BIBLIOTHECA ALEXANDRINA
مكتبة الإسكندرية

SIMAR
Serageldin Institute for Multi-disciplinary Advanced Research

Sentiment Analysis for Arabic Social Media

Noha Adly
Sara El Shobaky
Bibliotheca Alexandrina

1

What is sentiment Analysis?

Discovering people opinions, emotions and feeling about a topic being a product, a service, etc, …

The movie is great

The movie is horrible

The movie is 90 minutes

Measure Public Opinion about Elections

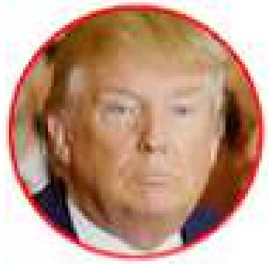
Hillary Clinton

Democratic Party

Mentions	
Total	785
No. of positive	151
No. of negative	155
No. of neutral	479

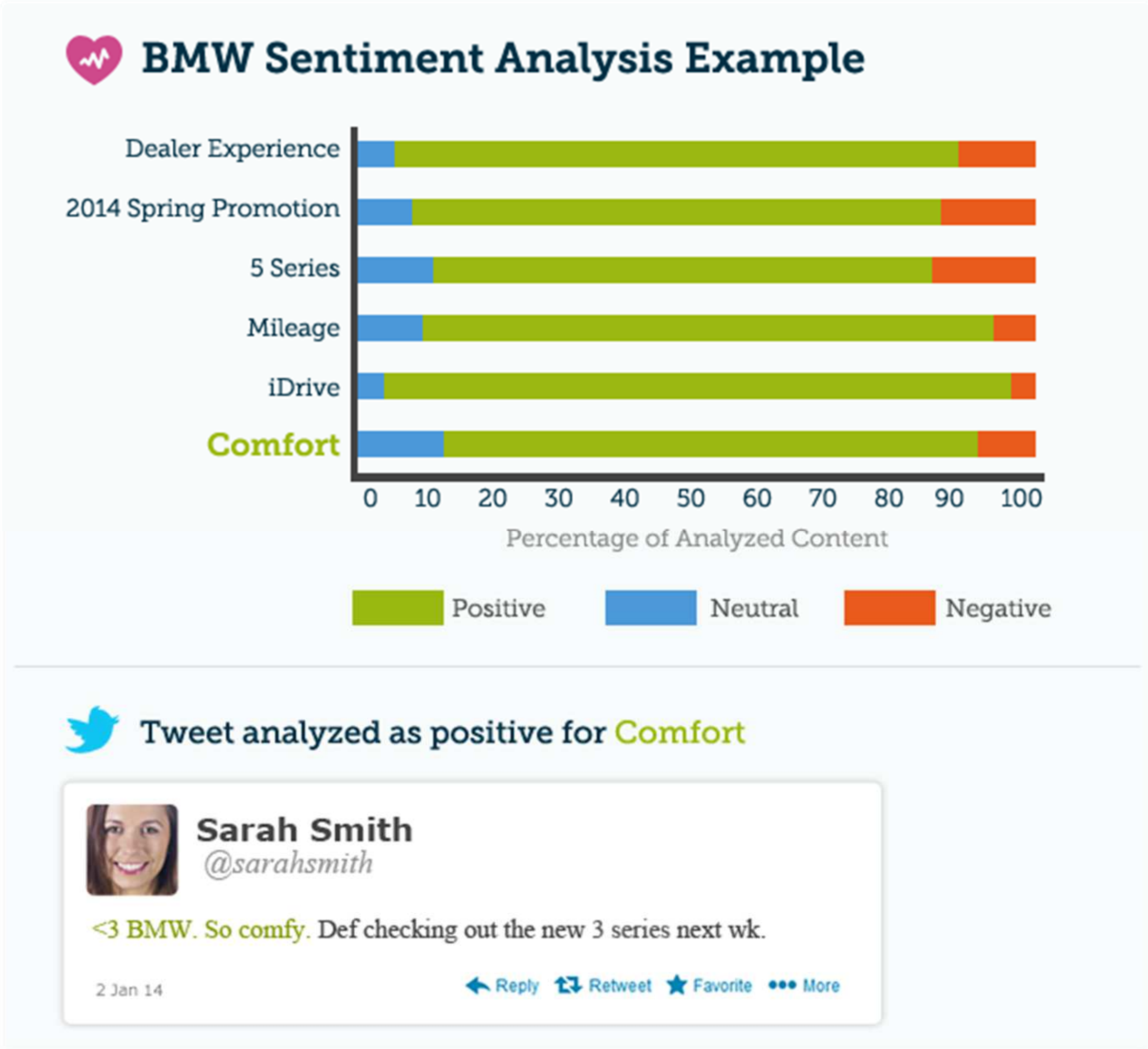
Donald Trump

Republican Party



	Mentions
1,412	Total
255	No. of positive
402	No. of negative
755	No. of neutral

Brand Monitoring



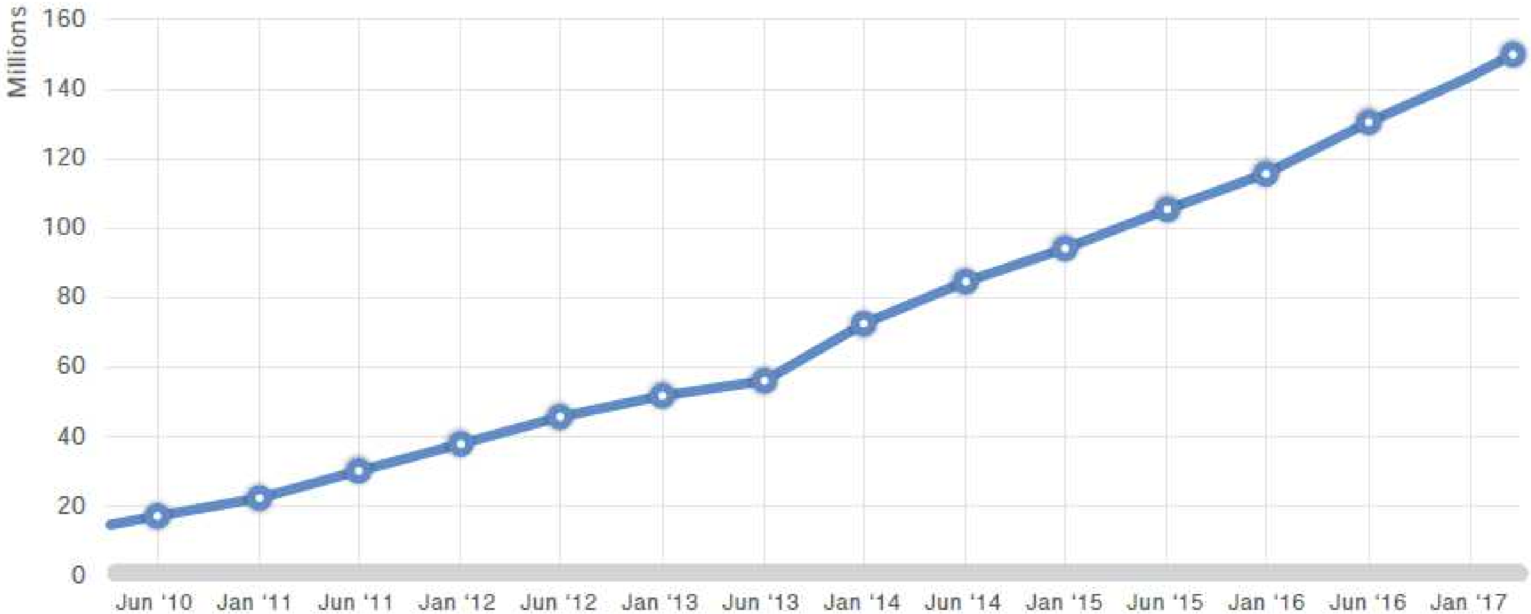
Source: https://www.crowdsource.com/solutions/content-moderation/sentiment-analysis/

Where we can find the public opinion?

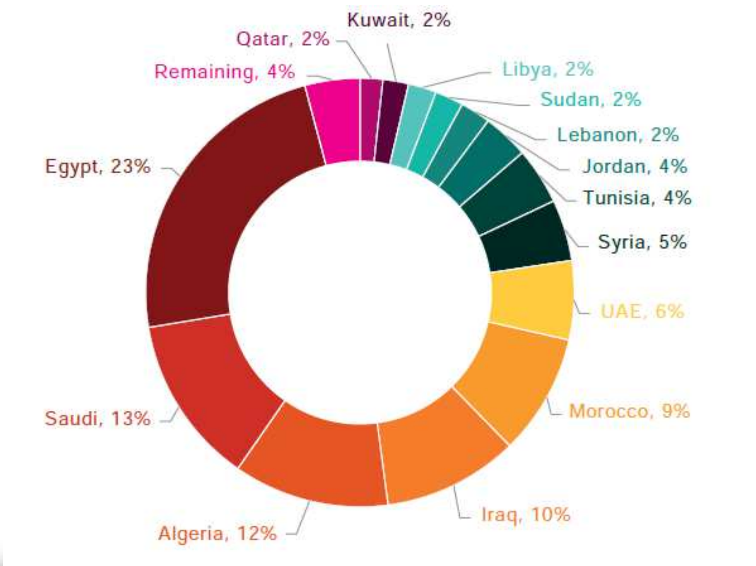
Why Social Media?

https://www.smartinsights.com/internet-marketing-statistics/happens-online-60-seconds/

Social Media in the Arab Region



Growth of Facebook Users in the Arab Region 2010-17



Distribution of Facebook Users in Arab Region (2017)

Salem, F. (2017). The Arab Social Media Report 2017. Arab Social Media Report Series (Vol. 7). Dubai: Governance and Innovation Program, MBR School of Government.

Why Arabic Social Media?

- Lack of research on sentiment analysis in Arabic
- Colloquial Arabic is challenging

Colloquial Arabic is challenging

Modern Standard Arabic vs Egyptian Colloquial Arabic

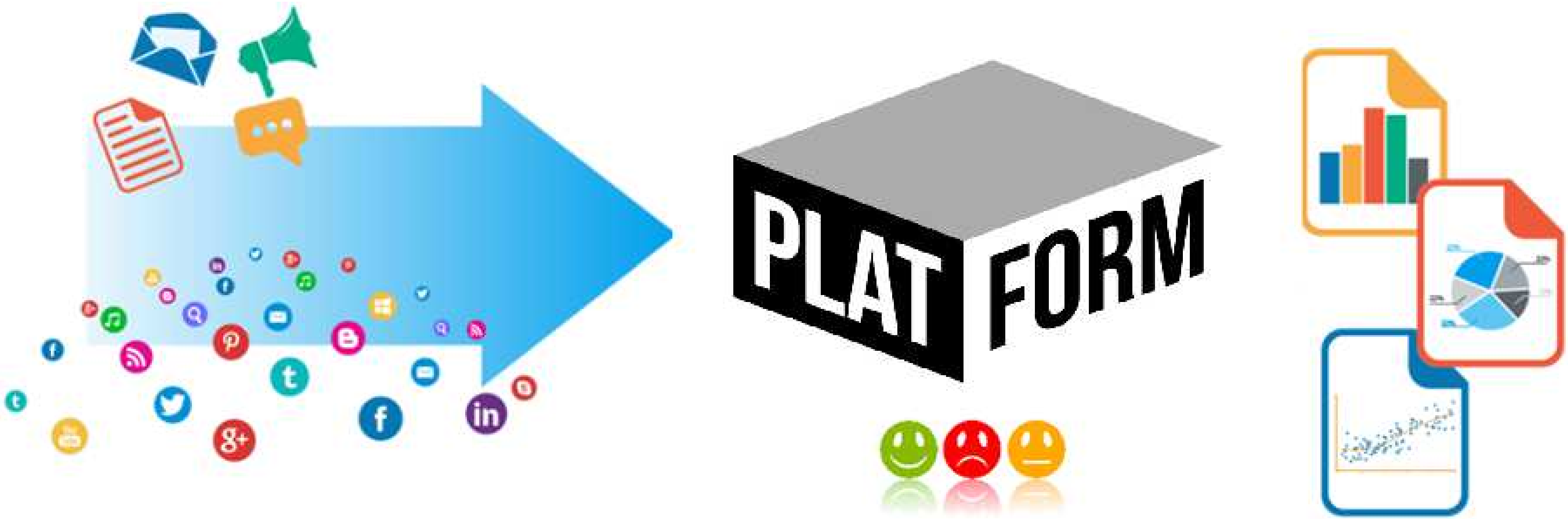
This Masculine singular noun	هذا	ده
This Feminine singular noun	هذه	دي
This Plural noun	هؤلاء	دول

