

ANALYSIS OF TWEETS ON DEMONETIZATION IN INDIA USING SAS ENTERPRISE MINER

PAPER NUMBER: 135



JITAL PATEL and NARMADA PANNEERSELVAM

MS IN BUSINESS ANALYTICS
OKLAHOMA STATE UNIVERSITY

ABSTRACT

The citizens of a country often face the benefits or the brunt of various policies adopted by the government. Social platforms then become sounding boards for them to express their displeasures or concerns for the matter at hand. One such policy that sent Twitter into a frenzy was the demonetization order that was announced rather abruptly by the Indian Prime Minister on November 8th 2016 without any form of prior intimation to public. In this paper, we have analyzed the tweets that helped us recognize whether demonetization was perceived positively or negatively by the citizens.

About 15 days after the demonetization decision was announced, a dataset of 8,000 tweets spanning two days, was collected from a publicly available data source. Using the commonly used terms in the tweets and studying the strength of their relations using concept links, in general, show a positive feedback to the demonetization implementation policy. Using text clustering and text topic to group people with similar thoughts based on the tweets, reveal that demonetization was positively supported by a large number of people. Thus, our overall analysis showed that a vast majority of Indians accepted the demonetization policy positively while some of them expressed their displeasure over it.

INTRODUCTION

Demonetization is the act of stripping a currency unit of its status as legal tender. It occurs whenever there is a change of national currency: The current form or forms of money is pulled from circulation and retired, often to be replaced with new notes or coins. Sometimes, a country completely replaces the old currency with new currency. Many reasons could lead to this act of demonetization of currency in various countries all over the world. For example, to discourage cash-dependent economy, United States declared The Coinage Act of 1873 as a legal tender. According to the Act, silver was demonetized to fully adopt to gold standards. Another example would be The Nations of European Union, who demonetized the old national currencies, such as German mark, French franc, and Italian lira and introduced euro bills and coins in 2002. The major

reason for demonetization being trade purposes. Similarly, the Zimbabwean government demonetized its dollar in 2015, as a means to fight country's hyperinflation. (Staff, 2017)

The most recent demonetization ensued in India on 8th November, 2016. The major objective being, to battle against corruption and crime. The Indian Government's decision to demonetize all 500 and 1000 Rupee notes of the Mahatma Gandhi series as a form of legal tender. These were the two biggest denominations in its currency system accounting for about 86% of India's circulating cash. The announcement was made by the Prime Minister of India, Mr. Narendra Modi with no warning, that those notes were worthless and they had to be exchanged with the newly introduced 2000 and 500 Rupee bills by the end of the year. (Staff, 2017)

Public opinion is everything and Twitter is an extensive platform to discover the latest news and world events. Therefore, analyzing tweets on demonetization in India and deriving high-quality information from Text Mining is the scope of this paper. One of the objectives of this paper is, determining the most commonly used terms in the tweets and studying the strength of their relationship with the other terms. Another goal of this paper is, using Text Clustering to group people with similar thoughts based on the tweets and extracting Text Topics.

LITERATURE REVIEW

Various research studies and peer-viewed articles show different effects of demonetization.

"Sentiment Analysis of demonetization of 500 & 1000 rupee banknotes by Indian Government" by Prabhsimran Singh, Ravinder Singh Sawhney, Karanjeet Singh Kahlon. In this paper, they have analyzed the effect of implementing demonetization policy using sentiment analysis. Analysis shows that large share of Indian people was happy with this policy. During initial days, the sentiment was more negative due to hardships but after the release of new bank notes, overall sentiment of the people became positive.

Mohasin A. Tamboli in his paper, “Demonetization: Retrospective Analysis,” concludes that all the gains of Demonetization are contingent on government’s ability to control corruption. If corruption is not controlled, all the short term gains of demonetization will be lost and the only thing that would change for the better is look and feel of new currency.

“Demonetization and Complete Financial Inclusion” by S. Vijay Kumar and T. Shiva Kumar, also determines that rewards of demonetization are encouraging and in the long term interest of the country but governments need to ensure smooth flow of currency. It will take India ten steps ahead and will influence corruption, elections and terrorism.

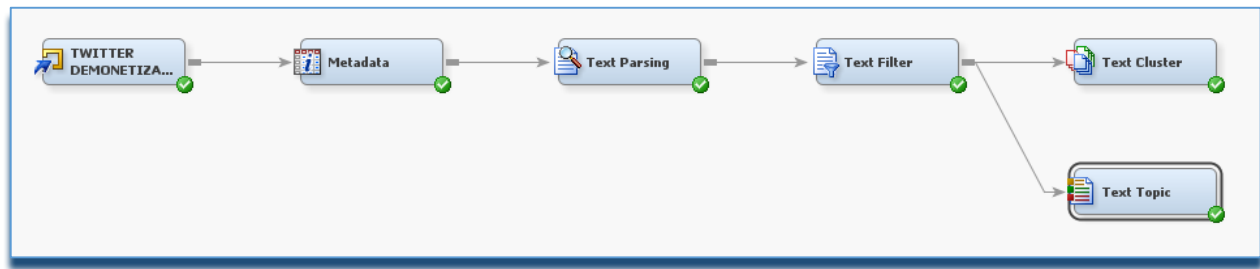
Another paper by Dr. Pawan Kumar about “Demonetization and its impact on Employment in India,” describes that the employment scenario is not conducive enough to face any challenge of demonetization of currency. In fact, the decision of demonetization will further destabilize the already volatile labor market in India.

