**Analisis dan Penerapan *Algoritma Machine Learning* di dalam Prediksi Cuaca**

**Wahyudi**

Ilmu Komputer, Universitas PGRI Wiranegara

[wahyuyudi2801@gmail.com](mailto:wahyuyudi2801@gmail.com)

**Abstrak**

Prediksi cuaca merupakan suatu hal yang penting bagi kehidupan sehari-hari. Prediksi cuaca akan memberikan informasi mengenai iklim yang akan datang di masa depan. Memprediksi cuaca membantu perencanaan kegiatan atau acara yang sudah dijadwalkan, baik menguntungkan ataupun merugikan. Kita harus tetap waspada terhadap cuaca yang buruk dengan menggunakan teknologi prediksi cuaca untuk mendapatkan peringatan bahaya. Banyak prediksi cuaca seperti prediksi curah hujan, badai petir, kondisi awan yang merupakan sebuah tantangan untuk meneliti atmosfer.

Paper ini menganalisis penerapan teknologi *machine learning* dalam prediksi cuaca. Saya menjelaskan beberapa prinsip dasar *machine learning* dan metode apa saja yang digunakan. Data cuaca yang berkualitas sangat diperlukan untuk pelatihan model. Proses pelatihan dan penilaian model, dengan fokus pada penyesuaian parameter dan penilaian performa menggunakan penyelidikan yang bersangkut-paut. Hasil penelitian menunjukkan potensi teknologi ini dalam meningkatkan akurasi prediksi cuaca, meskipun masih ada tantangan seperti kekurangan data historis dan ketidakpastian faktor eksternal. Penerapan teknologi *machine learning* dapat memberikan manfaat yang penting di berbagai lingkungan suatu usaha yang bergantung pada prediksi cuaca yang akurat.

**Kata Kunci:** *machine learning, prediksi cuaca, data cuaca*

**Pendahuluan**

Cuaca merupakan kondisi udara yang diamati dalam kurun waktu yang relatif singkat pada suatu wilayah yang terbatas. *World Climate Conference* menyatakan bahwa cuaca merupakan keadaan atmosfer yang diukur dengan kompleks mencakup perubahan, perkembangan, dan muncul atau hilangnya suatu fenomena udara. Kondisi cuaca dapat mempengaruhi pada kegiatan sehari-hari seperti menjemur pakaian. Meninjau potensi perubahan yang bersifat labil (Luthfiarta et al., 2020).

Saat ini iklim di negara Indonesia tidak menentu. Ketika cuaca di suatu wilayah terlihat cerah, tiba-tiba bisa berubah menjadi hujan ataupun badai. Perubahan iklim yang tidak menentu dapat menyebabkan kesulitan dalam memprediksi kondisi cuaca yang akan mendatang(Suma, 2021). *Badan Meteorologi Klimatologi dan Geofisika* (BMKG) telah menginformasikan prediksi musim hujan 2022/2023 pada 31 Agustus 2022. Hingga akhir agustus 2022, hasil pemantauan perkembangan musim telah terdapat 15% Zona Musim (ZOM) yang sudah memasuki musim hujan mencakup sebagian besar wilayah Sumatera bagian tengah dan selatan, Kalimantan, Sulawesi bagian selatan, Maluku Utara, Maluku dan Papua Barat bagian timur (Dr. Ir. Dodo Gunawan et al., 2023).

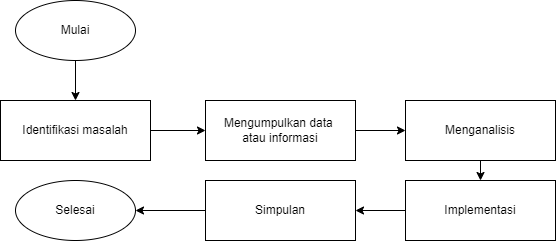
Supaya bisa memprediksi kondisi cuaca di suatu wilayah tertentu secara akurat, maka diperlukan sebuah teknologi modern yang dapat menganalisa dan memprediksi kondisi cuaca di wilayah tersebut. Salah satu teknologi yang dapat melakukan semua itu adalah *machine learning* (ML)(Suma, 2021). *Machine Learning* merupakan bagian dari bidang pengetahuan kecerdasan buatan yang banyak digunakan untuk membantu menyelesaikan berbagai masalah (Roihan et al., 2020). *Machine learning* mengolah data yang sudah diriset, kemudian akan mendapatkan hasil yaitu pola cuaca. Pola tersebut dapat membantu menganalisa cuaca yang berubah-ubah sehingga dapat memprediksi kondisi cuaca pada saat itu dengan tingkat akurasi yang tepat. Tingkat akurasi tersebut dapat dilihat dari data yang sudah diriset. Jika data yang digunakan sebagai riset yang menyesatkan dan menghasilkan *overfitting*, maka tingkat akurasi yang dihasilkan akan sangat buruk.

