

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Dataset diperoleh dari *website* University California Irvine (UCI) Machine Learning Repository yaitu *dataset Cervical Cancer (Risk Factor)* yang berisikan data pasien ‘Hospital Universitario de Caracas’ yang bertempat di Caracas, Venezuela. *Dataset* terdiri dari informasi demografi, kebiasaan dan rekam medis dari 858 pasien dengan 36 atribut. *Dataset* yang digunakan memiliki label lebih dari satu atau multi-label sehingga setiap pasiennya dapat terlibat dalam beberapa label di waktu yang bersamaan. Pada *dataset* tersebut terdapat *missing value* karena pasien memutuskan tidak menjawab beberapa pertanyaan karena alasan *privacy*. Berikut ini *dataset* yang digunakan.

Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes	Smokes (years)	...	STDs: Time since first diagnosis	STDs: Time since last diagnosis	Dx:Cancer	Dx:CIN	Dx:HPV	Dx	Hinselmann	Schiller	Citology	Biopsy
18	4	15	1	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
15	1	14	1	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
52	5	16	4	1	37	...	NaN	NaN	1	0	1	0	0	0	0	0
46	3	21	4	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
42	3	23	2	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
...
34	3	18	0	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
32	2	19	1	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
25	2	17	0	0	0	...	NaN	NaN	0	0	0	0	0	0	1	0
33	2	24	2	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0
29	2	20	1	0	0	...	NaN	NaN	0	0	0	0	0	0	0	0

Gambar 4.1 Dataset

4.2 Pra-proses Data

Pada tahap ini dilakukan dengan melakukan beberapa proses, yaitu pemeriksaan *missing value*, pemeriksaan *duplicate data*, pemeriksaan tipe data dan *resample data*.

4.2.1 Pemeriksaan Missing Value

Pada *dataset* yang digunakan terdapat beberapa baris yang memiliki *missing value* lebih dari 2 atribut sehingga baris-baris tersebut dihilangkan atau dihapus menggunakan *tools Microsoft excel*. Saat dilakukan pemeriksaan kembali terhadap *dataset* yang digunakan, ternyata masih terdapat *missing value* seperti yang terlihat pada gambar 4.1, sehingga dilakukan penghilangan/penghapusan terhadap kolom yang

memiliki *missing value* lebih dari 50% dari total jumlah baris *dataset* yaitu pada kolom *STDs: Time since first diagnosis* dan *STDs: Time since last diagnosis* seperti yang terlihat pada gambar 4.2 di bawah ini.

```

Age 0
Number of sexual partners 0
First sexual intercourse 0
Num of pregnancies 0
Smokes 0
Smokes (years) 0
Smokes (packs/year) 0
Hormonal Contraceptives 0
Hormonal Contraceptives (years) 0
IUD 0
IUD (years) 0
STDs 0
STDs (number) 0
STDs:condylomatosis 0
STDs:cervical condylomatosis 0
STDs:vaginal condylomatosis 0
STDs:vulvo-perineal condylomatosis 0
STDs:syphilis 0
STDs:pelvic inflammatory disease 0
STDs:genital herpes 0
STDs:molluscum contagiosum 0
STDs:AIDS 0
STDs:HIV 0
STDs:Hepatitis B 0
STDs:HPV 0
STDs: Number of diagnosis 0
STDs: Time since first diagnosis 609
STDs: Time since last diagnosis 609
Dx:Cancer 0
Dx:CIN 0
Dx:HPV 0
Dx 0
Hinselmann 0
Schiller 0
Citology 0
Biopsy 0
dtype: int64

```

Gambar 4.2 Atribut yang Memiliki *Missing Value*

Sehingga *dataset* yang digunakan ada sebanyak 668 data pasien dengan 34 atribut yang 4 di antaranya merupakan target label dan *dataset* yang digunakan seperti yang terlihat pada gambar 4.3 di bawah ini.

Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes	Smokes (years)	Smokes (packs/year)	...	Hinselmann	Schiller	Citology	Biopsy
18.0	4.0	15.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
15.0	1.0	14.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
52.0	5.0	16.0	4.0	1.0	37.0	37.0	...	0.0	0.0	0.0	0.0
46.0	3.0	21.0	4.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
42.0	3.0	23.0	2.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
...
34.0	3.0	18.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
32.0	2.0	19.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
25.0	2.0	17.0	0.0	0.0	0.0	0.0	...	0.0	0.0	1.0	0.0
33.0	2.0	24.0	2.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
29.0	2.0	20.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0

Gambar 4.3 *Dataset* setelah Menghilangkan *Missing Value*

4.2.2 Pemeriksaan *Duplicate Data*

Setelah menghilangkan *missing value* dari *dataset*, tahap selanjutnya yaitu melakukan pemeriksaan *duplicate data* pada tiap baris *dataset*. Pemeriksaan ini dilakukan agar tidak ada baris data yang *double* saat menggunakan *dataset*. Setelah dilakukan pemeriksaan *duplicate data*, tidak ditemukan adanya *duplicate row data* pada *dataset* yang digunakan.

Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes	Smokes (years)	Smokes (packs/year)	...	Hinselmann	Schiller	Citology	Biopsy
18.0	4.0	15.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
15.0	1.0	14.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
52.0	5.0	16.0	4.0	1.0	37.0	37.0	...	0.0	0.0	0.0	0.0
46.0	3.0	21.0	4.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
42.0	3.0	23.0	2.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
...
34.0	3.0	18.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
32.0	2.0	19.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
25.0	2.0	17.0	0.0	0.0	0.0	0.0	...	0.0	0.0	1.0	0.0
33.0	2.0	24.0	2.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0
29.0	2.0	20.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0

Gambar 4.4 *Dataset* Setelah Dilakukan Pemeriksaan Data Duplikat

4.2.3 Pemeriksaan Tipe Data

Pada tahap ini dilakukan pemeriksaan tipe data pada *dataset* yang digunakan, terdapat beberapa kolom yang datanya memiliki tipe data yang tidak sesuai dengan deskripsi yang ada di *UCI Machine Learning* yang dapat juga dilihat pada lampiran 3 karena *google collaboration* kurang tepat saat membaca tipe data pada tiap kolom dari *dataset* yang di-*import* dari *google sheet*. Sehingga pada kolom-kolom tersebut diubah tipe datanya agar menyesuaikan dengan lampiran 3 tersebut. Kolom-kolom yang sudah diubah dapat dilihat pada lampiran 4. Sehingga tampilan *dataset* yang digunakan adalah seperti pada gambar 4.5.

Age	Number of sexual partners	First sexual intercourse	Num of pregnancies	Smokes	Smokes (years)	Smokes (packs/year)	...	Hinselmann	Schiller	Citology	Biopsy
18	4	15	1	0	0	0	...	0	0	0	0
15	1	14	1	0	0	0	...	0	0	0	0
52	5	16	4	1	37	37	...	0	0	0	0
46	3	21	4	0	0	0	...	0	0	0	0
42	3	23	2	0	0	0	...	0	0	0	0
...
34	3	18	0	0	0	0	...	0	0	0	0
32	2	19	1	0	0	0	...	0	0	0	0
25	2	17	0	0	0	0	...	0	0	1	0
33	2	24	2	0	0	0	...	0	0	0	0
29	2	20	1	0	0	0	...	0	0	0	0

Gambar 4.5 *Dataset* Setelah Tipe Data Tiap Kolom Disesuaikan

4.2.4 *Resample Data*

Setelah dilakukan pemeriksaan *missing value*, *duplicate data*, dan tipe data, tahap selanjutnya dilakukan *resample data*. *Resample data* dilakukan karena *dataset* yang digunakan memiliki rasio data kelas 0 dan kelas 1 pada tiap labelnya tidak seimbang. Rasio kelas 0 dan kelas 1 pada tiap label sebelum dilakukan *resample data* terlihat seperti pada tabel 4.1 di bawah ini.

Tabel 4.1 Rasio Sebelum Dilakukan *Resample*

No.	Label	Rasio
1.	<i>Hinselmann</i>	21,27 : 1
2.	<i>Schiller</i>	9,6 : 1
3.	<i>Citology</i>	16,13 : 1
4.	<i>Biopsy</i>	13,84 : 1

Banyaknya jumlah data pada masing-masing label dapat dilihat seperti pada tabel 4.2 berikut ini.

Tabel 4.2 Banyaknya Jumlah Data Pada Masing-masing Label

No.	Label	Kelas 0	Kelas 1
1.	<i>Hinselmann</i>	638	30

2.	<i>Schiller</i>	605	63
3.	<i>Citology</i>	629	39
4.	<i>Biopsy</i>	623	45

Karena memiliki data dengan jumlah kelas 0 dan 1 yang tidak seimbang, maka dilakukan percobaan *oversampling* pada ke empat kolom tersebut. Percobaan melakukan *oversampling* kelas 1 pada kolom *Hinselmann*, percobaan melakukan *oversampling* kelas 1 pada kolom *Schiller*, percobaan melakukan *oversampling* kelas 1 pada kolom *Citology*, dan percobaan melakukan *oversampling* kelas 1 pada kolom *Biopsy*. Dari percobaan-percobaan tersebut diperoleh performa terbaik yaitu saat melakukan *oversampling* kelas 1 pada kolom *Biopsy*, sehingga pada penelitian ini digunakan *dataset* yang dilakukan *oversampling* kelas 1 pada kolom *Biopsy* dengan besaran rasio antara kelas 0 dan kelas 1 dapat dilihat seperti di tabel 4.3 di bawah ini.

Tabel 4.3 Rasio Setelah Dilakukan *Oversampling* pada Kolom *Biopsy*

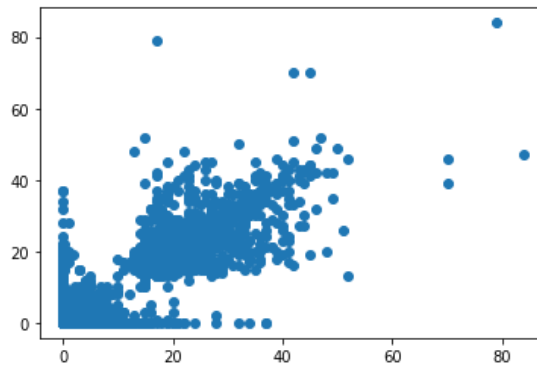
No.	Label	Rasio
1.	<i>Hinselmann</i>	3,58 : 1
2.	<i>Schiller</i>	1,18 : 1
3.	<i>Citology</i>	3,83 : 1
4.	<i>Biopsy</i>	1 : 1

Dengan banyaknya jumlah data pada masing-masing label seperti pada tabel 4.4 berikut ini.

Tabel 4.4 Banyaknya Jumlah Data Pada Masing-masing Label

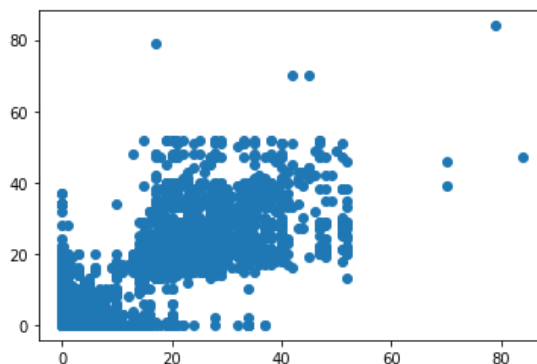
No.	Label	Kelas 0	Kelas 1
1.	<i>Hinselmann</i>	974	272
2.	<i>Schiller</i>	675	571
3.	<i>Citology</i>	988	258
4.	<i>Biopsy</i>	623	623

Oversampling pada kolom *Biopsy* pada kelas 1 yang merupakan kelas minoritas dilakukan secara *random* dengan data buatan yang terbentuk berdasarkan dari data yang sudah ada sehingga nilai-nilai dari data buatan mirip dengan nilai-nilai pada data yang sudah ada dan dapat dilihat seperti pada gambar 4.6 dan 4.7 di bawah ini.



Gambar 4.6 Sebelum Dilakukan *Oversampling*

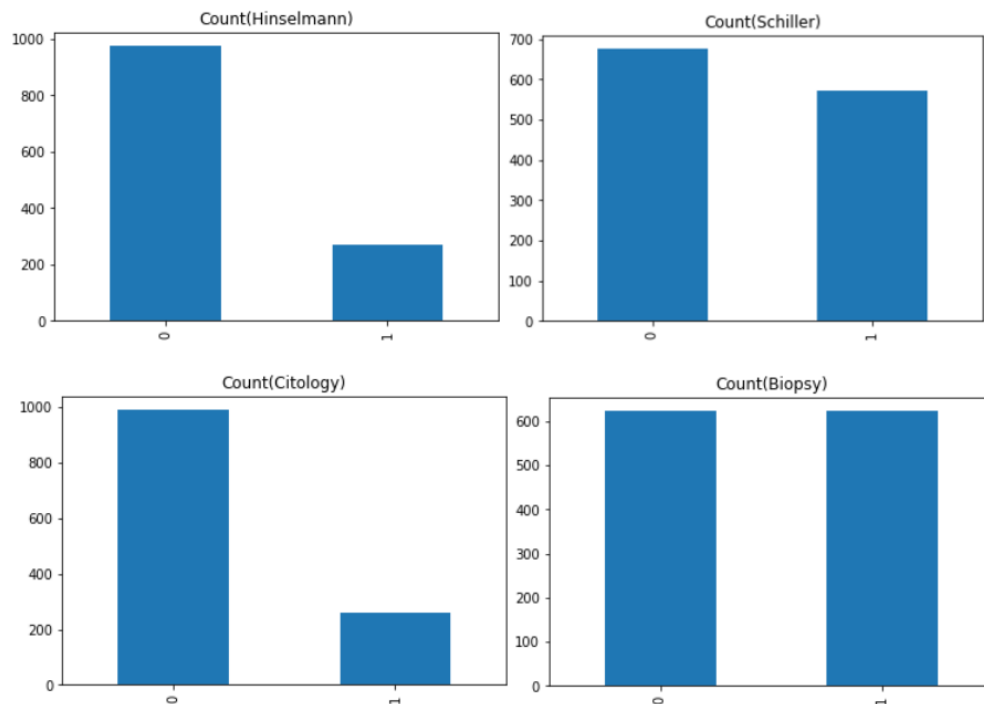
Pada gambar 4.6 terlihat persebaran data sebelum dilakukan *oversampling* pada kolom *Biopsy*. Untuk persebaran data setelah dilakukan *oversampling* dapat dilihat pada gambar 4.7 berikut ini.



Gambar 4.7 Setelah Dilakukan *Oversampling*

Pada gambar 4.7 terlihat persebaran data setelah dilakukan *oversampling* pada kolom *Biopsy* dan dapat dilihat bahwa terdapat data buatan yang muncul dan memiliki jarak yang dekat dengan data yang sudah ada. Hal ini menandakan nilai-nilai pada data buatan mirip dengan nilai-nilai pada data yang sudah ada yang merupakan kelas minoritas.

Visualisasi rasio dari kolom *Biopsy* yang sudah di-*oversampling* dapat dilihat pada gambar 4.8 berikut.



Gambar 4.8 Visualisasi Rasio Kelas 0 dan 1

Gambar 4.8 menunjukkan visualisasi rasio data kelas 0 dan 1 pada masing-masing label setelah dilakukan *oversampling* pada kolom *Biopsy*.

4.3 Pembagian Data

Setelah proses *resample data*, selanjutnya dilakukan proses pembagian data. Dilakukan percobaan dengan menggunakan beberapa rasio data latih dan data uji untuk mengetahui rasio mana yang menghasilkan performa yang lebih baik pada *dataset* yang digunakan. Beberapa rasio yang dicoba, yaitu 80:20, 75:25, 70:30, 65:35, dan 60:40. Pada data yang dilakukan *oversampling* terhadap kolom *Biopsy* diperoleh rasio data latih dan data uji dengan nilai evaluasi terbaik yaitu saat menggunakan rasio 80:20. Dari hasil percobaan tersebut maka pada penelitian ini rasio data latih dan data uji yang digunakan yaitu 80:20 sehingga diperoleh pembagian data untuk *dataset* yang

dilakukan *oversampling* pada kolom *Biopsy* yang dapat dilihat pada tabel 4.5 di bawah ini.

Tabel 4.5 Pembagian Data

Pembagian Data	Jumlah Data
Data Latih	996
Data Uji	250
Total	1.246

Dari tabel 4.5 di atas diketahui terdapat 996 baris data latih dan 250 baris data uji. Dengan jumlah kelas 0 dan kelas 1 pada masing-masing label adalah seperti pada tabel 4.6 berikut ini.

