

BAB 6

KESIMPULAN DAN SARAN

Pada bab ini akan dibahas kesimpulan yang didapatkan penulis terkait penelitian yang telah dilakukan. Kesimpulan pada penelitian ini akan diambil kelebihan dan kekurangan yang didapatkan dari hasil penelitian, proses penelitian dan tanggapan yang diambil dari peneliti. Bab ini peneliti juga memberikan saran yang dapat dipergunakan untuk penelitian selanjutnya yang akan menggunakan topik maupun metodologi yang relevan.

6.1 Kesimpulan

Berikut merupakan hasil kesimpulan yang didapatkan dari hasil dari penelitian ini:

1. Dari hasil analisis sentimen dan dari penentuan kelas atribut pada penelitian ini memperlihatkan bahwa opini pada masyarakat lebih banyak mengandung kata positif, di mana kelas tersebut berdampak baik bagi untuk penyelenggara Debat Pemilihan Presiden.
2. Hasil dari akurasi yang didapatkan dalam analisis sentimen dengan menggunakan metode *Naïve Bayes classifier* (NBC) mencapai 98% dibandingkan dengan menggunakan metode *One Rules (One-R)* dan *Zero-R* yang mencapai akurasi 46%. Pada penelitian ini membuktikan bahwa analisis sentimen dengan menggunakan metode *Naïve Bayes classifier* baik digunakan dalam pengklasifikasian analisis sentimen berbahasa Indonesia.
3. Dari pembuatan *N-gram* didapatkan kata-kata yang paling sering digunakan pada data *tweet* debat pilpres 2019. Kata-kata tersebut adalah gagapunicorn, debatcawapres, debatpilpres, Prabowo, Jokowi.

6.2 Saran

Berikut adalah hasil saran yang diberikan peneliti yang berkaitan dengan topik pembahasan, metodologi penelitian dan proses penelitian:

1. Penelitian ini dapat dijadikan sebagai acuan dan evaluasi bagi pihak panitia yang melaksanakan debat pilpres.
2. Kamus yang digunakan sebagai acuan dalam pelabelan data baik itu untuk kata positif maupun kata negatif dapat ditambahkan agar menjadi lebih baik lagi.



Daftar Pustaka

- [1] K. Bohang, Fatimah, “Berapa jumlah pengguna internet Indonesia.” p. <https://tekno.kompas.com/2018/02/22/>, 2018.
- [2] <https://nasional.kompas.com/read/2019/01/17/10354791/debat-diadakan-untuk-kali-pertama-di-Indonesia>
- [3] andi nurul Hidayat, “Analisis Sentimen Terhadap Wacana Politik Pada Media Masa Online Menggunakan Algoritma *Support Vector Machine* Dan *Naive Bayes*,” J. Elektron. Sist. Inf. dan Komput., vol. 1, no. 1, pp. 12–18, 2015.
- [4] F. N. Zuhri, A. Alamsyah, S. Si, and M. Sc, “Menggunakan *Naive Bayes Classifier* Di Forum Kaskus Public Sentimen Analysis of Smartfren Brand Using *Naive Bayes Classifier* on Kaskus Forum,” vol. 4, no. 1, pp. 242–251, 2017.
- [5] L. Oktasari, Y. H. Chrisnanto, and R. Yuniarti, “Text *Mining* Dalam Analisis Sentimen Asuransi Menggunakan Metode Niave Bayes Classifier,” Pros. SNST, vol. 7, pp. 37–42, 2016.
- [6] A. F. Hidayatullah and A. Sn, “Analisis Sentimen dan Klasifikasi Kategori Terhadap Tokoh Publik Pada *Twitter*,” Semin. Nas. Inform. 2014, vol. 2014, no. August 2013, pp. 0–8, 2014.
- [7] N. Saputra, Permanasari, “Analisis Sentimen Data Presiden Jokowi dengan Preprocessing Normalisasi dan Stemming menggunakan Metode *Naive Bayes*.”
- [8] F. Nurhuda, S. W. Sihwi, and A. Doewas, “Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari *Twitter* Menggunakan Metode *Naive Bayes Classifier*,” ITSmart J. Ilm. Teknol. dan Inf., pp. 35–42.
- [9] <http://makassar.tribunnews.com/2019/01/17/tribunwiki-kapankah-debat-Presiden-dimulai-pertama-kali-di-Indonesia-ini-sejarahny>
- [10] M. Ridwan, H. Suyono, and M. Sarosa, “Penerapan Data *Mining* Untuk

- Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma *Naive Bayes Classifier*,” *Eeccis*, vol. 7, no. 1, pp. 59–64, 2013.
- [11] L. R. Angga Ginanjar Mabur, “Penerapan Data *Mining* Untuk Memprediksi Kriteria Nasabah Kredit,” *J. Komput. dan Inform.*, vol. 1, no. 1, pp. 53–57, 2012.
- [12] M. K. Sari and I. Wisnubhadra, “Pembangunan Aplikasi Klasifikasi Mahasiswa Baru untuk Prediksi Hasil Studi Menggunakan Naïve Bayes Classifier,” *Buana Inform.*, vol. 7, no. April, pp. 135–142, 2015.
- [13] Andri, Y. N. Kunang, and S. Murniati, “Implementasi Teknik Data *Mining* Untuk Memprediksi Tingkat Kelulusan Mahasiswa Pada Universitas Bina Darma Palembang,” *Semin. Nas. Inform. 2013 (semnasIF 2013)*, vol. 2013, no. A, pp. 56–63, 2013.
- [14] L. Zhang and B. Liu, “Sentimen Analysis and Opinion *Mining*,” *Encycl. Mach. Learn. Data Min.*, no. May, pp. 1–10, 2016.
- [15] G. A. Buntoro, T. B. Adji, and A. E. Purnamasari, “Sentimen Analysis *Twitter* dengan Kombinasi Lexicon Based dan Double Propagation,” *Citee 2014*, vol. ISSN: 2085, no. OCTOBER 2014, pp. 7–8, 2014.
- [16] S. L. Ting, W. H. Ip, and A. H. C. Tsang, “Is Naïve bayes a good classifier for document classification?,” *Int. J. Softw. Eng. its Appl.*, vol. 5, no. 3, pp. 37–46, 2011.
- [17] A. W. Puspita, “No TitleANALISIS PENGGUNAAN MEDIA SOSIAL *TWITTER* OLE H PEJABAT PUBLIK DALAM PENERAPAN GOOD GOVERNANCE (Studi terhadap akun Gubernur Lampung, Gubernur Jawa Tengah dan Wali Kota Bandung),” 2016.
- [18] Xhemali, D., J. Hinde, C. & G. Stone, R., 2009. Naïve Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages. *International Journal of Computer Science Issues*, 4(1), pp. 16-23.
- [19] <http://www.bps.go.id/>





