GetSentiment

Using rules and machine learning on customer feedback

Mark Rogers

About us

GetSentiment

Mark Rogers

CEO

Specialisation in text analytics, image-recognition, AI

Ex-BBC Online, Amazon, now Oracle

Viktor Pekar, PhD

CTO

Research Fellow, Business School & School of Computer Science, University of Birmingham

NLP/Text Analytics, Machine Learning, Software Development

Our Clients

intuit

Insight360 by TruValue Labs

NYU

GAIAMTV.com

THE GOOD SCHOOLS GUIDE

calibrum

Waiting for the tech to deliver its promise

- ~~Sentiment analysis~~
- ~~Benchmarking~~
- ~~Insights from customer data~~
- ~~Quantitative metrics~~

- “My transcripts are a mess”
- “I don’t know what’s in there”
- “I haven’t got time to label it”

- If only all the transcripts were neatly categorised
- I could benchmark by subject-matter
- I could compare one brand to another

Example client data

	A	B	C	D	E	F	G	H	I	J
1	review_id	review_text								
2	2025622	I had these knobs on some cabinets I installed years ago. So when the old cabinets were paint								
3	6259496	has nice color for how long it last, price is a little high but not bad								
4	7195685	WELL, IT'S A NICE LOOKING FAN...BUT, THAT'S ABOUT ALL. I WASTED A COUPLE OF HOURS INST								
5	9320458	Unit arrived in good shape.Worked fine with various loads up to a continuous 1500 watt resist								
6	9603120	This two story scaffolding worked better than I expected. It is exactly what I needed to compl								
7	9810589	Purchased this product in July it has 13 hrs on it. The mower started making a loud noise unde								
8	10515279	I have owned my Dyson DC25 for 4 months now and absolutely love it. It maneuvers around o								
9	11210374	The motor blew with 21 hours on it took it in to get repaired they said no it was due to lack of								
10	11569140	I purchased this blower to replace a gasoline blower and it has lived up to its reputation. I rec								
11	11685148	I was concerned before purchasing by many of the negative reviews on related Altura fans, bu								
12	11891070	I love, love, love this fan! So glad it is available in 52. The delivery time was exceptional. Insta								
13	2078633	Bought this beautiful, HUGE 68" Ceiling fan...That is pretty much it... it is a nice piece of art, ho								
14	2230548	just purchased this item from home depot to use when i go camping i have a upright camper w								
15	6284718	I bought this mower this weekend. I was debating between this model and the John Deere 11								
16	6893034	Bought this unit to run a portable a/c 8000 BTU in my truck it worked perfectly No problems w								
17	10036830	I've owned this generator for 18 months now. I did use it for 5 days after an ice storm a year ag								
18	10858577	No reflection on [redacted] as their service is great! However this generator is junk. Starte								
19	10918312	I used it on the edges of my sidewalk and it works great.								
20	10918339	We have had this system for about 5 years and could not be more pleased. Our lab puppy was								

Solution 1: Generic, off-the-shelf

- Terms can convey positive sentiment in car reviews, but can be negative or neutral in other contexts:
- “expensive”: “The car has an expensive look.”
- “cheap”: “The car is cheap to run.”
- “rough”: “The car is suitable for rough terrain.”
- “high torque”, “low emissions”, “low mpg”, …

Solution 2: Supervised ML

- Supervised machine learning can be great if there is hand-labeled data for that particular domain
- But what if there isn't?
- And hand-labelled data is hard to come by?

Solution 3: Unsupervised ML

ID	Topics
0	payment late made due make appli month past day miss
1	receiv letter sent send mail state request email document notic
2	loan student borrow privat navient repay lender default defer forbear
3	car vehicl financ dealership dealer ticket book drive trade truck
4	servic custom repres manag transfer spoke depart supervisor cancel speak
5	check cash advanc clear return wrote flagstar seiz present payabl
6	file complaint cfpb case complain respond clerk district bsi compliant
7	home hous equiti repair inspect buy door damag sell valu
8	call phone number person stop time answer messag harass work
9	credit report remov bureau show neg correct inform agenc transunion
10	co program school class colleg signer enrol student region educ
11	account close open activ statu inaccuraci limit creat past restrict
12	purchas express store buy product order return replac item refund
13	consum financi protect practic decept abus institut mislead repossess factor
14	bill insur medic polici hospit cover coverag flood doctor health
15	disput item investig equifax verifi delet inform inaccur valid creditor

Late payments (?)
Communication (?)
Vehicles (?)
Communication (?)
Communication (?)
Healthcare (?)

