uji	<i>Black Box</i>
5	Tambah Data Uji	Masukan <i>file CV</i> data uji dengan benar	<i>Black Box</i>
		Masukan <i>file CV</i> data uji kosong	<i>Black Box</i>
		Menekan tombol <i>upload</i>	<i>Black Box</i>
6	Hasil ekstraksi	Menampilkan hasil ekstraksi dari data uji	<i>Black Box</i>
7	Mulai pengujian	Masukan label kelas data pengujian dengan benar	<i>Black Box</i>
		Masukan label kelas data pengujian kosong	<i>Black Box</i>
		Menekan tombol simpan	<i>Black Box</i>
8	Hasil pengujian akurasi	Menampilkan hasil akurasi dan <i>confusion matrix</i> dari proses pengujian	<i>Black Box</i>

4.2.2 Skenario dan Hasil Pengujian Fungsionalitas

Hasil pengujian ini menampilkan benar atau tidaknya fungsionalitas dari sistem yang dibangun. Adapun hasilnya adalah sebagai berikut.

Tabel 4.6 Hasil Pengujian Fungsionalitas

No	Komponen Yang Diuji	Skenario	Hasil Yang Diharapkan	Hasil Pengujian
1	Data Latih	Menampilkan data latih dan informasi hasil filter data latih	Tampil data latih beserta informasi hasil filter data latih	(√) Berhasil () Gagal
		Lihat <i>file CV</i> data latih	Menampilkan <i>file CV</i> data latih	(√) Berhasil () Gagal
2	Tambah data latih	Masukan <i>file CV</i> data latih dengan benar	Berhasil menambah <i>file CV</i> data latih	(√) Berhasil () Gagal
		Masukan <i>file CV</i> data latih kosong	Menampilkan pesan <i>file CV</i> data latih harus di masukan	(√) Berhasil () Gagal
		Menekan tombol <i>upload</i>	Berhasil menambah data latih dan masuk ke halaman pelatihan <i>SVM</i>	(√) Berhasil () Gagal
3	Pelatihan <i>SVM</i>	Masukan label kelas data latih dengan benar	Berhasil menyimpan label kelas data latih	(√) Berhasil () Gagal
		Masukan label kelas data latih kosong	Menampilkan pesan label kelas data latih harus di isi	(√) Berhasil () Gagal
		Lihat <i>file CV</i> data latih	Menampilkan <i>file CV</i> data latih	(√) Berhasil () Gagal

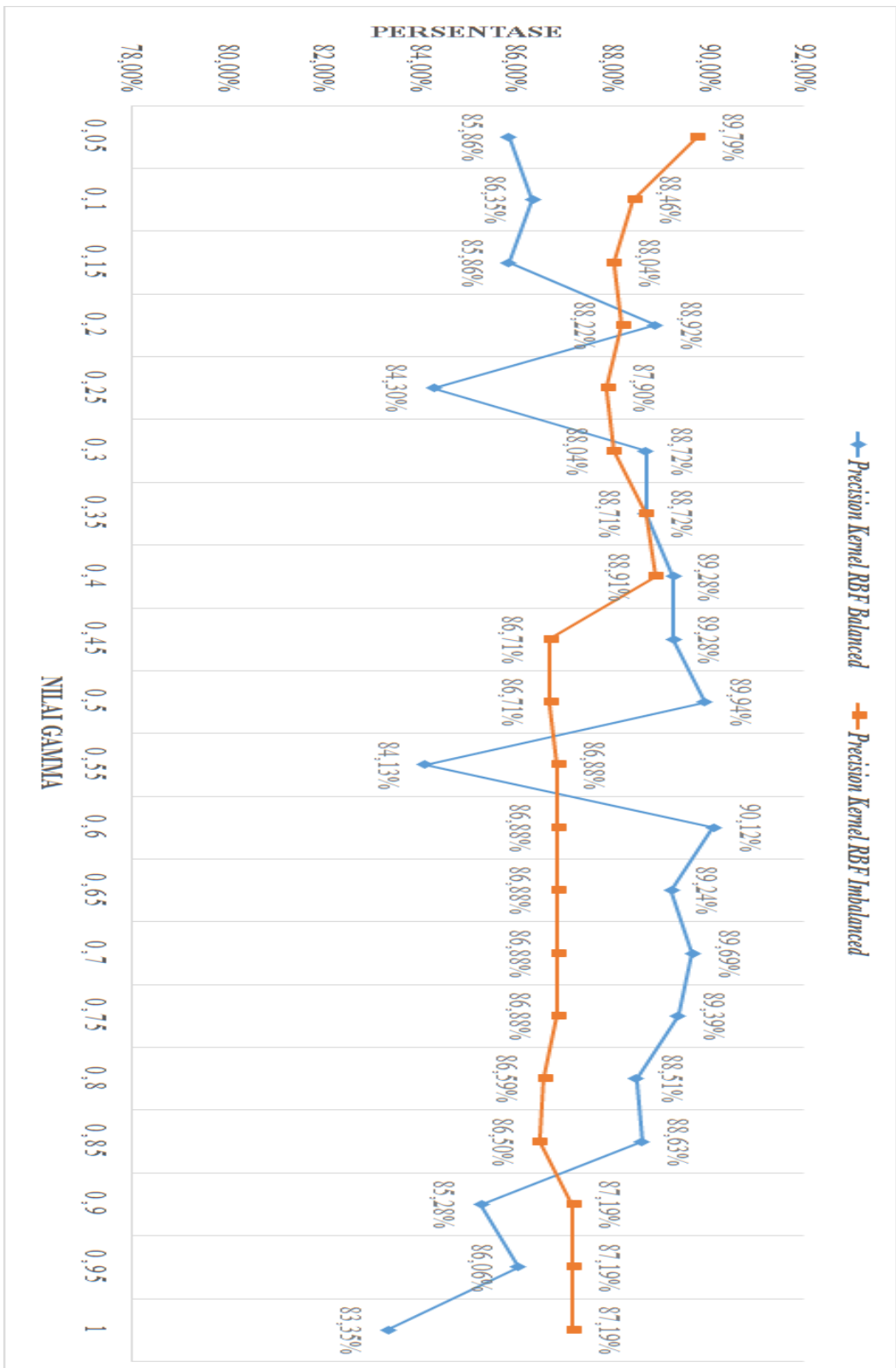
		Menekan tombol simpan	Berhasil menyimpan data latih hasil pelabelan dan hasil <i>SVM training</i>	(√) Berhasil () Gagal
4	Data Uji	Menampilkan data uji dan informasi hasil filter data uji	Tampil data uji beserta informasi hasil filter data uji	(√) Berhasil () Gagal
		Lihat <i>file CV</i> data uji	Menampilkan <i>file CV</i> data uji	(√) Berhasil () Gagal
5	Tambah Data Uji	Masukan <i>file CV</i> data uji dengan benar	Berhasil menambah <i>file CV</i> data uji	(√) Berhasil () Gagal
		Masukan <i>file CV</i> data uji kosong	Menampilkan pesan <i>file CV</i> data uji harus di masukan	(√) Berhasil () Gagal
		Menekan tombol <i>upload</i>	Berhasil menambah data uji	(√) Berhasil () Gagal
6	Hasil ekstraksi	Menampilkan hasil ekstraksi dari data uji	Tampil data uji hasil ekstraksi sehingga menjadi sebuah informasi terstruktur	(√) Berhasil () Gagal
7	Mulai pengujian	Masukan label kelas data pengujian dengan benar	Berhasil menyimpan label kelas data pengujian	(√) Berhasil () Gagal
		Masukan label kelas data pengujian kosong	Menampilkan pesan label kelas data pengujian harus di isi	(√) Berhasil () Gagal

