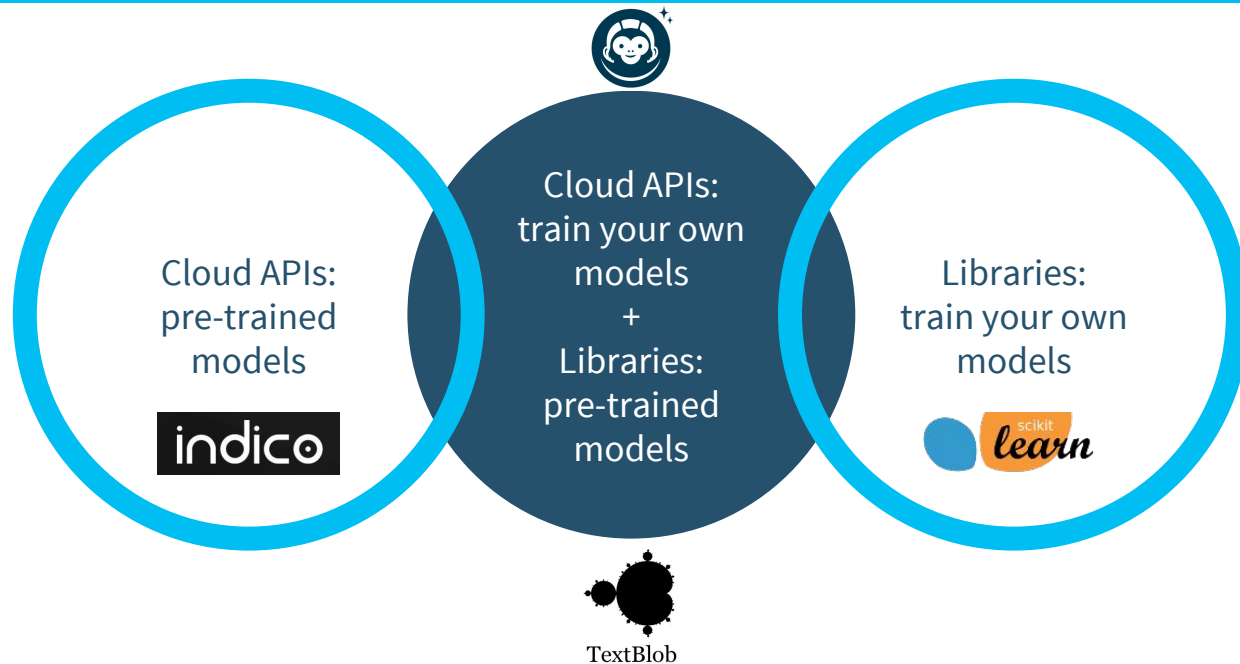


# Between Custom and Off-the-shelf NLP

**Yves Peirsman**



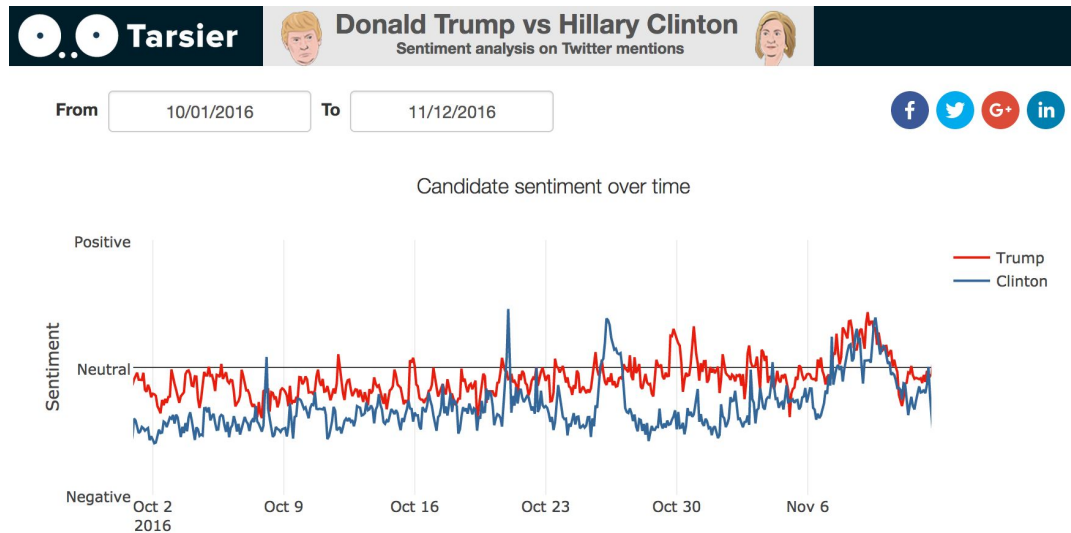


*“The process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc., is positive, negative, or neutral.”*



## SENTIMENT ANALYSIS: APPLICATIONS

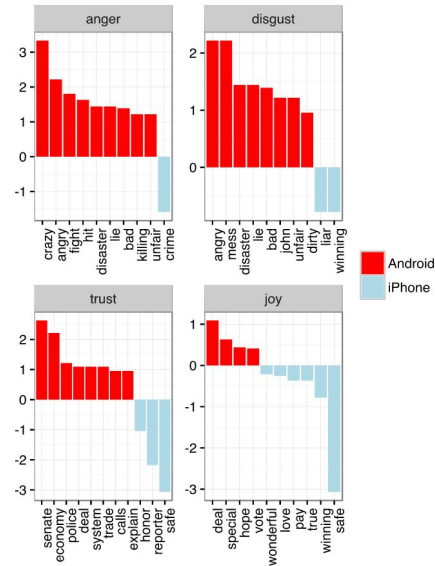
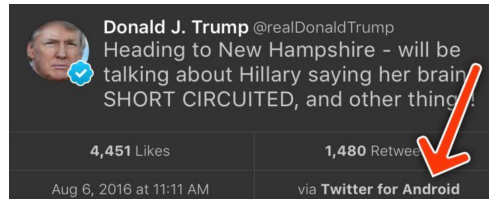
4



<http://tarsier.monkeylearn.com>

## SENTIMENT ANALYSIS: APPLICATIONS

5



<http://varianceexplained.org/r/trump-tweets/>

# SENTIMENT ANALYSIS: APPLICATIONS

6

StockFluence

Search

About StockFluence

## FINANCIAL SENTIMENT ANALYSIS


StockFluence.com provides financial sentiment analysis for investors to discover, react and respond to market opinions. We monitor (social) media channels and analyze the overall sentiment with our algorithms. Based on the sentiment, we make predictions with an accuracy level of 70%.

▼ -6.93% 127	▲ +115.71% 60	▲ +2.42% 67
Buffalo Wild Wings Inc.	B3's Restaurants, Inc.	Camblum Learning Group, Inc.

APPLE INC. >

145


-1.5%



ERICSSON >

125


30.94%



ING GROEP NV >

91

-8.89%



<http://www.stockfluence.com/>





# NLP LANDSCAPE

8

**Sentigem** BETA  
Sentiment Analysis API sentiment.vivekn.com



**HPE Haven**  
OnDemand



**IBM**

**AYLIEN**

text-processing.com

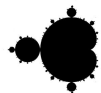


Google Cloud Platform

Cloud APIs:  
pre-trained  
models

**indico**

Cloud APIs:  
train your own  
models  
+  
Libraries:  
pre-trained  
models



TextBlob



Libraries:  
train your own  
models





## DATA SETS

9

Domain	Source	Categories	Baseline
Movie reviews	rottentomatoes.com	positive, negative	50.0%
Baby products	Amazon	positive (****, *****), negative (*, **)	81.5%
Android apps	Amazon	positive (****, *****), negative (*, **)	72.7%
Android apps	Amazon, Wikipedia	positive (****, *****), negative (*, **), neutral	51.7%
Hotels, restaurants	Yelp	positive (****, *****), negative (*, **)	70.8%

### Positive

If you're looking for something scary, this is the first great horror film of the spooky season.

### Neutral

Avernum is a series of demoware role-playing video games by Jeff Vogel of Spiderweb Software available for Macintosh and Windows-based computers. Several are available for iPad and Android tablet.

