

Identify the most important factors affecting the value added of industrial workshops with ten employees and more in Kermanshah province using data mining methods; during the years 1381 to 1397

Salah Ghorbani

Master of Statistics of Razi University, Kermanshah, Iran

salah_ghorbani@yahoo.com

Abstract

Value added is considered as one of the criteria for evaluating the performance of economic units. Statistics related to various economic indicators of industrial workshops in the provinces of the country, along with their value added, are collected every year in the form of "statistics of industrial workshops" by the Statistics Center of Iran. The present study for the first time uses one of the data mining and machine learning methods, as support vector machine (SVM) to find the most important factors affecting the added value of these factories.

The results of this study show that in the period of 1381 to 1397, the output value of industrial activities (9.9%), raw materials (9.7%), inventory value (7.8%), People with less than a diploma in manufacturing (7.03%), payment of non-industrial services (6.9%) and engineers of manufacturing sector (6.3%), respectively, in the highest degree of importance to create There are added value industrial factories.

Keyword: Value Added, Data Mining, Support Vector Machine, Industrial Workshopos

شناسایی مهم ترین عوامل تأثیر گذار بر ارزش افزوده کارگاه های صنعتی با ده نفر کارکن و بیش تر استان کرمانشاه با استفاده روش های داده کاوی

در طول سال های ۱۳۸۱ تا ۱۳۹۷

صلاح قربانی

کارشناسی ارشد آمار، دانشگاه رازی، کرمانشاه، ایران

salah_ghorbani@yahoo.com

چکیده

ارزش افزوده، به عنوان یکی از معیارهای ارزیابی عملکرد واحدهای اقتصادی محسوب می شود. آمارهای مربوط به شاخص های مختلف اقتصادی کارگاه های صنعتی استان های کشور به همراه ارزش افزوده آنها، هر سال در قالب طرح "آمارگیری از کارگاه های صنعتی" توسط مرکز آمار ایران جمع آوری می شود. مطالعه حاضر برای اولین بار، با استفاده از یکی از روش های داده کاوی و یادگیری ماشین، با عنوان ماشین بردار پشتیبان (SVM) در پی شناسایی مهم ترین عوامل مؤثر بر ارزش افزوده کارگاه های صنعتی دارای ده نفر کارکن و بیشتر در استان کرمانشاه است. فهرست این کارگاه ها توسط مرکز آمار ایران در قالب طرح "آمارگیری از کارگاه های صنعتی"، تعیین شده است. تجزیه و تحلیل داده ها با استفاده از نرم افزار آماری R انجام شد. نتایج این تحقیق نشان می دهد که در بازه زمانی ۱۳۸۱ تا ۱۳۹۷ ارزش ستانده فعالیت های صنعتی (۹/۹ درصد)، مواد خام و اولیه (۹/۷ درصد)، ارزش موجودی انبار (۷/۸ درصد)، افراد کمتر از دیپلم بخش شاغلان تولیدی (۷/۰۳ درصد)، پرداختی خدمات غیرصنعتی (۶/۹ درصد) و مهندسین بخش شاغلان تولیدی (۶/۳ درصد) به ترتیب در بالاترین درجه اهمیت برای ایجاد ارزش افزوده کارگاه های صنعتی دارای ده نفر کارکن و بیشتر در استان کرمانشاه قرار دارند.

کلید واژه ها: ارزش افزوده، داده کاوی، ماشین بردار پشتیبان، کارگاه های صنعتی

۱- مقدمه:

ارزش افزوده را می‌توان ثروت تولید شده توسط بنگاه‌های فعال اقتصادی یک کشور در نظر گرفت. یکی از مهم‌ترین تفاوت‌های ارزش افزوده با درآمد بنگاه‌ها، هزینه‌های مربوط به خرید مواد اولیه و خدمات دریافتی است که از آن با عنوان داده یاد می‌شود. ارزش افزوده، معیاری برای اندازه‌گیری مازاد ارزشی است که بنگاه بر روی مواد و خدمات اولیه مصرفی ایجاد می‌کند. این ارزش اضافی تولید شده به صورت دستمزد، بهره، مالیات و سود سهام در میان عوامل تولید تقسیم شده و بخشی از آن نیز صرف سرمایه‌گذاری مجدد در مسیر تولید می‌شود.

سرمایه‌گذاران همواره نیازمند اطلاعاتی هستند تا بتوانند از آن در تصمیمات خود برای سرمایه‌گذاری به طور مناسبی استفاده کنند. سهامداران به عنوان مالکان واحدهای اقتصادی در پی افزایش سود خود هستند و با توجه به اینکه افزایش سود، نتیجه عملکرد مطلوب واحدهای اقتصادی می‌باشد ارزیابی این واحدها برای سرمایه‌گذاران دارای اهمیت فراوانی است. تاکنون برای ارزیابی عملکرد واحدهای اقتصادی معیارهای مختلفی نظیر سود عملیاتی و سود هر سهم رایج شده است. ارزش افزوده جدیدترین معیار در این رابطه است (استوارت^۱، ۱۹۹۱).