DATA ACCESS

The dataset for this paper contains 8,000 observations, that is the number of tweets collected over a span of two days and about 15 days after the demonetization decision was announced. The dataset is taken from Kaggle (kaggle.com. 2017). The links to the main source of the dataset is also provided in the dataset. A data dictionary for the variables used in the analysis are stated below.

Variable	Role	Description
ID	ID	Observation number
Text	Text	Demonetization Tweets

METHODOLOGY



1. Text Import

The dataset is available in the excel file format. Using Enterprise Miner, the Excel file is imported using the File Import Node. The file import node converts the excel file format to SAS file format. SAS format file can be used for further analysis. Metadata node is used to restrict only the ID and Tweets as inputs for the advance study.

2. Text Parsing

After importing and selecting variables, Text Parsing node is used to identify terms and their instances in the data, which contains text. The results of Text Parsing node display the most frequently occurring term, the number of documents it has occurred in, and the terms that are rarely used. Usually, the terms that are used moderately, are the ones that are the most helpful in exploration and modeling.

Term	Role	Attribute	Freq	# Docs	Keep	Parent/Child Status	Parent ID	Rank for Variable numdocs
demonetization	Miscellaneous P...	Entity	3854	3849Y			21	1
https		Alpha	3847	3556Y			46	2
t		Alpha	3701	3432N			47	3
demonetization		Alpha	3680	3375Y			45	4
rt	Miscellaneous P...	Entity	2938	2930Y			4	5
s		Alpha	2196	2119N			12	6
amp		Alpha	1686	1656Y			18	7
+ be		Alpha	1622	1411N	+		721	8
+ have		Alpha	1126	1119N	+		73	9
+ support		Alpha	1086	1076Y	+		91	10
narendramodi		Alpha	928	928Y			88	11
+ people		Alpha	866	833Y	+		118	12
+ bank		Alpha	767	764Y	+		67	13
u		Alpha	10157	728N			127	14
+ not		Alpha	846	718N	+		31	15
on		Alpha	641	640N			23	16
such		Alpha	652	624N			70	17
+ lakh		Alpha	604	602Y	+		65	18
ed		Alpha	3785	583Y			126	19
a0	Miscellaneous P...	Entity	1894	582Y			122	20
k		Alpha	577	577N			68	21
bd	Miscellaneous P...	Entity	1835	576Y			123	22
+ terrorist		Alpha	571	569Y	+		8820	23
j	Miscellaneous P...	Entity	556	556N			63	24
+ loot		Alpha	545	545Y	+		9828	25
kishtwar	Miscellaneous P...	Entity	544	544Y			62	26
incident		Alpha	542	542Y			71	27
third		Alpha	542	542Y			69	27
third such incide...	Noun Group	Alpha	542	542Y			61	27
gauravcsawant		Alpha	541	541Y			64	30
rt @modibharos...	Miscellaneous P...	Entity	539	539Y			386	31
+ go		Alpha	505	503N	+		623	32
+ do		Alpha	554	460N	+		303	33
b8	Miscellaneous P...	Entity	1618	449Y			124	34
+ question		Alpha	442	442Y	+		8	35

The most frequently used term is demonetization which has a total frequency of 7,534 (3,854+3,680) since it has appeared twice. It has appeared in total 7,224 (3,849+3,375) documents. Similarly, the next most frequently used terms are app (1,686), support(1,086), narendramodi(928), people(866), bank(767).

3. Text Filtering

A Text Filter node is added to the Text Parsing node to eliminate the terms that occur least number of times in all the documents and to perform spell check to suggest potential synonyms.