On Social Media

Loooovely	جميييل	@, #, up	tag, hashtag, follow
Wow	واااو (Transliteration)	XOXO, OMG	Hugs and kisses, Oh my God
Dangerous	خطير	Lol, ههههههه	Laugh out Loud

Objective

Build a platform for applying sentiment analysis on Social media in Arabic

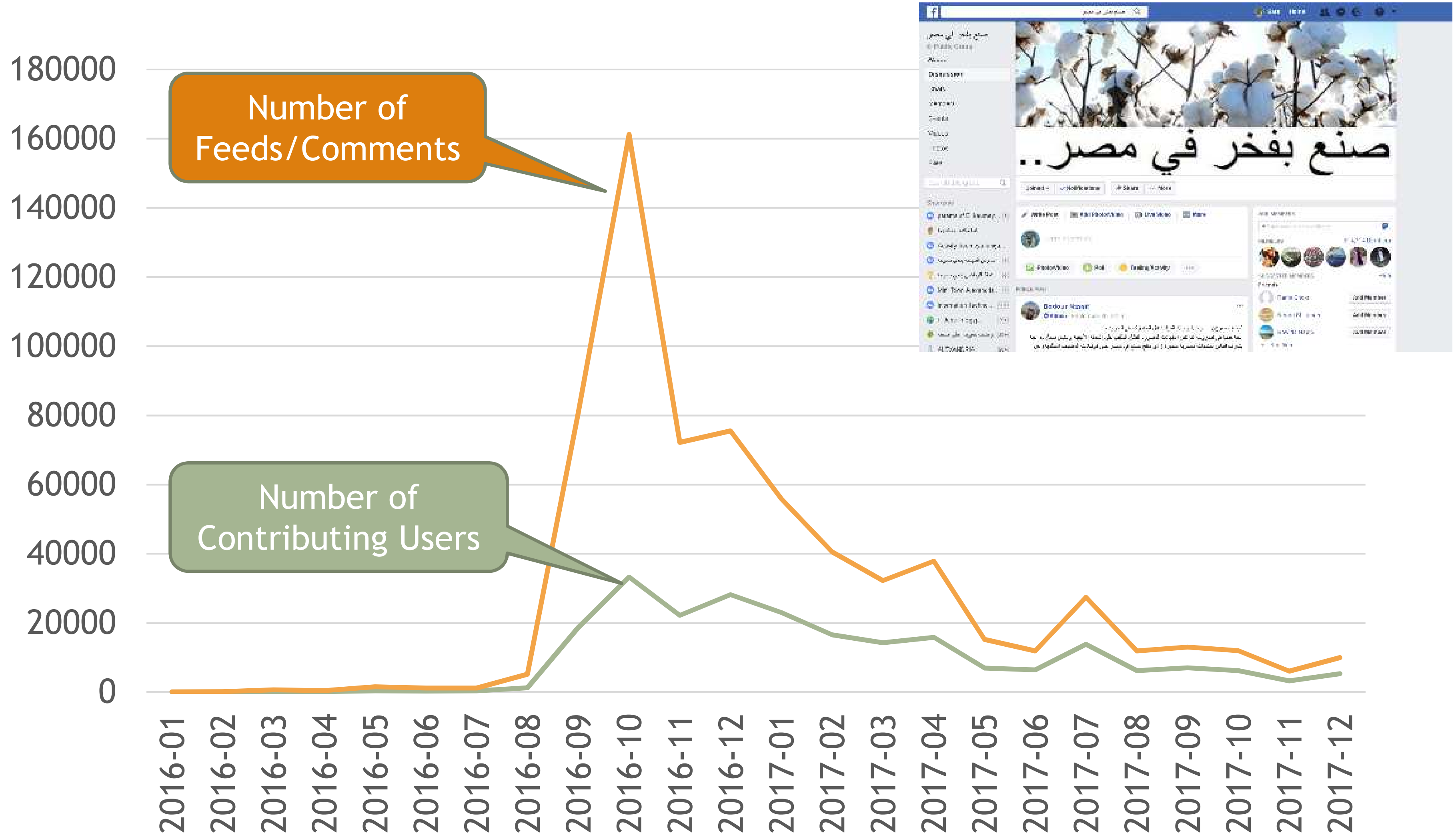


10

Case Study: Products Made in Egypt

Measure opinion toward products made in Egypt

Example: a group on Facebook



614,197 Members, 676,151 Feeds/comments, 155,926 Shares

Study the effect of special events

- Media: e.g. TV shows

- Policies: e.g. صاحبة السعادة
  - Floating Egyptian pound ,
  - Increase fuel price,
  - Importing ban.

13

SIMAR

Study the effect of special events

- Media: e.g. TV shows
- Policies: e.g.
  - Floating Egyptian pound
  - Increase fuel price
  - Importing ban

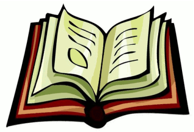
14

SIMAR

Study sentiment towards different aspects

SIMAR

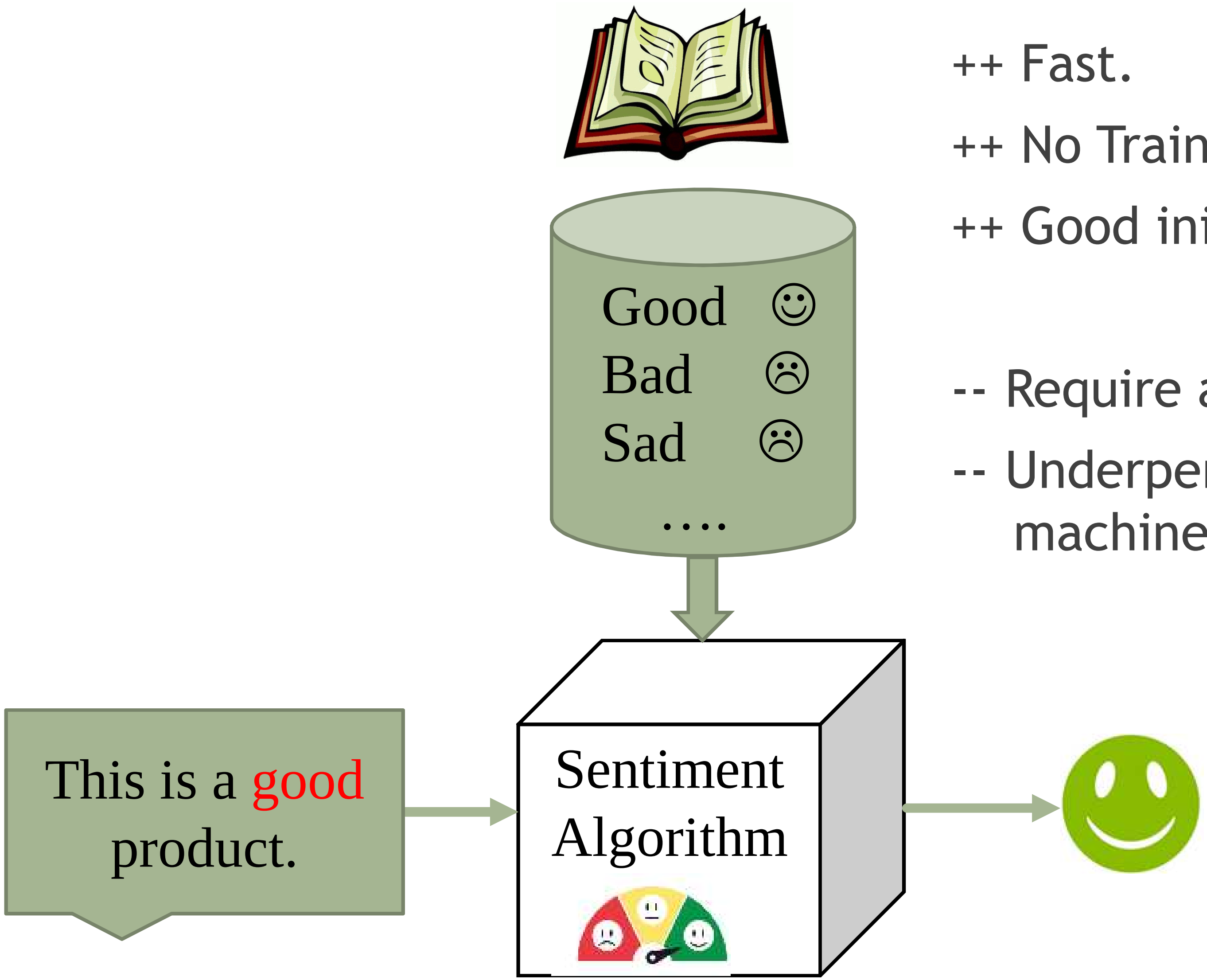
Sentiment analysis approaches

- Lexicon-based 
- Machine-Learning 

Each has advantages and disadvantages…

16

Lexicon Based Approach



++ Fast.

++ No Training data necessary.

