

OPINION MINING POPULARITAS TEMPAT WISATA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) DAN PARTICLE SWARM OPTIMIZATION DARI INSTAGRAM

Esti Ika Maharani¹, Naim Rochmawati²

^{1,2}Jurusan Teknik Informatika, Universitas Negeri Surabaya

¹esti.17051204019@mhs.unesa.ac.id

²naimrochmawati@unesa.ac.id

Abstrak— Sosial media sekarang mempunyai efek besar pada pengambilan keputusan, termasuk pada industri wisata untuk mengetahui reputasi dan popularitas sebuah destinasi wisata. Pada era modern seperti sekarang penggunaan media sosial tidak sekedar media hiburan akan tetapi juga menjadi media buat mengembangkan informasi, baik informasi yang dipercaya maupun tidak. Dari media sosial tersebut, mengizinkan penggunanya untuk bisa mengetahui reaksi pengguna lain terhadap informasi yang sedang hangat dibicarakan. Analisa sentiment atau *opinion mining* adalah satu dari berbagai solusi untuk mengatasi masalah dalam mengelompokkan opini atau *review* sebagai opini positif, negatif, atau netral secara otomatis. Metode yang dipakai merupakan *Support Vector Machine* (SVM) dan penerapan seleksi fitur *Particle Swarm Optimization*. Keduanya memanfaatkan library yang telah tersedia dalam python. *Support Vector Machine* (SVM) merupakan metode untuk menganalisa data dan mengenali pola yang mampu dipakai untuk pengklasifikasian.. Sementara *particle swarm* berpola *optimizer* jadi tidak mampu berdiri sendiri, wajib diikuti algoritma lain yang harus melakukan proses pembelajaran terlebih dahulu. Nilai akurasi yang didapatkan sebagai tolak ukur dalam mencari contoh pengujian terbaik untuk masalah klasifikasi sentimen. Hasil penelitian memperlihatkan peningkatan nilai akurasi sebanyak 0.07 untuk algoritma *Support Vector Machine* dari 0.90 menjadi 0.97 sesudah penerapan seleksi fitur *Particle Swarm Optimization*.

Kata Kunci— Tempat wisata, *Analysis Sentiment*, *Support Vector Machine*, Instagram, *Particle Swarm Optimization*

I. PENDAHULUAN

Sosial media adalah wahana prasarana untuk menyediakan kebutuhan kelangsungan hidup manusia supaya memudahkan pada pengerjaannya.

Pemanfaatan *sentiment analysis* sudah diterapkan dalam berbagai kebutuhan, bahkan banyak perusahaan merogoh nilai lebih menurut penerapannya untuk menerima fungsi baru. Banyak algoritma untuk sentiment analysis, dari dalam output pembagian terstruktur mengenai data pelatihan yang memakai data set sebelum data dilatih maka diekstraksi lebih dulu.

Crawling data menurut Instagram adalah pengambilan data baik melalui *posting* juga komentar yang dianalisis untuk bisa membuat *opinion mining* tempat wisata. Ada 2 komponen yang akan dievaluasi untuk membuat hasil menurut peringkat popularitas yaitu *Instagram comment* dan *Instagram post*. Komponen tersebut dianalisis memakai *sentiment analysis* yang bertujuan memilih apakah komentar bernilai positif, negatif, atau netral.

Particle Swarm Optimization (PSO) adalah teknik meningkatkan secara optimal dan mudah untuk menerapkan serta merubah sebagian parameter. Dalam teknik ini ada beberapa cara melakukan pengoptimalan antara lain menaikkan bobot atribut (*attribute weight*) pada seluruh atribut yang digunakan, *attribute selection* dan *feature selection*.

SVM merupakan metode penjabaran yang digunakan dalam mencari *Maximum Marginal Hyperplane* (MMH) atau pemisah paling baik untuk membentuk pemisahan aporisma seluruh kelas. Tahap penjabaran yaitu dilakukan dengan data positif, negatif, dan netral berdasarkan *output review* yang sudah di *scraping* terlebih dahulu, didalamnya berisi data ulasan positif, data ulasan negatif, dan data ulasan netral yang dipakai oleh algoritma SVM untuk mempelajari pola data menurut karakteristik-karakteristik data dalam masing-masing kelas. Hasil dari algoritma SVM tersebut lalu bisa diuji memakai data uji, sebagai akibatnya bisa diukur taraf akurasi pada memprediksi kelas dalam data baru, proses itu disebut *machine learning*.