از این رو شناسایی عوامل مؤثر بر ارزش افزوده، یکی از مهم‌ترین راهکارهای بررسی ایجاد شرایط مناسب در راستای جهش اقتصادی بشمار می‌رود. شناسایی این عوامل، علاوه بر فراهم کردن چشم انداز بلندمدت برای سیاست‌گذاران اقتصادی، سرمایه‌گذاران و مالکان بنگاه‌های خصوصی کشور را نیز در یافتن عوامل مهم و تأثیرگذار بر ارزش افزوده یاری خواهد کرد. مرکز آمار ایران در قالب طرح "آمارگیری از کارگاه‌های صنعتی"، همه ساله شاخص‌های مختلف اقتصادی کارگاه‌های صنعتی استان‌های کشور را جمع‌آوری و منتشر می‌کند. مطالعه حاضر با استفاده از اطلاعات جمع‌آوری شده از طرح مذکور، در پی شناسایی مهم‌ترین عوامل مؤثر بر ارزش افزوده کارگاه‌های صنعتی با ده نفر کارکن و بیشتر است.

در بسیاری از مسایل عملی، محقق با تعداد زیادی از متغیرهای توضیحی مواجه است که بنا به ضرورت باید تنها تعداد محدودی از آنها را برای حضور در مدل انتخاب کند. زیرا اگر مدل، حاوی متغیرهای زاید و بی‌اهمیت باشد، یا به اصطلاح دچار بیش برآوردی شود، علاوه بر آنکه تعبیر و تفسیر چنین مدلی مشکل است، جمع‌آوری مشاهدات نیز برای چنین مدلی مستلزم صرف وقت و هزینه زیاد می‌باشد. بر عکس، اگر متغیرهای مهم از مدل کنار گذاشته شوند، یا به اصطلاح مدل دچار کم برآوردی شود، در این صورت ممکن است داده‌های موجود به خوبی توصیف نشوند. با توجه به مطالب فوق، هدف تحقیق حاضر را می‌توان به صورت زیر بیان کرد: "شناسایی مهم‌ترین عوامل مؤثر بر ارزش افزوده کارگاه‌های صنعتی با ده نفر کارکن و بیشتر"

۲- پیشینه تحقیق:

گسترده‌گی کاربرد ارزش افزوده سبب شده است تا این مفهوم، طیف وسیعی از مطالعات انجام شده در بخش‌های مختلف اقتصادی و سیاست‌گذاری‌های مرتبط با آنها را به خود اختصاص دهد. بسیاری از این مطالعات، پیرامون عوامل مؤثر بر ارزش افزوده بخش صنعت و معدن است. نادران (۱۳۸۳)، با استفاده از روش مدل‌سازی، 2VAR به برآورد اثر سیاست‌های اعتباری دولت بر ارزش افزوده بخش صنعت پرداخت. نتایج به دست آمده نشان می‌دهد که مقدار ارتباط بلند مدت ارزش افزوده با اعتبارات بانکی ۰/۲۹، با مخارج دولت ۰/۱۶-، با مالیات ۰/۱۸ و با شوک تکنولوژیکی ۰/۲۰ است

ابزری و همکاران (۱۳۸۳)، به بررسی تأثیر نیروی کار بر ارزش افزوده کارخانجات سیمان پرداختند. آنها سهم هزینه سرمایه‌ای در ارزش افزوده کارخانجات سیمان را معادل ۵/۲، سهم هزینه‌های تسهیلات اعتباری را معادل ۲/۵ و سهم نیروی کار را برابر ۴۵/۷ به دست آوردند. همچنین عمادزاده و بکتاش (۱۳۸۴)، اثر نیروی انسانی را بر ارزش افزوده بخش صنعت با استفاده از داده‌های سری زمانی، در فاصله ۱۳۴۵ تا ۱۳۸۰ مورد بررسی قرار دادند. نتایج به دست آمده از این تحقیق نشان می‌دهد که یک درصد افزایش در سرمایه فیزیکی، نیروی کار متخصص و غیرمتخصص، به ترتیب سبب افزایش، ۳۴، ۲۹ و ۱۱ درصدی در ارزش افزوده بخش صنعت می‌شود.

ماسون^۱ با بررسی بسیاری از شرکت‌های کشور انگلستان از چهار بخش مختلف شامل صنعت چاپ، پلاستیک، بیمه و پخش و توزیع، نشان داد که، مهارت کارکنان، پیش‌نیاز دستیابی به ارزش افزوده بالاست. با ذکر این نکته که مهارت کارکنان در این چهاربخش به صورت‌های مختلف بروز می‌کند، مثلاً در صنعت پلاستیک، این مهارت شامل نوآوری‌های مداوم و متناسب با نیاز مشتری است.

در سایر بخش‌های اقتصادی نیز مطالعات مشابهی انجام گرفته است، که از میان آنها می‌توان به مطالعه اکبری و همکاران (۱۳۸۲)، اشاره کرد، که در آن تأثیر هزینه‌های دولت بر ارزش افزوده بخش کشاورزی بررسی شده است.

آنها که در تحقیق خود از سیستم مطالعات همزمان استفاده کردند به این نتیجه رسیدند که مؤثرترین متغیرهای تأثیرگذار بر ارزش افزوده بخش کشاورزی به ترتیب عبارتند از: هزینه‌های تحقیقاتی، آموزشی و هزینه‌های عمرانی دولت.

فان و همکاران^۴ (۲۰۰۲)، اثرات مخارج دولت چین را بر تولیدات کشاورزی بررسی کردند. بر اساس این تحقیق نیز سرمایه‌گذاری دولت در تحقیقات و توسعه کشاورزی، بیشترین تأثیر را بر روی رشد تولیدات

کشاورزی داشته است. آراجی و وایت^۱، تأثیر تحقیقات کشاورزی بر روی تولید محصولات کشاورزی ایالات متحده در فاصله سال‌های ۱۹۵۰ تا ۱۹۸۸ را بررسی کردند.

