

عنوان مقاله:

مروری بر مدل های داده های متفاوت در مواجهه با شناسایی داده های پرت

محل انتشار:

چهارمین کنفرانس ملی مهندسی برق و کامپیوتر سیستم های توزیع شده و شبکه های هوشمند و نخستین کنفرانس ملی مهندسی پزشکی (سال: 1396)

تعداد صفحات اصل مقاله: 9

نویسندگان:

مریم صمیمیت - دانشجوی کارشناسی ارشد کامپیوتر، دانشگاه آزاد اسلامی واحد نراق،

مهدی اسماعیلی - عضو هیئت علمی دانشگاه آزاد اسلامی واحد کاشان،

خلاصه مقاله:

شناسایی داده های پرت فرآیندی است برای یافتن اشیایی که رفتار آنها بسیار متفاوت تر از انتظار است. چنین اشیایی داده های پرت یا آنومالیا نامیده میشوند. تشخیص داده های پرت علاوه بر شناسایی کلاهبرداریها، در بسیاری از کاربردها نظیر پزشکی، امنیت، آشکارسازی خرابی های صنعتی، پردازش تصویر، نظارت شبکه های حسگر/ویدیویی و آشکارسازی مزاحمت استفاده میشود. داده های پرت از این نظر مورد توجه قرار گرفته اند که شک داریم توسط مکانیزمی مشابه آنچه که مابقی داده ها تولید شده اند آنها نیز تولید شده باشند. از اینرو در شناسایی داده های پرت این موضوع مهم است که مشخص کنیم چرا داده های پرت تشخیص داده شده با مکانیزم های دیگری تولید شده اند. اغلب برای پاسخ به این پرسش فرضیات متعددی برای مابقی داده ها در نظر گرفته میشود و مهمترین آن مدل داده هایی است که مجموعه ما از آن پیروی میکند. در این مقاله مدلهای متفاوتی برای داده ها بیان میشود و طبیعی است که بر اساس مدل مزبور و همچنین پارامترهای آن، داده های پرت از میان داده های نرمال تشخیص داده میشوند.

کلمات کلیدی:

خوشه بندی، داده پرت، داده کاوی، دسته بندی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/768159>

