UNIVERSITAS
GADJAH MADA
Analisis Komparatif Multi-Task Learning dan Bayesian GAM untuk Prediksi Multi-Biomarker dengan Pengaruh Spot Size dan Kondisi Lingkungan
Luffy Ardiansyah, Dr. Adhitya Ronnie Effendie, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Analisis Komparatif *Multi-Task Learning* dan Bayesian GAM untuk Prediksi *Multi-Biomarker* dengan Pengaruh *Spot Size* dan Kondisi Lingkungan

Oleh

LUFFY ARDIANSYAH

24/533246/PPA/06765

Dried blood spot (DBS) merupakan metode pengambilan sampel darah yang praktis dan banyak digunakan dalam survei kesehatan berskala besar. Namun, akurasi pengukuran biomarker dipengaruhi oleh ukuran bercak darah (*spot size*) serta kondisi pra-analitik dan lingkungan yang hubungannya dengan kadar biomarker sering bersifat *nonlinier*. Penelitian ini bertujuan membandingkan *Multi-Task Learning* (MTL) dengan regularisasi *group sparsity* dan *Bayesian Generalized Additive Model* (Bayesian GAM) dalam memprediksi tujuh biomarker berdasarkan delapan variabel prediktor pra-analitik. Data sekunder berasal dari pengukuran DBS SHARE *Wave* 6 pada responden berusia 50 tahun ke atas. MTL diformulasikan dalam kerangka *Single-Input Multi-Output*, sedangkan Bayesian GAM dibangun secara terpisah untuk setiap biomarker. Untuk memastikan perbandingan yang adil, kedua model menggunakan transformasi logaritma dan ekspansi basis *B-spline* yang sama. Kinerja model dievaluasi menggunakan RMSE dan koefisien determinasi (R^2). Hasil penelitian menunjukkan bahwa kedua model memiliki kinerja yang tidak berbeda secara signifikan pada seluruh biomarker karena perbedaan RMSE dan R^2 lebih kecil daripada variasi antar-*seed* (± 2 simpangan baku), sehingga tidak ada model yang secara meyakinkan mengungguli model lainnya. Nilai R^2 seluruh biomarker berada di bawah 0,25, yang menunjukkan bahwa *spot size* dan kondisi lingkungan memiliki daya prediksi terbatas, dengan penggunaan desikan, jenis kelamin, dan usia sebagai faktor yang paling berpengaruh. MTL lebih unggul dalam interpretabilitas, sedangkan Bayesian GAM lebih baik dalam kuantifikasi ketidakpastian dan pemodelan hubungan *nonlinier*.

Kata kunci: *dried blood spot*, *multi-task learning*, Bayesian GAM, *group sparsity*, prediksi *multi-biomarker*, *spot size*

UNIVERSITAS
GADJAH MADA
Analisis Komparatif Multi-Task Learning dan Bayesian GAM untuk Prediksi Multi-Biomarker dengan Pengaruh Spot Size dan Kondisi Lingkungan
Luffy Ardiansyah, Dr. Adhitya Ronnie Effendie, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Comparative Analysis of Multi-Task Learning and Bayesian GAM for Multi-Biomarker Prediction Considering Spot Size and Environmental Effects

By

LUFFY ARDIANSYAH

24/533246/PPA/06765

Dried blood spot (DBS) sampling is a practical blood collection method widely used in large-scale health surveys. However, the accuracy of biomarker measurements is affected by blood spot size as well as pre-analytical and environmental conditions, whose relationships with biomarker levels are often non-linear. This study compared *Multi-Task Learning* (MTL) with group sparsity regularisation and a *Bayesian Generalized Additive Model* (Bayesian GAM) for predicting seven biomarkers based on eight pre-analytical predictor variables. Secondary data were obtained from DBS measurements in the SHARE Wave 6 survey of respondents aged 50 years and older. The MTL model was formulated within a *Single-Input Multi-Output* framework, whereas Bayesian GAM models were fitted separately for each biomarker. For a fair comparison, both models used the same logarithmic transformation and *B-spline* basis expansion. Model performance was evaluated using RMSE and the coefficient of determination (R^2). The results showed that the two models achieved statistically comparable performance across all seven biomarkers because the differences in test-set RMSE and R^2 were smaller than the variation across random seeds (± 2 standard deviations), indicating that neither model convincingly outperformed the other. The R^2 values for all biomarkers remained below 0.25, suggesting that spot size and environmental conditions had limited predictive power, with desiccant use, sex, and age being the most influential predictors. MTL provided greater interpretability, whereas Bayesian GAM offered better uncertainty quantification and modelling of non-linear relationships.

Keywords: dried blood spot, multi-task learning, Bayesian GAM, group sparsity, multi-biomarker prediction, spot size