Penelitian mengenai *sentiment analysis* yang sudah dilakukan terkait wisata antara lain dalam penelitian [11] yang dilakukan oleh Muhammad Azmi, Amiruddin Khairul Huda, dan Arief Setiyanto menggunakan data Instagram untuk *searching* lokasi wisata yang mempunyai *rating* tinggi pada Pulau Lombok memakai *keywords* pantai, gili, dan gunung menggunakan *Naïve Bayes Classifier*. Dalam penelitian tadi dihasilkan akurasi *naïve bayes* sebesar 59%, wisata tersebut yaitu Pantai Kuta, Gili Air dan Gunung Rinjani.

Pada penelitian [8] yang dilakukan Murnawan dan Ardiles Sinaga memakai metode *Naïve Bayes Classifier*, menggunakan *test data* Instagram terhadap 50 *Point of Interest* (POI) tujuan wisata di Pulau Bali, hasilnya menggunakan destinasi wisata di Pulau Bali tadi membentuk *rating* yang bisa digunakan sebagai media pertimbangan wisatawan untuk menjadi tujuan wisata mereka.

Kemudian pada penelitian [3] yang dilakukan oleh Achmad Rofi'i, Totok Wahyu Wibowo, Sudaryanto, dan Nur Mohammad Farda menggunakan metode *Support Vector Machine*, dengan *hashtag* "wisata" dalam rentang 3 bulan, diperoleh total 277.638 *posting*. Namun, setelah melalui *pre-processing*, tersisa 56.525 kiriman. Ini menunjukkan keragaman data di Provinsi Jawa Tengah dan DIY. Ekstraksi pengetahuan menggunakan data besar Instagram untuk wisatawan dapat diterapkan dengan menggunakan geovisualisasi. Namun demikian, penelitian untuk ekstraksi kuantitatif kebutuhan wisatawan komprehensif karena jangkauan data wisatawan statistik antara Instagram dan turis sangat besar.

Penerapan *Support Vector Machine* dalam penelitian [10] yang dilakukan Elly Indrayuni membuktikan algoritma SVM adalah salah satu prosedur pemecahan yang akurat sebesar 91.33% dan nilai AUC sebanyak 0.988. Penggunaan seleksi fitur PSO dalam algoritma *Support Vector Machine* bisa menaikkan akurasi. Penerapan fitur *Particle Swarm Optimization* dalam algoritma SVM menciptakan nilai akurasi semakin tinggi sampai 5.61%, yaitu menjadi 96.94% dan AUC sebanyak 0.992.

Pada website ini bisa digunakan untuk sentiment analisa tempat wisata di seluruh Indonesia dan menggunakan Bahasa Indonesia. Untuk *range* waktu pengambilan data pada Instagram berkisar 1 bulan terakhir atau satu bulan dari sekarang karena pengambilan data bersifat otomatis.

Instagram menjadi media sebagian besar masyarakat untuk mengemukakan informasi yang patut diabadikan melalui *posting* maupun berpendapat mengenai sesuatu hal yang sedang hangat dibicarakan melalui *caption* atau komentar. Komentar maupun *posting* pada Instagram dapat menentukan sebuah *rating* pada tempat wisata. Untuk komentar dapat diklasifikasikan menjadi komentar positif, komentar negatif, dan komentar netral.

Klasifikasi *sentiment analysis* tingkat popularitas tujuan wisata melalui *crawling* komentar dan *posting* dari Instagram dengan menggunakan algoritma SVM serta penerapan seleksi fitur *Particle Swarm Optimization*. Keduanya memanfaatkan library yang sudah tersedia pada python. SVM merupakan metode untuk menganalisis data serta mengenali pola yang dapat dipergunakan untuk *classification*.

Algoritma utama di sini memakai *Support Vector Machine* sebagai *classifier* karena SVM memiliki tingkat akurasi kurang sempurna maka dari itu pasca *processing*

diperlukan *optimizer* menggunakan pola *particle swarm* untuk meningkatkan tingkat akurasi nya. Karena sama-sama berbasis *machine learning*, maka 2 algoritma ini cocok untuk digabungkan dan bisa saling menyesuaikan pembelajaran dengan populasi data tinggi. Untuk output dari aplikasi ini berupa hasil *scraping* atau *text mining* dari Instagram dan hasil probabilitas *sentiment analysis*.

