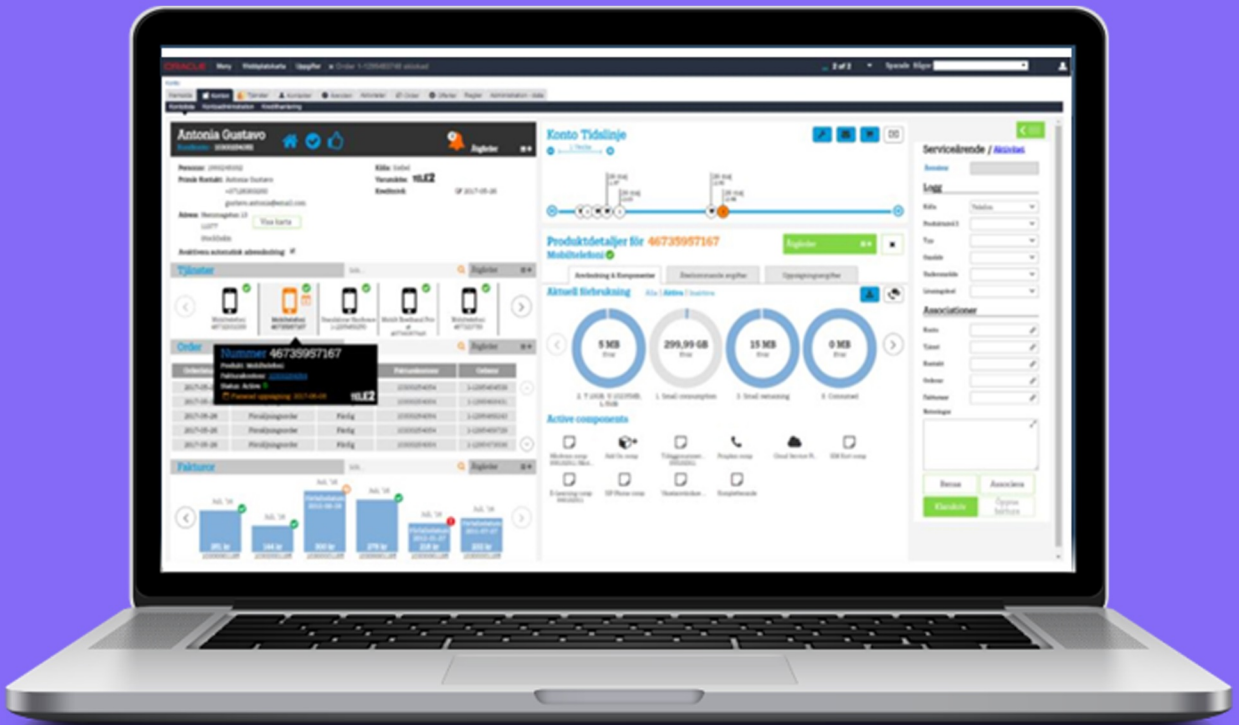


# داده کاوی

## در سیستم مدیریت آموزشی



مریم سادات خامسی نیا  
سیدتی مدیری بیدگلی



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



# داده کاوی

## در سیستم مدیریت آموزشی

تالیف:

مریم سادات خامسی نیا

کارشناسی ارشد مهندسی کامپیوتر / نرم افزار

سید تقی مدیری بیدگلی

کارشناسی ارشد مدیریت بازرگانی / استراتژیک



انتشارات موجک



سرشناسه : خامسی نیا، مریم سادات، ۱۳۵۹-

عنوان و نام پدید آور: داده کاوی در سیستم مدیریت آموزشی / تالیف مریم سادات خامسی نیا، سید تقی مدیری بیدگلی.

مشخصات نشر: تهران: انتشارات موجک، ۱۴۰۰.

مشخصات ظاهری: ۱۱۴ ص.

