

# Penerapan Algoritma KNN untuk Klasifikasi dalam Weather Forecasting dengan Optimasi Ensemble

*by Miftahudin Faiz*

---

**Submission date:** 12-Jun-2023 04:53PM (UTC+0700)

**Submission ID:** 2114393179

**File name:** incomtechii\_after\_pembahasan.docx (157.07K)

**Word count:** 3212

**Character count:** 20127



# Penerapan Algoritma KNN untuk Klasifikasi dalam Weather Forecasting dengan Optimasi Ensemble

Miftahudin Faiz<sup>1</sup>

<sup>1</sup>*Informatika, Universitas AMIKOM Yogyakarta,  
Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa  
Yogyakarta 55281, Indonesia*

\*Email Penulis Koresponden: [miftahudinfaiz@students.amikom.ac.id](mailto:miftahudinfaiz@students.amikom.ac.id)

## Abstrak :

Cuaca merupakan perubahan keadaan udara di atmosfer yang sifatnya tidak menentu. Perubahan yang terjadi dipengaruhi oleh beberapa variabel yang berbeda seperti temperatur, penyerapan kecepatan angin, dan faktor eksternal lainnya. Prediksi cuaca dapat dilakukan dengan mempelajari bagaimana sebuah atmosfer bekerja. Prediksi tersebut dapat dilaksanakan dengan perkembangan teknologi dan model matematika untuk mengetahui tingkah laku alam. Dengan model komputer serta kombinasi data yang telah diobservasi maka akan didapatkan hasil prediksi cuaca. Penelitian ini menggunakan menggunakan algoritma KNN dan *Ensemble learning* yang ditulis dengan Google Colab untuk mendapatkan hasil analisa prediksi cuaca berdasarkan dataset yang telah diberikan. Berdasarkan *evaluation metrics* yang didapat dalam penelitian ini, menunjukkan bahwa metode *Ensemble* terutama *Ensemble Boosting* bekerja sangat baik dan memiliki akurasi sebesar 98.9%. Maka dapat disimpulkan bahwa penerapan algoritma KNN dengan optimasi *Ensemble* mampu mengklasifikasikan *Weather Forecasting* terutama pada prediksi hujan dengan akurat.

30

This is an open access article under the [CC BY-NC license](#)



## Kata Kunci:

*Prediksi Cuaca;  
Klasifikasi;  
Optimasi;*

## Riwayat Artikel:

Received Jun x, 20xx  
Revised Nov x, 20xx  
Accepted Dec x, 20xx

## DOI:

[10.22441/incomtech.v10i3.7777](https://doi.org/10.22441/incomtech.v10i3.7777)

## 1. PENDAHULUAN

Cuaca merupakan kondisi udara yang diamati dalam periode yang relatif singkat atau pendek pada teritori yang terbatas [1]. Perubahan cuaca ini dapat dinyatakan dengan nilai parameter seperti suhu, kecepatan angin, tekanan, curah hujan sebagai komponen utamanya [5]. Dengan berkembangnya teknologi, komputer memiliki kemampuan untuk melakukan berbagai macam hal. Komputer dapat meramal cuaca

22

dengan melakukan pengamatan terhadap gambar satelit dan menentukan cuaca pada hari tersebut dan melakukan ramalan untuk cuaca selanjutnya. [2].

. Dalam memprediksi cuaca, algoritma data mining dapat digunakan untuk membantu mengidentifikasi pola dan tren data cuaca historis dan data cuaca saat ini untuk meramalkan cuaca di masa depan. Salah satu pendekatan umum dalam penggunaan algoritma data mining untuk prediksi cuaca adalah dengan memanfaatkan teknik Machine Learning, seperti regresi linear, regresi logistik, dan jaringan saraf tiruan untuk mempelajari hubungan antara berbagai variabel cuaca seperti suhu, kelembaban, tekanan udara, dan kecepatan angin dengan hasil prediksi cuaca. Data mining adalah suatu proses untuk menemukan informasi yang menarik dan tersembunyi dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, data warehouse atau tempat penyimpanan data lainnya [8]. Metode data mining lainnya yang dapat digunakan yaitu clustering dan decision tree. Clustering digunakan untuk mengelompokkan data cuaca menjadi kelompok-kelompok tertentu berdasarkan karakteristik yang serupa, sedangkan decision tree digunakan untuk membuat keputusan berdasarkan aturan logika berdasarkan data cuaca.

