Quantifying Influence in Collaborative Social Networks

Elizabeth Bradley
Adviser: Thomas Funkhouser

Abstract

In this paper, we attempt to isolate quantifiable features of a Wikipedia editor's user history that contributes to their success or failure in the sitewide Request for Adminship (RfA) elections. Unsurprisingly, we find that the most influential features are those most central to the Wikipedian mission: sustained and voluminous collaboration and individual contribution. As an additional conclusion, we find that decision tree class of classifiers is particularly effective at representing the thought process of voters in the RfA elections as they parse through leveled priorities and self-imposed quotas.

1. Introduction

Online social networks dominate the experience of the internet's users. Micro-blogs (e.g. Twitter) filter our news, sharing sites (e.g. Instagram, Facebook) determine our social connections, and collaborative sites (e.g. Wikipedia) frame our knowledge of the world. Users on online social networks do not exist in a vacuum, but instead in an interwoven web of user-on-user influence. In order to understand the mechanics of online social networks, the expression of influence in these webs must be understood as well.

Identifying influencers is a matter of both practical concern for online social networks. Advertisers benefit from hiring influencers on social sharing sites to hock product, while online social networks benefit from identifying influencers capable of taking on formal duties in the community.

This paper analyzes a social network not often covered in the literature: Wikipedia. Wikipedia is an online collaborative encyclopedia whose five million articles[16] are the product of volunteer editors. There is a low cost of participation to edit as users do not even need to register an account. In this goal, Wikipedia is unique among social networks. The problem of identifying influence is complicated by this emphasis on collaboration, i.e. users interact through corporate effort instead of socialization, and the anonymity of its users.

The goal of this report is to identifying quantitative features of a user's history, unique to the characteristics of Wikipedia, that increase that user's perceived influence. Our method of identifying influence will be the success of Request for Adminship elections, for — as Kittur et. all conclude — "[Wikipedia Admins] have been peer-selected as belonging to a class trusted with more power than a normal user." [6]

2. Related Work

2.1. Influence

Since the advent of online social networks, numerous parties have invested effort into programmatically identifying influencers in various shapes of social graphs. This identification, although a modified problem for every network, invokes the spirit of Katz's 1955 defining explanation of social influence: "the power exerted by a minority of people, called opinion leaders. An opinion leader is a subject which is very informed about a topic, well-connected with other people in the society and well-respected."[4] [5]

The quantitative indicators of influence tend to platform-specific. For instance Bakshy et. all understood an "opinion leader" on Twitter to be a user with an undue number of retweets and social connections [1], the retweet as a measure of word-of-mouth diffusion — a unique consequence of the ephemerality of micro-blogs.

2.2. Wikipedia

Wikipedia has over 27 million registered users, with 115 thousand active (having contributed in December 2015). Yet, even with these large user counts, very few users are considered "power users:" as of February 2015, 12,000 users had over 600 edits with activity in the previous six months [17]. Panciera et. all [11] concluded that this set of users has quantifiably different user activity than lay users, from even their first day of editing. It is the hope of this report to identify even finer distinctions in the user activity between another two classes of editor: those who are successful admin candidates and those who are unsuccessful.

3. Approach

Unlike other online social networks more prevalent in the literature — e.g. Twitter, Facebook — Wikipedia's users are wholly anonymous. Whereas a pop artist such as Lady Gaga may exhibit large amounts of influence on Twitter, this influence extends not from the quality of the user's activity (tweets) but from the user's real-world social capital. The challenge of identifying influence in Wikipedia is complicated by this level playing-field.

In addition to the lack of real-world connections (superficial or personal) between users, it is unlikely that any two users, even power users, might encounter one another on the site due to Wikipedia's aforementioned size. Therefore, we must redefine influence from our conceptual definition based off association and interaction to a baser meaning: "perceived levels of trustworthiness." On a site full of vandals, trustworthiness is oft-admired characteristic.

In order to identify a set of users wherein distinguishing influence was nontrivial, a corpus of Request for Adminship (RfA) elections was examined. RfAs are bureaucratic feature of site, purposely hidden from non-editors. Although any user with a registered account may become a candidate in these elections, knowledge of the RfA elections' existence itself provides a barrier to entry. Administrators, when elected are entrusted with a potent set of tools that includes banning users and deleting pages. This trust falls in line with our revised definition of influence — "perceived trustworthiness" — i.e. users would only grant these tools to users they trusted to wisely wield them.

4. Implementation

The Stanford Network Analysis Project (SNAP)[14][13] published the underlying dataset for this report. The dataset contains a record of each vote in a Request for Adminship (RfA) election from January 2003 to May 2013 (an example vote is included in Figure 1). Using Python, the

```
SRC:Guettarda TGT:Lord Roem VOT:1 RES:1 YEA:2013 DAT:19:53, 25 January 2013
TXT:'''Support''' per [[WP:DEAL]]: clueful, and unlikely to break Wikipedia.
```

Figure 1: Sample Vote from the SNAP dataset. Each record contains the usernames of the admin candidate (TGT) and the voter (SRC), the direction of the vote (VOT), the outcome of the election (RES), the date of voting (DAT), and a written text of the voting (TXT) (median/mean: 19/34 tokens).