Example topic models from a corpus of Customer Financial Protection Bureau complaints (Bastani et al 2018)

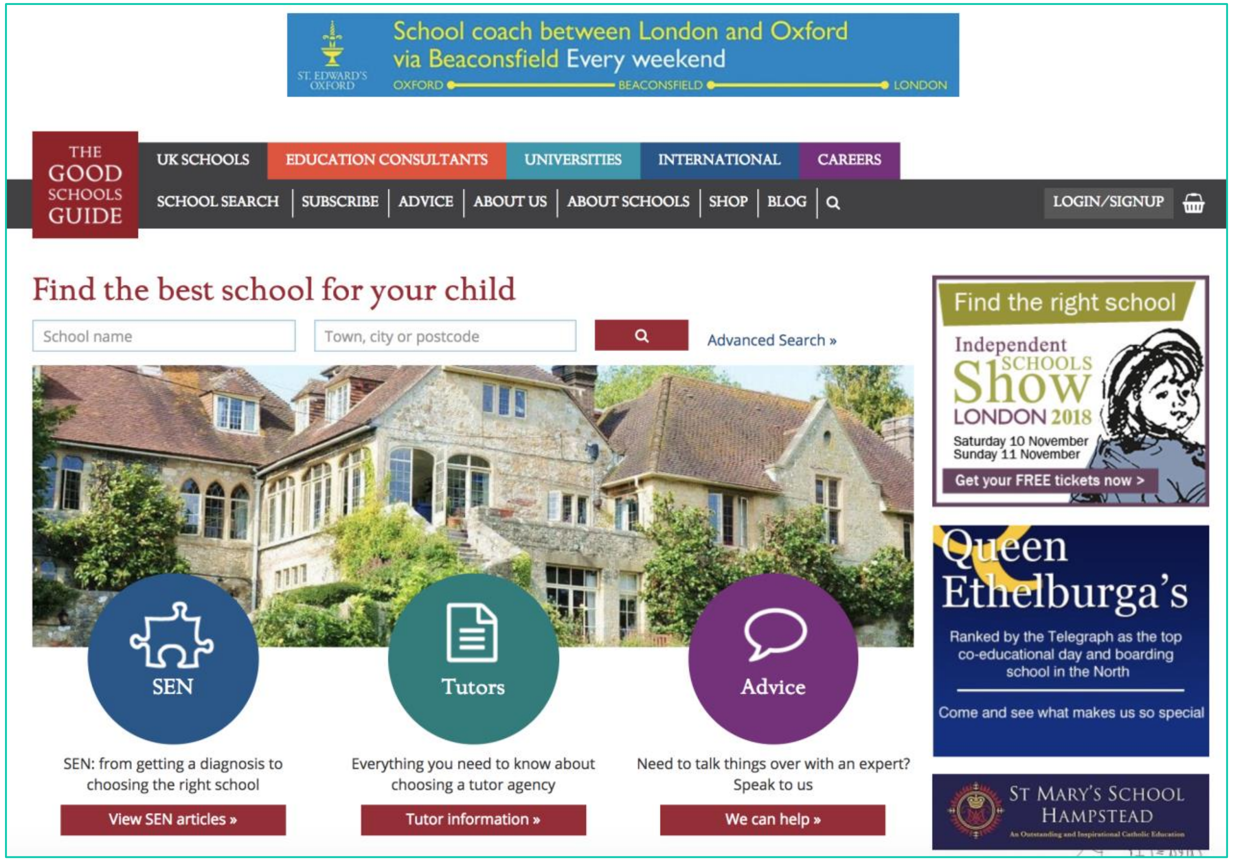
Possible solutions: Unsupervised ML

- Unsupervised discovery of topics only works if very large amounts of data are available
- Automatically induced categories seldom correspond to business requirements

Good Schools Guide

GetSentiment

- Can we scrape and categorise listings for schools?



Our Approach

- **Step 1**: Automatic keyword extraction to find prominent multi-word terms

Our **Approach**

Step 2: Semi-supervised construction of domain lexicons from unlabelled text:

- The client suggests categories and gives examples (“seeds”)
- We expand those seed samples based on the data
- The client reviews and post-edits the domain lexicon

Lexicon Editor

GetSentiment

REFRESH | FILTER BY CATEGORY ... | Search term ... | SEARCH

	DEFINE CATEGORY ... (ADD NEW)
sirloin	FOOD
cordial	SERVICE
condiment	FOOD
harrassment	SERVICE
cake	FOOD
pool table	AMBIENCE
called back	SERVICE
threaten	SERVICE

Our Approach

Step 3: Text analysis

- We look for relations between terms and sentiment phrases
- Our domain lexicons allow us to recognize categories and sentiment in the text
- We score sentiment by sentence and aggregate for the document as a whole

Aspect-Based Sentiment Analysis

- An research field in NLP that emerged around 2010
- Aims to:
  - Recognize issues talked about (“terms”),
  - Classify them into categories (“aspects”),
  - and recognize the author’s sentiment with respect to the terms and categories of terms (“sentiment”)

Aspect Term **Polarity**

GetSentiment

Task: Given a sentence, identify terms related to a relevant aspect and identify sentiment associated with the terms:

For example:

"I loved their **fajitas**" → {fajitas: *positive*}

"I hated their **fajitas**, but their **salads** were great" → {fajitas: *negative*, salads: *positive*}

"The **fajitas** are their first plate" → {fajitas: *neutral*}

"The **fajitas** were great to taste, but not to see" → {fajitas: *conflict*}

- Relevant terms (**fajitas** and **salads**) represent the Food aspect
- Different sentiment can be present in the same sentence

Aspect-Based Sentiment Analysis

Aggregated sentiment scores

A restaurant review

tripadvisor.co.il: Palio's

2 days 6 hrs 12 mins ago

I have had dinner at Palio's multiple times in the past (over 5 years) and thought our meals tonight were great. Mussels app was really good, dipping bread in it was fantastic. Two of us had filets, trout special was awesome and one scallop for our entrees. Desert -creme brûlée and cannoli were a great finish. I will be back

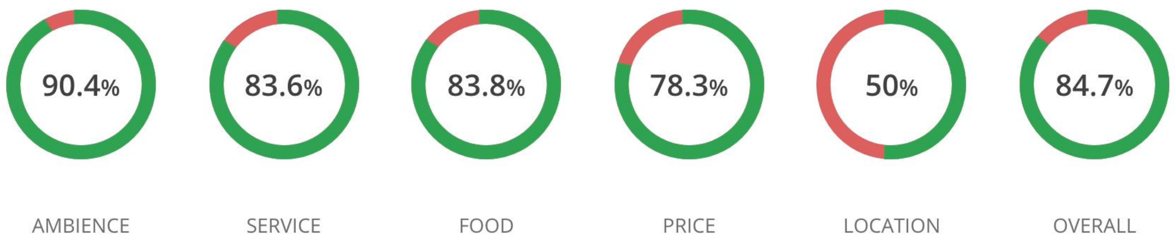
Multiple aspects identified

Each aspect is scored for sentiment

GetSentiment

FEATURES DEMO PRICING GET STARTED SERVICES CONTACT

Restaurant Reviews



We just got back and my toddler is lactose intolerant disney do offer rice dream or tofutti dairy free icecreams though not sure if its available at all places and the pineapple dole whips are also dairy free the table service resturants we ate at were quite helpful at the buffets the chef walked us through the options .at chef mickeys & cape may they cooked him dairyfree options though the portions were huge esp for a toddler. at others they provided an allergy menu to chose from. depends on ho....

Ambience Service Food Price Location Overall

Not being fans of buffet dining, we ate at the a la cartes every night: - TOSCANA - had dinner here our first night (I'd made reservations via e-mail prior to leaving home). We both ordered the filet steak which was very disappointing. The people at the next table over had ordered the lobster and burger; they weren't overly impressed with either. I can't remember what we had for dessert, so it obviously wasn't anything wonderful. However, I do remember the appetizer which consisted of a few ite....

Ambience Service Food Price Location Overall

The second is that we ate at the deck 2x during our stay and were also disappointed by service there. Anyways not the end of the world. I too had read the reviews here on the forum and everyones excitement was the reason we booked. Maybe my above post doesn't quite capture how bad the expierience was but I can assure

Evaluation

- Participated at SemEval-2014, Task 4 (Aspect-Based Sentiment Analysis)
- SemEval competitions: academic teams develop solutions to NLP tasks, which are evaluated on the same reference dataset
- Around 30 teams competed at the ABSA task at SemEval
- Two evaluation datasets: Restaurant and Laptop reviews
- Pontiki M. (2014). Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 27–35.

Results

GetSentiment

Laptops		Restaurants	
Team	Acc.	Team	Acc.
DCU	70.48	DCU	80.95
NRC-Can.	70.48	NRC-Can.	80.15†
SZTE-NLP	66.97	UWB	77.68*
UBham	66.66	XRCE	77.68
UWB	66.66*	SZTE-NLP	75.22
lsis_lif	64.52	UNITOR	74.95*
USF	64.52	UBham	74.6
SNAP	64.06	USF	73.19
UNITOR	62.99	UNITOR	72.48
...			
Blinov	52.29*	V3	59.78
iTac	51.83*	SINAI	58.73
DLIREC	36.54	DLIREC	42.32*
DLIREC	36.54*	DLIREC	41.71
IITP	66.97	IITP	67.37
Baseline	51.37	Baseline	64.28
Majority	52.14	Majority	64.19

Came 4th out of ~30 teams on the Laptops, 7th on the Restaurants dataset

Baseline: 51% (laptops) and 64% (Restaurants)

Achieved accuracy of 66% (Laptops, best team 70%), 74% (Restaurants, best team: 80%)

CONTACT US

GetSentiment

MARK ROGERS
Phone: +44 (0) 7866 369181
Email: mark@getsentiment.io

www.getsentiment.io

@getsentiment_io

GetSentiment

THANKS FOR WATCHING