

**KLASIFIKASI SENTIMEN KOMENTAR INSTAGRAM
BERBAHASA INDONESIA MENGGUNAKAN SUPPORT
VECTOR MACHINE**

Skripsi



Diajukan oleh:

ADIEL JULIO

71150017

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
YOGYAKARTA

2019

**KLASIFIKASI SENTIMEN KOMENTAR INSTAGRAM
BERBAHASA INDONESIA MENGGUNAKAN SUPPORT
VECTOR MACHINE**

Skripsi



Diajukan kepada Fakultas Teknologi Informasi Program Studi Informatika

Universitas Kristen Duta Wacana

Sebagai Salah Satu Syarat dalam Memperoleh Gelar Sarjana Komputer

Diajukan oleh:

ADIEL JULIO

71150017

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS KRISTEN DUTA WACANA YOGYAKARTA

2019

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS
SECARA ONLINE UNIVERSITAS KRISTEN DUTA WACANA**

Saya yang bertanda tangan di bawah ini:

NIM : 71150017
Nama : Adiel Julio
Prodi / Fakultas : Informatika / Teknologi Informasi
Judul Tugas Akhir : Klasifikasi Sentimen Komentar Instagram
Berbahasa Indonesia Menggunakan Support
Vector Machine

bersedia menyerahkan Tugas Akhir kepada Universitas melalui Perpustakaan untuk keperluan akademis dan memberikan **Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalty-free Right*)** serta bersedia Tugas Akhirnya dipublikasikan secara online dan dapat diakses secara lengkap (*full access*).

Dengan Hak Bebas Royalti Noneksklusif ini Perpustakaan Universitas Kristen Duta Wacana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk database, merawat, dan memublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenar-benarnya.

Yogyakarta, 25 Juni 2019

Yang menyatakan,

Adiel Julio

71150017

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

KLASIFIKASI SENTIMEN KOMENTAR INSTAGRAM BERBAHASA INDONESIA MENGGUNAKAN SUPPORT VECTOR MACHINE

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 24 Juni 2019



ADIEL JULIO

71150017

HALAMAN PERSETUJUAN

Judul Skripsi : KLASIFIKASI SENTIMEN KOMENTAR
INSTAGRAM BERBAHASA INDONESIA
MENGUNAKAN SUPPORT VECTOR MACHINE

Nama Mahasiswa : ADIEL JULIO

N I M : 71150017

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Genap

Tahun Akademik : 2018/2019

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 24 Juni 2019

Dosen Pembimbing I



Danny Sebastian, S.Kom., M.M., M.T.

Dosen Pembimbing II



Kristian Adi Nugraha, S.Kom., M.T.

HALAMAN PENGESAHAN

KLASIFIKASI SENTIMEN KOMENTAR INSTAGRAM BERBAHASA INDONESIA MENGGUNAKAN SUPPORT VECTOR MACHINE

Oleh: ADIEL JULIO / 71150017

Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 12 Juni 2019


Yogyakarta, 24 Juni 2019
Mengesahkan,

Dewan Penguji:


1. Danny Sebastian, S.Kom., M.M., M.T.
2. Kristian Adi Nugraha, S.Kom., M.T.
3. Antonius Rachmat C., S.Kom., M.Cs.
4. Aditya Wikan Mahastama, S.Kom., M.Cs.



Dekan


(Bidi Susanto, S.Kom., M.T.)

Ketua Program Studi,


(Gloria Virginia, Ph.D.)

