

BAB I

PENDAHULUAN

1.1 Latar Belakang

Campak merupakan sebuah penyakit yang masuk ke dalam kategori 10 penyakit paling menular pada negara berkembang termasuk negara Indonesia (Bottmer dkk., 2022). Pada tahun 2022, Indonesia memiliki kasus campak sebesar 3.341 yang tersebar pada 223 kota dan kabupaten. Data tersebut memiliki peningkatan sebesar 32 kali lipat dibandingkan tahun sebelumnya (Pulungan & Sari, 2023a). Salah satu penyebabnya adalah penurunan cakupan imunisasi yang turun dari 87% menjadi 84% pada vaksin campak dosis pertama antara tahun 2013 dan 2022 (Chacko dkk., 2023).

Pada penelitian (Widyaningsih dkk., 2022) yang berjudul *"Measles free prediction and control strategy to achieve measles elimination target in Indonesia based on SVEIR model"* mengatakan bahwa seluruh dunia menargetkan untuk mengeliminasi penyakit campak pada tahun 2030, namun penelitian ini memprediksi bahwa Indonesia akan bebas campak pada tahun 2046, strategi pengendalian tersebut mencakup mengurangi kontak fisik dengan penderita dan meningkatkan cakupan vaksinasi campak. Namun, upaya memutus rantai penularan sering kali terkendala oleh proses diagnosis konvensional yang memakan waktu, terutama di wilayah dengan fasilitas laboratorium yang terbatas. Selain itu, gejala klinis campak sering kali mirip dengan penyakit lain seperti demam berdarah, rubella, atau infeksi pernapasan akut, yang menyebabkan tantangan dalam penegakan diagnosis dini secara akurat.

Selain pemilihan algoritma yang tepat, kinerja model pembelajaran mesin sangat bergantung pada inisialisasi parameter yang digunakan. Penentuan parameter secara manual sering kali tidak menghasilkan performa optimal. Oleh karena itu, penerapan metode optimasi parameter yang sistematis seperti *Grid Search* menjadi krusial untuk mengeksplorasi kombinasi *learning rate*, *epoch*, dan jumlah *neuron* terbaik guna memastikan stabilitas dan akurasi model.

Di sisi lain, dalam konteks diagnosis medis, akurasi prediksi saja tidak cukup; interpretasi terhadap bagaimana model mengambil keputusan juga sangat dibutuhkan. Analisis *Feature Importance* (kepentingan fitur) diperlukan untuk memahami gejala klinis mana yang paling dominan mempengaruhi keputusan model. Hal ini bertujuan untuk memvalidasi apakah pola yang dipelajari oleh algoritma selaras dengan pengetahuan medis, serta untuk mengidentifikasi fitur yang benar-benar relevan dalam membedakan suspek campak dan non-campak.

Oleh karena itu, pengenalan pola terhadap gejala campak diperlukan sebagai metode skrining awal (*rapid assessment*) yang cepat. Pendekatan komputasi ini dapat menjadi alternatif solusi ketika hasil laboratorium tertunda, sehingga keputusan isolasi dapat dilakukan lebih dini untuk mencegah penyebaran virus yang lebih luas.

Pola merupakan suatu set objek atau konsep yang dimana elemen yang ada mirip satu dengan yang lainnya (Tinani & Kandpal, 2019). Pengenalan pola dapat diautomatisasi menggunakan Algoritma pembelajaran mesin maupun pembelajaran mendalam. Pembelajaran mendalam menawarkan keuntungan signifikan dalam

prediksi penyakit, meningkatkan akurasi diagnostik dan efisiensi di berbagai kondisi medis (Tito dkk., 2024).