LAMPIRAN

1. Dataset

Tabel 1 sebagian dari data mentah yang telah di dapat dengan menggunakan *GetOldTweets-python-master*

| NO | Username | Tweet |
|----|-----------------|--|
| 1 | CNNIndonesia | KPU Ragukan Keaslian Surat yang Sempat Tercoblos di Malaysia #DebatPilpres2019 |
| 2 | serbaserbi_indo | Hmmmm.... Bikin rusuhhhh ajahhhh !!! #SepertiDuriDalamDaging #BPNPecah #2019tetapjokowi |
| 3 | deenandep | debat capres berasa nonton stand up comedy ngakak mulu gua. jawabannya kocak ugha y pak. Kaga #DebatPilpres2019 https://twitter.com/deenandep/status/1117078942060429312 |
| 4 | Yudi79130082 | Ni capresnya pk sandi apa pk prbwo... .. #DebatPilpres2019 https://twitter.com/Yudi79130082/status/1117078933898289158 |
| 5 | jarkiyo | Ditanya tentang e a sport dan mobil legend jawabnya malah permainan harvest moon . Ra masyook. #DebatTerakhirPilpres2019 #DebatPilpres2019 #OptimisIndonesiaMaju https://twitter.com/jarkiyo/status/1117078860988686337 |
| 6 | martaariesty | Turunkan harga listrik#DebatPilpres2019 |
| 7 | Ria52123551 | Duit Pribadi ?? Bayar karyawannya saja ngk bisa !!! #01JOKOWIAMIN |
| 8 | DianArdiantiP1 | Hanya ingin membangun bangsa ini. #JokowiUpgradeIndonesia #02GagapUnicorn |
| 9 | Ardiclay1 | malu maluin negara aja nih...berita sampe dunia internasional tentang unicorn dan ekonomi digital aja gak ngerti..gimana mau membangun negara kalau kaya gitu aja gak ngerti...udahlah..usahamu sia sia...malu maluin negara aja.. #02GagapUnicorn pic.twitter.com/bDIBuGS9Yv |
| 10 | alimunnasrun17 | bro lu harus belajar banyak di organisasi supaya jangan bego seperti ini. mana ada prabowo selamatkan tanah negara. baca sejarah 1997 soal HGU yang di rampas oleh prabowo di aceh. #2019NenoGantiTuhan #02GagapUnicorn #JokowiOrangnyaBaik #ManaKeluargamu |
| 11 | sendirianbanget | bapak salah pak gimana mau punya visi misi kalau yang di maksud saja tidak tahu #DebatCawapres2019 |

| | | |
|----|-----------------|--|
| 12 | bralbrilbrul | Pak sandi senyumnya yaAllahhhhh :(#DebatTerakhirPilpres2019 #DebatCawapres2019 pic.twitter.com/S8w8PIB6Zm |
| 13 | akbarfatts | jangan sampai mbah maruf amin afk lagi #DebatCawapres2019 |
| 14 | chairol_saleh02 | mantap pak Prabowo. Menang telak. #dabatacapres #DebatCawapres2019 #PrabowoBentengNKRI #18HariLagiCoblosCapres02 #PrabowoMenangDebat #JokowiDiambangKekalahan #2019gantipresiden pic.twitter.com/MVLRi8Zehr |
| 15 | muhammadrid_ | Pak Jokowi kurang persiapan dan sepertinya dia mengakui kekalahan di debat ini. Wkwk #DebatCawapres2019 #DebatKeempatCapres2019 #PrabowoMenangDebat #PrabowoBentengNKRI #2019GantiPresiden |
| 16 | tejmaul | Nonton debat cm mau blg wkwkwk.. #DebatCawapres2019 #DebatAsyikJokowi |
| 17 | serbaserbi_indo | KPU TIDAK TUNDUK PADA PIHAK MANAPUN#debatpilpres2019 |
| 18 | Saratogain | Lalu gimana ya nasib Sandi selanjutnya #SandiKembaliKeDKI #IndonesiaFraudElection #2019tetapjokowi #jokowi #pilpres2019 #nkri #jokowimaruf #jokowi2periode #umkm #ingatjokowi #nawacita #jokowijk #salamsatujempol #prabowo #debatpilpres2019 #prabowosandi2019 #kubuhhoax pic.twitter.com/18YZOX7KUG |
| 19 | MetroTVToday | Cara yang normal adalah dengan cara memperkenalkan diri dan menargetkan pemilih. Mungkin yang melakukan politik uang ini orang yang menggunakan dana terbatas." M. Qodari #DebatPilpres2019 #MetroTVIndonesiaMemilih #KnowledgeToElevate" |
| 20 | Vincensiatan | Baru tahap debat aja sudah jadi bahan pembicaraan media asing karena Gagap Unicorn" |
| 21 | HusnahQoriatul | Prabowo aja mengakui kehebatan jokowi. #JokowiUpgradeIndonesia #02GagapUnicorn #jokowilagi #01IndonesiaMaju #JokowiMenangPrabowoNyerah pic.twitter.com/Mk2GGg2tX8 |
| 22 | DianArdiantiP1 | sama sekali ngga bermanfaat!!! #JokowiUpgradeIndonesia #02GagapUnicorn #jokowilagi #01IndonesiaMaju #JokowiMenangPrabowoNyerah pic.twitter.com/NLl2ORevWb |
| 23 | DianSalamah | kita punya menteri dengan level dunia #JokowiUpgradeIndonesia #JKWMerahPutih #01IndonesiaMaju #02GagapUnicorn pic.twitter.com/06PsIiuaJs |
| 24 | serbaserbi_indo | Ternyata ada yg lagi caper nihhh yeeeeee..... #MassaBayaranKivlan #KedaulatanMilikRakyat |
| 25 | seputarbetting | Jokowi Pembawa Keberuntungan https:// sabungovo.com/2019/04/24/aya m-hias-pembawa-keberuntungan-jika-dipelihara/#ayambangkokimport #JokowiPutihkanGBK #debatpilpres2019 |

Tabel 2 sebagian dari data yang sudah melalui proses *cleansing* menggunakan bahasa pemrograman R

| NO | Username | Tweet |
|----|-----------------|--|
| 1 | CNNIndonesia | kpu ragukan keaslian Surat yang Sempat tercoblos malaysia debatpilpres2019 |
| 2 | serbaserbi_indo | Hmmmm bikin rusuhhhh ajahhhh SepertiDuriDalamDaging bnbpecah 2019tetapjokowi |
| 3 | deenandepp | debat capres berasa nonton stand up comedy ngakak mulu gua. jawabannya kocak juga pak. kaga debatPilpres2019 |
| 4 | Yudi79130082 | ni capresnya pk sandi apa pk prbwo debatpilpres2019 |
| 5 | jarkiyo | ditanya tentang ea sport dan mobil lagend jawabnya malah permainan harvest moon . da masyook. debatterakhirPilpres2019 debatpilpres2019 optimisIndonesiaMaju |
| 6 | martaariesty | turunkan harga listrik debatpilpres2019 |
| 7 | Ria52123551 | duit Pribadi Bayar karyawannya saja ngak bisa 01jokowiamin |
| 8 | DianArdiantiP1 | Hanya ingin membangun bangsa ini. JokowiUpgradeIndonesia 02GagapUnicorn |
| 9 | Ardiclay1 | malu maluin negara aja berita sampe dunia internasional tentang unicorn ekonomi digital aja gak ngerti..gimana mau membangun negara kalau kaya gitu aja gak ngerti udahlah usahamu malu maluin negara aja.. 02gagapunicorn |
| 10 | alimunnasrun17 | bro lu harus belajar banyak di organisasi supaya jangan bego seperti . mana ada prabowo selamatkan tanah negara. baca sejarah 1997 soal hgu yang rampas oleh prabowo aceh. 2019nenogantituhan 02gagapunicorn jokowiorangnyabaik manakeluargamu |
| 11 | sendirianbanget | bapak salah pak gimana mau punya visi misi kalau yang maksud saja tidak tahu DebatCawapres2019 |
| 12 | bralbrilbrul | pak sandi senyum yallahhhhh debatterakhirpilpres2019 debatcawapres2019 |
| 13 | akbarfatts | jangan sampai mbah maruf amin afk lagi debatcawapres2019 |
| 14 | chairol_saleh02 | mantap pak prabowo. menang telak. dabatcapres debatcawapres2019 prabowobentengkri 18Harilagicobloscapres02 prabowomenangdebat jokowidiambangkalah 2019gantipesiden |
| 15 | muhammadrid_ | pak jokowi kurang persiapan dan sepertinya dia mengakui kekalahan di debat ini. wkwk debatcawapres2019 debatkeempatcapres2019 prabowomenangdebat prabowobentengkri |

