



ABOITIZ  
SCHOOL OF INNOVATION,  
TECHNOLOGY, & ENTREPRENEURSHIP

# HMM...INTENSE!

Emotion Distribution Recognition of News Articles  
using Rappler's Crowd-Annotated Mood Metrics



From Rappler.com

---

---

BAYHON, Tonee L.

MADRIGAL, Rafael G.

MAULION, Matthew R.

PEREZ, Adrian M.

TABONG, Kristoffer Dave A.



From Google .com

INTRODUCTION

PROBLEM

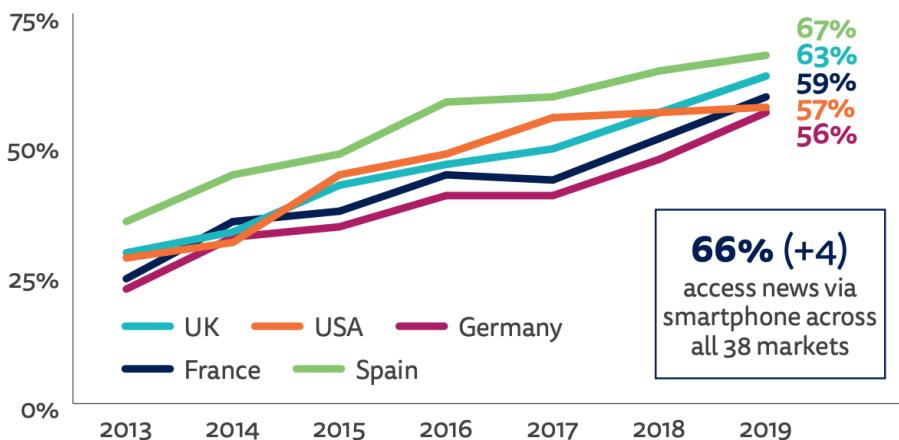
METHODOLOGY

RESULTS

CONCLUSION

# SIGNIFICANT SHIFT TOWARDS THE USE OF SMARTPHONES FOR NEWS IN THE LAST DECADE<sup>1</sup>

PROPORTION THAT USED A SMARTPHONE FOR NEWS IN THE LAST WEEK (2013-19) – SELECTED MARKETS



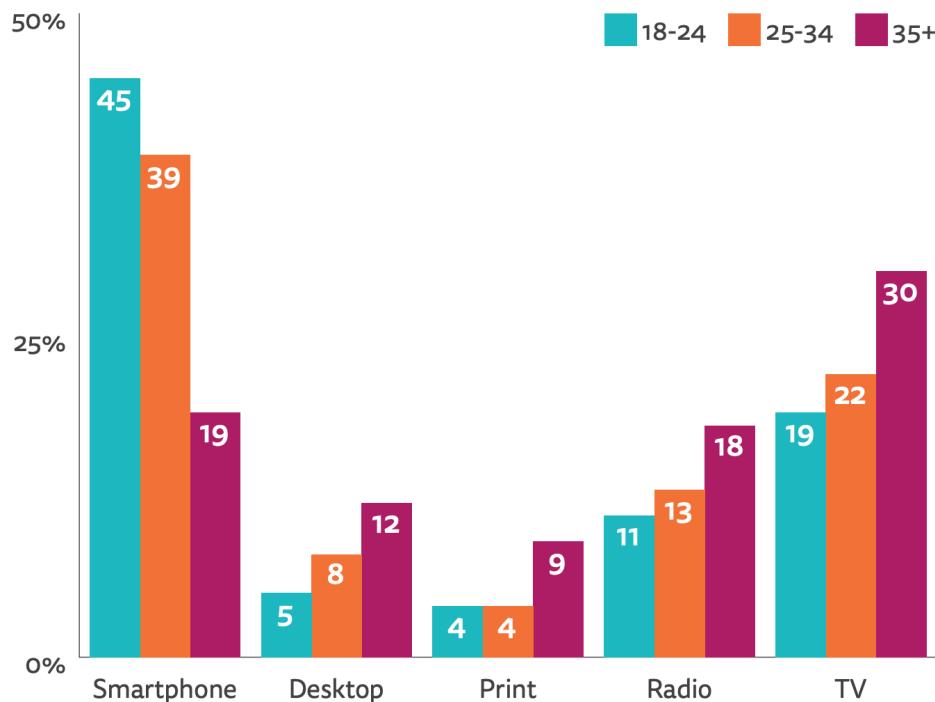
**Q8B.** Which, if any, of the following devices have you used to access news in the last week?

