

ABSTRAK

Nama : Dika Putra Maulana

NIM : 1227010010

Judul : **Pembobotan Adaptif Fungsi Objektif W2VPred Berbasis *Joint Maximum A Posteriori* pada *Word Embedding* Multi-Domain**

Model W2VPred mempelajari representasi kata multi-domain dan struktur kedekatan antar domain secara simultan melalui tiga komponen fungsi objektif: *Fidelity*, *Structure*, dan *Regularization*. Namun, penyeimbangan ketiganya bergantung pada koefisien statis yang dicari melalui *grid search* yang mahal secara komputasi. Penelitian ini mengusulkan W2VPred-JMAP, sebuah reformulasi probabilistik berbasis *Joint Maximum A Posteriori* (JMAP) yang memanfaatkan estimasi ketidakpastian tugas (*task uncertainty*) untuk menggantikan bobot statis. *Fidelity* dimodelkan sebagai *likelihood* Gaussian, sementara *Structure* dan *Regularization* diformulasikan sebagai *prior* Gibbs, di mana parameter ketidakpastian setiap distribusi dipelajari secara dinamis sebagai bobot adaptif terhadap besaran galat masing-masing komponen. Eksperimen pada korpus WikiFoS (16 disiplin ilmu) dan New York Times (20 tahun) menunjukkan bahwa W2VPred-JMAP menyamai akurasi analogi *baseline* terbaik (69,05% pada Top-10) dan mereduksi kebutuhan *grid search* sebanyak 25 kombinasi (λ, τ) menjadi satu parameter skalar (η_s). Model ini terbukti sama efektifnya dengan *baseline* dalam membedakan makna kata polisemik lintas domain tanpa mengorbankan representasi kata, serta meningkatkan prediksi struktur (Recall@3) dari 47,92% menjadi 50,00%.

Kata Kunci: *Word Embedding*, Multi-Domain, W2VPred, *Joint Maximum A Posteriori*, Pembobotan Adaptif.

ABSTRACT

Name : Dika Putra Maulana

NIM : 1227010010

Title : **Adaptive Weighting of W2VPred Objective Functions Based on
Joint Maximum A Posteriori in Multi-Domain Word Embedding**

The W2VPred model simultaneously learns multi-domain word representations and inter-domain affinity structures through three objective function components: *Fidelity*, *Structure*, and *Regularization*. However, balancing them relies on static coefficients searched through computationally expensive *grid search*. This study proposes W2VPred-JMAP, a probabilistic reformulation based on *Joint Maximum A Posteriori* (JMAP) that leverages task uncertainty estimation to replace static weights. *Fidelity* is modeled as a Gaussian *likelihood*, while *Structure* and *Regularization* are formulated as Gibbs *priors*, where the uncertainty parameter of each distribution is dynamically learned as an adaptive weight relative to the error magnitude of each component. Experiments on the WikiFoS (16 disciplines) and New York Times (20 years) corpora show that W2VPred-JMAP matches the best *baseline* analogy accuracy (69.05% at Top-10) and reduces the *grid search* requirement of 25 combinations (λ, τ) to a single scalar parameter (η_s). This model is proven to be as effective as the *baseline* in distinguishing the meaning of polysemic words across domains without sacrificing word representation, and improves structure prediction (Recall@3) from 47.92% to 50.00%.

Keywords: *Word Embedding*, Multi-Domain, W2VPred, *Joint Maximum A Posteriori*, Adaptive Weighting.