

KLASIFIKASI PRODUKSI BIBIT KAYU MENGGUNAKAN ALGORITMA NAIVE BAYES STUDI KASUS : PT. MUSI HUTAN PERSADA (MHP)

Syotian Barokah¹, Rusmin Syafari, M.Kom², Devi Udariansyah, M.Kom³

Fakultas Teknik Ilmu Komputer, Universitas Bina Darma

Email: mujairpratama@gmail.com¹, rusmin.syafari@binadarma.ac.id², devi.udariansyah@binadarma.ac.id³

ABSTRACT

Of information technology and cause the data collection mounting leading to the need application of technology data analysis that the data can be information the application of technology will analysis data can be help an enterprise in obtain information hidden from data supersized to avoid the waste of seeds, to know seeds needs, and see the percentage the target. To companies the absence of the technique or a particular method pengklasifikasian seeds to the increased availability of land in accordance, but in the process all the computerized pengklasifikasiannya only have still not optimal. So necessary the application of the classification methods and tools with naif algorithms bayes ' rapidminer studio. Accuracy the algorithms % 98,97 and analysis class b (0,997), c), class (0,001 and a class (0,001).

Keywords: nursery, naive bayes , rapidminer classifications seeds to land.

ABSTRAK

Perkembangan teknologi informasi yang semakin pesat menyebabkan data yang dihasilkan menggunung sehingga diperlukan penerapan teknologi analisis data agar data-data tersebut dapat menjadi informasi yang bermanfaat dengan adanya penerapan teknologi analisis data dapat membantu perusahaan dalam menggali informasi tersembunyi dari data yang berukuran besar untuk menghindari pemborosan bibit, untuk mengetahui kebutuhan bibit, dan melihat presentase pencapaian target. Pada perusahaan belum adanya penerapan teknik atau metode tertentu untuk pengklasifikasian bibit sesuai ketersediaan lahan, tetapi dalam proses pengklasifikasiannya sudah terkomputerisasi hanya saja hasilnya masih kurang optimal. Sehingga perlu dilakukan penerapan metode klasifikasi dengan algoritma *naive bayes* dan *tools rapidminer studio*. Akurasi penerapan algoritma 98,97% dan hasil analisis class B (0,997), class C (0,001), dan class A (0,001).

Kata kunci: *Nursery, Naive Bayes, Rapidminer* Klasifikasi Bibit ke Lahan.

1. PENDAHULUAN

Perkembangan teknologi informasi disegala bidang kehidupan yang semakin pesat. Menyebabkan data yang dihasilkan oleh teknologi informasi dari berbagai bidang menggunung. Sehingga dalam teknologi-teknologi tersebut perlu adanya penerapan teknologi analisis data agar data-data tersebut menjadi sebuah informasi yang berguna bagi perusahaan.

Pada perusahaan Musi Hutan Persada data penyapihan dan pengiriman bibit ke lahan dapat menghasilkan informasi yang berlimpah dimana informasi tersebut sangat berguna untuk evaluasi presentase pencapaian target perusahaan dan sebagai bahan acuan perusahaan dalam mengambil keputusan untuk mempersiapkan pengiriman bibit ke lahan berikutnya sehingga dapat meminimalisir pemborosan bibit serta menghemat biaya pengiriman ke lahan. Berdasarkan berlimpahnya data penyapihan dan pengiriman bibit ke lahan informasi yang tersembunyi dapat diketahui dengan cara melakukan pengolahan data tersebut menggunakan teknik atau metode tertentu sehingga berguna bagi perusahaan. Pengolahan data

tersebut perlu dilakukan untuk mengetahui informasi penting berupa pengetahuan baru misalnya informasi mengenai pengklasifikasian pencapaian target berdasarkan bibit siap tanam, penyapihan bibit, dan pengiriman bibit sesuai dengan ketersediaan lahan. Pengetahuan baru tersebut dapat membantu perusahaan untuk melakukan klasifikasi mengenai pencapaian target penyapihan dan pengiriman bibit ke lahan guna menentukan strategi untuk meningkatkan bibit siap tanam sesuai dengan lahanyang tersedia pada tahun-tahun berikutnya.

