

سورة

پردیس دانشگاهی  
گروه فناوری اطلاعات  
گرایش تجارت الکترونیک

به کارگیری داده‌کاوی در بهبود مدیریت ارتباط با مشتری

از :

گل‌چیا جنابی

استاد راهنما :

جناب آقای دکتر میرروشندل

شهریور ۱۳۹۳

تقدیم و تشکر

تقدیم به بهترین‌های زندگی‌ام

معلمان عشق و ایثار

تقدیم به

مادرم که می‌ستایم

صبرش را، عشقش را

و پدرم که سپاس می‌دارم

حمایتش را، فکرش را

سر تعظیم فرود می‌آورم

بر لطف بی‌پایانش

اعجاز علم بی‌کرانش

زیبایش

بر خداوند

خداوند عزیز دانا را سپاس می‌گذارم که از لطف خویش به من فرصت کاوش و جستجو در گوشه‌ای از این بیکران زیبای‌های آفرینش را عطا فرمودند تا آغازی باشد برای درک زیبایی‌های او، راهی باشد برای زندگی بهتر.

از استاد گرامی، جناب آقای دکتر میر روشندل که در مراحل مختلف انجام رساله حاضر، از نظرات و راهنمایی‌های ارزشمندشان بهره برده‌ام، صمیمانه سپاسگذارم.

بعلاوه از زحمات خانواده بزرگوارم، مشوقین و راهنمایان همیشگی زندگی‌ام و دوستان عزیزم سپاسگزارم.

## فهرست مطالب

### فصل ۱

- ۱-۱ مقدمه ..... ۲
- ۲-۱ بیان مسئله ..... ۳
- ۳-۱ دلایل اهمیت مطالعه تحقیقاتی ..... ۴
- ۴-۱ هدف پژوهش ..... ۴
- ۵-۱ رویه انجام تحقیق ..... ۴
- ۶-۱ ساختار گزارش ..... ۵

### فصل ۲

- ۱-۲ مقدمه ..... ۷
- ۲-۲ داده کاوی ..... ۷
- ۱-۲-۲ تعریف داده کاوی ..... ۸
- ۲-۲-۲ کاربردهای داده کاوی ..... ۹
- ۳-۲-۲ مراحل داده کاوی ..... ۱۱
- ۳-۲-۲ روش ها و فن های داده کاوی ..... ۱۳
- ۳-۲ مدیریت ارتباط با مشتری ..... ۲۲
- ۱-۳-۲ تعریف مدیریت ارتباط با مشتری ..... ۲۳
- ۲-۳-۲ مدیریت ارتباط با مشتری و دسته بندی های عملکردی ..... ۲۵
- ۳-۳-۲ اهداف و مزایای مدیریت ارتباط با مشتری ..... ۲۹
- ۴-۳-۲ ابعاد مدیریت ارتباط با مشتری ..... ۳۲

### فصل ۳

- ۱-۳ مقدمه ..... ۳۶
- ۲-۳ رابطه داده کاوی و مدیریت ارتباط با مشتری ..... ۳۷
- ۳-۳ مروری بر پژوهش های انجام شده ..... ۴۱

- ۳-۳-۱ ارائه مدل جدید برای ارزیابی رفتار مشتری ..... ۴۱
- ۳-۳-۲ یکپارچه‌سازی فناوری داده‌کاوی برای آنالیز ارزش مشتری ..... ۴۲
- ۳-۳-۳ مدیریت ارتباط با مشتری و واسط کاربردی ..... ۴۳
- ۳-۳-۴ اجرای داده‌کاوی در مدیریت ارتباط با مشتری ..... ۴۴
- ۳-۳-۵ اندازه‌گیری اثر مدیریت ارتباط با مشتری ..... ۴۴
- ۳-۳-۶ بهینه‌سازی الگوریتم K-means و واسط کاربردی سیستم مدیریت ارتباط با مشتری ..... ۴۷
- ۳-۳-۷ استفاده از روش‌های داده‌کاوی برای توسعه مدیریت ارتباط با مشتری به منظور دستیابی به وفاداری مشتری ..... ۴۸
- ۳-۴ جمع‌بندی ..... ۵۱