Term	Role	Attribute	Status	Weight	Imported Frequency	Freq	Number of Imported Documents	# Docs	Rank	Parent/Child Status	Parent ID
+ demonetizatio...	Miscellaneous P...	Entity	Keep	0.081	3854	3860	3849	3854	1+		21
https	...	Alpha	Drop	0.000	3847	3847	3556	3556	2		46
t	...	Alpha	Drop	0.000	3701	3701	3432	3432	3		47
demonetization	...	Alpha	Keep	0.099	3680	3680	3375	3375	4		45
rt	Miscellaneous P...	Entity	Keep	0.112	2938	2938	2930	2930	5		4
s	...	Alpha	Drop	0.000	2196	2196	2119	2119	6		12
amp	...	Alpha	Drop	0.000	1686	1686	1656	1656	7		18
+ be	...	Alpha	Drop	0.000	1622	1629	1411	1418	8+		721
+ have	...	Alpha	Drop	0.000	1126	1134	1119	1127	9+		73
+ support	...	Alpha	Keep	0.223	1086	1091	1076	1080	10+		91
+ narendramodi	...	Alpha	Keep	0.239	928	929	928	929	11+		88
+ people	...	Alpha	Keep	0.253	866	872	833	839	12+		118
+ bank	...	Alpha	Keep	0.261	767	767	764	764	13+		67
+ u	...	Alpha	Drop	0.000	10157	10161	728	730	14+		127
+ not	...	Alpha	Drop	0.000	846	846	718	718	15+		31
on	...	Alpha	Drop	0.000	641	641	640	640	16		23
+ such	...	Alpha	Drop	0.000	652	655	624	627	17+		70
+ lakh	...	Alpha	Keep	0.288	604	605	602	603	18+		65
ed	...	Alpha	Drop	0.000	3785	3785	583	583	19		126
a0	Miscellaneous P...	Entity	Drop	0.000	1894	1894	582	582	20		122
+ k	...	Alpha	Drop	0.000	577	579	577	579	21+		68
bd	Miscellaneous P...	Entity	Drop	0.000	1835	1835	576	576	22		123
+ terrorist	...	Alpha	Keep	0.294	571	572	569	570	23+		8820
+ j	Miscellaneous P...	Entity	Drop	0.000	556	557	556	557	24+		63
+ loot	...	Alpha	Keep	0.299	545	545	545	545	25+		9828
kishtwar	Miscellaneous P...	Entity	Drop	0.000	544	544	544	544	26		62
third such incide	Noun Group	Alpha	Keep	0.299	542	542	542	542	27		61
third	...	Alpha	Keep	0.299	542	542	542	542	27		69
incident	...	Alpha	Keep	0.299	542	542	542	542	27		71
+ rt @modibharo	Miscellaneous P...	Entity	Keep	0.299	539	542	539	542	27+		386
gauravcsawant	...	Alpha	Drop	0.000	541	541	541	541	31		64
+ go	...	Alpha	Drop	0.000	505	505	503	503	32+		623
+ do	...	Alpha	Drop	0.000	554	558	460	463	33+		303
b8	Miscellaneous P...	Entity	Drop	0.000	1618	1618	449	449	34		124

The results of the Text Filtering node show that terms like https, t, s, and have, have been dropped in the Terms table, as they do not add any meaning to the analysis. Only the Terms that add meaning to the analysis of demonetization are retained.

22	568.0	tertorists	1.0	terrorists	PROP_MISC		10.0		2552.0	72.0
23	759.0	hav	2.0	have			10.0		2432.0	73.0
24	103.0	nationalism	2.0	nationalists			12.0		6383.0	77.0
25	102.0	mask	1.0	mas			10.0		2436.0	82.0
26	123.0	opposition	1.0	opposition	PROP_MISC		0.0		2759.0	86.0
27	123.0	oppositio	1.0	opposition			3.0		92.0	86.0
28	928.0	narednramodi	1.0	narendramodi	PROP_MISC		4.0		8517.0	88.0
29	7.0	respectd	1.0	respect			4.0		2677.0	89.0
30	636.0	supprt	2.0	support			8.0		439.0	91.0
31	636.0	sport	1.0	support			15.0		2053.0	91.0
32	33.0	givin	1.0	giving			6.0		4476.0	103.0
33	20.0	tough	1.0	though			10.0		3234.0	105.0
34	150.0	wher	1.0	where			8.0		8554.0	115.0
35	141.0	takng	1.0	taking			10.0		7727.0	116.0
36	826.0	people	5.0	people	PROP_MISC		0.0		3710.0	118.0
37	826.0	peopl	1.0	people			6.0		6050.0	118.0