Namun, penting untuk diingat bahwa prediksi cuaca melibatkan banyak faktor kompleks, termasuk perubahan kondisi cuaca secara tiba-tiba dan faktor lain seperti topografi dan kondisi geografis. Oleh karena itu, meskipun algoritma data mining dapat memberikan prediksi cuaca yang baik, namun hasilnya tidak selalu akurat dan dapat berubah-ubah sesuai dengan perubahan kondisi cuaca aktual.

Optimasi algoritma dibutuhkan untuk mendapatkan hasil yang lebih akurat. Optimasi merupakan permasalahan komputasional yang bertujuan untuk menemukan solusi terbaik dari beberapa solusi yang mungkin dari sejumlah alternatif solusi dengan memenuhi sejumlah batasan (constraints) [13]. Dalam memilih teknik optimasi algoritma dalam prediksi data mining untuk prediksi cuaca, diperlukan pemilihan teknik yang tepat dan sesuai dengan karakteristik data cuaca yang ada. Selain itu, parameter dari teknik optimasi juga harus diatur dengan baik untuk mencapai hasil yang optimal. Teknik optimasi algoritma yang akan digunakan penulis yaitu Algoritma *Ensemble Bagging* dan *Boosting*. Algoritma Ensemble melibatkan penggabungan beberapa model yang berbeda untuk meningkatkan kinerja prediksi. Ensemble dapat meningkatkan akurasi dan keandalan prediksi dengan mengurangi efek overfitting dan underfitting.

## 2. METODE PENELITIAN

### 2.1. K-Nearest Neighbor

Metode K-Nearest Neighbor (KNN) adalah sebuah metode supervised yang berarti membutuhkan data training untuk mengklasifikasikan objek yang jaraknya paling dekat. Prinsip kerja K-Nearest Neighbor adalah mencari jarak terdekat antara data yang akan di evaluasi dengan k tetangga (neighbor) dalam data pelatihan [3].

Metode penelitian yang digunakan dalam penelitian ini adalah pengembangan algoritma K-Nearest Neighbor (KNN) yang dioptimasi dengan teknik Ensemble untuk memprediksi kondisi cuaca di suatu daerah pada waktu tertentu.

14

## 2.2. Algoritma Decision Tree

Decision tree atau pohon keputusan adalah pohon yang setiap cabangnya terdapat pilihan diantara sejumlah alternatif dan daunnya menunjukkan hasil keputusan dari apa yang dipilih [4]. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami [6].

## 2.3. Ensemble Learning

Ensemble Learning adalah metode dimana algoritma mempelajari data dengan cara menggabungkan beberapa algoritma secara bersamaan untuk mendapatkan hasil pemodelan yang kuat daripada hanya menggunakan satu algoritma saja [12]. Terdapat beberapa tipe Ensemble learning, pada penelitian ini penulis menggunakan *Ensemble Bagging* dan *Boosting*.

### 2.3. Ensemble Bagging

Bagging adalah salah satu teknik dari Ensemble method dengan cara memanipulasi data training, data training di duplikasi sebanyak  $d$  kali dengan pengembalian (sampling with replacement), yang akan menghasilkan sebanyak  $d$  data training yang baru, kemudian dari data  $d$  training tersebut akan dibangun classifier yang disebut sebagai bagged classifier [6]. Secara ringkas Bagging merupakan metode yang memungkinkan pengurangan nilai varians secara signifikan [10].

### 2.4. Ensemble Boosting

Boosting merupakan salah satu metode Ensemble yang cukup populer dan masih banyak di minati kalangan peneliti. Ide dasar dari Boosting yaitu bobot disetiap learning diatur memiliki bobot nonnegatif. Konsep sederhana dari Boosting yaitu memberikan bobot yang sama pada data training. Kemudian setelah itu, proses dilanjutkan dengan menentukan base learner (weak learner) yang merupakan suatu fungsi yang mengklasifikasikan data sampel yang telah diboboti [7]. Konsep Ensemble dengan Boosting bekerja melalui persiapan kumpulan model secara berurutan dan kemudian menggabungkan seluruh model untuk membuat harapan, model berikutnya mendapatkan keuntungan dari kesalahan model sebelumnya [11].