حسینی‌نسب و همکاران (۱۳۸۹)، عوامل مؤثر بر ارزش‌افزوده کارگاه‌های صنعتی را در طول سال‌های ۱۳۷۶ تا ۱۳۸۵ بررسی کردند که نتایج آن شناسایی عوامل میزان تحصيلات نیروی کار، نوع مدیریت کارگاه‌ها، پرداختی بابت خدمات حسابرسي، هزینه تحقیقات، هزینه تبلیغات و ارزش زمین بود. اسماعیلی و رحمتی (۱۳۸۷)، اثر آزادسازی تجاری بر ارزش افزوده بخش کشاورزی را مورد مطالعه قرار دادند. آنها در تحقیق خود از شاخص‌هایی نظیر شاخص ادغام تجاری و شاخص سطح تجارت جهانی استفاده کردند و با توجه به ضرایب به دست آمده نتیجه گرفتند که ایران در بر همکشی مثبت با اقتصاد جهانی قرار دارد. بیضایی (۱۳۸۴)، به بررسی عوامل مؤثر بر ارزش افزوده بخش حمل و نقل پرداخت. نتیجه کار وی نشان می‌دهد که ارزش افزوده بخش حمل و نقل تا حد زیادی به درآمدهای نفتی وابسته است.

آل عمران و همکاران (۱۳۹۹)، به بررسی عوامل اقتصادی مؤثر بر ارزش افزوده بخش کشاورزی ایران پرداختند که نتایج به دست آمده از پژوهش آنها دلالت بر این دارد که اثرگذاری ضرایب متغیرهای تسهیلات پرداختی بانک کشاورزی، باز بودن تجاری و موجودی سرمایه خالص بخش کشاورزی براساس مبانی نظری مورد انتظار بوده و از نظر آماری نیز معنی‌دار بوده‌اند.

۳- مقدمه‌ای بر داده‌کاوی:

داده‌کاوی یکی از پیشرفت‌های اخیر در راستای فناوری‌های مدیریت داده‌هاست. داده‌کاوی مجموعه‌ای از فنون است که به شخص امکان می‌دهد تا ورای داده‌پردازی معمولی حرکت کند و به استخراج اطلاعاتی که در انبوه داده‌ها پنهان است، کمک می‌کند. تاریخچه کشف دانش از پایگاه‌های اطلاعاتی قدمت چندانی ندارد و امروزه به داده‌کاوی مشهور است. اصطلاح کشف دانش برای نخستین بار در دهه ۱۹۹۰ توسط پیاتسکی و همکارانش^۲ مطرح شد و توجه پژوهشگران را به سمت الگوریتم‌های داده‌کاوی معطوف کرد. هدف داده‌کاوی، کشف دانش جدید، معتبر و قابل پیگیری با استفاده از ابزارهای هوش مصنوعی و آماری در حجم بالایی از داده‌ها است.

داده‌کاوی یکی از زمینه‌های مهم و مؤثر از علوم کامپیوتر است. این زمینه در اواخر دهه ۸۰ میلادی با استفاده از مفاهیم و روش‌های هوش مصنوعی، شناسایی الگو، آمار و سیستم‌های پایگاه داده به وجود آمد و هدف آن کشف و تشخیص اطلاعات مخفی، غیرواضح و پیچیده در میان مقدار زیادی از داده‌ها است. به همین دلیل یک نام معادل برای داده‌کاوی، اکتشاف دانش در پایگاه داده (Kdd)^۳ است که به صورت معمول در ادبیات به کار می‌رود.

1. Arajji, and White

2. Piatetsky-Shapiro et al

3. Knowledge Discovery in Databases

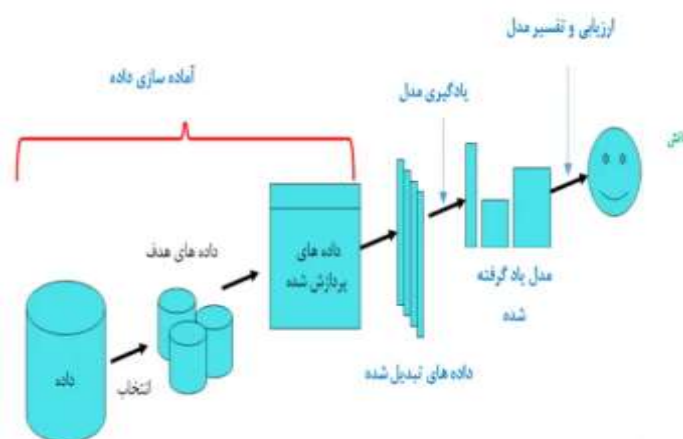
در داده کاوی از بخشی از علم آمار به نام تحلیل اکتشافی داده‌ها استفاده می‌شود که در آن بر کشف اطلاعات نهفته و ناشناخته از درون حجم انبوه داده‌ها تأکید می‌شود. علاوه بر این داده کاوی با هوش مصنوعی و یادگیری ماشین نیز ارتباط تنگاتنگی دارد، بنابراین می‌توان گفت در داده کاوی تئوری‌های پایگاه داده‌ها، هوش مصنوعی، یادگیری ماشین و علم آمار را در هم می‌آمیزند تا زمینه کاربردی فراهم شود.

هر چه حجم داده‌ها بیشتر و روابط میان آنها پیچیده‌تر باشد دسترسی به اطلاعات نهفته در میان داده‌ها مشکل‌تر می‌شود و نقش داده کاوی به‌عنوان یکی از روش‌های کشف دانش، روشن‌تر می‌گردد.

