Machine Learning Best Practices with Alfresco & Activiti

Jason Jolley

Goal

Empower everyone to use Machine Learning in Content and Process Services.

Want to build your own Cognitive Process?

1) Download an Activiti Enterprise Trial
2) Sign up for a free IBM BlueMix Account [No Credit Card Required]
3) Configure an Activiti Endpoint to the BlueMix Watson service of your choosing.
4) Enjoy!

You will be able to create your very own Cognitive Process POC without writing any code!

Back to this near the end of the presentation

Agenda

Agenda

1) Machine Learning Overview
2) Common Tools & Services
3) Patterns Applied to Alfresco and Activiti

Machine Learning?

What is Machine Learning?

- ✔ An Overused Buzzword
- ✔ A Transformative Technology
- ✔ A Confusing Mess
- ✔ All of the above

What is Machine Learning?

“A field of study that gives computers the ability to learn without being explicitly programmed.”

-Dr. Arthur Samuel

Think: “The algorithms to accomplish a task.”

What about “Cognitive Computing”?

For most of us – Machine Learning and Cognitive Computing are analogous.

Cognitive computing leverages machine learning and other AI to emulate Human Cognition.

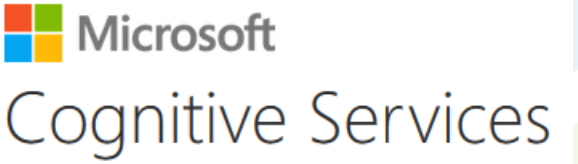
Most of the major vendors now brand themselves with ‘Cognitive Services’

How does it relate to Content & Process Services?

Content & Process Services deal with unstructured content! **Sometimes, A LOT of unstructured content!**

Today’s cognitive processes **LOVE** unstructured content.

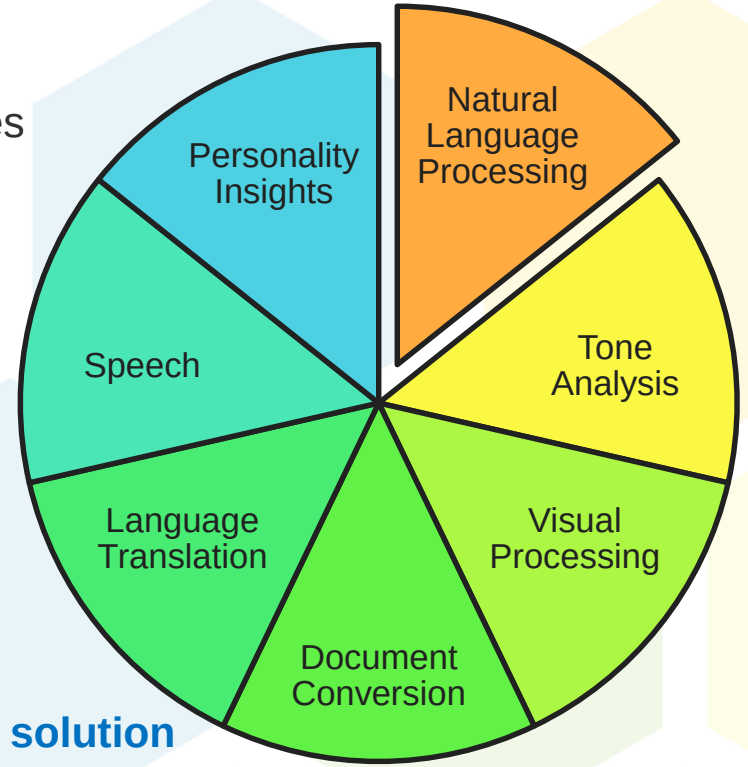
“Ready to Go” Cognitive Services API



Common APIs

There are commonalities in various Cognitive Services

Each API is independent.



Best Practice: **Mix and Match the APIs for your the solution**

Language Translator

Translate text from one language to another for specific domains.