## 2.3. Ensemble

Ensemble adalah teknik di mana beberapa model dipadukan untuk mencapai performa yang lebih baik dibandingkan dengan menggunakan model tunggal. Pada pendekatan algoritma dan Ensemble memiliki tujuan yang sama, yaitu memperbaiki algoritma pengklasifikasi tanpa mengubah data, sehingga dapat dianggap ada 2 pendekatan saja, yaitu pendekatan level data dan pendekatan level algoritma [9]. Dalam hal ini, beberapa model KNN dibuat dengan memvariasikan nilai parameter  $K$ , sehingga setiap model memiliki sensitivitas yang berbeda terhadap data masukan.

Selanjutnya, algoritma KNN dan algoritma Decision Tree dioptimasi dengan teknik Ensemble. Ensemble adalah teknik di mana beberapa model dipadukan untuk mencapai performa yang lebih baik dibandingkan dengan menggunakan

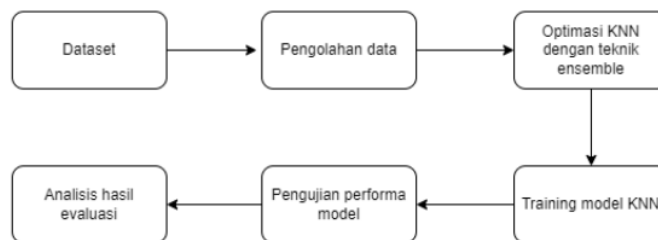
model tunggal. Dalam hal ini, beberapa model KNN dan Decision Tree dibuat dengan memvariasikan nilai parameter K, sehingga setiap model memiliki sensitivitas yang berbeda terhadap data masukan.

Setelah itu, data cuaca yang sudah diolah dan diformat digunakan untuk melatih model KNN yang sudah dioptimasi dengan teknik Ensemble. Data tersebut dibagi menjadi data latih dan data uji. Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk menguji performa model.

Terakhir, performa model dievaluasi menggunakan metrik seperti akurasi, presisi, *recall*, dan F1-score. Hasil evaluasi tersebut digunakan untuk menentukan apakah model KNN dan Decision Tree yang dioptimasi dengan teknik Ensemble dapat menghasilkan perkiraan cuaca yang akurat atau tidak. Jika hasil evaluasi menunjukkan bahwa model tersebut dapat menghasilkan perkiraan cuaca yang akurat, maka model tersebut dapat digunakan untuk memprediksi kondisi cuaca di suatu daerah pada waktu tertentu dengan menggunakan data cuaca yang sudah ada. Berikut adalah bagan alur penelitian mengenai perkiraan cuaca dengan algoritma KNN dan algoritma Decision Tree yang dioptimasi dengan Ensemble:

1. Pengumpulan data cuaca dari berbagai sumber seperti stasiun meteorologi, satelit, dan sensor cuaca lainnya.
2. Pengolahan dan pemformatan data cuaca agar dapat digunakan oleh algoritma KNN dan algoritma Decision Tree.
3. Optimasi algoritma KNN dan algoritma Decision Tree dengan teknik Ensemble dengan memvariasikan nilai parameter K untuk menciptakan beberapa model KNN.
4. Pelatihan model KNN dan algoritma Decision Tree dengan data cuaca yang sudah diolah dan diformat, dengan menggunakan data latih.
5. Pengujian performa model KNN dengan menggunakan data uji dan metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-score.
6. Analisis hasil evaluasi untuk menentukan apakah model KNN dan Decision Tree yang dioptimasi dengan teknik Ensemble dapat menghasilkan perkiraan cuaca yang akurat atau tidak.
7. Jika hasil evaluasi menunjukkan bahwa model tersebut dapat menghasilkan perkiraan cuaca yang akurat, maka model tersebut dapat digunakan untuk memprediksi kondisi cuaca di suatu daerah pada waktu tertentu dengan menggunakan data cuaca yang sudah ada.

Berikut adalah ilustrasi bagan alur penelitian tersebut:



Gambar 1. Bagan Alur Penelitian



Metode yang akan digunakan dalam penelitian ini adalah melakukan verifikasi prakiraan curah hujan secara statistik, prediksi curah hujan dengan model cuaca numerik, dan pengujian akurasi hasil prediksi.

Kejadian hujan yang diperkirakan dinilai memiliki sifat eksklusif, yaitu tidak boleh lebih dari satu kejadian yang terjadi, dan sedikitnya satu kejadian akan terjadi dalam satu waktu, sifat ini dikenal juga dengan istilah MECE (*mutually exclusive and collective exclusive*) [14,15]. Sifat eksklusif ini apabila diterapkan dalam prakiraan curah hujan yang dikeluarkan oleh BMKG, maka kejadian hujan akan terdiri dari 5 (lima) kategori, seperti diperlihatkan dalam Tabel 1.

