

IMPLEMENTASI METODE KLASIFIKASI NAIVE BAYES DALAM MEMPREDIKSI PRODUKTIVITAS HASIL PERTANIAN BAWANG MERAH DIBREBES MENGGUNAKAN RAPIDMINER

Satria Prayoga Imaniar¹, Eko Budiraharjo²

Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer

Universitas Pancasakti Tegal

satriaimaniarr@gmail.com

Abstrak

Penelitian ini bertujuan untuk memprediksi produktivitas panen bawang merah menggunakan metode Naive Bayes, mengingat sektor pertanian merupakan pilar penting ekonomi bangsa. Metode ini diterapkan untuk menghitung tingkat keberhasilan prediksi produktivitas hasil panen di tengah tantangan berupa perubahan dan perkembangan yang meningkatkan kompleksitas data serta informasi. Data penelitian diperoleh melalui website Badan Pusat Statistik Kabupaten Brebes. Hasil analisis menunjukkan bahwa variabel seperti jenis benih, pengobatan, pengairan, penyakit, dan pupuk berpengaruh signifikan terhadap produktivitas bawang merah. Metode Naive Bayes terbukti efektif dalam melakukan klasifikasi data dan menghasilkan prediksi yang akurat, sehingga dapat diimplementasikan untuk mendukung pengambilan keputusan di bidang pertanian.

Kata Kunci: Bawangmerah, Klasifikasi, Naivebayes, Produktivitas, Rapidminer

Abstract

This study aims to predict the productivity of onion harvest using Naive Bayes method, considering that the agricultural sector is an important pillar of the nation's economy. This method is applied to calculate the success rate of crop productivity prediction in the midst of challenges in the form of changes and developments that increase the complexity of data and information. Research Data obtained through the website of the Central Bureau of Statistics of Brebes regency. The results showed that variables such as seed type, treatment, irrigation, disease, and fertilizer significantly affect onion productivity. Naive Bayes method proved effective in classifying data and producing accurate predictions, so it can be implemented to support decision making in agriculture.

Keywords: Red Onion, Classification, Naivebayes, Productivity, Rapidminer

Pendahuluan

Pertanian merupakan salah satu sektor ekonomi yang sangat penting bagi negara, dan peningkatan produktivitas lahan pertanian dapat berkontribusi pada kesejahteraan petani. Bawang merah adalah salah satu komoditas hortikultura yang memiliki nilai strategis dan ekonomis. Tanaman bawang merah merupakan bahan penting dalam berbagai masakan Asia Tenggara dan dunia. Umbi bawang merah paling banyak dimanfaatkan, meskipun beberapa tradisi kuliner juga menggunakan daun dan tangkai

bunga sebagai bahan penyedap masakan sehari-hari. Produksi bawang merah dapat mengalami risiko yang bersifat fluktuatif, baik yang meningkatkan maupun mengurangi hasil, dan faktor lingkungan dapat mempengaruhi hasil produksi yang lebih sulit dikontrol.

Provinsi Jawa Tengah memberikan kontribusi terbesar, yaitu 33,5% dari total produksi bawang merah Indonesia yang berjumlah 1,82 juta ton. Di Kabupaten Brebes, penanaman bawang merah telah tersebar di seluruh kecamatan, dengan beberapa kecamatan mengalami peningkatan produksi yang signifikan, seperti di Kecamatan Tanjung, Sirampog, Bumiayu, Banjarharjo, Ketanggungan, Jatibarang, Paguyangan, dan Di Kecamatan Wanasari, penanaman bawang merah terdapat di Desa Sisalam, Desa Lengkong, dan Desa Glonggong, dengan Desa Krandon menjadi pusat utama penanaman bawang merah.

Tujuan dari penelitian ini adalah untuk mengimplementasikan metode klasifikasi Naive Bayes dan menguji kelayakan algoritma Naive Bayes dalam memprediksi produktivitas hasil pertanian bawang merah. Naive Bayes adalah metode yang digunakan untuk memprediksi probabilitas dengan asumsi bahwa antar kelas dalam data bersifat independen.

