e-ISSN: 2550-0813 | p-ISSN: 2541-657X | Vol 11 No 5 Tahun 2024 Hal. : 2046-2053

NUSANTARA: Jurnal Ilmu Pengetahuan Sosial

available online http://jurnal.um-tapsel.ac.id/index.php/nusantara/index

PENERAPAN MULTINOMIAL NAÏVE BAYES DENGAN PARTICLE SWARM OPTIMIZATION PADA KLASIFIKASI BERITA HOAX KONFLIK PALESTINA-ISRAEL

Salma Dian Aprilia, Basuki Rahmat, Hendra Maulana

Prodi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur Indonesia

Abstrak

Berita merupakan sumber informasi penting dalam masyarakat. Namun, dengan adanya media sosial yang semakin marak saat ini, banyak kasus penyebaran berita hoax yang kevalidannya tidak dapat dibuktikan. Salah satu topik yang hangat diperbincangkan oleh satu dunia adalah mengenai konflik Palestina-Israel yang terjadi akhir-akhir ini. BBC World sebagai media internasional pun juga menyebarkan berita terkait topik ini. Akan tetapi, dalam rubrik BBC World tidak ada penyaringan kategori berita yang termasuk hoax. Oleh karena itu, diperlukan suatu pengklasifikasian berita hoax dengan algoritma Multinomial Naïve Bayes. Untuk semakin meningkatkan akurasi, pada penelitian ini terdapat percobaan tambahan dengan Particle Swarm Optimization. Hasil pengujian membuktikan jika percobaan klasifikasi dengan menggunakan Multinomial Naïve Bayes hanya menghasilkan akurasi paling tinggi sebesar 76%, sedangkan hasil klasifikasi yang ditambah menggunakan Particle Swarm Optimization akan meningkat hasil akurasinya menjadi 84%. Dalam hal ini, metode optimassi dari Particle Swarm Optimization terbukti mampu dalam meningkatkan hasil akurasi penelitian klasifikasi berita hoax konflik Palestina-Israel ini.

Kata Kunci: Hoax, Optimasi, PSO, Multinomial Naïve Bayes, Akurasi.

PENDAHULUAN

Berita berperan sebagai suatu informasi yang memiliki peran penting dalam menyebarkan informasi kepada masyarakat secara cepat. Dalam era globalisasi yang didukung oleh kemajuan teknologi, media online seperti Kompas.com, Detik.com, dan sejenisnya menjadi salah satu sarana cepat untuk menyebarkan berita. Penggunaan

*Correspondence Address : basukirahmat.if@upnjatim.ac.id
DOI : 10.31604/jips.v11i5.2024. 2046-2053

internet di Indonesia mencapai 63 juta pengguna yang 95%-nya menggunakan internet untuk mengakses informasi (Sahid, 2023).

Meskipun media online memfasilitasi akses secara cepat tanpa batasan ruang dan waktu, tetapi seringkali berita yang tersebar pada media online adalah berita hoax yang menyesarkan. Berita hoax merupakan berita yang kebenarannya tidak dapat dibuktikan dan berpotensi akan membawa dampak negatif untuk manusia (Batoebara et al., 2020). Kementrian Komunikasi dan Informatika (Kominfo) mencatat bahwa hingga awal tahun 2022, terdapat 9546 berita hoax yang beredar di Indonesia (Muttaqin et al., 2023).

Topik berita hoax biasanya merupakan topik yang sedang hangat diperbincangkan oleh masyarakat Indonesia, salah satu contoh topiknya adalah konflik Palestina-Israel. Contoh berita hoax mengenai topik tersebut yang beredar dalam platform media sosial adalah video pemberitaan mengenai perang Hamas dan Israel, tetapi faktanya video tersebut merupakan video palsu yang sebenarnya adalah rekaman kegiatan lain (Wulandari & Candra, 2024). Topik mengenai konflik Palestina-Israel ini telah banyak diangkat di portal berita berbagai negara, salah satunya adalah media internasional British Broadcasting Corporation.