Jika hanya menggunakan SVM maka tingkat akurasi nya rendah. Sementara *particle swarm* berupa *optimizer* jadi tidak bisa berdiri sendiri, harus diikuti algoritma lain yang harus melakukan proses pembelajaran terlebih dahulu.

II. METODOLOGI

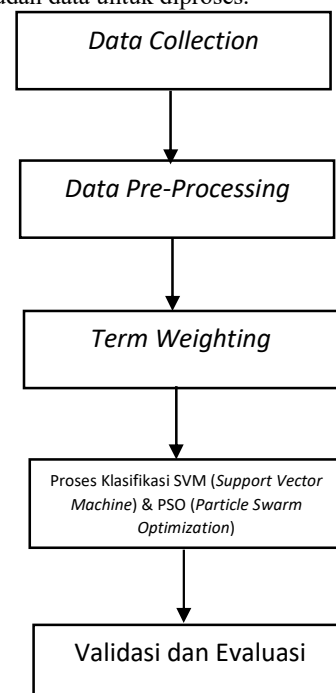
Pada pelaksanaan penelitian harus memiliki pedoman berupa langkah-langkah atau alur penelitian agar hasil pelaksanaan penelitian sesuai dengan tujuan awal yang akan dicapai, oleh karenanya dibutuhkan metodologi dalam pelaksanaan penelitian. Metodologi merupakan rancangan alur penelitian yang terdiri dari beberapa tahapan penelitian yang terstruktur dan sesuai dengan tujuan penelitian yang diinginkan. Beberapa tahapan pada peneltisn ini, yaitu :

1. Langkah Pertama

Pada langkah ini terjadi *data collecion* yang mengambil data Instagram dari *keywords* yang berasal dari *input* pengguna. Untuk memperoleh data ini digunakan library dari python yaitu selenium, yang dimana bekerja untuk memparser konten html dari Instagram. Data kemudian akan dikoleksi ke dalam bentuk dokumen dan dipisah menjadi data *training* dan data *testing*.

2. Langkah Kedua

Dalam langkah ini terjadi *pre-processing* data dengan beberapa tahapan diantaranya, *case folding*, *steaming*, *tokenize*, *slang words*, *negation*, dan *stopwords*. *Pre-processing* terjadi supaya data bisa diklasifikasi serta agar mempermudah data untuk diproses.



Gambar. 1 Tahapan Penelitian

a. Case folding

Case folding yaitu melakukan perubahan semua alfabet sebagai alfabet kecil, hanya alfabet 'a' hingga menggunakan 'z' yang diterima. Karakter selain alfabet dihapus serta disebut delimiter.

b. Tokenize

Tokenizing merupakan tahap pemisahan teks sebagai potongan-potongan yang diklaim menjadi token lalu dianalisa. Kata, angka, simbol, tanda baca, juga entitas krusial lain bisa disebut menjadi token. Pada NLP, token bisa disebut menjadi "kata" meskipun *tokenize* bisa dilakukan dalam paragraf juga kalimat.

c. Stemming

Stemming merupakan proses *searching* root kata menurut per kata *output filtering*. Dalam proses ini dilakukan pengambilan sebuah istilah ke dalam suatu representasi yang sama.

d. Slang words

Slang words berarti mengganti istilah-istilah formal yang terkadang disederhanakan menggunakan istilah lain yang lebih nyaman untuk diucapkan pada komunikasi verbal.

e. Negation

Negation adalah bentuk penyangkalan atau pertentangan akan sesuatu. *Negation* mampu diklaim sebagai perubahan bentuk pernyataan afirmatif menjadi bersifat bertolak belakang.

f. Stopwords

Stopword merupakan istilah generik yang umumnya ada pada jumlah besar dan dipercaya tidak mempunyai makna.

