

# **APLIKASI METODE *SUPPORT VECTOR MACHINE* UNTUK PREDIKSI KUALITAS GLISEROL MONOOLEAT**

*(Application of Support Vector Machine Method for Prediction of Glycerol Monooleat Quality)*

**Iwan Aang Soenandi**

**Fakultas Teknologi dan Ilmu Komputer Jurusan Teknik Industri  
Universitas Kristen Krida Wacana  
iwan.as@ukrida.ac.id**

## **Abstrak**

Faktor kualitas merupakan faktor yang penting untuk diperhatikan pada proses produksi Gliserol MonoOleat sehingga kepercayaan konsumen terhadap produsen tetap terjaga. Untuk menjaga kualitas tersebut dapat menggunakan metode *Data Mining* dari parameter hasil produksi, yang diolah dengan metode SVM, untuk mengklasifikasikan nilai Kualitas tersebut, yang diuji dengan analisis ROC. Dari hasil pengolahan data pada penelitian ini dengan menggunakan visualisasi grafik dari ROC, dapat dibuktikan bahwa untuk mengklasifikasi kualitas tersebut dapat menggunakan metode SVM. Berdasarkan pengujian menggunakan perbandingan dengan perangkat lunak, dengan nilai untuk SVM Kualitas *Pass* adalah 0.522 dan untuk *Not Pass* adalah 0.524 dengan nilai *TP rate (Sensitivity) Pass* adalah 0.95 dan *TP Rate (Sensitivity) Not Pass* adalah 0,92.

**Kata kunci:** kualitas, *data mining*, *support vector machine*, analisis *ROC*

## **Abstract**

*Quality is an important factor to consider in the glycerol monooleate production process so that the consumers' trust is retained. To maintain the quality of Glycerol MonoOleat, data mining method can be used based on the parameters of production output processed with SVM method. This method serves to classify the quality value being tested by using ROC analysis. The data processing result of this study, in which ROC graphical visualization was used, suggested that SVM method could be applied to classify quality. Based on the test done using comparative software, the value of SVM Quality Pass is 0.522 and 0.524 for Not Pass. TP Rate (Sensitivity) Pass is 0.95 and TP Rate (Sensitivity) Not Pass is 0.92.*

**Keywords:** *Quality, Data Mining, Support Vector Machine, ROC Analysis*

**Tanggal Terima Naskah : 05 Juni 2014**  
**Tanggal Persetujuan Naskah : 01 Juli 2014**

## **1. PENDAHULUAN**

Untuk dapat bersaing dan berkembang secara berkelanjutan dalam setiap industri, faktor kualitas sangat menentukan. Faktor kualitas merupakan faktor pengukuran kinerja dari hasil produksi selain kapasitas produksi. Oleh karena itu, pada sistem produksi dari Gliserol Ester akan dianalisis kualitas dari produk Gliserol MonoOleat yang akan diolah menggunakan metode *Support Vector Machine (SVM)*. *Pattern recognition* merupakan salah satu bidang dalam *computer science*, yang memetakan suatu data ke dalam konsep

