HELPING COMPUTERS UNDERSTAND US: MACHINE LEARNING FOR PERSONALITY TRAIT RECOGNITION

Scott Nowson

- Dublin Machine Learning Meetup
- March 27, 2017

THE FUTURE BECOMES REAL

accenture
THE DOCK

COMING UP, TONIGHT

The Academics

- Why personality is popular; and what is it?
- A brief (incomplete) history of personality and language
  - From development to classification

The Applications

- Techniques for recognising personality traits
  - From shallow to deep
- Business applications

INSIGHTS FROM PERSONALITY

Personality Matters: How one company doubled its ROI by customizing ads based on personality

6 Ways Introverts Vacation Differently

This personality trait predicts your tendency to lie and cheat

What Your Personality Says About Your Career Path

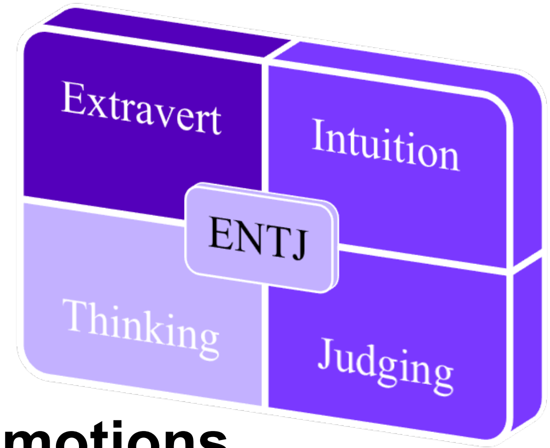
Psychologists found the personality traits that make people fat

How Facebook 'likes' could be used to make personality-based hiring decisions

12 Ways 'Type A' People Love Differently

Will you have to take a personality test to get a loan? The ten questions that you could be asked

PERSONALITY … & LANGUAGE



Each of us is unique; but not in random ways

- Determines our characteristic patterns of **behaviour**, **thoughts**, and **emotions**

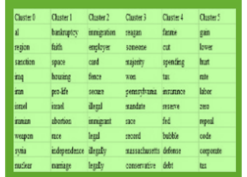
O	• Openness
C	• Conscientiousness
E	• Extraversion
A	• Agreeableness
N	• Neuroticism

1880s
Galton
The Lexical Hypothesis

Aspects of personality will become part of language

1930s
Allport

Psycholexical studies of traits, 1000s of terms identified



1981
Goldberg
Big Five

Neuroticism
Extraversion
Open to Experience
Agreeable
Conscientious

1999
Pennebaker & King
LIWC

Psychologically validated lexica; semantic labels; relationship with personality

2005/6
Argamon
Nowson
Mairesse

First published work on computational personality recognition

2007
TSWG

First US govt sponsored program on personality classification

2013
WCPR

First Workshop on Computational Personality Recognition (shared task)

LINGUISTICS = FEATURE ENGINEERING

Domains studied

Essays, email, blogs, social media, speech, vlogs

Linguistic Level	<u>EXTRAVERTS</u>	introverts
Lexical choice	Social words; references to others; positive emotions; lexically repetitive	References to self; negations; negative emotions; lexically diverse
Syntax	Pronouns, adverbs, verbs; More informal	Nouns, modifiers, prepositions
Construction	multiple topics, more aggregation	single topics
Fluency	Higher speech rate; more errors	More hesitant; longer pauses
Behaviour	Greater desire to initiate; shorter bursts	Less words overall

AUTHOR PROFILING CHALLENGE

First shared evaluation task, 2015

- Five personality traits (O, C, E, A, N)
- Four languages (En, Es, It, Nl)
- Anonymised Twitter data

Features vary, techniques less so

- SVMs 2006-2015
- Language representation
  - Surface forms – word, lemma and character n-grams
  - Syntactic features – POS tags and dependency relations
  - Feature curation – punctuation and emoticon use; topic modelling; sentiment; psychological dictionary

Francisco Rangel, Fabio Celli, Paolo Rosso, Martin Potthast, Benno Stein, and Walter Daelemans (2015). Overview of the 3rd Author Profiling Task at PAN 2015. In *Working Notes Papers of the CLEF 2015 Evaluation Labs.*

BUT WAIT!