UCAPAN TERIMA KASIH

Dalam penyusunan skripsi ini tidak terlepas dukungan dari berbagai pihak. Peneliti secara khusus mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu. Peneliti banyak menerima bimbingan, petunjuk dan bantuan serta dorongan dari berbagai pihak baik yang bersifat moral maupun material. Pada kesempatan ini penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Tuhan dengan segala rahmat serta karunia-Nya yang memberikan kekuatan bagi peneliti dalam menyelesaikan skripsi ini.
2. Kepada kedua orang tua tercinta yang selama ini telah membantu peneliti dalam bentuk perhatian, kasih sayang, semangat, serta doa demi kelancaran dan kesuksesan peneliti dalam menyelesaikan skripsi ini. Kemudian terima kasih banyak untuk kakak tercinta Setia Budi dan Novita Triana Putri yang telah memberikan dukungan serta perhatian kepada peneliti.
3. Kepada Bapak Danny Sebastian, S.Kom., M.M., M.T. dan Bapak Kristian Adi Nugraha, S.Kom., M.T. selaku dosen pembimbing yang selalu memberikan bimbingan, arahan, dorongan, dan semangat kepada peneliti, sehingga skripsi ini dapat terselesaikan.
4. Kepada Bapak Budi Susanto, S.Kom., M.T. selaku dekan Fakultas Teknologi Informasi Universitas Kristen Duta Wacana.
5. Segenap dosen dan seluruh staf akademik yang selalu membantu dalam memberikan fasilitas, ilmu, serta pendidikan pada peneliti hingga dapat menunjang dalam penyelesaian skripsi ini.
6. Kepada Aldo Saputra, Nathania Saphira, Joshua Gibeon, Setya Budi, dan Adrean Raphael yang telah membantu dalam proses pelabelan data.
7. Teman-teman yang telah memberikan banyak masukan serta dukungan kepada peneliti.

Serta masih banyak lagi pihak-pihak yang sangat berpengaruh dalam proses penyelesaian skripsi yang yang tidak bisa peneliti sebutkan satu persatu. Semoga

Tuhan senantiasa membalas semua kebaikan yang telah diberikan yang telah diberikan. Semoga penelitian ini dapat bermanfaat bagi peneliti umumnya kepada para pembaca.

Yogyakarta, 25 Juni 2019.....

Adiel Julio

©UKDWN

INTISARI

Klasifikasi Sentimen Komentar Instagram Berbahasa Indonesia Menggunakan Support Vector Machine

Media sosial adalah salah satu kebutuhan wajib bagi sebagian besar kalangan masyarakat. Salah satu media sosial yang populer adalah Instagram. Namun dari manfaat penggunaan Instagram tersebut, masih banyak pengguna yang belum memahami etika dalam bersosialisasi pada dunia maya. Hal tersebut tentunya menjadi salah satu peringatan bagi pengguna agar dapat mengurangi perbuatan yang merugikan tersebut.

Dari masalah tersebut, analisis sentimen sangat diperlukan dalam menyaring komentar di media sosial, untuk mengetahui komentar yang bersifat negatif, positif, maupun netral. Dari analisis sentimen tersebut kita dapat mengetahui bagaimana sentimen seseorang berlaku terhadap akun tersebut. Dalam melakukan analisis sentimen diperlukan metode klasifikasi yang menunjang.

Dalam penelitian ini, Penulis akan membuat sistem yang menerapkan analisis sentimen terhadap komentar di Instagram dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM). Penulis meneliti akurasi metode SVM dalam mengklasifikasi sentimen. Penulis juga meneliti faktor yang mempengaruhi tingkat akurasi. Faktor yang akan diteliti berupa jumlah k dalam k -fold, penggunaan kernel, dan penggunaan stemmer dalam *preprocessing*.

Berdasarkan hasil penelitian Penulis, didapatkan rata-rata akurasi terbaik dengan fitur kombinasi unigram dan bigram sebesar 76.307 %. Serta beberapa pengujian mendapatkan kesimpulan bahwa kernel polynomial kurang sesuai dalam mengklasifikasi dengan fitur *tf-idf*. Selain itu, penggunaan stemmer cukup mempengaruhi hasil pada kasus tertentu.