tertentu, yang telah didefinisikan sebelumnya. Konsep tertentu ini disebut sebagai *class* atau *category*. Aplikasi *pattern recognition* sangat luas, di antaranya mengenali suara dalam sistem keamanan, membaca huruf dalam OCR, mengklasifikasikan penyakit secara otomatis berdasarkan hasil diagnosis kondisi medis pasien, dan sebagainya. Berbagai metode dikenal dalam *pattern recognition*, seperti *linear discrimination analysis*, *hidden markov model* hingga metode kecerdasan buatan, seperti *artificial neural network*. Salah satu metode yang saat ini banyak mendapat perhatian sebagai *state of the art* dalam *pattern recognition* adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) dikembangkan oleh Boser, Guyon, Vapnik, dan pertama kali dipresentasikan pada tahun 1992 di *Annual Workshop on Computational Learning Theory*. Konsep dasar SVM sebenarnya merupakan kombinasi harmonis dari teori-teori komputasi yang telah ada puluhan tahun sebelumnya, seperti *margin hyperplane*, dan demikian juga dengan konsep-konsep pendukung yang lain dalam tulisan dari Vijayakumar [1]. Akan tetapi, hingga tahun 1992 belum pernah ada upaya merangkaikan komponen-komponen tersebut. Berbeda dengan strategi *neural network* yang berusaha mencari *hyperplane* pemisah antar-*class*, SVM berusaha menemukan *hyperplane* yang terbaik pada *input space*. Prinsip dasar SVM adalah *linear classifier*, selanjutnya dikembangkan agar dapat bekerja pada *problem non-linear* dengan memasukkan konsep *kernel trick* pada ruang kerja berdimensi tinggi[2]. Perkembangan ini memberikan rangsangan minat penelitian di bidang *pattern recognition* untuk melakukan investigasi potensi kemampuan SVM secara teoritis maupun dari segi aplikasi[3]. Dewasa ini SVM telah berhasil diaplikasikan dalam persoalan dunia nyata (*real-world problems*), dan secara umum memberikan solusi yang lebih baik dibandingkan metode konvensional, seperti *artificial neural network* [4]. Tulisan ini memperkenalkan konsep dasar SVM serta membahas aplikasinya pada bidang produksi Gliserol MonoOleat untuk menentukan kualitas produk. Untuk menganalisis sensitivitas dari SVM akan digunakan analisis ROC (*Receiver Operating Characteristic*).

## 2. PERMASALAHAN

Di industri Gliserol Monooleat faktor kualitas merupakan faktor yang penting untuk diperhatikan agar hasil produksi dapat terpakai serta untuk mengurangi keluhan pelanggan. Untuk menjaga faktor kualitas tersebut akan diolah data dari lima atribut, yaitu Suhu Esterifikasi, Berat Molekul, Viskositas, Densitas, dan Waktu Esterifikasi. Dari kelima atribut tersebut dengan data percobaan yang dilakukan pada proses produksi akan dikelompokkan berdasarkan standar kualitas yang baik (*pass*) dan yang tidak baik (*not pass*). Dalam hal ini analisis untuk dapat memprediksi kualitas sangat diperlukan agar gejala perubahan atau penurunan kualitas dapat diketahui secepat mungkin sehingga memperkecil faktor produksi yang gagal atau harus diproduksi kembali.

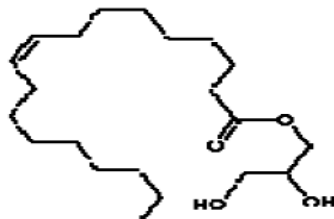
## 3. PROSES ESTERIFIKASI GLISEROL

Gliserol adalah produk samping produksi biodiesel dari reaksi transesterifikasi yang merupakan senyawa alkohol dengan gugus hidroksil berjumlah tiga buah. Gliserol (1,2,3 propanetriol) merupakan cairan yang tidak berwarna, tidak berbau, dan merupakan cairan kental yang memiliki rasa manis. Gliserol dapat dimurnikan dengan proses destilasi agar dapat digunakan pada industri makanan, farmasi, atau dapat juga digunakan untuk pengolahan air. Sebagai produk samping industri biodiesel, gliserol belum banyak diolah sehingga nilai jualnya masih rendah. Dalam bidang makanan, ester dari gliserin, terutama ester parsial (mono-dan di-gliserida) telah menjadi komponen yang sangat khusus. Produk emulsi memberikan kontribusi pengendalian atas kelembutan dari

kecantikan, juga untuk margarin. Proses esterifikasi gliserol adalah salah satu metode yang banyak digunakan memproduksi produk turunan gliserol. Dalam reaksi esterifikasi dihasilkan bermacam-macam ester yang mempunyai banyak kegunaan dan bernilai lebih tinggi. Produk dari konversi gliserol ini bersifat ramah lingkungan dan terbarukan karena bukan merupakan turunan dari minyak bumi.

