



پایان‌نامه‌ی کارشناسی ارشد: اعظم خلقی، ۱۳۹۷

ترکیب الگوریتم‌های تکاملی با مجموعه‌های راف جهت انتخاب ویژگی بهینه

امروزه مجموعه داده‌ها حجیم هستند و به دلیل صرفه‌جویی در منابع انرژی نیاز داریم که بار داده‌ها را کم کنیم. در داده‌کاوی روش‌های زیادی برای کاهش ابعاد داده‌ها وجود دارد. در این تحقیق یک استراتژی جدید انتخاب ویژگی را بر مبنای مجموعه‌های راف و بهینه‌سازی ازدحام ذرات پیشنهاد می‌کنیم. مجموعه‌های راف به عنوان یک روش انتخاب ویژگی بسیار موفق بوده، اساس ریاضی آن استفاده از ماتریس تمایز می‌باشد و همچنین تکنیکی برای رفع ابهام در داده‌ها است. اما رویکردهای تپه‌نوردی مجموعه‌های راف برای انتخاب ویژگی در پیدا کردن روش‌های کاهش ویژگی که به عنوان یک روش هیوریستیک بهینه بتواند مورد استفاده واقع شود، کافی نیست. از طرف دیگر جستجوهای کامل برای هر دیتاست دارای اندازه‌ی متوسط امکان‌پذیر نیست. بنابراین رویکردهای تصادفی یک مکانیزم انتخاب ویژگی قابل قبول را فراهم می‌کنند. مانند الگوریتم‌های ژنتیک و PSO که در حل یک مسئله بالقوه به عنوان تکنیک محاسبه تکاملی جدید است، به طوری که یک ذره با سرعت مشخصی در بین فضای مسئله در حرکت می‌باشد. ذرات پراکنده ناحیه‌های بهینه را از بین فضای جستجوی پیچیده از طریق تعامل ذرات در جمعیت پیدا می‌کنند. در این تحقیق انبوه ذرات بهترین ترکیب ویژگی که در فضای زیرمجموعه معلق باشند را پس از اعتبارسنجی توسط تابع هزینه که در اینجا مجموعه‌های راف است، پیدا خواهند کرد. آزمایشات با استفاده از 11 دیتاست UCI انجام شده که الگوریتم پیشنهادی را با رویکرد بر مبنای ژنتیک و راف مقایسه می‌کند. نتایج نشان می‌دهد که PSO برای کاهش ویژگی بر مبنای مجموعه‌های راف مناسب می‌باشد.

کلیدواژه‌ها: انتخاب ویژگی، مجموعه‌های راف، بهینه‌سازی ازدحام ذرات، الگوریتم ژنتیک، کاهش ویژگی

شماره‌ی پایان‌نامه: ۱۲۷۴۱۰۰۶۹۶۱۰۰۲

تاریخ دفاع: ۱۳۹۷/۰۳/۱۰

رشته‌ی تحصیلی: مهندسی کامپیوتر - نرم‌افزار

دانشکده: فنی و مهندسی

استاد راهنما: مهندس حسام حسن‌پور

استاد مشاور: مهندس کبری بهروان

M.A. Thesis:

combination the evolutionary algorithms with rough sets for optimal attribute selection



Nowadays datasets are huge; and for saving the energy we need to less the capacity of information. In datamining there are lots of methods for reducing data dimensionality. We propose a new feature selection strategy based on rough sets and Particle Swarm Optimization (PSO). Rough sets has been used as a feature selection method with much success, that its mathematical basis is discernibility matrix; There is also a technique for removing vagueness and uncertainty in data. but current hill-climbing rough set approaches to feature selection are inadequate at finding optimal reductions as no perfect heuristic can guarantee optimality. On the other hand, complete searches are not feasible for even medium-sized datasets. So, stochastic approaches provide a promising feature selection mechanism. Like Genetic Algorithms, PSO is a new evolutionary computation technique, in which each potential solution is seen as a particle with a certain velocity flying through the problem space. The Particle Swarms find optimal regions of the complex search space through the interaction of individuals in the population. PSO is attractive for feature selection in that particle swarms will discover best feature combinations as they fly within the subset space. In this work particle swarm will discover the best combination of features flying in the feature space, after evaluating by cost function named rough set. Experimentation is carried out, using UCI data, which compares the proposed algorithm with a GA-based approach and other deterministic rough set reduction algorithms. The results show that PSO is efficient for rough set-based feature selection.