Tantangan utama dalam membangun prediksi yang akurat adalah kompleksitas data yang meningkat seiring dengan perubahan dan perkembangan yang terjadi. Sebagai contoh, dalam penerapan algoritma Naive Bayes pada klasifikasi tebu produktif di Dinas Kehutanan dan Perkebunan, sistem menunjukkan kinerja dengan akurasi 73,3% dari 28 data (Anandita, 2015). Keuntungan menggunakan Naive Bayes adalah kemampuan untuk menghasilkan klasifikasi dengan error yang lebih kecil meskipun dataset yang digunakan cukup besar (Berry, 2006). Selain itu, Han dan Kamber (2006) menyatakan bahwa metode Naive Bayes terbukti efektif dalam menghasilkan akurasi dan kecepatan tinggi saat diterapkan pada basis data besar.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengimplementasikan metode klasifikasi Naive Bayes dalam memprediksi produktivitas hasil pertanian bawang merah. Diharapkan penelitian ini dapat memberikan informasi yang berguna dan membantu petani, khususnya petani bawang merah di Kecamatan Wanasari, dalam memprediksi hasil panen mereka.

Landasan Teori

1. Bawang Merah dan Faktor Kegagalan Panen

Bawang merah merupakan salah satu komoditas sayuran unggulan yang memiliki arti penting bagi masyarakat, baik ditinjau dari tingginya nilai ekonomi maupun dari kandungan gizinya[1]. Hampir setiap masakan menggunakan tanaman hortikultura ini sebagai pelengkap bumbu penyedap, sehingga komoditas ini memberikan kontribusi yang cukup tinggi terhadap perkembangan ekonomi suatu wilayah. Namun, dalam proses usahatannya, petani kerap menghadapi berbagai kendala yang berisiko pada kegagalan panen. Terdapat lima faktor utama yang menyebabkan petani mengalami gagal panen, yaitu: (1) cuaca ekstrem yang cenderung hujan terus-menerus atau memicu kabut dan serangan ulat saat panas berlebihan, (2) kesalahan dalam pemilihan bibit tanaman, (3) takaran obat tanaman yang tidak sesuai, (4) kualitas tanah yang buruk, serta (5) harga pasar bawang merah yang rendah.

2. Data Mining dan Klasifikasi

Untuk meminimalisir risiko kegagalan dan memprediksi produktivitas hasil pertanian, diperlukan teknologi komputasi melalui pengolahan data. Salah satu metode dalam pengolahan

data adalah klasifikasi, yang merupakan salah satu tugas penting dalam *data mining* di mana sebuah pengklasifikasi dibangun dari sekumpulan data latih yang kelasnya telah ditentukan sebelumnya [2]. Proses klasifikasi ini pada dasarnya adalah pengelompokan fitur ke dalam kelas yang sesuai, yang di dalam pengenalan pola disebut sebagai metode terbimbing (*supervised*).

3. Algoritma Naive Bayes

Salah satu algoritma dalam teknik klasifikasi yang mudah diimplementasikan dan cepat prosesnya adalah Naive Bayes [3]. *Naive Bayes Classifier* merupakan metode klasifikasi probabilistik yang berakar pada teorema Bayes dengan ciri utama berupa asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi atau kejadian. Metode ini dirancang untuk memprediksi probabilitas dengan asumsi bahwa antar satu kelas dengan kelas lainnya tidak saling bergantung (*independen*).

4. Evaluasi Model

Dalam pengujian algoritma klasifikasi (seperti penerapannya menggunakan RapidMiner), data yang telah disiapkan umumnya dikelompokkan menjadi dua bagian menggunakan teknik *split validation*, yakni data latih (*training data*) dan data uji (*testing data*) untuk memastikan model memiliki generalisasi yang baik [4]. Evaluasi kinerja dari model prediksi tersebut kemudian diukur menggunakan perhitungan *confusion matrix* guna mendapatkan nilai ukur performa, di antaranya: *accuracy* (tingkat kedekatan hasil prediksi terhadap fakta sebenarnya), *precision* (tingkat ketepatan antara informasi yang diminta dengan jawaban sistem), dan *recall* (tingkat keberhasilan sistem dalam menemukan kembali informasi).