Tabel 1. Kategori Prakiraan Hujan BMKG

No.	Kategori	Intensitas Curah Hujan	
		Setiap jam (mm/jam)	Setiap hari (mm/hari)
1.	Tidak hujan (berawan)	-	-
2.	Hujan ringan	0.1 – 4.9	0.1 – 19.9
3.	Hujan sedang	5.0	20.0 – 49.9
4.	Hujan lebat	10.0 – 20	50.0 - 100
5.	Hujan sangat lebat	>20.0	>100.0

### 2.3. Dataset

Data curah hujan yang digunakan adalah hasil pengukuran penakar hujan otomatis dan konvensional di wilayah Jabodetabek. Pengamatan hujan secara otomatis dengan menggunakan penakar hujan Tipe Hellman, *Automatic Weather Station* (AWS) dan *Automatic Rain- Gauge* (ARG) dengan interval waktu 10 menit, yaitu tujuh titik pengamatan (Bekasi, Tomang, Kelapa Gading, Manggarai, Pulomas, Ciomas, dan Ciganjur) namun dalam pengolahan data diakumulasi menjadi curah hujan harian (24 jam). Data dari penakar hujan konvensional berupa data curah hujan harian dari 18 stasiun BMKG dan pos pengamatan (Tabel 1). Verifikasi terhadap prediksi curah hujan yang sudah dibuat oleh BMKG untuk wilayah Jabodetabek menggunakan arsip prediksi yang lampau. Prediksi tersebut terbagi dalam tiga interval waktu, yaitu pagi (00.05 - 12.00 WIB), siang (12.05 – 18.00 WIB), dan malam (18.05 – 24.00 WIB). Penelitian prediksi numerik menggunakan model WRF dilakukan dengan syarat awal dan syarat batas menggunakan keluaran dari model *Global Forecast System* (GFS) dengan resolusi 0.5 x 0.5 setiap tiga jam yang disediakan oleh *National Ocean and Atmospheric*.

Berikut adalah alur pengolahan data cuaca dalam penelitian ini:

1. Pengumpulan data cuaca dari berbagai sumber seperti stasiun meteorologi, satelit, dan sensor cuaca lainnya. Dalam penelitian ini peneliti menggunakan data dari BMKG.
2. Preprocessing data cuaca dengan menghilangkan data yang tidak diperlukan, seperti data yang kosong, duplikat, atau tidak relevan.
3. Konversi data cuaca ke dalam format numerik, seperti mengubah kondisi cuaca menjadi angka yang merepresentasikan kondisi cuaca tersebut.
4. Pemisahan data menjadi data latih dan data uji.
5. Penerapan teknik scaling pada data latih dan data uji, yaitu melakukan normalisasi data sehingga setiap fitur memiliki rentang nilai yang sama.
6. Pemilihan fitur atau variabel yang akan digunakan dalam model KNN, seperti suhu, kelembapan, dan tekanan udara.

7. Pemodelan keterkaitan antar fitur atau variabel dengan menerapkan teknik korelasi untuk menentukan fitur yang memiliki korelasi tinggi dengan kondisi cuaca.
8. Data latih dan data uji disimpan dalam format yang sesuai untuk dapat digunakan dalam model KNN.

#### 2.4. Penelitian Eksperimental.

Data pengamatan curah hujan yang dapat digunakan dalam melakukan verifikasi masih terbatas, karena sebagian besar data yang tersedia adalah data curah hujan harian (jumlah curah hujan dalam 24 jam), sedangkan prakiraan curah hujan harian BMKG dibuat dalam beberapa interval waktu (pagi, siang dan malam). Sehingga verifikasi dilakukan pada curah hujan harian dan sebelum dilakukan verifikasi, prakiraan curah hujan tersebut dimodifikasi menjadi prakiraan curah hujan harian dengan mengambil prakiraan dengan jumlah curah hujan terbesar dari masing-masing interval dalam satu hari sebagai prakiraan yang akan diverifikasi. Hal ini penting diperhatikan agar sistem prediksi yang dibandingkan berada dalam kondisi yang identik seperti situasi prakiraan, variabel cuaca, dan lokasi. Contoh perubahan klasifikasi prakiraan hujan dari beberapa interval waktu menjadi prakiraan hujan harian. Untuk verifikasi prakiraan dimulai dengan membandingkan prediksi hujan harian dengan hasil pengamatan.