BBC merupakan portal media internasional yang jaringan wartawannya tersebar di banyak negara, tetapi belum ada rubrik tersendiri untuk mendeteksi berita hoax atau non hoax mengenai konflik Palestina-Israel. Banyaknya berita simpang siur mengenai topik ini sehingga diperlukan suatu pengklasifikasian hoax dan non hoax dengan algoritma yang cocok untuk klasifikasi teks, yaitu *Naïve Bayes Classifier.* Jenis tipe algoritma *Naïve Bayes* yang digunakan dalam penelitian ini adalah *Multinomial Naïve Bayes* karena memiliki kelebihan dalam menghasilkan tingkat akurasi yang tinggi dalam hal klasifikasi teks.

Penelitian sebelumnya yang dilakukan oleh Nur dan kawan-kawannya membahas mengenai perbandingan penggunaan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM). Studi kasus yang diangkat ialah pendeteksian berita hoax dari portal berita Indonesia dari bulan Januari 2020 hingga Juli 2023. Hasil penelitian menunjukan jika *Naïve Bayes* memiliki tingkat akurasi yang lebih unggul dibandingkan metode SVM dengan perolehan akurasi sebesar 85% untuk *Naïve Bayes* dan 75,5% untuk SVM (Febriyanty et al., 2023).

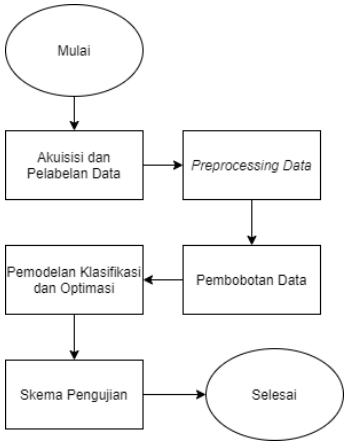
Pada beberapa penelitian lainnya memanfaatkan metode optimasi agar hasil akurasi dari penelitian dapat meningkat. Salah satu contoh metode optimasi adalah *Particle Swarm Optimization* (PSO). Metode optimasi PSO telah dilakukan oleh Syafira dan kawan-kawannya terkait analisis kecacatan produk menggunakan *Naïve Bayes* dan PSO. Hasil dari penelitian tersebut menunjukkan jika tingkat akurasi mencapai angka 84,38% untuk percobaan menggunakan *Naïve Bayes*, tetapi jika ditambah dengan PSO, tingkat akurasi bertambah naik sehingga menjadi 88,62%. Pada penelitian tersebut membuktikan jika PSO dapat membantu dalam menaikkan tingkat akurasi dari algoritma *Naïve Bayes* (Syafira et al., 2021).

Dari penjelasan tersebut, maka penulis berencana akan meneliti mengenai Klasifikasi Berita Hoax Konflik Palestina-Israel menggunakan *Multinomial Naïve Bayes* dengan optimasi *Particle Swarm Optimization*. Dataset yang diambil sebanyak 250 data berbahasa Inggris dari website berita online internasional BBC World. Data

yang diambil berkisar pada bulan Oktober 2023 hingga bulan Januari 2024.

METODE PENELITIAN

Penelitian ini melibatkan beberapa tahapan penelitian yang digambarkan dalam diagram alur di bawah ini:



Gambar 1. Diagram alur tahapan penelitian

Pada gambar 1 di atas merupakan tahapan penelitian yang dilakukan pada penelitian ini, dimulai dengan akuisisi dan pelabelan data hingga diakhiri pada tahap skema pengujian.

a. Akuisisi dan Pelabelan Data

Tahap akuisisi atau pengambilan data merupakan tahapan awal dari penelitian ini. Pengambilan data dilakukan dengan metode *web scraping* melalui *library* BeautifulSoup yang fungsi dari *library* tersebut adalah mengambil data dari website.

