

به کارگیری داده کاوی در تشخیص الگوی رفتار تأمین کنندگان قطعات خودرویی

مورد مطالعه: شرکت صنعتی مهد خودرو توس

بهناز ترابی مقدم^۱؛ محمد لگزیان^۲؛ شمس الدین ناظمی^۳؛ محسن کاهانی^۴

چکیده

در محیط رقابتی و پرتلاطم دنیای کسب و کار امروز، به دلیل اینکه هیچ سازمانی نمی تواند به تنهایی بقا داشته باشد، مقوله مدیریت زنجیره تأمین به عنوان استراتژی کلیدی سازمان ها از اهمیت مضاعفی برخوردار گردیده است؛ از دیگر سو تمرکز بر استراتژی های تولید درست به موقع، منجر به افزایش اهمیت مسأله ارزیابی و انتخاب پیمانکاران شده است. از آنجا که دقیق ترین ارزیابی از تأمین کنندگان با بررسی داده های تأمین آنها در بازه های گذشته میسر می گردد، داده کاوی به عنوان ابزار پیمودن این راه انتخاب گردیده است.

تحقیق حاضر گامی است مؤثر در جهت اینکه از بانک های داده سازمان در اجرای فرآیند داده کاوی بهره گرفته و از نتایج آن در جهت تصمیم سازی های مفید برای مدیران ارشد استفاده شود. به این منظور در این پژوهش الگوی رفتار تأمین کنندگان در تأمین قطعات خودرویی جهت یک شرکت قطعه ساز خودرو، از طریق داده کاوی مورد بررسی قرار گرفته است. متدولوژی مورد استفاده CRISP-DM و نرم افزار مورد استفاده SPSS Clementine انتخاب گردیده است. با اعمال کارکرد دسته بندی روی داده های تأمین قطعات در بازه زمانی سال ۸۷، ۸۸ و نیمه اول ۸۹، مهم ترین عوامل مؤثر بر رفتار تأمین کننده از منظر تحقق تأمین سفارش به عنوان خروجی مدل شناسایی شده و همچنین درجه اهمیت هر یک از این صفات در ارزش گذاری متغیر هدف استخراج گردیده است. خروجی مدل حاکی از اهمیت بالای صفات میزان سفارش ماهیانه، روز سفارش گذاری، وزن قطعه؛ و عدم اهمیت فاکتورهای نوع انبار مرتبط با قطعه و مشتری قطعه بوده است.

واژگان کلیدی

داده کاوی، CRISP-DM، رفتار تأمین کنندگان، دسته بندی، تحقق تأمین سفارش.

کنفرانس داده کاوی ایران

۱. نویسنده مسئول: کارشناس ارشد مدیریت صنعتی، دانشگاه فردوسی مشهد، Btorabimoghadam@yahoo.com

۲. دانشیار گروه مدیریت، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی مشهد، M-Lagzian@um.ac.ir

۳. دانشیار گروه مدیریت، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی مشهد، Nazemi_shm@um.ac.ir

۴. دانشیار گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه فردوسی مشهد، Kahani@um.ac.ir

۱. مقدمه

تشدید رقابت جهانی در دهه ۱۹۸۰ و مطرح شدن فلسفه تولید بهنگام^۱ سازمان های در تراز جهانی را روز به روز بیشتر به اهمیت و منافع بالقوه تعامل در روابط خریدار-فروشنده متوجه ساخت(جعفرنژاد، ۱۳۸۵). از دهه ۱۹۹۰ به بعد با مطرح شدن مدیریت زنجیره تأمین، نگرش شرکت های برتر از تمرکز بر بهبود وضعیت سایت های تولیدی خود، به بررسی و بهبود کل زنجیره تأمین تعمیم پیدا کرد(پراون^۲، ۱۳۸۳). ایده اصلی زنجیره تأمین با هدف استقرار یک رویکرد سیستمی و جامع(والمن^۳ و دیگران، ۲۰۰۵)، موجب گردید نقش تأمین کنندگان با هدف بهبود شبکه تأمین به وضعیت برنده - برنده(چیس^۴ و دیگران، ۲۰۰۴) پررنگ تر گردیده(هوها و کریشنان^۵، ۲۰۰۸)؛ و لذا مسأله انتخاب تأمین کننده مناسب یکی از مسائل مهم پیش روی صاحبان کسب و کار قرار گیرد.

از سوی دیگر همزمان با رشد رقابت بازارهای جهانی، انقلاب در تکنولوژی و اطلاعات منجر به معرفی دانش به عنوان مهم ترین دارایی با ارزش سازمان گردید. در این میان، وجود انباره های حجیم داده، توان بالای ذخیره سازی و محاسبات و الگوریتم های جدید سازمان های رقابتی را ناگزیر از ورود به عرصه استفاده از جدیدترین تکنولوژی های مرتبط در زمینه اطلاعاتی از جمله داده کاوی نموده است(شهرابی، ۱۳۸۶، شهرابی و شکورنیا، ۱۳۸۸، باکائو^۶، ۲۰۰۸). در این مقاله ابتدا با بررسی پیشینه مسأله انتخاب تأمین کنندگان، به روشن سازی مفهوم ارزیابی و انتخاب پرداخته شده و سپس مصادیق استفاده از داده کاوی در این مسأله بررسی می شود. در نهایت با طی گام های فرآیند داده کاوی طبق چرخه CRISP-DM به بررسی الگوی رفتار تأمین کنندگان در تحقق تأمین قطعات پرداخته شده و دانش نهفته در لایه های داده سازمان استخراج و از آن در مدیریت روابط با تأمین کنندگان بهره جسته می شود.

