

مبانی داده کاوی و آموزش کار در کلمنتاین

مؤلف:

حامد آقاییارزاده
(مدرس دانشگاه)

آقای پتی



آقایارزاده، حامد، ۱۳۶۲ :	سرشناسه
مبانی داده کاوی و آموزش کار در کلمنتاین/مولف حامد آقایارزاده.	عنوان و نام پدیدآور
تهران: آفتاب گیتی، ۱۳۹۹.	مشخصات نشر
۱۴۳ص: مصور، جدول.	مشخصات ظاهری
۳۵۰۰۰ ریال ۹۷۸-۶۲۲-۹۷۶-۹۲۴۵۱۷۶ :	شابک
فیفا :	وضعیت فهرست نویسی
کتابنامه.	یادداشت
نرم افزار آی. بی. ام. اس. پی. اس. اس. مودلر :	موضوع
IBM SPSS Modeler (Computer software) :	موضوع
داده کاوی :	موضوع
Data mining :	موضوع
داده کاوی — نرم افزار :	موضوع
Data mining -- Software :	موضوع
۹۷۶ / ۹ :	رده بندی کنگره
۳۱۲ / ۰۰۶ :	رده بندی دیویی
۹۶۹۲ / ۰۰۱ :	شماره کتابشناسی ملی

مولف : حامد آقايارزاده

ناشر : آفتاب گیتی

صفحه آرایي و تنظيم: ابراهيم علائي

تیراڑ : ۱۰۰ جلد

نوبت چاپ : اول ۱۳۹۹

قیمت : ۳۵۰۰۰ تومان

شایک: ۹۷۸-۶۲۲-۲۴۵-۱۷۶-۹

تلفن های مرکزی پخش: ارسال پستی به سراسر کشور ۰۹۱۲۳۳۴۲۳۶۲ - ۰۲۱۶۶۹۶۹۸۳۷.

دفترپخش: تهران، میدان انقلاب، تقاطع ۱۲ فروردین ساختمان ولی عصر طبقه ۳ واحد ۹

این اثر مشمول قانون حمایت از مولفان و مصنفان مصوب ۱۳۴۸ است. هر کس تمام یا قسمتی از این اثر را بدون اجازه مولفین نشر یا پخش کند. مورد پیگرد قانونی قرار خواهد گرفت.
حق چاپ و نشر برای ناشر محفوظ است.

فهرست مطالب

عنوان

صفحه

۱-۱	تعریف داده کاوی	۱۱
۱-۲	مفهوم داده کاوی	۱۲
۱-۳	تاریخچه داده کاوی	۱۳
۱-۴	چه چیزی باعث پیدایش داده کاوی شده است؟	۱۵
۱-۵	جایگاه داده کاوی در میان علوم مختلف	۱۷
۱-۵	داده کاوی چه کاری انجام می دهد؟	۱۹
۱-۶	داده کاوی چه کارهایی نمیتواند انجام دهد؟	۱۹
۱-۷	داده کاوی و ابزارهای داده کاوی	۲۰
۱-۸	چالش های داده کاوی	۲۱
۱-۹	محدودیت های داده کاوی	۲۲
۱-۱۰	پیشرفت در تکنولوژیهای داده کاوی	۲۳
۱-۱۱	مفاهیم پایه در داده کاوی	۲۳
۱-۱۲	مفاهیم اساسی در داده کاوی	۲۴
۱-۱۳	مراحل فرایند کشف دانش از پایگاه داده ها	۲۴
۱-۱۳-۱	انبارش داده ها	۲۵
۱-۱۳-۲	انتخاب داده ها	۲۶
۱-۱۳-۳	تبدیل داده ها	۲۶
۱-۱۳-۴	کاوش در داده ها	۲۶
۱-۱۳-۵	تفسیر نتیجه	۲۶
۱-۱۴	گام های اصلی داده کاوی جهت کشف دانش عبارتند از:	۲۶
۱-۱۴-۱	تعریف مسأله	۲۷
۱-۱۴-۲	ساختن یک پایگاه داده برای داده کاوی	۲۷
۱-۱۴-۳	جستجوی داده	۲۹
۱-۱۴-۴	آماده سازی داده برای مدل سازی	۲۹
۱-۱۵	ساختن مدل داده کاوی	۲۹
۱-۱۶	تائید اعتبار ساده	۳۰