Tabel 4.6 Jumlah Kelas 0 dan 1 Pada Masing-Masing Label

Pembagian Data	Label	Kelas	Jumlah Data	Total Data
Data Latih	<i>Hinselmann</i>	Negatif (0)	780	996
		Positif (1)	216	
	<i>Schiller</i>	Negatif (0)	530	996
		Positif (1)	466	
	<i>Citology</i>	Negatif (0)	780	996
		Positif (1)	216	
	<i>Biopsy</i>	Negatif (0)	490	996
		Positif (1)	506	
Data Uji	<i>Hinselmann</i>	Negatif (0)	194	250
		Positif (1)	56	
	<i>Schiller</i>	Negatif (0)	145	250
		Positif (1)	105	
	<i>Citology</i>	Negatif (0)	208	250
		Positif (1)	42	
	<i>Biopsy</i>	Negatif (0)	133	250

		Positif (1)	117	
--	--	-------------	-----	--

Karena *dataset* yang digunakan merupakan *dataset* multi-label, sehingga ada kemungkinan terdapat 1 data atau lebih yang data tersebut memiliki lebih dari 1 label.

4.4 Klasifikasi Menggunakan ML-KNN

Selanjutnya data latih yang telah dibagi sebelumnya akan dilakukan pelatihan metode klasifikasi multi-label dengan menggunakan algoritma *Multi Label K-Nearest Neighbor* (ML-KNN). Data latih tersebut dicari kedekatan setiap data yang dihitung menggunakan rumus *euclidean distance*. Kedekatan setiap data dicari dengan nilai k yang sudah ditentukan yaitu $K=1$, $K=3$, $K=5$, $K=7$, dan $K=9$.

Agar lebih jelas dengan cara kerja dari algoritma ML-KNN yang digunakan, maka akan dibuat contoh cara kerjanya. Pada contoh ini akan digunakan sebanyak 10 data latih yang telah dibagi menggunakan *Python* dan diambil dari *dataset* yang dilakukan *oversampling* pada kolom *Hinselmann*. Selanjutnya akan dilakukan perhitungan jarak antara ke-10 data latih dengan data baru. Data latih yang digunakan seperti yang terlihat pada lampiran 5. Selanjutnya dilakukan penentuan label dari 1 contoh data uji yang belum diketahui labelnya, seperti yang terlihat pada lampiran 6.

Pada lampiran 6, belum diketahui label untuk contoh data uji tersebut, maka labelnya akan dicari menggunakan perhitungan *euclidean distance* dengan rumus (2.2) seperti yang terdapat pada bab 2. Langkah-langkah algoritma ML-KNN adalah sebagai berikut.

1. Hitung *prior* probabilitasnya menggunakan rumus (2.5) dan (2.6). Perhitungan dari rumus-rumus tersebut akan diterapkan pada kolom *Hinselmann* seperti berikut ini.

$$P(H_1^l) = \frac{s + \sum_{j=1}^n \vec{y}_{x_i}}{s \times 2 + n}$$

$$P(H_1^l) = \frac{1 + 8}{1 \times 2 + 10}$$

$$P(H_1^l) = \frac{9}{12}$$

$$P(H_1^l) = 0,75$$

$$P(H_0^l) = 1 - P(H_1^l)$$

$$P(H_0^l) = 1 - 0,75$$

$$P(H_0^l) = 0,25$$

Begitu seterusnya sampai ke-4 label sudah dihitung. Hasil perhitungan dapat dilihat pada tabel 4.7 di bawah ini.

Tabel 4.7 Hasil Perhitungan *Prior Probability*

Keterangan	<i>Hinselmann</i>	<i>Schiller</i>	<i>Citology</i>	<i>Biopsy</i>
Probabilitas	$P(H_1^l) = 0,75$	$P(H_1^l) = 0,66$	$P(H_1^l) = 0,42$	$P(H_1^l) = 0,42$
	$P(H_0^l) = 0,25$	$P(H_0^l) = 0,33$	$P(H_0^l) = 0,58$	$P(H_0^l) = 0,58$

2. Tentukan nilai K sebagai banyaknya jumlah tetangga terdekat dari data uji terhadap data latih. Pada contoh ini akan menggunakan nilai K=3. Hitung jaraknya menggunakan rumus (2.2). Perhitungan dari algoritma ML-KNN yang diterapkan pada data ke-1 dan ke-2 adalah sebagai berikut.

$$E(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$$

$$E(x, y) = \sqrt{\begin{matrix} (28 - 20)^2 + (3 - 1)^2 + (15 - 18)^2 + (3 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (1 - 1)^2 + (3 - 14)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \end{matrix}}$$

$$E(x, y) = \sqrt{\begin{matrix} 64 + 4 + 9 + 4 + 0 + 0 + 0 + 0 + 121 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \\ + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \end{matrix}}$$

$$E(x, y) = \sqrt{202}$$

$$E(x, y) = 14,2$$

$$E(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$$

$$E(x, y) = \sqrt{\begin{matrix} (28 - 19)^2 + (3 - 1)^2 + (15 - 17)^2 + (3 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (1 - 1)^2 + (3 - 2)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \end{matrix}}$$

$$E(x, y) = \sqrt{\begin{matrix} 81 + 4 + 4 + 4 + 0 + 0 + 0 + 0 + 1 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \\ + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \end{matrix}}$$

$$E(x, y) = \sqrt{94}$$

$$E(x, y) = 9,6$$

Begitu seterusnya sampai data ke 10. Setelah itu diurutkan berdasarkan jarak terkecil ke jarak terbesar. Hasil pengurutan dapat dilihat pada tabel 4.8 berikut.

Tabel 4.8 Hasil Pengurutan

Urutan Ke -	Jarak	Hinselmann	Schiller	Citology	Biopsy
1	4,7	1	1	1	1
2	6,5	1	1	0	1
3	7,6	1	1	0	1
4	8,0	1	1	1	0
5	9,6	0	0	0	0
6	9,7	0	0	0	0
7	10,5	1	0	0	0
8	11,7	1	1	1	0
9	14,2	1	1	0	1
10	24,0	1	1	1	0

Dari tabel 4.7 dapat diketahui untuk nilai K=3, penentuan tetangga terdekatnya berdasarkan 3 data yang memiliki jarak terkecil dengan data uji, seperti yang terlihat di tabel 4.9 di bawah ini.

Tabel 4.9 Penentuan Tetangga Terdekat

Urutan Ke -	Jarak	K=3	Hinselmann	Schiller	Citology	Biopsy
1	4,7	Ya	1	1	1	1
2	6,5	Ya	1	1	0	1
3	7,6	Ya	1	1	0	1
4	8,0	Tidak	1	1	1	0
5	9,6	Tidak	0	0	0	0
6	9,7	Tidak	0	0	0	0
7	10,5	Tidak	1	0	0	0
8	11,7	Tidak	1	1	1	0
9	14,2	Tidak	1	1	0	1
10	24,0	Tidak	1	1	1	0

3. Selanjutnya hitung vektor keanggotaan untuk mengetahui banyaknya data tetangga dari data uji yang berlabel 1 dengan persamaan (2.5). Perhitungan rumus akan diterapkan pada label *Hinselmann* seperti berikut ini.

$$\vec{C}_x(l) = \sum_{\alpha \in N(x)} \vec{y}_\alpha(l)$$

$$\vec{C}_x(l) = 3$$

Begitu seterusnya sampai ke-4 label sudah dihitung. Hasil perhitungan dapat dilihat pada tabel 4.10 di bawah ini.

Tabel 4.10 Hasil Perhitungan Vektor Keanggotaan

Keterangan	<i>Hinselmann</i>	<i>Schiller</i>	<i>Citology</i>	<i>Biopsy</i>
Probabilitas	$P(H_1^l) = 0,75$	$P(H_1^l) = 0,66$	$P(H_1^l) = 0,42$	$P(H_1^l) = 0,42$
	$P(H_0^l) = 0,25$	$P(H_0^l) = 0,33$	$P(H_0^l) = 0,58$	$P(H_0^l) = 0,58$
Tetangga pertama	1	1	1	1
Tetangga kedua	1	1	0	1

Tetangga ketiga	1	1	0	1
Vektor keanggotaan	$\vec{C}_x(l) = 3$	$\vec{C}_x(l) = 3$	$\vec{C}_x(l) = 1$	$\vec{C}_x(l) = 3$

4. Selanjutnya hitung *posterior probability* dari tiap label pada data tetangga yang bernilai 1 menggunakan rumus (2.6) dan 0 menggunakan rumus (2.7) seperti berikut.

$$P(E_j^l | H_1^l) = \frac{s + c[j]}{s \times (k + 1) + \sum_{j=1}^k c[j]}$$

$$P(E_j^l | H_1^l) = \frac{1 + 3}{1 \times (3 + 1) + 3}$$

$$P(E_j^l | H_1^l) = \frac{4}{7}$$

$$P(E_j^l | H_1^l) = 0,57$$

$$P(E_j^l | H_0^l) = \frac{s + c'[j]}{s \times (k + 1) + \sum_{j=1}^k c'[j]}$$

$$P(E_j^l | H_0^l) = \frac{1 + 0}{1 \times (3 + 1) + 0}$$

$$P(E_j^l | H_0^l) = \frac{1}{4}$$

$$P(E_j^l | H_0^l) = 0,25$$

Begitu seterusnya sampai ke-4 label sudah dihitung. Hasil perhitungan dapat dilihat pada tabel 4.11 di bawah ini.

Tabel 4.11 Hasil Perhitungan *Posterior Probability*

Keterangan	Label 1	Label 2	Label 3	Label 4
Prior	$P(H_1^l) = 0,75$	$P(H_1^l) = 0,66$	$P(H_1^l) = 0,42$	$P(H_1^l) = 0,42$
Probabilitas	$P(H_0^l) = 0,25$	$P(H_0^l) = 0,33$	$P(H_0^l) = 0,58$	$P(H_0^l) = 0,58$
Tetangga pertama	1	1	1	1
Tetangga kedua	1	1	0	1

Tetangga ketiga	1	1	0	1
Vektor keanggotaan	$\vec{C}_x(l) = 3$	$\vec{C}_x(l) = 3$	$\vec{C}_x(l) = 1$	$\vec{C}_x(l) = 3$
Posterior	$P(E_j^l H_1^l) = 0,57$	$P(E_j^l H_1^l) = 0,57$	$P(E_j^l H_1^l) = 0,4$	$P(E_j^l H_1^l) = 0,57$
Probabilitas	$P(E_j^l H_0^l) = 0,25$	$P(E_j^l H_0^l) = 0,25$	$P(E_j^l H_0^l) = 0,5$	$P(E_j^l H_0^l) = 0,25$

5. Selanjutnya untuk menentukan label dari data uji dengan mencari *Maximum a posteriori (MAP)* digunakan rumus (2.9) seperti berikut.

$$\vec{y}_t(\text{Hinselmann}) = \arg \max_{b \in \{1,0\}} P(H_b^l) \cdot P(E_{\vec{C}_x(l)}^l | H_b^l)$$

$$\vec{y}_t(\text{Hinselmann}) = \arg \max P(H_1^l) \cdot P(E_{\vec{C}_x(l)}^l | H_1^l)$$

$$\vec{y}_t(\text{Hinselmann}) = 0,75 \times 0,57$$

$$\vec{y}_t(\text{Hinselmann}) = 0,42$$

$$\vec{y}_t(\text{Hinselmann}) = \arg \max_{b \in \{1,0\}} P(H_b^l) \cdot P(E_{\vec{C}_x(l)}^l | H_b^l)$$

$$\vec{y}_t(\text{Hinselmann}) = \arg \max P(H_0^l) \cdot P(E_{\vec{C}_x(l)}^l | H_0^l)$$

$$\vec{y}_t(\text{Hinselmann}) = 0,25 \times 0,25$$

$$\vec{y}_t(\text{Hinselmann}) = 0,06$$

Dari perhitungan di atas diperoleh hasil *maximum a posteriori (MAP)* dari data yang bernilai 0 lebih besar dari hasil *maximum a posteriori (MAP)* dari data yang bernilai 1 dan untuk penentuan label dari data uji menggunakan nilai MAP yang paling besar, maka data uji kolom *Hinselmann* bernilai 0. Seperti itu seterusnya sampai keseluruhan label dari data uji selesai ditentukan. Hasilnya seperti yang terlihat pada tabel 4.12 berikut ini.

Tabel 4.12 Hasil Penentuan Label dari Data Uji

Keterangan	Label 1	Label 2	Label 3	Label 4
Prior	$P(H_1^l) = 0,75$	$P(H_1^l) = 0,66$	$P(H_1^l) = 0,42$	$P(H_1^l) = 0,42$
Probabilitas	$P(H_0^l) = 0,25$	$P(H_0^l) = 0,33$	$P(H_0^l) = 0,58$	$P(H_0^l) = 0,58$
Tetangga pertama	1	1	1	1
Tetangga kedua	1	1	0	1

Tetangga ketiga	1	1	0	1
Vektor keanggotaan	$\vec{C}_x(l) = 3$	$\vec{C}_x(l) = 3$	$\vec{C}_x(l) = 1$	$\vec{C}_x(l) = 3$
Posterior Probabilitas	$P(E_j^l H_1^l) = 0,57$ $P(E_j^l H_0^l) = 0,25$	$P(E_j^l H_1^l) = 0,57$ $P(E_j^l H_0^l) = 0,25$	$P(E_j^l H_1^l) = 0,4$ $P(E_j^l H_0^l) = 0,5$	$P(E_j^l H_1^l) = 0,57$ $P(E_j^l H_0^l) = 0,25$
Maximum A Posteriori (MAP) Probabilitas	$\vec{y}_t(l)1 = 0,42$ $\vec{y}_t(l)0 = 0,06$	$\vec{y}_t(l)1 = 0,38$ $\vec{y}_t(l)0 = 0,08$	$\vec{y}_t(l)1 = 0,16$ $\vec{y}_t(l)0 = 0,29$	$\vec{y}_t(l)1 = 0,23$ $\vec{y}_t(l)0 = 0,14$
Label untuk data uji	1	1	0	1

Karena waktu perhitungan yang lama, maka klasifikasi *Multi-Label K-Nearest Neighbor* diterapkan menggunakan *Python* untuk mempersingkat waktu komputasi dengan menggunakan nilai $K=1$, $K=3$, $K=5$, $K=7$, dan $K=9$.

4.5 Evaluasi Klasifikasi Menggunakan ML-KNN

Evaluasi dilakukan dengan melihat nilai *hamming loss* dan akurasi. Selain itu juga digunakan *confusion matrix* untuk dilihat nilai *precision*, *recall*, dan *f-measure* atau *f1-score*. Seperti yang sudah dijelaskan pada subbab 2.7 terdapat rata-rata *micro*, *macro*, dan *weighted* untuk perhitungan pada nilai *precision*, *recall*, dan *f-measure* atau *f1-score*. Karena *dataset* yang digunakan tidak seimbang untuk kelas 0 dan kelas 1-nya pada kolom *Hinselmann*, *Schiller*, *Citology*, dan *Biopsy* maka digunakan rata-rata *weighted* pada *precision*, *recall*, dan *f1-score* untuk evaluasinya. Evaluasi dilakukan dengan melihat perbandingan nilai K untuk mengetahui perbedaan performa dari setiap nilai K . Perbandingan nilai K bertujuan untuk mendapatkan nilai K yang optimal terhadap *dataset*.

Tabel *confusion matrix* terdiri dari nilai prediksi dan nilai sebenarnya. Bentuk matriks yang didapatkan dari penerapan klasifikasi ML-KNN pada penelitian ini berbentuk matriks 3x3. Berikut ini masing-masing matriks dari setiap nilai K yang

digunakan dengan menggunakan *dataset* yang di-*oversampling* kelas 1 pada kolom *Biopsy*.

4.5.1 Nilai K = 1

Berikut ini *confusion matrix* label *Hinselmann* untuk nilai K = 1 pada tabel 4.13.

Tabel 4.13 *Confusion Matrix* Label *Hinselmann* K = 1

	Nilai Sebenarnya	
Nilai	53 (TP)	3 (FP)
Prediksi	3 (FN)	191 (TN)

Dari tabel di atas, diketahui pada label *Hinselman* saat nilai K = 1 terdapat 53 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 191 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 3 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 3 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Schiller* untuk nilai K = 1 pada tabel 4.14.

Tabel 4.14 *Confusion Matrix* Label *Schiller* K = 1

	Nilai Sebenarnya	
Nilai	101 (TP)	8 (FP)
Prediksi	4 (FN)	137 (TN)

Dari tabel di atas, diketahui pada label *Schiller* saat nilai K = 1 terdapat 101 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 137 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 8 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 4 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Citology* untuk nilai $K = 1$ pada tabel 4.15.