| | | |
|----|-----------------|---|
| | | 2019gantipresiden |
| 16 | tejmaul | nonton debat cm mau blg wkwwk.. debatcawapres2019 debatasyikJokowi |
| 17 | serbaserbi_indo | Kpu tudak tunduk pada pihak mana debatpilpres2019 |
| 18 | Saratogain | Lalu gimana ya nasib Sandi selanjutnya SandiKembaliKeDKI #IndonesiaFraudElection 2019tetapjokowi jokowi pilpres2019 nkri jokowimaruf jokowi2periode umkm ingatjokowi nawacita jokowijk salamsatujempol prabowo debatpilpres2019 prabowosandi2019 kubuhoax |
| 19 | MetroTVToday | cara yang normal adalah dengan cara memperkenalkan diri dan menargetkan pemilih mungkin yang melakukan politik uang ini orang yang menggunakan dana terbatas m qodari debatpilpres2019 MetrotvIndonesiaMemilih KnowledgeToElevate" |
| 20 | Vincensiatan | baru tahap debat aja sudah jadi bahan pembicaraan media asing karena Gagap Unicorn" 02gagapunicorn |
| 21 | HusnahQoriatul | Prabowo aja mengakui kehebatan jokowi. #JokowiUpgradeIndonesia 02GagapUnicorn jokowilagi 01IndonesiaMaju JokowiMenangPrabowoNyerah |
| 22 | DianArdiantiP1 | sama sekali ngga bermanfaatJokowiUpgradeIndonesia 02GagapUnicorn jokowilagi 01IndonesiaMaju JokowiMenangPrabowoNyerah |
| 23 | DianSalamah | kita punya menteri dengan level dunia #JokowiUpgradeIndonesia JKWMerahPutih 01IndonesiaMaju 02GagapUnicorn |
| 24 | serbaserbi_indo | Ternyata ada yg lagi caper nihhh yeeeeee..... MassaBayaranKivlan KedaulatanMilikRakyat |
| 25 | seputarbetting | Jokowi Pembawa Keberuntungan thefightingcock ayambangkokimport JokowiPutihkanGBK debatpilpres2019 |

Tabel 3 sebagian dari data yang sudah melalui proses *Stopwords* dan *Stemming* menggunakan program *Jsastrawi* dan *Exude* pada bahasa pemrograman *Java Netbeans*.

| NO | Username | Tweet |
|----|-----------------|--|
| 1 | CNNIndonesia | kpu ragu keaslian durat yang dempat tercoblos malaysia debatpilpres2019 |
| 2 | serbaserbi_indo | hmmmm bikin rusuhhhh ajahhhh SepertiDuriDalamDaging bnbpecah 2019tetapjokowi |

| | | |
|----|-----------------|--|
| 3 | deenandepp | debat capres berasa nonton stand up comedy ngakak mulu gua. jawabannya kocak juga pak. kaga debatPilpres2019 |
| 4 | Yudi79130082 | ni capresnya pk sandi apa pk prbwo debatpilpres2019 |
| 5 | jarkiyo | ditanya tentang ea sport dan mobil lagend jawabnya malah permainan harvest moon . da masyook. debatterakhirPilpres2019 debatpilpres2019 optimisIndonesiaMaju |
| 6 | martaariesty | turun harga listrik debatpilpres2019 |
| 7 | Ria52123551 | duit Pribadi Bayar karyawannya saja ngak bisa 01jokowiamin |
| 8 | DianArdiantiP1 | hanya ingin membangun bangsa ini. JokowiUpgradeIndonesia 02GagapUnicorn |
| 9 | Ardiclay1 | malu maluin negara aja berita sampe dunia internasional tentang unicorn ekonomi digital aja gak ngerti.. gimana mau membangun negara kalau kaya gitu aja gak ngerti udahlah usahamu malu maluin negara aja.. 02gagapunicorn |
| 10 | alimunnasrun17 | bro lu harus belajar banyak di organisasi supaya jangan bego seperti . mana ada prabowo selamatkan tanah negara. baca sejarah 1997 soal hgu yang rampas oleh prabowo aceh. 2019nenogantituhan 02gagapunicorn jokowiorangnyabaik manakeluargamu |
| 11 | sendirianbanget | bapak salah pak gimana mau punya visi misi kalau yang maksud saja tidak tahu DebatCawapres2019 |
| 12 | bralbrilbrul | pak sandi senyum yallahhhhh debatterakhirpilpres2019 debatcawapres2019 |
| 13 | akbarfatts | jangan sampai mbah maruf amin afk lagi debatcawapres2019 |
| 14 | chairol_saleh02 | mantap pak prabowo. menang telak. dabatcapres debatcawapres2019 prabowobentengkri 18Harilagicobloscapres02 prabowomenangdebat jokowidiambangkalah 2019gantipresiden |
| 15 | muhammadrid_ | pak jokowi kurang persiapan dan sepertinya dia mengakui kekalahan di debat ini. wkwk debatcawapres2019 debatkeempatcapres2019 prabowomenangdebat prabowobentengkri 2019gantipresiden |
| 16 | tejmaul | nonton debat cm mau blg wkwkwk.. debatcawapres2019 debatasyikJokowi |
| 17 | serbaserbi_indo | Kpu tudak tunduk pada pihak mana debatpilpres2019 |
| 18 | Saratogain | lalu gimana ya nasib sandi selanjutnya sandikembalikedki Indonesiafraudelection 2019tetapjokowi jokowi pilpres2019 nkri jokowimaruf jokowi2periode umkm ingatjokowi nawacita jokowijk salamsatujempol prabowo debatpilpres2019 prabowosandi2019 kubuhoax |
| 19 | MetroTVToday | cara yang normal adalah dengan cara memperkenalkan diri dan menargetkan pemilih mungkin yang |

| | | |
|----|-----------------|---|
| | | melakukan politik uang ini orang yang menggunakan dana terbatas m qodari debatpilpres2019 metrotvIndonesiamemilih knowledgetoelevate" |
| 20 | Vincensiatan | baru tahap debat aja sudah jadi bahan pembicaraan media asing karena Gagap Unicorn" 02gagapunicorn |
| 21 | HusnahQoriatul | Prabowo aja mengakui kehebatan jokowi. jokowiupgradeIndonesia 02gagapunicorn jokowilagi 01IndonesiaMaju jokowimenangprabowoNyerah |
| 22 | DianArdiantiP1 | sama sekali ngga bermanfaatJokowiUpgradeIndonesia 02GagapUnicorn jokowilagi 01IndonesiaMaju JokowiMenangPrabowoNyerah |
| 23 | DianSalamah | kita punya menteri dengan level dunia jokowiupgradeIndonesia jkwmerahputih 01indonesiamaju 02gagapunicorn |
| 24 | serbaserbi_indo | ternyata ada yg lagi caper nihhh yeeeeee massabayarankivlan kedaulatanmilikrakyat |
| 25 | seputarbetting | jokowi pembawa keberuntungan thefightingcock ayambangkokimport jokowiputihkangbk debatpilpres2019 |

Tabel 4 sebagian hasil penentuan kelas sentimen pada data yang sudah bersih.