Kata kunci: [*Support Vector Machine*, Analisis Sentimen, *k-fold cross-validation*, *tf-idf*, stemmer, n-gram]

DAFTAR ISI

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS SECARA ONLINE UNIVERSITAS KRISTEN DUTA WACANA.....	ii
PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
UCAPAN TERIMA KASIH.....	iv
INTISARI.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metode / Pendekatan.....	3
1.6.1 Tahap Studi Literatur.....	3
1.6.2 Tahap Pengumpulan Data.....	3
1.6.3 Tahap Perancangan Sistem.....	4
1.6.4 Tahap Validasi dan Evaluasi.....	4
1.6.5 Tahap Penarikan Kesimpulan.....	4
1.7 Sistematika Penulisan.....	4

BAB 2	TINJAUAN PUSTAKA	6
2.1	Tinjauan Pustaka	6
2.2	Landasan Teori	7
2.2.1	Text Mining.....	7
2.2.2	Analisis Sentimen	7
2.2.3	Stopword Removal.....	8
2.2.4	Stemming	8
2.2.5	TF-IDF	8
2.2.6	Support Vector Machine	9
2.2.7	Validasi	16
2.2.8	Confusion Matrix	17
BAB 3	PERANCANGAN SISTEM	20
3.1	Analisis Kebutuhan Sistem	20
3.1.1	Kebutuhan Perangkat Lunak	20
3.1.2	Kebutuhan Perangkat Keras	20
3.2	Arsitektur Sistem	20
3.3	Sistem Pengambilan Data.....	21
3.4	Sistem Pelabelan.....	21
3.5	Sistem Klasifikasi.....	22
3.5.1	Preprocessing	22
3.5.2	Ekstraksi Fitur	23
3.5.3	Klasifikasi	23
3.5.4	Validasi dan Evaluasi.....	24
BAB 4	IMPLEMENTASI DAN ANALISIS SISTEM	25
4.1	Implementasi Sistem Pengambilan Data.....	25

4.2	Implementasi Sistem Pelabelan.....	26
4.3	Implementasi Sistem Klasifikasi.....	27
4.3.1	Tampilan Antarmuka	27
4.3.2	Input	27
4.3.3	Preprocessing	28
4.3.4	Ekstraksi fitur	29
4.3.5	Klasifikasi	29
4.3.6	Validasi dan Evaluasi.....	30
4.4	Pengujian Sistem	31
4.5	Analisis Sistem	32
BAB 5	KESIMPULAN DAN SARAN	39
5.1	Kesimpulan.....	39
5.2	Saran.....	39
	DAFTAR PUSTAKA	40
	LAMPIRAN.....	42
LAMPIRAN A	Listing Program Utama.....	42
LAMPIRAN B	Listing Sistem Klasifikasi.....	56
LAMPIRAN C	Kartu Konsultasi	71
LAMPIRAN D	Lembar Perbaikan Skripsi.....	73

DAFTAR TABEL

Tabel 2.1. Pembobotan tf-idf	14
Tabel 2.2. Persamaan Support Vector Machine.....	14
Tabel 2.3. Confusion Matrix.....	17
Tabel 2.4. Contoh data confusion matrix.....	18
Tabel 4.1. Hasil Pengujian k-fold	32
Tabel 4.2. Hasil Pengujian data positif dan negatif dengan tf-idf unigram	33
Tabel 4.3. Hasil Pengujian data positif dan negatif dengan tf-idf unigram dan bigram	34
Tabel 4.4. Hasil Pengujian data positif dan negatif dengan tf-idf bigram	35
Tabel 4.5. Frekuensi kemunculan fitur pada komentar.....	37
Tabel 4.6. Beberapa fitur yang digunakan	37
Tabel 4.7. Hasil pengujian data positif dan negatif untuk validasi klasifikasi.....	38

DAFTAR GAMBAR

Gambar 2.1. Contoh beberapa hyperplanes	9
Gambar 2.2. Maximum-margin hyperplane.....	10
Gambar 2.3. Contoh iterasi k-fold cross-validation	16
Gambar 3.1. Arsitektur Sistem.....	20
Gambar 3.2. Sistem Pelabelan	21
Gambar 3.3. Entity Relationship Diagram (ERD) Sistem Pelabelan.....	22
Gambar 3.4. Diagram sistem.....	22
Gambar 3.5. Diagram Alur Preprocessing	23
Gambar 4.1. Format Hasil.....	25
Gambar 4.2. Pseudocode Sistem Pengambilan Data	26
Gambar 4.3. Antarmuka Sistem Pelabelan	26
Gambar 4.4. Tampilan antarmuka.....	27
Gambar 4.5. Data input	28
Gambar 4.6. Pseudocode Preprocessing	29
Gambar 4.7. Pseudocode Training	30
Gambar 4.8. Pseudocode Validasi dan Evaluasi	31
Gambar 4.9. Perbandingan hasil antara fitur.....	36

