Analysis of goodreads.com

-Ankita Sawant

-Shradha Nayak

-Muskaan Mulchandani

MOTIVATION

Encouraging authors to write books which will be liked by goodread readers

GOAL OF OUR PROJECT

- Predicting if reviewers will like a book or not.

ASSUMPTION

- Ratings given by reviewers for books, their genre, book authors, past history of the reviewers and the authors are related to them liking a particular book.

DATA COLLECTION

- The website that we have crawled for doing the prediction is https://www.goodreads.com

Features scraped from the website

- Avg Book Rating
- Book published date
- Reviewer Name
- Reviewer Rating
- Book added date
- Genre
- Book
- ISBN
- Author
- Book read date
- Count of people interaction
- Review

Features used for prediction

- Genre
- Reviewer Ratings
- Past Count Genre
- Past Rating Genre
- Book Length
- Past Author Book count
- Past Author Book rating

TOOLS

- Python
- BeautifulSoup
- Chrome Web Driver
- Matplotlib and Pandas for plotting graphs

FILTERING AND CLEANING DATA

- We removed the books which were not in English language.(As no reviews were found)
- We removed the reviews which were not in English Language.
- We removed the records which did not have any reviewer rating
- We made the date format uniform

Book read date
Jan-16
Feb25,2015
Feb27,2016
Nov20,2014
Sep19,2015
Apr06,2014
Mar-16

WEBSITE USED FOR SCRAPING

karen > Books

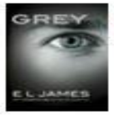
Search and add books | Compare Books | Settings | Stats | Print

bookshelves

- all (7576)
- read (4055)
- currently-reading (3)
- to-read (1695)
- amd (470)
- ask-greg (288)
- at-my-desk (74)
- books-about-greg-s-mom (4)
- books-i-hate-more-than-most-other-t (1)
- buy-for-me-thanks (103)
- ceci-n-est-ce-pas-un-compte-rendu (18)
- continuing-ed-cookbooks (3)
- do-i-own-you (20)
- holy-grail-unicorn-tamerlane (90)
- i-am-the-one-percent (1)
- in-the-pipeline (15)
- nook-tbr (171)
- oh-dear (8)
- released-into-the-wild (127)
- romance-covers-i-have-loved (30)
- soon-to-buy (175)
- to-get-from-liberry (225)

- aaaahhrrrtt (128)
- americas-favorite-president (7)
- and-so-this-is-grad-school (63)
- animal-butts (3)
- appalachian-noir-southern-gothic (14)
- awkward-age-books-forgotten-til-now (113)
- babys-first-manga (15)

« previous 1 2 3 4 5 6 7 8 9 ... 378 379 next »

cover	title	author	avg rating	rating	my rating	review	date read	date added	
	Divergent (Divergent, #1)	Roth, Veronica *	4.28	★★★★★	add to shelves	i need to make something perfectly clear. i am well aware that i gave 4 stars to Daughter of Smoke and Bone. and i am giving 5 stars to this one. the ...more	Jun 18, 2011	Jun 17, 2011	view
	Life and Death: Twilight Reimagined	Meyer, Stephenie	3.34		add to shelves	e.l. james runs out of ideas and desperately rewrites Fifty Shades of Grey from christian's perspective. stephenie meyers "celebrates" Twilight by swap ...more	not set	Oct 06, 2015	view
	Infinite Jest	Wallace, David Foster	4.32	★★★★★	add to shelves	this book... i think it is time to write a proper review for this book, as it is one of my all-time favorites and deserves way more than two words. bac ...more	not set	May 08, 2008	view
	Grey (Fifty Shades, #4)	James, E.L. *	3.75		add to shelves	my inner goddess is screaming out the safe word, but e.l. james is acting like we never set ground rules. rutabaga RUTABAGA!! **RUTABAGA!!!!!**	not set	Jun 02, 2015	view

STATISTICS OF DATA SCRAPED

Attributes	Count
Total number of rows scraped	1760
After filtering(Removing rows with no rating)	1671
Original dataset number of features scraped	12
Final dataset number of features	57
Total number of reviewers scraped	98
Total number of reviews scraped per reviewer	Approx 20 (First page of each reviewer)

STATISTICS OF DATA SCRAPED

Raw Features

Groups	Features
Genre	Genre Type
Author	Author Name
Book	1)Book Name
	2)Book added date
	3)Book published date
	4)Book read date
	5)ISBN
Reviews	1)Review
	2)Reviewer Name
	3)Reviewer Rating
	4)Average book rating
	5)Count of people interactio with reviewer

Processed Features

Groups	Features
Genre	1) Different genres: 50
	2)Past Count of book reviews based on Genre
	3)Past Rating of book reviews based on Genre
Author	1)Past Count of book reviews based on Author
	2)Past Rating of book reviews based on Author
Book	1)Length of the book
Reviews	1)Length of reviews
	2)Reviewer ratings