Base: Total sample 2013-19 sample in each country ≈ 2000.

<sup>1</sup> Newman, N., Fletcher, R. Kalogeropoulos, A., Nielsen, R. (2019) Reuters Institute Digital News Report 2019. Reuters Institute for the Study of Journalism. Accessed 31 August 2020. Retrieved From [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR\\_2019\\_FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR_2019_FINAL.pdf)

# YOUNGER GENERATIONS MOSTLY USE SMARTPHONES TO CHECK THE NEWS<sup>1</sup>

PROPORTION THAT SAID EACH WAS THEIR FIRST CONTACT WITH NEWS IN THE MORNING BY AGE – SELECTED MARKETS



**Q9c\_new2016.** What is the FIRST way you typically come across news in the morning? Base: 18-24/25-34/35+: selected countries = 1863/3019/15247. Note: Data from US, UK, France, Italy, Spain, Ireland, Norway, Finland, Netherlands, Japan.

<sup>1</sup> Newman, N., Fletcher, R. Kalogeropoulos, A., Nielsen, R. (2019) Reuters Institute Digital News Report 2019. Reuters Institute for the Study of Journalism. Accessed 31 August 2020. Retrieved From [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR\\_2019\\_FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR_2019_FINAL.pdf)

# NEWS ARTICLES DO NOT ONLY DESCRIBE REAL EVENTS BUT ALSO EVOKE COMPLEX FEELINGS

## Macron to mark Lebanon's centenary as nation teeters on brink

Raya Jalabi

BEIRUT (Reuters) - French President Emmanuel Macron marked Lebanon's centenary on Tuesday by planting a cedar tree, a gesture that is collapsing under the weight of a crippling

## IMF predicts worst Middle East downturn in half century

The International Monetary Fund sees the Middle East and North Africa regional economy shrinking by 5.7% in 2020



Agence France-Presse  
@afp  
Published 8:30 PM, July 13, 2020  
Updated 8:50 PM, July 13, 2020