SNAP dataset was collated into unique elections, indexed by candidate's username and the date of the earliest vote. This paradigm — treating elections as instances — offered a natural evaluation procedure for the experiment: predicting a given election's success through features of its candidate's activity.

4.1. Data Collection

Between the literature review and a manual examination of the textual content of votes in the SNAP dataset, four categories of potentially significant features were identified: general activity, roles assumed, quality of edits, and voting and/or editing habits. In total, 44 features were collected for each user. Each features was collected through one of three pipelines:

1. **API Request** MediaWiki[10] is free server-based software built to support Wikipedia and related Wikimedia sites. Maintained by volunteers, the MediaWiki API has a robust feature set used for as varied purposes as creating vandalism-revision bots to maintaining internal statistics. The Python package *Requests*[2] was used to interface with the API and parse json. The routes /Users[8] and /Usercontribs[9] were hit for the user's registration data and editing behavior, respectively.
2. **Webscraping - HTML** The Tools Labs project [] hosts nearly 1250 community-maintained tools for Wikipedia. Several tools — including *afdstats*[15] which provided the basis for features concerning voting behavior — output simple HTML charts. The lightweight Python package *Beautiful Soup*[] was used to extract data from these tables.
3. **Webscraping - JavaScript** A majority of tools hosted by the Tools Labs project render data inside figures and charts in JavaScript (invisible to traditional scraping libraries). To address both the render complication and the slow response of the tools, 30+ Virtual Machines (1 CPU, 1 GB Ram, 5GB HDD) were hosted on Cloud9[]. Each used the Python package *Selenium*[3]

interfaced with the Node.js package *Phantom.js*[12] to emulate a browser, enabling JavaScript-rendered objects to be viewed. Features gathered in this manner include those in the category "roles assumed," collected from [].

4.2. Machine Learning

All experimentation was performed in Weka, a free machine learning workbench developed by the University of Waikato [7]. The classifiers used were ZeroR, Logistic Regression, MultilayerPerceptron, J48 Decision Tree, Random Forest, and AdaBoosted Random Forest. Plot.ly was used to generate the bar graphs.

Standardization was performed on all data to minimize the relative effects of scaling (note: the data distributions presented in Section 5 have not been standardized). Due to the small size of the dataset, 10-fold cross-validation was performed instead of the more typical procedure of isolating a training set. Accepting the variance introduced by attempting to predicting human decisions, no outliers were removed from the dataset.

Principal Component Analysis was performed in Python using code provided by COS IW05 Teaching Assistant Theodore Brundage.

5. Data

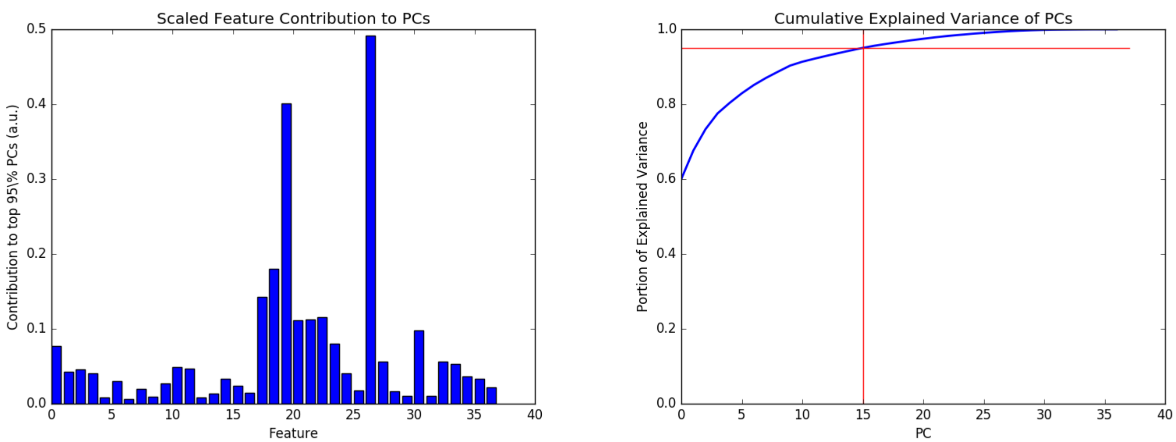
The initial dataset recorded votes from approximately 3,500 elections. After removing malformed data (e.g. missing names, nonsensical dates), complete or partial records existed for 3,388 elections. Of these elections, 50.94% (1,726) were successful — the candidate became an admin — and 49.06 % (1,662) were unsuccessful — the candidate did not become an admin.

In these elections, 3,039 unique users ran for adminship. As there are no formal restrictions on re-running, some users who are unsuccessful choose to run again months or years later (even as many as five additional times). For the purposes of prediction, each election in which a user — unsuccessful in a previous attempt — runs again we record as a separate instance from that initial attempt, for the "new" user has a different, albeit not wholly unique, user history.

In total, 44 features were collected for each user, each a quantitative aspect of that user's history, broken into four categories: general activity, roles assumed, quality of edits, and voting and/or editing habits. Each category of features will be discussed at length in later subsections.