Adapun batasan dari data yang diambil adalah data berita dari website BBC.com dalam lingkup *world* internasional. Kata kunci yang digunakan ialah 'gaza', 'israel', dan 'hamas' dalam rentang waktu bulan Oktober 2023 – bulan Januari 2024. Data yang diambil khusus data berbahasa inggris sebanyak 250 data.

Setelah berhasil dilakukan *scraping* dataset, selanjutnya data yang ada dilakukan pelabelan secara manual dengan website *fact checker*, seperti Snopes.com dan Politifact.com. Cara lain melakukan pelabelan secara manual adalah dengan menggunakan gambar di dalam berita kemudian dicek melalui platform Google Lens, maka seluruh informasi berita lain yang menggunakan gambar yang sama akan muncul. Pelabelan dibagi menjadi 2 kategori, yaitu hoax dan non hoax.

b. Preprocessing Data

Tahapan selanjutnya dari penelitian adalah tahap *preprocessing* data yang menjadi tahap untuk membersihkan data mentah menjadi data yang siap dilakukan proses klasifikasi. Terdapat beberapa tahapan *preprocessing* dalam klasifikasi teks, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*.

1. Cleaning
   Tahap pertama dari tahapan *preprocessing* data adalah tahap *cleaning*, yaitu membersihkan seluruh data mentah dari bentuk-bentuk tanda baca, simbol, dan elemen lain selain alfabet a-z.
2. Case Folding
   Tahapan kedua adalah *case folding* yang berarti mengganti seluruh huruf kapital menjadi *lowercase*.
3. Tokenizing
   Tahap *tokenizing* merupakan tahap untuk memisahkan kalimat menjadi kata per kata.
4. Stopword Removal
   Tahap *stopword removal* merupakan tahapan keempat dari tahap *preprocessing* data. Tahapan ini berarti tahap untuk menghapus kata yang tidak mencakup informasi penting dan biasanya merupakan kata konjungsi yang tidak memiliki hubungan dengan tahapan selanjutnya.
5. Stemming

Tahap ini menjadi tahap terakhir dari *preprocessing* data. *Stemming* merupakan tahap untuk mengubah kata berimbuhan menjadi kata dasar.

c. Pembobotan Data

Tahapan pembobotan data merupakan tahapan penting pada penelitian ini. Pembobotan data berupa teks atau *word embedding* merupakan proses untuk memberikan bobot pada setiap elemen kata. Data hasil pra proses diubah menjadi bentuk numerik agar dapat diproses oleh komputer. Metode yang digunakan dalam tahapan ini adalah *Term Frequency – Inverse Document Frequency* (TF-IDF). Rumus dari metode TF-IDF dapat dilihat pada persamaan 1.

Perhitungan bobot suatu kata harus melewati perhitungan awal nilai TF sebuah kata. Cara menghitung nilai TF didasarkan pada bobot t dalam dokumen. Untuk perhitungan IDF dapat dilakukan melalui perhitungan banyak dokumen yang memiliki suatu kata. Perhitungan bobot dapat dilakukan setelah melalui perkalian dari TF dan IDF.

d. Pemodelan Klasifikasi dan Optimasi

Tahapan ini merupakan tahap klasifikasi model menggunakan algoritma *Multinomial Naïve Bayes* ditambah dengan metode optimasi dari *Particle Swarm Optimization*. Dalam hal ini, metode optimasi dari PSO digunakan sebagai hyperparameter tuning yang memiliki pengaruh terhadap hasil akurasi suatu penelitian. Berikut adalah langkah-langkah untuk melakukan klasifikasi dan optimasi dalam penelitian ini:

1. Menginisialisasi parameter PSO, yaitu *lower bound, upper bound*. Menginisialisasi juga terhadap kecepatan dan posisi partikel secara acak.
2. Perhitungan nilai *fitness* yang didasarkan pada perhitungan dari algoritma *Multinomial Naïve Bayes*.
3. Memperbarui nilai Pbest yang didasarkan pada hasil dari suatu partikel dan Gbest yang didasarkan pada hasil keseluruhan partikel.
4. Memperbarui posisi partikel dan kecepatan partikel berdasarkan nilai inersia, nilai koefesien akselerasi (c1 dan c2), serta nilai r1 dan r2-nya. Berikut adalah rumus dari pembaruan kecepatan masing-masing partikel:

$$v = \omega vn + c1r1\,(pbest - xn\,) + c2r2\,(gbest - xn\,) \quad (1)$$

Berikut adalah rumus persamaan matematika dari pembaruan posisi partikel:

$$xn{+}1 = xn + vn{+}1 \quad (2)$$

5. Melakukan pengecekan terhadap *stopping criteria* yang didasarkan pada maksimum iterasi, nilai *fitness* yang telah mencapai target atau nilai *fitness* pahda iterasi sebelumnya dan iterasi saat ini tidak memiliki selisih yang signifikan. Apabila *stopping criteria* belum terpenuhi, maka proses PSO akan melakukan perulangan dari poin B.

e. Skenario Pengujian

Tahapan ini merupakan tahap terakhir dari penelitian. Pengujian dilakukan dengan evaluasi *confusion matrix* yang akan dibagi menjadi dua tipe, yaitu dengan *Multinomial Naïve Bayes* dan algoritma *Multinomial Naïve Bayes* ditambah dengan metode optimasi

dari *Particle Swarm Optimization*. Penulis juga membagi skenario pengujian berdasarkan pembagian dari data latih dan data uji. Berikut adalah skenario pembagian untuk pengujiannya.

Tabel 1. Skenario Pengujian

No	Aspek	Data Latih	Data Uji
1.	*Multinomial Naïve Bayes*	90%	10%
2.		80%	20%
3.		70%	30%
4.	*Multinomial Naïve Bayes + Particle Swarm Optimization*	90%	10%
5.		80%	20%
6.		70%	30%

Pada tabel 1 di atas merupakan pembagian dari skenario pengujian pada penelitian ini. Skenario melibatkan perhitungan *confusion matrix* untuk menentukan skenario mana yang memiliki tingkat akurasi paling tinggi. Dari evaluasi tersebut juga dapat diketahui nilai *precision*, *recall, f1-score,* dan *accuracy* masing-masing skenario pengujian.

HASIL DAN PEMBAHASAN

a. Akuisisi dan Pelabelan Data

Tahap pertama yang dilakukan adalah akuisisi dan pelabelan data. Pengambilan data diambil menggunakan teknik *web scraping* dengan kata kunci ‘hamas’, ‘israel’, dan ‘gaza. Atribut data yang diambil hanya berupa data judul berita dan tanggal berita tersebut diterbitkan. Berita yang diambil dibatasi pada kurun waktu bulan Oktober 2023 hingga Januari 2024 sebanyak 250 data.

Dataset yang telah berhasil diambil kemudian diunduh dalam bentuk format CSV yang kemudian akan dilakukan pelabelan secara menual oleh penulis. Masing-masing data diberi keterangan hoax atau non hoax. Berikut adalah contoh dari pelabelan manual:

Tabel 2. Skenario Pengujian

Judul Berita	Pelabelan
IDF says 24 soldiers killed in Gaza in one day	Non hoax
Israel Gaza: Hamas raped and mutilated women on 7 October, BBC hears	Hoax
McDonald's hit by Israel-Gaza 'misinformation'	Hoax
More than 25,000 now killed in Gaza since Israel offensive began, Hamas-run health ministry says	Non hoax
Hamas Accused Israel of Hitting Gaza ‘safe zone’ killing 14	Non hoax

