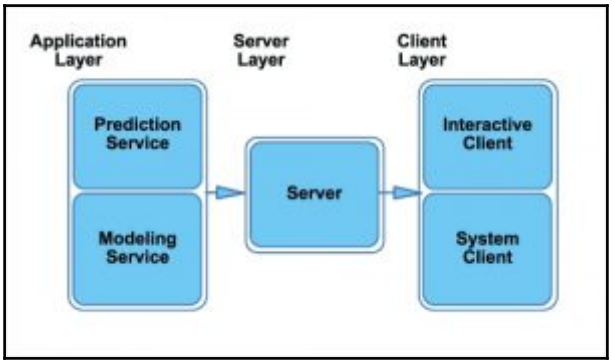
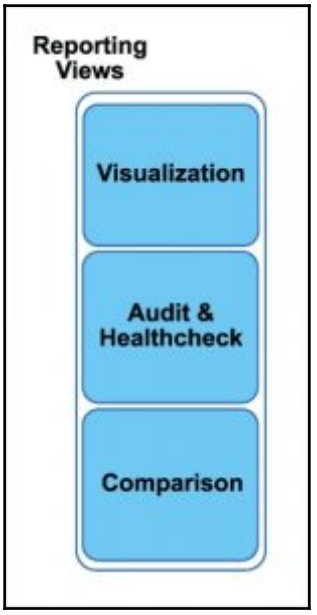
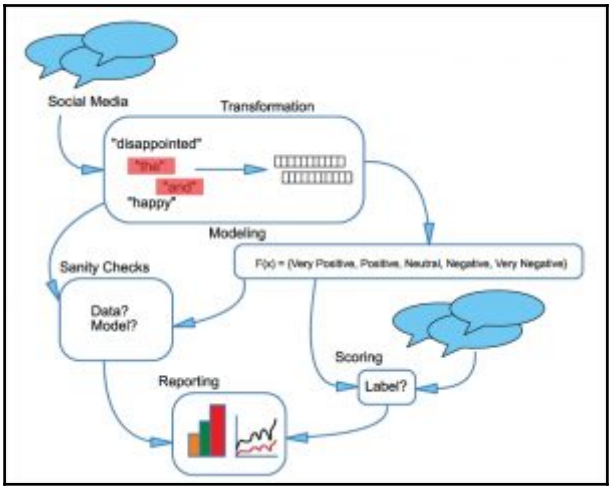
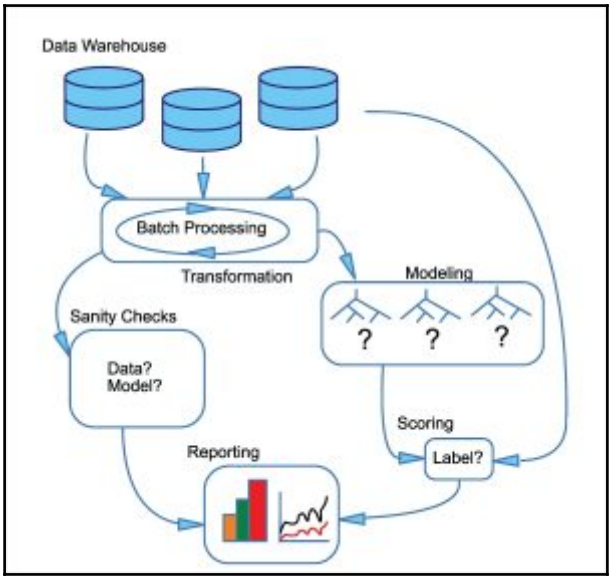
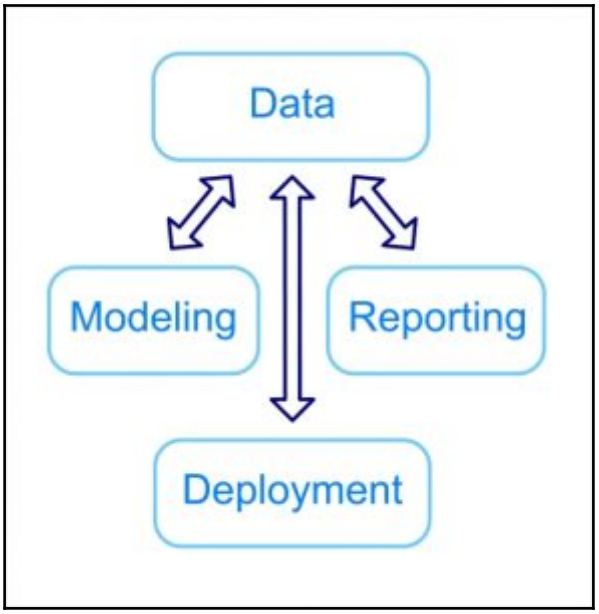
Chapter 1: From Data to Decisions – Getting Started with Analytic Applications

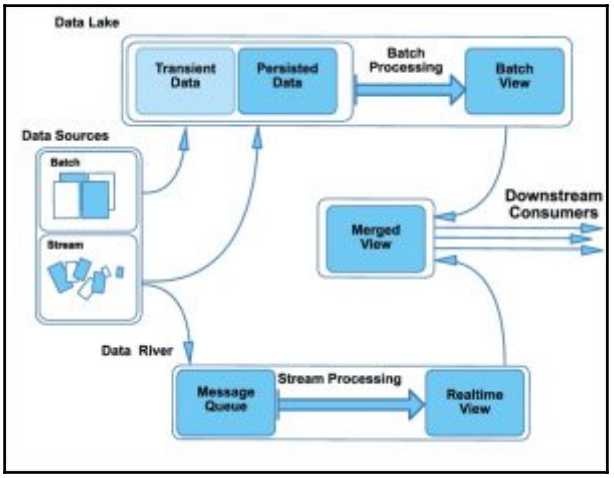


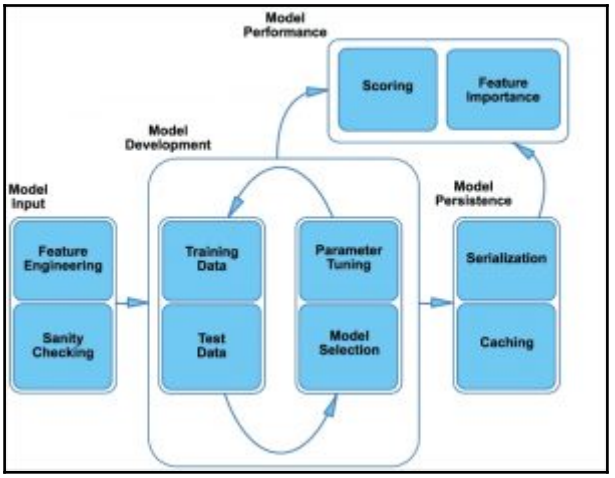



Social Media
Transformation
"disappointed"
"this"
"and"
"happy"
Modeling
F(x) = (Very Positive, Positive, Neutral, Negative, Very Negative)
Sanity Checks
Data?
Model?
Scoring
Label?
Reporting


Data Warehouse
Batch Processing
Transformation
Modeling
? ? ?
Sanity Checks
Data?
Model?
Scoring
Label?
Reporting


Data
Modeling
Reporting
Deployment


Data Lake
Transient Data
Persisted Data
Batch Processing
Batch View
Data Sources
Batch
Stream
Merged View
Downstream Consumers
Data River
Message Queue
Stream Processing
Realtime View



Chapter 2: Exploratory Data Analysis and Visualization in Python

	Unnamed: 0	year	length	budget	rating	votes	r1
count	58788.00000	58788.000000	58788.000000	5.215000e+03	58788.000000	58788.000000	58788.000000
mean	29394.50000	1976.130582	82.337875	1.341251e+07	5.932850	632.130384	7.014382
std	16970.77815	23.735125	44.347717	2.339068e+07	1.553031	3629.821413	10.936759
min	1.00000	1893.000000	1.000000	0.000000e+00	1.000000	5.000000	0.000000
25%	14697.75000	1958.000000	74.000000	2.500000e+05	5.000000	11.000000	0.000000
50%	29394.50000	1983.000000	90.000000	3.000000e+06	6.100000	30.000000	4.500000
75%	44091.25000	1997.000000	100.000000	1.500000e+07	7.000000	112.000000	4.500000
max	58788.00000	2005.000000	5220.000000	2.000000e+08	10.000000	157608.000000	100.000000

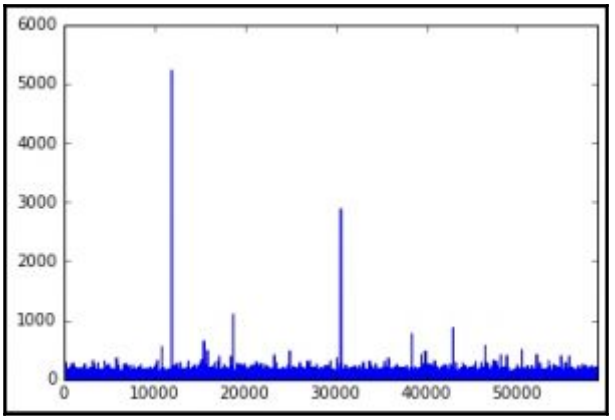
```
Unnamed: 0       int64
title           object
year             int64
length           int64
budget         float64
rating         float64
votes            int64
r1             float64
r2             float64
r3             float64
r4             float64
r5             float64
r6             float64
r7             float64
r8             float64
r9             float64
r10            float64
mpaa            object
Action           int64
Animation        int64
Comedy           int64
Drama            int64
Documentary      int64
Romance          int64
Short            int64
dtype: object
```

```
0    1971
1    1939
2    1941
3    1996
4    1975
Name: year, dtype: int64
```

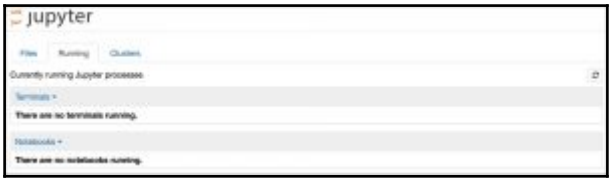
2002	2168
2003	2158
2001	2121
2000	2048
2004	1945
1999	1927
1998	1705
1997	1568
1996	1390
1995	1248
1994	1199
1993	1016
1987	957
1992	948
1989	944
1988	944
1990	899
1991	888
1985	792
1986	792
1984	749