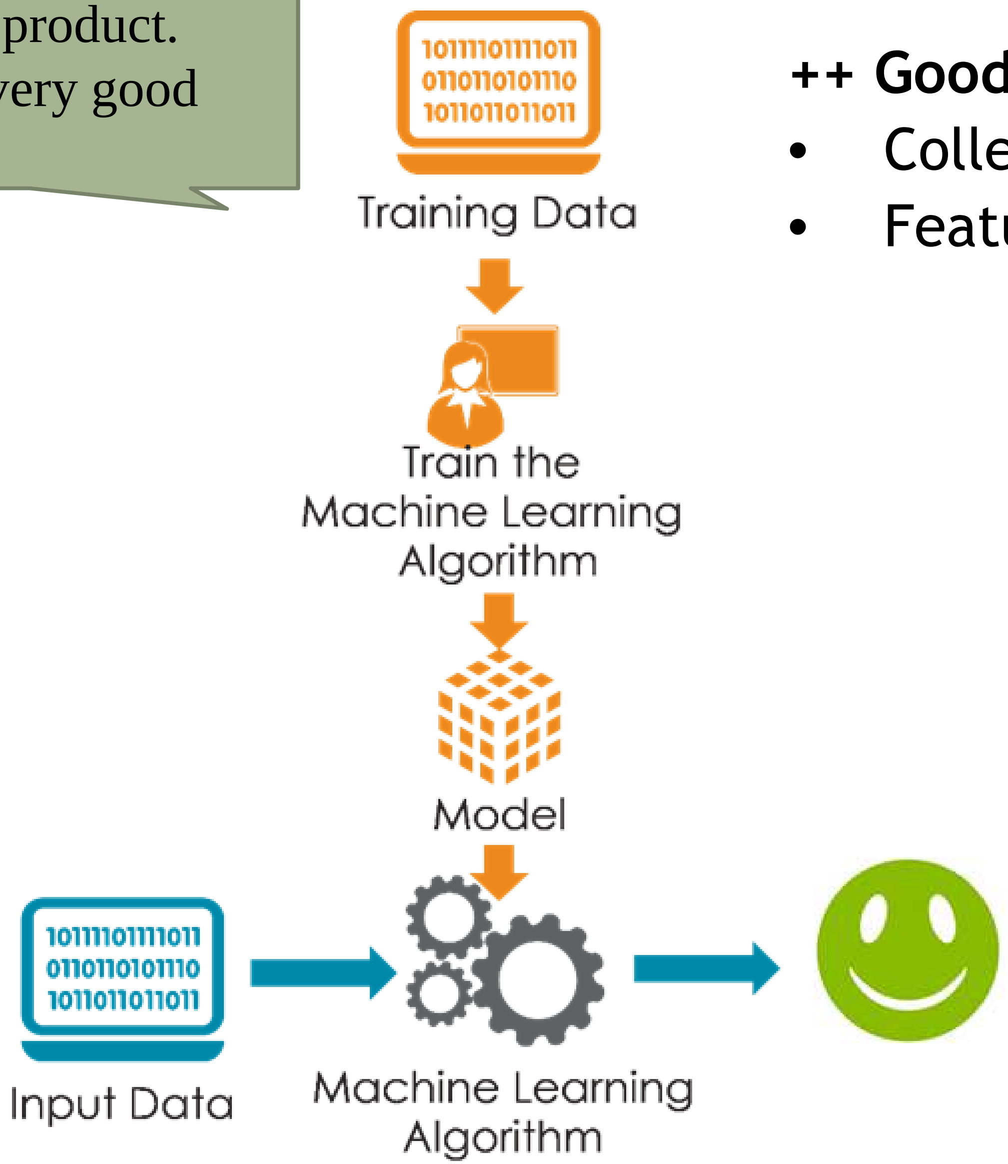
++ Good initial accuracy.

-- Require a polar lexicon.

-- Underperform well trained machine learning.

Machine Learning Approach

- This is a bad product.
- The price is very good
- ...



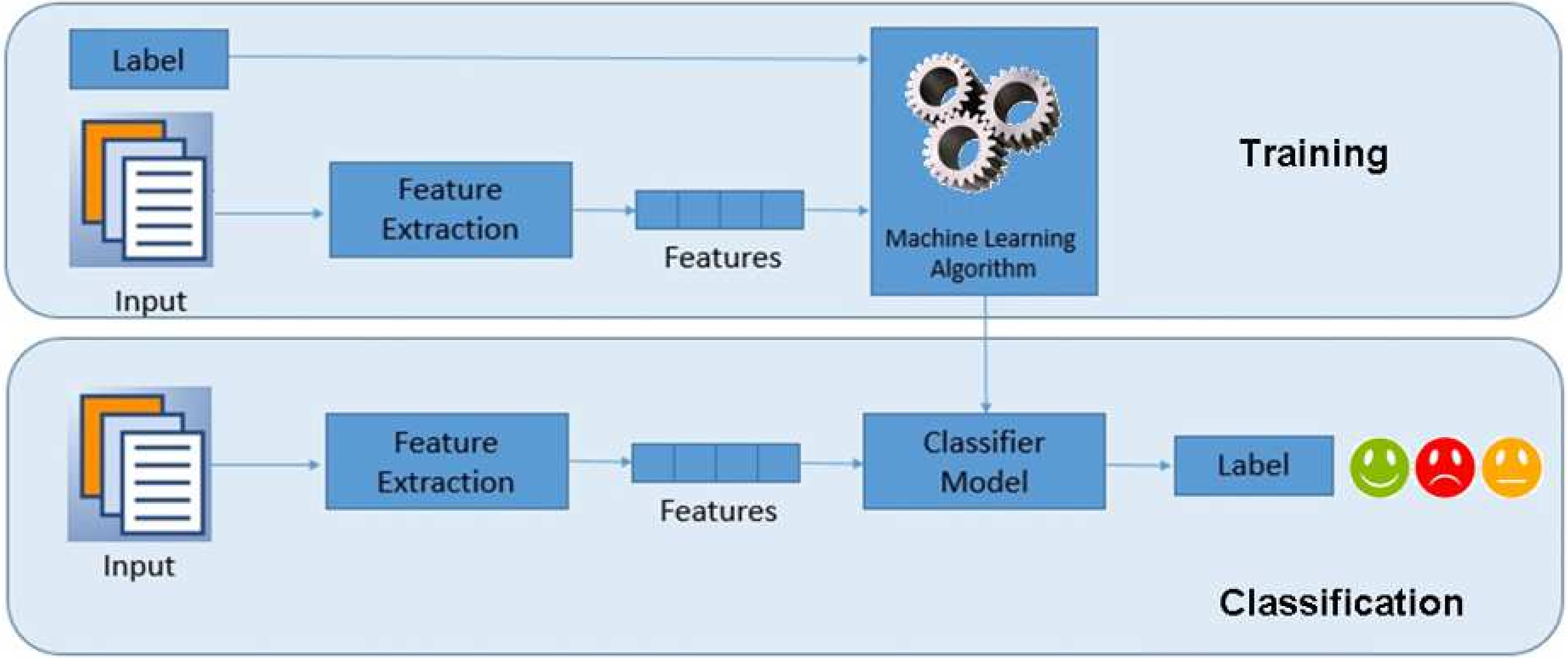
++ Good accuracy based on:

- Collected Training data
- Features extraction

This is a
good
product.

Feature Extraction

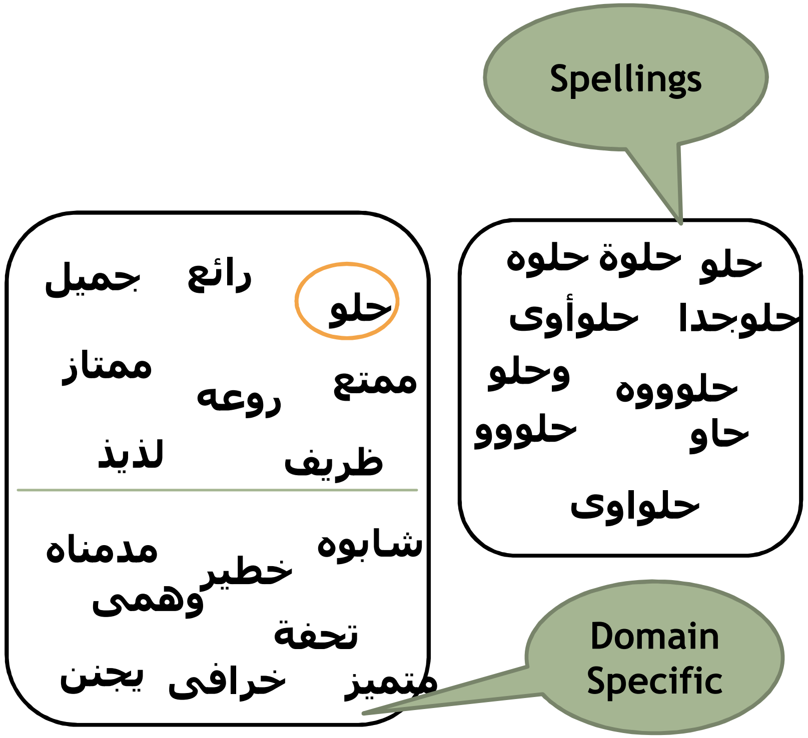
Transforming raw data into features that can be used as input for a learning algorithm



Hand-crafted features

- Unigram, bigram,…

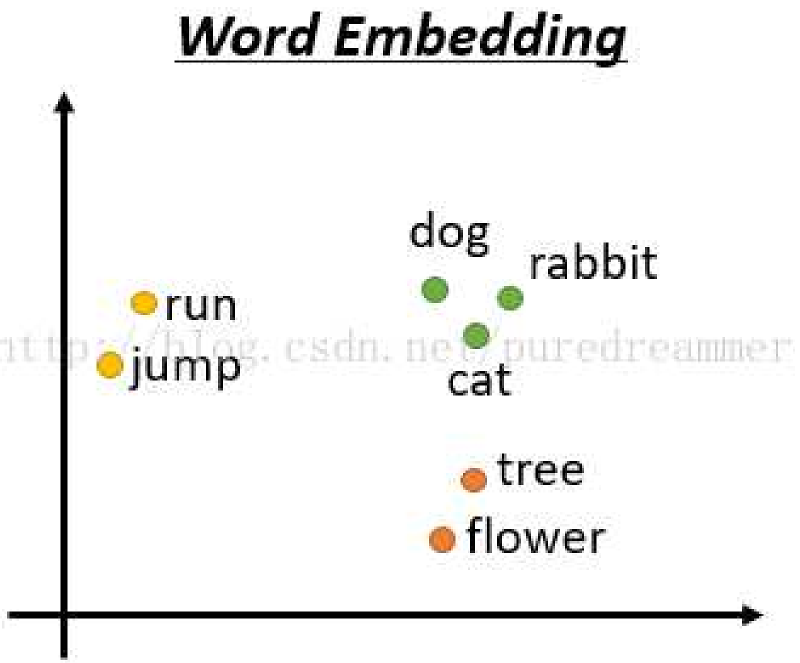
- Stemming
  - وبيمشي → يمش
  - بيوكلينا -> وكل
  - تابليت → تبل



- POS-Tagging (Names, verbs, adjectives,…)
  - انا بحبه وبجيبه ع طول
  - اللي بيسئلو عليه طعمه حلو جدااااااا
  - ولا ليه اى لازمة وغالى على الفاضى

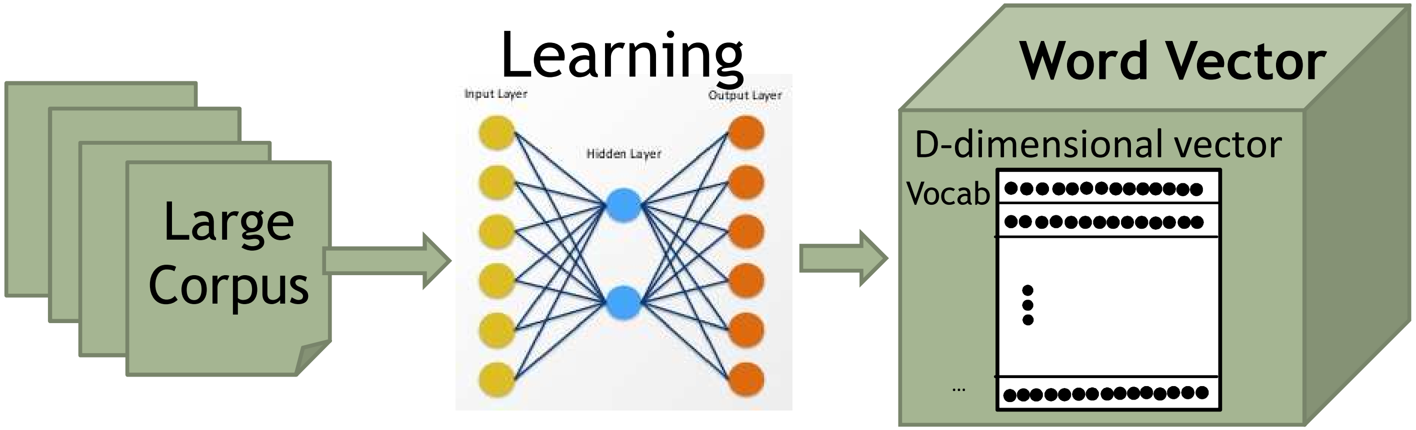
Word Embedding for Feature Extraction

- Using Neural word embedding to represent each word with a low-dimensional vector
- Similarity in meaning = Similarity in vectors
  - Two words have close meanings if their local neighborhoods are similar



