BAB 4
IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Implementasi Sistem

Tahap implementasi sistem merupakan tahap pembangunan aplikasi, tahap lanjut dari perancangan sistem. Tahap ini dilakukan untuk menerjemahkan perancangan berdasarkan hasil analisis dalam bahasa yang dimengerti oleh komputer. Adapun implementasi sistem terdiri dari implementasi perangkat keras, implementasi perangkat lunak, implementasi basis data dan implementasi antarmuka.

4.1.1 Implementasi Perangkat Keras

Spesifikasi perangkat keras yang digunakan untuk membangun aplikasi ini yaitu sebagai berikut:

1. *Processor* 2,27 GHz
2. RAM 4 GB
3. Layar dengan resolusi 1366 x 768 pixel
4. *Keyboard* standar
5. *Mouse* standar

4.1.2 Implementasi Perangkat Lunak

Spesifikasi perangkat lunak yang digunakan untuk membangun aplikasi ini yaitu sebagai berikut:

1. Sistem operasi Windows 8.1 Professional 32 Bit
2. Bahasa pemrograman PHP
3. Web server XAMPP 1.8.3
4. Editor Sublime Text 3
5. DBMS MySQL 5.5
6. Framework Twitter Bootstrap
7. Web Browser Google Chrome

4.1.3 Implementasi Basis Data

Pembuatan basis data, dilakukan dengan menggunakan MySQL. Implementasi basis data dalam bahasa SQL seperti yang ditampilkan pada tabel 4.1 berikut.

Tabel 4.1 Implementasi basis data

No	Proses	SQL
1	Pembuatan basis data	`CREATE DATABASE analisis_sentimen2;`
2	Pembuatan tabel Tweet (tbl_tweet)	`CREATE TABLE IF NOT EXISTS `tbl_tweet` ( `id_tweet` char(18) NOT NULL, `tweet` varchar(140) NOT NULL, `sentimen` enum('0','1') NOT NULL, `dataset` int(2) NOT NULL, PRIMARY KEY (`id_tweet`) ) ENGINE=MyISAM;`
3	Pembuatan tabel Training (tbl_training)	`CREATE TABLE IF NOT EXISTS `tbl_training` ( `id_fitur` int(11) NOT NULL AUTO_INCREMENT, `fitur` varchar(10) NOT NULL, `jml_positif` int(4) NOT NULL, `jml_negatif` int(4) NOT NULL, `prob_positif` double NOT NULL, `prob_negatif` double NOT NULL, `n_gram` int(2) NOT NULL, PRIMARY KEY (`id_fitur`) ) ENGINE=MyISAM;`
8	Pembuatan tabel Testing (tbl_testing)	`CREATE TABLE IF NOT EXISTS `tbl_testing` ( `id_tweet` char(18) NOT NULL, `tweet` varchar(140) NOT NULL, `sentimen_manual` enum('0','1') NOT NULL, `sentimen_sistem0` enum('0','1','2') NOT NULL, `sentimen_sistem2` enum('0','1','2') NOT NULL, `sentimen_sistem3` enum('0','1','2') NOT NULL, `sentimen_sistem4` enum('0','1','2') NOT NULL, `sentimen_sistem5` enum('0','1','2') NOT NULL, `sentimen_sistem6` enum('0','1','2') NOT NULL, `sentimen_sistem7` enum('0','1','2') NOT NULL, `sentimen_sistem8` enum('0','1','2') NOT NULL, `sentimen_sistem9` enum('0','1','2') NOT NULL,`