IBM

Natural Language Classifier

Natural Language Classifier performs natural language

IBM

Natural Language Understanding

Analyze text to extract meta-data from content such as concepts,

IBM

Personality Insights

The Watson Personality Insights derives insights from transactio

IBM

Retrieve and Rank

Add machine learning enhanced search capabilities to your

IBM

Speech to Text

Low-latency, streaming transcription

IBM

Text to Speech

Synthesizes natural-sounding speech from text.

IBM

Tone Analyzer

Tone Analyzer uses linguistic analysis to detect three types of

IBM

Tradeoff Analytics

Helps make better choices under multiple conflicting goals.

IBM Deprecated

Visual Recognition

Find meaning in visual content! Analyze images for scenes,

IBM

Watson Services (Part of IBM BlueMix)

Vision	Speech	Language	Knowledge	Search
Computer Vision	Bing Speech	Bing Spell Check	Academic	Bing Autosuggest
Content Moderator	Custom Speech Service	Language Understanding	Entity Linking	Bing Image Search
Emotion	Speaker Recognition	Linguistic Analysis	Knowledge Exploration	Bing News Search
Face		Text Analytics	QnA Maker	Bing Video Search
Video		Translator	Recommendations	Bing Web Search
		WebLM		

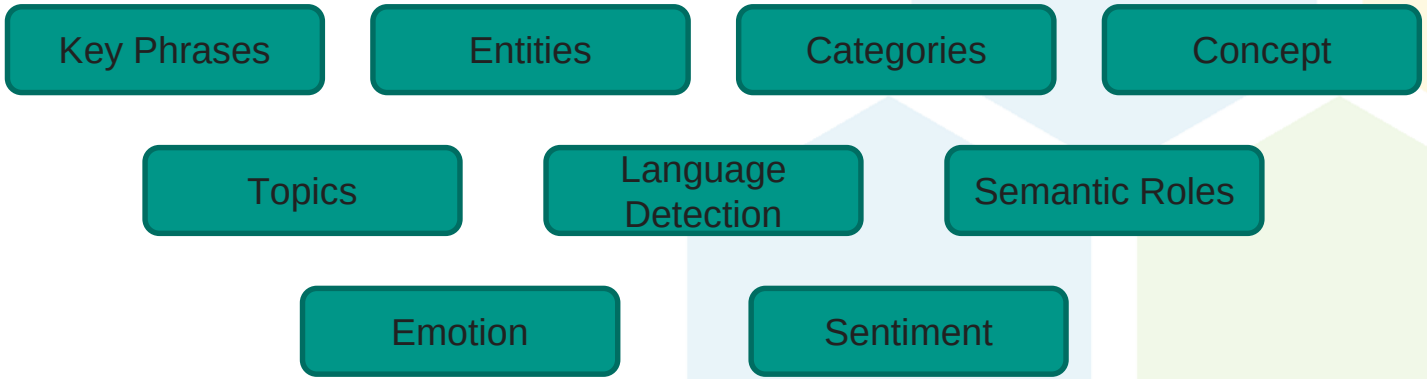
Microsoft Cognitive Services

Natural Language Processing (NLP)


NLP

Natural Language Processing

Analyzes **Unstructured Text** to extract items like:



NLP Example: http://www.beecon.buzz

Sentiment

Positive 0.88

Entities

Alfresco Company 0.94

Emotion

Joy 0.47 Anger 0.26 Disgust 0.03 Sadness 0.08 Fear 0.01

Categories

/ art and entertainment / shows and events / conference 0.60

/ technology and computing / software / net conferencing 0.24

/ business and industrial / business operations / business plans 0.23

Note: All Scores are between 0.00 and 1.00

NLP Example: http://www.beecon.buzz

Keywords	
new Alfresco developers	0.94
Alfresco expertise	0.78
Alfresco community	0.78
collaboration itch	0.56
perfect session	0.55
experimental addon	0.54
BeeCon Conference	0.54

Keywords	
newest release	0.53
essential factors	0.51
wide array	0.51
potential activities	0.50
community project	0.50
new partners	0.49
attendee	0.37

NLP Use Cases with Content & Process Services

Too Many Uses Cases!