Pada tabel 2 di atas merupakan contoh beberapa dataset yang telah dilakukan pelabelan secara manual. Setiap data berbahasa inggris yang telah diambil akan dilakukan pelabelan yang masuk ke dalam dua kategori, yaitu hoax dan non hoax.

b. Preprocessing Data

Pada tahapan *preprocessing* ini seluruh data mentah akan dibersihkan akan dapat diproses oleh komputer. Tahapan ini memanfaatkan *library* dari Python, seperti nltk, string, numpy, dan lainnya. Berikut adalah tahapan dari *preprocessing* data:

1. Cleaning
   Tahap *cleaning* merupakan tahap membersihkan seluruh simbol selain alfabet a-z.
2. Case Folding
   Pada tahap ini seluruh dataset yang memiliki huruf kapital akan diubah menjadi *lower case*.
3. Tokenizing
   Pada tahap *tokenizing*, seluruh kalimat akan dipisahkan kata per kata.
4. Stopword Removal
   Tahapan *stopword removal* berarti seluruh kata yang

tidak memiliki hubungan atau pengaruh dengan tahapan selanjutnya akan dihapus.

5. Stemming
   Stemming berarti seluruh kata berimbuhan dalam dataset akan dikembalikan ke kata dasar.

	Title	Date	Label
0	[unrwa, claim, un, chief, aid, plea, staff, ac...	1/27/2024	non hoax
1	[israel, gaza, negoti, delay, un, vote, urg, p...	12/20/2023	non hoax
2	[sunak, israelgaza, war, mani, civilian, die]	12/19/2023	non hoax
3	[israel, turn, water, back, gaza, lord, cameron]	1/9/2024	hoax
4	[unrwa, claim, uk, halt, aid, un, agenc, alleg...	1/26/2024	non hoax
...	...	...	...
245	[uk, scientist, stay, gaza, famili, lifechang,...	1/3/2024	non hoax
246	[israel, fight, south, africa, gaza, genocid, ...	1/2/2024	non hoax
247	[israelgaza, war, hezbollah, leader, say, alar...	1/2/2024	non hoax
248	[israel, say, war, gaza, expect, continu, thro...	1/1/2024	non hoax
249	[israel, gaza, un, warn, aid, system, could, c...	1/31/2024	non hoax

Gambar 2. Hasil *Preprocessing*

Pada gambar 2 di atas merupakan hasil dari *preprocessing* setelah melewati lima tahapan, yaitu *cleaning, case folding, tokenizing, stopword removal*, dan *stemming*.

c. Pembobotan Data

Metode pembobotan data yang digunakan dalam penelitian adalah TF-IDF yang akan digunakan pada kolom *'Title'*, sedangkan kolom 'Label' akan diproses sebagai *label encoding* yang tujuannya untuk mengubah data label dari string menjadi numerik. Berikut adalah contoh output dari *label encoding*:

```
array([1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1,
       0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0,
       1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1,
       1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1,
       0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1,
       0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1,
       1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0,
       1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0,
       1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1,
       0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1,
       1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0,
       0, 0, 1, 1, 1, 1, 1, 1])
```

Gambar 3. Output Proses *Label Encoding*

Setelah seluruh kolom label diubah menjadi numerik, maka dilakukan proses pembobotan terhadap token yang telah dipisahkan. Proses pembobotan ini bertujuan agar kata memiliki jumlah bobotnya masing-masing dan dapat dilanjutkan ke proses klasifikasi. Berikut adalah output proses dari pembobotan per kata dalam dataset:

	abc	abl	absolut	accus	act	action	admit	advoc	africa	african	...	written	yahel	yahya	year	yearold	yemen	young	youni	yousaf	zone
0	0.0	0.0	0.0	0.331052	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
245	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
246	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.437078	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
247	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
248	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
249	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Gambar 4. Output Proses Pembobotan Data

Pada gambar 4 di atas merupakan hasil dari pembobotan data dengan metode TF-IDF. Setiap kata yang telah melewati tahapan *preprocessing* akan memiliki bobot.

d. Pembagian Data

Tahapan selanjutnya adalah pembagian dataset ke dalam dua kategori, yaitu data uji dan data latih. Fungsi data uji untuk proses pengujian, sedangkan fungsi data latih untuk melatih model algoritma yang dipakai agar dapat membuat prediksi akurat.