مروری بر ادبیات موضوع

۱-۲. ارزیابی و انتخاب تأمین کنندگان

۱-۱-۲. پیشینه و لزوم ارزیابی و انتخاب تأمین کنندگان

مطالعات اولیه در حوزه تأمین کنندگان اکثراً معطوف به انگیزه ها از برون سپاری بوده، در حالی که در سال های اخیر این مطالعات بیشتر معطوف بر روابط سازمان- تأمین کننده گردیده است(کنان و تان^۷، ۲۰۰۶). جذابیت این موضوع برای صاحبان کسب و کار، باعث جلب توجه محققان آکادمیک به این حوزه گردیده است(دونگ لی^۸ و دیگران، ۲۰۰۷). همچنین مطرح شدن اثر شلاق چرمی^۹ نیز به پررنگ شدن نقش تأمین کننده کمک نموده(والمن و دیگران، ۲۰۰۵). که بر اثر آن نوسانات تقاضای مشتری با فشار

۱ . Just In Time

۲ . Brown

۳ . Vollman

۴ . Chase

۵ . Ho Ha & Krishna

۶ . Bacao

۷ . Kanan & Tan

۸ . Dong Li

۹ . Bull weap effect

مضاعف بر شرکت و سپس بر تأمین کنندگان آن وارد می گردد. انتخاب تأمین کننده یک فرآیند تصمیم گیری جمعی (شرفی، ۱۳۸۸) و به دلیل نیل به چند هدف به طور همزمان، جزو دسته مسائل تصمیم گیری چندمعیاره^۱ محسوب می گردد (راماناتان^۲، ۲۰۰۷).

۲-۱-۲. فرآیند انتخاب تأمین کنندگان

از مطالعه ادبیات موضوع، دو دسته بندی مطرح ذیل استخراج شده که در جدول ۱ همراه با انطباق بین گام های فرآیند آورده شده است. لازم به ذکر است بیشترین تحقیقات در این حوزه بر گام ۳ و ۴ از این فرآیند متمرکز بوده است (هولین و دیگران، ۲۰۰۹).

جدول ۱ - مقایسه دو مدل مطرح در فرآیند انتخاب تأمین کنندگان

مدل هوها و کریشنا (۲۰۰۸)	مدل مونسزکا و دیگران ^۳ (۱۹۹۸)
گام ۱. تعریف مسأله اصلی	گام ۱. طرح مسأله نیاز به انتخاب تأمین کننده
گام ۲. فرموله کردن ضابطه ارزیابی	گام ۲. تعریف و احصاء معیارهای حیاتی منبع یابی
	گام ۳. انتخاب استراتژی خرید یا برون سپاری مناسب
	گام ۴. انتخاب روش ارزیابی تأمین کننده
گام ۳. بررسی اولیه پتانسیل و توان تأمین کنندگان	گام ۴. شناخت مجموعه منابع ممکن
	گام ۵. کاهش گستردگی مجموعه تأمین کنندگان در دسترس و محدود نمودن جامعه آن ها
گام ۴. انتخاب نهایی بهترین تأمین کننده	گام ۷. انتخاب نهایی تأمین کننده مناسب

۲-۲. ورود داده کاوی به مدیریت زنجیره تأمین

هان و کمبر^۴ (۲۰۰۶) عامل اصلی گرایش سازمان ها به استفاده از تکنیک های داده کاوی را سرعت بالای رشد داده های ذخیره شده در پایگاه های داده سازمان ها می دانند. بدیهی است پایگاه های داده مرتبط با مدیریت زنجیره تأمین در سازمان هم از این قاعده مستثنی نبوده اند. از این منظر داده کاوی را می توان فرآیندی ارزیابی کننده تلقی نمود که با هدف استخراج دانش مفید و سودمند از داده های سازمان وارد حوزه ارزیابی و انتخاب تأمین کنندگان گردیده است (گارگانو و رگاد^۵، ۱۹۹۹).

تاکنون کاربردهای داده کاوی در انتخاب تأمین کنندگان بیشتر شامل روش شبکه های عصبی و سیستم های خبره شده است. رویکردهای داده کاوی مثل درخت تصمیم و شبکه های عصبی ابزار خوبی برای تقریب مسائل ناپارامتریک و غیرخطی محسوب می شوند (وو^۶، ۲۰۰۹). هلت^۷ (۱۹۹۸) پس از بررسی چند روش تصمیم گیری از قبیل تحلیل خوشه، رگرسیون چندگانه و تحلیل چند متغیره به این نتیجه رسید که برای مسأله انتخاب تأمین کنندگان، تحلیل خوشه در بین سایر روش ها، از پتانسیل بالاتری برخوردار است. زیرا احتمال رشدن تأمین کننده "خوب" را در مراحل اولیه کار داده کاوی کاهش می دهد (هنگ هونگ و دیگران،

۱ . MCDM: Multi Criteria Decision Making

۲ . Ramanathan

۳ . Monczka

۴ Han & Kamber

۵ .Gargano & Raggad

۶ . Wu

۷ . Holt

۲۰۰۵). کومار و دیگران در سال ۲۰۰۴ برای داده کاوی روی داده های مربوط به تأمین کنندگان که با پارامترهای ماهیتاً فازی سروکار داشت، رویکرد برنامه ریزی هدف فازی را به کار گرفتند که به طور همزمان چند هدف را ارضا می نمود (یائو کنگ و لی، ۲۰۱۰، وو، ۲۰۰۹). در این خصوص پس از آن در سال ۲۰۰۷ رویکرد دیگری توسط دونگ لی و دیگران پیشنهاد گردید؛ محققان با ارائه مثال عددی واقعی نشان دادند این رویکرد علاوه بر در نظر گرفتن رویکرد فازی، در شرایط عدم اطمینان نیز کارایی بیشتری دارد.

۲. فرآیند داده کاوی و متدولوژی مورد استفاده در پژوهش

داده کاوی یک فرآیند تکرار پذیر است و به همین دلیل می تواند بارها تکرار گردد (کانتاردزیک، ۱۳۸۵). برای اطمینان از اجرای موفق یک پروژه داده کاوی همواره باید یک متدولوژی مشخص جهت اجرا وجود داشته باشد (بری و لینف^۱، ۲۰۰۰). در بین متدولوژی هایی که به عنوان فرآیندهای استاندارد داده کاوی تعریف شده اند، می توان از CRISP-DM^۲ و SEMMA^۳، چرخه ویلیامز و هوانگ^۴ (۱۹۹۶)، ژو و دیگران^۵ (۱۹۹۸)، ویتن و فرانک^۶ (۲۰۰۰) و گائو^۷ (۲۰۰۰) نام برد (لی و روآن^۸، ۲۰۰۷). در این تحقیق با تحقیق با توجه به ماهیت داده ها، متدولوژی CRISP-DM انتخاب گردیده است. بر اساس این متدولوژی، پروژه داده کاوی شامل یک چرخه عمر شش مرحله ای است که مراحل باهم در تعاملند (لی و روآن، ۲۰۰۷)^۸.