## فصل ۴

- ۴-۱ مقدمه ..... ۵۳
- ۴-۲ شناخت کسب‌وکار ..... ۵۳
- ۴-۳ داده‌های اولیه مورد استفاده در مورد مطالعه ..... ۵۴
- ۴-۴ داده‌کاوی ..... ۵۶
- ۴-۴-۱ پیش‌پردازش ..... ۵۶
- ۴-۴-۲ مدل‌سازی ..... ۵۹
- ۴-۵ ارزیابی ..... ۶۷
- ۴-۶ جمع‌بندی ..... ۷۳

## فصل ۵

- ۵-۱ جمع‌بندی ..... ۷۵
- ۵-۲ نتیجه‌گیری ..... ۷۶
- ۵-۳ مشکلات و محدودیت‌های موجود ..... ۷۷
- ۵-۴ مطالعات آتی ..... ۷۷

## منابع و ماخذ

پیوست

## فهرست جداول

جدول ۲-۱. تعاریف داده‌کاوی.....	۹
جدول ۲-۲. تعاریف مدیریت ارتباط با مشتری.....	۲۴
جدول ۲-۳. دسته‌بندی‌های عملکردی مدیریت ارتباط با مشتری.....	۲۶
جدول ۳-۱. خلاصه کارهای پیشین.....	۳۶
جدول ۳-۲. نتایج مدل ارائه‌شده.....	۴۳
جدول ۳-۳. مقایسه زمان اجرای دو روش.....	۴۸
جدول ۳-۴. نتایج خوشه‌بندی.....	۴۸
جدول ۴-۱. خلاصه مدل‌های حاصل از اجرای الگوریتم‌های درخت تصمیم.....	۶۵
جدول ۴-۲. خلاصه معیارهای حاصل از اجرای الگوریتم‌ها.....	۷۰
جدول ۴-۳. خلاصه معیارهای حاصل از اجرای الگوریتم‌های درخت تصمیم.....	۷۰
جدول ۴-۴. خلاصه نتایج حاصل از اجرای الگوریتم‌های درخت تصمیم با استفاده از feature selection.....	۷۰
جدول. بسته‌های ارائه‌شده برای خودرو تیبیا.....	۸۸

## فهرست اشکال

- شکل ۱-۲. مراحل داده‌کاوی ..... ۱۱
- شکل ۲-۲. شبکه عصبی با دولایه پنهان ..... ۱۸
- شکل ۳-۲. داده‌ها و خط جداکننده ..... ۲۰
- شکل ۴-۲. ساختار عملیاتی مدیریت ارتباط با مشتری ..... ۳۴
- شکل ۱-۳. چارچوب دسته‌بندی گرافیکی در مدیریت ارتباط با مشتری ..... ۳۹
- شکل ۲-۳. مدل ارائه‌شده ..... ۴۶
- شکل ۳-۳. مدل ارائه‌شده ..... ۵۰
- شکل ۱-۴. نمونه‌ای از داده‌ها پس از پیش‌پردازش ..... ۵۹
- شکل ۲-۴. نمونه جریان مدل‌سازی ..... ۶۰
- شکل ۳-۴. خروجی شبکه بیزین ..... ۶۱
- شکل ۴-۴. نمودار توزیعی شبکه بیزین ..... ۶۱
- شکل ۵-۴. نمودار توزیعی شبکه بیزین ..... ۶۲
- شکل ۶-۴. خروجی الگوریتم k-means ..... ۶۳
- شکل ۷-۴. خروجی الگوریتم Kohenen ..... ۶۳
- شکل ۸-۴. قوانین حاصل از اجرای الگوریتم QUEST ..... ۶۶
- شکل ۸-۴. قوانین حاصل از اجرای الگوریتم CART ..... ۶۶
- شکل ۹-۴. قوانین حاصل از اجرای الگوریتم CART ..... ۶۷
- شکل ۱۰-۴. گراف مربوط به برچسب ..... ۷۱
- شکل ۱۱-۴. گراف مربوط به جنسیت ..... ۷۱
- شکل ۱۲-۴. گراف مربوط به کد استان ..... ۷۲
- شکل ۱۳-۴. گراف مربوط به نوع خودرو ..... ۷۳