ORIGINAL DATASET

A	B	C	D	E	F	G	H	I	J	K	L
Author	Average book Rating	Book	Book added date	Book published date	Book read date	Count of people interaction	Genre	ISBN	Review	Reviewer Name	Reviewer Rating
Fisher, Tarryn	4.18	F*ck Love	Jul08,2015	Dec31,2015	Jan-16	125	Romance	-	5 + STARS â€	9916227-jennifer-kyle	it was amazing
Kennedy, Elle	4.33	The Deal(Off-Campus, #1)	Feb24,2015	Feb24,2015	Feb25,2015	112	New Adult	-	5 WINNING S	9916227-jennifer-kyle	it was amazing
Webster, K.	4.23	This is War, Baby(War & Peace, #1)	Feb26,2016	Feb29,2016	Feb27,2016	116	Dark	-	5 + I WAR S	9916227-jennifer-kyle	it was amazing
Harmon, Amy	4.39	The Law of Moses(The Law of Moses, #1)	Sep24,2014	Nov18,2014	Nov20,2014	119	Romance	-	5 *Greats* S	9916227-jennifer-kyle	it was amazing
Cherry, Brittainy C.	4.29	The Air He Breathes(Elements, #1)	Jul31,2015	Sep25,2015	Sep19,2015	102	Romance	-	5 + STARS "H	9916227-jennifer-kyle	it was amazing
Fisher, Tarryn	4.09	Mud Vein	Jul28,2013	Mar08,2014	Apr06,2014	110	Dark	-	The writing i	9916227-jennifer-kyle	it was ok
Zapata, Mariana	4.36	The Wall of Winnipeg and Me	Dec26,2015	Feb29,2016	Mar-16	112	Romance	-	5 "TEAM GR	9916227-jennifer-kyle	it was amazing
Williams, Nicole	4.08	Collared	Feb19,2016	Mar22,2016	Mar03,2016	118	Romance	-	5 + STARS â€	9916227-jennifer-kyle	it was amazing
Cole, Tillie	4.39	A Thousand Boy Kisses	Jul16,2015	Mar15,2016	Mar15,2016	102	Young Adult	-	This one was	9916227-jennifer-kyle	it was ok
Walters, A. Meredith	4.11	One Day Soon(One Day Soon, #1)	Nov04,2015	Feb18,2016	Feb15,2016	99	Romance	-	5 + STARSItâ	9916227-jennifer-kyle	it was amazing
Hoover, Colleen	4.27	Confess	Jan29,2015	Mar10,2015	Jan 31,2015	84	Romance	-	3.75 ~ 4 OM	9916227-jennifer-kyle	really liked it
Holden, Kim	4.25	So Much More	Feb27,2016	Mar29,2016	Mar29,2016	113	Romance	-	5 â€œWEâ€	9916227-jennifer-kyle	it was amazing
Torre, Alessandra	4.14	Hollywood Dirt(Hollywood Dirt, #1)	Jun27,2015	Sep07,2015	Sep03,2015	89	Romance	-	4.5 Coca Col	9916227-jennifer-kyle	it was amazing
Frazier, T.M.	4.25	King(King, #1)	Jan22,2014	Jun15,2015	Jun15,2015	101	Dark	-	I have been a	9916227-jennifer-kyle	liked it
Ward, Penelope	4.03	Sins of Sevin	Mar23,2015	Sep14,2015	Sep21,2015	112	Romance	-	2.75 Stars "H	9916227-jennifer-kyle	liked it
Martinez, Aly	4.1	The Fall Up(The Fall Up, #1)	Aug12,2015	Oct26,2015	Oct13,2015	95	Romance	1518711391	5 "And fallin	9916227-jennifer-kyle	it was amazing
Lilley, R.K.	4.28	Breaking Him(Love is War, #1)	Aug25,2015	Oct03,2015	Oct14,2015	84	Romance	-	4.5 STARS â€	9916227-jennifer-kyle	it was amazing
Harmon, Amy	4.4	Making Faces	Oct19,2013	Oct12,2013	Oct21,2013	69	Romance	-	This isnâ€™t	9916227-jennifer-kyle	it was amazing
Sheridan, Mia	3.76	Midnight Lily	Mar02,2016	Mar01,2016	Mar02,2016	101	Romance	-	3.5 StarsSom	9916227-jennifer-kyle	liked it
Shen, L.J.	4.14	Sparrow	Feb27,2016	Mar01,2016	Mar09,2016	98	Romance	-	4.5 Stars â€	9916227-jennifer-kyle	really liked it
Moriarty, Liane	4.16	Big Little Lies	May28,2014	Jul29,2014	Jul20,2014	20	Fiction	399167064	This one was	4159922-diane-s	really liked it
Kalanithi, Paul	4.37	When Breath Becomes Air	Jan14,2016	Jan12,2016	Jan31,2016	43	Nonfiction	-	As I finished	4159922-diane-s	really liked it
Doerr, Anthony	4.31	All the Light We Cannot See	Dec11,2013	May-14	Apr03,2014	66	Historical	1476746583	For me, this	4159922-diane-s	it was amazing

FINAL DATASET USED FOR PREDICTION

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Action	Adult	Adult Fiction	Adventure	Autobiography	Biography	Childrens	Classics	Comics	Contemporary	Cultural	Dark	Disability	Economics	Erotica	Fantasy	Fiction	Glbt	Historical	Historical I	History	Horror	Humor	Literature	Magical Realism
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU
Media Tie	Military History	Music	Mystery	New Adult	Nonfiction	Paranormal	Philosophy	Plays	Poetry	Politics	Race	Realistic Fiction	Romance	Science	Science Fiction	Sequential	Short Stories	Social Issues	Sports and Games	Suspense	Thriller
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE
War	Western	Young Adult	Length_Of_Review	Reviewer_Ratings	Past_Count_Genre	Past_Rating_Genre	Book_Length	Past_Author_Book_Count	Past_Author_Book_Rating
0	0	0	150	1	0	0	28	0	0
0	0	0	150	1	1	1	27	0	0
0	0	0	150	3	2	1	45	0	0
0	0	0	152	1	3	1.67	36	0	0
0	0	0	150	3	4	1.5	23	0	0
0	0	0	150	3	5	1.8	21	0	0
0	0	0	150	3	6	2	4	0	0
0	0	0	154	3	7	2.14	15	0	0
0	0	0	150	2	8	2.25	15	0	0
0	0	0	156	3	9	2.22	24	0	0
0	0	0	150	3	10	2.3	10	0	0
0	0	0	150	3	11	2.36	14	0	0
0	0	0	150	3	12	2.42	45	1	3
0	0	0	150	1	13	2.46	55	0	0
0	0	0	150	3	14	2.36	13	0	0
0	0	0	150	3	15	2.4	18	0	0
0	0	0	152	3	0	0	49	0	0
0	0	0	150	3	0	0	16	0	0
0	0	0	150	3	1	3	16	1	3
0	0	0	150	3	0	0	6	0	0
0	0	0	149	1	1	3	36	0	0
0	0	0	150	1	2	2	36	1	1
0	0	0	152	3	3	1.67	12	0	0
0	0	0	150	1	4	2	17	0	0
0	0	0	150	3	5	1.8	38	0	0

