

عنوان مقاله:

مطالعه ی اهمیت پیش پردازش داده ها در پردازش سری زمانی با استفاده از؟؟؟ با استفاده از یک مورد مطالعه

محل انتشار:

پنجمین کنفرانس بین المللی توسعه فناوری در مهندسی برق ایران (سال: 1400)

تعداد صفحات اصل مقاله: 15

نویسندگان:

زهرا نریمانی - استادیار، دانشکده علوم رایانه و فن آوری اطلاعات، دانشگاه تحصیلات تکمیلی علوم پایه زنجان

امیر حسین آدینه - فارغ التحصیل دوره کارشناسی، دانشکده علوم رایانه و فن آوری اطلاعات، دانشگاه تحصیلات تکمیلی علوم پایه زنجان

خلاصه مقاله:

تجزیه و تحلیل داده های سری زمانی در دهه های اخیر، از اهمیت ویژه ای برخوردار بوده است. حوزه های مختلف مانند تجزیه و تحلیل داده های مالی، تجزیه و تحلیل داده های بیولوژیکی و تشخیص گفتار ذاتا با سیگنال های وابسته به زمان سر و کار دارند. نظارت بر رفتار گذشته سیگنال ها کلید پیش بینی دقیق رفتار یک سیستم در آینده نزدیک است. در سناریوهایی مانند پیش بینی داده های مالی، سیگنال غالب دارای یک رفتار تناوبی (از ابتدای ماه، هفته و غیره) است و یک روند کلی و رفتار فصلی را نیز می توان برای آن در نظر گرفت. مدل میانگین متحرک خودهمبسته یکپارچه؟؟؟ و همچنین نوع فصلی آن؟؟؟ در پیش بینی داده های سری زمانی بطور گسترده ای مورد استفاده قرار گرفته و همچنین قادر به مدلسازی روند و رفتار فصلی در داده ها است. اگرچه ممکن است رفتار داده ها به صورت خودهمبسته باشد و روند و رفتار فصلی توسط؟؟؟ شناسایی و اداره شود، اما ویژگی های داده های واقعی لزوما دقیقا با مفروضات (؟؟؟ یا به طور کلی؟؟؟) سازگار نیست. علاوه بر این، در این مدلها داده کامل (بدون دادهی گمشده) فرض شده در حالی که در دنیای واقعی، همیشه می توان به دلایل مختلف مانند تعطیلات که برای آنها امکان ثبت داده وجود ندارد، داده های گمشده وجود داشته باشد. همچنین ساعات کاری مختلف در روزهای متفاوت هفته می تواند الگوی نامنظمی را در دادهها ایجاد کند که در؟؟؟ در نتیجه در تنظیم پارامترها دچار گمراهی شویم. در این مقاله، اثربخشی استفاده از؟؟؟ در داده های دنیای واقعی را با استفاده از یک مورد کاربردی مورد بررسی قرار داده و نشان می دهیم چگونه روش های پیش پردازش می توانند به منظور سازگار کردن داده ها برای بهبود مدل؟؟؟ اعمال شوند. داده های پژوهش حاضر از معاملات یک شرکت سرمایه گذاری صندوق می باشد که حاوی مقادیر از دست رفته (تک نقطه های و بازهای) و همچنین بی نظمی در تعداد ساعات کار در روز در هفته است. بدون پیش پردازش، داده های بی کیفیت منجر به تضعیف مدل ایجاد شده می شود. علاوه بر مشکل وجود دادهی بیکیفیت، تعداد نقاط داده در زمان تجزیه و تحلیل برای ساختن مدل؟؟؟ کافی نبود. در این پژوهش، مراحل پیش پردازش مانند پر کردن مقادیر گمشده و ترفندهایی برای سازگاری داده ها با مدل؟؟؟ ارائه شده است. نتایج نشان می دهد که عملکرد پیش بینی؟؟؟ در این مجموعه از داده های دنیای واقعی با استفاده از چندین مرحله پیش پردازش معرفی شده به منظور مقابله با شرایط ذکر شده به طور قابل توجهی بهبود می یابد. مراحل پیش پردازش پیشنهادی را می توان در دیگر تجزیه و تحلیل داده های سری زمانی واقعی نیز استفاده کرد.

کلمات کلیدی:

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1232089>