Tabel 4.15 *Confusion Matrix* Label *Citology* $K = 1$

	Nilai Sebenarnya	
Nilai	36 (TP)	6 (FP)
Prediksi	6 (FN)	202 (TN)

Dari tabel di atas, diketahui pada label *Citology* saat nilai $K = 1$ terdapat 36 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 202 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 6 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 6 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Biopsy* untuk nilai $K = 1$ pada tabel 4.16.

Tabel 4.16 *Confusion Matrix* Label *Biopsy* $K = 1$

	Nilai Sebenarnya	
Nilai	117 (TP)	9 (FP)
Prediksi	0 (FN)	124 (TN)

Dari tabel di atas, diketahui pada label *Biopsy* saat nilai $K = 1$ terdapat 117 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 124 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 9 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 0 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Nilai *hamming loss* dicari menggunakan persamaan (2.10), jika $|I(y_j^{(i)} \neq \hat{y}_j^{(i)})| = 39$, yaitu jumlah data yang terprediksi tidak benar yang dihitung sesuai dengan data yang ada pada lampiran 8 dan dibandingkan dengan data yang ada pada lampiran 7, maka perhitungannya seperti berikut ini.

$$hloss(h) = \frac{1}{NL} \sum_{i=1}^N \sum_{j=1}^L |I(y_j^{(i)} \neq \hat{y}_j^{(i)})|$$

$$hloss(h) = \frac{39}{250 \cdot 4}$$

$$hloss(h) = \frac{39}{1000}$$

$$hloss(h) = 0,039$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 1$, nilai *hamming loss* yang diperoleh sebesar 0,039 atau 3,9%.

Nilai akurasi dicari menggunakan persamaan (2.11), jika $\sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right| = 230 + 0,34$ yang dihitung sesuai dengan data yang ada pada lampiran 8 dan dibandingkan dengan data yang ada pada lampiran 7 menggunakan tabel kebenaran seperti pada tabel 2.12 dan hasilnya seperti pada lampiran 9, maka perhitungannya seperti berikut ini.

$$Accuracy = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right|$$

$$Accuracy = \frac{1}{250} |230 + 0,34|$$

$$Accuracy = \frac{230,34}{250}$$

$$Accuracy = 0,92$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 1$, nilai akurasi yang diperoleh sebesar 0,92 atau 92%.

Nilai *precision*, *recall*, dan *f1-score* pada masing-masing label dicari menggunakan persamaan (2.12) untuk mencari nilai *precision*, persamaan (2.16) untuk mencari nilai *recall*, persamaan (2.20) untuk mencari nilai *f1-score*. Untuk perhitungan *precision* untuk label *Hinselmann* saat nilai $K = 1$ adalah seperti berikut ini.

$$Precision\ Hinselmann = \frac{TP}{TP+FP}$$

$$Precision\ Hinselmann = \frac{53}{53+3}$$

$$Precision\ Hinselmann = \frac{53}{56}$$

$$Precision\ Hinselmann = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Hinselmann* saat nilai $K = 1$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *recall* untuk label *Hinselmann* saat nilai $K = 1$ adalah seperti berikut ini.

$$Recall\ Hinselmann = \frac{TP}{TP+FN}$$

$$Recall\ Hinselmann = \frac{53}{53+3}$$

$$Recall\ Hinselmann = \frac{53}{56}$$

$$Recall\ Hinselmann = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Hinselmann* saat nilai $K = 1$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *f1-score* untuk label *Hinselmann* saat nilai $K = 1$ adalah seperti berikut ini.

$$F1 - Score\ Hinselmann = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score\ Hinselmann = \frac{2 \times 0,94 \times 0,94}{0,94 + 0,94}$$

$$F1 - Score\ Hinselmann = \frac{1,77}{1,88}$$

$$F1 - Score\ Hinselmann = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Hinselmann* saat nilai $K = 1$ diperoleh sebesar 0,94 atau 94%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Hinselmann* saat nilai $K = 1$ dapat dilihat pada tabel 4.17 berikut ini.

Tabel 4.17 *Precision*, *Recall*, dan *F1-Score* Label *Hinselmann* $K = 1$

	<i>Hinselmann</i>
<i>Precision</i>	0,94
<i>Recall</i>	0,94

<i>F1-Score</i>	0,94
-----------------	------

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Hinselmann* saat nilai $K = 1$ diperoleh nilai yang sama yaitu sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *precision* untuk label *Schiller* saat nilai $K = 1$ seperti berikut ini.

$$\textit{Precision Schiller} = \frac{TP}{TP+FP}$$

$$\textit{Precision Schiller} = \frac{101}{101+8}$$

$$\textit{Precision Schiller} = \frac{101}{109}$$

$$\textit{Precision Schiller} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Schiller* saat nilai $K = 1$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall* untuk label *Schiller* saat nilai $K = 1$ adalah seperti berikut ini

$$\textit{Recall Schiller} = \frac{TP}{TP+FN}$$

$$\textit{Recall Schiller} = \frac{101}{101+4}$$

$$\textit{Recall Schiller} = \frac{101}{105}$$

$$\textit{Recall Schiller} = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Schiller* saat nilai $K = 1$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk perhitungan *f1-score* untuk label *Schiller* saat nilai $K = 1$ adalah seperti berikut ini.

$$\textit{F1 - Score Schiller} = \frac{2 \times \textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

$$\textit{F1 - Score Schiller} = \frac{2 \times 0,93 \times 0,96}{0,93 + 0,96}$$

$$F1 - Score Schiller = \frac{1,78}{1,89}$$

$$F1 - Score Schiller = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Schiller* saat nilai $K = 1$ diperoleh sebesar 0,94 atau 94%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Schiller* saat nilai $K = 1$ dapat dilihat pada tabel 4.18 berikut ini.

Tabel 4.18 *Precision*, *Recall*, dan *F1-Score* Label *Schiller* $K = 1$

	<i>Schiller</i>
<i>Precision</i>	0,93
<i>Recall</i>	0,96
<i>F1-Score</i>	0,94

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Schiller* saat nilai $K = 1$ diperoleh nilai sebesar 0,93 atau 93% untuk *precision*, 0,96 atau 96% untuk *recall*, dan 0,94 atau 94% untuk *f1-score*. Selanjutnya untuk perhitungan *precision* untuk label *Citology* saat nilai $K = 1$ seperti berikut ini.

$$Precision\ Citology = \frac{TP}{TP+FP}$$

$$Precision\ Citology = \frac{36}{36+6}$$

$$Precision\ Citology = \frac{36}{42}$$

$$Precision\ Citology = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Citology* saat nilai $K = 1$ diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *recall* untuk label *Citology* saat nilai $K = 1$ adalah seperti berikut ini.

$$Recall\ Citology = \frac{TP}{TP+FN}$$

$$Recall\ Citology = \frac{36}{36+6}$$

$$\text{Recall Citology} = \frac{36}{42}$$

$$\text{Recall Citology} = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Citology* saat nilai $K = 1$ diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *f1-score* untuk label *Citology* saat nilai $K = 1$ adalah seperti berikut ini.

$$F1 - \text{Score Citology} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Citology} = \frac{2 \times 0,86 \times 0,86}{0,86 + 0,86}$$

$$F1 - \text{Score Citology} = \frac{1,48}{1,72}$$

$$F1 - \text{Score Citology} = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Citology* saat nilai $K = 1$ diperoleh sebesar 0,86 atau 86%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Citology* saat nilai $K = 1$ dapat dilihat pada tabel 4.19 berikut ini.

Tabel 4.19 *Precision, Recall, dan F1-Score Label Citology K = 1*

	<i>Citology</i>
<i>Precision</i>	0,86
<i>Recall</i>	0,86
<i>F1-Score</i>	0,86

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Citology* saat nilai $K = 1$ diperoleh nilai yang sama yaitu sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *precision* untuk label *Biopsy* saat nilai $K = 1$.

$$\text{Precision Biopsy} = \frac{TP}{TP+FP}$$

$$\text{Precision Biopsy} = \frac{117}{117+9}$$

$$\text{Precision Biopsy} = \frac{117}{126}$$

$$\text{Precision Biopsy} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 1$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall* untuk label *Biopsy* saat nilai $K = 1$ adalah seperti berikut ini.

$$\text{Recall Biopsy} = \frac{TP}{TP+FN}$$

$$\text{Recall Biopsy} = \frac{117}{117+0}$$

$$\text{Recall Biopsy} = \frac{117}{117}$$

$$\text{Recall Biopsy} = 1$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 1$ diperoleh sebesar 1 atau 100%. Selanjutnya untuk perhitungan *f1-score* untuk label *Biopsy* saat nilai $K = 1$ adalah seperti berikut ini.

$$F1 - \text{Score Biopsy} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Biopsy} = \frac{2 \times 0,93 \times 1}{0,93 + 1}$$

$$F1 - \text{Score Biopsy} = \frac{1,86}{1,93}$$

$$F1 - \text{Score Biopsy} = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Biopsy* saat nilai $K = 1$ diperoleh sebesar 0,96 atau 96%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Biopsy* saat nilai $K = 1$ dapat dilihat pada tabel 4.20 berikut ini.

Tabel 4.20 *Precision*, *Recall*, dan *F1-Score* Label *Biopsy* $K = 1$

	<i>Biopsy</i>
<i>Precision</i>	0,93
<i>Recall</i>	1
<i>F1-Score</i>	0,96

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Biopsy* saat nilai $K = 1$ diperoleh nilai sebesar 0,93 atau 93% untuk *precision*, 1 atau 100% untuk *recall*, dan 0,96 atau 96% untuk *f1-score*.

Untuk nilai *precision micro* dicari menggunakan persamaan (2.13), *precision macro* menggunakan persamaan (2.14), dan *precision weighted* menggunakan persamaan (2.15). Untuk perhitungan *precision micro* adalah seperti berikut ini.

$$Precision\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FP_1 + \dots + FP_n}$$

$$Precision\ Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FP_{Hi} + FP_{Sc} + FP_{Ci} + FP_{Bi}}$$

$$Precision\ Micro = \frac{53 + 101 + 36 + 117}{53 + 101 + 36 + 117 + 3 + 8 + 6 + 9}$$

$$Precision\ Micro = \frac{307}{333}$$

$$Precision\ Micro = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *precision micro* saat nilai $K = 1$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk perhitungan *precision macro* adalah seperti berikut ini.

$$Precision\ Macro = \frac{Precision_1 + \dots + Precision_n}{n}$$

$$Precision\ Macro = \frac{Precision_{Hinselmann} + Precision_{Schiller} + Precision_{Cytology} + Precision_{Biopsy}}{4}$$

$$Precision\ Macro = \frac{0,94 + 0,92 + 0,86 + 0,92}{4}$$

$$Precision\ Macro = \frac{3,64}{4}$$

$$Precision\ Macro = 0,91$$

Dari perhitungan di atas diketahui bahwa nilai *precision macro* saat nilai $K = 1$ diperoleh sebesar 0,91 atau 91%. Selanjutnya untuk perhitungan *precision weighted* adalah seperti berikut ini.

$$Precision\ Weighted = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FP_i}}{\sum_{i=1}^n |y_i|}$$

$$Precision\ Weighted = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FP_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FP_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FP_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FP_{Bi}}}{|y_{total}|}$$

$$Precision\ Weighted = \frac{56 \frac{53}{53+3}}{320} + \frac{105 \frac{101}{101+8}}{320} + \frac{42 \frac{36}{36+6}}{320} + \frac{117 \frac{117}{117+9}}{320}$$

$$Precision\ Weighted = \frac{56 \times 0,94}{320} + \frac{105 \times 0,93}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 0,93}{320}$$

$$Precision\ Weighted = \frac{52,64}{320} + \frac{97,65}{320} + \frac{36,12}{320} + \frac{108,81}{320}$$

$$Precision\ Weighted = \frac{295,22}{320}$$

$$Precision\ Weighted = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *precision weighted* saat nilai $K = 1$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk nilai *recall micro* dicari menggunakan persamaan (2.17), *recall macro* menggunakan persamaan (2.18), dan *recall weighted* menggunakan persamaan (2.19). Untuk perhitungan *recall micro* adalah seperti berikut ini.

$$Recall\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FN_1 + \dots + FN_n}$$

$$Recall\ Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FN_{Hi} + FN_{Sc} + FN_{Ci} + FN_{Bi}}$$

$$Recall\ Micro = \frac{53 + 101 + 36 + 117}{53 + 101 + 36 + 117 + 3 + 4 + 6 + 0}$$

$$Recall\ Micro = \frac{307}{320}$$

$$Recall\ Micro = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *recall micro* saat nilai $K = 1$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk perhitungan *recall macro* adalah seperti berikut ini.

$$\text{Recall Macro} = \frac{\text{Recall}_1 + \dots + \text{Recall}_n}{n}$$

$$\text{Recall Macro} = \frac{\text{Recall}_{Hi} + \text{Recall}_{Sc} + \text{Recall}_{Ci} + \text{Recall}_{Bi}}{4}$$

$$\text{Recall Macro} = \frac{0,94 + 0,96 + 0,86 + 1}{4}$$

$$\text{Recall Macro} = \frac{3,76}{4}$$

$$\text{Recall Macro} = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *recall macro* saat nilai $K = 1$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *recall weighted* adalah seperti berikut ini.

$$\text{Recall Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$\text{Recall Weighted} = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FN_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$\text{Recall Weighted} = \frac{56 \frac{53}{53+3}}{320} + \frac{105 \frac{101}{101+4}}{320} + \frac{42 \frac{36}{36+6}}{320} + \frac{117 \frac{117}{117+0}}{320}$$

$$\text{Recall Weighted} = \frac{56 \times 0,94}{320} + \frac{105 \times 0,96}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 1}{320}$$

$$\text{Recall Weighted} = \frac{52,64}{320} + \frac{100,8}{320} + \frac{36,12}{320} + \frac{117}{320}$$

$$\text{Recall Weighted} = \frac{306,56}{320}$$

$$\text{Recall Weighted} = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *recall weighted* saat nilai $K = 1$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk nilai *f1-score micro* dicari menggunakan persamaan (2.21), *f1-score macro* menggunakan persamaan (2.22), dan *f1-score weighted* menggunakan persamaan (2.23). Untuk perhitungan *f1-score micro* adalah seperti berikut ini.

$$F1 - Score Micro = \frac{2 \times Precision Micro \times Recall Micro}{Precision Micro + Recall Micro}$$

$$F1 - Score Micro = \frac{2 \times 0,92 \times 0,96}{0,92 + 0,96}$$

$$F1 - Score Micro = \frac{1,77}{1,88}$$

$$F1 - Score Micro = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score micro* saat nilai K = 1 diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *f1-score macro* adalah seperti berikut ini.

$$F1 - Score Macro = \frac{2 \times Precision Macro \times Recall Macro}{Precision Macro + Recall Macro}$$

$$F1 - Score Macro = \frac{2 \times 0,91 \times 0,94}{0,91 + 0,94}$$

$$F1 - Score Macro = \frac{1,72}{1,85}$$

$$F1 - Score Macro = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score macro* saat nilai K = 1 diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *f1-score weighted* adalah seperti berikut ini.

$$F1 - Score Weighted = \frac{\sum_{i=1}^n |y_i| \frac{2TP_i}{2TP_i + FP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$F1 - Score Weighted = \frac{|y_{Hi}| \frac{2TP_{Hi}}{2TP_{Hi} + FP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{2TP_{Sc}}{2TP_{Sc} + FP_{Sc} + FN_{Sc}}}{|y_{total}|} +$$

$$\frac{|y_{Ci}| \frac{2TP_{Ci}}{2TP_{Ci} + FP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{2TP_{Bi}}{2TP_{Bi} + FP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$F1 - Score Weighted = \frac{56 \frac{2 \times 53}{2 \times 53 + 3 + 3}}{320} + \frac{105 \frac{2 \times 101}{2 \times 101 + 8 + 4}}{320} + \frac{42 \frac{2 \times 36}{2 \times 36 + 6 + 6}}{320} + \frac{117 \frac{2 \times 117}{2 \times 117 + 9 + 0}}{320}$$

$$F1 - Score Weighted = \frac{56 \frac{106}{112}}{320} + \frac{105 \frac{202}{214}}{320} + \frac{42 \frac{72}{84}}{320} + \frac{117 \frac{234}{243}}{320}$$

$$F1 - Score Weighted = \frac{56 \times 0,94}{320} + \frac{105 \times 0,94}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 0,96}{320}$$

$$F1 - Score Weighted = \frac{52,64}{320} + \frac{98,7}{320} + \frac{36,12}{320} + \frac{112,32}{320}$$

$$F1 - Score Weighted = \frac{299,78}{320}$$

$$F1 - Score Weighted = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score weighted* saat nilai K = 1 diperoleh sebesar 0,94 atau 94%.

