# CHAPTER 5
# IMPLEMENTATION AND RESULTS

## 5.1. Implementation

This research uses the Python programming language, in this chapter we will discuss how the program works. The first step in this program is to retrieve the data first. The data retrieval process must have an API Key. What you will get is the consumer key, consumer secret, access token, and access secret.

```
1. consumer_key = 'Gq6DQpjxV4tMo3BH2zZmmD51T'
2. consumer_secret = 'Bg7XsPSmnlWyKeZLtZrsjW34WdULkGcUWHxwACIwa1J4exDlsK'
3. access_token = '1456889597883912228-CePiSfi4SZWXVGPOkDRLa9n8WDxKJc'
4. access_secret = 'yTNTjituna3Xa2pEAaatbhgmgt8GRbcW5nh8GiyM68WP5'
```

if you have got the API key, next is the process of retrieving data.

```
5. for tweet in tweepy.Cursor(api.search,q="vaksin, covid", lang="id", since="2021-11-01").items(10000):
6. print (tweet.created_at,tweet.author.screen_name,tweet.text)
7. csvWriter.writerow([tweet.created_at, tweet.author.screen_name, tweet.text])
```

In rows 5 and 6 is the data collection process, the data taken is data that queries the covid vaccine, in Indonesian, the data is taken starting November 1, 2021, and takes about 10000 data. The data taken from twitter is only the date, name, and tweet only. If it will be saved as a datacrawling.csv file.

Before calculating using the algorithm, the data must be processed by going through case folding, tokenization and filtering stopwords. The following is a case folding process that functions to convert all letters to lowercase. In addition to characters a to z will be omitted and will remove numbers and punctuation that have nothing to do with analysis.

```
1. df['TWEET'] = df['tweet'].str.lower()
2. df.head()
3.
4. def remove_punct(text):
```

```
5. #text    =    "".join([char    for    char    in    text    if    char    not    in
   string.punctuation])
6.      text = re.sub(r'[^a-zA-Z0-9]', ' ', str(text))
7.      text = re.sub(r'\b\w{1,2}\b', '', text) #menghilangkan 2 kata
8.      text = re.sub(r'\s\s+', ' ', text)
9.      text = re.sub(r"\d+", "", text)
10.     return text
11. df['tweet_clean'] = df['TWEET'].apply(lambda x: remove_punct(x))
12. df.head()
```

After the case folding process is complete, the next step is the tokenization process. Tokenization is the process of separating text into chunks of words so that they can then be analyzed.

```
1. def tokenization(text):
2.     text = re.split('\W+', text)
3.     return text
4.
5. df['TOKENIZATION'] = df['tweet_clean'].apply(lambda x:
   tokenization(x.lower()))
6. df.head()
```

If the tokenization process is complete, the next is the filtering process which functions to retrieve important words from the tokenization results, however, if common words that have no meaning in the data will be deleted and classified as general words.

```
1. Stopword = nltk.corpus.stopwords.words('indonesian')
2. def remove_stopwords(text):
3. text = [word for word in text if word not in stopword]
4. return text
5.
6. df['STOP_REMOVAL'] = df['TOKENIZATION'].apply
   (lambda x:remove_stopwords(x))
7. df.head()
8.
9. Stop_removal = df[['STOP_REMOVAL']]
10.
11.  def fit_stopwords(text):
12.     text = np.array(text)
13.     text = ' '.join(text)
14.     return text
15.
16.  df['STOP_REMOVAL'] = df['STOP_REMOVAL'].apply
    (lambda x: fit_stopwords(x))
17.  df.head()
```

12

Then this is the last process for the calculation of the algorithm. First, files that have gone through a pre-processing process will be subjected to a labelling process whose function is to separate negative and positive sentiments.

```
1. df['label'] = ''
2. for i,x in df. tweet_akhir.iteritems():
3. label = TextBlob(x)
4. df['label'][i] = label.sentiment.polarity
5. print("Index: ",i, "label", label.sentiment.polarity)
6.
7. def polarity_to_label(x):
8.     if(x >= -1 and x < 0):
9.         return 'negatif'
10.     if(x == 0):
11.         return 'neutral'
12.     if(x > 0 and x <=1):
13.         return 'positif'
14. df.label = df.label.apply(polarity_to_label)
```

In line 2 df. tweet_akhir is the column that will be labelled sentiment. If it is less than 0 then the sentence is included in the negative sentiment label. Meanwhile, if it is more than 0 then the sentence is included in the positive sentiment label. When finished, the results of the labelling process will be saved into a csv file with the name resultlabelling.csv.