3. Langkah Ketiga

Tahap ini melakukan *term weighting* yang dipergunakan untuk mengubah dari data opini ke dalam data numerik. Dalam penelitian ini menggunakan *Multiclass Confusion Matrix* yang dapat digambarkan sebagai berikut :

		Prediksi		
		Positif	Negatif	Netral
Aktual	Positif	TPos	FPosNeg	FPosNet
	Negatif	FNegPos	TNeg	FNegNet
	Netral	FNetPos	FNetNeg	TNet

Tabel. 1 *Multiclass Confusion Matrix*

Tidak jauh beda dengan *confusion matrix* berdimensi 2x2, *multiclass confusion matrix* mempunyai elemen TP (*True Positive*), FN (*False Negative*), FP (*False Positive*), dan TN (*True Negative*). Ketentuan dalam menetapkan elemen tersebut adalah sebagai berikut :

- TP adalah banyaknya data yang kelas aktualnya sama seperti kelas prediksinya.
- FN adalah jumlah dari semua baris yang ditunjuk kecuali TP yang dicari.
- FP adalah jumlah dari semua kolom yang ditunjuk kecuali TP yang dicari.
- TN adalah jumlah semua kolom dan baris selain yang ditunjuk.

Dari keterangan di atas maka diperoleh rumus :

Accuracy

$$Accuracy = \frac{N_{benar}}{N} \times 100\%$$

N = jumlah

Precision

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

Recall

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

F-1 score

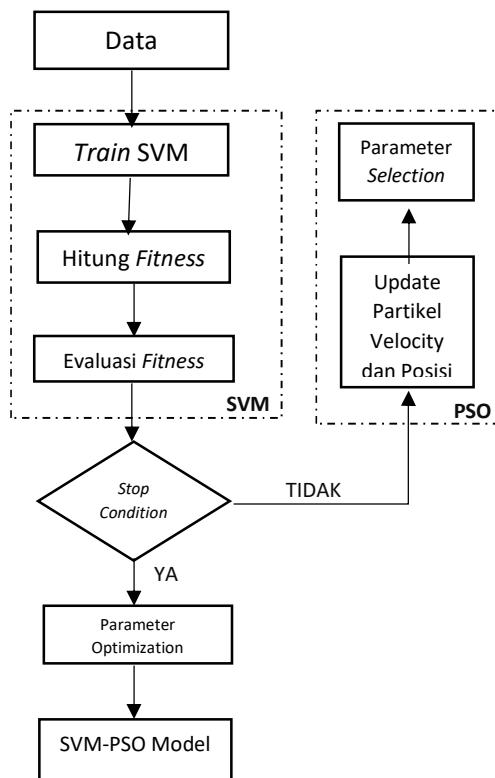
$$F-1 \text{ score} = \frac{2 \times Recall \times Precision}{(Recall + Precision)}$$

4. Langkah Keempat

Langkah ini merupakan penjabaran dengan metode SVM dan SVM-PSO. Fungsinya adalah membandingkan *output* sentimen positif, negatif, atau netral, akurasi, dan AUC menurut metode klasifikasi SVM dan SVM-PSO. Data *training* dipakai untuk melatih algoritma sedangkan data *testing* dipergunakan untuk melihat kinerja algoritma yang telah sebelumnya dilatih. Proses analisis memakai SVM mulanya yaitu mengganti teks sebagai data vektor. Akibat sulit dalam memilih parameter yang cocok menggunakan SVM, dalam artikel ini PSO dipakai untuk mengatur parameter SVM guna dioptimalkan. Tetapi, dalam artikel ini setiap parameter dalam SVM dan PSO diatur default menggunakan nilai 0. Output yang didapat merupakan kelas positif, negatif, dan netral.

5. Langkah Kelima

Tahap ini merupakan proses validasi memakai 10 *k-fold cross validation* yang berfungsi untuk menghapus *noise*, sebagai akibatnya bisa mempertinggi akurasi. Evaluasi pengukuran akurasi diukur memakai *confusion matrix*.



Gambar. 2 Model SVM-PSO

III. HASIL DAN PEMBAHASAN

Pada penelitian mengenai *Opinion Mining* Popularitas Tempat Wisata dari Instagram ini menggunakan algoritma *Support Vector Machine* (SVM) serta penerapan seleksi fitur *Particle Swarm Optimization*. Tujuannya untuk menentukan tingkat popularitas tujuan wisata dari hasil *crawling* Instagram.

