Evaluating Automatic Support for Writing by Example

Katie Kuksenok & Hao Lu

CSE 574 3/14/2011

Outline

- NLP for supporting better writing
- Writing by example as a **mutliclass classification problem**
- Two different classification approaches
- Experimental results

Support for better writing

- Common tools: spell and grammar checking

Support for better writing

- Common tools: spell and grammar checking
- “Rigid-language tools” … alternatives?
  - n-grams+input=support for non-native speakers

Park, T., Lank, E., Poupart, P., & Terry, M. " *Is the Sky Pure Today ?” AwkChecker : An Assistive Tool for Detecting and Correcting Collocation Errors.* UIST 2008.

Support for better writing

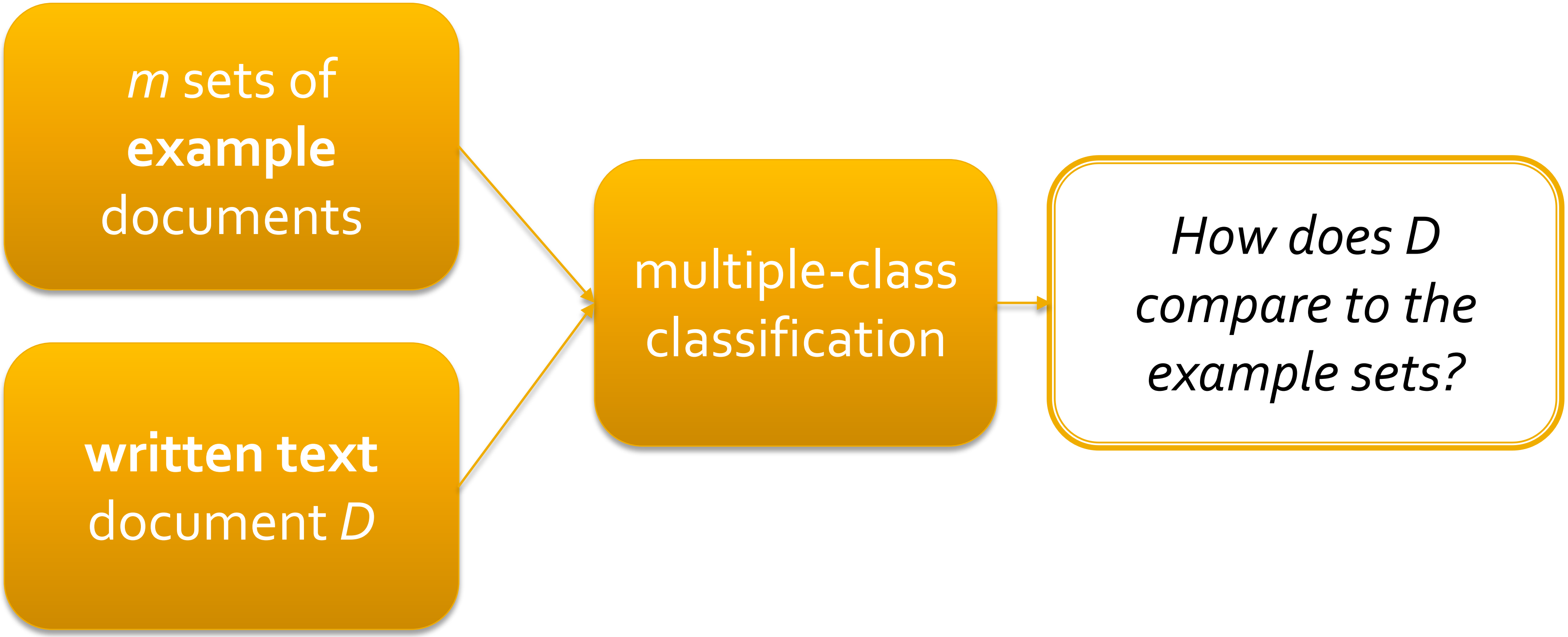
- Common tools: spell and grammar checking
- “Rigid-language tools” … alternatives?
  - n-grams+input=support for non-native speakers

Park, T., Lank, E., Poupart, P., & Terry, M. “ *Is the Sky Pure Today ?” AwkChecker : An Assistive Tool for Detecting and Correcting Collocation Errors.* UIST 2008.

- Leaving decision-making to the user?
  - Visual feedback without valence/judgment

Keim, D. a, & Oelke, D. (2007). Literature Fingerprinting: *A New Method for Visual Literary Analysis.* IEEE Symposium on Visual Analytics Science and Technology, 115-122. 2007.