As an initial exploration, principle component analysis (PCA), a naïve statistical procedure, was performed to attempt an explanation at the variance between instances. As PCA is sensitive to the scaling of the data, each feature was re-scaled by its range. As a quirk of the dataset wherein the minimum of every feature is zero, this produced the atypical result of rescaling every feature to the range [0,1]. In doing so, binary features (i.e. features that were either true, "1," or false, "0") had to be excluded from the PCA as to not skew feature importance in their favor. [1]

Figure 2: Principal Component Analysis of all Features



As Figure 2 reflects, the PCA concludes that the features contributing most prominently to the principle components are the number of days of activity (17), voting habits in the sitewide Article for Deletion polls (18-26), and number of administrator elections — including the current one — the user has participated in (30). These features, like those that contribute least — e.g. number of images uploaded (31) or number of edits on the user's own homepage (4) — are members of varied categories, i.e. represent different voting paradigms. This suggests that the human calculus that goes into voting for administrators is more nuanced than examining only one facet of a user's activity.

[1]To explain the "quirk" in the dataset. A handful of users ran for administrator election on their first day as registered users and thus possessing no user activity.

Even from the shallow analysis provided by PCA, it is apparent that although some features contribute more prominently, no single feature (or even handful of features) is solely able to explain the variance between instances. Therefore, the remainder of this section dissects each category of user history (general activity, roles assumed, quality of edits, and voting and/or editing habits) in further detail.

5.1. General Activity

As the most accessible features, made available by numerous tools[18], general measures of activity are often the first user characteristics analyzed by voter unfamiliar with the candidate. Of all the feature categories, general user activity is mentioned the most often in the SNAP dataset. For example, in Figure 3, excellent edit activity is reason enough for a positive vote.

As listed in Figure 1, the measures selected capture varied means of assessing a user's activity:

1. **Total edits** — the most straightforward, if unnuanced measure. A count of all a user's activity. Included due to overwhelming reference in the text of RfA votes and Panciera et. all [11] conclusion that Wikipedia administrators are consistently more prolific than lay users from their first day as a registered user.
2. **Average Daily Edits** — a reweighted edit count that grants leniency to younger users who, although active, have smaller total edit counts.
3. **Days Active** — a record of longevity. Inspired by Panciera's further conclusion that Wikipedia administrators maintain consistent activity until "burn-out" as compared to lay users[11].
4. **Pages Created / Images Uploaded** — a measure of a user's interest in bring new content to the website. Included due to constant reference in the text of RfA votes about the importance of content creation.
5. **Accounts** — a representation of a user's summed involvement in Wikimedia (umbrella group of sites).
6. **Votes in RfA / Votes in AfD** — a measure of a user's involvement in "meta" activities of the site, uninteresting and all-but-hidden from lay users.

Figure 3: Excerpt from Pharaoh of the Wizards's vote for Ynhockey (2008)

```
TXT:'''Support''' User has been around since Nov 2004 and had the first RFA
in 2006 and has waited for 2 years to try again after overcoming the points
raised in previous RFA. User has more than 19000 mainspace edits and over
25000 overall.
```

	Pages Created	Total Edits	Av. Daily Edits	Days Active
Mean	165.7	4,760.0	12.8	524.5
Std. Dev.	472.4	9153.3	25.1	495.6
Min	0	0	0	0
25%	7	646.3	1.9	176
50%	40	2421.5	6.1	388
75%	134.3	5551.8	14.3	725
Max	9828	238686	690	3877
	Accounts	**Votes in RfA**	**Votes in AfD**	**Images Uploaded**
Mean	13.6	10.0	56.4	9.0
Std. Dev.	53.1	26.3	62.5	60.7
Min	0	0	0	0
25%	1	0	2	0
50%	1	1	29	1
75%	7	8	100	4
Max	700	316	199	2511

Table 1: Distribution of General Activity Feature Set

5.2. Roles Assumed

As exploratory evaluation suggested, the unnuanced feature "Total Edit Count" (from the previous category) in isolation failed to adequately distinguish successful users from unsuccessful ones. In response, the general edit count was dissected into ten separate edits counts for distinct areas of the site. These new are more explanatory of a candidate's preferred "role" or "roles" on the site, e.g. "Content Creator" or "Collaborator."

Total edit counts in each of these areas of the site was chosen in lieu of percentages (e.g. "Percentage of Total Edits on User Pages"), as users — especially as they grow more active —

switch fluidly between roles. Although, these total counts still suffer from the undue influence of sheer volume, it enables users to adopt several roles over time in the calculation.

As it also became clear through an examination of the textual content of votes, several voters had standards as to how many edits they wished a candidate to have in specific categories. For instance in Figure 4.