	Year	Oil prices in constant 1997 dollars. 1870-1997
123	1993	17.15
124	1994	18.27
125	1995	19.40
126	1996	20.52
127	Oil prices in constant 1997 dollars. 1870-1997	NaN

	Unnamed: 0	title	year	length	budget	rating	votes	r1	r2	r3	...	r9	r10	mpaa	Action	Animation
6	7	$windle	2002	93	NaN	5.3	200	4.5	0.0	4.5	...	4.5	14.5	R	1	0
42	43	'R Xmas	2001	83	NaN	4.9	288	14.5	4.5	4.5	...	4.5	4.5	R	0	0
122	123	100 Girls	2000	90	NaN	5.8	3349	4.5	4.5	4.5	...	4.5	4.5	R	0	0
123	124	100 Mile Rule	2002	98	1100000	5.6	181	4.5	4.5	4.5	...	4.5	14.5	R	0	0
152	153	11:11	2004	95	NaN	4.3	222	14.5	14.5	4.5	...	4.5	14.5	R	0	0



	Year	Oil_Price_1997_Dollars
0	1870-01-01	58.53
1	1871-01-01	49.09
2	1872-01-01	24.68
3	1873-01-01	16.71
4	1874-01-01	19.86



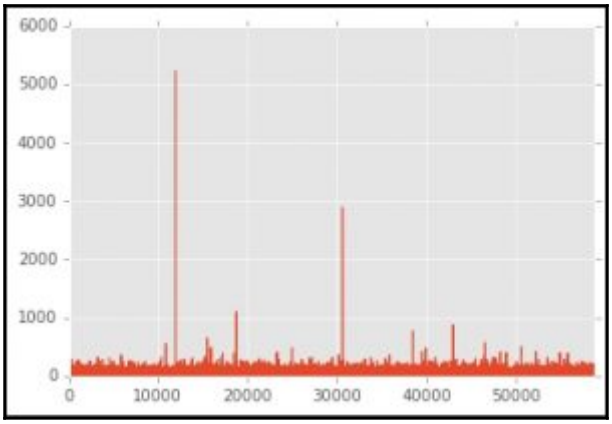
2005	349
2004	1945
2003	2158
2002	2168
2001	2121
2000	2048
1999	1927
1998	1705
1997	1568
1996	1390
1995	1248
1994	1199
1993	1016
1992	948
1991	888
1990	899
1989	944
1988	944
1987	957
1986	792
1985	792
1984	749
1983	698
1982	689

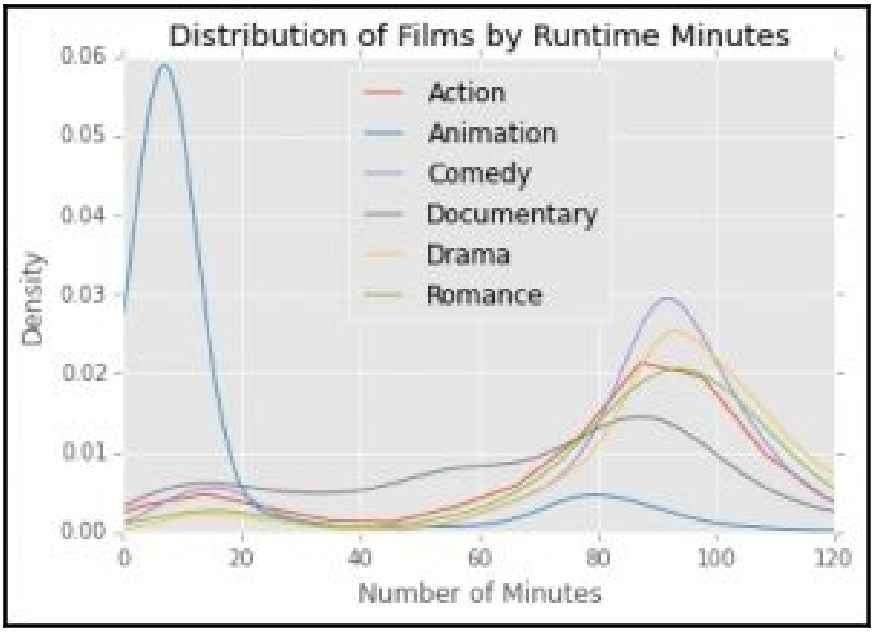
	mean	std
year		
1893	7.000000	NaN
1894	4.888889	0.727056
1895	5.500000	0.624500
1896	5.269231	1.325635
1897	4.677778	0.732765
1898	5.040000	0.950263
1899	4.277778	0.713754
1900	4.731250	1.358783
1901	4.682143	1.081513
1902	4.900000	1.615549
1903	4.808108	1.334662
1904	4.223810	1.291588
1905	5.047059	1.196410
1906	5.676471	1.274034

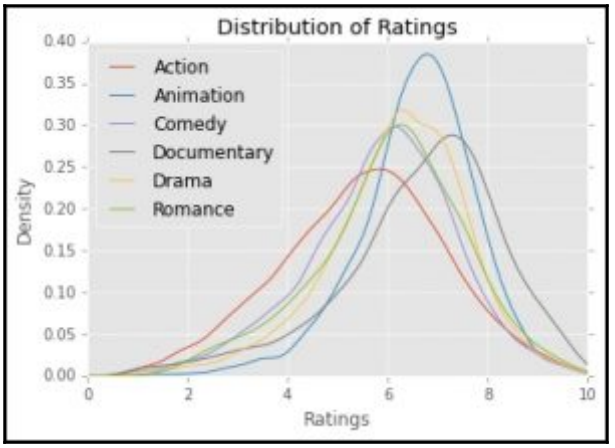
```
0       Comedy
1       Comedy
2    Animation
3       Comedy
4       Action
```

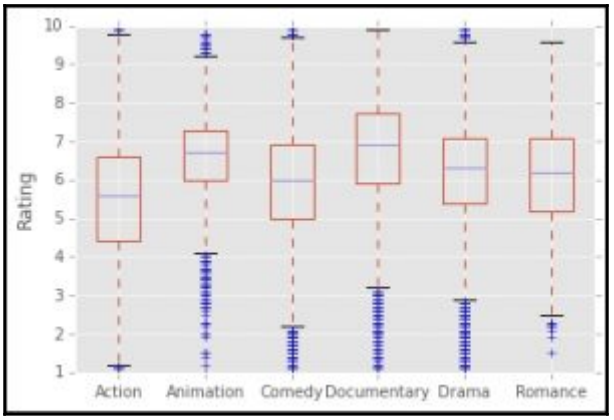
```
0    yellow
1    yellow
2      blue
3    yellow
4       red
```

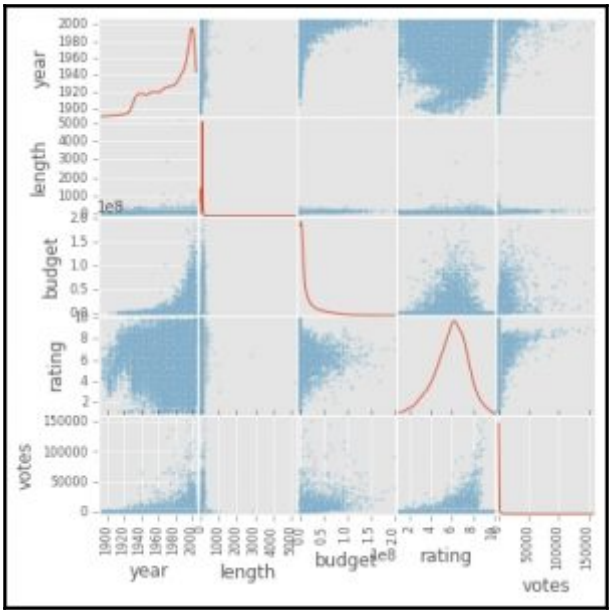
genre	Action	Animation	Comedy	Documentary	Drama	Romance
year						
1893	7.000000	NaN	NaN	NaN	NaN	NaN
1894	5.100000	NaN	NaN	4.720000	NaN	NaN
1895	5.700000	NaN	NaN	5.400000	NaN	NaN
1896	5.875000	NaN	3.900000	5.571429	2.100000	NaN
1897	5.900000	NaN	5.200000	4.300000	NaN	NaN
1898	6.000000	NaN	5.050000	5.500000	3.600000	NaN
1899	4.533333	NaN	3.600000	3.866667	4.850000	NaN

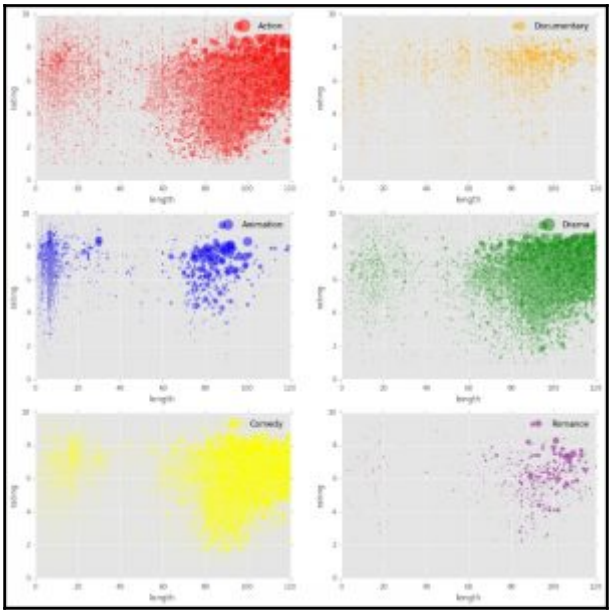



Distribution of Films by Runtime Minutes
Action
Animation
Comedy
Documentary
Drama
Romance
Density
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
20
40
60
80
100
120
Number of Minutes

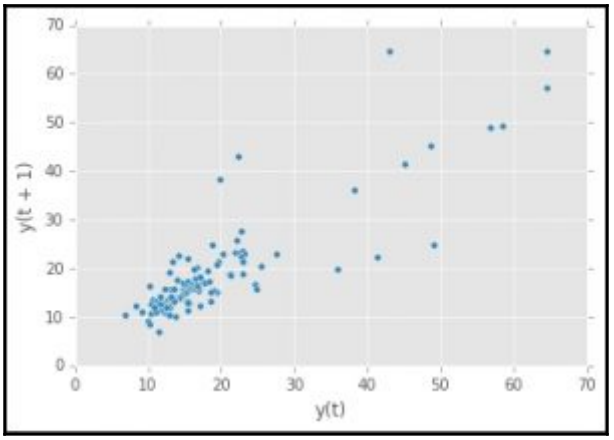

Distribution of Ratings
Action
Animation
Comedy
Documentary
Drama
Romance
Density
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
Ratings