## به کارگیری داده کاوی در بهبود مدیریت ارتباط با مشتری

گل چیا جنابی

از دهه ۱۹۸۰ مدیریت ارتباط با مشتری به یک مفهوم مهم در کسب و کار و بازاریابی تبدیل شده است اما تعریف جامعی از مدیریت ارتباط با مشتری وجود ندارد. مدیریت ارتباط با مشتری یک روش استراتژیک است که از اطلاعات، پردازش‌ها، مردم و فناوری برای مدیریت ارتباطات مشتریان با بنگاه استفاده می‌کند و منظور از ارتباطات مشتریان، بازاریابی، فروش، خدمات و پشتیبانی در طول حیات مشتری است. مدیریت ارتباط با مشتری تلاش‌های مدیریتی در فناوری‌ها و پردازش‌هایی است که به شناخت مشتریان بنگاه کمک می‌کند. به این منظور محققین از روش‌های داده کاوی برای استخراج دانش و اطلاعات پنهان در داده‌ها بهره می‌گیرند. در این تحقیق، تمرکز بر صنعت نگهداری ماشین در ایران بوده و از روش‌های داده کاوی گوناگون به منظور بخش‌بندی مشتریان استفاده گردیده است. هدف آن شناسایی مشتریان بالقوه است که تمایل بیشتری به خرید خدمات اختیاری ارائه شده پس از دوران گارانتی دارند. از آنجا که داده‌های استفاده شده در این تحقیق داده‌های واقعی شرکت سایپادک است، چند مرحله پیش‌پردازش انجام شده و رکوردهای مجموعه داده به کمک تخصیص برچسب به رکوردها، به دو دسته تقسیم شده است. پس از انجام مراحل پیش‌پردازش، روش‌های C5.0، CART CHAID و QUEST که از جمله الگوریتم‌های درخت تصمیم هستند، برای دسته‌بندی مشتریان و کمک به سازمان مورد نظر در اتخاذ تصمیم اجرا شده است. علاوه بر آن، از شبکه‌های عصبی و شبکه بیزین نیز برای دسته‌بندی مشتریان استفاده شده است و الگوریتم‌های K-means و Kohonen از جمله روش‌های خوشه‌بندی اجرا شده در این تحقیق هستند. با توجه به متعادل نبودن مجموعه داده و به منظور ارزیابی مدل‌های اجرا شده، از متغیرهای بسیاری مانند recall، precision و F-measure استفاده گردید. نتیجه الگوریتم‌های درخت تصمیم اجرا شده نشان می‌دهد که ویژگی‌های مهم‌تری برای اتخاذ تصمیم بنگاه وجود داشته و نتیجه حاکی از این است که خروجی دو الگوریتم CART و QUEST قوانین کارآمدتری به منظور جذب و حفظ مشتری و در نهایت بدست آوردن سود بیشتر به سازمان ارائه می‌کند.

**کلمات کلیدی:** مدیریت ارتباط با مشتری، وفاداری، داده کاوی، درخت تصمیم، شبکه عصبی، شبکه بیزین، K-means،

Kohonen

## Abstract

### Using Data Mining Techniques for Improving Customer Relationship Management

Golchia Jenabi

Since 1980s, customer relationship management (CRM) has become an important concept in business and marketing. There is no general definition of CRM. It can be described as a strategic approach of using information, processes, technology and people to manage the customers' relationship with the company which means marketing, sales, services, and support across the whole customer life cycle. Customer relationship management refers to the managerial efforts to technologies and processes that helped to understand firms' customers. For this reason data mining techniques have an important role to extract the hidden knowledge and information which is inherit in the data used by researchers. This investigation focuses on the current automotive maintenance industry in Iran and apply various data mining technologies to partitioning customers and its purpose is to determine the group of potential customers who are more likely to purchase optional services after grantee time. Whereas the dataset used in this study is the real data of company, many steps of preprocess were applied and dataset records have been divided into two categories by attributing labels to the records. After preprocessing steps, CAID, C5.0, CART and QUEST methods of decision tree has been applied to classify customers and help the desired organization to make decision. In addition Neural Network and Bayes Network has been used to classify customers. K-means and Kohonen algorithms are clustering algorithm has been applied. For evaluating tested models, many variables such as recall, precision and F-measure have been used. By the results of decision tree methods, there are some more important features for the firm to making decision and the result shows that two algorithms of CART and QUEST conduct more useful rules for the organization to attract and maintain more customers and finally gain more profit.