As listed in Figure 2, the measures selected capture varied roles a user might undertake on the site:

1. **Article / Template Edits** — indicative of a "Content Creator." Edits on Article Pages and Templates are called "mainspace edits" as they are seen in the most-trafficked areas of the site. This area was included because it was highly regarded by voters in the textual content of votes. Some voters went as far as to assert quotas on these edits.
2. **User Edits/User Talk Edits** — indicative of a "Communicator." Wikipedia lacks user inboxes, instead User pages and User Talk pages are the vehicle for user-to-user communication.
3. **File Edits** — indicative of a "License Servant." File edits only appear if editing the license of a file, a necessary but thankless and tedious task.
4. **Wikipedia Edits** — indicative of "Meta-Activity Participant." The "Wikipedia namespace" includes both elections and discussions of editing procedures and is invisible to common readers of the site.
5. **Talk Edits** — indicative of a "Collaborator." Talk pages are for on-topic user discussion about "mainspace edits."

```
TXT:'''Neutral'''. I'd support in a heartbeat if you had two or three times
as much experience in the Wikipedia namespace.
```

Figure 4: Excerpt from Useight's vote for AlbertHerring (2008)

5.3. Quality

Other voters prized quality over quantity. An article's quality on Wikipedia is community-determined along the following scale: stub, start, C, B, A, and Good. No records exist of users who

	Article Edits	User Edits	User Talk	File Edits	File Talk
Mean	2573.2	187.6	809.4	23.6	2.0
Std. Dev.	5556.2	765.6	2097.6	248.9	23.5
Min	0	0	0	0	0
25%	259.5	17	42	0	0
50%	1156	69	246	1	0
75%	2793.75	190.25	802.5	9	0
Max	154680	39526	66109	12574	626

	Template Edits	Template Talk	Wikipedia	Wikipedia Talk	Talk Edits
Mean	72.0	25.1	513.5	78.5	424.7
Std. Dev.	520.2	114.4	935.2	179.9	2129.8
Min	0	0	0	0	0
25%	0	0	38	2	23
50%	8	1	241	19	115
75%	36	10	634.5	76	331
Max	25787	2808	20915	3820	104667

Table 2: Distribution of Roles Assumed Feature Set

submit articles for Good consideration (as each of these rankings is determined by each WikiProject individually) and this assesment does not abide by a universal standard. However, one universal standard of quality does exist: "Featured Articles" and "Featured Lists." Every day, one Good article appears prominently on the Wikipedia homepage. Of the five million articles on English Wikipedia, nearly 4,700 have been featured (approximately $\tilde{0}.1\%$). A record does exist of users who submit content for Featured consideration (traditionally a key author). Therefore, a user “with” featured articles and lists may be considered a content creator, and a very good one at that. The distribution can be viewed in Figure 3.

```
TXT:'''S'''upport. Graham is a content beast. I think all content beasts,
especially in the field of GA/FA content, should have the option of admin
tools provided they are unlikely to abuse them. This content beast can be
trusted.
```

Figure 5: Excerpt from Jfdwolff’s posisive vote for GrahamColm (2008)

	# Featured Lists	# Featured Articles
Mean	0.1	0.2
Std. Dev.	1.0	1.2
Min	0	0
25%	0	0
50%	0	0
75%	0	0
Max	30	36

Table 3: Distribution of Quality Feature Set

5.4. Voting / Editing Habits

The distributions visible in Figure 4 and Figure 5 are two different means of representing user habits and/or behaviors in voting in Article for Deletion polls (a contentious portion of the "meta" portion of the website) and and editing habits respectively.

	Delete Votes	Speedy Delete Votes	Keep Votes	Speedy Keep Votes
Mean	56.4	2.7	1.5	0.2
Std. Dev.	62.5	4.7	3.1	0.6
Min	0	0	0	0
25%	2	0	0	0
50%	29	1	0	0
75%	100	3	2	0
Max	199	61	51	12
	Total AfD Votes	**Merge Votes**	**Redirect Votes**	**Transwiki Votes**
Mean	37.24026	1.218123	0.431523	13.110094
Std. Dev.	45.475373	2.831234	1.161471	21.891147
Min	0	0	0	0
25%	1	0	0	0
50%	16	0	0	5
75%	62	1	0	16
Max	191	34	14	192

Table 4: Distribution of Voting Habits Feature Set

6. Evaluation

Our dataset was constructed with the following paradigm: each of the 3,388 instances is a unique RfA election. An instance is "correctly classified" if its assigned label matches the election's real-

	Minor Edits	Non-Minor Edits	Reverts	Large Edits
Mean	2192.6	4532.8	200.0	1040.8
Std. Dev.	5537.9	9285.4	715.4	2618.5
Min	0	0	0	0
25%	94.8	512	3	73.8
50%	702	2114.5	25	388
75%	2147.3	5216	146	1082.3
Max	105122	204859	16278	82567

Table 5: Distribution of Edit Habits Feature Set

world outcome. A successful classifier is one which maximizes the number of correctly classified instance.