FEATURE TYPE

- Author → Text
- Avg book Rating→ Float
- Book → Alpha Numeric
- Book added date→ Date
- Book published date → Date
- Book read date →Date
- Count of people interaction → Numeric
- ISBN → Alpha Numeric
- Review→ Text
- Reviewer Name→ Text

Reviewer Rating →Numeric

Genre → Categorical

Past Count of Genre→ Numeric

Past Rating of Genre → Float

Length Of Review → Numeric

Book Length → Numeric

Past Author Book Count → Numeric

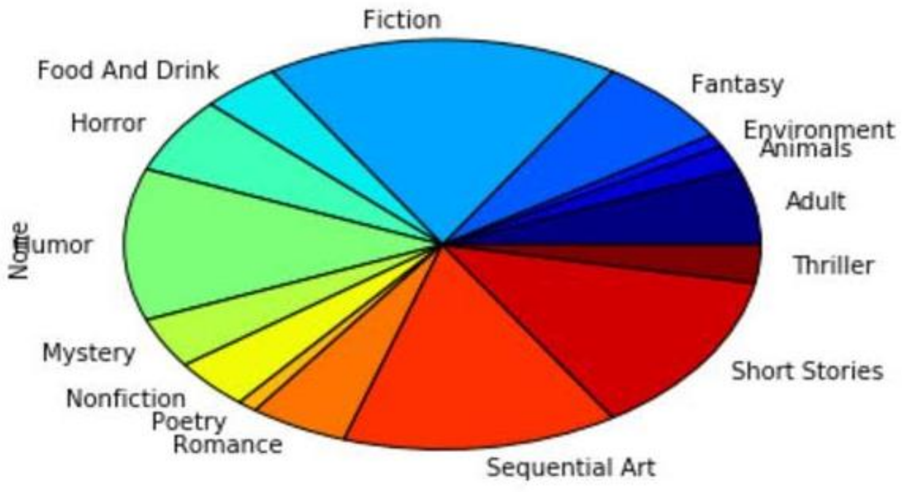
Past Author Book Rating →Float

DISTRIBUTION OF FEATURES

Distribution of Genre according to mid term data set

```
data.groupby('Genre').size().plot(kind='pie', colormap='jet' )
```

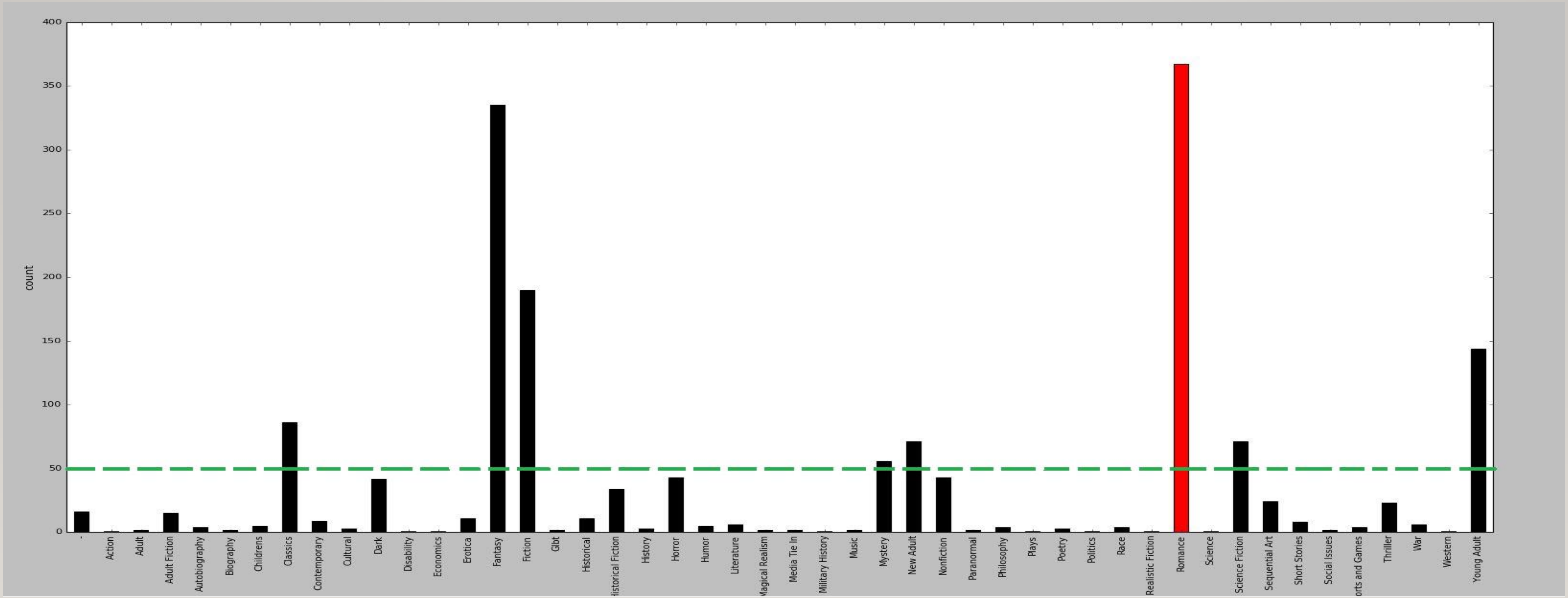
```
<matplotlib.axes._subplots.AxesSubplot at 0xe683cc0>
```