		Menekan tombol simpan	Berhasil menyimpan data hasil pengujian	(√) Berhasil () Gagal
8	Hasil pengujian akurasi	Menampilkan hasil akurasi dan <i>confusion matrix</i> dari proses pengujian	Tampil akurasi pengujian berupa <i>chart</i> serta penjelasannya dari data uji beserta tabel <i>confusion matrix</i>	(√) Berhasil () Gagal

4.2.3 Pengujian *Confusion Matrix*

Pengujian *confusion matrix* dilakukan pada ekstraksi daftar riwayat hidup yang didapatkan secara umum maupun dari internet. Daftar riwayat hidup atau *Curriculum Vitae (CV)* yang digunakan dalam penelitian ini berjumlah 50 *file* dan dibagi menjadi data latih dan data uji. 30 *file CV* sebagai data latih dan 20 *file CV* sebagai data uji. Setelah dilakukan *preprocessing* menjadi bentuk tabel daftar riwayat hidup tersebut akan menghasilkan 718 token kalimat pada data latih dan 462 token kalimat pada data uji. Data uji tersebut kemudian diklasifikasikan menggunakan algoritma *SVM* menggunakan pendekatan *one against all* dan akan digunakan fungsi *kernel RBF*, *linear* dan *polynomial* sebagai pembanding. Selain penggunaan fungsi *kernel* sebagai pembanding, parameter pengujian lain yang digunakan adalah masukan nilai *gamma* untuk *kernel RBF* dan *polynomial*. Dari hasil klasifikasi algoritma *SVM* akan ditampilkan *confusion matrix* dan kemudian akan dihitung tingkat akurasi secara keseluruhan, nilai *precision*, nilai *recall*, dan *f1-score*. Pengujian *confusion matrix* akan terbagi menjadi pengujian dengan menggunakan *balanced data set* dan tanpa menggunakan *balanced data set*. Tujuan dari pengujian *confusion matrix* untuk mengetahui bagaimana pengaruh komposisi data antara menggunakan *balanced data set* atau tidak *balanced data set* terhadap akurasi, *precision*, *recall* dan *f1-score* dari *SVM*, karena dalam penelitian sebelumnya yang berjudul pemodelan *Support Vector Machine* data tidak seimbang keberhasilan studi mahasiswa magister IPB oleh Octavia, dkk[31] mengatakan bahwa pemodelan *SVM* pada data tidak seimbang menghasilkan kinerja yang

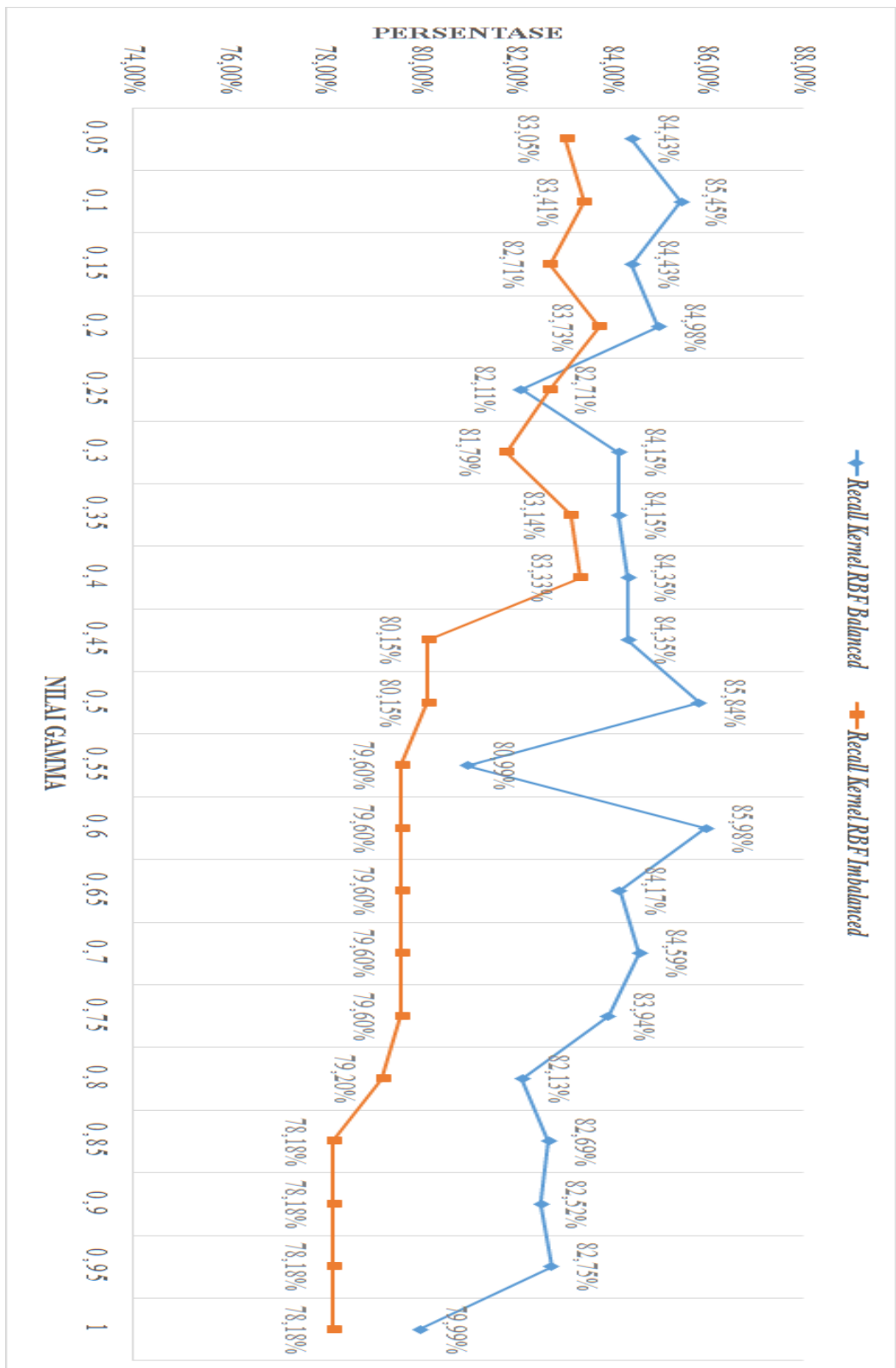
kurang baik sedangkan dengan data yang seimbang berhasil meningkatkan kinerja *SVM* dengan hasil yang paling baik terdapat pada *kernel RBF*. Pada Gambar 4.10 sampai Gambar 4.19 adalah hasil grafik pada masing masing fungsi *kernel* yang telah dilakukan proses perhitungan. Untuk Gambar 4.10 hingga Gambar 4.13 yaitu grafik hasil perhitungan *precision*, *recall*, *f1-score* dan akurasi menggunakan fungsi *kernel RBF* dengan menggunakan *balanced data set* dan *imbalanced data set* dengan parameter masukan berupa nilai *gamma* kelipatan 0,05 sebanyak 20 kali. Sedangkan untuk lebih detail mengenai hasil perhitungannya terdapat pada Lampiran E. Berikut pada Gambar 4.10 hasil perhitungan *precision* dalam bentuk grafik baik menggunakan *kernel RBF balanced data set* dan *imbalanced data set*.



