wikigrader

Towards automatic article assessment

Wikipedia

The Free Encyclopedia

Assessment grades

FA
A
GA
B
C
Start
Stub

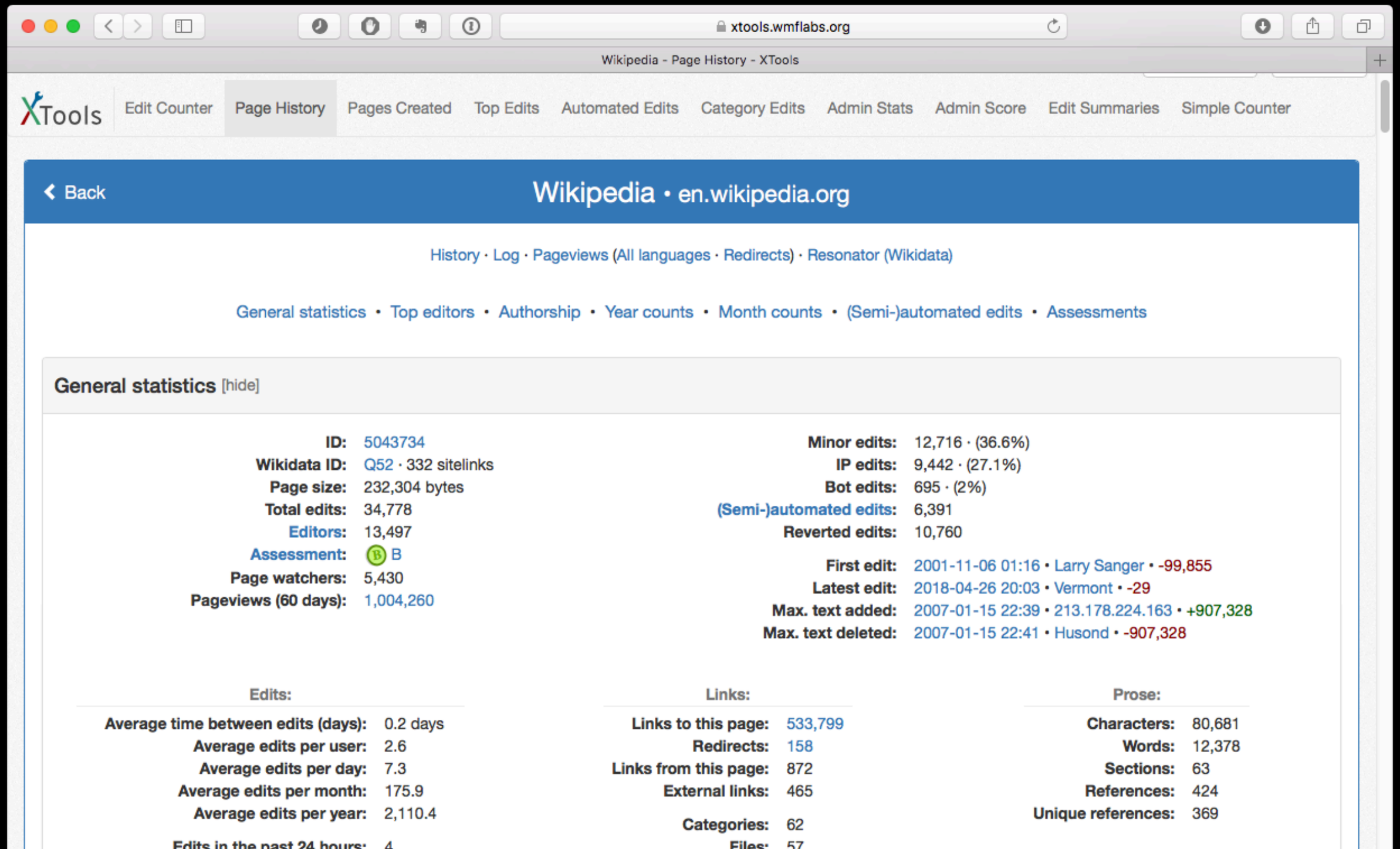
A challenge

- 574,103 articles haven't been assessed.
- About 800 new articles added every day.
- About 10 edits per second.

