

چکیده

رسانه اجتماعی به عنوان بستری برای سخنرانی آنلاین گسترش یافته است که افراد محتوا را ایجاد می کنند، آن را به اشتراک می گذارند، آن را نشانه گذاری می کنند، و با نرخ حیرت آوری آن را شبکه ای می کنند. فیس بوک و توئیتر مثال هایی از رسانه اجتماعی می باشند. در رسانه اجتماعی، حجم و واریانس بالای اطلاعاتی که از طریق انجمن های با کاربران زیاد منتشر می شود، یک فرصت طلایی برای کنترل این داده ها مهیا می کند و امکان پیش بینی نتایج خاص را بدون نیاز به تأسیس مکانیزم های بازاریابی فراهم می سازد. بنابراین می توان مدل هایی را برای تجمیع عقاید و افکار جمعیت جمعی و حصول بینش های درونی سودمند در رفتار آن ها ایجاد کرد، در حالی که امکان پیش بینی گرایش آینده آن ها نیز وجود داشته باشد. این پایان نامه به بررسی چنین موضوعاتی می پردازد. هدف اصلی این پایان نامه بررسی تکنیک های داده کاوی در شبکه های اجتماعی با رویکرد احساسات می باشد. به طور خاص کار پیش بینی درآمد باکس-آفیس برای فیلم ها را با استفاده از گردآوری اطلاعات از توئیتر که یکی از بزرگترین و رو به رشدترین شبکه های اجتماعی در اینترنت می باشد، انجام می دهیم. برای تحلیل احساسات معیارهای مختلف ذهنیت و قطبیت را بکار می گیریم و همچنین الگوریتمی را برای طبقه بندی فیلم ها پیشنهاد می نماییم. با استفاده از مدل رگرسیون خطی مدل های مختلفی برای پیش بینی درآمد فیلم ها ایجاد شده که هر یک از آنها از متغیرهای مختلفی استفاده می کنند. آزمایشات انجام شده نقش قابل توجه تحلیل احساسات در پیش بینی دقیق درآمد فیلم و بهبود نتایج نسبت به کارهای پیشین را نشان می دهند.

کلمات کلیدی: رسانه اجتماعی، تحلیل احساسات، داده کاوی، توئیتر، مدل رگرسیون خطی