هدف داده کاوی جستجو برای یافتن اطلاعاتی با ارزش تجاری در یک پایگاه داده است. به‌طور کلی، از داده کاوی به‌منظور کشف اطلاعات نهفته در داده‌ها استفاده می‌شود. علم داده کاوی در سلامت عمومی، تحقیقات بازار خرید مشتریان، آموزش، ساخت و عمران، مدیریت ارتباطات مشتریان، هواشناسی و بسیاری از حوزه‌های دیگر کاربرد دارد.

۴- مراحل (فرآیند) داده کاوی:

با نگاهی اجمالی به مراحل داده کاوی، می‌توان نحوه انجام آن را به‌صورت نمودار زیر بیان کرد:



شکل ۱. مراحل داده کاوی

۴-۱- آماده‌سازی داده‌ها

مرحله اول داده کاوی، آماده‌سازی داده‌هاست که در آن اقداماتی انجام می‌شود. در این مرحله سلسله فرایندهایی صورت می‌پذیرد که باعث برطرف شدن مشکلات مختلف داده مورد بررسی، خواهد شد. به این ترتیب داده برای انجام فرایند یادگیری مدل، پالایش شده و آماده می‌شود.

۴-۲- یادگیری مدل

پس از آماده‌سازی داده‌ها در فرایند داده‌کاوی، داده آماده اعمال به مرحله یادگیری مدل است. در مرحله یادگیری مدل، نظم حاکم بر داده‌های پیش‌پردازش شده، با توجه به روش کاوش داده‌ای که انتخاب می‌شود، شناسایی شده و مدل تولید شده برای ارزیابی به مرحله بعد یعنی ارزیابی و تفسیر مدل منتقل خواهد شد. روش‌های مختلفی برای یادگیری مدل وجود دارد که در این مقاله از روش ماشین بردار پشتیبان استفاده شده است.

۴-۳- ارزیابی و تفسیر مدل

در این مرحله دانش تولید شده در مرحله قبل ارزیابی شده و مورد تفسیر قرار می‌گیرد. منظور از ارزیابی دانش آن است که می‌بایست میزان صحت دانش تولید شده مشخص شود تا بتوان به آن اعتماد نمود و به صورت عملی از آن استفاده کرد. تفسیر مدل به معنای آن است که دانش تولید شده را مورد بررسی قرار داده و توجیهی معنایی جهت تبیین منطق آن ارائه نماییم.

داده‌کاوی یک فعالیت میان رشته‌ای و ماحصل رویارویی و هم‌افزایی علوم مختلفی چون آمار و کامپیوتر است. در عین حال این محصول مشترک وجوه افتراق مشخصی با پدیدآورندگان خود دارد. مشکانی (۱۳۸۸).

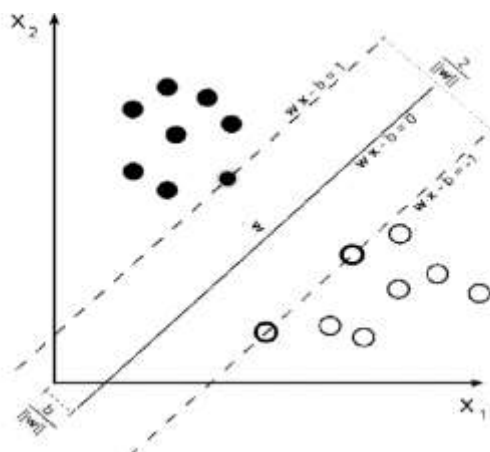
در ادامه به ارائه تعریف و مفاهیم روش ماشین بردار پشتیبان خواهیم پرداخت که با استفاده از آن قصد شناسایی مهم‌ترین عوامل مؤثر بر ارزش افزوده را داریم.

۵- ماشین بردار پشتیبان:

داده‌کاوی و یادگیری ماشین دو علمی هستند که برای توسعه کسب و کارهای کلان امروزی بسیار ضروری هستند. با داده‌کاوی داده‌ها و اطلاعات، می‌توان الگوریتم‌های مناسبی برای یادگیری ماشین در نظر گرفت. در واقع ابتدا باید مراحل داده‌کاوی انجام بگیرد و سپس الگوریتم‌های مناسب روش‌های یادگیری ماشین را استفاده کرد. ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است که هم برای دسته‌بندی و هم رگرسیون قابل استفاده است. ماشین بردار پشتیبان در اصل یک دسته‌بندی کننده دو و یا چند کلاسی است که کلاس‌ها را توسط یک مرز خطی از هم جدا می‌کند. در این روش نزدیکترین نمونه‌ها به مرز تصمیم‌گیری را بردارهای پشتیبان می‌نامند. این بردارها معادله مرز تصمیم‌گیری را مشخص می‌کنند. این روش به دلیل استفاده از اصل کمینه سازی ریسک ساختاری که از طریق بیشینه کردن فاصله بین دو ابر صفحه گذرا از بردارهای پشتیبان هر دو کلاس، اعمال می‌شود، برخلاف حالت کمینه سازی ریسک تجربی که سعی در کمینه کردن خطای آموزش را دارد عملکرد بهتری بر روی داده‌هایی که مدل با آنها ساخته نشده است، از خود نشان می‌دهند. به منظور سادگی در فهم، برای بیان تئوری ماشین بردار پشتیبان از ساده‌ترین حالت ممکن یعنی دسته‌بندی دو کلاسی در حالت جدایی‌پذیر به صورت خطی در شکل زیر نشان داده شده است.