Source: en.wikipedia.org/wiki/Wikipedia:Statistics

Can we build a tool to help?

Data


xtools.wmflabs.org
Wikipedia - Page History - XTools
XTools
Edit Counter
Page History
Pages Created
Top Edits
Automated Edits
Category Edits
Admin Stats
Admin Score
Edit Summaries
Simple Counter
Back
Wikipedia • en.wikipedia.org
History · Log · Pageviews (All languages · Redirects) · Resonator (Wikidata)
General statistics • Top editors • Authorship • Year counts • Month counts • (Semi-)automated edits • Assessments
General statistics [hide]
ID: 5043734
Wikidata ID: Q52 · 332 sitelinks
Page size: 232,304 bytes
Total edits: 34,778
Editors: 13,497
Assessment: B
Page watchers: 5,430
Pageviews (60 days): 1,004,260
Minor edits: 12,716 · (36.6%)
IP edits: 9,442 · (27.1%)
Bot edits: 695 · (2%)
(Semi-)automated edits: 6,391
Reverted edits: 10,760
First edit: 2001-11-06 01:16 • Larry Sanger • -99,855
Latest edit: 2018-04-26 20:03 • Vermont • -29
Max. text added: 2007-01-15 22:39 • 213.178.224.163 • +907,328
Max. text deleted: 2007-01-15 22:41 • Husond • -907,328
Edits:
Average time between edits (days): 0.2 days
Average edits per user: 2.6
Average edits per day: 7.3
Average edits per month: 175.9
Average edits per year: 2,110.4
Links:
Links to this page: 533,799
Redirects: 158
Links from this page: 872
External links: 465
Categories: 62
Prose:
Characters: 80,681
Words: 12,378
Sections: 63
References: 424
Unique references: 369

Data

FA	1321 samples
A	490 samples
GA	1978 samples
B	1841 samples
C	2102 samples
Start	2448 samples
Stub	2248 samples

Data

5.0	1321 samples
4.0	2468 samples
3.0	1841 samples
2.0	2102 samples
1.0	2448 samples
0.0	2248 samples

Features

	coef	P>|t|
const	-4.6743	0.000
Account edits (ratio)	2.4272	0.000
Age (days) (log)	-0.0754	0.000
Average edits per day (log)	-0.1285	0.000
Bugs	-0.0852	0.000
Categories	-0.0068	0.000
Words (log)	0.7011	0.000
Editors (log)	0.1244	0.000
File links (log)	0.2212	0.000
Redirects	-0.0079	0.000
References	0.0067	0.000
Reverted edits (ratio)	0.8645	0.000
Sections	-0.0312	0.000
Semi-automated edits (ratio)	-0.8501	0.000
Unique references	-0.0060	0.000
External links (log)	0.0899	0.000

Linear model results

(training set)

Dep. Variable:	Rating	**R-squared:**	0.644
Model:	OLS	**Adj. R-squared:**	0.644
Method:	Least Squares	**F-statistic:**	1198.

Linear model results

(test set)

