

مقدمه‌ای بر داده‌کاوی: مفاهیم، روش‌ها و کاربردها در تحلیل بازار برق

نویسندگان:

دکتر معصومه رستم نیاکان کلهری

دکتر مسعود حسنی مرزونی

انتشارات پژوهشگاه نیرو

۱۴۰۱

QA ۷۶/۹

رستم نیاکان کلهری، معصومه، ۱۳۶۳-

مقدمه‌ای بر داده‌کاوی: مفاهیم، روش‌ها و کاربرد در تحلیل بازار برق / نویسندگان: معصومه رستم نیاکان کلهری، مسعود حسنی مرزونی. ویراستار ادبی: نازنین محمدی‌نام. تهران: پژوهشگاه نیرو، انتشارات، ۱۴۰۱.

۱۹۰ ص.: مصور، جدول

ISBN: ۹۷۸-۶۲۲-۵۶۷۳-۰۲-۱ ؛ ریال ۱,۱۰۰,۰۰۰

فهرست‌نویسی بر اساس اطلاعات فیپا

کتابنامه: ص. ۱۶۷

الف. حسنی مرزونی، مسعود، ۱۳۵۷ - ۱. داده‌کاوی ۲. Data miningt

شماره کتابشناسی ملی: ۸۸۶۹۲۷۴ رده‌بندی دیویی: ۰۰۶/۳۱۲



انتشارات پژوهشگاه نیرو

چاپ اول: بهار ۱۴۰۱

کلیه حقوق قانونی این اثر متعلق به پژوهشگاه نیرو است.

نشانی: تهران، شهرک غرب، انتهای بلوار شهید دادمان، پژوهشگاه نیرو، کدپستی: ۱۴۶۸۶۱۳۱۱۳

تلفن: ۸۸۰۷۹۴۰۱-۹

نمبر: ۸۸۰۷۸۲۹۶

وبگاه: www.nri.ac.ir

پست الکترونیک انتشارات: publications@nri.ac.ir

تقدیم به خانواده‌های عزیزمان، به پاس مهر و عشق بی‌اندازه‌شان

و

به یاد مهندس شهریار رستم نیاکان کلهری

معصومه رستم نیاکان کلهری

مسعود حسنی مرزونی

پیشگفتار

در سازمان‌های پیچیده دولتی یا خصوصی، تصمیم‌ها پی‌درپی اتخاذ می‌شوند. درجه اهمیت، دوره زمانی و سطح اثرگذاری این تصمیم‌ها در سازمان‌ها متفاوت است. توانایی دانش‌کاران^۱ برای تصمیم‌گیری، از عوامل اثرگذار بر روی عملکرد و رقابت‌پذیری یک سازمان است. بسیاری از این افراد با استفاده از روش‌های ساده و مبتنی بر تجربه، دانش خود و اطلاعات در دسترس، تصمیم‌گیری می‌کنند. این رویکرد ایستا در تصمیم‌گیری با تغییرات سریع و شرایط پویای محیط اقتصادی سازگار نیست. فرایند تصمیم‌گیری در سازمان‌های امروزی، در واقع بسیار پیچیده، پویا و نیازمند رویکردهای تحلیلی و مدل‌های ریاضیاتی است. هوش تجاری^۲ فرایندی است که ابزارهای مناسب تصمیم‌گیری را در اختیار تصمیم‌گیران قرار می‌دهد. یکی از این ابزارها داده‌کاوی^۳ است. با بروز پیشرفت‌های سریع در حوزه جمع‌آوری داده‌ها^۴ و تکنولوژی‌های ذخیره آنها، سازمان‌ها قادر به گردآوری حجم عظیمی از داده‌ها شده‌اند. با این وجود، استخراج اطلاعات مفید از این داده‌ها به چالشی برای آنها تبدیل شده است. چرا که تنها داده‌ها برای تصمیم‌گیری کافی نیستند. داده‌ها باید تحلیل و به اطلاعات تبدیل شوند. هنگامی که این اطلاعات در تصمیم‌گیری در یک حوزه خاص مورد استفاده قرار می‌گیرند، می‌توان آنها را دانش^۵ نامید. اغلب، روش‌ها و ابزارهای سنتی تحلیل داده‌ها را نمی‌توان برای تحلیل داده‌های حجیم مورد استفاده قرار داد. گاهی اوقات ساختار و نوع غیرسنتی داده‌ها، مانعی برای استفاده از روش‌های سنتی برای تحلیل آنها هستند. در مواقعی نیز سؤالی که با تحلیل داده‌ها باید به آن پاسخ داده شود، با روش‌های سنتی قابل بررسی نیست. در نتیجه باید روش‌هایی نوین برای این موضوع توسعه داده شود. داده‌کاوی یک تکنولوژی است که روش‌های سنتی تحلیل داده‌ها را با الگوریتم‌های پیچیده برای پردازش حجم عظیم داده‌ها ترکیب می‌کند.