Where does data come from?

- How do you collect such data
- What are gold standard labels

Ask people to complete personality questionnaires

- High quality data; choice of inventory
- Time consuming; potentially expensive; small scale

Find people who have published their personality

- Memes FTW
  - Nowson & Oberlander 2007, Plank & Hovy, 2015
- Orders of magnitude bigger; less control; self selection bias

DEEP LEARNING FOR PERSONALITY

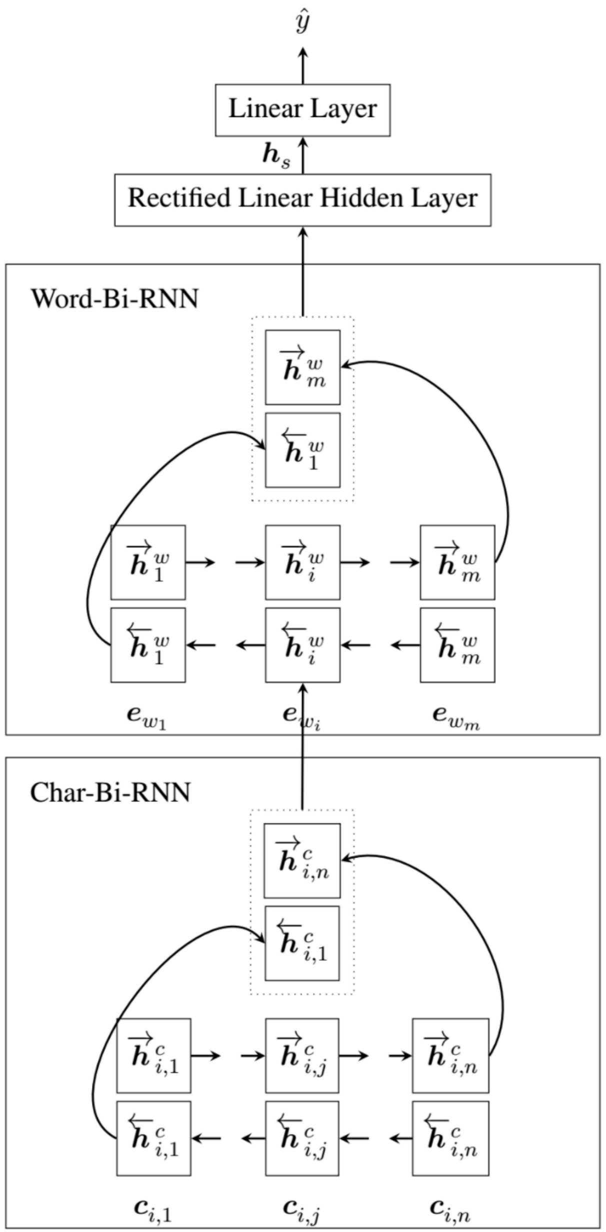
Previous approaches use shallow learning

- Alternative – deep-learning on atomic units, characters
- Combine Bi-directional recurrent neural networks
- Character to Word to Sentence for Personality Traits – `C2W2S4PT`

Results outperform baseline

- PAN challenge results; SVM
- Multi-task models – leverage known trait relationships
  – Offered minimal improvement