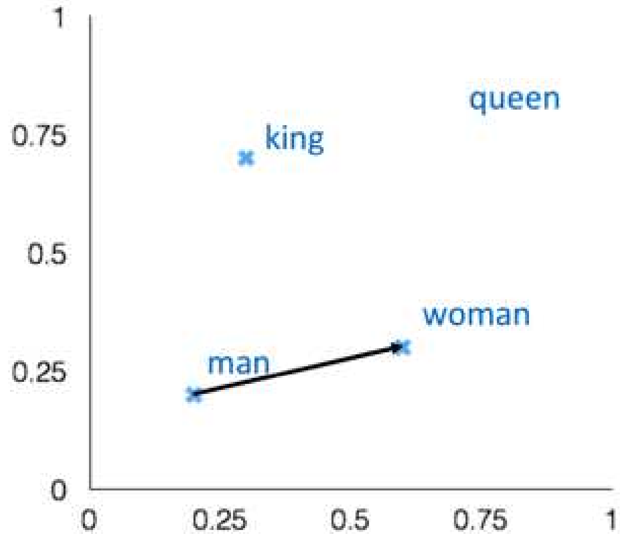
Word Embedding for Feature Extraction

- Trains a Neural Network Model from a large corpus

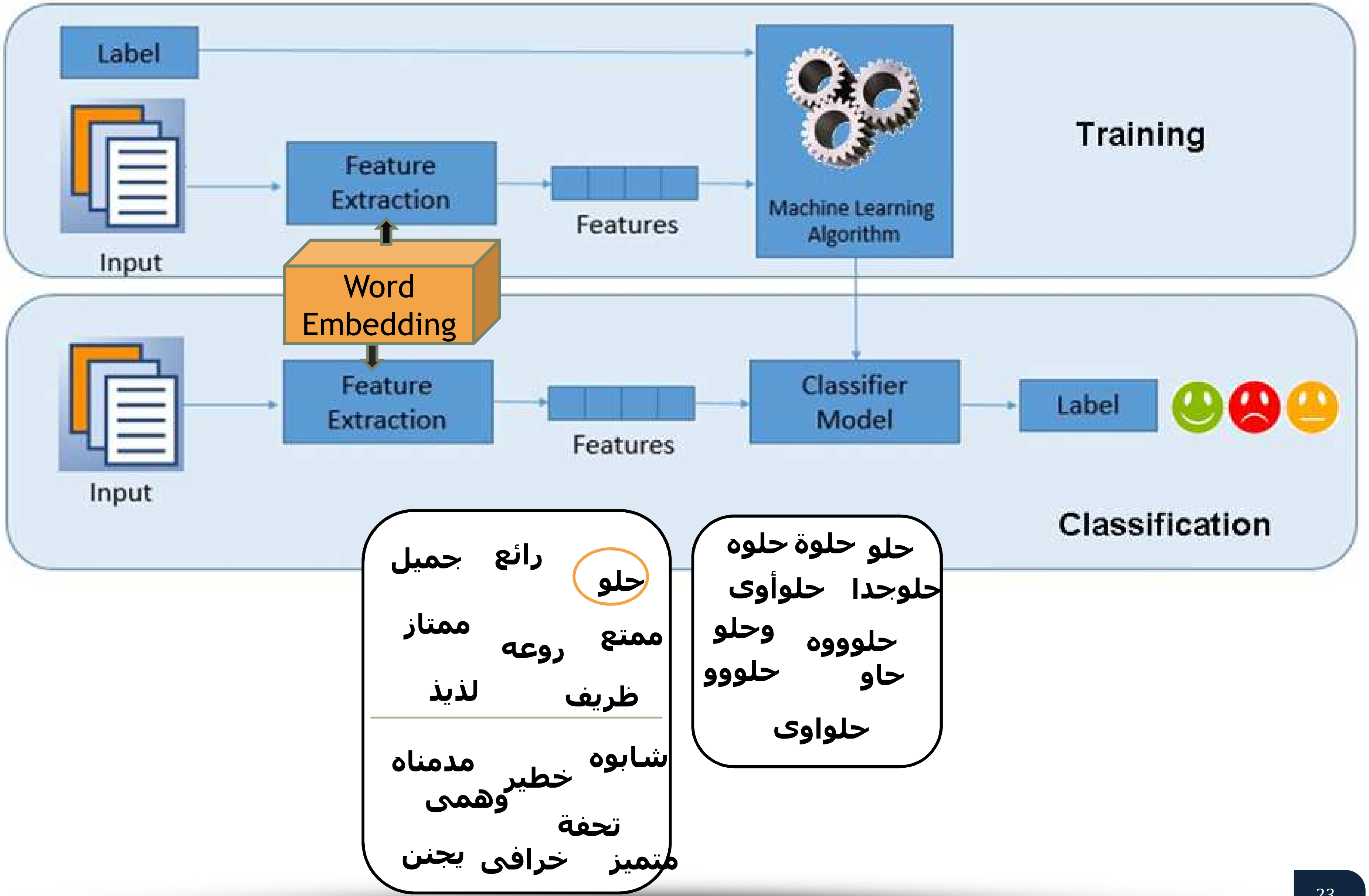


man:woman :: king:?

+	king	[0.30 0.70]
-	man	[0.20 0.20]
+	woman	[0.60 0.30]
	queen	[0.70 0.80]



Using the embedding in feature extraction



SIMAR

How to generate a Word Embedding?

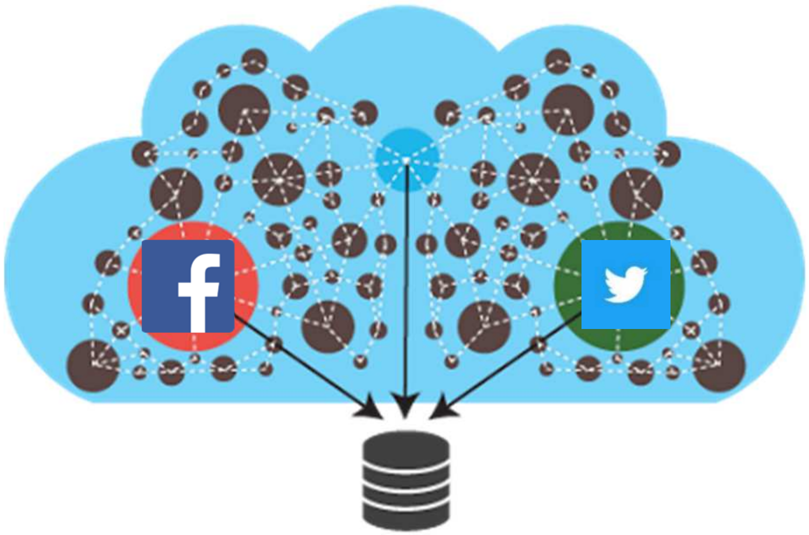
1. Collect a larges Corpus

2. Preprocessing the corpus

3. Learning the Embedding

1. Collect a large Corpus

- We collected a large Arabic corpus from social media
  - Source: Facebook and Twitter
  - Size: more than 50 Million feeds/comments/tweets
  - Domain: Products domain
  - Number of tokens (476 Million token)



2. Preprocessing the corpus

- Decreasing number of vocabularies in a corpus decreases the computation complexity while learning the embedding

Filtration	
Foreign language	English, French, …
Repeated characters	طويييييييييييل
Punctuations	“ : ; , ‘
Diactrice	◌َ◌ً◌ّ◌ُ
Tatweel	طويــــــــــل

Characters Standardization	
أ إ ا	ا
ى ي	ي
ة ه	ه

Unification	
? ؟	__QUES
http://www.....	__URL
@PersonName	__TAG
م f up	__FOllOW
1234 567890	__NUM
☺ :-D LOL, ….. ☹ : (… 😐 : | …	__EMOT_POS, __EMOT_NEG, __EMOT_NEU

3. Learning the Word Embedding

- **Learning Models**
  - CBOW (Continuous Bag of Words)
    Predict the word given its context.
  - Skip Gram
    Predict the context given a word.
  - Glove (Global Vectors)
    Tries to capture the counts of overall statistics.

Word2Vec
Mikolov, 2013

Google

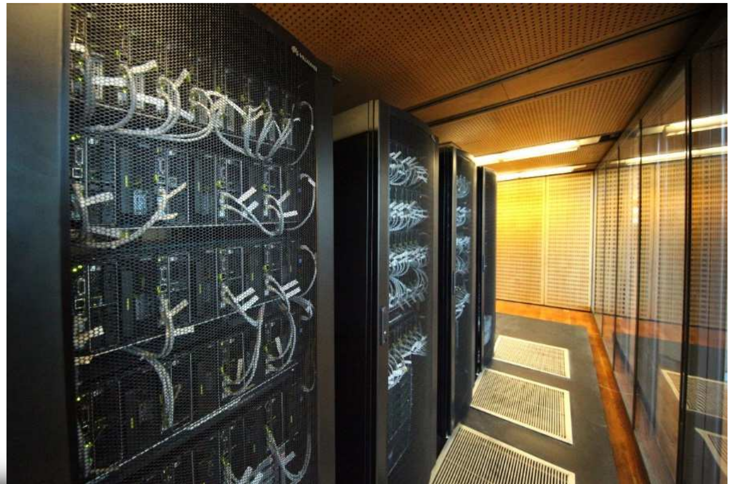
Pennington
2014

- **Learning Parameters**
  - *Vector Dimension*
  - Window-size to be included as the context of a target word
  - Number of training Iterations over the corpus

How to choose the best model?

- Learning Word Embeddings on BA HPC
- More than 45 different embedding have been generated with different models and parameters

- On BA HPC (92 nodes, 128G RAM, 24 threads)
  - Preprocessing the corpus
  - Learning one embedding took up to 13 hours.

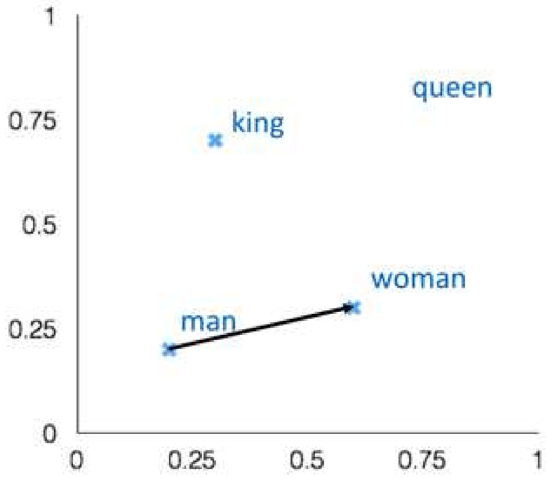


Analogies tests

Accuracy based on Arabic Analogies test

Man is to woman as king is to ____ ?