-
1. Knowledge workers.
 2. Business intelligence.
 3. Data mining.
 4. Data.
 5. Knowledge.

در طول سه دهه گذشته، شبکه‌های برق تجدیدساختاریافته^۱ با چالش‌های پیچیده اقتصادی و مهندسی مواجه بوده‌اند؛ اما در مجموع توانسته‌اند به موفقیت‌هایی در جهت تأمین برق با کمترین هزینه و بیشترین قابلیت اطمینان دست یابند. این کار ساده‌ای نیست. چرا که در هر ثانیه باید تعادل عرضه و تقاضا رعایت شود و هزاران محدودیت مربوط به منابع و شبکه برقرار شوند. همچنین بازار باید سیگنال‌های درست قیمت را به‌منظور تولید مؤثر و همچنین ایجاد انگیزه برای سرمایه‌گذاری در منابع ارسال کند.

در صنعت برق، تحلیل سیستماتیک بازار برق^۲ از جمله پیش‌بینی و شناسایی الگوهای موجود در تقاضا و قیمت و همچنین پیش‌رفتار بازیگران بازار برق، از اهمیت ویژه‌ای برای برنامه‌ریزی و تأمین پایدار این کالای استراتژیک و در عین حال شناسایی رفتارهای غیررقابتی بازیگران برخوردار است. تحلیل دقیق میزان تقاضا و قیمت، امکان برقراری تعادل عرضه و تقاضا را فراهم می‌سازد. به‌علاوه، تجربه پیاده‌سازی بازارهای برق در کشورهای مختلف نشان داده است که خروجی آنها در بسیاری از موارد یک بازار کاملاً رقابتی نبوده است. این نتایج نشانه‌هایی از وجود جریان‌هایی در طراحی بازار است که به بازیگران بازار مجالی برای رفتارهای استراتژیکی می‌دهد که منجر به ایجاد پدیده‌هایی مانند قدرت بازار^۳ و یا حتی ایجاد اختلال در بازار می‌شود. بنابراین رگولاتورها مکانیزم‌هایی مانند پیش‌بازار را به‌عنوان راهی برای حصول اطمینان از اینکه قیمت‌های بازار متأثر از رفتارهای غیررقابتی و سوء استفاده از بازار نباشد، ایجاد کردند. در این میان، داده‌کاوی با در اختیار گذاشتن ابزارهایی متنوع می‌تواند نقش مهمی در تحلیل این بازارها ایفا کند.

در کتاب حاضر، نخست به مفاهیم و روش‌های داده‌کاوی و سپس مثال‌هایی از کاربردهای آنها در تحلیل بازار برق پرداخته می‌شود. بدین ترتیب، فصل اول کتاب به مبانی داده‌کاوی اختصاص داده شده است. با توجه به دو کارکرد اصلی داده‌کاوی شامل پیش‌بینی و تحلیل‌های توصیفی، در فصول دوم و سوم به شرح این کارکردها و روش‌های مربوط به آنها پرداخته شده است. به‌منظور آشنایی بیشتر مخاطبان با مطالب فصول اول تا سوم، در فصل چهارم مثال‌هایی از کاربردهای روش‌های داده‌کاوی در تحلیل بازارهای برق شرح داده شده است.

-
1. Restructured electricity network.
 2. Electricity market.
 3. Market power.