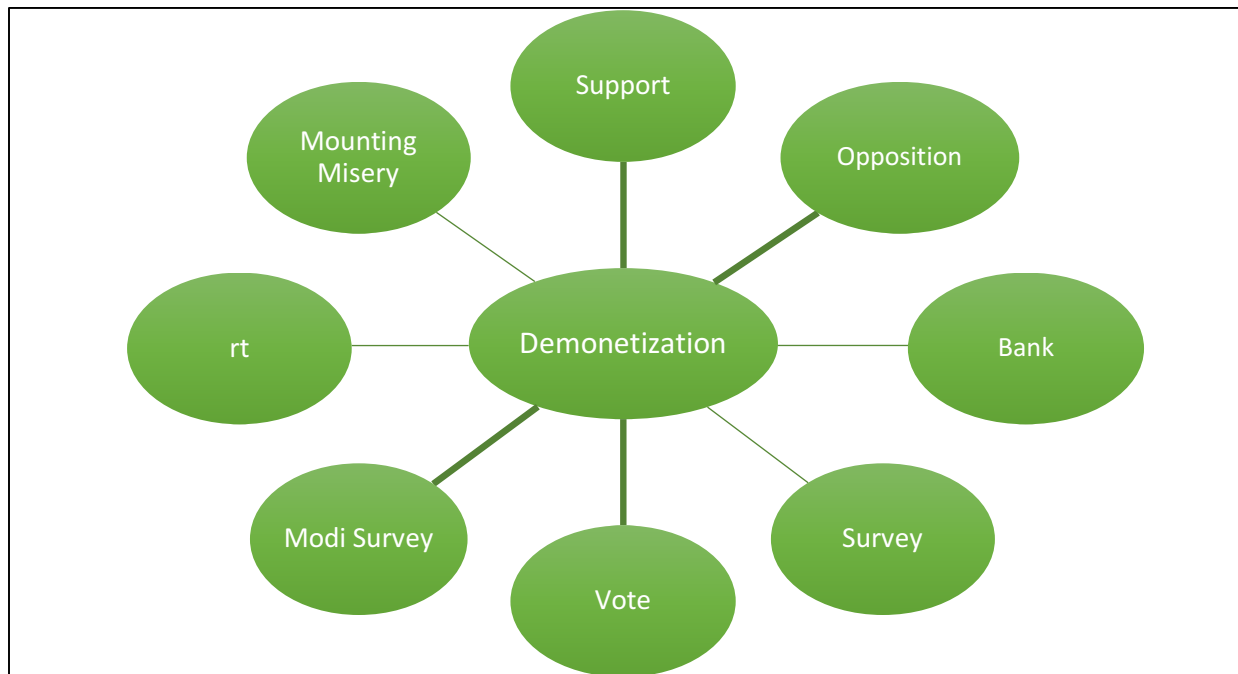
The Text Filter node, autocorrects the misspelt terms like 'hav' is corrected to 'have', 'tertorists' is corrected to 'terrorists' and 'oppositio' is corrected to 'opposition'.

	demonetization	3860	3854	☒	0.081	Miscellaneous P...	Entity
	demonization	3	3	☐			Alpha
	demoditization	1	1	☐			Alpha
	demonitization	1	1	☐		Miscellaneous P...	Entity
	demonetization	3854	3849	☐		Miscellaneous P...	Entity
	dmonetization	1	1	☐		Miscellaneous P...	Entity

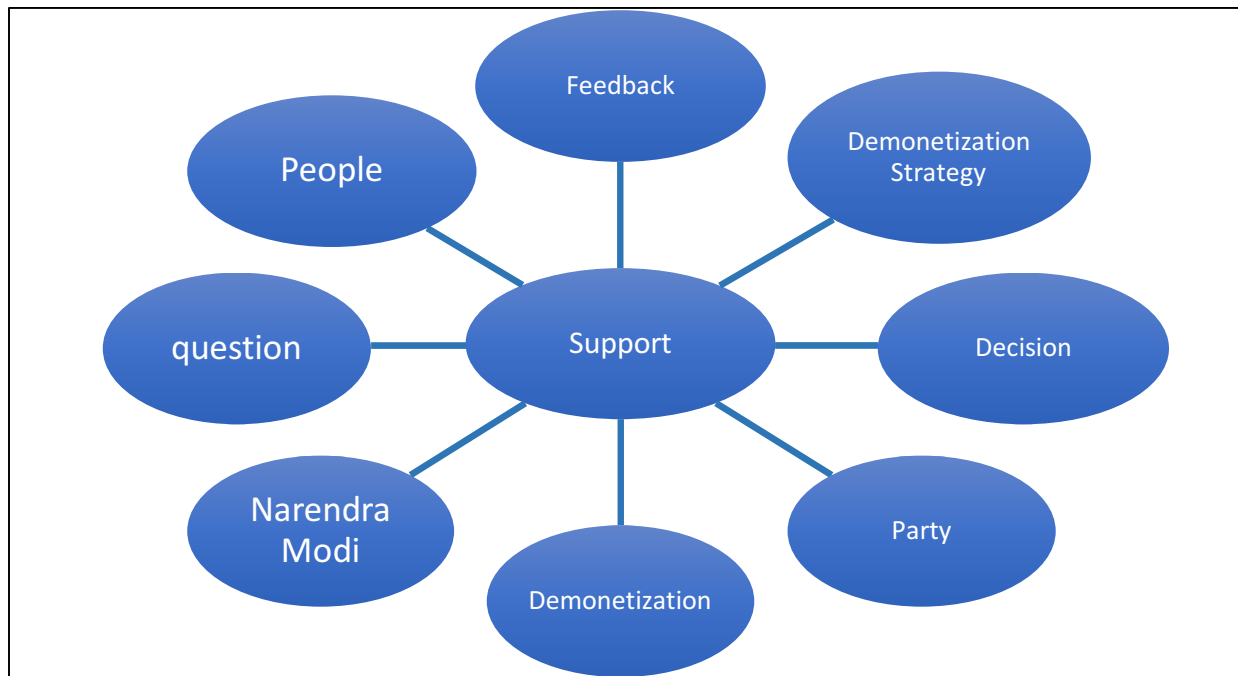
The node also, groups the synonyms together based on the synonyms file that is imported or the terms are manually dragged and dropped into each other. For example, terms like 'demonization', 'demoditization', 'demonitization', and 'dmonetization' are all grouped into 'demonetization' by the Text Filter node.