After the labelling process is complete, the next step is to split the data. This is the process of dividing the data into 2 namely test data and train data.

```
1. Train_X, test_X, train_X, test_Y=model_selection.train_test_split
   (df[' tweet_akhir '], df['label'], test_size = 0.1, random_state = 0)
2.
3. df_train = pd.DataFrame()
4. df_train[' tweet_akhir '] = train_X
5. df_train['label'] = train_Y
6. df_test = pd.DataFrame()
7. df_test[' tweet_akhir '] = test_X
8. df_test['label'] = test_Y
```

In the data split process in line 1 is the process to separate the data that has been labelled into test and training data. In the random_state section, 0 is made so that there is no randomization in the split data, which means the sequence is still the same as before it was split. Lines 4 and 5 are splitting data into train data and lines 7 and 8 are splitting test data. The data taken is data from the tweet_akhir and label columns. When finished, the file will be saved in a csv file with the names df_train.csv and df_test.csv.

If the data split process is complete, the next step is the TF-IDF process to analyze the relationship between a sentence and a set of documents. After that, the last SVM algorithm testing process will be carried out and later accuracy results and others will appear.

```
1. predictions_SVM = model.predict(test_X_tfidf)
2. test_prediction = pd.DataFrame()
3. test_prediction[' tweet_akhir '] = test_X
4. test_prediction['label'] = predictions_SVM
5. SVM_accuracy = accuracy_score(predictions_SVM, test_Y)*100
6. SVM_accuracy = round(SVM_accuracy,1)
7.
8. test_prediction
9.
10.   SVM_accuracy
11.
12.   From sklearn.metrics import classification_report
```

## 5.2. Results

In this study, it showed that 92.4% of Twitter application users gave a neutral response to the topic of the Covid vaccine. But there are 0.8% of the people who gave a negative response to the covid vaccine. Meanwhile, 6.7% of people in Indonesia gave a positive response to the topic of the Covid vaccine in Indonesia.

In this study, before classifying the data, the thing that must be done is to retrieve data using the crawling method by having a twitter account to get the twitter API which will be used to retrieve data from twitter using the crawling method. Here are some data results that have been successfully retrieved from Twitter.



**Figure 5.1** some data results taken from twitter

The picture above is some of the data results that were successfully retrieved through the crawling method by having an API key that contains consumer_key, consumer_secret, access_token, and access_secret. The data taken are datetime, name, and tweet. The data that was successfully retrieved were 9981 data.

After the data has been successfully retrieved by the crawling method, then the data will go through a pre-processing process. In this process, case folding will be carried out which serves to convert all letters into lowercase letters, only letters a to z are accepted. Other than that character will be omitted and considered delimiter. At this stage only use the modules available in python and do not use any external libraries. In addition to changing the letters to lowercase. The result of this process is in the TWEET column. The following is the result of case folding which converts letters into lower case.