Writing by Example



m sets of **example** documents

written text document *D*

multiple-class classification

How does D compare to the example sets?

Two datasets

ACM DL

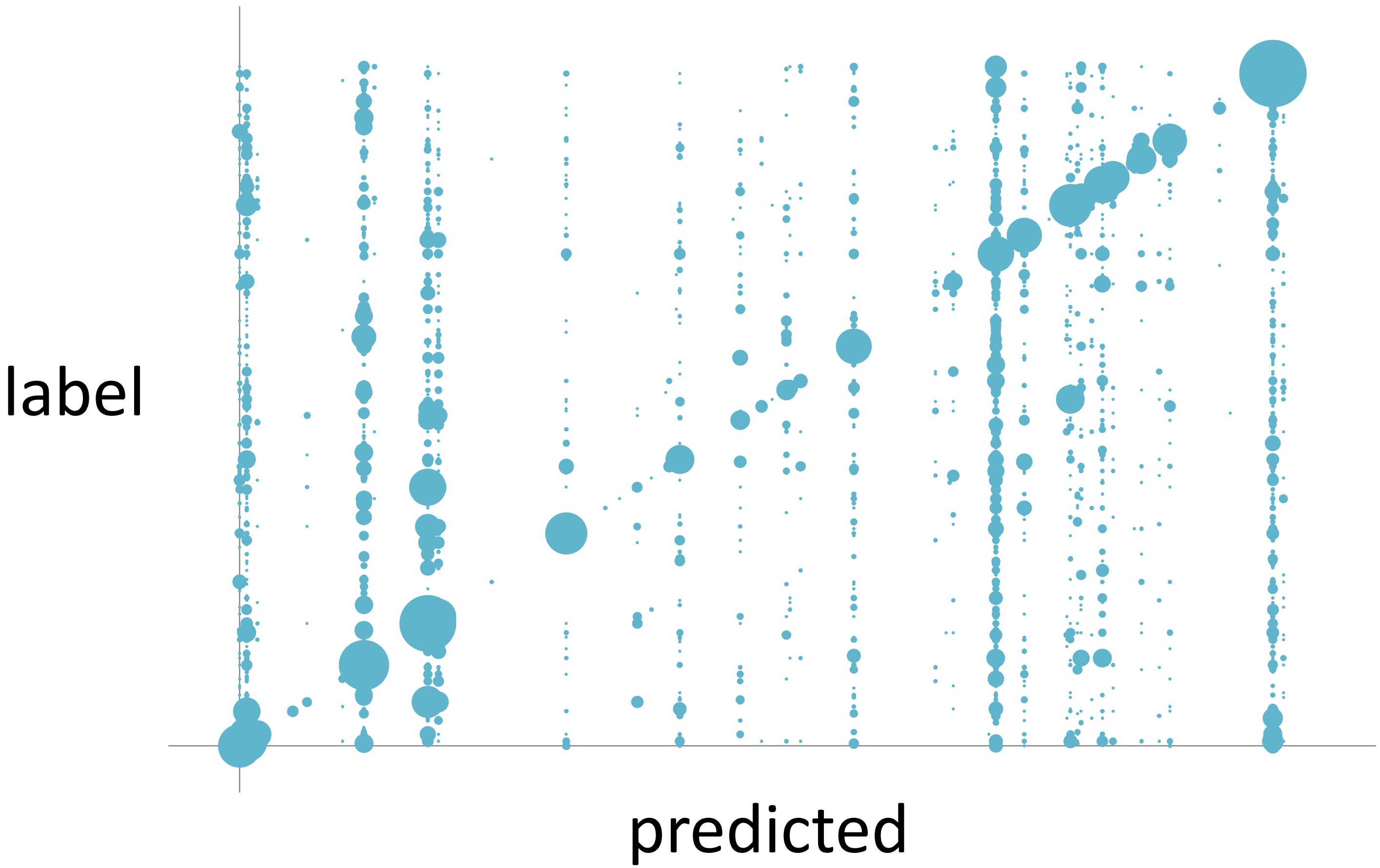
117,823	abstracts
296	venues

Project Gutenberg

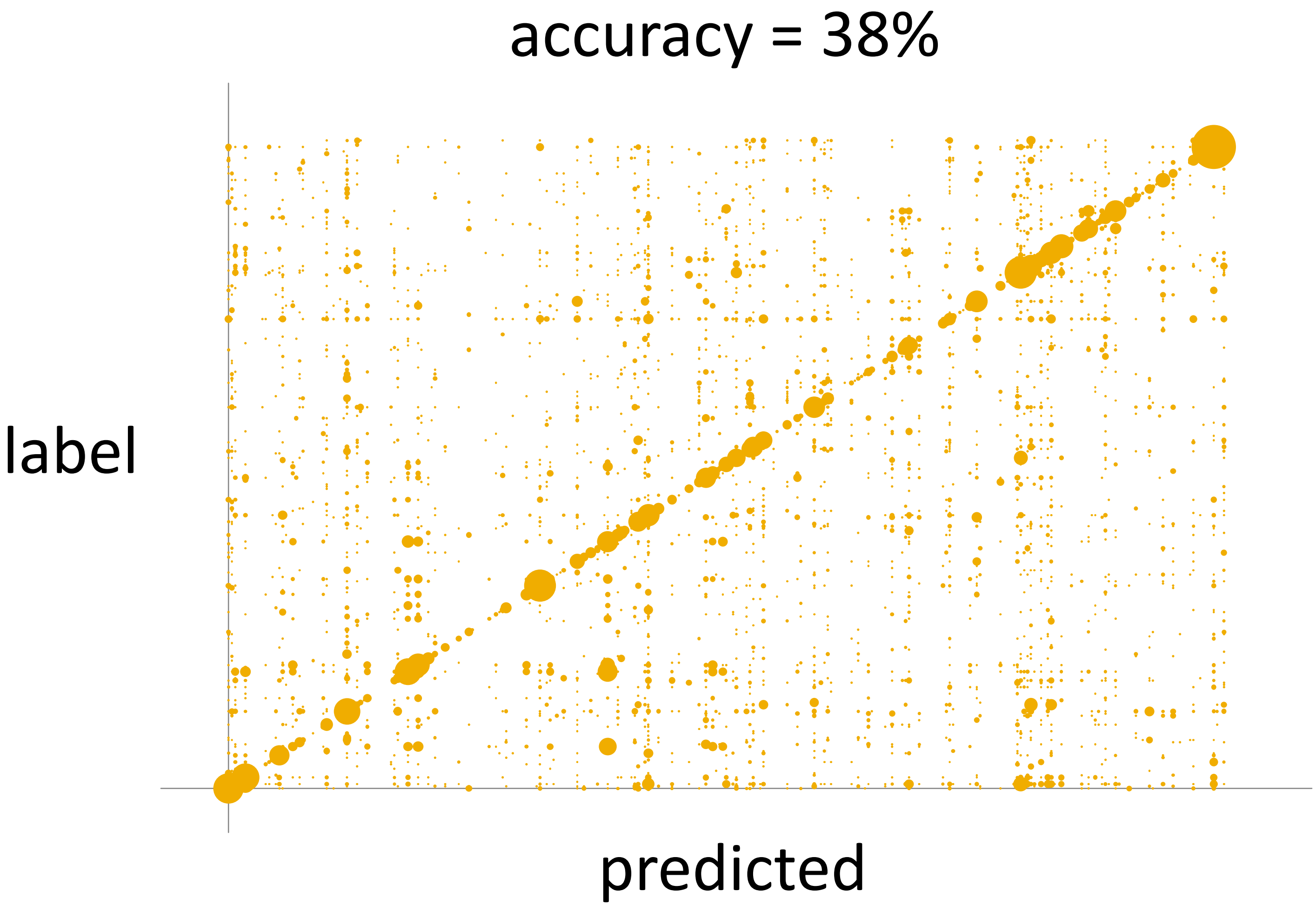
12,912,372	segments
3654	authors

Multiclass with bag of words

accuracy = 29%



Multiclass with bag of words

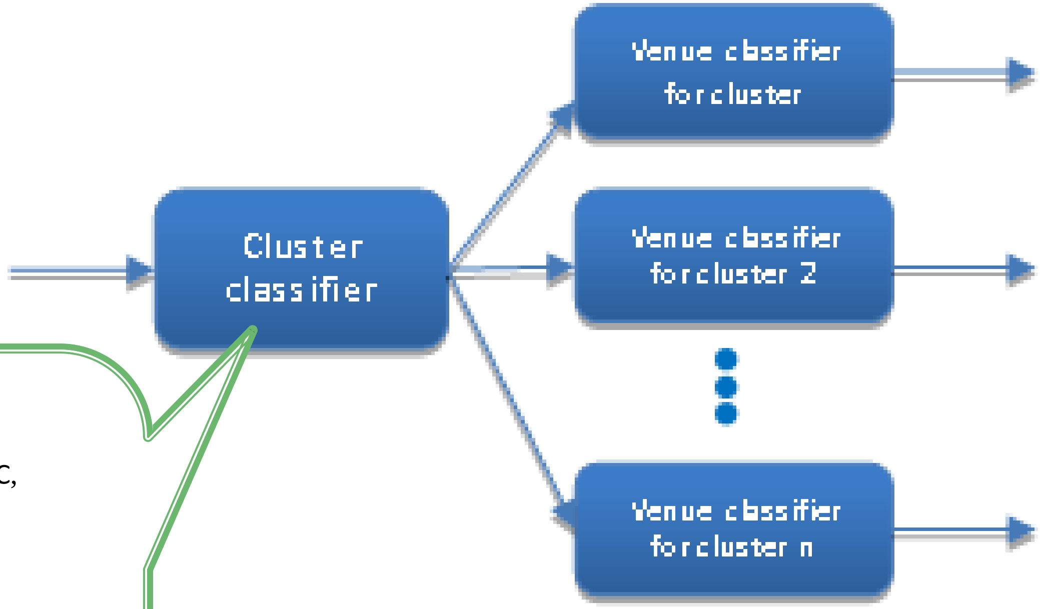


Multiclass with bag of words

DATE, ASPDAC, ISSS ISPD, CODES, GLSVLSI, DAC, EURO-DAC, ICCAD, ISLPED
SIGIR, WWW, CIKM, KDD, PODS, SIGMOD
PODC, SoCG, SPAA, STOC, SODA
CHI, UIST, CSCW
POPL, PLDI

Clusters of Venues

Multiclass with bag of words


Cluster classifier
Venue classifier for cluster
Venue classifier for cluster 2
Venue classifier for cluster n

DATE, ASPDAC, ISSS ISPD, CODES, GLSVLSI, DAC, EURO-DAC, ICCAD, ISLPED

SIGIR, WWW, CIKM, KDD, PODS, SIGMOD

PODC, SoCG, SPAA, STOC, SODA

CHI, UIST, CSCW

POPL, PLDI

Figure 4: two-stage classification

Multiclass with bag of words

- No improvement
- Clusters as classes
  - accuracy: 49%
- Improve clustering?