	Score
R^2	0.631885
RMSE	1.000740
MAE	0.815706

Quadratic model results

Training set

	Scores (3 folds)
R^2	0.704022
RMSE	0.807266
MAE	0.700995

Test set

	Score
R^2	0.680508
RMSE	0.932310
MAE	0.729501

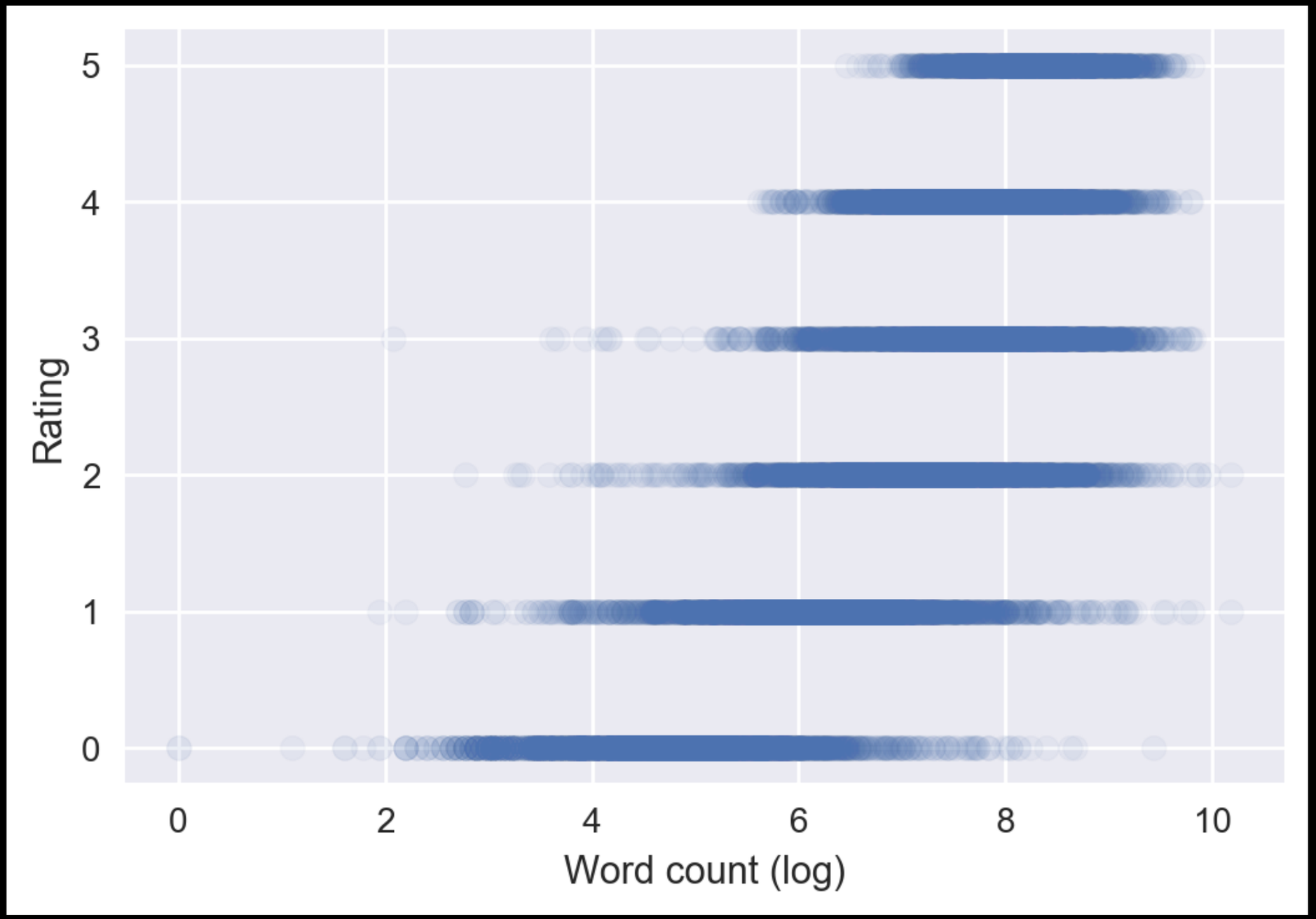
Single regressor models

	R^2
Words (log)	0.542076
External links (log)	0.316510
Total edits (log)	0.299527
References	0.297694
File links (log)	0.263080
Unique references	0.241305