شابک: ۹۷۸-۶۰۰-۹۹۴-۲۷۶-۳

وضعیت فهرست نویسی: فیپا

موضوع: داده کاوی -- ایران -- نمونه پژوهی

Data mining -- Iran-- Case studies: موضوع

موضوع: آمار آموزشی -- داده پردازگی -- نمونه پژوهی

Educational statistics -- Data processing -- Case studies: موضوع

شناسه افزوده: مدیری بیدگلی، سید تقی، ۱۳۵۴-

رده بندی کنگره: QA ۷۶/۹

رده بندی دیویی: ۰۰۶/۳۱۲

شماره کتابشناسی ملی: ۷۵۷۲۶۹۵

انتشارات موجک

واتساپ: ۰۹۳۶۳۰۳۱۲۵۸ کانال: telegram.me/mojak1

تلفن مرکز پخش: ۰۲۶۳۲۷۰۵۳۱۸ - ۰۲۶۳۲۷۲۱۸۱۹ - ۰۲۱۶۶۱۲۷۵۹۳ - ۰۲۱۶۶۴۲۹۷۳۳

ایمیل: mojakpublication@yahoo.com

سایت: www.mojak.ir



عنوان: داده کاوی در سیستم مدیریت آموزشی

تالیف: مریم سادات خامسی نیا، سید تقی مدیری بیدگلی

مشخصات ظاهری: ۱۱۴ صفحه، قطع وزیری

چاپ اول: بهار ۱۴۰۰، تیراژ: ۵۰۰ جلد

قیمت: ۴۷۰۰۰۰ ریال، شابک: ۹۷۸-۶۰۰-۹۹۴-۲۷۶-۳

کلیه حقوق مادی و معنوی این اثر برای انتشارات موجک محفوظ است. هیچ شخص حقیقی و حقوقی حق

چاپ و تکثیر این اثر را به هر شکل و صورت اعم از فتوکی، چاپ کتاب و ... را ندارد. متخلفین به موجب

بند ۵ ماده قانون حمایت از ناشرین تحت پیگرد قانونی قرار می گیرند.

# فهرست مطالب

صفحه

عنوان

پیش گفتار..... ۱

**فصل اول: کلیات** ..... ۵

۱-۱ تشریح موضوع کتاب ..... ۵

۲-۱ اهمیت داده کاوی آموزشی ..... ۶

۳-۱ اهداف ..... ۷

۴-۱ تعریف واژه‌های کلیدی ..... ۸

**فصل دوم: داده کاوی آموزشی** ..... ۹

۱-۲ مفهوم داده کاوی آموزشی ..... ۹

۲-۲ طبقه‌بندی داده‌ها ..... ۱۲

۳-۲ اندازه‌گیری خطای روش تخمین ..... ۱۴

۴-۲ ارزیابی مدل ..... ۱۵

۱-۴-۲ تکنیک Holdout ..... ۱۶

۲-۴-۲ تکنیک Random Subsampling ..... ۱۶

۳-۴-۲ تکنیک k-fold cross-validation ..... ۱۷

۴-۴-۲ تکنیک Bootstrap ..... ۱۷

۵-۲ معیارهای دیگر جهت ارزشیابی روش‌ها ..... ۱۹

## فصل سوم: روش‌های طبقه‌بندی داده‌ها ..... ۲۱

- ۱-۳ استفاده از چند مدل ..... ۲۱
- ۲-۳ روش‌های طبقه‌بندی و تخمین ..... ۲۳
- ۱-۲-۳ درخت‌های تصمیم ..... ۲۳
- ۲-۲-۳ معیارهای انتخاب صفت خاصه ..... ۲۵
- ۳-۲-۳ چند الگوریتم درخت تصمیم ..... ۲۸
- ۳-۳ طبقه‌بندی با کمک قانون بیز ..... ۳۰
- ۴-۳ روش‌های طبقه‌بندی مبتنی بر یافتن شروط ..... ۳۳