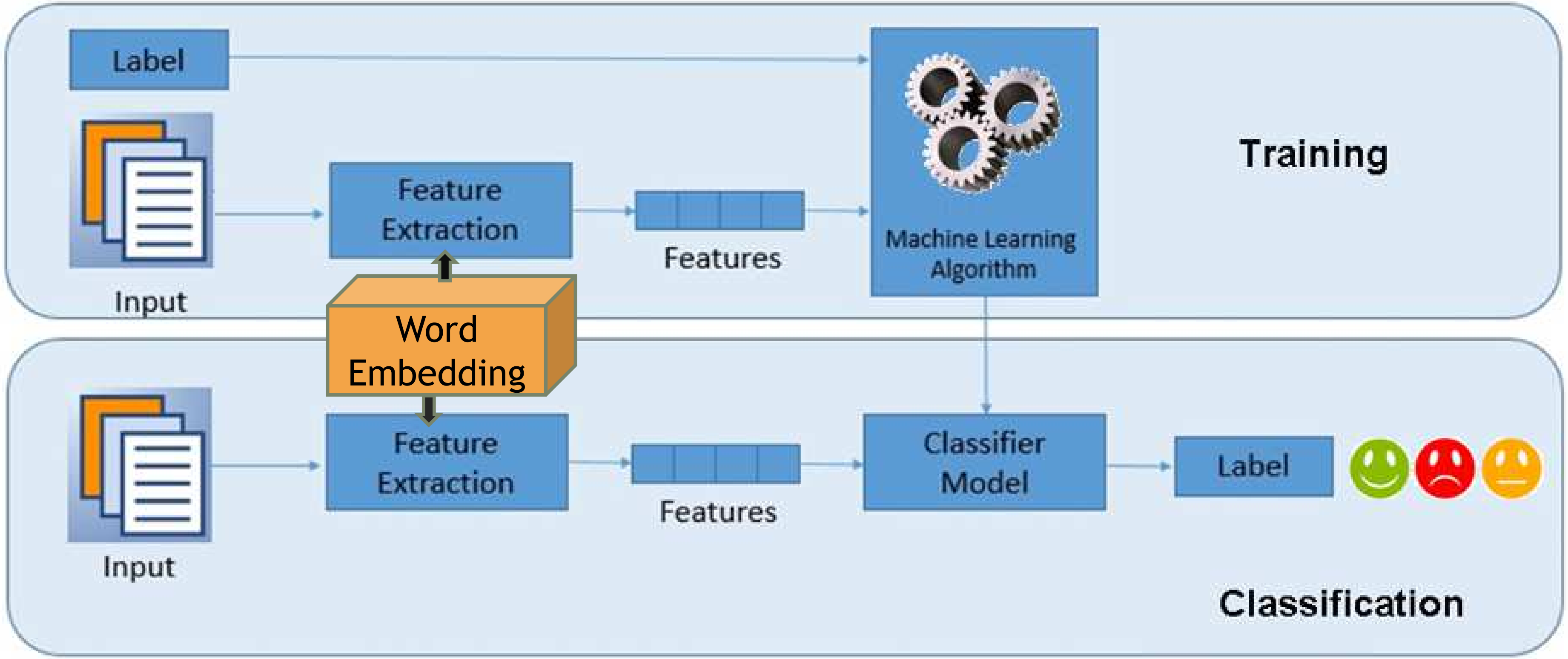
+	king	[0.30 0.70]
-	man	[0.20 0.20]
+	woman	[0.60 0.30]
	queen	[0.70 0.80]



Dim.	Win.	Iter.	CBOW Accu.(%)	SG Accu(%)	GloVe Accu(%)
100	10	15	59.98	47.27	49.93
200	10	15	65.77	57.80	57.10
300	10	15	67.70	59.49	58.58
400	10	15	67.77	60.56	58.10
500	10	15	64.35	61.15	0.04?
300	5	15	66.49	62.84	56.50
300	10	15	67.70	59.49	58.58
300	15	15	67.46	56.83%	59.16
300	20	15	67.19	54.58	60.37
300	10	5	64.78	57.40	0.00?
300	10	10	66.61	59.16	58.10
300	10	15	67.70	59.49	58.58
300	10	20	67.60	60.10	58.31
300	10	25	63.50	59.28	58.60
300	10	30	63.65	59.98	58.29

- Glove model has the worst accuracy.
- Increasing dimensions increase the accuracy.
- Window-size 10 generates best results.
- 15 to 20 Iterations has the best results.

Sentiment Classification test



Split training data for evaluation

- 90% training data
- 10% test data

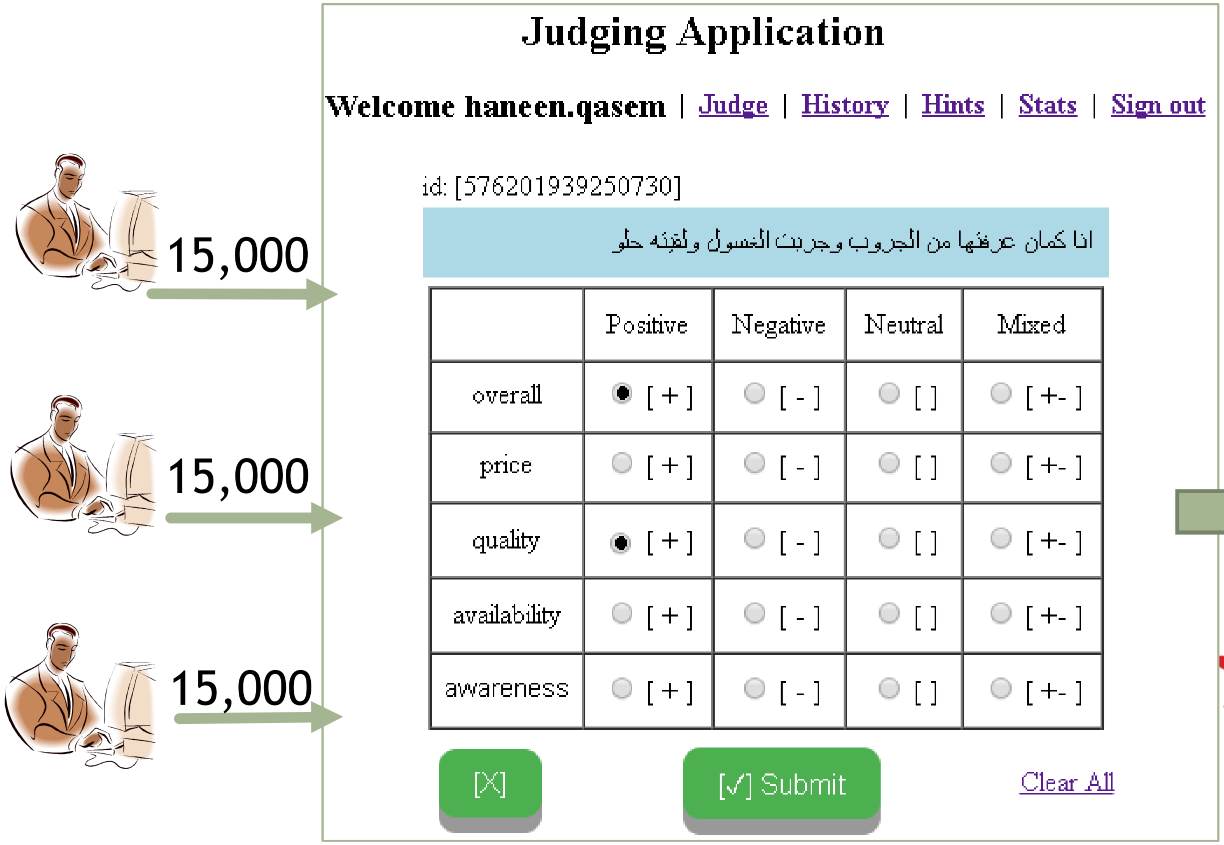
Classification Algorithm

- Support Vector Machine

Sentiment Classification in 2 steps:

1. Subjectivity Classification
   (Subjective/Objective)
2. Polarity Classification
   (Positive/Negative)

Training data



Overall	5,127	3,453	5,304
Price	469	744	536
Quality	4,238	2,018	151
Availability	1,082	476	183
Awareness	93	39	3

Best embedding for Sentiment Classification task

Parameters			CBOW Polarity			CBOW Subjectivity			SG Polarity			SG Subjectivity		
Dim	Win	Iter.	F	Pr	Rec	F	Pr	Rec	F	Pr	Rec	F	Pr	Rec
100	10	15	0.912	0.911	0.913	0.863	0.853	0.873	0.9142	0.912	0.917	0.859	0.849	0.869
200	10	15	0.921	0.924	0.918	0.868	0.858	0.878	0.925	0.926	0.923	0.867	0.856	0.879
300	10	15	0.927	0.929	0.924	0.877	0.871	0.885	0.9223	0.925	0.92	0.874	0.868	0.88
400	10	15	0.932	0.935	0.928	0.874	0.865	0.883	0.935	0.938	0.933	0.875	0.862	0.888
500	10	15	0.929	0.93	0.928	0.875	0.866	0.884	0.932	0.936	0.928	0.876	0.871	0.881
300	5	15	0.928	0.93	0.927	0.875	0.868	0.882	0.936	0.937	0.935	0.875	0.867	0.883
300	10	15	0.927	0.931	0.923	0.876	0.87	0.884	0.925	0.926	0.924	0.877	0.866	0.889
300	15	15	0.929	0.929	0.93	0.874	0.871	0.876	0.929	0.928	0.929	0.87	0.866	0.874
300	20	15	0.927	0.928	0.927	0.878	0.865	0.89	0.935	0.938	0.932	0.873	0.865	0.883
300	10	5	0.924	0.929	0.919	0.876	0.868	0.885	0.931	0.934	0.929	0.875	0.866	0.884
300	10	10	0.93	0.933	0.926	0.876	0.869	0.884	0.929	0.933	0.924	0.869	0.864	0.875
300	10	15	0.929	0.935	0.923	0.874	0.868	0.88	0.93	0.931	0.928	0.873	0.865	0.88
300	10	20	0.928	0.936	0.921	0.875	0.873	0.878	0.925	0.93	0.92	0.877	0.869	0.885
300	10	25	0.928	0.928	0.928	0.876	0.872	0.879	0.927	0.931	0.923	0.872	0.866	0.878
300	10	30	0.923	0.926	0.92	0.873	0.865	0.882	0.93	0.932	0.928	0.871	0.866	0.876

Skip-Gram Model
400 Dimensions
10 Window-Size
15 Iterations

ACCURACY

87.5% Subjectivity Classification
93.5 % Polarity Classification

Category classification Approach

The same machine learning approach using the word embedding is used to classify the Dataset.

	F	Precision	Recall
Price	0.722	0.832	0.638
Quality	0.849	0.867	0.835
Availability	0.768	0.821	0.722

SIMAR

Experiments setup

- Dataset:
  - 645 Facebook page/group of products made in Egypt
  - Period from 1-1- 2016 till 31-12-2017
  - 4,104,538 Arabic feeds and comments out of 5,161,247
    - 3,409,461 Arabic text
    - 695,077 tags, follows,..