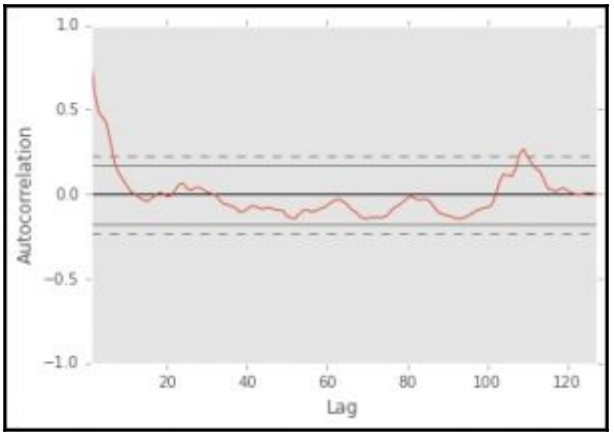

Rating
10
9
8
7
6
5
4
3
2
1
Action
Animation
Comedy
Documentary
Drama
Romance


year
length
budget
rating
votes
1e8
year
length
budget
rating
votes

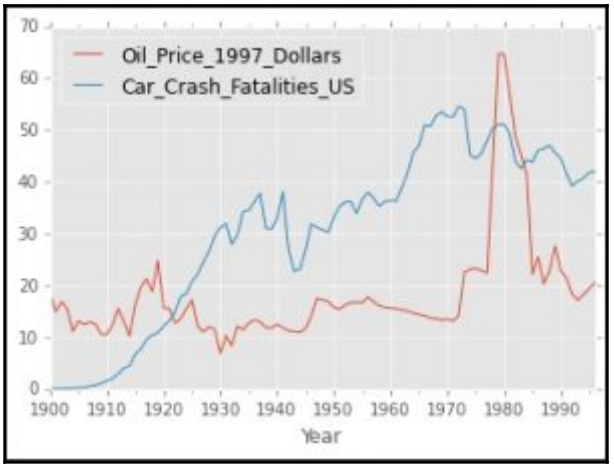

Action
Documentary
Animation
Drama
Comedy
Romance
length

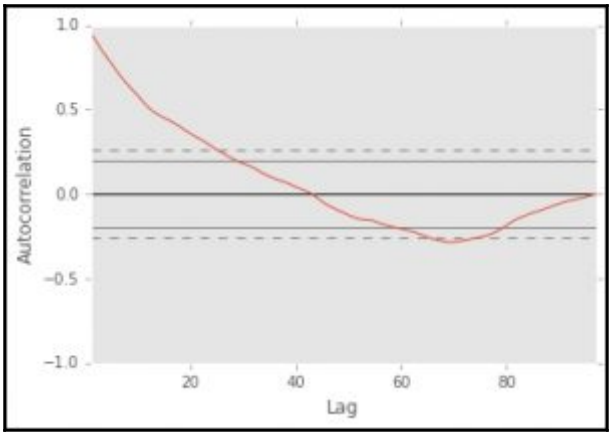

Oil_Price_1997_Dollars
70
60
50
40
30
20
10
0
1870
1895
1920
1945
1970
1995
Year



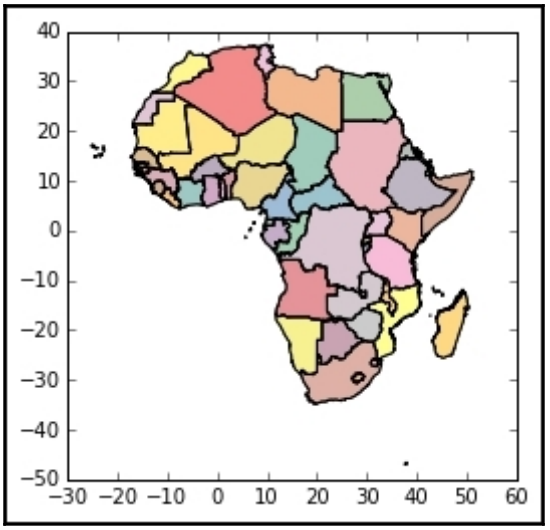


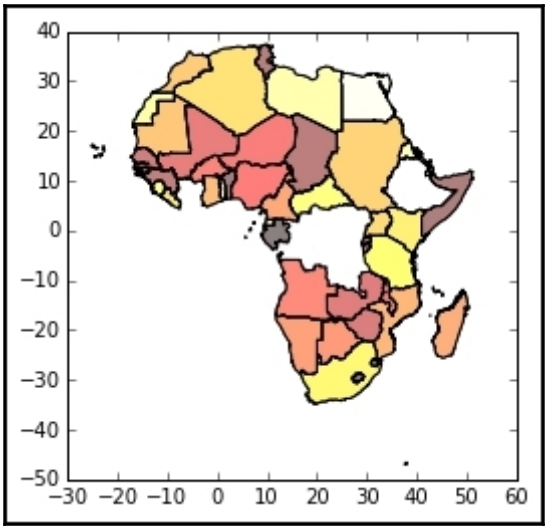
	Year	Car_Crash_Fatalities_US
0	1900-01-01	36
1	1901-01-01	54
2	1902-01-01	79
3	1903-01-01	117
4	1904-01-01	172





	CODE	COUNTRY	ID	geometry
0	ALG	Algeria	1	POLYGON ((-5.7636199999979 25.5862499999302, ...
1	ANG	Angola	2	POLYGON ((13.36632442474365 -8.32172966003418,...
2	ANG	Angola	3	POLYGON ((12.80576000000292 -4.80649000002668...
3	ANG	Angola	4	POLYGON ((11.76834011077681 -16.79932975769043...
4	ANG	Angola	5	POLYGON ((12.89840030670166 -5.988018989562988...


40
30
20
10
0
−10
−20
−30
−40
−50
−30 −20 −10 0 10 20 30 40 50 60


40
30
20
10
0
−10
−20
−30
−40
−50
−30 −20 −10 0 10 20 30 40 50 60

```
15/09/17 20:58:36 INFO SparkContext: Running Spark version 1.5.0
15/09/17 20:58:36 INFO SecurityManager: Changing view acls to: jbabcock
15/09/17 20:58:36 INFO SecurityManager: Changing modify acls to: jbabcock
15/09/17 20:58:36 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permission
s: Set(jbabcock); users with modify permissions: Set(jbabcock)
15/09/17 20:58:36 INFO Slf4jLogger: Slf4jLogger started
15/09/17 20:58:36 INFO Remoting: Starting remoting
15/09/17 20:58:36 INFO Remoting: Remoting started; listening on addresses :[akka.tcp://sparkDriver@localhost:50237]
15/09/17 20:58:36 INFO Utils: Successfully started service 'sparkDriver' on port 50237.
15/09/17 20:58:36 INFO SparkEnv: Registering MapOutputTracker
15/09/17 20:58:36 INFO SparkEnv: Registering BlockManagerMaster
15/09/17 20:58:36 INFO DiskBlockManager: Created local directory at /private/var/folders/mw/gf7_2mds3dd_01hq50r4bbgrqjxrx8/T/b
lockmgr-827f83ed8-c704-4247-b9a0-8c2e1332d615
15/09/17 20:58:36 INFO MemoryStore: MemoryStore started with capacity 504.3 MB
15/09/17 20:58:36 INFO HttpFileServer: HTTP File server directory is /private/var/folders/mw/gf7_2mds3dd_01hq50r4bbgrqjxrx8/T/
spark-0a59ff77-5730-4801-ad5d-e38d32404c36/httpd-6ec28393-7c40-42dd-956a-454026d1102#
15/09/17 20:58:36 INFO HttpServer: Starting HTTP Server
15/09/17 20:58:36 INFO Server: jetty-8.y.z-SNAPSHOT
15/09/17 20:58:36 INFO AbstractConnector: Started SocketConnector@0.0.0.0:50238
15/09/17 20:58:36 INFO Utils: Successfully started service 'HTTP file server' on port 50238.
15/09/17 20:58:36 INFO SparkEnv: Registering OutputCommitCoordinator
15/09/17 20:58:36 INFO Server: jetty-8.y.z-SNAPSHOT
15/09/17 20:58:36 INFO AbstractConnector: Started SelectChannelConnector@0.0.0.0:4040
15/09/17 20:58:36 INFO Utils: Successfully started service 'SparkUI' on port 4040.
15/09/17 20:58:36 INFO SparkUI: Started SparkUI at http://localhost:4040
15/09/17 20:58:36 WARN MetricsSystem: Using default name DAGScheduler for source because spark.app.id is not set.
15/09/17 20:58:36 INFO Executor: Starting executor ID driver on host localhost
15/09/17 20:58:36 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port
```


```
Index(['Unnamed: 0', 'title', 'year', 'length', 'budget', 'rating', 'votes',
       'r1', 'r2', 'r3', 'r4', 'r5', 'r6', 'r7', 'r8', 'r9', 'r10', 'mpaa',
       'Action', 'Animation', 'Comedy', 'Drama', 'Documentary', 'Romance',
       'Short'],
      dtype='object')
```

$$K(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)}{2\sigma^2}}$$

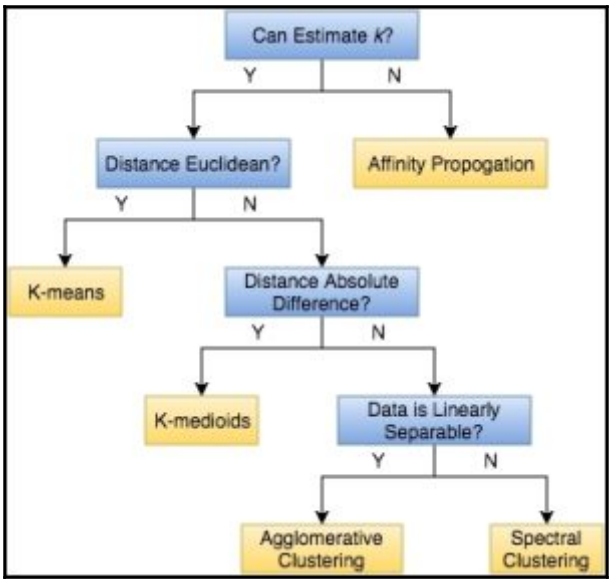
$$f(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

```
Year                                              object
Oil prices in constant 1997 dollars. 1870-1997   float64
dtype: object
```

