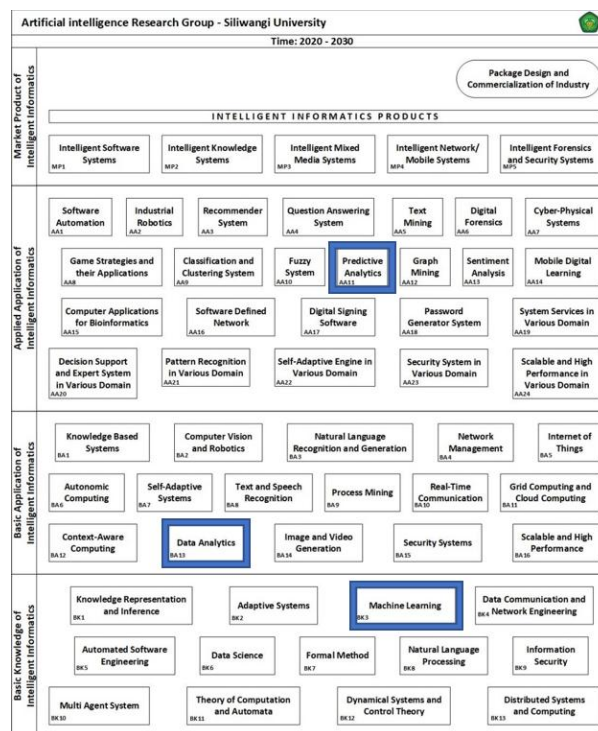


BAB III

METODOLOGI PENELITIAN

3.1 Roadmap Penelitian

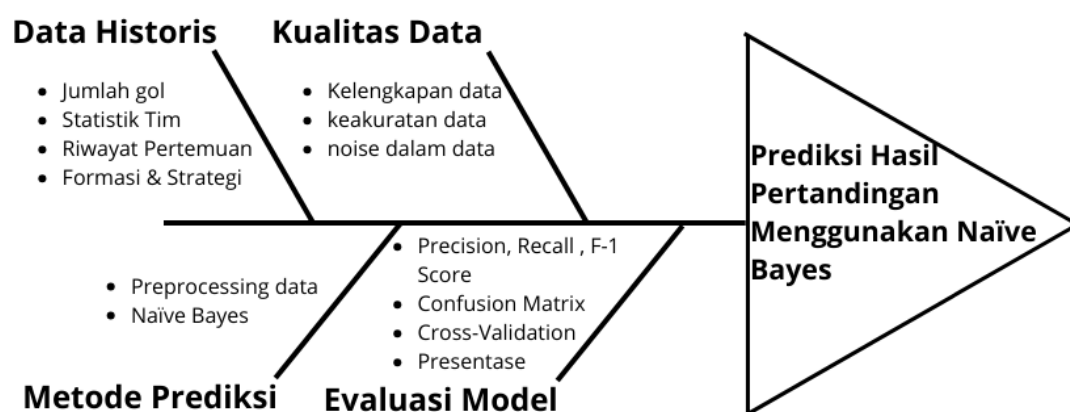
Penelitian ini berfokus pada analisis data dan prediksi hasil pertandingan sepak bola menggunakan algoritma *Naïve Bayes*. Metode ini dipilih karena kemampuannya dalam menghitung probabilitas berdasarkan data historis. Penelitian yang dilakukan sejalan dengan *Roadmap Artificial Intelligence Research Group*. Peta tersebut dapat dilihat pada gambar 3.1 menampilkan struktur penelitian yang ditandai dengan kotak biru.



Gambar 3.1 Roadmap Penelitian (AIS Universitas Siliwangi, 2020)

Berdasarkan Gambar 3.1 penelitian ini diawali dengan pengumpulan dan pengolahan data, yang bertujuan untuk menyusun informasi yang relevan sebelum digunakan dalam proses prediksi. Selanjutnya, algoritma *Naïve Bayes* diterapkan untuk memperkirakan peluang kemenangan, hasil imbang, atau kekalahan dalam bentuk persentase. Model yang dikembangkan kemudian dievaluasi untuk memastikan hasil prediksi yang akurat dan dapat diandalkan. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan yang lebih jelas mengenai faktor-faktor yang memengaruhi hasil pertandingan serta mendukung pengambilan keputusan dalam analisis sepak bola.

