

ایجاد یک سیستم خبره به منظور شناسایی مدل مناسب برای پیش‌بینی سری‌های زمانی

مهدی لطفی^۱ و حمیده رضوی*^۲

^۱ دانشجوی کارشناسی ارشد مهندسی صنایع - دانشگاه فردوسی مشهد

^۲ دانشیار گروه مهندسی صنایع - دانشگاه فردوسی مشهد

چکیده

وجود الگوهای متنوع برای مدل‌سازی و پیش‌بینی سری‌های زمانی، باعث می‌شود انتخاب ساختار و تحلیل این گونه مدل‌ها با سعی و خطا، صرف زمان زیاد و مبتنی بر نظر افراد خبره انجام شود. با توجه به ماهیت شرطی رویه‌های شناسایی مدل پیش‌بینی سری‌های زمانی، در این مقاله سعی می‌شود با ایجاد تعدادی موتور جستجو، تکنیک تجزیه و تحلیل مشخص شود و در مرحله بعد، با فرض معین بودن تکنیک مناسب، پایگاه دانش به گونه‌ای توسعه داده شود که در نهایت امکان نسبت دادن مدل مناسب به داده‌های در دست مطالعه، فراهم شود. پس از تعیین نوع روش، میزان برازش مدل مورد نظر بر داده‌های تاریخی به وسیله شاخص‌های مناسب اندازه‌گیری می‌شود. در صورت مناسب بودن این شاخص‌ها، با به کارگیری مدل انتخاب‌شده، پیش‌بینی دوره‌های بعدی به دست آمده و می‌توان تناسب مدل پیش‌بینی را سنجید. سپس با تکرار این فرآیند و تغییر پاسخ سؤالات در مراحل که حالت قطعی برای جواب وجود ندارد، به تعداد کافی مدل انتخاب کرده و از بین آنها برترین مدل برگزیده می‌شود. در انتها، پیاده‌سازی سیستم خبره و اجرای مدل روی داده‌های نمونه به صورت مطالعه موردی، کارآیی و اعتبار روش پیشنهادی را تأیید می‌کند.

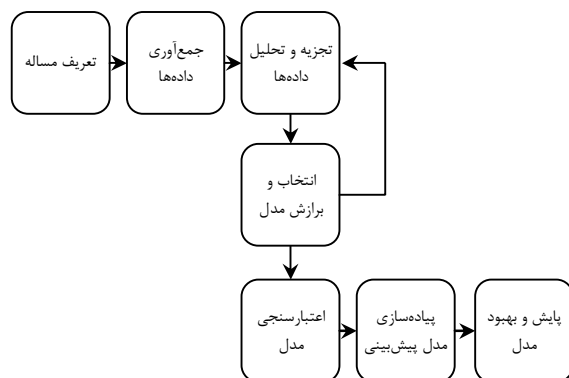
واژه‌های کلیدی: سیستم خبره، پیش‌بینی، سری زمانی، خطای پیش‌بینی

مقدمه

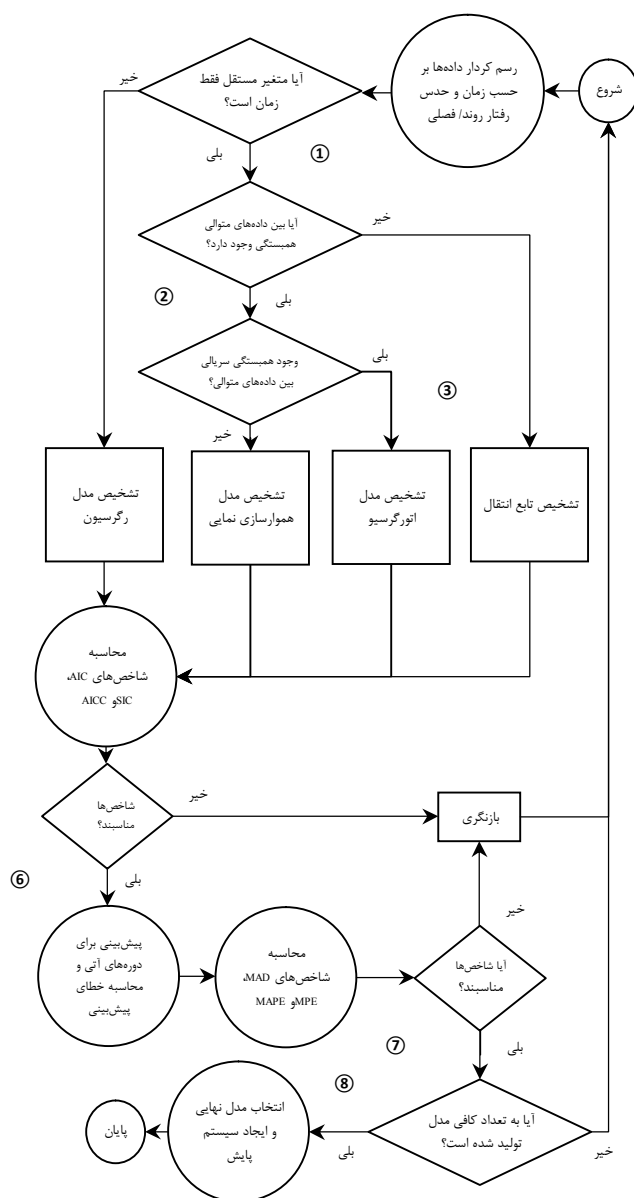
به عمل آورد و با طرح سؤال‌های معین، در مورد مناسب بودن روش فرضی تصمیم‌گیری کرد. واضح است که در صورت مناسب نبودن روش پیشنهادی، باید به مرحله بعدی رفت و روش دیگری را برای مطالعه داده‌های سری زمانی به کار گرفت. شناسایی مدل مناسب، با رویکردهای مختلفی مورد مطالعه قرار گرفته است که از آن جمله می‌توان به روش‌های مبتنی بر الگوریتم ژنتیک [۴] و شبکه عصبی [۵] اشاره کرد [۶].