## فصل چهارم: الگوریتم‌های داده کاوی ..... ۳۵

- ۱-۴ الگوریتم‌های یادگیری قوانین ..... ۳۵
- ۲-۴ معیارهای سنجش قوانین ..... ۳۵
- ۳-۴ الگوریتم‌های SVM ..... ۳۶
- ۴-۴ طبقه‌بندی بر اساس تشابه نزدیک ..... ۳۷
- ۵-۴ K نزدیکترین همسایه (KNN) ..... ۳۷
- ۶-۴ رگرسیون ..... ۳۷
- ۷-۴ شبکه عصبی ..... ۳۸
- ۱-۷-۴ ایده شبکه‌های عصبی مصنوعی ..... ۳۹
- ۲-۷-۴ مفاهیم پایه‌ای شبکه عصبی ..... ۳۹
- ۳-۷-۴ آموزش در پرسپترون تک لایه ..... ۴۱
- ۴-۷-۴ الگوریتم آموزش پرسپترون تک لایه ..... ۴۲
- ۵-۷-۴ نحوه اصلاح پارامترهای آزاد ..... ۴۲
- ۶-۷-۴ شبکه پرسپترون چند لایه ..... ۴۳
- ۷-۷-۴ آموزش شبکه MLP ..... ۴۴
- ۸-۷-۴ روش پس‌انتشار خطا ..... ۴۵



- ۴-۷-۹ افزودن واحد گشتاور به روابط اصلاح وزن ها ..... ۴۷
- ۴-۸ الگوریتم‌های ژنتیک ..... ۴۷
- ۴-۹ مجموعه‌های فازی ..... ۴۸

## فصل پنجم: تحقیقات داده کاوی آموزشی ..... ۵۱

### فصل ششم: مطالعه موردی ..... ۶۳

- ۶-۱ مقدمه ..... ۶۳
- ۶-۲ چارچوب نظری ..... ۶۳
- ۶-۳ مدل تحلیلی ..... ۶۵
- ۶-۴ روش پیشنهادی ..... ۶۶
- ۶-۵ مرحله تنظیم داده‌های ورودی ..... ۶۷
- ۶-۶ استخراج ویژگی‌های مهم ..... ۶۷
- ۶-۷ پیش‌بینی لینک با استفاده از شبکه عصبی ..... ۶۷
- ۶-۸ جامعه آماری ..... ۶۷
- ۶-۹ تنظیم داده‌های ورودی ..... ۶۹
- ۶-۹-۱ آماده‌سازی داده‌ها ..... ۶۹
- ۶-۹-۲ نرمال‌سازی داده‌ها ..... ۶۹
- ۶-۱۰ استخراج ویژگی‌های مهم با استفاده از الگوریتم ژنتیک چند هدفه نامغلوب ..... ۷۰
- ۶-۱۱ انتخاب مجموعه شاخص‌ها (جمعیت اولیه) ..... ۷۰
- ۶-۱۲ عملگر انتخاب و تابع بهینگی (برازش) ..... ۷۱
- ۶-۱۳ عملگر ترکیب ..... ۷۶
- ۶-۱۴ عملگر جهش ..... ۷۷
- ۶-۱۵ ارزیابی فرزندان ..... ۷۸
- ۶-۱۶ انتخاب نسل بعد ..... ۷۸

- ۶-۱۷ فاصله ازدحامی ..... ۸۳
- ۶-۱۸ ساختار شبکه عصبی در روش پیشنهادی ..... ۸۵
- ۶-۱۹ نتیجه گیری ..... ۸۶