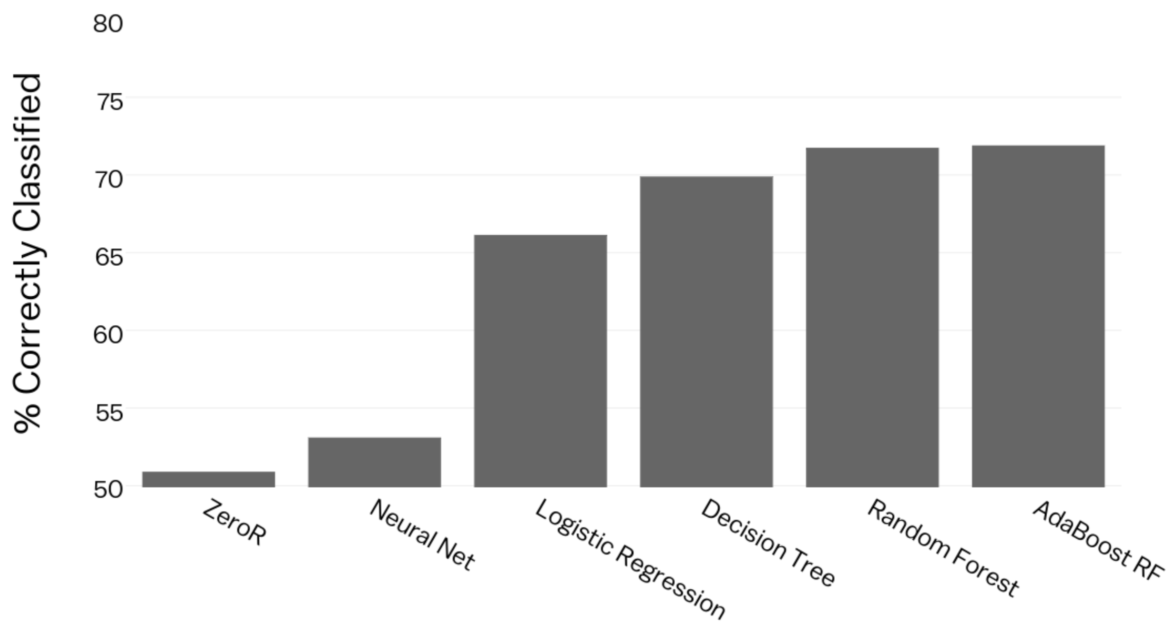
As a baseline for the evaluation of different classifiers, we use Weka's ZeroR Classifier. The ZeroR classifier labels every instance with the value of the most common class. Of the 3,388 elections, 1,726 (50.9%) were successful. Any classifier with a classification accuracy above this threshold can be said to have successfully "learned."

In the following section, we apply several different machine learning algorithms:

1. Multilayer Perceptron — a simple neural network
2. Logistic Regression — the appropriate regression for a predicting binary class
3. J48 Decision Tree — the open-source implementation of the C4.5 decision tree algorithm. Listed by Wu et. all [19] in 2008 to be the most influential data mining algorithm.
4. Random Forest — a randomized variation on the decision tree.
5. AdaBoosted Random Forest — an attempt at improving the accuracy of the Random Forest algorithm which consistently out-performed the other algorithms.

6.1. General Activity

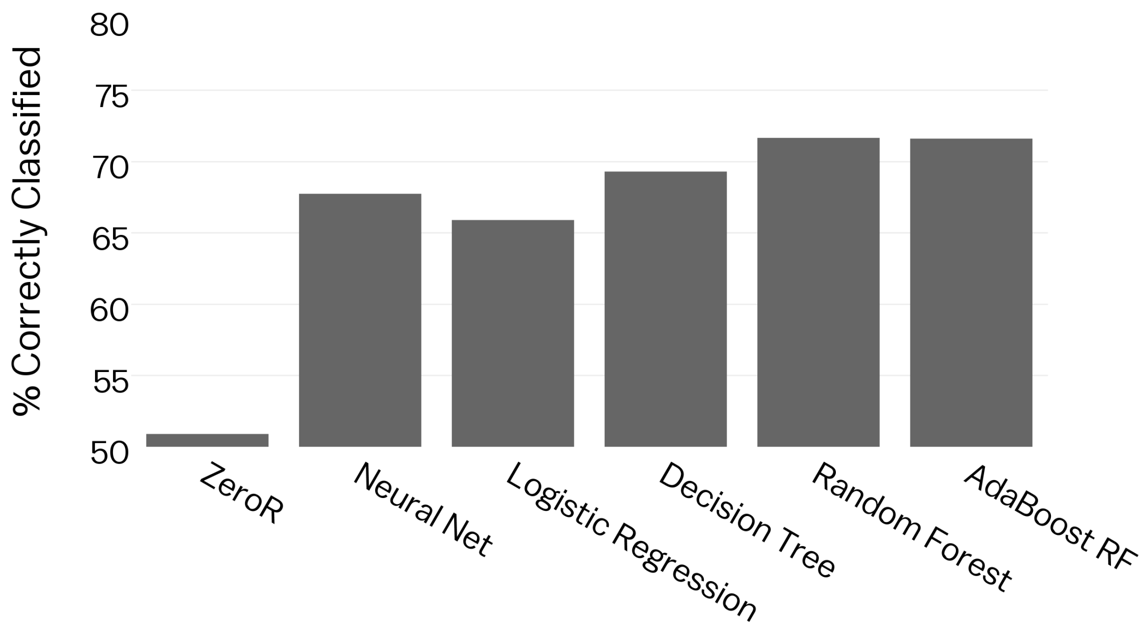
Figure 6: Algorithm Performance on General Activity Features



Algorithm	**# Correctly Classified**	**% of Classifications Correct**
ZeroR	1726	50.9 %
Neural Net	1800	53.13 %
Logistic Regression	2241	66.15%
Decision Tree	2369	69.92 %
Random Forest	2431	71.75 %
AdaBoost Random Forest	2437	71.93 %