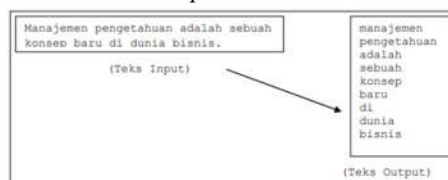
1. Data Collection

Dalam hal ini, digunakan data dari *posting* Instagram berupa *hashtag* dan *comment* pengguna Instagram seluruh Indonesia dan menggunakan Bahasa Indonesia.

2. Pengolahan Data Awal (Preprocessing)

a. Case folding

Case folding yaitu merubah seluruh alfabet sebagai alfabet kecil, hanya alfabet 'a' hingga menggunakan 'z' yang diterima. Selain alfabet dihapus dan disebut delimiter.



Gambar. 3 Proses *folding*

b. Tokenize

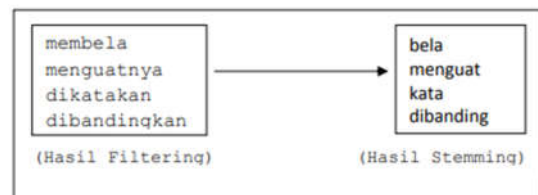
Tokenizing merupakan tahapan pemisahan teks sebagai potongan-potongan yang diklaim menjadi token lalu dianalisa. Angka, kata, simbol, tanda baca, juga entitas krusial lainnya bisa dianggap menjadi token. Pada NLP, token bisa disebut menjadi "kata" meskipun *tokenize* bisa dilakukan dalam paragraf juga kalimat. Pada python fungsi `split ()` bisa digunakan untuk memisahkan teks. Sebagai contoh, perhatikan kalimat di bawah ini :

```
kalimat = "tempat wisata indah Indonesia"
pisah = kalimat.split()
print(pisah)
#output
#[ 'tempat', 'wisata', 'indah', 'Indonesia' ]
```

Hasil dari kode di atas kemudian diolah, kata "indah" dan "indah" sebagai 2 entitas yang berbeda. Akibatnya dapat berpengaruh pada hasil dari analisa teks itu sendiri.

c. Stemming

Stemming merupakan proses *searching* root kata menurut tiap kata *output filtering*. Dalam proses ini dilakukan pengambilan bentuk istilah ke dalam suatu representasi yang sama. Tahap ini memang banyak digunakan untuk teks yang menggunakan Bahasa Inggris dan lebih sulit diimplementasikan pada teks yang menggunakan Bahasa Indonesia. Ini karena Bahasa Indonesia tidak memiliki rumus bentuk baku yang permanen.



Gambar. 4 Proses *stemming*

d. Slang words

Slang words berarti pengganti istilah-istilah formal yang terkadang disederhanakan menggunakan istilah lain yang lebih nyaman untuk diucapkan pada komunikasi verbal.

e. Negation

Negation yaitu suatu penyangkalan atau kontradiksi akan sesuatu. *Negation* mampu diklaim sebagai perubahan bentuk pernyataan afirmatif menjadi bersifat bertolak belakang.

f. Stopwords

Stopword merupakan istilah generik yang umumnya ada pada jumlah besar serta dipercaya tidak mempunyai arti. Contoh *stopword* pada Bahasa Indonesia adalah "yang", "dan", "di", "dari", dan lain-lain. Makna dibalik penggunaan *stopword* yaitu dengan menghilangkan kata-

sosial. Terbukti dari *scraping* website banyaknya pengunjung yang mengomentari laman Instagram terkait pantai ini, meskipun komentar bernilai positif namun hanya sedikit komentar yang didapat.

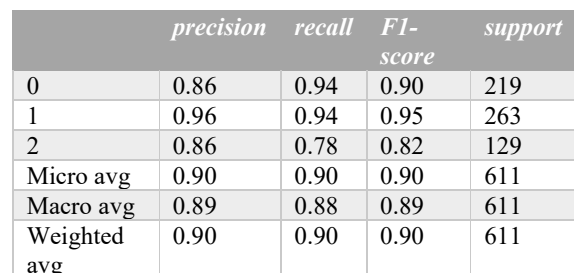
INFORMATION ABOUT pantalapcar					
Total Comment: 2					
Positive Comment: 2					
Negative Comment: 0					
New Comment: 0					
Alkuran: 0.97428574285714					
Caption	Hashtag	Mention	Link		
Chara m'la berbau, adabir e m'ak reginistral, 🤔🤔			https://www.instagram.com/p/CQhUW4WU/		
Kendira tepungtu arana, k'at rebaku,			https://www.instagram.com/p/CQD...J380W/		
Post Link	Username	Comment	Sentiment	Predict Probability	Accurasion
1.	pantalapcar	Chara m'la berbau, adabir e m'ak reginistral, 🤔🤔	positive	0.995	
2.					
3.					