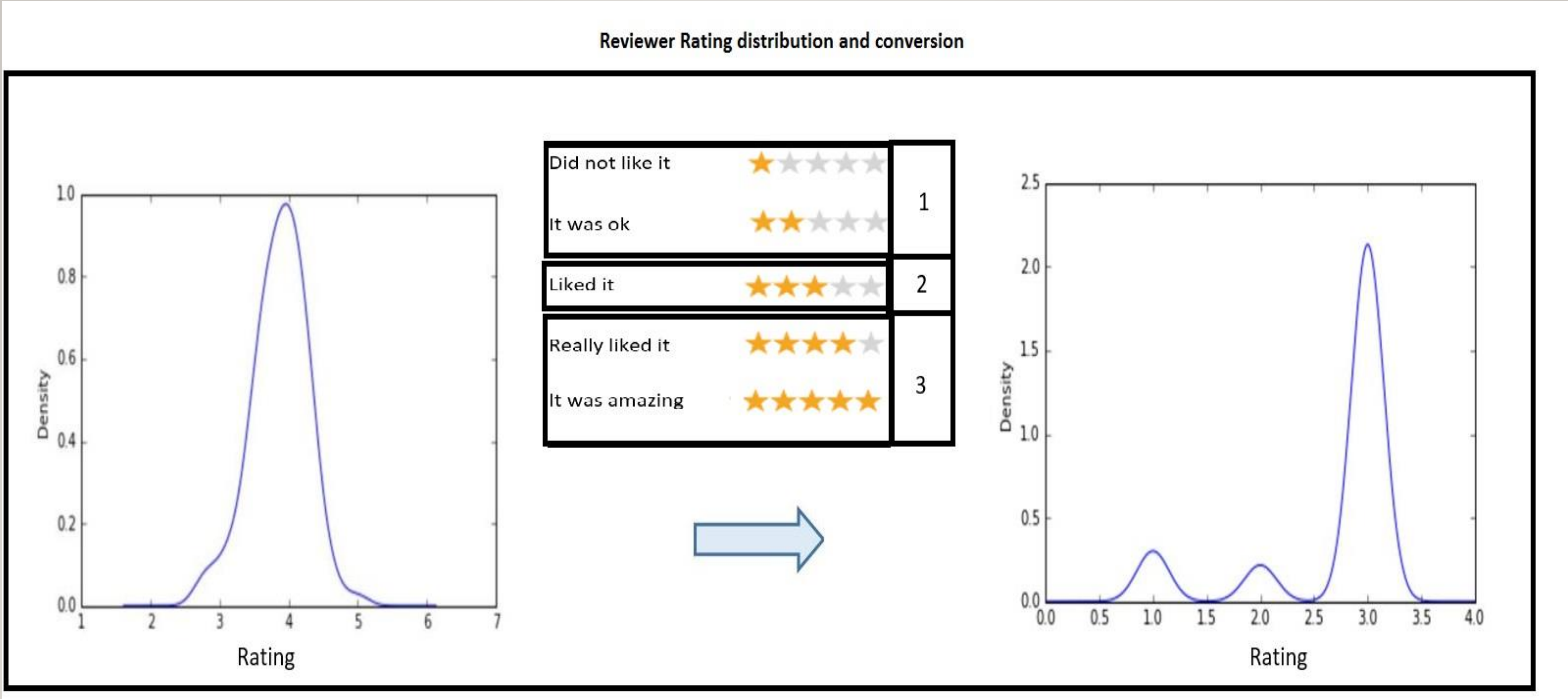
Genre

Distribution of Genre according to final data set

- The Genres: Romance, Fantasy and Fiction are the most prominent in our data set.

