

BAB III

METODOLOGI

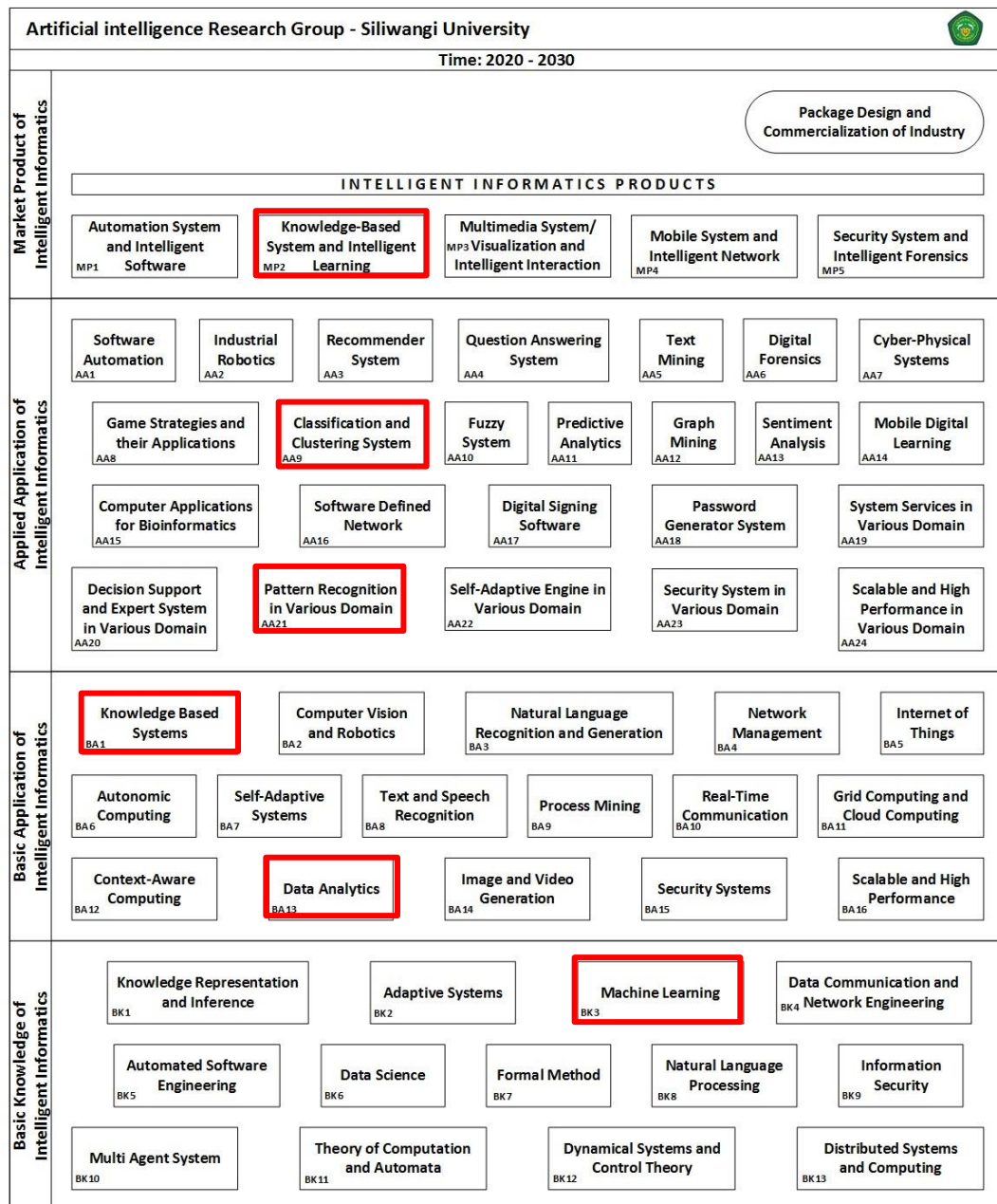
3.1 Metodologi Penelitian

Penelitian ini termasuk ke dalam penelitian dengan pendekatan kuantitatif. Pendekatan kuantitatif merupakan pendekatan yang digunakan untuk menguji teori-teori objektif dengan memeriksa hubungan antar variabel (Subhaktiyasa, 2024). Penelitian dilakukan dengan menggunakan sumber data yang ada untuk mencapai hasil akhir bersifat statistik.

Pada penelitian ini, objek yang digunakan untuk dikembangkan adalah suspek campak berdasarkan data campak Kota Tasikmalaya pada tahun 2024 dan 2025. Penelitian yang dilakukan adalah dengan melakukan proses *preparation* data dan pemodelan dengan Algoritma yang telah ditentukan. Pada tahap *preparation* data dilakukan normalisasi data dan penghapusan fitur yang tidak diperlukan. Setelah itu dilakukan pemodelan menggunakan *Algoritma Learning Vector Quantization* (LVQ) yang hasil dari pengujiannya berupa evaluasi *confusion matrix* untuk mengukur kemampuan dari Algoritma LVQ.