Gambar 4.10 Grafik Hasil Perhitungan *Precision Kernel RBF Balanced Data Set dan Imbalanced*

Pada grafik Gambar 4.10 menunjukkan bahwa *precision* dengan *kernel RBF balanced* dari gamma 0,5 sampai 1 terlihat persentase mengalami naik turun yang cukup signifikan. Sedangkan *precision* dengan *kernel RBF imbalanced* dari gamma 0,5 sampai 1 mengalami naik turun yang tidak terlalu signifikan dan persentase cenderung stabil pada gamma 0,55 hingga 0,75. Untuk nilai *precision* yang paling tinggi terdapat pada gamma 0,6 dengan nilai persentase sebesar 90,12% menggunakan fungsi *kernel RBF* dengan *balanced data set*.

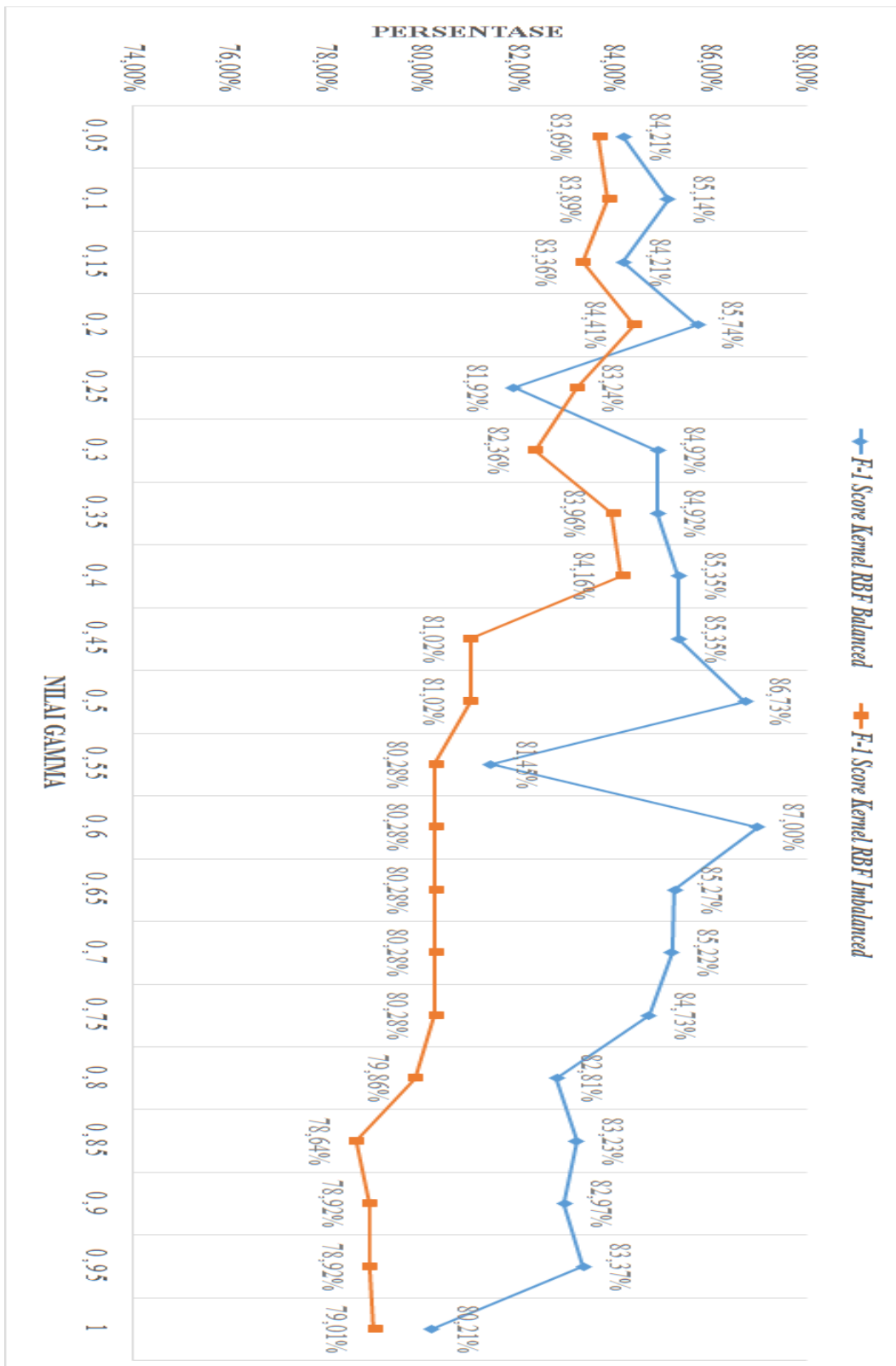
Berikut pada Gambar 4.11 hasil perhitungan *recall* dalam bentuk grafik baik menggunakan *kernel RBF balanced data set* dan *imbalanced data set*.



