



Analisis performa metode Support Vector Machine untuk klasifikasi dataset aroma tahu berformalin

La Saiman^{a,1}, Ramdan Satra^{a,2}

^a Universitas Muslim Indonesia, Jl. Urip Sumoharjo KM.05, Makassar 90231, Indonesia

¹ lasaimanode8@gmail.com; ² ramdan@umi.ac.id;

INFORMASI ARTIKEL	ABSTRAK
Diterima : 15 – 05 – 2021 Direvisi : 26 – 06 – 2021 Diterbitkan : 31 – 07 – 2021	Bermunculan adanya kasus penyalahgunaan formalin sebagai bahan pengawet untuk tahu dan ketersediaan cara deteksi kandungan formalin yang tergolong tidak praktis dan membutuhkan tenaga ahli khusus. Penelitian ini bertujuan untuk melakukan analisis performa suatu metode. Pengolahan data tersebut digunakan untuk memprediksi terdapatnya kandungan formalin pada tahu melalui aroma yang dikeluarkan. Teknik klasifikasi yang digunakan yaitu dengan menggunakan pengukuran performa (<i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , dan <i>F-Measure</i>) menggunakan Metode <i>Support Vector Machine</i> . Inputan yang digunakan yaitu atribut data aroma tahu berformalin, dan tidak berformalin. Hasil penelitian ini menunjukkan metode SVM mempunyai kinerja yang cukup baik. Berdasarkan hasil pengujian menggunakan confusion matrix didapatkan hasil uji performa Metode SVM <i>accuracy</i> 67%, <i>precision</i> 69%, <i>recall</i> 67%, dan <i>f-measure</i> 66% dengan Nilai K=5.
Kata Kunci: SVM Klasifikasi <i>Accuracy</i> <i>Precision</i> <i>F-Measure</i>	 

I. Pendahuluan

Makanan merupakan salah satu yang menjadi kebutuhan manusia, Tahu merupakan makanan yang mengandung gizi yang banyak, terbuat dari olahan yang kaya akan protein nabati yang banyak dikonsumsi oleh masyarakat Indonesia. Tahu adalah makanan yang dibuat dari kacang kedelai yang difermentasikan dan diambil sarinya. Berbeda dengan tempe yang asli Indonesia, tahu berasal dari negara China. Tahu sangat populer di kalangan masyarakat Indonesia bukan hanya karena harganya yang murah, tetapi juga mudah diolah menjadi makanan lain.

Berdasarkan data yang diperoleh dari Biro Pusat Statistik dan Survei Sosial Ekonomi Nasional.[1], tingkat konsumsi tahu di Indonesia mencapai 0,152 kg/kapita/minggu, dan tahu masuk kategori bahan makanan penting. Jumlah ini lebih dari empat kali lipat jika dibandingkan dengan tingkat konsumsi daging ayam dan daging sapi yang merupakan sumber protein hewani. Hal tersebut disebabkan harga tahu dan tempe jauh lebih terjangkau jika dibandingkan dengan harga daging.[2].

Tahu merupakan bahan pangan dengan kandungan protein yang tinggi dan kadar air mencapai 85%, sehingga tahu tidak dapat bertahan lama. Kerusakan tahu ditandai dengan bau asam dan berlendir. Berdasarkan hal tersebut membuat beberapa oknum pedagang dan pembuat tahu yang mulai memasang formalin pada tahu buatannya, karena proses pembusukan tahu yang cepat.[3] Tahu mulai mengalami kerusakan sehari setelah produksi. Perendaman tahu dalam air yang diberi formalin akan membuat tahu menjadi lebih keras dan kenyal, sehingga tidak mudah hancur dan tahan terhadap mikroorganisme, sehingga tahan lama dan dapat bertahan hingga tujuh hari.[4] Formalin dapat bereaksi cepat dengan lapisan lendir saluran pencernaan dan saluran pernafasan. Pemakaian pada makanan dapat mengakibatkan keracunan pada tubuh manusia, yaitu rasa sakit perut yang akut disertai muntah-muntah, timbulnya depresi susunan syaraf atau kegagalan peredaran darah dan dalam Peraturan Menteri Kesehatan No. 1168/MenKes/PER/X/1999.[5]–[8].