Pada PT. Musi Hutan Persada data pesuplaian kayu sudah disimpan dalam sistem basis data yang sudah terdistribusi sehingga mudah diakses oleh karyawan dan pimpinan sesuai dengan hak aksesnya masing-masing dari semua unit dan cabang untuk melihat lahan yang siap tanam dan data penyapihan bibit, akan tetapi belum adanya penerapan teknik atau metode tertentu untuk pengklasifikasian bibit siap tanam sesuai dengan lahan yang tersedia sehingga hasil pengklasifikasian yang dilakukan oleh perusahaan masih kurang efektif. Pada proses pengklasifikasian yang dilakukan perusahaan sudah terkomputerisasi hanya saja hasil pengklasifikasiannya masih kurang optimal dan maksimal. Karena besarnya jumlah data yang ada sekarang, sehingga diperlukan teknik khusus untuk menggali informasi bibit siap tanam dari data penyapihan dan pengiriman bibit ke lahan. Teknik yang paling tepat digunakan untuk pemanfaatan data yang besar tersebut yaitu menggunakan teknik *data mining* dengan metode klasifikasi dan algoritma *naive bayes* serta *tools rapidminer studio* menggunakan formata *excel* atau *CSV*.

Dengan adanya penerapan data mining diharapkan dapat mempermudah perusahaan dalam menggali informasi tersembunyi dari sekumpulan data yang besar seperti untuk melihat bibit siap tanam sesuai dengan lahan yang tersedia dengan cara menganalisa data penyapihan dan pengiriman bibit ke lahan menggunakan teknik *data mining* dengan metode klasifikasi dan algoritma *naive bayes* juga menggunakan *tools rapidminer studio*.

Berdasarkan masalah dan solusi tersebut sehingga penulis akan menuangkan penelitiannya dalam bentuk penulisan skripsi yang berjudul “klasifikasi produksi bibit kayu menggunakan algoritma *naive bayes* studi kasus : PT. Musi Hutan Persada (MHP)”.

2. METODOLOGI PENELITIAN

2.1 Metode Deskriptif

Metode penelitian yang digunakan adalah Deskriptif , Deskriptif adalah penelitian yang dilakukan untuk mengetahui variabel mandiri, baik hanya pada satu variabel atau lebih tanpa membuat perbandingan atau menghubungkan dengan variabel lainnya (variabel mandiri atau variabel yang berdiri sendiri, bukan variabel independen, karena kalau variabel independen selalu dipasangkan dengan variabel dependen (Sugyono, 2015).

Adapun langkah-langkah yang diikuti dalam melakukan penelitian deskriptif menurut (Sugyono, 2015) yaitu sebagai berikut :

- 1) Perumusan masalah. Metode penelitian manapun harus diawali dengan adanya masalah, yakni pengajuan pertanyaan-pertanyaan penelitian yang jawabannya harus dicari dengan menggunakan data lapangan. Pertanyaan masalah mengandung variabel-variabel yang menjadi kajian dalam studi ini. Dalam penelitian deskriptif peneliti dapat menentukan status variabel atau mempelajari hubungan antara variabel.
- 2) Menentukan jenis informasi yang diperlukan. Dalam hal ini peneliti perlu perlu menetapkan informasi apa yang diperlukan untuk menjawab pertanyaan atau masalah yang telah dirumuskan . apakah informasi kualitatif atau kuantitatif. Informasi kuantitatif berkenaan dengan data atau informasi dalam bentuk bilangan atau angka.
- 3) Menentukan prosedur pengumpulan data. Ada dua unsur penelitian yang diperlukan, yakni instrumen atau alat pengumpul data dan sumber data atau sampel yakni dari mana sebaiknya informasi diperoleh. Dalam penelitian ada sejumlah alat pengumpul data antara lain tes, wawancara, observasi, kuisioner, sosimetri. Alat-alat tersebut lazim digunakan dalam penelitian deskriptif.
- 4) Menentukan prosedur pengolahan informasi atau data. Data dan informasi yang telah diperoleh dengan instrumen yang dipilih dan sumber data atau sampel tertentu masih merupakan informasi atau data kasar. Informasi dan data tersebut perlu diolah agar dapat dijadikan bahan untuk menjawab pertanyaan penelitian.