Future directions

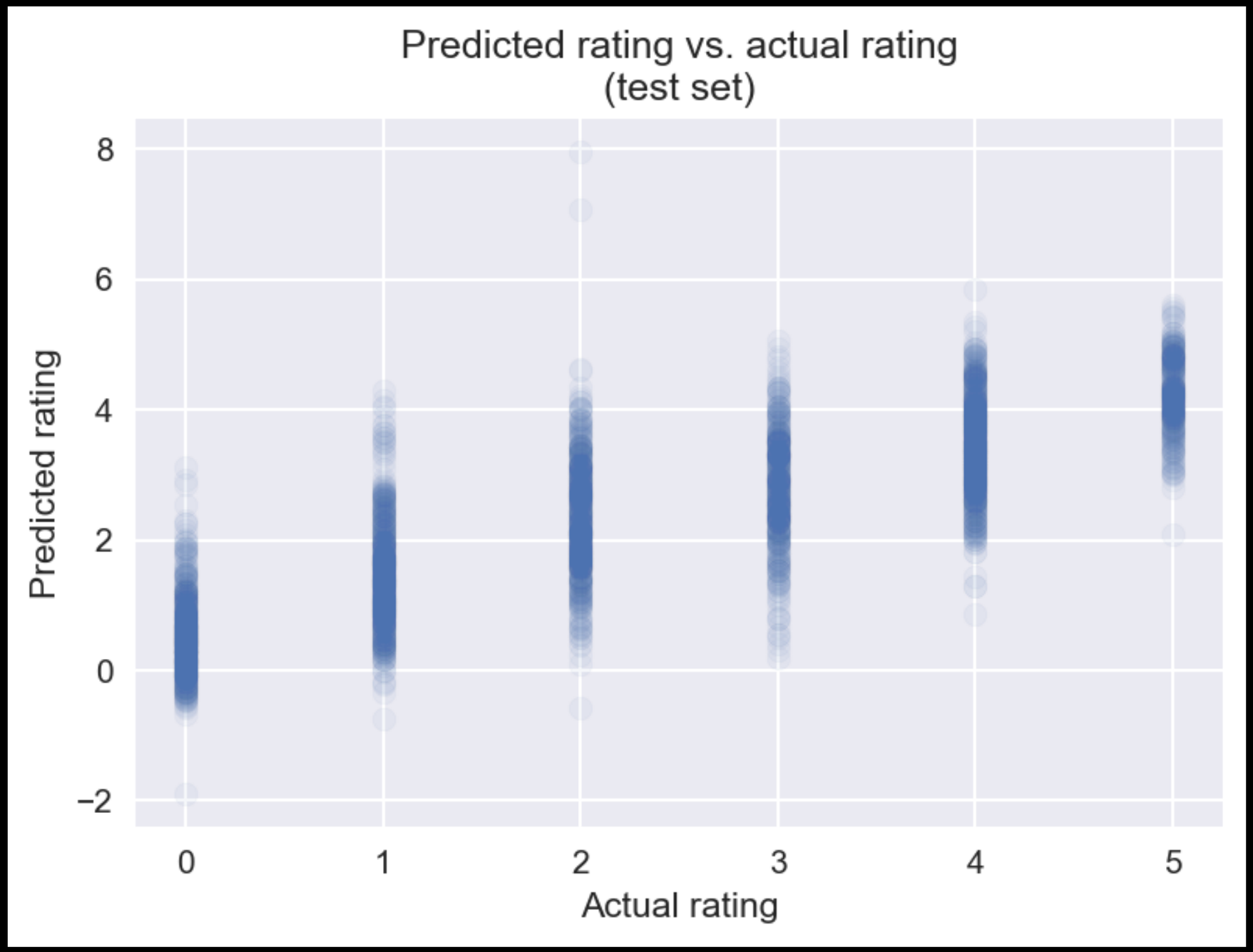
- Consider complex regression models.
- Simple neural net gets you to about 0.72 R^2, 0.65 MAE (test set).
- But to really improve, we'll probably have to actually look at the words in the articles.