Ini sangat berpengaruh buruk terhadap seluruh konsumen tahu yakni masyarakat Indonesia yang menjadikan tahu sebagai makanan utama dalam keseharian. Kasus penyalahgunaan formalin sebagai bahan pengawet makanan banyak dilakukan di Indonesia. Hasil uji Badan Pengawas Obat dan Makanan (BPOM) menyatakan dari 700 sampel produk makanan yang diambil dari Jawa, Sulawesi Selatan dan Lampung, 56% mengandung formalin. Bahaya dari penyalahgunaan formalin sebagai bahan pengawet makanan perlu adanya perhatian khusus, karena penggunaan formalin sebagai bahan pengawet makanan dapat menyebabkan

beberapa penyakit, diantaranya efek kesehatan manusia langsung terlihat akut seperti (iritasi, alergi, mual, muntah, sakit perut dan pusing), dan efek kronik yaitu efek pada kesehatan manusia terlihat terkena dalam jangka waktu yang lama dan berulang, seperti gangguan pencernaan, hati, ginjal, pankreas, sistem saraf pusat.[4] perlu dilakukan analisis lebih lanjut terkait dengan tahu yang menggunakan formalin. Proses analisa data dapat menggunakan salah satu metode dalam *data mining*. *Data mining* adalah teknik yang memanfaatkan data dalam jumlah yang besar untuk memperoleh informasi berharga yang sebelumnya tidak diketahui dan dapat dimanfaatkan untuk pengambilan keputusan.[9], [10] Salah satu metode statistik dalam *data mining* yang dapat diterapkan untuk melakukan klasifikasi adalah SVM. SVM merupakan suatu teknik untuk menemukan *hyperplane* yang bisa memisahkan dua set data dari dua kelas yang berbeda.[11], [12] SVM memiliki kelebihan diantaranya adalah dalam menentukan jarak menggunakan *support vector* sehingga proses komputasi menjadi cepat.[13]

Pada penelitian sebelumnya telah dilakukan penelitian mengenai klasifikasi tingkat keganasan kanker dengan menggunakan metode regresi logistik dan SVM yang akhirnya diperoleh hasil bahwa tingkat akurasi menggunakan SVM lebih tinggi, yaitu sebesar 98,11%.[13] penelitian lain juga melakukan analisa dataset multiclass yang membandingkan hasil performa tujuh metode klasifikasi yaitu K-Nearest Neighbor (KNN), Naive Bayes Classifier (NBC), Support Vector machine (SVM), Neural Network (NN), Random Forest Classifier (RFC), Ada Boost Classifier (ABC) dan Quadratic Discriminant Analysis (QDC) dan disimpulkan bahwa meskipun secara keseluruhan semua metode tidak mendapatkan nilai performa yang sangat baik, namun metode SVM memiliki nilai performa yang lebih baik dibandingkan metode klasifikasi lainnya dalam studi kasus klasifikasi dataset multiclass busur panah dengan akurasi 73%.

Pada penelitian ini penulis tertarik untuk menganalisis performa metode SVM dengan menerapkan metode tersebut untuk mengklasifikasikan tahu berformalin dengan tahu yang tidak berformalin, data yang digunakan adalah data rekaman dari sensor gas (Sensor MQ-2, MQ-3, MQ-4, MQ-5, MQ-7, MQ-8, MQ-9, DHT11), performa yang diukur dalam penelitian ini yaitu *Accuracy*, *Precision*, *Recall*, dan *F-Measure*. [14] dengan sampel yaitu tahu putih. Pengujian dilakukan menggunakan *cross validation*, *Cross validation* merupakan teknik validasi dari pengembangan model split validation dimana validasinya mengukur training error dengan menguji dengan test data atau data uji.[15]

II. Metode

Tahapan Penelitian.