- 5) Menarik kesimpulan penelitian. Berdasarkan hasil pengolahan diatas, peneliti menyimpulkan hasil penelitian deskriptif dengan cara menjawab pertanyaan-pertanyaan penelitian dan mensintesis semua jawaban tersebut dalam satu kesimpulan yang merangkum permasalahan penelitian secara keseluruhan.

2.2 Metode Pengumpulan Data

Adapun metode pengumpulan data yang digunakan dalam penelitian ini sebagai berikut :

- 1) Pengamatan (Observasi) merupakan metode pengumpulan data dengan cara mengunjungi langsung PT. Musi Hutan Persada dengan cara melakukan pengamatan langsung proses produksi bibit kayu eucalyptus pellitia.
- 2) Wawancara (Interview) merupakan metode pengumpulan data dengan cara melakukan proses tanya jawab seputar produksi pembibitan dan pengiriman bibit kayu eucalyptus pellitia di PT Musi Hutan Persada dengan head unit nursery, dengan tujuan untuk mendapatkan gambaran produksi dan pengiriman bibit serta masalah yang berhubungan dengan penelitian.
- 3) Studi pustaka merupakan proses pengumpulan data dengan cara mencari dan mempelajari data-data dari buku-buku ataupun dari referensi lain yang berhubungan dengan penulisan skripsi ini.

2.3 Metode Analisa Data

Metode analisis data yang digunakan adalah klasifikasi. Klasifikasi dapat didefinisikan sebagai proses untuk menyatukan suatu objek data sebagai salah satu kategori (kelas) yang telah didefinisikan sebelumnya Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018), (Zaki & Meira, 2013). Adapun tahapan yang digunakan untuk menganalisis data dalam penerapan data mining ini menggunakan knowledge discovery in databases (KDD) yang terdiri dari beberapa tahapan (Astuti Hermawati, 2013)

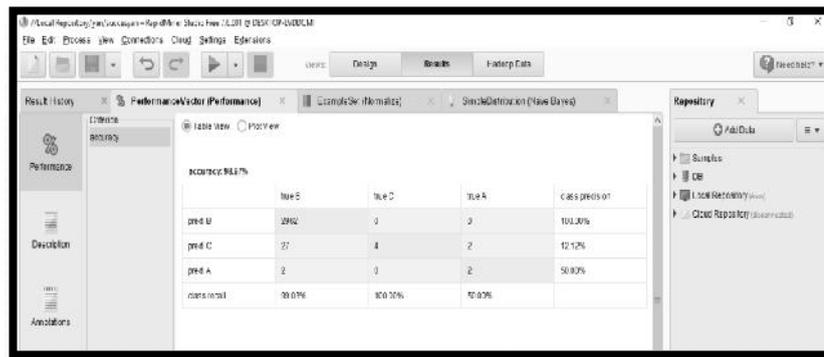
- 1) Data Cleaning (Pembersihan Data) adalah proses cleaning atau pembersihan data dari data yang mengandung missing value dan data yang mengandung redundan. Karena seleksi data relevan yang berarti atribut-atributnya ditentukan oleh output yang dihasilkan.
- 2) Data Integration (Integrasi Data) yaitu proses penggabungan data dari beberapa database produksi dan pengiriman bibit kayu di PT. Musi Hutan Persada menjadi satu kedalam satu database yang baru.
- 3) Data Selection (Seleksi Data) pada tahap ini penulis melakukan penyeleksian atau pemilihan atribut-atribut dalam database produksi dan pengiriman bibit kayu eucalyptus pellitia yang akan digunakan didalam tahap penerapan data mining.
- 4) Data Transformation (Transformasi Data) setelah data dibersihkan dari data-data yang mengandung missing value dan data yang mengandung redundan selanjutnya data akan ditransformasi menjadi data yang siap diminingkan. Karena software yang digunakan untuk data mining rapidminer studio maka dataset diatas tidak perlu lagi diekspor karena software tersebut bisa membaca format excel.
- 5) Proses Mining merupakan tahapan untuk menemukan pola atau informasi dari sekumpulan data menggunakan teknik dan algoritma tertentu. Pada penelitian penulisan menggunakan metode klasifikasi dan algoritma naive bayes untuk menemukan informasi mengenai data produksi dan pengiriman bibit kayu. Setelah dijelaskan proses penerapan data mining dengan teknik klasifikasi maka akan dijelaskan proses data mining secara aplikatif menggunakan rapidminer studio pada data pensuplaian kayu tersebut.
- 6) Pattern Evaluation (Evaluasi Pola) yaitu proses mengidentifikasi pola-pola yang menarik kedalam knowledge base yang ditemukan didalam data produksi dan pengiriman bibit kayu eucalyptus pellitia.
- 7) Knowledge Presentation (Presentasi pengetahuan) pada tahap ini penulis melakukan proses visualisasi dan penyajian data kedalam bentuk yang mudah dimengerti oleh user seperti dalam bentuk histogram, grafik, dan sebagainya.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