Gambar 4.11 Grafik Hasil Perhitungan Recall Kernel RBF Balanced Data Set dan Imbalanced

Pada grafik Gambar 4.11 menunjukkan bahwa *recall* dengan *kernel RBF balanced* dari gamma 0,5 sampai 1 terlihat persentase mendominasi jika dibandingkan dengan *imbalanced data set* dan persentase mengalami naik turun yang cukup signifikan. Sedangkan *recall* dengan *kernel RBF imbalanced* dari gamma 0,5 sampai 1 mengalami naik turun yang tidak terlalu signifikan diawal lalu mengalami penurunan yang signifikan pada gamma 0,45 dan persentase cenderung stabil pada gamma 0,55 hingga 0,75 dan gamma 0,85 hingga 1. Untuk nilai *recall* yang paling tinggi terdapat pada gamma 0,6 dengan nilai persentase sebesar 85,98% menggunakan fungsi *kernel RBF* dengan *balanced data set*.

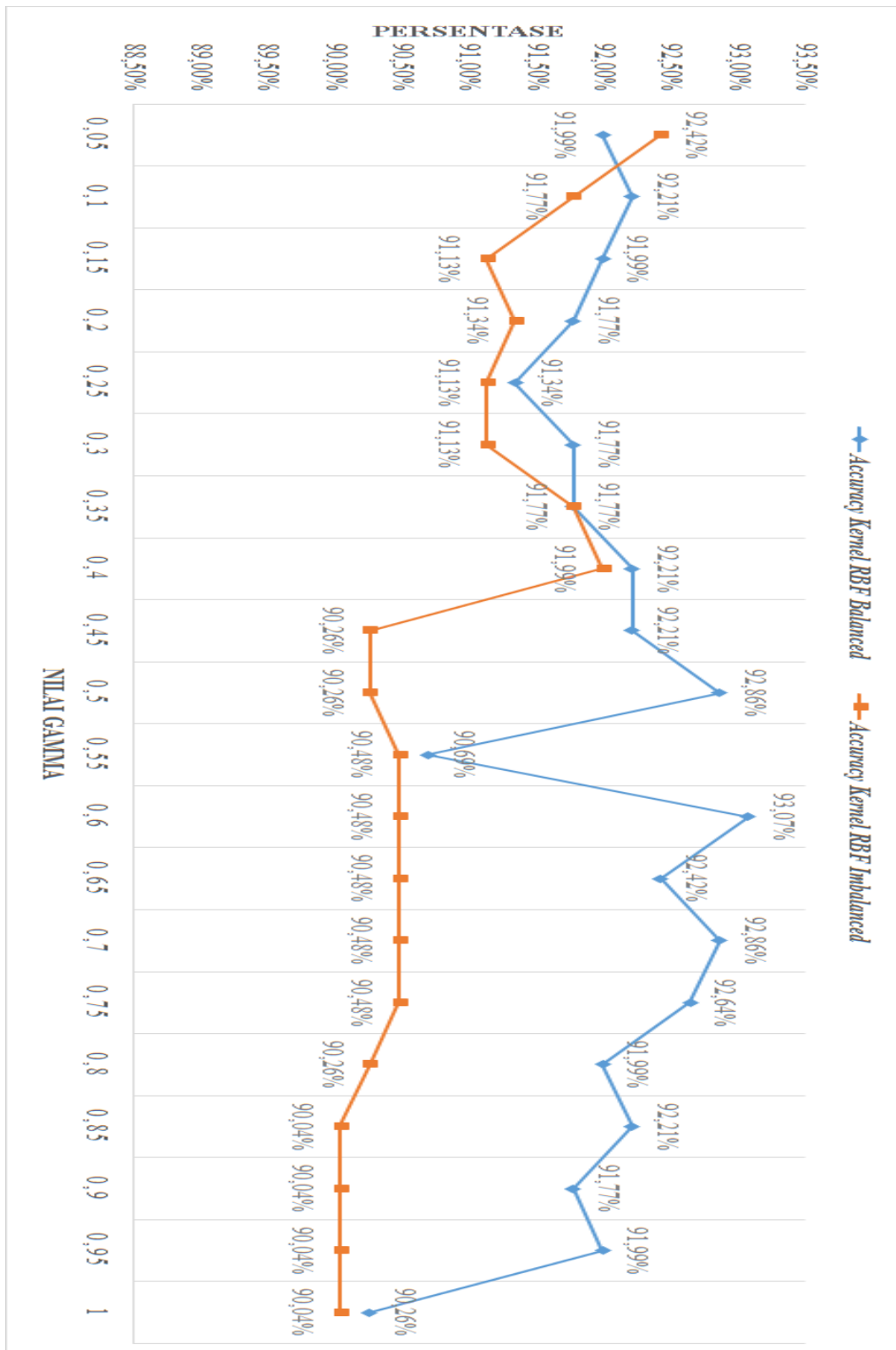
Berikut pada Gambar 4.12 hasil perhitungan *f1-score* dalam bentuk grafik baik menggunakan *kernel RBF balanced data set* dan *imbalanced data set*.



Gambar 4.12 Grafik Hasil Perhitungan *F1-score Kernel RBF Balanced Data Set dan Imbalanced*

Pada grafik Gambar 4.12 menunjukkan bahwa *f1-score* dengan *kernel RBF balanced* dari gamma 0,5 sampai 1 terlihat persentase mendominasi jika dibandingkan dengan *imbalanced data set* dan persentase mengalami naik turun yang cukup signifikan. Sedangkan *f1-score* dengan *kernel RBF imbalanced* dari gamma 0,5 sampai 1 mengalami naik turun yang tidak terlalu signifikan diawal lalu mengalami penurunan yang signifikan pada gamma 0,45 dan persentase cenderung stabil pada gamma 0,55 hingga 0,75 dan gamma 0,9 hingga 0,95. Untuk nilai *f1-score* yang paling tinggi terdapat pada gamma 0,6 dengan nilai persentase sebesar 87,00% menggunakan fungsi *kernel RBF* dengan *balanced data set*.

Berikut pada Gambar 4.13 hasil perhitungan akurasi dalam bentuk grafik baik menggunakan *kernel RBF balanced data set* dan *imbalanced data set*.



Gambar 4.13 Grafik Hasil Perhitungan Akurasi *Kernel RBF Balanced Data Set* dan *Imbalanced*

Pada grafik Gambar 4.13 menunjukkan bahwa akurasi dengan *kernel RBF balanced* dari gamma 0,5 sampai 1 terlihat akurasi mendominasi jika dibandingkan dengan *imbalanced data set* tetapi akurasi mengalami naik turun yang cukup signifikan. Sedangkan akurasi dengan *kernel RBF imbalanced* dari gamma 0,5 sampai 1 mengalami naik turun yang tidak terlalu signifikan diawal lalu mengalami penurunan yang signifikan pada gamma 0,45 dan akurasi cenderung stabil pada gamma 0,55 hingga 0,75 dan gamma 0,9 hingga 1. Untuk nilai akurasi yang paling tinggi terdapat pada gamma 0,6 dengan nilai sebesar 93,07% menggunakan fungsi *kernel RBF* dengan *balanced data set*.