Metode Penelitian

Data yang digunakan dalam penelitian ini diperoleh dari situs web Badan Pusat Statistik Kabupaten Brebes, yang mencakup data pertanian dari Kecamatan Wanasari pada tahun 2018. Dataset tersebut terdiri dari beberapa variabel yang dianggap relevan dalam memprediksi produktivitas hasil pertanian bawang merah, yaitu: jenis bibit, pupuk, penyakit, pengobatan, pengairan, dan hasil produktivitas. Berikut adalah contoh dataset yang digunakan dalam penelitian ini:

No	Bibit	Pupuk	Penyakit	Pengobatan	Pengairan	Hasil
1	Bagus	Cukup	Normal	Sangat cukup	Cukup	Panen
2	Bagus	Kurang	Mboler	Sedang	Kurang	Gagal
3	Bagus	Kurang	Ulat	Sangat cukup	Kurang	Gagal
4	Kurang Bagus	Cukup	Normal	Kurang	Cukup	Panen
5	Kurang Bagus	Cukup	Ulat	Kurang	Cukup	Gagal
6	Kurang Bagus	Kurang	Normal	Kurang	Kurang	Gagal
7	Bagus	Cukup	Ulat	Sedang	Cukup	Panen

8	Bagus	Cukup	Mboler	Sangat Cukup	Cukup	Panen
9	Bagus	Cukup	Mboler	Kurang	Cukup	Gagal
10	Bagus	Cukup	Ulat	Kurang	Cukup	Panen
.....						
200	Kurang Bagus	Kurang	Mboler	Sedang	Kurang	Gagal

Kerangka kerja dalam penelitian ini melibatkan serangkaian langkah yang terstruktur dan sistematis untuk mengatasi masalah yang ada. Proses dimulai dengan pengumpulan data yang relevan, diikuti dengan analisis pra-pemrosesan data untuk memastikan kualitas dan kelayakan dataset. Setelah itu, teknik analisis yang sesuai diterapkan untuk membangun model prediksi. Dalam penelitian ini, teknik yang digunakan adalah metode klasifikasi Naive Bayes, yang diharapkan dapat menghasilkan model yang efektif dalam memprediksi produktivitas hasil pertanian bawang merah berdasarkan variabel-variabel yang ada. Setiap tahapan dilakukan dengan pendekatan yang jelas dan metodologi yang tepat untuk memastikan hasil yang valid dan dapat diandalkan.

1. Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh langsung dari informasi yang akurat dan relevan, yang mendukung proses penelitian, Data diambil dari website Badan Pusat Statistik Kabupaten Brebes, yang mencakup data dari Kecamatan Wanasari pada tahun 2018.

2. Metode Analisis Data

Penelitian ini menggunakan metode klasifikasi Naive Bayes untuk memprediksi produktivitas hasil pertanian bawang merah. Naive Bayes adalah metode probabilistik yang digunakan untuk memprediksi kelas data berdasarkan asumsi bahwa antar kelas tidak saling bergantung atau independen. Bentuk umum dari rumus Naive Bayes dalam klasifikasi adalah sebagai berikut:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Dimana:

- XXX: Data dengan kelas yang belum diketahui
- HHH: Hipotesis yang merujuk pada kelas spesifik
- $P(H|X)P(H|X)P(H|X)$: Probabilitas hipotesis HHH berdasarkan kondisi XXX (probabilitas posterior)
- $P(X|H)P(X|H)P(X|H)$: Probabilitas XXX berdasarkan kondisi pada hipotesis HHH
- $P(H)P(H)P(H)$: Probabilitas hipotesis HHH (probabilitas prior)
- $P(X)P(X)P(X)$: Probabilitas data XXX

3. Rencana Pengujian

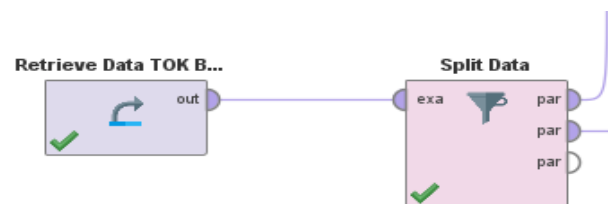
Pengujian dilakukan untuk memastikan bahwa sistem berfungsi dengan baik serta untuk mengidentifikasi potensi kesalahan yang mungkin terjadi. Pendekatan

pengujian yang digunakan dalam penelitian ini adalah uji angket kelayakan untuk mengevaluasi kinerja sistem.