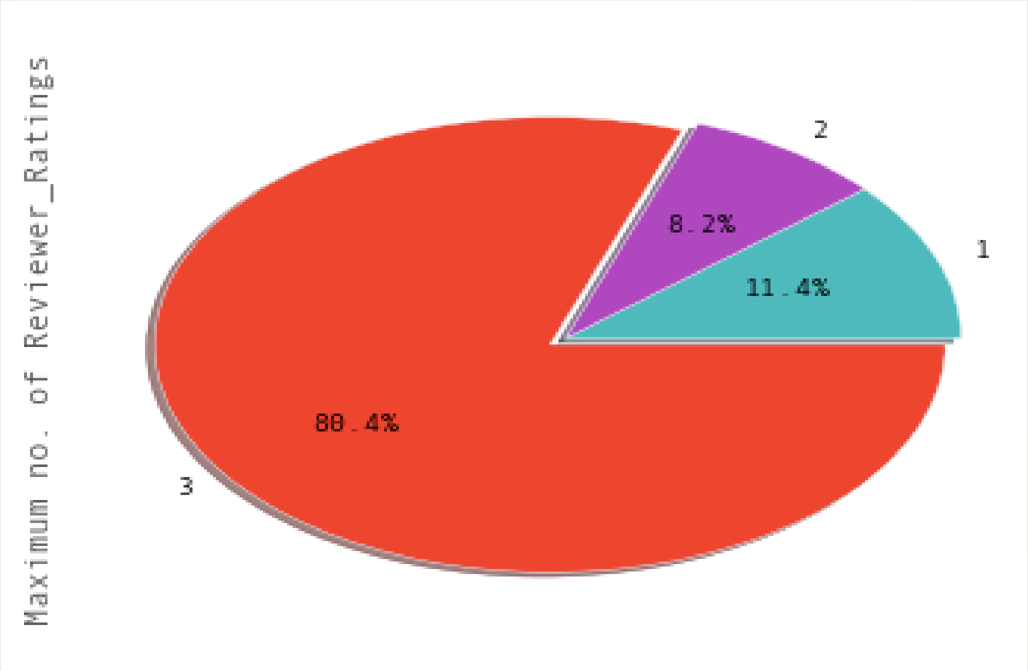


Genre



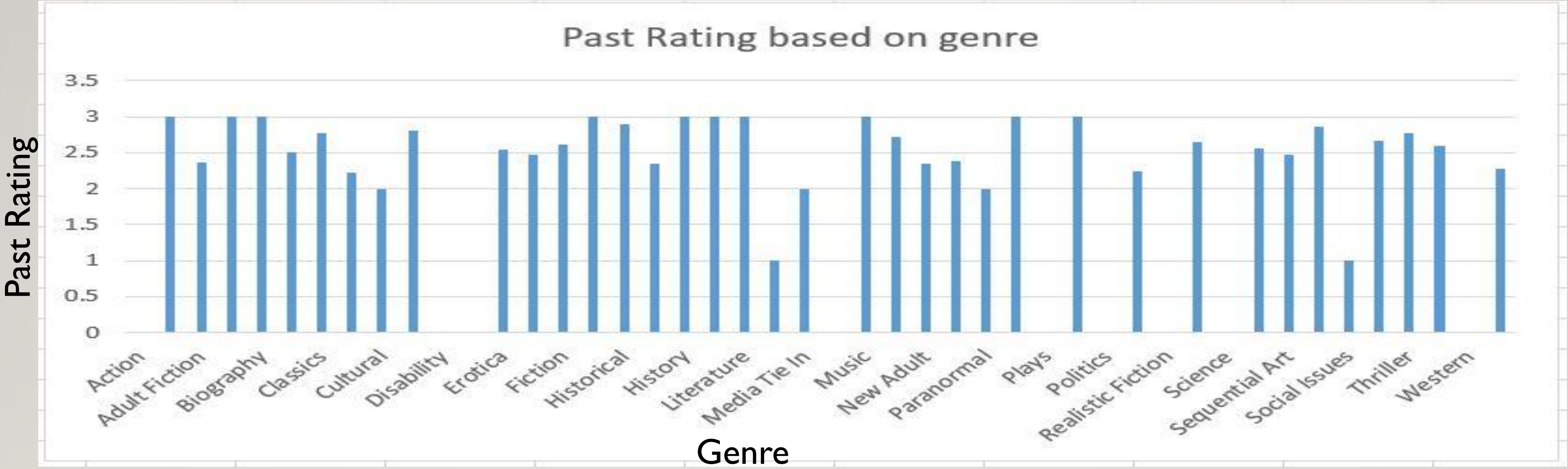
- The Reviewer Ratings were in the range 1 to 5. We converted them to 1,2,3.

Distribution of Reviewer Ratings



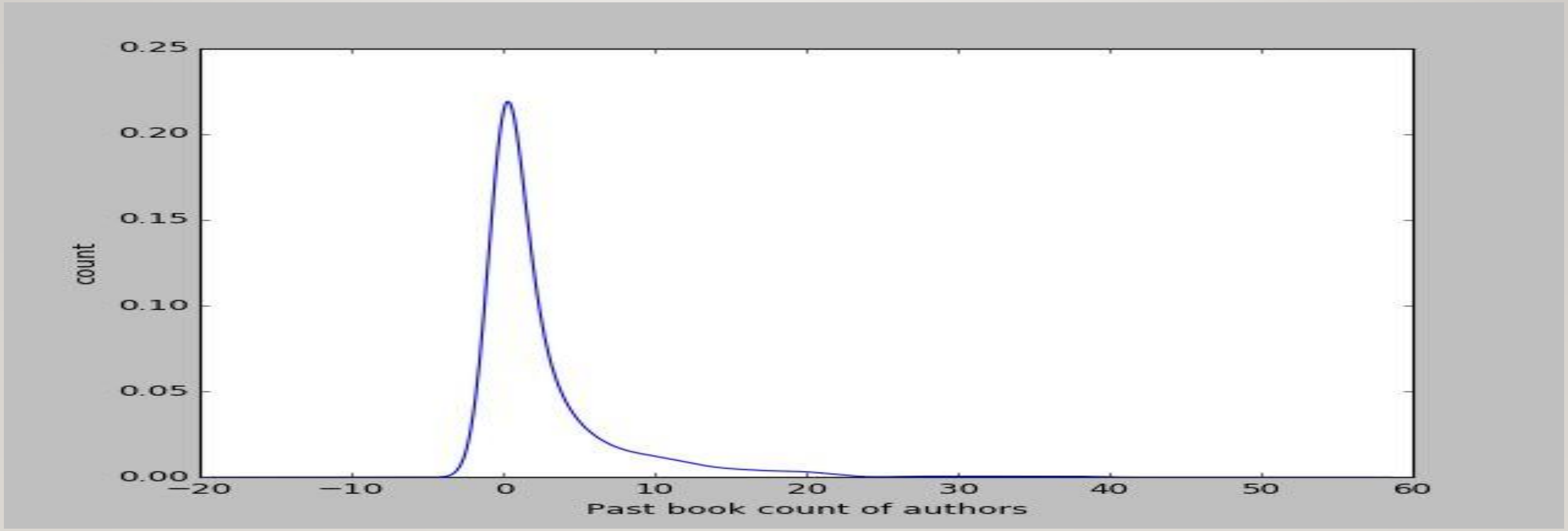
- Majority of the reviewer ratings were 3.

