

MENGIDENTIFIKASI POLA RAINFALL MENGGUNAKAN ANALISIS KOMPONEN PRINCIPAL ROBUST YANG DIPASANG DENGAN ANALISIS BICLUSTER

Oleh: Dr. Sri Andayani, S.Si., M.Kom, Kismiantini, S.Si, M.Si, Ph.D, Nikenasih Binatari, S.Si., M.Si

ABSTRAK

Data curah hujan merupakan nilai yang paling signifikan dalam pemodelan hidrologi dan klimatologi. Namun, kumpulan data rentan terhadap nilai yang hilang karena berbagai masalah. Penelitian ini bertujuan untuk menghitung nilai curah hujan yang hilang dengan menggunakan berbagai metode imputasi seperti Replacing by Mmean (RM), Nearest Neighbor (NN), Random Forest (RF), Non-linear Interactive Partial Least-Square (NIPALS) dan Markov Chain Monte Carlo (MCMC). Dataset curah hujan bulanan dari 24 stasiun curah hujan di Yogyakarta, Indonesia digunakan dalam studi ini. Dataset kemudian digunakan untuk bootstrap untuk mendapatkan perkiraan kesalahan standar dalam imputasi untuk setiap set data yang diperhitungkan. Performa kelima metode dievaluasi dengan metode root mean square (RMSE). Hasil eksperimen menunjukkan bahwa pendekatan RF-Bootstrap (RF-B) adalah yang paling cocok untuk data curah hujan yang hilang di Yogyakarta, Indonesia.

Kata kunci: MCMC, Missing value, nearest neighbor, NIPALS, random forest, replace by mean, bootstrap