| No | Username | Tweet | Positif | Negatif | Skor | Kelas |
|----|-----------------|--|---------|---------|------|---------|
| 1 | CNNIndonesia | kpu ragu keaslian surat yang sempat tercoblos malaysia debatpilpres2019 | 0 | 0 | 0 | Netral |
| 2 | serbaserbi_indo | hmmmm bikin rusuhhhh ajahhhh sepertifuridalamfaging bnbpecah 2019tetapjokowi | 0 | 1 | -1 | Negatif |
| 3 | deenandepp | debat capres berasa nonton stand up comedy ngakak mulu gua. jawabannya kocak juga pak. kaga debatPilpres2019 | 0 | 0 | 0 | Netral |

| | | | | | | |
|----|----------------|--|---|---|----|---------|
| 4 | Yudi79130082 | ni capresnya pk sandi apa pk prbwo debatpilpres2019 | 0 | 0 | 0 | Netral |
| 5 | jarkiyo | ditanya tentang ea sport dan mobil lagend jawabnya malah permainan harvest moon . da masyook. debatterakhirPilpres2019 debatpilpres2019 optimisIndonesiaMaju | 0 | 0 | 0 | Netral |
| 6 | martaariestya | turun harga listrik debatpilpres2019 | 0 | 1 | -1 | Netral |
| 7 | Ria52123551 | duit Pribadi bayar karyawannya saja ngak bisa 01jokowiamin | 0 | 1 | -1 | Netral |
| 8 | DianArdiantiP1 | hanya ingin membangun bangsa ini. JokowiUpgradeIndonesia 02GagapUnicorn | 1 | 0 | 1 | Positif |
| 9 | Ardiclay1 | malu maluin negara aja berita sampe dunia internasional tentang unicorn ekonomi digital aja gak ngerti..gimana mau membangun negara kalau kaya gitu aja gak ngerti udahlah usahamu malu maluin negara aja.. 02gagapunicorn | 0 | 3 | -3 | Negatif |
| 10 | alimunnasrun17 | bro lu harus belajar banyak di organisasi supaya jangan bego seperti . mana ada prabowo selamatkan tanah negara. baca sejarah 1997 soal hgu yang rampas oleh prabowo aceh. 2019nenogantituhan 02gagapunicorn jokowiorangnyabaik manakeluargamu | 0 | 1 | -1 | Negatif |

| | | | | | | |
|----|-----------------|--|---|---|----|---------|
| 11 | sendirianbanget | bapak salah pak gimana mau punya visi misi kalau yang maksud saja tidak tahu debatcawapres2019 | 0 | 0 | 0 | Netral |
| 12 | bralbrilbrul | pak sandi senyum yallahhhh debatterakhirpilpres2019 debatcawapres2019 | 1 | 0 | 1 | Positif |
| 13 | akbarfatts | jangan sampai mbah maruf amin afk lagi debatcawapres2019 | 0 | 1 | -1 | Negatif |
| 14 | chairol_saleh02 | mantap pak prabowo. menang telak. dabatcapres debatcawapres2019 prabowobentengkri 18Harilagicobloscapres02 prabowomenangdebat jokowidiambangkalah 2019gantipresiden | 2 | 0 | 2 | Positif |
| 15 | muhammadrid_ | pak jokowi kurang persiapan dan sepertinya dia mengakui kekalahan di debat ini. wkwk debatcawapres2019 debatkeempatcapres2019 prabowomenangdebat prabowobentengkri 2019gantipresiden | 0 | 2 | -2 | Negatif |
| 16 | tejmaul | nonton debat cm mau blg wkwkwk.. debatcawapres2019 debatasyikJokowi | 0 | 0 | 0 | Netral |
| 17 | serbaserbi_indo | kpu tudak tunduk pada pihak mana debatpilpres2019 | 0 | 0 | 0 | Netral |
| 18 | Saratogain | lalu gimana ya nasib sandi selanjutnya Sandikembalikedki | | | | |

| | | | | | | |
|----|----------------|---|---|---|----|---------|
| | | Indonesiafraudelection 2019tetapjokowi jokowi pilpres2019 nkri jokowimaruf jokowi2periode umkm ingatjokowi nawacita jokowijk salamsatu jempol prabowo debatpilpres2019 prabowosandi2019 kubuhox | 0 | 0 | 0 | Netral |
| 19 | MetroTVToday | cara yang normal adalah dengan cara memperkenalkan diri dan menargetkan pemilih mungkin yang melakukan politik uang ini orang yang menggunakan dana terbatas m qodari debatpilpres2019 metrotvIndonesiamemilih knowledgetoelevate" | 0 | 0 | 0 | Netral |
| 20 | Vincensiatan | baru tahap debat aja sudah jadi bahan pembicaraan media asing karena Gagap Unicorn" 02gagapunicorn | 0 | 0 | 0 | Netral |
| 21 | HusnahQoriatul | prabowo aja mengakui kehebatan jokowi. jokowiupgradeIndonesia 02gagapunicorn jokowilagi 01IndonesiaMaju jokowimenangprabowoNyerah | 1 | 0 | 1 | Positif |
| 22 | DianArdiantiP1 | sama sekali ngga bermanfaatJokowiUpgradeIndonesia 02GagapUnicorn jokowilagi 01IndonesiaMaju JokowiMenangPrabowoNyerah | 0 | 1 | -1 | Negatif |

| | | | | | | |
|----|-----------------|--|---|---|---|---------|
| 23 | DianSalamah | kita punya menteri dengan level dunia jokowiupgradeIndonesia jkwmerahputih 01indonesiamaju 02gagapunicorn | 1 | 0 | 1 | Positif |
| 24 | serbaserbi_indo | ternyata ada yg lagi caper nihhh yeeeeee massabayarankivlan kedaulatanmilikrakyat | 0 | 0 | 0 | Netral |
| 25 | seputarbetting | jokowi pembawa keberuntungan thefightingcock ayambangkokimport jokowiputihkangbk debatpilpres2019 | 1 | 0 | 1 | Positif |

2. Hasil Perangkat Lunak WEKA

2.1. Metode *Naïve Bayes Classifier*

```

precision

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      45240          98.4784 %
Incorrectly Classified Instances    699            1.5216 %
Kappa statistic                    0.9761
Mean absolute error                 0.0442
Root mean squared error             0.1026
Relative absolute error             10.3497 %
Root relative squared error         22.2232 %
Total Number of Instances          45939

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
          1        0.029   0.968     1      0.984     0.992    Positif
          0.959    0       1         0.959  0.979     0.994    Netral
          0.985    0       1         0.985  0.993     0.994    Negatif
Weighted Avg.   0.985    0.013   0.985     0.985  0.985     0.993

=== Confusion Matrix ===

  a    b    c  <-- classified as
21415  0    0 |  a = Positif
 525 12250  0 |  b = Netral
 174   0 11575 |  c = Negatif

```

Gambar 1 Hasil Klasifikasi Oleh Metode *Naïve Bayes Classifier* pada 5-fold Cross Validation

```

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      45276          98.5568 %
Incorrectly Classified Instances    663            1.4432 %
Kappa statistic                    0.9773
Mean absolute error                 0.0442
Root mean squared error             0.1025
Relative absolute error             10.3636 %
Root relative squared error         22.1972 %
Total Number of Instances          45939

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
          1        0.027   0.97     1      0.985     0.992    Positif
          0.962    0       1         0.962  0.981     0.994    Netral
          0.985    0       1         0.985  0.992     0.994    Negatif
Weighted Avg.   0.986    0.013   0.986     0.986  0.986     0.993

=== Confusion Matrix ===

  a    b    c  <-- classified as
21415  0    0 |  a = Positif
 488 12287  0 |  b = Netral
 175   0 11574 |  c = Negatif