| | datetime | name | tweet | TWEET |
|---|---|---|---|---|
| 0 | 2021-11-23 13:06:35 | polresjembrana | Bhabinkamtibmas Kelurahan Tegalcangkring, Pols... | bhabinkamtibmas kelurahan tegalcangkring, pols... |
| 1 | 2021-11-23 13:05:54 | kemitris | @AldoBabeh Anies gagal dalam urusan vaksin, ma... | @aldobabeh anies gagal dalam urusan vaksin, ma... |
| 2 | 2021-11-23 13:05:37 | tvOneNews | Mobil Vaksin Covid-19 Keliling, Ridwan Kamil: ... | mobil vaksin covid-19 keliling, ridwan kamil: ... |
| 3 | 2021-11-23 13:05:29 | VvVwWwLq | RT @TedHilbert: Mau tahu kenapa orang2 seperti... | rt @tedhilbert: mau tahu kenapa orang2 seperti... |
| 4 | 2021-11-23 13:04:18 | jeff_ikhlas | RT @TedHilbert: Mau tahu kenapa orang2 seperti... | rt @tedhilbert: mau tahu kenapa orang2 seperti... |
| 5 | 2021-11-23 13:04:06 | StopPlandemit | RT @mas_prasetiyo: Tidak usah sok superior sep... | rt @mas_prasetiyo: tidak usah sok superior sep... |
| 6 | 2021-11-23 12:59:21 | kab_upp | UPP Kab.Bantaeng melaksanakan giat pengamanan ... | upp kab.bantaeng melaksanakan giat pengamanan ... |
| 7 | 2021-11-23 12:59:08 | Primaditaaaa | RT @NarasiNewsroom: Sebelas jenis vaksin COVID... | rt @narasinewsroom: sebelas jenis vaksin covid... |
| 8 | 2021-11-23 12:58:21 | kab_upp | UPP Kab.Bantaeng melaksanakan giat pengamanan ... | upp kab.bantaeng melaksanakan giat pengamanan ... |
| 9 | 2021-11-23 12:56:54 | apien_k19 | RT @TedHilbert: Must watch!\nIni Dokter sejati... | rt @tedhilbert: must watch!\nini dokter sejati... |
| 10 | 2021-11-23 12:56:45 | BKabsumedang | Sobat BUMN,\n \nSebagai bentuk kepedulian terh... | sobat bumn,\n \nsebagai bentuk kepedulian terh... |
| 11 | 2021-11-23 12:55:38 | DinkesOfficial | Data vaksin dan pasien covid-19 hari ini https... | data vaksin dan pasien covid-19 hari ini https... |
| 12 | 2021-11-23 12:55:30 | ubsibekasi | SENTRA VAKSIN COVID-19 UNIVERSITAS BSI KAMPUS ... | sentra vaksin covid-19 universitas bsi kampus ... |
| 13 | 2021-11-23 12:55:27 | Abughazy_2007 | RT @TedHilbert: Peningkatan 40% Kematian akib... | rt @tedhilbert: peningkatan 40% kematian akib... |
| 14 | 2021-11-23 12:54:16 | apien_k19 | RT @TedHilbert: Akhirnya mereka melakukan surv... | rt @tedhilbert: akhirnya mereka melakukan surv... |

**Figure 5.2** case folding converts all letters to lowercase

Furthermore, in the case folding process, it is also done to remove numeric characters and punctuation marks that have nothing to do with what will be analyzed. In this process python uses the re module to remove this character. The result of this process is in the tweet_clean column. The following is the result of case folding that removes numbers and punctuation marks.

| | datetime | name | tweet | TWEET | tweet_clean |
|---|---|---|---|---|---|
| 0 | 2021-11-23 13:06:35 | polresjembrana | Bhabinkamtibmas Kelurahan Tegalcangkring, Pols... | bhabinkamtibmas kelurahan tegalcangkring, pols... | bhabinkamtibmas kelurahan tegalcangkring polse... |
| 1 | 2021-11-23 13:05:54 | kemitris | @AldoBabeh Anies gagal dalam urusan vaksin, ma... | @aldobabeh anies gagal dalam urusan vaksin, ma... | aldobabeh anies gagal dalam urusan vaksin mas... |
| 2 | 2021-11-23 13:05:37 | tvOneNews | Mobil Vaksin Covid-19 Keliling, Ridwan Kamil: ... | mobil vaksin covid-19 keliling, ridwan kamil: ... | mobil vaksin covid keliling ridwan kamil kami ... |
| 3 | 2021-11-23 13:05:29 | VvVwWwLq | RT @TedHilbert: Mau tahu kenapa orang2 seperti... | rt @tedhilbert: mau tahu kenapa orang2 seperti... | tedhilbert mau tahu kenapa orang seperti koko... |
| 4 | 2021-11-23 13:04:18 | jeff_ikhlas | RT @TedHilbert: Mau tahu kenapa orang2 seperti... | rt @tedhilbert: mau tahu kenapa orang2 seperti... | tedhilbert mau tahu kenapa orang seperti koko... |
| 5 | 2021-11-23 13:04:06 | StopPlandemit | RT @mas_prasetiyo: Tidak usah sok superior sep... | rt @mas_prasetiyo: tidak usah sok superior sep... | mas prasetiyo tidak usah sok superior seperti... |
| 6 | 2021-11-23 12:59:21 | kab_upp | UPP Kab.Bantaeng melaksanakan giat pengamanan ... | upp kab.bantaeng melaksanakan giat pengamanan ... | upp kab bantaeng melaksanakan giat pengamanan ... |
| 7 | 2021-11-23 12:59:08 | Primaditaaaa | RT @NarasiNewsroom: Sebelas jenis vaksin COVID... | rt @narasinewsroom: sebelas jenis vaksin covid... | narasinewsroom sebelas jenis vaksin covid yan... |
| 8 | 2021-11-23 12:58:21 | kab_upp | UPP Kab.Bantaeng melaksanakan giat pengamanan ... | upp kab.bantaeng melaksanakan giat pengamanan ... | upp kab bantaeng melaksanakan giat pengamanan ... |
| 9 | 2021-11-23 12:56:54 | apien_k19 | RT @TedHilbert: Must watch!\nIni Dokter sejati... | rt @tedhilbert: must watch!\nini dokter sejati... | tedhilbert must watch ini dokter sejati yang ... |
| 10 | 2021-11-23 12:56:45 | BKabsumedang | Sobat BUMN, \n \nSebagai bentuk kepedulian terh... | sobat bumn, \n \nsebagai bentuk kepedulian terh... | sobat bumn sebagai bentuk kepedulian terhadap ... |
| 11 | 2021-11-23 12:55:38 | DinkesOfficial | Data vaksin dan pasien covid-19 hari ini https... | data vaksin dan pasien covid-19 hari ini https... | data vaksin dan pasien covid hari ini https xo... |
| 12 | 2021-11-23 12:55:30 | ubsibekasi | SENTRA VAKSIN COVID-19 UNIVERSITAS BSI KAMPUS ... | sentra vaksin covid-19 universitas bsi kampus ... | sentra vaksin covid universitas bsi kampus bog... |
| 13 | 2021-11-23 12:55:27 | Abughazy_2007 | RT @TedHilbert: Peningkatan 40% Kematian akib... | rt @tedhilbert: peningkatan 40% kematian akib... | tedhilbert peningkatan kematian akibat vaksin... |
| 14 | 2021-11-23 12:54:16 | apien_k19 | RT @TedHilbert: Akhirnya mereka melakukan surv... | rt @tedhilbert: akhirnya mereka melakukan surv... | tedhilbert akhirnya mereka melakukan survei s... |