A. Pengumpulan Data

Pengumpulan data nantinya akan dilakukan di Laboratorium Terpadu Fakultas Ilmu Komputer Universitas Muslin Indonesia selama 15 hari, berikut rincian proses pengumpulan data pada penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. Rincian Proses Pengumpulan Data

Kategori	Perlakuan
Berformalin	Tahu direndam dalam formalin
	Tahu diolesi formalin
	Tahu direnam larutan formalin bercampur air
Tidak Berformalin	Tahu direndam dengan air diganti tiap hari
	Tahu direndam dengan air tidak diganti
	Tahu tidak direndam

Data yang dikumpulkan dikelompokkan berdasarkan kategori tahu mengandung formalin dan tahu yang tidak mengandung formalin, dan diperlakukan menjadi 6 variabel/sampel yakni: Tahu yang direndam keseluruhan bagian dengan formalin, Tahu yang hanya diolesi bagian luarnya dengan formalin dan tidak direndam, Tahu direndam larutan yang telah dicampurkan air dan formalin dengan perbandingan 50:50, Tahu yang hanya direndam dalam air dan diganti airnya tiap satu hari, Tahu yang hanya direndam dalam air tanpa diganti airnya sama sekali, Tahu yang dibiarkan tanpa air.

Setelah data dikelompokkan, maka data tersebut masuk ke tahap preprocessing data, tahap ini dijalankan rutin selama 15 hari.

B. Preprocessing Data

Data yang di peroleh langsung dari penelitian selama 15 hari, diolah terlebih dahulu melalui tahap-tahap data cleaning, data integration, data selection, dan data transformation:

1) Data Cleaning

- 2) *Data Integration*
- 3) *Data Selection*
- 4) *Data Transformation*

C. Analisis Performa Metode Support Vector Machine

Melakukan pengukuran performa Algoritma SVM pada dataset tahu berformalin sebagai berikut:

- 1) *Melakukan transformasi data.*
- 2) *Menentukan fungsi pemisah dengan metode multi kelas.*
- 3) *Menentukan nilai-nilai parameter $K=5$.*
- 4) *Memilih nilai parameter K terbaik.*
- 5) *Menghitung nilai ketepatan klasifikasi terbaik*

D. Pengambilan Keputusan

Kesimpulan diambil berdasarkan hasil analisis performa metode SVM pada dataset tahu berformalin.

III. Hasil dan Pembahasan

Pada tahap implementasi menggambarkan beberapa tahap yang terdiri dari implementasi kebutuhan perangkat lunak, perangkat keras, pengumpulan data, *preprocessing data*, metode *support vector machine*, pengujian performa metode.

A. Implementasi Kebutuhan Perangkat Lunak

- 1) *Arduino IDE*
- 2) *Browser*




B. Implementasi Kebutuhan Perangkat Keras














- 1) *Laptop*
- 2) *Processor Intel® Core™ i5.*
- 3) *RAM 8,00 GB.*
- 4) *Sensor MQ-2 (H, LPG, CO, -OH, C3H8), MQ-3 (CH4, LPG, CO, -OH, C6H14, C6H6), MQ-4 (CH4, LPG, CO, -OH, SMOKE), MQ-5 (H2, LPG, CH4, CO, SMOKE), MQ-7 (H2, LPG, CH4, CO, -OH), MQ-8 (H2, LPG, CH4, CO, -OH), MQ-9 (LPG, CH4, CO,), DHT11 (RH, ERH).*
- 5) *Arduino MEGA 2560*
- 6) *Fun Motherboard*
















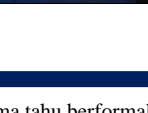
C. Implementasi Pengumpulan Data














Pengumpulan data dilakukan di Laboratorium Terpadu Fakultas Ilmu Komputer Universitas Muslin Indonesia selama 15 hari, berikut rincian proses pengumpulan data pada penelitian ini ditunjukkan pada [Tabel 2](#).













Tabel 2. Pengumpulan Data







Hari	Kategori	Perlakuan	Dokumentasi
Ke-1	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	

Hari	Kategori	Perlakuan	Dokumentasi
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	
		Tahu direndam dengan air tidak diganti	
		Tahu tidak direndam	
Ke-2	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	
		Tahu direndam dengan air tidak diganti	
		Tahu tidak direndam	
Ke-3	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	

Hari	Kategori	Perlakuan	Dokumentasi
		Tahu direndam dengan air tidak diganti	
		Tahu tidak direndam	
Ke-4	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	
		Tahu direndam dengan air tidak diganti	
		Tahu tidak direndam	
Ke-5	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	
		Tahu direndam dengan air tidak diganti	
		Tahu tidak direndam	
Ke-6	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	