فهرست مطالب

عنوان	شماره صفحه
فصل اول: مبانی داده کاوی	۱
مقدمه	۲
۱-۱- داده	۵
۲-۱- فرایند داده کاوی	۷
۱-۲-۱- مدل CRISP-DM	۷
۲-۲-۱- مدل SEMMA	۹
۳-۱- پیش پردازش داده ها	۱۲
۱-۳-۱- پاکسازی داده ها	۱۳
۱-۳-۱-۱- شناسایی و تحلیل مقادیر گم شده	۱۳
۲-۳-۱- تبدیل و گسسته سازی داده	۱۶
۳-۳-۱- انتخاب مشخصه	۱۸
۴-۱- کارکردهای داده کاوی	۲۳
جمع بندی	۲۵
فصل دوم: روش های داده کاوی برای پیش بینی	۲۶
مقدمه	۲۷
۱-۲- طبقه بندی داده ها	۲۷
۱-۱-۲- درخت تصمیم	۳۱
۱-۱-۱-۲- رویکرد تقسیم و غلبه برای ساخت درخت تصمیم	۳۳
۲-۱-۱-۲- انتخاب شاخصه برای ساخت درخت تصمیم	۳۶
۳-۱-۱-۲- روش جنگل تصادفی	۴۱
۴-۱-۱-۲- هرس کردن درخت تصمیم	۴۲
۲-۱-۲- بیز نایو و شبکه های باور بیزی	۴۴

۴۹	۲-۱-۳- طبقه‌بندی مبتنی بر قاعده.....
۵۲	۲-۱-۴- شبکه‌های عصبی مصنوعی.....
۵۴	۲-۱-۴-۱- شبکه عصبی پیشخور.....
۵۹	۲-۱-۴-۲- شبکه‌های عصبی بازگشتی.....
۶۱	۲-۱-۵- ماشین بردار پشتیبان.....
۶۵	۲-۱-۶- روش‌های مبتنی بر یادگیری از همسایه‌ها.....
۶۷	۲-۲- پیش‌بینی عددی.....
۶۷	۲-۲-۱- رگرسیون آماری.....
۶۹	۲-۲-۲- تحلیل سری‌های زمانی.....
۷۰	۲-۲-۳- سایر روش‌های پیش‌بینی عددی.....
۷۱	۲-۳- ارزیابی مدل پیش‌بینی.....
۷۳	جمع‌بندی.....
۷۴	۳. فصل سوم: داده‌کاوی توصیفی.....
۷۵	مقدمه.....
۷۵	۳-۱- استخراج الگوهای پرتکرار، وابستگی‌ها و همبستگی‌ها.....
۷۸	۳-۲- خوشه‌بندی داده‌ها.....
۸۰	۳-۲-۱- روش‌های بخش‌بندی.....
۸۲	۳-۲-۲- خوشه‌بندی سلسله‌مراتبی.....
۸۵	۳-۲-۳- خوشه‌بندی مبتنی بر چگالی.....
۸۶	۳-۲-۴- خوشه‌بندی شبکه‌ای.....
۸۶	۳-۲-۵- خوشه‌بندی مبتنی بر مدل.....
۸۹	۳-۳- شناسایی و تحلیل داده‌های پرت.....
۹۳	۳-۳-۱- رویکردهای آماری شناسایی داده‌های پرت.....
۹۴	۳-۳-۱-۱- روش‌های پارامتریک شناسایی داده‌های پرت.....
۹۶	۳-۳-۱-۲- روش‌های ناپارامتریک شناسایی داده‌های پرت.....
۹۷	۳-۳-۲- رویکردهای مبتنی بر مجاورت برای شناسایی داده‌های پرت.....
۹۷	۳-۳-۲-۱- روش‌های مبتنی بر فاصله.....