6.2. Roles Assumed

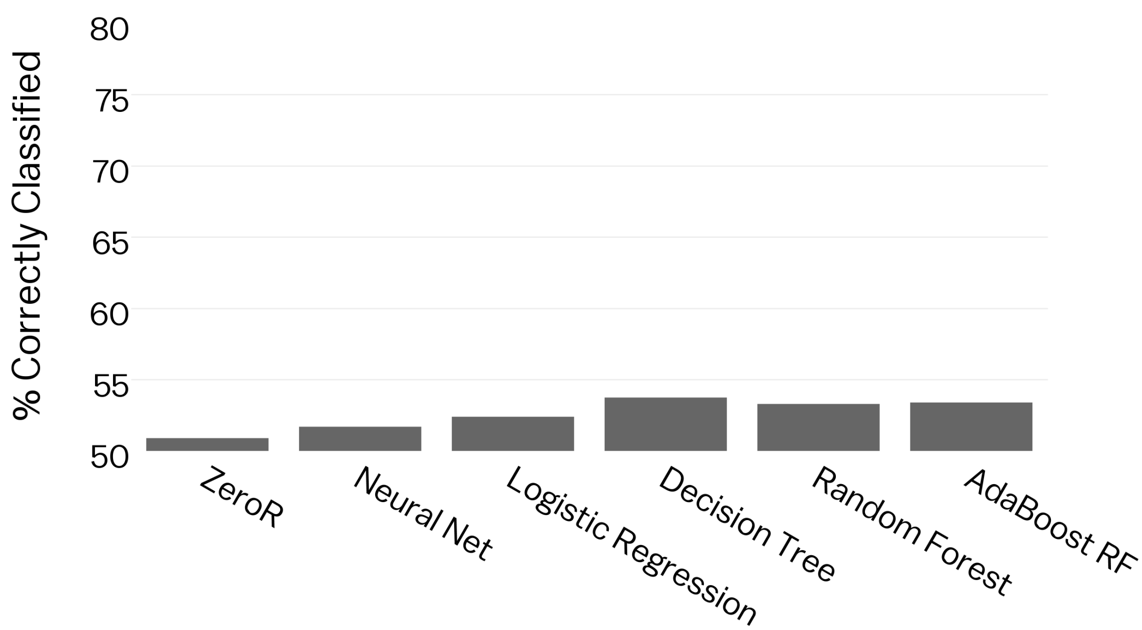
Figure 7: Algorithm Performance on Roles Assumed Features



Algorithm	# Correctly Classified	% of Classifications Correct
ZeroR	1726	50.9 %
Neural Net	2295	67.74 %
Logistic Regression	2233	65.90%
Decision Tree	2348	69.30 %
Random Forest	2428	71.66 %
AdaBoost Random Forest	2426	71.61 %

6.3. Quality

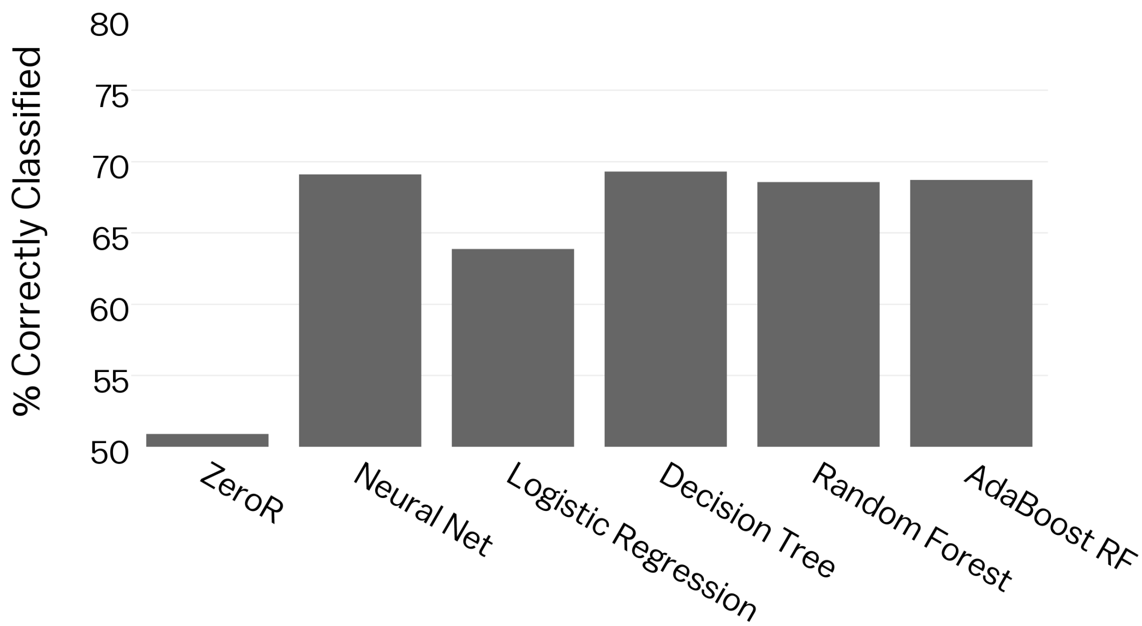
Figure 8: Algorithm Performance on Quality Features



