



Prediksi Ketebalan Powder Coating Menggunakan Algoritma SVM Dan Naïve Bayes

Zaenur Rozikin, Ahmad Turmudi Zy, Antika Zahrotul Kamalia

Fakultas Teknik, Program Studi Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia

Email: zaenurrozikin@pelitabangsa.ac.id, turmudi@pelitabangsa.ac.id, antika.kamalia@pelitabangsa.ac.idEmail Penulis Korespondensi: antika.kamalia@pelitabangsa.ac.id

Abstrak— Data Mining merupakan suatu metode yang telah banyak digunakan untuk melakukan penemuan ilmu pengetahuan dari kumpulan Dataset yang selama ini hanya disimpan tanpa dikelola lebih lanjut. Dalam dunia industri penggunaan metode data mining telah membantu masalah – masalah yang sering ditemukan dalam bidang industri. Data mining membantu dalam membuat prediksi mengenai masalah kualitas ketebalan (Thickness) pada sebuah produk box panel. Data mining sangat berguna untuk menemukan pola dalam proses pengelolahan data manufaktur yang kompleks. Apalagi jika kita berbicara konsumen atau pemakai jasa dari panel produk kita yang menginginkan *panel* tersebut bagus kualitas *powder coating* nya. Hal ini membuat peneliti melakukan penelitian guna mencari nilai *accuracy* yang nantinya akan dijadikan acuan yang pasti mengenai ketebalan powder coating. Hasil dari pengujian ini algoritma svm lebih baik dari *naïve bayes* karena data secara umum dapat dikategorikan sebagai hasil baik yang memiliki hasil *accuracy* 97.60 %, *precision* 99.56% dan *recall* 96.03 %. Hasil ini sebagai gambaran bagi konsumen untuk menyakinkan bahwa panel yang akan dibeli adalah kualitas terbaik. Dengan ditunjukkan data yang sudah diolah maka konsumen yakin dengan pembeliannya yang benar-benar nyata keabsahannya.

Kata Kunci: Data mining; SVM; Naïve Bayes; Data produk; powder coating; Rapid Miner

Abstract— *Data Mining is a method that has been widely used to make scientific discoveries from a collection of datasets which so far have only been stored without further management. In the industrial world the use of data mining methods has helped with problems that are often found in the industrial field. Data mining helps in making predictions regarding thickness quality problems in a panel box product. Data mining is very useful for finding patterns in complex manufacturing data processing processes. Especially when we talk about consumers or service users of our product panels who want the panel to have good powder coating quality. This made the researchers conduct research to find the accuracy value which would later be used as a definite reference regarding the thickness of the powder coating. The results of this test the svm algorithm is better than naïve Bayes because the data in general can be categorized as a good result which has an accuracy of 97.60%, precision 99.56% and 96.03% recall. This result is an illustration for consumers to ensure that the panels to be purchased are of the best quality. By showing the data that has been processed, the consumer is sure that the purchase is really valid.*

Keywords : Data mining; SVM; Naïve Bayes; Product data; powder coating; Rapid Miner

1. PENDAHULUAN

Dalam bidang industri yang semakin hari semakin berkembang pesat berdampak pada munculnya persaingan antar perusahaan baik swasta dan negeri, baik perusahaan bergerak manufaktur maupun perusahaan jasa. Akibatnya perusahaan akan berlomba untuk menciptakan pelayanan yang terbaik kepada pelanggan. Dan salah satunya yaitu memperhatikan kualitas dari produk yang dihasilkan. Aspek yang menunjang agar konsumen puas dengan kualitas *box panel* ialah *box panel* tersebut dapat berfungsi sesuai kebutuhan permintaan, dan kualitas dari ketebalan powder coating. Dimana dalam ketebalan powder coating ini terdapat data yang bervariasi yang menunjukkan ketidakpastian dalam hasil yang didapat. Maka dari para maker *box panel* khususnya bagian line spray powder coating berusaha semaksimal mungkin untuk menghasilkan kualitas produk *box panel* yang baik. Banyak percobaan secara manual untuk memperbaikinya baik dari pengaturan keluaran spray powder coating dan tekanan arus listrik, bahkan cara untuk menggerakkan tangan yang secara manual ditujukan kesebuah *box panel* tersebut. Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini yaitu untuk mengetahui nilai *accuracy* dari data yang diolah oleh algoritma support Vector Machine dan Naïve Bayes dalam menyelesaikan yang dialami dengan permintaan dari konsumen.

