Winning the I-COM Datascience Hackathon 2016

L.F. Chiroque, R. Cuevas, J.M. Carrascosa, C. Iordanou

2nd Annual Workshop on Complex Sociotechnical Systems València 8-10 June 2016

valencia 6-10 June 2010

I-COM Global Summit 2016

- A global trade body focused on the use of marketing data & measurement to obtain business advantage.
- A meeting venue (Seville 2016) for around 100 associations in 40 countries representing marketers, agencies and media owners.
- 2nd edition of the I-COM Data Science Hackathon (Madrid)





The Competition

Two categories: Scientist & Master Levels



Personal challenge: academia vs. industry

The UC3M Team

COSTAS

JUANMI RUBÉN

LUISFO

The Challenge

- In 24 hours, the participant teams have to be able to predict trends, provided with a one-year twitter dataset.
 - Use case: 'hair care' trends
- Input Data
 - ~3 M tweets from Jan. to Dec. 2015 containing the keyword 'hair'
 - We have got a 2013 sample ~10 days before
 - JSON objects
 - tweet | id | user info | timestamp | device info | hashtags | FAV count | RT count | gnip/klout info | language | ...
- Prediction: #tweets / hashtag during Q1-2016 [log(#tweets + 1)]
 - Only hashtags with >= 9 tweets during Q4-2015 considered ~11K hashtags

The Challenge

Evaluation

- Quantitative 50%
 - Pearson correlation coefficient
- Qualitative 50%
 - Business value, Story telling and Art & Technology
- Two Rounds
 - Round 1: Two finalists per category
 - Round 2: Presentation during the I-COM Global Summit in Seville
 - Bonus: I-COM Audience Award

The summary

PROGRESS



Considerations

- Evaluation: correlation coefficient on #tweets/hashtag estimation
- Supervised method: using Q4-2015 as ground truth
 - Only hashtags with >= 5 #tweets during Q3-2015 considered
- Random split 70% Training 30% Test
- Models
 - Time Series models
 - Regression models
 - Machine Learning



Preliminary results

PREDICTION OF #TWEETS PER HASHTAG Q4-2015 (METRIC: PEARSON CORRELATION)







PREDICTION BASED ON MONTHLY AGGREGATED INFO

Chosen model

- Generalized Linear Model
- PCA —> 12 components & Factor Analysis
- #TW_Q4 ~ #TW_Q3 + #TW_Q2 + #TW_Q1 + #TW_sep + #TW_ago
- Test set correlation coefficient ~0.77



Hackathon Results

- Correlation Coefficient: 0.81
 - 1st result in our category and 2nd overall



The Bussiness Application

HAIR RELATED For Unilever **KEYWORDS OF** Dictionary INTEREST TO UNILEVER **TOP TREND** HASHTAGS Q1-2016 GreysAnatomy 80sHairBands HairTips HairStyle 9 hairdressers Sbeauty 문 영 freewebcams NaturesBountyBeauty Size LoveYourCurls freewebcam Health g Barber NaturesBountyBeauty HealthyHair CARE I HairExtensions MakeUp HairGoals glamour greyeyes Repost Repost Fashion Hairstyles Fashion Hairstyles Model and Control of the second and the WomenHairStyle ponytail HairStylist product treatment HairColor CurlyHair loss HAIRBeautyTipseyebrows of haircare BEAUTY redcare lashes HairCut HairStylesLeedsNaturalHair LongHair hairdresser GrayHair hairstylist productshairdressing OhNoHarrysHair Care TheMadness Halloween GetWeird love a haircare RT kanzashi handmade DomesticShortHair barbersBadHairDay HairCare barber, Hairstylist HairAccessories XFactor Win MTVSTARS FreebieFriday VideoMTV2015 GOPDebate SEMANTIC TOP TREND HAIR SIMILARITY VSFashionShow TheWalkingDead beauty

RELATED HASHTAGS Q1-2016

Hair

The Bussiness Application

Engagement

- measure share of voice and competitor hashtags
- amplify key trending contents with social sharing and real-time bidding
- create content, polls, and paid tweets around trending hashtags

Sharing

- amplify key trends and key influencer activity
- focus participation in the most relevant conversations with greatest impact potential

The Bussiness Application

ENGAGEMENT



SHARING



I-COM Gala Awards Dinner || Seville • 2016

i-com i-com

om i-com

n-com

1-com

I-COM

- Audience Award

i-ci

GRACIAS

Winning the I-COM Datascience Hackathon 2016

L.F. Chiroque, R. Cuevas, J.M. Carrascosa, C. Iordanou

2nd Annual Workshop on Complex Sociotechnical Systems València 8-10 June 2016

valencia 6-10 June 2010