Prakiraan dibuat berdasarkan batasan kabupaten (atau kota) dan disebut juga wilayah prakiraan, maka curah hujan harian dari masing-masing stasiun dijadikan dasar untuk menyatakan hujan dan tidak-hujan di suatu wilayah prakiraan, daftar stasiun pengamatan yang digunakan untuk verifikasi. Suatu wilayah prakiraan didefinisikan terjadi hujan apabila salah satu titik atau stasiun pengamatan yang berada dalam wilayah tersebut mencatat hujan diatas 1 mm, 50 mm, dan 100 mm dalam satu hari, masing-masing untuk hujan dikotomi, hujan lebat, dan hujan sangat lebat. Selanjutnya prediksi yang benar (*hits*, dan *correct negatives*), dan yang salah (*false alarms* dan *misses*) dimasukkan dalam tabel kontijensi. Verifikasi juga dilakukan dengan menggunakan nilai ambang (*threshold*) 50 mm dan 100 mm per hari, untuk memperlihatkan kemampuan (akurasi dan *skill*) metode atau model dalam memprediksi hujan yang berpotensi terjadinya banjir dan ekstrim (sangat lebat).

### 3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan algoritma KNN optimasi *Ensemble Bagging* (KNN) dan *Ensemble Boosting* (Decision Tree) bertujuan untuk mencari metode dan algoritma terbaik untuk melakukan perkiraan cuaca. Proses klasifikasi, optimasi, dan perhitungan skor akurasi dilakukan menggunakan Google Colab dan mendapatkan hasil penelitian sebagai berikut:

#### 3.1. KNN

Pada klasifikasi KNN didapatkan *confusion matrix* yang dapat dilihat pada Tabel 2.

Tabel 2. *Confusion Matrix*

	Predicted 0	Predicted 1
Actual 0		
Actual 1		

Actual 0	187	18
Actual 1	1	71

Kemudian akan dikalkulasi untuk mendapatkan *classification report table* yang dapat dilihat pada Tabel 3.

Tabel 3. *Classification Report*

Class	Precision	Recall	F1-score	Support
0	0.99	0.91	0.95	205
1	0.80	0.99	0.88	72
Avg/total	0.94	0.93	0.93	277

### 3.2. Ensemble Bagging

Pada *Ensemble Bagging* didapatkan *confusion matrix* yang dapat dilihat pada Tabel 4.

Tabel 4. *Confusion Matrix*

	Predicted 0	Predicted 1
Actual 0	189	16
Actual 1	1	71

Kemudian akan dikalkulasi untuk mendapatkan *classification report table* yang dapat dilihat pada Tabel 5.

Tabel 5. *Classification Report*

Class	Precision	Recall	F1-score	Support
0	0.99	0.92	0.96	205
1	0.95	0.99	0.89	72
Avg/total	0.95	0.93	0.94	277

### 3.3. Ensemble Boosting

Pada *Ensemble Bagging* didapatkan *confusion matrix* yang dapat dilihat pada Tabel 6.

Tabel 6. *Confusion Matrix*

	Predicted 0	Predicted 1
Actual 0	202	3
Actual 1	0	72

Kemudian akan dikalkulasi untuk mendapatkan *classification report table* yang dapat dilihat pada Tabel 7.

Tabel 7. *Classification Report*

Class	Precision	Recall	F1-score	Support
0	1.00	0.99	0.99	205
1	0.96	1.00	0.98	72
Avg/total	0.99	0.99	0.99	277

### 3.4. Pembahasan

Berdasarkan data dari *classification report table*, *evaluation metrics* yang didapat dapat dilihat pada tabel 8 dan tabel 9.

