

خلاصه‌سازی چکیده‌ای نظرات بر پایه ساختار داده گراف

پروانه خسروی‌زاده

استادیار گروه زبان‌شناسی رایانشی

دانشگاه صنعتی شریف

khosravizadeh@sharif.ir

سارا بازرگان

دانشجوی کارشناسی ارشد زبان‌شناسی رایانشی

دانشگاه صنعتی شریف

bazargan_sara@mehr.sharif.ir

نشانی و تلفن:

چکیده

با رشد سریع تکنولوژی‌های تبادل اطلاعات، حجم زیادی داده در دنیای وب تولید و جمع‌آوری شده و دائماً حجم این داده‌های الکترونیکی نیز در حال افزایش است. یکی از این منابع داده، متون حاوی نظرات کاربران راجع به اخبار، محصول یا موضوعی خاص است. در این پژوهش یک روش خلاصه‌سازی چکیده‌ای بر اساس ساختار داده گراف برای فشرده‌سازی نظرات ارائه می‌گردد. در این پژوهش سعی بر این است که موضوعات مهم داخل متن با استفاده از افزونگی موجود در این‌گونه متون و ویژگی‌های گراف کشف شوند، جملات با هم ترکیب شوند، اطلاعات تکراری حذف شوند و در نهایت نیز جملات خلاصه، متناسب با اندازه خروجی مورد نیاز کاربر، استخراج شوند. این روش، بدون نظارت است و کمتر از اطلاعات وابسته به زبان استفاده می‌کند.

کلیدواژه‌ها: خلاصه‌سازی نظرات، چکیده‌ای، افزونگی، گراف