```

Gambar 2 Hasil Klasifikasi Oleh Metode *Naïve Bayes Classifier* pada 10-fold Cross Validation

```

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      45245          98.4893 %
Incorrectly Classified Instances    694            1.5107 %
Kappa statistic                     0.9763
Mean absolute error                 0.0442
Root mean squared error            0.1024
Relative absolute error             10.3671 %
Root relative squared error        22.1708 %
Total Number of Instances          45939

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                1       0.028   0.969      1       0.984      0.992    Positif
                0.959   0       1         0.959   0.979      0.994    Netral
                0.985   0       1         0.985   0.993      0.994    Negatif
Weighted Avg.   0.985   0.013   0.985     0.985   0.985      0.993

=== Confusion Matrix ===

  a    b    c  <-- classified as
21415  0    0 |  a = Positif
 521 12254  0 |  b = Netral
 173  0 11576 |  c = Negatif

```

Gambar 3 Hasil Klasifikasi Oleh Metode *Naïve Bayes Classifier* pada 15-fold Cross Validation

```

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      45311          98.633 %
Incorrectly Classified Instances    628            1.367 %
Kappa statistic                     0.9785
Mean absolute error                 0.0442
Root mean squared error            0.1023
Relative absolute error             10.368 %
Root relative squared error        22.1563 %
Total Number of Instances          45939

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                1       0.026   0.972      1       0.986      0.992    Positif
                0.964   0       1         0.964   0.982      0.994    Netral
                0.985   0       1         0.985   0.993      0.994    Negatif
Weighted Avg.   0.986   0.012   0.987     0.986   0.986      0.993

=== Confusion Matrix ===

  a    b    c  <-- classified as
21415  0    0 |  a = Positif
 456 12319  0 |  b = Netral
 172  0 11577 |  c = Negatif

```

Gambar 4 Hasil Klasifikasi Oleh Metode *Naïve Bayes Classifier* pada 20-fold Cross Validation

Time taken to build model: 0.05 seconds

=== Stratified cross-validation ===

=== Summary ===

| | | |
|----------------------------------|-----------|-----------|
| Correctly Classified Instances | 45258 | 98.5176 % |
| Incorrectly Classified Instances | 681 | 1.4824 % |
| Kappa statistic | 0.9767 | |
| Mean absolute error | 0.0442 | |
| Root mean squared error | 0.1023 | |
| Relative absolute error | 10.3704 % | |
| Root relative squared error | 22.1533 % | |
| Total Number of Instances | 45939 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | ROC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|----------|---------|
| | 1 | 0.028 | 0.969 | 1 | 0.984 | 0.992 | Positif |
| | 0.96 | 0 | 1 | 0.96 | 0.98 | 0.994 | Netral |
| | 0.985 | 0 | 1 | 0.985 | 0.993 | 0.994 | Negatif |
| Weighted Avg. | 0.985 | 0.013 | 0.986 | 0.985 | 0.985 | 0.993 | |

=== Confusion Matrix ===

| | | | |
|-------|-------|-------|-------------------|
| a | b | c | <-- classified as |
| 21415 | 0 | 0 | a = Positif |
| 510 | 12265 | 0 | b = Netral |
| 171 | 0 | 11578 | c = Negatif |

Gambar 5 Hasil Klasifikasi Oleh Metode *Naïve Bayes Classifier* pada 25-fold Cross Validation

2.2. Metode *One-R*

Time taken to test model on test split: 0.76 seconds

=== Summary ===

| | | |
|----------------------------------|------------|-----------|
| Correctly Classified Instances | 20361 | 46.6546 % |
| Incorrectly Classified Instances | 23281 | 53.3454 % |
| Kappa statistic | 0 | |
| Mean absolute error | 0.3556 | |
| Root mean squared error | 0.5964 | |
| Relative absolute error | 83.1859 % | |
| Root relative squared error | 129.1199 % | |
| Total Number of Instances | 43642 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-----|----------|----------|---------|
| | 1.000 | 1.000 | 0.467 | 1.000 | 0.636 | ? | 0.500 | 0.467 | Positif |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.278 | Netral |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.256 | Negatif |
| Weighted Avg. | 0.467 | 0.467 | ? | 0.467 | ? | ? | 0.500 | 0.360 | |

=== Confusion Matrix ===

| | | | |
|-------|---|---|-------------------|
| a | b | c | <-- classified as |
| 20361 | 0 | 0 | a = Positif |
| 12118 | 0 | 0 | b = Netral |
| 11163 | 0 | 0 | c = Negatif |

Gambar 6 Hasil Klasifikasi Oleh Metode *One-R* pada 5-fold Cross Validation

Time taken to test model on test split: 0.05 seconds

=== Summary ===

| | | |
|----------------------------------|------------|-----------|
| Correctly Classified Instances | 19293 | 46.6634 % |
| Incorrectly Classified Instances | 22052 | 53.3366 % |
| Kappa statistic | 0 | |
| Mean absolute error | 0.3556 | |
| Root mean squared error | 0.5963 | |
| Relative absolute error | 83.2493 % | |
| Root relative squared error | 129.1216 % | |
| Total Number of Instances | 41345 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-----|----------|----------|---------|
| | 1.000 | 1.000 | 0.467 | 1.000 | 0.636 | ? | 0.500 | 0.467 | Positif |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.278 | Netral |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.256 | Negatif |
| Weighted Avg. | 0.467 | 0.467 | ? | 0.467 | ? | ? | 0.500 | 0.360 | |

=== Confusion Matrix ===

| a | b | c | <-- classified as |
|-------|---|---|-------------------|
| 19293 | 0 | 0 | a = Positif |
| 11478 | 0 | 0 | b = Netral |
| 10574 | 0 | 0 | c = Negatif |

Gambar 7 Hasil Klasifikasi Oleh Metode *One-R* pada *10-fold Cross Validation*

Time taken to test model on test split: 0.05 seconds

=== Summary ===

| | | |
|----------------------------------|------------|-----------|
| Correctly Classified Instances | 18194 | 46.5939 % |
| Incorrectly Classified Instances | 20854 | 53.4061 % |
| Kappa statistic | 0 | |
| Mean absolute error | 0.356 | |
| Root mean squared error | 0.5967 | |
| Relative absolute error | 83.4817 % | |
| Root relative squared error | 129.1815 % | |
| Total Number of Instances | 39048 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-----|----------|----------|---------|
| | 1.000 | 1.000 | 0.466 | 1.000 | 0.636 | ? | 0.500 | 0.466 | Positif |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.278 | Netral |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.256 | Negatif |
| Weighted Avg. | 0.466 | 0.466 | ? | 0.466 | ? | ? | 0.500 | 0.360 | |

=== Confusion Matrix ===

| a | b | c | <-- classified as |
|-------|---|---|-------------------|
| 18194 | 0 | 0 | a = Positif |
| 10854 | 0 | 0 | b = Netral |
| 10000 | 0 | 0 | c = Negatif |

Gambar 8 Hasil Klasifikasi Oleh Metode *One-R* pada *15-fold Cross Validation*

Time taken to test model on test split: 0.04 seconds

=== Summary ===

| | | |
|----------------------------------|------------|-----------|
| Correctly Classified Instances | 17111 | 46.5593 % |
| Incorrectly Classified Instances | 19640 | 53.4407 % |
| Kappa statistic | 0 | |
| Mean absolute error | 0.3563 | |
| Root mean squared error | 0.5969 | |
| Relative absolute error | 83.5547 % | |
| Root relative squared error | 129.2064 % | |
| Total Number of Instances | 36751 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-----|----------|----------|---------|
| | 1.000 | 1.000 | 0.466 | 1.000 | 0.635 | ? | 0.500 | 0.466 | Positif |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.278 | Netral |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.257 | Negatif |
| Weighted Avg. | 0.466 | 0.466 | ? | 0.466 | ? | ? | 0.500 | 0.360 | |