3.1.1 Hasil Akurasi Pengujian

Hasil akurasi model naive bayes menunjukkan bahwa nilai tingkat akurasi yang dihasilkan oleh algoritma naive bayes memiliki tingkat keakuratan yang sangat tinggi. Hal ini dibuktikan dengan hasil akurasi perhitungan yang mencapai 98,97%, hal ini membuktikan bahwa model yang dibangun bisa digunakan untuk proses klasifikasi potensi pencapaian target bibit sesuai kebutuhan lahan, nilai itu juga disebabkan oleh kompleksnya data yang mengakibatkan model bisa memprediksi dengan tepat. Untuk lebih jelasnya dapat dilihat pada gambar 3.1 hasil akurasi dibawah ini.



Block	Type B	Type C	Type A	Class Prediction	Accuracy
pre-B	1092	0	0	100.0%	100.0%
pre-C	27	4	2	12.12%	12.12%
pre-A	2	0	2	50.0%	50.0%
cases result	99.0%	100.0%	10.0%		

Gambar 3.1 Hasil *Performance Vektor*

3.1.2 Hasil Evaluasi Pola

Dari penelitian yang sudah penulis lakukan, sudah didapatkan sebuah informasi dan pengetahuan baru dari proses data mining menggunakan tools rapidminer studio dalam menentukan pencapaian target perusahaan dalam pengiriman bibit ke lahan, di blok mana saja yang berpotensi tercapai target, SOP, dan tidak tercapai target. Adapun yang didapat setelah dilakukan tahapan knowledge discovery in database (KDD) dari data bibit PT Musi Hutan Persada (MHP) tahun 2015-2017. Dari proses perhitungan data mining menggunakan algoritma naive bayes dengan menerapkan teknik classification dan tingkat keakurasiannya menghasilkan suatu informasi berdasarkan perhitungan data bibit PT MHP tahun 2015-2017 menunjukkan kelas SOP (B) dengan prior probability senilai 0.997, sedangkan kelas target tercapai (A) dan kelas tidak tercapai (C) senilai 0.001, berdasarkan perhitungan tools rapidminer studio blok yang berpotensi tidak tercapai target dengan data sebanyak 0.154 adalah Tebing Indah 2, Serai, Deras, dan Sodong Barat dengan akurasi sebesar 98,97% yang diakibatkan karena jumlah bibit yang dikirim lebih banyak dari jumlah bibit siap tanam (BST) sehingga tingginya jumlah bibit yang terbuang sia-sia.