در این مقاله سعی بر آن شده تا با استفاده از الگوریتم ماشین بردار پشتیبان، مدل مناسبی روی داده‌های طرح کارگاه‌های صنعتی مرکز آمار ایران برآزش داده شود و به دنبال آن، متغیرهایی که در مدل برآزش داده شده بیشترین اهمیت در تغییرات ارزش افزوده را داشته‌اند، شناسایی شوند.

شکل ۲. ابر صفحه جدایش و بردارهای پشتیبان در حالت جدایی پذیر به صورت خطی



۶- روش شناسی تحقیق:

یکی از بهترین روش‌های محاسبه عوامل مؤثر بر ارزش افزوده، استفاده از علم داده کاوی است. در این مقاله که برای اولین بار در بحث ارزش افزوده کارگاه‌های صنعتی از این علم استفاده شده، نتایجی حاصل شده که می‌توان آن را پایه و اساس تحقیقات و پژوهش‌های علاقه‌مندان به این موضوع در نظر گرفت. چرا که از علم داده کاوی می‌توان بدون در نظر گرفتن توزیع آماری خاصی برای داده‌ها برای برآورد و ارائه مدل مناسب استفاده کرد. روشی که استفاده شده برای شناسایی عوامل مؤثر بر ارزش افزوده می‌باشد که نحوه محاسبه و دستیابی به مفهوم و مقدار ارزش افزوده کارگاه‌های صنعتی از طرف مرکز آمار ایران در ادامه تشریح خواهد شد.

برای انجام این مطالعه نمونه‌ها توسط مرکز آمار ایران در قالب طرح "آمارگیری از کارگاه‌های صنعتی"، انتخاب و سپس شاخص‌های مختلف اقتصادی کارگاه‌های صنعتی استان را جمع‌آوری و منتشر کرده است.

تجزیه و تحلیل داده‌ها با استفاده از نرم‌افزار آماری R انجام شده است. هم‌چنین مشخصات داده‌های استفاده شده در جدول ۱ آمده است.

۶-۱- مفهوم ارزش افزوده:

ارزش افزوده، اولین بار توسط اقتصاددانان کلاسیک مطرح شد؛ چرا که آنان معتقد بودند که فرایند تولید تحت تأثیر سه عامل کار، سرمایه و زمین بوده و سود به‌دست آمده نیز بین این عوامل تقسیم می‌شود. تأکید بر

خالص دریافتی‌ها در مقابل خالص پرداختی‌ها، اساس مفهوم ارزش افزوده را تشکیل می‌دهد؛ به عبارت دیگر ارزش افزوده را می‌توان خلق مازاد ارزش توسط هریک از عوامل تولید در نظر گرفت. با توجه به گستردگی فعالیت‌های اقتصادی، مشکلات و پیچیدگی‌های زیادی در مسیر محاسبه ارزش افزوده وجود دارد.

یکی از مهم‌ترین این مشکلات، بیش برآوردی مقدار ارزش افزوده بخش‌های مختلف اقتصادی است. امروزه در مباحث مهم مربوط به بنگاه‌های اقتصادی، از صورت‌های مالی مرتبط با ارزش افزوده استفاده‌های زیادی می‌شود. اطلاعات به دست آمده از این صورت‌های مالی، در اختیار افرادی که به نوعی با این بنگاه‌ها در ارتباط هستند، قرار می‌گیرد. بررسی عوامل مؤثر بر ارزش افزوده بنگاه‌های مختلف اقتصادی و شناسایی مهم‌ترین این عوامل، یکی از اساسی‌ترین نیازهای هر اقتصاد پویا محسوب می‌شود. رودپشتی و محمودی (۱۳۸۹).

۶-۲- طرح آمارگیری از کارگاه‌های صنعتی

نقش و اهمیتی که بخش‌های صنعت و معدن در جریان توسعه اقتصادی کشور ایفا می‌کنند و نیز کمبود آمارهای مناسب در این زمینه موجب شد با تأکید بر تأمین اطلاعات مورد نیاز بخش‌های مزبور، طرح سرشماری عمومی صنعت و معدن در سال ۱۳۷۳ و سرشماری عمومی کارگاهی در سال ۱۳۸۱ با هدف تهیه چارچوب فعالیت‌های اقتصادی کشور انجام پذیرد. با استفاده از این چارچوب در سال‌های بعد، طرح آمارگیری از کارگاه‌های صنعتی با هدف جمع‌آوری آمار و اطلاعات تفصیلی مربوط به بخش صنعت به اجرا در آمده و نتایج آن منتشر شده است. بخش صنعت در برنامه‌های توسعه اقتصادی به عنوان تأمین‌کننده بسیاری از نیازهای برنامه توسعه، جایگاه ویژه‌ای را به خود اختصاص داده است.

شناخت دقیق، رفع مشکلات و نارسایی‌ها و انجام برنامه‌ریزی‌های آتی برای این بخش، مستلزم جمع‌آوری اطلاعات جامع و به‌هنگام از کارگاه‌های صنعتی است. به منظور شناخت ساختار صنعتی کشور و اتخاذ سیاست‌های متناسب با آن، طرح آمارگیری از کارگاه‌های صنعتی کشور از سال ۱۳۵۱ به عنوان یکی از برنامه‌های مرکز آمار ایران منظور شده و همه ساله (به جز سال‌های ۱۳۵۶ و ۱۳۵۷) به اجرا درآمده و نتایج آن در اختیار استفاده‌کنندگان قرار گرفته است. در این طرح، تعداد زیادی از متغیرهای اقتصادی به همراه ارزش افزوده کارگاه‌های صنعتی اندازه‌گیری می‌شوند.

