

۵-۳- کارهای آتی

اگر چه نتایج ارزیابی روش پیشنهادی به منظور شناسایی رویداد شخصی آینده در مقایسه با روش‌های مرجع رضایت‌بخش و قابل قبول بوده است، اما برای بهبود و توسعه‌ی آن راه‌کارهایی وجود دارد. در این بخش، مهمترین اقداماتی که در ادامه مسیر این رساله قابل انجام است ذکر شده است.

۵-۳-۱- استفاده از ویژگی تعاملات در شناسایی رویداد شخصی آینده

در روش پیشنهادی، از ویژگی WMD برای آموزش رده‌بند رویداد شخصی استفاده شده است که قابلیت ضبط ویژگی‌های معنایی و ساختاری توییت‌ها را دارد، به عبارت دیگر برای ساخت تاریخچه کاربر تنها از دنباله توییت‌های خود کاربر استفاده شده است. با توجه به ماهیت شبکه‌های اجتماعی و ویژگی‌های آن می‌توان به مدل‌سازی مبتنی بر تعاملات اشاره کرد که در این کار تحقیقاتی بکار گرفته نشده است و با توجه به مطالعات انجام شده، ویژگی تعاملات می‌تواند نقش بسزایی در شناسایی رویداد شخصی داشته باشد [۸۹]. بنابراین می‌توان روش پیشنهادی را به یک روش ترکیبی گسترش داد که در آن رده‌بند رویداد شخصی علاوه بر WMD از ویژگی تعاملات نیز برای آموزش استفاده نماید.

۵-۳-۲- استفاده از ویژگی‌های شخصی کاربر

در روش پیشنهادی تنها از داده‌های متنی که کاربر در خط- زمانی خود منتشر کرده است، استفاده شده است. در حالی که می‌توان اطلاعات شخصی مانند تاریخ تولد، مکان تولد و یا جنسیت را از پروفایل کاربر استخراج و در مسأله شناسایی رویداد شخصی آنها را بکار برد. مثلاً دانستن سن کاربر در شناسایی نوع رویداد شخصی مانند ازدواج و یا ورود به دانشگاه و یا اطمینان از صحت شناسایی می‌تواند موثر باشد. علاوه بر اطلاعات شخصی، می‌توان از سایر اطلاعات توییت همانند مکان نوشتن توییت نیز استفاده کرد.

۵-۳-۳- رفع خطاهای گزارش شده در مدل دنباله به دنباله چند-به-چند

بر اساس سه نوع خطای شناسایی شده از مدل دنباله به دنباله چند-به-چند که جهت تولید محتوا پیشنهاد شده است، می‌توان موارد زیر را در جهت رفع آنها لحاظ کرد و مدل گسترش داد:

الف) کارایی مدل دنباله به دنباله چند-به-چند وابسته به کارایی رده‌بندی است که در لایه زیرین به منظور شناسایی رویداد از توییت بکار رفته است. در حالی که می‌توان در کنار معماری دنباله به دنباله، مکانیزمی برای یادگیری بازنمایی رویداد شخصی بکار برد. به بیان دیگر، بجای اعمال یک مرحله پیش‌پردازش برای تعیین رویداد ذکر شده در توییت، از یک معماری رمزگذار خودکار^۱ برای یادگیری بازنمایی چگال‌تر رویداد شخصی استفاده کرد که این خود دو مزیت به همراه دارد: (۱) وابستگی به رده‌بند رویداد شخصی از بین می‌رود، (۲) فرایند یادگیری بازنمایی رویداد شخصی توسط داده‌های آموزشی موجود انجام می‌شود.

ب) توییت‌های فراوانی وجود دارند که بخش‌های اطلاعاتی مشابه با رویداد شخصی دارند در حالی که رویداد شخصی (خود-گزارشی) Self-Report را گزارش نمی‌کنند. استفاده از یک رده‌بند دودویی جهت تعیین خود-گزارشی هر توییت گذشته کاربر می‌تواند به تعیین این موضوع کمک کند که آیا کاربر درگیر رویداد شخصی (خود-گزارشی) Self-Report خواهد بود یا نه؟ در واقع اگر کاربر در توییت‌های گذشته خود در یک الگو (خود-گزارشی) Self-Report شرکت داشته باشد، احتمال آنکه در آینده نیز درگیر رویداد (خود-گزارشی) Self-Report باشد افزایش می‌یابد. از این رده‌بند دودویی می‌توان به صورت مستقل به منظور اعمال فیلترینگ استفاده کرد.

ج) شبکه‌های عصبی بازگشتی، وابستگی‌های کوتاه مدت را بهتر مدل می‌کنند هرچند در تئوری نشان داده شده است که این شبکه‌ها قادرند هر نوع وابستگی را ضبط و نگهداری کنند [۱۳۵]. بر همین اساس، در مدل

^۱ Autoencoder

پیشنهادی، آخرین توییت‌ها اثر بیشتری در دنباله تولیدی خروجی دارند که از آن برای تعیین رویداد شخصی آینده کاربر استفاده می‌شود. روش‌های جدیدی همچون Hacks [۱۳۶] برای حفظ ترتیب ورودی‌ها و مدل کردن وابستگی‌های طولانی مدت در یک وابستگی کوتاه مدت معرفی شده‌اند. همچنین استفاده از روش‌های سیستماتیک دیگری مانند مدل‌های توجه [۱۳۷] نیز بسیار رایج است که می‌تواند تعیین کند چه بخشی از توییت‌های گذشته یا عبارت‌های موجود در آنها بهتر است در شبکه عصبی بازگشتی در نظر گرفته شوند و با کم توجه کردن به اطلاعات نامرتبط دقت شناسایی رویداد آینده را افزایش داد.

۵-۳-۴- تعمیم روش پیشنهادی برای شناسایی رویدادهای محلی

رویداد محلی مانند فروش و یا ارتقا در یک فروشگاه محلی که خبر آن در رسانه‌های اجتماعی مانند توییتر منعکس می‌شود، لزوماً شخصی نیستند و توسط یک کاربر تعیین نمی‌شوند. با توجه به شباهت رویدادهای محلی و رویدادهای شخصی (حجم کم پیام‌ها و مسأله عدم تعادل)، می‌توان از روش پیشنهادی برای شناسایی آنها استفاده کرد.