	Unnamed: 0	title	year	length	budget	rating	votes	r1	r2	r3	...	r9	r10	mpaa	Action
0	1	$	1971	121	NaN	6.4	348	4.5	4.5	4.5	...	4.5	4.5	NaN	0
1	2	$1000 a Touchdown	1939	71	NaN	6.0	20	0.0	14.5	4.5	...	4.5	14.5	NaN	0
2	3	$21 a Day Once a Month	1941	7	NaN	8.2	5	0.0	0.0	0.0	...	24.5	24.5	NaN	0
3	4	$40,000	1996	70	NaN	8.2	6	14.5	0.0	0.0	...	34.5	45.5	NaN	0
4	5	$50,000 Climax Show, The	1975	71	NaN	3.4	17	24.5	4.5	0.0	...	0.0	24.5	NaN	0

	Unnamed: 0	title	year	length	budget	rating	votes	r1	r2	r3
58773	58774	deadend.com	2002	120	NaN	6.9	53	64.5	4.5	0.0
58774	58775	e-Dreams	2001	94	NaN	6.8	86	4.5	0.0	0.0
58775	58776	eMale	2001	17	NaN	7.3	15	0.0	0.0	0.0
58776	58777	eRATicate	2003	9	NaN	6.0	5	0.0	0.0	0.0
58777	58778	eXXXorcismos	2002	78	NaN	4.2	11	34.5	0.0	0.0
58778	58779	eXistenZ	1999	97	NaN	6.7	14742	4.5	4.5	4.5
58779	58780	f2point8	2002	20	NaN	5.0	13	4.5	0.0	4.5
58780	58781	f8	2001	13	NaN	7.6	7	0.0	0.0	0.0
58781	58782	pLiRe kILLjoy	1998	87	NaN	5.2	6	0.0	14.5	14.5
58782	58783	sIDney	2002	15	NaN	7.0	8	14.5	0.0	0.0
58783	58784	tom thumb	1958	98	NaN	6.5	274	4.5	4.5	4.5
58784	58785	www.XXX.com	2003	105	NaN	1.1	12	45.5	0.0	0.0
58785	58786	www.hellssoapopera.com	1999	100	NaN	6.6	5	24.5	0.0	24.5

Chapter 3: Finding Patterns in the Noise – Clustering and Unsupervised Learning



$$r(i,k) = s(i,k) - \max_{k' \neq k} \left\{ a(i,k') + s(i,k') \right\}$$

$$a(i,k) = \min\left(0, r(k,k) + \sum_{i' \notin \{i,k\}} \max\left(0, r(i',k)\right)\right)$$

$$a(k,k) = \sum_{i' \notin \{i,k\}} \max\left(0, r(i',k)\right)$$

$$e^{-\gamma D\left(x_i, x_j\right)^2}$$

$$L = I - D^{1/2} K D^{1/2}$$

$$D_{ii} = \sum_{j=1}^{n} K_{ij}$$

$$Av = \lambda v$$

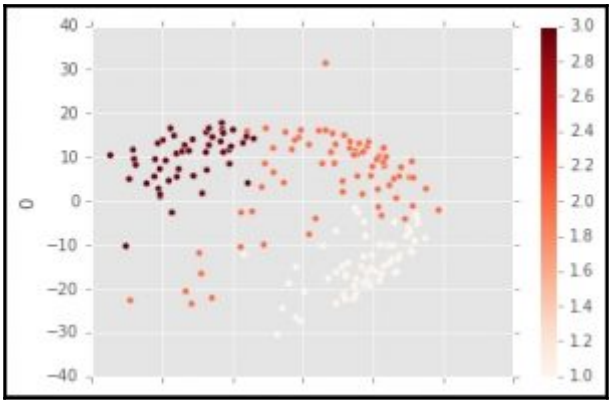
$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	1	14.23	1.71	2.43	15.6	127	2.80	3.06	0.28	2.29	5.64	1.04	3.92	1065
1	1	13.20	1.78	2.14	11.2	100	2.65	2.76	0.26	1.28	4.38	1.05	3.40	1050
2	1	13.16	2.36	2.67	18.6	101	2.80	3.24	0.30	2.81	5.68	1.03	3.17	1185
3	1	14.37	1.95	2.50	16.8	113	3.85	3.49	0.24	2.18	7.80	0.86	3.45	1480
4	1	13.24	2.59	2.87	21.0	118	2.80	2.69	0.39	1.82	4.32	1.04	2.93	735

	Class	Alcohol	Malic	Ash	Alcalinity	Magnesium	Total	Flavanoids
count	178.000000	178.000000	178.000000	178.000000	178.000000	178.000000	178.000000	178.000000
mean	1.938202	13.000618	2.336348	2.366517	19.494944	99.741573	2.295112	2.029270
std	0.775035	0.811827	1.117146	0.274344	3.339564	14.282484	0.625851	0.998859
min	1.000000	11.030000	0.740000	1.360000	10.600000	70.000000	0.980000	0.340000
25%	1.000000	12.362500	1.602500	2.210000	17.200000	88.000000	1.742500	1.205000
50%	2.000000	13.050000	1.865000	2.360000	19.500000	98.000000	2.355000	2.135000
75%	3.000000	13.677500	3.082500	2.557500	21.500000	107.000000	2.800000	2.875000
max	3.000000	14.830000	5.800000	3.230000	30.000000	162.000000	3.880000	5.080000

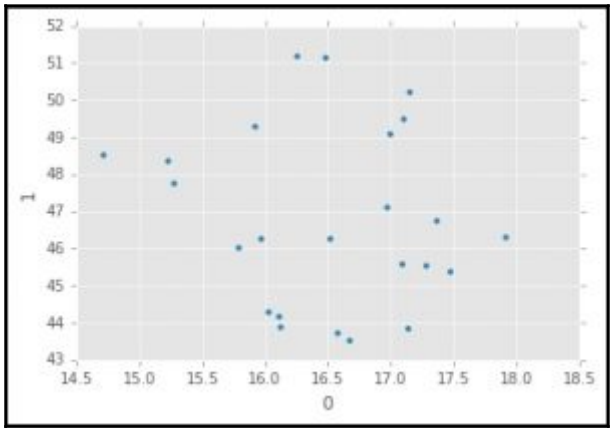
	Alcohol	Malic	Ash	Alcalinity	Magnesium	Total	Flavanoids
count	1.780000e+02	1.780000e+02	1.780000e+02	1.780000e+02	1.780000e+02	178.000000	1.780000e+02
mean	-6.382608e-16	-1.197544e-16	-6.370333e-16	-3.991813e-17	-3.991813e-17	0.000000	-3.991813e-16
std	1.002821e+00	1.002821e+00	1.002821e+00	1.002821e+00	1.002821e+00	1.002821	1.002821e+00
min	-2.434235e+00	-1.432983e+00	-3.679162e+00	-2.671018e+00	-2.088255e+00	-2.107246	-1.695971e+00
25%	-7.882448e-01	-6.587486e-01	-5.721225e-01	-6.891372e-01	-8.244151e-01	-0.885468	-8.275393e-01
50%	6.099988e-02	-4.231120e-01	-2.382132e-02	1.518295e-03	-1.222817e-01	0.095960	1.061497e-01
75%	8.361286e-01	6.697929e-01	6.861086e-01	6.020883e-01	5.096384e-01	0.808997	8.490851e-01
max	2.259772e+00	3.109192e+00	3.156325e+00	3.154511e+00	4.371372e+00	2.539515	3.062832e+00

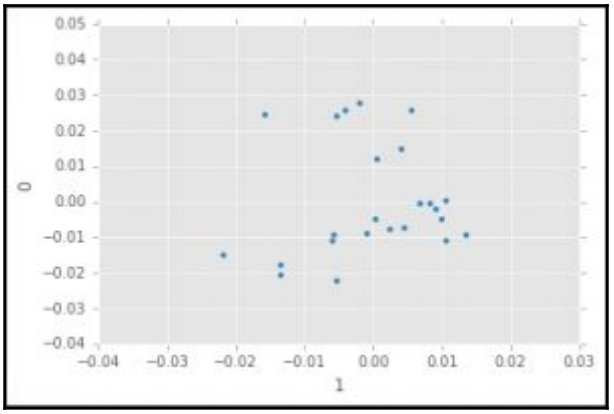
$$D(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

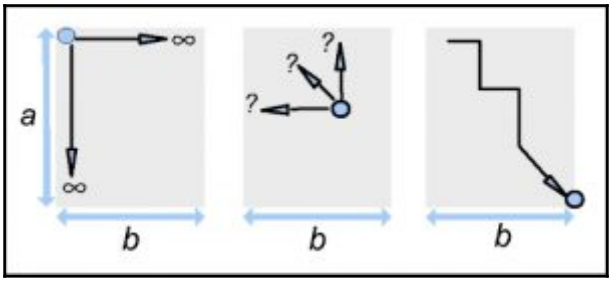


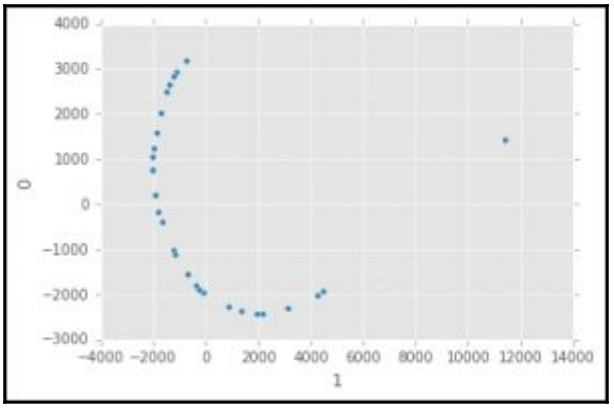
	quarter	stock	date	open	high	low	close	volume	percent_change_price
0	1	AA	1/7/2011	$15.82	$16.72	$15.78	$16.42	239655616	3.79267
1	1	AA	1/14/2011	$16.71	$16.71	$15.64	$15.97	242963398	-4.42849
2	1	AA	1/21/2011	$16.19	$16.38	$15.60	$15.79	138428495	-2.47066
3	1	AA	1/28/2011	$15.87	$16.63	$15.82	$16.13	151379173	1.63831
4	1	AA	2/4/2011	$16.18	$17.39	$16.18	$17.14	154387761	5.93325

date	stock	2011-01-07 00:00:00	2011-01-14 00:00:00	2011-01-21 00:00:00	2011-01-28 00:00:00	2011-02-04 00:00:00	2011-02-11 00:00:00	2011-02-18 00:00:00	2011-02-25 00:00:00	2011-0 00:00:
0	AA	16.42	15.97	15.79	16.13	17.14	17.37	17.28	16.68	16.58
1	AXP	44.36	46.25	46.00	43.86	43.82	46.75	45.53	43.53	43.72
2	BA	69.38	70.07	71.68	69.23	71.38	72.14	73.04	72.30	71.80
3	BAC	14.25	15.25	14.25	13.60	14.29	14.77	14.75	14.20	14.12
4	CAT	93.73	94.01	92.75	95.68	99.59	103.54	105.86	102.00	103.04