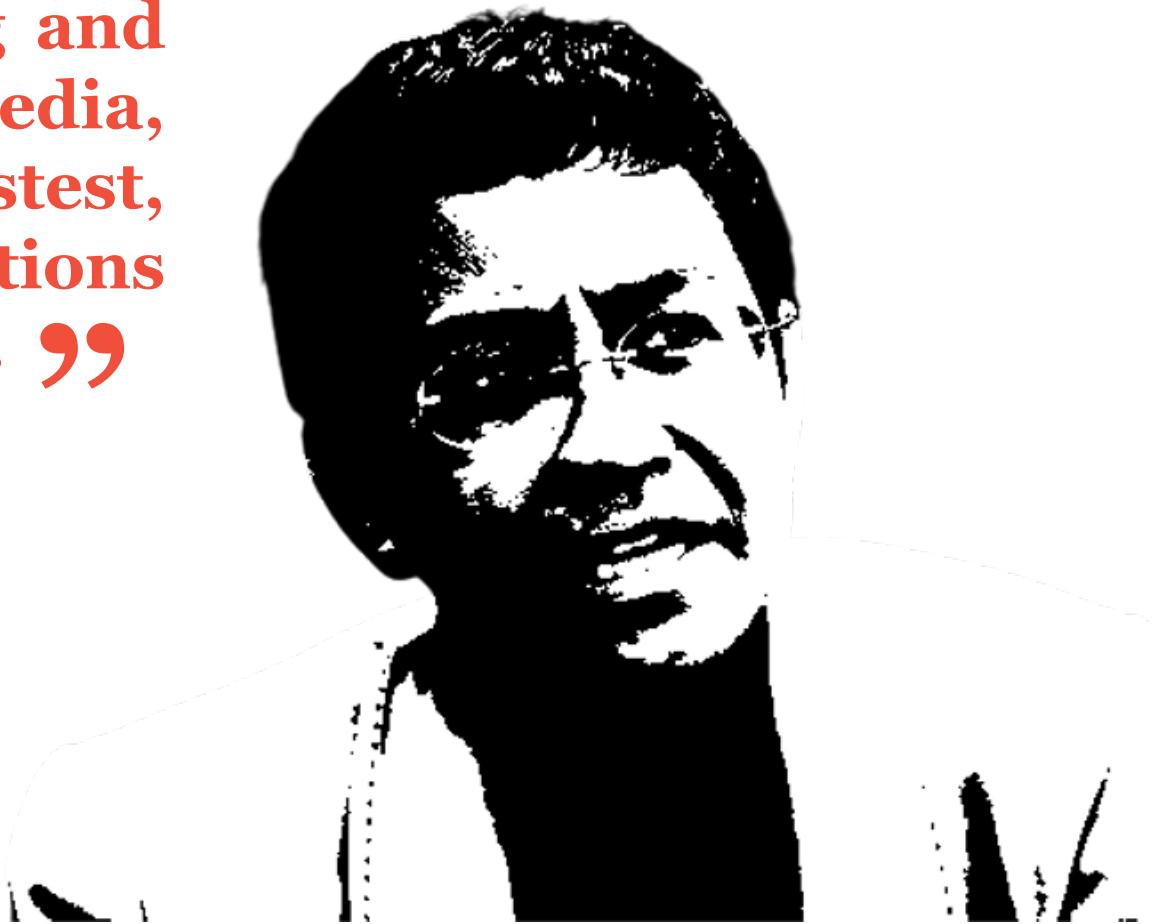
Saudi Arabia. Cars drive down King Fahad Road after authorities eased some of the lockdown measures that had been imposed in the Saudi capital Riyadh on June 21, 2020. Photo by Fayez Nureldine/ARD

<sup>1</sup> Agence France-Presse (2020) IMF predicts worse Middle East downturn in half century. Retrieved from <https://r3.rappler.com/business/266455-imf-regional-economic-outlook-middle-east-north-africa-july-2020>

<sup>2</sup> Jalabi, R., Nader, Y. (2020) Macron tells Lebanese leaders to reform swiftly or face sanctions. Retrieved from <https://in.reuters.com/article/lebanon-crisis/macron-to-mark-lebanons-centenary-as-nation-teeters-on-brink-idINKBN25S45L>

**“ [...] emotions are important in decision making and on social media, what spreads fastest, it’s actually emotions more than ideas. ”**

*– Maria Ressa*



<sup>1</sup> Lafrance, A. (2012) In the Philippines, Rappler is trying to figure out the role of emotion in news. Nieman Lab. Retrieved From: <https://www.niemanlab.org/2012/08/in-the-philippines-rappler-is-trying-to-figure-out-the-role-of-emotion-in-the-news/>

# PREVIOUS STUDIES ARE LIMITED TO EMOTION CLASSIFICATION AND SENTIMENT POLARITY IDENTIFICATION

## Sentiment topic models for social emotion mining

Yanghui Rao<sup>a,b</sup>, Qing Li<sup>a</sup>, Xudong Mao<sup>a</sup>, Liu Wenyin<sup>b</sup>

<sup>a</sup>Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong

<sup>b</sup>College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, China



### ARTICLE INFO

Article history:  
Received 12 June 2013  
Received in revised form 7 November 2013  
Accepted 30 December 2013  
Available online 11 January 2014

Keywords:  
Social emotion mining  
Sentiment topic model  
Social emotion classification  
Social emotion lexicon

### ABSTRACT

The rapid development of social media services has brought many opportunities through online news, blogs, microblogs/tweets, etc. This article concentrates on the mining of readers' emotions from social media. Compared to the classical sentiment analysis from the text, mining of readers' emotions from social media is sometimes more meaningful. In social media, readers' emotions are often evoked by some topics or events. Compared to the classical sentiment analysis from the text, mining of readers' emotions from social media is sometimes more meaningful. In social media, readers' emotions are often evoked by some topics or events. Compared to the classical sentiment analysis from the text, mining of readers' emotions from social media is sometimes more meaningful. In social media, readers' emotions are often evoked by some topics or events.

