

Literary quality through the eyes of readers, (gender,) and machines

Andreas van Cranenburgh

University of Groningen

February 14, 2019

Outline

1. What is literature?
2. Reader survey
3. Computational experiments
 - Cliche expressions
 - Deviation from the norm?
 - Literary writing style
 - Metadata: genre and gender
4. Reflection

Section 1

What is literature?

Characterizing Literary Language:

- ▶ What makes a literary novel *literary*?
 - ▶ What do readers think of this?
 - ▶ Can a model predict this?

Characterizing Literary Language:

- ▶ What makes a literary novel *literary*?
 - ▶ What do readers think of this?
 - ▶ Can a model predict this?

Specifically ...

Research Question

are there particular **textual conventions** in literary novels that contribute to readers judging them to be **literary**?

Background

Definition

Literature is the body of work with the most artistic or imaginative fine writing (Britannica, 1911).

Background

Definition

Literature is the body of work with the most artistic or imaginative fine writing (Britannica, 1911).

- ▶ **Demarcation** problem
- ▶ Compare with philosophy of science

Background

Definition

Literature is the body of work with the most artistic or imaginative fine writing (Britannica, 1911).

- ▶ **Demarcation** problem
- ▶ Compare with philosophy of science

Can we give necessary and sufficient conditions for something to be literature?

NB: not the same as success, popularity, quality, &c.

Historical background

Modern, Western notion of literature (and art)
is a late 18th century invention, within Romanticism

Historical background

Modern, Western notion of literature (and art)
is a late 18th century invention, within **Romanticism**

Important features:

Heumakers (2015)



- ▶ autonomous (Kant): not in service of moral or political goals
- ▶ originality
- ▶ not just entertainment
- ▶ not limited by genre, may be about anything
- ▶ social criticism
- ▶ reveals something not revealed by science or philosophy
- ▶ artist is ahead of his time
- ▶ art changes life, pushes boundaries

Who/what determines literary prestige?

Readers make aesthetic value-judgments

Kant: *sensus communis*, subjective but normative

Who/what determines literary prestige?

- Readers make aesthetic value-judgments
 - Kant: *sensus communis*, subjective but normative
- Text: intrinsic (objective?) property of *literariness*;
 - Russian formalism: foregrounding, defamiliarization distinguish literary from standard language

Who/what determines literary prestige?

Readers make aesthetic value-judgments
Kant: *sensus communis*, subjective but normative

Text: intrinsic (objective?) property of *literariness*;
Russian formalism: foregrounding, defamiliarization distinguish literary from standard language
e.g. 'Thou still unravished bride of quietness'
vs 'Don't you know the drivers are on strike?'
(Eagleton 1983, p. 2)

Who/what determines literary prestige?

Readers make aesthetic value-judgments
Kant: *sensus communis*, subjective but normative

Text: intrinsic (objective?) property of *literariness*;
Russian formalism: foregrounding, defamiliarization distinguish literary from standard language
e.g. 'Thou still unravished bride of quietness'
vs 'Don't you know the drivers are on strike?'
(Eagleton 1983, p. 2)

Culture: External social, cultural factors. Critics and publishers determine prestige of novels.
Bourdieu: cultural capital

Meyer (1997): prototype definition

Prototypical Western literary works:

- ▶ are written texts
- ▶ are marked by careful use of language, including features such as
 - ▶ creative metaphors,
 - ▶ well-turned phrases,
 - ▶ elegant syntax,
 - ▶ rhyme, alliteration, meter
- ▶ are in a literary genre (poetry, prose fiction, or drama)
- ▶ are read aesthetically and intended by the author to be read aesthetically
- ▶ contain many weak implicatures

Meyer (1997). What is literature.

Work Paper of the Summer Institute of Linguistics, 41:40–53

Computational humanities approach

Research Question

are there particular **textual conventions** in literary novels that contribute to readers judging them to be **literary**?

