



پایان‌نامه‌ی کارشناسی ارشد: جلال طیبی، ۱۳۹۸

یافتن مجموعه بهینه از ویژگی‌ها با استفاده از روش‌های یادگیری تقویتی

تشخیص بهترین ویژگی‌ها برای یک مجموعه داده، تاثیر زیادی در به حداقل رساندن خطای دسته بندی دارد. انتخاب ویژگی، فرآیند انتخاب زیرمجموعه‌ای کوچک از مجموعه‌ی بزرگ ویژگی‌ها است. در بسیاری از روش‌های انتخاب ویژگی، با جستجو در فضای حالت، بهترین زیرمجموعه از ویژگی‌ها انتخاب می‌شوند. در این مقاله، از روش یادگیری تقویتی، به منظور انتخاب ویژگی‌ها استفاده می‌شود. این روش، فضای حالت مساله را جستجو کرده و از تجربیات گذشته استفاده می‌کند. با استفاده از روش اختلاف زمانی، فضای حالت مساله پیمایش شده و بهترین زیرمجموعه از ویژگی‌ها انتخاب می‌شود. در ابتدا، مساله مانند یک فرآیند تصمیم مارکوف در نظر گرفته می‌شود. سپس از یک گراف جستجو بهینه برای حل مشکل پیچیدگی مساله استفاده می‌شود. برای بررسی حالات ممکن، ارائه یک تابع ارزیابی با پیچیدگی کم ضروری است. در این مقاله از روش SVM، k-means و کنزدیک ترین همسایه استفاده شده است.

کلیدواژه‌ها: انتخاب ویژگی، یادگیری تقویتی، اختلاف زمانی، SVM، گراف تصمیم

شماره‌ی پایان‌نامه: ۱۲۷۴۱۰۰۶۹۶۱۰۰۷

تاریخ دفاع: ۱۳۹۸/۱۰/۲۵

رشته‌ی تحصیلی: مهندسی کامپیوتر - نرم‌افزار

دانشکده: فنی و مهندسی

استاد راهنما: دکتر حسام حسن پور

استاد مشاور: مهندس علی اکبر نقابی

M.A. Thesis:

using reinforcement learning to find an optimal set of features

Detecting the best features for a data set has a great impact in minimizing the classification error. Feature selection is the process of selecting a small sub set of a large set of features. In many feature selection methods, the best sub set of features are selected by searching in the state space. In this paper, the reinforcement learning method is used to select features. This method searches for problem state space and draws on past experiences. Using the time difference method, the problem state space is scrolled and the best subset of features is selected. Initially, the problem is seen as a Markov decision process (MDP). Then an optimal search graph is used to solve the problem of complexity. To evaluate possible scenarios, it is necessary to provide an evaluation function with low complexity. In this paper,



SVM, k-means and k nearest neighbor methods are used. The k nearest neighbor was the best