Concept Links

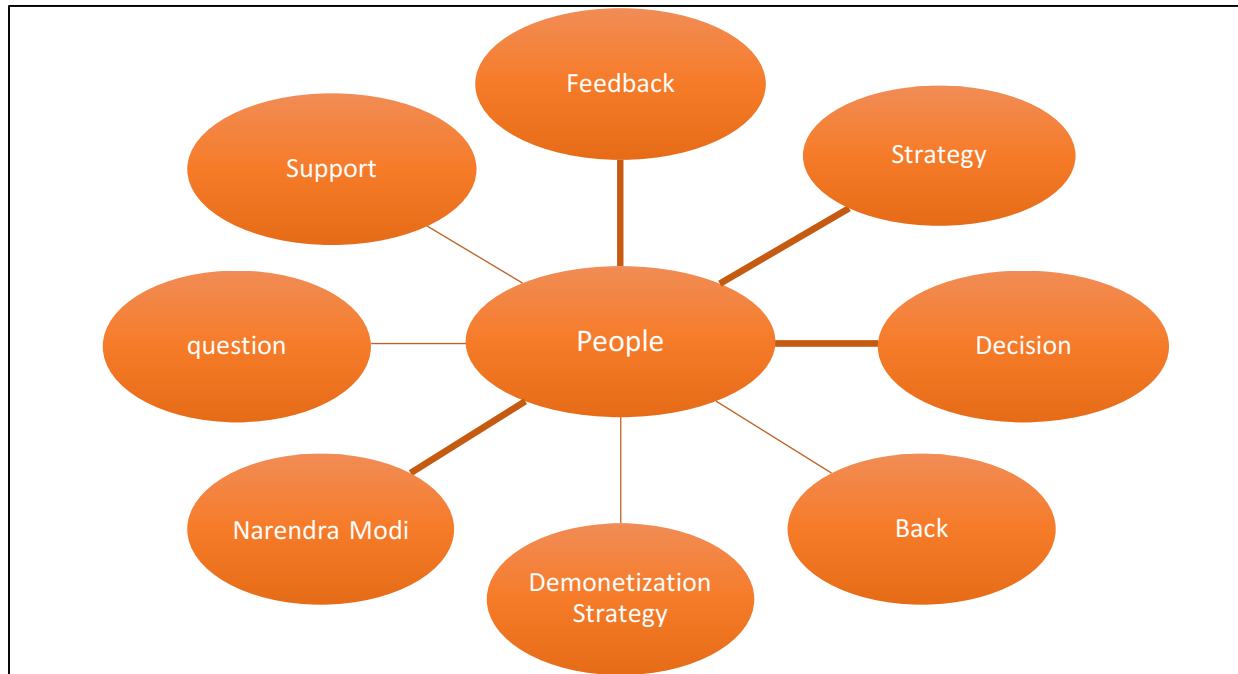
Concept Link, displays the terms that are highly associated with the selected term in the Terms table. The selected term is surrounded by the terms that correlate the strongest with the selected term. The width of the link depicts the strength of association between the two terms.



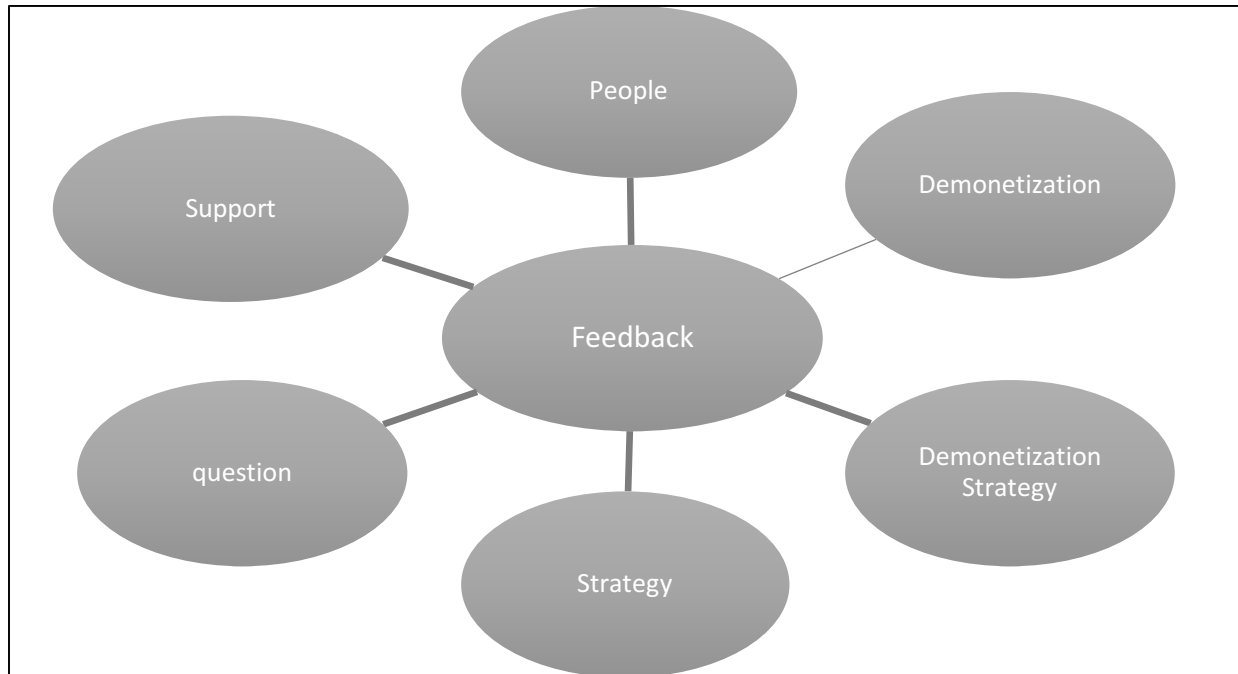
The above concept link is for the highest frequency term, Demonetization. The term demonetization is associated with bank (associated with demonetization of currency), survey (associated with demonetization survey), vote (associated with voting for demonetization survey app), modi survey (associated with demonetization survey app initiated by Modi), rt (associated with retweets about demonetization), mounting misery (associated with misery due to monetization) support (associated with the support of people for demonetization) opposition (associated with opposition party's view with respect to monetization). Out of all the terms that Demonetization is associated with, it is strongly related with support, opposition, modi survey and vote. The associated terms depict positive (support, rt) as well as negative (opposition, misery) responses of demonetization.