Metode yang akan digunakan penelitian ini merupakan penelitian eksperimen yaitu peneliti yang dengan sengaja dan secara sistematis mengadakan perlakuan atau tindakan pengamatan suatu variabel. Menurut Arikunto (2006:3). Memilih algoritma SVM dan Naïve Bayes untuk melakukan perbandingan dari data yang telah didapat kemudian data tersebut akan diolah dengan menggunakan algoritma svm dan naïve bayes mencari nilai *accuracy* atau memprediksi yang terbaik dari algoritma tersebut. Data yang bervariasi menjadi kesatuan data diolah dan di klarifikasi sehingga menghasilkan *output* data yang baik dari perbandingan algoritma svm dan naïve bayes.

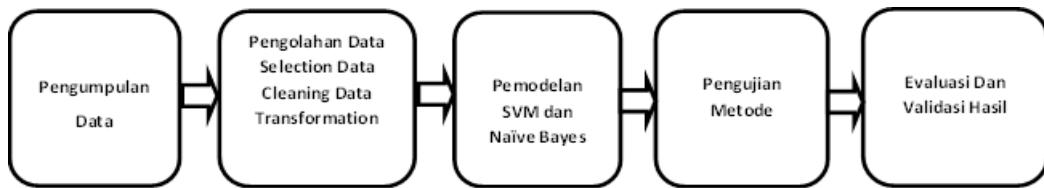
Klasifikasi ialah proses pembelajaran suatu fungsi yang digunakan untuk tiap atribut yang merupakan salah satu objek ke satu dari label kelas tertentu. Klarifikasi ini digunakan untuk mendeskripsikan data dari data nominal. Selanjutnya data mining mengintegrasikan dari sejumlah data yang memiliki nilai variable yang berbeda dengan serangkaian langkah memiliki tahapan pemilihan data, integrasi data, transformasi data, aplikasi teknik data mining, evaluasi pola yang ditemukan dan merupakan tahap dari proses interpretasi pola menjadi

pengetahuan yang dapat digunakan untuk mendukung pengambil keputusan dari *hipotesa* yang diprediksi, ditahapan terakhir mempresentasikan pengetahuan dengan teknik *visualisasikan* hasil data. Dalam paper ini peneliti menggunakan algoritma Support Vector Machine dalam pemodelan pembelajaran menganalisis suatu data yang digunakan untuk mengklarifikasi dan menganalisis data. Dan peneliti akan menambahkan algoritma *naïve bayes* untuk pengenalan pola yang fundamental yang memiliki hubungan dengan atribut yang diperlukan sebagai *variable*[2] .

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Dalam melakukan penelitian seorang peneliti memiliki pedoman yang secara bertahap yang akan dilakukan sebagai berikut :



Gambar 1. Tahapan Pelaksanaan Penelitian

1. Pengumpulan Data
Selanjutnya sebelum melakukan proses data mining menggunakan SVM, Naïve Bayes maka perlu dilakukan pengolahan data untuk mendapatkan nilai *accuracy* yang diinginkan dalam akhir penelitian ini. Data yang didapat dalam bentuk file spreadsheet berformat excel sehingga data masih tak beraturan pada beberapa kolom dan sheets. Sumber data yang peneliti di temukan secara langsung atas permintaan konsumen yang awalnya melihat nilai thickness dari box panel yang dipesan memiliki beragam variasi nilainya.
2. Data Cleaning
Dalam tahap data cleaning ialah proses pembersihan dari data yang akan diteliti untuk penghapusan data dengan cara membuang data yang tidak diperlukan saat penelitian atau data yang tidak ada pengaruhnya terhadap penelitian. Proses pada tahap ini dilakukan secara manual dengan bantuan *software spreadsheet*.
3. Data Selection
Didefinisikan sebagai proses pemilihan atau penentuan data yang akan diteliti dari sumber data yang akan diteliti oleh peneliti. Tujuan utama dari pemilihan data adalah menentukan jenis data, sumber dan serta variabel yang tepat yang memungkinkan peneliti untuk menjawab semua pertanyaan penelitian yang memadai. Penentuan ini merupakan disiplin khusus terutama didorong oleh sifat penyelidikan, literatur yang ada dan *aksesibilitas* ke sumber data yang diperlukan. Pada tahap ini dilakukan langkah-langkah sebagai berikut :
 - a. Diambil sampel data secara random dengan parameter atribut pada data *quality control powder coating painting area Document* yang memiliki jumlah data terbanyak untuk dijadikan dataset dan memastikan data yang dipilih layak dipakai untuk dilakukan proses pemodelan.
 - b. Dataset ini setelah dikelompokkan maka didapatkan jumlah data testing.
 - c. Dilakukan seleksi atribut yang akan dipakai dan dianalisis, karena pada data awal terdapat beberapa atribut yang tidak dibutuhkan seperti atribut.
4. Data Transformation
Dalam penelitian transformasi data ialah proses pengubahan data dari satu format ke format lain. Ini adalah aspek fundamental dari integrasi data dan pengolahan data, dapat sederhana dan kompleks berdasarkan perubahan yang akan diperlukan dalam data penelitian, transformasi data biasanya dilakukan melalui langkah manual dan otomatis. Dan transformasi data bisa saja merupakan perombakan sumber data dimana seluruh database yang nilai-nilai nya dapat ditransformasikan dengan nilai-nilai berpengaruh yang lainnya. Berikut adalah hasil pengolahan data awal setelah melewati tahapan diatas untuk dijadikan dataset pada tahap selanjutnya.
5. Pemodelan
Pada tahap pemodelan ini akan dilakukan pemodelan menggunakan metode Algoritma Support Vector Machine dan Naive Bayes untuk pengolahan dataset. Berikut merupakan tahapan proses dari data tersebut.
6. Evaluasi dan validasi hasil
Proses pengujian dilakukan dengan tujuan mengetahui hasil akhir perhitungan yang dianalisa dan mengukur seakurat manakah algoritma yang digunakan dalam penelitian yang dilakukan serta melihat apakah berfungsi dengan baik atau tidak, sehingga muncullah perbandingan yang baik dari hasil perhitungan masing-masing algoritma. Proses pengujian menggunakan tool *rapidminer* dan mengamati data apakah sesuai dengan hasil yang diperoleh melalui tool tersebut.