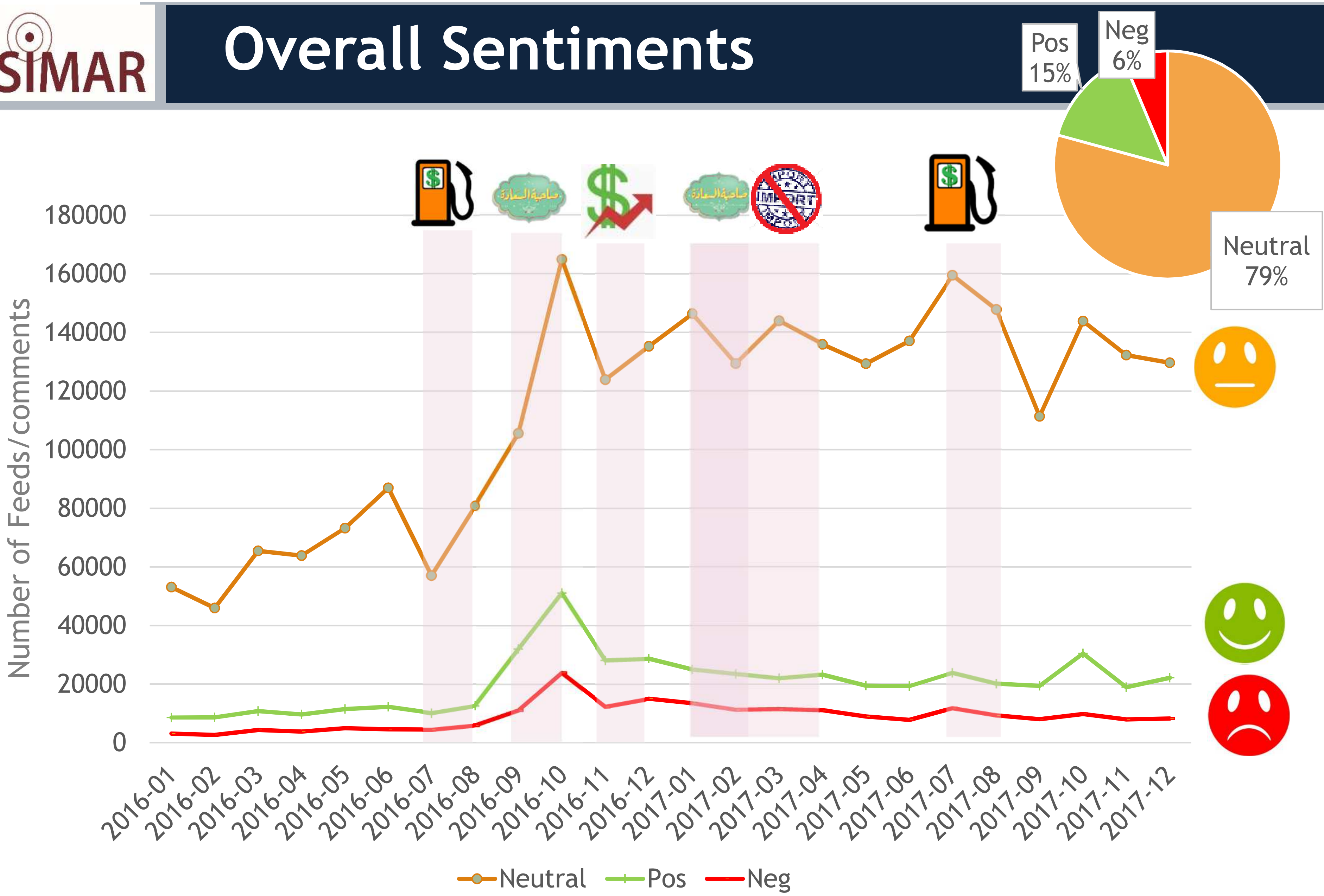
PROUDLY MADE IN EGYPT SEAL OF QUALITY G-0200-F-12-16

Proudly ... Made in Egypt

- Events

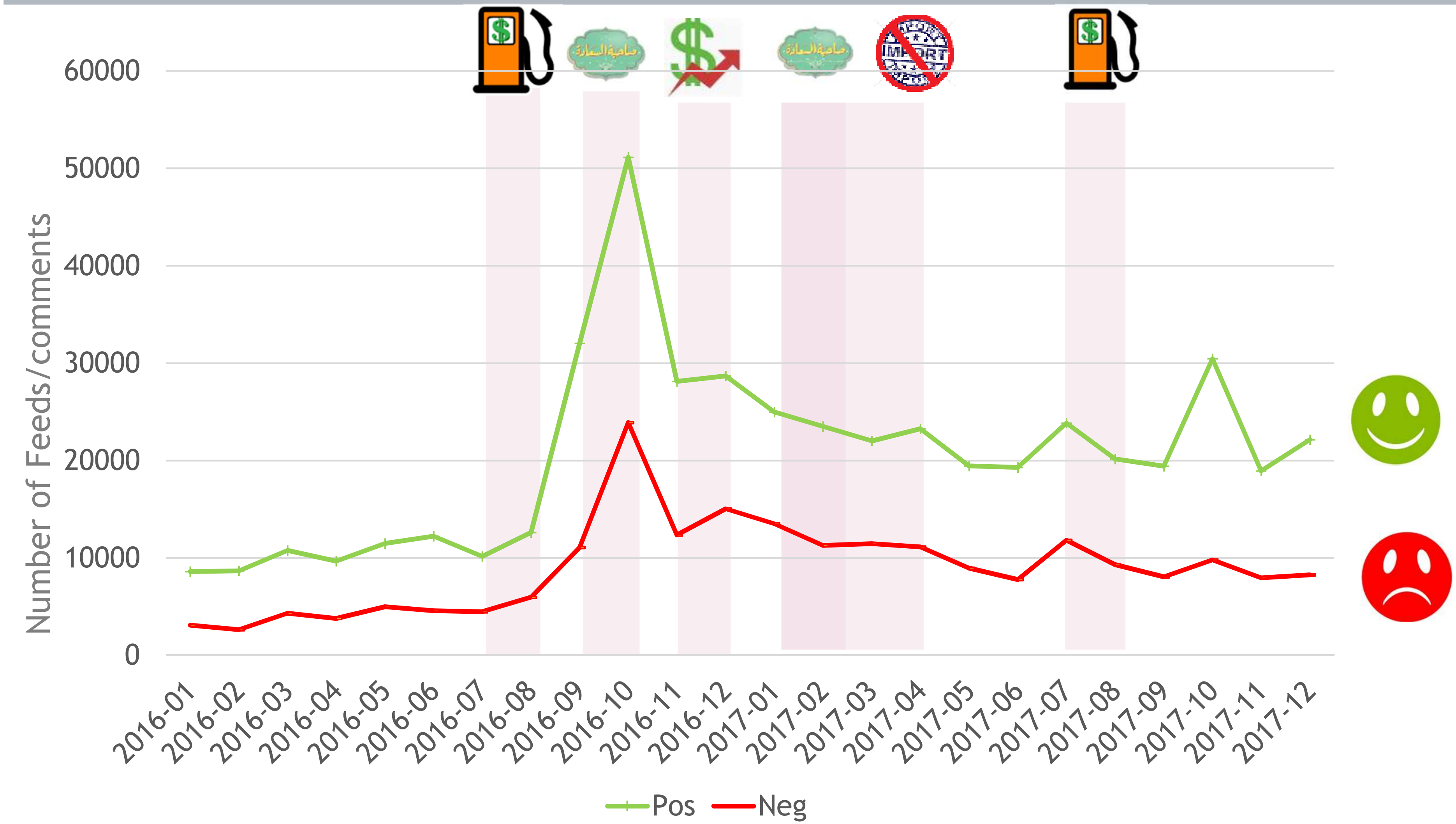
Date	Event
7/2016	Increase in fuel prices
9/2016	TV show 'Made in Egypt' episode 1.
11/2016	Floating the Egyptian pound
1/2017	TV show 'Made in Egypt' episode 2.
1-3/2017	Egyptian decision to stop importing
7/2017	Increase in fuel prices

Overall Sentiments



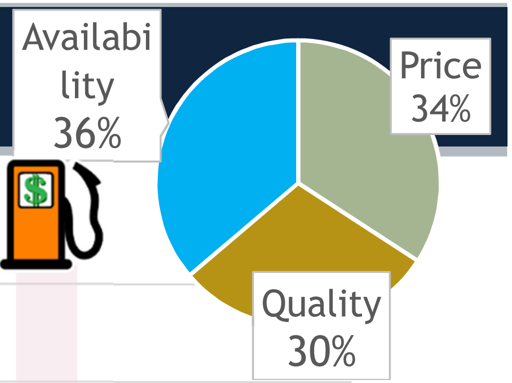
- 79% Neutral 14% positive posts 7% Negative posts

Overall Sentiments



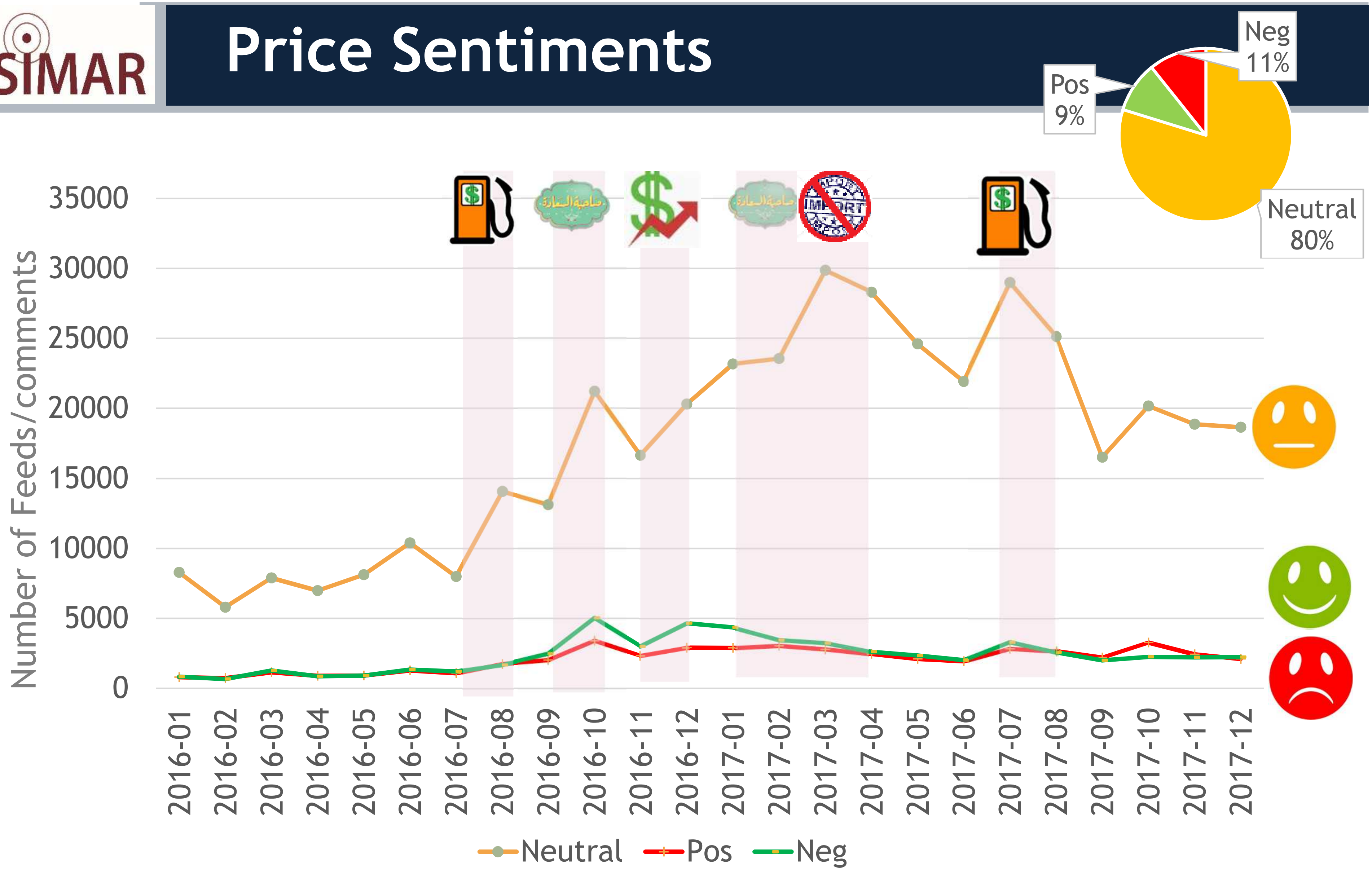
- Positive and Negative sentiments following the same pattern across time, with positivity almost double negativity

Interested aspects



- People are equally concerned to the 3 aspects.
- Quality was hot topic after 1st media show
- Availability was of interest when pound floated
- Prices became main interest during the importing ban and increase in fuel prices