Maka hasil perhitungan pada Gambar 4.10 hingga Gambar 4.13 menggunakan *kernel RBF* dengan *balanced data set* dan *imbalanced data set* dalam bentuk grafik, menunjukkan bahwa nilai akurasi, *precision*, *recall* dan *f1-score* yang paling tinggi sebesar 93,07%, 90,12%, 85,98% dan 87,00% terdapat pada gamma 0,6 dengan *balanced data set*.

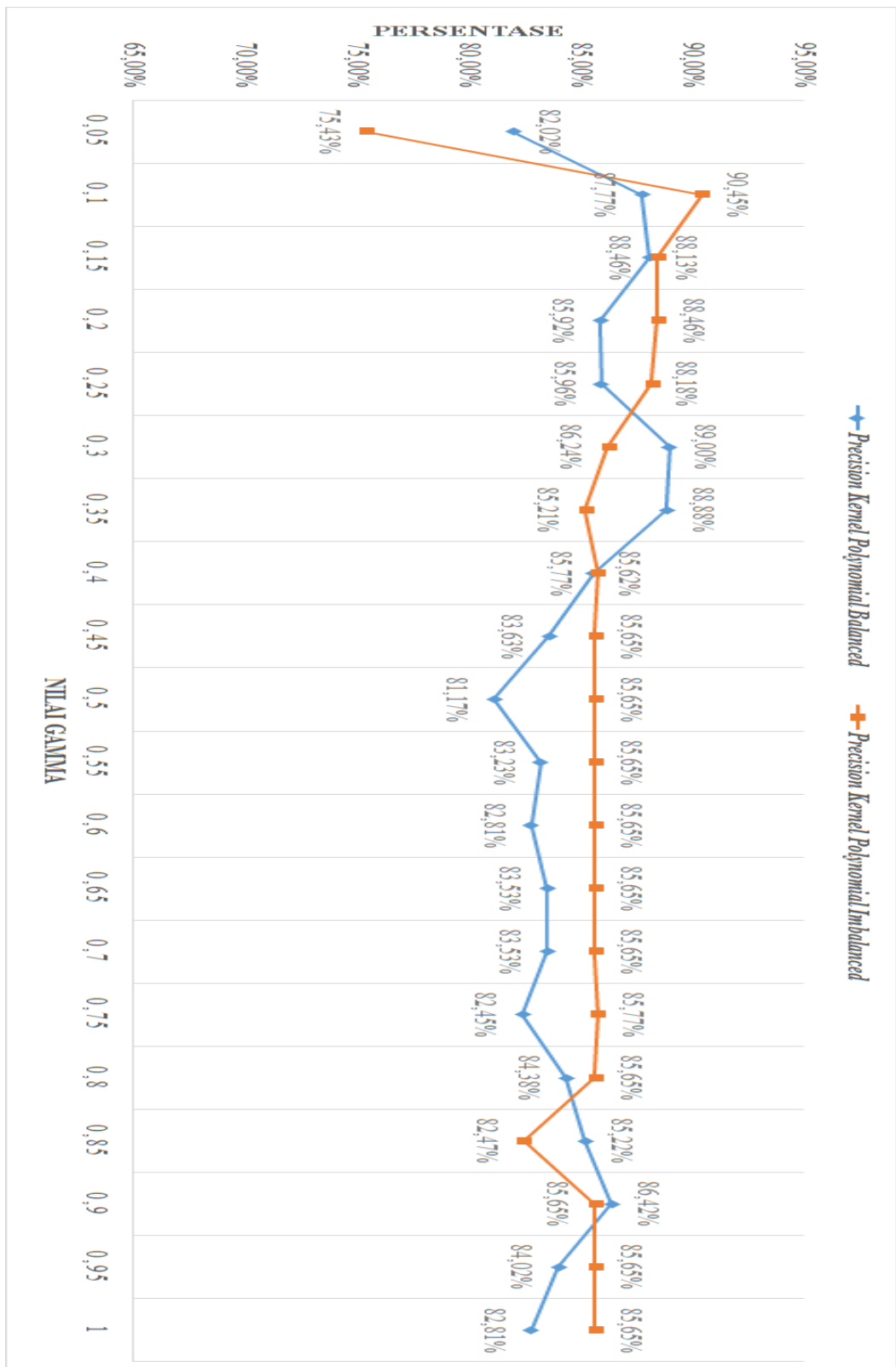
Selanjutnya merupakan tabel *confusion matrix* dengan nilai yang paling tinggi berdasarkan Gambar 4.10 sampai Gambar 4.13 Untuk hasil perhitungan tabel *confusion matrix* dengan *kernel RBF balanced data set* serta *imbalanced data set* lainnya, bisa dilihat pada Lampiran E. Berikut pada Tabel 4.7 merupakan nilai yang paling tinggi terdapat pada fungsi *kernel RBF balanced data set* dengan nilai gamma=0,6 sebagai berikut.

Tabel 4.7 Hasil Confusion Matrix Kernel RBF Nilai Gamma=0,6 Balanced

		Hasil Klasifikasi <i>Kernel RBF</i> (Gamma=0,6) Dengan <i>Balanced</i>						
		1	2	3	4	5	6	7
Nilai Sebenarnya	1	110	1	0	0	0	0	0
	2	0	140	0	0	0	0	2
	3	0	0	70	0	3	3	0
	4	0	0	2	8	2	2	0
	5	0	0	5	0	26	5	0
	6	0	0	0	0	0	56	5
	7	1	0	0	0	0	1	20

Nilai Sebenarnya	1	110	1	0	0	0	0	0
	2	0	138	0	0	0	1	3
	3	0	0	72	0	0	4	0
	4	0	0	1	8	2	2	1
	5	0	0	5	0	26	5	0
	6	0	3	0	0	0	53	5
	7	1	0	0	0	0	0	21