## فصل هفتم: یافته‌ها ..... ۸۷

- ۷-۱ مقدمه ..... ۸۷
- ۷-۲ تکنیک اعتبارسنجی K-FOLD ..... ۸۷
- ۷-۳ محیط آزمایش ..... ۸۸
- ۷-۴ شاخص‌های ارزیابی ..... ۸۸
- ۷-۵ ماتریس در هم ریختگی ..... ۸۹
- ۷-۶ فراخوانی مجدد ..... ۹۱
- ۷-۷ ارزیابی و مقایسه نتایج ..... ۹۲
- ۷-۷-۱ آزمایش شماره یک ..... ۹۲
- ۷-۷-۲ آزمایش شماره دو ..... ۹۶
- ۷-۷-۳ آزمایش شماره سوم ..... ۱۰۰
- ۷-۷-۴ آزمایش نهایی ..... ۱۰۴

## فصل هشتم: جمع‌بندی ..... ۱۰۹

- ۸-۱ نتیجه گیری ..... ۱۰۹
- ۸-۲ پیشنهادها ..... ۱۱۰

## منابع ..... ۱۱۱

- منابع فارسی ..... ۱۱۱
- منابع انگلیسی ..... ۱۱۳

## پیش‌گفتار

امروزه به دلیل وجود ابزارهای مختلف برای جمع‌آوری داده‌ها و پیشرفت قابل قبول تکنولوژی پایگاه داده، حجم انبوهی از اطلاعات در انبار داده‌های مختلف ذخیره شده است. این رشد انفجاری داده‌ها، احتیاج به یک سری تکنیک‌ها و ابزارهای جدید که توانایی پردازش هوشمندانه اطلاعات را دارا باشند، نمایان می‌سازد. داده‌کاوی با پیدا کردن مجموعه‌ای از الگوهای جالب از دل داده‌های موجود در انبارها، می‌تواند چنین نیازی را مرتفع کند.

محققان در بسیاری از رشته‌ها نظیر پایگاه داده‌ها، یادگیری ماشین و آمار، این موضوع را پیگیری کرده و تکنیک‌ها و روش‌های مختلف ارائه شده در این زمینه را معرفی کرده که آنها را طبقه‌بندی می‌کند.

به دلیل رقابت در عرصه‌های سیاسی، نظامی، اقتصادی و علمی و اهمیت دستیابی به اطلاعات در کمترین زمان بدون دخالت انسان علم و تجزیه و تحلیل داده‌ها یا داده‌کاوی پا به عرصه گذاشت. داده‌کاوی<sup>۱</sup> فرآیندی است که در آغاز دهه ۹۰ مطرح شد و با نگرشی نو، به مسئله استخراج اطلاعات از پایگاه داده‌ها می‌پردازد.

آموزش رکن اصلی در توسعه پایدار هر کشور است و یکی از حوزه‌های تحقیقاتی اخیر در موضوع داده‌کاوی، داده‌کاوی آموزشی است. داده‌کاوی آموزشی به استفاده از تکنیک‌های داده‌کاوی در مجموعه داده‌های آموزش می‌پردازد. از طریق داده‌های آموزش و با استفاده از تکنیک‌های داده‌کاوی می‌توان به انواع استراتژی‌هایی که می‌بایست از آنها اجتناب نمود یا به آنها تاکید کرد، پی برد و به اتخاذ تصمیمات در خصوص بهبود شیوه‌های آموزش در موسسات آموزشی پرداخت.

از مهمترین سیاست‌های آموزشی کشور می‌توان به بازنگری و اصلاح ساختار برنامه‌ریزی و برنامه‌های آموزشی و روزآمد نمودن آنها در جهت پاسخگویی به نیازهای بومی و رهیافت‌های