۳-۱. مرحله درک مسئله کسب و کار

۳-۱-۱. تعیین اهداف کسب و کار: در فاز شناخت محیط مورد مطالعه، هدف کلیدی سازمان به صورت "بهبود مدیریت زنجیره تأمین به جانب تأمین کنندگان" تعریف می شود. با توجه به اینکه بهبود هر سیستم، ابتدا نیاز به اندازه گیری وضعیت موجود دارد، نیاز به اجرای روشی جهت ارزیابی رفتار تأمین قطعات در سازمان احساس شد. همچنین با توجه به اینکه مهم ترین عامل در تأمین قطعات تأمین کنندگان، میزان تحقق تولید آنها شناسایی شده، در این مقاله به بررسی رفتار تأمین از این منظر پرداخته می شود.

۳-۱-۲. ارزیابی موقعیت: بررسی سوابق گذشته تأمین در بانک های اطلاعاتی مورد نظر پژوهش، حاکی از ممکن بودن اجرای فرآیند داده کاوی با هدف فوق الذکر می باشد. فاز بررسی داده های مورد نظر در گام شناخت داده به تفصیل مورد بحث قرار خواهد گرفت. همچنین برای این پژوهش نیاز به تشکیل تیمی متشکل از داده کاو، چند تن از خبرگان صنعت مورد نظر، کارشناس فناوری اطلاعات، و نیز از نظر سیستمی نیاز به استفاده از نرم افزارهای Excel، MySQL و SPSS Clementine می باشد.

۳-۱-۳. تعیین اهداف داده کاوی: با توجه به اهداف سازمان که در مرحله اول از گام درک مسأله کسب و کار مورد بررسی قرار گرفت، هدف عمده داده کاوی در این پژوهش تحت عنوان "بررسی الگوی رفتار تأمین کنندگان از منظر وضعیت تحقق تأمین سفارش قطعات خودروبی" تعریف شده است.

۱ . Berry & Linoff

۲ . SEMMA: Sample, Explore, Modify, Model, Assess

۳ . Willialms & Huang

۴ . Zhu et al

۵ . Witten & Frank

۶ . Gao

۷ . Li & Ruan

۸ WWW.CRISP-DM.ORG

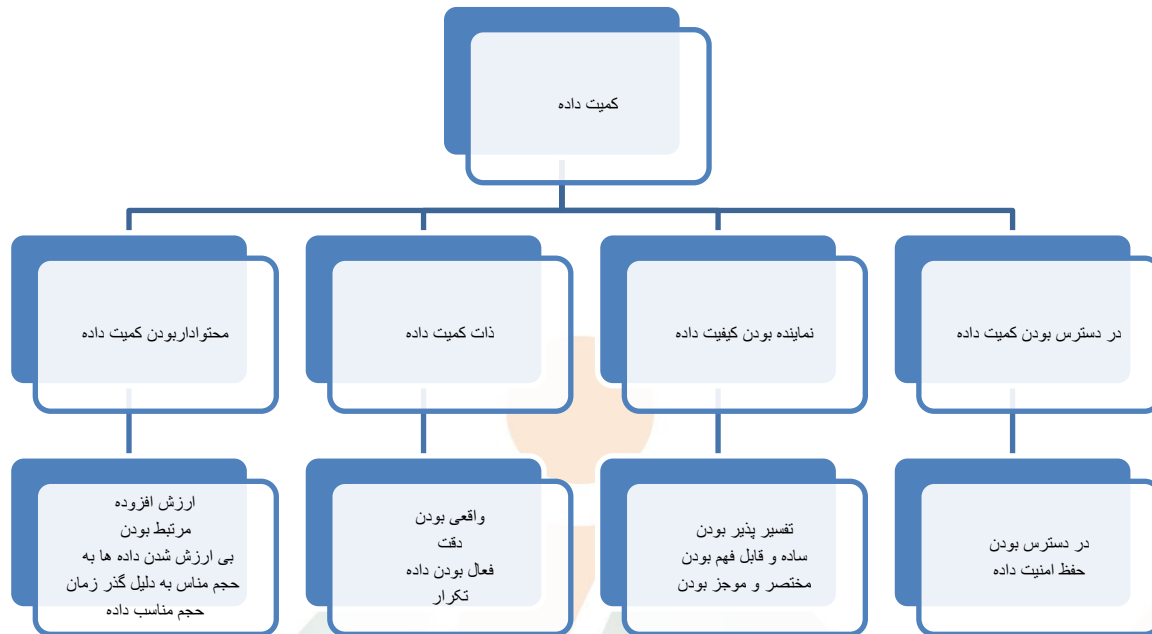
۲-۲. مرحله درک داده

۲-۲-۱. جمع آوری داده اولیه: با توجه به آنکه اجرای موفق پروژه داده کاوی وابستگی مستقیم به وضعیت داده های پروژه دارد، این گام را می توان یکی از مهم ترین گام های فرآیند داده کاوی دانست. داده های مورد نیاز این پژوهش در قالب جداول ذیل گردآوری شده اند:

- **جدول تأمین قطعات توسط تأمین کنندگان:** این جدول که مستقیماً از نرم افزار انبار سازمان استخراج گردیده، شامل ۱۸ فیلد و در بازه زمانی مورد نظر این تحقیق شامل ۱۸۴۰۰ رکورد می باشد؛
- **جدول برنامه ها یا سفارشات برای تأمین کنندگان:** این جدول شامل ۱۸ فیلد و ۱۰۵۹۰ رکورد می باشد که هر رکورد معرف یک سفارش مجزا برای تأمین کننده است.
- **جدول مشخصات تأمین کنندگان:** با توجه به اینکه در سازمان هیچ منبع یکپارچه و جامعی جهت استخراج فاکتورهای مرتبط با هر تأمین کننده وجود نداشت، در خلال تحقیق ناگزیر از استخراج این اطلاعات از منابع پراکنده موجود در سازمان، استعلام از تأمین کنندگان و رجوع به کتب راهنمای صنعت بوده ایم. این جدول شامل ۵۶ رکورد یعنی به تعداد تأمین کنندگان مورد نظر پژوهش و ۹ فیلد توصیف کننده جهت هر تأمین کننده می باشد.
- **جدول مشخصات قطعات:** در خصوص هر کد قطعه مورد سفارش به تأمین کننده، از طریق بررسی اسناد پراکنده موجود در سازمان، مراجعه به آدرس اینترنتی سازمان های سایپا و سازه گستر و نیز برگزاری جلسه با افراد آگاه فیلدهای توصیف کننده قطعه استخراج گردیده است. لذا این جدول شامل ۴۴۸ رکورد یعنی به تعداد کد قطعات مورد نظر پژوهش و ۵ فیلد می باشد.
- **جدول مشخصات نیروی انسانی:** این صفت از آن جهت از نظر محقق با اهمیت پیش بینی می گردد که در بازه زمانی مورد نظر تحقیق، سه دسته افراد مسئولیت برنامه ریزی و تأمین قطعات پیمانکاری را به عهده داشته اند. ورود داده ها به نرم افزار Clementine در قالب فایل Excel ۲۰۰۳ صورت گرفته است.

۲-۲-۲. تشریح داده: بانک داده مورد نظر این پژوهش شامل ۱۰۴۷۸ رکورد بوده که مربوط به سفارشات صادره توسط شرکت مهدخودرو توس برای تأمین کنندگان در بازه زمانی فروردین ۸۷ تا شهریور ۸۹ می باشد. در مورد هر رکورد، ۴۳ فیلد شناسایی و استخراج شده است.

۲-۲-۳. بررسی داده: به منظور بررسی اولیه داده ها، پس از ورود داده ها به نرم افزار از نمودارهای توزیع و هیستوگرام استفاده شده است. در این گام جهت کنترل حجم و کمیت داده از پارامترهای مطرح شده در مدل ونگ و دیگران (۱۹۹۴) استفاده شده که در شکل شماره ۱ مشاهده می گردد.



شکل شماره ۱- مدل کنترلی در خصوص کمیت داده در داده کاوی(ونگ و دیگران، ۱۹۹۴)

بر اساس مدل فوق، برخی عوامل که احتمالاً در خلال تحقیق ایجاد مشکل می کرده اند، مورد کنترل قرار گرفته اند که به شرح ذیل می باشد:

- **واقعی بودن:** با توجه به اینکه داده ها از نرم افزار انبار و یا از طریق فایل های موجود در سازمان استخراج شده اند، اطمینان از صوری و ساختگی نبودن داده ها حاصل شده است.
- **تکرارپذیری:** به عنوان مثال با وجود آنکه کدهای قطعات در سازمان، حدود ۲۰۰ مورد بوده است، در گام آماده سازی داده که در ادامه آورده خواهد شد، سعی شده است تا تعدد این فیلد از طریق گنجاندن داده های این فیلد در ۳ دسته به تفکیک نوع قطعه(مواد اولیه، نیم ساخته، محصول) پوشش داده شود. در مورد سایر صفات نیز طبقه بندی های مشابهی روی مقادیر موجود اعمال گردیده است.
- **مرتبط بودن داده:** با توجه به اینکه استخراج داده ها تحت کنترل محقق بوده، این عامل خود به خود کنترل شده است. در مواردی که گرفتن گزارش از نرم افزار انبار، منجر به وارد شدن داده های زاید در فایل شده، در نهایت از ورود فیلدهای نالازم به مدل جلوگیری گردیده است.
- **ساده و قابل فهم بودن داده:** سعی بر آن بوده است تا نامگذاری فیلدها قابل فهم باشد.
- **در دسترس بودن داده ها:** این موضوع در گام اول فرآیند و هنگام تعریف مسأله توسط محقق کنترل شده و اطمینان نسبی از در دسترس بودن داده مورد نیاز حاصل شده است.

۳-۲-۴. **تأیید کیفیت داده:** جهت بررسی کیفیت داده ورودی به نرم افزار، از **Table node** استفاده شده و بارها از طریق اصلاح داده های ورودی و اعمال آن، اطمینان از کیفیت داده های تحقیق حاصل شده است. برخی خطاهای احتمالی که در این قسمت مد نظر قرار گرفته و مرتفع گردیدند، عبارتند از: مقادیر خارج از رفتار یا مقادیر حدی، ارزش های گم شده، صفات تکراری و داده هایی که در فرم مناسب برای مدل سازی نیستند.

۳-۳. مرحله پیش پردازش داده

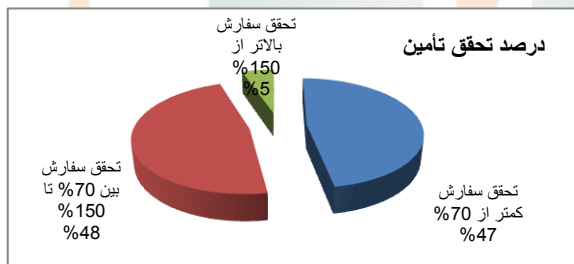
پیش پردازش داده ها به تنهایی ۶۰ درصد از حجم کار داده کاوی است (نای^۱ و دیگران، ۲۰۰۹). عمده اقدامات لازم در این مرحله از دو منظر ذیل تحلیل گردیده است:

۳ ۴ ۱. **تهیه جدول انطباقی داده های تأمین با سفارشات:** با استفاده از جداول مطرح شده در قسمت "تشریح داده" پس از تعیین صفاتی که مؤثر بر الگوی رفتار تأمین به نظر می رسیدند، اکنون نیاز به جدولی است که کلیه صفات مورد نظر را به رکورد مورد نظر تخصیص دهد. تهیه این جدول تطابقی طی دو فاز صورت پذیرفته است:

- **فاز اول:** در این فاز تطابق صفات مربوط به مشخصات قطعه، مشخصات تأمین کننده و مشخصات تیم نیروی انسانی از طریق برقراری تناظرهای لازم صورت گرفته است. این کار در محیط Excel انجام شده و با توجه به اینکه از ابتدای اجرای پروژه صفات ورودی به مدل ها دقیقاً مشخص نبود، بارها در خلال تحقیق ناگزیر از برگشت به عقب، بازنگری صفات مؤثر و اعمال اصلاحات در جدول بوده ایم.
- **فاز دوم:** در این فاز به منظور تطبیق داده های مربوط به سفارشات تأمین با مقادیر واقعی تأمین همان طور که بیان گردید، کدنویسی در محیط MySQL انجام شده و هر سابقه تأمین به سابقه برنامه سفارش تخصیص یافته است. پس از الحاق این صفت به مجموعه داده برای حذف تأثیر تعداد قطعه سفارش داده شده، نیاز به هنجارسازی صفت "مقدار تخصیص یافته" می باشد؛ که این امر از طریق ایجاد صفت الحاق شده درصد تحقق تأمین میسر گردیده است. پس از هنجارسازی، به منظور ارزش گذاری متغیر هدف، وجود طیف وسیع اعداد این فیلد نیاز به افراز مقادیر در دسته هایی مطابق جدول ۳ و نمودار ۱ گردیده است.

نمودار ۱: وضعیت درصد تحقق تأمین سفارش (حالات A, B و C)

جدول ۳: وضعیت درصد تحقق تأمین سفارش (حالات A, B و C)



وضعیت درصد تحقق سفارش	وضعیت درصد تحقق سفارش
A = درصد تحقق سفارش کمتر از ۷۰ درصد	وضعیت درصد تحقق سفارش
B = درصد تحقق سفارش بین ۷۰ درصد تا ۱۵۰ درصد	
C = درصد تحقق سفارش بالاتر از ۱۵۰ درصد	

۳ ۴ ۲. **بررسی نیاز به کاهش ابعاد داده:** مشکانی و ناظمی (۱۳۸۸) معتقدند با وجود اینکه مجموعه داده های حجیم استعداد استخراج اطلاعات بهتری دارند، اما هیچ تضمینی وجود ندارد که داده های حجیم، دانش بهتری از مجموعه داده های کوچک فراهم کنند. بنابراین موضوع اصلی در ساده کردن داده ها در این گام، کاهش بُعد است و سؤال عمده اینجاست که آیا برخی داده های پیش پردازش شده را می توان بدون فدا کردن کیفیت کنار گذاشت؟

در عمل لازم است که تناسبی بین تعداد صفات و تعداد رکوردهای در اختیار برقرار باشد. در غیر این صورت به عنوان مثال با داشتن ۱۰۰ بعد و تنها ۲۰۰ یا ۳۰۰ نمونه، داده های با بار زیاد به علت بعد بالا، ممکن است سبب غیر قابل اعمال شدن برخی الگوریتم های داده کاوی گردند. سه بعد عمده مجموعه داده های پیش پردازش شده که معمولاً به شکل مسطح نمایش داده می شوند، عبارتند از: سطرها یا نمونه ها، ستون ها یا صفات، مقادیر صفات.

همچنین به منظور اعمال کارکرد دسته بندی، متغیر هدف "تحقق سفارش تأمین" از حالت پیوسته به گسسته تبدیل گردیده است. برای این کار از طریق نرم افزار توزیع متغیرها بررسی و با استفاده از خوشه بندی های متعدد، نحوه رتبه بندی مناسبتر انتخاب گردیده است. رتبه بندی نمودن صفت در مورد سایر فیلدها از قبیل قیمت فرآیند، وزن قطعه، آدرس تأمین کننده، سابقه همکاری تأمین کننده با سازمان، سابقه کار مدیریت عامل شرکت تأمین کننده و ... نیز اعمال گردیده است.

۳-۴. مرحله مدل سازی

با توجه به شناخت حاصل شده در مرحله درک کسب و کار و نظرات افراد خبره سازمان، صفات ورودی به مدل انتخاب شدند؛ انتخاب دستی فیلدهای ورودی در مورد صفاتی صورت گرفته که مؤثر بر هر یک از متغیرهای هدف به نظر می رسیدند. انتخاب خودکار (حذف خودکار) صفات نیز توسط نرم افزار صورت گرفته است. مطابق هدف تعیین شده، وضعیت درصد تأمین به عنوان خروجی مدل انتخاب گردید.

در ادامه با استفاده از Record operations / Partition node مجموعه داده در دسترس به دو دسته آموزشی و آزمایشی افزای می گردد. در این پژوهش درصد داده های آموزشی به آزمایشی ۸۰ به ۲۰ انتخاب گردیده که بر اساس آن تعداد ۸۳۸۹ رکورد در مجموعه آموزشی و تعداد ۲۰۸۹ رکورد در مجموعه آزمایشی قرار می گیرند. الگوریتم های مختلف دسته بندی از جمله C^{۵,۰}، Quest، C&R Tree، و Chaid روی مجموعه داده اعمال شده و در هر مورد با استفاده از Output / Analysis node دقت پیش بینی سنجیده شده است. با توجه به نوع داده ها و هدف تحقیق، بهترین دقت با استفاده از الگوریتم C^{۵,۰} حاصل شده و لذا جهت استفاده در ساخت مدل انتخاب می گردد.

پس از انتخاب الگوریتم مذکور هنوز می توان با تغییر پارامترهای مؤثر در این الگوریتم دقت خروجی دسته بندی را افزایش داده و مدل بهتری ایجاد نمود. پارامترهای مؤثر در این مدل عبارتند از:

- **درجه هرس نمودن درخت تصمیم:** با افزایش این پارامتر دقت مدل کاهش و عمومیت آن افزایش می یابد. عمل هرس نمودن درخت تصمیم ساخته شده در اصطلاح همان هرس کردن درخت است که از ایجاد شاخه هایی که افزایش چندانی در دقت مدل ندارند، جلوگیری می نماید.
- **کمترین تعداد رکورد در هر شاخه فرزند:** پارامتر دیگر این مدل همان گونه که از نامش پیداست، کمینه تعداد فرزند در شاخه والد را تعیین می کند. لذا بدیهی است با کاهش این پارامتر، درخت رشد بیشتری نموده، دقت افزایش و عمومیت مدل تا حدودی کاهش می یابد.

