

عنوان مقاله:

طبقه بندی مجموعه داده های بزرگ با استفاده از SVMs

محل انتشار:

اولین کنفرانس ملی نوآوری در مهندسی کامپیوتر و فناوری اطلاعات (سال: 1392)

تعداد صفحات اصل مقاله: 37

نویسنده:

راضیه نصرت زاده

خلاصه مقاله:

ماشین بردار پشتیبان (SVMs) از روش های خوب برای طبقه بندی و تجزیه و تحلیل رگرسیون می باشند به دلیل توابع ریاضی که انتقال چند خواص برجسته است که روشهای دیگر به سختی فراهم می کند. با این حال، با وجود خواص برجسته SVMs آنها برای داده های بزرگ در مقیاس مورد علاقه یادگیری برای به رسمیت شناختن الگو و یا دستگاه به دلیل پیچیدگی آموزش SVMs به شدت به اندازه مجموعه داده وابسته است. بسیاری از برنامه های کاربردی داده کاوی در دنیای واقعی شامل میلیون ها یا میلیاردها پرونده داده که در آن حتی اسکن چندگانه تمام داده ها بیش از حد گران است انجام می شوند. این مقاله یک روش جدید، مبتنی بر خوشه SVM (CB-SVM) است که به طور خاص برای کار با مجموعه داده های بسیار بزرگ طراحی شده است. CB-SVM اعمال می شود یک الگوریتم میکرو خوشه بندی سلسله مراتبی است که تنظیم اسکن تمام داده ها فقط یک بار به ارائه SVM با نمونه های با کیفیت بالا که حمل خلاصه آماری از SVM داده به طوری که خلاصه به حداکثر رساندن سود از یادگیری CB-SVM. تلاش می کند به تولید بهترین مرز SVM را برای داده های بسیار بزرگ مجموعه گفته شده مقدار محدود از منابع است. آزمایشات ما بر روی داده های واقعی و مصنوعی مجموعه نشان می دهد که CB-SVM بسیار مقیاس پذیر است برای مجموعه داده های بسیار بزرگ و برای تولید طبقه بندی با دقت بالا.

کلمات کلیدی:

ماشین های بردار پشتیبان، کلاستر، خوشه بندی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/262603>