Sesuai dengan skenario pengujian yang telah tercantum pada tabel 1, terdapat perbandingan persentase data latih dan data uji. Oleh karena itu, pada beberapa skenario perlu diubah pembagian datanya agar dapat menemukan tingkat akurasi mana yang paling tinggi dalam persentase pembagian data tertentu.

e. Model Multinomial Naïve Bayes

Penelitian ini menggunakan model algoritma *Multinomial Naïve Bayes* untuk klasifikasi teks. Dibandingkan dengan tipe *Naïve Bayes* lainnya, *Multinomial Naïve Bayes* memiliki akurasi yang lebih tinggi. Penelitian ini tidak langsung menggunakan *library Multinomial Naïve Bayes* dari Python, tetapi melewati proses secara manual yang melibatkan beberapa langkah-langkah, yaitu inisialisasi parameter alpha, menghitung probabilitas label hoax dan non hoax, menghitung

probabilitas masing-masing fitur, dan menghitung probabilitas akhir.

Selain langkah-langkah tersebut, pada penelitian ini juga memanfaatkan *class* yang sebelumnya telah dibuat dengan cara memanggil untuk melatih modelnya berdasarkan data latih yang sebelumnya telah dibagi. Model *Multinomial Naïve Bayes* akan dijalankan atau diproses beberapa kali sesuai dengan skenario pengujian yang telah dijabarkan.

f. Model Particle Swarm Optimization

Selain menggunakan model *Multinomial Naïve Bayes*, pada penelitian ini percobaan juga akan ditambahkan dengan menggunakan metode optimasi dari *Particle Swarm Optimization*. PSO digunakan dalam pencarian nilai optimal parameter alpha dari model *Multinomial Naïve Bayes* agar mendapatkan hasil akurasi lebih tinggi.

Terdapat beberapa langkah-langkah dalam prosesnya, yaitu pembuatan fungsi objektif PSO yang berisi mengenai perhitungan akurasi metode *Multinomial Naïve Bayes*, kemudian inisialisasi posisi dan kecepatan partikel, perhitungan dari PSO hingga memperoleh iterasi maksimum yang diinginkan. Pada penelitian ini, penulis menggunakan batas iterasi maksimum sebesar 100 dengan jumlah partikel sebanyak 100 pula. Hasil dari output pada model *Particle Swarm Optimization* ini adalah pencarian nilai alpha paling terbaik di antara yang lain.

g. Hasil Pengujian

Setelah melewati tahapan akuisisi hingga klasifikasi, maka aspek-aspek berupa *precision, recall, f1-score,* dan akurasi dapat ditemukan. Berikut adalah perbandingan hasil sesuai skenario pengujian yang sebelumnya telah dijabarkan.

Tabel 3. Perbandingan Hasil Pengujian

Aspek	Data Latih	Data Uji	*Precision*	*Recall*	*F1-Score*	Akurasi
MNB	90%	10%	88%	57%	55%	76%
	80%	20%	51%	50%	48%	72%
	70%	30%	46%	49%	45%	69%
MNB + PSO	90%	10%	80%	80%	80%	84%
	80%	20%	76%	80%	77%	82%
	70%	30%	69%	65%	66%	76%

Pada tabel 3 di atas, skenario 1, 2, dan 3 menggunakan model algoritma *Multinomial Naïve Bayes* dengan pengujian pertama pada pembagian data latih dan data uji sebesar 90%:10% menghasilkan akurasi sebesar 76%, pembagian data sebesar 80%:20% menghasilkan akurasi sebesar 72%, dan pembagian data sebesar 70%:30% menghasilkan akurasi sebesar 69%

Pada skenario selanjutnya, yaitu skenario 4, 5, dan 6 menggunakan model algoritma *Multinomial Naïve Bayes* ditambah dengan optimasi dari *Particle Swarm Optimization*. Pengujian dengan pembagian data 90%:10% menghasilkan akurasi sebesar 84%, sedangkan pengujian dengan pembagian 80%:20% menghasilkan akurasi sebesar 82%, dan pengujian dengan pembagian 70%:30% menghasilkan akurasi sebesar 76%.

