

**PENERAPAN METODE *K-NEAREST NEIGHBORS* DAN
RANDOM FOREST UNTUK DETEKSI KESALAHAN PENYETELAN
KONSTANTA PENGENDALI PADA KALANG KONTROL**

Oleh

Aulannisa Isthafi
17/413542/TK/45982

Diajukan kepada Departemen Teknik Nuklir dan Teknik Fisika Fakultas Teknik
Universitas Gadjah Mada pada tanggal 28 September 2022
untuk memenuhi sebagian persyaratan untuk memperoleh derajat
Sarjana Program Studi Teknik Fisika

INTISARI

Kesalahan dalam penyetelan konstanta pengendali (*tuning error*) dapat menimbulkan kerugian industri, serta kecelakaan kerja. Karenanya, penting untuk melakukan deteksi kesalahan penyetelan konstanta pengendali secara dini. Tujuannya ialah untuk dapat melakukan perbaikan dan pencegahan secepatnya. *Disruption* yang terjadi saat ini ikut mengubah paradigma di industri proses. Program deteksi yang biasanya *rule based* atau berbasis aturan tertentu, berubah menjadi program cerdas berbasis *machine learning* dan beroperasi secara *realtime*.

Penelitian ini dilakukan untuk membangun program deteksi kesalahan penyetelan konstanta pengendali berbasis *machine learning* dengan menggunakan algoritma *K-Nearest Neighbors* (KNN), dan *Random Forest* (RF) yang dapat diimplementasikan secara *realtime*. Pembangunan model dilakukan dengan menggunakan *dataset South African Council for Automatic Control* (SACAC), dan *dataset International Stiction Data Base* (ISDB). Kedua *dataset* tersebut memiliki kelas *tuning error* dan kelas *non tuning error* yang tidak seimbang. Untuk itu, dilakukan *undersampling* terhadap kedua kumpulan data tersebut. Selanjutnya data disegmentasi pada ukuran jendela data tertentu dan dilakukan ekstraksi serta seleksi fitur *time series*. Setelah persiapan data selesai, model dibangun dengan menggunakan algoritma KNN dan RF. Model yang telah terbentuk lalu dioptimasi untuk mendapatkan model dengan performa terbaik. Setelah dioptimasi, model diimplementasikan secara *realtime* dengan bantuan *Message Queue Telemetry Transport* (MQTT).

Hasil penelitian menunjukkan bahwa model KNN mempunyai skor f1 terbaik sebesar 88,89% ketika ukuran jendela data sebesar 200 dengan rumus jarak *Euclidean* pada jumlah tetangga dekat (k) ialah 5 atau 7. Sementara itu, model RF memiliki skor f1 terbaiknya yaitu sebesar 95,65% pada saat jendela data bernilai 150 dengan n trees sebesar 25, *maximum features* bernilai 18, *minimum samples split* sebesar 50, dan *maximum depth* bernilai 10.

Kata kunci: *tuning error*, kalang kontrol, *random forest*, *k-nearest neighbors*

Pembimbing Utama : Dr.-Ing. Awang N. I. Wardana, S.T., M.T., M.Sc.

Pembimbing Pendamping : Ir. Nopriadi, S.T., M.Sc., Ph.D.



IMPLEMENTATION OF K-NEAREST NEIGHBORS AND RANDOM FOREST FOR DETECTION OF CONTROL LOOP TUNING ERROR

by

Aulannisa Isthafi
17/413542/TK/45982

Submitted to the Departement of Nuclear Engineering and Engineering Physics
Faculty of Engineering Universitas Gadjah Mada on September 28, 2022
in partial fulfillment of the requirement for the Degree of
Bachelor of Engineering in Engineering Physics

ABSTRACT

One of the aspects that affect the performance of the controller is controller tuning. Tuning errors can cause industrial losses, as well as work accidents. Therefore, it is important to detect tuning errors early so that repairs and prevention acts can be carried out as soon as possible. The current disruption era has also changed the paradigm in the process industry. The old detection program is usually a rule-based program or based on certain rules, has turned into an intelligent program based on machine learning, and operates in real-time.

This research was conducted to build a machine learning-based tuning errors detection program that can be implemented in real-time. The program was created using K-Nearest Neighbors (KNN) and Random Forest (RF) algorithms. The model was created and developed using the South African Council for Automatic Control (SACAC) dataset, and International Stiction Data Base (ISDB) dataset. Class labels on those data were not balanced, so under sampling is carried out on the data. Then the balanced data is segmented on a certain data window size, and then a time series feature extraction is performed. The extracted data is then used to build the model. Each model that has been formed is implemented online using Message Queue Telemetry Transport (MQTT).

The results of this study using ISDB dataset showed that the KNN model had the best f1 score at 88.89% when the data window size was 200 with the Euclidean distance formula on the number of close neighbors (k) being 5 or 7. While the RF model had the best f1 score at 95.65% when the data window was 150 with n trees was 25, maximum features were 18, minimum samples split was 50, and the maximum depth was 10.

Keywords: tuning error, control loop, random forest, k-nearest neighbors

Supervisor : Dr.-Ing. Awang N. I. Wardana, S.T., M.T., M.Sc.

Co-supevisor : Ir. Nopriadi, S.T., M.Sc., Ph.D.

