

**ANALISIS SENTIMENT MASYARAKAT TERHADAP  
*FINTECH* MENGGUNAKAN METODE *NAÏVE BAYES*  
*CLASSIFIER***

SKRIPSI

Untuk memenuhi sebagian persyaratan

mencapai derajat Sarjana S-1

Program Studi Teknik Informatika



Diajukan Oleh

NAFI'A NISA'UL HAMIDAH

15650018

Kepada

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA**

**YOGYAKARTA**

**2019**



KEMENTERIAN AGAMA  
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

## PENGESAHAN TUGAS AKHIR

Nomor : B-1885/Un.02/DST/PP.00.9/05/2019

Tugas Akhir dengan judul : Analisis Sentiment Masyarakat Terhadap FinTech Menggunakan Metode Naive Bayes Classifier

yang dipersiapkan dan disusun oleh:

Nama : NAFIA NISA'UL HAMIDAH  
Nomor Induk Mahasiswa : 15650018  
Telah diujikan pada : Selasa, 07 Mei 2019  
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

### TIM UJIAN TUGAS AKHIR

Ketua Sidang

Sumarsono, S.T., M.Kom.  
NIP. 19710209 200501 1 003

Penguji I

Nurochman, S.Kom., M.Kom.  
NIP. 19801223 200901 1 007

Penguji II

Rahmat Hidayat, S.Kom., M.Cs.  
NIP. 19850514 201503 1 002

Yogyakarta, 07 Mei 2019



Sumarsono, M.Si.  
NIP. 19591212 200003 1 001

## **SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR**

Hal :

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

*Assalamu'alaikum wr. wb.*

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Nafi'a Nisa'ul Hamidah

NIM : 15650018

Judul Skripsi : **Analisis Sentimen Masyarakat Terhadap *FinTech***

**Menggunakan Metode *Naive Bayes Classifier***

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

*Wassalamu'alaikum wr. wb.*

Yogyakarta, 29 April 2019

Pembimbing



Sumarsono, M.Kom.

NIP. 19710209 200501 1 00 3

## PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini :

Nama : NAFI'A NISA'UL HAMIDAH

NIM : 15650018

Jurusan : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi saya yang berjudul "**Analisis Sentimen Masyarakat Terhadap FinTech Menggunakan Metode Naïve Bayes Classifier**" merupakan hasil penelitian saya sendiri, tidak terdapat pada karya yang pernah di ajukan untuk memperoleh gelar kesarjana di suatu perguruan tinggi, dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 29 April 2019



Nafi'a Nisa'ul Hamidah  
NIM.15650018

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penelitian yang berjudul **Analisis Sentiment Masyarakat Terhadap *Fintech* Menggunakan Metode Naïve Bayes Classifier** sebagai salah satu syarat untuk mencapai gelar sarjana program studi Teknik INformatika UIN Sunan Kalijaga Yogyakarta. Sholawat serta salam selalu tercurahkan kepada junjungan kita Nabi Agung Muhammad SAW beserta seluruh keluarga dan sahabat beliau.

Penulis menyadari bahwa apa yang dilakukan dalam penyusunan laporan penelitian ini masih terlalu jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan kritik dan saran yang berguna dalam penyempurnaan analisa ini di masa yang akan datang. Semoga apa yang telah penulis lakukan dapat bermanfaat bagi pembaca.

Tidak lupa penulis juga mengucapkan terimakasih kepada pihak-pihak yang telah membantu dalam penyelesaian skripsi ini, baik secara langsung maupun tidak langsung. Ucapan terimakasih penulis sampaikan kepada:

1. Bapak Prof. Drs. K.H. Yudian Wahyudi, M.A., Ph.D., selaku Rektor UIN Sunan Kalijaga Yogyakarta.

2. Bapak Dr. Murtono, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Bapak Sumarsono, S.T., M.Kom., selaku Ketua Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta dan juga selaku Dosen Pembimbing Skripsi yang telah membimbing, memberikan waktu, motivasi, koreksi dan kritik saran kepada penulis sehingga skripsi ini dapat terselesaikan.
4. Bapak Dr. Bambang Sugiantoro, M.T., selaku Dosen Pembimbing Akademik.
5. Bapak dan Ibu Dosen Teknik Informatika, selaku dosen pengampu mata kuliah program studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta yang telah banyak membantu penulis hingga penulis dapat menyusun tugas akhir.
6. Seluruh staf dan karyawan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
7. Ucapan terimakasih yang terdalam untuk kedua orangtua, Ibu Siti Atifah dan Bapak Suheni yang selalu memberikan doa, perhatian, kasih sayang dan semua support yang telah diberikan.
8. Kedua adek-adekku Nazula dan Nisa, Uti Ah, Bude dan Pakde La-

sidi, Bude Rum, Ibu Kom dan Abah Nur, Elisa Verma dan Arvind Verma, dan seluruh keluarga besar dari bapak dan ibu. Terimakasih atas kasih sayang, semua doa dan segala bentuk support yang telah kalian berikan.

9. Sahabat-sahabatku tersayang Fauzan, Irsalina, Ozi, Annisa, Faisal, Muftia, Riko, Fahrul, Dani yang selalu menemani penulis serta senantiasa memberikan semangat dan dorongan agar menyelesaikan skripsi ini.
10. Umam dan Lana yang telah senantiasa membantu penulis dalam proses awal berjalannya skripsi hingga selesai.
11. Seluruh Teman-Teman Teknik Informatika 2015 yang tidak dapat penulis sebutkan satu-satu.
12. Teman-teman Himpunan Mahasiswa Teknik Informatika UIN Sunan Kalijaga dan Seluruh Keluarga Teknik Informatika UIN Sunan Kalijaga Yogyakarta
13. Pak Edwien dan semua team PT. Kreasi Kode Indonesia yang telah banyak memberikan pengalaman bekerja, support dan doa untuk penulis.
14. Ka Fahmi, Aca, TRI, serta teman-teman Ambassador Cicil.co.id dan

semua team PT. Cicil Solusi Mitra Teknologi yang telah banyak memberikan ilmu dan pengalaman baru untuk penulis.

15. Teman-teman FKMTIF PTAI, PERMIKOMNAS, FORSAMANTA, FKMT, RADIOACTIVE, KKN Angkatan 96 Kelompok 59.
16. Bapak, ibu Bambang Ananto dan semua teman-teman Kos BS-39
17. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu dan telah memberikan banyak campur tangan, doa, support sehingga penelitian ini dapat terselesaikan.

Yogyakarta, 29 April 2019

Penulis

## HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur, Karya sederhana ini saya persembahkan untuk :

1. Bapak dan Ibu tercinta yang senantiasa memberikan dukungan sampai sekarang. Terimakasih banyak atas semua doa, nasihat, semangat dan semua yang telah kalian berikan hingga nafi bisa sampai pada titik sekarang.
2. Adek Nisa dan adek zula, Terimakasih telah menjadi penyemangat kakak
3. Uti Ah, Bude dan Pakde Lasidi, Bude Rum, Ibu Kom dan Abah Nur, Elisa Verma dan Arvind Verma dan seluruh keluarga besar dari bapak dan ibu ku tercinta terimakasih banyak atas kasih sayang, semua doa dan segala bentuk support yang telah kalian berikan.
4. Fauzan, Fauzi Robbani, Irsalina, Muftia, Annisa Cibi, Faisal, Riko, Dani, Fahrul, Lana, Umam. Terimakasih telah menjadi sahabat, keluarga, seperjuangan, berbagi, menginspirasi, dan memotivasi.
5. Teman-teman seperjuangan Teknik Informatika 2015, Himpunan Mahasiswa Program Studi Teknik Informatika dan Seluruh Keluarga Teknik Informatika UIN Sunan Kalijaga Yogyakarta

6. Almamater tercinta UIN Sunan Kalijaga Yogyakarta
7. PT. Kreasi Kode Indonesia
8. PT. Cicil Solusi Mitra Teknologi
9. Bapak, Ibu Bambang Ananto dan semua teman-teman Kos BS-39
10. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu dan telah memberikan banyak campur tangan, doa, support sehingga penelitian ini dapat terselesaikan.

## HALAMAN MOTTO

*"Karena setiap kejadian yang kita alami adalah proses menuju puncak kemenangan"*

- Nafi -

*"Jika Kamu tidak dapat menahan lelahnya belajar, Maka kamu harus sanggup menahan perihnya Kebodohan."*

- Imam Syafi'i -

*"Aku melihat pemilik ilmu hidupnya mulia walau ia dilahirkan dari orangtua terhina. Ia terus menerus terangkat hingga pada derajat tinggi dan mulia.."*

- Imam Syafi'i -

*"Tidak pernah salah untuk menjadi seorang pemaaf"*

- Nafi -

## DAFTAR ISI

|                                               |              |
|-----------------------------------------------|--------------|
| <b>HALAMAN JUDUL</b>                          | <b>i</b>     |
| <b>LEMBAR PENGESAHAN SKRIPSI/TUGAS AKHIR</b>  | <b>ii</b>    |
| <b>LEMBAR PERSETUJUAN SKRIPSI/TUGAS AKHIR</b> | <b>iii</b>   |
| <b>HALAMAN PERNYATAAN KEASLIAN SKRIPSI</b>    | <b>iv</b>    |
| <b>KATA PENGANTAR</b>                         | <b>v</b>     |
| <b>HALAMAN PERSEMBAHAN</b>                    | <b>ix</b>    |
| <b>HALAMAN MOTTO</b>                          | <b>xi</b>    |
| <b>DAFTAR ISI</b>                             | <b>xii</b>   |
| <b>DAFTAR TABEL</b>                           | <b>xv</b>    |
| <b>DAFTAR GAMBAR</b>                          | <b>xvii</b>  |
| <b>INTISARI</b>                               | <b>xviii</b> |
| <b>ABSTRACT</b>                               | <b>xix</b>   |
| <b>I PENDAHULUAN</b>                          | <b>1</b>     |
| 1.1. Latar Belakang                           | 1            |
| 1.2. Rumusan Masalah                          | 5            |
| 1.3. Batasan Masalah                          | 5            |
| 1.4. Tujuan Penelitian                        | 6            |
| 1.5. Manfaat Penelitian                       | 6            |
| 1.6. Keaslian Penelitian                      | 7            |
| 1.7. Sistematika Penulisan                    | 8            |

|            |                                                      |           |
|------------|------------------------------------------------------|-----------|
| <b>II</b>  | <b>TINJAUAN PUSTAKA DAN LANDASAN TEORI . . . . .</b> | <b>10</b> |
| 2.1.       | Tinjauan Pustaka . . . . .                           | 10        |
| 2.2.       | Landasan Teori . . . . .                             | 22        |
| 2.2.1.     | <i>Text Mining</i> . . . . .                         | 22        |
| 2.2.2.     | Analisis Sentimen . . . . .                          | 31        |
| 2.2.3.     | <i>Naïve Bayes Classifier</i> . . . . .              | 32        |
| 2.2.4.     | TF dan TF-IDF . . . . .                              | 36        |
| 2.2.5.     | Klasifikasi . . . . .                                | 38        |
| 2.2.6.     | Python . . . . .                                     | 38        |
| 2.2.7.     | <i>Twitter</i> . . . . .                             | 39        |
| 2.2.8.     | <i>FinTech</i> . . . . .                             | 41        |
| 2.2.9.     | <i>Cross Validation</i> . . . . .                    | 44        |
| <b>III</b> | <b>METODE PENELITIAN . . . . .</b>                   | <b>48</b> |
| 3.1.       | Metode Penelitian . . . . .                          | 48        |
| 3.2.       | Tahapan-tahapan penelitian . . . . .                 | 48        |
| 3.2.1.     | Studi Pustaka . . . . .                              | 49        |
| 3.2.2.     | Pengumpulan Data . . . . .                           | 49        |
| 3.2.3.     | Seleksi Data . . . . .                               | 49        |
| 3.2.4.     | <i>Preprocessing</i> . . . . .                       | 50        |
| 3.2.5.     | Analisa Hasil . . . . .                              | 50        |
| 3.2.6.     | Implementasi . . . . .                               | 51        |
| <b>IV</b>  | <b>HASIL DAN PEMBAHASAN . . . . .</b>                | <b>52</b> |

|                                               |            |
|-----------------------------------------------|------------|
| 4.1. Pengumpulan Data . . . . .               | 52         |
| 4.2. Pembersihan dan Pelabelan Data . . . . . | 54         |
| 4.3. <i>Preprocessing</i> . . . . .           | 56         |
| 4.3.1. <i>Cleansing</i> . . . . .             | 56         |
| 4.3.2. <i>Case Folding</i> . . . . .          | 58         |
| 4.3.3. Tokenisasi . . . . .                   | 59         |
| 4.3.4. <i>Slangword</i> . . . . .             | 60         |
| 4.3.5. <i>StopWord Removal</i> . . . . .      | 61         |
| 4.3.6. <i>Stemming</i> . . . . .              | 62         |
| 4.4. Analisa dan Evaluasi . . . . .           | 63         |
| 4.4.1. Analisa . . . . .                      | 63         |
| 4.4.2. Evaluasi Model Klasifikasi . . . . .   | 80         |
| 4.5. Implementasi . . . . .                   | 83         |
| <b>V PENUTUP . . . . .</b>                    | <b>85</b>  |
| 5.1. Kesimpulan . . . . .                     | 85         |
| 5.2. Saran . . . . .                          | 86         |
| <b>DAFTAR PUSTAKA . . . . .</b>               | <b>87</b>  |
| <b>A DATA . . . . .</b>                       | <b>94</b>  |
| <b>B SKRIP PROGRAM PYTHON . . . . .</b>       | <b>97</b>  |
| <b>C CURICULUM VITAE . . . . .</b>            | <b>111</b> |