		`sentimen_sistem10` enum('0','1','2') NOT NULL, PRIMARY KEY (`id_tweet`)) ENGINE=MyISAM;
9	Pembuatan tabel Stopword (tbl_stopword)	CREATE TABLE IF NOT EXISTS `tbl_stopword` (`id` int(5) NOT NULL AUTO_INCREMENT, `stopword` varchar(10) NOT NULL, PRIMARY KEY (`id`)) ENGINE=MyISAM;
10	Pembuatan tabel Kata Dasar (tbl_katadasar)	CREATE TABLE IF NOT EXISTS `tbl_katadasar` (`id_katadasar` int(5) NOT NULL AUTO_INCREMENT, `katadasar` varchar(20) NOT NULL, PRIMARY KEY (`id_katadasar`)) ENGINE=MyISAM;
11	Pembuatan tabel Emoticon (tbl_emoticon)	CREATE TABLE IF NOT EXISTS `tbl_emoticon` (`id_emoticon` int(3) NOT NULL AUTO_INCREMENT, `emoticon` varchar(3) NOT NULL, `text` varchar(6) NOT NULL, PRIMARY KEY (`id_emoticon`)) ENGINE=MyISAM;
12	Pembuatan tabel Cross Validation (tbl_cross_validation)	CREATE TABLE IF NOT EXISTS `tbl_cross_validation` (`id_cross` char(3) NOT NULL, `tp` int(3) NOT NULL, `tn` int(3) NOT NULL, `fp` int(3) NOT NULL, `fn` int(3) NOT NULL, `akurasi` int(3) NOT NULL, PRIMARY KEY (`id_cross`)) ENGINE= MyISAM;

4.1.4 Implementasi Antarmuka

Implementasi antarmuka dilakukan untuk setiap tampilan aplikasi yang dibangun dalam bentuk file code. Deskripsi implementasi antarmuka pada aplikasi yang dibangun dapat dilihat pada tabel 4.2 berikut:

Tabel 4.2 Implementasi antarmuka

No	Nama antarmuka	Dekripsi	Nama file
1	Pengambilan Tweet	Menampilkan halaman Pengambilan Tweet.	index.php
2	Training	Menampilkan halaman Training	training.php

3	Testing	Menampilkan halaman Testing	testing.php
4	Penghitungan Akurasi	Menampilkan halaman Penghitungan Akurasi	akurasi.php
5	Cross Validation	Menampilkan halaman Cross Validation	cross_validation.php

4.2 Skenario Pengujian Sistem

Pengujian sistem ditujukan untuk mengukur tingkat akurasi klasifikasi dari penggunaan masing-masing jenis fitur (fitur kata dan fitur n-gram). Adapun pengujian yang dilakukan menggunakan *k-fold cross validation* dengan nilai k=10. Nilai 10 menunjukan jumlah dataset (D), dimana dataset (D) dibagi menjadi 10 bagian yaitu D_1, D_2,..,D_{10} dan iterasi (pengujian) dilakukan sebanyak sepuluh kali.

Jumlah data yang digunakan sebanyak 1000 data. Data ini merupakan data yang telah dikelompokan secara manual ke dalam kelas positif dan negatif dengan jumlah sebesar 500 data untuk masing-masing kelasnya. Tiap dataset (D) terdiri dari 100 data. Percobaan pengujian dilakukan sebanyak sepuluh kali, dimana tiap dataset akan secara bergantian menjadi data *training* dan data *testing*. Skenario pengujian yang akan dilakukan ditampilkan pada Tabel 4.3 berikut.

Tabel 4.3 Skenario pengujian *10-fold cross validation*

Pengujian (fold)	**Data *Testing***	**Data *Training***
1	D_1	$D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$
2	D_2	$D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$
3	D_3	$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$
4	D_4	$D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$
5	D_5	$D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$
6	D_6	$D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$
7	D_7	$D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$
8	D_8	$D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$
9	D_9	$D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$
10	D_{10}	$D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$

4.3 Pengujian Sistem

Pengujian sistem dilakukan sebanyak sepuluh kali dengan skenario yang telah ditampilkan pada Tabel 4.3. Berikut ini diuraikan masing-masing dari pengujian yang dilakukan.

4.3.1 Pengujian Sistem dengan Fitur Kata

Pada pengujian ini fitur yang digunakan adalah fitur kata. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.4 berikut.

Tabel 4.4 *Confussion matrix* pengujian satu fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	4
	Negatif	4	49

2. Pengujian Dua

Pada pengujian dua dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.5 berikut.

Tabel 4.5 *Confussion matrix* pengujian dua fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	47	4
	Negatif	0	49

3. Pengujian Tiga

Pada pengujian tiga dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.6 berikut.