INTISARI

Klasifikasi Sentimen Komentar Instagram Berbahasa Indonesia Menggunakan Support Vector Machine

Media sosial adalah salah satu kebutuhan wajib bagi sebagian besar kalangan masyarakat. Salah satu media sosial yang populer adalah Instagram. Namun dari manfaat penggunaan Instagram tersebut, masih banyak pengguna yang belum memahami etika dalam bersosialisasi pada dunia maya. Hal tersebut tentunya menjadi salah satu peringatan bagi pengguna agar dapat mengurangi perbuatan yang merugikan tersebut.

Dari masalah tersebut, analisis sentimen sangat diperlukan dalam menyaring komentar di media sosial, untuk mengetahui komentar yang bersifat negatif, positif, maupun netral. Dari analisis sentimen tersebut kita dapat mengetahui bagaimana sentimen seseorang berlaku terhadap akun tersebut. Dalam melakukan analisis sentimen diperlukan metode klasifikasi yang menunjang.

Dalam penelitian ini, Penulis akan membuat sistem yang menerapkan analisis sentimen terhadap komentar di Instagram dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM). Penulis meneliti akurasi metode SVM dalam mengklasifikasi sentimen. Penulis juga meneliti faktor yang mempengaruhi tingkat akurasi. Faktor yang akan diteliti berupa jumlah k dalam k -fold, penggunaan kernel, dan penggunaan stemmer dalam *preprocessing*.

Berdasarkan hasil penelitian Penulis, didapatkan rata-rata akurasi terbaik dengan fitur kombinasi unigram dan bigram sebesar 76.307 %. Serta beberapa pengujian mendapatkan kesimpulan bahwa kernel polynomial kurang sesuai dalam mengklasifikasi dengan fitur *tf-idf*. Selain itu, penggunaan stemmer cukup mempengaruhi hasil pada kasus tertentu.

Kata kunci: [*Support Vector Machine*, Analisis Sentimen, *k-fold cross-validation*, *tf-idf*, stemmer, n-gram]

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam dunia yang berkembang secara pesat seperti sekarang ini, teknologi membuat jarak tidak menjadi masalah dalam berkomunikasi. Media sosial adalah salah satu kebutuhan wajib bagi sebagian besar kalangan masyarakat. Pengguna media sosial tidak hanya bertambah, tetapi semakin banyak pula jenis dari media sosial yang ditawarkan.

Beberapa jenis media sosial yang ditawarkan saat ini adalah Instagram, Facebook, Twitter, LINE, Path, LinkedIn, dan masih banyak media sosial lainnya. Komunikasi yang lebih terbuka dan tak terbatas merupakan beberapa keunggulan dari media sosial. Pada media sosial, masyarakat dapat berpendapat ataupun membagikan cerita tentang apa yang mereka rasakan dalam kehidupannya.

Berdasarkan KOMINFO (2018) jumlah pengguna internet di Indonesia mencapai 143.26 juta orang. Angka tersebut setara dengan 54,68 persen dari total jumlah penduduk di Indonesia. Dari 143.26 juta orang ini sebagian besar mengakses media sosial. Media sosial yang populer di Indonesia salah satunya adalah Instagram. Instagram merupakan media sosial yang pada dasarnya berfungsi untuk berbagi foto maupun video pada sesama pengguna. Pengguna Instagram tentunya tidak hanya orang dewasa namun remaja dan anak-anak juga ikut menggunakan media sosial ini.

