

عنوان مقاله:

معرفی روش های نوین گروه بندی فازی و جستجوی فازی و ارائه راهکارهای کاربردی در پالایش داده های هویتی بیمه شدگان سازمان تامین اجتماعی

محل انتشار:

دومین کنفرانس داده کاوی ایران (سال: 1387)

تعداد صفحات اصل مقاله: 8

نویسندگان:

جمال شهربابی
ونوس شکورنیا
آناهیتا علم الهدی

خلاصه مقاله:

در دنیای واقعی بطور معمول بنا به دلایلی (غلط املائی و قطعی ناگهانی ارتباط، از دست دادن یا نداشتن نشانه، فیلهایی با محتوای Null، اختصارات غیر معمول و یا فیلهایی که به هر دلیلی دچار اختلالاتی شده اند) داده ها ناخالص هستند؛ در نتیجه درصد زیادی از زمان و پول در پروژه های انبار داده و داده کاوی، صرف فازهای استخراج، تبدیل و بارگذاری داده (ETL) می گردد. داده هایی که به منظور ساخت پایگاه داده های تحلیلی و فرایند داده کاوی مورد استفاده قرار می گیرند مستلزم این است که با کیفیت باشند؛ تا با استفاده از روشها و الگوریتم های پالایش داده که معمولاً 60% از کل فرایند داده کاوی را تشکیل می دهند؛ پس از پیاده سازی و اجرای فرایند ETL داده های جدید بصورتی پالایش شده، اصلاح شده، استاندارد شده و سازگار با داده موجود بدست آیند. روشهای جستجو و گروه بندی فازی، روشهای اصلی و سودمندی هستند که تنوع وظایف پالایش و آماده سازی داده را که می بایست مکرراً جهت نگهداری داده در انبار داده ها، انجام شوند را تسهیل می سازند. جستجوی فازی تطابق رکوردهای ورودی را با رکوردهای صحیح و استاندارد شده میسر می سازد. فرایند تطابق در مقابل خطاهای محتمل موجود در رکوردهای ورودی بسیار انعطاف پذیر است. مکانیزم جستجوی فازی در صورت مشاهده خطا، نزدیکترین تطابق درست را برمیگرداند و حتی قادر است میزان کیفیت تطابق را نیز مشخص نماید. روش گروه بندی فازی بدین صورت عمل می کند که با استفاده از مقادیر فیلهای رشته ای، موارد مشابه بین سطرهای ورودی داده را کشف کرده و مشخص نماید کدام سطرها بصورت فازی تکراری می باشد. در این تحقیق با معرفی روش های نوین گروه بندی فازی و جستجوی فازی در راستای تحقق الزامات فراسازمانی و اهداف سازمان تامین اجتماعی در احراز هویت منحصر به فرد بیمه شدگان به ارائه راهکارهای کاربردی در امر پالایش داده ها در این خصوص پرداخته شده است تا بستری مناسب و قابل اطمینان به منظور ساخت سیستمهای متمرکز، پایگاههای تحلیلی (انبار داده) و اجرای فرایند داده کاوی فراهم گردد

کلمات کلیدی:

پالایش داده، داده کاوی، گروه بندی فازی، جستجوی فازی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/70553>

