

عنوان مقاله:

کاهش بعد داده ها به شیوه غربالگری

محل انتشار:

چهارمین کنفرانس بین المللی محاسبات نرم (سال: 1400)

تعداد صفحات اصل مقاله: 7

نویسنده:

محمد کاظمی - استادیار گروه آمار، دانشگاه گیلان

خلاصه مقاله:

در تحلیل داده های با بعد بسیار بالا، شناسایی متغیرهای توضیحی موثر بر پاسخ، از اهمیت بسزایی برخوردار است. در این نوع داده ها، ابتدا با استفاده از یک روش غربالگری تعداد متغیرهای توضیحی را کاهش داده و سپس از روش های انتخاب متغیر مبتنی بر جریمه برای انتخاب مدل نهایی استفاده می کنیم. در این مقاله به مروری بر چند روش غربالگری رایج پرداخته می شود. این روش ها شامل غربالگری مستقل مطمئن (SIS)، غربالگری و رتبه بندی مستقل مطمئن (SIRS) و غربالگری مستقل مطمئن براساس همبستگی فاصله ای (DC-SIS) می باشند. پس از معرفی این روش ها، با استفاده شبیه سازی، توانایی آنها در انتخاب متغیرهای مهم با هم مقایسه می کنیم. نتایج مطالعات شبیه سازی نشان می دهد که روش DC-SIS در مقایسه با دو روش دیگر، از قابلیت بیشتری در تشخیص متغیرهای مهم برخوردار است، به ویژه هنگامی که برخی از متغیرهای توضیحی دارای اثر غیرخطی بر متغیر پاسخ باشند

کلمات کلیدی:

انتخاب متغیر، داده های بعد بالا، غربالگری مستقل مطمئن، کاهش بعد، رتبه بندی متغیرها

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1418689>