جدول ۱. متغیرهای مورد استفاده در تحقیق

کد متغیر	نام متغیر	تعریف عملی
V1	ارزش افزوده	ارزش افزوده کارگاه‌ها
V2	تعداد کارگاه	تعداد کارگاه‌ها
V3	کارگران ساده	تعداد کارگران ساد بخش شاغلان تولیدی

V4	کارگران ماهر	تعداد کارگران ماهر بخش شاغلان تولیدی
V5	تکنسین‌ها	تعداد تکنسین‌های بخش شاغلان تولیدی
V6	مهندسین	تعداد مهندسین بخش شاغلان تولیدی
V7	شاغلان غیرتولیدی	تعداد شاغلان غیرتولیدی
V8	کمتر از دیپلم	تعداد افراد باسواد کمتر از دیپلم
V9	دیپلم	تعداد افراد دارای مدرک دیپلم
V10	فوق دیپلم	تعداد افراد دارای مدرک فوق دیپلم
V11	لیسانس	تعداد افراد دارای مدرک لیسانس
V12	فوق لیسانس	تعداد افراد دارای مدرک فوق لیسانس
V13	دکتر	تعداد افراد دارای مدرک دکترا
V14	بی سواد	تعداد افراد بی سواد
V15	مواد خام و اولیه	ارزش مواد خام و اولیه، لوازم بسته بندی، ابزار و وسایل کار کم دوام برحسب میلیون ریال
V16	مواد مصرفی غذای طبخ شده توسط کارگاه	ارزش مواد مصرفی غذای طبخ شده توسط کارگاه‌ها برحسب میلیون ریال
V17	لوازم مصرفی برای تعمیر اساسی اموال سرمایه ای توسط کارگاه	ارزش لوازم مصرفی برای تعمیر اساسی اموال سرمایه‌ای توسط کارگاه‌ها برحسب میلیون ریال
V18	سوخت مصرف شده	ارزش سوخت مصرف شده برحسب میلیون ریال
V19	برق خریداری شده	ارزش برق خریداری شده برحسب میلیون ریال
V20	آب خریداری شده	ارزش آب خریداری شده برحسب میلیون ریال
V21	پرداختی بابت خدمات صنعتی	پرداختی بابت خدمات صنعتی برحسب میلیون ریال
V22	ارزش ستانده فعالیت های صنعتی	ارزش ستانده فعالیت های صنعتی برحسب میلیون ریال
V23	پرداختی خدمات غیرصنعتی	پرداختی خدمات غیرصنعتی برحسب میلیون ریال
V24	دریافتی خدمات غیرصنعتی	دریافتی خدمات غیرصنعتی برحسب میلیون ریال
V25	ارزش سرمایه گذاری کارگاه‌ها	ارزش سرمایه گذاری کارگاه‌ها برحسب میلیون ریال
V26	ارزش موجودی انبار	ارزش موجودی انبار برحسب میلیون ریال

در جدول فوق برخی متغیرهای اندازه‌گیری شده (V2 تا V26) به عنوان متغیرهای مستقل و متغیر ارزش افزوده (V1) به عنوان متغیر وابسته (پاسخ) در نظر گرفته شده است. در ادامه بر اساس داده‌های متغیرهای فوق روش ماشین بردار پشتیبان^۱ (SVM) برازش داده می‌شود. سپس براساس آن، مهم‌ترین عوامل تأثیرگذار روی ارزش افزوده مشخص می‌گردد.

¹. Support Vector Machine

۷- اجرای ماشین بردار پشتیبان:

برای استفاده و به‌کارگیری روش ماشین بردار پشتیبان که یکی از روش‌های پرکاربرد در زمینه داده‌کاوی می‌باشد، لازم هست که داده‌های جمع‌آوری شده به دو گروه آموزشی و آزمایشی تفکیک شوند که بتوان مدل خوبی روی داده‌ها برازش داده شود. در این تحقیق، ۷۵ درصد از مجموع داده‌ها به عنوان گروه آموزشی و ۲۵ درصد باقیمانده نیز به عنوان گروه آزمایشی در نظر گرفته شده است.

پس از اجرای دستورات زیر در نرم‌افزار R اهمیت متغیرهای ورودی موجود در مطالعه محاسبه می‌شود.

```
> data<- read.csv("data.csv")
> n <- nrow(data) # Number of observations
> ntrain <- round(n*0.75) # 75% for training set
> set.seed(314) # Set seed for reproducible results
> tindex <- sample(n, ntrain) # Create a random index
> train_data <- data[tindex,] # Create training set
> test_data <- data[-tindex,] # Create test set
> library(rminer)
> M <- fit(v1~., data=train_data, model="svm")
> svm.imp <- Importance(M, data=train_data, method="sensv")
> imp<-matrix(svm.imp$imp[-1], nrow=1, ncol=25)
> colnames(imp)<-colnames(data[, -1])
> imp1<-imp[,order(imp[1,],decreasing = T)]
> t1<-lapply(imp1, attributes)
> plot(imp1,type="p", col=2, pch=8, xlab="Variables", ylab="Importance")
> text(as.vector(imp1), labels=colnames(t(as.matrix(t1))), cex= 0.7, pos=3, offset = 0.7)
> write.table(imp1, file = "imp-svm.csv", sep = ",", row.names=TRUE,
+ qmethod = "double")
```

