

BAB III

METODE PENELITIAN

3.1. Metode *Machine Learning Life Cycle*

Metode *machine learning life cycle* adalah metode pengembangan arsitektur model jaringan syaraf tiruan yang digunakan pada penelitian ini.

3.1.1. Model Requirements

Pada tahap model *requirement* dilakukan Analisa mulai dari pengumpulan data yang akan di lakukan seperti:

- Observasi dilakukan mandiri dengan menganalisis kebutuhan data yang diperlukan untuk membuat model penjadwalan mata kuliah di PLPP IIB Darmajaya. Hasil observasi akan kebutuhan ini adalah sebagai berikut:
 1. Observasi untuk menentukan penggunaan dataset dilakukan selama lebih kurang 1 bulan, dilakukan dari (April – Mei 2023)
 2. Dataset yang akan digunakan untuk melatih arsitektur model jaringan syaraf tiruan adalah dengan menggunakan dataset yang diperoleh dari hasil pengumpulan data historis berupa jadwal mata kuliah IIB Darmajaya pada tahun ajaran 2021/2022 yang mencakup semester ganjil dan semester genap.
 3. Terdapat dua kelas variable output yang di jadikan sebagai predictor untuk melatih model yaitu, kelas regular dan kelas ekstensi.

3.1.2. Data Collection

Pada tahap data *collection* dilakukan pengumpulan data yang akan dijadikan sebagai dataset untuk membuat dan melatih model dengan arsitektur jaringan syaraf tiruan dimana data yang akan di kumpulkan berupa data historis yang berisikan penjadwalan matakuliah yang telah digunakan

sebelumnya atau sudah pernah dibuat oleh PLPP IIB Darnajaya. Data yang terkumpul kemudian akan diproses dan diolah untuk menghilangkan dupleks, mengurangi noise, dan menyesuaikan pola dari setiap atribut agar sesuai dengan kebutuhan pelatihan model.

3.1.3. Data Cleaning

Pada fase *Cleaning* data yang sudah dikumpulkan dilakukan cleaning data dengan menghapus data duplikat, menghapus data kosong/ null values, menghapus variable yang tidak dibutuhkan atau tidak relevan yang akan mempengaruhi kinerja pelatihan model serta menambahkan variable baru yang bertujuan untuk memudahkan proses pembuatan dan pelatihan model sehingga sesuai dengan tujuan awal penelitian ini dilakukan.

3.1.4. Data Labeling

Proses *Labeling* dilakukan dengan memberikan label pada setiap dataset. Sehingga pada tahap ini yang dilakukan adalah sesuai dengan tahap sebelumnya yaitu dengan menambah variable baru “Tipe Kelas” dimana dalam variable ini memiliki dua atribut yaitu “kelas regular dan kelas ekstensi” yang kemudian akan diubah menjadi label dalam bentuk angka atau numerik yaitu 0 dan 1.

3.1.5. Feature Engineering

Proses *Feature Engineering* dilakukan dengan memecah data menjadi data *training* dan *testing*. Untuk data *training* berjumlah 80% dari keseluruhan data sebagai data *training* dan 20% dari keseluruhan dari data *testing* serta menentukan data yang digunakan untuk validasi yang berbeda-beda untuk setiap kelasnya. Namun sebelum dilakukannya pembagian data *training* dan *testing* terlebih dahulu dilakukan standarisasi data dengan menerapkan teknik *MinMaxScaling* data dimana data secara keseluruhan akan diubah dalam bentuk numerik dengan batas ambang 0.1 samapi dengan 0.9. Setelah menerapkan standarisasi data kemudian diterapkannya teknik *SMOTE* untuk menangani kelas yang tidak seimbang atau yang sering disebut dengan *unbalanced dataset* dimana teknik *SMOTE* dapat menyeimbangkan kedua

kelas yang ada dalam dataset menjadi seimbang ang bertujuan untuk menghindari overfitting pada saat pelatihan model serta diharapkan dapat meningkatkan kualitas dataset

3.1.6. Model Training

Model training dilakukan untuk melatih model jaringan syaraf tiruan yang sudah dibuat. *Training* di lakukan agar model dapat bekerja dengan optimal dimana model ditraining berdasarkan arsitektur jaringan syaraf tiruan dengan menggunakan algoritma *backpropagation*.

3.1.7. Model Evaluation

Model evaluation pada tahapan ini akan mengkonfirmasi bahwa suatu model dapat memprediksi suatu data baru dengan cukup baik. Data yang akan dimasukkan ke dalam evaluasi ini adalah data testing di mana data testing memiliki 20% dari seluruh dataset. Setelah model selesai memprediksi data langkah selanjutnya adalah menggunakan dua jenis *evaluation matrix* untuk melihat performa dari model yang sudah di buat. Jenis matrix tersebut adalah *Mean Squarred Error (MSE)* dan *Root Mean Squarred. Error (RMSE)*.