

عنوان مقاله:

استفاده از دسته بندهای تجمیعی برای دادههای نامتوازن و کاربرد آنها در دادههای مامایی

محل انتشار:

دومین کنفرانس ملی مهندسی نرم افزار دانشگاه آزاد لاهیجان (سال: 1391)

تعداد صفحات اصل مقاله: 6

نویسنده:

خلاصه مقاله:

در این مقاله، مسئله یادگیری دادههای بهشدت نامتوازن، مورد مطالعه قرار گرفته است. دادههای نامتوازن به این معناست که دستههای هدف از دادهها در جدول توزیع منحرف میشوند؛ حداقل، نمونههای یک دسته بهطور قابل توجهی بیشتر از دستههای دیگر وجود دارد. روشهای دستهبندی سنتی، بهدنبال بهدافلسانیدن میزان کلی خطا از کل مجموعهی آموزش، بر روی دادههای نامتوازن بهخوبی عمل نمیکند، زیرا آنها معمولاً فرض میکنند توزیع یک کلاس نسبتاً متوازن است و انرژی زیادی بر روی یادگیری دادههای نامتوازن کلاس اکثریت میگذارند که این موضوع اهمیت زیادی دارد و در بسیاری از برنامههای کاربردی چالشبرانگیز است. مجموعه داده مورد استفاده در این مقاله از 7272 نوزاد که در بیمارستان شهید رجایی شهرستان آران و بیدگل استان اصفهان که در بین سال های 1832 و 1837 متولد شدهاند، جمعآوری شده است. بهطور کلی، برخورد با کلاس اقلیت نیاز به مفاهیم جدید، بررسیها و راهحلهایی بهمنظور درک کامل از مدلهای اساسی پیچیده دارد. در این مقاله، توجه بخصوصی به مسئله دستهبندی دودویی نامتوازن شده است و چندین روش یادگیری تجمیعی موثر برای حل این مسئله پیشنهاد شده است

کلمات کلیدی:

داده های نامتوازن، داده های مامایی، دسته بندی، روشهای تجمیعی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/184831>