**Key words:** Customer Relationship Management, Loyalty, Data mining, Decision tree, Neural Network, Bayes Network, K-means, Kohonen

# فصل اوّل

## مقدمه



داده‌کاوی مفهومی نوین در دهه اخیر است به عنوان روشی جهت استخراج دانش استفاده می‌شود که به صورت فراگیر بر کاربردهای آن در زمینه‌های مختلف صنعتی، پژوهشی و تجاری اضافه می‌شود. داده‌کاوی در واقع الگوها و روابط پنهان در میان حجم بسیار بزرگ داده‌های درون پایگاه‌های داده را بازیابی می‌کند که ارزیابی آن‌ها با استفاده از روش‌های مرسوم تحلیل داده‌ها ممکن نیست. داده‌کاوی به صورت توأم مفاهیمی نظیر هوش مصنوعی، یادگیری ماشین و روش‌های آماری و الگوریتم‌های پیشرفته فرا ابتکاری را در کنار روش‌های بهینه‌سازی به کار می‌بندد تا بتواند رهیافتی جدید برای استخراج دانش و الگوهای پنهان در میان داده‌های عظیم ارائه نماید.

داده‌کاوی در زمینه‌های مختلفی کاربرد دارد و پیوسته در حال گسترش است. یکی از کاربردی‌ترین زمینه‌های داده‌کاوی در حوزه مدیریت ارتباط با مشتریان است و کاربردهای زیادی برای شناخت مشتریان و نیازهایشان را ارائه می‌کند. فرآیند مدیریت ارتباط با مشتریان شامل چهار بخش شناخت، جذب، حفظ و نگهداری مشتری و درنهایت گسترش مشتری است. مدیریت ارتباط با مشتری را می‌توان به‌عنوان استراتژی جامعی شناخت که شامل فرآیندهای اندوختن، نگهداری و همکاری با مشتریان منتخب به‌منظور ایجاد ارزش افزوده است. مدیریت ارتباط با مشتری، استراتژی کسب‌وکار و بازاریابی فراگیر است که تکنولوژی، پردازش و تمام فعالیت‌های کسب‌وکار مشتریان را یکپارچه می‌سازد. به بیان دیگر، یک استراتژی کلیدی برای تمرکز بر نیازهای مشتریان و یکپارچه‌سازی رویکردهایی است که مشتری در سازمان با آن‌ها مواجه می‌شود. موضوع شناخت نیازهای مشتریان و جهت‌دهی فعالیت‌های مختلف واحدهای صنعتی و خدماتی به‌منظور رفع نیاز مشتریان، از اهمیت ویژه‌ای برخوردار است و هریک از این ابعاد داده‌کاوی می‌تواند به‌عنوان یک ابزار حمایت‌کننده و پشتیبان تصمیم‌گیری عمل کند.

حجم زیاد داده‌هایی که در پایگاه داده‌های بنگاه‌ها در مورد مشتریان ثبت و نگهداری می‌شود، امکان انجام فرآیند داده‌کاوی و استخراج دانش پنهان از داده‌ها را فراهم می‌کند. به این جهت مباحث مربوط به مدیریت ارتباط با مشتری مانند حفظ مشتری، افزایش ارزش و سودآوری آن‌ها برای شرکت‌ها نیاز به استفاده از فن‌های داده‌کاوی را تقویت نموده‌است.

با گسترش فناوری اطلاعات و امکان جمع‌آوری، ذخیره‌سازی و تحلیل سریع حجم بسیار زیادی داده، موضوع داده‌کاوی به‌عنوان ابزاری قدرتمند در اختیار تحلیل‌گران و کارشناسان قرار گرفته است. هم‌اکنون در بسیار از جنبه‌های مدیریت ارتباط با مشتریان، تحلیل‌گران با استفاده از فن‌های داده‌کاوی سعی بر آن دارند که الگوی رفتاری یا نحوه عملکرد مشتریان را شناسایی و بر اساس این الگو به پیش‌بینی رفتار آنان و تعیین خواسته‌ها و درنهایت وفاداری ایشان دست یابند.

## ۲-۱ بیان مسئله

در محیط B2C<sup>۱</sup>، تأمین‌کنندگان و ارائه‌کنندگان خدمات نیاز به شناخت ویژگی‌ها و نیازهای مشتریان خود دارند. جذب مشتری و جلب رضایتمندی آنان یکی از موضوعات مهم در هدایت بنگاه‌ها است و به این منظور می‌توان از روش‌های داده‌کاوی برای تجزیه و تحلیل داده‌ها و رفتار مشتری استفاده کرد.