Pada penelitian yang dilakukan oleh (Gyebi dkk., 2023a) yang berjudul ”*Prediction of measles patients using machine learning classifiers: a comparative study*”, menunjukkan bahwa Algoritma pembelajaran mesin *random forest* menjadi Algoritma yang paling efektif dalam memprediksi penyakit campak dibandingkan dengan Algoritma pembelajaran mesin yang lainnya, dengan hasil akurasi 92,11%. Namun, penelitian tersebut masih melibatkan variabel demografi pasien dalam struktur modelnya. Hal ini membuka ruang penelitian lebih lanjut untuk memvalidasi apakah variabel demografi, seperti umur, benar-benar signifikan dalam meningkatkan akurasi diagnosis atau apakah pengenalan pola berbasis gejala klinis murni justru dapat memberikan hasil yang lebih efektif dan tidak dipengaruhi oleh bias variabel demografis. Oleh karena itu, penelitian ini mengkaji pengaruh pelibatan fitur demografi (umur) terhadap performa Algoritma dalam mengenali pola suspek campak. Selanjutnya pada penelitian yang dilakukan oleh (Franciska dkk., 2022a) yang berjudul ”Jaringan Syaraf Tiruan Untuk Memprediksi Penyakit Campak Menggunakan Algoritma *Backpropagation*” Algoritma *backpropagation* dapat digunakan dalam penggunaan prediksi menggunakan gejala dari campak menggunakan beberapa sampel rekam medis. Namun penelitian ini tidak menjelaskan mengenai data sampel yang diambil dan tidak adanya evaluasi mengenai tingkat akurasi dari Algoritma yang digunakan sehingga hanya menghasilkan 1 prediksi saja yaitu positif. Berdasarkan panduan dari Kementerian Kesehatan Republik Indonesia (Kemenkes RI) tahun 2023, gejala yang ada pada

campak adalah demam yang disertai salah satu atau lebih gejala batuk, pilek, mata merah, terdapat ruam pada tubuh yang dimulai dari belakang telinga, namun pada penelitian yang dilakukan terdapat variabel lain yang dimasukkan kedalam proses pelatihan model seperti rasa gatal, sakit tenggorokan, muka sembab, dan lain lain. Oleh karena itu, penting untuk mendapatkan dataset yang tepat dan model pembelajaran mesin yang tepat untuk dapat dievaluasi sehingga mendapatkan hasil yang optimal.

Pada penelitian yang dilakukan oleh (Begum dkk., 2024a) yang berjudul "*Medical Diagnosis Using Artificial Neural Networks*" menjelaskan bahwa penerapan dari berbagai jenis Algoritma ANN yaitu *multilayer neural network*, *probabilistic neural network*, *Learning Vector Quantization neural network*, *generalized regression neural network*, dan *radial basis function neural network* terhadap diagnosa penyakit autisme, kanker payudara, dan *haberman's survival data*. Data yang didapat dibagi menjadi 2 bagian yaitu dataset training dan dataset testing. Pelatihan dan pengujian dilakukan dengan teknik *5-fold cross-validation*. Hasil dari penelitian ini menunjukkan bahwa *Learning Vector Quantization neural network* mendapat keunggulan pada rekognisi penyakit autisme dan *haberman's survival data* dengan hasil 99,5% dan 99,4%. Hasil terendah pada penyakit autisme terdapat pada *radial basis function neural network*. Untuk penyakit kanker payudara dan *haberman's survival data*, hasil terendah diperoleh Algoritma *multilayer neural network* dengan hasil 83,3% dan 80,6%. Keberhasilan ini mengindikasikan ketangguhan LVQ dalam memisahkan pola penyakit yang kompleks. Selain itu, sebagai Algoritma berbasis prototipe dengan arsitektur yang lebih sederhana, LVQ

dinilai lebih efisien dan memiliki risiko *overfitting* yang lebih rendah untuk menangani dataset medis dengan jumlah sampel terbatas dibandingkan pendekatan *Deep Learning* yang kompleks. Namun, potensi Algoritma ini belum dieksplorasi secara mendalam pada kasus campak, khususnya dalam memproses pola gejala klinis yang sesuai dengan panduan nasional.

Berdasarkan seluruh uraian pekerjaan terkait yang telah dibahas tersebut, terdapat kesenjangan penelitian yang masih tersisa, yaitu penggunaan data demografi suspek campak yang dimasukkan kedalam penelitian sebelumnya sebagai salah satu variabel pengenalan pola tidak berdasarkan panduan dari Kemenkes RI, kurangnya evaluasi pada hasil model yang digunakan serta belum adanya penelitian menggunakan Algoritma *Learning Vector Quantization* menjadi faktor dilakukan penelitian ini. Oleh karena itu, penelitian mengenai pengenalan pola suspek campak menggunakan Algoritma *Learning Vector Quantization* ini menggunakan dataset dari gejala campak yang sesuai dengan panduan Kemenkes RI. Pada penelitian sebelumnya terdapat kekurangan pada dataset yang digunakan dan evaluasi Algoritma yang diterapkan, sehingga pada penelitian ini melibatkan pengukuran menggunakan *confusion matrix* sebagai evaluasi dari Algoritma yang dilatih serta diuji menggunakan dataset campak yang sesuai dengan panduan Kemenkes RI tahun 2023. Untuk memperoleh konfigurasi model yang paling optimal, penelitian ini menerapkan *grid search* yang mengkombinasikan berbagai nilai *learning rate*, jumlah *epochs*, dan jumlah *neuron*, sehingga proses pelatihan dapat mengeksplorasi ruang parameter secara sistematis dan menyeluruh. Penelitian ini bertujuan untuk mengevaluasi performa Algoritma LVQ,