=== Confusion Matrix ===

| a | b | c | <-- classified as |
|-------|---|---|-------------------|
| 17111 | 0 | 0 | a = Positif |
| 10211 | 0 | 0 | b = Netral |
| 9429 | 0 | 0 | c = Negatif |



Gambar 9 Hasil Klasifikasi Oleh Metode *One-R* pada *20-fold Cross Validation*

Time taken to test model on test split: 0.04 seconds

=== Summary ===

| | | |
|----------------------------------|------------|-----------|
| Correctly Classified Instances | 16067 | 46.6332 % |
| Incorrectly Classified Instances | 18387 | 53.3668 % |
| Kappa statistic | 0 | |
| Mean absolute error | 0.3558 | |
| Root mean squared error | 0.5965 | |
| Relative absolute error | 83.3891 % | |
| Root relative squared error | 129.1378 % | |
| Total Number of Instances | 34454 | |

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class |
|---------------|---------|---------|-----------|--------|-----------|-----|----------|----------|---------|
| | 1.000 | 1.000 | 0.466 | 1.000 | 0.636 | ? | 0.500 | 0.466 | Positif |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.276 | Netral |
| | 0.000 | 0.000 | ? | 0.000 | ? | ? | 0.500 | 0.257 | Negatif |
| Weighted Avg. | 0.466 | 0.466 | ? | 0.466 | ? | ? | 0.500 | 0.360 | |

=== Confusion Matrix ===

| a | b | c | <-- classified as |
|-------|---|---|-------------------|
| 16067 | 0 | 0 | a = Positif |
| 9526 | 0 | 0 | b = Netral |
| 8861 | 0 | 0 | c = Negatif |

Gambar 10 Hasil Klasifikasi Oleh Metode *One-R* pada *25-fold Cross Validation*

2.3. Metode Zero-R

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.25 seconds

=== Summary ===

Correctly Classified Instances      20361          46.6546 %
Incorrectly Classified Instances    23281          53.3454 %
Kappa statistic                     0
Mean absolute error                 0.4275
Root mean squared error            0.4619
Relative absolute error             100 %
Root relative squared error        100 %
Total Number of Instances          43642

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          1.000   1.000   0.467     1.000   0.636     ?       0.500   0.467   Positif
          0.000   0.000   ?         0.000   ?         ?       0.500   0.278   Netral
          0.000   0.000   ?         0.000   ?         ?       0.500   0.256   Negatif
Weighted Avg.   0.467   0.467   ?         0.467   ?         ?       0.500   0.360

=== Confusion Matrix ===

      a    b    c  <-- classified as
20361  0    0 |  a = Positif
12118  0    0 |  b = Netral
11163  0    0 |  c = Negatif

```

Gambar 11 Hasil Klasifikasi Oleh Metode *Zero-R* pada *5-fold Cross Validation*

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.25 seconds

=== Summary ===

Correctly Classified Instances      19293          46.6634 %
Incorrectly Classified Instances    22052          53.3366 %
Kappa statistic                     0
Mean absolute error                 0.4271
Root mean squared error            0.4618
Relative absolute error             100 %
Root relative squared error        100 %
Total Number of Instances          41345

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          1.000   1.000   0.467     1.000   0.636     ?       0.500   0.467   Positif
          0.000   0.000   ?         0.000   ?         ?       0.500   0.278   Netral
          0.000   0.000   ?         0.000   ?         ?       0.500   0.256   Negatif
Weighted Avg.   0.467   0.467   ?         0.467   ?         ?       0.500   0.360

=== Confusion Matrix ===

      a    b    c  <-- classified as
19293  0    0 |  a = Positif
11478  0    0 |  b = Netral
10574  0    0 |  c = Negatif

```

Gambar 12 Hasil Klasifikasi Oleh Metode *Zero-R* pada *10-fold Cross Validation*

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.05 seconds

=== Summary ===

Correctly Classified Instances      18194          46.5939 %
Incorrectly Classified Instances    20854          53.4061 %
Kappa statistic                    0
Mean absolute error                 0.4265
Root mean squared error             0.4619
Relative absolute error              100           %
Root relative squared error         100           %
Total Number of Instances          39048

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          1.000    1.000    0.466     1.000    0.636     ?       0.500    0.466    Positif
          0.000    0.000    ?         0.000    ?         ?       0.500    0.278    Netral
          0.000    0.000    ?         0.000    ?         ?       0.500    0.256    Negatif
Weighted Avg.   0.466    0.466    ?         0.466    ?         ?       0.500    0.360

=== Confusion Matrix ===

  a    b    c  <-- classified as
18194  0    0 |  a = Positif
10854  0    0 |  b = Netral
10000  0    0 |  c = Negatif

```

Gambar 13 Hasil Klasifikasi Oleh Metode *Zero-R* pada *15-fold Cross Validation*

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.03 seconds

=== Summary ===

Correctly Classified Instances      17111          46.5593 %
Incorrectly Classified Instances    19640          53.4407 %
Kappa statistic                    0
Mean absolute error                 0.4264
Root mean squared error             0.462
Relative absolute error              100           %
Root relative squared error         100           %
Total Number of Instances          36751

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          1.000    1.000    0.466     1.000    0.635     ?       0.500    0.466    Positif
          0.000    0.000    ?         0.000    ?         ?       0.500    0.278    Netral
          0.000    0.000    ?         0.000    ?         ?       0.500    0.257    Negatif
Weighted Avg.   0.466    0.466    ?         0.466    ?         ?       0.500    0.360

=== Confusion Matrix ===

  a    b    c  <-- classified as
17111  0    0 |  a = Positif
10211  0    0 |  b = Netral
 9429  0    0 |  c = Negatif

```

Gambar 14 Hasil Klasifikasi Oleh Metode *Zero-R* pada *20-fold Cross Validation*

```

=== Evaluation on test split ===

Time taken to test model on test split: 0.03 seconds

=== Summary ===

Correctly Classified Instances      16067           46.6332 %
Incorrectly Classified Instances    18387           53.3668 %
Kappa statistic                     0
Mean absolute error                 0.4266
Root mean squared error             0.4619
Relative absolute error              100 %
Root relative squared error          100 %
Total Number of Instances          34454

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
1.000    1.000    0.466    1.000    0.636    ?      0.500    0.466    Positif
0.000    0.000    ?         0.000    ?         ?      0.500    0.276    Netral
0.000    0.000    ?         0.000    ?         ?      0.500    0.257    Negatif
Weighted Avg.  0.466    0.466    ?         0.466    ?         ?      0.500    0.360

=== Confusion Matrix ===

  a    b    c  <-- classified as
16067  0    0 |  a = Positif
 9526  0    0 |  b = Netral
 8861  0    0 |  c = Negatif

```

Gambar 15 Hasil Klasifikasi Oleh Metode *Zero-R* pada *25-fold Cross Validation*

3. Gambar code yang digunakan dalam penelitian

3.1. Code pada proses pengumpulan data *tweet*

```

# -*- coding: utf-8 -*-
import sys, getopt, datetime, codecs
if sys.version_info[0] < 3:
    import got
else:
    import got3 as got

def main(argv):

    if len(argv) == 0:
        print('You must pass some parameters. Use \'-h\' to help.')
        return

    if len(argv) == 1 and argv[0] == '-h':
        f = open('exporter_help_text.txt', 'r')
        print f.read()
        f.close()

        return

    try:
        opts, args = getopt.getopt(argv, "",
            ("username=", "near=", "within=", "since=", "until=",
            "querysearch=", "toptweets", "maxtweets=", "output="))

        tweetCriteria = got.manager.TweetCriteria()
        outputFileName = "output_got.csv"

        for opt, arg in opts:
            if opt == '--username':
                tweetCriteria.username = arg