Binary classification

Document 1 = *which class?*

Binary classification

Document 1 = *which class?*

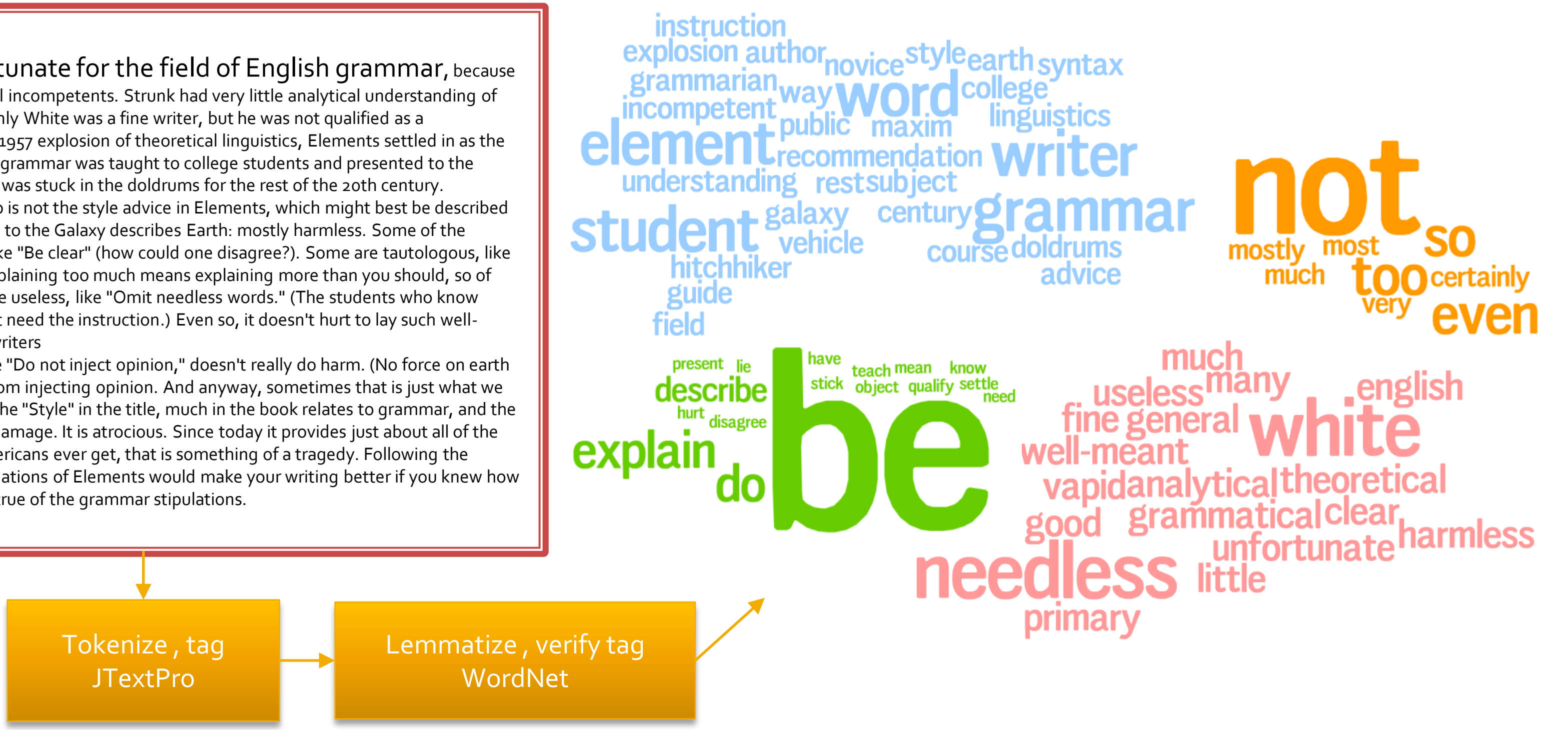
Document 1 x Document 2 = *same class?*

Binary classification

This was most unfortunate for the field of English grammar, because both authors were grammatical incompetents. Strunk had very little analytical understanding of syntax, White even less. Certainly White was a fine writer, but he was not qualified as a grammarian. Despite the post-1957 explosion of theoretical linguistics, Elements settled in as the primary vehicle through which grammar was taught to college students and presented to the general public, and the subject was stuck in the doldrums for the rest of the 20th century.

Notice what I am objecting to is not the style advice in Elements, which might best be described the way The Hitchhiker's Guide to the Galaxy describes Earth: mostly harmless. Some of the recommendations are vapid, like "Be clear" (how could one disagree?). Some are tautologous, like "Do not explain too much." (Explaining too much means explaining more than you should, so of course you shouldn't.) Many are useless, like "Omit needless words." (The students who know which words are needless don't need the instruction.) Even so, it doesn't hurt to lay such well-meant maxims before novice writers

Even the truly silly advice, like "Do not inject opinion," doesn't really do harm. (No force on earth can prevent undergraduates from injecting opinion. And anyway, sometimes that is just what we want from them.) But despite the "Style" in the title, much in the book relates to grammar, and the advice on that topic does real damage. It is atrocious. Since today it provides just about all of the grammar instruction most Americans ever get, that is something of a tragedy. Following the platitudinous style recommendations of Elements would make your writing better if you knew how to follow them, but that is not true of the grammar stipulations.



Tokenize , tag
JTextPro

Lemmatize , verify tag
WordNet

Binary classification