Price Sentiments

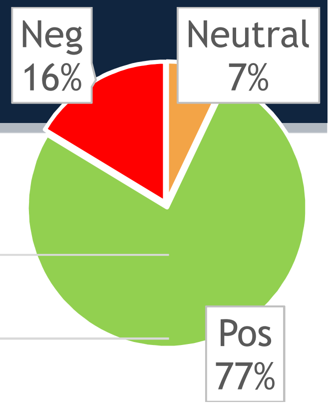


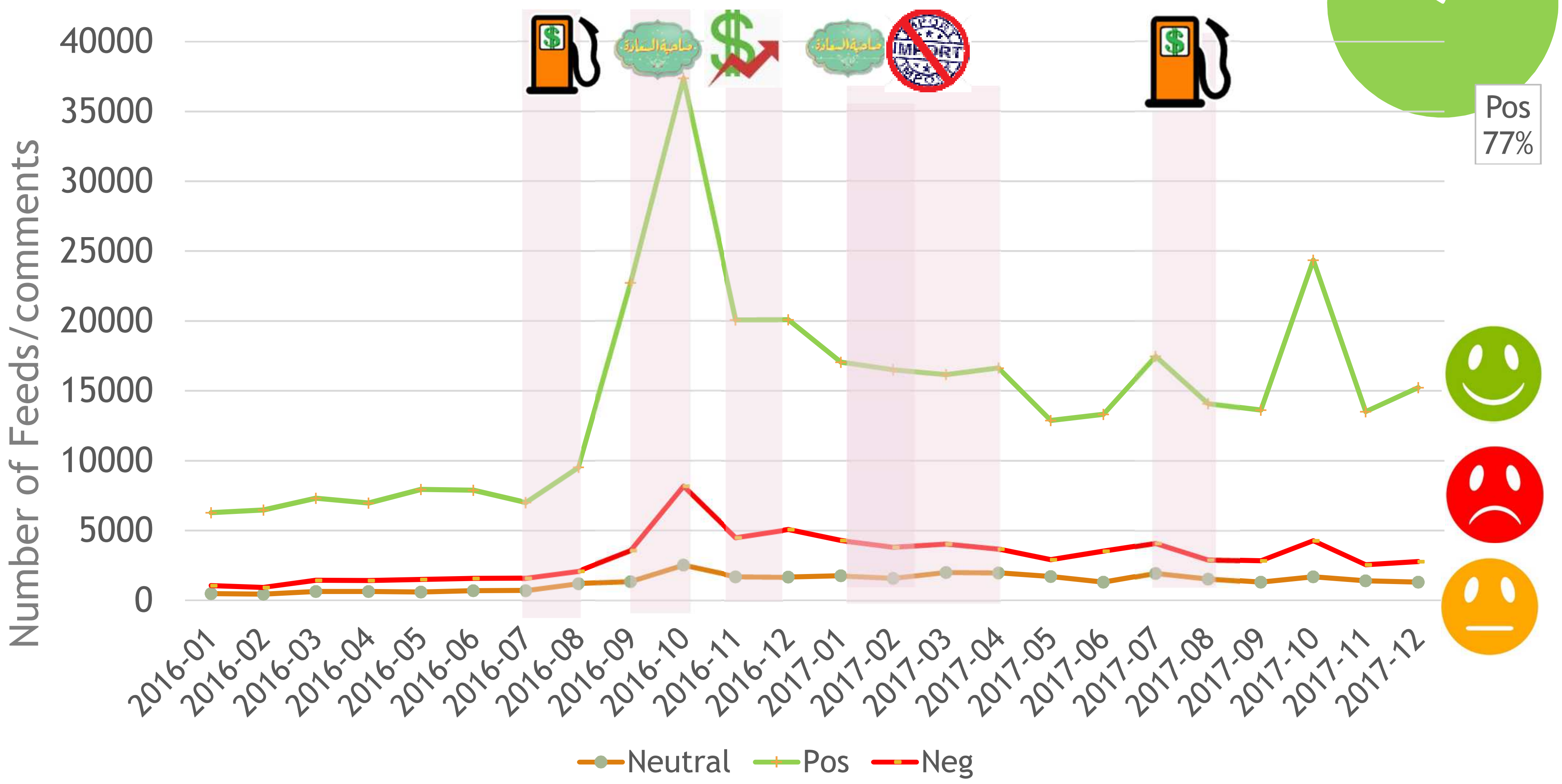
- 80 % of public is talking about price, but not saying an opinion

Price Sentiments

- People are more negative about price than positive
- Negativity is observed more i

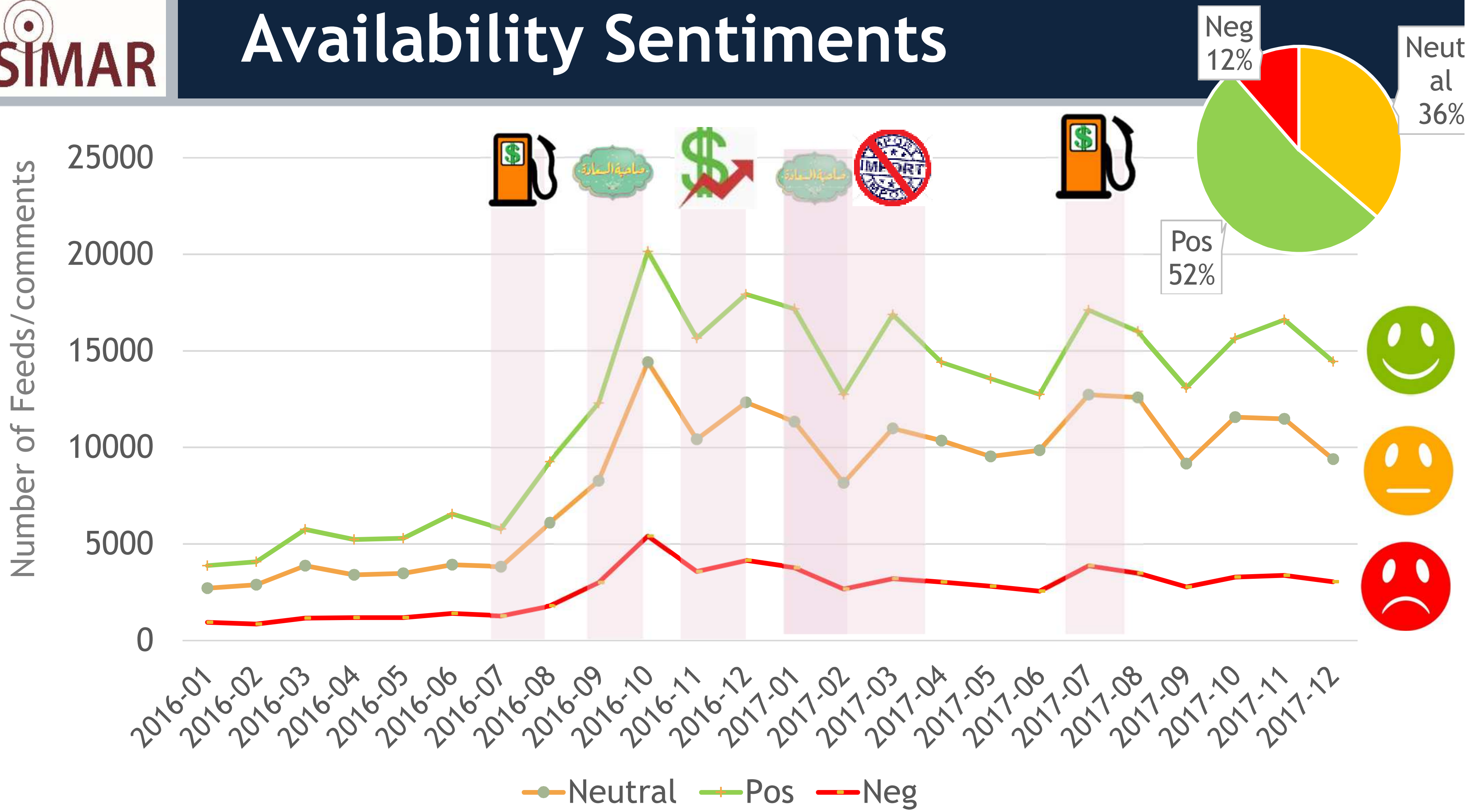
Quality Sentiments





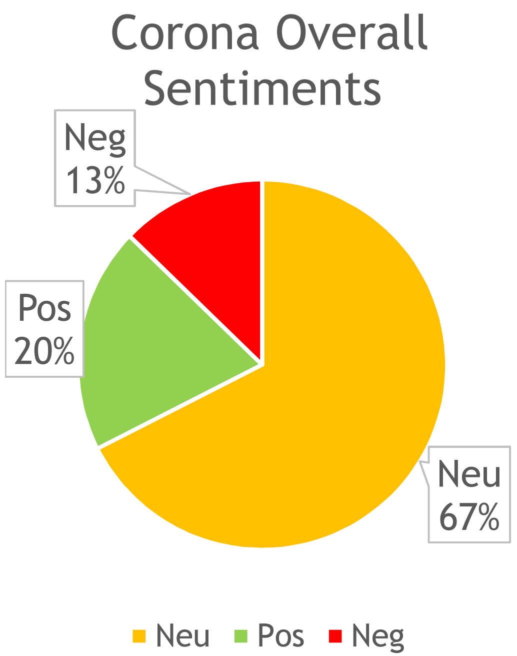
- People tend to express positive sentiments toward Egyptian products 77%
- The TV show boosted the positivity feeling towards quality.