- Bulk Document Classification & Re-organization
- Inbound Document Parsing
- Automatic Categorization & Tagging of Content
- Automatic Folder creation and document relocation
- Process Decisions
- Issue Escalation
- And many more….

NLP Demo

Visual Recognition

Visual Recognition

Visual Recognition uses deep learning algorithms to analyze images t
visual content.

- Object Classification
- Face Detection

Classes	Score
Chihuahua dog	0.97
small dog	0.98
dog	0.98
domestic animal	0.98
animal	0.98
pale yellow color	0.70
tan color	0.68

Type Hierarchy

/domestic animal/small dog/Chihuahua dog

Visual Recognition

Visual Recognition uses deep learning algorithms to analyze images that
visual content.

- Object Classification
- Face Detection

Faces	Score
age 18 - 24	0.40
female	0.92
age 18 - 24	0.51
female	0.99

Classes	Score
bridesmaid	0.77
woman	0.77
female	0.77
person	0.88
sister	0.67
big sister	0.53
claret red color	0.83
alizarine red color	0.66

Type Hierarchy

/person/female/woman/bridesmaid

/person/sister

/person/big sister

Beecon Hackathon

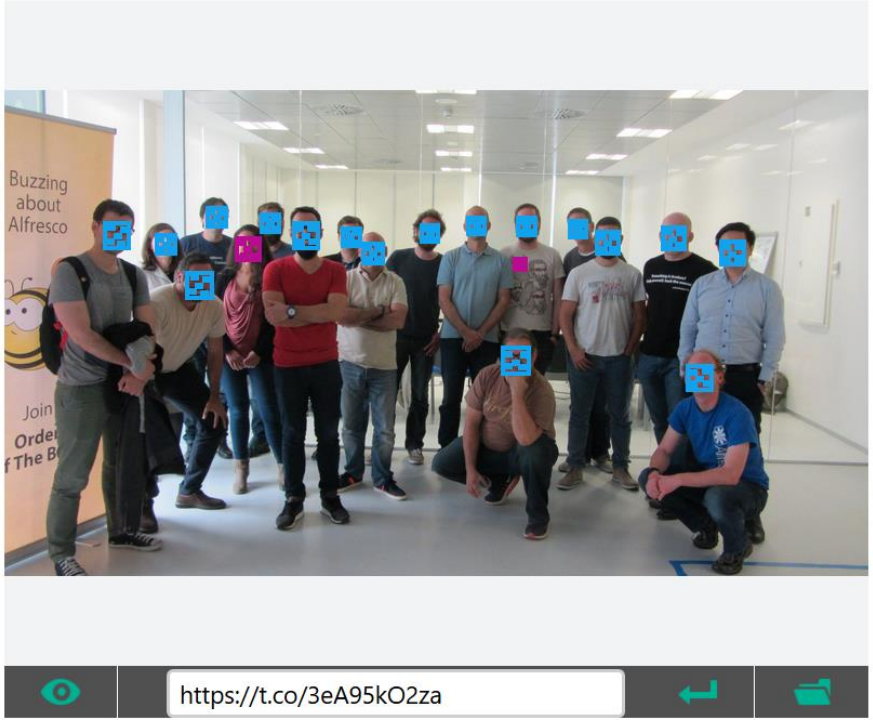
Beecon Hackathon

https://t.co/3eA95kO2za


```
Detection Result:
19 faces detected

JSON:
[
  {
    "faceId": "4675e0ec-bf79-483b-b396-95ce5b10c1e7",
    "faceRectangle": {
      "width": 73,
      "height": 73,
      "left": 1175,
      "top": 604
    },
    "faceLandmarks": {
      "pupilLeft": {
        "x": 1193.4,
        "y": 624.5
      },
      "pupilRight": {
        "x": 1231.9,
        "y": 626.4
      },
      "noseTip": {
```

