

عنوان مقاله:

ارائه تکنیکی کارآ به منظور آماده سازی داده ها در فرایند داده کاوی

محل انتشار:

چهارمین کنفرانس بین المللی مدیریت فناوری اطلاعات و ارتباطات (سال: 1386)

تعداد صفحات اصل مقاله: 9

نویسنده:

مهدی اسماعیلی - عضو هیئت علمی دانشگاه آزاد اسلامی واحد کاشان

خلاصه مقاله:

مدل های داده ای پیشرفته و روش های بهینه ساز پرس و جو از پیشرفت های چشمگیر در پایگاه داده ها به شمار می رود. اما این تمام خواسته های یک کاربر نیست. فرایند کشف و استخراج دانش از این داده های ذخیره شده بخشی از سیستم های اطلاعاتی نسل جدید است. معرفی مسئله، پیش پردازش و آماده سازی داده ها، اجرای الگوریتم های داده کاوی و در نهایت ارائه الگوهای مناسب مراحل فرایند مذکور را تشکیل می دهد. داده های بدون نقص الگوهای مناسبی را برای کاربر تولید می کنند اما داده های دنیای واقعی ناقص و نادرست هستند. از آنجا که سیستم های استخراج دانش به داده هایی با کیفیت بالا و ترجیحا کم حجم نیاز دارند بنابراین، آماده سازی داده ها می تواند یکی از گامهای بنیادی در تحلیل داده ها باشد. جمع اوری و تغییر شکل داده ها، پالایش داده ها و همچنین کاهش اندازه داده ها از جمله تکنیک هایی هستند که در مرحله آماده سازی داده ها مطرح یم شوند. برای هر یک از این تکنیک ها الگوریتم های متعددی وجود دارد که هر یک از آنها دارای مزایا و معایب و همچنین کاربرد خاصی هستند. در این مقاله یک روش برای کاهش داده ها به منظور آماده سازی داده ها جهت داده کاوی ارائه می شود که مبتنی بر نمونه گیری از داده هاست. در این روش مزایای همچون دقت و سرعت قابل قبول با نمونه گیری مناسب بدست می آید.

کلمات کلیدی:

آماده سازی داده ها ، پیش پردازش داده ها ، داده کاوی ، کاهش داده ها

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/36906>