---

<sup>۱</sup> data mining

جدید علمی، تقویت و توسعه دوره‌های تحصیلات تکمیلی و ارتقاء کیفیت علمی آن در جهت تأمین اعضای هیأت علمی مراکز آموزشی و پژوهشی، رفع نیازهای تخصصی بخش‌های مختلف و کاهش اعزام دانشجو به خارج از کشور، راه اندازی و تقویت دوره‌ها و مراکز جدید آموزشی در علوم پایه و پیشرفته از قبیل زمینه‌های خاص در علوم پایه، تکنولوژی اطلاعاتی و بیوتکنولوژی، ساماندهی و نهادینه نمودن سیاست آموزش استعدادهای خلاق در مراکز آموزش عالی و وضع مقررات و ضوابط ویژه آموزشی و پژوهشی، فراهم آوردن فرصت و افزایش دسترسی افراد با استعداد به مراکز دانشگاهی در سراسر کشور و حمایت از اقشار محروم جهت دستیابی به امکانات دانشگاهی و تکمیل تحصیلات عالی، تقویت ارتباط، همبستگی و تطابق میان نظام آموزش عالی، توسعه فناوری و بازارکار، گسترش و تقویت آموزش‌های علمی- کاربردی و فنی- حرفه‌ای با هماهنگی دستگاه‌های تولیدی و خدماتی جهت ارتقای کیفیت آموزشی اشاره کرد که در این راستا فعالیت‌هایی صورت گرفته از جمله گسترش کارائی درونی نظام آموزشی به واسطه شناسائی موانع و برطرف نمودن آن‌ها، تقویت عوامل تسهیل کننده آموزشی و بهره‌گیری از نتایج مربوطه در برنامه‌های آموزشی جهت افزایش میزان پیشرفت آموزشی و کاهش میزان تکرار مطالب درسی، گسترش کارائی بیرونی نظام آموزشی به واسطه تغییر و بررسی برنامه‌ها، عناوین و محتوای کتب درسی در جهت تناسب کمیت و کیفیت مواد آموزشی با نیازهای روزمره و استعدادها قابلیت‌های بالقوه دانش آموزان، رشد و گسترش خدمات مشاوره و هدایت تحصیلی متناسب با نیازهای خاص دانش آموزان از طریق گسترش و تقویت مراکز ویژه مشاوره تحصیلی، به کارگیری ابزار کمک آموزشی و راه بردهای جدید آموزشی در جهت حمایت از روش‌های ابتکاری و به کارگیری چنین روش‌هایی در مراکز آموزشی، اجرای تحقیقات کاربردی به منظور شناسائی عوامل مؤثر بر عدم یادگیری و عقب افتادگی تحصیلی دانش آموزان و تدارک طرح‌های آموزشی بر مبنای نتایج بدست آمده از تحقیقات مذکور و در نهایت طراحی سیستم مدیریت جامع آموزشی از طریق طراحی پایگاه داده‌های جامع.

با رویکرد فعالیت‌های انجام شده در این حیطه داده کاوی آموزشی به فهم بهتر فرایند یادگیری در موسسات آموزشی می‌پردازد. در نتیجه این امر منجر به بهبود کیفیت آموزش و کاهش هزینه می‌گردد. می‌توان با استفاده از تکنیک‌های داده کاوی به طراحی محتوای آموزشی و بهبود عملکرد

آموزشی دانش‌آموزان کمک می‌کند. به شناسایی حوزه‌های علاقه‌مندی دانش‌آموزان در امر آموزش کمک می‌کند. از این طریق می‌توان به اتخاذ تصمیمات مدیریتی و زیرساختی مناسب پرداخت.

هدف اصلی از تالیف کتاب حاضر روش پیشنهادی برای پردازش داده‌های آموزش با استفاده از الگوریتم ژنتیک چند هدفه با تابع برازش نزدیکترین همسایگی و بیز ارائه شده و ویژگی‌های مهم انتخاب شده می‌باشد. پس از انتخاب بهترین ویژگی‌ها، دسته‌بندی داده‌ها با استفاده از شبکه عصبی انجام می‌شود. مدل نهایی برای تشخیص تعداد مشروطی آینده دانشجوی مناسب خواهد بود.