Hasil dan Pembahasan

1. Pembagian Dataset

Sebelum dilakukan proses pra-pemrosesan data, dataset yang diperoleh dari website Badan Pusat Statistik Kabupaten Brebes dalam format Excel masih memerlukan penyesuaian untuk dapat diolah lebih lanjut. Dataset tersebut mencakup data pertanian dari Kecamatan Wanasari pada tahun 2018 dan disiapkan untuk dianalisis menggunakan metode klasifikasi Naive Bayes pada aplikasi RapidMiner.

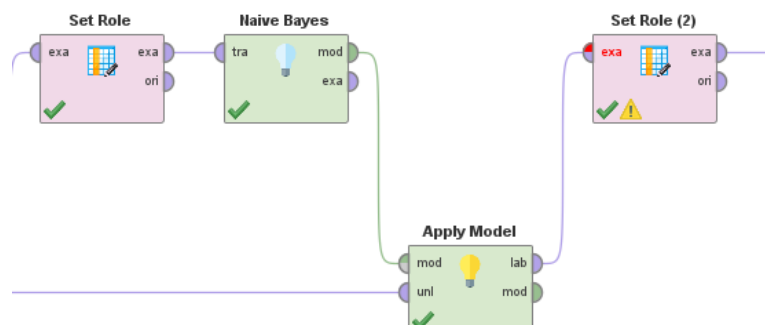


Gambar 1. Pembagian Dataset, Sumber : Dokumentasi Pribadi

Karena penelitian ini hanya menggunakan satu algoritma, yaitu Naive Bayes, pembagian dataset dilakukan dengan memisahkan data menjadi dua bagian: data latih (training data) dan data uji (testing data). Data latih digunakan untuk membangun model prediksi, sedangkan data uji digunakan untuk mengevaluasi performa model. Proses pembagian ini bertujuan untuk memastikan bahwa model yang dihasilkan mampu memprediksi produktivitas hasil pertanian bawang merah secara akurat.

2. Menentukan Label (Set Role)

Proses *Set Role* di RapidMiner digunakan untuk menentukan peran (role) dari atribut dalam dataset, seperti atribut mana yang akan digunakan sebagai label (target) dan atribut mana yang akan digunakan sebagai prediktor (features). Dalam penelitian ini, atribut **Hasil** yang berisi nilai kategori "Panen" dan "Gagal" diatur sebagai **label** karena menjadi target prediksi produktivitas bawang merah.



Gambar 2. Pe-labelan Data Set Role, Sumber: Dokumentasi Pribadi

Penelitian ini menggunakan dua kali proses *Set Role*, yaitu setelah tahap **Split Data** dan setelah **Apply Model**, yang masing-masing memiliki tujuan dan fungsi yang spesifik:

Set Role Setelah Split Data

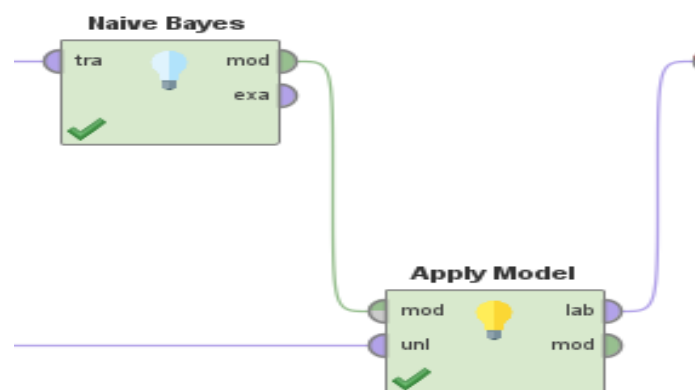
- **Tujuan:** Pada tahap ini, atribut target (**Hasil**) ditentukan sebagai label sebelum data diproses lebih lanjut. Hal ini memastikan bahwa algoritma pembelajaran mesin dapat mengenali variabel target yang akan diprediksi.
- **Konteks:** Proses ini dilakukan setelah pembagian dataset menjadi data pelatihan (*training data*) dan data pengujian (*test data*). Dengan menentukan *role* di awal, struktur data menjadi konsisten, sehingga memudahkan algoritma dalam mempelajari hubungan antara variabel prediktor dan target.