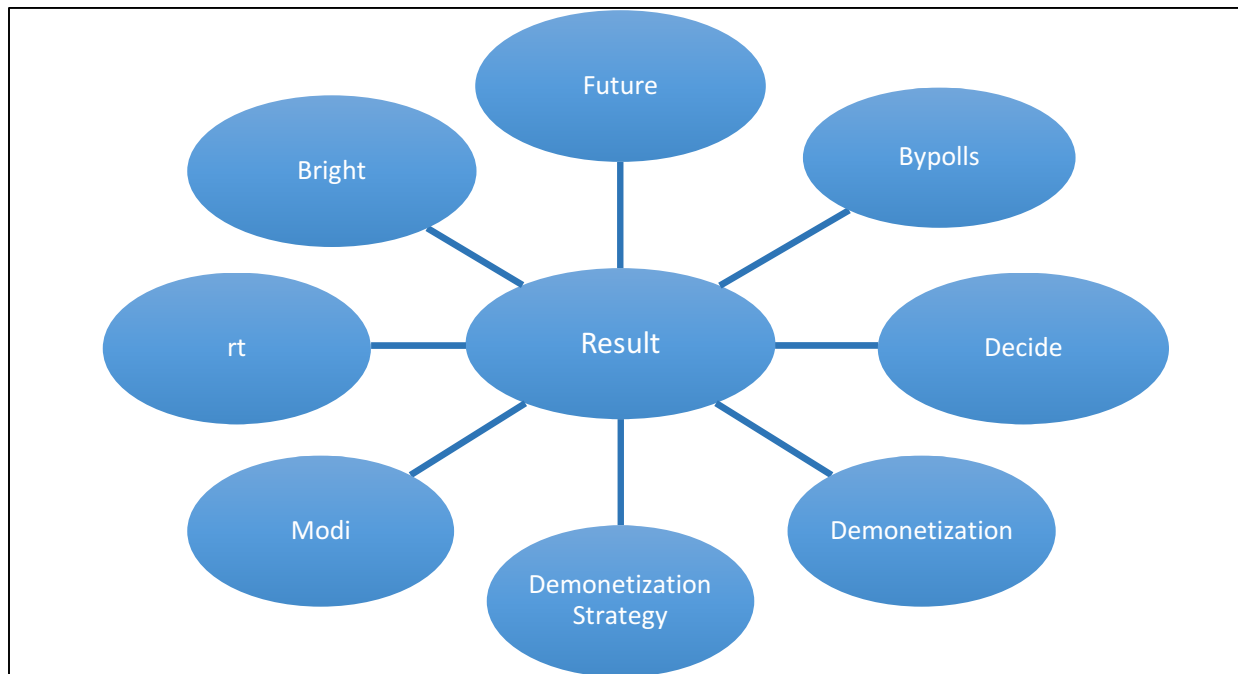
The above concept link is for the term Support, which is strongly associated with Demonetization. The important terms associated with the term support are demonetization strategy, party, decision, narendra modi, demonetization, question, people, and feedback. The term support is equally strongly associated with all the above terms. This concept link depicts that the feedback of majority people and party, is supporting the demonetization strategy or decision taken by Narendra Modi.



