

BAB III

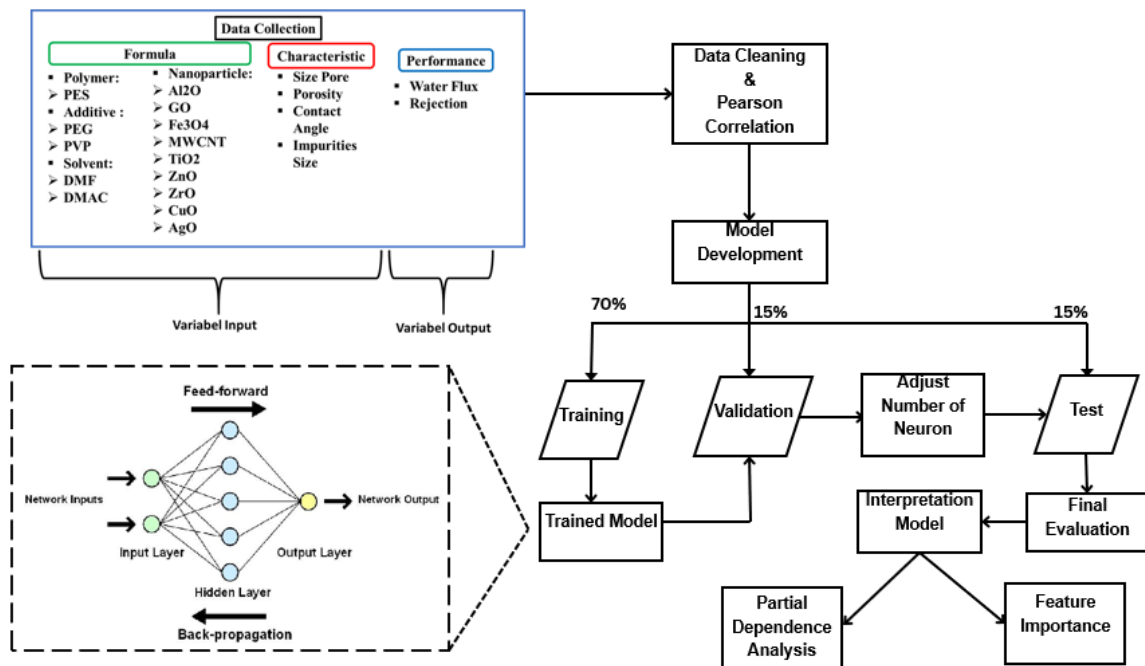
METODE PENELITIAN

3.1 Waktu dan Lokasi Penelitian

Semua tahap penelitian ini dilakukan di Program Studi Kimia, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Pendidikan Indonesia. Rentang waktu penelitian berlangsung mulai dari bulan Mei 2024 hingga Agustus 2024.

3.2 Tahap Penelitian

Pada tahap penelitian ini terdapat tiga tahap, yaitu pengumpulan data (*data collecting*), (*data preprocessing*), pengembangan model (*model development*), dan interpretasi model (*interpretation model*). Adapun bagan alir penelitian ditunjukkan pada gambar 3.1.



Gambar 3. 1. Bagan Alir Penelitian

3.3 Sumber Data Literatur

Pengumpulan data eksperimen diambil dari berbagai artikel ilmiah yang membahas topik riset terkait dari *website google scholar*: <https://scholar.google.com>. Spesifikasi data yang dikumpulkan merupakan parameter yang berpengaruh terhadap peningkatan sifat hidrofilitas membran komposit PES/aditif/nanopartikel. Data terdiri atas dua bagian data, yakni data input dan data output. Data input merupakan parameter yang berpengaruh terhadap performa kinerja membran, yakni parameter formulasi dan karakteristik membran. Sedangkan, data output merupakan parameter kinerja membran, yaitu fluks air dan rejeksi membran.

Database yang telah terkumpul berupa artikel ilmiah tentang membran. Lalu, Data ini akan dianalisis lebih lanjut untuk mendapatkan pemahaman mendalam tentang kinerja membran komposit PES/aditif/nanopartikel dalam berbagai formulasi dan karakteristik membran. Parameter formulasi membran komposit PES/aditif/nanopartikel mencakup komposisi PES serta variasi dan komposisi pelarut, polimer aditif, dan nanopartikel. Variasi pelarut yang dimasukkan adalah *Dimethylformamide* (DMF) dan *Dimethylacetamide* (DMAc). polimer aditif yang ditambahkan terdiri atas *polyvinylpyrrolidone* (PVP) dan *polyethyleneglycol* (PEG), sementara polimer utama yang dipakai adalah *polyethersulfone* (PES). Kemudian, nanopartikel yang ditambahkan meliputi Al_2O_3 , GO, Fe_3O_4 , MWCNT, TiO_2 , ZnO, CuO, dan AgO. Sedangkan, parameter karakteristik membran terdiri atas sudut kontak (*contact angle*), porositas (*porosity*), ukuran pori (*size pore*), dan ukuran molekul umpan (*feed molecule size*). Variabel output digunakan untuk melihat performa kinerja membran, yakni fluks air (*water flux*) dan rejeksi (*rejection*).