## DAFTAR TABEL

|      |                                                     |    |
|------|-----------------------------------------------------|----|
| 2.1  | Tinjaun Pustaka . . . . .                           | 17 |
| 2.2  | Lanjutan Tinjaun Pustaka . . . . .                  | 18 |
| 2.3  | Lanjutan Tinjaun Pustaka . . . . .                  | 19 |
| 2.4  | Lanjutan Tinjaun Pustaka . . . . .                  | 20 |
| 2.5  | Lanjutan Tinjaun Pustaka . . . . .                  | 21 |
| 2.6  | Contoh <i>Cleansing</i> . . . . .                   | 23 |
| 2.7  | Contoh <i>Case folding</i> . . . . .                | 24 |
| 2.8  | Contoh <i>Tokenizing</i> . . . . .                  | 25 |
| 2.9  | Contoh <i>Slangword</i> . . . . .                   | 25 |
| 2.10 | Contoh <i>StopWord Removal</i> . . . . .            | 27 |
| 2.11 | Contoh <i>stemming</i> . . . . .                    | 30 |
| 4.1  | Contoh Data <i>Tweet</i> . . . . .                  | 53 |
| 4.2  | Contoh Pelabelan Data . . . . .                     | 55 |
| 4.3  | Lanjutan Contoh Pelabelan Data . . . . .            | 56 |
| 4.4  | Contoh Hasil <i>Cleansing</i> . . . . .             | 57 |
| 4.5  | Contoh Hasil <i>Case Folding</i> . . . . .          | 58 |
| 4.6  | Lanjutan Contoh Hasil <i>Case Folding</i> . . . . . | 59 |
| 4.7  | Contoh Hasil Tokenisasi . . . . .                   | 60 |
| 4.8  | Contoh Hasil <i>Stopword</i> . . . . .              | 61 |
| 4.9  | Contoh Hasil <i>Stopword</i> . . . . .              | 61 |

|      |                                                                         |    |
|------|-------------------------------------------------------------------------|----|
| 4.10 | Lanjutan Contoh Hasil <i>Stopword</i> . . . . .                         | 62 |
| 4.11 | Contoh Hasil <i>Steming</i> . . . . .                                   | 62 |
| 4.12 | Lanjutan Contoh Hasil <i>Stemming</i> . . . . .                         | 63 |
| 4.13 | Contoh Hasil <i>Preprocessing</i> Data <i>Tweet</i> beserta label . . . | 64 |
| 4.14 | <i>Term Frequency</i> . . . . .                                         | 64 |
| 4.15 | <i>Lanjutan Term Frequency</i> . . . . .                                | 65 |
| 4.16 | Peluang <i>Term</i> Sentimen Negatif . . . . .                          | 67 |
| 4.17 | Lanjutan Peluang <i>Term</i> Sentimen Negatif . . . . .                 | 68 |
| 4.18 | Peluang <i>Trem</i> Sentimen Netral . . . . .                           | 69 |
| 4.19 | Lanjutan Peluang <i>Trem</i> Sentimen Netral . . . . .                  | 70 |
| 4.20 | Peluang <i>Trem</i> Sentimen Positif . . . . .                          | 71 |
| 4.21 | Lanjutan Peluang <i>Trem</i> Sentimen Positif . . . . .                 | 72 |
| 4.22 | Peluang <i>Trem</i> Sentimen <i>Irrelevant</i> . . . . .                | 72 |
| 4.23 | Lanjutan Peluang <i>Trem</i> Sentimen <i>Irrelevant</i> . . . . .       | 73 |
| 4.24 | Contoh Perhitungan <i>TF-IDF</i> . . . . .                              | 78 |
| 4.25 | Lanjutan Contoh Perhitungan <i>TF-IDF</i> . . . . .                     | 79 |

## DAFTAR GAMBAR

|     |                                                                              |    |
|-----|------------------------------------------------------------------------------|----|
| 2.1 | Metode Pembagian Data dengan <i>Random Subsampling</i> . .                   | 45 |
| 2.2 | Metode Pembagian Data dengan <i>K-Fold Cross Validation</i> .                | 46 |
| 2.3 | Metode Pembagian Data dengan <i>Leave-one-out Cross Validation</i> . . . . . | 47 |
| 3.1 | Tahapan Penelitian . . . . .                                                 | 48 |
| 4.1 | Proses Analisa Sentimen . . . . .                                            | 52 |
| 4.2 | Grafik Akurasi <i>K-fold cross validation</i> TF . . . . .                   | 81 |
| 4.3 | Grafik Akurasi <i>K-fold cross validation</i> TFIDF . . . . .                | 81 |
| 4.4 | Hasil Implementasi Model Klasifikasi pada Data Uji . . .                     | 84 |

# **Analisis Sentiment Masyarakat Terhadap *Fintech* Menggunakan Metode *Naïve Bayes Classifier***

NAFI'A NISA'UL HAMIDAH

15650018

## **INTISARI**

*Twitter* merupakan situs *microbloging* yang banyak digemari oleh masyarakat. Salah satu topik cuitan warganet di *Twitter* adalah *FinTech*, *Fin-tech* atau *financial technology* saat ini sedang menjadi pembicaraan yang sangat menarik. Maka dari itu dilakukan penelitian mengenai Analisis Sentimen terhadap *Fintech*.

Analisis sentimen dengan kata kunci *fintech* dengan metode *Naïve Bayes Classifier*. *Naïve Bayes Classifier* merupakan metode klasifikasi dengan memanfaatkan probabilitas dan statistik. Data yang digunakan sebanyak 6050 data dengan komposisi 2500 data latih dengan label dan 3550 data uji. Pada penelitian metode yang digunakan adalah *Naïve Bayes Classifier*.

Penelitian yang dilakukan menggunakan dua model pembobotan, yaitu TF dan TF-IDF dan menggunakan dua model akurasi, yaitu *K-Fold Cross Validation* dan *Leave One Out Cross Validation*. Pembobotan *term frequency* (TF) pada data latih *Fintech* dengan akurasi *K-Fold Cross Validation* mendapat nilai akurasi sebesar 76,02%. Sedangkan dengan pembobotan TF dengan akurasi *Leave One Out Cross Validation* mendapat nilai akurasi sebesar 71,68%. Menggunakan pembobotan *term frequency-invers document frequency* (TF-IDF) pada data latih *Fintech* dengan metode akurasi *K-Fold Cross Validation* mendapat nilai akurasi sebesar 67,24%. Sedangkan dengan pembobotan metode akurasi *Leave One Out Cross Validation* mendapat nilai akurasi sebesar 51,80%.

**Kata Kunci :** *analisis sentimen, twitter, fintech, k-fold cross validation, leave one out cross validation, naive bayes classifier, tf, tf-idf, klasifikasi*

# *Sentiment Analysis to Society towards Fintech Using Naïve Bayes*

## *Classifier Method*

NAFI'A NISA'UL HAMIDAH

15650018

## **ABSTRACT**

Twitter is a microblogging site that is popular with many people. One of the topics of warganet on Twitter is FinTech, Fintech or financial technology is now a very interesting conversation. So from that carried out research on Sentiment Analysis of Fintech.

Sentiment analysis with fintech keywords using the Naive Bayes Classifier method. Naive Bayes Classifier is a classification method by utilizing probabilities and statistics. The data used were 6050 data with a composition of 2500 data trained with labels and 3550 test data tweet. The researcher uses the Naive Bayes Classifier method.

Research carried out using two weighting models, namely TF and TF-IDF and using two models of accuracy, namely K-Fold Cross Validation and Leave One Out Cross Validation. Weighting term frequency (TF) on Fintech training data with the accuracy of K-Fold Cross Validation received an accuracy value of 76.02%. Whereas by weighting TF with textit accuracy Leave One Out Cross Validation gets an accuracy value of 71.68%. Using weighting term frequency-invers document frequency (TF-IDF) on Fintech training data with accuracy method K-Fold Cross Validation gets a value accuracy of 67.24%. Whereas by weighting the accuracy method Leave One Out Cross Validation gets an accuracy value of 51.80%.

**Key Word** : *analisis sentimen, twitter, fintech, k-fold cross validation, leave ono out cross validation, naive bayes classifier, tf, tf-idf, klasifikasi*

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Adanya revolusi Industri 4.0, yang menggunakan teknologi internet sebagai penggerak utama ekonomi, merombak kondisi ekonomi diberbagai belahan dunia tak lain juga Indonesia. Inovasi ini seharusnya jadi momentum penting perkembangan konsep ekonomi kerakyatan di Indonesia. Karena dengan adanya ekonomi digital, rakyat dapat berpartisipasi menuju pasar dunia tanpa harus mengeluarkan modal besar (Bintarto 2018). Saat ini perkembangan Teknologi yang pesat serta didukung oleh infrastruktur komunikasi yang semakin kuat dan stabil telah membawa dampak baru terhadap masyarakat. Inovasi perkembangan teknologi dan kecerdasan tak luput mempengaruhi industri keuangan dan tata kelola. Munculnya *financial technology* dalam berbagai platfrom bisa di rasakan oleh masyarakat. Menurut Bank Indonesia, *financial technology* merupakan hasil gabungan antara jasa keuangan dengan teknologi yang akhirnya mengubah model bisnis dari konvensional menjadi moderat (Akmal 2019) .

*Financial Technology* atau *FinTech* telah menyerap banyak perhatian pada akhir-akhir ini. Industri layanan finansial yang menyediakan berbagai

macam layanan dari finansial, perbankan dan asuransi telah menyimpan perhatian yang lebih pada keputusan untuk mengadopsi teknologi dalam rangka untuk berinovasi dan berkembang (Abyan 2018). Menurut Dhar (Dhar & Stein 2017), definisi dari *financial technology* merupakan sebuah inovasi dari sektor finansial yang melibatkan model bisnis yang telah terintegrasi dengan teknologi yang dapat memfasilitasi peniadaan perantara: dapat mengubah cara perusahaan yang sudah ada dalam membuat dan menyediakan produk dan layanan; dapat menangani masalah privasi, regulasi, dan tantangan hukum, memberikan peluang untuk pertumbuhan yang inklusif .

Informasi mengenai *FinTech* begitu banyak dijumpai di sosial media, salah satunya di *twitter*. Saat ini sosial media memiliki pengaruh yang besar terhadap persebaran informasi dalam kehidupan. Di Indonesia sendiri *Twitter* adalah salah satu sosial media yang cukup menarik banyak perhatian pengguna internet. *Twitter* merupakan sosial media *microblogging* yang dipergunakan sebagai media komunikasi dan penyampaian informasi yang didalamnya memungkinkan terjadinya pertukaran informasi dalam sosial media tersebut (Hendriyani 2013). Indonesia menempati peringkat 5 pengguna *twitter* di dunia. Posisi Indonesia hanya kalah dari USA, Brazil, Jepang, dan Inggris. Berdasarkan data PT. Bakrie Telecom, Indonesia memiliki 19,5 juta pengguna *twitter* dari total 500 juta pengguna global

(Kominfo 2019).

Chief Executive Officer *twitter*, Dick Costolo, mengatakan bahwa Indonesia menjadi salah satu negara dengan pengguna *Twitter* terbanyak. Karena itulah *twitter* pun akhirnya mendirikan kantor di Jakarta. Berdasarkan laporan *Twitter* di kuartal IV 2014, total pengguna aktifnya mencapai 288 juta per bulan. Menurut Costolo pengguna media sosial di Indonesia memiliki pengetahuan yang baik dengan dunia digital. Disamping itu pengguna *twitter* di Indonesia dinilai sangat atraktif dan bersemangat dalam menuliskan cuitannya. Tidak jarang, obrolan di lini masa menjadi *Trending Topic* atau topik yang paling banyak di bicarakan di seluruh dunia (Cnn Indonesia 2019).