2.2 Pengujian Metode

Proses pengujian dilakukan dengan tujuan mengetahui hasil akhir perhitungan yang dianalisa dan mengukur seakurat manakah algoritma yang digunakan dalam penelitian yang dilakukan serta melihat apakah berfungsi dengan baik atau tidak, sehingga muncullah perbandingan yang baik dari hasil perhitungan masing-masing algoritma. Proses pengujian menggunakan tool *rapidminer* dan mengamati data apakah sesuai dengan hasil yang diperoleh melalui tool tersebut. Sedangkan validasi Algoritma *Support Vector Machine* dan *Naïve bayes* dilakukan dengan mengukur nilai *accuracy*, *Precision* dan *recall* dan dapat dihitung dengan menggunakan *confusion matrix* sebagai berikut:

1. Algoritma Support Vector Machine

Nilai *Accuracy* dihitung dengan cara menjumlah data benar yang bernilai positif (*True Positive*) ditambah dengan nilai Negatif (*True Negatif*) dibagi dengan jumlah data benar yang bernilai positif (*True Positive*), Negatif (*True Negatif*) dan ditambah dengan data salah yang bernilai positif (*False Positif*), Negatif (*False Negatif*).

Tabel 1. Perhitungan Nilai *Accuracy* SVM *Confusion Matrix*

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP	FP
	FALSE	FN	TN
$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\%$			

Nilai *precision* dihitung dengan cara membagi jumlah data benar yang bernilai *positif* (*True Positive*) dibagi dengan jumlah data benar yang bernilai *positif* (*True Positive*) dan data salah yang bernilai positif (*False Positive*).

Tabel 2. Perhitungan Nilai *Precision* SVM *Confusion Matrix*

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP	FP
	FALSE	FN	TN

Nilai *recall* dihitung dengan cara membagi data benar yang bernilai *positive* (*True Positive*) dengan hasil penjumlahan dari data benar yang bernilai *positive* (*True Positive*) dan data salah yang bernilai negatif (*False Negative*).

Tabel 3. Nilai *Recall* SVM *Confusion Matrix*

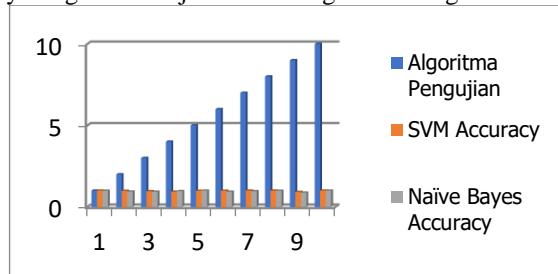
		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP	FP
	FALSE	FN	TN
$Recall = \frac{TP}{TP + FN} * 100\%$			