Beecon Hackathon



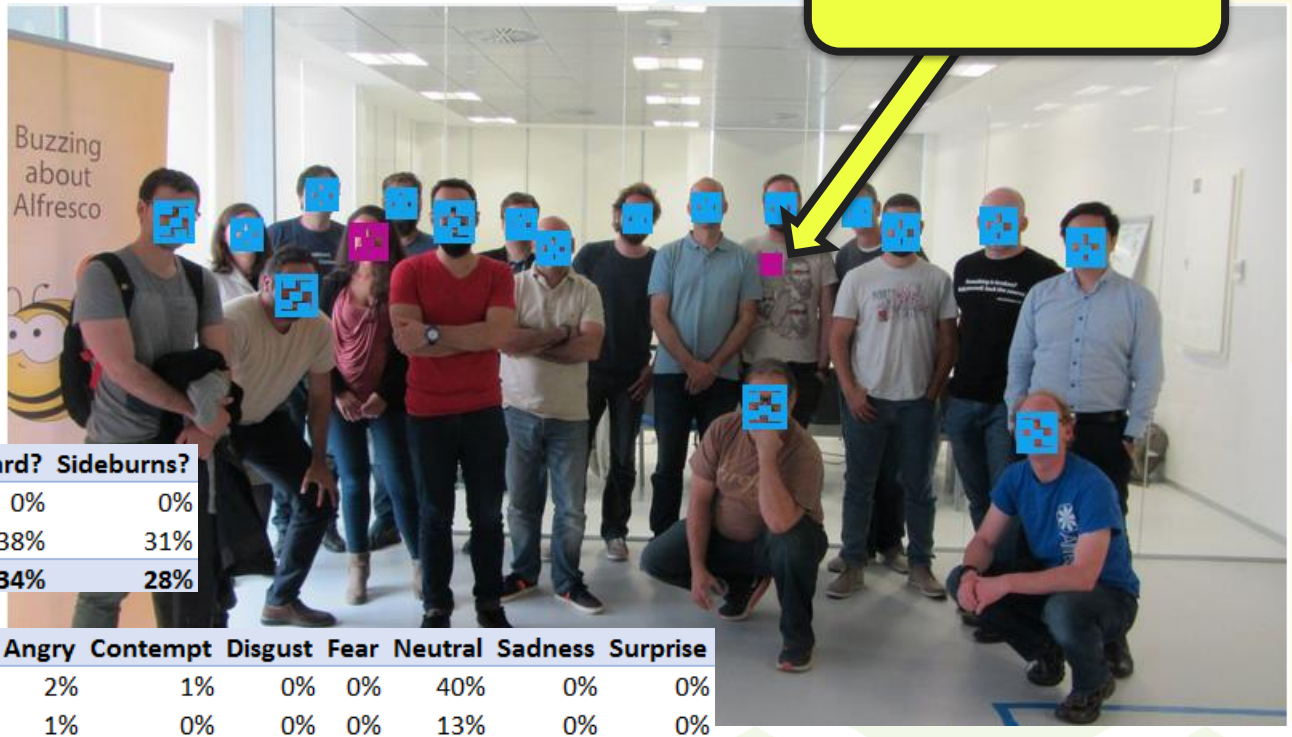

```
    },
    "faceAttributes": {
      "age": 47.2,
      "gender": "male",
      "headPose": {
        "roll": -11.1,
        "yaw": 3.2,
        "pitch": 0
      },
      "smile": 1,
      "facialHair": {
        "moustache": 0.1,
        "beard": 0,
        "sideburns": 0
      },
      "glasses": "NoGlasses",
      "emotion": {
        "anger": 0,
        "contempt": 0,
        "disgust": 0,
        "fear": 0,
        "happiness": 1,
        "neutral": 0,
```

Some Hackathon Stats

Gender	Count	Smiling?
female		
NoGlasses	2	56%
female Total	**2**	**56%**
male		
NoGlasses	9	89%
ReadingGlasses	7	81%
Sunglasses	1	87%
male Total	**17**	**86%**
Grand Total	**19**	**83%**

