

**COMPARATIVE ANALYSIS OF LATENT DIRICHLET ALLOCATION (LDA)
AND NON-NEGATIVE MATRIX FACTORIZATION (NMF) FOR TOPIC
MODELING**

FADEL THARIQ GIFARI
18/423103/PA/18186

The increasing pressure for controlling global warming and the speed at which new renewable energy technologies are being developed has made it necessary to efficiently organize and extract information from millions of renewable energy articles. This research investigates the use of topic modeling as a potent means of finding latent themes in renewable energy. The study seeks to compare Latent Dirichlet Allocation with Non-Negative Matrix Factorization by identifying latent themes and analyzing its respective computational time. The results of this experiment stated that the evaluation of LDA with NMF for topic modeling and discovered that LDA consistently performed better than NMF in coherence scores, with a total score difference of 3.23%. NMF showed a 19.0% gain in computational efficiency whereas LDA showed a 50% decrease in computing time. LDA is the recommended option for coherent topic modeling as it preserves greater coherence without incurring a large computational time.

Keywords: renewable energy, LDA, NMF, solutions, computational time, coherent

ANALISIS KOMPARATIF LATENT DIRICHLET ALLOCATION (LDA) DAN NON-NEGATIVE MATRIX FACTORIZATION (NMF) UNTUK PEMODELAN TOPIK

FADEL THARIQ GIFARI

18/423103/PA/18186

Tekanan yang meningkat untuk mengendalikan pemanasan global dan kecepatan pengembangan teknologi energi terbarukan baru telah membuat perlu untuk mengorganisir dan mengekstrak informasi secara efisien dari jutaan artikel energi terbarukan. Penelitian ini menyelidiki penggunaan pemodelan topik sebagai sarana yang ampuh untuk menemukan tema laten dalam energi terbarukan. Studi ini berupaya membandingkan Latent Dirichlet Allocation dengan Non-Negative Matrix Factorization dengan mengidentifikasi tema-tema laten dan menganalisis waktu komputasi masing-masing. Hasil eksperimen ini menyatakan bahwa evaluasi LDA dengan NMF untuk pemodelan topik menemukan bahwa LDA secara konsisten berkinerja lebih baik daripada NMF dalam skor koherensi, dengan total perbedaan skor 3,23%. NMF menunjukkan peningkatan 19,0% dalam efisiensi komputasi sedangkan LDA menunjukkan penurunan 50% dalam waktu komputasi. LDA adalah pilihan yang direkomendasikan untuk pemodelan topik yang koheren karena mempertahankan koherensi yang lebih besar tanpa menimbulkan waktu komputasi yang besar..

Kata kunci: energi terbarukan, LDA, NMF, solusi, waktu komputasi, koheren