Algorithm	# Correctly Classified	% of Classifications Correct
ZeroR	1726	50.9 %
Neural Net	1752	51.7 %
Logistic Regression	1775	52.40 %
Decision Tree	1821	53.75 %
Random Forest	1807	53.33 %
AdaBoost Random Forest	1809	53.40 %

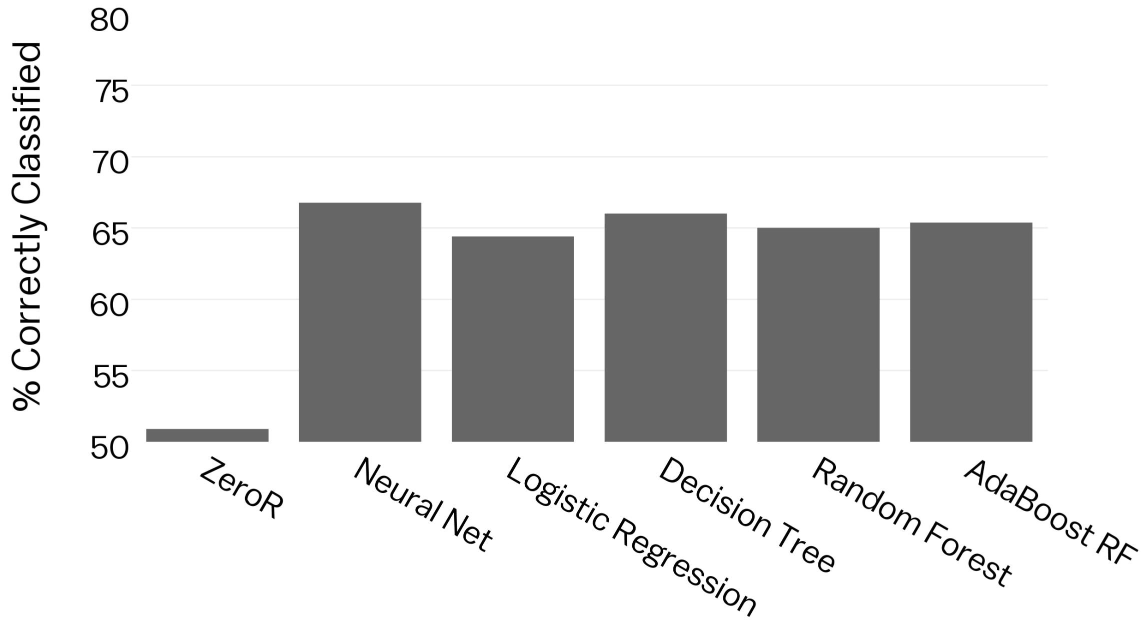
6.4. Voting / Editing Habits

Figure 9: Algorithm Performance on Editing Behavior Features



Algorithm	**# Correctly Classified**	**% of Classifications Correct**
ZeroR	1726	50.9 %
Neural Net	2341	69.1 %
Logistic Regression	2164	63.87 %
Decision Tree	2348	69.3 %
Random Forest	2323	68.56 %
AdaBoost Random Forest	2328	68.71 %

Figure 10: Algorithm Performance on Voting Behavior Features



Algorithm	# Correctly Classified	% of Classifications Correct
ZeroR	1726	50.9 %
Neural Net	2262	66.76 %
Logistic Regression	2183	64.40 %
Decision Tree	2235	66.0 %
Random Forest	2202	65.0 %
AdaBoost Random Forest	2215	65.38 %

7. Summary

7.1. Conclusions

Although no classifier outperformed all the others, the logistic regression classifier performed (or very nearly performed) the worst in every feature set. Of the five algorithms, the logistic regression classifier was the only parameterized classifier as well as the only linear classifier. This failure to perform is telling: the data is likely to not be linearly separable.

To explain further, we examine the set of algorithms that performed well: the decision tree and its variants. A decision tree is a set of consecutive decisions, moving from root to leaves, resulting in an classification. This structure distinct resembles the human calculus in the RfA voting process. As exhibited in excerpts from the textual content of votes (e.g. Figure 3), voters often have strict numerical criteria (if not quotas) they desire in a candidate. That most important characteristic, e.g. "mainspace edits," would be the root of that voters decision tree.

Despite the parallels of the voter decision-making process to the decision tree classifier and its variants, these classifiers were still never able to classify more than 72 % of users correctly. As discussed in greater detail in the following section, there are many intangibles and non-quantifiable factors influencing the decision such as user disposition or unpopular opinions.

As for the feature sets, general measures of activity and roles assumed outperformed the two behavioral measures which outperformed the quality measures. Returning to the distribution of the quality data, which was very sparse, it is logical that those two features alone could distinguish many users — especially considering most successful candidates did not featured content. See Figure ?? for an sample decision tree which merely divided users into two classes: "has featured content" and "does not have featured content."