۸- تعیین اهمیت متغیرها در ماشین بردار پشتیبان:

برای به‌دست آوردن اهمیت هر یک از متغیرهای موجود در مدل ماشین بردار پشتیبان، با استفاده از پکیج آماری rminer نتایجی به صورت زیر حاصل شد:

جدول ۲. اهمیت متغیرها در SVM

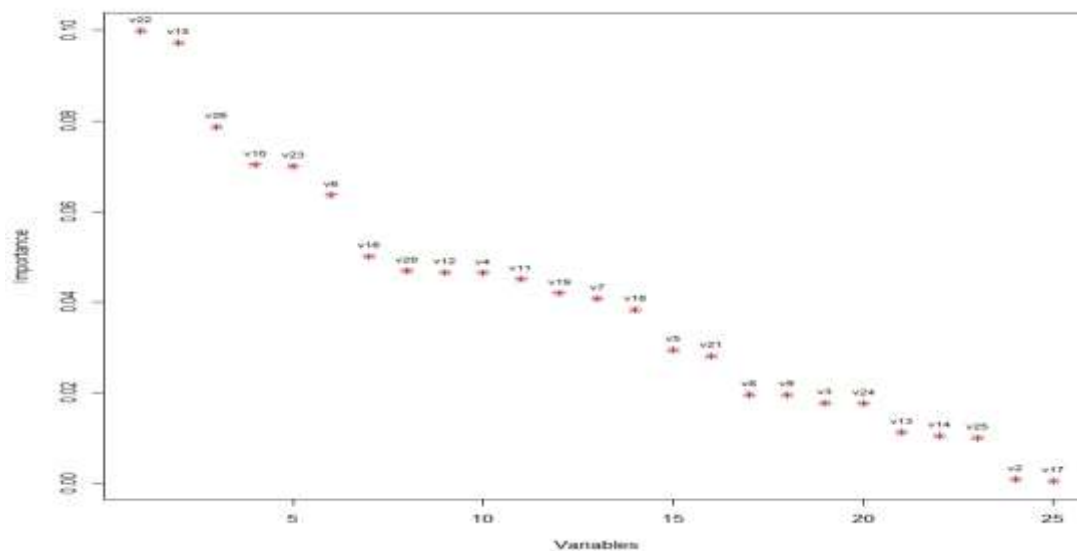
ترتیب	نام متغیر	اهمیت
۱	v22	0.099888
۲	v15	0.097269
۳	v26	0.078676
۴	v10	0.070316
۵	v23	0.069931
۶	v6	0.063633
۷	v16	0.050036

0.046876	v20	۸
0.046557	v12	۹
0.046417	v4	۱۰
0.045033	v11	۱۱
0.041931	v19	۱۲
0.040763	v7	۱۳
0.038269	v18	۱۴
0.029413	v5	۱۵
0.028055	v21	۱۶
0.01953	v8	17
0.019447	v9	18
0.017643	v3	19
0.017586	v24	20
0.011173	v13	21
0.0104	v14	22
0.009907	v25	23
0.000836	v2	24
0.000417	v17	25

با توجه به جدول فوق نتیجه می‌گیریم که متغیرهای ارزش ستانده فعالیت‌های صنعتی (V22)، مواد خام و اولیه (V15)، ارزش موجودی انبار (V26)، افراد کمتر از دیپلم بخش شاغلان تولیدی (V10)، پرداختی خدمات غیرصنعتی (V23) و مهندسین (V6) بخش شاغلان تولیدی به ترتیب از بالاترین درجه اهمیت برای ایجاد ارزش افزوده کارگاه‌های صنعتی در مدل پیشنهادی بر اساس ماشین بردار پشتیبان برخوردار هستند.

همچنین با توجه به نمودار زیر، اهمیت تمام متغیرها در ایجاد ارزش افزوده در طول سال‌های ۱۳۸۱ تا ۱۳۹۷ به صورت واضح دیده می‌شود.

شکل ۳. اهمیت متغیرها در روش ماشین بردار پشتیبان (SVM)



۹- بحث و نتیجه‌گیری:

به طور کلی بر اساس نتایج به دست آمده، می‌توان مهم‌ترین مباحثی را که ارزش افزوده کارگاه‌های صنعتی بالای ده نفر کارکن را در طول دوره مورد بررسی تحت تأثیر قرار داده‌اند، شناسایی کرد. بر اساس نتایج حاصل از تحقیق انجام گرفته، می‌توان گفت متغیرهایی که بیشترین تأثیر را در ارزش افزوده کارگاه‌های صنعتی استان کرمانشاه در طول سال‌های ۱۳۸۱ تا ۱۳۹۷ داشته، به ترتیب ارزش ستانده فعالیت‌های صنعتی، مواد خام و اولیه، ارزش موجودی انبار، افراد کمتر از دیپلم بخش شاغلان تولیدی، پرداختی خدمات غیرصنعتی و مهندسی بخش شاغلان تولیدی بوده است.