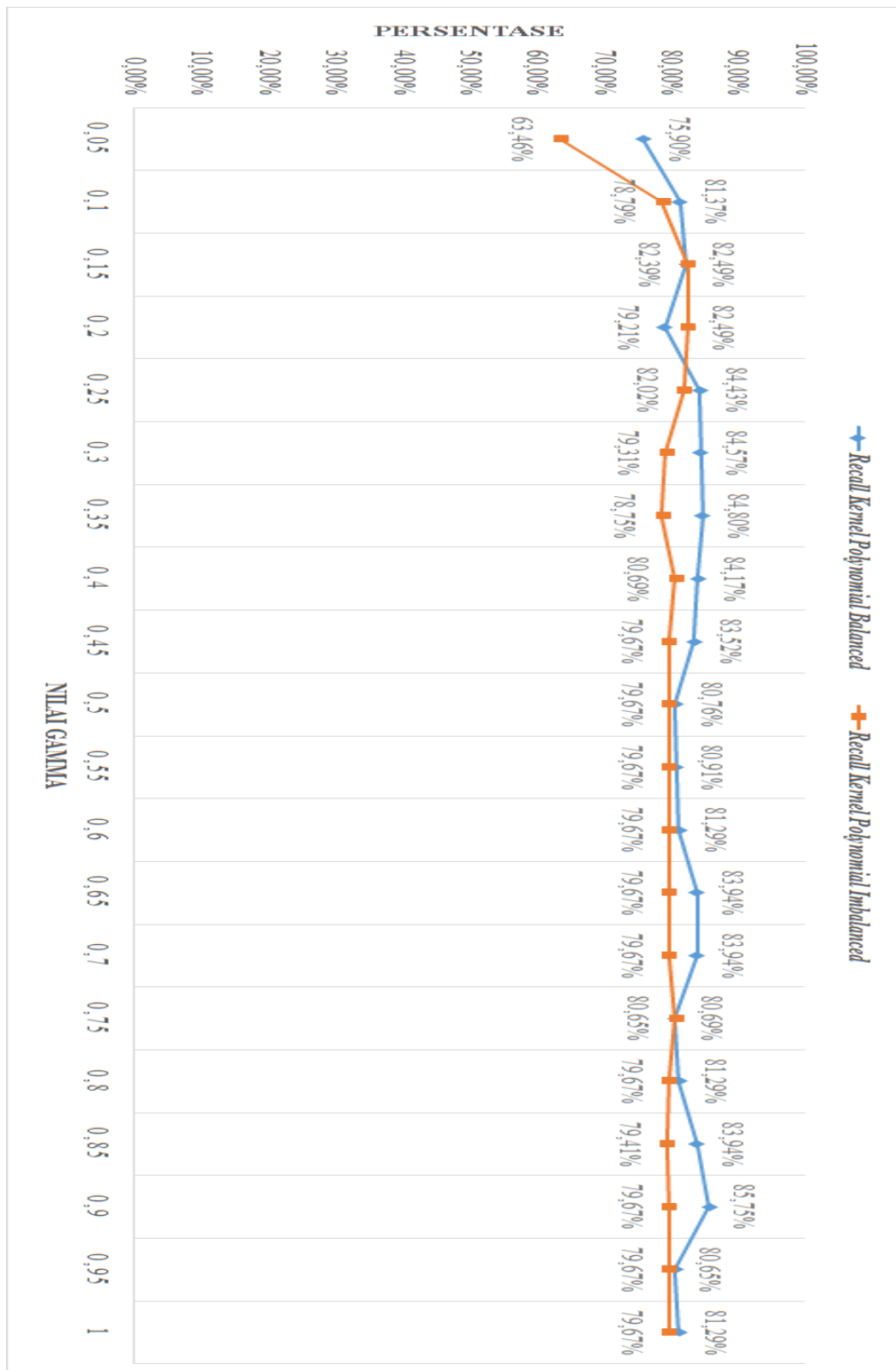
Terakhir untuk Gambar 4.15 sampai 4.18 yaitu grafik hasil perhitungan *precision*, *recall*, *f1-score* dan akurasi menggunakan fungsi *kernel polynomial* dengan menggunakan *balanced data set* dan *imbalanced data set* dengan parameter masukan berupa nilai gamma kelipatan 0,05 sebanyak 20 kali. Mengenai hasil perhitungan yang lebih detail dapat dilihat pada Lampiran E. Berikut pada Gambar 4.15 hasil perhitungan *precision* dalam bentuk grafik baik menggunakan *kernel polynomial balanced data set* dan *imbalanced data set*.



Gambar 4.15 Grafik Hasil Perhitungan *Precision Kernel Polynomial* *Balanced Data Set* dan *Imbalanced*

Pada grafik Gambar 4.15 menunjukkan bahwa *precision* dengan *kernel polynomial balanced* dari gamma 0,5 sampai 1 terlihat persentase mengalami naik turun yang cukup signifikan. Sedangkan *precision* dengan *kernel polynomial imbalanced* dari gamma 0,5 sampai 1 mengalami kenaikan yang cukup signifikan pada awal lalu mengalami penurunan hingga pada gamma 0,4 persentase cenderung stabil hingga gamma 0,8 dan gamma 0,9 sampai gamma 1. Untuk nilai *precision* yang tertinggi terdapat pada gamma 0,1 dengan nilai persentase sebesar 90,45% menggunakan fungsi *kernel polynomial* dengan *imbalanced data set*.

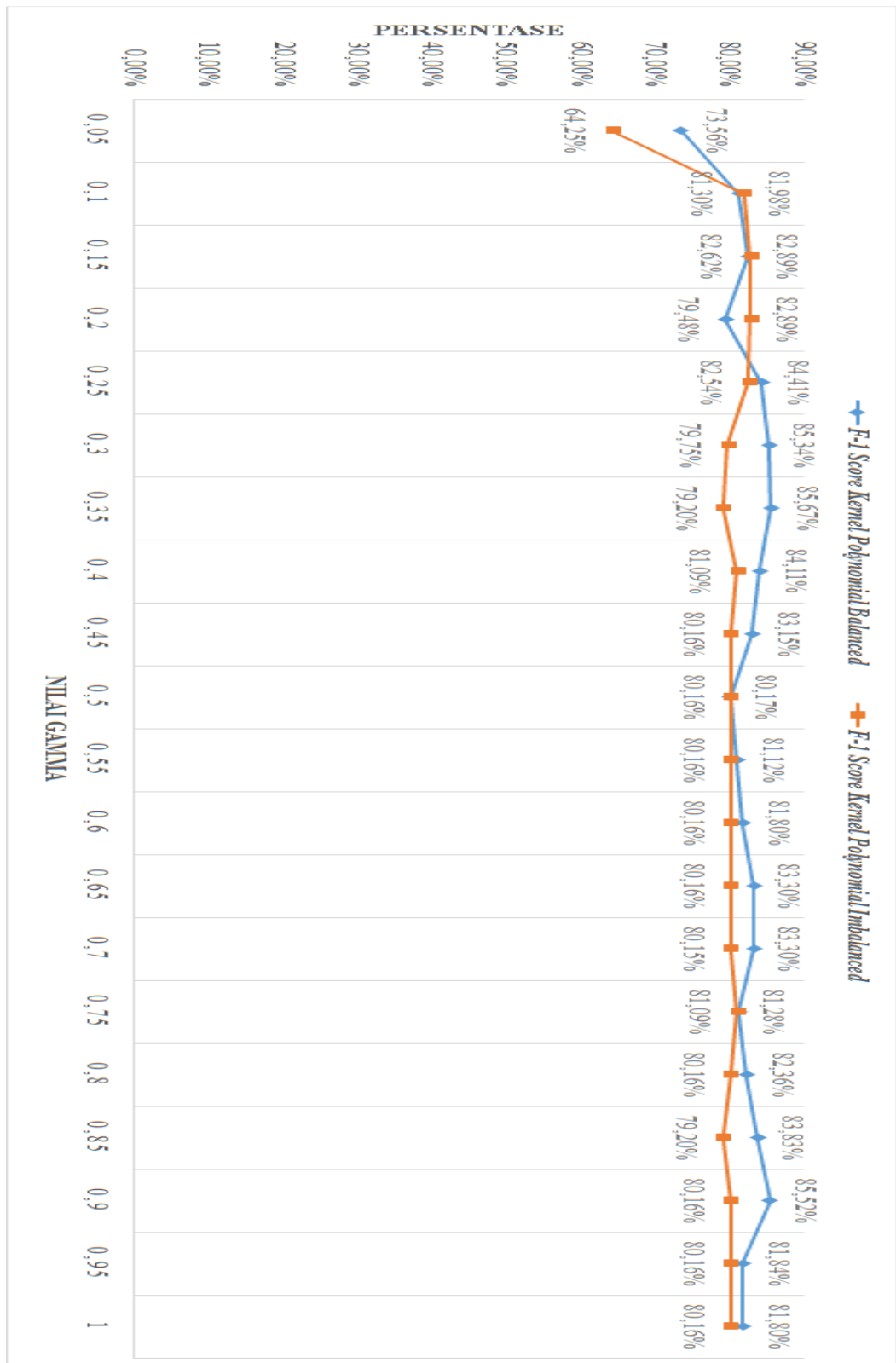
. Berikut pada Gambar 4.16 hasil perhitungan *recall* dalam bentuk grafik baik menggunakan *kernel polynomial balanced data set* dan *imbalanced data set*.