#### 4. GLISEROL MONOOLEAT

Gliserol monooleat (monoolein) adalah sintesis kimia aktif permukaan yang secara luas digunakan sebagai surfaktan non-ionik dan pengemulsi. Hal ini dihasilkan oleh reaksi gliserol dan asam oleat dengan katalis untuk membentuk ester monogliserida [5]. Gliserol monooleat tidak berbahaya bagi lingkungan. Semua bahan pembuatnya dihasilkan oleh alam. Asam oleat diekstrak dari produk alami, sedangkan gliserol dapat dibuat dari lemak alami atau petrokimia [6]. Gliserol monooleat secara fisik berwujud cair berwarna jernih kekuning-kuningan atau kuning pucat. Gliserol monooleat ( $C_{21}H_{40}O_4$ ) memiliki berat molekul 356,54 gr/mol, tidak larut dalam air, sedikit larut dalam alkohol dingin, dan larut dalam alkohol panas, kloroform, eter, petroleum eter, serta larut dalam minyak serta bersifat sebagai pengemulsi pada makanan [7]. Gliserol monooleat larut baik dalam minyak [8]. Rumus struktur gliserol monooleat dapat dilihat pada gambar 1.



Gambar 1. Rumus struktur gliserol monooleat

Identik dengan surfaktan-surfaktan lain, sifat gliserol monooleat ini ditentukan oleh sifat reaktan pembentuknya. Gliserol monooleat dapat membentuk suatu mikro-emulsi di dalam air. Gliserol larut baik dalam air [5], sedangkan asam oleat tidak larut dalam air. Hydrophilic-Lipophilic Balance (HLB) dari gliserol monooleat adalah 3,8 [9].

#### 5. METODE PERCOBAAN

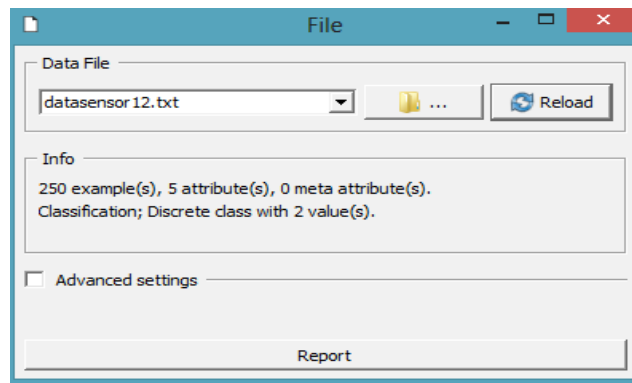
Pada tulisan ini percobaan yang dilakukan adalah reaksi esterifikasi diawali dengan mereaksikan antara asam oleat dan gliserol dengan menggunakan katalis asam sulfat pada labu leher tiga yang dilengkapi dengan kondensor dan pengaduk magnet pada rentang suhu  $220^{\circ}\text{C}$  hingga  $250^{\circ}\text{C}$ . Proses dilanjutkan dengan reaksi polimerisasi sampai suhu  $260^{\circ}\text{C}$ . Rasio asam oleat dan gliserol merupakan variabel tetap 1:3 dengan berat katalis asam sulfat 0,1% dan variasi suhu  $220^{\circ}\text{C}$ ,  $240^{\circ}\text{C}$ , dan  $250^{\circ}\text{C}$  untuk reaksi esterifikasi. Selanjutnya dilakukan reaksi polimerisasi dengan menggunakan katalis benzoil peroksida sebesar 0,1 % berat sampai suhu mencapai  $260^{\circ}\text{C}$  dalam waktu dua jam. Karakteristik hasil reaksi diketahui dengan menganalisis densitas, viskositas berat molekul, dan waktu esterifikasi dari larutan pada setiap rentang waktu reaksi untuk menentukan kualitas dari Gliserol Monooleat yang dihasilkan. Dari kelima atribut dengan data percobaan yang dilakukan pada proses produksi akan dikelompokkan berdasarkan standar kualitas yang baik (*pass*) dan yang tidak baik (*not pass*).