### Negative

It's Starbucks only with bad customer service. Baristas with attitude that don't know their own product. If I'm paying \$7.00 for a coffee at least drop the 'tude.

# 1.

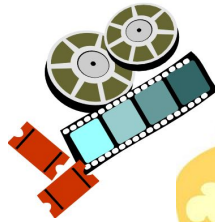
## OFF-THE-SHELF NLP

Is it OK to be lazy?

### Variables

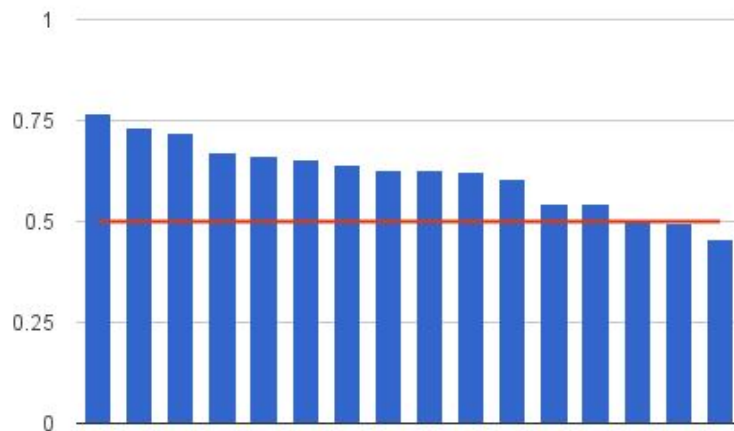


### Data



## OFF-THE-SHELF MODELS: MOVIE REVIEWS

13



Indico

76.8%

IBM AlchemyAPI

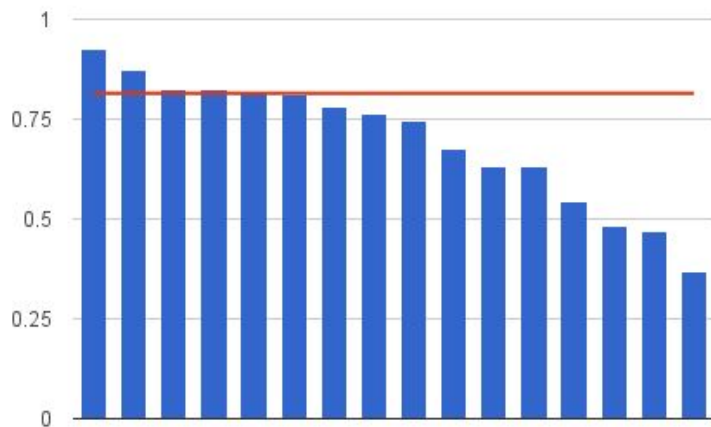
73.4%

Stanford CoreNLP

71.9%

## OFF-THE-SHELF MODELS: BABY PRODUCTS

14



Indico

92.6%

MonkeyLearn  
(Product)