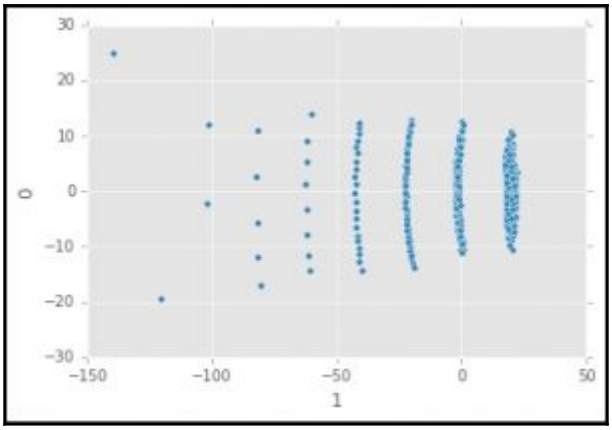


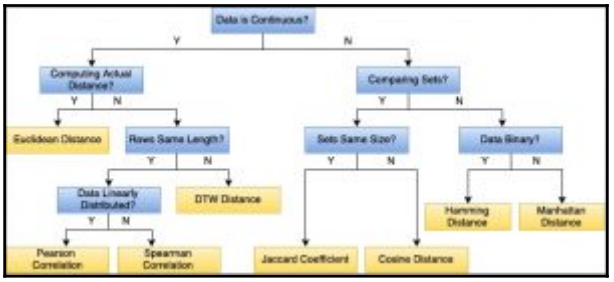
$$P(a,b)=\frac{Cov(a,b)}{\sigma(a)\sigma(b)}=\frac{\frac{1}{n}\sum_{i=1}^{n}(a_i-\mu(a))(b_i-\mu(b))}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i-\mu(a))^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(b_i-\mu(b))^2}}$$

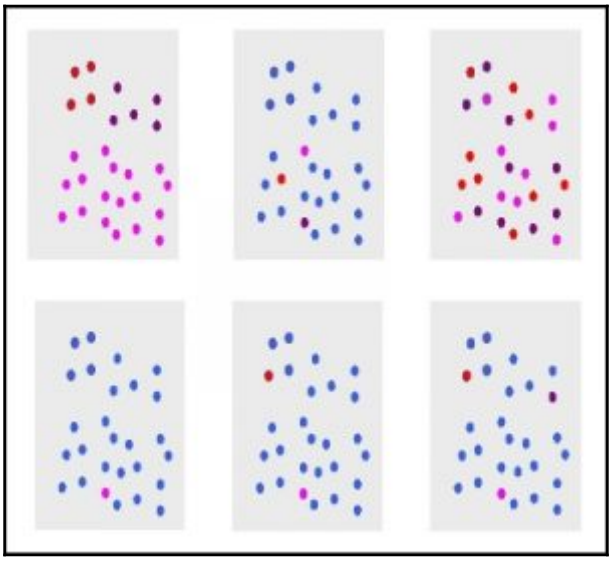
	title	authors	groups	keywords	topics	abstract
0	Kernelized Bayesian Transfer Learning	Mehmet Gönen and Adam A. Margolin	Naval Machine Learning Algorithms (NMLA)	cross-domain learning\ndomain adaptation\nkern...	APP: Biomedical / Bioinformatics\nNMLA: Bayesi...	Transfer learning considers related but distin...
1	"Source Free" Transfer Learning for Text Class...	Zhongqi Lu, Yin Zhu, Sinno Pan, Evan Xiang, Yu...	AI and the Web (AIW)\nNaval Machine Learning A...	Transfer Learning\nAuxiliary Data Retrieval\nT...	AIW: Knowledge acquisition from the web\nAIW: ...	Transfer learning uses relevant auxiliary data...
2	A Generalization of Probabilistic Serial to Ra...	Haris Aziz and Paul Stursberg	Game Theory and Economic Paradigms (GTEP)	social choice theory\nvoting\nfair division\nra...	GTEP: Game Theory\nGTEP: Social Choice / Voting	The probabilistic serial (PS) rule is one of t...

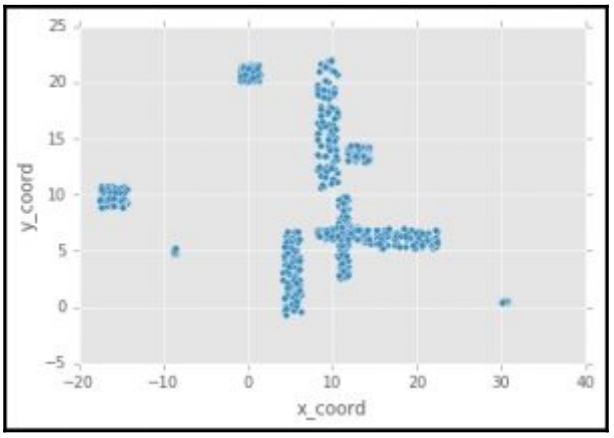
	keyword_422	keyword_1174	keyword_640	keyword_1287	keyword_312	keyword_378	keyword_826	keyword_9
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0

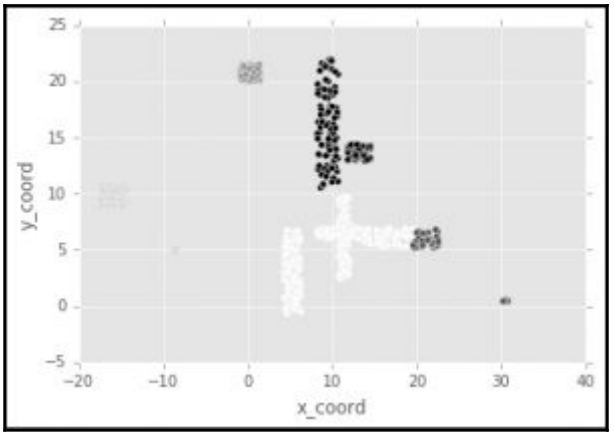
$$\rho(a,b) = \frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}$$

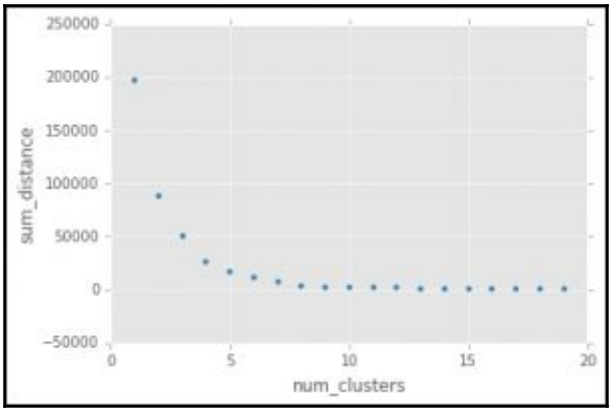


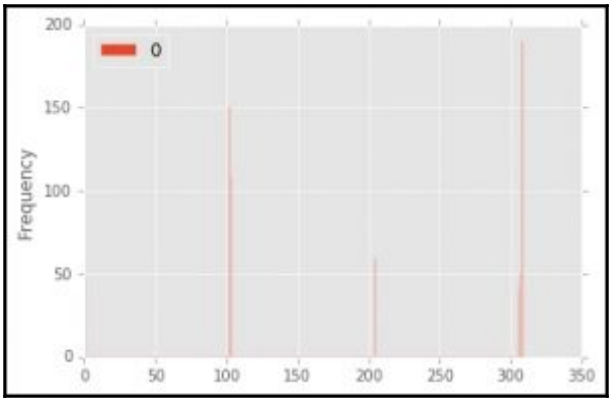

Data is Continuous?
Y
N
Computing Actual Distance?
Comparing Sets?
Y
N
Y
N
Euclidean Distance
Rows Same Length?
Sets Same Size?
Data Binary?
Y
N
Y
N
Y
N
Data Linearly Distributed?
DTW Distance
Hamming Distance
Manhattan Distance
Y
N
Pearson Correlation
Spearman Correlation
Jaccard Coefficient
Cosine Distance

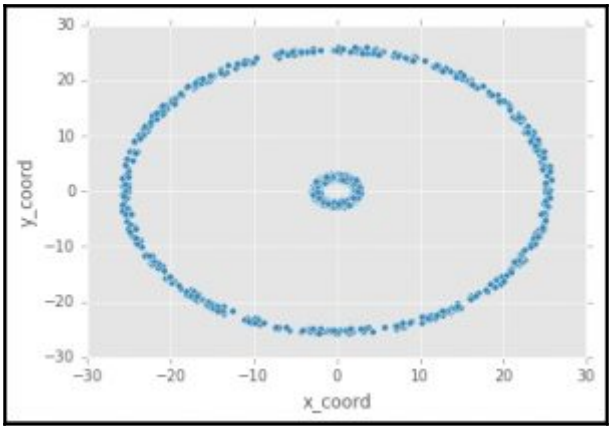


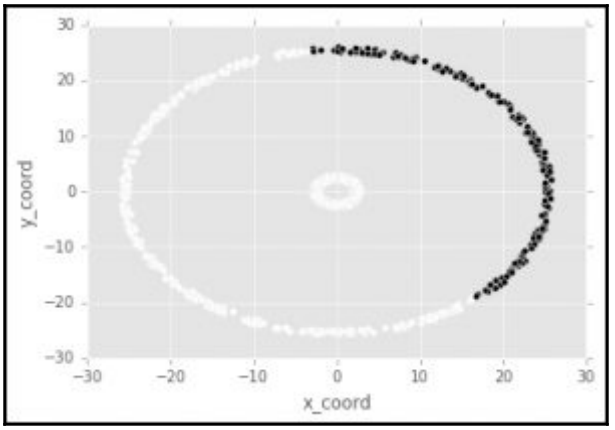

25
20
15
10
5
0
−5
y_coord
−20
−10
0
10
20
30
40
x_coord


y_coord
x_coord
25
20
15
10
5
0
−5
−20
−10
0
10
20
30
40

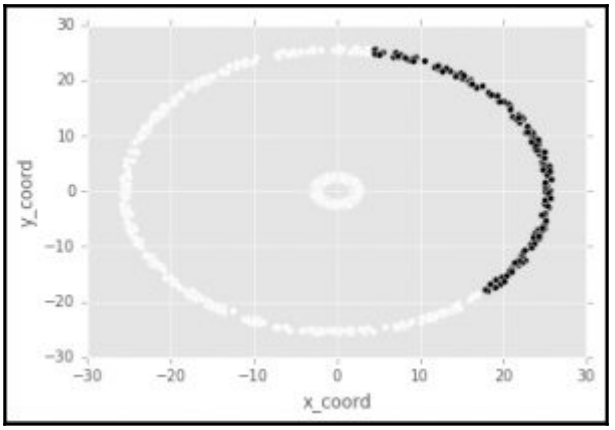

sum_distance
num_clusters
250000
200000
150000
100000
50000
0
−50000
0
5
10
15
20