Distribution of Past Rating according to Genre



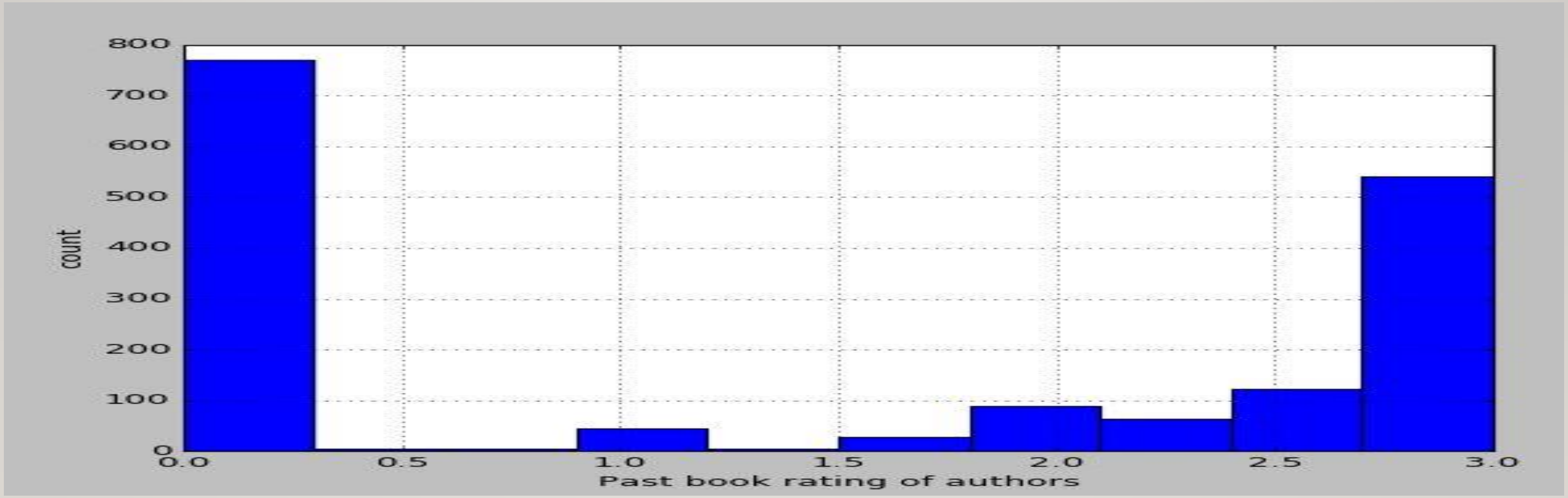
- We calculated the past rating for each record based on the ratings of the books read before that book and belonging to the same genre.

Distribution of Past book count of authors



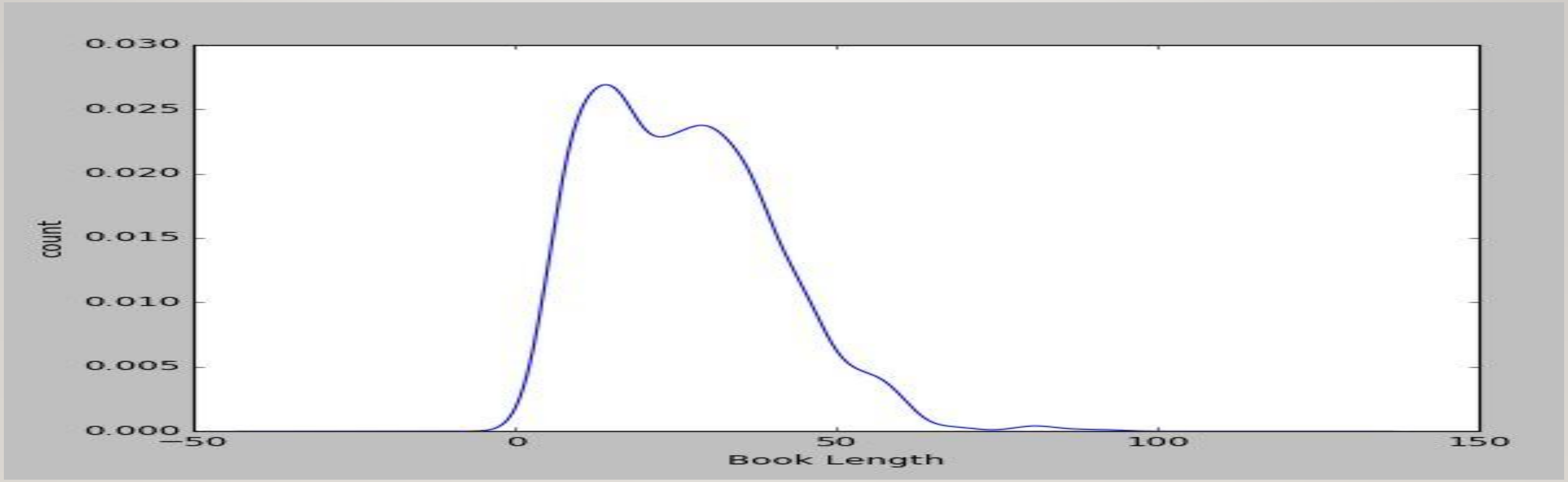
- We calculated the past book count for each record based on the books read before that book and belonging to the same author.