Pengguna menggunakan Instagram sebagai akun pribadi baik untuk orang biasa maupun artis, hingga menjadi sarana bisnis perseorangan. Namun dari manfaat penggunaan Instagram tersebut, masih banyak pengguna yang belum memahami etika dalam bersosialisasi pada dunia maya. Banyak terjadi kasus dimana beberapa pengguna berkomentar yang tidak sopan hingga menghina pengguna lain yang tidak mereka sukai. Hal tersebut tentunya menjadi salah satu peringatan bagi pengguna maupun pemerintah agar dapat mengurangi perbuatan yang merugikan tersebut.

Analisis sentimen sangat diperlukan dalam menyaring komentar di media sosial, untuk mengetahui komentar yang bersifat negatif, positif, maupun netral. Dari analisis sentimen tersebut kita dapat mengetahui bagaimana sentimen seseorang berlaku terhadap akun tersebut. Dalam melakukan analisis sentimen diperlukan metode klasifikasi yang menunjang.

Berdasarkan hasil penelitian analisis sentimen yang sebelumnya dilakukan oleh Rachmat C. dan Lukito (2016) untuk analisis sentimen pada Facebook, dengan mengimplementasikan metode Naïve Bayes dan Support Vector Machine, menghasilkan akurasi masing-masing 82.23% dan 84.82%. Sementara berdasarkan penelitian yang dilakukan oleh Jadav dan Vaghela (2016) dengan mengimplementasikan SVM untuk Twitter menghasilkan akurasi sebesar 76.92%. Angka ini membuktikan bahwa SVM memiliki tingkat akurasi yang tinggi untuk digunakan pada sosial media. Serta penelitian yang dilakukan oleh Rachmat C. dan Lukito (2017) untuk identifikasi spam pada komentar Instagram, dengan mengimplementasikan metode Naïve Bayes dan Support Vector Machine menghasilkan akurasi tertinggi sebesar 76.1% untuk Naïve Bayes dan 78.30% untuk Support Vector Machine yang menggunakan kernel RBF. Penelitian tersebut belum meneliti kernel lain yang dapat digunakan pada Support Vector Machine. Maka penelitian ini akan meneliti penggunaan kernel linear dan polynomial pada metode Support Vector Machine untuk klasifikasi pada komentar Instagram.

1.2 Perumusan Masalah

Berdasarkan penjabaran latar belakang di atas, rumusan masalah pada penelitian ini adalah:

- a) Bagaimana menerapkan metode *Support Vector Machine* menggunakan kernel *linear* dan *polynomial* untuk melakukan klasifikasi sentimen pada data komentar Instagram.
- b) Bagaimana akurasi metode *Support Vector Machine* menggunakan kernel *linear* dan *polynomial* dalam melakukan klasifikasi sentimen pada data komentar Instagram.

1.3 Batasan Masalah

Dalam penelitian kali ini, yang tidak akan dilakukan penulis adalah :

- a) Data set yang akan digunakan akan sejumlah 3000.
- b) Mengambil data dari akun Instagram brand smarphone di Indonesia.
- c) *Stemmer* akan menggunakan *library* sastrawi.
- d) Mengabaikan komentar netral, sarkasme dan *spam*.
- e) Tidak sampai semantik

1.4 Tujuan Penelitian

Dalam penelitian ini, tujuan yang ingin dicapai oleh penulis adalah untuk mendapatkan akurasi dan parameter yang tepat dalam sebuah sistem analisis sentimen menggunakan metode *Support Vector Machine* (SVM) terhadap komentar pada *post* di Instagram.

1.5 Manfaat Penelitian

Dengan adanya penelitian ini diharap dapat menghasilkan sistem yang dapat mengklasifikasi sentimen komentar pada *post* di Instagram.

1.6 Metode / Pendekatan

Metode yang akan digunakan dalam penelitian kali ini akan dibagi menjadi sebagai berikut:

1.6.1 Tahap Studi Literatur

Pada tahap ini penulis akan mengumpulkan beberapa teori, artikel-artikel dan penelitian lain yang berkaitan dengan Analisis sentimen, *Text Mining*, dan *Support Vector Machine* yang dapat membantu penulis dalam menyelesaikan penelitian ini.