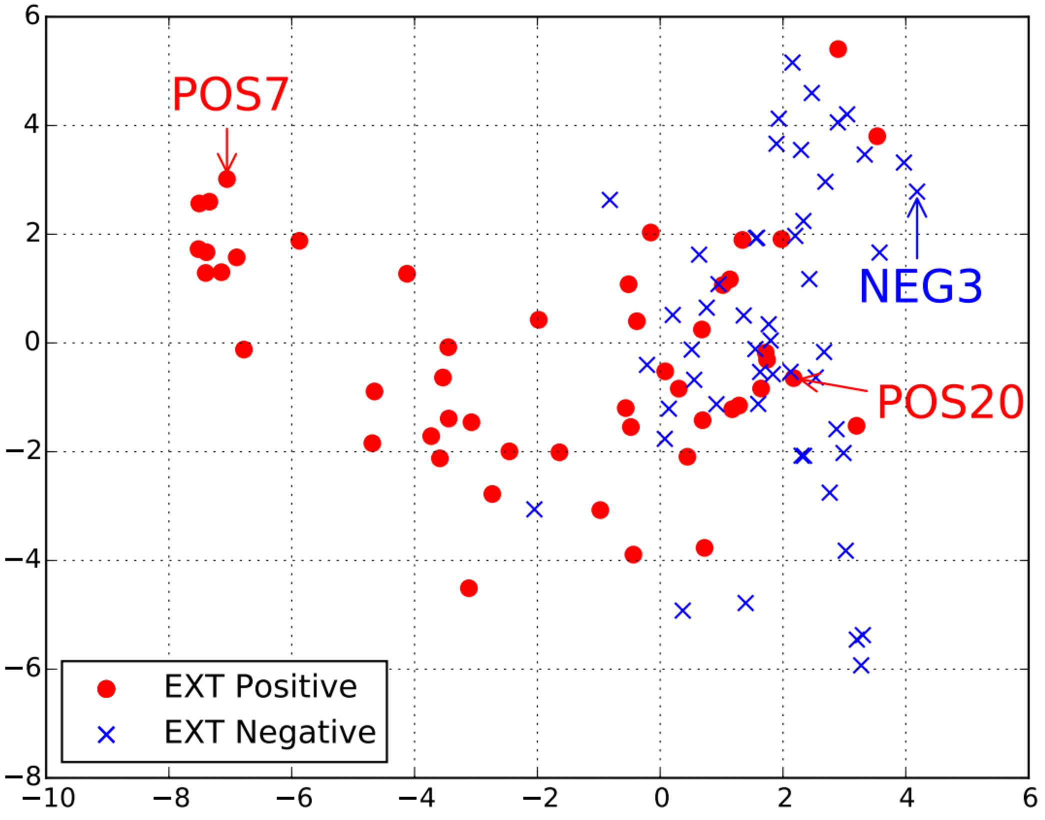
Fei Liu, Julien Perez and Scott Nowson (2016). A Recurrent and Compositional Model for Personality Trait Recognition from Short Texts. Workshop on Computational Modelling of Peoples Opinions, Personality, and Emotions in Social Media (PEOPLES); colocated with Coling 2016; Osaka, Japan.



DEEP LEARNING INSIGHTS?

Query the model, visualise the data

- Rendered with PCA; t-SNE non-interpretable
- 100 random tweets
  - 50 high extraverts; 50 high introverts
- Reduce to 2D representation



- POS7: *“@username: Feeling like you’re not good enough is probably the worst thing to feel.”*
- NEG3: *“Being good ain’t enough lately.”*
- POS20: *“o.O Lovely.”*