در توضیح نتایج حاصل از روش‌های داده‌کاوی در این مقاله نیاز به توضیح است که تأمین مواد اولیه برای تولیدکنندگان بسیار حائز اهمیت است و چنانچه تولیدکنندگان چشم‌انداز روشنی از وضعیت تأمین مواد اولیه مورد نیاز خود داشته باشند، به‌طورقطع برنامه‌ریزی دقیق‌تری برای تولید خواهند داشت. همچنین در بحث مواد اولیه و ارزش موجودی انبار باید به این نکته توجه شود که نگهداری موجودی بیشتر از حد بهینه، باعث حبس سرمایه ناشی از تأمین اقلام غیرضروری و تحمیل هزینه‌های مالی حمل و انبارداری و کاهش سودآوری خواهد شد. از طرف دیگر، موجودی کمتر از حد بهینه نیز موجب عدم تأمین به‌موقع مواد اولیه می‌شود و باعث ایجاد تأخیر در تکمیل فرآیند تولید و کاهش سودآوری و بهره‌وری خواهد شد. لذا باید مدیریت صحیحی برای موجودی مواد اولیه برقرار شود به‌طوری که از خام‌فروشی جلوگیری شود که در نهایت باعث توسعه صادرات

مبتنی بر کالاهای دارای ارزش افزوده خواهد شد. همچنین باید دقت داشت که سرمایه انسانی از نوع آموزش، مهارت و تخصص می‌تواند تأثیر مثبت و معناداری بر بهره‌وری کل عوامل مؤثر در ارزش افزوده کارگاه‌های صنعتی داشته باشد. با توجه به مطالب گفته شده و نتایج این تحقیق، چنانچه توجه ویژه‌ای به اهمیت این عوامل در بحث اقتصاد استان صورت بگیرد، شاهد رشد نرخ اشتغال و اقتصاد خواهیم بود و در آینده می‌توان اثرات مثبت آن را در میزان ارزش افزوده کارگاه‌های صنعتی استان شاهد بود.

فهرست منابع:

- آل عمران، رویا؛ آل عمران سیدعلی (۱۳۹۹). **بررسی عوامل اقتصادی مؤثر بر ارزش افزوده بخش کشاورزی ایران**، تحقیقات اقتصاد کشاورزی ۱۳، ۱، ۲۰۶-۱۹۱
- ابزری، مهدی؛ عمادزاده، مصطفی و رفیعی، عباس (۱۳۸۳). **عوامل مؤثر در تعیین سهم هزینه نیروی کار در ارزش افزوده صنایع**، دانش و توسعه ۲۴، ۱۵.
- اسماعیلی عبدالکریم؛ رحمتی، داریوش (۱۳۸۷). **اثر آزادسازی تجاری بر بخش کشاورزی ایران**، اقتصاد و کشاورزی ۱۲۸، ۱، ۲-۱۱۹.
- اکبری، نعمت‌الله؛ سامت، مرتضی و هادی، ولی‌اله (۱۳۸۲). **بررسی تأثیر هزینه‌های دولت بر ارزش افزوده بخش کشاورزی**، اقتصاد کشاورزی و توسعه ۱۶۶، ۴۱، ۱۱-۱۳۷.
- بیضایی، ابراهیم (۱۳۸۴). **ارائه مدل‌های تعیین عوامل مؤثر بر ارزش افزوده بخش حمل و نقل و تولید ناخالص داخلی کشور**، پژوهش‌نامه حمل و نقل، ۲، ۱، ۶۵-۷۷.
- حسینی‌نسب، محمد ابراهیم؛ موقری، هاد؛ باسرخا، مهدی (۱۳۸۹). **عوامل مؤثر بر ارزش افزوده کارگاه‌های صنعتی ایران با ده نفر کارکن و بیش‌تر**. تحقیقات اقتصادی. ۹۲، ۸۹، ۴۷-۶۴.
- رهنمای رودپشتی، فریدون؛ محمودی، محمد (۱۳۸۹). **بررسی و ارزیابی ارزش افزوده اقتصادی و ارزش افزوده بازار در مدیریت ارزش افزوده سهامداران**. پژوهشگر پاییز ۱۳۸۹ شماره ویژه.
- عمادزاده بکتاش، فروزان (۱۳۸۴). **اثر آموزش بر ارزش افزوده بخش صنعت**، دانش و توسعه ۵۰، ۱۶-۳۷.
- مشکانی علی؛ ناظمی، عبدالرضا (۱۳۸۸). **مقدمه‌ای بر داده‌کاوی**. مشهد: مؤسسه چاپ و انتشارات دانشگاه فردوسی مشهد.
- موقری، هادی (۱۳۸۷). **انتخاب مدل در داده‌های طولی**، پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس تهران.
- نادران، الیاس (۱۳۸۳). **اثر سیاست‌های اعتباری بر ارزش افزوده بخش صنعت ایران**، جستارهای اقتصادی، ۱، ۹-۴۲.

Araji, A. A. and White, F. C. (1996), **The Impact of Agricultural Research on United States Exports**, Idaho Agricultural Experiment Station Bulletin, 155.

Dugan, T.M., et al., **Machine learning techniques for prediction of early childhood obesity**. Applied clinical informatics, ۲۰۱۵. ۶(۰۳): p. ۵۲۰-۵۰۶

Fan, S., Zhang, L. and Zhang, X. (2002), **Growth, inequality, and poverty in rural China: The role of**

Lan, L. (2006), **Variable Selection for Linear Mixed Models for Longitudinal Data**, Ph.D. thesis, North Carolina State University.

Matloff, N., **Statistical regression and classification: from linear models to machine learning**. 2017: Chapman and Hall/CRC.

Stewart, G. B. (1991), **The Quest for Value: A Guide for Senior Managers**, Harper Business Publisher: New York.

Tibshirani, R. J. (1996), **Regression Shrinkage and Selection Via the LASSO**, Journal of the Royal Statistical Society Ser B, 58, 267 – 288.