

IMPLEMENTASI MACHINE LEARNING UNTUK MEMPREDIKSI CUACA MENGGUNAKAN SUPPORT VECTOR MACHINE

Oleh:

Moch. Fauzi Raja Mahendra

Nuril Lutvi Azizah

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Agustus, 2023

Pendahuluan

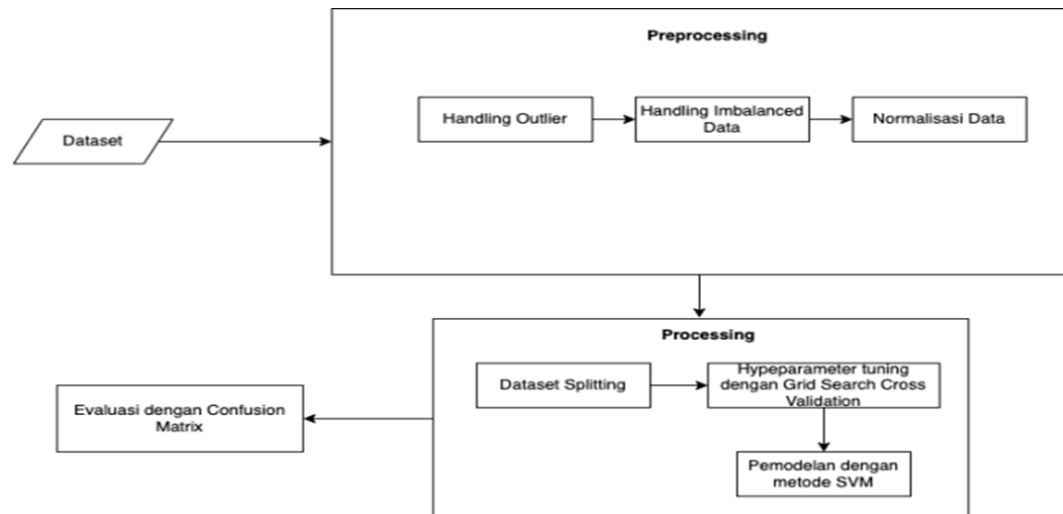
- Cuaca ialah keseluruhan dari peristiwa yang berlangsung di atmosfer bumi selama beberapa hari. Cuaca yang terjadi dalam jangka waktu lebih lama disebut sebagai iklim. Hal yang dapat memengaruhi keadaan suatu cuaca yakni antara lain: suhu, tekanan udara, kecepatan angin, udara serta curah hujan.
- Kini iklim di Indonesia sendiri kadang-kadang tidak menentu. Bahkan, ketika cuaca di suatu wilayah cerah, dalam waktu yang tidak lama bisa berubah menjadi hujan atau bahkan badai. Tidak menentunya perubahan pada iklim ini dapat mengakibatkan kesulitan dalam pemrediksian suatu cuaca.
- Badan Meteorologi Klimatologi Geofisika (BMKG) Indonesia menyampaikan bahwasanya suhu udara rata-rata bulan juli di Indonesia periode 1981 – 2010 ialah $26,39^{\circ}\text{C}$ sedangkan pada tahun 2020 bulan februari ialah 27.22°C dengan demikian anomali peningkatan suhu udara rata-rata sebanyak $0,83^{\circ}\text{C}$.

Pendahuluan

- Machine learning adalah salah satu aplikasi yang menjadi bagian dari Artificial intelligence. Machine learning dalam bahasa Indonesia disebut juga pembelajaran mesin. Machine Learning di peruntukkan dalam pengembangan sebuah sistem atau mesin cerdas yang dapat belajar sendiri tanpa harus di program oleh manusia secara berulang kali .
- Metode yang bisa dipergunakan dalam proses pemrediksian salah satunya yaitu support vector machine. Metode ini ialah tehnik dengan basis machine learning di mana masih jarang di gunakan dalam proses pemrediksian, baik dipergunakan pada kasus klasifikasi ataupun regresi
- Berdasarkan latar belakang yang diuraikan di atas, maka peneliti tertarik untuk membahas dan mengkaji lebih dalam mengenai penelitian yang berkaitan dengan Prediksi cuaca dengan judul, “Implementasi Machine Learning Untuk Memprediksi Cuaca Menggunakan Support Vector Machine”.

Metode Penelitian

- Tahapan penelitian merupakan gambaran umum terkait alur penelitian yang akan dilakukan dalam pengerjaan penelitian ini dari awal hingga akhir. Tahapan yang dilakukan dalam penelitian ini dapat di paparkan melalui diagram alir seperti pada gambar 1. Seluruh tahapan ini dikerjakan menggunakan Bahasa pemrograman python. Dengan rincian untuk tahapan preprocessing hingga pemodelan menggunakan software jupyter notebook.