Tabel 8. *Evaluation Metrics*

Model	Accuracy	Precision	Recall	F1-Score
KNN	0.931408	0.797753	0.986111	0.881988



Ensemble Bagging	0.938628	0.816092	0.986111	0.893082
Ensemble Boosting	0.989169	0.96	1.0	0.979592

Tabel 9. *Evaluation Metrics*

Model	Specificity	Average Precision	AUC
KNN	0.912195	0.790283	0.949153
Ensemble Bagging	0.921951	0.921951	0.954031
Ensemble Boosting	0.985366	0.96	0.992683

Berdasarkan *evaluation metrics* yang didapat, kedua KNN dan *Ensemble Bagging* memiliki *accuracy* sekitar 0.93 dan 0.94, sedangkan untuk *Ensemble Boosting* mempunyai *accuracy* yang lebih tinggi yaitu 0.99. Pada umumnya *accuracy* yang tinggi mengindikasikan bahwa model yang dibuat membuat prediksi yang akurat. *Precision* pada penelitian ini juga mendapat nilai yang tinggi sekitar 0.80 hingga 0.96 yang mengindikasikan bahwa model dapat mengidentifikasi *positive instances* secara benar dan memiliki *false positive* yang rendah. *Recall* pada penelitian ini juga mendapat nilai yang tinggi bahkan mendapat nilai 1.0 *F1-Score* pada penelitian ini mendapat nilai sekitar 0.88 hingga 0.98 yang mengindikasikan bahwa *precision* dan *recall* terdapat keseimbangan. *Specificity* pada penelitian ini mendapat nilai sekitar 0.91 hingga 0.99 yang mengindikasikan bahwa model dalam penelitian ini sangat mampu mengidentifikasi *negative instances* dengan benar. *Average precision* didapat dari rata-rata *precision-recall* dari perbedaan semua level pada *recall*. Pada penelitian ini mendapat nilai sekitar 0.79 hingga 0.96 yang mengindikasikan performa yang baik dalam *positive instance* yang lebih tinggi. AUC (Area Under the Curve) pada penelitian ini mendapat nilai sekitar 0.95 hingga 0.99 yang berarti bahwa model dalam penelitian ini memiliki kemampuan yang tinggi untuk membedakan antara *positive* dan *negative instances*. Dari *evaluation metrics* ini didapat bahwa model berperforma sangat baik, memiliki nilai *accuracy*, *precision*, *recall*, dan AUC yang tinggi. Namun nilai yang tinggi ini bukan berarti hal yang baik karena mungkin saja dalam model ini terjadi *overfitting*, maka akan dilakukan verifikasi untuk memastikan bahwa model bekerja dengan baik dalam tes yang berbeda. Maka akan dilakukan *cross-validation* yang dapat dilihat pada Tabel 10.

Tabel 10. *Cross-validation*

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average Score
KNN	0.961538	0.937984	0.953488	0.937984	0.96124	0.950447
Ensemble Bagging	0.961538	0.945736	0.953488	0.930233	0.96124	0.950447
Ensemble Boosting	0.946154	0.953488	0.96124	0.992248	0.96124	0.962874

Berdasarkan data yang didapat dari tabel *cross-validation*, model dalam penelitian ini tidak terjadi *overfitting* karena nilai *average score* dari ketiga model relatif cukup dekat yang dapat mengindikasikan bahwa model tidak terjadi *overfitting*.

#### 4. KESIMPULAN

Pada penelitian ini didapatkan bahwa perbandingan ketiga model, baik KNN maupun *Ensemble Bagging* dan *Ensemble Boosting*, memiliki akurasi yang tinggi. Ini menunjukkan bahwa ketiga model mampu dengan baik dalam memprediksi label dengan benar. Dengan *Ensemble Boosting* yang konsisten memiliki performa yang lebih baik di antara model lainnya. Namun pada model KNN dan *Ensemble Bagging* sedikit menampakkan tanda-tanda terjadinya *overfitting* bahkan pada *Ensemble Boosting* dapat berkemungkinan terjadi *overfitting* dalam beberapa kondisi dan selanjutnya melakukan penelitian lebih lanjut.