3.1.3 Hasil Analisis

Dari proses perhitungan data bibit PT Musi Hutan Persada (MHP) tahun 2015-2017 menggunakan tools rapidminer studio penulis mendapatkan hasil pencapaian target di setiap blok PT. MHP rata-rata pencapaian target SOP. Sedangkan blok yang berpotensi tidak mencapai target terdapat pada sebagian blok yaitu tebing indah 2, serai, deras, dan sodong barat dengan jumlah data masing-masingnya 154 pengiriman bibit juga blok Keruh 2, Ibul, Subat, Merbau 3 dan Sodong Selatan dengan jumlah masing-masing pengiriman 77. Sehingga dengan hasil penelitian ini diharapkan bisa menjadi bahan referensi perusahaan dalam memberikan arahan terhadap blok-blok tersebut agar bisa mencapai target dalam pengiriman bibit ke lahan sehingga dapat mengurangi kerugian perusahaan dalam biaya operasional pembibitan. Berikut ini data pencapaian target yang

tercapai target dapat dilihat pada tabel 3.1 dibawah ini. Jumlah pada tabel dibawah ini didapat dari jumlah seluruh estimasi target tercapai (A) pada data set yang digunakan.

Tabel 3.1 Hasil Pengujian Estimasi (A)

Estimasi	Blok	Jumlah
Target Tercapai (A)	Sungai Tuha	0.231
	Deras	0.231
	Merbau 1	0.154
	Keruh 1	0.077
	Subat	0.077
	Banding Anyar	0.077
	Caban Utara	0.077
	Tebing Indah 2	0.077

Dari tabel 3.1 hasil pengujian diatas dengan jumlah data sebanyak 1001 pencapaian target blok yang kelas target tercapai adalah blok deras dan sungai tuha dengan masing-masing probabilitas 0.662 sedangkan data bibit pencapaian target dengan estimasi pencapaian target SOP dapat dilihat pada tabel 3.2 hasil SOP. Jumlah pada tabel dibawah ini didapat dari jumlah seluruh estimasi SOP (B) pada data set yang digunakan, dari tabel 3.2 hasil pengujian dibawah ini dengan jumlah data sebanyak 1048 pencapai target blok yang kelas SOP adalah blok Tebing Indah 2 dengan jumlah sebanyak 0.162.

Tabel 3.2 Hasil Estimasi (B)

estimasi	Blok	Jumlah
SOP	Keruh 2	0.013
	Lanting Barat	0.014
	Lagan	0.045
	Serdang Selatan	0.005
	Tebing Indah 2	0.162
	Serai	0.018
	Ipuh	0.057
	Deras	0.019
	Teras	0.009
	Lanting Timur	0.005
	Cawang	0.014
	Selibing	0.046
	Baung Selatan	0.003
	Tebing Indah 1	0.010
	Ibul	0.016
	Serdang Utara	0.005
	Keruh 1	0.023
	Koneng	0.083
	Merbau 1	0.020
	Subat	0.057
	Niru	0.083
	Jernih	0.006
	Banding Anyar	0.024
	Merbau 2	0.013
Subanjeriji	0.071	
Merbau 3	0.009	
Lubuk Guci	0.008	
Toman 2	0.010	

Keterangan Gambar :

Angka 0,000 – 0,275 adalah range probabilitas data untuk menentukan jumlah probabilitas pencapaian target setiap blok berdasarkan range pencapaian target.

-  Range Pencapaian Target
-  Pencapaian target melebihi target (A)
-  Pencapaian target SOP (B)
-  Pencapaian target tidak tercapai target (C)

3.2 Pembahasan

3.2.1 Proses Pengolahan Informasi atau Data

Berikut ini langkah-langkah implementasi algoritma naive bayes menggunakan rapidminer studio.

- 1) Menambahkan operator set role, Set role merupakan proses untuk menambahkan sebuah atribut khusus, pada penelitian ini field data yang di set yaitu pencapaian target sebagai label.
- 2) Menambahkan operator split data Split data berfungsi untuk membagi class estimasi sesuai dengan hasil perhitungan dari rapidminer tentang jumlah dari class.
- 3) Menambahkan algoritma naive bayes Tahapan ini merupakan bagian terpenting dengan menggunakan algoritma naive bayes sebagai pemodelan yang akan dilakukan dalam penelitian ini.
- 4) Menambahkan operator apply model Penambahan operator apply model digunakan untuk menjalankan model naive bayes yang dilakukan pada tahapan sebelumnya untuk mendapatkan knowledge yang diinginkan dan dimengerti oleh user.
- 5) Menambahkan operator performance classification Operator ini digunakan untuk evaluasi kinerja statistik tugas klasifikasi. Operator ini memberikan daftar nilai kriteria kinerja tugas klasifikasi.