با این توضیحات در ابتدا عددی را به عنوان پارامتر اولیه وارد نموده و بعد از اجرای الگوریتم، ساخت مدل آغاز می گردد. پس از آن Field Operations / Analysis را به جریان^۳ وارد نموده و مدل مورد ارزیابی قرار می گیرد. با تغییر پارامترهای مؤثر بر دقت پیش بینی، مدل های ذیل خروجی گردیده که در جدول ۴ آورده شده است:

۱ . Pruning Severity

۲ . Minimum records per child branch

۳ . Stream

جدول ۴: دقت مدل دسته بندی ساخته شده با تغییر پارامترهای مؤثر (حالات A, B و C)

حالات مختلف	درجه هرس نمودن درخت تصمیم	کمترین تعداد رکورد در هر شاخه فرزند	دقت در مجموعه داده آموزشی	دقت در مجموعه داده آزمایشی
۱	۱۰۰	۵۰	۵۷,۵۹	۵۸,۱۱
۲	۳۰	۱۰	۷۹,۵۳	۷۲,۱۹
۳	۳۵	۱۰	۷۹,۵۳	۷۲,۱۹
۴	۴۵	۱۰	۷۹,۵۳	۷۲,۱۹
۵	۵۵	۱۰	۷۹,۵۳	۷۲,۱۹
۶	۶۵	۱۰	۷۹,۵۳	۷۲,۱۹
۷	۷۰	۱۰	۷۸,۸۲	۷۱,۹
۸	۶۵	۲۰	۷۳,۳۶	۶۹,۵۵
۹	۱۰	۲	۹۱,۲۵	۷۷,۷۹

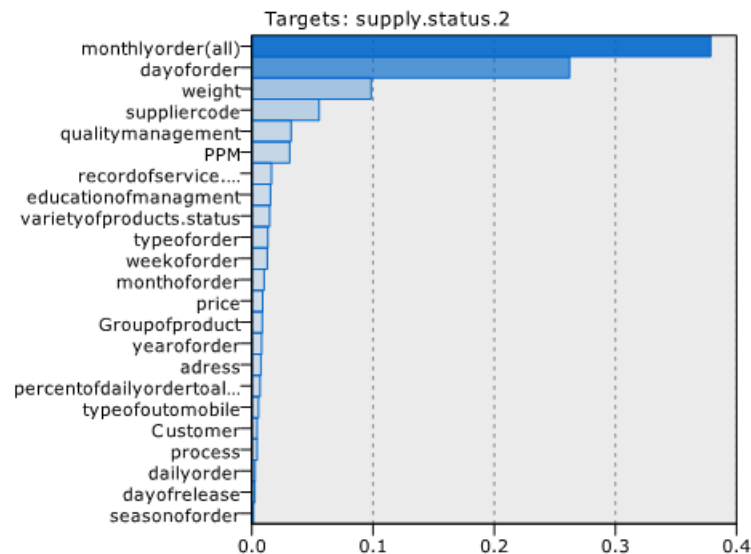
۳. یافته ها

۴-۱. خروجی مدل

پس از انتخاب ردیف ۶ از جدول ۴، الگوریتم با این پارامترها اجرا شده که نتایج آن در نمودار ۲ و جدول ۵ آورده شده است.

نمودار ۲: ضرایب اهمیت صفات ورودی به مدل

Variable Importance



بران

کنفر

جدول ۵: مشخصات مدل ساخته شده و ضریب اهمیت صفات ورودی

ردیف	صفت ورودی	اهمیت متغیر در تعیین ارزش متغیر هدف	ورود به مدل
۱	میزان سفارش ماهانه	۰,۳۷۹	*
۲	روز سفارش	۰,۲۶۲	*
۳	نام تأمین کننده	۰,۱۴۸	*
۴	وزن	۰,۰۹۹	*
۵	کد تأمین کننده	۰,۰۵۶	*
۶	دارا بودن گواهی نامه کیفی	۰,۰۳۲	*
۷	وضعیت کیفی قطعه (PPM)	۰,۰۳۱	*
۸	سابقه کار مدیریت عامل	۰,۰۱۶	*
۹	تحصیلات مدیریت عامل	۰,۰۱۵	*
۱۰	تنوع تعداد کد نزد تأمین کننده	۰,۰۱۵	*
۱۱	نوع سفارش	۰,۰۱۳	*
۱۲	هفته سفارش	۰,۰۱۳	*
۱۳	ماه سفارش	۰,۰۱۰	*
۱۴	قیمت	۰,۰۰۹	*
۱۵	گروه قطعه نزد مشتری	۰,۰۰۹	*
۱۶	سال سفارش	۰,۰۰۸	*
۱۷	آدرس	۰,۰۰۷	*
۱۸	نسبت سفارش به کل سفارش ماهانه	۰,۰۰۷	*
۱۹	نوع خودرو	۰,۰۰۵	*
۲۰	مشتری	۰,۰۰۴	*
۲۱	فرآیند	۰,۰۰۴	*
۲۲	سفارش روزانه	۰,۰۰۳	*
۲۳	روز صدور سفارش	۰,۰۰۲	*
۲۴	فصل سفارش	۰,۰۰۱	*
۲۵	کد انبار		

متغیر هدف: وضعیت تحقق تولید تأمین کننده
 حالات ممکن برای متغیر هدف: A, B و C
 الگوریتم مورد استفاده: C5.0
 پارامترهای الگوریتم:

Pruning severity = ۶۵

Minimom records per child branch = ۱۰

درصد داده های آموزشی به آزمایشی: ۸۰ به ۲۰
 تعداد رکورد مجموعه آموزش: ۸۳۸۹
 تعداد رکورد مجموعه آزمایش: ۲۰۸۹

۲۶	نام قطعه	
۲۷	مشتری	
۲۸	تیم نیروی انسانی مؤثر	
۲۹	جنسیت مدیریت عامل	

همان طور که در جدول ۵ مشاهده می شود، اولین نتیجه قابل تعمق از الگوی استخراج شده، شناسایی صفات با اهمیت در ورود به مدل و تحلیل ضرایب به دست آمده برای اهمیت این صفات است. صفات بدون اهمیت که در سطرهای آخر جدول آمده، به دلیل عدم تأثیر در ارزش گذاری متغیر هدف از ورود به مدل حذف گردیده و صفت میزان سفارش ماهانه و روز سفارش دهی دارای بیشترین تأثیر در این ارزش گذاری تعیین شده است. خروجی اصلی مدل دسته بندی ساخته شده، درخت تصمیمی با ۸ لایه می باشد.