3. HASIL DAN PEMBAHASAN

Dari hasil pengujian dengan melakukan 10 pengujian secara random menghasilkan tingkat recall tertinggi yaitu algoritma. Dari data yang telah diuji, kemudian hasil dari data tersebut menyatakan tingkat *accuracy*, *Recall* dan *Precision*. Berikut grafik dari keseluruhan data yang telah di uji dengan rapid mener :

1. Accuracy

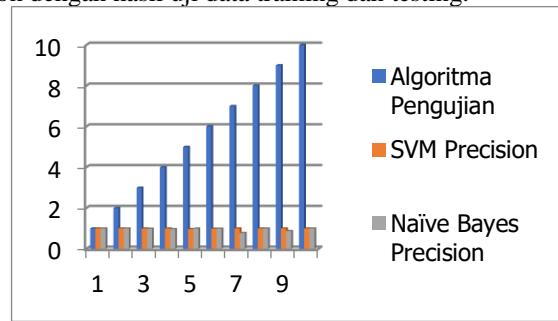
Grafik hasil dari nilai accuracy dengan hasil uji data training dan testing.



Gambar Grafik 2. Accuracy

2. Precision

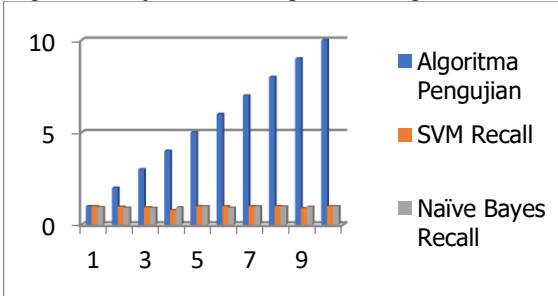
Grafik hasil dari nilai precision dengan hasil uji data training dan testing.



Gambar Grafik 3. Precision

3. Recall

Grafik hasil dari nilai recall dengan hasil uji data training dan testing.



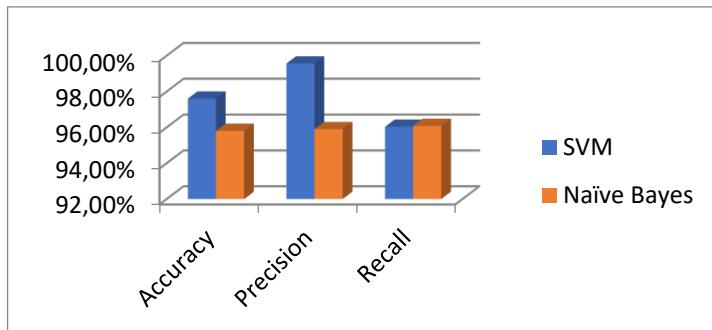
Gambar Grafik 4. Recall

Hasil pengujian dari data produksi powder coating menghasilkan tingkat accuracy, precision dan recall

Tabel 4. Nilai Rata – Rata Accuracy, Precision Dan Recall

Algoritma	Accuracy	Precision	Recall
SVM	97.60%	99.56%	96.03%
Naïve Bayes	95.80%	95.90%	96.08%

Dari hasil pengujian dengan melakukan 10 kali pengujian secara random dengan menghasilkan nilai rata – rata *accuracy* 97.60 %, *precision* 99.56% dan *recall* 96.03, *accuracy* tertinggi adalah *SVM*. Berdasarkan tabel 4 diatas, diketahui bahwa sebagian besar data yang saat ini dapat mempermudah dalam pengambilan keputusan dan upaya memberikan informasi dan pengetahuan berdasarkan hasil yang sudah didapatkan dalam penelitian ini maka algoritma yang menghasilkan tingkat *accuracy* yang baik *SVM* atau semua atribut dalam meningkatkan hasil *accuracy*, *recall* dan *precision*, berikut ini adalah grafik dari pengujian algoritma.