Berdasarkan data-data diatas dapat disimpulkan bahwa *twitter* merupakan sosial media yang akrab dengan masyarakat Indonesia. Tidak hanya perorangan yang memanfaatkannya namun juga institusi-institusi seperti perusahaan-perusahaan untuk mempromosikan produknya. *Twitter* menjadi salah satu situs *microbloging* yang dapat menjadi sumber informasi atau data pendapat dan sentimen masyarakat. Data tersebut dapat digunakan secara efisien untuk pemasaran atau studi sosial (Pak & Paroubek 2010).

Pembahasan mengenai *FinTech* pun banyak di cuitkan oleh warga *twitter* dan dapat menjadi salah satu media untuk melakukan Sosial Media

Analitik (SMA) yang memuat mengenai kecenderungan informasi mengenai suatu topik apakah cenderung positif, negatif, netral atau *irrelevant*. Oleh sebab itu, penelitian ini akan melakukan analisa terhadap opini *tweet* mengenai pandangan masyarakat terhadap *FinTech*. Data yang akan di analisis adalah postingan *tweet* hasil pencarian kata kunci *FinTech*. Pengambilan atau pengumpulan data dilakukan dengan menggunakan API *twitter* yang kemudian akan diklasifikasikan menjadi empat kelas sentimen, yaitu positif, negatif, netral, serta *irrelevant*.

Berikutnya data tersebut akan diolah menggunakan bahasa pemrograman python dengan metode *Naïve Bayes Classifier*. *Naïve Bayes Classifier* bekerja sangat baik dibanding dengan model classifier lainnya. Hal ini dibuktikan pada jurnal Xhemali, Daniela, Chris J. Hinde, and Roger G. Stone. "Naive Bayes vs. decision trees vs. neural networks in the classification of training web pages." (2009), mengatakan bahwa Nave Bayes Classifier memiliki tingkat akurasi yg lebih baik dibanding model classifier lainnya (Daniela dkk 2009).

Berdasarkan pemaparan diatas, dapat dilakukan penelitian mengenai "*Analisis Sentimen Masyarakat Terhadap FinTech Menggunakan Metode Naïve Bayes Classifier*". Dalam penelitian ini akan menghasilkan informasi dan diharapkan dapat membantu pihak-pihak terkait dalam menyikapi

keberadaan *FinTech* saat ini.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah diuraikan diatas maka dapat diambil rumusan yang akan menjadi pembahasan penelitian ini, yaitu :“Bagaimana sentimen masyarakat terhadap perkembangan *FinTech* di lingkup Indonesia pada media *twitter* secara otomatis menggunakan algoritma *Naïve Bayes Classifier*”.

## 1.3. Batasan Masalah

Adapun yang menjadi batasan masalah dari penelitian ini adalah :

1. Data yang digunakan berasal dari postingan di *twitter* dengan kata kunci “*FinTech* ”.
2. Data yang digunakan adalah berbahasa Indonesia dengan jumlah data yang digunakan sebanyak 6050 cuitan
3. Data yang diambil adalah data bulan Januari s/d Februari 2019.
4. Pengambilan data menggunakan API *twitter* dengan search dengan kata kunci “*FinTech* ” atau cuitan yang mengandung kata “*FinTech* ”.
5. Pengolahan data menggunakan bahasa pemrograman Python.

6. Algoritma yang digunakan dalam pengklasifikasian data adalah metode *Naïve Bayes Classifier* dan tidak ada perbandingan dengan metode lain.

#### **1.4. Tujuan Penelitian**

Adapun tujuan yang ingin dicapai oleh penulis dalam penelitian tugas akhir ini adalah :

1. Melihat sejauh mana tingkat akurasi analisis sentimen menggunakan algoritma *Naïve Bayes Classifier*.
2. Dengan metode *Naïve Bayes Classifier* dapat mempercepat proses klasifikasi dan mendapatkan kategori sentimen yang positif, negatif, netral dan *irrelevant*.
3. Mendapatkan hasil analisa dan informasi yang dapat membantu pihak-pihak terkait dalam menyikapi keberadaan *FinTech* saat ini.

#### **1.5. Manfaat Penelitian**

Adapun manfaat dari penelitian ini adalah :

1. Menambah wawasan dalam penerapan *text mining* dalam kehidupan sehari-hari.

2. Dengan metode *Naïve Bayes Classifier* dapat diperoleh hasil analisa yang dapat memberikan informasi khalayak umum terkait Finansial Teknologi.
3. Bagi UIN Sunan Kalijaga, dapat digunakan sebagai tambahan referensi yang terkait dengan peningkatan pengguna sosial media seiring kemajuan teknologi.
4. Bagi penulis, dapat memberikan pengetahuan bahwa apa yang *posting* di sosial media dapat digunakan sebagai data analitik yang dapat menjadi informasi berharga.

## **1.6. Keaslian Penelitian**

Penelitian mengenai sentimen dan klasifikasi pada sosial media *twitter* menggunakan metode *Naïve Bayes Classifier*, sampai saat ini sudah banyak dilakukan oleh peneliti sebelumnya. Namun, berdasarkan referensi dan tinjauan pustaka, penelitian Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta mengenai analisis sentimen menggunakan sosial media *twitter* dengan metode *Naïve Bayes Classifier* mengenai pandangan masyarakat terhadap finansial teknologi atau *FinTech* belum pernah ditemukan oleh peneliti.

## **1.7. Sistematika Penulisan**

Untuk memberikan gambaran dan kerangka yang jelas mengenai pokok bahasan setiap bab dalam penelitian ini, maka diperlukan sistematika penulisan. Berikut ini adalah gambaran sistematika penulisan pada masing-masing bab :

### **BAB I            PENDAHULUAN**

Bab *satu* pendahuluan, berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian, dan sistematika penulisan skripsi.

### **BAB II            TINJAUAN PUSTAKA DAN LANDASAN TEORI**

Bab *dua* tinjauan pustaka dan landasan teori, menjelaskan tentang tinjauan pustaka dan landasan teori yang berhubungan dengan topik yang akan di bahas dalam penelitian ini.

### **BAB III           METODE PENELITIAN**

Bab *tiga* metode penelitian, membahas mengenai metode yang dilakukan dan atau yang digunakan dalam menyelesaikan penelitian ini.

## **BAB IV HASIL DAN PEMBAHASAN**

Bab *empat* hasil dan pembahasan, menjelaskan hasil dari penelitian yang dicapai. Dari proses pengambilan data hingga hasil yang diperoleh dari penelitian ini.

## **BAB V KESIMPULAN DAN SARAN**

Bab *lima* penutup dan kesimpulan, menjelaskan kesimpulan yang dihasilkan serta saran yang akan diberikan berdasarkan hasil yang telah dicapai sehingga dapat digunakan sebagai bahan pertimbangan bagi oihak-pihak yang berkepentingan serta untuk perkembangan penelitian-penelitian selanjutnya.

dengan jumlah 6050 data *tweet* dengan kata kunci *fintech*. Contoh data *tweet* disajikan dalam Tabel 4.1.

**Tabel 4.1** Contoh Data *Tweet*

| No. | <i>Tweet</i>                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1   | Selama ini, fasilitas pembiayaan bagi kalangan mahasiswa sangat terbatas. Hal itu lantaran kredit konvensional menetapkan persyaratan yang sulit dipenuhi kalangan mahasiswa. Kini hadir @cicilcoid sebagai solusinya Yuk pakai untuk kebutuhan kuliahmu #tentangcicil #fintech #CicilDuluAja                                                                                                                                               |
| 2   | Tim Direktorat Tindak Pidana Siber Bareskrim Polri menyatakan cara ilegal yang dilakukan fintech pinjaman online adalah dengan mengakses seluruh data yang ada di HP nasabah #PinjamanUangOnline                                                                                                                                                                                                                                            |
| 3   | - bismillah... SUKSES SENDIRI ITU BIASA.. SUKSES BERJAMAAH BARU LUAR BIASA ....Mau cari usaha atau BISNIS SAMPINGAN yang tidak mengganggu pekerjaan ataupun ibadah dan tiap minggunya <a href="https://t.co/FuFnZ15JBX">https://t.co/FuFnZ15JBX</a> <a href="https://t.co/E8QjEWTy4l">https://t.co/E8QjEWTy4l</a>                                                                                                                           |
| 4   | Tidak semua fintech memberikan uang pinjaman 100%. Banyak diantaranya melakukan potongan biaya administrasi yang cukup besar #PinjamUangOnline.                                                                                                                                                                                                                                                                                             |
| 5   | Penyaluran Pinjaman Fintech Mencapai Rp 22,67 Triliun <a href="https://t.co/lcKDpD2xKh">https://t.co/lcKDpD2xKh</a> #Fintech                                                                                                                                                                                                                                                                                                                |
| 6   | ojk sebanyak 22 fintech berpotensi naik status jadi berizin fintech                                                                                                                                                                                                                                                                                                                                                                         |
| 7   | Hai telkomsel harus bagaimana cara mencegah sms spam fintech2 jahat dan jasa keuangan lain nya sudah bosan juga saya ngeblock dan melaporkan nya sms spam fintech udah riba nyepam pula. <a href="https://t.co/9jxPh031oC">https://t.co/9jxPh031oC</a>                                                                                                                                                                                      |
| 8   | Salah satu penyedia layanan fintech, Ovo merespon baik standarisasi QR code yang akan dilaksanakan oleh Bank Indonesia. <a href="https://t.co/z1x9MiQGxR">https://t.co/z1x9MiQGxR</a> <a href="https://t.co/NBq6XntFSV">https://t.co/NBq6XntFSV</a>                                                                                                                                                                                         |
| 9   | Kurang ajarnya lagi, orang yg tidak ada urusan dng pinjam meminjam juga ditelp dan diteror hanya karena no HP nya ada di phonebook HP si Peminjam yg diakses oleh Pengelola Fintech Online...                                                                                                                                                                                                                                               |
| 10  | Pinjaman online atau pinjaman dalam jaringan akhir-akhir ini ramai dibilakan salah satunya karena penagihan yang dianggap tidak etis sampai kebocoran data yang terjadi. OJK mempunyai penilaian terhadap tingginya kredit macet fintech lending. #ValidNih <a href="https://www.validnews.id/Kredit-Macet-Meningkat-Bisnis-Fintech-Tetap-Memikat-cfa">https://www.validnews.id/Kredit-Macet-Meningkat-Bisnis-Fintech-Tetap-Memikat-cfa</a> |

## 4.2. Pembersihan dan Pelabelan Data

Data yang didapatkan dari *Twitter* merupakan data mentah yang perlu dilakukan pembersihan dengan membuang data yang berasal dari *boot* atau duplikasi, kosong, dan tidak sesuai. Kemudian data dibagi menjadi 2 bagian, yaitu diambil data sebanyak 2500 data latih dan 3550 sebagai data uji untuk klasifikasi otomatis.

Proses pelabelan dilakukan secara manual. Dengan proses pelabelan berdasarkan Data latih kemudian diberi label sebagai model klasifikasi. Sebanyak 2500 data diberi label secara manual. Setiap *tweet* akan dikelompokkan ke dalam empat kelas yaitu negatif (-1), netral (0), positif (1), *irrelevant* (2).

Pemberian label negatif didasari oleh opini atau tindakan yang negatif atau merugikan dan meresahkan. Pemberian label netral didasari oleh opini atau tindakan yang tidak menunjukkan ke dalam negatif dan positif namun masih dalam konteks *Fintech*. Pemberian label positif didasari oleh opini atau tindakan yang positif, manfaat, menguntungkan. Pemberian label *irrelevant* didasari oleh opini atau *tweet* tersebut sama sekali tidak ada hubungan dengan *Fintech*.

Contoh data yang telah diberi label disajikan dalam Tabel 4.2.