ماهیت بسیاری از مسایل سری زمانی، به گونه‌ای است که تصمیم‌گیری در مورد مدل مناسب را دشوار می‌کند. زیرا این انتخاب، مستلزم آگاهی از رفتار قبلی و نسبت دادن الگویی مناسب به داده‌های تاریخی در محدوده مشخصی از زمان است؛ با این فرض که بتوان با تعمیم این الگو، وضعیت پدیده مورد مطالعه را برای مقاطع معینی از آینده پیش‌بینی و بر آن مبنا تصمیم‌گیری کرد. همچنین وجود مدل‌های مختلف و شاخص‌های متنوع برای سنجش کارآیی مدل بر پیچیدگی آن می‌افزاید [۱]. به عنوان مثال، می‌توان به پیش‌بینی وضعیت فروش و میزان تقاضا اشاره کرد. به طور کلی یک سری زمانی، مطابق فرآیند شکل ۱ مورد مطالعه قرار می‌گیرد [۲].

برای تعیین مدل سری زمانی مربوط به مجموعه‌ای از داده‌های مورد مطالعه، تکنیک‌ها و روش‌های مختلفی ارائه شده است. بسته به رفتار سری زمانی مورد مطالعه، روش‌هایی نظیر تحلیل رگرسیون، هموارسازی‌نمایی، مدل‌های اتورگرسیو، توابع انتقال و ... توسعه یافته‌اند [۳]. در هر یک از روش‌های کلاسیک گفته شده، باید تحلیل‌های مشخصی در مورد داده‌های در دست مطالعه



شکل ۱: فرآیند کلی در مطالعات سری زمانی



شکل ۲: منطق کلی سیستم خبره

در مورد نحوه تعیین پارامترها در هر یک از روش‌های تحلیل، انجام شده است [۸، ۹]. اما رویکرد اکثر مطالعات انجام‌شده و نرم‌افزارهای موجود، بر این مبنا بوده که، کاربر، روش‌ها و تکنیک‌های تحلیل سری‌های زمانی را به طور کامل می‌شناسد و به آسانی قادر است روش مناسب را انتخاب کند. به عبارت دیگر، آنچه کمتر مورد توجه قرار گرفته است، ارائه شیوه‌ای برای انتخاب و تعیین روش تجزیه و تحلیل است. در این پژوهش، سعی می‌شود با طراحی یک سیستم خبره، در انتخاب مدل مناسب پیش‌بینی سری زمانی به کاربر عادی کمک شود. از سوی

از آنجا که بخش مهمی از تکنیک‌ها و روش‌های تجزیه و تحلیل سری‌های زمانی، بر پایه ابزار و فنون آماری بنا شده است، نرم‌افزارهای متعددی نیز در زمینه برازش مدل پیش‌بینی به وجود آمده‌اند که از آن جمله می‌توان به نرم‌افزارهای Minitab، JMP و SAS اشاره کرد [۷]. این نرم‌افزارها، متناظر با روش تحلیلی مورد نظر کاربر، امکانات لازم را برای انتخاب مقادیر مختلف پارامترهای ورودی مدل و مقایسه نتایج حاصل از آن در اختیار می‌گذارند. در مقالات مختلف، مطالعات بسیاری نیز

متغیر وجود داشته باشد، باید مدل‌های رگرسیونی یا روش حداکثر درست‌نمایی را آزمون کرد. اما در غیر این صورت روش هموارسازی نمایی، مدل‌های اتورگرسیو و یا توابع انتقال، تکنیک‌های مناسب‌تری به نظر می‌رسند. در این محدوده، تصمیم‌گیری در مورد انتخاب تکنیک مناسب از بین سه تکنیک گفته شده، منوط به تشخیص نوع همبستگی بین داده‌های متوالی است؛ به این ترتیب که تکنیک توابع انتقال در صورت نبود همبستگی بین داده‌های متوالی، تکنیک مدل‌های اتورگرسیو در صورت وجود همبستگی سریالی بین داده‌های متوالی و تکنیک هموارسازی نمایی در صورت وجود همبستگی غیرسریالی بین داده‌های متوالی، پیشنهاد می‌شود.

به هر حال پس از تعیین نوع روش، باید میزان برازش مدل مورد نظر بر داده‌های تاریخی را به وسیله شاخص‌های مناسب نظیر AIC^6 ، SIC^7 ، $AICC^8$ اندازه‌گیری کرد.

در صورت مناسب بودن شاخص‌های ذکر شده، در گام سوم لازم است با به کارگیری مدل انتخابی، پیش‌بینی دوره‌های بعدی را به دست آورد تا بتوان با استفاده از شاخص‌های موجود مانند MAD^9 ، MPE^10 ، $MAPE^11$ مناسب بودن مدل پیش‌بینی را هم سنجید. بدیهی است با تکرار این فرآیند و تغییر دادن پاسخ‌های سؤالات (در مراحلی که حالت قطعی برای جواب وجود ندارد) می‌توان به تعداد کافی، مدل انتخاب کرده و سپس از بین آنها بهترین را برگزید.

به هر حال پس از انتخاب مدل نهایی، باید در گام چهارم، سیستم مناسبی برای پایش مدل طراحی کرد تا امکان تشخیص انحراف‌های احتمالی مدل پیش‌بینی نسبت به داده‌های در حال تولید نیز فراهم شده و در صورت لزوم، کل فرآیند بازنگری شود.

تشخیص مدل اتورگرسیو

پس از اطمینان از اینکه بین داده‌های متوالی همبستگی سریالی وجود دارد، باید نوع مدل اتورگرسیو را مطابق فلوجارت شکل ۳ تشخیص داد. ابزار مناسب برای این امر، اغلب کردارهای ACF^12 و $PACF^13$ است [۱۱]. مطابق تعریف، سه حالت زیر در بررسی کردارهای ذکر شده قابل مشاهده است.