## Deep Feelings: A Massive Cross-Lingual Study on the Relation between Emotions and Virality

Marco Guerini  
Trento Rise  
Via Sommarive 18  
Trento, Italy  
marco.guerini@trentorise.eu

### ABSTRACT

This article provides a comprehensive investigation on the relations between virality of news articles and the emotions they are found to evoke. Virality, in our view, is a phenomenon with many facets, i.e., under this generic term several different effects of persuasive communication are comprised. By exploiting a high-coverage and bilingual corpus of documents containing metrics of their spread on social networks as well as a massive affective annotation provided by readers, we present a thorough analysis of the interplay between evoked emotions and viral facets.

We highlight and discuss our findings in light of a cross-lingual approach: while we discover differences in evoked emotions and corresponding viral effects, we provide preliminary evidence of a generalizable explanatory model rooted in the deep structure of emotions: the Valence-Arousal-Dominance (VAD) circumplex. We find that viral facets appear to be consistently affected by particular VAD configurations, and these configurations indicate a clear connection with distinct phenomena underlying persuasive communication.

## Multi-label Text Classification Approach for Sentence Level News Emotion Analysis

n Basu, Pabitra Mitra, and Abhishek Prasad

Technology, Kharagpur, India-721302

Jacopo Staiano  
LIP6, UPMC - Sorbonne Universités  
and FBK MobS Lab  
4 place Jussieu, Paris, France  
jacopo.staiano@lip6.fr

people to explicitly tag their emotional states as they browse the web.

Although the adoption of the latter is not yet prominent, we found two very popular news outlets, one in English and one in Italian, embedding similar interfaces for affective feedback in each article page. We thus became interested in investigating the relation of this newly available affective data with virality of news articles in a cross-lingual experimental setting. The two online websites which made this research possible are:

1. Rappi ([rappi.it](http://rappi.it)), an English-written "social news" portal, which makes extensive use of its distinctive Mood Meter feature;
2. the online version of the Italian newspaper Corriere della Sera ([corriere.it](http://corriere.it)), one of the most popular daily newspapers in Italy.

In our previous work [30], we have shown the impressive value of such reader-provided affective data: we used Rappi data to automatically build an emotion lexicon and a system for automatic af-

Articles were sourced from Google Scholar. These papers are referenced in the manuscript



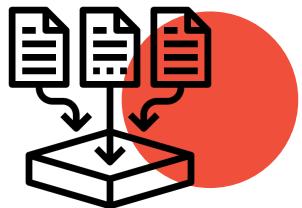
## What emotions are evoked by articles and how intense are these relative to each other?



Icons for each emotions were obtained from Rappler.com

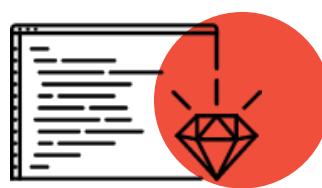
# DATA SCIENCE METHODOLOGY

## Collection and Exploratory Analysis



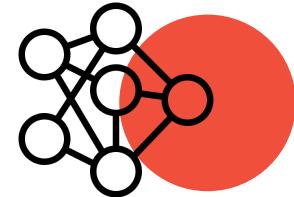
17000+ articles  
August 2019 – July 2020

## Cleaning and Preprocessing



Stop words & punctuations removal,  
lowercasing, word2vec Implementation,  
Data Balancing using SMOTE

## Machine Learning



Training and Hyperparameter Tuning  
of Machine Learning Algorithms

Icons from Flaticon.com

# Tisoy weakens into tropical depression as it leaves PAR

Though Tisoy (Kammuri) is no longer inside the Philippine Area of Responsibility, there may still be heavy rain due to the surge of the northeast monsoon and the tail-end of a cold front