Set Role Setelah Apply Model

- **Tujuan:** Proses *Set Role* pada tahap ini bertujuan untuk mengatur ulang atribut yang dihasilkan oleh model prediksi. Sebagai contoh, atribut **Prediction(Hasil)**, yang merupakan hasil prediksi model, dapat ditetapkan sebagai target tambahan atau prediktor baru untuk evaluasi kinerja model.
- **Konteks:** Proses ini penting untuk membandingkan hasil prediksi dengan nilai aktual dalam dataset pengujian. Dengan cara ini, evaluasi model dapat dilakukan secara lebih terstruktur, seperti melalui perhitungan akurasi atau metrik kinerja lainnya.

3. Naive Bayes & Apply Model

Setelah peran atribut ditetapkan, algoritma **Naive Bayes** digunakan untuk membangun model prediksi. Naive Bayes adalah algoritma berbasis probabilitas yang sederhana tetapi sangat efektif untuk data kategorikal. Dalam kasus ini, Naive Bayes akan mempelajari hubungan antara variabel input (Bibit, Pupuk, Penyakit, Pengobatan, dan Pengairan) dengan label target (Hasil).



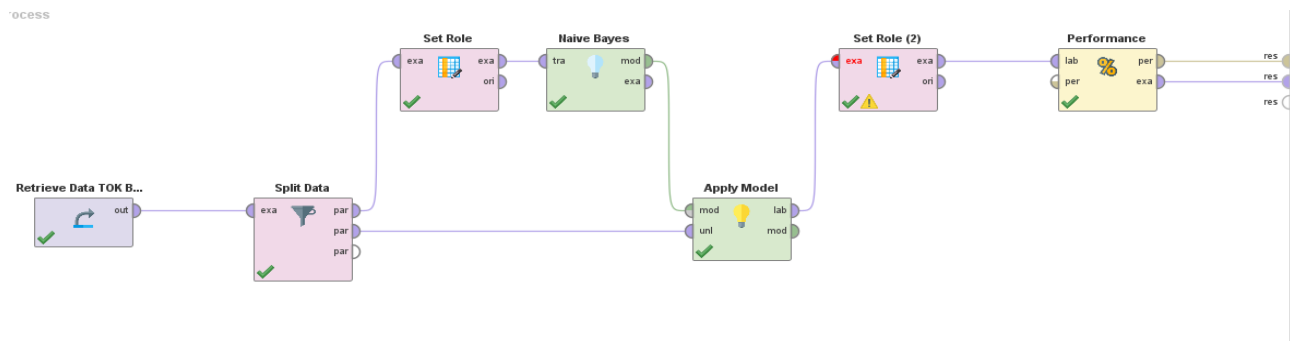
Gambar 3. Pe-labelan Naivebayes dan Apply Model, Sumber: Dokumentasi Pribadi

Operator **Apply Model** digunakan untuk menerapkan model yang telah dihasilkan dari proses pelatihan (training) pada dataset uji (testing data). Langkah ini merupakan tahapan penting dalam evaluasi model, di mana model yang telah dibuat akan diuji untuk menghasilkan prediksi berdasarkan data baru.

4. Performance

Operator **Performance** digunakan untuk mengevaluasi kualitas model prediksi yang dihasilkan. Evaluasi kinerja model dilakukan dengan menggunakan berbagai metrik evaluasi, di antaranya:

- **Akurasi:** Mengukur sejauh mana prediksi model sesuai dengan data sebenarnya. Akurasi dihitung sebagai rasio antara jumlah prediksi yang benar dengan total jumlah data.
- **Presisi:** Mengukur proporsi prediksi positif yang benar-benar positif terhadap semua data yang diprediksi positif.
- **Recall (Sensitivitas):** Mengukur kemampuan model dalam mendeteksi data positif dari keseluruhan data positif yang ada.
- **F1-Score:** Menggabungkan presisi dan recall untuk memberikan gambaran kinerja model secara keseluruhan, terutama jika terdapat ketidakseimbangan kelas dalam dataset.