```

Gambar 16 Code Untuk Pengumpulan Data *Tweet* Menggunakan Program *GetOldTweets-python-master* Pada File *Exporter.py*


```
for opt, arg in opts:
    if opt == '--username':
        tweetCriteria.username = arg

    elif opt == '--since':
        tweetCriteria.since = arg

    elif opt == '--until':
        tweetCriteria.until = arg

    elif opt == '--querysearch':
        tweetCriteria.querySearch = arg

    elif opt == '--toptweets':
        tweetCriteria.topTweets = True

    elif opt == '--maxtweets':
        tweetCriteria.maxTweets = int(arg)

    elif opt == '--near':
        tweetCriteria.near = "" + arg + ""

    elif opt == '--within':
        tweetCriteria.within = "" + arg + ""

    elif opt == '--within':
        tweetCriteria.within = "" + arg + ""

    elif opt == '--output':
        outputFileName = arg
```

Gambar 17 Code Untuk Pengumpulan Data *Tweet* Menggunakan Program *GetOldTweets-python-master* Pada File *Exporter.py*

```

        elif opt == '--output':
            outputFileName = arg

outputFile = codecs.open(outputFileName, "w+", "utf-8")

outputFile.write
('username;date;retweets;favorites;text;geo;mentions;hashtags;id;permalink')

print('Searching...\n')

def receiveBuffer(tweets):
    for t in tweets:
        outputFile.write((' \n%s;%s;%d;%d;"%s";%s;%s;%s;"%s";%s' %
            (t.username, t.date.strftime("%Y-%m-%d %H:%M"),
            t.retweets, t.favorites, t.text, t.geo,
            t.mentions, t.hashtags, t.id, t.permalink)))
        outputFile.flush()
        print('More %d saved on file...\n' % len(tweets))

got.manager.TweetManager.getTweets(tweetCriteria, receiveBuffer)

except arg:
    print('Arguments parser error, try -h' + arg)
finally:
    outputFile.close()
    print('Done. Output file generated "%s".' % outputFileName)

if __name__ == '__main__':
    main(sys.argv[1:])

```

Gambar 18 Code Untuk Pengumpulan Data *Tweet* Menggunakan Program *GetOldTweets-python-master* Pada File *Exporter.py*

```

import sys
if sys.version_info[0] < 3:
    import got
else:
    import got3 as got

def main():
    def printTweet(descr, t):
        print(descr)
        print("Username: %s" % t.username)
        print("Retweets: %d" % t.retweets)
        print("Text: %s" % t.text)
        print("Mentions: %s" % t.mentions)
        print("Hashtags: %s\n" % t.hashtags)
        # Example 1 - Get tweets by username
        tweetCriteria = got.manager.TweetCriteria()
        .setUsername('barackobama').setMaxTweets(1)
        tweet = got.manager.TweetManager.getTweets(tweetCriteria) [0]
        printTweet("### Example 1 - Get tweets by username [barackobama]", tweet)
        # Example 2 - Get tweets by query search
        tweetCriteria = got.manager.TweetCriteria().setQuerySearch
        ('europe refugees').setSince("2015-05-01").
        setUntil("2015-09-30").setMaxTweets(1)
        tweet = got.manager.TweetManager.getTweets(tweetCriteria) [0]
        printTweet("### Example 2 - Get tweets by query search [europe refugees]", tweet)
        # Example 3 - Get tweets by username and bound dates
        tweetCriteria = got.manager.TweetCriteria().setUsername
        ('barackobama').setSince("2015-09-10").
        setUntil("2015-09-12").setMaxTweets(1)
        tweet = got.manager.TweetManager.getTweets(tweetCriteria) [0]

        printTweet("### Example 3 - Get tweets by username and bound dates [barackobama, '2015-09-10', '2015-09-12']", tweet)

if __name__ == '__main__':
    main()

```

Gambar 19 Code Untuk Pengumpulan Data *Tweet* Menggunakan Program *GetOldTweets-python-master* Pada File *Main.py*

```
C:\Windows\system32\cmd.exe - python Exporter.py --querysearch "bukalapak" --since 2019-03-01 --until 2019-03-31 --maxtweets 100000
Microsoft Windows [Version 10.0.15063]
(c) 2017 Microsoft Corporation. All rights reserved.

C:\Users\asus>D:

D:\>cd D:\GetOldTweets-python-master
D:\GetOldTweets-python-master>python Exporter.py --querysearch "bukalapak" --since 2019-03-01 --until 2019-03-31 --maxtweets 100000
Searching...

More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
More 100 saved on file...
```

Gambar 120 proses pengambilan data tweet menggunakan Command Prompt untuk menjalankan program *GetOldTweets-python-master*

3.2. Code pada proses *Cleansing*

```
setwd("F:/")
mydata <- read.csv(file.choose(),header=T)
class(mydata)
library("NLP")
library("tm")
df<- do.call("rbind", lapply(mydata$status,as.data.frame))
df<- sapply(df,function(row) iconv(row,"latin1","ASCII",sub=""))

df <- gsub("\\\\w+(?:\\.\\.\\w+)*\\/\\s+", "", df)

df=gsub("http[^:a1num:]*", "", df)

df=gsub("www[[:a1num:][:punct:]]*", "", df)

df=gsub("[^0-9A-ZA-z//]", " ", df)

df=gsub("[0-9]", " ", df)

df=gsub("http[^:a1num:]*", " ", df)

df<- stripwhitespace(df)

df<- tolower(df)

write.table(df,file="cobaDatacleansing.csv",sep=" ",row.names=F)
getwd()
```

Gambar 21 Code Untuk Proses *Cleansing* Menggunakan *Library NLP* Dalam Bahasa Pemrograman R

3.3.Code pada proses *Stopwords*

```
package com.uttesh.exude.stemming;

import com.uttesh.exude.ExudeData;
import com.uttesh.exude.exception.InvalidDataException;
import static com.uttesh.exude.stemming.Stemmer.c;
import static com.uttesh.exude.stemming.Stemmer.path;
import static com.uttesh.exude.stemming.Stemmer.url;
import java.io.File;
import java.io.FileWriter;
import java.io.IOException;
import java.sql.DriverManager;
import java.sql.PreparedStatement;
import java.sql.ResultSet;
import java.sql.Statement;

/**...4 lines */
public class Stopwords {

    /**...3 lines */
    public static void main(String[] args) throws InvalidDataException {
        // TODO code application logic here
        String sql = "SELECT * FROM Twitter ";
        System.out.println("Sedang diproses....");
        try
        {
            c=DriverManager.getConnection(url+path);
            System.out.println("Berhasil konekk database");
            Statement state = c.createStatement();
            ResultSet rs = state.executeQuery(sql);
            if(rs!=null)
            {
                while(rs.next())
```

Gambar 22 Code Untuk Proses *Stopwords* Menggunakan Bahasa Pemrograman *Java NetBeans*

```

while(rs.next())
{
    int id = rs.getInt("ID");
    String username= rs.getString("username");
    String inputData = rs.getString("Status");
    String tanggal = rs.getString("tanggal");
    int retweet = rs.getInt("retweet");
    int favorites=rs.getInt("favorites");
    String mention = rs.getString("mentions");
    String hastags =rs.getString("hastags");
    String geo_location=rs.getString("geo_location");

    String output = ExudeData.getInstance().filterStoppingsKeepDuplicates(inputData);
    //String sql2="UPDATE Stopwords set Status = ? where ID=?"; //versi edit
    String sql2 = "INSERT into stopwords values(?,?,?,?,?,?,?,?)"; //versi insert
    System.out.println("output ID "+id+ " : "+output);
    PreparedStatement pStmt = c.prepareStatement(sql2);
    // pStmt.setInt(1, c.GetRowDataTwitter());
    pStmt.setInt(1, id);
    pStmt.setString(2, username);
    pStmt.setString(3, output);
    pStmt.setString(4, tanggal);
    pStmt.setInt(5, retweet);
    pStmt.setInt(6, favorites);
    pStmt.setString(7, mention);
    pStmt.setString(8, hastags);
    pStmt.setString(9, geo_location);

    pStmt.executeUpdate();
}
}
rs.close();