**Tabel 4.2** Contoh Pelabelan Data

| No. | Data <i>tweet</i>                                                                                                                                                                                                                                                                                                   | Label |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| 1   | Selama ini, fasilitas pembiayaan bagi kalangan mahasiswa sangat terbatas. Hal itu lantaran kredit konvensional menetapkan persyaratan yang sulit dipenuhi kalangan mahasiswa. Kini hadir @cicilcoid sebagai solusinya Yuk pakai untuk kebutuhan kuliahmu #tentangcicil #fintech #CicilDuluAja                       | 1     |
| 2   | Tim Direktorat Tindak Pidana Siber Bareskrim Polri menyatakan cara ilegal yang dilakukan fintech pinjaman online adalah dengan mengakses seluruh data yang ada di HP nasabah #PinjamanUangOnline                                                                                                                    | -1    |
| 3   | - bismillah... SUKSES SENDIRI ITU BIASA.. SUKSES BERJAMAAH BARU LUAR BIASA ....Mau cari usaha atau BISNIS SAMPINGAN yang tidak mengganggu pekerjaan ataupun ibadah dan tiap minggunya <a href="https://t.co/FuFnZ15JBX">https://t.co/FuFnZ15JBX</a> <a href="https://t.co/E8QjEWTy4l">https://t.co/E8QjEWTy4l</a> . | 2     |
| 4   | Tidak semua fintech memberikan uang pinjaman 100%. Banyak diantaranya melakukan potongan biaya administrasi yang cukup besar #PinjamUangOnline.                                                                                                                                                                     | -1    |
| 5   | Penyaluran Pinjaman Fintech Mencapai Rp 22,67 Triliun <a href="https://t.co/lcKDpD2xKh">https://t.co/lcKDpD2xKh</a> #Fintech                                                                                                                                                                                        | 0     |
| 6   | ojk sebanyak 22 fintech berpotensi naik status jadi berizin fintech                                                                                                                                                                                                                                                 | 1     |
| 7   | Hai telkomsel harus bagaimana cara mencegah sms spam fintech2 jahat dan jasa keuangan lain nya sudah bosan juga saya ngeblock dan melaporkan nya sms spam fintech udah riba nyepam pula <a href="https://t.co/9jxPh031oC">https://t.co/9jxPh031oC</a>                                                               | -1    |
| 8   | Salah satu penyedia layanan fintech, Ovo merespon baik standarisasi QR code yang akan dilaksanakan oleh Bank Indonesia. <a href="https://t.co/z1x9MiQGxR">https://t.co/z1x9MiQGxR</a> <a href="https://t.co/NBq6XntFSV">https://t.co/NBq6XntFSV</a>                                                                 | 0     |
| 9   | Kurang ajarnya lagi, orang yg tidak ada urusan dng pinjam meminjam juga ditelp dan diteror hanya karena no HP nya ada di phonebook HP si Peminjam yg diakses oleh Pengelola Fintech Online...                                                                                                                       | -1    |

**Tabel 4.3** *Lanjutan Contoh Pelabelan Data*

| No. | Data <i>tweet</i>                                                                                                                                                                                                                                                                                                                                                                                                                             | Label |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| 10  | Pinjaman online atau pinjaman dalam jaringan akhir-akhir ini ramai dibicarakan salah satunya karena penagihan yang dianggap tidak etis sampai kebocoran data yang terjadi. OJK mempunyai penilaian terhadap tingginya kredit macet fintech lending. #ValidNih <a href="https://www.validnews.id/Kredit-Macet-Meningkat-Bisnis-Fintech-Tetap-Memikat-cfa">https://www.validnews.id/Kredit-Macet-Meningkat-Bisnis-Fintech-Tetap-Memikat-cfa</a> | 0     |

### 4.3. *Preprocessing*

Pada tahap *preprocessing* dilakukan untuk melakukan pembersihan terhadap data *tweet*. Adapun langkah-langkah dalam pembersihan data adalah sebagai berikut:

#### 4.3.1. *Cleansing*

Data *tweet* umumnya mengandung *hashtag*, *url*, *username*, tanda baca ataupun karakter-karakter non huruf lainnya. Pada tahap *cleansing* ini, dilakukan pembersihan terhadap komponen-komponen tersebut. Pada tabel 4.4 disajikan perbandingan data sebelum dan sesudah proses *cleansing* seperti berikut:

**Tabel 4.4** Contoh Hasil *Cleansing*

| Data | Sebelum                                                                                                                                                                                                                                                                                                                                                                                                                                          | Sesudah                                                                                                                                                                                                                                                                        |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DT1  | Selama ini, fasilitas pembiayaan bagi kalangan mahasiswa sangat terbatas. Hal itu lantaran kredit konvensional menetapkan persyaratan yang sulit dipenuhi kalangan mahasiswa. Kini hadir @cicilcoid sebagai solusinya Yuk pakai untuk kebutuhan kuliahmu #tentangcicil #CicilDuluAja                                                                                                                                                             | Selama ini fasilitas pembiayaan bagi kalangan mahasiswa sangat terbatas Hal itu lantaran kredit konvensional menetapkan persyaratan yang sulit dipenuhi kalangan mahasiswa Kini hadir cicilcoid sebagai solusinya Yuk pakai untuk kebutuhan kuliahmu tentangcicil CicilDuluAja |
| DT2  | Tim Direktorat Tindak Pidana Siber Bareskrim Polri menyatakan cara ilegal yang dilakukan fintech pinjaman online adalah dengan mengakses seluruh data yang ada di HP nasabah #PinjamanUangOnline                                                                                                                                                                                                                                                 | Tim Direktorat Tindak Pidana Siber Bareskrim Polri menyatakan cara ilegal yang dilakukan fintech pinjaman online adalah dengan mengakses seluruh data yang ada di HP nasabah PinjamanUangOnline                                                                                |
| DT3  | Salah satu penyedia layanan fintech, Ovo merespon baik standarisasi QR code yang akan dilaksanakan oleh Bank Indonesia. <a href="https://t.co/z1x9MiQGxR">https://t.co/z1x9MiQGxR</a><br><a href="https://t.co/NBq6XntFSV">https://t.co/NBq6XntFSV</a>                                                                                                                                                                                           | Salah satu penyedia layanan fintech Ovo merespon baik standarisasi QR code yang akan dilaksanakan oleh Bank Indonesia                                                                                                                                                          |
| DT4  | Pinjaman online atau pinjaman dalam jaringan akhir-akhir ini ramai dibicarakan salah satunya karena penagihan yang dianggap tidak etis sampai kebocoran data yang terjadi. OJK mempunyai penilaian terhadap tingginya kredit macet fintech lending. #ValidNih<br><a href="https://www.validnews.id/Kredit-Macet-Meningkat-Bisnis-Fintech-Tetap-Memikat-cfa">https://www.validnews.id/Kredit-Macet-Meningkat-Bisnis-Fintech-Tetap-Memikat-cfa</a> | Pinjaman online atau pinjaman dalam jaringan akhir-akhir ini ramai dibicarakan salah satunya karena penagihan yang dianggap tidak etis sampai kebocoran data yang terjadi OJK mempunyai penilaian terhadap tingginya kredit macet fintech lending ValidNih                     |

4.3.2. Case Folding

Tidak semua dokumen teks konsisten dalam penggunaan huruf kapital. oleh karena itu, peran *case folding* dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen. Pada tahap ini semua huruf pada *tweet* diubah menjadi huruf kecil (*lowercase*). Adanya huruf kapital akan mempengaruhi perhitungan *TF* dan *TF-IDF*. Maka dari itu dilakukan proses *Case Folding* untuk menyamaratakan huruf menjadi huruf kecil. Pada tabel 4.5 disajikan perbandingan data sebelum dan sesudah proses *case folding* seperti berikut:

Tabel 4.5 Contoh Hasil Case Folding

| Data | Sebelum                                                                                                                                                                                                                                                                                                              | Sesudah                                                                                                                                                                                                                                                                                                              |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DT1  | Selama ini fasilitas pembiayaan ba-<br>gi kalangan mahasiswa sangat ter-<br>batas Hal itu lantaran kredit kon-<br>vensional menetapkan persyaratan<br>yang sulit dipenuhi kalangan ma-<br>hasiswa Kini hadir cicilcoid seba-<br>gai solusinya Yuk pakai untuk ke-<br>butuhan kuliahmu tentangcicil Ci-<br>cilDuluAja | selama ini fasilitas pembiayaan ba-<br>gi kalangan mahasiswa sangat ter-<br>batas hal itu lantaran kredit kon-<br>vensional menetapkan persyaratan<br>yang sulit dipenuhi kalangan ma-<br>hasiswa Kini hadir cicilcoid seba-<br>gai solusinya yuk pakai untuk ke-<br>butuhan kuliahmu tentangcicil cici-<br>lduluaja |

**Tabel 4.6** Lanjutan Contoh Hasil *Case Folding*

|     |                                                                                                                                                                                                                                                               |                                                                                                                                                                                                                                                           |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DT2 | Tim Direktorat Tindak Pidana Siber Bareskrim Polri menyatakan cara ilegal yang dilakukan fintech pinjaman online adalah dengan mengakses seluruh data yang ada di HP nasabah PinjamanUangOnline                                                               | tim direktorat tindak pidana siber bareskrim polri menyatakan cara ilegal yang dilakukan fintech pinjaman online adalah dengan mengakses seluruh data yang ada di hp nasabah pinjamanuangonline                                                           |
| DT3 | Salah satu penyedia layanan fintech Ovo merespon baik standarisasi QR code yang akan dilaksanakan oleh Bank Indonesia                                                                                                                                         | salah satu penyedia layanan fintech ovo merespon baik standarisasi qr code yang akan dilaksanakan oleh bank indonesia                                                                                                                                     |
| DT4 | Pinjaman online atau pinjaman dalam jaringan akhir-akhir ini ramai dibicarakan salah satunya karena penagihan yang dianggap tidak etis sampai kebocoran data yang terjadi di OJK mempunyai penilaian terhadap tingginya kredit macet fintech lending ValidNih | pinjaman online atau pinjaman dalam jaringan akhirakhir ini ramai dibicarakan salah satunya karena penagihan yang dianggap tidak etis sampai kebocoran data yang terjadi ojk mempunyai penilaian terhadap tingginya kredit macet fintech lending validnih |

### 4.3.3. Tokenisasi

Tokenisasi merupakan proses memecah atau memisahkan kata atau token. Tahap tokenisasi atau pemisahan token-token ini biasanya mengacu

pada karakter spasi atau tanda baca. Pada tabel 4.7 disajikan perbandingan data sebelum dan sesudah proses Tokenisasi seperti berikut :

**Tabel 4.7** Contoh Hasil Tokenisasi

| Data | Tokenisasi                                                                                                                                                                                                                                                                                                                                                                        |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DT1  | ['selama', 'ini', 'fasilitas', 'pembiayaan', 'bagi', 'kalangan', 'mahasiswa', 'sangat', 'terbatas', 'hal', 'itu', 'lantaran', 'kredit', 'konvensional', 'menetapkan', 'persyaratan', 'yang', 'sulit', 'dipenuhi', 'kalangan', 'mahasiswa', 'kini', 'hadir', 'cicilcoid', 'sebagai', 'solusinya', 'yuk', 'pakai', 'untu', 'kebutuhan', 'kuliahmu,' 'tentangcicil', 'cicilduluaja'] |
| DT2  | ['tim', 'direktorat', 'tindak', 'pidana', 'siber', 'bareskrim', 'polri', 'menyatakan', 'cara', 'ilegal', 'yang', 'dilakukan', 'fintech', 'pinjaman', 'online', 'adalah', 'dengan', 'mengakses', 'seluruh', 'data', 'yang', 'ada', 'di', 'hp', 'nasabah', 'pinjamanuangonline']                                                                                                    |
| DT3  | ['salah', 'satu', 'penyedia', 'layanan', 'fintech', 'ovo', 'merespon', 'baik', 'standarisasi', 'qr', 'code', 'yang', 'akan', 'dilaksanakan', 'oleh', 'bank', 'indonesia']                                                                                                                                                                                                         |

#### 4.3.4. Slangword

Data *tweet* yang di posting terkadang mengandung kata-kata yang tidak baku dan tidak sesuai dengan standart Bahasa Indonesia. Oleh karena itu perlu adanya proses *Convert Slangword*. Pada prosesnya peneliti membuat kamus kata tidak baku disertai bentuk baku dalam bentuk *file .txt* sebagai

bahan *convert slangword*. Pada tabel 4.8 disajikan perbandingan data sebelum dan sesudah proses *Slangword* seperti berikut :

**Tabel 4.8** Contoh Hasil *Stopword*

| Dokumen | Sebelum                                                        | Sesudah                                                           |
|---------|----------------------------------------------------------------|-------------------------------------------------------------------|
| DT1     | jgn gunakan fintech itu riba                                   | jangan gunakan itu fintech riba                                   |
| DT2     | asn meminta untk waspada dg investasi fintech                  | asn meminta untuk waspada dengan investasi fintech                |
| DT3     | sdh tau kalau cicil membantu mahasiswa cicil tnpa kartu kredit | sudah tau kalau cicil membantu mahasiswa cicil tanpa kartu kredit |
| DT4     | bgini rasanya menginap di hotel dg pelayanan robot             | begini rasanya menginap di hotel dengan pelayanan robot           |