در این کتاب جامعه آماری مورد استفاده تعداد ۱۲۰ نفر از دانشجویان دانشکده فنی دخترانه فدک کاشان است. داده‌های مورد استفاده ترکیبی از سوابق تحصیلی دانشجویان و اطلاعات شخصی آنها است. تعداد ویژگی‌های استخراج و دریافت شده ۲۰ ویژگی است. برچسب یا کلاس، تعداد دفعات مشروطی دانشجوی است. در روش پیشنهادی با توجه به نتایج حاصل شده، میزان صحت روش پیشنهادی ۹۵.۸۳ درصد، میزان دقت ۹۴.۱۶۸ و میزان فراخوانی مجدد ۹۵.۸۳ درصد است.

شیوه سازماندهی کتاب حاضر به صورت زیر می‌باشد.

فصل اول: کلیات

فصل دوم: داده کاوی آموزشی

فصل سوم: روش‌های طبقه‌بندی داده‌ها

فصل چهارم: الگوریتم‌های داده کاوی

فصل پنجم: تحقیقات داده کاوی آموزشی

فصل ششم: مطالعه موردی

فصل هفتم: یافته‌ها

فصل هشتم: جمع‌بندی

در پایان، بر خود لازم می‌دانیم از همه عزیزانی که در مراحل گوناگون آماده‌سازی این کتاب ما را یاری نموده‌اند، صمیمانه تقدیر و تشکر نمائیم.

مریم سادات خامسی‌نیا، سیدتقی مدیری بیدگلی

بهار ۱۴۰۰

# Data Mining in Educational Management System

Maryam Sadat Khamesinia

Seyed Taqi Modiri Bidgoli

آموزش رکن اصلی در توسعه پایدار هر کشور است و یکی از حوزه‌های تحقیقاتی اخیر در موضوع داده کاوی، داده کاوی آموزشی است. داده کاوی آموزشی به استفاده از تکنیک‌های داده کاوی در مجموعه داده‌های آموزش می‌پردازد. از طریق داده‌های آموزش و با استفاده از تکنیک‌های داده کاوی می‌توان به انواع استراتژی‌هایی که می‌بایست از آن‌ها اجتناب نمود یا به آن‌ها تاکید کرد، پی برد و به اتخاذ تصمیمات در خصوص بهبود شیوه‌های آموزش در موسسات آموزشی پرداخت.

از مهمترین سیاست‌های آموزشی کشور می‌توان به بازنگری و اصلاح ساختار برنامه‌ریزی و برنامه‌های آموزشی و روزآمد نمودن آنها در جهت پاسخگویی به نیازهای بومی و رهیافت‌های جدید علمی، تقویت و توسعه دوره‌های تحصیلات تکمیلی و ارتقاء کیفیت علمی آن در جهت تأمین اعضای هیأت علمی مراکز آموزشی و پژوهشی، رفع نیازهای تخصصی بخش‌های مختلف و کاهش اعزام دانشجو به خارج از کشور، راه اندازی و تقویت دوره‌ها و مراکز جدید آموزشی در علوم پایه و پیشرفته از قبیل زمینه‌های خاص در علوم پایه، تکنولوژی اطلاعاتی و بیوتکنولوژی، ساماندهی و نهادینه نمودن سیاست آموزش استعدادهای خلاق در مراکز آموزش عالی و وضع مقررات و ضوابط ویژه آموزشی و پژوهشی، فراهم آوردن فرصت و افزایش دسترسی افراد با استعداد به مراکز دانشگاهی در سراسر کشور و حمایت از اقشار محروم جهت دستیابی به امکانات دانشگاهی و تکمیل تحصیلات عالی، تقویت ارتباط، همبستگی و تطابق میان نظام آموزش عالی، توسعه فناوری و بازارکار، گسترش و تقویت آموزش‌های علمی - کاربردی و فنی - حرفه ای با هماهنگی دستگاه‌های تولیدی و خدماتی جهت ارتقای کیفیت آموزشی اشاره کرد.