menganalisis kelebihan serta kekurangan dari Algoritma LVQ, dan menghasilkan rekomendasi mengenai Algoritma LVQ pada pengenalan pola suspek campak berdasarkan gejala.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, dapat dirumuskan masalah pada penelitian ini, yaitu:

- a. Bagaimana penerapan algoritma *Learning Vector Quantization* (LVQ) yang dioptimasi menggunakan metode *Grid Search* untuk pengenalan pola suspek campak berdasarkan gejala klinis dan riwayat imunisasi?
- b. Bagaimana pengaruh fitur demografi (umur) terhadap kinerja model LVQ dalam mengenali pola suspek campak dibandingkan dengan model tanpa fitur demografi?

1.3 Tujuan Penelitian

Dari uraian masalah yang sudah dirumuskan, didapat tujuan penelitian sebagai berikut:

- a. Menerapkan algoritma *Learning Vector Quantization* (LVQ) dengan optimasi parameter melalui *Grid Search* untuk mengenali pola suspek campak.
- b. Mengevaluasi kinerja model LVQ dan menganalisis dampak penggunaan fitur demografi terhadap akurasi, *precision*, dan *recall* dalam deteksi suspek campak.

1.4 Manfaat Penelitian

Adapun beberapa manfaat yang didapatkan dari penelitian ini, sebagai berikut:

- a. Menambah literatur terkait pengenalan pola pada suspek campak berdasarkan gejala menggunakan Algoritma *learning vector quantization*.
- b. Memberikan kontribusi akademik pada pengembangan model pembelajaran mesin untuk pengenalan pola suspek campak.
- c. Memberikan rekomendasi Algoritma *Learning Vector Quantization* untuk pengenalan pola suspek campak berdasarkan gejala, sehingga dapat diterapkan pada departemen kesehatan yang ada.
- d. Membantu departemen kesehatan untuk memahami pola suspek campak berdasarkan gejala.

1.5 Batasan Masalah

Berikut merupakan batasan masalah yang terdapat pada penelitian ini:

- a. Data yang digunakan merupakan data kasus campak yang diambil dari Dinas Kesehatan Kota Tasikmalaya tahun 2024-2025 tahun berjalan.
- b. Penelitian ini hanya menggunakan Algoritma *Learning Vector Quantization*.
- c. Fokus penelitian terbatas pada evaluasi model meliputi *accuracy*, *precision*, *recall*, dan *f1-score* dari Algoritma *Learning Vector Quantization*.
- d. Implementasi Algoritma dilakukan menggunakan *python* dengan *library* tertentu yang dibutuhkan.

1.6 Sistematika Penulisan

Pada penelitian ini sistematika penulisan yang digunakan terdiri dari 5 BAB, sistematika penulisan yang digunakan meliputi:

BAB I Pendahuluan

Berisi tentang Latar Belakang, Rumusan Masalah, Tujuan Penelitian, Manfaat Penelitian, Batasan Masalah, dan Sistematika Penulisan.

BAB II Landasan Teori

Berisi Landasan teori yang memuat teori relevan terhadap penelitian, *State of the art* yang memuat penelitian sebelumnya serta hasil penelitian sebelumnya yang relevan terhadap penelitian ini

BAB III Metodologi Penelitian

Menjelaskan mengenai metode yang diterapkan pada penelitian ini, mencakup roadmap penelitian serta rincian prosedur setiap tahapan penelitian.

BAB IV Hasil dan Pembahasan

Memaparkan dan menganalisis hasil penelitian, analisis data, evaluasi model hingga mendapatkan hasil akhir dari penelitian yang dilakukan

BAB V Penutup

Berisi kesimpulan akhir dari penelitian yang dilakukan dan saran-saran yang direkomendasikan oleh peneliti berdasarkan pengalaman yang telah dilalui sebagai perbaikan untuk penelitian selanjutnya.