

امروزه با افزایش استفاده از اینترنت و گسترش شبکه‌های اجتماعی، انجمن‌ها، صفحات وب و ... خبرگزاری‌های اینترنتی نیز افزایش یافته‌اند این خبرگزاری‌ها برخلاف روزنامه‌ها که فقط نظر نویسندگان را منتشر می‌کردند، دیدگاه‌های کاربران در مورد هر موضوع را نیز در پایان مطلب منتشر می‌کنند. دیدگاه‌ها در مورد مطالب منتشر شده، منابع اطلاعاتی ارزشمندی محسوب می‌شوند زیرا با توجه به تعداد دیدگاه‌های منتشر شده که با جمله‌ای در خبر ارتباط دارند، اهمیت و بحث‌برانگیز بودن جمله خبری را در می‌یابیم. در این تحقیق قصد داریم خلاصه‌ای از خبرها را تولید کنیم به همین منظور علاوه بر متن اصلی از دیدگاه‌های کاربران نیز برای خلاصه‌سازی استفاده شده است .

در این پژوهش علاوه بر دیدگاه‌های منتشر شده در انتهای خبر از دیدگاه‌های کاربران شبکه اجتماعی توییتر که مطالب مرتبط با خبر منتشر کرده‌اند نیز استفاده می‌کنیم .

دو روش خلاصه‌سازی در این پژوهش ارائه شده است، روش اول که مبتنی بر تکرار کلمات است و از دیدگاه‌ها جهت تعیین مهم‌ترین جملات استفاده شده است، در روش دوم که روشی مبتنی بر گراف است از الگوریتم PageRank برای امتیازدهی به جملات و تعیین اهمیت هر جمله استفاده کردیم. در هر دو الگوریتم مذکور از موقعیت جمله و شباهت کسینوسی برای نتیجه بهتر استفاده می‌کنیم، همچنین به دیدگاه‌ها با توجه به اهمیت اجتماعی فرد منتشرکننده و کلمات مورد استفاده آن امتیاز داده شده است.

خروجی روش‌های ارائه شده در این پژوهش از دقت بهتری نسبت به سایر روش‌های موجود برخوردار است.