

## INTISARI

# KALIBRASI NONPARAMETRIK BERBASIS PROSES GAUSSIAN UNTUK MENINGKATKAN KEANDALAN KLASIFIKASI MACHINE LEARNING

Oleh

Mahardi Nalendra Syafa

22/502515/PA/21558

Model *machine learning* seringkali menghasilkan estimasi probabilitas yang tidak terkalibrasi (*miscalibrated*), di mana skor kepercayaan menyimpang dari probabilitas kejadian yang sebenarnya, sehingga menurunkan kualitas pengambilan keputusan pada aplikasi yang sensitif terhadap ketidakpastian. Meskipun metode *post-hoc* konvensional seperti *Platt Scaling* dan *Isotonic Regression* merupakan standar yang umum digunakan, metode tersebut sering menghadapi keterbatasan terkait asumsi parametrik dan skalabilitas pada kasus multi-kelas. Penelitian ini mengevaluasi *Gaussian Process Calibration* (GP-Calib), sebuah kerangka kerja non-parametrik, untuk klasifikasi multi-kelas pada data terstruktur. Metodologi penelitian mencakup eksperimen pada lima *benchmark dataset* dan delapan *classifier*, dengan memanfaatkan *variational inference* dan *inducing variables* untuk memaksimalkan *Evidence Lower Bound* (ELBO) dengan tetap menjaga efisiensi komputasi. Hasil penelitian menunjukkan bahwa GP-Calib mencapai rata-rata nilai *Expected Calibration Error* (ECE) yang unggul sebesar 0,0412. Performa ini memberikan perbaikan sebesar 60,6% dibandingkan *baseline* tanpa kalibrasi dan 31,1% dibandingkan *Isotonic Regression*. Pemilihan kernel terbukti krusial, dengan kernel RBF menjadi konfigurasi optimal pada 25% kombinasi *dataset-classifier*. Secara keseluruhan, GP-Calib menyediakan alternatif yang tangguh dan andal untuk meningkatkan estimasi probabilitas dalam analisis data terstruktur multi-kelas.

**Kata Kunci:** *Gaussian Process Calibration*, Kalibrasi *Post-hoc*, Klasifikasi Multi-kelas, *Expected Calibration Error*, Kuantifikasi Ketidakpastian Model

## ABSTRACT

### **Gaussian Process–Based Nonparametric Calibration for Reliable Machine Learning Classification**

By

Mahardi Nalendra Syafa

22/502515/PA/21558

Machine learning models frequently produce miscalibrated probability estimates, where confidence scores deviate from actual likelihoods, thereby compromising decision-making in uncertainty-sensitive applications. While conventional post-hoc methods like Platt Scaling and Isotonic Regression are standard, they often face limitations regarding parametric assumptions and multi-class scalability. This study evaluates Gaussian Process Calibration (GP-Calib), a non-parametric framework, for multi-class classification on structured data. The methodology involves experiments across five benchmark datasets and eight classifiers, utilizing variational inference and inducing variables to maximize the Evidence Lower Bound (ELBO) while maintaining computational efficiency. Results demonstrate that GP-Calib achieves a superior average Expected Calibration Error (ECE) of 0.0412. This performance yields an improvement of 60.6% over uncalibrated baselines and 31.1% over Isotonic Regression. Kernel selection proved critical, with the RBF kernel emerging as the optimal configuration in 25% of dataset-classifier combinations. Overall, GP-Calib provides a robust and reliable alternative for enhancing probabilistic estimates in multi-class structured data analysis.

**Keywords:** Gaussian Process Calibration, Post-hoc Calibration, Multi-class Classification, Expected Calibration Error, Model Uncertainty Quantification