Distribution of Past book rating of authors



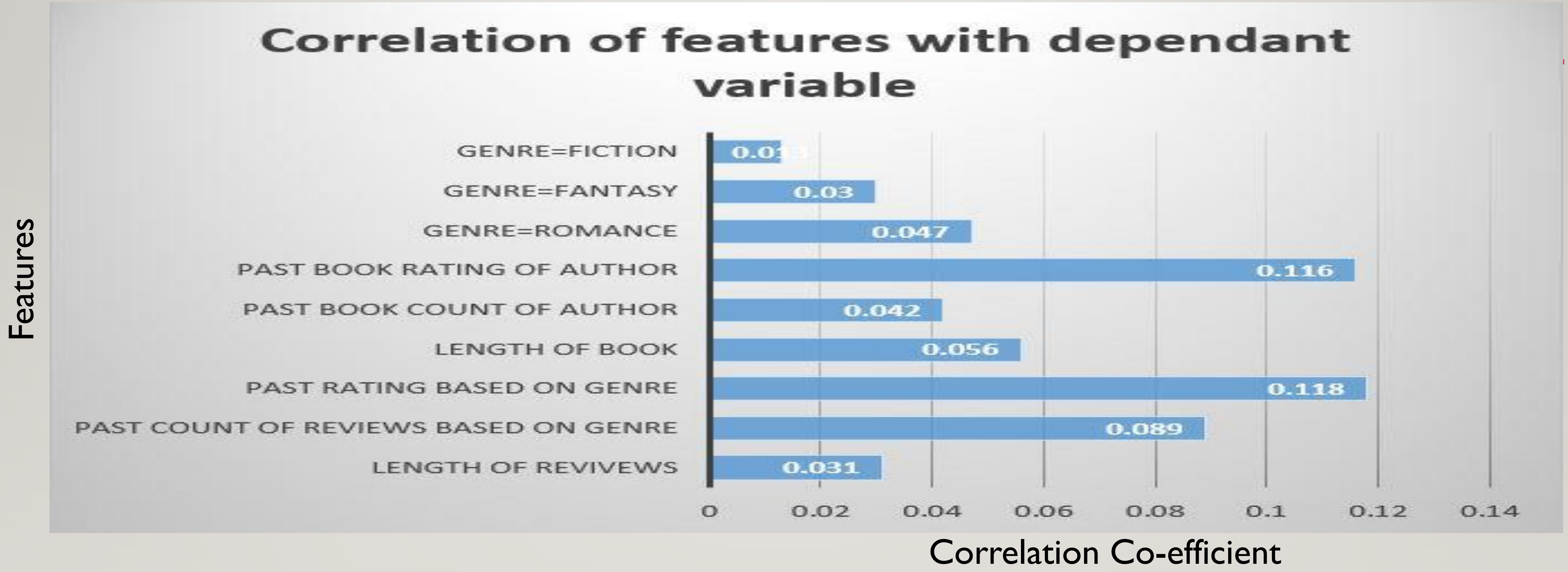
- We calculated the past rating for each record based on the ratings of the books read before that book and belonging to the same author. Most of the past book ratings were 0 as the authors did not repeat much in our dataset.

Distribution of Book Title Length



- We calculated the book title length for each record. We considered this feature on the assumption that sequels would have the same book title length.

FEATURES CORRELATION



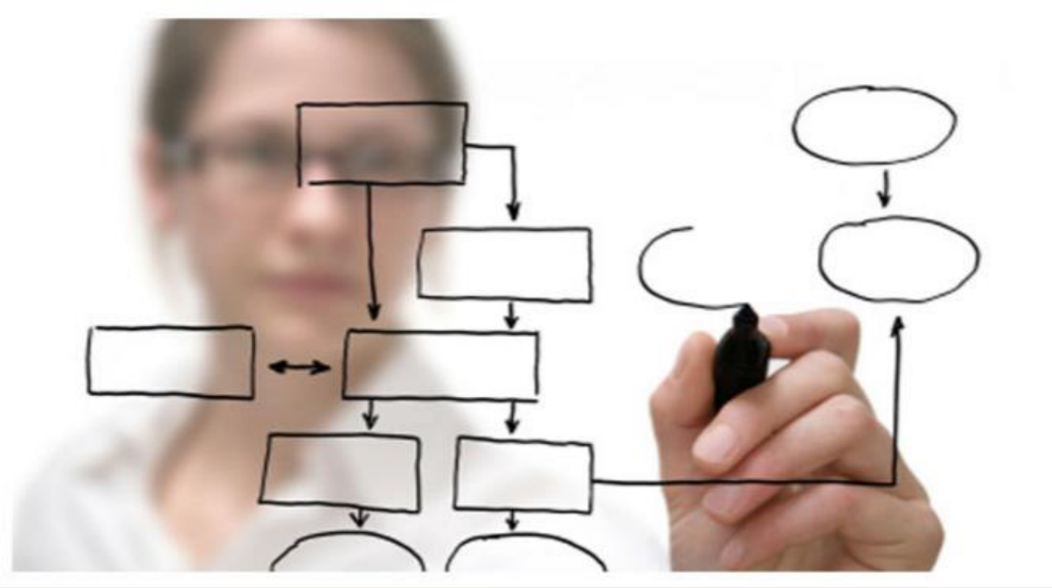
- The ‘Past book rating of author’ and ‘Past rating based on genre’ had a high correlation with our dependent variable ‘Reviewer Ratings’.

METHODS TO ACHIEVE OUR GOAL

After website scraping and data cleaning we used the following:

- Knn
- Logistic regression
- Naive Bayes
- Random Forest

We also used Voting Classifier on our 3 best algorithms.



PRE-PROCESSING

Making the dataset random

```
df1=csv.reindex(np.random.permutation(csv.index))
```

Training and Test data set

```
df1
count=len(csv)
trainCount=int((70*count)/100)
testCount=(count-trainCount)
rev_train=df1[0:trainCount]
rev_test=df1[trainCount+1:count]
```

Normalisation

```
import pandas as pd
def norm(x):
    print x
    df_max_aa=x.max()
    df_mean_aa=x.mean()
    df_min_aa=x.min()
    x=(x-df_min_aa)/(df_max_aa-df_min_aa)
    return x
```

ACCURACY OF MODELS

KNN

```
KNN=KNeighborsClassifier(5)
KNN.fit(rev_train_Norm.astype(int),labels_train.astype(int))
predicted=KNN.predict(rev_test_Norm)
print accuracy_score(predicted,labels_test)
```

```
0.802395209581
```

```
{'n_neighbors': 1, 'weights': 'uniform'} 0.51497005988
{'n_neighbors': 1, 'weights': 'distance'} 0.51497005988
{'n_neighbors': 3, 'weights': 'uniform'} 0.526946107784
{'n_neighbors': 3, 'weights': 'distance'} 0.526946107784
{'n_neighbors': 5, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 5, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 7, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 7, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 9, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 9, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 11, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 11, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 13, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 13, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 15, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 15, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 17, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 17, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 19, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 19, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 21, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 21, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 23, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 23, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 25, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 25, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 27, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 27, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 29, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 29, 'weights': 'distance'} 0.800684345595
{'n_neighbors': 31, 'weights': 'uniform'} 0.800684345595
{'n_neighbors': 31, 'weights': 'distance'} 0.800684345595

Best parameters {'n_neighbors': 5, 'weights': 'uniform'}
0.812375249501
```