مدیریت ارتباط با مشتری شامل مجموعه‌ای از فرآیندها و سیستم‌هایی است که از استراتژی‌های کسب‌وکار به منظور ایجاد ارتباط طولانی‌مدت و سودمند با مشتریان خاص پشتیبانی می‌کند. داده و اطلاعات مشتریان و ابزارهای فناوری اطلاعات از پایه و اساس موفقیت هر استراتژی مدیریت ارتباط با مشتری است. علاوه بر این، رشد سریع اینترنت و فناوری‌های وابسته، امکانات کسب‌وکارها را افزایش داده و در نحوه ارتباط بنگاه‌ها با یکدیگر و مشتریانشان تغییر ایجاد کرده و آن را مدیریت می‌کند. بنابراین موضوع اصلی این پایان‌نامه ارائه روشی کارآمد و مؤثر برای رسیدن بنگاه‌ها به اهدافشان به بهترین شکل ممکن است.

در این پژوهش به مروری بر مفاهیم و تعاریف و رویکردهای مدیریت ارتباط با مشتری و داده‌کاوی پرداخته می‌شود و سپس فن‌ها و کاربردهای داده‌کاوی در مدیریت ارتباط با مشتریان در مطالعات پیشین بررسی می‌شود. تأکید این پژوهش بر کاربردهای داده‌کاوی در مدیریت ارتباط با مشتریان و با توجه به مبحث مدیریت فروش خدمات اختیاری ارائه‌شده در صنعت خودرو است. با استفاده از داده‌های مربوط به مراجعات مشتریان، خدمات دریافتی، خرید آنان و همچنین روش‌ها و ابزارهای مختلف داده‌کاوی، مدل‌های مختلفی برای بخش‌بندی، دسته‌بندی مشتریان و پیش‌بینی رفتار آتی آن‌ها ارائه می‌شود.

## ۳-۱ دلایل اهمیت مطالعه تحقیقاتی

شرکت‌های خودروسازی داده‌های فراوانی از خودروهای تولیدی و صاحبان آن‌ها جمع‌آوری می‌کنند. به این ترتیب حجم زیادی از داده‌ها در پایگاه داده انباشته می‌شود و امروزه با استفاده از الگوریتم‌های پیچیده، استخراج دانش از میان داده‌های حجیم فراهم شده است که با کمک آن‌ها می‌توان دانش نهفته در این منابع را کشف و از آن به منظور بهبود تصمیم‌گیری‌های تجاری استفاده کرد.

شرکت‌ها خودروسازی تمایل دارند دانش بیشتری از ویژگی‌های رفتاری و نیازهای مشتریان خود بدست آورند تا بتوانند پیش‌بینی‌های دقیق‌تری از رفتارها، درخواست‌ها و نیازهای مشتریان داشته باشند. به این جهت شرکت‌ها می‌خواهند نیازهای

---

<sup>1</sup> Business to Customer

مشتریان خود را بشناسند تا به آن‌ها پاسخ مناسبی بدهند و از این طریق مشتریان را جذب و حفظ کنند و آنان را به سازمان وفادار کنند.

#### ۴-۱ هدف پژوهش

هدف از انجام این پژوهش، بررسی کاربرد ابزارها و روش‌های داده‌کاوی در مدیریت ارتباط با مشتری است. تمرکز این مطالعه تحقیقاتی بر مدیریت ارتباط با مشتری در شرکت‌های خودروسازی و خدمات پس از فروش است. با توجه به این که در اجرای پروژه از داده‌های یک شرکت ارائه‌کننده خدمات پس از فروش خودرو استفاده می‌شود، اهداف مطالعه موردی بر روی موارد زیر متمرکز شده‌است:

- دسته‌بندی مشتریان بر اساس ویژگی‌های رفتاری و داده‌های ثبت‌شده آنان.
- تحلیل ویژگی‌های گروه‌های مختلف مشتریان به منظور تصمیم‌گیری سازمان در مورد نحوه ارتباط با مشتریان.
- پیش‌بینی رفتار آتی مشتریان با توجه به عملکرد آنان در گذشته.
- تمرکز بر خدمات ارائه‌شده به مشتریان سازمان و ترغیب مشتریان به خرید خدمات

برای دست یافتن به اهداف ذکر شده، از روش‌های مختلف داده‌کاوی مانند دسته‌بندی و خوشه‌بندی استفاده شده‌است.