۴-۲. مرحله ارزیابی الگوها

جهت ارزیابی کیفیت مدل خروج شده از نرم افزار، از ماتریس انطباقی ارزیابی استفاده شده که در جداول ۶ و ۷ آورده شده است. با توجه به اینکه هزینه پیش بینی نادرست برای دسته های تعیین شده از نظر مفهومی برای این پژوهش یکسان بوده است، در ماتریس هزینه خطا کلیه هزینه ها یک در نظر گرفته شده است.

جدول ۶: ماتریس انطباقی پیش بینی کلاس رکوردها براساس متغیر وضعیت تحقق سفارش تأمین کننده / مجموعه داده آموزشی

کلاس واقعی	A = درصد تحقق سفارش کمتر از ۷۰ درصد	B = درصد تحقق سفارش بین ۷۰ تا ۱۵۰ درصد	C = درصد تحقق سفارش بالاتر از ۱۵۰ درصد
کلاس پیش بینی شده			
A = درصد تحقق سفارش کمتر از ۷۰ درصد	۳۴۵۲	۶۸۸	۲۹
B = درصد تحقق سفارش بین ۷۰ تا ۱۵۰ درصد	۶۴۶	۳۱۰۰	۳۳
C = درصد تحقق سفارش بالاتر از ۱۵۰ درصد	۱۱۵	۲۰۶	۱۲۰

جدول ۷: ماتریس انطباقی پیش بینی کلاس رکوردها براساس متغیر وضعیت تحقق سفارش تأمین کننده / مجموعه داده آزمایشی

کلاس واقعی	A = درصد تحقق سفارش کمتر از ۷۰ درصد	B = درصد تحقق سفارش بین ۷۰ تا ۱۵۰ درصد	C = درصد تحقق سفارش بالاتر از ۱۵۰ درصد
کلاس پیش بینی شده			
A = درصد تحقق سفارش کمتر از ۷۰ درصد	۷۷۲	۲۲۱	۸
B = درصد تحقق سفارش بین ۷۰ تا ۱۵۰ درصد	۲۵۲	۷۲۰	۱۵
C = درصد تحقق سفارش بالاتر از ۱۵۰ درصد	۳۰	۵۵	۱۶

بررسی جداول ارزیابی بالا منتج به استخراج جدول ۸ گردیده که خلاصه وضعیت پیش بینی مدل را نشان می دهد.

جدول ۸: دقت پیش بینی مدل ساخته شده برای مجموعه داده آموزشی و آزمایشی

مجموعه آموزشی		مجموعه آزمایشی		درست
تعداد کل	درست	تعداد کل	درست	
۱۷۱۷	۶۶۷۲	۱،۵۰۸	۷۲،۱۹ درصد	
۸،۳۸۹	۲۰،۴۷ درصد	۵۸۱	۲۷،۸۱ درصد	
		۲،۰۸۹		

ملاحظه می گردد که دقت مدل در پیش بینی مجموعه داده آزمایشی ۷۹،۵۳ و در مجموعه داده جدید آزمایشی ۷۲،۱۹ درصد می باشد.

۴. بحث و نتیجه گیری

همان طور که از خروجی مدل بر می آید، یکی از مهم ترین عوامل مؤثر بر متغیر هدف وضعیت تحقق تأمین، روز سفارش شناسایی شده است. خروجی درخت تصمیم نشان می دهد برنامه سفارشات که در روزهای مختلف ماه صادر شده اند، برای تأمین کنندگان مختلف دارای وضعیت تحقق تأمین کاملاً متفاوتی بوده اند. با توجه به این موضوع در برنامه ریزی سفارشات آتی تأمین کنندگان سازمان می توان سفارشات صادره را به روزهای هر ماه به نحوی تخصیص داد که به هر تأمین کننده در روزهای بهینه خود برنامه سفارش تعلق گیرد.

همچنین خروجی مدل بیانگر تأثیر قابل توجه میزان سفارش ماهیانه بر وضعیت تحقق سفارش است؛ با توجه به اینکه این صفت تا حدودی تابع وزن قطعه می باشد، درجه اهمیت به دست آمده برای فیلد وزن قطعه نیز مؤید همین مطلب می باشد. دانش به دست آمده در این زمینه را می توان در تصمیم سازی و تصمیم گیری برنامه های لجستیکی چه در مبحث حمل و نقل قطعات و چه در مبحث تعیین اندازه بچ های سفارش دهی به کار گرفت.

همچنین یکی دیگر از صفاتی که در این تحقیق بر وضعیت تحقق تأمین مؤثر شناخته شد، دارا بودن گواهی نامه کیفی تأمین کننده است. با توجه به بررسی صفت مذکور در بین کلیه تأمین کنندگان و نظر به اینکه عموم آنها از این منظر دارای نقطه ضعف می باشند، به نظر می رسد سازمان باید برنامه های ارتقاء تأمین کنندگان خود را در صدر برنامه های ارتباط با تأمین کنندگان قرار دهد. سابقه همکاری بالای تأمین کننده با سازمان مورد نظر در این تحقیق نیز از جمله صفاتی بوده است که روی وضعیت تحقق تأمین آنها تأثیر مثبت داشته است.