Tabel 4.6 *Confussion matrix* pengujian tiga fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	47	3
	Negatif	2	48

4. Pengujian Empat

Pada pengujian empat dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.7 berikut.

Tabel 4.7 *Confussion matrix* pengujian empat fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	45	4
	Negatif	7	44

5. Pengujian Lima

Pada pengujian lima dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.8 berikut.

Tabel 4.8 *Confussion matrix* pengujian lima fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	3
	Negatif	8	41

6. Pengujian Enam

Pada pengujian enam dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.9 berikut.

Tabel 4.9 *Confussion matrix* pengujian enam fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	56	1
	Negatif	4	39

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.10 berikut.

Tabel 4.10 *Confussion matrix* pengujian tujuh fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	5
	Negatif	8	44

8. Pengujian Delapan

Pada pengujian delapan dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.11 berikut.

Tabel 4.11 *Confussion matrix* pengujian delapan fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	40	4
	Negatif	2	50

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.12 berikut.

Tabel 4.12 *Confussion matrix* pengujian sembilan fitur kata

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	52	4
	Negatif	5	39

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur kata, digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur kata menggunakan *10-fold cross validation* disajikan pada Tabel 4.13 berikut.

Tabel 4.13 *Confussion matrix* pengujian sepuluh fitur kata

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	43	0
	Negatif	4	53

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.14 berikut.

Tabel 4.14 Rekap hasil pengujian fitur kata

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	43	4	49	4	92
2	47	4	49	0	96
3	47	3	48	2	95
4	45	4	44	7	89
5	48	3	41	8	89
6	56	1	39	4	95
7	43	5	44	8	87
8	44	4	50	2	94
9	52	4	39	5	91
10	43	0	53	4	96
Rata-rata akurasi					**92 %**

4.3.2 Pengujian Sistem dengan Fitur n-gram(n=2)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=2. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.15 berikut.

Tabel 4.15 *Confussion matrix* pengujian satu fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	45	2
	Negatif	6	47

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.16 berikut.

Tabel 4.16 *Confussion matrix* pengujian dua fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	2
	Negatif	4	45

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.17 berikut.

Tabel 4.17 *Confussion matrix* pengujian tiga fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	1
	Negatif	7	43

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.18 berikut.

Tabel 4.18 *Confussion matrix* pengujian empat fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	45	4
	Negatif	7	44

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=2) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.19 berikut.

Tabel 4.19 *Confussion matrix* pengujian lima fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	2
	Negatif	5	44

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.20 berikut.

Tabel 4.20 *Confussion matrix* pengujian enam fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	56	1
	Negatif	5	38

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.21 berikut.

Tabel 4.21 *Confussion matrix* pengujian tujuh fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	2
	Negatif	12	40

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.22 berikut.

Tabel 4.22 *Confussion matrix* pengujian delapan fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	47	1
	Negatif	6	46

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.23 berikut.

Tabel 4.23 *Confussion matrix* pengujian sembilan fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	55	1
	Negatif	10	34

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=2), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=2) menggunakan *10-fold cross validation* disajikan pada Tabel 4.24 berikut.

Tabel 4.24 *Confussion matrix* pengujian sepuluh fitur n-gram(n=2)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	6	51

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.25 berikut.

Tabel 4.25 Rekap hasil pengujian fitur n-gram(n=2)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	45	2	47	6	92
2	49	2	45	4	94
3	49	1	43	7	92
4	45	4	44	7	89

5	49	2	44	5	93
6	56	1	38	5	94
7	46	2	40	12	86
8	47	1	46	6	93
9	55	1	34	10	89
10	43	0	51	6	94
Rata-rata akurasi					**92 %**

4.3.3 Pengujian Sistem dengan Fitur n-gram(n=3)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=3. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.26 berikut.

Tabel 4.26 *Confussion matrix* pengujian satu fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	1
	Negatif	5	48

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.27 berikut.

Tabel 4.27 *Confussion matrix* pengujian dua fitur n-gram(n=3)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	50	1
	Negatif	2	47

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.28 berikut.