#### ۵-۱ رویه انجام تحقیق

در بخش ابتدایی پژوهش به بررسی ادبیات موضوع از دیدهای مختلف پرداخته می‌شود. ابتدا به بررسی مفهوم و ادبیات مدیریت ارتباط با مشتری، داده‌کاوی و کاربردهای آن پرداخته می‌شود. سپس مطالعات انجام‌شده در گذشته بررسی و مرور می‌شود. بخش دوم این تحقیق به مطالعه و بررسی مورد مطالعه و کاربرد ابزارهای داده‌کاوی در استخراج دانش از مجموعه داده‌های حاصل از ذخیره مراجعات مشتریان شرکت سایپا پدک به نمایندگی‌های این شرکت است. انجام مراحل داده‌کاوی در فصل بعد به تفصیل آمده‌است.

## ۶-۱ ساختار گزارش

همانطور که اشاره شد، داده‌کاوی و مدیریت ارتباط با مشتری دو محور اصلی این پژوهش هستند. لذا در فصل دوم این گزارش به تعریف و بررسی مفاهیم مدیریت ارتباط با مشتری، داده‌کاوی و روش‌های آن پرداخته می‌شود و در فصل سوم مروری بر کارهای مرتبط انجام‌شده در حوزه مدیریت ارتباط با مشتری و داده‌کاوی در صنایع مختلف صورت گرفته است. فصل چهارم پایان‌نامه شامل توضیح مورد مطالعه این پژوهش، روش‌ها و متدهای استفاده‌شده مانند دسته‌بندی و خوشه‌بندی به منظور مدل‌سازی است و در نهایت به مرور نتایج حاصل از انجام پژوهش و ارائه پیشنهادها برای تحقیقات آتی در فصل پایانی پرداخته می‌شود.

فصل دوم

مفاهیم

پیش زمینه





## ۱-۲ مقدمه

پس از آنکه در فصل قبل به کلیات تحقیق اشاره شد و نحوه کار و شمای کلی تحقیق ارائه گردید، در این فصل مبانی نظری تحقیق و پیشینه تحقیق مورد بررسی قرار می‌گیرد. در این فصل، اطلاعات جمع‌آوری شده از کتاب‌ها و مجلات معتبر دانشگاهی به منظور مروری بر ادبیات پژوهش ارائه می‌شود. ابتدا به ادبیات نظری ارائه شده درباره موضوع پژوهش پرداخته می‌شود به این ترتیب که اول داده کاوی و تعاریف مفاهیم اساسی مربوط به آن معرفی می‌شود. در بخش دوم به بسط مفهوم مدیریت ارتباط با مشتری پرداخته شده و تعاریف متعدد، اهداف و مزایای آن ارائه گردیده است.

حال به منظور جمع‌آوری اطلاعات در زمینه ادبیات پژوهش از روش کتابخانه‌ای استفاده شده است. لذا با مطالعه کتاب‌ها، مقالات و تحقیقات دیگر پژوهشگران در این حوزه اطلاعات مورد نیاز جمع‌آوری گردیده است. میزان دسترسی به اطلاعات به زبان فارسی در ارتباط با این موضوع، بسیار محدود در حد یک یا دو مقاله و پایان‌نامه است و برخی اطلاعات موجود در این مجموعه برآمده از مطالعه پایان‌نامه‌هایی در زمینه داده‌کاوی و مدیریت ارتباط با مشتری است اما در ارتباط با این موضوع به زبان انگلیسی اطلاعات بسیار مفید و جامع وجود دارد.