• چنانچه کردار $PACF$ میرا و کردار ACF بعد از

دیگر، این سیستم از جهاتی مانند یک سیستم پشتیبانی تصمیم عمل می‌کند و انتخاب نهایی، به هر حال، بر عهده کاربر خواهد بود. در ابتدا با توجه به ماهیت اگر-آنگاه^۱ بودن رویه انتخاب روش، مجموعه‌های قاعده-بنیاد^۲ ایجاد شده و نحوه انتخاب تکنیک تجزیه و تحلیل مشخص می‌شود. در مرحله بعد، با فرض معین بودن تکنیک مناسب، مجموعه‌ای از بسته‌های دانش-بنیاد^۳ پیشنهاد می‌شود که امکان نسبت دادن مدل مناسب به داده‌های در دست مطالعه را فراهم می‌آورد. در انتها، شاخص‌های کارایی مدل، از طریق مقایسه خطاها، تناسب مدل انتخابی با داده‌های موجود را می‌سنجند.

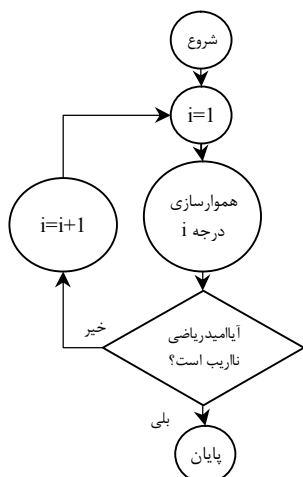
منطق کلی سیستم خبره

سیستم‌های خبره، برنامه‌های کامپیوتری هستند که نحوه تفکر یک متخصص در یک زمینه خاص را شبیه‌سازی می‌کنند [۱۰]. پایگاه دانش و موتور تصمیم‌گیری، دو بخش مجزای هر سیستم خبره به شمار می‌روند. پایگاه دانش یک سیستم خبره از هر دو نوع دانش مبتنی بر حقایق^۴ و نیز دانش غیرقطعی^۵ استفاده می‌کند. دانش حقیقی یا قطعی، نوعی از دانش است که می‌توان آن را در حیطه‌های مختلف به اشتراک گذاشت و تعمیر داد؛ چرا که درستی آن قطعی است. در سوی دیگر، دانش غیر قطعی قرار دارد که بیشتر مبتنی بر برداشت‌های شخصی است. سیستم خبره‌ای که در این مقاله به آن پرداخته می‌شود، به دنبال آن است که با طرح سؤال‌های متوالی مبتنی بر دانش قطعی، کاربر را به سمت انتخاب مدل مناسب پیش‌بینی سری‌های زمانی راهنمایی کند. منطق کلی این سیستم، مطابق شکل ۲ است. بدیهی است انتخاب مقادیر پارامتری در هر مدل پیشنهادی، همچنان بر دانش غیرقطعی کاربر استوار است.

مطابق شکل ۲، اولین گام برای تشخیص مدل مناسب، رسم کردار داده‌ها بر حسب زمان و حدس رفتار روند یا فصلی است. این گام در عین سادگی و کلی بودن، می‌تواند بسیار راهگشا و جهت‌دهنده باشد. واضح است فرضیاتی که در این نخستین گام پایه‌گذاری می‌شوند، در مراحل بعدی باید به طور دقیق‌تر محک زده شوند.

در گام دوم، یک سؤال کلیدی و اساسی مطرح می‌شود و آن اینکه آیا متغیر مستقل فقط زمان است یا بیش از یک متغیر در مدل تأثیرگذار است. چنانچه بیش از یک

ابتدا با قراردادن $i=1$ باید روش هموارسازی‌نمایی مرتبه i را پیاده کرد و سپس امید ریاضی را به دست آورد. اگر امید ریاضی حاصل، نآریب نباشد، باید با قراردادن $i=i+1$ آن قدر این فرآیند را تکرار کرد تا در نهایت امید ریاضی دیگر آریب نباشد. به این ترتیب، آخرین مقدار i ، نشان‌دهنده درجه (مرتبه) مدل هموارسازی‌نمایی خواهد بود.



شکل ۴: فلوجارت تشخیص مدل هموارسازی

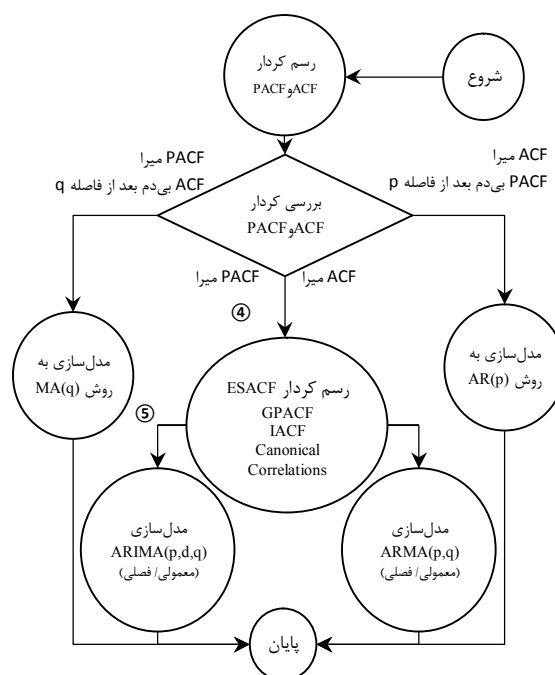
تشخیص مدل رگرسیون

چنانچه متغیرهای مستقل بیش از تنها متغیر مرسوم (یعنی زمان) باشد، استفاده از مدل‌های رگرسیونی، انتخاب بهتری به نظر می‌رسد. مطابق شکل ۴، فرض اولیه آن خواهد بود که مدل از نوع OLS^{23} است. به عبارت دیگر باید کردار باقیمانده‌ها را رسم کرده و در مورد نویز سفید بودن آن‌ها تجزیه و تحلیل کرد. اگر نتیجه تحلیل، نویز سفید بودن باقیمانده‌ها باشد، حالت ایده‌آلی است و می‌توان با برقرار کردن یک آزمون فرض و محاسبه ضریب تعیین، نسبت به مناسب بودن مدل (اگر مقدار ضریب تعیین در سطح مورد انتظار باشد) یا افزودن به / کاستن از تعداد متغیرها به روش پله‌ای رو به جلو و یا پله‌ای رو به عقب (در حالتی که ضریب تعیین نسبت به مقدار مطلوب، کوچک/ بزرگ باشد) تصمیم‌گیری کرد و فرآیند را تا رسیدن به وضعیت مطلوب تکرار کرد. البته اگر باقیمانده‌ها نویز سفید نباشند، فرضیات مدل OLS برقرار نبوده و باید در آن تجدید نظر کرد. ولی قبل از آن، باید با انجام آزمون دوربین-واتسون و محاسبه QLB و QBP فرض خودهمبستگی باقیمانده‌ها را بررسی کرد. در صورت وجود

فاصله q بی‌دُم^{۱۴} شود، مدل‌سازی به روش $MA(q)$ ^{۱۵} انجام می‌شود.