## REFERENSI

- [1] A. Luthfiara, A. Febriyanto, H. Lestiawan and W. Wicaksono, "Journal of Information System," Analisa Prakiraan Cuaca dengan Parameter Suhu, vol. 5, 2020.
- [2] M. F. Naufal, "Jurnal Teknologi Informasi dan Ilmu Komputer," ANALISIS PERBANDINGAN ALGORITMA SVM, KNN, DAN CNN UNTUK KLASIFIKASI CITRA CUACA, vol. 8, pp. 311-318, 2021.
- [3] R. N. Devita, H. W. Herwanto and A. P. Wibawa, "Jurnal Teknologi Informasi dan Ilmu Komputer," PERBANDINGAN KINERJA METODE NAIVE BAYES DAN K-NEAREST, vol. 5, pp. 427-434, 2018.
- [4] K. F. L. Simanjuntak, A. C. B. Barus and A. , "JOISIE Journal Of Information System And Informatics Engineering," IMPLEMENTASI METODE DECISION TREE DAN ALGORITMA C4.5 UNTUK KLASIFIKASI KEPRIBADIAN MASYARAKAT, vol. 5, pp. 51-59, 2021 [1].
- [5] A. M. Siregar, S. Faisal, Y. Cahyana and B. Priyatama, "Accounting Information System," PERBANDINGAN ALGORITME KLASIFIKASI UNTUK PREDIKSI CUACA, vol. 3, 2020.
- [6] A. Kurniawan and A. Prihandono, "JURNAL BISNIS DIGITAL DAN SISTEM INFORMASI," PENERAPAN TEKNIK BAGGING UNTUK MENINGKATKAN AKURASI KLASIFIKASI PADA ALGORITMA NAIVE BAYES DALAM MENENTUKAN BLOGGER PROFESIONAL, vol. 1, 2020.
- [7] L. Hakim, A. Saefuddin and S. Nisrina, "Statistika," Klasifikasi Varietas Unggul Padi Menggunakan Metode Bagging, Boosting, dan Extremely Randomized Trees, vol. 22, pp. 127-132, 2022.
- [8] N. A. A. Lestari, M. N. Hayati and F. D. T. Amijaya, "Jurnal EKSPONENSIAL," Analisis Cluster Pada Data Kategorik dan Numerik dengan Pendekatan Cluster Ensemble, vol. 11, 2020.
- [9] A. and A. Subekti, "Analisis Sentiment pada Ulasan Film Dengan Optimasi Ensemble Learning," JURNAL INFORMATIKA, vol. 7, pp. 15-23, 2020.
- [10] Y. Pristyanto and A. A. Zein, "Model Balanced Bagging Berbasis Decision Tree Pada Dataset Imbalanced Class," Jurnal SISFOKOM (Sistem Informasi dan Komputer), vol. 12, pp. 9-15, 2023.
- [11] S. E. H. Yulianti, O. Soesanto and Y. Sukmawaty, "Penerapan Metode Extreme Gradient Boosting(XGBOOST) pada Klasifikasi Nasabah Kartu Kredit," Journal of Mathematics: Theory and Applications, vol. 4, 2022.
- [12] M. L. Mu'tashim, A. Zaidiah and B. S. Yulistiawan, "Klasifikasi Ketepatan Lama Studi Mahasiswa Dengan Algoritma Random Forest Dan Gradient Boosting," Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA), 2023.
- [13] R. N. Devita and A. P. Wibawa, "Teknik-teknik optimasi knapsack problem," Sains, Aplikasi, Komputasi dan Teknologi Informasi, vol. 2, pp. 35-40, 2020.
- [14] V. K. Shankarnarayan and H. Ramakrishna, "Data Science and Management," Comparative study of three stochastic future weather forecast approaches: a case study, pp. 3-12, 2021.
- [15] I. Gustari, T. W. Hadi, S. Hadi and F. Renggono, "AKURASI PREDIKSI CURAH HUJAN HARIAN OPERASIONAL DI JABODETABEK : PERBANDINGAN DENGAN MODEL WRF".

# Penerapan Algoritma KNN untuk Klasifikasi dalam Weather Forecasting dengan Optimasi Ensemble

## ORIGINALITY REPORT

48%

SIMILARITY INDEX

46%

INTERNET SOURCES

15%

PUBLICATIONS

20%

STUDENT PAPERS

## PRIMARY SOURCES

1

[puslitbang.bmkg.go.id](http://puslitbang.bmkg.go.id)

Internet Source

15%

2

Submitted to Institut Teknologi Kalimantan

Student Paper

5%

3

[jtiik.ub.ac.id](http://jtiik.ub.ac.id)

Internet Source

3%

4

[gabya06.github.io](https://github.com/gabya06)

Internet Source

2%

5

[jurnal.fmipa.unmul.ac.id](http://jurnal.fmipa.unmul.ac.id)

Internet Source

1%

6

[ejr.stikesmuhkudus.ac.id](http://ejr.stikesmuhkudus.ac.id)

Internet Source

1%

7

[ojs.amikom.ac.id](http://ojs.amikom.ac.id)

Internet Source

1%

8

[docplayer.info](http://docplayer.info)

Internet Source

1%

9

[ejurnal.stmik-budidarma.ac.id](http://ejurnal.stmik-budidarma.ac.id)