Penelitian ini akan menggunakan teknologi *machine learning* karena *machine learning* dapat menangani pekerjaan yang kompleks dan dapat menyesuaikan *inputan* data sehingga memberikan *output* atau solusi yang baik untuk masalah prediksi cuaca.

**Metode**

*Metodologi* merujuk pada kerangka dasar atau diagram alur yang digunakan untuk menyelesaikan langkah-langkah tugas akhir. *Metodologi* penulisan pada tugas akhir ini mencakup seluruh kegiatan atau aktifitas yang dilaksanakan untuk menganalisis dan memecahkan masalah pada kasus ini. Dalam tugas akhir ini, penerapan yang dilaksanakan adalah tentang menganalisis dan bagaimana cara menerapkan *machine learning* dalam prediksi cuaca.

Berikut merupakan gambar bagan alur atau *flow chart metodologi* penyelesaian tugas akhir yang digunakan untuk penerapan *machine learning* dalam prediksi cuaca.

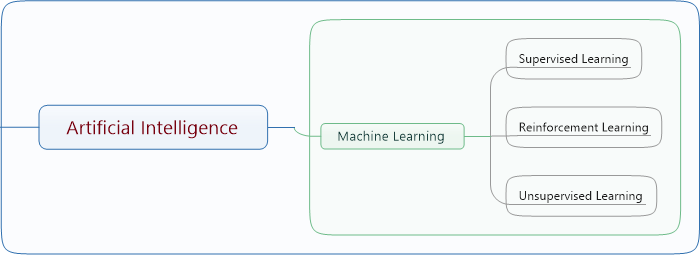


Gambar 1. *Metodologi* analisis dan penerapan *machine learning* dalam prediksi cuaca

**Pembahasan**

*Machine learning* atau biasa disingkat ML adalah ilmu komputer yang dapat bekerja tanpa harus diprogram secara jelas dan spesifik. Secara definisi, *machine learning* merupakan ilmu yang mempelajari tentang algoritma dan model statistik yang menggunakan sistem komputer untuk menyelesaikan tugas-tugas yang ada tanpa perintah yang jelas dan spesifik. Pemakaian teknologi ini berkaitan dengan pembelajaran mesin dan AI. Oleh karena itu, pastikan untuk terus berinteraksi dengan data (Dinata et al., 2020).

Prinsip bekerjanya *machine learning* sebenarnya tidak sama dengan apa yang dipergunakan *machine learning*. Namun secara prinsip, cara kerja dan pembelajaran mesin tetap serupa, yang melibatkan pengumpulan data, analisis data, dan pemilihan model atau teknik yang sesuai (Jamaaluddin & Indah, 2021). Penelitian terkini menyatakan bahwa ada tiga kategori dalam *machine learning*, antara lain *Supervised Learning*, *Unsupervised Learning*, dan *Reinforcement Learning* (Roihan et al., 2020). Struktur hubungan antara *artificial intelligence* dan *machine learning* dapat dilihat pada gambar 2.



Gambar 2. Skema AI dan ML (sumber: (Roihan et al., 2020))

Teknik *Supervised Learning* adalah metode klasifikasi di mana kumpulan data sepenuhnya diberi label untuk mengklasifikasi kelas yang tidak dikenal. Sedangkan, teknik *Unsupervised Learning* sering disebut sebagai *clustering* karena tidak memerlukan pemberian label pada kumpulan data dan hasilnya tidak mengidentifikasi kelas yang sudah ditentukan sebelumnya. Sedangkan, *Reinforcement Learning* biasanya berada diantara kedua teknik yang sudah dijelaskan sebelumnya, teknik ini bekerja secara dinamis di mana konsepnya harus menyelesaikan tujuan tanpa adanya pemberitahuan dari komputer secara jelas jika tujuan sudah tercapai (Roihan et al., 2020).