4.5.2 Nilai K = 3

Berikut ini *confusion matrix* label *Hinselmann* untuk nilai K = 3 pada tabel 4.21.

Tabel 4.21 *Confusion Matrix* Label *Hinselmann* K = 3

	Nilai Sebenarnya	
Nilai	50 (TP)	1 (FP)
Prediksi	6 (FN)	193 (TN)

Dari tabel di atas, diketahui pada label *Hinselman* saat nilai K = 3 terdapat 50 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 193 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 1 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 6 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Schiller* untuk nilai K = 3 pada tabel 4.22.

Tabel 4.22 *Confusion Matrix* Label *Schiller* K = 3

	Nilai Sebenarnya	
Nilai	98 (TP)	6 (FP)
Prediksi	7 (FN)	139 (TN)

Dari tabel di atas, diketahui pada label *Schiller* saat nilai $K = 3$ terdapat 98 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 139 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 6 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 7 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Citology* untuk nilai $K = 3$ pada tabel 4.23.

Tabel 4.23 *Confusion Matrix* Label *Citology* $K = 3$

	Nilai Sebenarnya	
Nilai	36 (TP)	5 (FP)
Prediksi	6 (FN)	203 (TN)

Dari tabel di atas, diketahui pada label *Citology* saat nilai $K = 3$ terdapat 36 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 203 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 5 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 6 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Biopsy* untuk nilai $K = 3$ pada tabel 4.24.

Tabel 4.24 *Confusion Matrix* Label *Biopsy* $K = 3$

	Nilai Sebenarnya	
Nilai	114 (TP)	8 (FP)
Prediksi	3 (FN)	125 (TN)

Dari tabel di atas, diketahui pada label *Biopsy* saat nilai $K = 3$ terdapat 114 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 125 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 8 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif

dan sebanyak 3 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Nilai *hamming loss* dicari menggunakan persamaan (2.10), jika $|I(y_j^{(i)} \neq \hat{y}_j^{(i)})| = 42$, yaitu jumlah data yang terprediksi tidak benar yang dihitung sesuai dengan data yang ada pada lampiran 10 dan dibandingkan dengan data yang ada pada lampiran 7, maka perhitungannya seperti berikut ini.

$$hloss(h) = \frac{1}{NL} \sum_{i=1}^N \sum_{l=1}^L |I(y_j^{(i)} \neq \hat{y}_j^{(i)})|$$

$$hloss(h) = \frac{42}{250 \cdot 4}$$

$$hloss(h) = \frac{42}{1000}$$

$$hloss(h) = 0,042$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 3$, nilai *hamming loss* yang diperoleh sebesar 0,042 atau 4,2%.

Nilai akurasi dicari menggunakan persamaan (2.11), jika $\sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right| = 230 + 0,34$ yang dihitung sesuai dengan data yang ada pada lampiran 10 dan dibandingkan dengan data yang ada pada lampiran 7 menggunakan tabel kebenaran seperti pada tabel 2.12 dan hasilnya seperti pada lampiran 11, maka perhitungannya seperti berikut ini.

$$Accuracy = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right|$$

$$Accuracy = \frac{1}{250} |230 + 0,34|$$

$$Accuracy = \frac{230,34}{250}$$

$$Accuracy = 0,92$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 3$, nilai akurasi yang diperoleh sebesar 0,92 atau 92%.

Nilai *precision*, *recall*, dan *f1-score* pada masing-masing label dicari menggunakan persamaan (2.12) untuk mencari nilai *precision*, persamaan (2.16) untuk mencari nilai *recall*, persamaan (2.20) untuk mencari nilai *f1-score*. Untuk perhitungan *precision* untuk label *Hinselmann* saat nilai $K = 3$ adalah seperti berikut ini.

$$\textit{Precision Hinselmann} = \frac{TP}{TP+FP}$$

$$\textit{Precision Hinselmann} = \frac{50}{50+1}$$

$$\textit{Precision Hinselmann} = \frac{50}{51}$$

$$\textit{Precision Hinselmann} = 0,98$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Hinselmann* saat nilai $K = 3$ diperoleh sebesar 0,98 atau 98%. Selanjutnya untuk perhitungan *recall* untuk label *Hinselmann* saat nilai $K = 3$ adalah seperti berikut ini.

$$\textit{Recall Hinselmann} = \frac{TP}{TP+FN}$$

$$\textit{Recall Hinselmann} = \frac{50}{50+6}$$

$$\textit{Recall Hinselmann} = \frac{50}{56}$$

$$\textit{Recall Hinselmann} = 0,89$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Hinselmann* saat nilai $K = 3$ diperoleh sebesar 0,89 atau 89%. Selanjutnya untuk perhitungan *f1-score* untuk label *Hinselmann* saat nilai $K = 3$ adalah seperti berikut ini.

$$\textit{F1 - Score Hinselmann} = \frac{2 \times \textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

$$\textit{F1 - Score Hinselmann} = \frac{2 \times 0,98 \times 0,89}{0,98 + 0,89}$$

$$\textit{F1 - Score Hinselmann} = \frac{1,74}{1,87}$$

$$\textit{F1 - Score Hinselmann} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Hinselmann* saat nilai $K = 3$ diperoleh sebesar 0,93 atau 93%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Hinselmann* saat nilai $K = 3$ dapat dilihat pada tabel 4.25 berikut ini.

Tabel 4.25 *Precision, Recall, dan F1-Score* Label *Hinselmann* $K = 3$

	<i>Hinselmann</i>
<i>Precision</i>	0,98
<i>Recall</i>	0,89
<i>F1-Score</i>	0,93

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Hinselmann* saat nilai $K = 3$ diperoleh nilai *precision* sebesar 0,94 atau 94%, *recall* sebesar 0,89 atau 89%, dan *f1-score* sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *precision* untuk label *Schiller* saat nilai $K = 3$ seperti berikut ini.

$$\textit{Precision Schiller} = \frac{TP}{TP+FP}$$

$$\textit{Precision Schiller} = \frac{98}{98+6}$$

$$\textit{Precision Schiller} = \frac{98}{104}$$

$$\textit{Precision Schiller} = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Schiller* saat nilai $K = 3$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *recall* untuk label *Schiller* saat nilai $K = 3$ adalah seperti berikut ini

$$\textit{Recall Schiller} = \frac{TP}{TP+FN}$$

$$\textit{Recall Schiller} = \frac{98}{98+7}$$

$$\textit{Recall Schiller} = \frac{98}{105}$$

$$\textit{Recall Schiller} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Schiller* saat nilai $K = 3$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *f1-score* untuk label *Schiller* saat nilai $K = 3$ adalah seperti berikut ini.

$$F1 - Score Schiller = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score Schiller = \frac{2 \times 0,94 \times 0,93}{0,94 + 0,93}$$

$$F1 - Score Schiller = \frac{1,75}{1,87}$$

$$F1 - Score Schiller = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Schiller* saat nilai $K = 3$ diperoleh sebesar 0,92 atau 92%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Schiller* saat nilai $K = 3$ dapat dilihat pada tabel 4.26 berikut ini.

Tabel 4.26 *Precision, Recall, dan F1-Score Label Schiller K = 3*

	<i>Schiller</i>
<i>Precision</i>	0,94
<i>Recall</i>	0,93
<i>F1-Score</i>	0,92

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Schiller* saat nilai $K = 3$ diperoleh nilai sebesar 0,94 atau 94% untuk *precision*, 0,93 atau 93% untuk *recall*, dan 0,92 atau 92% untuk *f1-score*. Selanjutnya untuk perhitungan *precision* untuk label *Citology* saat nilai $K = 3$ seperti berikut ini.

$$Precision Citology = \frac{TP}{TP + FP}$$

$$Precision Citology = \frac{36}{36 + 5}$$

$$Precision Citology = \frac{36}{41}$$

$$Precision Citology = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Citology* saat nilai $K = 3$ diperoleh sebesar 0,88 atau 88%. Selanjutnya untuk perhitungan *recall* untuk label *Citology* saat nilai $K = 3$ adalah seperti berikut ini.

$$\text{Recall Citology} = \frac{TP}{TP+FN}$$

$$\text{Recall Citology} = \frac{36}{36+6}$$

$$\text{Recall Citology} = \frac{36}{42}$$

$$\text{Recall Citology} = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Citology* saat nilai $K = 3$ diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *f1-score* untuk label *Citology* saat nilai $K = 3$ adalah seperti berikut ini.

$$F1 - \text{Score Citology} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Citology} = \frac{2 \times 0,88 \times 0,86}{0,88 + 0,86}$$

$$F1 - \text{Score Citology} = \frac{1,51}{1,74}$$

$$F1 - \text{Score Citology} = 0,87$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Citology* saat nilai $K = 3$ diperoleh sebesar 0,87 atau 87%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Citology* saat nilai $K = 3$ dapat dilihat pada tabel 4.27 berikut ini.

Tabel 4.27 *Precision*, *Recall*, dan *F1-Score* Label *Citology* $K = 3$

	<i>Citology</i>
<i>Precision</i>	0,86
<i>Recall</i>	0,86
<i>F1-Score</i>	0,86

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Citology* saat nilai $K = 3$ diperoleh nilai yang sama yaitu sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *precision* untuk label *Biopsy* saat nilai $K = 3$.

$$Precision\ Biopsy = \frac{TP}{TP+FP}$$

$$Precision\ Biopsy = \frac{114}{114+8}$$

$$Precision\ Biopsy = \frac{114}{122}$$

$$Precision\ Biopsy = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 3$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall* untuk label *Biopsy* saat nilai $K = 3$ adalah seperti berikut ini.

$$Recall\ Biopsy = \frac{TP}{TP+FN}$$

$$Recall\ Biopsy = \frac{114}{114+3}$$

$$Recall\ Biopsy = \frac{114}{117}$$

$$Recall\ Biopsy = 0,97$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 3$ diperoleh sebesar 0,97 atau 97%. Selanjutnya untuk perhitungan *f1-score* untuk label *Biopsy* saat nilai $K = 3$ adalah seperti berikut ini.

$$F1 - Score\ Biopsy = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score\ Biopsy = \frac{2 \times 0,93 \times 0,97}{0,93 + 0,97}$$

$$F1 - Score\ Biopsy = \frac{1,80}{1,90}$$

$$F1 - Score\ Biopsy = 0,95$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Biopsy* saat nilai $K = 3$ diperoleh sebesar 0,95 atau 95%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Biopsy* saat nilai $K = 3$ dapat dilihat pada tabel 4.28 berikut ini.

Tabel 4.28 *Precision, Recall, dan F1-Score Label Biopsy K = 3*

	<i>Biopsy</i>
<i>Precision</i>	0,93
<i>Recall</i>	0,97
<i>F1-Score</i>	0,95

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Biopsy* saat nilai $K = 3$ diperoleh nilai sebesar 0,93 atau 93% untuk *precision*, 0,97 atau 97% untuk *recall*, dan 0,95 atau 95% untuk *f1-score*.

Untuk nilai *precision micro* dicari menggunakan persamaan (2.13), *precision macro* menggunakan persamaan (2.14), dan *precision weighted* menggunakan persamaan (2.15). Untuk perhitungan *precision micro* adalah seperti berikut ini.

$$Precision\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FP_1 + \dots + FP_n}$$

$$Precision\ Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FP_{Hi} + FP_{Sc} + FP_{Ci} + FP_{Bi}}$$

$$Precision\ Micro = \frac{50 + 98 + 36 + 114}{50 + 98 + 36 + 114 + 1 + 6 + 5 + 8}$$

$$Precision\ Micro = \frac{298}{318}$$

$$Precision\ Micro = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *precision micro* saat nilai $K = 3$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *precision macro* adalah seperti berikut ini.

$$Precision\ Macro = \frac{Precision_1 + \dots + Precision_n}{n}$$

$$Precision\ Macro = \frac{Precision_{Hinselmann} + Precision_{Schiller} + Precision_{Citology} + Precision_{Biopsy}}{4}$$

$$Precision\ Macro = \frac{0,98 + 0,94 + 0,86 + 0,93}{4}$$

$$Precision\ Macro = \frac{3,71}{4}$$

$$Precision\ Macro = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision macro* saat nilai $K = 3$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *precision weighted* adalah seperti berikut ini.

$$Precision\ Weighted = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FP_i}}{\sum_{i=1}^n |y_i|}$$

$$Precision\ Weighted = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FP_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FP_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FP_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FP_{Bi}}}{|y_{total}|}$$

$$Precision\ Weighted = \frac{56 \frac{50}{50+1}}{320} + \frac{105 \frac{98}{98+6}}{320} + \frac{42 \frac{36}{36+5}}{320} + \frac{117 \frac{114}{114+8}}{320}$$

$$Precision\ Weighted = \frac{56 \times 0,98}{320} + \frac{105 \times 0,94}{320} + \frac{42 \times 0,88}{320} + \frac{117 \times 0,94}{320}$$

$$Precision\ Weighted = \frac{54,88}{320} + \frac{98,7}{320} + \frac{36,96}{320} + \frac{109,98}{320}$$

$$Precision\ Weighted = \frac{300,52}{320}$$

$$Precision\ Weighted = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *precision weighted* saat nilai $K = 3$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk nilai *recall micro* dicari menggunakan persamaan (2.17), *recall macro* menggunakan persamaan (2.18), dan *recall weighted* menggunakan persamaan (2.19). Untuk perhitungan *recall micro* adalah seperti berikut ini.

$$Recall\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FN_1 + \dots + FN_n}$$

$$\text{Recall Micro} = \frac{TP_{Hi}+TP_{Sc}+TP_{Ci}+TP_{Bi}}{TP_{Hi}+TP_{Sc}+TP_{Ci}+TP_{Bi}+FN_{Hi}+FN_{Sc}+FN_{Ci}+FN_{Bi}}$$

$$\text{Recall Micro} = \frac{50+98+36+114}{50+98+36+114+6+7+6+3}$$

$$\text{Recall Micro} = \frac{298}{320}$$

$$\text{Recall Micro} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *recall micro* saat nilai K = 3 diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall macro* adalah seperti berikut ini.

$$\text{Recall Macro} = \frac{\text{Recall}_1+\dots+\text{Recall}_n}{n}$$

$$\text{Recall Macro} = \frac{\text{Recall}_{Hi}+\text{Recall}_{Sc}+\text{Recall}_{Ci}+\text{Recall}_{Bi}}{4}$$

$$\text{Recall Macro} = \frac{0,89+0,93+0,86+0,97}{4}$$

$$\text{Recall Macro} = \frac{3,65}{4}$$

$$\text{Recall Macro} = 0,91$$

Dari perhitungan di atas diketahui bahwa nilai *recall macro* saat nilai K = 3 diperoleh sebesar 0,91 atau 91%. Selanjutnya untuk perhitungan *recall weighted* adalah seperti berikut ini.

$$\text{Recall Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i+FN_i}}{\sum_{i=1}^n |y_i|}$$

$$\text{Recall Weighted} = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi}+FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc}+FN_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci}+FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi}+FN_{Bi}}}{|y_{total}|}$$

$$\text{Recall Weighted} = \frac{56 \frac{50}{50+6}}{320} + \frac{105 \frac{98}{98+7}}{320} + \frac{42 \frac{36}{36+6}}{320} + \frac{117 \frac{114}{114+3}}{320}$$

$$\text{Recall Weighted} = \frac{56 \times 0,89}{320} + \frac{105 \times 0,93}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 0,97}{320}$$

$$Recall\ Weighted = \frac{49,84}{320} + \frac{97,65}{320} + \frac{36,12}{320} + \frac{113,49}{320}$$

$$Recall\ Weighted = \frac{294,1}{320}$$

$$Recall\ Weighted = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *recall weighted* saat nilai $K = 3$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk nilai *f1-score micro* dicari menggunakan persamaan (2.21), *f1-score macro* menggunakan persamaan (2.22), dan *f1-score weighted* menggunakan persamaan (2.23). Untuk perhitungan *f1-score micro* adalah seperti berikut ini.

$$F1 - Score\ Micro = \frac{2 \times Precision\ Micro \times Recall\ Micro}{Precision\ Micro + Recall\ Micro}$$

$$F1 - Score\ Micro = \frac{2 \times 0,94 \times 0,93}{0,94 + 0,93}$$

$$F1 - Score\ Micro = \frac{1,75}{1,87}$$

$$F1 - Score\ Micro = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score micro* saat nilai $K = 3$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *f1-score macro* adalah seperti berikut ini.

$$F1 - Score\ Macro = \frac{2 \times Precision\ Macro \times Recall\ Macro}{Precision\ Macro + Recall\ Macro}$$

$$F1 - Score\ Macro = \frac{2 \times 0,93 \times 0,91}{0,93 + 0,91}$$

$$F1 - Score\ Macro = \frac{1,69}{1,84}$$

$$F1 - Score\ Macro = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score macro* saat nilai $K = 3$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk perhitungan *f1-score weighted* adalah seperti berikut ini.

$$F1 - \text{Score Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{2TP_i}{2TP_i + FP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$F1 - \text{Score Weighted} = \frac{|y_{Hi}| \frac{2TP_{Hi}}{2TP_{Hi} + FP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{2TP_{Sc}}{2TP_{Sc} + FP_{Sc} + FN_{Sc}}}{|y_{total}|} +$$

$$\frac{|y_{Ci}| \frac{2TP_{Ci}}{2TP_{Ci} + FP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{2TP_{Bi}}{2TP_{Bi} + FP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$F1 - \text{Score Weighted} = \frac{56 \frac{2 \times 50}{2 \times 50 + 1 + 6}}{320} + \frac{105 \frac{2 \times 98}{2 \times 98 + 6 + 7}}{320} + \frac{42 \frac{2 \times 36}{2 \times 36 + 5 + 6}}{320} + \frac{117 \frac{2 \times 114}{2 \times 114 + 8 + 3}}{320}$$

$$F1 - \text{Score Weighted} = \frac{56 \frac{100}{107}}{320} + \frac{105 \frac{196}{209}}{320} + \frac{42 \frac{72}{83}}{320} + \frac{117 \frac{228}{239}}{320}$$

$$F1 - \text{Score Weighted} = \frac{56 \times 0,93}{320} + \frac{105 \times 0,94}{320} + \frac{42 \times 0,87}{320} + \frac{117 \times 0,95}{320}$$

$$F1 - \text{Score Weighted} = \frac{52,08}{320} + \frac{98,7}{320} + \frac{36,54}{320} + \frac{111,15}{320}$$

$$F1 - \text{Score Weighted} = \frac{298,45}{320}$$

$$F1 - \text{Score Weighted} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score weighted* saat nilai K = 3 diperoleh sebesar 0,93 atau 93%.