## ۲-۲ داده کاوی

استخراج خودکار دانش جدید و مفید از منابع داده‌ای حجیم موجود طی یک فرآیند غیر بدیهی مشخص، داده‌کاوی نامیده می‌شود [۱]. هدف اصلی در داده‌کاوی کشف دانش و نظمی است که در داده‌ها وجود دارد. پس از کشف دانش ممکن است دو حالت وجود داشته باشد: حالت اول آن است که افراد خبره در دامنه داده مورد کاوش، آگاه به دانش استخراج شده باشند. در این صورت آن دانش به عنوان یک قانون صحیح تلقی خواهد شد. در حالت دوم ممکن است دانش کشف شده، یک دانش جدید بوده و در بین افراد متخصص در آن زمینه شناخته شده نباشد. حال این دانش بررسی شده و در صورت منطقی بودن، تبدیل به فرضیه شده و در نهایت درست یا غلط بودن این فرضیه با آزمایش‌ها و بررسی‌های متعدد اثبات می‌شود و در نهایت این فرضیه در صورت درست بودن به قانون تبدیل خواهد شد.

## ۱-۲-۲ تعریف داده‌کاوی

داده‌کاوی می‌تواند با توجه به شرکتی که کسب‌وکار رو دایر کرده، استخراج و آنالیز حجم زیاد داده بنگاه به منظور آشکار کردن قوانین پنهان و ناشناخته یا بررسی و ممیزی قانون شناخته‌شده و همچنین مدل کردن آن با روش‌های پیشرفته و کارآمد تعریف شود. داده‌کاوی به عنوان فرآیند استخراج اطلاعات کارآمد به نام داده‌ی محصول شناخته می‌شود و در واقع کاوش ضمنی ناشناخته‌ها، ارزش بالقوه دانش تصمیم‌گیری و قوانین از یک پایگاه داده بزرگ است. به علاوه پیش‌بینی نتایج رفتاری که رخ نداده با توجه به اطلاعات، می‌تواند اساس و اصولی را برای مدیریت تصمیم و برنامه‌ریزی بازار را برای شرکت‌ها ارائه کند [۲].

اطلاعات داده‌کاوی سه ویژگی اصلی ناشناخته، مؤثر و کاربردی بودن را داراست. منظور این است که داده‌کاوی برای یافتن اطلاعات و دانشی است که به صورت مستقیم قابل‌درک نیستند و حتی با یکدیگر در تضاد هستند. به بیان دیگر داده‌کاوی کشف اطلاعات غیر منتظره است [۲].

افراد خبره در زمینه داده‌کاوی تعاریف متنوعی از یک مفهوم دارند که این تعاریف گوناگون با پیش‌زمینه تخصصی ایشان مرتبط است. برخی از تعاریف داده‌کاوی در جدول زیر آمده‌است.

جدول ۲-۱. تعاریف داده‌کاوی

تعریف داده‌کاوی	محقق
فرآیند شناخت الگوهای معتبر، جدید، مفید و قابل فهم از داده‌ها می‌باشد.	فیاد [۳]
به فرآیند استخراج اطلاعات نهفته، قابل فهم، قابل پیگیری از پایگاه داده‌های بزرگ و استفاده از آن در تصمیم‌گیری‌های تجاری مهم اطلاق می‌شود.	زکولین [۳]
مجموعه‌ای از روش‌ها در فرآیند کشف دانش است که برای تشخیص الگوها و روابط نامعلوم در داده‌ها مورد استفاده قرار می‌گیرد.	فروزا [۳]
فرآیند کشف الگوهای مفید از داده‌ها را داده‌کاوی گویند.	جان [۳]
فرآیند انتخاب، کاوش و مدل‌بندی داده‌های حجیم، جهت کشف روابط نهفته با هدف به دست آوردن نتایج واضح و مفید، برای مالک پایگاه داده را داده‌کاوی گویند.	ناظمی و مشکانی [۳]
فرآیند کشف دانش از حجم زیادی داده که شامل روش‌ها و فن‌های منظم علمی مانند یادگیری ماشین، ریاضیات و هوش مصنوعی برای دستیابی به هدف است.	وانگ [۲].
فرآیند استخراج و شناسایی الگوها یا اطلاعات از حجم داده بسیار زیاد است.	برسون و همکاران [۴]

داده‌کاوی مجموعه‌ای از فرآیندها شامل آماده‌سازی داده، داده‌کاوی و ارزیابی نتایج است اما به طور کلی به فرآیند اساسی اشاره دارد که دانش از داده موجود بدست می‌آید [۲].

