

PENERAPAN ALGORITMA *RANDOM FOREST* UNTUK KLASIFIKASI KAB KOTA PROVINSI JAWA BARAT BERDASARKAN PENDAPATAN SEKTOR PERTANIAN

¹Yana Cahyana
²Amril Mutoi Siregar

Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Buana Perjuangan
Karawang

[¹yana.cahyana@ubpkarawang.ac.id](mailto:yana.cahyana@ubpkarawang.ac.id)

[²amrilmutoi@ubpkarawang.ac.id](mailto:amrilmutoi@ubpkarawang.ac.id)

ABSTRAK

Kesejahteraan masyarakat pedesaan yang ada Provinsi Jawa Barat, yang jauh dari perkotaan, tidak lepas dari pendapatan utamanya adalah di sektor pertanian. Sektor pertanian meliputi beberapa sektor, kondisinya masih banyak ditemukan dibawah garis kemiskinan. Karena fasilitas dan pembiayaan masih minim dari pemerintah setempat. Mengingat akar permasalahannya adalah hampir semua desa yang ada di kab kota tidak mempunyai data yang benar, akurat dan tepat tentang kondisi permasalahan dan potensi desa yang dimiliki. Dengan penelitian ini dapat diharapkan salah satu cara untuk menganalisa data pendapatan dari sektor pertanian, baik potensi yang dimiliki, agar dapat lebih baik dalam pengambilan keputusan. Dan teknik pengolahan data ini bisa diimplemetasikan bagi pemerintah setempat untuk mengukur tingkat keberhasilan pertaniannya. Untuk seleksi fitur dalam penelitian ini adalah menggunakan algoritma *Random Forest* untuk mengklasifikasikan data secara otomatis dengan hasil akhir keberhasilan algoritmanya di peroleh tingkat akurasi 87,50 %.

Kata kunci: Datamining, Klasifikasi, *Random forest*, Pertanian

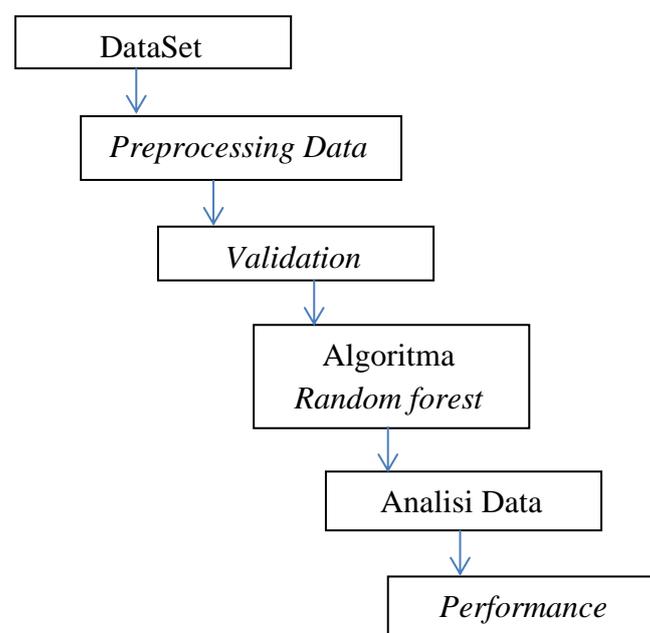
PENDAHULUAN

Potensi pertanian di Provinsi Jawa Barat tersebar secara merata di seluruh daerah, yang meliputi komoditas padi, palawija, dan hortikultural. Selain itu, jenis sayuran dan buah-buahan di daerah Provinsi Jawa Barat memiliki potensi yang sangat menjanjikan. hasil pertanian utama di daerah Provinsi Jawa Barat 1997 adalah sebagai berikut: hasil produksi padi mencapai 10.352.650 ton dengan luas panen sawah dan ladang 2.040.680 ha; hasil produksi jagung 336.014 ton dengan luas panen 127.994 ha, dan hasil produksi ubi kayu 1.816.487 ton dengan luas panen 141.637 ha. (<https://www.pikiran-rakyat.com/bandung-raya/2019/03/14/lahan-pertanian-di-jabar-terus-menyusut>). Lahan pertanian di Provinsi Jawa Barat terus mengalami penyusutan seiring dengan kebutuhan sarana dan prasarana bagi penduduk di Provinsi Jawa Barat yang begitu banyak. Apalagi, di Provinsi Jawa Barat terdapat banyak proyek infrastruktur yang praktis turut menggerus lahan pertanian. Kemandirian pangan merupakan upaya pemerintah untuk meningkatkan kemajuan di sektor pertanian. Akan tetapi penyusutan lahan