Gambar 4.16 Grafik Hasil Perhitungan *Recall Kernel Polynomial Balanced* Data Set dan Imbalanced

Pada grafik Gambar 4.16 menunjukkan bahwa *recall* dengan *kernel polynomial balanced* dari gamma 0,5 sampai 1 persentase mendominasi dibandingkan dengan *kernel polynomial imbalanced*. Sedangkan *recall* dengan *kernel polynomial imbalanced* dari gamma 0,5 sampai 1 pada awal mengalami persentase paling rendah lalu mengalami kenaikan hingga pada gamma 0,3 sampai gamma 1 persentase cenderung stabil. Untuk nilai *recall* yang tertinggi terdapat pada gamma 0,9 dengan nilai persentase sebesar 85,75% menggunakan fungsi *kernel polynomial* dengan *balanced data set*.

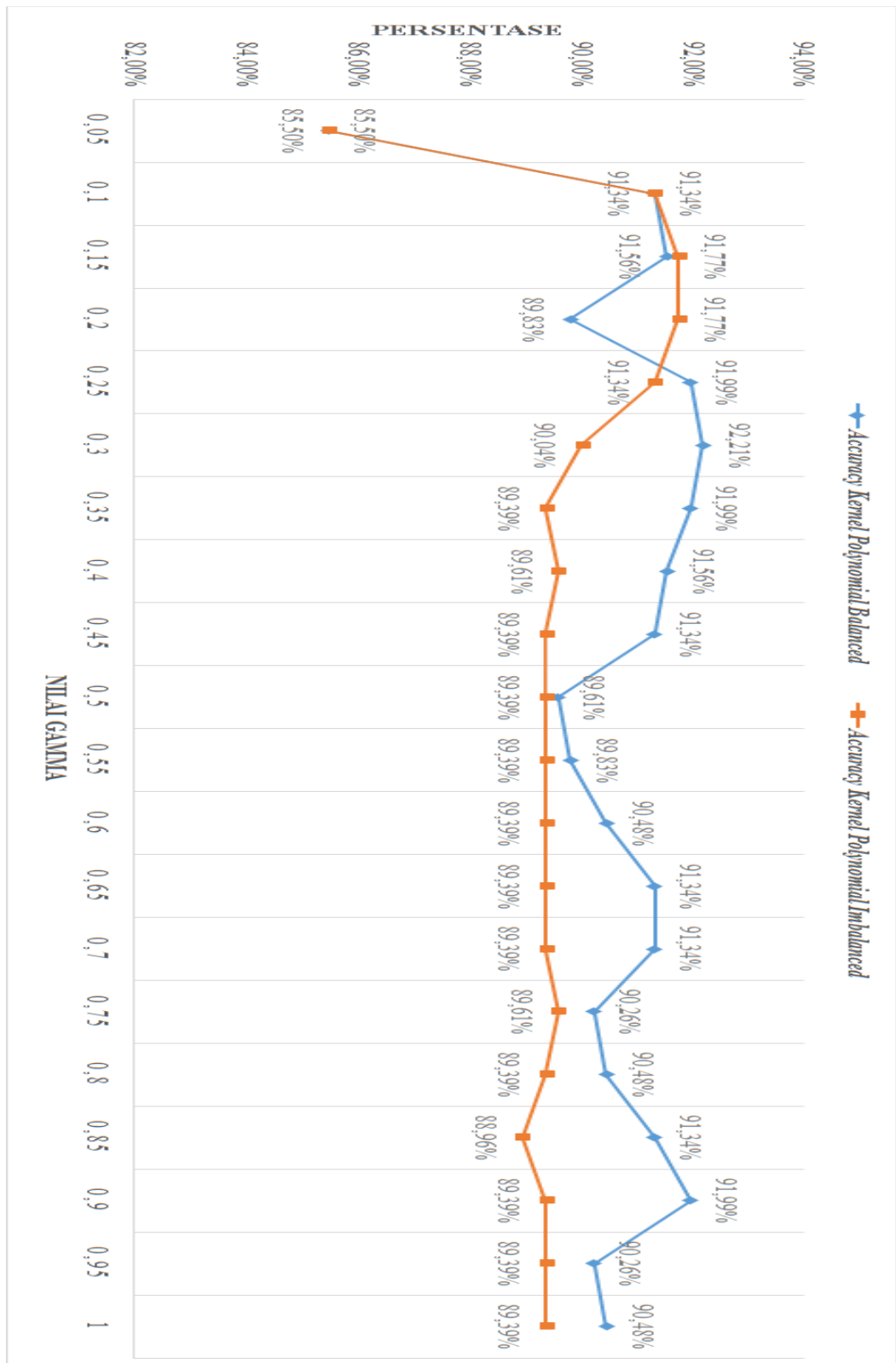
. Berikut pada Gambar 4.17 hasil perhitungan *f1-score* dalam bentuk grafik baik menggunakan *kernel polynomial balanced data set* dan *imbalanced data set*.