#### 4.3.5. *StopWord Removal*

Proses selanjutnya adalah proses *stopword*, proses ini diperlukan dalam *preprocessing* karena mampu meningkatkan peforma klasifikasi karena dapat meningkatkan nilai akurasi. Pada tabel 4.9 disajikan perbandingan data sebelum dan sesudah proses *Stopword* seperti berikut :

**Tabel 4.9** Contoh Hasil *Stopword*

| Dokumen | Sebelum                                            | Sesudah                               |
|---------|----------------------------------------------------|---------------------------------------|
| DT1     | jangan gunakan fintech itu riba                    | jangan gunakan fintech riba           |
| DT2     | asn meminta untuk waspada dengan investasi fintech | asn meminta waspada investasi fintech |

**Tabel 4.10** Lanjutan Contoh Hasil *Stopword*

| Dokumen | Sebelum                                                           | Sesudah                                                     |
|---------|-------------------------------------------------------------------|-------------------------------------------------------------|
| DT3     | sudah tau kalau cicil membantu mahasiswa cicil tanpa kartu kredit | sudah tau kalau cicil membantu mahasiswa tanpa kartu kredit |
| DT4     | begini rasanya menginap di hotel dengan pelayanan robot           | begini rasanya menginap hotel pelayanan robot               |

#### 4.3.6. *Stemming*

*Stemming* adalah proses mengubah bentuk kata yang berimbuhan ke dalam bentuk kata dasar. Proses *stemming* dilakukan untuk menghilangkan imbuhan pada suatu kata dengan memanfaatkan library *stemming* satriwi Bahasa Indonesia di *python*. Contohnya, yaitu kata “memanfaatkan” menjadi kata “manfaat”. Pada tabel 4.11 disajikan perbandingan data sebelum dan sesudah proses *stemming* seperti berikut:

**Tabel 4.11** Contoh Hasil *Steming*

| Dokumen | Sebelum                                            | Sesudah                             |
|---------|----------------------------------------------------|-------------------------------------|
| DT1     | jangan gunakan fintech itu riba                    | jangan guna fintech riba            |
| DT2     | asn meminta untuk waspada dengan investasi fintech | asn minta waspada investasi fintech |

**Tabel 4.12** Lanjutan Contoh Hasil *Stemming*

| Dokumen | Sebelum                                                           | Sesudah                                                  |
|---------|-------------------------------------------------------------------|----------------------------------------------------------|
| DT3     | sudah tau kalau cicil membantu mahasiswa cicil tanpa kartu kredit | sudah tau kalau cicil bantu mahasiswa tanpa kartu kredit |
| DT4     | begini rasanya menginap di hotel dengan pelayanan robot           | gini rasa inap hotel layan robot                         |

#### 4.4. Analisa dan Evaluasi

##### 4.4.1. Analisa

Data latih sebanyak 2500 data yang telah di *preprocessing* dan diberi label secara manual akan dilakukan ekstraksi fitur. Ekstraksi fitur digunakan untuk mencari *term frequency (tf)* dan *term frequency-invers Document frequency*. Ekstraksi fitur ini kemudian digunakan dalam perhitungan probabilitas menggunakan *Naïve Bayes Classifier* sebagai metode klasifikasi data yang telah dilakukan *preprocessing* dan pelabelan manual. Dalam pelabelan ini terbagi menjadi empat label, yaitu positif(1), negatif(-1), neutral(0), *irrelevant*(2). Contoh label hasil preprocessing disajikan dalam tabel 4.12.

**Tabel 4.13** Contoh Hasil *Preprocessing* Data *Tweet* beserta label

| Dokumen | Data <i>Tweet</i>                                        | Label |
|---------|----------------------------------------------------------|-------|
| DT1     | jangan guna fintech riba                                 | -1    |
| DT2     | asn minta waspada investasi fintech                      | 0     |
| DT3     | sudah tau kalau cicil bantu mahasiswa tanpa kartu kredit | 1     |
| DT4     | gini rasa inap hotel layan robot                         | 2     |

### ***Term Frequency (TF)***

*Term Frequency*(tf) merupakan suatu metode yang digunakan untuk mencari bobot suatu kata dalam dokumen *text* di setiap kategori dan mencari kata kunci yang hampir mirip dengan kategori yang tersedia

Berikut ini disajikan tabel *Term Frequency* dari data *tweet* ditunjukkan pada tabbel 4.14.

**Tabel 4.14** *Term Frequency*

| No. | <i>Term</i> | Dokumen |     |     |     |
|-----|-------------|---------|-----|-----|-----|
|     |             | DT1     | DT2 | DT3 | DT4 |
| 1.  | asn         | 0       | 1   | 0   | 0   |
| 2.  | bantu       | 0       | 0   | 1   | 0   |
| 3.  | cicil       | 0       | 0   | 1   | 0   |
| 4.  | fintech     | 1       | 1   | 0   | 0   |

**Tabel 4.15** *Lanjutan Term Frequency*

| No.    | Term      | Dokumen |     |     |     |
|--------|-----------|---------|-----|-----|-----|
|        |           | DT1     | DT2 | DT3 | DT4 |
| 5.     | gini      | 0       | 0   | 0   | 1   |
| 6.     | guna      | 1       | 0   | 0   | 0   |
| 7.     | hotel     | 0       | 0   | 0   | 1   |
| 8.     | inap      | 0       | 0   | 0   | 1   |
| 9.     | investasi | 0       | 1   | 0   | 0   |
| 10.    | jangan    | 1       | 0   | 0   | 0   |
| 11.    | kalau     | 0       | 0   | 1   | 0   |
| 12.    | kartu     | 0       | 0   | 1   | 0   |
| 13.    | kredit    | 0       | 0   | 1   | 0   |
| 14.    | layan     | 0       | 0   | 0   | 1   |
| 15.    | mahasiswa | 0       | 0   | 1   | 0   |
| 16.    | minta     | 0       | 1   | 0   | 0   |
| 17.    | rasa      | 0       | 0   | 0   | 1   |
| 18.    | riba      | 1       | 0   | 0   | 0   |
| 19.    | robot     | 0       | 0   | 0   | 1   |
| 20.    | sudah     | 0       | 0   | 1   | 0   |
| 21.    | tanpa     | 0       | 0   | 1   | 0   |
| 22.    | tau       | 0       | 0   | 1   | 0   |
| 23.    | wasapada  | 0       | 1   | 0   | 0   |
| Jumlah |           | 4       | 5   | 9   | 6   |

Selanjutnya dilakukan proses klasifikasi *Naïve Bayes Classifier* ber-

dasarkan data latih sebanyak 2500 data dan pembobotan TF dengan cara menghitung nilai peluang pada masing-masing kelas berdasarkan contoh dokumen data *tweet* diatas dengan persamaan 4.1:

$$P(i) = \frac{N(i)}{N} \quad (4.1)$$

Mencari nilai peluang pada setiap *term* pada label sentimen tertentu dengan persamaan 4.2:

$$P(t|i) = \frac{Count(t,i)+1}{count(i)+|V|} \quad (4.2)$$

Mencari nilai klasifikasi dengan persamaan 4.3

$$P(i|A) = P(i) * P(t|i) * ... * P(t|i) \quad (4.3)$$

Keterangan :

$P(i)$  = peluang sentimen i

$N(i)$  = jumlah dokumen yang masuk dalam sentimen i

$N$  = jumlah keseluruhan dokumen

$t$  = *term*

$|V|$  = jumlah *term*

$P(t|i)$  = varibel kelas

Berdasarkan contoh pada tabel 4.9, dengan menggunakan persamaan 4.1 maka :

$$P(negatif) = \frac{N(negatif)}{N} = \frac{1}{4} = 0,25$$

$$P(netral) = \frac{N(netral)}{N} = \frac{1}{4} = 0,25$$

$$P(positif) = \frac{N(positif)}{N} = \frac{1}{4} = 0,25$$

$$P(irrelevant) = \frac{N(irrelevant)}{N} = \frac{1}{4} = 0,25$$

Langkah berikutnya adalah dengan mencari nilai peluang pada setiap *term* pada dokumen label sentimen tertentu. Misalkan menghitung peluang *term* 'jangan' pada label sentiment negatif menggunakan persamaan 4.2 adalah sebagai berikut :

$$P(jangan|negatif) = \frac{Count(jangan,negatif)+1}{count(negatif)+|V|} = \frac{1+1}{4+23} = \frac{2}{27} = 0,07407407$$

Pada tabel 4.16 berikut merupakan perhitungan setiap *term* pada label sentimen negatif.

**Tabel 4.16** Peluang *Term* Sentimen Negatif

| No. | <i>Term</i> | Dokumen<br>DT1 | DF | Peluang <i>Term</i> Negatif |
|-----|-------------|----------------|----|-----------------------------|
| 1.  | asn         | 0              | 0  | 0,03703704                  |
| 2.  | bantu       | 0              | 0  | 0,03703704                  |

**Tabel 4.17** Lanjutan Peluang *Term* Sentimen Negatif

| No.    | <i>Term</i> | Dokumen<br>DT1 | DF | Peluang <i>Term</i> Negatif |
|--------|-------------|----------------|----|-----------------------------|
| 3.     | cicil       | 0              | 0  | 0,03703704                  |
| 4.     | fintech     | 1              | 1  | 0,07407407                  |
| 5.     | gini        | 0              | 0  | 0,03703704                  |
| 6.     | guna        | 1              | 1  | 0,07407407                  |
| 7.     | hotel       | 0              | 0  | 0,03703704                  |
| 8.     | inap        | 0              | 0  | 0,03703704                  |
| 9.     | investasi   | 0              | 0  | 0,03703704                  |
| 10.    | jangan      | 1              | 1  | 0,07407407                  |
| 11.    | kalau       | 0              | 0  | 0,03703704                  |
| 12.    | kartu       | 0              | 0  | 0,03703704                  |
| 13.    | kredit      | 0              | 0  | 0,03703704                  |
| 14.    | layan       | 0              | 0  | 0,03703704                  |
| 15.    | mahasiswa   | 0              | 0  | 0,03703704                  |
| 16.    | minta       | 0              | 0  | 0,03703704                  |
| 17.    | rasa        | 0              | 0  | 0,03703704                  |
| 18.    | riba        | 1              | 1  | 0,07407407                  |
| 19.    | robot       | 0              | 0  | 0,03703704                  |
| 20.    | sudah       | 0              | 0  | 0,03703704                  |
| 21.    | tanpa       | 0              | 0  | 0,03703704                  |
| 22.    | tau         | 0              | 0  | 0,03703704                  |
| 23.    | wasapada    | 0              | 0  | 0,03703704                  |
| Jumlah |             |                | 4  |                             |

Untuk *term* label sentimen netral juga dilakukan perhitungan yang sama seperti label sentimen negatif dengan contoh term investasi sebagai berikut :

$$P(\text{investasi}|\text{netral}) = \frac{\text{Count}(\text{investasi}|\text{netral})+1}{\text{count}(\text{netral})+|V|} = \frac{1+1}{5+23} = \frac{2}{28} = 0,07142857$$

Pada tabel 4.18 berikut merupakan perhitungan setiap *term* pada label sentimen netral.

**Tabel 4.18** Peluang *Trem* Sentimen Netral

| No. | <i>Term</i> | Dokumen<br>DT2 | DF | Peluang <i>Term</i> Netral |
|-----|-------------|----------------|----|----------------------------|
| 1.  | asn         | 1              | 1  | 0,07142857                 |
| 2.  | bantu       | 0              | 0  | 0,03571429                 |
| 3.  | cicil       | 0              | 0  | 0,03571429                 |
| 4.  | fintech     | 1              | 1  | 0,07142857                 |
| 5.  | gini        | 0              | 0  | 0,03571429                 |
| 6.  | guna        | 0              | 0  | 0,03571429                 |
| 7.  | hotel       | 0              | 0  | 0,03571429                 |
| 8.  | inap        | 0              | 0  | 0,03571429                 |
| 9.  | investasi   | 1              | 1  | 0,07142857                 |
| 10. | jangan      | 0              | 0  | 0,03571429                 |
| 11. | kalau       | 0              | 0  | 0,03571429                 |
| 12. | kartu       | 0              | 0  | 0,03571429                 |
| 13. | kredit      | 0              | 0  | 0,03571429                 |

**Tabel 4.19** Lanjutan Peluang *Trem* Sentimen Netral

| No.    | <i>Term</i> | Dokumen<br>DT2 | DF | Peluang <i>Term</i> Netral |
|--------|-------------|----------------|----|----------------------------|
| 14.    | layan       | 0              | 0  | 0,03571429                 |
| 15.    | mahasiswa   | 0              | 0  | 0,03571429                 |
| 16.    | minta       | 1              | 1  | 0,03571429                 |
| 17.    | rasa        | 0              | 0  | 0,03571429                 |
| 18.    | riba        | 0              | 0  | 0,03571429                 |
| 19.    | robot       | 0              | 0  | 0,03571429                 |
| 20.    | sudah       | 0              | 0  | 0,03571429                 |
| 21.    | tanpa       | 0              | 0  | 0,03571429                 |
| 22.    | tau         | 0              | 0  | 0,03571429                 |
| 23.    | wasapada    | 1              | 1  | 0,07142857                 |
| Jumlah |             |                | 5  |                            |

Untuk *term* label sentimen positif juga dilakukan perhitungan yang sama seperti label sentimen negatif dan netral dengan contoh term *cicil* sebagai berikut :

$$P(cicil|positif) = \frac{Count(cicil,positif)+1}{count(positif)+|V|} = \frac{1+1}{9+23} = \frac{2}{32} = 0,0625$$

Pada tabel 4.20 berikut merupakan perhitungan setiap *term* pada label sentimen positif.