pertanian akibat modernisasi dan alih fungsi lahan menjadi properti ataupun industri, masih terjadi. Tidak terkecuali di Provinsi Jawa Barat, yang digadang – gadang sebagai wilayah lumbung padi terbesar nasional. Lahan pertanian seluas 139 hektare di Provinsi Jawa Barat (Jabar) gagal panen selama musim kemarau tahun ini. Pemerintah Provinsi Jawa Barat memastikan kondisi tersebut tidak berdampak terhadap pasokan beras. Ketika data sudah lengkap disektor pertanian, perlu kebijakan baik dari pemerintah Provinsi dan kab kota untuk bersama sama untuk mengatasi permasalahan, baik dari segi lahan yang menyusut dan keuangan yang dibutuhkan oleh petani. Dalam penelitian dapat, membantu mengolah data untuk klasifikasi kab kota berdasarkan pendapatan dari sektor pertanian dengan algoritma *random forest*. Agar data di olah untuk mendapatkan informasi yang berguna untuk mengambil keputusan.

METODE PENELITIAN

Langkah yang dilakukan pada penelitian yang akan dilakukan adalah bagaimana klasifikasi daerah menggunakan algoritma *random forest*. Objek penelitian ini adalah data penghasilan sektor pertanian di Provinsi Jawa Barat, pada tahun 2010-2019. Tersedia di laman dinas pertanian: <https://jabar.bps.go.id>. Berikut langkah langkah penelitian yang digunakan dibawah ini.



Gambar 1 Diagram alir yang digunakan

Diagram alir penelitian adalah sebagai berikut:

Data set yang digunakan adalah database dari pemerintah bagian statistik yang memiliki banyak atribut dibidang pertanian.

Preprocessing Data adalah proses setelah didapatkan dataset mentah dari situs pemerintah. Akan diproses sesuai dengan kebutuhan pada penelitian ini.

Validasi dari penelitian ini adalah memastikan data yang digunakan sudah valid dan kesesuaian atribut yang digunakan untuk mengambil informasi yang dibutuhkan seperti transformasi data sehingga akan didapatkan data yang sudah dikurangi tingkat kompleksitasnya yang kemudian akan memudahkan proses pengolahan karena ukuran data yang diproses sudah menjadi lebih kecil.

Algoritma Random forest secara otomatis untuk *random forest* bantuan *software RapidMiner*.

Analisa data adalah melihat hasil yang didapat apa saja yang dapat diambil sebuah informasi yang berharga. Dan permasalahan yang ditemui pada saat penelitian ini

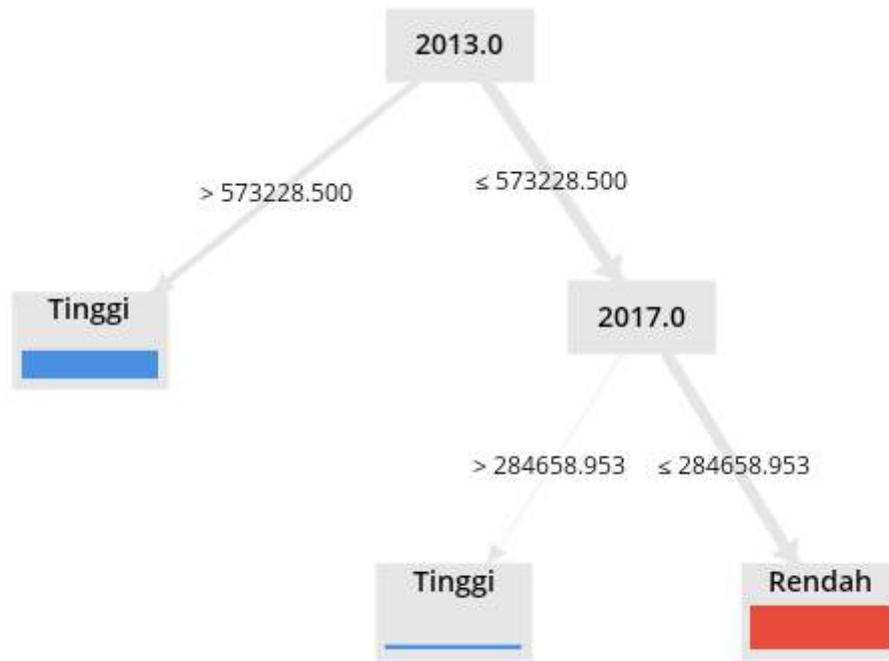
Hasil / performance adalah hasil dari pengolahan data yang sudah dilakukan akan ditentukan rule yang terbentuk untuk memberikan suatu keputusan dari algoritma.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil penelitian yang didapatkan merupakan hasil klasifikasi menggunakan *RapidMiner* sebagai tools perangkat lunak untuk pengolahan data. Dengan menggunakan prinsip dalam algoritma *data mining*, *RapidMiner* dapat mengekstrak pola-pola dari data set yang besar dengan mengkombinasikan kecerdasan buatan, metode statistika, dan *database*.