Tabel 4.28 *Confussion matrix* pengujian tiga fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	1
	Negatif	3	47

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.29 berikut.

Tabel 4.29 *Confussion matrix* pengujian empat fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	0
	Negatif	6	45

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=3) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5.

Hasil dari pengujian lima dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.30 berikut.

Tabel 4.30 *Confussion matrix* pengujian lima fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	0
	Negatif	6	43

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.31 berikut.

Tabel 4.31 *Confussion matrix* pengujian enam fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	56	1
	Negatif	6	37

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.32 berikut.

Tabel 4.32 *Confussion matrix* pengujian tujuh fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	0
	Negatif	10	42

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.33 berikut.

Tabel 4.33 *Confussion matrix* pengujian delapan fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	2
	Negatif	4	48

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.34 berikut.

Tabel 4.34 *Confussion matrix* pengujian sembilan fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	56	0
	Negatif	7	37

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=3), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=3) menggunakan *10-fold cross validation* disajikan pada Tabel 4.35 berikut.

Tabel 4.35 *Confussion matrix* pengujian sepuluh fitur n-gram(n=3)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	7	50

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.36 berikut.

Tabel 4.36 Rekap hasil pengujian fitur n-gram(n=3)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	46	1	48	5	94
2	50	1	47	2	97
3	49	1	47	3	96
4	49	0	45	6	94
5	51	0	43	6	94
6	56	1	37	6	93
7	48	0	42	10	90
8	46	2	48	4	94
9	56	0	37	7	93
10	43	0	50	7	93
Rata-rata akurasi					**94 %**

4.3.4 Pengujian Sistem dengan Fitur n-gram(n=4)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=4. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.37 berikut.

Tabel 4.37 *Confussion matrix* pengujian satu fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	1
	Negatif	4	49

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.38 berikut.

Tabel 4.38 *Confussion matrix* pengujian dua fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	50	1
	Negatif	1	48

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.39 berikut.

Tabel 4.39 *Confussion matrix* pengujian tiga fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	1
	Negatif	3	47

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 3.40 berikut.

Tabel 4.40 *Confussion matrix* pengujian empat fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	1
	Negatif	4	47

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=4) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.41 berikut.

Tabel 4.41 *Confussion matrix* pengujian lima fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	0
	Negatif	5	44

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.42 berikut.

Tabel 4.42 *Confussion matrix* pengujian enam fitur n-gram(n=4)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	57	0
	Negatif	4	39

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.43 berikut.

Tabel 4.43 *Confussion matrix* pengujian tujuh fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	0
	Negatif	6	46

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.44 berikut.

Tabel 4.44 *Confussion matrix* pengujian delapan fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	2
	Negatif	1	51

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9.

Hasil dari pengujian sembilan dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.45 berikut.

Tabel 4.45 *Confussion matrix* pengujian sembilan fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	56	0
	Negatif	6	38

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=4), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=4) menggunakan *10-fold cross validation* disajikan pada Tabel 4.46 berikut.

Tabel 4.46 *Confussion matrix* pengujian sepuluh fitur n-gram(n=4)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	6	51

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.47 berikut.

Tabel 4.47 Rekap hasil pengujian fitur n-gram(n=4)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	46	1	49	4	95
2	50	1	48	1	98
3	49	1	47	3	96
4	48	1	47	4	95
5	51	0	44	5	95
6	57	0	39	4	96
7	48	0	46	6	94
8	46	2	51	1	97

9	56	0	38	6	94
10	43	0	51	6	94
Rata-rata akurasi					**95 %**

4.3.5 Pengujian Sistem dengan Fitur n-gram(n=5)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=5. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.48 berikut.

Tabel 4.48 *Confussion matrix* pengujian satu fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	1
	Negatif	4	49

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.49 berikut.

Tabel 4.49 *Confussion matrix* pengujian dua fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	50	1
	Negatif	1	48

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.50 berikut.

Tabel 4.50 *Confussion matrix* pengujian tiga fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	50	0
	Negatif	4	46

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.51 berikut.