**Tabel 4.20** Peluang *Trem* Sentimen Positif

| No. | <i>Term</i> | Dokumen<br>DT3 | DF | Peluang <i>Term</i> Positif |
|-----|-------------|----------------|----|-----------------------------|
| 1.  | asn         | 0              | 0  | 0,03125                     |
| 2.  | bantu       | 1              | 1  | 0,0625                      |
| 3.  | cicil       | 1              | 1  | 0,0625                      |
| 4.  | fintech     | 0              | 0  | 0,03125                     |
| 5.  | gini        | 0              | 0  | 0,03125                     |
| 6.  | guna        | 0              | 0  | 0,03125                     |
| 7.  | hotel       | 0              | 0  | 0,03125                     |
| 8.  | inap        | 0              | 0  | 0,03125                     |
| 9.  | investasi   | 1              | 1  | 0,0625                      |
| 10. | jangan      | 0              | 0  | 0,03125                     |
| 11. | kalau       | 1              | 1  | 0,0625                      |
| 12. | kartu       | 1              | 1  | 0,0625                      |
| 13. | kredit      | 1              | 1  | 0,0625                      |
| 14. | layan       | 0              | 0  | 0,03125                     |
| 15. | mahasiswa   | 1              | 1  | 0,0625                      |
| 16. | minta       | 0              | 0  | 0,03125                     |
| 17. | rasa        | 0              | 0  | 0,03125                     |
| 18. | riba        | 0              | 0  | 0,03125                     |
| 19. | robot       | 0              | 0  | 0,03125                     |
| 20. | sudah       | 1              | 1  | 0,0625                      |
| 21. | tanpa       | 1              | 1  | 0,0625                      |
| 22. | tau         | 1              | 1  | 0,0625                      |

**Tabel 4.21** Lanjutan Peluang *Trem* Sentimen Positif

| No.    | <i>Term</i> | Dokumen<br>DT3 | DF | Peluang <i>Term</i> Positif |
|--------|-------------|----------------|----|-----------------------------|
| 23.    | wasapada    | 0              | 0  | 0,03125                     |
| Jumlah |             |                | 9  |                             |

Untuk *term* label sentimen *irrelevant* juga dilakukan perhitungan yang sama seperti label sentimen positif dengan contoh term rasa sebagai berikut :

$$P(rasa|irrelevant) = \frac{Count(rasa|irrelevant)+1}{count(irrelevant)+|V|} = \frac{1+1}{6+23} = \frac{2}{29} = 0,06896552$$

Pada tabel 4.22 berikut merupakan perhitungan setiap *term* pada label sentimen *Irrelevant*.

**Tabel 4.22** Peluang *Trem* Sentimen *Irrelevant*

| No. | <i>Term</i> | Dokumen<br>DT4 | DF | Peluang <i>Term irrelevant</i> |
|-----|-------------|----------------|----|--------------------------------|
| 1.  | asn         | 0              | 0  | 0,03448276                     |
| 2.  | bantu       | 0              | 0  | 0,03448276                     |
| 3.  | cicil       | 0              | 0  | 0,03448276                     |
| 4.  | fintech     | 0              | 0  | 034482759                      |
| 5.  | gini        | 1              | 1  | 0,06896552                     |

**Tabel 4.23** Lanjutan Peluang *Trem* Sentimen *Irrelevant*

| No.    | <i>Term</i> | Dokumen<br>DT4 | DF | Peluang <i>Term irrelevant</i> |
|--------|-------------|----------------|----|--------------------------------|
| 6.     | guna        | 0              | 0  | 0,03448276                     |
| 7.     | hotel       | 1              | 1  | 0,06896552                     |
| 8.     | inap        | 1              | 1  | 0,06896552                     |
| 9.     | investasi   | 0              | 0  | 0,03448276                     |
| 10.    | jangan      | 0              | 0  | 0,03448276                     |
| 11.    | kalau       | 0              | 0  | 0,03448276                     |
| 12.    | kartu       | 0              | 0  | 0,03448276                     |
| 13.    | kredit      | 0              | 0  | 0,03448276                     |
| 14.    | layan       | 1              | 1  | 0,06896552                     |
| 15.    | mahasiswa   | 0              | 0  | 0,03448276                     |
| 16.    | minta       | 0              | 0  | 0,03448276                     |
| 17.    | rasa        | 1              | 1  | 0,06896552                     |
| 18.    | riba        | 0              | 0  | 0,03448276                     |
| 19.    | robot       | 1              | 1  | 0,06896552                     |
| 20.    | sudah       | 0              | 0  | 0,03448276                     |
| 21.    | tanpa       | 0              | 0  | 0,03448276                     |
| 22.    | tau         | 0              | 0  | 0,03448276                     |
| 23.    | wasapada    | 0              | 0  | 0,03448276                     |
| Jumlah |             |                | 6  |                                |

Setelah dilakukan perhitungan *term frequency* pada masing-masing la-

bel sentimen, kemudian dilakukan proses klasifikasi dengan contoh dokumen baru. Proses klasifikasi ini dengan cara mengalikan semua variabel kelas pada masing-masing label sentimen. Seperti pada persamaan 4.3.

$A$  = cicil beri fasilitas pinjam mahasiswa tanpa kartu kredit

### 1. Label Sentimen Negatif

$$\begin{aligned}
 P(Negatif|A) &= P(Negatif) * P(cicil|Negatif) * P(beri|Negatif) * \\
 &P(fasilitas|Negatif) * P(pinjam|Negatif) * P(mahasiswa|Negatif) * \\
 &P(tanpa|Negatif) * P(kartu|Negatif) * P(kredit|Negatif) \\
 P(Negatif|A) &= 0,25 * 0,0370 * 0,0370 * 0,0370 * 0,0370 * 0,0370 * \\
 &0,0370 * 0,0370 * 0,0370 \\
 P(Negatif|A) &= 0,000000000000008781199 \text{ atau } 8,781199e - 13
 \end{aligned}$$

### 2. Label Sentimen Netral

$$\begin{aligned}
 P(Netral|A) &= P(Netral) * P(cicil|Netral) * P(beri|Netral) * \\
 &P(fasilitas|Netral) * P(pinjam|Netral) * P(mahasiswa|Netral) * \\
 &P(tanpa|Netral) * P(kartu|Netral) * P(kredit|Netral) \\
 P(Netral|A) &= 0,25 * 0,03571 * 0,03571 * 0,03571 * 0,03571 * \\
 &0,03571 * 0,03571 * 0,03571 * 0,03571 \\
 P(Netral|A) &= 0,000000000000006610871 \text{ atau } 6,610871e - 13
 \end{aligned}$$

### 3. Label Sentimen Positif

$$P(Positif|A) = P(Positif) * P(cicil|Positif) * P(beri|Positif) *$$

$$P(fasilitas|Positif)*P(pinjam|Positif)*P(mahasiswa|Positif)*$$

$$P(tanpa|Positif) * P(kartu|Positif) * P(kredit|Positif)$$

$$P(Positif|A) = 0,25 * 0,0625 * 0,03125 * 0,03125 * 0,03125 *$$

$$0,0625 * 0,03125 * 0,0625 * 0,0625$$

$$P(Positif|A) = 0,0000000000003637979 \text{ atau } 3,637979e - 12$$

#### 4. Label Sentimen Irrelevant

$$P(Irrelevant|A) = P(Irrelevant)*P(cicil|Irrelevant)*P(beri|Irrelevant)*$$

$$P(fasilitas|Irrelevant)*P(pinjam|Irrelevant)*P(mahasiswa|Irrelevant)*$$

$$P(tanpa|Irrelevant) * P(kartu|Irrelevant) * P(kredit|Irrelevant)$$

$$P(Irrelevant|A) = 0,25 * 0,03448 * 0,03448 * 0,03448 * 0,03448 *$$

$$0,03448 * 0,03448 * 0,03448 * 0,03448$$

$$P(Irrelevant|A) = 0,00000000000049943 \text{ atau } 4,9943e - 13$$

Dari hasil perhitungan dari keempat label kemudian dilakukan perbandingan bobot dari masing-masing sentimen. Besaran nilai perhitungan menentukan label sentimennya. Pada dokumen A setelah dilakukan perhitungan nilai yang didapatkan  $8,781199e - 13$  (negatif),  $6,610871e - 13$  (netral),  $3,637979e - 12$  (positif), dan  $4,9943e - 13$  (irrelevant). Dari hasil nilai yang didapatkan, nilai tertinggi terdapat pada label Positif. Maka dari itu dokumen A termasuk dalam label Positif

Selanjutnya dilakukan proses *training* menggunakan klasifikasi *Naïve*

*Bayes Classifier* dengan pembobotan *term frequency* terhadap data latih yang berjumlah 2500 data dengan komposisi label negatif, netral, positif, dan *irrelevant*. Dalam proses *training* ini akan menghasilkan model klasifikasi untuk memberi label pada data uji.

### ***Term Frequency-Invers Document Frequency (TF-IDF)***

Semakin banyak suatu kata dalam suatu kelas yang terdapat dalam suatu dokumen, maka nilai  $tf(\text{term frequency})$  akan menjadi lebih tinggi pada kelas tersebut. Namun, apa bila suatu kata muncul pada semua dokumen, maka tidak bisa dipastikan bahwa suatu kata tersebut memiliki makna khusus yang relevan. Untuk mengatasi masalah tersebut, salah satu cara yang umum digunakan adalah dengan menggunakan  $tf-idf$  *Term Frequency-Invers Document Frequency* dengan menggunakan metode  $tf$  (*term frequency*) dan  $idf$  (*invers document frequency*) Saputro (Michael & Tyas).

Rumus TF-IDF bisa dilihat sebagai berikut :

$$idf(t) = \log \frac{N_d}{df_t} + 1 \quad (4.4)$$

$$tf-idf = tf_{t,d} * idf_t \quad (4.5)$$

Keterangan :

$N_d$  = Jumlah dokumen

$df_t$  = Jumlah dokumen yang mengandung *term* tersebut

$tf_{t,d}$  = Frekuensi kemunculan *term* dalam data tersebut

Selanjutnya dilakukan proses *training* menggunakan klasifikasi *Naïve Baye Classifier* dengan pembobotan *term frequency-invers document frequency* terhadap data latih yang berjumlah 2500 data, yang memiliki komposisi label negatif, netral, positif, dan *irrelevant*. Proses pelatihan / training ini akan menghasilkan model klasifikasi untuk memberi label data uji yang belum memiliki label. Penggunaan pembobotan TF-IDF dilakukan sebagai pembandingan dengan menggunakan pembobotan TF dengan melihat akurasi masing-masing pembobotan.