first small modest low

question query

set circle circuit

part contribution

verification chip check

configuration shape form contour

procedure operation function process

track course class path row line

instance example illustration

frame sort variant shape kind phase figure configuration descriptor variety pattern form class contour

translation adaptation variation variant rendering interpretation version

agree check fit match correspond hold accord harmonize

prove see test try analyze examine study

search seek look explore

sample prove stress test attempt try render seek examine

integrate comprise incorporate contain

see affect consider regard involve

distinguish recognize realize know

execute do perform

search see face expect calculate appear look depend seem

gentle soft blue easy

yet still even

so then

problem trouble job

figure finger digit

sketch study view sight survey

condition status position

lot set circle band ring stripe

care help aid assistance attention

leg branch stage peg

trace shadow darkness phantom tail phantasm vestige dark

don father

captain master chieftain

mouthpiece lip mouth

precept principle rule

vision imagination sight

account chronicle history story

finish stop end terminate cease

note name remark mention cite refer observe

propose extend tender volunteer bid provide offer

Binary classification

- >50% accuracy, with up to 75%
- 44 experiments

Binary classification

- >50% accuracy, with up to 75%
- 44 experiments
  - ACM: between 50 and 60
  - Gutenberg: high 60s and mid 70s

Final thoughts

- A clearly difficult problem…

Final thoughts

- A clearly difficult problem…
  - …but maybe there is promise at a smaller scale?

Final thoughts

- A clearly difficult problem…
  - …but maybe there is promise at a smaller scale?
- Other features…
  - …like features
    provided by the user?

Final thoughts

- A clearly difficult problem…
  - …but maybe there is promise at a smaller scale?
- Other features…
  - …like features provided by the user?

- Demo + questions?