The above concept link is for the term People and the important terms associated with it are demonetization strategy, support, decision, narendra modi, question, back, strategy, and feedback. The term people is strongly associated only with strategy, feedback, narendramodi and decision terms. This concept link also, depicts that the feedback of majority people is to support the demonetization strategy or decision taken by Narendra Modi.



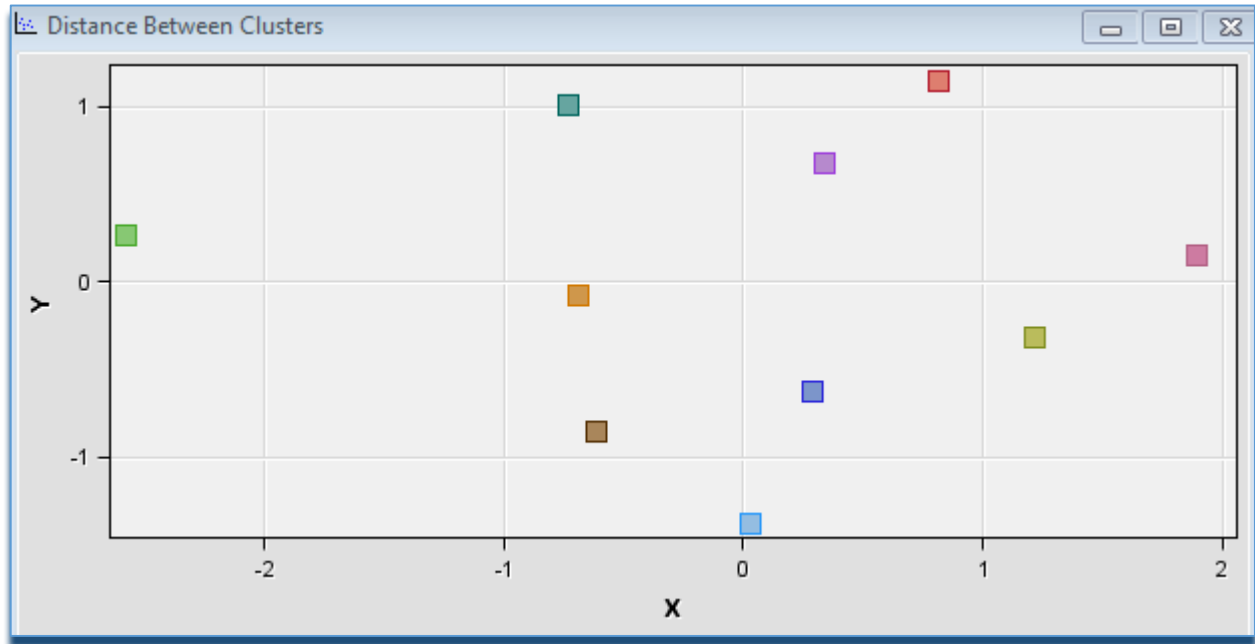
The above concept link is for the term Feedback and the important terms associated with it are strategy, people, demonetization strategy, support, question, and demonetization. The term feedback is strongly associated only with strategy, people, demonetization strategy, support, and question terms. This concept link also, depicts that the feedback of majority people is to support the demonetization strategy.



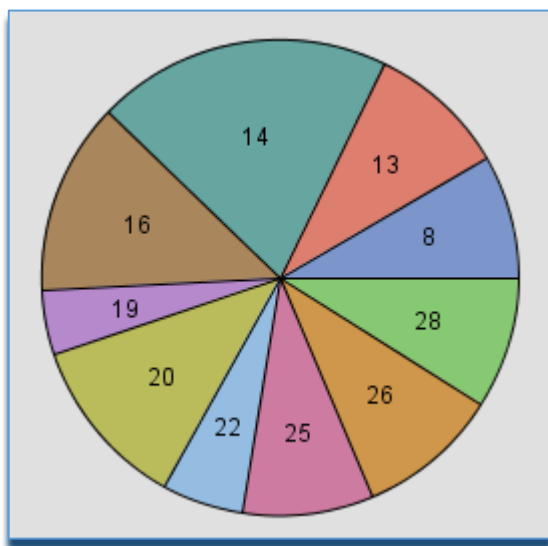
The above concept link is for the term Result and the important terms associated with it are demonetization, decide, bypolls, modi, demonetization strategy, rt, bright, and future. The term result is equally strongly associated only with all the above terms. This concept link also, depicts that the result of the survey is that modi's demonetization decision will lead to a bright future is retweeted many times.

4. Text Clustering

A Text Cluster node is added to the text filtering node. The text clustering node clusters the documents into disjointed sets of documents and reports on descriptive terms of those clusters. Hierarchical clustering algorithm is used to group clusters into a tree hierarchy. The approach of hierarchical clustering relies on singular value decomposition (SVD) to transform original weighted, term-document frequency matrix into a dense but low dimensional representation.



The above plot, shows the distance between the clusters. All the clusters are distributed well apart. The pie chart below, shows the distribution of the cluster frequencies. Apart from the cluster number 14 and cluster number 16 the frequencies are well distributed among all the clusters.