**Gambar Grafik 5.** Nilai Hasil Dari pengujian

4. KESIMPULAN

Berdasarkan analisa data testing pengujian dari data training yang dihasilkan terdapat hasil yang relevan dengan komposisi data yang akurat, pengujian algoritma naive bayes dan svm menghasilkan hasil optimal, berdasarkan hasil yang sudah didapatkan dalam penelitian ini maka algoritma yang menghasilkan tingkat *accuracy*, *recall* dan *precision* yang baik dan semua atribut, dalam meningkatkan hasil *accuracy*, *recall* dan *precision* hasil yang di dapat dari pengujian dengan melakukan 10 kali pengujian secara random dengan menghasilkan nilai rata – rata *accuracy* 97.60 %, *precision* 99.56% dan *recall* 96.03, *accuracy* teringgi adalah *SVM*. Dan hasil pengujian algoritma svm tidak memiliki nilai missing atau dalam pengujian menghasilkan hasil yang baik dalam memprediksi data, data ok menghasilkan prediksi ok, sedangkan algoritma naïve bayes dari pengujian dengan melakukan 10 kali pengujian secara random dengan menghasilkan nilai rata – rata *accuracy* 95.80%, *precision* 95.90 % dan *recall* 96.08 %, hasil dari pengujian algortitma memiliki nilai missing data atau dalam proses prediksi data ok akan prediksi ng. Hasil yang didapatkan hasil dari pengujian ini algoritma svm lebih baik dari naïve bayes karena data secara umum dapat dikatagorikan sebagai hasil baik yang memiliki hasil *accuracy* 97.60 %, *precision* 99.56% dan *recall* 96.03 %.

REFERENCES

- [1] M. Jeihouni, A. Toomanian, and A. Mansourian, "Decision Tree-Based Data Mining and Rule Induction for Identifying High Quality Groundwater Zones to Water Supply Management: a Novel Hybrid Use of Data Mining and GIS," *Water Resources Management*, vol. 34, no. 1, pp. 139–154, Jan. 2020, doi: 10.1007/s11269-019-02447-w.
- [2] M. S. Aliero, B. I. Shamaki, A. M. Bello, and ..., "CLASSIFICATION ALGORITHM: WEB FIREWALL," *International Journal Of ...*, 2020, [Online]. Available: <http://www.ijciras.com/Users/ManuScript/ManuScriptDetails/b2a8147a-ed8b-4dc1-9eb7-f33b69339f6a>
- [3] "PROSES POWDER COATING | POWDER COATING APPLICATOR." <http://powdercoating.web.id/proses-powder-coating/> (accessed Apr. 12, 2023).
- [4] I. I. Hazmi and D. Wicaksono, "PENGARUH METODE HOT DIPPING ALUMUNIZING COATING PADA MATERIAL STAINLESS STEEL 304 DALAM KOMPONEN FIREWALL PESAWAT CESSNA 150," *Teknika STTKD: Jurnal Teknik ...*, 2021, [Online]. Available: <https://www.jurnal.sttkd.ac.id/index.php/ts/article/view/313>
- [5] M. Jeihouni, A. Toomanian, and A. Mansourian, "Decision Tree-Based Data Mining and Rule Induction for Identifying High Quality Groundwater Zones to Water Supply Management: a Novel Hybrid Use of Data Mining and GIS," *Water Resources Management*, vol. 34, no. 1, pp. 139–154, Jan. 2020, doi: 10.1007/s11269-019-02447-w.
- [6] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, p. 219, Dec. 2021, doi: 10.30865/jurikom.v8i6.3655.
- [7] Retno Tri Vulandari, 2017. Data mining. Gava Media
- [8] Pandiyan, P, R. Sitharthan, S. Saravanan, Natarajan Prabaharan, M. Ramji Tiwari, T. Chinnadurai, T. Yuvaraj, and K.R. Devabalaji. 2022."A Comprehensive Review of the Prospects for rural electrification using stand-alone and hybrid energy technologies." Sustainable energy technologies and assessments.
- [9] AM Rahat, A. Kahir, dan AKM Masum, "Perbandingan Algoritma Naïve Bayes dan SVM Berdasarkan analisis sentimen menggunakan dataset review," di int ke-8 konf. Sistem. Model. Lanjut res. Trends, 2019 hlm. 266-270.
- [10] L.X.Hùng : Détection des émotions dans des énoncés audio multilingues. Institut polytechnique de Grenoble, 2009.



BULLETIN OF INFORMATION TECHNOLOGY (BIT)

Vol 4, No 2, Juni 2023, Hal 226 - 231

ISSN 2722-0524 (media online)

DOI [10.47065/bit.v3i1.687](https://doi.org/10.47065/bit.v3i1.687)

<https://journal.fkpt.org/index.php/BIT>

- [11] Y. Ben Ayed : Détection de mots clés dans un flux de parole. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications ENST, 2003.
- [12] <http://powdercoating.web.id/proses-powder-coating/>
- [13] Nofriansyah, Dicky & Gunadi Widi Nurcahyo. 2015. Algoritma Data Mining dan Pengujian. Yogyakarta:Deepublish
- [14] Zaki J. Muhammed, 2019. Data Mining And Analysis. Cambridge University Press
- [15] Dr. Suyanto. 2019. Data Mining untuk klarifikasi dan klasterifikasi Data. Bandung:Informatika