Meode yang diusulkan yaitu penerapan algoritma *Support Vector Machine* (SVM) dan menggunakan seleksi fitur *Particle Swarm Optimization* (PSO).

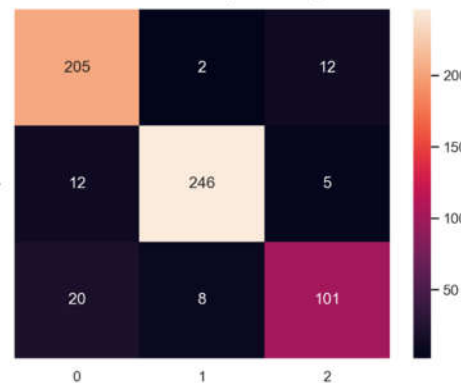
5. Hasil Eksperiment dan Pengujian Metode *Support Vector Machine* (SVM)

- ## Classification Report SVM

(*'Accuracy:'*, 0. 9034369885433715)

Tabel. 2 *Classification Report SVM*

- Pantai Klayar adalah sebuah wisata yang terletak di Pacitan Jawa Timur dan berbatasan dengan Wonogiri Jawa Tengah. Tiket masuk yang terbilang murah membuat wisatawan berantusias untuk mendatangi wisata ini. Terbukti dengan komentar positif yang lumayan tinggi pada laman Instagram yang dapat terlihat pada *scraping* website penelitian ini.



Gambar. 8 *Convolution* matrix SVM

- Pantai Pacar adalah sebuah pantai yang masih alami dan terletak di Pucanglaban Tulungagung. Meskipun tidak ada tiket masuk, namun pantai ini terbilang masih sepi pengunjung dikarenakan belum dikelola oleh pemerintah daerah secara resmi dan belum terlalu ter *ekspose* ke media

Pada Tabel 2 dan Gambar 8 di atas diketahui bahwa nilai *accuracy* yaitu sebesar 0.90, komentar positif (0) menghasilkan *precision* 0.86, *recall* 0.94, *f1-score* 0.90, dan *support* 219, komentar negatif (1) menghasilkan *precision* 0.96, *recall* 0.94, *f1-score* 0.95, dan *support* 263, sedangkan komentar netral (2) menghasilkan *precision* 0.86, *recall* 0.78, *f1-score* 0.82, dan *support* 129. Dari data di atas maka diperoleh *micro average* untuk *precision* 0.90, *recall* 0.90, *f1-score* 0.90, dan *support* 611, kemudian *macro average*

untuk *precision* 0.89, *recall* 0.88, *f1-score* 0.89, dan *support* 611, sedangkan *weighted average* untuk *precision* 0.90, *recall* 0.90, *f1-score* 0.90, dan *support* 611.

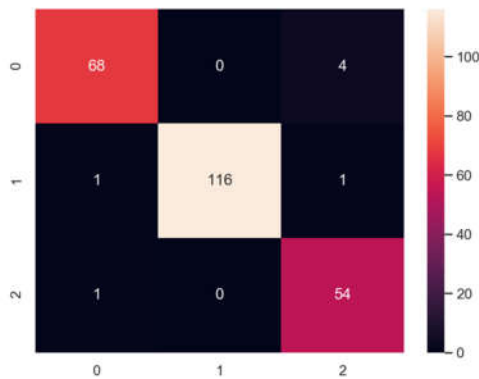
6. Hasil serta Pengujian Metode Support Vector Machine dan menggunakan seleksi fitur Particle Swarm Optimization

Classification Report SVM-PSO

(‘Accuracy’: 0.9714285714285714)

	<i>precision</i>	<i>recall</i>	<i>F1-score</i>	<i>support</i>
0	0.97	0.94	0.96	72
1	1.00	0.98	0.99	118
2	0.92	0.98	0.95	55
Micro avg	0.97	0.97	0.97	245
Macro avg	0.96	0.97	0.97	245
Weighted avg	0.97	0.97	0.97	245