Perhitungan TF-IDF menggunakan persamaan 4.4 berdasarkan dokumen contoh pada tabel 4.13 ditunjukkan pada tabel 4.24 sebagai berikut :

Tabel 4.24 Contoh Perhitungan TF-IDF

| No. | Term      | Dokumen |     |     |     | IDF | IDF+1 | TF * IDF |       |       |       |
|-----|-----------|---------|-----|-----|-----|-----|-------|----------|-------|-------|-------|
|     |           | DT1     | DT2 | DT3 | DT4 |     |       | DT1      | DT2   | DT2   | DT4   |
| 1.  | asn       | 0       | 1   | 0   | 0   | 1   | 0,602 | 0        | 1,602 | 0     | 0     |
| 2.  | bantu     | 0       | 0   | 1   | 0   | 1   | 0,602 | 0        | 0     | 1,602 | 0     |
| 3.  | cicil     | 0       | 0   | 1   | 0   | 1   | 0,602 | 0        | 0     | 1,602 | 0     |
| 4.  | fintech   | 1       | 1   | 0   | 0   | 2   | 0,301 | 1,301    | 1,301 | 0     | 0     |
| 5.  | gini      | 0       | 0   | 0   | 1   | 1   | 0,602 | 0        | 0     | 0     | 1,602 |
| 6.  | guna      | 1       | 0   | 0   | 0   | 1   | 0,602 | 1,602    | 0     | 0     | 0     |
| 7.  | hotel     | 0       | 0   | 0   | 1   | 1   | 0,602 | 0        | 0     | 0     | 1,602 |
| 8.  | inap      | 0       | 0   | 0   | 1   | 1   | 0,602 | 0        | 0     | 0     | 1,602 |
| 9.  | investasi | 0       | 1   | 0   | 0   | 1   | 0,602 | 0        | 1,602 | 0     | 0     |
| 10. | jangan    | 1       | 0   | 0   | 0   | 1   | 0,602 | 1,602    | 0     | 0     | 0     |
| 11. | kalau     | 0       | 0   | 1   | 0   | 1   | 0,602 | 0        | 0     | 1,602 | 0     |
| 12. | kartu     | 0       | 0   | 1   | 0   | 1   | 0,602 | 0        | 0     | 1,602 | 0     |
| 13. | kredit    | 0       | 0   | 1   | 0   | 1   | 0,602 | 0        | 0     | 1,602 | 0     |

Tabel 4.25 Lanjutan Contoh Perhitungan TF-IDF

| No.    | Term      | Dokumen |     |     |     | IDF   | IDF+1 | TF * IDF |       |        |       |
|--------|-----------|---------|-----|-----|-----|-------|-------|----------|-------|--------|-------|
|        |           | DT1     | DT2 | DT3 | DT4 |       |       | DT1      | DT2   | DT2    | DT4   |
| 14.    | layan     | 0       | 0   | 0   | 1   | 0,602 | 1,602 | 0        | 0     | 0      | 1,602 |
| 15.    | mahasiswa | 0       | 0   | 1   | 0   | 0,602 | 1,602 | 0        | 0     | 1,602  | 0     |
| 16.    | minta     | 0       | 1   | 0   | 0   | 0,602 | 1,602 | 0        | 1,602 | 0      | 0     |
| 17.    | rasa      | 0       | 0   | 0   | 1   | 0,602 | 1,602 | 0        | 0     | 0      | 1,602 |
| 18.    | riba      | 1       | 0   | 0   | 0   | 0,602 | 1,602 | 1,602    | 0     | 0      | 0     |
| 19.    | robot     | 0       | 0   | 0   | 1   | 0,602 | 1,602 | 0        | 0     | 0      | 1,602 |
| 20.    | sudah     | 0       | 0   | 1   | 0   | 0,602 | 1,602 | 0        | 0     | 1,602  | 0     |
| 21.    | tanpa     | 0       | 0   | 1   | 0   | 0,602 | 1,602 | 0        | 0     | 1,602  | 0     |
| Jumlah |           |         |     |     |     |       |       | 6,107    | 7,709 | 14,418 | 9,612 |

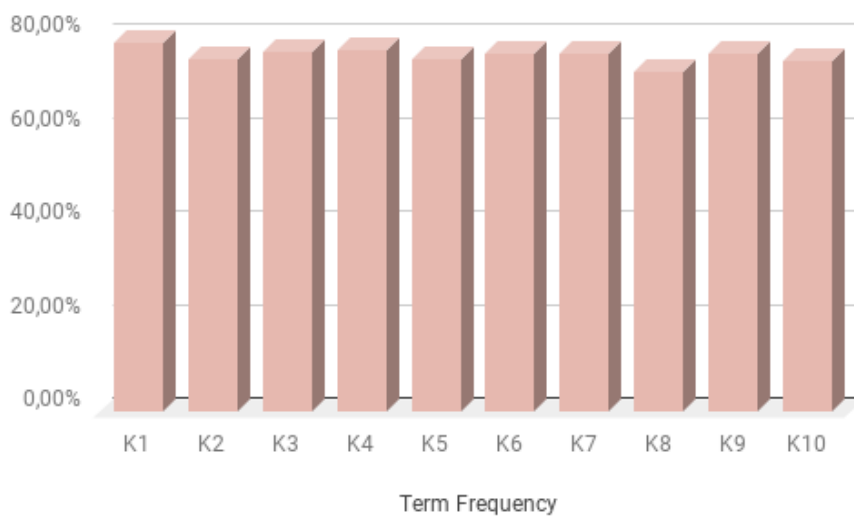
#### 4.4.2. Evaluasi Model Klasifikasi

Diperlukan evaluasi model klasifikasi untuk menghitung akurasi dari klasifikasi sentimen menggunakan metode *Naïve Bayes Classifier*. Evaluasi model klasifikasi ini dilakukan dengan metode *cross validation*, yaitu metode statistik yang dapat digunakan untuk mengevaluasi kinerja model atau algoritma dimana data dipisahkan menjadi dua subset yaitu data proses pembelajaran dan data validasi / evaluasi. Dalam metode *cross validation* penulis menggunakan dua metode, yaitu *K-fold* dan *Leave One Out*.

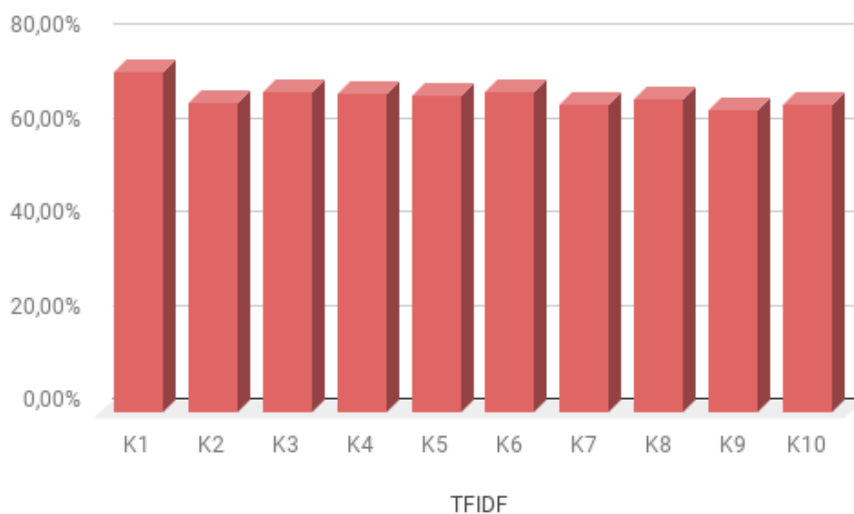
*Pertama, k-fold cross validation* yang berfungsi menentukan data training dan data testing yang baik. *Fold* yang digunakan adalah *10-fold*, yaitu dengan membagi data dengan sepersepuluh bagian. *K-Fold Cross Validation* pada iterasi ke 1 menjadikan data ke- 1 hingga 250 data pertama sebagai data uji, berikutnya untuk data ke- 251 hingga 2500 menjadi data *training*. Setelah semua iterasi dilakukan maka akan didapatkan rata-rata akurasi dari klasifikasi sentimen. Rata-rata akurasi analisis sentimen dengan pembobotan *term frequency* (TF) pada data latih *Fintech* sebesar 76,02%. Sedangkan rata-rata akurasi sentimen menggunakan pembobotan *term frequency-invers document frequency* (TF-IDF) pada data latih *Fintech* sebesar 67,24%.

Adapun masing-masing akurasi dapat dilihat pada Gambar 4.2 dan Gam-

bar 4.3 berikut :



**Gambar 4.2** Grafik Akurasi *K-fold cross validation* TF



**Gambar 4.3** Grafik Akurasi *K-fold cross validation* TFIDF

*Kedua*, menggunakan *Leave One Out Cross Validation*. Perhitungan

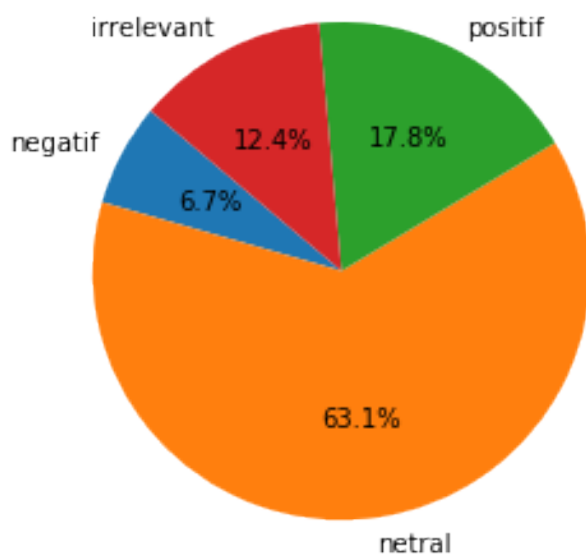
*Leave One Out Cross Validation* menggunakan seluruh data kecuali satu data digunakan untuk latih dan sisanya sebagai data uji. Proses perhitungan berulang sejumlah banyaknya objek data. Dengan demikian, setiap data pernah digunakan untuk data latih dan juga sebagai data uji. Setelah semua data di proses maka akan didapatkan rata-rata akurasi klasifikasi sentimen. Rata-rata akurasi *Leave One Out Cross Validation* dengan pembobotan *term frequency* (TF) pada data latih *Fintech* sebesar 71,68%. Sedangkan rata-rata akurasi sentimen menggunakan sentimen dengan pembobotan *term frequency-invers document frequency* (TF-IDF) pada data latih *Fintech* sebesar 51,80%.

Akurasi dengan pembobotan *term frequency* (TF) dalam perhitungan validasi *K-fold* sebesar 76,02% dan *LOOCV* sebesar 71,68%, kedua validasi ini mendapatkan nilai akurasi lebih tinggi dibandingkan dalam perhitungan pembobotan TF-IDF. Pembobotan TF-IDF menggunakan *K-Fold* mendapat nilai 71,68% dan *LOOCV* sebesar 51,80% hasilnya cenderung lebih rendah dibandingkan dengan akurasi pada *term frequency*. Hal ini menunjukkan bahwa akurasi *k-fold croos validation* dengan pembobotan TF lebih besar daripada dengan pembobotan TF-IDF. Penggunaan *k-fold croos validation* dengan 10 fold direkomendasikan untuk pemilihan model terbaik karena cenderung memberikan estimasi akurasi yang lebih baik dibandingkan dengan *leave one out cross validation*.

Akurasi dengan pembobotan *term frequency* (TF) dalam perhitungan nilai peluang frekuensi kata dari setiap kelas bekerja dengan baik, dimana nilai peluang frekuensi kata tersebut disimpan dalam tabel *Naïve Bayes* yang menyimpan nilai peluang frekuensi kata dalam data latih dan *term* atau kata. Data uji dilakukan pengecekan kata dalam daftar data latih, apabila kata dalam data uji tidak ada dalam daftar data latih maka akan diabaikan.

#### 4.5. Implementasi

Pada tahap implementasi ini dilakukan klasifikasi secara otomatis pada data uji yang belum memiliki label. Klasifikasi ini berdasarkan model klasifikasi dari data latih atau training sebelumnya dengan menggunakan klasifikasi *Naïve Bayes Classifier* dan pembobotan TF-IDF. Data uji sebanyak 3550 data dengan kata kunci *fintech*. Setelah dilabeli secara otomatis, maka didapatkan hasil klasifikasi sentimen negatif sebanyak 6,7% dengan data 242 data, sentimen netral sebanyak 63,1% dengan 2290 data, sentimen positif sebanyak 17,8% dengan 648 data, dan sentiment *irrelevant* sebanyak 12,4% dengan 452 data. Data grafik pada gambar 4.4 menunjukkan bahwa data *tweet* dengan kata kunci *Fintech* pada bulan Januari hingga Februari mendapatkan sentimen netral.



**Gambar 4.4** Hasil Implementasi Model Klasifikasi pada Data Uji

## BAB V

### PENUTUP

#### 5.1. Kesimpulan

Data yang digunakan adalah data yang bersal dari twitter yang di ambil menggunakan *API search* dengan kata kunci *FinTech*. Data yang didapatkan sebesar 8249 data, kemudian dilakukan proses pembersihan data dan mendapatkan data sebanyak 6050 data dan dibagi dua kategori, yaitu sebanyak 2500 sebagai data latih dan 3550 sebagai data uji. Data diolah menggunakan *text mining* dengan metode *Nave Bayes Classifier* menggunakan pembobotan TF dan TF-IDF serta akurasi *K-Fold Cross Validation* dan *Leave One Out Cross Validation*.