**Figure 5.3** case folding results remove numbers and punctuation marks

After the case folding process is complete, then in the tweet_clean column the tokenization process will be carried out. This process serves to separate the text into pieces called tokens so that they can then be analyzed. The result of this process is in the TOKENIZATION column. The following is the result of tokenization.

| tweet_clean | TOKENIZATION |
|---|---|
| bhabinkamtibmas kelurahan tegalcangkring polse... | [bhabinkamtibmas, kelurahan, tegalcangkring, p... |
| aldobabeh anies gagal dalam urusan vaksin mas... | [, aldobabeh, anies, gagal, dalam, urusan, vak... |
| mobil vaksin covid keliling ridwan kamil kami ... | [mobil, vaksin, covid, keliling, ridwan, kamil... |
| tedhilbert mau tahu kenapa orang seperti koko... | [, tedhilbert, mau, tahu, kenapa, orang, seper... |
| tedhilbert mau tahu kenapa orang seperti koko... | [, tedhilbert, mau, tahu, kenapa, orang, seper... |
| mas prasetiyo tidak usah sok superior seperti... | [, mas, prasetiyo, tidak, usah, sok, superior,... |
| upp kab bantaeng melaksanakan giat pengamanan ... | [upp, kab, bantaeng, melaksanakan, giat, penga... |
| narasinewsroom sebelas jenis vaksin covid yan... | [, narasinewsroom, sebelas, jenis, vaksin, cov... |
| upp kab bantaeng melaksanakan giat pengamanan ... | [upp, kab, bantaeng, melaksanakan, giat, penga... |
| tedhilbert must watch ini dokter sejati yang ... | [, tedhilbert, must, watch, ini, dokter, sejat... |

**Figure 5.4** Result of tokenization process

After the tokenization process is complete, then in the TOKENIZATION column, a stopwords filtering process will be carried out. In this process, it functions to retrieve important words from the tokenization results by using a stoplist algorithm that functions to remove less important words and a wordlist to store important words. Words that often appear but have no meaning will be deleted and classified as general words. The result of this process is in the STOP_REMOVAL column. The following is the result of filtering stopwords.