Gender	Moustache?	Beard?	Sideburns?
female	0%	0%	0%
male	35%	38%	31%
Grand Total	**32%**	**34%**	**28%**

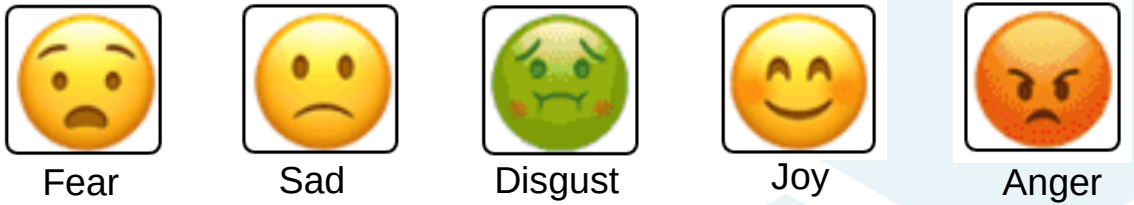
Gender	Smiling?	Happy	Angry	Contempt	Disgust	Fear	Neutral	Sadness	Surprise
female	56%	56%	2%	1%	0%	0%	40%	0%	0%
male	86%	86%	1%	0%	0%	0%	13%	0%	0%
Grand Total	**83%**	**83%**	**1%**	**0%**	**0%**	**0%**	**16%**	**0%**	**0%**



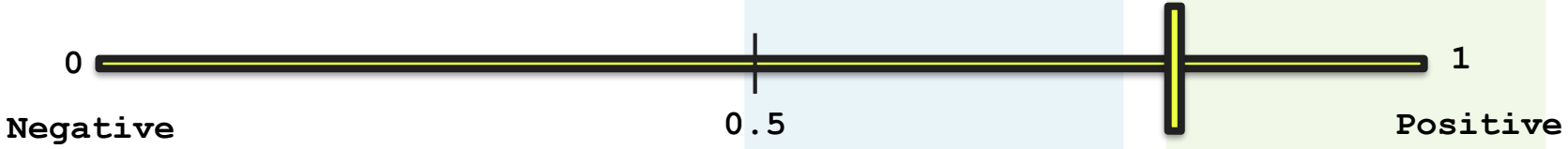
Tone Analysis

Tone Analysis

Detect and interpret emotions, social tendencies, and language style cues found in text.



Sentiment Analysis



Use Case – JIRA & Customer Service

We use JIRA to help service our clients.

Customer Response Time is VERY Important.

BUT – Even More Important is **Customer Satisfaction**!

Response Time is easy to measure.

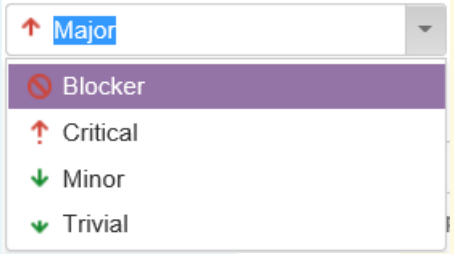
How do you automatically measure Customer Satisfaction?

Use Case – JIRA & Customer Service

Customers can choose Priority…

Priority != Satisfaction

A Customer could log a Trivial issue, but still be very dis-satisfied.



Use Case – JIRA & Customer Service

Customer Enters JIRA Ticket

Activiti Calls Watson

Watson Analyzes Sentiment & Emotion:
- Summary
- Description
- Comments

Activiti notifies Stakeholders if thresholds are exceeded.
Stores Data for reporting.

Use Case – JIRA & Customer Service

Some tips:

- Start Small – a Test Project, then a small project
- Configurable Thresholds
- Save historical data for trends. Consider using trends for alerts instead of individual tickets.

Configuration - Demo

THANK YOU!!!

Jason Jolley

jjolley@microstrat.com
@jasonjolley