Hari	Kategori	Perlakuan	Dokumentasi
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	
		Tahu direndam dengan air tidak diganti	
		Tahu tidak direndam	Membusuk dan dibuang
Ke-7	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	
		Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
Ke-8	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	Membusuk dan dibuang
		Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
Ke-9	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	Membusuk dan dibuang
		Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan

Hari	Kategori	Perlakuan	Dokumentasi
			dibuang
Ke-10	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	Membusuk dan dibuang
		Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
Ke-11	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	Membusuk dan dibuang
		Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
Ke-12	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	Membusuk dan dibuang
		Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
Ke-13	Berformalin	Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
	Tidak Berformalin	Tahu direndam dengan air diganti tiap hari	Membusuk dan dibuang

Hari	Kategori	Perlakuan	Dokumentasi
Ke-14	Berformalin	Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
		Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
		Tidak Berformalin	Tahu direndam dengan air diganti tiap hari
Ke-15	Berformalin	Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang
		Tahu direndam dalam formalin	
		Tahu diolesi formalin	
		Tahu direndam larutan formalin bercampur air	
		Tidak Berformalin	Tahu direndam dengan air diganti tiap hari
Ke-15	Tidak Berformalin	Tahu direndam dengan air tidak diganti	Membusuk dan dibuang
		Tahu tidak direndam	Membusuk dan dibuang

Data yang dikumpulkan dikelompokkan berdasarkan kategori tahu mengandung formalin dan tahu tidak mengandung formalin, dan diperlakukan menjadi 6 variabel/sampel yakni: Tahu yang direndam keseluruhan bagian dengan formalin, Tahu yang hanya diolesi bagian luarnya dengan formalin dan tidak direndam, Tahu direndam larutan yang telah dicampurkan air dan formalin dengan perbandingan 50:50, Tahu yang hanya direndam dalam air dan diganti airnya tiap satu hari, Tahu yang hanya direndam dalam air tanpa diganti airnya sama sekali, Tahu yang dibiarkan tanpa air.

Setelah data dikelompokkan, maka data tersebut masuk ke tahap preprocessing data, tahap ini dijalankan rutin selama 15 hari dan menghasilkan 91.500 data, yang ditunjukkan pada Gambar 1.

1	CO_MQ2(ppm)	Alcohol_MQ2(ppm)	Propane_MQ2(ppm)	CH4_MQ4(ppm)	Smoke_MQ4(ppm)	Temperature(C)	label
91481	3.6	1.3	1.1	1.7	0.2	1.00	0
91482	3.5	1.2	1	1.7	0.2	1.00	0
91483	3.6	1.3	1.1	1.8	0.2	1.00	1
91484	3.8	1.4	1.1	1.8	0.2	1.00	1
91485	3.6	1.3	1.1	1.8	0.2	1.00	1
91486	3.6	1.3	1.1	1.7	0.2	1.00	0
91487	3.6	1.3	1.1	1.8	0.2	1.00	1
91488	3.5	1.2	1	1.8	0.2	1.00	1
91489	4	1.4	1.1	1.8	0.2	1.00	1
91490	3.6	1.3	1.1	1.8	0.2	1.00	1
91491	3.8	1.4	1.1	1.8	0.2	1.00	1
91492	3.6	1.3	1.1	1.8	0.2	1.00	1
91493	3.8	1.4	1.1	1.8	0.2	1.00	1
91494	3.8	1.4	1.1	1.8	0.2	1.00	1
91495	3.6	1.3	1.1	1.8	0.2	1.00	1
91496	4	1.4	1.1	1.8	0.2	1.00	1
91497	3.8	1.4	1.1	1.7	0.2	1.00	1
91498	3.6	1.3	1.1	1.8	0.2	1.00	1
91499	3.8	1.4	1.1	1.8	0.2	1.00	1
91500	3.6	1.3	1.1	1.7	0.2	1.00	1
91501	3.8	1.4	1.1	1.7	0.2	1.00	1
91502							

Gambar 1. Jumlah Dataset Hingga Hari Ke-15

D. Implementasi Preprocessing

Data yang di peroleh langsung dari penelitian selama 15 hari, diolah terlebih dahulu melalui tahap-tahap data cleaning, data integration, data selection, dan data transformation:

1) Data Cleaning

Dilakukan proses *cleaning* dari data awal yang diperoleh untuk mengatasi nilai yang hilang[16]–[23], *noise* dan data yang tidak konsisten. Maka dari data awal yang ditemukan 39 kolom perbaris yang ditunjukkan pada Gambar 2. selanjtnya dilakukan proses cleaning maka didapat data 9 kolom perbaris seperti yang ditunjukkan pada Gambar 3.