## 6. HASIL DAN PEMBAHASAN

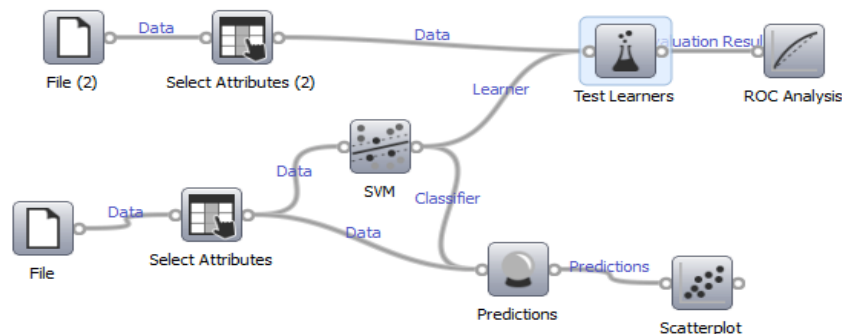
### 6.1 Data Mining Dengan Python Orange

Salah satu *software* yang dapat digunakan untuk *data mining* adalah *Software Orange* dengan Skema seperti pada Gambar 3. Tahapan proses pengolahan data adalah sebagai berikut:

1. Penentuan *file* data yang akan digunakan untuk pengolahan seperti pada Gambar 2
2. Data kemudian diolah dengan pemilihan atribut data
3. Pengolahan dengan Metode SVM
4. Dibandingkan hasil antara Prediksi dengan tes
5. Hasil tes dianalisis dengan Analisis ROC
6. Hasil prediksi di-plot dengan *scatter*



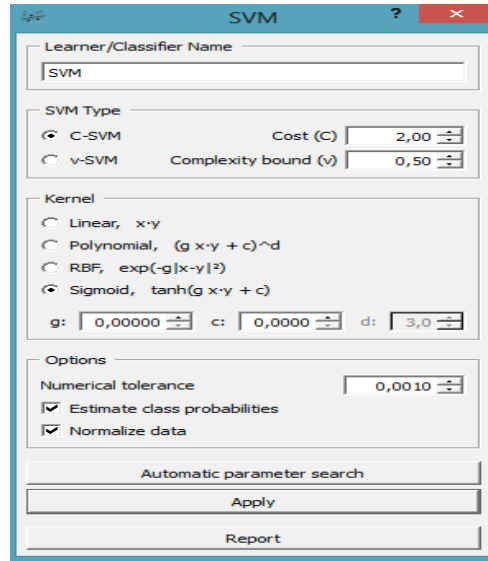
Gambar 2. Data file input



Gambar 3. Skema *Software Orange* Metode SVM dan Analisis ROC

Tabel 1. Perbandingan Hasil Parameter Untuk C-SVM Kernel Sigmoid with Cross Validation folds=5

C	Classification Accuracy	Sensitivity	Specivicity	Area Under Curve	Quality
4	0,9800	0,9647	0,9879	0,9982	NOT PASS
	0,9800	0,9879	0,9647	0,9982	PASS
3	0,9840	0,9765	0,9879	0,9979	NOT PASS
	0,9840	0,9879	0,9765	0,9979	PASS
2	0,9800	0,9647	0,9879	0,9968	NOT PASS
	0,9800	0,9879	0,9647	0,9968	PASS
1	0,9760	0,9765	0,9758	0,9961	NOT PASS
	0,9760	0,9758	0,9765	0,9961	PASS
0.8	0,9640	0,9765	0,9576	0,9950	NOT PASS
	0,9640	0,9576	0,9765	0,9950	PASS
0.5	0,9560	0,9882	0,9394	0,9964	NOT PASS
	0,9560	0,9294	0,9882	0,9964	PASS
	0,9760	0,9765	0,9758	0,9961	NOT PASS