17

| TOKENIZATION | STOP_REMOVAL |
|---|---|
| [bhabinkamtibmas, kelurahan, tegalcangkring, p... | bhabinkamtibmas kelurahan tegalcangkring polse... |
| [, aldobabeh, anies, gagal, dalam, urusan, vak... | aldobabeh anies gagal urusan vaksin mumpuni p... |
| [mobil, vaksin, covid, keliling, ridwan, kamil... | mobil vaksin covid keliling ridwan kamil menga... |
| [, tedhilbert, mau, tahu, kenapa, orang, seper... | tedhilbert orang koko blogdokter tonangardyan... |
| [, tedhilbert, mau, tahu, kenapa, orang, seper... | tedhilbert orang koko blogdokter tonangardyan... |
| [, mas, prasetiyo, tidak, usah, sok, superior,... | mas prasetiyo sok superior percaya vaksin cov... |
| [upp, kab, bantaeng, melaksanakan, giat, penga... | upp kab bantaeng melaksanakan giat pengamanan ... |
| [, narasinewsroom, sebelas, jenis, vaksin, cov... | narasinewsroom sebelas jenis vaksin covid izi... |
| [upp, kab, bantaeng, melaksanakan, giat, penga... | upp kab bantaeng melaksanakan giat pengamanan ... |

**Figure 5.5** stopwords filtering process

Furthermore, after the case folding process has been completed, a labeling process is carried out in the tweet_final column which functions to label sentiments which include negative comments, positive comments, and neutral comments. In this process there are 9981 data labeled. The results of the labeling process are in the label column. The following is the result of the labeling process.

| | datetime | name | tweet_akhir | label |
|---|---|---|---|---|
| 0 | 2021-11-23 13:06:35 | polresjembrana | bhabinkamtibmas kelurahan tegalcangkring polse... | neutral |
| 1 | 2021-11-23 13:05:54 | kemitris | aldobabeh anies gagal urusan vaksin mumpuni pe... | neutral |
| 2 | 2021-11-23 13:05:37 | tvOneNews | mobil vaksin covid keliling ridwan kamil menga... | neutral |
| 3 | 2021-11-23 13:05:29 | VVwWwLq | tedhilbert orang koko blogdokter tonangardyant... | neutral |
| 4 | 2021-11-23 13:04:18 | jeff_ikhlas | tedhilbert orang koko blogdokter tonangardyant... | neutral |
| ... | ... | ... | ... | ... |
| 9976 | 2021-11-17 6:53:51 | CapikWite | fahmibinhassan dos penggalak vaksin solusi ide... | positif |
| 9977 | 2021-11-17 6:53:09 | _mfdks | fahmibinhassan dos penggalak vaksin solusi ide... | positif |
| 9978 | 2021-11-17 6:52:52 | FirmanSupriadi7 | mediaindonesia pakar penyakit menular terkemuk... | neutral |
| 9979 | 2021-11-17 6:52:52 | EzrilF | fahmibinhassan dos penggalak vaksin solusi ide... | positif |
| 9980 | 2021-11-17 6:52:25 | khaalisz | fahmibinhassan dos penggalak vaksin solusi ide... | positif |

9981 rows × 4 columns

**Figure 5.6** the result of the labeling process

Furthermore, after the labeling process, the data will be split. This data will be divided into test data and train data. The test data will be used to determine the performance of the previously trained algorithm and when finding new data that has never been seen before.



**Figure 5.7** data test result

While the train data is used to train the algorithm used. After splitting the data, the test data has 999 data while the train data has 8982 data. The following are the results of the test data and train data that have been separated.

[10] df_train

| | tweet_akhir | label |
|---|---|---|
| 8786 | jpenerangan suntikan dos penggalak vaksin covi... | neutral |
| 6024 | cek fakta istri ceo pfizer meninggal akibat ko... | neutral |
| 5800 | indonesia terima juta dosis vaksin astrazeneca... | neutral |
| 6435 | sobat bumn bentuk kepedulian percepatan penang... | neutral |
| 2168 | kuala lumpur november malaysia menerima juta t... | neutral |
| ... | ... | ... |
| 9225 | kkmputrajaya kenyataan akhbar kementerian kesi... | neutral |
| 4859 | vaksin covid dikawal ahli bpom mui https uuwkgzvc | neutral |
| 3264 | alexcg hanux tedhilbert micmedsos aldobabeh sh... | neutral |
| 9845 | kkmputrajaya pelanjutan jangka hayat produk va... | neutral |
| 2732 | terbuka warga kalimantan utara ayo warga bulun... | neutral |

8982 rows × 2 columns

**Figure 5.8** data training result

After splitting the data, then a test is carried out using tf-idf which serves to reduce the weight of a term if its occurrence is spread throughout the document. Here are some of the results.