• چنانچه کردار ACF میرا و کردار $PACF$ بعد از فاصله p بی‌دُم شود، مدل‌سازی به روش $AR(q)$ ^{۱۶} انجام می‌شود.

• چنانچه کردار ACF و $PACF$ هر دو میرا باشند، آنگاه این کردارها دیگر مناسب نبوده و باید با رسم کردارهای $ESACF$ ^{۱۷}، $GPACF$ ^{۱۸}، $IACF$ ^{۱۹} و یا روش همبستگی کانونی^{۲۰} مقادیر p و q را در قالب یک مدل $ARMA$ ^{۲۱} و با در نظر گرفتن یک دیفرانسیل مرتبه d ام به شکل یک مدل $ARIMA$ ^{۲۲} در نظر گرفت که البته هر یک از این مدل‌ها می‌توانند فصلی (با پارامتر S) هم باشند.



شکل ۳: فلوجارت تشخیص مدل اتورگرسیو

تشخیص مدل هموارسازی‌نمایی

چنانچه همبستگی بین داده‌های متوالی به شکل سریالی نباشد، می‌توان روش هموارسازی‌نمایی را پیشنهاد کرد. مطابق شکل ۳ در این مرحله، هدف، یافتن پارامتر i به عنوان درجه (یا مرتبه) هموارسازی است. به این منظور،

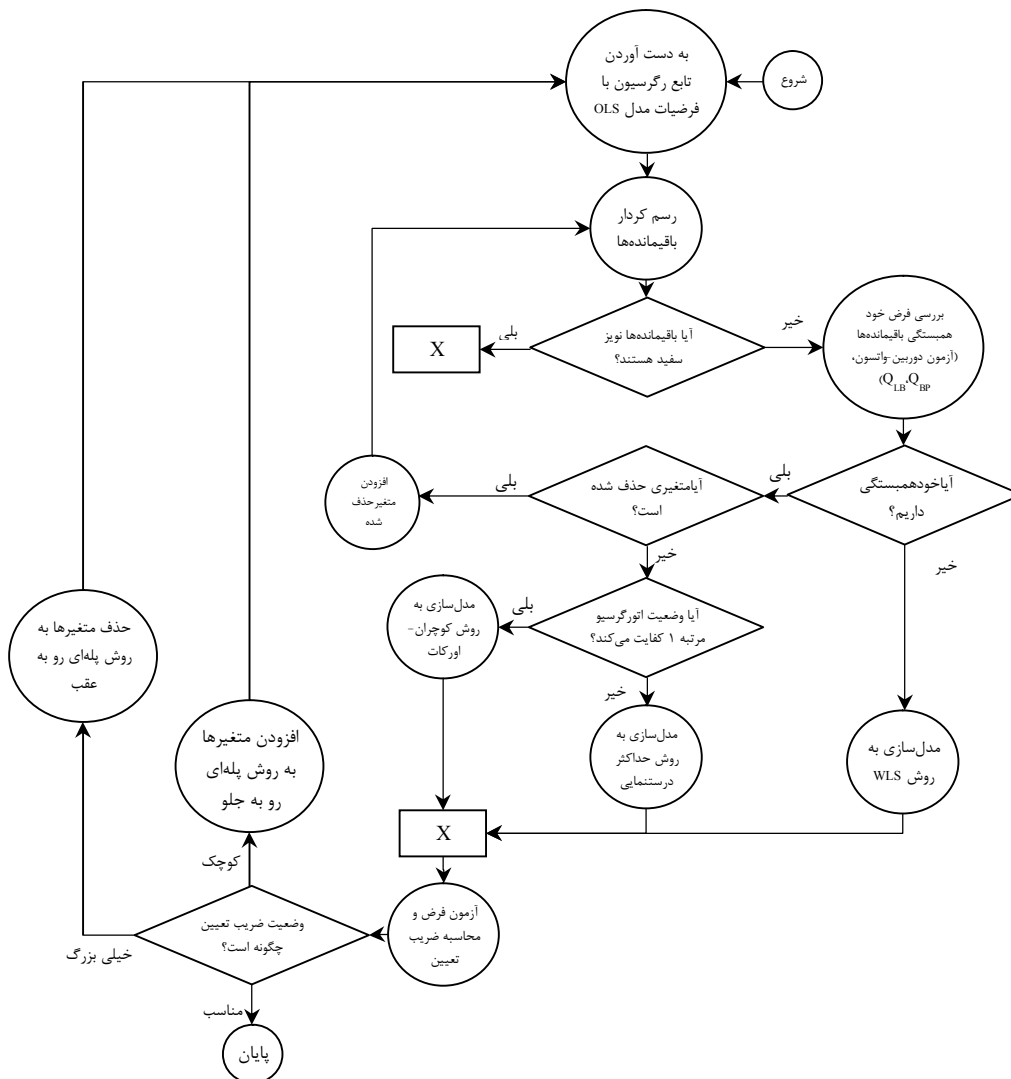
در عین حال اگر نتیجه بررسی مسئله در مرحله برقرار کردن آزمون دوربین-واتسون، حاکی از نبود همبستگی بین داده‌های سری زمانی باشد، مدل‌سازی به روش WLS^{۲۴} انجام می‌شود.