Gambar 4. Paramater metode SVM

Parameter yang digunakan pada metode SVM adalah C-SVM dengan C=2,00, *Kernel method Sigmoid* dengan nilai g:0 dan c:0

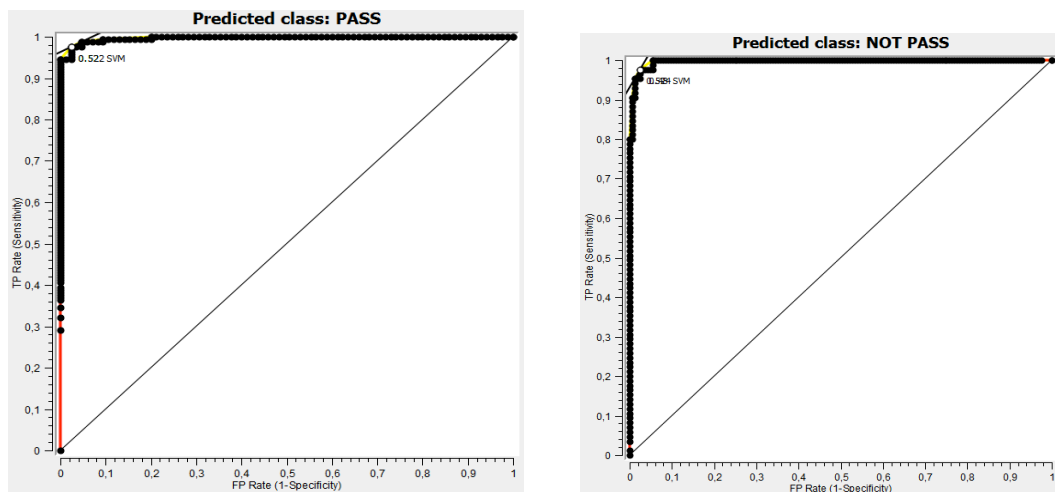
Tabel 1. Prediksi SVM

Predictions					
BeratMolekul	Viskositas	Densitas	Waktu	kualitas	SVM
19639	0.000529	1.03	64	NOT PASS	NOT PASS
19964	0.000517	1.02	67	PASS	PASS
19504	0.000534	0.97	70	NOT PASS	NOT PASS
19984	0.000517	1.03	75	PASS	PASS
19950	0.000526	1.04	86	PASS	PASS
19735	0.000515	0.99	90	PASS	PASS
19917	0.000515	0.95	67	PASS	PASS
19720	0.000521	0.97	73	PASS	PASS
19884	0.000532	1.03	70	PASS	PASS
19782	0.000516	0.95	63	PASS	PASS
19668	0.000514	1.02	77	NOT PASS	NOT PASS
19740	0.000529	0.95	84	PASS	PASS
20008	0.000518	1.00	80	PASS	PASS
19557	0.000524	1.03	66	NOT PASS	NOT PASS
19746	0.000518	1.00	60	PASS	PASS
19854	0.000534	0.96	89	PASS	PASS
19611	0.000527	1.03	72	NOT PASS	NOT PASS
19634	0.000515	1.00	71	NOT PASS	NOT PASS
19935	0.000518	1.01	61	PASS	PASS
19939	0.000531	1.00	61	PASS	PASS
19983	0.000524	0.97	69	PASS	PASS
19593	0.000517	0.98	79	NOT PASS	NOT PASS
19538	0.000514	0.99	62	NOT PASS	NOT PASS
19557	0.000515	1.00	69	NOT PASS	NOT PASS
19732	0.000534	0.95	68	PASS	PASS
19888	0.000532	1.03	79	PASS	PASS
19556	0.000517	0.95	69	NOT PASS	NOT PASS
19600	0.000527	1.03	78	NOT PASS	NOT PASS
19767	0.000529	0.95	76	PASS	PASS

Pengujian prediksi dengan metode SVM dengan menunjukkan hasil kualitas Baik (*pass*) dan Tidak Baik (*not pass*) ditunjukkan pada Tabel 1 yang sangat sesuai dengan *data testing*.