**Table 5.1.** Result from tf-idf

```
{'kelurahan': 1746, 'polsek': 3305, 'pengamanan': 3003, 'pemantauan': 2919,
'vaksin': 4321, 'covid': 695, 'door': 982, 'https': 1360, 'gagal': 1166,
'urusan': 4299, 'wab': 4501, 'mobil': 2515, 'keliling': 1736, 'ridwan':
3555, 'orang': 2730, 'koko': 1889, 'blogdokter': 521, 'tonangardyanto':
4197, 'sridiana': 3895, 'va': 4314, 'drpriono': 1005, 'meny': 2433, 'mas':
2150, 'prasetiyo': 3364, 'sok': 3868, 'superior': 3956, 'percaya': 3070,
'silakan': 3823, 'melindungi': 2214, 'upp': 4296, 'kab': 1602, 'bantaeng':
286, 'melaksanakan': 2197, 'giat': 1214, 'amp': 111, 'pendampingan': 2981,
'kantor': 1653, 'desa': 788, 'pattaneteang': 2830, 'kec': 1705, 'tompob':
4194, 'narasinewsroom': 2597, 'sebelas': 3680, 'jenis': 1552, 'izin': 1492,
'penggunaan': 3020, 'darurat': 743, 'badan': 239, 'pengawas': 3010, 'obat':
2706, 'makanan': 2105, 'bpom': 548, 'lurah': 2079, 'bonto': 537, 'sunggu':
3947, 'bissa': 508, 'must': 2572, 'watch': 4549, 'dokter': 972, 'sejati':
3697, 'mengerti': 2352, 'sales': 3618, 'indonesia': 1420, 'malu': 2120,
'dir': 913, 'sobat': 3864, 'bumn': 580, 'bentuk': 350, 'kepedulian': 1793,
'percepatan': 3073, 'penanganan': 2970, 'pusri': 3456, 'penyerahan': 3053,
'ban': 267, 'data': 748, 'pasien': 2806, 'sentra': 3745, 'universitas':
4286, 'bsi': 560, 'kampus': 1639, 'bogor': 528, 'cilebut': 651, 'hai':
1265, 'catat': 613, 'informasi': 1436, 'peningkatan': 3034, 'kematian':
1760, 'akibat': 63, 'gangguan': 1176, 'jantung': 1519, 'berusia': 463,
'melonjak': 2217, 'hingg': 1335, 'survei': 3962, 'serologi': 3774,
'berbahaya': 363, 'endemik': 1066, 'flu': 1133, 'pandemi': 2777, 'selesai':
```

```
   3727,  'varian':   4340,  'covovax':   702,  'bohong':   529,  'rakyat':   3487,
  'menghentikan':  2366,  'pdip':  2862,  'bonekanya':   536,  'negara':   2616,
  'sukangetweet': 3927, 'puanmaharani': 3433
```

Furthermore, in the process of classifying this data, the process of testing the SVM algorithm is carried out. and the final result of the test is the SVM algorithm has an accuracy of 98.6.

```
Classification report:
               precision   recall  f1-score   support

      negatif       0.80     0.50      0.62         8
      neutral       0.99     1.00      0.99       909
      positif       1.00     0.89      0.94        82

     accuracy                          0.99       999
    macro avg       0.93     0.80      0.85       999
 weighted avg       0.99     0.99      0.99       999

Confusion Matrix:
[[  4   4   0]
 [  1 908   0]
 [  0   9  73]]
```

**Figure 5.9** : classification report and confusion matrix

Below is the confusion matrix table.

**Table 5.2.** Confusion matrix

| Actual | Prediction | | |
|---|---|---|---|
| | negative | neutral | Positive |
| Negative | 4 | 4 | 0 |
| Neutral | 1 | 908 | 0 |
| Positive | 0 | 9 | 73 |

From the confusion matrix table above, the data is tested to get the recall, precision, and accuracy values. Here is how to get the accuracy value.

TP = 4 + 908 + 73 = 985

Total data = 999

$$\frac{TP}{\text{total data}} = \frac{985}{999} = 0{,}98$$

Next is calculating precision.

$$\frac{TP}{(TP + FP)}$$

If all are calculated, the result of all precision is

$$\frac{precision\ A+B+C}{jumlah\ kelas} = \frac{0,5+0,99+0,89}{3} = 1,78$$

Next is to calculate recall.

$$\frac{TP}{TP + FN}$$

If all are counted, the result of all recall is

$$\frac{Recall\ A+B+C}{jumlah\ kelas} = \frac{0,8+0,98+1}{3} = 2,11$$

So the results from the confusion matrix table above get an accuracy value of 0.98 while the results from all precision are 1.78, while the results from all recall are 2.11.