0
Frequency
200
150
100
50
0
0
50
100
150
200
250
300
350





$$J(a,b) = \frac{\|a \cap b\|}{\|a \cap b\|}$$



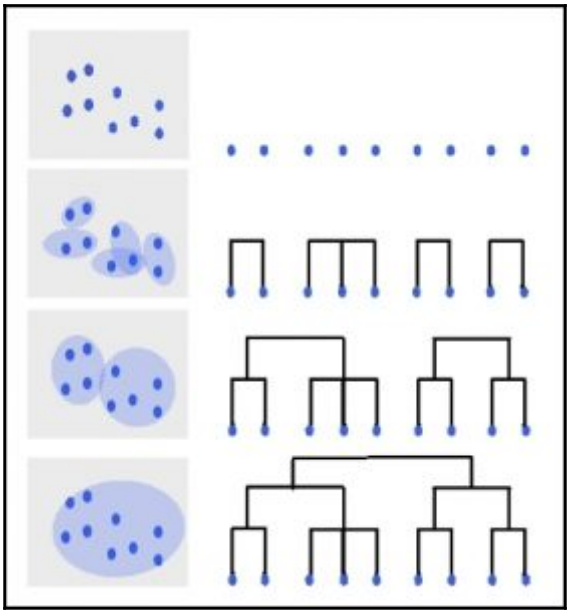
$$Cos(a,b) = \frac{a \sum b}{\|a\| \|b\|}$$

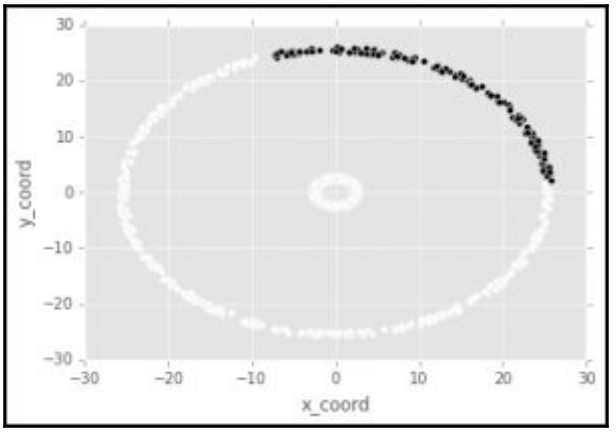
$$\|a\| = \sqrt{\sum_{i=1}^{n} a_i^2}$$

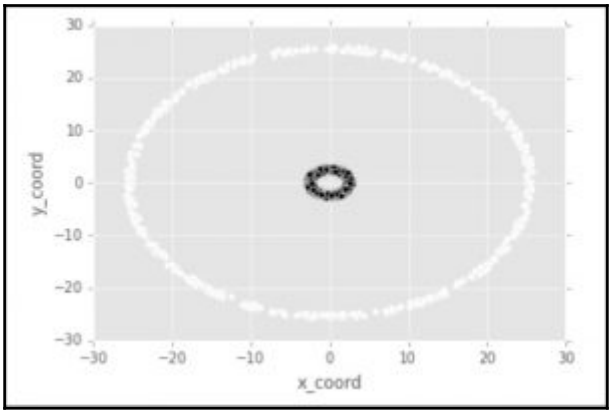
$$a \sum b = \sum_{i=1}^{n} a_i b_i$$

$$H(a,b) = \sum_{i=1}^{n} 1 if\ a_i = b_i\ else\ 0$$

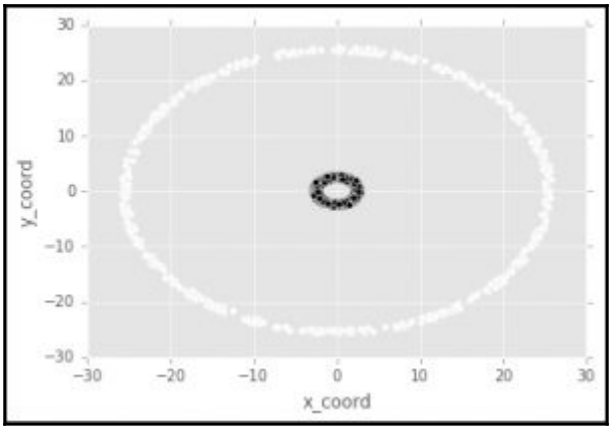
$$M(a,b) = \sum_{i=1}^{n} |a_i - b_i|$$




30
20
10
0
−10
−20
−30
y_coord
−30
−20
−10
0
10
20
30
x_coord



$$SSE = \sum_{i=1}^{n} D\left(x_i, c_i\right)^2$$



Chapter 4: Connecting the Dots with Models – Regression Methods

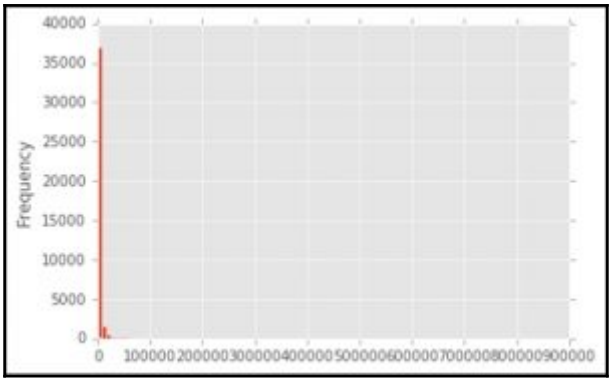
$$VR(L,R) = \frac{1}{|A^2|}\sum_{i\in A}\sum_{j\in A}\frac{1}{2}\left(y_i - y_j\right)^2 - \left(\frac{1}{|L^2|}\sum_{i\in L}\sum_{j\in L}\frac{1}{2}\left(y_i - y_j\right)^2 + \frac{1}{|R^2|}\sum_{i\in R}\sum_{j\in R}\frac{1}{2}\left(y_i - y_j\right)^2\right)$$

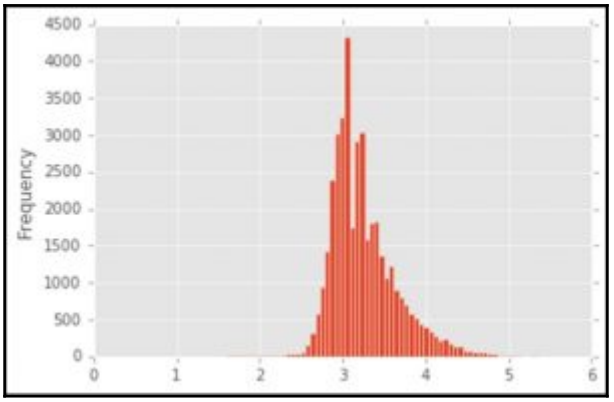
$$IG(L,R) = -\sum_{k=1}^{K} f_{A_k} \log_2 f_{A_k} - \left(-\alpha \sum_{k=1}^{K} f_{L_k} \log_2 f_{L_k} + (1-\alpha) \sum_{k=1}^{K} f_{R_k} \log_2 f_R \right)$$

$$GI(L,R) = 1 - \left(\alpha \sum_{k=1}^{K} f_{L_k}^2 + (1-\alpha) \sum_{k=1}^{K} f_{R_k}^2 \right)$$

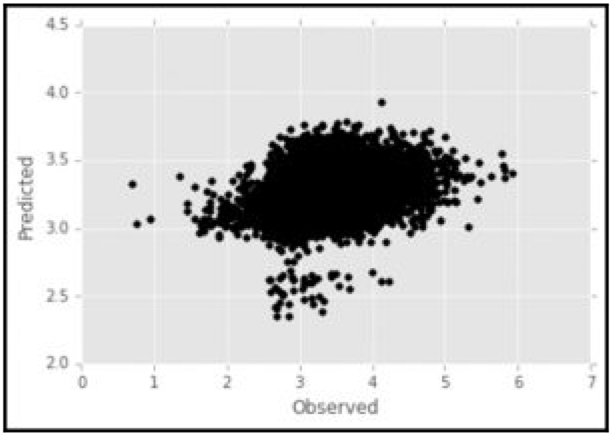
$$GI(L,R) = 1 - \left(\alpha \sum_{k=1}^{K} f_{L_k}^2 + (1-\alpha) \sum_{k=1}^{K} f_{R_k}^2 \right)$$