Ada beberapa algoritma yang digunakan *machine learning*, seperti *decision tree*, *k-Nearest Neighboars*(KNN), dan *naive bayes*. Sebenarnya masih ada algoritma lain yang digunakan *machine learning*, tapi di sini hanya menjelaskan 3 algoritma yang telah disebutkan tadi. Setiap *algoritma* memiliki prinsip kerja, kelebihan, dan kelemahannya masing-masing.

*Algoritma decision tree* adalah sebuah *diagram alir* yang berbentuk seperti struktur pohon di mana setiap *internal node*-nya menyatakan pengujian pada suatu *atribut*, setiap cabang menyatakan keluaran atau *output* dari pengujian tersebut dan *leaf node* menyatakan kelas-kelas atau *distribusi* kelas. *Node* yang paling atas disebut *root node* atau *node* akar. Proses *decision tree* melibatkan transformasi data dalam bentuk tabel menjadi model pohon, kemudian mengonversi model pohon tersebut menjadi aturan dan menyederhanakannya (Firdaus & C, 2019).

*Algoritma k-Nearest Neighboars*(KNN) merupakan metode yang digunakan untuk melakukan pengelompokan data berdasarkan jarak terpendek pada objek data. Penentuan nilai K yang terbaik pada *algoritma* ini adalah dengan berdasarkan data yang ada. Nilai K yang tinggi dapat mengurangi efek *noise* pada pengelompokan, bisa juga membuat batasan untuk pengelompokan supaya menjadi lebih samar. *Algoritma* ini termasuk dalam metode *supervised learning*, karena *instance query* perlu dikelompokkan berdasarkan mayoritas kategori pada *algoritma* K-NN (Cholil et al., 2021).

*Algoritma naive bayes* merupakan suatu metode pengelompokan kemungkinan yang sederhana dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari *basis* data yang sudah didapatkan. Suatu *algoritma* memanfaatkan *teorema bayes* dan memperkirakan *atribut* yang bebas yang dapat diberikan pada suatu nilai *variabel*. *Naive bayes* adalah sebuah *algoritma* yang ditemukan berdasarkan konsep teori *probabilitas Bayes* oleh seorang matematikawan dan pendeta yang berasal dari Inggris bernama Thomas Bayes. Beliau berhasil menghasilkan prediksi peluang yang akan datang berdasarkan pengalaman sebelumnya (Rachman & Handayani, 2021).

Dari beberapa kategori *algoritma machine learning* yang sudah dijelaskan sebelumnya. Penelitian ini hanya akan berfokus menganalisis penerapan atau implementasi *algoritma K-Nearest Neighbor*(KNN). Pada *algoritma* KNN, data yang sudah diteliti akan dihubungkan dengan faktor-faktor yang bisa mempengaruhi cuaca seperti *Fenomena La Nina El Nino* dan daerah pertemuan angin antar tropis. Data tersebut kemudian dikategorisasikan sehingga menghasilkan sebuah kesimpulan bahwa dengan *algoritma* K-NN dapat menghasilkan informasi mengenai kondisi cuaca (Rangkuti et al., 2021).

Penggambaran penggunaan salah satu *algoritma machine learning* yaitu K-NN misalnya kita akan memprediksi pada tahun 2023 apakah termasuk tahun yang akan mengalami penurunan atau peningkatan suhu. Dengan menggunakan *algoritma* K-NN, kita dapat memprediksi kondisi cuaca dengan melibatkan beberapa faktor contohnya seperti curah hujan, suhu minimum, kelembapan udara, suhu maksimum yang terjadi pada setiap harinya. Hal ini dapat dilakukan dengan menghitung nilai rata-rata pada suhu tahunan di sekitar tahun tersebut (Rangkuti et al., 2021).