تشخیص تابع انتقال

چنانچه بین داده‌های متوالی، همبستگی وجود نداشته باشد، روش پیشنهادی برای تجزیه و تحلیل داده‌های سری زمانی، استفاده از توابع انتقال است که خود شامل توابع انتقال خانواده توانی، پاسخ ضربه و پاسخ پله‌ای است [۱۲].

خودهمبستگی، اولین موضوعی که باید بررسی شود آن است که آیا همه متغیرهای مستقل در نظر گرفته شده‌اند. بدیهی است در صورت منفی بودن جواب این سؤال، افزودن متغیر حذف شده به مدل و بررسی مجدد نتایج تحلیل، ضروری خواهد بود.

اما چنانچه در بررسی‌های تکمیلی، نشانه‌ای دال بر حذف متغیر مشاهده نشود، پیشنهاد می‌شود، مدل‌سازی با یکی از روش‌های کوچران-آرکات یا حداکثر درست‌نمایی انجام شود. به این ترتیب که اگر مدل اتورگرسیو از مرتبه یک، کفایت لازم را داشته باشد، باید از روش کوچران-آرکات و در غیر این صورت باید از روش حداکثر درست‌نمایی استفاده شود.



شکل ۵: فلوجارت تشخیص مدل رگرسیون

پیاده‌سازی سیستم خبره

می‌توان با زبان‌هایی مانند VisualC++ به بخش‌های محاسباتی افزود.

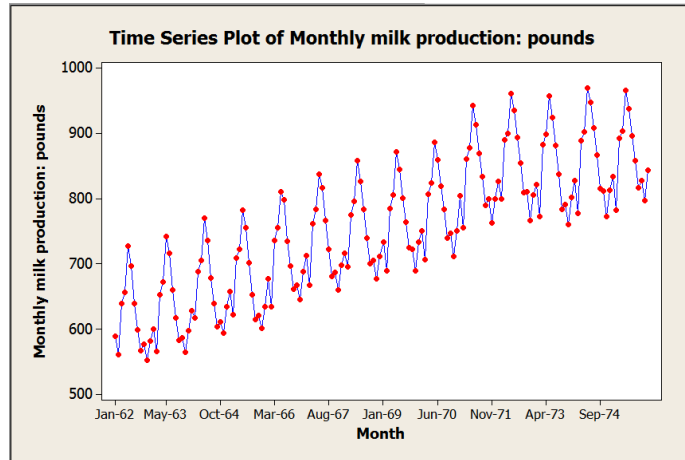
برای پیاده‌سازی سیستم خبره، نرم‌افزارهای متنوعی از جمله Prolog, Clips, DEX, Doctus و JESS توسعه یافته است [۱۳]. منطق کلی این نرم‌افزارها یکسان بوده و اغلب بر اساس جستجوی پاسخ از بین اطلاعات سازماندهی شده قبلی استوار است. بنابراین، برنامه‌نویسی و اجرای پوسته آن اغلب نیاز به ابتکار خاصی ندارد. از سوی دیگر، در سیستم مورد نیاز برای این پروژه، امکان ارتباط پوسته سیستم با نرم‌افزارهای تخصصی پیش‌بینی مانند Minitab, SAS و JMP که بتواند همزمان مدل‌ها را اجرا کرده و شاخص‌ها را برآورد کند، الزامی است. خوشبختانه این مهم از طریق کدنویسی در نرم‌افزارهای یاد شده، فراهم بوده و پوسته خارجی را نیز

مطالعه موردی

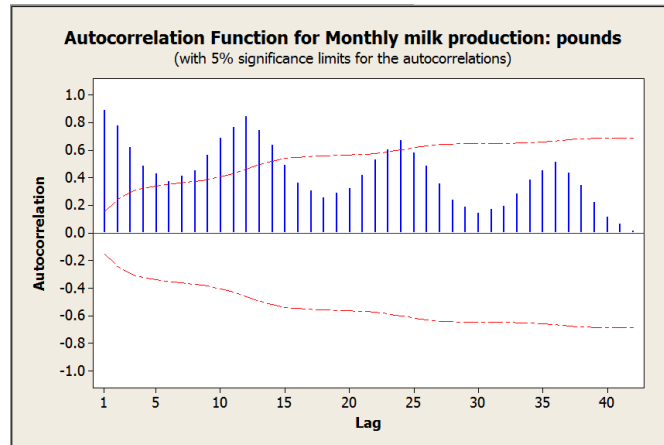
برای بررسی میزان توانمندی سیستم خبره پیشنهادی، یک مجموعه از داده‌های سری زمانی مطابق جدول ۱، مربوط به تولید ماهانه شیر (بر حسب پوند به ازای هر راس گاو) در نظر گرفته شده است که مربوط به دوره ۱۶۸ ماهه از ژانویه ۱۹۶۲ تا دسامبر ۱۹۷۵ است [۱۴]. شکل ۶ کردار دو بعدی داده‌ها در طول زمان را نشان می‌دهد که شکل کلی سری زمانی، حاکی از روند افزایشی و در عین حال رفتار دوره‌ای داده‌ها است.

جدول ۱: تولید ماهانه شیر (بر حسب پوند به ازای هر راس گاو) از ژانویه ۱۹۶۲ تا دسامبر ۱۹۷۵