Logistic Regression

```
LR=LogisticRegression ()
LR.fit(rev_train,labels_train)
#use the classifier to predict
predicted=LR.predict(rev_test)
#print the accuracy
print accuracy_score(predicted,labels_test)
```

```
0.818363273453
```

Naïve Bayes

```
rev_train_Norm = sklearn.preprocessing.normalize(rev_train, norm='l2')
rev_test_Norm=sklearn.preprocessing.normalize(rev_test, norm='l2')
clf3 = MultinomialNB()
clf3.fit(rev_train_Norm,labels_train)
pred=clf3.predict(rev_test_Norm)
print accuracy_score(pred,labels_test)
```

```
0.796407185629
```

Random Forest

```
clf3 = RandomForestClassifier(n_estimators=100)
clf3.fit(rev_train_Norm,labels_train)
pred=clf3.predict(rev_test_Norm)
print accuracy_score(pred,labels_test)
```

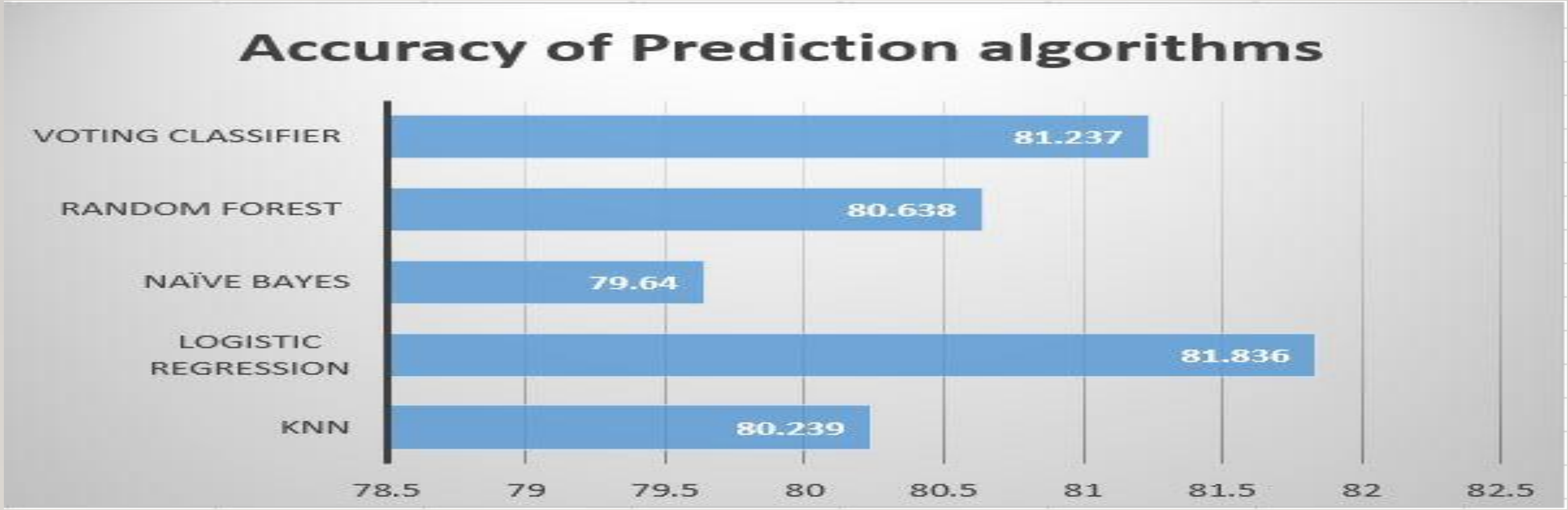
```
0.806387225549
```

Voting Classifier

```
clf1 = LogisticRegression()
clf2 = KNeighborsClassifier(5)
clf3 =RandomForestClassifier(n_estimators=100)
eclf = VotingClassifier(estimators=[('lr', clf1), ('knn', clf2), ('rf',clf3)], voting='hard')
eclf.fit(rev_train_Norm,labels_train)
pred=eclf.predict(rev_test_Norm)
print accuracy_score(pred,labels_test)
```

```
0.812375249501
```

SUMMARY



- The ‘Logistic Regression’ model proved to be the best for our prediction.

DETAILED SUMMARY

With length of review				
Algos	Accuracy without normalisation and reviewer ratings=1,2,3	Accuracy with normalisation & reviewer ratings =1,2,3	Accuracy without normalisation and reviewer ratings=1,2,3,4,5	Accuracy with normalisation & reviewer ratings =1,2,3,4,5
KNN	77.45	80.39	45.5	56.45
Logistic Regression	79.44	80.83	55.28	55.88
NB	32.5	79.24	17.77	57.89
Random Forest	78.04	80.03	52.89	52.09
Combination of 3(KNN,Random Forest,Logistic)	78.64	81.83	51.49	55.48

Without length of review				
Algos	Accuracy without normalisation and reviewer ratings=1,2,3	Accuracy with normalisation & reviewer ratings =1,2,3	Accuracy without normalisation and reviewer ratings=1,2,3,4,5	Accuracy with normalisation & reviewer ratings =1,2,3,4,5
KNN	76.44	81.43	46.03	55.88
Logistic Regression	79.44	80.63	56.8	57.28
NB	54.08	81.63	33.93	55.88
Random Forest	78.84	79.24	49.9	51.49
Combination of 3(KNN,Random Forest,Logistic)	80.83	80.638	52.09	53.89

- Length of review feature will not be used for the prediction.


Thank You