Tabel. 3 Classification Report SVM-PSO



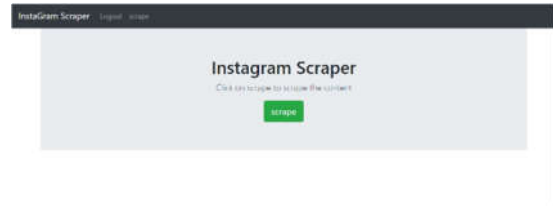
Gambar. 9 Convolution matrix SVM-PSO

Dari Tabel 3 dan Gambar 9 di atas diketahui bahwa nilai *accuracy* yaitu sebesar 0.97, komentar positif (0) menghasilkan *precision* 0.97, *recall* 0.94, *f1-score* 0.96, dan *support* 72, komentar negatif (1) menghasilkan *precision* 1.00, *recall* 0.98, *f1-score* 0.99, dan *support* 118, sedangkan komentar netral (2) menghasilkan *precision* 0.92, *recall* 0.98, *f1-score* 0.95, dan *support* 55. Dari data di atas maka diperoleh *micro average* untuk *precision* 0.97, *recall* 0.97, *f1-score* 0.97, dan *support* 245, kemudian *macro average* untuk *precision* 0.96, *recall* 0.97, *f1-score* 0.97, dan *support* 245, sedangkan *weighted average* untuk *precision* 0.97, *recall* 0.97, *f1-score* 0.97, dan *support* 245.

7. Implementasi

Dari hasil pengujian dengan *Confusion Matrix*, maka dilakukan implementasi algoritma dengan membuat suatu website yang berfungsi mempermudah wisatawan dalam mencari tempat wisata dengan *review* yang bagus pada Instagram.

Desainnya dibuat *user friendly* agar mempermudah pengguna dalam menganalisa *review* tempat wisata. Karena adanya website ini, wisatawan yang membutuhkan informasi tentang tempat wisata dapat dengan mudah memperolehnya.



Gambar. 10 Tampilan Halaman Home

Apabila pengguna sudah melakukan *login* sebelumnya, maka pengguna akan masuk pada halaman *home* dan terdapat *button scrape* untuk melanjutkan ke tahap selanjutnya.



Gambar. 11 Tampilan Halaman Login

Namun apabila pengguna belum melakukan *login*, maka pengguna akan masuk pada halaman *login* untuk mengisi *username* dan *password* terlebih dahulu.



Gambar. 12 Tampilan Halaman Enter Content

Setelah melakukan *login*, pengguna akan masuk pada halaman *enter content* dan mengisi destinasi wisata yang ingin mereka cari tahu.



Gambar. 13 Tampilan Halaman Scrape Result

Setelah pengguna mengisi destinasi wisata yang ingin mereka cari tahu, maka akan masuk pada halaman *scrape result* yang berisi jumlah komentar positif, negatif, netral, serta tingkat akurasinya. Selain jumlahnya, juga terdapat penjelasan seperti *post link*, *username*, *comment*, *sentiment*, dan *predict probability accuracy*.

8. Implikasi Penelitian

Dalam penelitian ini terdapat dua aspek, yaitu :

a. Aspek Sistem

Akibat adanya *opinion mining*, pengguna Instagram terkait popularitas tempat wisata bisa mempermudah wisatawan untuk menemukan wisata yang tepat dengan keinginan mereka dan tidak perlu membaca *review* panjang sehingga tidak memakan waktu.

b. Aspek Manajerial

Secara manajerial hasil *opinion mining* pengguna Instagram terkait popularitas tempat wisata dapat digunakan pihak pengelola tempat wisata yaitu melihat *review* pengunjung bersifat positif, negatif, atau netral untuk dijadikan evaluasi dalam peningkatan fasilitas dan pelayanan (untuk yang bersifat wisata buatan).

IV. KESIMPULAN

Adapun kesimpulan dari hasil pengujian algoritma *Support Vector Machine* dalam penelitian yang sudah dilakukan yaitu bahwa dalam penelitian ini terbukti algoritma *Support Vector Machine* adalah salah satu algoritma yang paling akurat menghasilkan akurasi 0.90, lalu dengan menggunakan seleksi fitur *Particle Swarm Optimization* dalam algoritma *Support Vector Machine* bisa menaikkan akurasi dan menciptakan nilai akurasi semakin tinggi sampai 0.07 yang termasuk *Excellent Classification*.