Gambar 4.17 Grafik Hasil Perhitungan *F1-score Kernel Polynomial Balanced Data Set dan Imbalanced*

Pada grafik Gambar 4.17 menunjukkan bahwa *f1-score* dengan *kernel polynomial balanced* dari gamma 0,5 sampai 1 persentase mendominasi dibandingkan dengan *kernel polynomial imbalanced*. Sedangkan *f1-score* dengan *kernel polynomial imbalanced* dari gamma 0,5 sampai 1 pada awal mengalami persentase paling rendah lalu mengalami kenaikan hingga pada gamma 0,4 sampai gamma 0,7 dan gamma 0,9 hingga gamma 1 persentase cenderung stabil. Untuk nilai *f1-score* yang tertinggi terdapat pada gamma 0,35 dengan nilai persentase sebesar 85,67% menggunakan fungsi *kernel polynomial* dengan *balanced data set*.

. Berikut pada Gambar 4.18 hasil perhitungan akurasi dalam bentuk grafik baik menggunakan *kernel polynomial balanced data set* dan *imbalanced data set*.



Gambar 4.18 Grafik Hasil Perhitungan Akurasi Kernel Polynomial Balanced Data Set dan Imbalanced

Pada grafik Gambar 4.18 menunjukkan bahwa akurasi dengan *kernel polynomial balanced* dari gamma 0,5 sampai 1 pada awal akurasi mendapatkan nilai rendah lalu mengalami kenaikan yang signifikan dan grafik mengalami kenaikan dan penurunan dari gamma 0,1 hingga gamma 1. Sedangkan *f1-score* dengan *kernel polynomial imbalanced* dari gamma 0,5 sampai 1 pada awal mendapatkan nilai rendah lalu mengalami kenaikan dan penurunan pada gamma 0,1 sampai gamma 0,35 ,dilanjutkan dengan nilai akurasi yang stabil terdapat gamma 0,45 hingga gamma 0,7. Untuk nilai akurasi yang tertinggi terdapat pada gamma 0,3 dengan nilai persentase sebesar 92,21% menggunakan fungsi *kernel polynomial* dengan *balanced data set*.

Maka grafik hasil perhitungan pada Gambar 4.15 sampai Gambar 4.18 menggunakan *kernel polynomial* dengan *balanced data set* dan *imbalanced data set* dalam bentuk grafik, menunjukkan bahwa nilai akurasi, *precision*, *recall* dan *f1-score* yang paling tinggi sebesar 91,99%, 88,88%, 84,80% dan 85,67% terdapat pada gamma 0,35 dengan *balanced data set*. Untuk gamma 0,1 dengan *imbalanced data set* mempunyai *precision* tertinggi 90,45% tidak disebut paling tinggi karena nilai *recall* pada gamma 0,1 memiliki nilai yang rendah sebesar 78,79% dan nilai *f1-score* sebesar 81,98% jika dibandingkan pada gamma 0,35 menggunakan *balanced data set*. Lalu gamma 0,9 dengan *balanced data set* mempunyai *recall* tertinggi sebesar 85,75% tidak disebut paling tinggi karena nilai *precision* pada gamma 0,9 memiliki nilai yang rendah sebesar 86,42% dan nilai *f1-score* sebesar 85,52% jika dibandingkan pada gamma 0,35 menggunakan *balanced data set*. Sedangkan gamma 0,3 dengan *balanced data set* mempunyai akurasi tertinggi sebesar 92,21% tidak disebut paling tinggi karena nilai *precision*, *recall* dan *f1-score* pada gamma 0,3 memiliki nilai yang rendah sebesar 89,00%, 84,57% dan 85,34% jika dibandingkan pada gamma 0,35 menggunakan *balanced data set*.

Selanjutnya merupakan tabel *confusion matrix* dengan nilai yang paling tinggi berdasarkan Gambar 4.15 sampai Gambar 4.18. Untuk hasil perhitungan tabel *confusion matrix* dengan *kernel polynomial balanced data set* serta *imbalanced data set* lainnya, bisa dilihat pada Lampiran E. Berikut pada Tabel 4.9

merupakan nilai yang paling tinggi terdapat pada fungsi *kernel polynomial balanced data set* dengan nilai $\gamma=0,35$ sebagai berikut.

**Tabel 4.9 Hasil Confusion Matrix Kernel Polynomial Nilai Gamma=0,35
Balanced**

		Hasil Klasifikasi <i>Kernel Polynomial</i> (Gamma=0,35) Dengan <i>Balanced</i>						
		1	2	3	4	5	6	7
Nilai Sebenarnya	1	110	1	0	0	0	0	0
	2	0	138	0	0	0	1	3
	3	0	0	69	0	3	4	0
	4	0	0	2	8	2	2	0
	5	0	0	7	0	24	5	0
	6	0	0	0	0	0	56	5
	7	1	0	0	0	0	1	20

4.2.4 Kesimpulan Pengujian

Berdasarkan hasil perhitungan *confusion matrix* dengan 462 token kalimat data uji dan 718 token kalimat data latih, nilai yang paling tinggi diantara fungsi *kernel RBF*, *linear* dan *polynomial* baik menggunakan *balanced data set* dan *imbalanced data set* terdapat pada fungsi *kernel RBF* dengan nilai $\gamma=0,6$ menggunakan *balanced data set*, serta nilai akurasi, *precision*, *recall* dan *f1-score* adalah sebesar 93,07%, 90,12%, 85,98% dan 87,00%. Dari analisis yang dilakukan pada hasil pengujian *confusion matrix* ditemukan bahwa data yang seimbang atau *balanced data set* berhasil meningkatkan kinerja *SVM*, berbanding terbalik dengan data yang tidak seimbang atau *imbalanced data set* yang dapat mengurangi kinerja *SVM*. Serta pada penelitian ini hasil yang paling baik dimiliki oleh fungsi *kernel RBF* dengan *balanced data set*.

Dapat dilihat pula penurunan akurasi, *precision*, *recall* dan *f1-score* disebabkan oleh proses konversi dari *PDF* ke *HTML*, data yang seharusnya angka tetapi terkonversi menjadi huruf berpengaruh besar terhadap proses pembobotan karena dalam penelitian ini terdapat ekstraksi fitur untuk mengenali angka. Selain itu data kelas 6 (Kemampuan) terklasifikasi menjadi kelas 7 (Lain-lain) dan kelas 3 (Pendidikan formal) serta kelas 4 (Pendidikan non formal) terklasifikasi menjadi

kelas 5 (Pengalaman kerja) disebabkan nilai pembobotan yang memiliki bobot yang hampir sama. Oleh karena itu perlunya penambahan fitur yang lebih spesifik untuk menghindari nilai pembobotan antar kelas yang hampir sama, sehingga dapat meminimalisir penurunan akurasi.