

INTISARI

ESTIMASI MODEL CAMPURAN DUA DISTRIBUSI DENGAN ALGORITMA EKSPEKTASI MAKSIMISASI PADA DATA SURVIVAL

Oleh

Deny Budi Astuti

11/312932/PA/13610

Model campuran merupakan model gabungan dari dua atau lebih distribusi. Distribusi yang digunakan dalam model campuran bisa distribusi yang sama dengan parameter berbeda maupun distribusi yang berbeda. Model ini mengelompokkan data di dalam suatu dataset menjadi kelompok-kelompok data yang sebelumnya tidak terdefiniskan.

Dalam model campuran dua distribusi, distribusi eksponensial dan distribusi Weibull dapat membentuk tiga model campuran yang berbeda. Model yang terbentuk adalah model campuran eksponensial-eksponensial, model campuran eksponensial-weibull dan model campuran Weibull-Weibull. Dalam model campuran dua distribusi akan diberikan bobot atau berat pada dua distribusi yang digunakan. Jumlahan dari bobot harus sama dengan satu. Untuk mencari estimasi dari model campuran dua distribusi digunakan metode algoritma ekspektasi maksimisasi (algoritma EM) karena adanya data tidak lengkap (tersensor) dalam kasus data survival. Studi kasus menggunakan data pasien kanker paru-paru. Menggunakan pemrograman R, diperoleh kesimpulan bahwa tiga model campuran dua distribusi yang dibentuk sesuai digunakan untuk data pasien kanker paru-paru. Model campuran Weibull-Weibull merupakan model yang paling baik untuk data kanker paru-paru.

Kata kunci: analisis survival, model campuran dua distribusi, algoritma EM

ABSTRACT

ESTIMATED MIXTURE MODELS OF TWO DISTRIBUTION WITH EXPECTATION MAXIMIZATION ALGORITHM FOR SURVIVAL DATA

By:

Deny Budi Astuti

11/312932/PA/13610

Mixture model is a combination model of two or more distributions. Distribution in the model mixture can consist of distribution identic with different parameter or non-identic. This model classify data within dataset into groups of data that previously were not defined.

In the mixture two distribution models, exponential and Weibull distribution make three different models. This models are mixture models of exponential-exponential, exponential-Weibull, Weibull-Weibull. In the mixture two distribution models will be give weight of the two distributions. The sum of the weights should be one. To look the estimation of exponential Weibull mixture model used method algorithm EM because asumsy missing data. A case study on patient-lung-data by programming R, we concluded that the three models that form a mixture of two distributian coresponding to the dataset lung. Mixture model of Weibull-Weibull is the best model for dataset the lung.

Keywords : survival analysis, model mixture two distribution, algorithm EM.