```

Gambar 23 Code Untuk Proses *Stopwords* Menggunakan Bahasa Pemrograman *Java NetBeans*

```

    rs.close();
    state.close();
    c.close();
    System.out.println("database ditutup");
}
catch(Exception EX)
{
    System.out.println("Error Reading From database. . .");
    System.out.println(EX);
}
}
}

```

Gambar 24 Code Untuk Proses *Stopwords* Menggunakan Bahasa Pemrograman *Java NetBeans*

3.4. Code pada proses *Stemming*

```
package jsastrawi.cli;

import java.io.BufferedReader;
import java.io.File;
import java.io.FileReader;
import java.io.FileWriter;
import jsastrawi.cli.output.Output;
import jsastrawi.cli.output.BufferedOutput;
import java.io.IOException;
import java.io.InputStream;
import java.io.InputStreamReader;
import java.sql.Connection;
import java.util.ArrayList;
import java.util.Arrays;
import java.util.HashSet;
import java.util.List;
import java.util.Set;
import java.util.StringTokenizer;
import jdk.nashorn.internal.parser.TokenStream;
import jsastrawi.morphology.DefaultLemmatizer;
import jsastrawi.morphology.Lemmatizer;
import jsastrawi.tokenization.HeuristicTokenizer;
import java.io.IOException;
import java.sql.DriverManager;
import java.sql.PreparedStatement;
import java.sql.ResultSet;
import java.sql.Statement;

/**
 * The main class for Command Line Interface.
 */
public class Main
```

Gambar 25 Code Untuk Proses *Stemming* Menggunakan Bahasa Pemrograman *Java NetBeans*

```

public class Main
{
    private static HeuristicTokenizer tokenizer;
    public static Connection c;
    public static final String driver = "sun.jdbc.odbc.JdbcOdbcDriver";

    public static String url = "jdbc:ucanaccess://";
    //public static final String path = "D:\\Tweet\\dbGetDataTwitter.mdb";
    public static final String path = "F:\\preprocessing jav - Copy.mdb";

    public void setUp()
    {
        tokenizer = new HeuristicTokenizer();
    }

    /** The main entry point function which is called first from the CLI ...6 lines */
    public static void main(String[] args) throws IOException
    {
        Set<String> dictionary = new HashSet<String>();

        // Memuat file kata dasar dari distribusi JSastrawi
        // Jika perlu, anda dapat mengganti file ini dengan kamus anda sendiri
        InputStream in = Lemmatizer.class.getResourceAsStream("/root-words.txt");
        BufferedReader br = new BufferedReader(new InputStreamReader(in));

        String line;
        while ((line = br.readLine()) != null)
        {
            dictionary.add(line);
        }

        Lemmatizer lemmatizer = new DefaultLemmatizer(dictionary);
    }
}

```

Gambar 26 Code Untuk Proses *Stemming* Menggunakan Bahasa Pemrograman *Java NetBeans*

```

Lemmatizer lemmatizer = new DefaultLemmatizer(dictionary);
// Selesai setup JSastrawi
// lemmatizer bisa digunakan berkali-kali
String sql = "SELECT * FROM Stopwords ";
System.out.println("Sedang diproses....");
try
{
    c = DriverManager.getConnection(url + path);
    System.out.println("Berhasil konekk database");
    Statement state = c.createStatement();
    ResultSet rs = state.executeQuery(sql);
    if (rs != null)
    {
        while (rs.next())
        {
            int id = rs.getInt("ID");
            String username = rs.getString("username");
            String inputData = rs.getString("Status");
            String tanggal = rs.getString("tanggal");
            int retweet = rs.getInt("retweet");
            int favorites = rs.getInt("favorites");
            String mention = rs.getString("mentions");
            String hastags = rs.getString("hastags");
            String geo_location = rs.getString("geo_location");

            StringTokenizer st = new StringTokenizer(inputData);
            System.out.println("---- Split by space -----");
            StringBuilder sb = new StringBuilder();
            String temp = "";
            while (st.hasMoreElements())
            {
                //System.out.println(lemmatizer.lemmatize(st.nextToken()));
                temp = temp + lemmatizer.lemmatize(st.nextToken()) + " ";
            }
        }
    }
}

```

Gambar 27 Code Untuk Proses *Stemming* Menggunakan Bahasa Pemrograman *Java NetBeans*


```

//System.out.println(lemmatizer.lemmatize(st.nextToken()));
temp = temp + lemmatizer.lemmatize(st.nextToken()) + " ";

//sb.append( lemmatizer.lemmatize(st.nextToken())+" ");
}
//String sql2="UPDATE Stopwords set Status = ? where ID =?"; //versi edit
String sql2 = "INSERT into stemming values(?,?,?,?); //versi insert
System.out.println("output ID " + id + " : " + temp);
PreparedStatement pStmt = c.prepareStatement(sql2);
// pStmt.setInt(1, c.getRowDataTwitter());
pStmt.setInt(1, id);
pStmt.setString(2, username);
pStmt.setString(3, temp);
pStmt.setString(4, tanggal);
pStmt.setInt(5, retweet);
pStmt.setInt(6, favorites);
pStmt.setString(7, mention);
pStmt.setString(8, hastags);
pStmt.setString(9, geo_location);

pStmt.executeUpdate();
}
}
rs.close();
state.close();
c.close();
System.out.println("database ditutup");
}
catch (Exception EX)

```

Gambar 28 Code Untuk Proses *Stemming* Menggunakan Bahasa Pemrograman *Java NetBeans*

```

catch (Exception EX)
{
    System.out.println("Error Reading From database. . .");
    System.out.println(EX);
}
if (args.length == 0)
{
    printHelp();
}
else if (args[0].toLowerCase().equals("lemmatize"))
{
    Output bufferedOutput = new BufferedOutput();
    LemmatizeCmd lemmatizeCmd = new LemmatizeCmd(bufferedOutput);
    lemmatizeCmd.handle(removeCommandFromArgs(args));
    System.out.print(bufferedOutput.toString());
}
else
{
    printHelp();
}
}

/**
 * Print Command Line usage
 */
private static void printHelp()
{
    System.out.println("JSastrawi");
    System.out.println("usage: command [arguments]\n");
    System.out.println("Available commands:");
    System.out.println("lemmatize Determbine the lemma (base form) for a given word.");
}

```

Gambar 29 Code Untuk Proses *Stemming* Menggunakan Bahasa Pemrograman *Java NetBeans*

```

System.out.println("lemmatize Determinine the lemma (base form) for a given word.");
}

/** Remove the command (which is the first argument) from an array of ...7 lines */
static String[] removeCommandFromArgs(String[] args)
{
    List<String> largs = new ArrayList<>(Arrays.asList(args));
    if (largs.size() > 0)
    {
        largs.remove(0);
    }
    return largs.toArray(new String[0]);
}
}
}

```

Gambar 30 Code Untuk Proses *Stemming* Menggunakan Bahasa Pemrograman *Java NetBeans*