1. Acquire empirical data: texts + judgments
2. Operationalize hypothesized characteristics
3. Confirm role of textual conventions with models

Section 2

Reader survey

The Riddle of Literary Quality

Corpus:

- ▶ 401 recent Dutch novels (translated & original)
- ▶ Published 2007–2012
- ▶ Most bought & borrowed books

The Riddle of Literary Quality

Corpus:

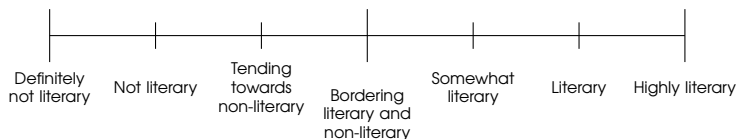
- ▶ 401 recent Dutch novels (translated & original)
- ▶ Published 2007–2012
- ▶ Most bought & borrowed books

Motivation:

- 401 bestsellers
 - ▶ Avoid bias of researcher or experts
 - ▶ Representative, not a sample of convenience
 - ▶ Ensure that enough judgments can be collected
- recent novels
 - ▶ Ignore diachronic dimension
 - ▶ Avoid settled opinions about classics, canon

Reader survey

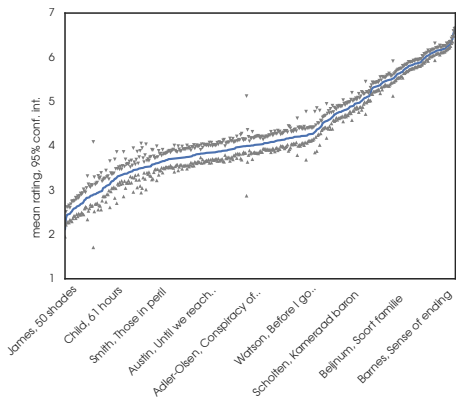
Respondents select books they have read,
and rate them ...



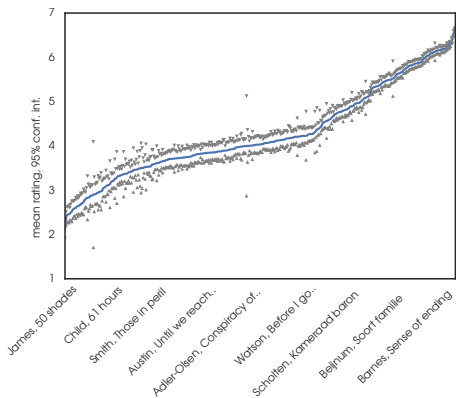
Considerations:

- ▶ Influence of text is not assumed
- ▶ Therefore, rating is only on the basis of author and title
- ▶ Survey is open to anyone, aim is to survey general reading public

Survey ratings: 401 novels; N=14k



Survey ratings: 401 novels; N=14k



Constraints:

- ▶ ≥ 50 ratings
- ▶ ≥ 2000 sent.

369 novels remain

91 % novels
conf. int. < 0.5

Survey responses

Some motivations by respondents to the question:
“Why did you give this book this literary rating?”

- ▶ The writing style
- ▶ Great, suspenseful and surprising book. Writing style not that surprising.
- ▶ I did not like the writing style
- ▶ The book has a lot of depth and multiple layers
- ▶ It is suspenseful, the storyline is perfect, but in a literary novel I expect a deeper layer
- ▶ Shallow story, one-dimensional characters, no deeper layers

Quiz: Guess the Literary Novel

You will see two fragments,

- ▶ one from a literary novel,
- ▶ one from a non-literary novel.

Which is which?

Fragment 1

Na een dag of tien in Italië hebben Depressie en Eenzaamheid me gevonden. Op een avond loop ik na een heerlijke dag op school door de Villa Borghese, terwijl de zon in een waas van goud boven de Sint-Pietersbasiliek ondergaat. Ik voel me gelukkig bij deze romantische aanblik, ook al ben ik helemaal alleen, terwijl voor de rest iedereen in het park ofwel aan zijn geliefde zit te friemelen ofwel met een lachend kind speelt. Ik sta echter bij een balustrade stil om naar de zonsondergang te kijken en begin dan iets te veel na te denken, en dan verandert al dat denken in gepieker, en dan staan ze ineens voor mijn neus.

Stilletjes en dreigend komen ze op me af (net Pinkerton-detectives) en gaan aan weerszijden van me staan—Depressie links van me, Eenzaamheid rechts. (etc.)

Fragment 2

‘Het was niet jouw schuld.’

S. hield mijn handen tussen die van haar.

‘Alsjeblieft, B.,’ zei ze. ‘Ik moet het weten.’

Ik zuchtte. Ik was boos. Ik wilde er niet over praten, maar als deze vrouw me ertoe dwong dan zou ik het snel doen en ik zou haar niet sparen.