V. SARAN

Penelitian ini mampu diluaskan lagi untuk penjabaran teks. Data juga tidak harus dari *review* yang berisi opini, tetapi bisa menggunakan *review* yang bersifat *summary*, ataupun menurut status langsung misalnya *facebook* atau bisa menggunakan media sosial lain. Tidak hanya itu, algoritma nya pun juga bisa dikembangkan menggunakan algoritma lainnya misalnya algoritma *Neural Network*, KNN, dan juga bisa menggunakan seleksi fitur lain misalnya *Distinguishing Feature Selector*.

UCAPAN TERIMA KASIH

Puji syukur saya panjatkan atas rahmat dan ridho Allah SWT yang telah diberikan kepada saya, sehingga saya dapat menyelesaikan penelitian ini dengan baik, dan juga shalawat serta salam tercurahkan kepada junjungan kita Nabi Muhammad SAW. Terimakasih juga kepada kedua orangtua, orang-orang terdekat, dan teman-teman di sekeliling saya yang selalu memberikan semangat dalam

menuntut ilmu. Serta terimakasih kepada Bu Naim Rochmawati, S.Kom., MT. selaku dosen pembimbing saya yang sudah membimbing penelitian ini dari awal sampai akhir. Terakhir, terimakasih juga kepada seluruh pihak yang telah mendukung jalannya penelitian ini.

REFERENSI

- [1] Alaei, Ali Reza, dkk. 2017. Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*. 58(9). [Diakses 20 Januari 2021].
- [2] Alita, Debby, dkk. 2019. Analysis of Emoticon and Sarcasm Effect on Sentiment Analysis of Indonesian Language on Twitter. *JISEBI*. 5(2) :100-109.
- [3] Azmi, Muhammad, dkk. 2020. Pemanfaatan Data Instagram untuk Mengetahui Reputasi Tempat Wisata di Lombok. *Teknimedia*. 1(1): 39-46. [Diakses 1 Januari 2021].
- [4] Buntoro, Ghulam Asrofi. 2017. Sentiment Analysis to Prediction DKI Jakarta Governor 2017 on Indonesian Twitter. *IJSEIT*. 2(3).
- [5] Buntoro, Ghulam Asrofi. 2019. Sentiments Analysis for Governor of East Java 2018 in Twitter. *Journal Publications & Informatics Engineering Research*. 3(2).
- [6] Fanissa, Shima, dkk. 2018. Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naïve Bayes dan Seleksi Fitur Query Expansion Ranking. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. 2(8). [Diakses 14 Desember 2020].
- [7] H.N Nadya, R. Naim. 2021. Penerapan NBC dan SVM Terhadap kKlasifikasi Opini Sistem Pendidikan Menggunakan Data Twitter.
- [8] Indrayuni, Elly. 2016. Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization. *Jurnal Evolusi* 4(2). [Diakses 20 Januari 2021].
- [9] O. Somantri, Dairoh. 2019. Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining. [Diakses 14 Desember 2020].
- [10] Rofi'i, Achmad, dkk. 2019. Tourist Geovisualization Analysis Utilizing Instagram Data in Central Java Province and Special Region of Yogyakarta. [Diakses 4 Januari 2021].
- [11] S. Ardiles, Murnawa. 2017. Implementasi Sentiment Analysis untuk Menentukan Tingkat Popularitas Tujuan Wisata. [Diakses 1 Januari 2021].
- [12] W. Shafqat, Y.C Byun. 2019. A Recommendation Mechanism for Under Emphasized Tourist Spots Using Topic Modeling and Sentiment Analysis. [Diakses 14 Desember 2020].
- [13] Wilianto, Lio, dkk. 2017. Analisis Sentimen Terhadap Tempat Wisata dari Komentar Pengunjung dengan Menggunakan Metode Navie Bayes Classifier Studi Kasus Jawa Barat. [Diakses 14 Desember 2020].
- [14] Y.W Syaifudin, R.A Irawan. 2018. Implementasi Analisis Clustering dan Sentimen Data Twitter pada Opini Wisata Pantai Menggunakan Metode K-Means. *Jurnal Informatika Polinema*. 4(3): 189. [Diakses 14 Desember 2020].