

## INTISARI

# **IMPLEMENTASI *MACHINE LEARNING* UNTUK MEMPREDIKSI ENERGI ATOMISASI MOLEKUL DENGAN ALGORITMA *BOOSTED REGRESSION TREE* DAN *NEURAL NETWORK***

Oleh

Carlo Abimanyu

17/409389/PA/17696

Seiring dengan peningkatan daya komputasi serta tersedianya berbagai *database* terkait fenomena Fisika, membuat analisis berbasis *statistical learning* menjadi layak dilakukan untuk menyelesaikan permasalahan Fisika secara numerik. Pada penelitian ini, telah dibuat 2 model *machine learning* dengan algoritma *Extreme Gradient Boosted (XGB) Regression Tree* dan *Neural Network* untuk memprediksi energi atomisasi molekul. Fitur yang digunakan untuk *training* model dikonstruksi dari definisi Matriks Coulomb. Adapun dataset yang digunakan diperoleh dari *database* PubChem yang memuat 16,242 molekul dengan unsur penyusun C, H, N, O, P, dan S. Untuk model XGB diperoleh hasil  $MAE = 16,69$  kcal/mol dengan durasi *training* 86,7 detik. Sedangkan untuk model *Neural Network* diperoleh hasil  $MAE = 77,14$  kcal/mol dengan durasi pelatihan 195,0 detik. Sehingga pada penelitian ini model XGB dipilih sebagai model yang optimal. Pendekatan *machine learning* ini dapat menjadi salah satu alternatif untuk memperoleh penyelesaian persamaan Schrödinger pada sistem molekul.

Kata kunci: Matriks Coulomb, *machine learning*, energi atomisasi.

## ABSTRACT

# **MACHINE LEARNING IMPLEMENTATION TO PREDICT MOLECULAR ATOMIZATION ENERGY USING BOOSTED REGRESSION TREE AND NEURAL NETWORK ALGORITHM**

By

Carlo Abimanyu  
17/409389/PA/17696

Along with improvement in computing power and the availability of various databases related to physical phenomena, it makes statistical learning-based analysis feasible to solve physics problems numerically. In this research, 2 machine learning models have been made with Extreme Gradient Boosted (XGB) Regression Tree and Neural Network algorithm to predict the atomization energy of molecules. Features that used to train the models are constructed from the definition of Coulomb Matrix. The data set obtained from PubChem database which contained 16242 molecules made up of C, H, N, O, P, and S. For the XGB model, the results obtained  $MAE = 16.69$  kcal/mol with 86.7 seconds fitting duration. Meanwhile, for the Neural Network model, the results obtained  $MAE = 77.14$  kcal/mol with 195.0 seconds fitting duration. So that in this study the XGB model was chosen as the optimal model. This machine learning approach can be an alternative for solving the Schrödinger's equation in the molecular system.

Keywords: Coulomb Matrix, machine learning, atomization energy.