Rating vs. word count



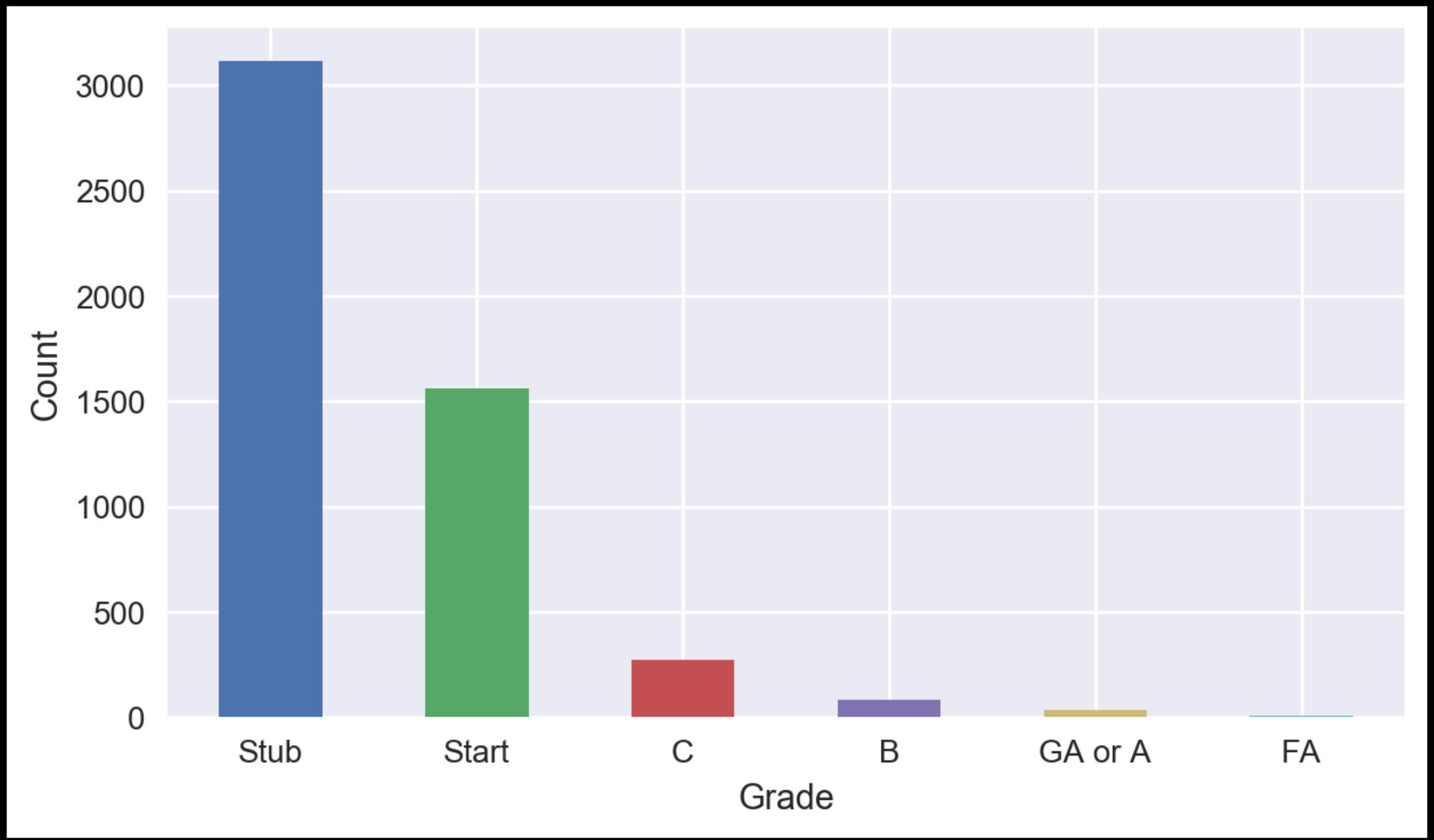
Predicted vs. actual rating

(test set)