- ۹۸.....۳-۳-۲- روش‌های مبتنی بر چگالی.....
- ۹۹.....۳-۳-۳- رویکردهای مبتنی بر خوشه‌بندی برای شناسایی داده‌های پرت.....
- ۱۰۰.....۳-۴- یادگیری نیمه‌نظارتی.....
- ۱۰۲..... جمع‌بندی.....
- ۱۰۵.....۴. فصل چهارم: داده‌کاوی و تحلیل بازار برق.....
- ۱۰۶..... مقدمه.....
- ۱۰۶-۱-۴- روش‌های داده‌کاوی برای پیش‌بینی در بازار برق.....
- ۱۰۶-۱-۴- پیش‌بینی قیمت برق.....
- ۱۰۷-۱-۱-۴- رویکردهای آماری برای پیش‌بینی قیمت برق.....
- ۱۰۸-۱-۱-۴- شبکه‌های عصبی برای پیش‌بینی قیمت برق.....
- ۱۱۲-۱-۱-۳- ماشین بردار پشتیبان برای طبقه‌بندی قیمت برق.....
- ۱۱۳-۱-۱-۴- شبکه‌های عصبی، درخت تصمیم، شبکه‌های بیزی و KNN برای طبقه‌بندی قیمت برق.....
- ۱۱۴-۱-۴- پیش‌بینی جهش قیمت.....
- ۱۱۴-۱-۲- ماشین بردار پشتیبان برای پیش‌بینی جهش قیمت.....
- ۱۱۴-۱-۲-۲- شبکه‌های عصبی، درخت تصمیم و شبکه‌های بیزی، برای طبقه‌بندی جهش قیمت.....
- ۱۱۵-۱-۴- پیش‌بینی تقاضای برق.....
- ۱۱۸-۱-۳-۱-۴- رویکردهای آماری برای پیش‌بینی تقاضا.....
- ۱۲۱-۱-۳-۲- شبکه‌های عصبی برای پیش‌بینی تقاضا.....
- ۱۲۴-۱-۴- پیش‌بینی میزان تولید تکنولوژی‌های باد و خورشید.....
- ۱۲۵-۱-۴-۱- سیستم فازی-عصبی ANFIS برای پیش‌بینی تولید تکنولوژی بادی.....
- ۱۲۸-۱-۴-۲- شبکه باور عمیق، مبتنی بر خوشه‌بندی برای پیش‌بینی کوتاه‌مدت تولید تکنولوژی بادی.....
- ۱۲۸-۱-۴-۳- مقایسه روش‌های مختلف داده‌کاوی برای پیش‌بینی تولید تکنولوژی خورشیدی.....
- ۱۲۹-۱-۴-۵- پیش‌بینی رفتار بازیگران بازار برق.....
- ۱۲۹-۱-۵-۱- شبکه‌های عصبی فازی برای برآورد رقابت‌گریزی در بازار کالیفرنیا.....
- ۱۳۰-۱-۴-۵- استخراج قواعد اگر-آنگاه در سیستم قیمت‌دهی نیروگاه‌ها در ایران.....

- ۲-۴- روش‌های داده‌کاوی توصیفی برای تحلیل بازار برق ۱۳۶
- ۱-۲-۴- شناسایی الگوهای قیمت برق ۱۳۶
- ۱-۱-۲-۴- روش‌های شناسایی داده‌های پرت برای تحلیل قیمت‌ها ۱۳۶
- ۲-۱-۲-۴- رویکرد آماری برای تحلیل قیمت‌ها ۱۳۸
- ۳-۱-۲-۴- خوشه‌بندی فازی برای تحلیل قیمت‌ها ۱۴۱
- ۲-۲-۴- شناسایی الگوهای تقاضای برق ۱۴۱
- ۱-۲-۲-۴- تحلیل همبستگی سرعت باد و میزان تقاضا ۱۴۱
- ۲-۲-۲-۴- خوشه‌بندی برای تحلیل تقاضا ۱۴۲
- ۳-۲-۴- تحلیل شبکه برق ۱۴۳
- ۱-۳-۲-۴- خوشه‌بندی برای شناسایی مناطق قابل تحویل محلی ۱۴۳
- ۳-۴- روش‌های یادگیری ماشین در بهینه‌سازی ۱۴۷
- ۴-۴- سیستم‌های مبتنی بر عامل و نرم‌افزارهای تحلیلی ۱۵۲
- ۱-۴-۴- سیستم‌های مبتنی بر عامل ۱۵۲
- ۲-۴-۴- نرم‌افزارهای تحلیل داده بازار برق ۱۵۸
- ۱-۲-۴-۴- نرم‌افزار SAS برای تحلیل بازار کالیفرنیا ۱۵۸
- ۲-۲-۴-۴- نرم‌افزارهای تحلیل بازار استرالیا ۱۵۹
- ۳-۲-۴-۴- نرم‌افزارهای تحلیل بازار PJM ۱۶۱
- فهرست مراجع ۱۶۶