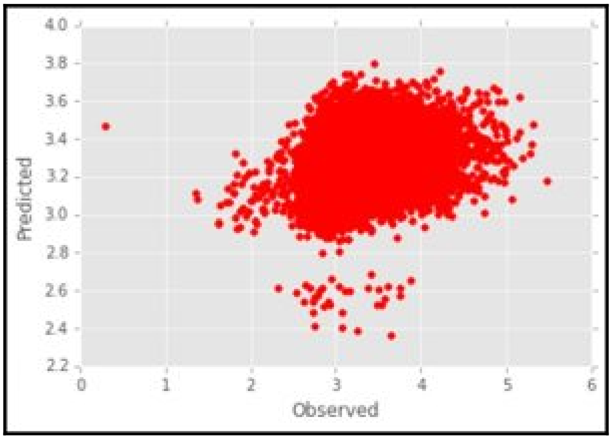
	timedelta	n_tokens_title	n_tokens_content	n_unique_tokens	n_non_stop_words	n_non_stop_unique_tok
count	39644.000000	39644.000000	39644.000000	39644.000000	39644.000000	39644.000000
mean	354.530471	10.398749	546.514731	0.548216	0.996469	0.689175
std	214.163767	2.114037	471.107508	3.520708	5.231231	3.264816
min	8.000000	2.000000	0.000000	0.000000	0.000000	0.000000
25%	164.000000	9.000000	246.000000	0.470870	1.000000	0.625739
50%	339.000000	10.000000	409.000000	0.539226	1.000000	0.690476
75%	542.000000	12.000000	716.000000	0.608696	1.000000	0.754630
max	731.000000	23.000000	8474.000000	701.000000	1042.000000	650.000000





	timedelta	n_tokens_title	n_tokens_content	n_unique_tokens	n_non_stop_words	n_non_stop_unique_tokens	num_hrefs
count	39644.000000	39644.000000	39644.000000	39644.000000	39644.000000	39644.000000	39644.000000
mean	0.531735	0.311212	0.542389	0.072853	0.110895	0.086453	0.279180
std	0.053246	0.017781	0.102187	0.016811	0.018587	0.017908	0.080195
min	0.266976	0.169416	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.507516	0.301030	0.530545	0.067268	0.114287	0.083162	0.230168
50%	0.547667	0.309627	0.557642	0.074561	0.114287	0.089202	0.290876
75%	0.572287	0.325083	0.586083	0.081518	0.114287	0.094866	0.337877
max	0.587095	0.378616	0.692083	0.565047	0.804041	0.581333	0.542116



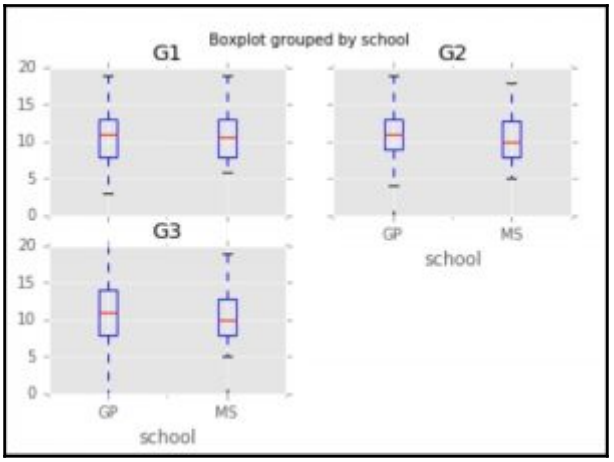


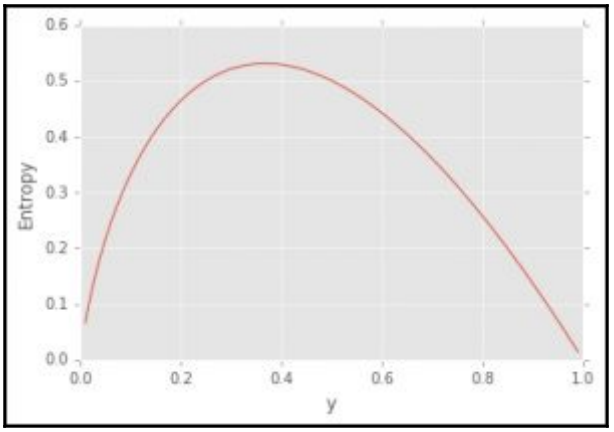
OLS Regression Results

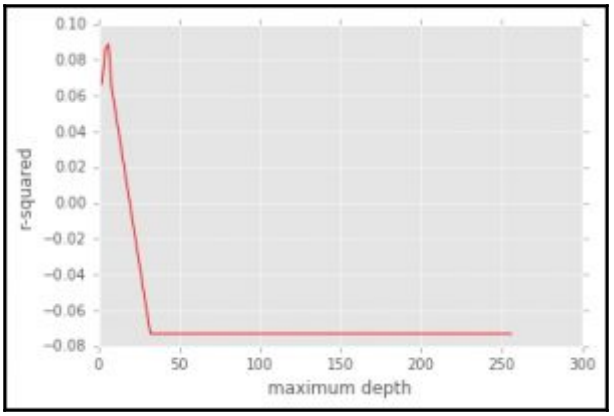
Dep. Variable:	shares	**R-squared:**	0.123
Model:	OLS	**Adj. R-squared:**	0.121
Method:	Least Squares	**F-statistic:**	98.80
Date:	Mon, 21 Mar 2016	**Prob (F-statistic):**	0.00
Time:	00:03:53	**Log-Likelihood:**	-17704.
No. Observations:	39644	**AIC:**	3.552e+04
Df Residuals:	39587	**BIC:**	3.601e+04
Df Model:	56		
Covariance Type:	nonrobust		

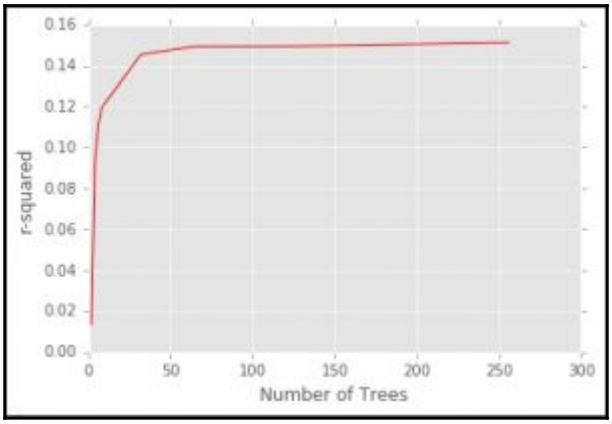
	coef	std err	t	P>|t|	[95.0% Conf. Int.]
timedelta	0.0142	0.006	2.322	0.020	0.002 0.026
n_tokens_title	0.0640	0.024	2.626	0.009	0.016 0.112
n_tokens_content	-0.0897	0.020	-4.451	0.000	-0.129 -0.050
n_unique_tokens	-1.3315	0.320	-4.164	0.000	-1.958 -0.705
n_non_stop_words	0.8054	0.215	3.740	0.000	0.383 1.228
n_non_stop_unique_tokens	0.5612	0.233	2.411	0.016	0.105 1.017
num_hrefs	0.0973	0.008	11.474	0.000	0.081 0.114
num_self_hrefs	-0.1083	0.011	-9.649	0.000	-0.130 -0.086
num_imgs	0.0487	0.006	7.981	0.000	0.037 0.061
num_videos	0.0667	0.008	8.620	0.000	0.052 0.082
average_token_length	-0.2459	0.103	-2.390	0.017	-0.448 -0.044
num_keywords	0.1209	0.024	5.130	0.000	0.075 0.167
data_channel_is_lifestyle	-0.0592	0.013	-4.545	0.000	-0.085 -0.034
data_channel_is_entertainment	-0.1166	0.008	-14.059	0.000	-0.133 -0.100

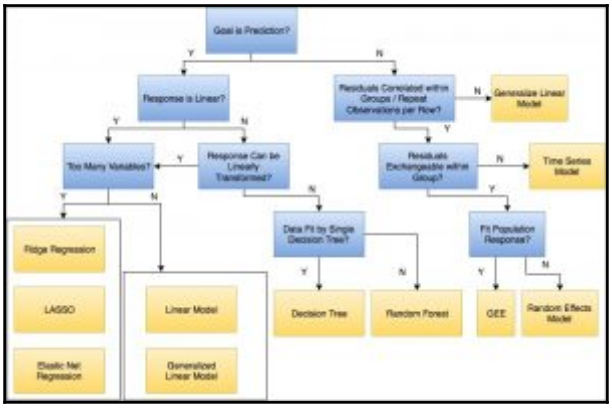
Omnibus:	7432.190	**Durbin-Watson:**	1.940
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	19103.164
Skew:	1.030	**Prob(JB):**	0.00
Kurtosis:	5.705	**Cond. No.**	3.16e+03

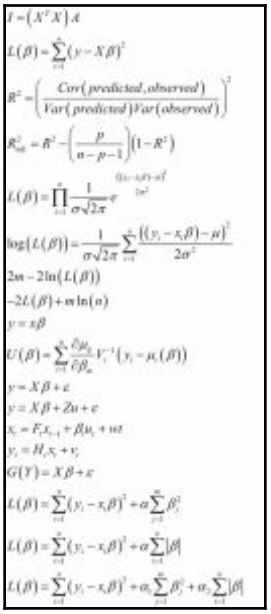

Boxplot grouped by school
G1
G2
G3
20
15
10
5
0
GP
MS
school


Entropy
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.2
0.4
0.6
0.8
1.0
y


r-squared
0.10
0.08
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.08
0
50
100
150
200
250
300
maximum depth


0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
r-squared
0
50
100
150
200
250
300
Number of Trees