Gambar 2. Dataset Sebelum Proses *Cleaning*

Gambar 3. Dataset Setelah Proses *Cleaning*

2) Data Integration

Dilakukan proses *Integration* untuk menggabungkan data yang diambil sejak Rabu, 7 April 2021 – Kamis, 22 April 2021, sehingga terampung total 91.500 baris data yang ditunjukkan pada Gambar 4.

Gambar 4. Dataset Setelah Proses *Integration*

3) Data Selection

Data selection yang digunakan yaitu proses sampling, dari jumlah data 91.500 menjadi 45.000 data yang ditunjukkan pada Gambar 5.

	H2_MQ2(ppm)	LPG_MQ2(ppm)	CO_MQ2(ppm)	Alcohol_MQ2(ppm)	Propane_MQ2(ppm)	CH4_MQ2(ppm)	Smoke_MQ2(ppm)	Temperature(C)	label
44973	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44974	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44975	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44976	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44977	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44978	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44979	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44980	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44981	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44982	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44983	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44984	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44985	1.4	0.7	3	1.1	0.9	1	0	1.00	1
44986	1.6	0.8	3.5	1.3	1	1	0	1.00	1
44987	1.5	0.7	3.3	1.2	1	1.1	0	1.00	1
44988	1.5	0.7	3.2	1.1	1	1	0	1.00	1
44989	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44990	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44991	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44992	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44993	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44994	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44995	1.5	0.7	3.3	1.2	1	1	0	1.00	1
44996	1.6	0.8	3.5	1.3	1	1	0	1.00	1
44997	1.5	0.7	3.2	1.1	1	1.1	0	1.00	1
44998	1.5	0.7	3.2	1.1	1	1.1	0	1.00	1
44999	1.5	0.7	3.3	1.2	1	1.1	0	1.00	1
45000	1.5	0.7	3.3	1.2	1	1.1	0	1.00	1
45001	1.5	0.7	3.3	1.2	1	1.1	0	1.00	1

Gambar 5. Dataset Setelah Proses Selection

4) Data Transformation

Data transformation yang digunakan yaitu proses scaling untuk membuat numerical data pada dataset dengan tetap mempresentasikan nilai yang sama.

E. Implementasi Metode SVM

Implementasi metode SVM dilakukan pada dataset yang telah melewati proses *preprocessing data* dengan jumlah total 45.000 data. dataset akan dibagi terdiri dari 80% *data training* dan 20% *data testing*.

Kesimpulan diambil berdasarkan hasil pengujian performa (*Accuracy, Precision, Recall, and F-Measure*) metode Support Vector Machine pada dataset aroma tahu berformalin.

Pembahasan

Pada penelitian ini menggunakan bahasa pemrograman Python dan library scikit learn. Mulai dari tahap load dataset aroma tahu berformalin, implementasi metode support vector machine, visual boxplot, visual diagram batang, visual diagram pencar, visual diagram lingkaran, Visual digram grafus, hingga proses perhitungan nilai performa (*Accuracy, Precision, Recall, and F-Measure*).

Dataset yang digunakan sebagai obyek penelitian ini merupakan dataset aroma tahu berformalin. Dimana datanya berupa nilai yang didapat dari beberapa sensor yang digunakan dalam penelitian. Dataset tersebut memiliki 2 label yaitu 1 yang berarti tahu berformalin dan 0 yang berarti tahu tidak berformalin, yang mana data tersebut diambil pada tanggal 7 April 2021 – 22 April 2021, berjumlah 45.000 data, dan diunggah pada tanggal 11 Desember 2020 di repository Kaggle, berikut adalah tampilan dataset yang telah diunggah dan telah diberikan penamaan untuk mempermudah saat implementasi ke bahasa pemrograman, sampel dataset yang diunggah dapat dilihat pada Tabel 3.