4.5.3 Nilai K = 5

Berikut ini *confusion matrix* label *Hinselmann* untuk nilai K = 5 pada tabel 4.29.

Tabel 4.29 *Confusion Matrix* Label *Hinselmann* K = 5

	Nilai Sebenarnya	
Nilai	53 (TP)	2 (FP)
Prediksi	3 (FN)	192 (TN)

Dari tabel di atas, diketahui pada label *Hinselman* saat nilai K = 5 terdapat 53 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 192 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif.

Sebanyak 2 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 3 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Schiller* untuk nilai $K = 5$ pada tabel 4.30.

Tabel 4.30 *Confusion Matrix* Label *Schiller* $K = 5$

	Nilai Sebenarnya	
Nilai	101 (TP)	7 (FP)
Prediksi	4 (FN)	138 (TN)

Dari tabel di atas, diketahui pada label *Schiller* saat nilai $K = 5$ terdapat 101 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 138 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 7 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 4 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Citology* untuk nilai $K = 5$ pada tabel 4.31.

Tabel 4.31 *Confusion Matrix* Label *Citology* $K = 5$

	Nilai Sebenarnya	
Nilai	36 (TP)	5 (FP)
Prediksi	6 (FN)	203 (TN)

Dari tabel di atas, diketahui pada label *Citology* saat nilai $K = 5$ terdapat 36 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 203 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 5 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 6 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Biopsy* untuk nilai $K = 5$ pada tabel 4.32.

Tabel 4.32 *Confusion Matrix* Label *Biopsy* K = 5

	Nilai Sebenarnya	
Nilai	117 (TP)	9 (FP)
Prediksi	0 (FN)	124 (TN)

Dari tabel di atas, diketahui pada label *Biopsy* saat nilai K = 5 terdapat 117 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 124 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 9 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 0 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Nilai *hamming loss* dicari menggunakan persamaan (2.10), jika $|I(y_j^{(i)} \neq \hat{y}_j^{(i)})| = 36$, yaitu jumlah data yang terprediksi tidak benar yang dihitung sesuai dengan data yang ada pada lampiran 12 dan dibandingkan dengan data yang ada pada lampiran 7, maka perhitungannya seperti berikut ini.

$$hloss(h) = \frac{1}{NL} \sum_{i=1}^N \sum_{j=1}^L |I(y_j^{(i)} \neq \hat{y}_j^{(i)})|$$

$$hloss(h) = \frac{36}{250 \cdot 4}$$

$$hloss(h) = \frac{36}{1000}$$

$$hloss(h) = 0,036$$

Dari perhitungan di atas, diketahui bahwa saat nilai K = 5, nilai *hamming loss* yang diperoleh sebesar 0,036 atau 3,6%.

Nilai akurasi dicari menggunakan persamaan (2.11), jika $\sum_{i=1}^n \frac{|Y_i \cap \hat{Y}_i|}{|Y_i \cup \hat{Y}_i|} = 232 + 0,34$ yang dihitung sesuai dengan data yang ada pada lampiran 12 dan dibandingkan dengan data yang ada pada lampiran 7 menggunakan tabel kebenaran seperti pada tabel 2.12 dan hasilnya seperti pada lampiran 13, maka perhitungannya seperti berikut ini.

$$Accuracy = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right|$$

$$Accuracy = \frac{1}{250} |232 + 0,34|$$

$$Accuracy = \frac{232,34}{250}$$

$$Accuracy = 0,93$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 5$, nilai akurasi yang diperoleh sebesar 0,93 atau 93%.

Nilai *precision*, *recall*, dan *f1-score* pada masing-masing label dicari menggunakan persamaan (2.12) untuk mencari nilai *precision*, persamaan (2.16) untuk mencari nilai *recall*, persamaan (2.20) untuk mencari nilai *f1-score*. Untuk perhitungan *precision* untuk label *Hinselmann* saat nilai $K = 5$ adalah seperti berikut ini.

$$Precision\ Hinselmann = \frac{TP}{TP+FP}$$

$$Precision\ Hinselmann = \frac{53}{53+2}$$

$$Precision\ Hinselmann = \frac{53}{55}$$

$$Precision\ Hinselmann = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Hinselmann* saat nilai $K = 5$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk perhitungan *recall* untuk label *Hinselmann* saat nilai $K = 5$ adalah seperti berikut ini.

$$Recall\ Hinselmann = \frac{TP}{TP+FN}$$

$$Recall\ Hinselmann = \frac{53}{53+3}$$

$$Recall\ Hinselmann = \frac{53}{56}$$

$$Recall\ Hinselmann = 0,95$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Hinselmann* saat nilai $K = 5$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *f1-score* untuk label *Hinselmann* saat nilai $K = 1$ adalah seperti berikut ini.

$$F1 - Score\ Hinselmann = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score\ Hinselmann = \frac{2 \times 0,96 \times 0,94}{0,96 + 0,94}$$

$$F1 - Score\ Hinselmann = \frac{1,80}{1,90}$$

$$F1 - Score\ Hinselmann = 0,95$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Hinselmann* saat nilai $K = 5$ diperoleh sebesar 0,95 atau 95%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Hinselmann* saat nilai $K = 5$ dapat dilihat pada tabel 4.33 berikut ini.

Tabel 4.33 *Precision, Recall, dan F1-Score Label Hinselmann K = 5*

	<i>Hinselmann</i>
<i>Precision</i>	0,96
<i>Recall</i>	0,95
<i>F1-Score</i>	0,95

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Hinselmann* saat nilai $K = 5$ diperoleh nilai *precision* sebesar 0,96 atau 96%, nilai *recall* sebesar 0,94 atau 94%, dan nilai *f1-score* sebesar 0,95 atau 95%. Selanjutnya untuk perhitungan *precision* untuk label *Schiller* saat nilai $K = 5$ seperti berikut ini.

$$Precision\ Schiller = \frac{TP}{TP + FP}$$

$$Precision\ Schiller = \frac{101}{101 + 7}$$

$$Precision\ Schiller = \frac{101}{108}$$

$$Precision\ Schiller = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Schiller* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall* untuk label *Schiller* saat nilai $K = 5$ adalah seperti berikut ini

$$\text{Recall Schiller} = \frac{TP}{TP+FN}$$

$$\text{Recall Schiller} = \frac{101}{101+7}$$

$$\text{Recall Schiller} = \frac{101}{108}$$

$$\text{Recall Schiller} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Schiller* saat nilai $K = 1$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk perhitungan *f1-score* untuk label *Schiller* saat nilai $K = 1$ adalah seperti berikut ini.

$$F1 - \text{Score Schiller} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Schiller} = \frac{2 \times 0,93 \times 0,93}{0,93 + 0,93}$$

$$F1 - \text{Score Schiller} = \frac{1,73}{1,86}$$

$$F1 - \text{Score Schiller} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Schiller* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Schiller* saat nilai $K = 5$ dapat dilihat pada tabel 4.34 berikut ini.

Tabel 4.34 *Precision*, *Recall*, dan *F1-Score* Label *Schiller* $K = 5$

	<i>Schiller</i>
<i>Precision</i>	0,93
<i>Recall</i>	0,93
<i>F1-Score</i>	0,93

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Schiller* saat nilai $K = 5$ diperoleh nilai yang sama yaitu sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *precision* untuk label *Citology* saat nilai $K = 5$ seperti berikut ini.

$$Precision\ Citology = \frac{TP}{TP+FP}$$

$$Precision\ Citology = \frac{36}{36+5}$$

$$Precision\ Citology = \frac{36}{41}$$

$$Precision\ Citology = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Citology* saat nilai $K = 5$ diperoleh sebesar 0,88 atau 88%. Selanjutnya untuk perhitungan *recall* untuk label *Citology* saat nilai $K = 5$ adalah seperti berikut ini.

$$Recall\ Citology = \frac{TP}{TP+FN}$$

$$Recall\ Citology = \frac{36}{36+6}$$

$$Recall\ Citology = \frac{36}{42}$$

$$Recall\ Citology = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Citology* saat nilai $K = 5$ diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *f1-score* untuk label *Citology* saat nilai $K = 5$ adalah seperti berikut ini.

$$F1 - Score\ Citology = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score\ Citology = \frac{2 \times 0,88 \times 0,86}{0,88 + 0,86}$$

$$F1 - Score\ Citology = \frac{1,51}{1,74}$$

$$F1 - Score\ Citology = 0,87$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Citology* saat nilai $K = 5$ diperoleh sebesar 0,87 atau 87%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Citology* saat nilai $K = 5$ dapat dilihat pada tabel 4.35 berikut ini.

Tabel 4.35 *Precision, Recall, dan F1-Score Label Citology K = 5*

	<i>Citology</i>
<i>Precision</i>	0,88
<i>Recall</i>	0,86
<i>F1-Score</i>	0,87

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Citology* saat nilai $K = 5$ diperoleh nilai *precision* sebesar 0,88 atau 88%, nilai *recall* sebesar 0,86 atau 86%, dan nilai *f1-score* sebesar 0,87 atau 87%. Selanjutnya untuk perhitungan *precision* untuk label *Biopsy* saat nilai $K = 5$.

$$Precision\ Biopsy = \frac{TP}{TP+FP}$$

$$Precision\ Biopsy = \frac{117}{117+9}$$

$$Precision\ Biopsy = \frac{117}{126}$$

$$Precision\ Biopsy = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall* untuk label *Biopsy* saat nilai $K = 5$ adalah seperti berikut ini.

$$Recall\ Biopsy = \frac{TP}{TP+FN}$$

$$Recall\ Biopsy = \frac{117}{117+0}$$

$$Recall\ Biopsy = \frac{117}{117}$$

$$Recall\ Biopsy = 1$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 5$ diperoleh sebesar 1 atau 100%. Selanjutnya untuk perhitungan *f1-score* untuk label *Biopsy* saat nilai $K = 5$ adalah seperti berikut ini.

$$F1 - Score Biopsy = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score Biopsy = \frac{2 \times 0,93 \times 1}{0,93 + 1}$$

$$F1 - Score Biopsy = \frac{1,86}{1,93}$$

$$F1 - Score Biopsy = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Biopsy* saat nilai $K = 5$ diperoleh sebesar 0,96 atau 96%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Biopsy* saat nilai $K = 5$ dapat dilihat pada tabel 4.36 berikut ini.

Tabel 4.36 *Precision*, *Recall*, dan *F1-Score* Label *Biopsy* $K = 5$

	<i>Biopsy</i>
<i>Precision</i>	0,93
<i>Recall</i>	1
<i>F1-Score</i>	0,96

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Biopsy* saat nilai $K = 5$ diperoleh nilai sebesar 0,93 atau 93% untuk *precision*, 1 atau 100% untuk *recall*, dan 0,96 atau 96% untuk *f1-score*.

Untuk nilai *precision micro* dicari menggunakan persamaan (2.13), *precision macro* menggunakan persamaan (2.14), dan *precision weighted* menggunakan persamaan (2.15). Untuk perhitungan *precision micro* adalah seperti berikut ini.

$$Precision Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FP_1 + \dots + FP_n}$$

$$Precision Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FP_{Hi} + FP_{Sc} + FP_{Ci} + FP_{Bi}}$$

$$Precision\ Micro = \frac{53+101+36+117}{53+101+36+117+2+7+5+9}$$

$$Precision\ Micro = \frac{307}{330}$$

$$Precision\ Micro = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision micro* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *precision macro* adalah seperti berikut ini.

$$Precision\ Macro = \frac{Precision_1 + \dots + Precision_n}{n}$$

$$Precision\ Macro = \frac{Precision_{Hinselmann} + Precision_{Schiller} + Precision_{Cytology} + Precision_{Biopsy}}{4}$$

$$Precision\ Macro = \frac{0,96+0,93+0,88+0,93}{4}$$

$$Precision\ Macro = \frac{3,7}{4}$$

$$Precision\ Macro = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision macro* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *precision weighted* adalah seperti berikut ini.

$$Precision\ Weighted = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FP_i}}{\sum_{i=1}^n |y_i|}$$

$$Precision\ Weighted = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FP_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FP_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FP_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FP_{Bi}}}{|y_{total}|}$$

$$Precision\ Weighted = \frac{56 \frac{53}{53+2}}{320} + \frac{105 \frac{101}{101+7}}{320} + \frac{42 \frac{36}{36+5}}{320} + \frac{117 \frac{117}{117+9}}{320}$$

$$Precision\ Weighted = \frac{56 \times 0,96}{320} + \frac{105 \times 0,93}{320} + \frac{42 \times 0,88}{320} + \frac{117 \times 0,93}{320}$$

$$Precision\ Weighted = \frac{53,76}{320} + \frac{97,65}{320} + \frac{36,96}{320} + \frac{108,81}{320}$$

$$Precision\ Weighted = \frac{297,18}{320}$$

Precision Weighted = 0,93

Dari perhitungan di atas diketahui bahwa nilai *precision weighted* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk nilai *recall micro* dicari menggunakan persamaan (2.17), *recall macro* menggunakan persamaan (2.18), dan *recall weighted* menggunakan persamaan (2.19). Untuk perhitungan *recall micro* adalah seperti berikut ini.

$$\text{Recall Micro} = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FN_1 + \dots + FN_n}$$

$$\text{Recall Micro} = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FN_{Hi} + FN_{Sc} + FN_{Ci} + FN_{Bi}}$$

$$\text{Recall Micro} = \frac{53 + 101 + 36 + 117}{53 + 101 + 36 + 117 + 3 + 4 + 6 + 0}$$

$$\text{Recall Micro} = \frac{307}{320}$$

Recall Micro = 0,96

Dari perhitungan di atas diketahui bahwa nilai *recall micro* saat nilai $K = 5$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk perhitungan *recall macro* adalah seperti berikut ini.

$$\text{Recall Macro} = \frac{\text{Recall}_1 + \dots + \text{Recall}_n}{n}$$

$$\text{Recall Macro} = \frac{\text{Recall}_{Hi} + \text{Recall}_{Sc} + \text{Recall}_{Ci} + \text{Recall}_{Bi}}{4}$$

$$\text{Recall Macro} = \frac{0,95 + 0,93 + 0,86 + 1}{4}$$

$$\text{Recall Macro} = \frac{3,74}{4}$$

Recall Macro = 0,94

Dari perhitungan di atas diketahui bahwa nilai *recall macro* saat nilai $K = 5$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *recall weighted* adalah seperti berikut ini.