**Published:** 2019-12-05

**Written by:** Acor Arceo

**Link:** <https://rappler.com/nation/weather/tropical-depression-tisoy-pagasa-forecast-december-5-2019-11am>

Emotion Distribution



## TEXT PREPROCESSING AND VECTORIZATION

# Tisoy weakens into tropical depression as it leaves PAR

Though Tisoy (Kammuri) is no longer inside the Philippine Area of Responsibility, there may still be heavy rain due to the surge of the northeast monsoon and the tail-end of a cold front

Published: 2019-12-05

Written by: Acor Arceo

Link: <https://rappler.com/nation/weather/tropical-depression-tisoy-pagasa-forecast-december-5-2019-11am>

## REMOVAL OF STOPWORDS AND NON-RELEVANT METADATA

# Tisoy weakens tropical depression leaves PAR

Though Tisoy (Kammuri) longer inside  
Philippine Area Responsibility,  
heavy rain due surge  
northeast monsoon tail-end cold  
front

## REMOVAL OF PUNCTUATION MARKS

# Tisoy weakens tropical depression leaves PAR

Though Tisoy Kammuri longer inside  
Philippine Area Responsibility  
heavy rain due surge  
northeast monsoon tail end cold  
front

## LOWERCASING AND LEMMATIZATION WITH POS TAGGING

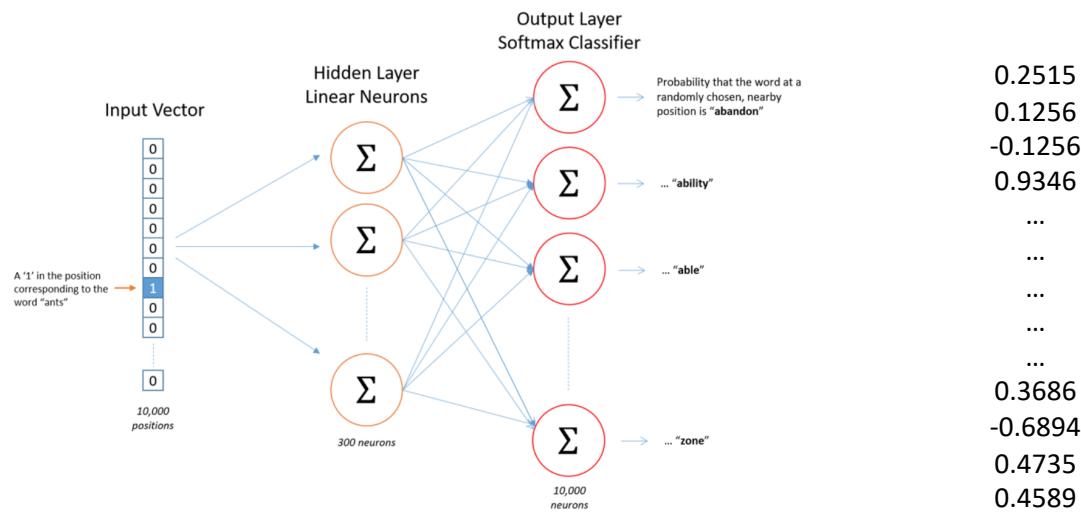
tisoy weakens  
tropical depression  
leave par

though tisoy kammuri              longer inside  
philippine area              responsibility  
            heavy rain due              surge  
northeast monsoon              tail end      cold  
front

# VECTORIZATION USING CORPUS-TRAINED WORD EMBEDDINGS FROM WORD2VEC\*

Words per Article → Word Embeddings Using Word2Vec\* → Vector Representation

tisoy	area
weakens	responsibility
tropical	heavy
depression	rain
leave	due
par	surge
though	northeast
tisoy	monsoon
kammuri	tail
longer	end
inside	cold
philippine	front



Architecture for skip-gram model. Source: [McCormickml tutorial](#)

Processed text data

Using a **skip-gram word2vec model** allows the prediction of context given a word.