3.4 Pembersihan Data dan Analisis Korelasi Pearson

Sebelum dilakukan proses latihan, pembersihan data (*data cleaning*) akan diterapkan untuk menghilangkan duplikasi data dan menangani data yang hilang, sehingga memastikan bahwa data yang digunakan dalam penelitian adalah akurat dan dapat diandalkan. Selanjutnya, keterhubungan antara data variabel input dan output akan dianalisis

menggunakan koefisien korelasi pearson untuk mendeteksi kolinearitas antara variabel input dan output (Shrestha, 2020). Metode analisis pra proses ini menggunakan paket pemrograman Python 3.12 untuk memvisualisasikannya. Adapun persamaan matematika untuk analisis korelasi pearson adalah sebagai berikut (Profillidis & Botzoris, 2019).

$$r_{XY} = \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{\sqrt{n \cdot \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \cdot \sqrt{n \cdot \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}}$$

Berdasarkan persamaan 3.1, bahwasannya keterangan simbol dari r adalah koefisien korelasi pearson, n adalah jumlah pasangan data, serta x_i dan y_i adalah variabel yang ingin dianalisis. Variabel yang berkorelasi positif berarti hubungan antar data berbanding lurus dan variabel yang berkorelasi negatif berarti hubungan antar data berbanding terbalik.

3.5 Model Artificial Neural Network

Artificial neural network (ANN) digunakan sebagai algoritma untuk pengembangan model prediksi. Selama tahap proses pengembangan model, data dibagi menjadi tiga bagian yang terdiri atas training dengan persentase 70%, validation sebanyak 15%, dan test sebanyak 15%. Selama proses machine learning digunakan *software* Matlab R2023b. Pada dataset training dilakukan proses *learning* dengan menggunakan algoritma ANN. Kemudian, model hasil latihan tersebut diuji kembali dengan data yang baru, proses ini dinamakan validasi model. Pada tahap validasi model dilakukan optimasi jumlah neuron sebagai hyperparameter. Neuron berfungsi untuk memproses informasi dari pemasukan data input atau neuron sebelumnya, optimasi jumlah neuron bertujuan mengetahui agar jumlah neuron yang optimal bisa menghindari *underfitting* dan *overfitting*. Setelah, model dilakukan uji validitas, dilakukan evaluasi akhir menggunakan dataset tes yang berbeda dari dataset training dan validasi, hasil evaluasi akhir dari model prediksi ANN menggunakan parameter akurasi regresi dan kesalahan standar berupa *root mean squared error* (RMSE) (Yeo dkk., 2020). Adapun persamaan matematika untuk regresi adalah sebagai berikut.

$$R^2 = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (3.1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (3.2)$$

Berdasarkan persamaan 3.1 dan 3.2 yang merupakan persamaan regresi, keterangan simbol y_i adalah nilai output eksperimen, \bar{y} adalah nilai rata-rata sampel, dan \hat{y} adalah nilai output prediksi.

3.6 Interpretasi Model

Setelah pengembangan model, interpretasi dilakukan menggunakan analisis feature importance dan partial dependence analysis untuk memahami kontribusi setiap variabel input terhadap variabel output dan bagaimana variasi input memengaruhi performa membran. Analisis ini memberikan wawasan tentang faktor-faktor kunci yang menentukan kinerja membran komposit PES/polimer aditif/nanopartikel.

3.6.1 Analisis Feature Importance

Analisis feature importance adalah metode yang digunakan untuk menentukan seberapa besar kontribusi setiap variabel input terhadap hasil prediksi model. Dalam penelitian ini, analisis feature importance membantu mengidentifikasi variabel mana yang paling mempengaruhi performa membran, seperti fluks air dan rejeksi.

3.6.2 Analisis Partial Dependence

Analisis partial dependence adalah teknik yang digunakan untuk memahami bagaimana satu atau dua fitur input memengaruhi output model prediksi, sementara variabel lain tetap pada nilai rata-rata. Analisis ini sangat berguna untuk memvisualisasikan efek marjinal dari satu atau dua fitur (Yeo dkk., 2020). Adapun persamaan matematika untuk analisis partial dependence univariat adalah sebagai berikut:

$$\hat{f}(X_j) = \frac{1}{n} \sum_{i=1}^n f(X_j, X_{C(i)}) \quad (3.2)$$

Berdasarkan persamaan tersebut, bahwasanya simbol n adalah jumlah data, X_j adalah variabel yang dianalisis, $X_{C(i)}$ adalah variabel lainya yang tetap pada nilai rata-rata, $f(X_j, X_{C(i)})$ adalah model prediksi ketika X_j yang diubah dan variabel lainnya tetap. Sedangkan, persamaan matematika untuk analisis partial dependence bivariat adalah sebagai berikut:

$$\hat{f}(X_j, X_k) = \frac{1}{n} \sum_{i=1}^n f(X_j, X_k, X_{C(i)}) \quad (3.3)$$

Berdasarkan persamaan tersebut, bahwasanya simbol n adalah jumlah data, X_j dan X_k adalah variabel yang dianalisis, $X_{C(i)}$ adalah variabel lainya yang tetap pada nilai rata-rata, $f(X_j, X_{C(i)})$ adalah model prediksi ketika X_j dan X_k yang diubah dan variabel lainnya tetap.