$$\text{Recall Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$\text{Recall Weighted} = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FN_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$\text{Recall Weighted} = \frac{56 \frac{53}{53+3}}{320} + \frac{105 \frac{101}{101+4}}{320} + \frac{42 \frac{36}{36+6}}{320} + \frac{117 \frac{117}{117+0}}{320}$$

$$\text{Recall Weighted} = \frac{56 \times 0,94}{320} + \frac{105 \times 0,96}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 1}{320}$$

$$\text{Recall Weighted} = \frac{52,64}{320} + \frac{100,8}{320} + \frac{36,12}{320} + \frac{117}{320}$$

$$\text{Recall Weighted} = \frac{306,56}{320}$$

$$\text{Recall Weighted} = 0,96$$

Dari perhitungan di atas diketahui bahwa nilai *recall weighted* saat nilai $K = 5$ diperoleh sebesar 0,96 atau 96%. Selanjutnya untuk nilai *f1-score micro* dicari menggunakan persamaan (2.21), *f1-score macro* menggunakan persamaan (2.22), dan *f1-score weighted* menggunakan persamaan (2.23). Untuk perhitungan *f1-score micro* adalah seperti berikut ini.

$$F1 - \text{Score Micro} = \frac{2 \times \text{Precision Micro} \times \text{Recall Micro}}{\text{Precision Micro} + \text{Recall Micro}}$$

$$F1 - \text{Score Micro} = \frac{2 \times 0,93 \times 0,96}{0,93 + 0,96}$$

$$F1 - \text{Score Micro} = \frac{1,78}{1,89}$$

$$F1 - \text{Score Micro} = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score micro* saat nilai $K = 5$ diperoleh sebesar 0,94 atau 94%. Selanjutnya untuk perhitungan *f1-score macro* adalah seperti berikut ini.

$$F1 - \text{Score Macro} = \frac{2 \times \text{Precision Macro} \times \text{Recall Macro}}{\text{Precision Macro} + \text{Recall Macro}}$$

$$F1 - \text{Score Macro} = \frac{2 \times 0,93 \times 0,94}{0,93 + 0,94}$$

$$F1 - \text{Score Macro} = \frac{1,75}{1,87}$$

$$F1 - \text{Score Macro} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score macro* saat nilai $K = 5$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *f1-score weighted* adalah seperti berikut ini.

$$F1 - \text{Score Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{2TP_i}{2TP_i + FP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$F1 - \text{Score Weighted} = \frac{|y_{Hi}| \frac{2TP_{Hi}}{2TP_{Hi} + FP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{2TP_{Sc}}{2TP_{Sc} + FP_{Sc} + FN_{Sc}}}{|y_{total}|} +$$

$$\frac{|y_{Ci}| \frac{2TP_{Ci}}{2TP_{Ci} + FP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{2TP_{Bi}}{2TP_{Bi} + FP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$F1 - \text{Score Weighted} = \frac{56 \frac{2 \times 53}{2 \times 53 + 2 + 3}}{320} + \frac{105 \frac{2 \times 101}{2 \times 101 + 7 + 4}}{320} + \frac{42 \frac{2 \times 36}{2 \times 36 + 5 + 6}}{320} + \frac{117 \frac{2 \times 117}{2 \times 117 + 9 + 0}}{320}$$

$$F1 - \text{Score Weighted} = \frac{56 \frac{106}{111}}{320} + \frac{105 \frac{202}{213}}{320} + \frac{42 \frac{72}{83}}{320} + \frac{117 \frac{234}{243}}{320}$$

$$F1 - \text{Score Weighted} = \frac{56 \times 0,95}{320} + \frac{105 \times 0,95}{320} + \frac{42 \times 0,87}{320} + \frac{117 \times 0,96}{320}$$

$$F1 - \text{Score Weighted} = \frac{53,2}{320} + \frac{99,75}{320} + \frac{36,54}{320} + \frac{112,32}{320}$$

$$F1 - \text{Score Weighted} = \frac{301,81}{320}$$

$$F1 - \text{Score Weighted} = 0,94$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score weighted* saat nilai $K = 5$ diperoleh sebesar 0,94 atau 94%.

4.5.4 Nilai K = 7

Berikut ini *confusion matrix* label *Hinselmann* untuk nilai K = 7 pada tabel 4.37.

Tabel 4.37 *Confusion Matrix* Label *Hinselmann* K = 7

	Nilai Sebenarnya	
Nilai	43 (TP)	2 (FP)
Prediksi	13 (FN)	192 (TN)

Dari tabel di atas, diketahui pada label *Hinselman* saat nilai K = 7 terdapat 43 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 194 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 2 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 13 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Schiller* untuk nilai K = 7 pada tabel 4.38.

Tabel 4.38 *Confusion Matrix* Label *Schiller* K = 7

	Nilai Sebenarnya	
Nilai	88 (TP)	7 (FP)
Prediksi	17 (FN)	138 (TN)

Dari tabel di atas, diketahui pada label *Schiller* saat nilai K = 7 terdapat 88 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 138 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 7 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 17 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Citology* untuk nilai K = 7 pada tabel 4.39.

Tabel 4.39 *Confusion Matrix* Label *Citology* K = 7

	Nilai Sebenarnya	
Nilai	34 (TP)	5 (FP)
Prediksi	8 (FN)	203 (TN)

Dari tabel di atas, diketahui pada label *Citology* saat nilai K = 7 terdapat 34 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 203 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 5 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 8 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Biopsy* untuk nilai K = 7 pada tabel 4.40.

Tabel 4.40 *Confusion Matrix* Label *Biopsy* K = 7

	Nilai Sebenarnya	
Nilai	104 (TP)	9 (FP)
Prediksi	13 (FN)	124 (TN)

Dari tabel di atas, diketahui pada label *Biopsy* saat nilai K = 7 terdapat 104 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 124 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 9 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 13 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Nilai *hamming loss* dicari menggunakan persamaan (2.10), jika $\left| I(y_j^{(i)} \neq \hat{y}_j^{(i)}) \right| = 74$, yaitu jumlah data yang terprediksi tidak benar yang dihitung sesuai dengan data yang ada pada lampiran 14 dan dibandingkan dengan data yang ada pada lampiran 7, maka perhitungannya seperti berikut ini.

$$hloss(h) = \frac{1}{NL} \sum_{i=1}^N \sum_{j=1}^L \left| I(y_j^{(i)} \neq \hat{y}_j^{(i)}) \right|$$

$$hloss(h) = \frac{74}{250 \cdot 4}$$

$$hloss(h) = \frac{74}{1000}$$

$$hloss(h) = 0,074$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 7$, nilai *hamming loss* yang diperoleh sebesar 0,074 atau 7,4%.

Nilai akurasi dicari menggunakan persamaan (2.11), jika $\sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right| = 219 + 0,34$ yang dihitung sesuai dengan data yang ada pada lampiran 14 dan dibandingkan dengan data yang ada pada lampiran 7 menggunakan tabel kebenaran seperti pada tabel 2.12 dan hasilnya seperti pada lampiran 15, maka perhitungannya seperti berikut ini.

$$Accuracy = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i \cap \hat{Y}_i}{Y_i \cup \hat{Y}_i} \right|$$

$$Accuracy = \frac{1}{250} |219 + 0,34|$$

$$Accuracy = \frac{219,34}{250}$$

$$Accuracy = 0,88$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 7$, nilai akurasi yang diperoleh sebesar 0,88 atau 88%.

Nilai *precision*, *recall*, dan *f1-score* pada masing-masing label dicari menggunakan persamaan (2.12) untuk mencari nilai *precision*, persamaan (2.16) untuk mencari nilai *recall*, persamaan (2.20) untuk mencari nilai *f1-score*. Untuk perhitungan *precision* untuk label *Hinselmann* saat nilai $K = 7$ adalah seperti berikut ini.

$$Precision\ Hinselmann = \frac{TP}{TP+FP}$$

$$Precision\ Hinselmann = \frac{43}{43+2}$$

$$Precision\ Hinselmann = \frac{43}{45}$$

$$\text{Precision Hinselmann} = 0,95$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Hinselmann* saat nilai $K = 7$ diperoleh sebesar 0,95 atau 95%. Selanjutnya untuk perhitungan *recall* untuk label *Hinselmann* saat nilai $K = 5$ adalah seperti berikut ini.

$$\text{Recall Hinselmann} = \frac{TP}{TP+FN}$$

$$\text{Recall Hinselmann} = \frac{43}{43+13}$$

$$\text{Recall Hinselmann} = \frac{43}{56}$$

$$\text{Recall Hinselmann} = 0,76$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Hinselmann* saat nilai $K = 7$ diperoleh sebesar 0,76 atau 77%. Selanjutnya untuk perhitungan *f1-score* untuk label *Hinselmann* saat nilai $K = 7$ adalah seperti berikut ini.

$$F1 - \text{Score Hinselmann} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Hinselmann} = \frac{2 \times 0,95 \times 0,77}{0,95 + 0,77}$$

$$F1 - \text{Score Hinselmann} = \frac{1,46}{1,72}$$

$$F1 - \text{Score Hinselmann} = 0,85$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Hinselmann* saat nilai $K = 7$ diperoleh sebesar 0,85 atau 85%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Hinselmann* saat nilai $K = 7$ dapat dilihat pada tabel 4.41 berikut ini.

Tabel 4.41 *Precision, Recall, dan F1-Score Label Hinselmann K = 7*

	<i>Hinselmann</i>
<i>Precision</i>	0,95
<i>Recall</i>	0,76
<i>F1-Score</i>	0,85

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Hinselmann* saat nilai $K = 7$ diperoleh nilai *precision* sebesar 0,95 atau 95%, nilai *recall* sebesar 0,76 atau 76%, dan nilai *f1-score* sebesar 0,85 atau 85%. Selanjutnya untuk perhitungan *precision* untuk label *Schiller* saat nilai $K = 7$ seperti berikut ini.

$$\text{Precision Schiller} = \frac{TP}{TP+FP}$$

$$\text{Precision Schiller} = \frac{88}{88+7}$$

$$\text{Precision Schiller} = \frac{88}{95}$$

$$\text{Precision Schiller} = 0,93$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Schiller* saat nilai $K = 7$ diperoleh sebesar 0,93 atau 93%. Selanjutnya untuk perhitungan *recall* untuk label *Schiller* saat nilai $K = 7$ adalah seperti berikut ini

$$\text{Recall Schiller} = \frac{TP}{TP+FN}$$

$$\text{Recall Schiller} = \frac{88}{88+17}$$

$$\text{Recall Schiller} = \frac{88}{105}$$

$$\text{Recall Schiller} = 0,83$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Schiller* saat nilai $K = 7$ diperoleh sebesar 0,83 atau 83%. Selanjutnya untuk perhitungan *f1-score* untuk label *Schiller* saat nilai $K = 7$ adalah seperti berikut ini.

$$F1 - \text{Score Schiller} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Schiller} = \frac{2 \times 0,93 \times 0,84}{0,93 + 0,84}$$

$$F1 - \text{Score Schiller} = \frac{1,56}{1,77}$$

$$F1 - \text{Score Schiller} = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Schiller* saat nilai $K = 7$ diperoleh sebesar 0,88 atau 88%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Schiller* saat nilai $K = 7$ dapat dilihat pada tabel 4.42 berikut ini.

Tabel 4.42 *Precision*, *Recall*, dan *F1-Score* Label *Schiller* $K = 7$

	<i>Schiller</i>
<i>Precision</i>	0,93
<i>Recall</i>	0,83
<i>F1-Score</i>	0,88

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Schiller* saat nilai $K = 7$ diperoleh nilai sebesar 0,93 atau 93% untuk *precision*, 0,83 atau 83% untuk *recall*, dan 0,88 atau 88% untuk *f1-score*. Selanjutnya untuk perhitungan *precision* untuk label *Citology* saat nilai $K = 7$ seperti berikut ini.

$$\text{Precision Citology} = \frac{TP}{TP+FP}$$

$$\text{Precision Citology} = \frac{36}{36+5}$$

$$\text{Precision Citology} = \frac{36}{41}$$

$$\text{Precision Citology} = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Citology* saat nilai $K = 7$ diperoleh sebesar 0,88 atau 88%. Selanjutnya untuk perhitungan *recall* untuk label *Citology* saat nilai $K = 7$ adalah seperti berikut ini.

$$\text{Recall Citology} = \frac{TP}{TP+FN}$$

$$\text{Recall Citology} = \frac{36}{36+6}$$

$$\text{Recall Citology} = \frac{36}{42}$$

$$\text{Recall Citology} = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Citology* saat nilai $K = 7$ diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *f1-score* untuk label *Citology* saat nilai $K = 7$ adalah seperti berikut ini.

$$F1 - Score\ Citology = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score\ Citology = \frac{2 \times 0,88 \times 0,86}{0,88 + 0,86}$$

$$F1 - Score\ Citology = \frac{1,51}{1,74}$$

$$F1 - Score\ Citology = 0,87$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Citology* saat nilai $K = 7$ diperoleh sebesar 0,87 atau 87%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Citology* saat nilai $K = 7$ dapat dilihat pada tabel 4.43 berikut ini.

Tabel 4.43 *Precision, Recall, dan F1-Score Label Citology K = 7*

	<i>Citology</i>
<i>Precision</i>	0,88
<i>Recall</i>	0,86
<i>F1-Score</i>	0,87

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Citology* saat nilai $K = 1$ diperoleh nilai *precision* sebesar 0,88 atau 88%, nilai *recall* sebesar 0,86 atau 86%, dan nilai *recall* sebesar 0,87 atau 87%. Selanjutnya untuk perhitungan *precision* untuk label *Biopsy* saat nilai $K = 7$.

$$Precision\ Biopsy = \frac{TP}{TP + FP}$$

$$Precision\ Biopsy = \frac{104}{104 + 9}$$

$$Precision\ Biopsy = \frac{104}{113}$$

$$Precision\ Biopsy = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 7$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk perhitungan *recall* untuk label *Biopsy* saat nilai $K = 7$ adalah seperti berikut ini.

$$\text{Recall Biopsy} = \frac{TP}{TP+FN}$$

$$\text{Recall Biopsy} = \frac{104}{104+13}$$

$$\text{Recall Biopsy} = \frac{104}{117}$$

$$\text{Recall Biopsy} = 0,89$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai $K = 7$ diperoleh sebesar 0,89 atau 89%. Selanjutnya untuk perhitungan *f1-score* untuk label *Biopsy* saat nilai $K = 7$ adalah seperti berikut ini.

$$F1 - \text{Score Biopsy} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Biopsy} = \frac{2 \times 0,92 \times 0,89}{0,92 + 0,89}$$

$$F1 - \text{Score Biopsy} = \frac{1,64}{1,81}$$

$$F1 - \text{Score Biopsy} = 0,91$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Biopsy* saat nilai $K = 7$ diperoleh sebesar 0,91 atau 91%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Biopsy* saat nilai $K = 7$ dapat dilihat pada tabel 4.44 berikut ini.

Tabel 4.44 *Precision*, *Recall*, dan *F1-Score* Label *Biopsy* $K = 7$

	<i>Biopsy</i>
<i>Precision</i>	0,92
<i>Recall</i>	0,89
<i>F1-Score</i>	0,91

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Biopsy* saat nilai $K = 7$ diperoleh nilai sebesar 0,92 atau 92% untuk *precision*, 0,89 atau 89% untuk *recall*, dan 0,91 atau 91% untuk *f1-score*.

