

خلاصه کتاب: داده‌کاوی کاربردی

نویسندگان: دکتر محمد صنیعی آبادی، سینا محمودی، محدثه طاهرپرور ناشر: انتشارات نیاز دانش

مقدمه

کتاب «داده‌کاوی کاربردی» که توسط دکتر محمد صنیعی آبادی و همکارانش نگاشته شده و انتشارات نیاز دانش آن را منتشر کرده است، پاسخی به نیاز روزافزون جامعه علمی و صنعتی به رویکردهای عملی در حوزه داده‌کاوی است. این اثر با هدف پر کردن خلأ موجود در منابع فارسی که به جنبه‌های کاربردی و عملی داده‌کاوی کمتر پرداخته‌اند، تدوین شده است. تمرکز اصلی کتاب بر ارائه مباحث نظری و سپس پیاده‌سازی عملی آن‌ها با استفاده از ابزارهای موجود است تا خواننده بتواند دانش کسب شده را به صورت ملموس در پروژه‌های واقعی به کار گیرد.

بخش اول: داده‌کاوی نظری

این بخش از کتاب به مبانی نظری و مفاهیم اساسی داده‌کاوی می‌پردازد و چهار فصل کلیدی را شامل می‌شود:

فصل 1: معرفی داده‌کاوی و کشف دانش

در این فصل، خواننده با تعریف داده‌کاوی، اهداف، کاربردها و مراحل مختلف فرآیند کشف دانش از داده‌ها (KDD) آشنا می‌شود. اهمیت داده‌کاوی در دنیای امروز و نقش آن در تصمیم‌گیری‌های هوشمندانه در حوزه‌های مختلف کسب‌وکار، پزشکی، مالی و غیره مورد بحث قرار می‌گیرد.

فصل 2: آماده‌سازی و پویش داده

این فصل به یکی از مهم‌ترین مراحل در هر پروژه داده‌کاوی، یعنی آماده‌سازی داده‌ها، اختصاص دارد. مباحثی نظیر پاک‌سازی داده‌ها (حذف نویز و داده‌های پرت)، مدیریت داده‌های گمشده، ادغام داده‌ها از منابع مختلف، کاهش ابعاد و تبدیل داده‌ها برای بهبود کیفیت و کارایی مدل‌ها، به تفصیل شرح داده می‌شوند. همچنین، تکنیک‌های پویش داده برای درک اولیه و اکتشافی ساختار داده‌ها معرفی می‌گردند.

فصل 3: یادگیری مدل (الگوریتم‌های یادگیری)

در این فصل، انواع الگوریتم‌های یادگیری ماشین که هسته اصلی داده‌کاوی را تشکیل می‌دهند، معرفی می‌شوند. الگوریتم‌های طبقه‌بندی (مانند درخت تصمیم، شبکه‌های عصبی، ماشین بردار پشتیبان)، خوشه‌بندی (مانند K-Means، خوشه‌بندی سلسله‌مراتبی) و قواعد انجمنی (مانند Apriori) به همراه اصول کاری و کاربردهایشان توضیح داده می‌شوند. این بخش به خواننده کمک می‌کند تا با ابزارهای اصلی تحلیل داده آشنا شود.

فصل 4: ارزیابی و تفسیر مدل

پس از ساخت مدل‌های داده‌کاوی، ارزیابی و تفسیر صحیح نتایج از اهمیت بالایی برخوردار است. این فصل به معیارهای ارزیابی مدل‌های طبقه‌بندی (مانند دقت، صحت، فراخوانی، F-Measure)، خوشه‌بندی و قواعد

انجمنی می‌پردازد. همچنین، روش‌های تفسیر مدل‌ها و استخراج بینش‌های قابل درک از آن‌ها برای اتخاذ تصمیمات موثر، مورد بررسی قرار می‌گیرد.

بخش دوم: داده‌کاوی عملی (نرم‌افزار RapidMiner)

این بخش از کتاب، جنبه کاربردی و عملی داده‌کاوی را با تمرکز بر نرم‌افزار RapidMiner پوشش می‌دهد. RapidMiner به عنوان یکی از محبوب‌ترین و قدرتمندترین ابزارهای داده‌کاوی، امکان پیاده‌سازی الگوریتم‌های مختلف را بدون نیاز به کدنویسی فراهم می‌کند. در این بخش، خواننده به صورت گام‌به‌گام با محیط RapidMiner، نحوه وارد کردن داده‌ها، انجام مراحل پیش‌پردازش، ساخت و ارزیابی مدل‌ها، و تفسیر نتایج آشنا می‌شود. این رویکرد عملی، به خواننده کمک می‌کند تا مهارت‌های لازم برای اجرای پروژه‌های داده‌کاوی در دنیای واقعی را کسب کند.

نتیجه‌گیری

کتاب «داده‌کاوی کاربردی» با تلفیق مباحث نظری و رویکرد عملی، منبعی ارزشمند برای دانشجویان، پژوهشگران و متخصصانی است که به دنبال درک عمیق و کاربرد عملی تکنیک‌های داده‌کاوی هستند. این کتاب با ارائه مثال‌های کاربردی و آموزش گام‌به‌گام نرم‌افزار RapidMiner، پلی میان تئوری و عمل ایجاد می‌کند و خواننده را برای حل مسائل پیچیده دنیای واقعی با استفاده از داده‌کاوی توانمند می‌سازد.