4. KESIMPULAN

Dari uraian-uraian pada bab-bab yang sudah dibahas sebelumnya sehingga penulis dapat menarik kesimpulan dari penelitian ini yaitu sebagai berikut :

- 1) Dari perhitungan data mining menggunakan algoritma naive bayes serta teknik classification dan tools rapidminer studio bisa diambil keputusan bahwa kelas estimasi pencapaian target pengiriman bibit ke lahan. Kelas estimasi SOP (B) dominan lebih besar dari kelas estimasi target tercapai (A) dan tidak tercapai (C).
- 2) Dari hasil observasi terhadap data bibit PT Musi Hutan Persada (MHP) Niru menggunakan teknik classification dan algoritma naive bayes dengan bantuan tools rapidminer studio didapatkan blok yang berpotensi tidak tercapai target dalam pengiriman bibit kelahan pada sebagian pengiriman blok tersebut yaitu blok Keruh 2, Tebing Indah 2, Serai, Deras, Ibul, Subat, Merbau 3, Sodong Barat, dan Sodong Selatan yang diakibatkan oleh bibit yang dikirim lebih dominan dari pada bibit siap tanam sehingga bibit tersebut akan terbuang sia-sia dan mengakibatkan kerugian bagi perusahaan karena mereka membuat biaya produksi bibit yang lebih besar dari sebelumnya.
- 3) Tools rapidminer studio sukses menghasilkan potensi pencapaian target pengiriman bibit ke lahan per blok yang berpotensi tidak tercapai, SOP, dan target tercapai pada data pengiriman bibit PT Musi Hutan Persada tahun 2015-2017 dengan akurasi sebesar 98,97% dari sejumlah data pengiriman bibit sebesar 25.970.
- 4) Dari hasil analisa data bibit PT Musi Hutan Persada tahun 2015-2017 menggunakan tools rapidminer studio serta menggunakan algoritma naive bayes dan teknik klasifikasi dapat disimpulkan bahwa kinerja bagian nursery perusahaan tersebut cukup baik dikarenakan estimasi pencapaian target pengiriman bibitnya 0,997 sesuai standar operasional (SOP).

DAFTAR PUSTAKA

- Aprilia C, D., Aji Baskoro, D., Ambarwati, L., & Simri Wicaksana, I. W. (2013). Belajar Data Mining Dengan Rapidminer. 8-9.
- Astuti Hermawati, F. (2013). *Data Mining*. Yogyakarta: CV Andi Offset.
- Bustami. (2013). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah . TECHSI Jurnal Penelitian Teknik Informatika , 127-146
- Dr. Suyanto, S. M. (2017). Data Mining untuk Klasifikasi dan Klasterisasi Data. Bandung: Informatika Bandung.
- Kusrini, & Luthfi, E. T. (2010). Algoritma Data Mining. Yogyakarta: Andi Yogyakarta.
- Sugyono. (2015). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif dan R & D)*. Bandung: CV Alfabeta.
- Meilina, P. (2015). Penerapan Data Mining Dengan Metode Klasifikasi Menggunakan Decision Tree Dan Regresi. Jurnal Teknoogi .
- Mujiasih, S. (2011). pemanfaatan Data Mining Untuk Prakiraan Cuaca. Jurnal Meteorologi dan Geofisika , 189-195
- Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018). Improving Naïve Bayes in Sentiment Analysis For Hotel Industry in Indonesia. In 2018 Third International Conference on Informatics and Computing (ICIC) (pp. 1-6). IEEE.
- Velishya, A. (2014). Klasifikasi Data Mining untuk Penerimaan Pegawai menggunakan Algoritma C4.5. *Academia* .
- Zaki, M., & Meira, M. (2013). Data Mining and Analysis : Fundamental Concepts and Algorithms.