Internet Source

1%

10	Submitted to Universitas Pertamina Student Paper	1 %
11	doaj.org Internet Source	1 %
12	Submitted to STIE Perbanas Surabaya Student Paper	1 %
13	www.ejr.stikesmuhkudus.ac.id Internet Source	1 %
14	filisaoctary.wordpress.com Internet Source	1 %
15	Submitted to Universitas Amikom Student Paper	1 %
16	Submitted to Universitas Raharja Student Paper	1 %
17	e-journals.unmul.ac.id Internet Source	1 %
18	ejournal.gunadarma.ac.id Internet Source	1 %
19	www.ejournal.pelitaindonesia.ac.id Internet Source	1 %
20	journal.stkom.ac.id Internet Source	1 %
21	www.mdpi.com Internet Source	1 %

22	publikasi.dinus.ac.id Internet Source	1 %
23	ejournal.uika-bogor.ac.id Internet Source	<1 %
24	scholar.smu.edu Internet Source	<1 %
25	eprints.uny.ac.id Internet Source	<1 %
26	jurnal.atmaluhur.ac.id Internet Source	<1 %
27	ojs.unud.ac.id Internet Source	<1 %
28	Maruli Tua Silaen. "KLASIFIKASI KARAKTERISTIK KEPERIBADIAN SISWA BERDASARKAN THE BIG FIVE PERSONALITY DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR (KNN)", Jurnal Informatika dan Rekayasa Elektronik, 2023 Publication	<1 %
29	Submitted to University College London Student Paper	<1 %
30	eprints.perbanas.ac.id Internet Source	<1 %
31	repo.undiksha.ac.id Internet Source	<1 %



32	<a href="http://mikeprastuti.blogspot.com">mikeprastuti.blogspot.com</a> Internet Source	<1 %
33	Sulthan Rafif, Pramana Yoga Saputra, Moch Zavaruddin Abdullah. "Classification of Trends in Lecturer Research Fields Using Naive Bayes Method", 2021 International Conference on Electrical and Information Technology (IEIT), 2021 Publication	<1 %
34	<a href="http://ejurnalunsam.id">ejurnalunsam.id</a> Internet Source	<1 %
35	<a href="http://journals.unisba.ac.id">journals.unisba.ac.id</a> Internet Source	<1 %
36	<a href="http://adoc.pub">adoc.pub</a> Internet Source	<1 %
37	<a href="http://www.researchgate.net">www.researchgate.net</a> Internet Source	<1 %
38	Nisa Hanum Harani. "Penerapan Adaboost Berbasis Pohon Keputusan Guna Menentukan Pola Masuknya Calon Mahasiswa Baru", Jurnal Transformatika, 2020 Publication	<1 %
39	<a href="http://journal.ubpkarawang.ac.id">journal.ubpkarawang.ac.id</a> Internet Source	<1 %
40	<a href="http://libraryproceeding.telkomuniversity.ac.id">libraryproceeding.telkomuniversity.ac.id</a> Internet Source	<1 %

41	media.neliti.com Internet Source	<1 %
42	repository.uinjkt.ac.id Internet Source	<1 %
43	Submitted to Sriwijaya University Student Paper	<1 %
44	core.ac.uk Internet Source	<1 %
45	d-nb.info Internet Source	<1 %
46	repository.ub.ac.id Internet Source	<1 %
47	www.neliti.com Internet Source	<1 %
48	www.scribd.com Internet Source	<1 %
49	Lian Ardiani, Herry Sujaini, Tursina Tursina. "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak", Jurnal Sistem dan Teknologi Informasi (Justin), 2020 Publication	<1 %
50	Dina Meilida Meliala, Penda Hasugian. "Perbandingan Algoritma K-Nearest Neighbor Dengan Decision Tree Dalam Memprediksi	<1 %

# Penjualan Makanan Hewan Peliharaan Di Petshop Dore Vet Clinic", Respati, 2020

Publication

51

Sumarni Sumarni, Suhardi Rustam. "Klasifikasi  
Topik Tugas Akhir Mahasiswa menggunakan  
Algoritma Particle Swarm Optimization dan K-  
Nearest Neighbor", ILKOM Jurnal Ilmiah, 2020

Publication

<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off

# Penerapan Algoritma KNN untuk Klasifikasi dalam Weather Forecasting dengan Optimasi Ensemble

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9