3.2 Peta Jalan (*Road Map*) Penelitian

Penelitian yang diajukan merupakan bagian dari peta jalan riset Kelompok Keahlian Informatika dan Sistem Inteligen (KK ISI) berkolaborasi dengan komunitas *AI Research Group Siliwangi University* yang dimana kajian sebelumnya dijadikan dasar pengetahuan untuk penyempurnaan bagi capaian akhir dari peta jalan penelitian secara keseluruhan.



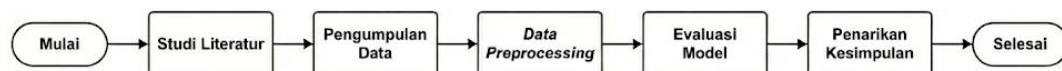
Gambar 3. 1 *Roadmap/Peta Jalan Penelitian* (AIS Universitas Siliwangi, 2020)

Gambar 3.1 menunjukkan beberapa kajian penelitian yang dilakukan dalam penelitian ini, yaitu merupakan bagian kajian yang ditandai di dalam garis persegi panjang berwarna merah yaitu terkait *machine learning*, *data analytics*, *knowledge based systems*, *classification and clustering system*, *pattern recognition in various*

domain untuk mencapai target berupa *knowledge-based system and intelligent learning* tentang pengenalan pola terhadap penyakit campak.

Machine learning yang digunakan dengan menggunakan Algoritma LVQ, *data analytics* sebagai fokus utama menganalisis data gejala untuk menemukan pola, *knowledge based systems* untuk memecahkan masalah berdasarkan pengetahuan yang ada, *classification and clustering system* untuk mengelompokkan data ke dalam kategori yang sudah ditentukan, *pattern recognition in various domain* sebagai tugas utama untuk secara otomatis mengenali pola yang ada dalam data, *knowledge-based system and intelligent learning* merupakan tujuan akhir yang menggabungkan semua elemen sebelumnya menjadi produk jadi.

Tahapan penelitian terdiri dari studi literatur, pengumpulan data, data *preprocessing*, pemodelan pengenalan pola dengan Algoritma LVQ, evaluasi dengan menggunakan *confussion matrix* dan kesimpulan. Tahapan dari penelitian ini dapat dilihat pada gambar 3.2.



Gambar 3. 2 Tahapan Penelitian

Tahapan penelitian yang ada pada Gambar 3.2 dijelaskan lebih lanjut pada sub bab 3.3 hingga 3.8.

3.3 Studi Literatur

Studi literatur merupakan tahapan eksplorasi dan analisis area penelitian untuk memperoleh pemahaman serta dasar teori dalam mendukung penelitian. Studi literatur dilakukan dengan mencari informasi dari buku, hasil penelitian dan artikel terkait topik utama. Indikator tahapan ini adalah memahami terhadap area

penelitian sehingga dapat menghasilkan BAB I Pendahuluan dan BAB II Tinjauan Pustaka.

3.4 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan dataset penyakit campak di Kota Tasikmalaya tahun 2024 dan 2025 dari Dinas Kesehatan Kota Tasikmalaya (DINKES Kota Tasikmalaya) yang merupakan rekap dari *form* penyelidikan epidemiologi (PE).

3.5 Data *Preprocessing*

Aktivitas yang dilakukan pada tahap ini adalah menghilangkan data yang tidak diperlukan dan transformasi data menjadi bentuk data numerikal. Sehingga data tersebut dapat digunakan dalam proses pengujian model. Fitur yang digunakan merupakan gejala demam, batuk, pilek, mata merah, dan imunisasi pertama serta kedua. Data tersebut dirubah menjadi skala 0 dan 1 sehingga dapat digunakan untuk perhitungan tahap pelatihan dan pengujian. Selain itu, normalisasi diterapkan pada fitur umur dengan pendekatan adaptif berdasarkan keberadaan *outlier*. Deteksi *outlier* dilakukan terlebih dahulu menggunakan metode *Interquartile Range* (IQR). Apabila data terindikasi memiliki *outlier*, normalisasi dilakukan menggunakan teknik *Robust Scaling* untuk menjaga integritas distribusi data. Sebaliknya, jika distribusi data cenderung normal tanpa outlier, teknik *Standard Scaler* diterapkan. Setelah tahapan normalisasi selesai, dataset dibagi menjadi dua subset utama, yaitu data latih (*training set*) dan data uji (*testing set*). Pembagian dilakukan dengan rasio 75:25, di mana 75% data digunakan untuk pelatihan model dan 25% sisanya dialokasikan untuk pengujian. Proses pembagian ini menerapkan teknik *Stratified*

Sampling untuk memastikan bahwa proporsi kelas suspek campak dan non-campak pada kedua subset tetap seimbang dan representatif terhadap populasi data asli.