‘Oké, S.,’ zei ik. ‘Nadat je was weggegaan, namen de mannen ons mee over het strand. (etc.)’

Fragment 1 “Na een dag of tien in Italië hebben
Depressie en Eenzaamheid me gevonden.”

Fragment 2 “Het was niet jouw schuld.’ S. hield mijn
handen tussen die van haar.”

Fragment 1 “Na een dag of tien in Italië hebben
Depressie en Eenzaamheid me gevonden.”
Elizabeth Gilbert, *Eat Pray Love*
(mean rating 3.5)

Fragment 2 “Het was niet jouw schuld.’ S. hield mijn
handen tussen die van haar.”
Chris Kleave, *Kleine Bij*
(mean rating 5.4)

Section 3

Computational experiments

Subquestion 1

What is the influence of unoriginal language use?

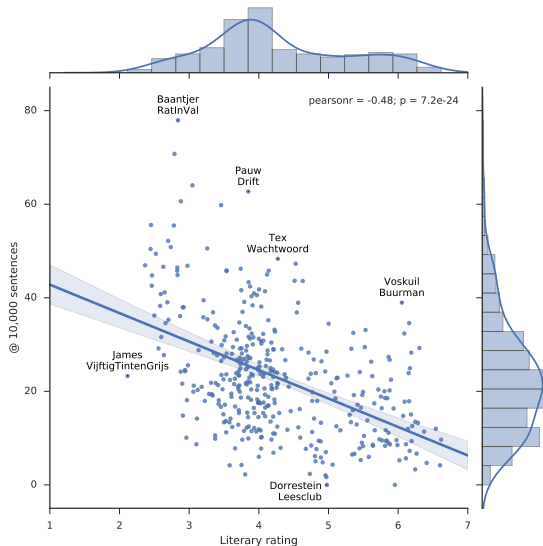
Definition

A **cliché expression** is a fixed, conventionalized multi-word expression which has become overused to the point of losing its original meaning or effect.



- ▶ Zo zie je maar weer.
Well there you go.
- ▶ Nou vooruit, nog eentje dan.
Alright, I'll have another one.
- ▶ Toen ik zo oud was als jij . . .
When I was your age . . .
- ▶ Zou je dat nou wel doen?
Do you really want to go there?

Correlating # cliches with literary ratings



van Cranenburgh (2018), Cliche expressions in literary and genre novels

Louwerse (2004, p. 220):

(...) the lack of internal homogeneity in one text, between texts and between authors can be explained by the (semantic) deviation from the norm the author tries to establish. These variations are exactly what makes the idiolect and sociolect of literary texts unique, and is in fact what makes those texts literary.

Subquestion 2

Is literature (characterized by) a semantic deviation from the norm?

Vector space representations

Represent each chunk of 1000 tokens with:

1. Topic model (LDA)
2. Paragraph vector (DBoW negative sampling)

Geometric characterizations of literariness

- ▶ Distances in vector space model represent semantic similarity and variance.
- ▶ A large distance from other texts can be used as a proxy for creativity, originality, etc.

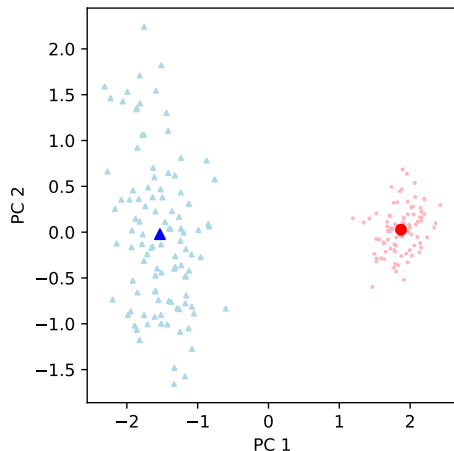
In next 4 slides, we operationalize the notion that literature is defined by a 'semantic deviation from the norm'

Intra-textual variance

Amount of variance within novel: mean of squared distances of chunks.

Intra-textual variance

Amount of variance within novel: mean of squared distances of chunks.



Left: Wieringa, Caesarion (high variance).

