

انتخاب ویژگی مبتنی بر همبستگی پیرسون با الگوریتم ژنتیک

مهدی صادق زاده^{۱۰۲}، محمد تشنه لب^{۱۰۳}

چکیده

انتخاب ویژگی اخیراً موضوع تحقیقات زیادی در داده کاوی خصوصاً برای مجموعه های داده با تعداد زیادی ویژگی یا صفات خاصه بوده است. کارهای اخیر نشان می دهد که انتخاب ویژگی می تواند تاثیر مثبتی بر روی کارایی الگوریتم های یادگیری ماشین داشته باشد. موفقیت بسیاری از الگوریتم های یادگیری در کوشش آنها برای ساخت مدلی از داده است که وابسته به شناسایی مطمئن مجموعه کوچکی از صفات خاصه تخمینی است. دارا بودن صفات خاصه نامربوط و زاید در مرحله فرآیند ساخت مدل می تواند به کارایی تخمین ضعیف و محاسبات زیاد منجر شود. این مقاله الگوریتمی مبتنی بر همبستگی پیرسون را توصیف می نماید که از الگوریتم های ژنتیک برای مسائل طبقه بندی جهت تعیین مطلوبیت زیر مجموعه های ویژگی ها استفاده می کند، و تاثیر آن را بر روی چندین الگوریتم یادگیری ماشین رایج ارزیابی می نماید.

کلمات کلیدی

انتخاب ویژگی، الگوریتم های ژنتیک، داده کاوی

¹⁰² دانشجوی دکترای تخصصی مهندسی کامپیوتر، گرایش نرم افزار، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، و عضو هیئت علمی دانشگاه آزاد اسلامی واحد ماهشهر،

sadegh_1999@yahoo.com

¹⁰³ استادیار دانشکده برق دانشگاه خواجه نصیر الدین طوسی