۳۰	۱-۱۷ تایید اعتبار مدل
۳۰	۱-۱۸ ایجاد معماری مدل و نتایج
۳۱	۱-۱۹ حلقه داده کاوی چیست؟
۳۱	۱-۲۰ شناسایی فرصت های تجاری
۳۳	۱-۲۱ عملیات های داده کاوی
۳۴	۱-۲۲ داده، اطلاعات و دانش
۳۵	۱-۲۲-۱ توصیف داده ها در داده کاوی
۳۵	۱-۲۲-۲ مقوله کاهش ابعاد داده
۳۶	۱-۲۳ روش های گردآوری اطلاعات
۳۶	۱-۲۴ سه راه مدیریت گردآوری اطلاعات
۳۶	۱-۲۴-۱ آماده سازی داده ها برای داده کاو
۳۷	۱-۲۵ اهمیت آماده سازی داده ها
۳۷	۱-۲۵-۱ کارهای محدود در آماده سازی یا پیش پردازش داده ها
۳۸	۱-۲۵-۲ پاک سازی داده ها
۳۸	۱-۲۵-۳ وظایف پاک سازی داده ها
۳۹	۱-۲۵-۴ حذف رکورد
۳۹	۱-۲۵-۵ حذف مشاهده
۴۰	۱-۲۵-۶ پر کردن به صورت دستی
۴۰	۱-۲۵-۷ پر کردن با مقدار ثابت سراسری
۴۰	۱-۲۵-۸ پر کردن با میانگین ویژگی
۴۰	۱-۲۵-۹ پر کردن با مقادیر با احتمال بیشتر
۴۱	۱-۲۵-۱۰ حل مشکل افزونگی (در عملیات تجمیع داده ها)
۴۱	۱-۲۵-۱۱ یکسان سازی فرمت ها
۴۲	۱-۲۵-۱۲ تصحیح داده های ناسازگار
۴۲	۱-۲۶ مواجهه با داده های نویز، داده های پرت و هموار کردن اغتشاشات داده ها
۴۳	۱-۲۶-۱ تلخیص توصیفی داده ها
۴۳	۱-۲۶-۲ گسسته سازی
۴۵	۱-۲۶-۳ حفاظت از حریم شخصی در سیستم های داده کاوی
۴۶	۱-۲۷ ابزارهای تجاری داده کاوی DM Commercial Tools
۴۷	۱-۲۸ آمار و داده کاوی
۴۸	۱-۲۸-۱ حجم داده ها

۴۹	۱-۲۸-۲ نوع داده ها
۴۹	۱-۲۸-۳ پردازش داده ها
۵۰	۱-۲۸-۴ وظایف معمول
۵۰	۱-۲۹ داده کاوی و OLAP
۵۱	۱-۳۰ دسته بندی
۵۲	۱-۳۰-۱ روش های دسته بندی
۵۲	۱-۳۰-۲ کلاسه بندی و پیشگویی داده ها
۵۴	۱-۳۰-۳ انواع روش های کلاسه بندی
۵۴	۱-۳۰-۴ مراحل یک الگوریتم کلاسه بندی
۵۵	۱-۳۰-۵ ارزیابی روش های کلاسه بندی
۵۵	۱-۳۰-۶ خوشه بندی
۵۹	۱-۳۰-۷ تعریف رده بندی پیش بینی
۶۲	۱-۳۰-۸ تعریف پیش بینی
۶۳	۱-۳۱ قواعد انجمنی
۶۵	۱-۳۲ نقاط دور افتاده
۶۷	۱-۳۳ الگوریتم های داده کاوی
۶۷	۱-۳۳-۱ درخت های انتخاب
۶۹	۱-۳۳-۲ انواع درخت های تصمیم
۶۹	۱-۳۳-۳ چگونه یک درخت تصمیم گیری رشد می کند؟
۶۹	۱-۳۳-۴ روش C4.5
۷۰	۱-۳۳-۵ روش C4.5
۷۱	۱-۳۳-۶ شبکه های عصبی
۷۲	۱-۳۳-۷ مزیت شبکه های عصبی
۷۲	۱-۳۳-۸ شباهت شبکه عصبی با مغز انسان
۷۵	۱-۳۳-۹ محاسبات تکاملی (الگوریتم ژنتیک)
۷۵	۱-۳۳-۱۰ الگوریتم ژنتیک
۷۷	۱-۳۳-۱۱ منطق فازی
۷۸	۱-۳۳-۱۲ تئوری بیز
۷۹	۱-۳۳-۱۳ الگوریتم K-Means
۷۹	۱-۳۳-۱۴ گام های الگوریتم k-means
۸۰	۱-۳۳-۱۵ ویژگی های الگوریتم k-means