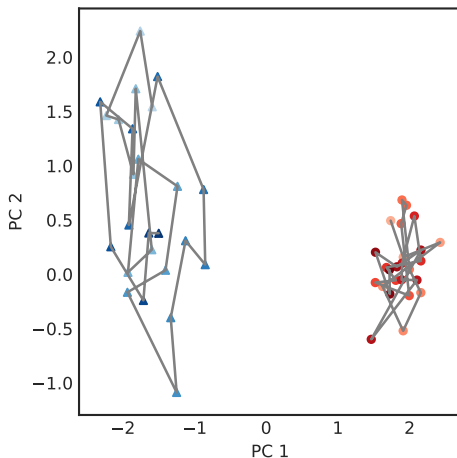
Right: Slee, Fatale Liefde (Fatal Attraction; low variance).

Stepwise distance

How gradual is the development in the novel? Mean of squared distances between consecutive chunks.

Stepwise distance

How gradual is the development in the novel? Mean of squared distances between consecutive chunks.



Left: Wieringa, Caesarion (large distances).

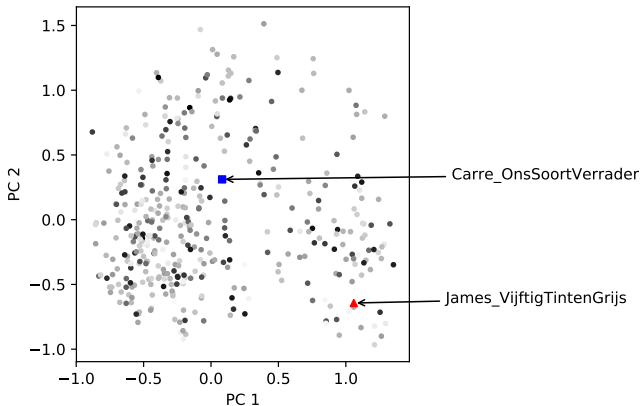
Right: Slee, Fatale Liefde (Fatal Attraction; small dist.)

Outlier score

How different is a novel from other novels? Distance from nearest neighboring novel.

Outlier score

How different is a novel from other novels? Distance from nearest neighboring novel.

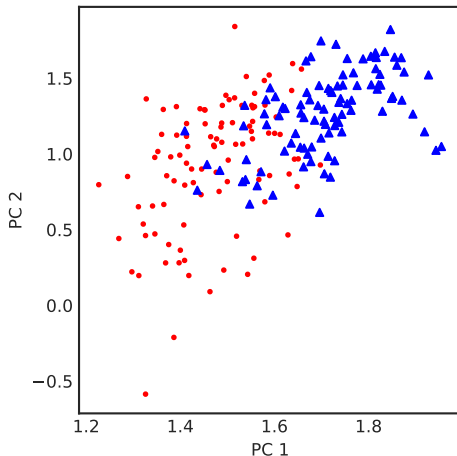


Overlap score

Overlap with other novels. How much of the k nearest neighboring chunks belong to other novels?

Overlap score

Overlap with other novels. How much of the k nearest neighboring chunks belong to other novels?



Red dots: Royen, Mannentester (Man Tester);
Blue triangles: Moelands, Weerloos (Defenseless).

Correlating scores with literary ratings

variable	correlation (r)
intra-textual variance	0.341*
stepwise distance	0.431*
outlier score	0.338*
overlap score	-0.200*

Yes, some support for the notion of literariness as semantic deviation (significant with $p < 0.05$)

Correlating scores with literary ratings

variable	correlation (r)
intra-textual variance	0.341*
stepwise distance	0.431*
outlier score	0.338*
overlap score	-0.200*

Yes, some support for the notion of literariness as semantic deviation (significant with $p < 0.05$)

But: far from a perfect correlation;
we cannot reduce literature to semantic deviation

Intermezzo: The Two Cultures

Explanation:

- ▶ Goal: confirm theory (significance, effect size)
- ▶ Method: hypothesis testing (t-test, χ^2 , OLS)
- ▶ Causal models
- ▶ Small number of predictors
- ▶ No held-out set
- ▶ Used by: social science, econometrics, &c.

Prediction

- ▶ Generalize over unseen data (accuracy)
- ▶ Machine Learning
- ▶ No pre-defined theory / hypothesis
- ▶ Often high-dimensional (many features)
- ▶ Held-out set, CV
- ▶ Industry, NLP

Breiman (2001) Statistical modeling: The two cultures. *Statistical Science*

Shmueli (2010) To explain or to predict? *Statistical Science*

Subquestion 3

Is there a recognizable literary writing style?