87.5%

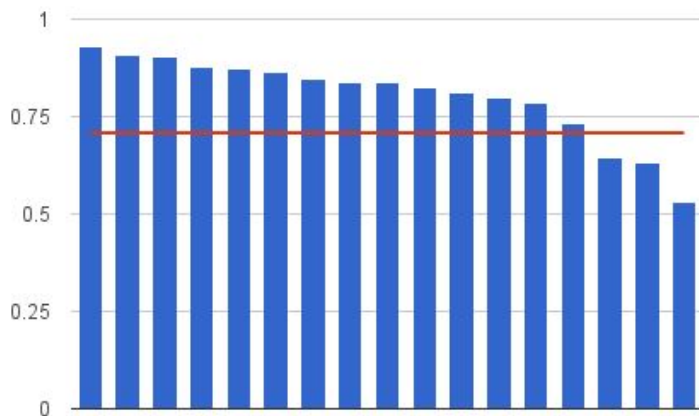
TextBlob Pattern

82.5%



## OFF-THE-SHELF MODELS: YELP REVIEWS

15



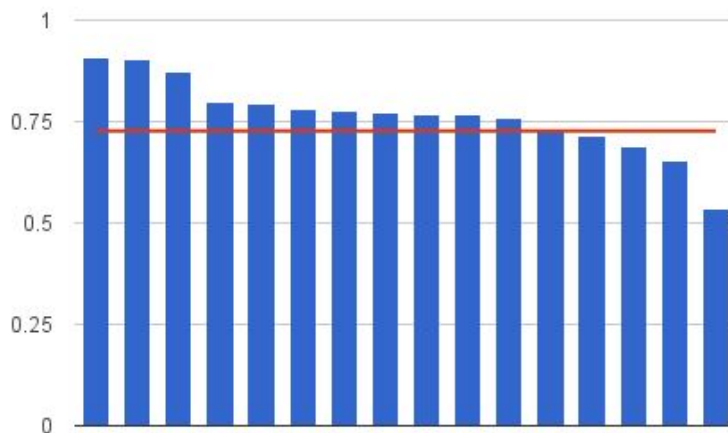
Indico 92.9%

Google 91.0%

IBM AlchemyAPI 90.4%

## OFF-THE-SHELF MODELS: ANDROID APPS (2-WAY)

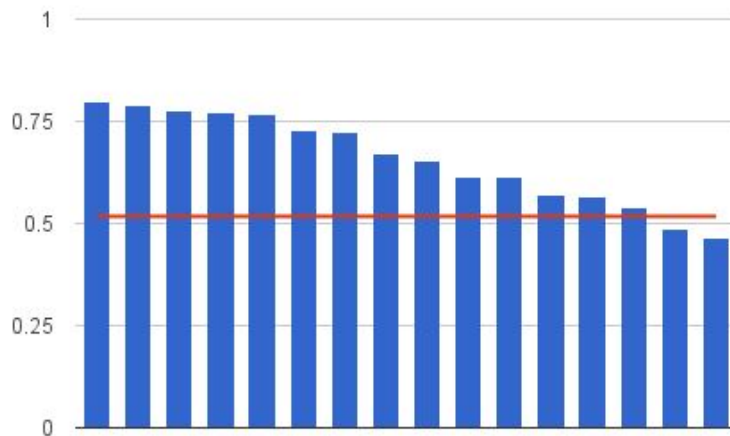
16



Indico	90.6%
Google	90.5%
MonkeyLearn (Product)	87.1%

## OFF-THE-SHELF MODELS: ANDROID APPS (3-WAY)

17



Indico

80.0%

HavenOnDemand

79.1%

Google

77.8%



There is enormous variation *between* and *within* off-the-shelf solutions.



High quality is possible, but not guaranteed.

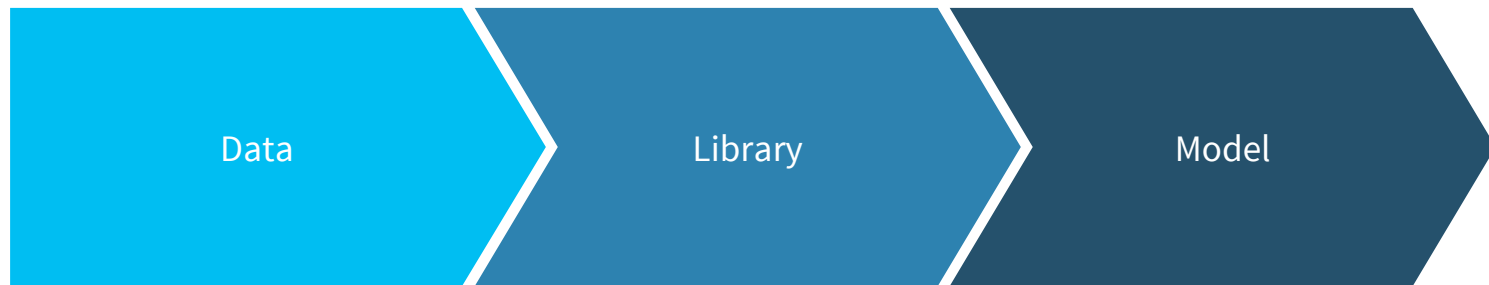


Comparing available solutions on your data is crucial.