منابع و مأخذ

- ۱) براون، جیمی، هارن، جان، شیونان، جیمز، (۱۳۸۳)، "سیستم های مدیریت تولید با نگرشی یکپارچه"، ترجمه: غضنفری، مهدی، صغیری، سروش، چاپ سوم، انتشارات دانشگاه علم و صنعت ایران.
- ۲) جعفرزاد، احمد، (۱۳۸۵)، "مدیریت تولید و عملیات نوین"، چاپ اول، انتشارات دانشکده مدیریت دانشگاه تهران.
- ۳) شهرابی، جمال، (۱۳۸۶)، "داده کاوی"، چاپ اول، موسسه پژوهشی داده پردازان گیتا، انتشارات جهاد دانشگاهی واحد صنعتی امیرکبیر.
- ۴) شهرابی، جمال، ذوالقدر شجاعی، علی، (۱۳۸۸)، "داده کاوی پیشرفته، مفاهیم و الگوریتم ها"، چاپ اول، انتشارات جهاد دانشگاهی واحد صنعتی امیرکبیر.
- ۵) شهرابی، جمال، شکورنیا، ونوس، (۱۳۸۸)، "داده کاوی کاربردی با مثال هایی در SQL Server"، چاپ اول، موسسه پژوهشی داده پردازان گیتا، انتشارات جهاد دانشگاهی واحد صنعتی امیرکبیر.
- ۶) کانتاردزیک، مهمد، (۱۳۸۵)، "داده کاوی"، مترجم: علیخانزاده، امیر، چاپ اول، نشر علوم رایانه.
- ۷) مشکانی، علی، ناظمی، عبدالرضا، (۱۳۸۸)، "مقدمه ای بر داده کاوی"، چاپ اول، نشر دانشگاه آزاد اسلامی واحد نیشابور.
- ۸) اسلایدهای دوره آموزشی آقای دکتر نخعی زاده، (۱۳۸۶)، مشهد.

- ۹) Bacao, F, (۲۰۰۸), "Data mining and knowledge discovery technologies", IGI publishing.
- ۱۰) Berry, M, Linoff, G, (۲۰۰۰), "Mastering data mining: The art and relationship of customer relationship management", John Wiley and sons.
- ۱۱) Chase, R, Jacobs, F, Aquilano, A, (۲۰۰۴), "Operation Management for competitive advantage", McGraw Hill Publication, ۱۰th edition.
- ۱۲) Dong Li, G, Yamaguchi, D, Nagai, M, (۲۰۰۷), "Application of grey-based rough decision-making approach to supplier selection", Journal of modeling in management, Vol ۲/۲, P ۱۳۱-۱۴۲.

- ۱۳) Gargano, M, Raggad, B, (۱۹۹۹), "Data mining, a powerful information creating tool", OCLC systems and services journal, Vol ۱۰/۲, P ۸۱-۹۰.
- ۱۴) Han, J, Kamber, M, (۲۰۰۶), "Data mining, concepts & techniques", ۲nd edition, Amsterdam, Morgan Kaufman publication.
- ۱۵) Hang Hong, G, Chan Park, S, Sik Jang, D, Min Rho, H, (۲۰۰۵), "An effective supplier selection method for constructing a competitive supply-relationship", Expert Systems with Applications, Vol ۲۸, P ۶۲۹-۶۳۹.
- ۱۶) Ho Ha, S, Krishnan, R, (۲۰۰۸), "A hybrid approach to supplier selection for the maintenance of a competitive supply chain", Expert systems with application, Vol ۳۴, P ۱۳۰۳-۱۳۱۱.
- ۱۷) Ho Lin, R, Chuang, C, Liou, J.H, Wu, G, (۲۰۰۹), "An integrated method for finding key suppliers in SCM", Expert systems with application, Vol ۳۶, P ۶۴۶۱-۶۴۶۵.
- ۱۸) Kanan, V, Tan, K, (۲۰۰۶), "Buyer-supplier relationships, The impact of supplier selection and buyer-supplier engagement on relationship and firm performance", International journal of physical distribution & logistics management, Vol ۳۶/ ۱۰, P ۷۵۵-۷۷۵.
- ۱۹) Larose, D, (۲۰۰۵), "Discovering Knowledge in data mining, An Introduction to Data Mining", John Wiley & sons publication.
- ۲۰) Li, T, Ruan, D, (۲۰۰۷), "An extended process model of knowledge discovery in database", Journal of enterprise information management, Vol ۲۰/۲, P ۱۶۹-۱۷۷.
- ۲۱) Monczka, R. M, Trent, R.J, Handfield, R. B, (۱۹۹۸), "Purchasing and supply chain management", Ohio, USA, South western college publishing co.
- ۲۲) Nie, G, Zhang, L, Liu, Y, Zheng, X, Shi, Y, (۲۰۰۹), "Decision analysis of data mining project based on Bayesian risk", Expert systems with applications, Vol ۳۶, P ۴۵۸۹-۴۵۹۴.
- ۲۳) Olson, D. L, Delen, D, (۲۰۰۸), "Advanced data mining techniques", Springer publication.
- ۲۴) Ramanathan, R, (۲۰۰۷), "Supplier selection problem: integrating DEA with the approaches of total cost of ownership and AHP", Supply chain management international journal, Vol ۱۲/۴, P ۲۵۸-۲۶۱.
- ۲۵) Vollmann, T, Berry, W, Whybark, D, Jacobs, F, (۲۰۰۵), "Manufacturing Planning & control for supply chain management", McGraw Hill Publication, ۵th edition.
- ۲۶) Wang, H, Wang, S, (۲۰۰۸), "A knowledge management approach to data mining process for business intelligence", Industrial management and data systems, Vol ۱۰۸/۵, P ۶۲۲-۶۳۴.
- ۲۷) Wang, J, (۲۰۰۳), "Data mining, Opportunities and challenges", IRM Press publication, ۱st edition.
- ۲۸) Wu, D, (۲۰۰۹), "Supplier selection: A hybrid model using DEA, decision tree and neural network", Expert Systems with Applications, Vol ۳۶, P ۹۱۰۵-۹۱۱۲.
- ۲۹) Yau K, He, L, Amy H, (۲۰۱۰), "A new supplier performance evaluation model, A case study of integrated circuit (IC) packaging companies", Kybernetes, Vol ۳۹/۱, P ۳۷-۵۴.