Subquestion 3

Is there a recognizable literary writing style?

Hypothesis

Literary authors employ a distinctive inventory of lexico-syntactic constructions (e.g., a register) that marks literary language.

Experimental setup

Task:	predict mean literary rating (1–7)
Training data:	texts of novels
Evaluation metric:	R^2 (\approx % variation explained, baseline=0.0, perfect=100 %)

Experimental setup

Task: predict mean literary rating (1–7)

Training data: texts of novels

Evaluation metric: R^2 (\approx % variation explained,
baseline=0.0, perfect=100 %)

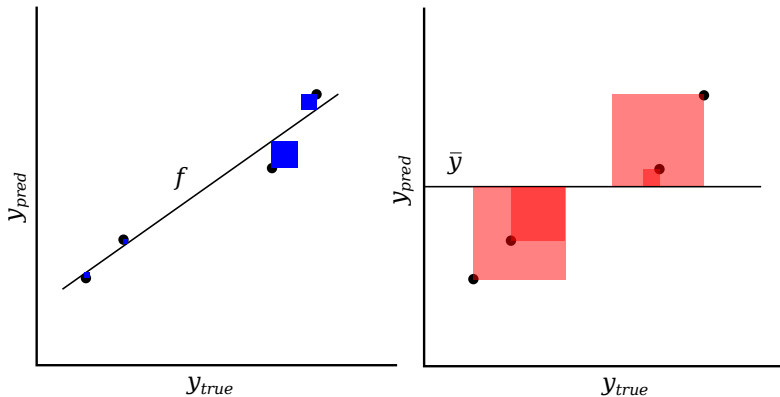
Linear model: simple, efficient;

prediction = sum of weighted features

Regularization: can use many features without overfitting

Crossvalidation: estimate predictive power with
out-of-sample evaluation

Metric: R^2 , % Variation explained



$$R^2 = 1 - \frac{\text{MSE}_{\text{model}}}{\text{MSE}_{\text{baseline}}}$$

Figure: by Orzetto, CC BY-SA 3.0, via Wikimedia commons

Strong lexical baselines

Setup: Linear Support Vector Machine,
5-fold cross-validation

R^2

Basic features

+ Word bigrams

+ Character 4-grams

Strong lexical baselines

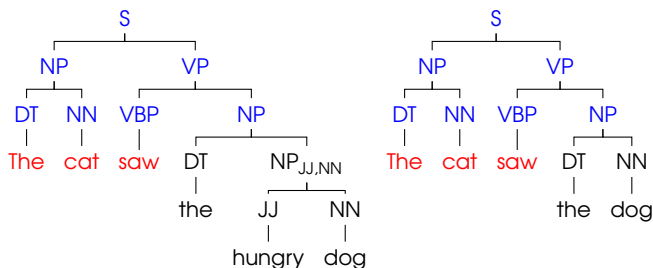
Setup: Linear Support Vector Machine,
5-fold cross-validation

	R^2
Basic features	30.0
+ Word bigrams	59.5
+ Character 4-grams	59.9

On average,

- ▶ 59.9 % of variation in ratings (R^2) is explained using basic and lexical features.
- ▶ the prediction is off by 0.64 (RMSE) out of 0–7.

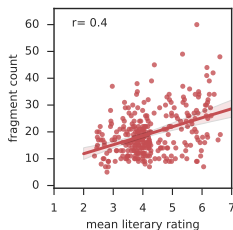
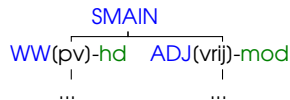
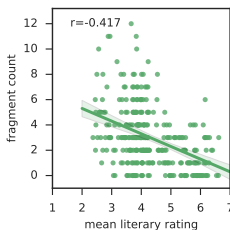
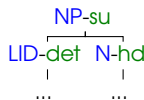
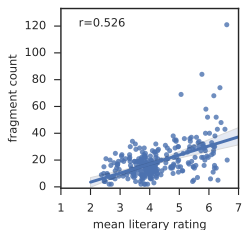
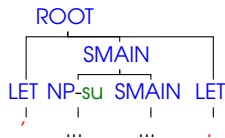
Syntactic Fragments



- ▶ **Data-Oriented Parsing** (Scha 1990, Bod 1992).
Double-DOP (Sangati & Zuidema 2011).
- ▶ Language use is memory-based, depending on arbitrary fragments of language experience
- ▶ Syntactic tree fragments of arbitrary size (connected subsets of tree productions)
- ▶ Extract automatically from **training data**:
find overlapping parts of parse trees

Which fragments are useful?