Gambar 4. Pe-labelan secara Keseluruhan, Sumber: Dokumentasi Pribadi

Dalam penelitian ini, metrik-metrik tersebut digunakan untuk menilai kinerja model Naive Bayes yang diterapkan pada dataset pertanian bawang merah. Evaluasi dilakukan untuk memastikan bahwa model mampu memberikan hasil prediksi yang valid dan andal dalam memprediksi produktivitas hasil pertanian.

Evaluasi dan Validasi

Evaluasi dan validasi pada penelitian ini dilakukan dengan menggunakan **confusion matrix** untuk menganalisis kinerja model prediksi, serta metrik lainnya seperti akurasi, presisi, dan recall. Berdasarkan hasil pengujian yang dilakukan, diperoleh hasil sebagai berikut

Metrik Evaluasi

1. Accuracy (Akurasi):

Akurasi dihitung dengan rumus:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

Dengan hasil:

$$\text{Accuracy} = \frac{21 + 25}{21 + 25 + 6 + 8} \times 100\% = 76.67\%$$

2. Class Recall:

- Recall untuk kelas "Panen": 72.41%

- Recall untuk kelas "Gagal": 80.65%

3. Class Precision:

- Presisi untuk kelas "Panen": 77.78%
- Presisi untuk kelas "Gagal": 75.76%

Hasil evaluasi menunjukkan bahwa model memiliki tingkat akurasi sebesar **76.67%**, dengan nilai presisi dan recall yang cukup baik untuk kedua kelas.

Kesimpulan

Penelitian ini menyimpulkan bahwa metode klasifikasi Naive Bayes adalah alat yang efektif untuk memprediksi produktivitas hasil pertanian bawang merah, khususnya di Kecamatan Wanasari, Kabupaten Brebes. Dengan menggunakan data yang relevan dari Badan Pusat Statistik Kabupaten Brebes, penelitian ini mengidentifikasi variabel-variabel signifikan seperti bibit, pupuk, penyakit, pengobatan, dan pengairan yang memengaruhi hasil panen. Implementasi metode Naive Bayes menunjukkan hasil prediksi dengan akurasi sebesar 76,67%, yang mengindikasikan performa yang cukup baik dalam memisahkan kategori "Panen" dan "Gagal."

Proses penelitian melibatkan pembagian dataset menjadi data pelatihan dan pengujian, penentuan label target, dan evaluasi model dengan metrik seperti akurasi, presisi, recall, serta F1-score. Operator di RapidMiner digunakan untuk membangun dan menerapkan model prediksi, serta mengevaluasi kinerjanya.

Hasil penelitian ini menunjukkan bahwa metode Naive Bayes tidak hanya dapat diimplementasikan untuk mendukung pengambilan keputusan, tetapi juga memiliki potensi untuk membantu petani dalam memprediksi hasil panen mereka dengan lebih baik. Dengan demikian, penelitian ini memberikan kontribusi praktis dalam bidang pertanian, khususnya dalam meningkatkan efisiensi dan keberlanjutan produksi bawang merah.

DAFTAR PUSTAKA

- [1] Natasya, F.M. (2021). Implementasi metode klasifikasi naive bayes dalam memprediksi produktivitas hasil pertanian bawang merah. *Jurnal Multidisciplinary Applications Of Quantum Information Science (Al-mantiq)* **1**, 36-42.
- [2] Permana, D.F.W., Mustofa, A.H., Nuryani, L., Kristiaputra, P.S., and Alamudin, Y. (2021). Budidaya bawang merah di Kabupaten Brebes. *Jurnal Bina Desa* **3**, 125-132.
- [3] Putri, R.Y., Marleni, and Akbar, W.K. (2021). Rasionalitas petani bawang merah saat gagal panen di Jorong Galagah Nagari Alahan Panjang Kecamatan Lembah Gumanti Kabupaten Solok. *Jurnal Pendidikan Tambusai* **5**, 5830-5839.
- [4] Surojudin, N., Ermanto, Danny, M., and Pratama, S. (2024). Implementation of the Naive Bayes algorithm for death due to heart failure using Rapid Miner. *Brilliance: Research Of Artificial Intelligence* **4**, 294-302.

- [5] Wibawa, A.P., Purnama, M.G.A., Akbar, M.F., and Dwiyanto, F.A. (2018). Metode-metode klasifikasi. *Prosiding Seminar Ilmu Komputer Dan Teknologi Informasi* **3**, 134-138.