Metode Penelitian

Preprocessing

- Data mentah tidak bisa dipergunakan langsung oleh sistem. Dengan begitu, perlu dilakukan beberapa preprocessing untuk sedikit melakukan modifikasi data supaya kualitas data yang dipergunakan meningkat. Preprocessing dalam penelitian dilakukan sebagai pembersihan data sebelum dilakukannya modeling.

Handling Outlier

- Handling outlier menggunakan teknik Z-score. Z-score dapat dipergunakan untuk membantu penentuan apakah suatu data memiliki nilai ekstrem, ataupun outlier. Data outlier ialah data dengan jauh dari rata-rata. Aturan umumnya ialah Z-score bernilai kurang dari -3 ataupun lebih dari $+3$ menjelaskan bahwasanya nilai data bernilai ekstrem. Dengan demikian data dengan nilai lebih dari batas bawah serta batas atas akan dihapuskan.

Metode Penelitian

Klasifikasi Dengan SVM

- Tahapan pertama dari pengklasifikasian menggunakan SVM ialah membuka data yang sudah diekstraksi fitur ke dalam jupyter notebook dengan library pandas. Saat data telah berhasil dimuat, selanjutnya data dibagi antara data X dan data y di mana data X ialah kolom fitur sedangkan data y ialah kolom target. Sesudah dibagi antara data X dengan y , berikutnya dibagi antara data train dengan data test pada data X memakai modul scikit-learn yakni `train_test_split`. Besaran pembagian data yakni 70% untuk data train sedangkan 30% untuk data test.

Metode Penelitian

Klasifikasi Dengan SVM

- Setelah dilakukan pembagian dataset, selanjutnya adalah tahapan modelling. Untuk memperoleh parameter teroptimal dalam kasus ini, peneliti memanfaatkan Teknik GridSearchCV yang mana Teknik ini mampu melakukan pencarian parameter optimal dari algoritma yang dipergunakan untuk kasus yang tengah dilakukan analisa. GridSearchCV ialah bagian dari modul scikit-learn di mana ditunjukan secara otomatis serta sistematis memvalidasi beberapa model dan setiap hyperparameter. Saat proses running GridSearchCV selesai, akan diperoleh model disertai score test dan score train.

Metode Penelitian

evaluasi

- Pada tahap evaluasi, peneliti menggunakan confusion matrix. Dapat dilihat pada gambar 6 bahwa nilai akurasi yang dihasilkan pada penelitian ini adalah 83%. sedangkan untuk nilai precisionnya adalah 79% dan F1 Score adalah 79%

Metode Penelitian

Penutup

- Penelitian ini telah berhasil melakukan klasifikasi mulai dari preprocessing sampai evaluasi. Berdasarkan penelitian yang telah dilakukan system dapat melakukan klasifikasi prediksi hujan dengan skor akurasi sebesar 83% dan precision yang dihasilkan 0.79% sehingga jika di lihat dari gap antara skor akurasi dan precision model tidak mengalami overfitting maupun underfitting. Dengan kata lain penelitian ini menghasilkan model yang fit dan mendapatkan hasil pengujian sesuai dengan realisasi yang di harapkan. Kelebihan dari metode SVM ini mampu menghasilkan model klasifikasi yang sangat baik.

Referensi

- **REFERENSI**
- [1] B. Suma, Implementasi Machine Learning Di Dalam Prediksi Cuaca. 2020. doi: 10.13140/RG.2.2.16086.47680.
- [2] F. R. Lumbanraja, R. S. Sani, D. Kurniawan, and A. R. Irawati, “Implementasi Metode Support Vector Machine Dalam Prediksi Persebaran Demam Berdarah Di Kota Bandar Lampung,” *J. Komputasi*, vol. 7, no. 2, 2019, doi: 10.23960/komputasi.v7i2.2426.
- [3] A. Sulthoni Akbar, C. Dewi, and R. C. Wihandika, “Prediksi Cuaca Kota Denpasar menggunakan Algoritma ELM dengan Optimasi Quantum Delta Particle Swarm Optimization,” vol. 5, no. 3, pp. 1126–1135, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] NUR ROCHMAN DARMAWAN, *Prediksi Kondisi Cuaca Kota Surabaya Menggunakan Metode Artificial Neural Network Prediction of Surabaya City Weather Conditions Using Artificial Neural Network Method*. 2019.
- [5] M. Y. R. Rangkuti, M. V. Alfansyuri, and W. Gunawan, “Penerapan Algoritma K-Nearest Neighbor (Knn) Dalam Memprediksi Dan Menghitung Tingkat Akurasi Data Cuaca Di Indonesia,” *Hexag. J. Tek. dan Sains*, vol. 2, no. 2, pp. 11–16, 2021, doi: 10.36761/hexagon.v2i2.1082.

