

ABSTRAK

Nama : Neneng Nurhayasih

NIM : 1227010048

Judul : Model *Bayesian Cox Proportional Hazards* dan Klasifikasi *Naive Bayes* untuk Analisis Faktor Risiko dan Status Mortalitas Pasien Kanker Payudara

Analisis survival menggunakan model *Cox Proportional Hazards* (CPH) tidak memerlukan asumsi distribusi tertentu pada *baseline hazard*. Penelitian ini menggunakan pendekatan Bayesian pada model CPH untuk mengidentifikasi faktor yang berpengaruh terhadap waktu survival pasien kanker payudara serta memanfaatkan variabel signifikan dalam klasifikasi status mortalitas menggunakan *Naive Bayes*. Estimasi parameter dilakukan menggunakan algoritma *Metropolis-Hastings* melalui pendekatan *Markov Chain Monte Carlo* (MCMC). Data yang digunakan merupakan data sekunder pasien kanker payudara dari Kaggle dengan variabel usia, stadium tumor, histologi, status HER2, dan jenis operasi. Hasil penelitian menunjukkan bahwa stadium tumor, histologi, dan jenis operasi memiliki kategori yang berpengaruh signifikan terhadap waktu survival, sedangkan usia dan status HER2 tidak menunjukkan pengaruh signifikan. Klasifikasi *Naive Bayes* dilakukan dalam dua skenario, yaitu menggunakan variabel signifikan berdasarkan model *Bayesian Cox Proportional Hazards* dan seluruh variabel prediktor. Hasil perbandingan menunjukkan bahwa skenario menggunakan variabel signifikan memberikan kinerja klasifikasi yang lebih baik dibandingkan skenario menggunakan seluruh variabel prediktor.

Kata Kunci: Analisis survival, *Bayesian Cox Proportional Hazards*, *Metropolis-Hastings*, *Naive Bayes*, kanker payudara.

ABSTRACT

Name : Neneng Nurhayasih

NIM : 1227010048

Title : *Bayesian Cox Proportional Hazards Model and Naive Bayes Classification for Analyzing Risk Factors and Predicting Mortality Status in Breast Cancer Patients*

Survival analysis using the *Cox Proportional Hazards* (CPH) model does not require a specific distributional assumption for the *baseline hazard*. This study applies a Bayesian approach to the CPH model to identify factors affecting the survival time of breast cancer patients and utilizes the significant variables for mortality status classification using *Naive Bayes*. Parameter estimation is performed using the *Metropolis-Hastings* algorithm within the *Markov Chain Monte Carlo* (MCMC) approach. The study uses secondary breast cancer patient data obtained from Kaggle, with age, tumor stage, histology, HER2 status, and type of surgery as predictor variables. The results show that tumor stage, histology, and type of surgery have categories that significantly affect survival time, while age and HER2 status do not show significant effects. The *Naive Bayes* classification is conducted using two scenarios: significant variables identified by the *Bayesian Cox Proportional Hazards* model and all predictor variables. The comparison results show that the scenario using significant variables provides better classification performance than the scenario using all predictor variables.

Keywords: *Survival Analysis, Bayesian Cox Proportional Hazards, Metropolis-Hastings, Naive Bayes, Breast Cancer.*