Tabel 4.51 *Confussion matrix* pengujian empat fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	0
	Negatif	4	47

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=5) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.52 berikut.

Tabel 4.52 *Confussion matrix* pengujian lima fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	0
	Negatif	6	43

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.53 berikut.

Tabel 4.53 *Confussion matrix* pengujian enam fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	57	0
	Negatif	4	39

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.54 berikut.

Tabel 4.54 *Confussion matrix* pengujian tujuh fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	0
	Negatif	8	44

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.55 berikut.

Tabel 4.55 *Confussion matrix* pengujian delapan fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	2
	Negatif	1	51

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.56 berikut.

Tabel 4.56 *Confussion matrix* pengujian sembilan fitur n-gram(n=5)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	56	0
	Negatif	7	37

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=5), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=5) menggunakan *10-fold cross validation* disajikan pada Tabel 4.57 berikut.

Tabel 4.57 *Confussion matrix* pengujian sepuluh fitur n-gram(n=5)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	43	0
	Negatif	7	50

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.58 berikut.

Tabel 4.58 Rekap hasil pengujian fitur n-gram(n=5)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	46	1	49	4	95
2	50	1	48	1	98
3	50	0	46	4	96
4	49	0	47	4	96
5	51	0	43	6	94
6	57	0	39	4	96
7	48	0	44	8	92
8	46	2	51	1	97
9	56	0	37	7	93
10	43	0	50	7	93
Rata-rata akurasi					**95 %**

4.3.6 Pengujian Sistem dengan Fitur n-gram(n=6)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=6. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.59 berikut.

Tabel 4.59 *Confussion matrix* pengujian satu fitur n-gram(n=6)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	50	1
	Negatif	14	35

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.60 berikut.

Tabel 4.60 *Confussion matrix* pengujian dua fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	1
	Negatif	6	42

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.61 berikut.

Tabel 4.61 *Confussion matrix* pengujian tiga fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	47	1
	Negatif	5	47

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4.

Hasil dari pengujian empat dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.62 berikut.

Tabel 4.62 *Confussion matrix* pengujian empat fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	0
	Negatif	4	50

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=6) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.63 berikut.

Tabel 4.63 *Confussion matrix* pengujian lima fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	0
	Negatif	4	47

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.64 berikut.

Tabel 4.64 *Confussion matrix* pengujian enam fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	38	0
	Negatif	5	57

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.65 berikut.

Tabel 4.65 *Confussion matrix* pengujian tujuh fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	55	1
	Negatif	9	35

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.66 berikut.

Tabel 4.66 *Confussion matrix* pengujian delapan fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	2	55

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.67 berikut.

Tabel 4.67 *Confussion matrix* pengujian sembilan fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	63	0
	Negatif	4	33

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=6), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=6) menggunakan *10-fold cross validation* disajikan pada Tabel 4.68 berikut.

Tabel 4.68 *Confussion matrix* pengujian sepuluh fitur n-gram(n=6)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	54	0
	Negatif	1	45

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.69 berikut.

Tabel 4.69 Rekap hasil pengujian fitur n-gram(n=6)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	50	1	35	14	85
2	51	1	42	6	93
3	47	1	47	5	94
4	46	0	50	4	96
5	49	0	47	4	96
6	38	0	57	5	95
7	55	1	35	9	90
8	43	0	55	2	98
9	63	0	33	4	96
10	54	0	45	1	99

Rata-rata akurasi	94 %

4.3.7 Pengujian Sistem dengan Fitur n-gram(n=7)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=7. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.70 berikut.

Tabel 4.70 *Confussion matrix* pengujian satu fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	50	1
	Negatif	11	38

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.71 berikut.

Tabel 4.71 *Confussion matrix* pengujian dua fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	1
	Negatif	6	42

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.61 berikut.

Tabel 4.72 *Confussion matrix* pengujian tiga fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	47	1
	Negatif	5	47

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.73 berikut.

Tabel 4.73 *Confussion matrix* pengujian empat fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	0
	Negatif	4	50

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=7) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.74 berikut.

Tabel 4.74 *Confussion matrix* pengujian lima fitur n-gram(n=7)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	49	0
	Negatif	4	47

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.75 berikut.