## 6.2 . Analisis ROC

Untuk pengujian pada penelitian ini menggunakan metode ROC (*Receiver Operating Characteristics*), yaitu sebuah metode untuk menguji kemampuan sistem prediksi kontingensi. ROC menampilkan *skill* sistem prediksi dengan membandingkan *hit rate* dan *false alarm rate*. Definisi *false alarm rate* harus dibedakan dengan *false alarm ratio*, tetapi dalam beberapa literatur keduanya mengacu pada *false alarm rate*. Derivasi dari ROC didasarkan pada tabel kontingensi memberikan *hit rate* dan *false alarm rate* untuk deterministik atau prediksi probabilistik. Suatu kejadian didefinisikan sebagai biner, yang berarti bahwa hanya ada dua hasil yang mungkin, suatu kejadian atau non-kejadian. Dari hasil visualisasi Grafik ROC untuk aplikasi metode SVM untuk penentuan kualitas Gliserol Monooleat, nilai untuk SVM dengan parameter C terbaik yaitu=3 untuk Akurasi Klasifikasi sebesar 0,9840.



Gambar 5. Analisis ROC Kualitas Baik (*Pass*) dan Tidak Baik (*Not Pass*)

## 7. KESIMPULAN

Dari hasil analisis *data testing* dan prediksi dari penentuan kualitas Gliserol monooleat dengan menggunakan metode SVM serta pengujian dengan analisis ROC, dapat diambil kesimpulan bahwa Metode SVM dapat mengklasifikasikan dengan baik kualitas yang dihasilkan oleh proses produksi dengan nilai untuk SVM dengan parameter terbaik C=3 untuk Akurasi Klasifikasi sebesar 0,9840. Metode ini dapat digunakan untuk penentuan parameter kualitas dalam produksi yang sedang berjalan bila menggunakan pendeteksian parameter pengujian dengan sensor dan perangkat pengolahan data akuisisi.

## REFERENSI

- [1]. Vijayakumar S, Wu S, 1999 “Sequential Support Vector Classifiers and Regression”, Proc.International Conference on Soft Computing (SOCO'99),Genoa, Italy, pp.610-619.

- [2]. Byun H., Lee S.W., 2003 “*A Survey on Pattern recognition Applications of Support Vector Machines*”, *International Journal of Pattern recognition and Artificial Intelligence*, Vol.17, No.3 , pp.459-486
- [3]. Ben-Hur, A., and Weston, 2010.J., A User's guide to support vector machines, *Data mining techniques for the life sciences*.Meth.Mol. Biol.609(2):223–239.
- [4]. Hamel, Lutz H. 2009. *Knowledge Discovery with Support Vector Machines*. John Wiley & Sons,Inc. New Jersey.
- [5]. Burdock,G.A.,1997. *Encyclopedia of food and Color additives*, Boca Raton, FL;CRC.
- [6]. Anonim. 2001.Liquid Density,Glycerine.<http://www.ici.go.id>
- [7]. Pardi, 2005.Optimasi Proses Produksi Gliserol MonoOleat dari Gliserol Hasil Samping Pembuatan Biodiesel.*Thesis*. Teknik Kimia Sekolah Pasca Sarjana Universitas Sumatera Utara Medan.
- [8]. Chen,J. and E.Dickinson, 1999. Effect of Monoglycerides and Diglycerol Ester on the Viscoelasticity of Heat-set Whey Protein Emulsion Gels. *International Journal of Food Science and Technology* 34:493-501.
- [9]. Griffin,W.C., 1979,*Emulsions*.Standen(ed) Kirk Othmer Encyclopedia of Technology (3<sup>rd</sup>ed) 8:117-154.