Hasil ditampilkan secara visual berupa grafik, penelitian berupa klasifikasi dengan dataset semua kab/kota yang ada di Provinsi Jawa Barat dengan menggunakan 27 atribut. Menampilkan hasil *performance* dari algoritma *Random Forest*. Hasil penelitian seperti dibawah ini:

Random Forest akan mengidentifikasi atribut mana yang paling besar atau dominan dari hasil pertanian di propinsi Jawa Barat, lalu hasilnya di ilustrasikan dalam bentuk pohon keputusan agar lebih mudah memahaminya.



Gambar 2 Model *Random Forest*

Berikut *rule* yang didapatkan dari dataset yang digunakan:

1. Pendapatan pertanian di tahun 2013 $> 573,228.500$, dikategorikan tinggi
2. Pendapatan pertanian di tahun 2013 $\leq 573,228.500$, Pendapatan pertanian di tahun 2017 $> 284,658.953$ dikategorikan tinggi
3. Pendapatan pertanian di tahun 2013 $\leq 573,228.500$, Pendapatan pertanian di tahun 2017 $\leq 284,658.953$, dikategorikan rendah.

Langkah selanjutnya adalah menampilkan *confusion matrix* untuk mengevaluasi algoritma *Random Forest* untuk klasifikasi. Dari tabel *confusion matrix* tersebut untuk mendapatkan nilai akurasi yang merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data.