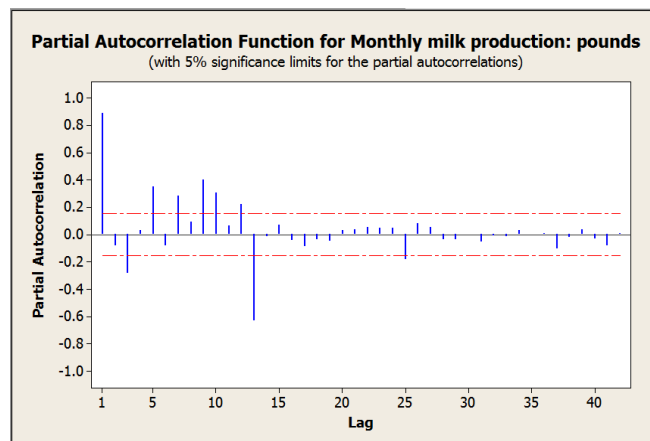
Monthly milk production: pounds per cow. Jan 62 – Dec 75									
Month	production	Month	production	Month	production	Month	production	Month	production
Jan-62	589	Nov-64	594	Sep-67	681	Jul-70	819	May-73	957
Feb-62	561	Dec-64	634	Oct-67	687	Aug-70	783	Jun-73	924
Mar-62	640	Jan-65	658	Nov-67	660	Sep-70	740	Jul-73	881
Apr-62	656	Feb-65	622	Dec-67	698	Oct-70	747	Aug-73	837
May-62	727	Mar-65	709	Jan-68	717	Nov-70	711	Sep-73	784
Jun-62	697	Apr-65	722	Feb-68	696	Dec-70	751	Oct-73	791
Jul-62	640	May-65	782	Mar-68	775	Jan-71	804	Nov-73	760
Aug-62	599	Jun-65	756	Apr-68	796	Feb-71	756	Dec-73	802
Sep-62	568	Jul-65	702	May-68	858	Mar-71	860	Jan-74	828
Oct-62	577	Aug-65	653	Jun-68	826	Apr-71	878	Feb-74	778
Nov-62	553	Sep-65	615	Jul-68	783	May-71	942	Mar-74	889
Dec-62	582	Oct-65	621	Aug-68	740	Jun-71	913	Apr-74	902
Jan-63	600	Nov-65	602	Sep-68	701	Jul-71	869	May-74	969
Feb-63	566	Dec-65	635	Oct-68	706	Aug-71	834	Jun-74	947
Mar-63	653	Jan-66	677	Nov-68	677	Sep-71	790	Jul-74	908
Apr-63	673	Feb-66	635	Dec-68	711	Oct-71	800	Aug-74	867
May-63	742	Mar-66	736	Jan-69	734	Nov-71	763	Sep-74	815
Jun-63	716	Apr-66	755	Feb-69	690	Dec-71	800	Oct-74	812
Jul-63	660	May-66	811	Mar-69	785	Jan-72	826	Nov-74	773
Aug-63	617	Jun-66	798	Apr-69	805	Feb-72	799	Dec-74	813
Sep-63	583	Jul-66	735	May-69	871	Mar-72	890	Jan-75	834
Oct-63	587	Aug-66	697	Jun-69	845	Apr-72	900	Feb-75	782
Nov-63	565	Sep-66	661	Jul-69	801	May-72	961	Mar-75	892
Dec-63	598	Oct-66	667	Aug-69	764	Jun-72	935	Apr-75	903
Jan-64	628	Nov-66	645	Sep-69	725	Jul-72	894	May-75	966
Feb-64	618	Dec-66	688	Oct-69	723	Aug-72	855	Jun-75	937
Mar-64	688	Jan-67	713	Nov-69	690	Sep-72	809	Jul-75	896
Apr-64	705	Feb-67	667	Dec-69	734	Oct-72	810	Aug-75	858
May-64	770	Mar-67	762	Jan-70	750	Nov-72	766	Sep-75	817
Jun-64	736	Apr-67	784	Feb-70	707	Dec-72	805	Oct-75	827
Jul-64	678	May-67	837	Mar-70	807	Jan-73	821	Nov-75	797
Aug-64	639	Jun-67	817	Apr-70	824	Feb-73	773	Dec-75	843
Sep-64	604	Jul-67	767	May-70	886	Mar-73	883		
Oct-64	611	Aug-67	722	Jun-70	859	Apr-73	898		



