



# کنگره بین المللی علوم و مهندسی

آلمان - هامبورگ

اسفند ماه ۱۳۹۶

## کاهش ابعاد ویژگی مبتنی بر تئوری راف، با استفاده از الگوریتم رقابت لیگ فوتبال

علی دلاکه<sup>۱</sup>، کاظم نیکفرجام<sup>۲\*</sup>، هادی چهکندی نژاد<sup>۳</sup>

۱- دانشجوی کارشناسی ارشد، گروه مهندسی کامپیوتر، واحد بیرجند، دانشگاه آزاد اسلامی، بیرجند، ایران، SA92dallakehali@iaubir.ac.ir

۲- دکترای کامپیوتر، گروه مهندسی کامپیوتر، واحد بیرجند، دانشگاه آزاد اسلامی، بیرجند، ایران، nikfarjam@iaubir.ac.ir

۳- دکترای برق، گروه مهندسی برق، واحد بیرجند، دانشگاه آزاد اسلامی، بیرجند، ایران، Hchahkandin@birjand.ac.ir

### چکیده

افزایش حجم پایگاه داده ها، مسئله کاهش ویژگی را ایجاد و پُر اهمیت میکند. از اینرو در علم داده کاوی یکی از زیر بخش های مهم مسئله کاهش ابعاد، همانا انتخاب ویژگی میباشد، که در آن دسته از ویژگی های ورودی که بیشترین تاثیر را بر روی خروجی کار دارند برگزیده و سایر ویژگی هایی که تاثیر چندانی در نتیجه ندارند در محاسبات لحاظ نمیشوند. از جمله تئوری هایی که در راستای حل مسائل کاهش ویژگی مورد استقبال قرار گرفته اند همانا تئوری راف میباشد؛ اما از آنجاییکه تئوری راف ذاتاً از دقت بالایی برخوردار بوده، این خصیصه سبب میشود که سرعت آن به طرز چشمگیری در حل مسائل بزرگ کاهش یابد، از اینرو مسئله کاهش ویژگی به وسیله تئوری راف، عملاً مسئله ای NP-Hard تلقی میشود، فلذا پژوهشگران این حوزه برای حل این معضل به روشهای فراابتکاری متوسل شده که تا حد امکان از این پیچیدگی زمانی کاسته و بتوانند گامی مؤثر در راستای حل اینگونه مسائل اتخاذ نمایند. برای این منظور با توجه به توانایی های بالقوه الگوریتم رقابت لیگ فوتبال در حل مسائل NP-Hard، چابکی بی نظیر، سرعت، پویایی و پتانسیل بالای آن در حل مسائل محاسباتی سنگین، ما را برآن داشت که از آن در راستای حل مسئله کاهش ویژگی بهره گیریم تا هم گامی مؤثر در راستای حل مسئله برداشته باشیم و هم محکی دیگر بر ظرفیت های بیشمار این الگوریتم داشته باشیم. در پایان بمنظور سنجش میزان قدرت واقعی این روش پیشنهادی و نیز داشتن قیاسی منصفانه و علمی؛ این الگوریتم را بر روی ۹ دیتاست معتبر مرسوم این حوزه پیاده سازی کرده و نتایج حاصله از کاهش ویژگی به کمک این روش را به لحاظ تعداد ویژگی های کاهش یافته با نتایج سایر الگوریتم های مدّعی در این زمینه مورد مقایسه قرار داده ایم تا قدرت و توان فوق العاده این روش در راستای حل مسئله کاهش ویژگی در برابر رقبایش کاملاً مشهود و مبرهن باشد.

**واژه های کلیدی:** داده کاوی، کاهش ویژگی، انتخاب ویژگی، تئوری مجموعه راف کلاسیک، الگوریتم رقابت لیگ فوتبال

### ۱- مقدمه

بخش عظیمی از بانک های اطلاعاتی، حاوی انواع مختلفی از داده ها می باشند. از جمله عملیات ضروری در این حوزه، کاهش ویژگی می باشد که با هدف کاهش تعداد ویژگی های دیتا و نیز کمترین داده از دست رفته صورت می پذیرد. علت استفاده از کاهش ویژگی؛ حذف ویژگی های غیر ضروری و ناکارآمد دیتا میباشد که تاثیر چندانی در خروجی کار ندارند. عبارتی، ویژگی های مهم و مؤثر در کاهش ابعاد پایگاه داده حفظ شده و سایر ویژگی های ناکارآمد حذف می شوند. بعلاوه اینکار به افزایش راندمان الگوریتم نیز کمک بسزایی خواهد کرد. شاید این موضوع این برداشت را در نگاه اول ایجاد کند که هرچه تعداد ویژگی ها بیشتر باشد، به همان میزان از دقت محاسبه بیشتری برخوردار خواهیم بود. حال اینکه افزایش تعداد ویژگی ها همیشه به معنی افزایش دقت محاسباتی نیست و در بسیاری مواقع به دلیل اهمیت پایین، دست و پا گیر محسوب