Cluster ID	Descriptive Terms	Frequency	Percentage	Example Tweets
8	Demonetization move + black money + black + government + govt + modi + money + move	658	8%	Demonetization is government's move against black money
13	Rt@shashitharoor + big + cash + country + demonetization + duty + time + understand	774	10%	Shashi Tharoor's tweets retweeted, that PM's duty to understand and address parliament about demonetization
14	Third such incident + bank + demonetization + lakh + paytm + question clearly critical	1584	20%	50 lakh customers responded in 24hours and 90% support demonetization. Was paytm informed about demonetization? 40 lakhs looted from a bank, third incident since demonetization.
16	Opposition mps + pmoindia address parliame + rt@drkumarvishwas +	1053	13%	Opposition MPs supporting demonetization. Demanding

	rt@shashitharoor + address + govt + line affected			PM of India to address parliament.
19	Abov + atm + byelection + people + state + support atms ippatel	347	4%	Delay replenishing ATMs. By election two days after demonetization has proved support.
20	Debate + decide + demonetization + Gandhi + impact + know + modi + note	950	12%	Easy exchange to new 2k note. Rahul Gandhi (opposition) issue with Demonetization.
22	Daughter's wedding + fund shortage + daughter +fund + life + prob +tweet ends	442	6%	Man ends life due to fund shortage for daughter's wedding.
25	Watch briefing + demonetization + effect +happiness + keep + pmoindia +stuff announcements	699	9%	Happiness about demonetization but announcement not made at parliament, so watch briefing
26	Modi survey + arunjaitley + blackmoney + decision + flight + move + response +survey	776	10%	Modi's demonetization decision survey showed positive response

28	Demonetization strategy huge support demonetization move + rt@modibharosa + back + c-voter + narendramodi + nation	717	9%	Demonetization strategy was supported on a large scale
----	---	-----	----	--

The above text clusters shows positive, negative as well as neutral tweet text review analysis. Out of the ten clusters, six clusters show positive reviews.

5. Text Topic

Text Topic node is connected to the Text Parsing node to combine the terms into topics so that they could be analyzed further. A list of topics is created to establish combinations of words, which could be of interest in analyzing.

Topic	Number of Terms	Number of Documents	Example Tweets
Third such incident, third , incident, loot, terrorist	10	542	40 lakhs looted from a bank, third incident since demonetization.
Critical question, demonetization edict, edict, inform, rsurjewala	15	289	Was paytm informed about demonetization?
Party politics, nitishkumar, putting	16	552	Putting nation over party politics, support demonetization by PM

nation, rt@modibharosa, narendramodi			
Huge support, huge, c- voter, demonetization move, nation	23	264	Huge support by people and nation for demonetization
Daughter's wedding, fund, shortage, daughter, end, life	20	159	Man ends life due to fund shortage for daughter's wedding
Ls, retain, byelection, seat, state	44	309	By election two days after demonetization has proved support.
rt@drkumarvishwas, demonetization, people, rt, huge, support	13	350	Huge support by people for demonetization
Join, corruptfreeindia, walk, nationalist, ma	31	189	Join the walk towards corruption free India

The above text topics show positive, negative as well as neutral tweet text review analysis. Out of the eight topics, five topics show positive reviews.

CONCLUSION

Radical changes are always faced with resistance of some kind, case in point being the policy to implement demonetization. The goal of this paper was to analyze the tweets by the Indian citizens about demonetization and their perception towards the policy, by using Text Mining. The results of Concept Links, in general, show a positive feedback to the demonetization implementation policy. Similarly, the results of our analysis of tweets using Text Clustering and Text Topics reveal that demonetization was positively supported by a large number of people.

Our overall analysis showed that a vast majority of Indians accepted the demonetization policy positively while some of them expressed their displeasure over it.

BIBLIOGRAPHY

[1] SAS Institute Inc. 2012. Getting Started with SAS® Text Miner 12.1. Cary, NC: SAS Institute Inc.

<https://support.sas.com/documentation/onlinedoc/txtminer/12.1/tmgs.pdf>

[2] Staff, I. (2017). Demonetization. [online] Investopedia. Available at:

<http://www.investopedia.com/terms/d/demonetization.asp?ad=dirN&qo=investopediaSiteSearch&qsrc=0&o=40186> [Accessed 5 Apr. 2017].

[3] “Sentiment Analysis of demonetization of 500 & 1000 rupee banknotes by Indian Government” by Prabhsimran Singh, Ravinder Singh Sawhney, Karanjeet Singh Kahlon

[4] “Demonetization: Retrospective Analysis,” by Mohasin A. Tamboli

[5] “Demonetization and Complete Financial Inclusion” by S. Vijay Kumar and T. Shiva Kumar

[6] “Demonetization and its impact on Employment in India,” by Dr. Pawan Kumar

[7] Kaggle.com. (2017). Demonetization in India Twitter Data | Kaggle. [online] Available at: <https://www.kaggle.com/arath33/demonetization-in-india-twitter-data> [Accessed 10 Feb. 2017].