- ▶ Apply cross-validation
- ▶ Feature selection using correlation with literary rating



Results w/Fragments

	R^2
Basic features	30.0
+ Word bigrams	59.5
+ Character 4-grams	59.9
+ Syntactic fragments	62.2

Results w/Fragments

	R^2
Basic features	30.0
+ Word bigrams	59.5
+ Character 4-grams	59.9
+ Syntactic fragments	62.2

- ▶ Syntax gives modest performance improvement
- ▶ However, features are linguistically more interesting

Analysis of tree fragments

Fragments positively correlated w/literary ratings:

- ▶ Many small fragments
- ▶ Indicators of more complex syntax, e.g.:

appositive NPs: His name was Adrian Finn, **a tall, shy boy**
who (...)
(Barnes, Sense of an ending)

complex, nested NPs/PPs: (...) **a whole storetank of**
existential rage
(Barnes, Sense of an ending)

discontinuous constituents: 'Miss Aibagawa,' declared
Ogawa, 'is a midwife.'
(Mitchell, Thousand autumns of J. Zoet)

Subquestion 4

What is the influence of external factors such as author gender and genre?

Metadata

Coarse genre: Fiction, Suspense, Romance, Other

Translated vs. originally Dutch

Author gender: male, female, mixed/unknown

Metadata

Coarse genre: Fiction, Suspense, Romance, Other

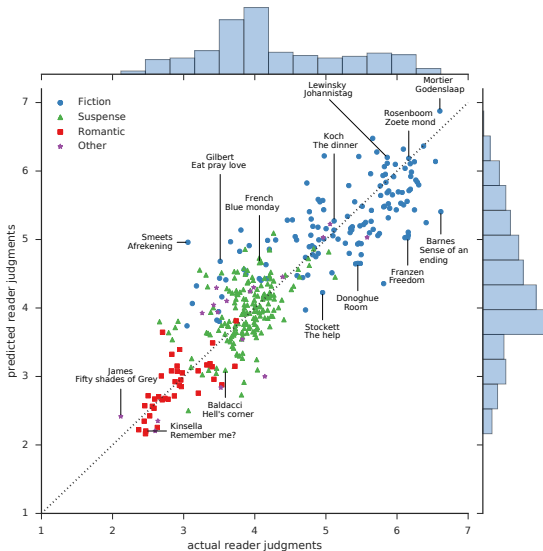
Translated vs. originally Dutch

Author gender: male, female, mixed/unknown

	R^2
BASIC FEATURES	30.0
+ LEXICAL, SYNTACTIC FEATURES	61.2
+ GENRE	74.3
+ TRANSLATED	74.0
+ AUTHOR GENDER	76.0

Table: Metadata features; incremental scores.

Prediction scatter plot: $R^2 = 76.0$



van Cranenburgh & Bod (EACL 2017),
A Data-Oriented Model of Literary Language

Summary

Research Question

are there particular **textual conventions** in literary novels that contribute to readers judging them to be **literary**?

Yes! Literature is characterized by ...

1. Lack of cliché expressions
2. Deviation from the norm
3. A **larger inventory of lexico-syntactic constructions**
4. Genre, author gender influence

Section 4

Reflection

What have we learned?

- ▶ Literary judgments are **not** arbitrary!
- ▶ The literary novel appears to be a genre with its own conventions
- ▶ Language use is highly predictive of literariness

What have we learned?

- ▶ Literary judgments are **not** arbitrary!
- ▶ The literary novel appears to be a genre with its own conventions
- ▶ Language use is highly predictive of literariness
- ▶ Results provide **lower bound** on how informative the text is wrt literariness
- ▶ What is the upper bound? Could we do better?
 - ▶ Yes: infinite number of more sophisticated textual features, models
 - ▶ **But**: noise in ratings, purely social factors, etc.