Tabel 4.75 *Confussion matrix* pengujian enam fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	38	0
	Negatif	5	57

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.76 berikut.

Tabel 4.76 *Confussion matrix* pengujian tujuh fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	55	1
	Negatif	9	35

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8.

Hasil dari pengujian delapan dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.77 berikut.

Tabel 4.77 *Confussion matrix* pengujian delapan fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	2	55

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.78 berikut.

Tabel 4.78 *Confussion matrix* pengujian sembilan fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	63	0
	Negatif	3	34

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=7), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=7) menggunakan *10-fold cross validation* disajikan pada Tabel 4.79 berikut.

Tabel 4.79 *Confussion matrix* pengujian sepuluh fitur n-gram(n=7)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	54	0
	Negatif	1	45

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.80 berikut.

Tabel 4.80 Rekap hasil pengujian fitur n-gram(n=7)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	50	1	38	11	88
2	51	1	42	6	93
3	47	1	47	5	94
4	46	0	50	4	96
5	49	0	47	4	96
6	38	0	57	5	95
7	55	1	35	9	90
8	43	0	55	2	98
9	63	0	34	3	97
10	54	0	45	1	99
Rata-rata akurasi					**95 %**

4.3.8 Pengujian Sistem dengan Fitur n-gram(n=8)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=8. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.81 berikut.

Tabel 4.81 *Confussion matrix* pengujian satu fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	0
	Negatif	11	38

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.82 berikut.

Tabel 4.82 *Confussion matrix* pengujian dua fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	1
	Negatif	6	42

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.83 berikut.

Tabel 4.83 *Confussion matrix* pengujian tiga fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	0
	Negatif	5	47

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.84 berikut.

Tabel 4.84 *Confussion matrix* pengujian empat fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	0
	Negatif	4	50

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=8) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.85 berikut.

Tabel 4.85 *Confussion matrix* pengujian lima fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	0
	Negatif	4	47

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.86 berikut.

Tabel 4.86 *Confussion matrix* pengujian enam fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	38	0
	Negatif	6	56

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.87 berikut.

Tabel 4.87 *Confussion matrix* pengujian tujuh fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	55	1
	Negatif	9	35

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.88 berikut.

Tabel 4.88 *Confussion matrix* pengujian delapan fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	2	55

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.89 berikut.

Tabel 4.89 *Confussion matrix* pengujian sembilan fitur n-gram(n=8)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	63	0
	Negatif	4	33

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=8), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=8) menggunakan *10-fold cross validation* disajikan pada Tabel 4.90 berikut.

Tabel 4.90 *Confussion matrix* pengujian sepuluh fitur n-gram(n=8)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	54	0
	Negatif	1	45

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.91 berikut.

Tabel 4.91 Rekap hasil pengujian fitur n-gram(n=7)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	51	0	38	11	89
2	51	1	42	6	93
3	48	0	47	5	95
4	46	0	50	4	96
5	49	0	47	4	96
6	38	0	56	6	94
7	55	1	35	9	90
8	43	0	55	2	98
9	63	0	33	4	96
10	54	0	45	1	99
Rata-rata akurasi					**95 %**

4.3.9 Pengujian Sistem dengan Fitur n-gram(n=9)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=9. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.92 berikut.

Tabel 4.92 *Confussion matrix* pengujian satu fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	0
	Negatif	11	38

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.93 berikut.

Tabel 4.93 *Confussion matrix* pengujian dua fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	1
	Negatif	6	42

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari

pengujian tiga dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.94 berikut.

Tabel 4.94 *Confussion matrix* pengujian tiga fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	48	0
	Negatif	5	47

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.95 berikut.

Tabel 4.95 *Confussion matrix* pengujian empat fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	46	0
	Negatif	4	50

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=9) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.96 berikut.

Tabel 4.96 *Confussion matrix* pengujian lima fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	0
	Negatif	4	47

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.97 berikut.

Tabel 4.97 *Confussion matrix* pengujian enam fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	38	0
	Negatif	6	56

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7. Hasil dari pengujian tujuh dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.98 berikut.