1.6.2 Tahap Pengumpulan Data

Pada tahap ini penulis akan mengumpulkan data-data yaitu berupa komentar yang ada di Instagram. Pengumpulan data akan menggunakan Instagram *private*

API. Serta pada tahap ini akan dilakukan pelabelan data. Untuk setiap akun *public figure* Indonesia di Instagram akan diambil post 10 hari terakhir. Setiap post akan di ambil 50 komentar.

1.6.3 Tahap Perancangan Sistem

Pada tahap ini penulis akan membuat sistem labeling untuk melabeli setiap dataset yang telah dikumpulkan. Sistem pelabelan akan dibuat menggunakan framework Laravel. Serta dalam tahap ini akan merancang sistem untuk *preprocessing* dan klasifikasi menggunakan SVM.

1.6.4 Tahap Validasi dan Evaluasi

Pada tahap ini, sistem yang telah dibuat pada tahap sebelumnya diuji dan dilakukan evaluasi sehingga dapat dikembangkan lagi untuk kedepannya. Validasi sistem menggunakan *k-fold cross-validation* sedangkan evaluasi sistem akan menggunakan *confusion matrix* yang mendapatkan nilai *accuracy, precision, recall dan f-measure*.

1.6.5 Tahap Penarikan Kesimpulan

Dari data yang didapat dari tahap evaluasi akan ditarik kesimpulan performa penerapan sistem yang dapat menjadi acuan untuk perbaikan pada penelitian selanjutnya.

1.7 Sistematika Penulisan

Sistematika penulisan laporan kerja praktik ini dibagi menjadi 5 bab, yaitu bagian pendahuluan, tinjauan pustaka, perancangan sistem, implementasi dan analisis sistem, kemudian kesimpulan dan saran.

Bab 1 adalah pendahuluan terdiri dari beberapa sub-bab yaitu berisikan latar belakang dari penelitian, perumusan masalah, tujuan penelitian, manfaat penelitian, metode atau pendekatan dan sistematika penulisan.

Bab 2 adalah tinjauan pustaka yang terdiri dari tujuan pustaka dan landasan teori. Tinjauan pustaka berisi tentang penelitian sebelumnya yang berkaitan dengan penelitian ini. Sedangkan landasan teori berisi tentang teori yang akan digunakan dalam penelitian ini.

Bab 3 adalah perancangan sistem, yang berisi tentang rancangan sistem yang akan digunakan untuk pengambilan data dan pengujian penelitian.

Bab 4 adalah implementasi dan analisis sistem yang berisi tentang training model serta evaluasi model pada sistem yang dibuat.

Bab 5 adalah kesimpulan dan saran berisikan tentang kesimpulan yang diperoleh berdasarkan analisis dan pengujian sistem.

©UKDW

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian ini dapat disimpulkan bahwa sistem klasifikasi Support Vector Machine dapat mengklasifikasi dengan baik komentar Instagram dengan data komentar positif. Untuk kernel linear mendapatkan rata-rata akurasi terbaik dengan fitur kombinasi unigram dan bigram sebesar 76.307 %. Sedangkan untuk kernel polinomial mendapatkan rata-rata akurasi terbaik dengan fitur unigram sebesar 73.1783 %. Serta fitur kombinasi unigram dan bigram mendapatkan hasil rata-rata akurasi tertinggi dibandingkan dengan fitur unigram ataupun fitur bigram saja.

Namun dari penelitian ini juga mendapat kesimpulan bahwa tidak semua kernel dapat mengklasifikasi fitur tf-idf dengan baik. Dikarenakan hasil rata-rata F1-measure yang cukup fatal yaitu pada rentang 0 – 50 % ketika menggunakan kernel polynomial. Sedangkan untuk penggunaan stemmer didalam preprocessing, dapat disimpulkan bahwa stemmer dapat meningkatkan akurasi walaupun tidak signifikan. Rata-rata akurasi yang didapatkan untuk penggunaan stemmer sebesar 76.307 % dan tanpa stemmer sebesar 75.4491 %. Stemmer tersebut membantu dengan memotong imbuhan pada kata yang ada di komentar menjadi kata yang baku.