Limitations

- Causality?** Cannot identify a direct causal link:
text → appreciation,
social & textual factors may overlap
- Scope** analyzed consensus on literariness of
particular time and place.
Results say nothing about aesthetics,
great literature etc.

Limitations

Causality? Cannot identify a direct causal link:
text → appreciation,
social & textual factors may overlap

Scope analyzed consensus on literariness of
particular time and place.
Results say nothing about aesthetics,
great literature etc.

Style

- ▶ Only assessed influence of (micro) style,
not plot and characters
- ▶ Did not disentangle style and topic
aspects
- ▶ No explicit notion of style as choice
between alternatives

Interpretation Machine learning models hard to interpret

Micro vs Macro Style

- ▶ Predictive models use 10,000+ features (**micro style**)
 - ▶ Models can be inspected to see the importance (weight*freq) of each feature
 - ▶ but there are too many, and too many interactions to understand the big picture.
- ▶ Readers seem to judge by an overall, holistic impression
 - ▶ How to describe such a **macro style**?
 - ▶ Can we operationalize this?

Explainable/interpretable ML: LIME, LRP, attention

atheism

christian



Text with highlighted words

From: johnchad@triton.unm.edu (jchadwic)
Subject: Another request for Darwin Fish
Organization: University of New Mexico, Albuquerque
Lines: 11
NNTP-Posting-Host: triton.unm.edu


Hello Gang,

There have been some notes recently asking where to obtain the DARWIN fish.

This is the same question I have and I have not seen an answer on the net. If anyone has a contact please post on the net or email me.

Figure: LIME explains a classification using salient words

Limits of explainability



Ted Underwood
@Ted_Underwood

Following

▼

Good stuff here from Hinton & Yarkoni. Explainability is fine & dandy. But people need to reflect that the barrier to explainability is sometimes the actual complexity of the world, or the limits of the human mind, not "the algorithm we happened to use to model things."

...can't explain how they work, for most of the things. When you hire somebody, the decision is based on things you can quantify, and then all sorts of things. People have no idea how they do that. If you want to explain their decision, you are forcing them to explain it.

...have a similar problem. When you train a neural network, it learns a billion members that represent the data it has extracted from the training data. If you ask it to explain the right decision, say, whether an image contains a pedestrian or not. But if you ask "why did it think this?" well, if there were any simple rules for deciding whether an image contains a pedestrian or not, it would be a solved problem ages ago.

Tal Yarkoni @talyarkoni
Geoff Hinton, on why the drive to "understand" how complex neural network models work is fundamentally misguided
[wired.com/story/googles-...](https://www.wired.com/story/googles-why-we-cant-understand-ai/)
Show this thread

12:24 PM - 15 Dec 2018

https://twitter.com/Ted_Underwood/status/1074037512673349633

Beyond writing style

Aside from writing style, **character and plot** are arguably the major aspects determining the quality of a novel

In future work, look beyond word usage and sentence structure:

Coreference: identify names, nominal phrases, and pronouns that refer to the same object or person

Events: identify events and participants

Separate narrative, description, dialogue

New Survey

Original survey presented respondents with “Author - Title”

- ▶ Respondents may be influenced by author prestige
- ▶ Don't have data on specific textual features that influence literariness
- ▶ General readers vs experts

New Survey

Original survey presented respondents with “Author - Title”

- ▶ Respondents may be influenced by author prestige
- ▶ Don't have data on specific textual features that influence literariness
- ▶ General readers vs experts

Idea for new survey:

- ▶ Present anonymized fragments without author, title
- ▶ Add manipulated versions to see the effect of particular stylistic variables

Conclusions

Research Question

are there particular **textual conventions** in literary novels that contribute to readers judging them to be **literary**?

- ▶ Yes! Literary conventions are **non-arbitrary** because they are associated with textual features
- ▶ Literature can be recognized from text to a large extent: **text-intrinsic literariness**

Conclusions

Research Question

are there particular **textual conventions** in literary novels that contribute to readers judging them to be **literary**?

- ▶ Yes! Literary conventions are **non-arbitrary** because they are associated with textual features
- ▶ Literature can be recognized from text to a large extent: **text-intrinsic literariness**
- ▶ Cumulative improvements with ensemble of features
- ▶ Robust result: both coarse & fine rating differences are predicted

THE END

Publications & code: <http://andreasvc.github.io>



Figure: Huff (1954). How to lie with statistics.