Dari penjelasan tersebut, dapat terlihat jika seluruh percobaan dengan menggunakan tambahan optimasi *Particle Swarm Optimization*, hasil akurasinya akan meningkat. Percobaan pembagian data latih dan data uji pun memengaruhi hasil tingkat akurasi dari percobaan. Dari keseluruhan percobaan, akurasi tertinggi dari *Multinomial Naïve Bayes* didapat pada pembagian data 90%:10% dengan akurasi sebesar 76%, sedangkan percobaan dengan tambahan optimasi, akurasi tertinggi juga didapat pada pembagian data 90%:10% sebesar 84%.

SIMPULAN

Setelah melakukan tahapan proses percobaan pada penelitian, terdapat beberapa kesimpulan yang bisa diambil, yaitu pembagian jumlah data latih dan data uji memiliki pengaruh dalam menghasilkan akurasi. Dari penelitian tersebut akurasi tertinggi didapatkan pada pembagian 90% untuk data latih dan 10% untuk data uji. Selanjutnya, metode optimasi *Particle Swarm Optimization* juga mampu dalam meningkatkan hasil akurasi. Tingkat akurasi tertinggi dalam percobaan *Multinomial Naïve Bayes* saja berada pada angka 76% dan percobaan yang ditambah dengan metode optimasi *Particle Swarm Optimization* berada pada angka 84%.

DAFTAR PUSTAKA

Batoebara, M. U., Suyani, E., & Nuraflah, C. A. (2020). Literasi Media dalam Menanggulangi Berita Hoaks (Studi Pada Siswa SMKN 5 Medan). *Jurnal Warta Edisi 63*, *14*, 34–41. http://jurnal.dharmawangsa.ac.id/index.php/juwarta/article/download/541/530

Febriyanty, N. E., Hariyadi, M. A., & Crysdian, C. (2023). Hoax Detection News Using Naïve Bayes and Support Vector Machine Algorithm. *International Journal of Advances in Data and Information Systems*, *4*(2), 191–200. https://doi.org/10.25008/ijadis.v4i2.1306

Muttaqin, M. F., Bukhori, T., Yanto, Y., Agustina, N., & Naseer, M. (2023). Sistem Prediksi Berita Palsu Tentang Virus Covid-19 Menggunakan Algoritma Support Vector Machine (Svm). *Naratif : Jurnal Nasional Riset, Aplikasi Dan Teknik Informatika*, *5*(1), 26–33. https://doi.org/10.53580/naratif.v5i1.187

Sahid, M. (2023). Penggunaan Media Sosial Dalam Peningkatan Pendaftar Mahasiswa Baru. *Jurnal Inovasi Penelitian*, *3*(8), 7417–7428.

Syafira, D., Suwilo, S., & Sihombing, P. (2021). *Naive Bayes Algorithm Implementation Based on Particle Swarm Optimization in Analyzing the Defect Product Naive Bayes Algorithm Implementation Based on Particle Swarm Optimization in Analyzing the Defect Product*. https://doi.org/10.1088/1742-6596/1845/1/012020

Wulandari, T., & Candra, M. F. (2024). KOMUNIKASI PROFETIK DI MEDIA SOSIAL DALAM KONFLIK PALESTINA ISRAEL. *LINIMASA : JURNAL ILMU KOMUNIKASI*, *VII*(I), 111–121.