۸۰	۱-۳۴ استنتاج قانون
۸۱	۱-۳۵ مدل فرآیند دو سویه
۸۱	۱-۳۶ مدل های پیش بینی داده ها
۸۱	۱-۳۷ یادگیری ماشین
۸۳	۱-۳۸ هدف
۸۴	۱-۳۹ ماشین یادگیرنده
۸۴	۱-۳۹-۱ روند تکاملی ماشین های یادگیرنده
۸۵	۱-۳۹-۲ انواع ماشین های یادگیرنده
۸۶	۱-۳۹-۳ اجزاء ماشین یادگیرنده
۸۷	۱-۳۹-۴ مشکلات ماشین های یادگیرنده
۸۹	۱-۳۹-۵ روش های یادگیری مدل در داده کاوی
۹۰	۱-۴۰ دسته بندی
۹۰	۱-۳۹-۷ گرس
۹۱	۱-۴۰ تشخیص انحراف
۹۱	۱-۴۱ خوشه بندی
۹۲	۱-۴۱-۱ کشف قوانین
۹۲	۱-۴۱-۲ کشف الگوهای ترتیبی
۹۲	۱-۴۲ نرم افزارهای داده کاوی
۹۴	فصل دوم آموزش کار با کلمنتاین
۹۴	۲-۱ مقدمه
۹۴	۲-۲ فرایند CRISP-DM
۹۶	۲-۳ درباره کلمنتاین (Clementine)
۹۶	۲-۳-۱ تعریف مسئله (Business Understanding)
۹۶	۲-۳-۲ تحلیل داده ها (Data Understanding)
۹۷	۲-۳-۳ آماده سازی داده ها (Data Preparation)
۹۷	۲-۳-۴ مدلسازی (Modeling)
۹۷	۲-۳-۵ ارزیابی (Evaluation)
۹۸	۲-۳-۶ توسعه (Deployment)
۹۸	۲-۴ انواع مدل ها در کلمنتاین (Types of Models)
۹۹	۲-۵ خوشه بندی (Clustering or Segmentation)
۱۰۰	۲-۶ الگوریتم های معروف خوشه بندی در کلمنتاین عبارتند از:

۲-۷	تحلیل همبستگی - وابستگی (Association Models)	۱۰۰
۲-۸	روش های آماری (statistical methods)	۱۰۱
۲-۹	شروع کار با کلمنتاین	۱۰۱
۲-۱۰	کار با کلمنتاین:	۱۱۰
۲-۱۱	برخی از داده های مورد استفاده در کلمنتاین	۱۱۶
۲-۱۱-۱	انواع داده ها:	۱۱۸
۲-۱۲	فیلتر کردن برخی از فیلدها:	۱۲۰
۲-۱۳	مرتب سازی	۱۲۶
۲-۱۴	جلوگیری از نمایش فیلدهای تکراری	۱۲۹
۲-۱۵	نمایش خروجی در جدول	۱۳۲
۲-۱۶	ادغام دو فایل اکسل	۱۳۳
۲-۱۷	نمایش خروجی برنانه ها در قالب نمودارها:	۱۳۷
	منابع	۱۴۳