Tabel 4.98 *Confussion matrix* pengujian tujuh fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	55	1
	Negatif	9	35

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.99 berikut.

Tabel 4.99 *Confussion matrix* pengujian delapan fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	2	55

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.100 berikut.

Tabel 4.100 *Confussion matrix* pengujian sembilan fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	63	0
	Negatif	4	33

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=9), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=9) menggunakan *10-fold cross validation* disajikan pada Tabel 4.101 berikut.

Tabel 4.101 *Confussion matrix* pengujian sepuluh fitur n-gram(n=9)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	54	0
	Negatif	1	45

Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.102 berikut.

Tabel 4.102 Rekap hasil pengujian fitur n-gram(n=9)

Pengujian	True Positive	False Positive	True Negative	False Negative	Akurasi
1	51	0	38	11	89
2	51	1	42	6	93
3	48	0	47	5	95
4	46	0	50	4	96
5	49	0	47	4	96
6	38	0	56	6	94
7	55	1	35	9	90
8	43	0	55	2	98
9	63	0	33	4	96
10	54	0	45	1	99
Rata-rata akurasi					**95 %**

4.3.10 Pengujian Sistem dengan Fitur n-gram(n=10)

Pada pengujian ini fitur yang digunakan adalah fitur n-gram dengan nilai n=9. Berikut ini ditampilkan 10 pengujian yang dilakukan berdasarkan skenario pada Tabel 4.3.

1. Pengujian Satu

Pada pengujian satu dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_1. Hasil dari pengujian satu dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.103 berikut.

Tabel 4.103 *Confussion matrix* pengujian satu fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	57	0
	Negatif	4	39

2. Pengujian Dua

Pada pengujian dua dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_2. Hasil dari pengujian dua dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.104 berikut.

Tabel 4.104 *Confussion matrix* pengujian dua fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	52	0
	Negatif	4	44

3. Pengujian Tiga

Pada pengujian tiga dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu$D_1, D_2, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_3. Hasil dari pengujian tiga dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.105 berikut.

Tabel 4.105 *Confussion matrix* pengujian tiga fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	51	0
	Negatif	4	45

4. Pengujian Empat

Pada pengujian empat dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_5, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_4. Hasil dari pengujian empat dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.106 berikut.

Tabel 4.106 *Confussion matrix* pengujian empat fitur n-gram(n=10)

	Terkategori oleh sistem

		Positif	Negatif
Kategori Sebenarnya	**Positif**	54	2
	Negatif	6	38

5. Pengujian Lima

Pada pengujian lima dengan fitur n-gram(n=10) , digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_6, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_5. Hasil dari pengujian lima dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.107 berikut.

Tabel 4.107 *Confussion matrix* pengujian lima fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	49	0
	Negatif	7	44

6. Pengujian Enam

Pada pengujian enam dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_7, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_6. Hasil dari pengujian enam dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.108 berikut.

Tabel 4.108 *Confussion matrix* pengujian enam fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	53	0
	Negatif	5	42

7. Pengujian Tujuh

Pada pengujian tujuh dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_8, D_9, D_{10}$, sedangkan data *testing* yaitu D_7.

Hasil dari pengujian tujuh dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.109 berikut.

Tabel 4.109 *Confussion matrix* pengujian tujuh fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	54	0
	Negatif	3	43

8. Pengujian Delapan

Pada pengujian delapan dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_9, D_{10}$, sedangkan data *testing* yaitu D_8. Hasil dari pengujian delapan dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.110 berikut.

Tabel 4.110 *Confussion matrix* pengujian delapan fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	47	0
	Negatif	6	47

9. Pengujian Sembilan

Pada pengujian sembilan dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_{10}$, sedangkan data *testing* yaitu D_9. Hasil dari pengujian sembilan dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.111 berikut.

Tabel 4.111 *Confussion matrix* pengujian sembilan fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	43	0
	Negatif	8	49

10. Pengujian Sepuluh

Pada pengujian sepuluh dengan fitur n-gram(n=10), digunakan 900 data sebagai data *training* dan 100 data sebagai data *testing*. Adapun data *training* yang digunakan yaitu $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$, sedangkan data *testing* yaitu D_{10}. Hasil dari pengujian sepuluh dengan fitur n-gram(n=10) menggunakan *10-fold cross validation* disajikan pada Tabel 4.112 berikut.