3.6 Pemodelan

Pemodelan dilakukan menggunakan Algoritma yang diajukan. Pemodelan menggunakan data yang sudah dilakukan *preprocessing*. Data tersebut kemudian diuji dengan Algoritma LVQ menggunakan bahasa pemrograman *python* dengan bantuan aplikasi *google colab*, *microsoft excel*, serta *library python* seperti *pandas*, *numpy*, dan *scikit-learn*. Pelatihan model dilakukan menggunakan metode *Grid Search* dengan teknik *5-Fold Cross-Validation*. Dalam mekanisme ini, data latih dibagi menjadi lima subset (*fold*), di mana pada setiap iterasi, empat *fold* digunakan untuk pelatihan dan satu *fold* untuk validasi. Pendekatan ini bertujuan untuk mengeksplorasi skenario kombinasi parameter terbaik sekaligus menguji konsistensi performa model (rata-rata dan standar deviasi) agar tidak bias terhadap subset data tertentu, sebelum akhirnya diterapkan pada data uji.

3.7 Evaluasi Model

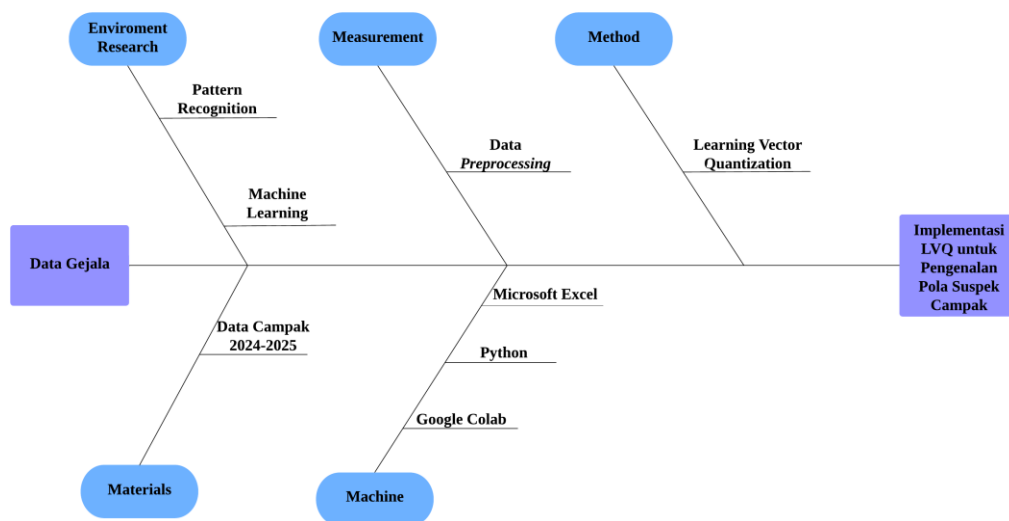
Setelah proses pemodelan menggunakan Algoritma LVQ, dilakukan tahapan evaluasi terhadap model yang telah dibuat. Evaluasi model terdiri dari 2 model yang sama sama menggunakan Algoritma LVQ yaitu dengan fitur demografi (umur) dan tanpa fitur demografi (tanpa umur), keduanya dibandingkan menggunakan *confusion matrix* untuk menilai kualitas objek. Evaluasi yang digunakan adalah evaluasi *accuracy*, *precision*, *recall*, dan *f1-score*.

3.8 Penarikan Kesimpulan

Tahapan terakhir yaitu melakukan penarikan kesimpulan berdasarkan hasil evaluasi dari Algoritma LVQ. Hasil tersebut dianalisis melalui temuan yang diperoleh sehingga mendapat luaran pada BAB V Kesimpulan dan Saran yang memuat ringkasan hasil penelitian serta rekomendasi untuk penelitian selanjutnya.

3.9 Fishbone Diagram

Tujuan dari penelitian ini juga direpresentasikan ke dalam diagram *fishbone* pada gambar 3.3. Diagram *fishbone* (sering juga disebut dengan *Ishikawa Diagram*, *Cause and Effect Diagram*) merupakan diagram yang menunjukkan penyebab dari sebuah masalah yang spesifik (Putu Widnyana dkk., 2022).



Gambar 3. 3 *Fishbone Diagram*

Gambar 3.3 menjelaskan bahwa kepala *fishbone* menggambarkan tujuan dari penelitian ini yakni implementasi Algoritma LVQ. Berikut merupakan keterangan dari gambar 3.3:

- a. *Environment Research*, merupakan area penelitian atau topik yang dijadikan fokus utama yaitu penelitian di bidang *machine learning* terutama pada bagian *pattern recognition*.
- b. *Material*, yaitu sumber atau data yang digunakan dalam penelitian. Data yang digunakan pada penelitian ini adalah data kasus campak Kota Tasikmalaya tahun 2024-2025.
- c. *Measurement*, merupakan tahapan untuk mempersiapkan data dengan proses *data preprocessing*.
- d. *Machine*, yaitu mesin atau alat bantu yang digunakan. Pada penelitian ini alat yang digunakan adalah *google colab* dan *microsoft excel*.
- e. *Method*, metode atau Algoritma yang digunakan dalam penelitian ini menggunakan Algoritma *Learning Vector Quantization*.