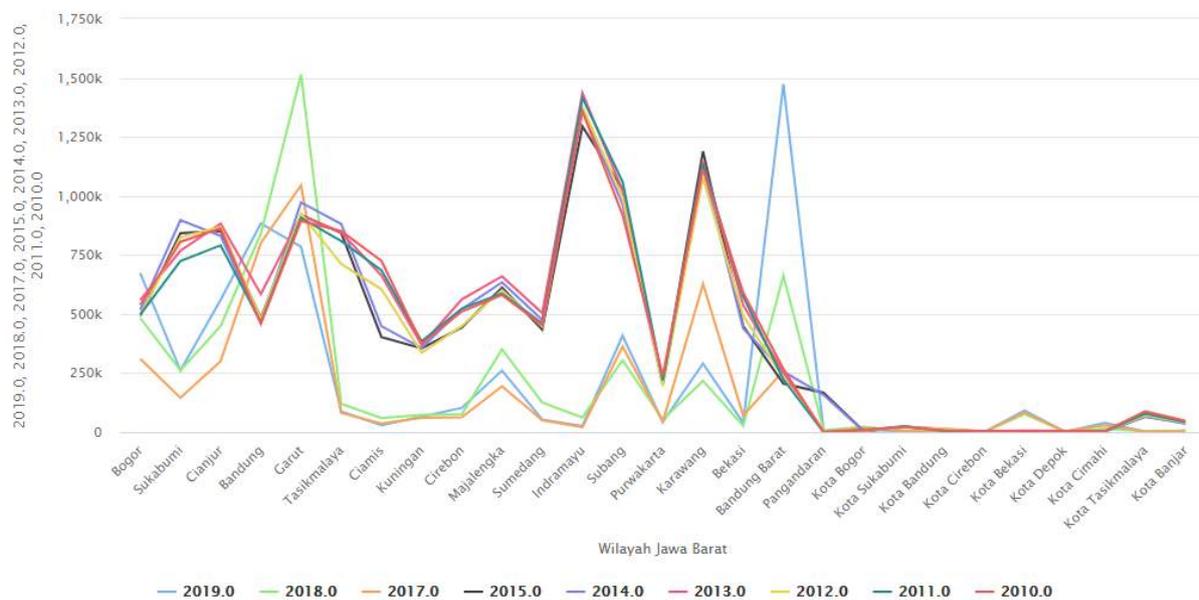
Tabel 1 *Performance Accuracy*

Akurasi 87,50%

	True Tinggi	True rendah	Class precision
Pred tinggi	3	1	75,00%
Pred rendah	0	4	100%
Class recall	100%	80,00%	

Berikut dibawah ini hasil *performance* penelitian ini, dengan menampilkan hasil setiap atribut untuk dapat dijadikan dasar ambil keputusan. Berikut atribut yang digunakan dalam ini adalah: Nama nama kabupaten/kota adalah Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar.

Dalam penelitian ini, beberapa point yang dapat diambil informasi penting dari 27 atribut yang digunakan adalah kategori pendapatan sektor pertanian yang tinggi dan rendah. Dari 27 atribut, pendapatan sektor pertanian 16 cluster rendah (Kuningan, Cirebon, Sumedang, Purwakarta, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar) dan 11 cluster tinggi (Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Majalengka, Indramayu, Subang, dan Karawang). Berikut hasil grafik pendapatan sektor pertanian di Provinsi Jawa Barat.



Gambar 3. Grafik Pendapatan Sektor Pertanian Provinsi Jawa Barat

KESIMPULAN DAN SARAN

Tahapan terakhir penelitian ini adalah tahap pengambilan kesimpulan dengan dataset sebanyak 27 atribut pada tahun 2010-2019. Hasil klasifikasi kabupaten kota berdasarkan hasil bidang pertanian dengan algoritma *random forest* dengan akurasi 87,50%. Algoritma *random forest* dapat mengklasifikasi dengan baik dengan menggunakan kelas tinggi dan rendah. Dengan hasil dan kesimpulan dari penelitian ini penulis memberikan beberapa saran adalah 1) Pengujian dengan data

sama perlu menggunakan algoritma yang lain. 2) Pengujian bisa digunakan *tools* yang lain. 3) Bisa jadi referensi bagi peneliti ke depannya dalam menganalisa data sejenis atau permasalahan yang mirip.

DAFTAR PUSTAKA

Safety and health in agriculture. International Labour Organization. 1999. ISBN 978-92-2-111517-5. Diakses tanggal 13 Februari 2019, Jam 19:30 Wib.

Agricultural Economics". University of Idaho. Diakses tanggal 2019-02-03. 88 ^ Runge, C. Ford (June 2006). "Agricultural Economics: A Brief Intellectual History" (PDF). Center for International Food and Agriculture Policy. hlm. 4. Diakses tanggal 2019-02-04.

Han, Jiawei, Micheline Kamber. 2001. *Data mining Concepts and. Techniques. Morgan Kaufmann Publisher.*

M. Sokolova dan G. Lapalme, 2009. "A systematic analysis of performance measures"

For classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, hal. 427–437, <http://perwakilan.jabarprov.go.id/artikel/16/pertanian>, diakses tanggal 13 Februari 2019

Ho, Tin Kam (1995). *Random Decision Forests . Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995.* pp. 278–282.

Breiman L (2001). *Random Forests. Machine Learning.* **45** (1): 5–32. [doi:1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).

<https://id.wikipedia.org/wiki/Pertanian>, diakses tanggal 10 Februari 2019, Jam 20:00 Wib.

<https://st2019.bps.go.id/dev2/index.php/site/tabel?tid=71&wid=3200000000>, diakses tanggal 06 Juni 2020

https://www.pikiran-rakyat.com/bandung-raja/2019/03/14/lahan-pertanian-di_jabar-terus-menyusut. Diakses tanggal 15 Maret 2019, Jam 19:00 Wib.