

## چکیده

امروز دسته بندی صفحات وب یکی از تکنیکهای پایه ای برای دسته بندی اطلاعات بر روی وب است اما مسئله اصلی این است که نمیتوان با روشهایی که بر اساس تحلیل و پردازش محتوا کار میکنند حجم زیادی از صفحات وب را دسته بندی نمود و دلیل اصلی این امر هم از آنجایی است که هزینه محاسباتی و پردازشی برای پردازش متن صفحات بسیار بالاست همین امر باعث میشود که روشهای ارائه شده هیچ کدام کارایی لازم را برای دسته بندی صفحات نداشته و از طرف دیگر یکی از اصلی ترین اهدافی که دسته بندی صفحات وب دنبال میکند ارائه اطلاعات بر اساس سلايق و علايق کاربران است.

به همین دلیل دقت در دسته بندی صفحات وب از مسایل بسیار مهم است و اثر مستقیم در دقت ارائه اطلاعات به کاربران را دارد به همین دلیل در این تحقیق رویکرد جدیدی برای دسته بندی صفحات وب با استفاده از معیار رفتارهای پیمایشی کاربران ارائه شده است. و بدین وسیله میزان هزینه های محاسباتی را برای اطلاعات مطلوب با تغییر فضای ویژگی کلمات موجود در صفحات با رفتارهای پیمایشی کاربران جهت دسته بندی صفحات در گروه های هم علاقه کاربران را انجام شده است و چون در این دسته بندی از شیوه دسته بندی بر اساس گروههای هم علايق کاربران انجام شده و ارائه اطلاعات به کاربران هم علاقه بسیار سریعتر و با هزینه محاسباتی و زمانی کمتر انجام میگيرد همانطور که در نتایج این تحقیق مشخص است دقت دسته بندی با استفاده از تابع هزینه جدید ارائه شده به وسیله الگوریتم ژنتیک و الگوریتم مورچگان موازی و سری حل شد که نتایج با استفاده از الگوریتم ژنتیک بسیار بهتر از نتایج با دو الگوریتم مورچگان موازی و سری است.

کلید واژه: وب کاوی، داده کاوی، دسته بندی صفحات وب، الگوریتم مورچگان، الگوریتم ژنتیک.