# 2.

## DIY MODELS

Are you better off building your own custom models?



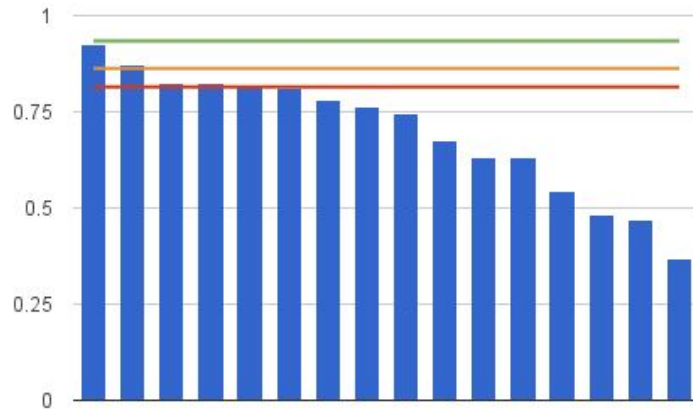




```
text_clf = Pipeline([('vectorizer', CountVectorizer()),
                     ('transformer', TfidfTransformer()),
                     ('classifier', MultinomialNB())])
text_clf.fit(data["train"]["texts"], data["train"]["labels"])
predicted_labels = text_clf.predict(data["test"]["texts"])
print(np.mean(predicted_labels == data["test"]["labels"]))
```