Penelitian ini berfokus pada prediksi hasil pertandingan menggunakan algoritma *Naïve Bayes*. Sebagai bagian dari penyusunan roadmap penelitian, diagram fishbone pada gambar 3.2 disajikan untuk membantu dalam mengidentifikasi dan mengorganisir faktor-faktor yang berpengaruh dalam penelitian ini gambar ini terinspirasi dari (Devi, Situmorang, dan Saragih 2025)

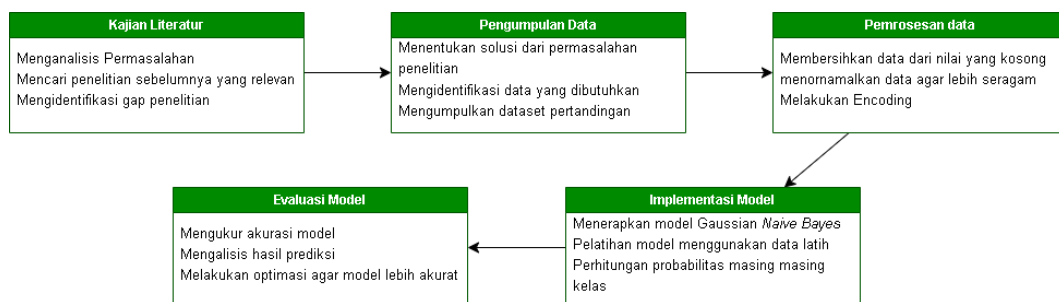


Gambar 3.2 Fishbone

Penelitian ini bertujuan memprediksi hasil pertandingan dengan menggunakan *Naïve Bayes*. Seperti terlihat pada gambar 3.2, data historis yang digunakan meliputi jumlah gol, statistik tim, riwayat pertemuan, formasi, dan strategi. Penting juga untuk memastikan data yang digunakan lengkap dan bersih dari gangguan. Sebelum menerapkan *Naïve Bayes*, data diproses terlebih dahulu agar model dapat bekerja dengan baik. Hasil prediksi dievaluasi melalui metrik seperti *precision*, *recall*, *F1-score*, *confusion matrix*, dan *cross-validation*, serta dilihat persentase peluang kemenangan, imbang, dan kekalahan.

3.2 Tahapan Penelitian

Dalam penelitian ini, penulis menerapkan serangkaian tahapan yang terstruktur untuk memastikan penelitian berjalan secara sistematis. Proses ini digambarkan dalam bagan alur pada Gambar 3.3, yang dilengkapi dengan penjelasan mengenai setiap tahap yang dilakukan.



Gambar 3. 3 Tahapan Penelitian

Berdasarkan gambar 3.3 tahapan penelitian terbagi menjadi 5 tahap sebagai berikut:

1. Tahapan awal penelitian ini adalah kajian literatur, yaitu melakukan analisis terhadap permasalahan yang akan diangkat melalui studi terhadap penelitian terdahulu yang relevan. Pada tahap ini dikaji berbagai pendekatan prediksi hasil pertandingan sepak bola yang telah dilakukan sebelumnya. Berdasarkan kajian tersebut, diperoleh gap penelitian berupa masih terbatasnya pendekatan yang menampilkan persentase probabilitas untuk setiap kemungkinan hasil pertandingan, yaitu menang, seri, dan kalah, karena sebagian besar penelitian sebelumnya hanya berfokus pada hasil klasifikasi akhir tanpa estimasi probabilitas.
2. Setelah kajian literatur, penelitian dilanjutkan ke tahap pengumpulan data. Data yang digunakan merupakan data historis pertandingan sepak bola yang mencakup informasi hasil pertandingan dan statistik pertandingan. Dataset dikumpulkan dari sumber terbuka dan dipilih dengan mempertimbangkan kelengkapan serta konsistensi data agar dapat digunakan secara optimal dalam proses pemodelan.
3. Pada tahap ini, data yang telah dikumpulkan diproses agar siap digunakan dalam pelatihan model. Proses yang dilakukan meliputi pembersihan data dari nilai yang kosong, normalisasi data untuk menyamakan skala antar fitur sehingga lebih seragam, serta proses encoding untuk mengubah nama tim menjadi bentuk numerik. Tahap ini bertujuan untuk memastikan data berada dalam format yang konsisten dan representatif sebelum diterapkan pada algoritma prediksi.

4. Setelah data siap, dilakukan implementasi model prediksi menggunakan algoritma Gaussian *Naïve Bayes*. Pada tahap ini, model dilatih menggunakan data latih untuk mempelajari pola statistik dari data historis pertandingan. Selanjutnya, model menghitung probabilitas masing-masing kelas hasil pertandingan, yaitu menang, imbang, dan kalah, berdasarkan fitur yang digunakan. Proses ini menghasilkan keluaran berupa nilai probabilitas untuk setiap kelas hasil pertandingan..
5. Tahap terakhir adalah evaluasi model untuk mengetahui performa prediksi yang dihasilkan. Evaluasi dilakukan dengan membandingkan hasil prediksi model dengan data aktual menggunakan metrik evaluasi seperti akurasi dan confusion matrix. Hasil evaluasi digunakan untuk menilai kemampuan model dalam memprediksi hasil pertandingan sepak bola.

Dengan mengikuti tahapan ini, penelitian ini diharapkan dapat menghasilkan model yang akurat dan bisa diandalkan dalam memprediksi hasil pertandingan sepak bola berdasarkan data historis.