Untuk nilai *precision micro* dicari menggunakan persamaan (2.13), *precision macro* menggunakan persamaan (2.14), dan *precision weighted* menggunakan persamaan (2.15). Untuk perhitungan *precision micro* adalah seperti berikut ini.

$$Precision\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FP_1 + \dots + FP_n}$$

$$Precision\ Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FP_{Hi} + FP_{Sc} + FP_{Ci} + FP_{Bi}}$$

$$Precision\ Micro = \frac{43 + 88 + 34 + 104}{43 + 88 + 34 + 104 + 2 + 7 + 5 + 9}$$

$$Precision\ Micro = \frac{269}{292}$$

$$Precision\ Micro = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *precision micro* saat nilai $K = 7$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk perhitungan *precision macro* adalah seperti berikut ini.

$$Precision\ Macro = \frac{Precision_1 + \dots + Precision_n}{n}$$

$$Precision\ Macro = \frac{Precision_{Hinselmann} + Precision_{Schiller} + Precision_{Cytology} + Precision_{Biopsy}}{4}$$

$$Precision\ Macro = \frac{0,95 + 0,93 + 0,88 + 0,92}{4}$$

$$Precision\ Macro = \frac{3,68}{4}$$

$$Precision\ Macro = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *precision macro* saat nilai $K = 7$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk perhitungan *precision weighted* adalah seperti berikut ini.

$$Precision\ Weighted = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i+FP_i}}{\sum_{i=1}^n |y_i|}$$

$$Precision\ Weighted = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi}+FP_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc}+FP_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci}+FP_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi}+FP_{Bi}}}{|y_{total}|}$$

$$Precision\ Weighted = \frac{56 \frac{43}{43+2}}{320} + \frac{105 \frac{88}{88+7}}{320} + \frac{42 \frac{34}{34+5}}{320} + \frac{117 \frac{104}{104+9}}{320}$$

$$Precision\ Weighted = \frac{56 \times 0,95}{320} + \frac{105 \times 0,93}{320} + \frac{42 \times 0,87}{320} + \frac{117 \times 0,92}{320}$$

$$Precision\ Weighted = \frac{53,2}{320} + \frac{97,65}{320} + \frac{36,54}{320} + \frac{107,64}{320}$$

$$Precision\ Weighted = \frac{295,03}{320}$$

$$Precision\ Weighted = 0,92$$

Dari perhitungan di atas diketahui bahwa nilai *precision weighted* saat nilai $K = 7$ diperoleh sebesar 0,92 atau 92%. Selanjutnya untuk nilai *recall micro* dicari menggunakan persamaan (2.17), *recall macro* menggunakan persamaan (2.18), dan *recall weighted* menggunakan persamaan (2.19). Untuk perhitungan *recall micro* adalah seperti berikut ini.

$$Recall\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FN_1 + \dots + FN_n}$$

$$Recall\ Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FN_{Hi} + FN_{Sc} + FN_{Ci} + FN_{Bi}}$$

$$Recall\ Micro = \frac{43+88+34+104}{43+88+34+104+13+17+8+13}$$

$$Recall\ Micro = \frac{269}{320}$$

$$Recall\ Micro = 0,84$$

Dari perhitungan di atas diketahui bahwa nilai *recall micro* saat nilai $K = 7$ diperoleh sebesar 0,84 atau 84%. Selanjutnya untuk perhitungan *recall macro* adalah seperti berikut ini.

$$\text{Recall Macro} = \frac{\text{Recall}_1 + \dots + \text{Recall}_n}{n}$$

$$\text{Recall Macro} = \frac{\text{Recall}_{Hi} + \text{Recall}_{Sc} + \text{Recall}_{Ci} + \text{Recall}_{Bi}}{4}$$

$$\text{Recall Macro} = \frac{0,76 + 0,83 + 0,86 + 0,89}{4}$$

$$\text{Recall Macro} = \frac{3,34}{4}$$

$$\text{Recall Macro} = 0,83$$

Dari perhitungan di atas diketahui bahwa nilai *recall macro* saat nilai $K = 7$ diperoleh sebesar 0,83 atau 83%. Selanjutnya untuk perhitungan *recall weighted* adalah seperti berikut ini.

$$\text{Recall Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$\text{Recall Weighted} = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FN_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$\text{Recall Weighted} = \frac{56 \frac{43}{43+13}}{320} + \frac{105 \frac{88}{88+17}}{320} + \frac{42 \frac{34}{34+8}}{320} + \frac{117 \frac{104}{104+13}}{320}$$

$$\text{Recall Weighted} = \frac{56 \times 0,76}{320} + \frac{105 \times 0,83}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 0,89}{320}$$

$$\text{Recall Weighted} = \frac{42,56}{320} + \frac{87,15}{320} + \frac{36,12}{320} + \frac{104,13}{320}$$

$$\text{Recall Weighted} = \frac{269,96}{320}$$

$$\text{Recall Weighted} = 0,84$$

Dari perhitungan di atas diketahui bahwa nilai *recall weighted* saat nilai $K = 7$ diperoleh sebesar 0,84 atau 84%. Selanjutnya untuk nilai *f1-score micro* dicari menggunakan persamaan (2.21), *f1-score macro* menggunakan persamaan (2.22), dan *f1-score weighted* menggunakan persamaan (2.23). Untuk perhitungan *f1-score micro* adalah seperti berikut ini.

$$F1 - Score Micro = \frac{2 \times Precision Micro \times Recall Micro}{Precision Micro + Recall Micro}$$

$$F1 - Score Micro = \frac{2 \times 0,92 \times 0,84}{0,92 + 0,84}$$

$$F1 - Score Micro = \frac{1,54}{1,76}$$

$$F1 - Score Micro = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score micro* saat nilai $K = 7$ diperoleh sebesar 0,88 atau 88%. Selanjutnya untuk perhitungan *f1-score macro* adalah seperti berikut ini.

$$F1 - Score Macro = \frac{2 \times Precision Macro \times Recall Macro}{Precision Macro + Recall Macro}$$

$$F1 - Score Macro = \frac{2 \times 0,92 \times 0,83}{0,92 + 0,83}$$

$$F1 - Score Macro = \frac{1,53}{1,75}$$

$$F1 - Score Macro = 0,87$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score macro* saat nilai $K = 7$ diperoleh sebesar 0,87 atau 87%. Selanjutnya untuk perhitungan *f1-score weighted* adalah seperti berikut ini.

$$F1 - Score Weighted = \frac{\sum_{i=1}^n |y_i| \frac{2TP_i}{2TP_i + FP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$F1 - Score Weighted = \frac{|y_{Hi}| \frac{2TP_{Hi}}{2TP_{Hi} + FP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{2TP_{Sc}}{2TP_{Sc} + FP_{Sc} + FN_{Sc}}}{|y_{total}|} +$$

$$\frac{|y_{Ci}| \frac{2TP_{Ci}}{2TP_{Ci} + FP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{2TP_{Bi}}{2TP_{Bi} + FP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$F1 - Score Weighted = \frac{56 \frac{2 \times 43}{2 \times 43 + 2 + 13}}{320} + \frac{105 \frac{2 \times 88}{2 \times 88 + 7 + 17}}{320} + \frac{42 \frac{2 \times 34}{2 \times 34 + 5 + 8}}{320} + \frac{117 \frac{2 \times 104}{2 \times 104 + 9 + 13}}{320}$$

$$F1 - Score Weighted = \frac{56 \frac{86}{101}}{320} + \frac{105 \frac{176}{200}}{320} + \frac{42 \frac{68}{81}}{320} + \frac{117 \frac{208}{230}}{320}$$

$$F1 - Score Weighted = \frac{56 \times 0,85}{320} + \frac{105 \times 0,88}{320} + \frac{42 \times 0,84}{320} + \frac{117 \times 0,90}{320}$$

$$F1 - Score Weighted = \frac{47,6}{320} + \frac{92,4}{320} + \frac{35,28}{320} + \frac{105,3}{320}$$

$$F1 - Score Weighted = \frac{280,58}{320}$$

$$F1 - Score Weighted = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score weighted* saat nilai K = 7 diperoleh sebesar 0,88 atau 88%.

4.5.5 Nilai K = 9

Berikut ini *confusion matrix* label *Hinselmann* untuk nilai K = 9 pada tabel 4.45.

Tabel 4.45 *Confusion Matrix* Label *Hinselmann* K = 9

	Nilai Sebenarnya	
Nilai Prediksi	41 (TP)	4 (FP)
	15 (FN)	190 (TN)

Dari tabel di atas, diketahui pada label *Hinselman* saat nilai K = 9 terdapat 41 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 190 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 4 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 15 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Schiller* untuk nilai K = 9 pada tabel 4.46.

Tabel 4.46 *Confusion Matrix* Label *Schiller* K = 9

	Nilai Sebenarnya	
Nilai Prediksi	86 (TP)	4 (FP)
	19 (FN)	134 (TN)

Dari tabel di atas, diketahui pada label *Schiller* saat nilai $K = 9$ terdapat 86 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 134 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 4 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 19 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Citology* untuk nilai $K = 9$ pada tabel 4.47.

Tabel 4.47 *Confusion Matrix* Label *Citology* $K = 9$

	Nilai Sebenarnya	
Nilai	32 (TP)	5 (FP)
Prediksi	10 (FN)	203 (TN)

Dari tabel di atas, diketahui pada label *Citology* saat nilai $K = 9$ terdapat 32 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 203 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak 5 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 10 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Berikut ini *confusion matrix* label *Biopsy* untuk nilai $K = 9$ pada tabel 4.48.

Tabel 4.48 *Confusion Matrix* Label *Biopsy* $K = 9$

	Nilai Sebenarnya	
Nilai	100 (TP)	12 (FP)
Prediksi	17 (FN)	121 (TN)

Dari tabel di atas, diketahui pada label *Biopsy* saat nilai $K = 9$ terdapat 100 data *true positive* (TP) atau banyaknya data positif yang tergolong positif. Sebanyak 121 data *true negative* (TN) atau banyaknya data negatif yang tergolong negatif. Sebanyak

12 data *false positive* (FP) atau banyaknya data negatif yang tergolong sebagai positif dan sebanyak 17 data *False Negative* (FN) atau banyaknya data positif yang tergolong sebagai negatif.

Nilai *hamming loss* dicari menggunakan persamaan (2.10), jika $|I(y_j^{(i)} \neq \hat{y}_j^{(i)})| = 93$, yaitu jumlah data yang terprediksi tidak benar yang dihitung sesuai dengan data yang ada pada lampiran 16 dan dibandingkan dengan data yang ada pada lampiran 7, maka perhitungannya seperti berikut ini.

$$hloss(h) = \frac{1}{NL} \sum_{i=1}^N \sum_{j=1}^L |I(y_j^{(i)} \neq \hat{y}_j^{(i)})|$$

$$hloss(h) = \frac{93}{250 \cdot 4}$$

$$hloss(h) = \frac{93}{1000}$$

$$hloss(h) = 0,093$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 9$, nilai *hamming loss* yang diperoleh sebesar 0,093 atau 9,3%.

Nilai akurasi dicari menggunakan persamaan (2.11), jika $\sum_{i=1}^n \frac{|Y_i \cap \hat{Y}_i|}{|Y_i \cup \hat{Y}_i|} = 211 + 0,34$ yang dihitung sesuai dengan data yang ada pada lampiran 16 dan dibandingkan dengan data yang ada pada lampiran 7 menggunakan tabel kebenaran seperti pada tabel 2.12 dan hasilnya seperti pada lampiran 17, maka perhitungannya seperti berikut ini.

$$Accuracy = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i \cap \hat{Y}_i|}{|Y_i \cup \hat{Y}_i|}$$

$$Accuracy = \frac{1}{250} |211 + 0,34|$$

$$Accuracy = \frac{211,34}{250}$$

$$Accuracy = 0,84$$

Dari perhitungan di atas, diketahui bahwa saat nilai $K = 9$, nilai akurasi yang diperoleh sebesar 0,84 atau 84%.

Nilai *precision*, *recall*, dan *f1-score* pada masing-masing label dicari menggunakan persamaan (2.12) untuk mencari nilai *precision*, persamaan (2.16) untuk mencari nilai *recall*, persamaan (2.20) untuk mencari nilai *f1-score*. Untuk perhitungan *precision* untuk label *Hinselmann* saat nilai K = 9 adalah seperti berikut ini.

$$\text{Precision Hinselmann} = \frac{TP}{TP+FP}$$

$$\text{Precision Hinselmann} = \frac{41}{41+4}$$

$$\text{Precision Hinselmann} = \frac{41}{45}$$

$$\text{Precision Hinselmann} = 0,91$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Hinselmann* saat nilai K = 9 diperoleh sebesar 0,91 atau 91%. Selanjutnya untuk perhitungan *recall* untuk label *Hinselmann* saat nilai K = 9 adalah seperti berikut ini.

$$\text{Recall Hinselmann} = \frac{TP}{TP+FN}$$

$$\text{Recall Hinselmann} = \frac{41}{41+15}$$

$$\text{Recall Hinselmann} = \frac{41}{56}$$

$$\text{Recall Hinselmann} = 0,73$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Hinselmann* saat nilai K = 9 diperoleh sebesar 0,73 atau 73%. Selanjutnya untuk perhitungan *f1-score* untuk label *Hinselmann* saat nilai K = 9 adalah seperti berikut ini.

$$F1 - \text{Score Hinselmann} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Hinselmann} = \frac{2 \times 0,91 \times 0,73}{0,91 + 0,73}$$

$$F1 - \text{Score Hinselmann} = \frac{1,33}{1,64}$$

$$F1 - \text{Score Hinselmann} = 0,81$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Hinselmann* saat nilai $K = 9$ diperoleh sebesar 0,81 atau 81%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Hinselmann* saat nilai $K = 9$ dapat dilihat pada tabel 4.49 berikut ini.

Tabel 4.49 *Precision*, *Recall*, dan *F1-Score* Label *Hinselmann* $K = 9$

	<i>Hinselmann</i>
<i>Precision</i>	0,91
<i>Recall</i>	0,73
<i>F1-Score</i>	0,81

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Hinselmann* saat nilai $K = 1$ diperoleh nilai *precision* sebesar 0,91 atau 91%, nilai *recall* sebesar 0,73 atau 73%, dan nilai *f1-score* sebesar 0,81 atau 81%. Selanjutnya untuk perhitungan *precision* untuk label *Schiller* saat nilai $K = 9$ seperti berikut ini.

$$\textit{Precision Schiller} = \frac{TP}{TP+FP}$$

$$\textit{Precision Schiller} = \frac{86}{86+11}$$

$$\textit{Precision Schiller} = \frac{86}{97}$$

$$\textit{Precision Schiller} = 0,89$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Schiller* saat nilai $K = 9$ diperoleh sebesar 0,89 atau 89%. Selanjutnya untuk perhitungan *recall* untuk label *Schiller* saat nilai $K = 9$ adalah seperti berikut ini

$$\textit{Recall Schiller} = \frac{TP}{TP+FN}$$

$$\textit{Recall Schiller} = \frac{86}{86+19}$$

$$\textit{Recall Schiller} = \frac{86}{105}$$

$$\textit{Recall Schiller} = 0,82$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Schiller* saat nilai $K = 9$ diperoleh sebesar 0,82 atau 82%. Selanjutnya untuk perhitungan *f1-score* untuk label *Schiller* saat nilai $K = 9$ adalah seperti berikut ini.

$$F1 - Score Schiller = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F1 - Score Schiller = \frac{2 \times 0,89 \times 0,82}{0,89 + 0,82}$$

$$F1 - Score Schiller = \frac{1,46}{1,71}$$

$$F1 - Score Schiller = 0,85$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Schiller* saat nilai $K = 9$ diperoleh sebesar 0,85 atau 85%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Schiller* saat nilai $K = 9$ dapat dilihat pada tabel 4.50 berikut ini.

Tabel 4.50 *Precision*, *Recall*, dan *F1-Score* Label *Schiller* $K = 9$

	<i>Schiller</i>
<i>Precision</i>	0,89
<i>Recall</i>	0,82
<i>F1-Score</i>	0,85

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Schiller* saat nilai $K = 1$ diperoleh nilai sebesar 0,89 atau 89% untuk *precision*, 0,82 atau 82% untuk *recall*, dan 0,85 atau 85% untuk *f1-score*. Selanjutnya untuk perhitungan *precision* untuk label *Citology* saat nilai $K = 9$ seperti berikut ini.

$$Precision Citology = \frac{TP}{TP + FP}$$

$$Precision Citology = \frac{32}{32 + 5}$$

$$Precision Citology = \frac{32}{37}$$

$$Precision Citology = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Citology* saat nilai $K = 9$ diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *recall* untuk label *Citology* saat nilai $K = 9$ adalah seperti berikut ini.

$$\text{Recall Citology} = \frac{TP}{TP+FN}$$

$$\text{Recall Citology} = \frac{32}{32+10}$$

$$\text{Recall Citology} = \frac{32}{42}$$

$$\text{Recall Citology} = 0,76$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Citology* saat nilai $K = 9$ diperoleh sebesar 0,76 atau 76%. Selanjutnya untuk perhitungan *f1-score* untuk label *Citology* saat nilai $K = 9$ adalah seperti berikut ini.

$$F1 - \text{Score Citology} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Citology} = \frac{2 \times 0,86 \times 0,76}{0,86 + 0,76}$$

$$F1 - \text{Score Citology} = \frac{1,30}{1,62}$$

$$F1 - \text{Score Citology} = 0,80$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score* untuk label *Citology* saat nilai $K = 9$ diperoleh sebesar 0,80 atau 80%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Citology* saat nilai $K = 9$ dapat dilihat pada tabel 4.51 berikut ini.

