

عنوان فارسی مقاله :

یادگیری بدون نظارت و تحت نظارت برای ارزیابی وابستگی رویداد بر مبنای  
محتوی کاوی از جریان های رسانه های اجتماعی

عنوان انگلیسی مقاله :

Unsupervised and supervised learning to evaluate event relatedness based on  
content mining from social-media streams



توجه !

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل  
با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

### 6.1.2. Case study (II): baseline event "Whitney Houston dead"

In the experiment, a total number of 575,438,311 Twitter posts were collected, dating from: October 1, 2010 to September 30, 2011 and January 1, 2012 to March 14, 2012. After filtering out non-ASCII tweets, 304,758,200 tweets had been utilized as our data source. Also, we utilized the dataset collected from October 1, 2010 to September 30, 2011 corpus as our dataset for training, and used the corpus dating from January 1, 2012 to March 14, 2012 as our test data. Subsequently, we partitioned messages into unigrams and all capital letters in each tweet were converted into lowercase for our experiments.



6.1.2 مطالعه ی موردی (II): رویداد پایه "وینتی هوستون (Whitney Houston dead)"

در این آزمایش، در مجموع تعداد 575438311 پست توییت از تاریخ 1 اکتبر 2010 تا 30 سپتامبر 2011 و 1 ژانویه 2012 تا 14 مارس 2012 جمع آوری شد. پس از فیلتر کردن توئیت های غیراسکی، 304758200 توئیت به عنوان منبع داده مان مورد استفاده قرار گرفته اند. همچنین از مجموعه داده جمع آوری شده از تاریخ 1 اکتبر 2010 تا 30 سپتامبر 2011 به عنوان مجموعه داده برای آموزش استفاده کردیم و مجموعه داده از تاریخ 1 ژانویه 2012 تا 14 مارس 2012 را به عنوان داده های آزمایش به کار بردیم. سپس، پیام ها را به یونی گرام تقسیم کردیم و همه ی حروف بزرگ را به حروف کوچک تبدیل کردیم.

## توجه !

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل با فرمت

ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

همچنین برای مشاهده سایر مقالات این رشته [اینجا](#) کلیک نمایید.