5073 graded articles

(aborted first scraping attempt)



NN model results

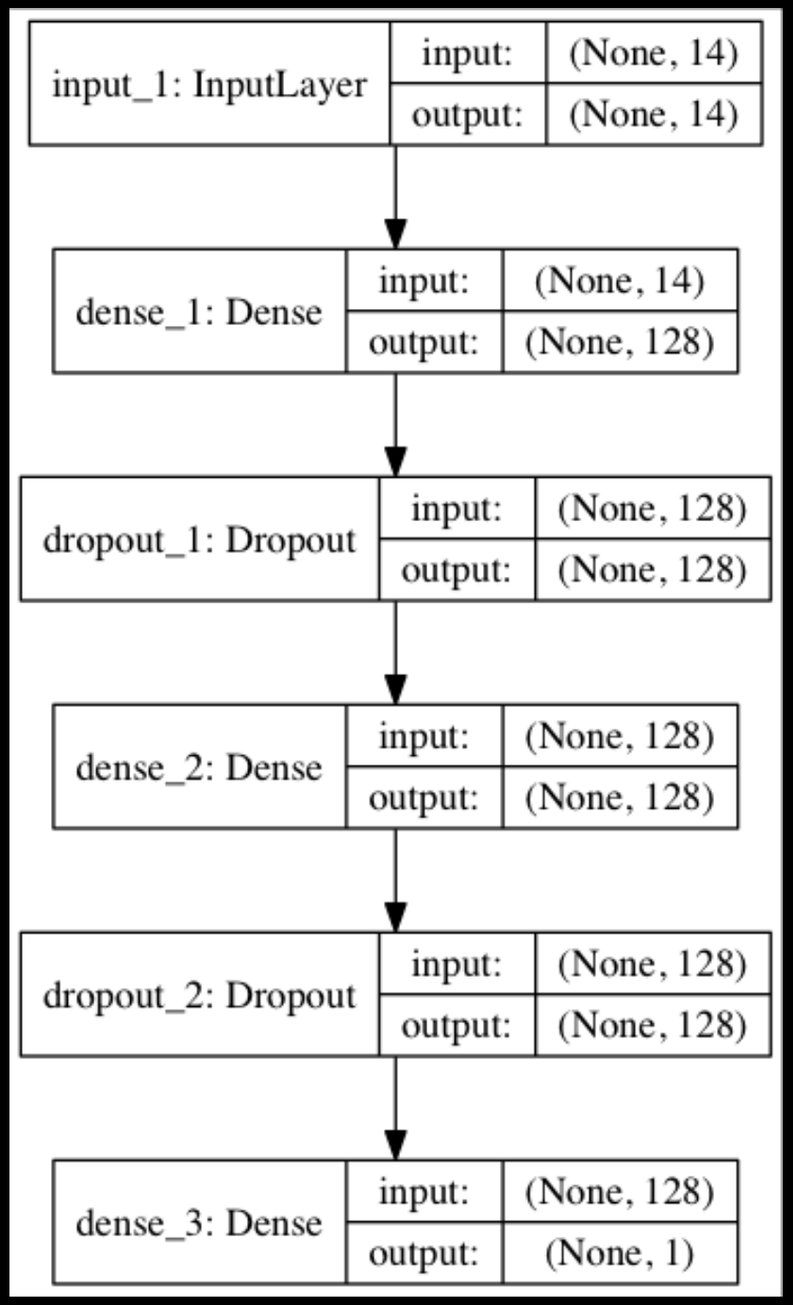
Training set

Score	
R^2	0.766819
RMSE	0.797531
MAE	0.588366

Test set

Score	
R^2	0.725147
RMSE	0.864729
MAE	0.645892

NN model details



Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 14)	0
dense_1 (Dense)	(None, 128)	1920
dropout_1 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 128)	16512
dropout_2 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 1)	129

Total params: 18,561
Trainable params: 18,561
Non-trainable params: 0