چکیده:

وسیع بودن حوزه سلامت و بیمه مراقبت‌های بهداشتی و حجم زیاد مالی باعث شده تا این حوزه برای کلاهبرداری مورد هدف قرار بگیرد. یکی از منابع هزینه‌های قابل توجه سازمان بهداشت و سلامت، پرداخت سهم بیمه داروهای تجویز شده برای بیماران تحت پوشش است. بطور كلي، هدف از تشخيص تقلب، به حداكثر رساندن پیشبيني‌هاي درست و حفظ پیشبيني‌هاي نادرست در يك سطح قابل قبول از هزينه ميباشد. با وجود تغييرات پيوسته رفتار متقلبين، مدل‌هايي كه براساس تحليل داده‌هاي گذشته ساخته ميشوند ممکن است نتوانند شكل‌هاي جديد تقلب را شناسايي كنند. همچنین، هيچ يك از سيستم‌هاي شناسايي تقلب به تنهايي نميتواند تمام شكل‌هاي تقلب را شناسايي و پوشش دهد. در این پایان‌نامه یک رویکرد مبتنی بر همسایگی گره برای تخمین احتمال تقلب در اسناد درماني مورد بررسي قرار گرفته است. مدل ارایه شده از دو ایده پایه که در نتیجه ساده‌سازی روش‌های مورد استفاده در بسیاری از تحقیق‌های دیگر است، بهره می‌برد: هر ارایه‌دهنده خدمات بهداشت و سلامت به کلیه اعضای مجموعه مرجع نگاشت و بعنوان یک گره از گراف درنظر گرفته شد وشباهت‌های رفتاری در دو دسته‌ی ارائه‌دهندگان مراقبت‌های بهداشتی کلاهبردار و غیرکلاهبردار با توجه به معیارهای قابل اندازه‌گیری فعالیت‌های مراقبت‌های بهداشتی مانند روش‌های پزشکی و تجویز داروها، محاسبه شد. به یال میان هر دو راس گراف وزنی بر اساس شباهت محاسبه شده اختصاص داده شد. در ادامه از k همسایگی برای حذف اثر داده پرت استفاده شد. ارزیابی مدل با توجه به توانایی آن‌ها در پیشبینی حضور یک ارائه‌دهنده در لیست ارائه‌دهندگان دفتر بازرسی کل (محرومیت از مشارکت در بیمه پزشکی سالمندان و سایر برنامه های مراقبت های بهداشتی فدرال)، انجام شده‌است.