



استفاده از تکنیک‌های متن کاوی برای تحلیل متون فنی: مطالعه موردی، تحلیل محتوای آیین‌نامه بتن آمریکا (ACI ۳۱۸-۰۸)

فاطمه حافظی مقدس^۱، مجتبی مغربی^{۲*}

۱- دانشجوی کارشناسی ارشد، مهندسی و مدیریت ساخت، دانشگاه فردوسی مشهد، ایران

۲- دانشیار، گروه عمران گرایش مهندسی و مدیریت ساخت، دانشگاه فردوسی مشهد، ایران

fatemeh.h.m@gmail.com

خلاصه

برای کنترل کیفیت ساخت و ساز در هر کشوری تعدادی اسناد فنی الزام آور وجود دارد که مبنای طراحی، اجرا و نگهداری می‌باشد. این اسناد به مرور زمان توسعه یافته و دقیق تر شده‌اند، به طوری که در آیین‌نامه‌ها برای جزئیات یک ساختمان بندی مرتبط با آن جزء وجود دارد. اسناد فعلی در عین دقیق بودن گاهی پیچیده بوده و نیاز به تفسیر دارند. با توجه به ضرورت فرآیند آموزش به مهندسين، در این تحقیق سعی شده است تا با استفاده از تکنیک‌های داده کاوی به بررسی محتوای آیین‌نامه‌های اجباری پرداخته شود و ACI ۳۱۸-۰۸ به عنوان نمونه موردی انتخاب شده است. بدین منظور ابتدا متن سند فوق مورد بررسی اولیه قرار گرفته و واژه‌نامه‌ی کلمات پرکاربرد (کلیدواژه‌ها) ایجاد شد. برای یافتن کلیدواژه‌ها پس از انجام مراحل پیش‌پردازش، از روش وزن‌دهی TF-IDF استفاده شده است. به منظور ارائه اطلاعات متنی به صورت تصویری (ابرکلمات)، باید روابط بین کلیدواژه‌های شناسایی شده، تعیین شود که این امر با تحلیل شاخص‌های مرکزیت برای کلیدواژه‌ها انجام می‌شود. در نهایت تصویری ارائه شده که در آن ارتباط بین کلمات مهم در هر بخش قابل شناسایی است. نتیجه این تحقیق می‌تواند برای مهندسين جوان مفید باشد و در شناسایی موضوعات مهم حوزه ساخت و ساز به آن‌ها کمک کند.

کلمات کلیدی: متن کاوی، بصری‌سازی متن، استخراج کلمات کلیدی، ابرکلمات، آیین‌نامه بتن آمریکا