Figure 11: Sample Decision Tree for Quality Features

```
featured_articles <= 0
|   featured_lists <= 0: 0 (3049.0/1477.0)
|   featured_lists > 0: 1 (52.0/20.0)
featured_articles > 0: 1 (287.0/70.0)
```

Fittingly, the features that were most successful at distinguishing between successful candidates for adminship and unsuccessful ones were those features most central to the mission of Wikipedia: contribution. These features (e.g. total edits, article edits, talk edits) were prone to influence by scale, thus emphasizing the accruement of a user's activity over time. As the activity grows more diverse and more substantial, the more likely a user is to be trusted as an elite member of the community and thus rewarded for the time spent as an editor.

7.2. Limitations

The scope of the problem was inherently limiting. Requests for Adminship can be too bureaucratic and infrequent to produce a robust statistical sample:

1. As discussed, Wikipedia has a limited set of power users (approx. 12,000) which at any time is an unofficial upper-bound for voters and candidates.
2. As a feature of the election, each candidate answers several survey questions to help voters gauge the candidate's beliefs. When a candidate gaffes or expresses an unpopular opinion and **this** is the reason for an unsuccessful election — statistical prediction is impotent in that case.
3. In general, user activity makes quantifying a user's disposition difficult, i.e. distinguishing a kind prolific user from a rude prolific user is difficult without out-of-scope textual analysis.

7.3. Future Work

As most of the features examined by this report are general quantitative factors from a user's history, there is still much room to examine more qualitative factors that enter into the human decision-making calculus.

1. User disposition was difficult to quantify, AfD voting records were a sparse indicator as they lacked ground truth. Any sort of semantic analysis of user comments looking for impoliteness, offensive behavior, and/or arrogance could be very effective at rooting out prolific, yet non-admin-like individuals.
2. A centrality feature, i.e. whether a user is a "maverick" spending a majority of their time building new content or a "populist" spending time bolstering already substantial articles, would be an interesting quality unlike any able to be examined in this report.

8. Acknowledgments

A sincere thanks must be extended to Professor Thomas Funkhouser without whom this report (or, frankly, any report at all) exists.

9. Honor Code

I pledge my Honor that I have not violated the Honor Code during the writing of this paper. /s/

Elizabeth Bradley

References

[1] E. Bakshy *et al.*, "Everyone's an influencer: quantifying influence on twitter," in *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 2011, pp. 65–74.

[2] Docs.python-requests.org, "Requests: Http for humans — requests 2.9.1 documentation," 2016. Available: http://docs.python-requests.org/en/latest/

[3] Docs.seleniumhq.org, "Selenium - web browser automation," 2016. Available: http://docs.seleniumhq.org/

[4] E. U. M. B. P. F. Eleonora Ciceri, Roman Fedorow, "Assessing online media content trustworthiness, relevance and influence: an introductory survey," in *1st International Workshop on Knowledge Discovery on the Web*, 2015.

[5] E. Katz and P. F. Lazarsfeld, *Personal Influence, The part played by people in the flow of mass communications*. Transaction Publishers, 1955.

[6] A. Kittur *et al.*, "Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie," *World wide web*, vol. 1, no. 2, p. 19, 2007.

[7] G. H. B. P. P. R. I. H. W. Mark Hall, Eibe Frank, "The weka data mining software: An update," in *SIGKDD Explorations, Volume 11, Issue 1.*, 2009.

[8] Mediawiki.org, "Api:usercontribs - mediawiki," 2016. Available: https://www.mediawiki.org/wiki/API:Usercontribs

[9] Mediawiki.org, "Api:users - mediawiki," 2016. Available: https://www.mediawiki.org/wiki/API:Users

[10] Mediawiki.org, "Mediawiki," 2016. Available: https://www.mediawiki.org/wiki/MediaWiki

[11] K. Panciera, A. Halfaker, and L. Terveen, "Wikipedians are born, not made: a study of power editors on wikipedia," in *Proceedings of the ACM 2009 international conference on Supporting group work*. ACM, 2009, pp. 51–60.

[12] Phantomjs.org, "Phantomjs | phantomjs," 2016. Available: http://phantomjs.org/

[13] S. N. A. Project, "Snap: Network datasets: Wikipedia requests for adminship (with text)," 2016. Available: https://snap.stanford.edu/data/wiki-RfA.html

[14] J. L. Robert West, Hristo S. Paskov and C. Potts, "Exploiting social network structure for person-to-person sentiment analysis," in *Transactions of the Association of Computational Linguistics*, 2014.

[15] Tools.wmflabs.org, "Afd stats," 2016. Available: https://tools.wmflabs.org/afdstats/

[16] Wikipedia, "Wikipedia:size of wikipedia — Wikipedia, the free encyclopedia," 2015, [Online; accessed 04-January-2015]. Available: https://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia

[17] Wikipedia, "Wikipedia:wikipedians — Wikipedia, the free encyclopedia," 2015, [Online; accessed 04-January-2015]. Available: https://en.wikipedia.org/wiki/Wikipedia:Wikipedians

[18] Wikipedia, "Wikipedia:wikiproject edit counters — Wikipedia, the free encyclopedia," 2015, [Online; accessed 04-January-2015]. Available: https://en.wikipedia.org/wiki/Wikipedia:WikiProject_edit_counters

[19] X. Wu *et al.*, "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.