## ۲-۲-۲ کاربردهای داده‌کاوی

امروزه داده‌کاوی به بحث داغی در میان میران تبدیل شده‌است. چرا داده‌کاوی می‌تواند هوش تجاری و دانش پنهانی را که در داده‌ها موجود است را ارائه کند. در گذشته تحلیلگران دانش و اطلاعات موجود در داده‌ها را استخراج می‌کردند. اما با رشد روزافزون حجم داده‌ها، باید اطلاعات مفید از آن‌ها استخراج و به صورت مختصر و مفید ارائه شود به این جهت روش‌های مبتنی بر رایانه به وجود آمده‌است. به این ترتیب داده‌کاوی در اکثر سازمان‌هایی که حجم عظیمی از داده وجود دارد و نیاز به کشف دانش نهفته دارند کاربرد دارد. مدیران سازمان‌ها می‌دانند که دانش حاصل از داده‌کاوی می‌تواند نقش بسزایی در تصمیم‌گیری‌های سازمانی

داشته باشد [۵]. به طور کلی روش‌های داده‌کاوی راهی برای استخراج دانش پنهان در پایگاه داده است [۶] و هر روش داده‌کاوی می‌تواند یک یا چند نوع مدل را ارائه کند که شامل روش‌های پیش‌بینی (دسته بندی<sup>۲</sup>، رگرسیون<sup>۳</sup>) و روش‌های توصیفی (خوشه‌بندی<sup>۴</sup>، کشف الگوهای ترتیبی<sup>۵</sup> و کشف قوانین انجمنی<sup>۶</sup>) است. روش‌های یادگیری ماشین بسیاری برای هر نوع مدل داده‌کاوی وجود دارد. انتخاب‌های روش‌های داده‌کاوی باید بر اساس ویژگی‌های داده و نیازهای کسب‌وکار باشد [۷].

از داده‌کاوی می‌توان برای مدیریت بحران، تحلیل شکست، مدیریت روابط با مشتری استفاده کرد. روش‌های داده‌کاوی در حیطه‌های گوناگون مانند بازاریابی، بخش‌بندی بازار، پیش‌بینی نیاز بازار، تشخیص خطا و تعیین استراتژی در صنایع مختلف اعم از بیمه، بانکداری، ارتباطات و گردشگری قابل اعمال است که برای رسیدن به مدیریت ارتباط با مشتری مؤثر اجرا می‌شود که به مدیران صنعتی برای اتخاذ تصمیمات بازار کمک می‌کند [۸]. اهمیت روش‌ها در ارتقا آنالیز مدیریت ارتباط با مشتری همراستا با رشد پایگاه داده است [۹]. یکی از کاربردهای داده‌کاوی تقسیم بازار مشتریان به زیر بخش‌های مجزا است به طوری که هر زیربخش می‌تواند به‌عنوان بازار هدف، برای دستیابی به بازاریابی متمایز انتخاب شود [۱۰].

امروزه در بیشتر صنایع، روش‌های دسته‌بندی داده‌کاوی، برتری قابل توجهی دارند و به طور گسترده در مدیریت ارتباط با مشتری اجرا می‌شوند که در این پژوهش نیز می‌شود. استفاده از الگوریتم‌های مبتنی بر کامپیوتر در قدرت پردازش به نسبت روش‌های آماری سنتی برتری قابل توجهی دارد که مبتنی بر روش‌های کلاس‌بندی است.

داده‌کاوی به تدریج اهمیت بیشتری برای بنگاه‌ها پیدا کرد. بنگاه‌ها می‌توانند با سرعت و به طور کارآمد، مشتریان پیچیده را با روش‌های داده‌کاوی از حجم زیادی داده به وسیله تدبیر بازاریابی مناسب، افزایش ارزش مشتری و بهبود مزایای رقابتی بنگاه، برای کمک به شناخت و ارتباط متقابل با آن‌ها بیابند [۱۱].

چنگ و لیو و چنگ دریافتند که فناوری داده‌کاوی می‌تواند به کسب‌وکار در مدیریت فناوری اطلاعات کمک کند به شرح زیر:

[۱۲]

- بهبود کارایی کسب‌وکار با حداقل بودجه.

- استفاده از بازاریابی پایگاه داده برای نگهداری ارتباط با مشتری.

---

<sup>2</sup> Classification

<sup>3</sup> Regression

<sup>4</sup> Clustering

<sup>5</sup> Sequential Pattern

<sup>6</sup> Association Rule