Dalam memprediksi dan menghitung tingkat keakuratan data cuaca di Indonesia ada beberapa tahap, mulai dari tahap persiapan data, tahap membersihkan data, tahap proses data, sampai dengan tahap pembanding mana tingkat keakuratan yang terbaik jika menggunakan *parameter* K (tetangga terdekat) secara acak (Rangkuti et al., 2021).

**Simpulan**

*Machine learning* merupakan ilmu yang mempelajari tentang *algoritma* dan model *statistik* yang menggunakan *sistem* komputer untuk menyelesaikan tugas-tugas yang ada tanpa perintah yang jelas dan spesifik. *Machine learning* dibagi menjadi 3 kategori, yaitu *Supervised Learning*, *Unsupervised Learning*,dan *Reinforcement Learning*. Ada beberapa algoritma yang digunakan *machine learning*, seperti *decision tree, k-Nearest Neighboars*(KNN), dan *naive bayes*.

Penggambaran penggunaan *algoritma* K-NN misal kita akan memprediksi pada tahun 2023 apakah termasuk tahun yang akan mengalami penurunan atau peningkatan suhu. Dengan menggunakan *algoritma* K-NN, kita dapat menentukan kondisi cuaca dengan melibatkan beberapa faktor antara lain curah hujan, kelembapan udara, suhu minimum, suhu maksimum yang terjadi setiap harinya. Salah satu pendekatan yang dapat dilakukan adalah dengan menghitung nilai rata-rata suhu tahunan di sekitar tahun tersebut sebagai referensi dalam pengelompokan data.

Ada beberapa tahap dalam memprediksi dan menghitung tingkat keakuratan data cuaca, yaitu (1) tahap persiapan data, (2) tahap membersihkan data, (3) tahap proses data, (4) tahap pembanding.

**Daftar Pustaka**

Cholil, S. R., Handayani, T., Prathivi, R., & Ardianita, T. (2021). IJCIT (Indonesian Journal on Computer and Information Technology) Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa. *IJCIT (Indonesian Journal on Computer and Information Technology)*, *6*(2), 118–127. http://ejournal.bsi.ac.id/ejurnal/index.php/ijcit

Dinata, R. K., Hasdyna, N., & Dr. Fajriana, S.Si., M. S. (2020). *Machine Learning.pdf*. Unimal Press. https://repository.unimal.ac.id/6707/1/Machine Learning.pdf

Dr. Ir. Dodo Gunawan, D., Sopaheluwakan, D. A., Haryoko, D. U., Dr. Supari, Setiawan, D. A. M., & Adi Ripaldi, M. S. (2023). BMKG Climate Outlook 2023. *Badan Meteorologi, Klimatologi, Dan Geofisika*, 1–23. https://iklim.bmkg.go.id/bmkgadmin/storage/buletin/BMKG Climate Outlook 2023.pdf

Firdaus, M., & C, T. A. (2019). *Implementasi algoritma decision tree untuk klasifikasi pola serangan pada log file*. http://repository.unmuhjember.ac.id/7150/1/JURNAL.pdf

Jamaaluddin, & Indah, S. (2021). Buku Ajar Kecerdasan Buatan. *Umsida Press*, 121. http://eprints.umsida.ac.id/9443/1/Bahan Ajar Jamal Indah.pdf

Luthfiarta, A., Febriyanto, A., Lestiawan, H., & Wicaksono, W. (2020). Analisa Prakiraan Cuaca dengan Parameter Suhu, Kelembaban, Tekanan Udara, dan Kecepatan Angin Menggunakan Regresi Linear Berganda. *JOINS (Journal of Information System)*, *5*(1), 10–17. https://doi.org/10.33633/joins.v5i1.2760