Tabel 4.112 *Confussion matrix* pengujian sepuluh fitur n-gram(n=10)

		Terkategori oleh sistem	
		Positif	**Negatif**
Kategori Sebenarnya	**Positif**	38	0
	Negatif	5	57

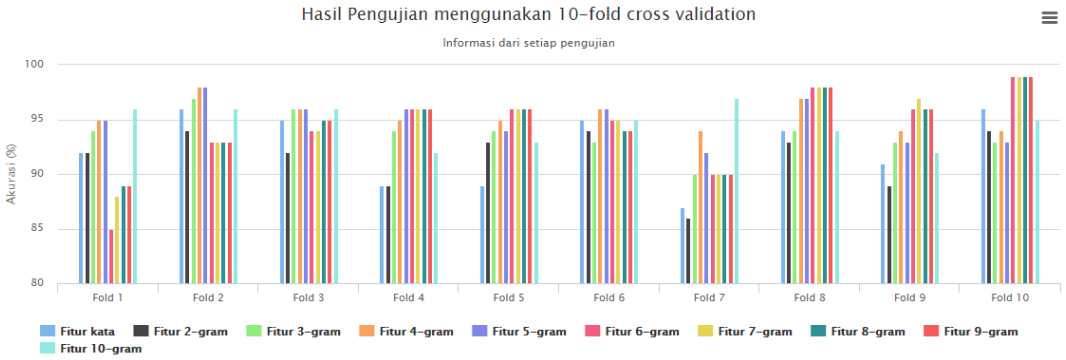
Setelah dilakukan pengujian sebanyak sepuluh kali, ditampilkan rekap hasil pengujian seperti yang ditampilkan pada Tabel 4.103 berikut.

Tabel 4.113 Rekap hasil pengujian fitur n-gram(n=9)

Pengujian	**True Positive**	**False Positive**	**True Negative**	**False Negative**	**Akurasi**
1	57	0	39	4	96
2	52	0	44	4	96
3	51	0	45	4	96
4	54	2	38	6	92
5	49	0	44	7	93
6	53	0	42	5	95
7	54	0	43	3	97
8	47	0	47	6	94
9	43	0	49	8	92
10	38	0	57	5	95
Rata-rata akurasi					**95 %**

4.4 Kesimpulan Pengujian Sistem

Semua pengujian yang telah dilakukan disajikan dalam bentuk grafik. Grafik hasil pengujian menggunakan 10-fold cross validation ditampilkan pada Gambar 4.1 dan Tabel 4.59 berikut:



Gambar 4.1 Grafik tingkat akurasi dari setiap pengujian

Dari hasil pengujian yang telah dilakukan, didapat rata-rata akurasi dari masing-masing penggunaan fitur. Informasi tersebut disajikan pada Tabel 4.57 berikut:

Tabel 4.114 Rata-rata akurasi dari masing-masing fitur

Fitur	Rata-rata akurasi
Fitur kata	92 %
Fitur n-gram (n=2)	92 %
Fitur n-gram (n=3)	94 %
Fitur n-gram (n=4)	95 %
Fitur n-gram (n=5)	95 %
Fitur n-gram (n=6)	94 %
Fitur n-gram (n=7)	95 %
Fitur n-gram (n=8)	95 %
Fitur n-gram (n=9)	95 %
Fitur n-gram (n=10)	95 %

Berdasarkan hasil pengujian yang ditampilkan pada Tabel 4.14, didapat informasi bahwa penggunaan fitur n-gram memiliki rata-rata akurasi lebih tinggi dari fitur kata. Jadi, bisa disimpulkan bahwa penggunaan fitur n-gram berpengaruh terhadap tingkat akurasi klasifikasi *tweet* dengan nilai n=4 sudah cukup mencapai rata-rata akurasi tertinggi yaitu sebesar 95%.