Tabel 4.51 *Precision, Recall, dan F1-Score Label Citology K = 9*

	<i>Citology</i>
<i>Precision</i>	0,86
<i>Recall</i>	0,76
<i>F1-Score</i>	0,80

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Citology* saat nilai K = 9 diperoleh nilai *precision* sebesar 0,86 atau 86%, nilai *recall* sebesar 0,76 atau 76%, dan nilai *f1-score* sebesar 0,80 atau 80%. Selanjutnya untuk perhitungan *precision* untuk label *Biopsy* saat nilai K = 9.

$$\text{Precision Biopsy} = \frac{TP}{TP+FP}$$

$$\text{Precision Biopsy} = \frac{104}{104+12}$$

$$\text{Precision Biopsy} = \frac{104}{116}$$

$$\text{Precision Biopsy} = 0,90$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai K = 9 diperoleh sebesar 0,90 atau 90%. Selanjutnya untuk perhitungan *recall* untuk label *Biopsy* saat nilai K = 1 adalah seperti berikut ini.

$$\text{Recall Biopsy} = \frac{TP}{TP+FN}$$

$$\text{Recall Biopsy} = \frac{104}{104+17}$$

$$\text{Recall Biopsy} = \frac{104}{121}$$

$$\text{Recall Biopsy} = 0,86$$

Dari perhitungan di atas diketahui bahwa nilai *precision* untuk label *Biopsy* saat nilai K = 9 diperoleh sebesar 0,86 atau 86%. Selanjutnya untuk perhitungan *f1-score* untuk label *Biopsy* saat nilai K = 9 adalah seperti berikut ini.

$$F1 - \text{Score Biopsy} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$F1 - \text{Score Biopsy} = \frac{2 \times 0,90 \times 0,86}{0,90 + 0,86}$$

$$F1 - \text{Score Biopsy} = \frac{1,55}{1,76}$$

$$F1 - \text{Score Biopsy} = 0,88$$

Dari perhitungan di atas diketahui bahwa nilai *recall* untuk label *Biopsy* saat nilai $K = 9$ diperoleh sebesar 0,88 atau 88%. Nilai *precision*, *recall*, dan *f1-score* untuk label *Biopsy* saat nilai $K = 9$ dapat dilihat pada tabel 4.52 berikut ini.

Tabel 4.52 *Precision, Recall, dan F1-Score Label Biopsy K = 9*

	<i>Biopsy</i>
<i>Precision</i>	0,90
<i>Recall</i>	0,86
<i>F1-Score</i>	0,88

Dari tabel di atas diketahui untuk nilai *precision*, *recall*, dan *f1-score* pada label *Biopsy* saat nilai $K = 9$ diperoleh nilai sebesar 0,90 atau 90% untuk *precision*, 0,86 atau 86% untuk *recall*, dan 0,88 atau 88% untuk *f1-score*.

Untuk nilai *precision micro* dicari menggunakan persamaan (2.13), *precision macro* menggunakan persamaan (2.14), dan *precision weighted* menggunakan persamaan (2.15). Untuk perhitungan *precision micro* adalah seperti berikut ini.

$$Precision\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FP_1 + \dots + FP_n}$$

$$Precision\ Micro = \frac{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi}}{TP_{Hi} + TP_{Sc} + TP_{Ci} + TP_{Bi} + FP_{Hi} + FP_{Sc} + FP_{Ci} + FP_{Bi}}$$

$$Precision\ Micro = \frac{41 + 86 + 32 + 104}{41 + 86 + 32 + 104 + 4 + 11 + 5 + 12}$$

$$Precision\ Micro = \frac{263}{295}$$

$$Precision\ Micro = 0,89$$

Dari perhitungan di atas diketahui bahwa nilai *precision micro* saat nilai $K = 9$ diperoleh sebesar 0,89 atau 89%. Selanjutnya untuk perhitungan *precision macro* adalah seperti berikut ini.

$$Precision\ Macro = \frac{Precision_1 + \dots + Precision_n}{n}$$

$$Precision\ Macro = \frac{Precision_{Hinselmann} + Precision_{Schiller} + Precision_{Cytology} + Precision_{Biopsy}}{4}$$

$$Precision\ Macro = \frac{0,91 + 0,89 + 0,86 + 0,90}{4}$$

$$Precision\ Macro = \frac{3,56}{4}$$

$$Precision\ Macro = 0,89$$

Dari perhitungan di atas diketahui bahwa nilai *precision macro* saat nilai $K = 9$ diperoleh sebesar 0,89 atau 89%. Selanjutnya untuk perhitungan *precision weighted* adalah seperti berikut ini.

$$Precision\ Weighted = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i + FP_i}}{\sum_{i=1}^n |y_i|}$$

$$Precision\ Weighted = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi} + FP_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc} + FP_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci} + FP_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi} + FP_{Bi}}}{|y_{total}|}$$

$$Precision\ Weighted = \frac{56 \frac{41}{41+4}}{320} + \frac{105 \frac{86}{86+11}}{320} + \frac{42 \frac{32}{32+5}}{320} + \frac{117 \frac{100}{100+12}}{320}$$

$$Precision\ Weighted = \frac{56 \times 0,91}{320} + \frac{105 \times 0,89}{320} + \frac{42 \times 0,86}{320} + \frac{117 \times 0,90}{320}$$

$$Precision\ Weighted = \frac{50,96}{320} + \frac{93,45}{320} + \frac{36,12}{320} + \frac{105,03}{320}$$

$$Precision\ Weighted = \frac{285,56}{320}$$

$$Precision\ Weighted = 0,89$$

Dari perhitungan di atas diketahui bahwa nilai *precision weighted* saat nilai $K = 9$ diperoleh sebesar 0,89 atau 89%. Selanjutnya untuk nilai *recall micro* dicari menggunakan persamaan (2.17), *recall macro* menggunakan persamaan (2.18), dan *recall weighted* menggunakan persamaan (2.19). Untuk perhitungan *recall micro* adalah seperti berikut ini.

$$Recall\ Micro = \frac{TP_1 + \dots + TP_n}{TP_1 + \dots + TP_n + FN_1 + \dots + FN_n}$$

$$\text{Recall Micro} = \frac{TP_{Hi}+TP_{Sc}+TP_{Ci}+TP_{Bi}}{TP_{Hi}+TP_{Sc}+TP_{Ci}+TP_{Bi}+FN_{Hi}+FN_{Sc}+FN_{Ci}+FN_{Bi}}$$

$$\text{Recall Micro} = \frac{41+86+32+104}{41+86+32+104+15+19+10+17}$$

$$\text{Recall Micro} = \frac{263}{324}$$

$$\text{Recall Micro} = 0,81$$

Dari perhitungan di atas diketahui bahwa nilai *recall micro* saat nilai K = 9 diperoleh sebesar 0,81 atau 81%. Selanjutnya untuk perhitungan *recall macro* adalah seperti berikut ini.

$$\text{Recall Macro} = \frac{\text{Recall}_1+\dots+\text{Recall}_n}{n}$$

$$\text{Recall Macro} = \frac{\text{Recall}_{Hi}+\text{Recall}_{Sc}+\text{Recall}_{Ci}+\text{Recall}_{Bi}}{4}$$

$$\text{Recall Macro} = \frac{0,73+0,82+0,76+0,86}{4}$$

$$\text{Recall Macro} = \frac{3,17}{4}$$

$$\text{Recall Macro} = 0,79$$

Dari perhitungan di atas diketahui bahwa nilai *recall macro* saat nilai K = 9 diperoleh sebesar 0,79 atau 79%. Selanjutnya untuk perhitungan *recall weighted* adalah seperti berikut ini.

$$\text{Recall Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{TP_i}{TP_i+FN_i}}{\sum_{i=1}^n |y_i|}$$

$$\text{Recall Weighted} = \frac{|y_{Hi}| \frac{TP_{Hi}}{TP_{Hi}+FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{TP_{Sc}}{TP_{Sc}+FN_{Sc}}}{|y_{total}|} + \frac{|y_{Ci}| \frac{TP_{Ci}}{TP_{Ci}+FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{TP_{Bi}}{TP_{Bi}+FN_{Bi}}}{|y_{total}|}$$

$$\text{Recall Weighted} = \frac{56 \frac{41}{41+15}}{320} + \frac{105 \frac{86}{86+19}}{320} + \frac{42 \frac{32}{32+10}}{320} + \frac{117 \frac{100}{100+17}}{320}$$

$$\text{Recall Weighted} = \frac{56 \times 0,73}{320} + \frac{105 \times 0,82}{320} + \frac{42 \times 0,76}{320} + \frac{117 \times 0,86}{320}$$

$$Recall\ Weighted = \frac{40,88}{320} + \frac{86,1}{320} + \frac{31,92}{320} + \frac{100,62}{320}$$

$$Recall\ Weighted = \frac{259,82}{320}$$

$$Recall\ Weighted = 0,81$$

Dari perhitungan di atas diketahui bahwa nilai *recall weighted* saat nilai $K = 9$ diperoleh sebesar 0,81 atau 81%. Selanjutnya untuk nilai *f1-score micro* dicari menggunakan persamaan (2.21), *f1-score macro* menggunakan persamaan (2.22), dan *f1-score weighted* menggunakan persamaan (2.23). Untuk perhitungan *f1-score micro* adalah seperti berikut ini.

$$F1 - Score\ Micro = \frac{2 \times Precision\ Micro \times Recall\ Micro}{Precision\ Micro + Recall\ Micro}$$

$$F1 - Score\ Micro = \frac{2 \times 0,89 \times 0,81}{0,89 + 0,81}$$

$$F1 - Score\ Micro = \frac{1,44}{1,70}$$

$$F1 - Score\ Micro = 0,85$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score micro* saat nilai $K = 9$ diperoleh sebesar 0,85 atau 85%. Selanjutnya untuk perhitungan *f1-score macro* adalah seperti berikut ini.

$$F1 - Score\ Macro = \frac{2 \times Precision\ Macro \times Recall\ Macro}{Precision\ Macro + Recall\ Macro}$$

$$F1 - Score\ Macro = \frac{2 \times 0,89 \times 0,79}{0,89 + 0,79}$$

$$F1 - Score\ Macro = \frac{1,40}{1,68}$$

$$F1 - Score\ Macro = 0,84$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score macro* saat nilai $K = 9$ diperoleh sebesar 0,84 atau 84%. Selanjutnya untuk perhitungan *f1-score weighted* adalah seperti berikut ini.

$$F1 - \text{Score Weighted} = \frac{\sum_{i=1}^n |y_i| \frac{2TP_i}{2TP_i + FP_i + FN_i}}{\sum_{i=1}^n |y_i|}$$

$$F1 - \text{Score Weighted} = \frac{|y_{Hi}| \frac{2TP_{Hi}}{2TP_{Hi} + FP_{Hi} + FN_{Hi}}}{|y_{total}|} + \frac{|y_{Sc}| \frac{2TP_{Sc}}{2TP_{Sc} + FP_{Sc} + FN_{Sc}}}{|y_{total}|} +$$

$$\frac{|y_{Ci}| \frac{2TP_{Ci}}{2TP_{Ci} + FP_{Ci} + FN_{Ci}}}{|y_{total}|} + \frac{|y_{Bi}| \frac{2TP_{Bi}}{2TP_{Bi} + FP_{Bi} + FN_{Bi}}}{|y_{total}|}$$

$$F1 - \text{Score Weighted} = \frac{56 \frac{2 \times 41}{2 \times 41 + 4 + 15}}{320} + \frac{105 \frac{2 \times 86}{2 \times 86 + 11 + 19}}{320} + \frac{42 \frac{2 \times 32}{2 \times 32 + 5 + 10}}{320} + \frac{117 \frac{2 \times 100}{2 \times 100 + 12 + 17}}{320}$$

$$F1 - \text{Score Weighted} = \frac{56 \frac{82}{101}}{320} + \frac{105 \frac{172}{202}}{320} + \frac{42 \frac{64}{79}}{320} + \frac{117 \frac{200}{229}}{320}$$

$$F1 - \text{Score Weighted} = \frac{56 \times 0,81}{320} + \frac{105 \times 0,85}{320} + \frac{42 \times 0,81}{320} + \frac{117 \times 0,87}{320}$$

$$F1 - \text{Score Weighted} = \frac{45,36}{320} + \frac{89,25}{320} + \frac{34,02}{320} + \frac{101,79}{320}$$

$$F1 - \text{Score Weighted} = \frac{270,42}{320}$$

$$F1 - \text{Score Weighted} = 0,85$$

Dari perhitungan di atas diketahui bahwa nilai *f1-score weighted* saat nilai K = 9 diperoleh sebesar 0,85 atau 85%. Untuk hasil evaluasi klasifikasi menggunakan ML-KNN dengan K = 1, K = 3, K = 5, K = 7, dan K = 9 dengan menggunakan *dataset* yang di-*oversampling* kelas 1 pada kolom *Biopsy* dapat dilihat pada tabel 4.53 berikut ini.

Tabel 4.53 Hasil Evaluasi Klasifikasi Menggunakan ML-KNN

		K = 1	K = 3	K = 5	K = 7	K = 9
<i>Hamming Loss</i> (%)		3,9	4,2	3,59	7,39	9,3
Akurasi (%)		92	92	93	88	84
<i>Precision</i> (%)	<i>Micro</i>	92	94	93	92	89
<i>Recall</i> (%)	<i>Micro</i>	96	93	96	84	81

<i>F1-score (%)</i>	<i>Micro</i>	94	93	94	88	85
<i>Precision (%)</i>	<i>Macro</i>	91	93	93	92	89
<i>Recall (%)</i>	<i>Macro</i>	94	91	94	83	79
<i>F1-score (%)</i>	<i>Macro</i>	93	92	93	87	84
<i>Precision (%)</i>	<i>Weighted</i>	92	94	93	92	89
<i>Recall (%)</i>	<i>Weighted</i>	96	93	96	84	81
<i>F1-score (%)</i>	<i>Weighted</i>	94	93	94	88	85

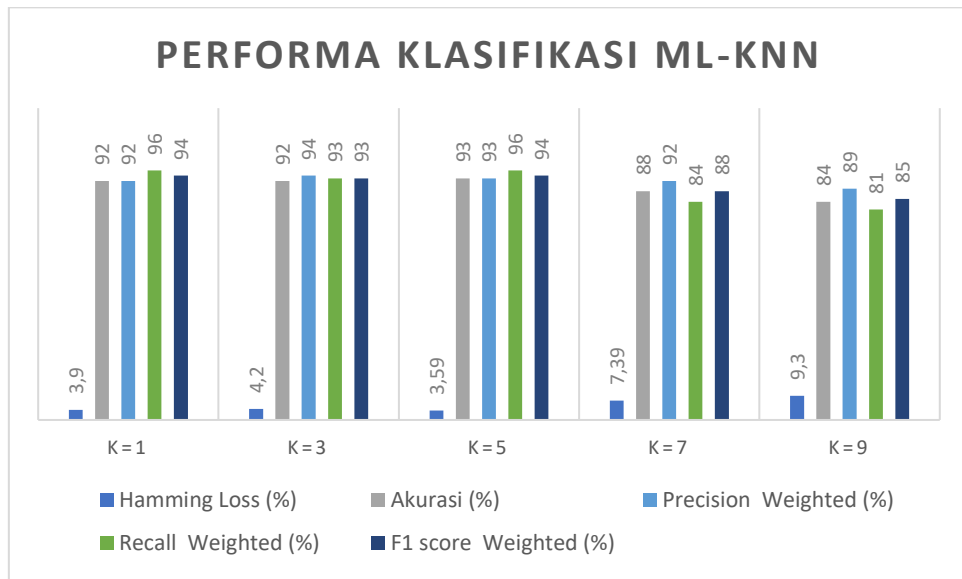
Keterangan Tabel:

Micro : Rata-rata yang digunakan ketika ada kebutuhan untuk menimbang setiap kejadian atau prediksi secara merata

Macro : Rata-rata yang digunakan ketika semua label perlu diperlakukan sama untuk mengevaluasi kinerja keseluruhan pengklasifikasi.

Weighted : Rata-rata di gunakan ketika terjadi ketidakseimbangan kelas

Dari tabel 4.53 diketahui bahwa nilai $K=5$ memperoleh nilai *hamming loss*, akurasi, *precision weighted*, *recall weighted*, dan *f1-score weighted* yang lebih baik daripada nilai k lainnya yaitu dengan nilai *hamming loss* sebesar 3,59%, akurasi sebesar 93%, *precision weighted* sebesar 93%, *recall weighted* sebesar 96%, dan *f1-score weighted* sebesar 94%.



Gambar 4.9 Visualisasi Evaluasi Klasifikasi ML-KNN

Dari gambar 4.9 terlihat bahwa klasifikasi ML-KNN dengan *oversampling* pada kolom *Biopsy* dengan nilai $K = 5$ menghasilkan nilai *hamming loss* yang lebih kecil dari nilai k lainnya dan nilai akurasi, *precision weighted*, *recall weighted*, dan *f1-score weighted* yang lebih besar dari pada nilai k lainnya.