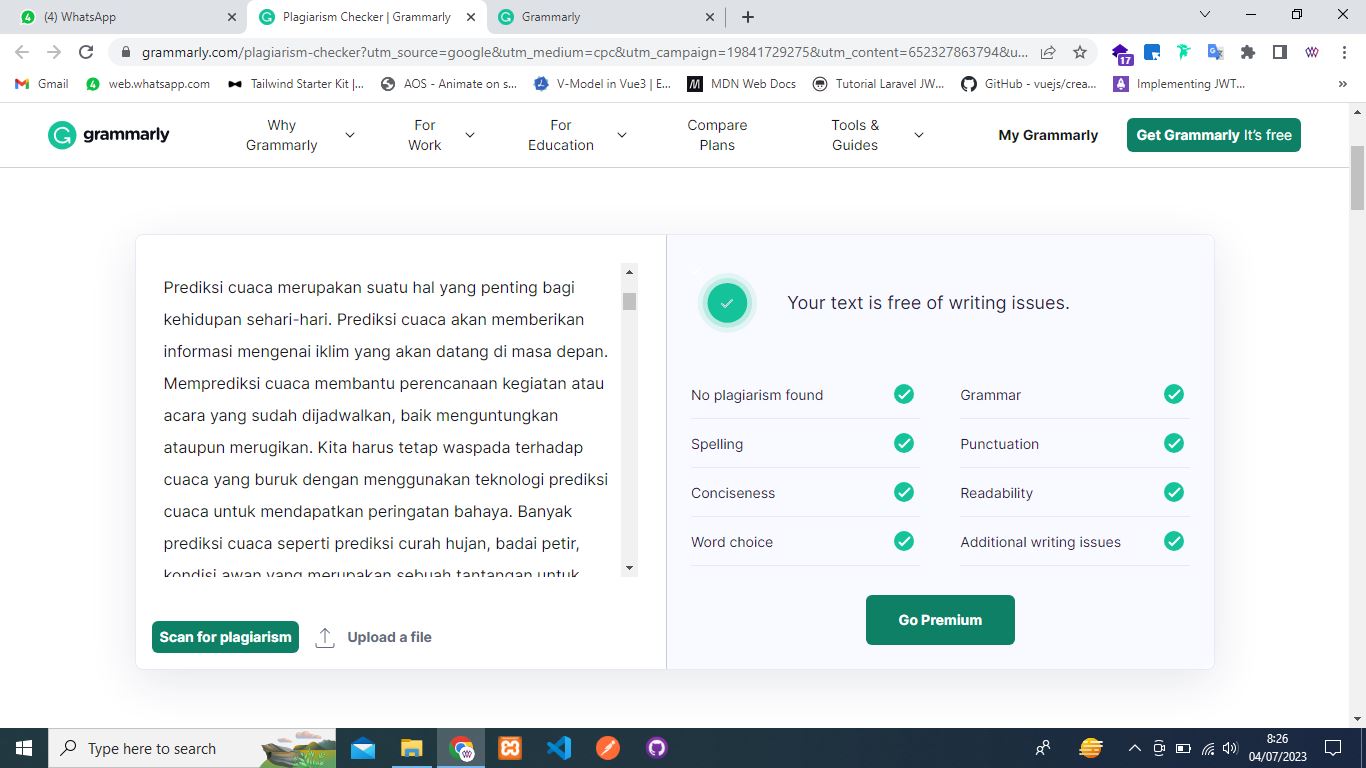
Rachman, R., & Handayani, R. N. (2021). Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM. *Jurnal Informatika*, *8*(2), 111–122. https://doi.org/10.31294/ji.v8i2.10494

Rangkuti, M. Y. R., Alfansyuri, M. V., & Gunawan, W. (2021). Penerapan Algoritma K-Nearest Neighbor (Knn) Dalam Memprediksi Dan Menghitung Tingkat Akurasi Data Cuaca Di Indonesia. *Hexagon Jurnal Teknik Dan Sains*, *2*(2), 11–16. https://doi.org/10.36761/hexagon.v2i2.1082

Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. *IJCIT (Indonesian Journal on Computer and Information Technology)*, *5*(1), 75–82. https://doi.org/10.31294/ijcit.v5i1.7951

Suma, B. (2021). *Implementasi Machine Learning Di Dalam Prediksi Cuaca Disusun sebagai salah satu syarat untuk kelulusan Program Strata 1 di Program Studi Teknik Informatika Universitas Pasundan Bandung oleh : BANDUNG SEPTEMBER 2020*. *September 2020*. https://doi.org/10.13140/RG.2.2.16086.47680

**Hasil Check Plagiasi**

****