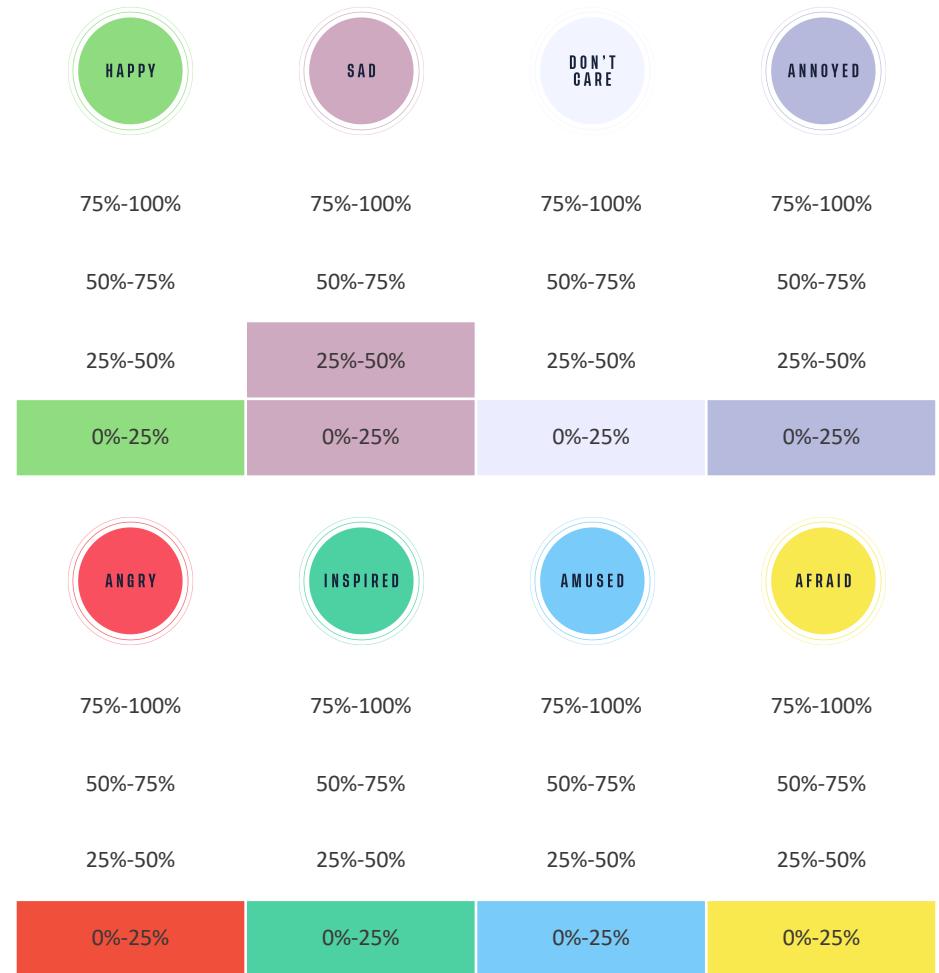
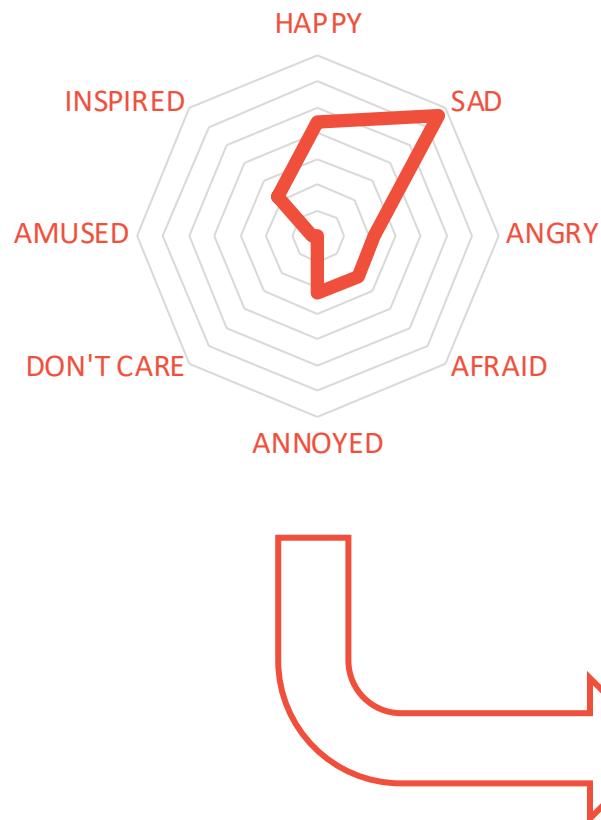
Parameters: Epochs = 5, Context size = 10 words