Berdasarkan penelitian yang telah dilakukan didapatkan kesimpulan bahwa Anilisis Sentimen dengan metode *Nave Bayes Classifier* dapat digunakan untuk mengklasifikasi data tweet dengan kata kunci *Fintech*. Pembobotan *term frequency* dengan akurasi *K-fold Cross Validation* mendapat nilai sebesar 76,02% dan pada akurasi *leave one out* mendapat nilai sebesar 71,68%. Sedangkan pada pembobotan *term frequency-invers document frequency* dengan akurasi *K-fold Cross Validation* mendapat nilai sebesar 67,24% dan pada akurasi *leave one out* mendapat nilai sebesar 51,80%.

Dari dua pembobotan dan dua metode akurasi yang digunakan, *term frequency* dengan metode *k-fold cross validation* menghasilkan nilai akurasi yang paling besar, yaitu 76,02%. Hasil implementasi pada data uji sebnayak 3550 data tweet menggunakan kata kunci *Fintech* dengan model

klasifikasi dari data latih menggunakan *Nave Bayes Classifier* dan pembobotan TF-IDF menghasilkan 6,7% sentimen negatif, 63,1% sentimen netral, 17,8% sentimen positif, dan 12,4% sentimen *irrelevant*.

## **5.2. Saran**

Penulis menyadari penelitian ini masih terdapat kekurangan, maka dari itu penulis berharap penelitian selanjutnya dapat memperbaiki kekurangan penelitian ini. Penelitian ini hanya dilakukan analisis dengan menggunakan satu metode dan tidak ada perbandingan metode. Diharapkan pada penelitian berikutnya dapat melakukan perbandingan metode analisis sentimen agar dapat mengetahui perbandingan dan diketahui metode apa yang lebih tepat digunakan dalam penelitian sejenis. Serta penelitian ini hanya analisis tanpa membuat rancang bangun sistem. Jadi diharapkan penelitian berikutnya dapat merancang sistem sentimen analisis yang dapat digunakan langsung oleh user.

## DAFTAR PUSTAKA

- A. Ghulam, A., Teguh, & E., Adhistya. (2014) “*Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation*”. Teknik Elektro dan Teknologi Informasi, FT UGM, 2014.
- A. Muhammad Alvin, “Konsep Penggunaan Financial Technology dalam Membantu Masyarakat Sub Urban di Indonesia dalam Melakukan Transaksi Finansial”. Dipublikasikan pada April 2018, tersedia di <https://www.researchgate.net>.
- A., Widiastuti, N., S., Santosa, & C., Supriyanto. (2014). “Algoritma Klasifikasi Data Mining *Naïve Bayes Classifier Particle SWARM Optimization* Untuk Deteksi Penyakit Jantung. Jurnal Pseudocode, Vol. 1 No. 1, pp 11-14.
- Akmal. “Inovasi *Financial Technology (Fintech)* Primadona Sistem Keuangan di Era Digital”. Dipublikasi 12 Maret 2019, Tersedia di <https://www.kompasiana.com>.
- Apriyani. “OJK: Waspada dampak teknologi perbankan. Infobanknews”. Dipublikasi 6 Oktober 2016, Tersedia di <http://infobanknews.com>.
- B., Dwi Pramita, Saptono, Ristu, & A, Rini. (2018) “Academica Articles Classification Using *Naïve Bayes Classifier (NBC)* Method”. ITS-MART: Jurnal Ilmiah Teknologi dan Informasi, Vol. 7, No.2.
- Bertalya. (2009). “Konsep Data Mining”. Universitas Gunadarma

- Bintarto, Elshabyta Auditya. (2018) “*Fintech dan Cashless Society: Sebuah REvolusi Pendongkrak Ekonomi Kerakyatan*, ACCESS 18, Himma Maksi UGM, pp 2-10.
- Chrismastinto, Imanuel Adhitya Wulanata. (2017). “Analisis *Swot* Implementasi Teknologi Finansial Terhadap Kualitas Layanan Perbankan di Indonesia”. Jurnal Ekonomi dan Bisnis, Vol. 20, No. 1, pp 133-144.
- Deviyanto, Akhmad dan Wahyudi, M. Didik R. (2018). “Penerapan Analisis Sentimen Pada Pengguna *Twitter* Menggunakan Metode *K-Nearest Neighbor*”. JISKa, Vol. 3, No. 1, pp 1-13.
- Dhar, V. & Stein, R. M.. (2017). “Economic and Business Dimensions Fin-Tech Platforms and Strategy”. Communications Of The ACM, Vol. 60, No 10., pp 32-35.
- Ernama, Budiharto, Hendro S. (2017). “Pengawasan Otoritas Jasa Keuangan Terhadap Financial Technology (Peraturan Otoritas Jasa Keuangan Nomor 77/POJK.01/2016)”Diponegoro Law Journal, Vol. 6, No.3, pp. 1-2.
- Faizin, AS Wahid. (2018). “Implementasi *K-Means Clustering* Pada Terjemah *Al-Qur'an* Berdasarkan Keterkaitan Topik”.Teknik Informatika UIN Sunan Kalijaga Yogyakarta.
- G., Ferly, F., Ali & A., Putra. (2017) “Analisis Sentimen Pada Ulasan Aplikasi *Mobile* Menggunakan *Naïve Bayes* Dan Normalisasi Kata Berbasis *Levenshten Distance*”. Jurnal Systemic, Vol. 03, No 02, pp 1-6.

- Hendriyani. (2013). "Penggunaan Jejaring Sosial Akun Twitter @Standupindobgd Dalam Penyampaian Informasi Seputar Kegiatan Standup Comedy Kepada Followers-nya di Bandung ". Digital library - Perpustakaan Pusat Unikom.
- Herwijayanti, Bening , Ratnawati, Dian Eka , dan Muflikhah Lailil. (2018). "Klasifikasi Berita *Online* dengan menggunakan Pembobotan *TF-IDF* dan *Cosine Similarity*". Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol.2, No. 1, pp 306-312.
- Hidayatullah, Ahamd Fathan. (2016). "Pengaruh *Stopword* Terhadap Performa Klasifikasi *Tweet* Berbahasa Indonesia". JISKa, Vol. 1, No.1, pp 1-4.
- Hidayatullah, Ahmad Fathan dan SN Azhari. (2014). "Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter ". SemnasIF UPN Yogyakarta.
- Juditha, Christiany. (2015). "Fenomena *Trending Topic* di *Twitter*: Analisis Wacana Twit #savehajilulung". Jurnal Penelitian Komunikasi dan Pengembangan, VOL. 16, No. 2, pp 138-152.
- Jurafsky, Daniel, and James H. Martin. (2000). *Speech and Processing Language: An Introduction to Natural Language Processing, Computational Linguistic, and Speech Recognition*. Prentice Hall.
- K. Bambang, Efendi S., & Opim Salin Sitompul. (2012). "Klasifikasi KOn-

- ten Berita Dengan Metode *Text Mining*". Jurnal Dunia Teknologi Informatika, Vol. 1, No. 1, pp 14-19.
- M. Ridwan. (2018). "Analisis SWOT Financial Technology (Fintech) Pembiayaan Perbankan Syariah Di Indonesia". AT-Tawassuth, Vol. III, No. 2, pp 335-357.
- Maarif, Abdul Aziz. (2015). "Penerapan Algoritma *TF-IDF* Untuk Pencarian Karya Ilmiah". Universitas Dian Nuswantoro Semarang.
- Mustaqhfiri, M., Abidin, Z., dan Kusumawati, R. (2011). "Peringkasan Teks Otomatis Berita Berbahasa Indonesia Menggunakan *Metode Maximum Marginal Relevance*". Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, UIN Malik Ibrahim.
- Nugroho, Agung. (2018). "Analisis Sentimen Pada Media Sosial *Twitter* Menggunakan *Naïve Bayes Classifier* Dengan Ekstraksi fitur *N-Gram*". Jurnal Sains Komputer & Informatika (J-SAKTI), Vol. 2, No. 2, pp 200-209.
- Nuriwan, B., Teguh, & P., Adhistya. (2015) "Analisis Sentimen Data Presiden Jokowi Dengan *Preprocessing* Normalisasi Dan *Stemming* Menggunakan Metode *Naïve Bayes* Dan SVM. Jurnal Dinamika Informatika, Vol. 5, No. 1.
- P. Riszki Wijayatun dan N., Yusuf Sulisty. (2016). "Prediksi *Rating* Film Menggunakan *Metode Nave Bayes Classifier*". Jurnal Teknik Elektro Vol. 8, No. 2.

- Pang, Bo and Lee, Lillian. (2004). “*A sentimental Education: Sentiment Analysis Using Subjectivity Aummarization Based on Minimum Cuts*”. *ACM Digital Library*, Article No. 271.
- Pak & Paroubek. (2010). “Twitter as a Corpus for Sentiment Analysis and Opinion Mining ”. *Universite de Paris-Sud, Laboratoire LIMSI-CNRS, Batiment 508*,
- R., Mujib, S., Hadi & Sarosa, M. (2013) “Penerapan Data Mining Untuk Evaluasi KInerja Akademik Mahasiswa Menggunakan Algoritma *Naïve Bayes Classifier* ”. *Jurnal EECCIS*, Vol. 7, No. 1, pp 59-64.
- R. Tri Indah F. “Persepsi Masyarakat Kota Medan Terhadap Penggunaan *Financial Teknologi (Fintech)*”. *At-Tawassuth*, Vol. 3, No. 1, PP 642-661.
- Ridwan Taufiq. (2016). “Pandangan Regulasi Secara Sosial Terhadap Pengaruh *Financial Technology*”. *Universitas Mercubuana*.
- S. Nirmala Fa‘izah, Indriati, & P. Rizal Setya. (2018). “Peringkasan Teks Otomatis Menggunakan Metode *Maximum Marginal Relevance* Pada Hasil Pencarian Sistem Temu Kembali Informasi Untuk Artikel Berbahasa Indonesia”. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol.2, No. 11, pp 5494-5502.
- S., Pipit Buana dan D., Handriyani. (2018). “Prospek *Financial Technology (FINTECH)* Di Sumatra Utara Dari Sisi Literasi Keuangan, Inklusi Keuangan Dan Kemiskinan ”. *Kajian Akulansi*, Vil. 19, No. 2, pp 9-17.

- Saputro, Michael & Tyas. (2017). “Klasifikasi Lagu Daerah Indonesia Berdasarkan Lirik Menggunakan Metode TF-IDF dan *Naïve Baye*”. Jurnal Teknologi Informatika dan Terapan, Vol. 04, No. 01, pp 47-52.
- Sevsa, Bama A. (2018). “Analisis Sentimen Komentar Pada Indeks Kinerja Dosen Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta menggunakan *Naïve Bayes Classifier*”. Yogyakarta: UIN Sunan Kalijaga.
- Suyanto. (2018). “*Machine Learning* Tingkat Dasar dan Lanjut). Bandung: Informatika.
- Wahid, Devid Haryalesmana dan SN, Azhari. (2016) “Peringkasan Sentimen Estraktif di *Twitter* Menggunakan *Hybrid TF-IDF* dan *Cosine Similarity*”. IJCCS, Vol. 10, No. 2, pp 207-208.
- Willi. (2015). “Distributed twitter crawler “. Universitas Pendidikan Indonesia
- Vinodhini, G., and Chandrasekaran, M. (2012). “Sentiment Analysis and Opinion Mining ”: A Survey. International Journal Of Advanced Research in Computer Science and Software Engineering, Vol.2, NO. 2.
- W. Rizki Tri, P. Dhidik, dan S., Eko. (2017) “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi”. Jurnal Teknik Elektro, Vol. 9, No. 1.
- Wibowo, Budi. (2016). “Analisa Regulasi Fintech Dalam Membangun

Pereko-nomian di Indonesia ”. Program Pascasarjana, Program Magister Teknik Elek-tro, Universitas Mercu Buana Jakarta, Indonesia.

Xhemali Daniela, Christopher, Roger. (2009). “Nave Bayes vs. Decision Treesvs. Neural Networks in the Classification of Training Web Pages”ComputerScience, Loughborough University, Loughborough, LE11 3TU, UK.

Yislam dan B. Indra. (2017). “Analisis Sentimen Masyarakat Terhadap Pemerintahan Jokowi Menggunakan Data Twitter ”. Fakultas Ilmu Komputer UI,Depok, Indonesia.

Zaman, B. dan Winarko, E. (2011). “Analisa Fitur Kalimat Peringkas Teks Oto-matis pada Bahasa Indonesia ”. IJCCS, Vol.5 No.2.

<https://www.cnnindonesia.com/teknologi/20150326141025-185-42076/jumlah-pengguna-twitter-di-indonesia-akhirnya-terungkap>  
Diak-ses pada 02 April 2019.

Kominfo<https://www.kominfo.go.id/content/detail/3415/kominfo-pengguna-internet-di-indonesia-63-juta-orang/0/beritasatker> Diakses pada02 April 2019.