شکل ۶: کردار دو بعدی داده‌ها بر حسب زمان



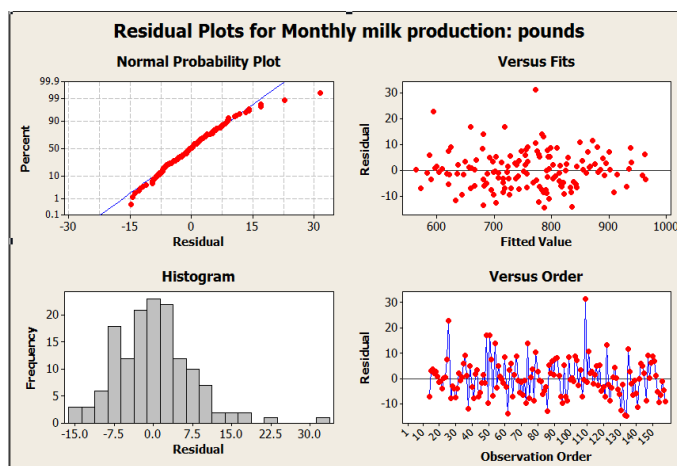
شکل ۷: تابع خودهمبستگی داده‌ها



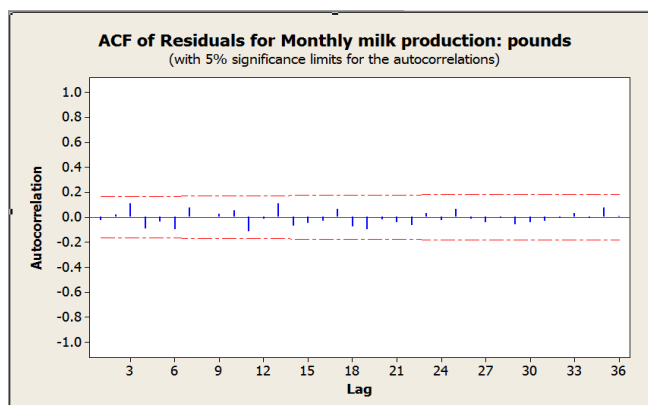
شکل ۸: تابع خودهمبستگی جزئی داده‌ها

جدول ۲: خروجی نرم افزار Minitab با فرض انتخاب مدل $ARIMA(0,1,1).(0,1,1)_{12}$

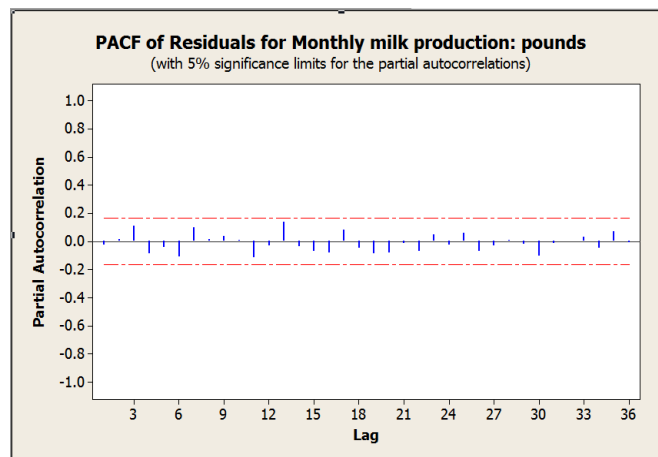
ARIMA Model: Monthly milk production: pounds				
Estimates at each iteration				
Iteration	SSE	Parameters		
0	10453.4	0.100	0.100	0.052
1	9242.8	0.166	0.250	-0.015
2	8420.7	0.214	0.400	-0.061
3	7918.1	0.248	0.550	-0.090
4	7793.9	0.261	0.662	-0.108
5	7793.0	0.254	0.657	-0.119
6	7793.0	0.254	0.657	-0.119
7	7793.0	0.254	0.657	-0.119
Relative change in each estimate less than 0.0010				
Final Estimates of Parameters				
Type	Coef	SE Coef	T	P
MA 1	0.2544	0.0815	3.12	0.002
SMA 12	0.6566	0.0703	9.33	0.000
Constant	-0.1187	0.1714	-0.69	0.490
Differencing: 1 regular, 1 seasonal of order 12				
Number of observations: Original series 158, after differencing 145				
Residuals: SS = 7619.92 (backforecasts excluded)				
MS = 53.66 DF = 142				
Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	8.3	16.5	19.9	31.2
DF	9	21	33	45
P-Value 0.502 0.742 0.964 0.941				



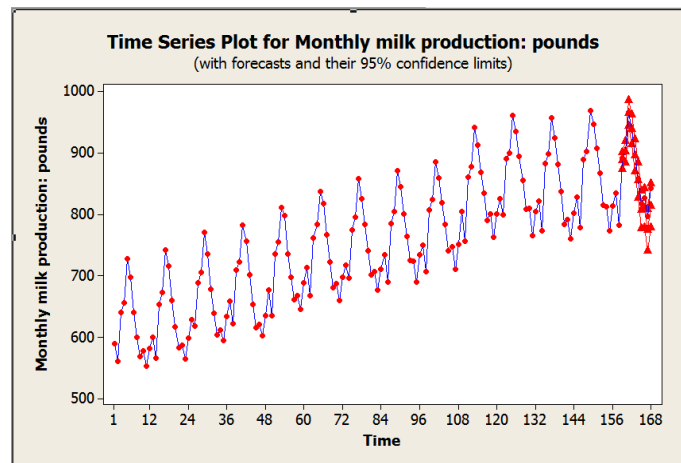
شکل ۹: کردارهای باقیماندهها برای بررسی پایداری مدل



شکل ۱۰: تابع خودهمبستگی باقیماندهها



شکل ۱۱: تابع خودهمبستگی جزئی باقیمانده‌ها



شکل ۱۲: کردار دو بعدی داده‌ها به همراه پیش‌بینی ۱۰ دوره آخر

جدول ۲، به صورت $ARIMA(0,1,1).(0,1,1)_{12}$ به دست آمده است.

پس از برازش مدل، نتیجه بررسی باقیمانده‌ها در شکل ۹ به تصویر در آمده است که نشان‌دهنده پایداری آن‌ها بوده و کردار فراوانی باقیمانده‌ها، وضعیت نرمال را نمایش می‌دهد. علاوه بر آن، کردار احتمال نرمال، انطباق قابل قبولی بر یک خط مستقیم را نشان داده و کردار باقیمانده‌ها بر حسب مقادیر برازش شده نیز، الگوی خاصی را نمایش نمی‌دهد. برای اطمینان از پایداری مدل، کردارهای ACF و PACF به ترتیب در شکل‌های ۱۰ و ۱۱ رسم شده است که این کردارها پایداری خودهمبستگی را به نحو مطلوبی به نمایش می‌گذارند. همچنین شاخص‌های برازش مدل، مناسب بودن آن را تأیید می‌کند. برای بررسی معیارهای کارآمدی مدل پیش‌بینی، ۱۰ داده آخر جدا شده و کردار دو بعدی به

از آنجا که تنها متغیر مستقل در این مسئله، زمان بوده و بین داده‌های متوالی همبستگی سریالی وجود دارد، پیشنهاد سیستم خبره، به کارگیری مدل‌های اتورگرسیو است (مسیر ③→②→① در شکل ۱). از طرف دیگر، کردار ACF در شکل ۷، نشان‌دهنده تغییرات دوره‌ای با طول دوره ۱۲ ماه است که به شکل میرا ادامه می‌یابد. هر چند مطابق شکل ۸، کردار PACF در مقطع زمانی $lag=1$ نشان‌دهنده cut-off است، اما در مقاطع بعدی (اغلب ۱۲، ۲۴، ۳۶) انحرافات از حدود کنترلی مشاهده می‌شود و حالت سینوسی میرا به وضوح مشهود است. در نتیجه پیشنهاد سیستم خبره، به کارگیری مدل فصلی $ARIMA(p,d,q).(P,D,Q)_s$ است (مسیر ⑤→④ در شکل ۳). لازم به توضیح است که پارامترهای مدل ذکرشده از طریق کدنویسی و با استفاده از نرم‌افزار Minitab مورد مطالعه قرار گرفته و مدل نهایی مطابق

افراد خبره همیشه در دسترس نیستند. بنابراین وجود یک سیستم پشتیبانی که بتواند گزینه‌های مختلف را مقایسه کرده و برترین گزینه‌ها را به کاربر عادی عرضه کند، بسیار راهگشا خواهد بود. با توجه به گستردگی مدل‌های پیش‌بینی، پس از مطالعه و دسته‌بندی مدل‌ها، در ابتدا تقسیم‌بندی کلی انجام شده و سپس در هر قسمت به کمک فلوچارت‌ها، مدل مناسب جستجو شده و چگونگی برآزش و شاخص‌های آن مورد تحلیل و مقایسه قرار گرفتند. در انتها با ارائه یک مثال موردی، کارایی سیستم اثبات شد.