Set to be a **300-dimension vector**

\*neural network architecture used for word embeddings is beyond the scope of the project

<sup>1</sup> Kulshrestha, R. (2019) NLP 101: Word2Vec -- Skip-gram and CBOW. Towards Data Science. Accessed August 31, 2020. Retrieved from <https://towardsdatascience.com/nlp-101-word2vec-skip-gram-and-cbow-93512ee24314>

# TIERING OF EMOTION INTENSITY PER EMOTION



# OVERSAMPLING TECHNIQUE USING SMOTE

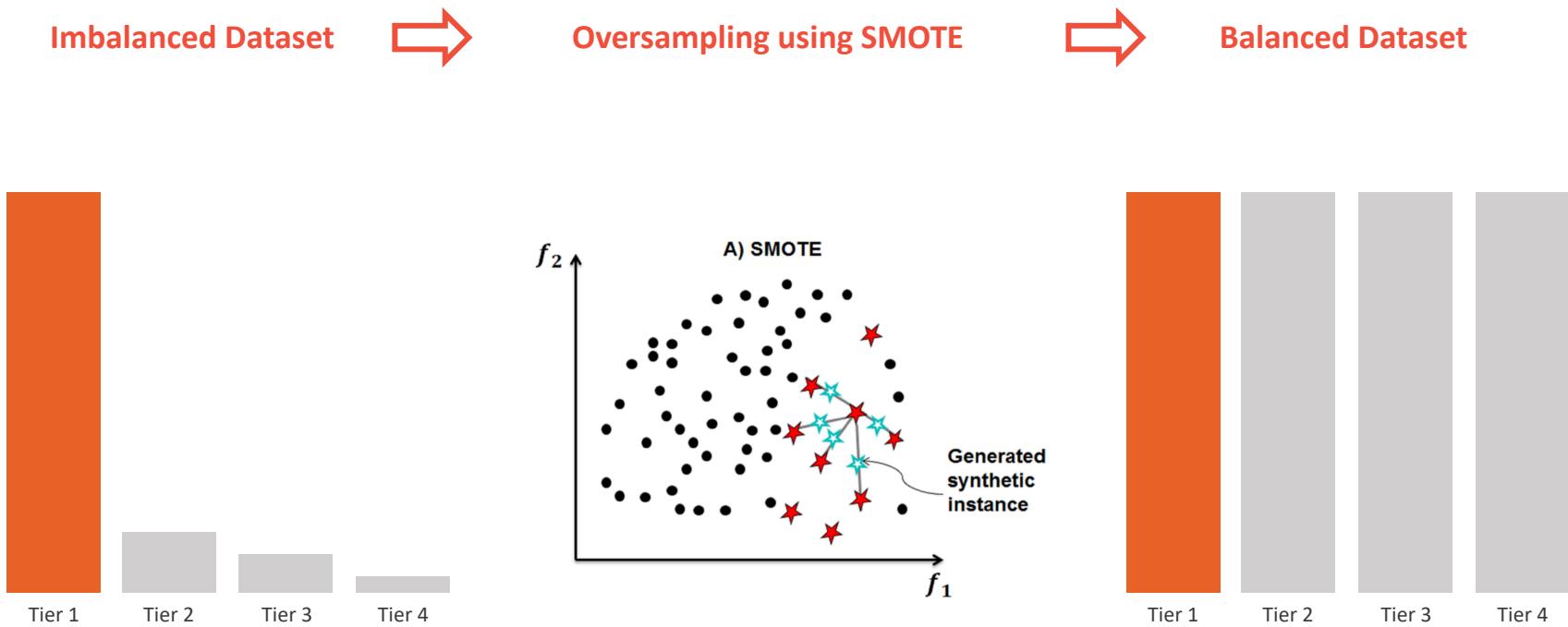


Image From: Soufan, Othman & Ba alawi, Wail & Afeef, Moataz & Essack, Magbubah & Rodionov, Valentin & Kalnis, Panos & Bajic, Vladimir. (2015). Mining Chemical Activity Status from High-Throughput Screening Assays. PloS one. 10. e0144426. 10.1371/journal.pone.0144426