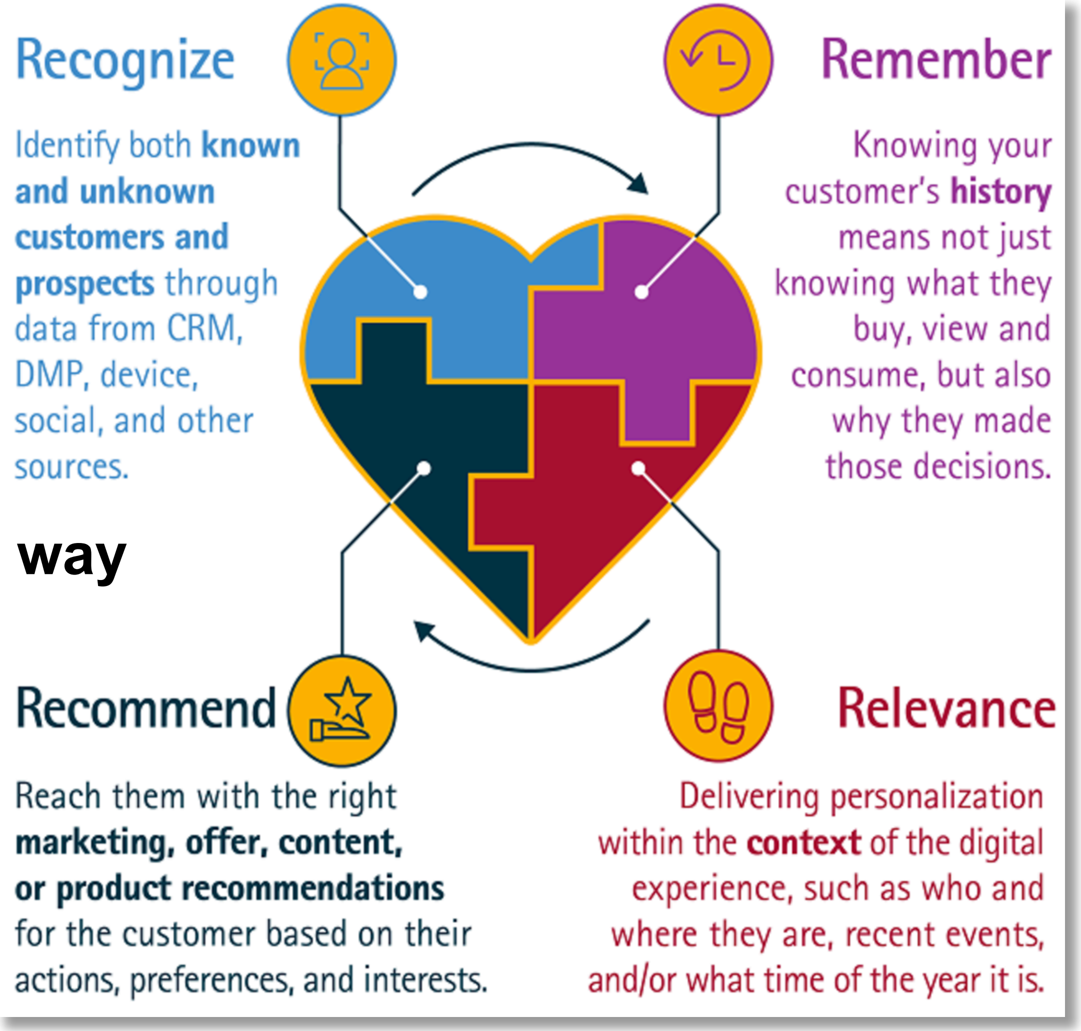
IN THE AGE OF THE CUSTOMER

Enabling personalisation

- **Understand** customer needs
- **Say/Do** the right thing at the right **time**, in the right **way**

Adaptive, automated communication

- Advertising, customer support, advice
- Personalised to improve effectiveness



Delivering on the Promise of Personalization @ Scale. Accenture Interactive. [link]

BROADER CUSTOMER INSIGHTS

Combining insights to feed analytics

- Profile **completion**; filling missing CRM/database values
- Enhanced **segmentation**
  - Predicting: age (21-25), life event (graduation) = **Millennials entering workforce** (interest modelling)

Automatic categorisation of individuals

- The **who** to complement **what** is being said
- Understanding customers:
  - Implicit CSAT measurement
  - Sales effort ROI

THUS … IN CONCLUSION

Personality is core to the human experience

- Machines can recognise it; we can use it

Machines develop better understanding of people

- Who they are; how they feel

Thank you …

… Questions