

۲۰۲۰۲۴

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

داده کاوی

مرتضی کا زرونی

www.ketab.ir

سرشناسه : کازرونی، مرتضی، ۱۳۶۱-
عنوان و نام پدیدآور : داده کاوی / مرتضی کازرونی.
مشخصات نشر : تهران: انتشارات ابداع، ۱۳۹۸.
مشخصات ظاهری : ۱۳۶ ص.
شابک : ۹۷۸-۶۲۲-۶۷۵۰-۰۱-۱
وضعیت فهرست نویسی : فیپا
موضوع : داده کاوی
موضوع : Data mining
رده بندی کنگره : QA۷۶۷/۹ :د۲ ک۲ ۱۳۹۸
رده بندی دیویی : ۰۰۶/۳۱۲
شماره کتابشناسی ملی : ۵۶۵۱۹۴۱

عنوان: داده کاوی
ناشر: ابداع
نوبت چاپ: اول ۱۳۹۸
مؤلف: مرتضی کازرونی
تیراژ: ۱۰۰۰ نسخه
شابک: ۹۷۸-۶۲۲-۶۷۵۰-۰۱-۱
لیتوگرافی: مینیاتور
چاپ: فرشیوه
قیمت: ۲۰۰۰۰ تومان

آدرس: تهران-میدان انقلاب-خیابان کارگر جنوبی بین خ نظری و رواسمهر- بن بست
گشتاسب- پلاک ۷- واحد ۱ - تلفن: ۶۶۴۷۲۲۳۰

فهرست مطالب

2	چکیده
4	مقدمه
6	۱- فصل اول: مقدمه ای بر داده کاوی (کلیات و پیشینه و بیان اهداف)
6	۱-۱- داده کاوی چیست؟
15	۱-۲- توصیف داده ها در داده کاوی
18	۱-۳- مدل ها: پیش بینی داده ها
18	۱-۳-۱- Classification
18	۱-۳-۲- Regression
19	۱-۳-۳- Time series
20	۲- فصل دوم: مدل ها و الگوریتم های داده کاوی
20	۲-۱- تکنیکهای داده کاوی
20	۲-۱-۱- شبکه های عصبی
22	۲-۱-۲- Feed-Forward Backpropagation
23	۲-۱-۳- Decision Trees
26	۲-۱-۴- Multivariate Adaptive Regression Splines (MARS)
27	۲-۱-۵- Rule Induction
27	۲-۱-۶- K-Nearest neighbour and memory-based reasoning (MBR)
28	۲-۲- مدل های داده کاوی
28	۲-۲-۱- رگرسیون منطقی
29	۲-۲-۲- تحلیل تفکیکی
30	۲-۲-۳- مدل افزودنی کلی (GAM)
30	۲-۲-۴- Boosting
30	۲-۳- سلسله مراتب انتخابها
33	۳- فصل سوم: وب کاوی
37	۳-۱- مقدمه ای برای وب کاوی

38	۳-۲ محتوا کاوی وب
38	۳-۳ ساختار کاوی وب
40	۳-۳-۱ تکنیکهای تحلیل ابرپیوندها
42	۳-۳-۲ الگوریتم رتبه بندی صفحه وب
43	۳-۳-۳ الگوریتم HITS
47	۳-۴ کاوش کاربردی وب
60	۳-۵ سیستم ب کاوی
66	۵-۱ الگوریتم KIM
71	۵-۲ ارزیابی آنتوانا
75	۴- فصل چهارم: معماری سرویس‌ها و وب سریس کاوی
75	۴-۱ سرویس چیست؟
80	۴-۲ وب سرویس استاندارد برای داده کاوی
88	۴-۳ وب سرویس کاوی
99	۵- فصل پنجم: نتیجه‌گیری و پیشنهادات
99	۵-۱ نتیجه‌گیری
100	۵-۲ بیان مسئله و پیشنهادات
105	پیوستها
106	پیوست ۱: لغت نامه
114	پیوست ۲: Service Oriented Architecture
120	پیوست ۳: وب سرویس
127	پیوست ۴: XML



چکیده

واژه داده کاوی برای توصیف مجموعه فعالیتهای وسیعی مورد استفاده قرار می گیرد. دنیای مدرن امروز، دنیای اطلاعات است و ما با حجم بسیار وسیعی از داده ها سرکار داریم. برای رسیدن به اطلاعات باید این داده ها تحلیل و پردازش شوند تا اطلاعات از آنها استخراج شود.

توسعه سریع و افزایش قابلیت های ذخیره سازی منجر به ذخیره بسیار وسیع داده با هزینه ذخیره سازی پائین شده است. وجود داده های زیادی بر روی وب سایتها، پایگاه داده ها، مالتی مدیا، وب سرویسها و ... اخیراً استفاده از داده کاوی را از این منابع مد نظر قرار داده است. کاوش اطلاعات مفید که دانش مناسب را از پایگاه داده های وسیع استخراج می کند، به یکی از زمینه های مهم تحقیقاتی تبدیل شده است.

وب کاوی یکی از زمینه های مهم تحقیقی در زمینه داده کاوی برای تعداد وسیعی از سرویسهای وب جهان گستر در سالهای اخیر تبدیل شده است. وب کاوی گستر بصورت گسترده ای هم در زمینه ترافیک و هم در زمینه اندازه و پیچیدگی وب سایتها در حال گسترش است. این پیچیدگی، نیاز ما را به تحلیل استفاده وب سایت، چگونگی ساختار وب و محتوای منابع وب وادار ساخته است.

تکنولوژی وب سرویس باعث افزایش نقش وب سرویسها در برنزه های اینترنتی بطور کلی و در زمینه تجاری به طور خاص شده است. از زمانی که وب سرویسها نقش مهم و فزاینده ای در تکنولوژی اطلاعات پیدا کرده اند، از سیستمهای سرویسگرا انتظار رشد و پیچیدگی روزافزون می رود. مانند بسیاری سیستمهای برای ابزارهای که اجازه تحلیل و نظارت بر سیستمهای سرویسگرا استفاده شده را می دهند.

روشهای مختلفی برای داده کاوی و فرایند کاوی از وب سرویسها و تعاملات بین آنها برای تحلیل تعاملات وب سرویسها بین استفاده کننده گان و مهیا کننده گان و کشف ترکیب سرویسهاست وجود دارد. از زمانیکه استفاده از وب سرویسها افزایش پیدا کرده، تحلیل رفتار کاربرانی که از وب سرویسها استفاده می کنند نیز به