5.2 Saran

Dalam penelitian ini terdapat beberapa hal yang perlu dikembangkan lebih lanjut terutama pada bagian preprocessing. Untuk meningkatkan tingkat akurasi bagian preprocessing dapat ditambahkan dengan penyelesaian masalah singkatan ataupun kata yang tidak baku. Diharap pada penelitian kedepannya, dapat menerapkan preprocessing yang lebih baik terutama pada singkatan, kata tidak baku dan kesalahan penulisan di komentar.

DAFTAR PUSTAKA

- Boser, B., Guyon, I., & Vapnik, V. (1992). A Training Algorithm for Optimal Margin Classifiers. *COLT '92 Proceedings of the fifth annual workshop on Computational learning theory*, 144-152.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 27:1--27:27.
- Duan, K., Keerthi, S., & Poo, A.-N. (2003). Evaluation of simple performance measures for tuning SVM hyperparameters. *Neurocomputing*, 41 – 59.
- Fan, R.-E., Chen, P.-H., & Lin, C.-J. (2005). Working Set Selection Using Second Order Information for Training Support Vector Machines. *Journal of Machine Learning Research* 6, 1889–1918.
- Ghag, K. (2015). Comparative analysis of effect of stopwords removal on sentiment. *International Conference on Computer, Communication and Control (IC4)*.
- Hamilton, H. (2018, July). *Confusion Matrix*. Diambil kembali dari [www2.cs.uregina.ca: http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html](http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html)
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques Third Edition*. Waltham: Morgan Kaufmann.
- Hidayatullah, A. F. (2016, Mei). Pengaruh Stopword Terhadap Performa Klasifikasi Tweet Berbahasa Indonesia. *JISKa*, 1(1), 1-4.
- Jadav, B. M., & Vaghela, V. B. (2016). Sentiment Analysis using Support Vector Machine based on Feature Selection and Semantic Analysis. *International Journal of Computer Applications (0975 – 8887)*, 146(13), 26-30.
- KOMINFO. (2018, Februari 19). *Jumlah Pengguna Internet 2017 Meningkat, Kominfo akan Terus Lakukan Percepatan Pembangunan Broadband [SIARAN PERS NO. 53/HM/KOMINFO/02/2018]*. Diambil kembali dari <https://kominfo.go.id/index.php/content/detail/12640/siaran-pers-no->

- 53hmkominfo022018-tentang-jumlah-pengguna-internet-2017-meningkat-kominfo-terus-lakukan-percepatan-pembangunan-broadband/0/siaran_pers
- Ling, J., Kencana, I. E., & Oka, T. B. (2014). Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier dengan Seleksi Fitur Chi Square. *E-Jurnal Matematika Vol. 3 (3)*, 92-99.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Chicago: Morgan & Claypool Publisher.
- Lukito, Y., & Rachmat C, A. (2016). SENTIPOL: Dataset Sentimen Komentar Pada Kampanye PEMILU Presiden Indonesia 2014 dari Facebook Page. *Konferensi Nasional Teknologi Informasi dan Komunikasi (KNASTIK 2016)*, (hal. 218-228). Yogyakarta.
- Rachmat C, A., & Lukito, Y. (2016). Klasifikasi Sentimen Komentar Politik dari Facebook Page Menggunakan Naive Bayes. *JUISI*, 26-34.
- Rachmat C, A., & Lukito, Y. (2017). Identifikasi Komentar Spam Pada Instagram. *LONTAR KOMPUTER VOL. 8, NO. 3*, 219-231.
- Santoso, V. I., Virginia, G., & Lukito, Y. (2017). Penerapan Sentiment Analisis pada Hasil Evaluasi Dosen. *JURNAL TRANSFORMATIKA*, 72-76.