

Jordan University of Science and Technology

Feature Reduction Techniques for Arabic Text Categorization

Authors: Rehab Duwairi, Mohammad Al-Refai, Natheer Khasawneh

Abstract: The problem of creating a global schema over a set of heterogeneous databases is important due the availability of multiple databases within organizations. The global schema should provide a unified representation of local heterogeneous schemas. In his paper, we provide a general framework that supports the integration of local schemas into a global one. The framework takes into consideration the fact that local schemas are autonomous and may evolve over time, which makes the definition of the global schema obsolete. We define a set of integration operators that integrates local schemas based on the semantic relevance of their classes, and provide a model independent representation of virtual classes of the global schema. We also define a set of modifications that can be applied to local schemas as a consequence of their local autonomy. For every local modification, we define a propagation rule that will automatically disseminate the effects of that modification to the global schema without having to regenerate it from scratch via integration .