همراه پیش‌بینی این ۱۰ دوره در شکل ۱۲ رسم شده که خروجی نرم‌افزار Minitab نیز در جدول ۳ قابل مشاهده است. در این جدول، حدود بالا و پایین مورد نیاز برای حصول اطمینان از پایش مناسب مدل پیشنهادی، مشخص شده است (مسیر ⑧→⑦→⑥ در شکل ۱).

نتیجه و جمع‌بندی

در این مقاله، شناسایی مدل پیش‌بینی برای سری‌های زمانی، به کمک یک سیستم خبره بررسی شد. در روش‌های سنتی، کارشناسان خبره با بررسی کردارها و نتایج محاسباتی به انتخاب مدل می‌پردازند. حال آنکه به دلیل کاربرد فراگیر سری‌های زمانی در حوزه‌های مختلف،

ARIMA Model: Monthly milk production: pounds

Estimates at each iteration

Iteration	SSE	Parameters		
0	10966.3	0.100	0.100	0.223
1	9748.6	0.151	0.250	0.133
2	8881.0	0.188	0.400	0.066
3	8312.8	0.214	0.550	0.013
4	8134.4	0.230	0.679	-0.025
5	8133.9	0.227	0.674	-0.028
6	8133.8	0.227	0.674	-0.027
7	8133.8	0.227	0.673	-0.027
8	8133.8	0.227	0.673	-0.027

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	0.2275	0.0796	2.86	0.005
SMA 12	0.6734	0.0641	10.51	0.000
Constant	-0.0273	0.1617	-0.17	0.866

Differencing: 1 regular, 1 seasonal of order 12

Number of observations: Original series 168, after differencing 155

Residuals: SS = 7956.41 (backforecasts excluded)
MS = 52.34 DF = 152

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	9.8	16.2	19.5	31.9
DF	9	21	33	45
P-Value	0.367	0.758	0.970	0.930

Forecasts from period 159

Period	Forecast	95% Limits		Actual
		Lower	Upper	
160	905.565	891.381	919.748	903.000
161	968.499	950.576	986.422	966.000
162	941.866	920.859	962.873	937.000
163	899.660	875.967	923.353	896.000
164	859.152	833.049	885.256	858.000
165	811.043	782.733	839.353	817.000
166	813.583	783.226	843.940	827.000
167	777.509	745.235	809.782	797.000
168	817.257	783.174	851.340	843.000
169	841.298	805.497	877.098	

جدول ۳: خروجی نرم‌افزار Minitab به همراه پیش‌بینی ۱۰ دوره آخر و حدود بالا و پایین آن‌ها

مراجع

- 1- Montgomery.C. Douglas, Jenniggs. L. Cherly, Kulahci. Muart, (2008). "Introduction to time series analysis and forecasting."
- 2- Cooray, T.M. (2008). "Applied time series, analysis and Forecasting." *Alpa science international Ltd.*
- 3- Brockwell, P. (1996). "Introduction to time series and forecasting", *Springer*.
- 4- Chorng .S. O, Jih-Jeng. H. and Gwo. H. T. (2005). "Model identification of ARIMA family using genetic algorithms." *Applied Mathematics and Computation*, No.164 , PP. 885-912.
- 5- Kun C. Lee, Sang B. Oh, July (1996). "An intelligent approach to time series identification by a neural network-driven decision tree classifier." *Decision Support Systems*, Volume 17, Issue 3, PP.183-197.
- 6- Piet M. T. Broersen and Stijn de Waele, (2005). "Automatic Identification of Time-Series Models From Long Autoregressive Models." *IEEE transactions on instrumentation and measurement*, Vol. 54, No. 5, PP.1862-1868.
- 7- www.jmp.com
www.minitab.com
www.sas.com
- 8- Monica A., Fred C., J.Scott A. and Miles K. (2001). "Automatic identification of time series features for rule-based forecasting." *International Journal of Forecasting*, Vol 17, Issue 2, PP. 143-157.
- 9- Pamela A. Texter, J.Keith Ord, (1989). "Forecasting using automatic identification procedures: A comparative analysis." *International Journal of Forecasting*, Vol. 5, Issue 2, PP. 209-215.
- 10- The Handbook of Applied Expert Systems, Edited by Jay Liebowitz, 1997.
- 11- Pole, A. (1994). "Applied Bayesian analysis and Forecsting." Chapman and Hall.
- 12- Bowerman, B.L. (1993). "Forecasting and time series:an applied approach",Duxbury.
www.jessrules.com
www.visual-prolog.com
www.dex.com
www.doctuskbs.com
- 13- <http://datamarket.com/data/set/22ox/monthly-milk-production-pounds-per-cow-jan-62-dec-75#!display=table&ds=22ox>

واژه‌های انگلیسی به ترتیب استفاده در متن

- 1- If-then
- 2- Rule Base
- 3- Knowledge Base
- 4- Factual Knowledge
- 5- Heuristic Knowledge
- 6- Akaike Information Criteria
- 7- Schwarz Information Criteria
- 8- Akaike Information Criteria-corrected
- 9- Mean Absolute Deviation
- 10- Mean Percent Error
- 11- Mean Absolute Percent Error
- 12- Auto Correlated Function
- 13- Partial Auto Correlated Function
- 14- Cut off

- 15- Moving Average
 - 16- Auto Regressive
 - 17- Extended Sample ACF
 - 18- Generalized Sample PACF
 - 19- Inverse ACF
 - 20- Canonical Correlations
 - 21- Auto Regressive Moving Average
 - 22- Auto Regressive Integrated Moving Average
 - 23- Ordinary Least Squares
 - 24- Weighted Least Squares
-