Tabel 3. Sampel Dataset Aroma Tahu

	H	H2	H3	H4	H5	H6	H7	H8
1	1,7	0,8	4	1,4	1,1	2,4	0	1
2	1,8	0,9	4,2	1,5	1,2	2,2	0	1
...
17554	1,6	0,8	3,7	1,3	1,1	2,5	0,5	0
17555	1,6	0,8	3,9	1,4	1,1	2,4	0,5	0
...
44.999	1,5	0,7	3,3	1,2	1	1,1	0	1
45.000	1,4	0,7	3	1,1	0,9	1	0	1

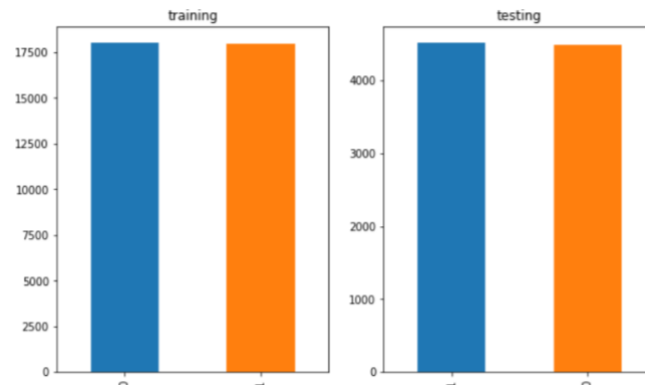
Untuk melihat rincian data berikut saya lampirkan transformasi dataset yang digunakan ditunjukkan pada Tabel 4.

Tabel 4. Rincian Dataset

	H	H2	H3	H4	H5	H6	H7	H8
Count	45.000	45.000	45.000	45.000	45.000	45.000	45.000	45.000
Mean	1,52	0,74	3,32	1,19	0,99	3,17	5,75	0,5

	H	H2	H3	H4	H5	H6	H7	H8
Std	0,14	0,08	0,43	0,14	0,09	2,89	21,86	0,5
Min	1,00	0,50	1,90	0,70	0,70	0,70	0,00	0,0
Max	1,90	0,90	4,50	1,50	1,20	20,40	262,80	1,0

Untuk memaksimalkan pengetahuan yang didapatkan *machine learning*, data yang digunakan haruslah data yang seimbang, berikut grafik jumlah data yang digunakan dalam bentuk diagram batang, ditunjukkan pada Gambar 6.



Gambar 6. Grafik Visualisasi Diagram Batang Pembagian Dataset

Pengujian performa Metode *Support Vector Machine* (SVM) dilakukan pada dataset aroma tahu berformalin yang berjumlah 45.000 data. Berdasarkan hasil pengujian performa pada metode SVM, untuk pengukuran *Accuracy*, *Precision*, *Recall*, and *F-Measure*, maka didapat hasil klasifikasi terbaik yakni *accuracy* 67%, *precision* 69%, *recall* 67%, dan *f-measure* 66%.

IV. Kesimpulan

Berdasarkan dari hasil penelitian dan pembahasan yang dilakukan maka diperoleh kesimpulan bahwa Metode *Support Vector Machine* dapat digunakan dalam proses klasifikasi terhadap dataset aroma tahu berformalin. Sebelum dilakukan proses klasifikasi data aroma terlebih dahulu melalui tahap preprocessing, dan ekstraksi fitur dan Berdasarkan hasil pengujian dengan menggunakan Metode *Support Vector Machine* pada dataset aroma tahu berformalin memperoleh nilai performa terbaik yaitu *accuracy* sebesar 67%, *precision* 69%, *recall* 67%, dan *f-measure* 66% Nilai K=5.