# TRAINING OF CLASSIFIERS FOR EACH EMOTION USING A 80-15-5 SPLIT



# EMOTION VECTOR: ENSEMBLE TREES PROVIDED BEST F1 SCORES



	HAPPY	SAD	DON'T CARE	ANNOYED
--	-------	-----	------------	---------

**Classifier Model**

ETS

ETS

RF

ETS

**Model Performance: F1 Score**

56.50%

77.68%

79.04%

88.56%

**Benchmark Performance: F1 Score  
Dummy Classifier**

45.33%

64.86%

73.42%

82.03%



	ANGRY	INSPIRED	AMUSED	AFRAID
--	-------	----------	--------	--------

**Classifier Model**

ETS

ETS

RF

RF

**F1 Score** is the harmonic mean of Precision and Recall. It provides a better measure of misclassified cases than Accuracy scores

**Model Performance: F1 Score**

76.77%

79.45%

83.67%

91.46%

**Benchmark Performance: F1 Score  
Dummy Classifier**

58.59%

73.03%

79.24%

87.68%

# PUTTING THE MODELS INTO ACTION



## NLEX plays spoiler, deals TNT first loss behind magnificent Harris

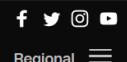
OCTOBER 26, 2019 1:50 PM PHT

DELFIN DIOQUINO



- Manny Harris unleashes 28 of his 40 points in the second half as NLEX ends TNT's seven-game unbeaten run
- 

HAPPY	SAD	DON'T CARE	ANNOYED
4	1	1	1
ANGRY	INSPIRED	AMUSED	AFRAID
3	1	1	1



Regional

### No damage to marine life in El Nido backhoe incident, official says

By CNN Philippines Staff

Published Sep 1, 2020 7:17:59 PM

Like Share Sign Up to see what your friends like.



LEONIDES ECHAGÜE/FACEBOOK

**Metro Manila (CNN Philippines, September 1)** — An environmental official assured that no marine life was harmed by the backhoe caught on video to be cruising along the shoreline in El Nido, Palawan on Aug.

HAPPY	SAD	DON'T CARE	ANNOYED
1	1	1	1
ANGRY	INSPIRED	AMUSED	AFRAID
4	1	1	1

## SUMMARY AND CONCLUSIONS



A simple way for news agencies, marketers, and corporates to **recognize reader-perspective emotion** distribution from text



Improve **writing style** by using the emotion vector to deliver more impactful stories



Generate an **emotion vector** using classifier models to measure relative strengths of emotions that can surpass benchmark measures

## OPPORTUNITIES FOR FUTURE WORK



Use Advanced Natural Language Processing Methods and Deep Learning Models



Compare **performance using finer intervals**



Extend the study to return continuous values using **Multi-Output Regression algorithms**

# RAPPLER'S MOOD METER



“

**There can be no transforming of darkness  
into light and of apathy into movement  
without emotion.**

*— Carl Jung*

“

**There can be no transforming of darkness  
into light and of apathy into movement  
without emotion.**

*— Carl Jung*



# THANK YOU!

Hmm... Intense! Emotion Distribution  
Recognition of News Articles using Rappler's  
Crowd-Annotated Mood Metrics

Join our **Zoom Parallel Meeting** going to  
<http://bit.ly/RapplerMoodAnalysis>  
or by scanning the QR code below



**Tonee Bayhon**



**Raf Madrigal**



**Matt Maulion**



**Aids Perez**



**Kris Tabong**