## DIY MODELS: BABY PRODUCTS

22



DIY SVM

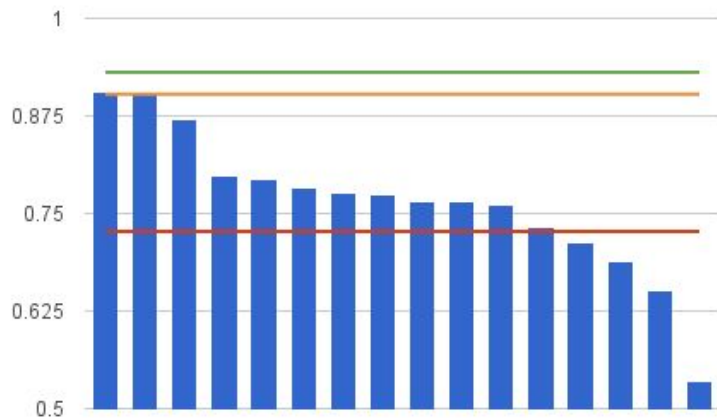
93.4%

Best off-the-shelf

92.6%

DIY Naive Bayes

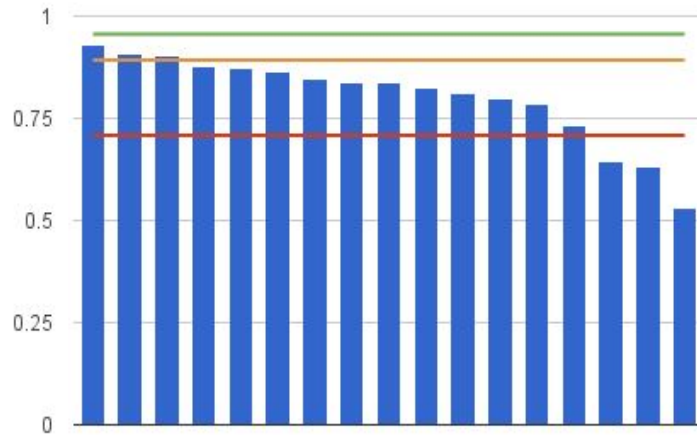
86.3%



DIY SVM	93.1%
Best off-the-shelf	90.6%
DIY Naive Bayes	90.3%

## DIY MODELS: YELP REVIEWS

24



DIY SVM

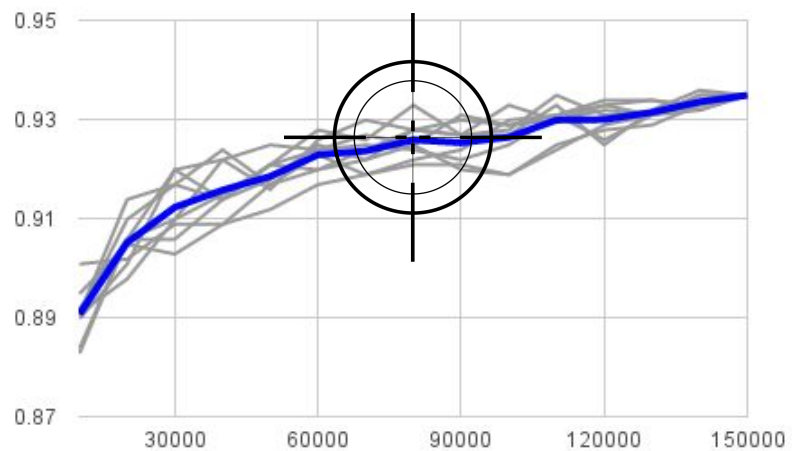
95.6%

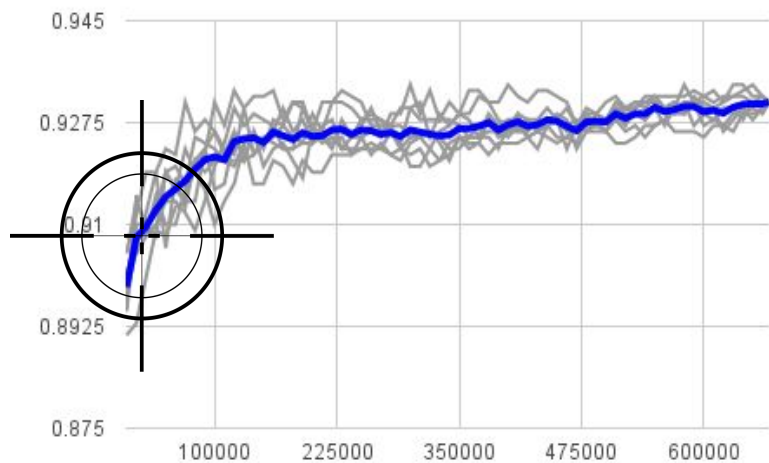
Best off-the-shelf

92.9%

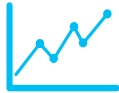
DIY Naive Bayes

89.2%









You need sufficient relevant data to build a good model.



DIY models built with sufficient data will typically outperform off-the-shelf solutions.



DIY models may or may not be worth the effort.

# 3.

## Conclusions

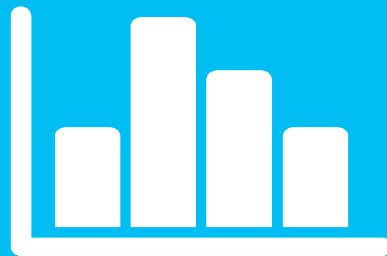
### Off-the-shelf

- ✓ no data
- ✓ little effort
- ✓ good quality possible, but
- ✗ not guaranteed
- ✗ no control

### DIY

- ✗ lots of data
- ✗ more effort
- ✓ superior quality
- ✓ full control

ALL MODELS ARE  
**WRONG**  
BUT SOME ARE  
**USEFUL**



- *George Box*



## Belgium NLP Meetup

nsors Photos Pages Discussions More

### Belgium NLP Kickoff Meetup

Edit Cancel Feature Copy Ticket Export

Tell a friend Share

**Wednesday, December 7, 2016**  
8:00 PM

**Faculty of Arts, KULeuven, Mgr. Sencie Instituut, Room 02.18**  
Erasmusplein 2, Leuven ([edit map](#))

- (Big) Business with NLP: dream or reality?

*Geert Adriaens, Dekimo Experts*

- The future of translation technology: shaken or stirred ?

*Vincent Vandeghinste & Tom Vanallemeersch, Centre for Computational Linguistics, KULeuven*

- Natural Language Understanding in E-Commerce

*Dirk Stevens, Enquire.ai*

# Any questions?

You can find me at:

- » @yvespeirsman
- » yves@nlp.town