Daftar Pustaka

- [1] BPS, "Statistics Indonesia," 2020.
- [2] L. Vinet and A. Zhedanov, "No Title," *J. Phys. A Math. Theor.*, vol. 44, no. 8, pp. 1–8, 2011, doi: 10.1088/1751-8113/44/8/085201.
- [3] S. Kesen, "Characterization of aroma and aroma-active compounds of Turkish turmeric (*Curcuma longa*)," *J. Raw Mater. to Process. Foods 1*, pp. 13–21, 2020.
- [5] S. Hastuti, "Analisis Kualitatif Dan Kuantitatif Formaldehid Pada Ikan Asin Di Madura," *J. AGROINTEK*, vol. 4, no. 2, pp. 132–137, 2010.
- [6] P. L. Terranova, *Peraturan Menteri Kesehatan Republik Indonesia Nomor 1168/Menkes/Per/X/1999*. 1999.
- [7] M. Syarfaini, & Rusmin, "Analisis Kandungan Formalin Pada Tahu di Pasar Tradisional Kota Makassar," *J. Kesehat. Masy.*, vol. 7, no. 1, pp. 1–11, 2014.
- [8] D. Arziyah and L. Yusmita, "Analisis mutu tahu dari beberapa produsen tahu di Kota Padang," *J. Teknol. Pertan. Andalas*, vol. 23, no. 2, pp. 143–148, 2019.
- [9] A. N. Khormarudin, "Teknik Data Mining: Algoritma K-Means Clustering," *J. Ilmu Komput.*, 2016.
- [10] S. Heni and A. I. Gufroni, "Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 299–305, 2017, doi: <https://doi.org/10.25077/TEKNOSI.v3i2.2017.299-305>.
- [11] N. H. Farhat, "Photonit neural networks and learning machines the role of electron-trapping materials," *IEEE Expert. Syst. Their Appl.*, vol. 7, no. 5, pp. 63–72, 1992.
- [12] P. A. Octaviani, Y. Wilandari, and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *J. Gaussian*,

- vol. 3, no. 8, pp. 811–820, 2014.
- [13] W. Purnami, A. M. Regresi, and L. Ordinal, “Perbandingan Klasifikasi Tingkat Keganasan Breast Cancer Dengan Menggunakan Regresi Logistik Ordinal Dan Support Vector Machine (SVM),” *J. Sains Dan Seni Its*, vol. 1, no. 1, 2012.
- [14] V. Sofia, I. Atsushi, and R. Anca, “Proceedings Of The Twenty-second Midwest Artificial Intelligence And Cognitive Science Conference,” *MAICS*, vol. 710, pp. 120–127, 2011.
- [15] C. Schaffer, “Selecting A Classification Method By Cross-Validation,” *Mach. Learn.*, vol. 13, no. 1, pp. 135–143, 1993.
- [16] A. Fitria and H. Azis, “Analisis Kinerja Sistem Klasifikasi Skripsi menggunakan Metode Naïve Bayes Classifier,” *Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 102–106, 2018.
- [17] H. Azis, F. T. Admojo, and E. Susanti, “Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah,” *Techno.Com*, vol. 19, no. 3, 2020.
- [18] L. Nurhayati and H. Azis, “Perancangan Sistem Pendukung Keputusan Untuk Proses Kenaikan Jabatan Struktural Pada Biro Kepegawaian,” *Semin. Nas. Teknol. Inf. dan Multimed.*, pp. 6–7, 2016.
- [19] M. M. Baharuddin, T. Hasanuddin, and H. Azis, “Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca,” *Ilk. J. Ilm.*, vol. 11, no. 28, pp. 269–274, 2019.
- [20] H. Azis, “Klasifikasi Penyakit Jantung Menggunakan Metode K-Nearest Neighbor,” *Indones. J. Data Sci.*, pp. 1–4, 2018, doi: <https://doi.org/10.33096/ijodas.v1i1.3>.
- [21] S. Sahar, “Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Clasiffier Pada Dataset Penyakit Jantung,” *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 79–86, 2020, doi: [10.33096/ijodas.v1i3.20](https://doi.org/10.33096/ijodas.v1i3.20).
- [22] L. B. C. Tanujayaa, B. Susanto, and A. Saragiha, “Perbandingan Metode Regresi Logistik dan Random Forest untuk Klasifikasi Fitur Mode Audio Spotify,” *Indones. J. data Sci.*, vol. 1, no. 3, pp. 68–78, 2020.
- [23] Nirmalasari, “Analisis Performa Metode Naïve Bayes Classifier Dan Cross Validation Pada Data Penyakit Cardiovascular,” *Indones. J. data Sci.*, 2020.