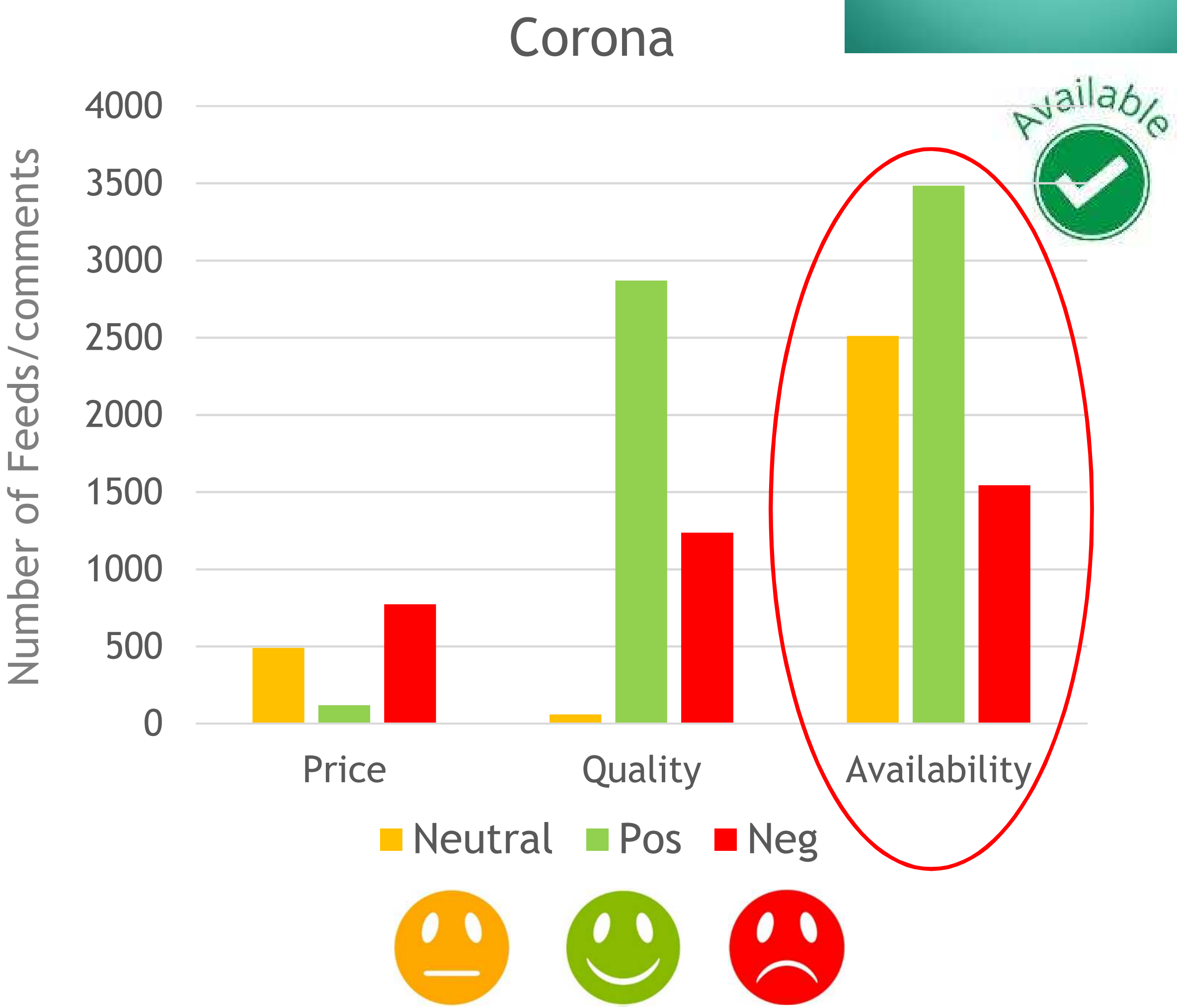
Availability Sentiments



- More than 50% of the people feel positively about availability
- 36% of the public discuss availability but do not express their opinion
- The negativity is only 12% and mostly not affected by events

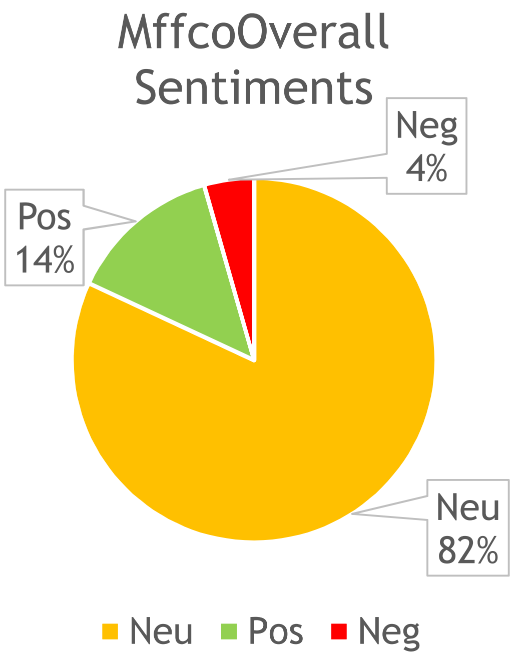
Corona Chocolate

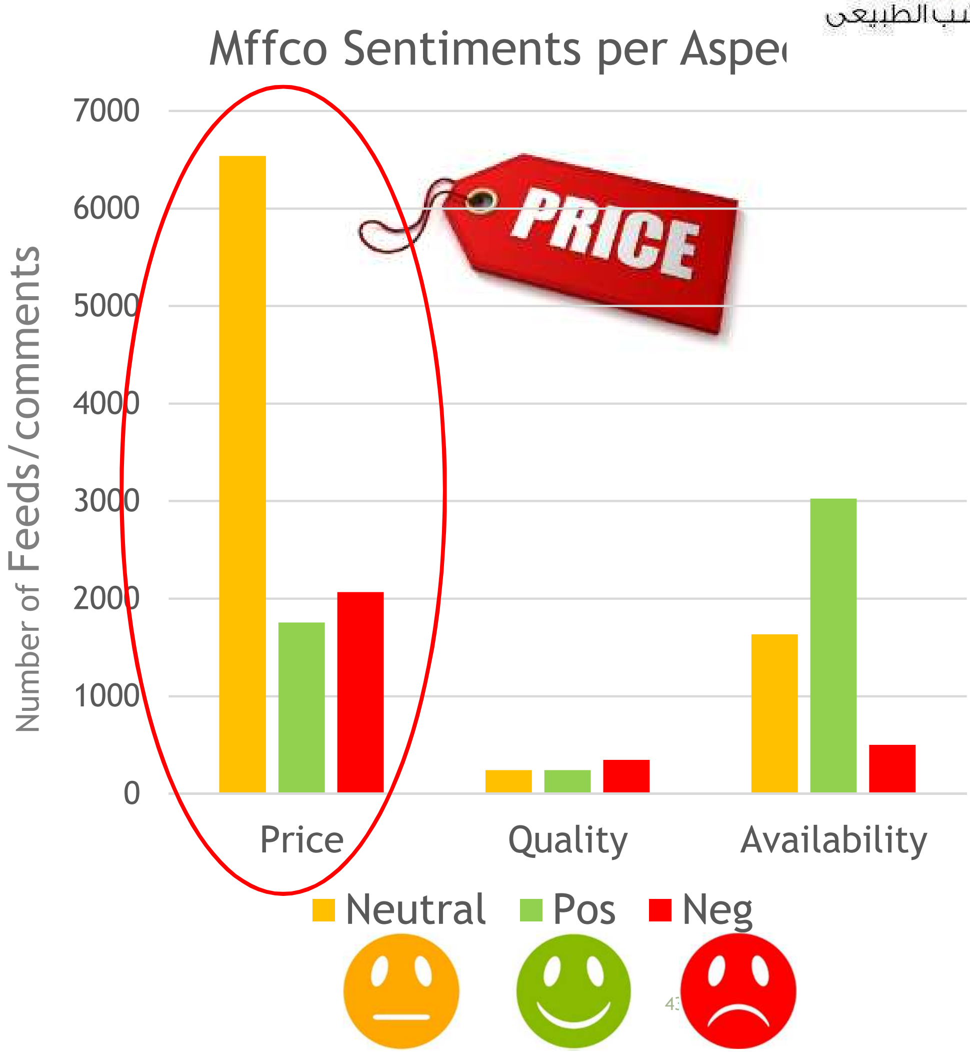




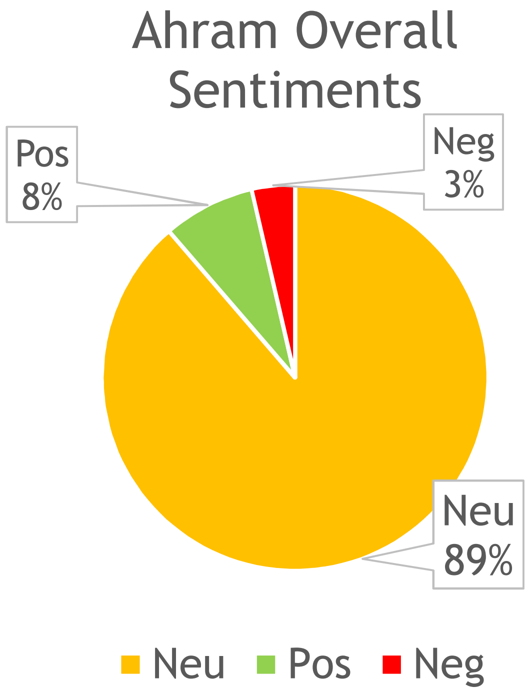
Mffco Helwan Furniture

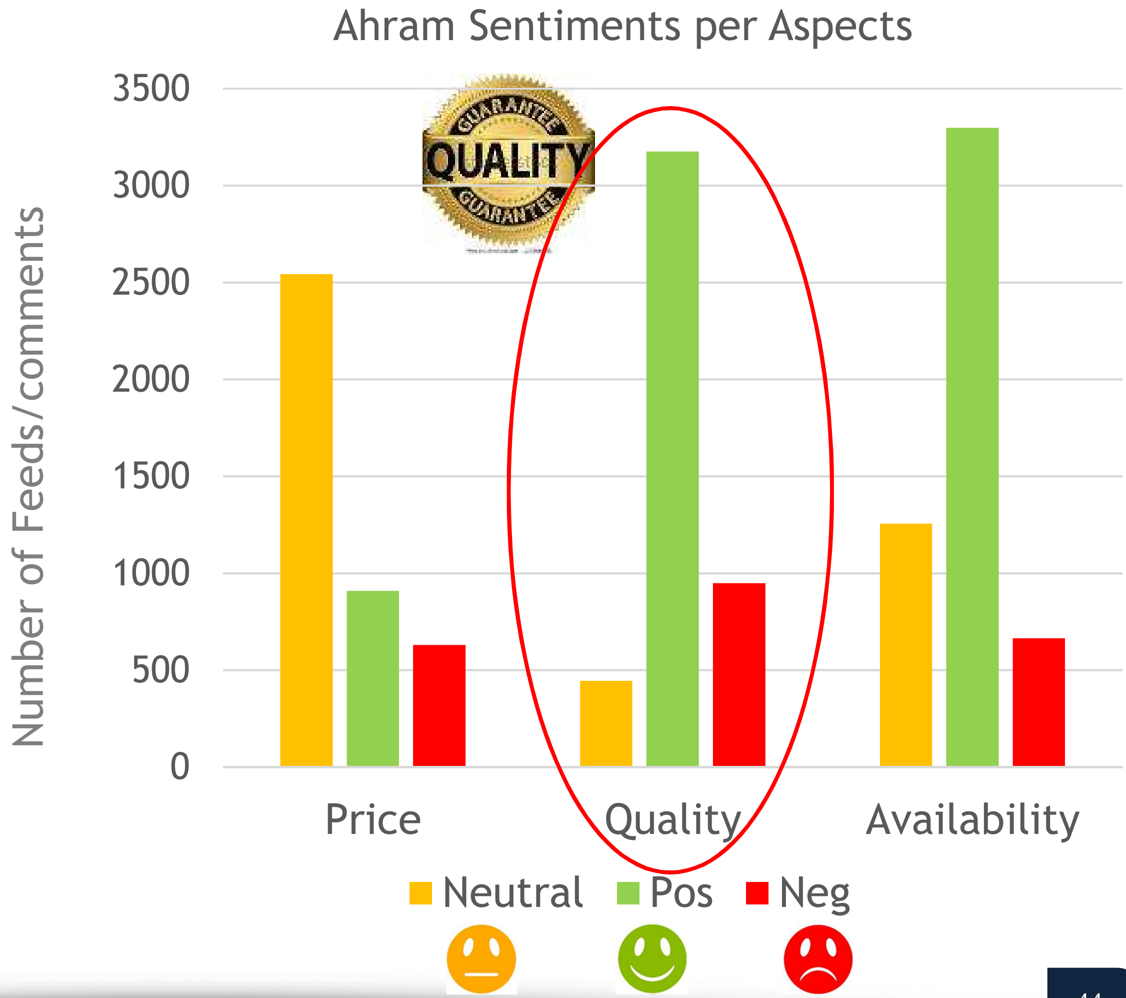
mffco
أثاث من الخشب الطبيعي





Al-Ahram Aluminum





Conclusions and Future Work

- Sentiment analysis in Arabic is still a fertile field for research, especially for social media.
- A platform has been established for sentiment analysis in Arabic social media.
- A case study has been applied analyzing sentiments towards products Made in Egypt.
- The results needs to be more carefully examined to draw insights.
- Future work will include:
  - Further tuning of parameters.
  - Enlarging the corpus and the data to include Twitter, Instagram and others.
  - Apply other case studies of interest to the platform.

45

Thank you

46