

Online News Popularity: Study of Global Subjectivity impact on Engagement through Shares

Ruchit Patel
Pace University
rp90622n@pace.edu
GitHub - <https://github.com/Ruchit42>

Targeted problem



One of the objective of this project is to find out why doe certain news stories are shared more than other. In the world where more and more people are getting their news from social media and application like reddit and WhatsApp, it is important than ever to understand what makes one new article more popular than others

- The object of this project is to understand the different features on an new article and create a machine learning model which predicts with high accuracy the number of Shares an a news article will get.

Research question



Do articles with a more positive or negative sentiment (global_sentiment_polarity) receive more shares? Is there a correlation between the sentiment of the title and the article's popularity?

How does Global Subjectivity impact the number of news Shares on social media? (News that are shared the most are subjective or Objective?)

Can we build a machine learning model which can predict the number of shares an article is going to get based on the features on the articles?

Dataset



- The dataset is about online News Popularity. With almost 40 thousand instances and 60 features, this dataset will help me understand why certain news or stories get shared more than other and in the processes understand how new stories are shared.
- Mashable (www.mashable.com), which published the articles, owns the rights to the articles' content and the reproduction of it. As a result, this dataset only shares portion of the statistics related to the original content.
- #rows - 39,644
- Parameters - 61

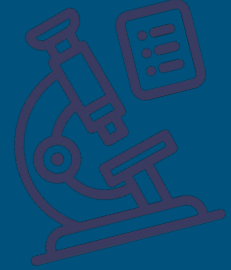
Dataset



<https://archive.ics.uci.edu/dataset/332/online+news+popularity>

	A	B	C	D	E	F	G	H	I	J	K	L
1	url	timedelta	n_tokens_title	n_tokens_content	n_unique_tokens	n_non_stop_words	n_non_stop_unique_tokens	num_hrefs	num_self_hrefs	num_imgs	num_videos	average_token_length
2	http://mashable.com/2013/01/07/amazon-instant-video-browser/	731	12	219	0.663594467	0.999999992	0.815384609	4	2	1	0	4.680365297
3	http://mashable.com/2013/01/07/ap-samsung-sponsored-tweets/	731	9	255	0.604743081	0.999999993	0.791946303	3	1	1	0	4.91372549
4	http://mashable.com/2013/01/07/apple-40-billion-app-downloads/	731	9	211	0.575129531	0.999999992	0.663865541	3	1	1	0	4.393364929
5	http://mashable.com/2013/01/07/astronaut-notre-dame-bcs/	731	9	531	0.503787878	0.999999997	0.665634673	9	0	1	0	4.404896422
6	http://mashable.com/2013/01/07/att-u-verse-apps/	731	13	1072	0.415645617	0.999999999	0.540889526	19	19	20	0	4.682835821
7	http://mashable.com/2013/01/07/beewi-smart-toys/	731	10	370	0.559888578	0.999999995	0.698198195	2	2	0	0	4.359459459
8	http://mashable.com/2013/01/07/bodymedia-armbandgets-update/	731	8	960	0.418162618	0.999999998	0.549833886	21	20	20	0	4.654166667
9	http://mashable.com/2013/01/07/canon-poweshot-n/	731	12	989	0.433573635	0.999999998	0.572107765	20	20	20	0	4.617795753
10	http://mashable.com/2013/01/07/car-of-the-future-infographic/	731	11	97	0.670103086	0.999999998	0.836734677	2	0	0	0	4.855670103
11	http://mashable.com/2013/01/07/chuck-hagel-website/	731	10	231	0.636363634	0.999999993	0.797101443	4	1	1	1	5.090909091
12	http://mashable.com/2013/01/07/cosmic-events-doomsday/	731	9	1248	0.490049751	0.999999999	0.731638417	11	0	1	0	4.617788462
13	http://mashable.com/2013/01/07/crayon-creatures/	731	10	187	0.666666663	0.999999991	0.799999993	7	0	1	0	4.657754011
14	http://mashable.com/2013/01/07/creature-cups/	731	9	274	0.6091954	0.999999994	0.707602335	18	2	11	0	4.233576642
15	http://mashable.com/2013/01/07/dad-jokes/	731	9	285	0.744186044	0.999999995	0.84153005	4	2	0	21	4.343859649
16	http://mashable.com/2013/01/07/downton-abbey-tumblrs/	731	8	259	0.562753034	0.999999994	0.644444441	19	3	9	0	5.023166023
17	http://mashable.com/2013/01/07/earth-size-planets-milky-way/	731	12	682	0.459541984	0.999999997	0.634961438	10	0	1	0	4.620234604
18	http://mashable.com/2013/01/07/echo-game/	731	8	1118	0.512396694	0.999999999	0.709770114	26	18	12	1	4.703935599
19	http://mashable.com/2013/01/07/entrepreneur-trends-2013/	731	8	397	0.624678662	0.999999996	0.805668013	11	0	1	0	5.445843829
20	http://mashable.com/2013/01/07/facebook-sick-app/	731	11	103	0.689320382	0.999999984	0.8064516	3	1	1	0	4.844660194
21	http://mashable.com/2013/01/07/felt-audio-pulse-speaker/	731	8	1207	0.410579345	0.999999999	0.548969071	24	24	42	0	4.716652858
22	http://mashable.com/2013/01/07/ford-glympse/	731	13	1248	0.390637611	0.999999999	0.523388116	21	19	20	0	4.686698718
23	http://mashable.com/2013/01/07/ftc-google-leaks/	731	9	391	0.510256409	0.999999996	0.649999997	9	2	1	1	5.296675192
24	http://mashable.com/2013/01/07/fujifilm-50x-superzoom/	731	11	1154	0.427304964	0.999999999	0.572815533	20	20	20	0	4.629982669
25	http://mashable.com/2013/01/07/hillary-clinton-helmet/	731	11	125	0.674796742	0.999999987	0.797468344	1	1	1	0	4.824
26	http://mashable.com/2013/01/07/htc-q1/	731	13	244	0.559999998	0.999999993	0.680272104	3	2	1	0	4.422131148
27	http://mashable.com/2013/01/07/huawei-ascend-mate/	731	8	266	0.572580643	0.999999993	0.72108843	5	2	1	0	4.259398496
28	http://mashable.com/2013/01/07/iheartradio-app-perfect-for/	731	8	331	0.56269113	0.999999995	0.724489792	5	3	1	0	4.782477341
29	http://mashable.com/2013/01/07/intel-awesome-laptop/	731	12	1225	0.385451505	0.999999999	0.508519003	22	22	28	0	4.635918367
30	http://mashable.com/2013/01/07/isp02-iphone-take-your-pulse/	731	10	243	0.619246859	0.999999992	0.824427475	1	1	0	0	4.382716049
31	http://mashable.com/2013/01/07/jobs-contently/	731	11	723	0.490934448	0.999999998	0.642241378	18	1	1	0	5.228215768
32	http://mashable.com/2013/01/07/land-a-job-at-spotify/	731	8	708	0.48221906	0.999999998	0.68811881	8	3	1	1	4.620056497

Motivation



I was shocked when I found out that according to an article on [pewresearch.org](https://www.pewresearch.org)

- 48% of U.S adults say they get news from social media “often”
- Out of all the social media, Facebook outpaces all other social media sites as nearly a third of Americans regularly get news on Facebook

After coming across this facts, I wanted to know what and how online news spreads and more specially, which type of stories are shared more than other.