Goal is Prediction?
Y
N
Response is Linear?
Residuals Correlated within Groups / Repeat Observations per Row?
N
Generalize Linear Model
Y
N
Y
Too Many Variables?
Y
Response Can be Linearly Transformed?
Residuals Exchangeable within Group?
N
Time Series Model
Y
N
N
Y
Data Fit by Single Decision Tree?
Fit Population Response?
Ridge Regression
LASSO
Elastic Net Regression
Linear Model
Generalized Linear Model
Y
N
Y
N
Decision Tree
Random Forest
GEE
Random Effects Model



$$y = X\beta$$

$$X^T y = X^T X\beta$$

$$\beta = \frac{X^T y}{\left(X^T X\right)^{-1}}$$

$$I = \left(X^T X\right) A$$

$$L(\beta) = \sum_{i=1}^{n} \left(y - \hat{y}\right)^2 = \sum_{i=1}^{n} \left(y - X\beta\right)^2$$

$$y = X\beta$$

$$\sum_{i=1}^{n}\left(y - X\beta\right)^2$$

$$R^2 = \left(\frac{Cov\left(predicted, observed\right)}{Var\left(predicted\right)Var\left(observed\right)}\right)^2$$

$$R^2_{adj} = R^2 - \left(\frac{p}{n-p-1}\right)\left(1-R^2\right)$$

$$L\left(\beta\right) = \prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{\left(\left(y_i - x_i\beta\right)-\mu\right)^2}{2\sigma^2}}$$

$$Log\left(L\left(\beta\right)\right) = \frac{1}{\sigma\sqrt{2\pi}}\sum_{i=1}^{n}\frac{\left(\left(y_i - x_i\beta\right)-\mu\right)^2}{2\sigma^2}$$

$$2m - 2\ln\left(L\left(\beta\right)\right)$$

$$-2L\left(\beta\right) + m\ln\left(n\right)$$

$$\frac{\sum_{t=2}^{T}\left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{T}e_t^2}$$

$$U(\beta) = \sum_{k=1}^{K} \frac{\partial \mu_k}{\partial \beta} V_k^{-1} \left(Y_k - \mu_k(\beta) \right)$$

$$y = X\beta + \varepsilon$$

$$y = X\beta + Zu + \varepsilon$$

$$x_t = F_t x_{t-1} + \beta_t u_t + w_t$$

$$y_t = H_t x_t + v_t$$

$$y = X\beta + \varepsilon$$

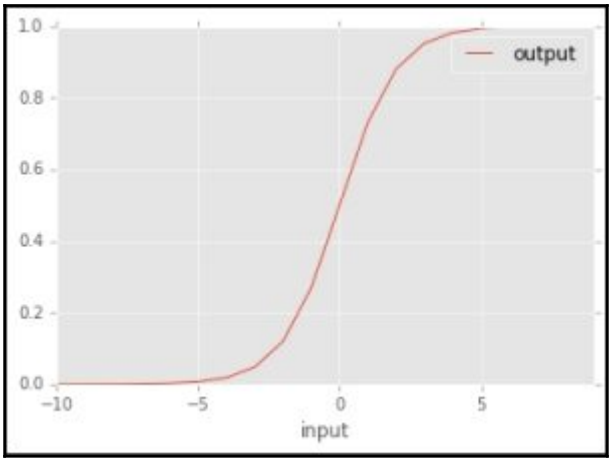
$$G(Y) = X\beta + \varepsilon$$

$$L(\beta) = \sum_{i=1}^{n} (y_i - x_i\beta)^2 + \alpha \sum_{j=1}^{m} \beta_j^2$$

$$L(\beta) = \sum_{i=1}^{n} (y_i - x_i\beta)^2 + \alpha \sum_{j=1}^{m} |\beta_j|$$

$$L(\beta) = \sum_{i=1}^{n} (y_i - x_i\beta)^2 + \alpha_1 \sum_{j=1}^{m} \beta_j^2 + \alpha_2 \sum_{i=1}^{n} |\beta|$$

Chapter 5: Putting Data in its Place – Classification Methods and Analysis



$$Y = 1 - logistic(z) = 1 - \frac{1}{1+e^{(-z)}} = \frac{1+e^{(-z)}}{1+e^{(-z)}} - \frac{1}{1+e^{(-z)}} = \frac{e^{(-z)}}{1+e^{(-z)}}$$

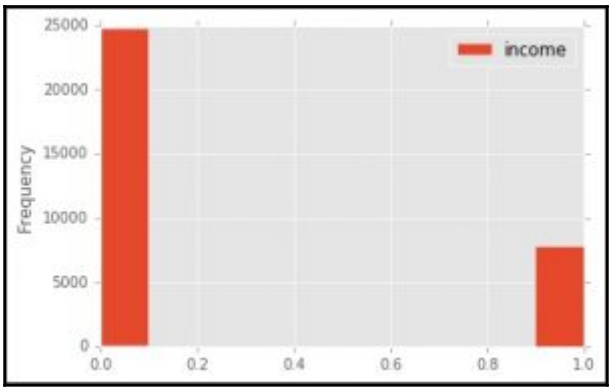
$$\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \log\left(\frac{\frac{1}{1+e^{(-z)}}}{\frac{e^{(-z)}}{1+e^{(-z)}}}\right) = \log(1) - \log\left(e^{(-z)}\right) = X\beta^T$$

$$e^{logit(z)} = \frac{logisitic(z)}{1-logisitic(z)} = \frac{\frac{1}{1+e^{(-z)}}}{\frac{e^{(-z)}}{1+e^{(-z)}}} = \frac{1}{e^{(-z)}} = e^{(z)} = e^{(X\beta^T)}$$

$$P(y_i = k) = \frac{e^{\left(x_i \beta_k^T\right)}}{\sum_{j=1}^{K} e^{\left(x_i \beta_j^T\right)}}$$

```
census = pd.read_csv('census.data',header=None)
census.head()
```

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	<=5
1	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	<=5
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	<=5
3	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40	United-States	<=5
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0	0	40	Cuba	<=5



$$P(y_i = 1) = \frac{1}{1+e^{(-z)}} = F(z_i)$$

$$P(y_i = 0) = \frac{e^{(-z)}}{1+e^{(-z)}} = 1 - F(z_i)$$

$$L(y_i) = \left(F(z_i)\right)^{y_i} \left(1 - F(z_i)\right)^{1-y_i}$$

$$L(Y) = \sum_{i=1}^{n} \left(y_i \log\left(F(z_i)\right) + (1-y_i)\log\left(1-F(z_i)\right)\right)$$

$$
\begin{aligned}
&\frac{\partial}{\partial\beta}\log\left(L\left(y_i\right)\right)=\frac{\partial}{\partial\beta}\left(y_i\log\left(F\left(z_i\right)\right)+\left(1-y_i\right)\log\left(1-F\left(z_i\right)\right)\right)=0\\
&=\frac{y_i}{F\left(z_i\right)}F'\left(z_i\right)+\frac{\left(1-y_i\right)}{\left(1-F\left(z_i\right)\right)}F'\left(z_i\right)\\
&=\left(\frac{y_i}{F\left(z_i\right)}-\frac{\left(1-y_i\right)}{\left(1-F\left(z_i\right)\right)}\right)F'\left(z_i\right)\\
&=\left(\frac{y_i}{F\left(z_i\right)}-\frac{\left(1-y_i\right)}{\left(1-F\left(z_i\right)\right)}\right)\left(\frac{1}{1+e^{\left(-x_i\beta^T\right)}}\right)\left(\frac{e^{\left(-x_i\beta^T\right)}}{1+e^{\left(-x_i\beta^T\right)}}\right)\left(-x_i\right)\\
&=\left(\frac{y_i}{\frac{1}{1+e^{\left(-x_i\beta^T\right)}}}-\frac{\left(1-y_i\right)}{\frac{e^{\left(-x_i\beta^T\right)}}{1+e^{\left(-x_i\beta^T\right)}}}\right)\left(\frac{1}{1+e^{\left(-x_i\beta^T\right)}}\right)\left(\frac{e^{\left(-x_i\beta^T\right)}}{1+e^{\left(-x_i\beta^T\right)}}\right)\left(-x_i\right)\\
&=\left(y_i\left(\frac{e^{\left(-x_i\beta^T\right)}}{1+e^{\left(-x_i\beta^T\right)}}\right)-\left(1-y_i\right)\left(\frac{1}{1+e^{\left(-x_i\beta^T\right)}}\right)\right)\left(-x_i\right)\\
&=\left(\frac{y_i\left(1+e^{\left(-x_i\beta^T\right)}\right)}{1+e^{\left(-x_i\beta^T\right)}}-\left(\frac{1}{1+e^{\left(-x_i\beta^T\right)}}\right)\right)\left(-x_i\right)\\
&=-\left(y_i-F\left(z_i\right)\right)x_i
\end{aligned}
$$

$$\beta_t=\beta_{t-1}+\alpha\left(y_i-F\left(z_i\right)\right)x_i$$

$$\frac{1}{2}\lambda\beta\beta^T$$

$$\beta_t=\beta_{t-1}+\alpha\left(y_i-F\left(z_i\right)\right)x_i-\lambda\beta_{t-1}$$

$$f\left(x^*\right)=f\left(x_t+\Delta x\right)$$

$$f\left(x^*\right)=f\left(x_t\right)+f'\left(x_t\right)\Delta x+\frac{1}{2}f''\left(x_t\right)\left(\Delta x\right)^2$$

$$\frac{df(x^*)}{d\Delta x} = f'(x) + 2\left(\frac{1}{2}\right) f''(x)(\Delta x) = 0$$

$$-f'(x) = f''(x)(\Delta x)$$

$$\Delta x = -\frac{f'(x)}{f''(x)}$$

$$\left(y_i - F(z_i)\right) x_i - \lambda\beta_t$$

$$\frac{\partial L(y_i)}{\partial\beta\partial\beta^T} = \frac{\partial}{\partial\beta}\left(\frac{\partial L(y_i)}{\beta^T}\right)$$

$$= \frac{\partial}{\partial\beta}\left(-\left(y_i - F(z_i)\right) x_i - \lambda\beta\right)$$

$$= \frac{\partial F(z_i) x_i}{\partial\beta} - \lambda = -\frac{\partial}{\partial\beta}\left(\frac{1}{1+e^{\left(-x_i\beta^T\right)}}\right) x_i - \lambda$$

$$= \left(\frac{1}{1+e^{\left(-x_i\beta^T\right)}}\right)\left(\frac{e^{\left(-x_i\beta^T\right)}}{1+e^{\left(-x_i\beta^T\right)}}\right)(x_i)(x_i)^T - \lambda$$

$$= (x_i) F(z_i)\left(1 - F(z_i)\right)(x_i)^T - \lambda$$

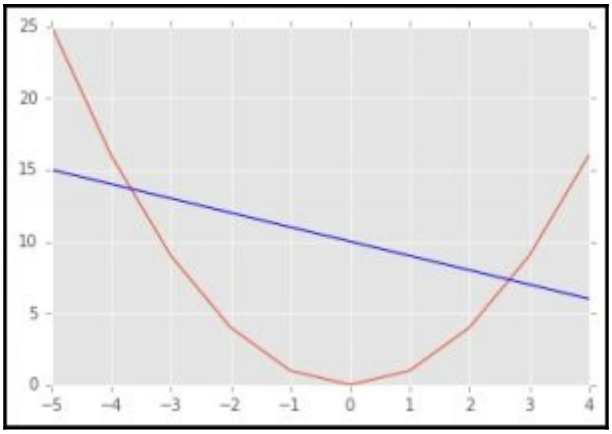
$$-\sum_{i=1}^{n} x_i \left(y_i - F(z_i)\right) - \lambda\beta$$

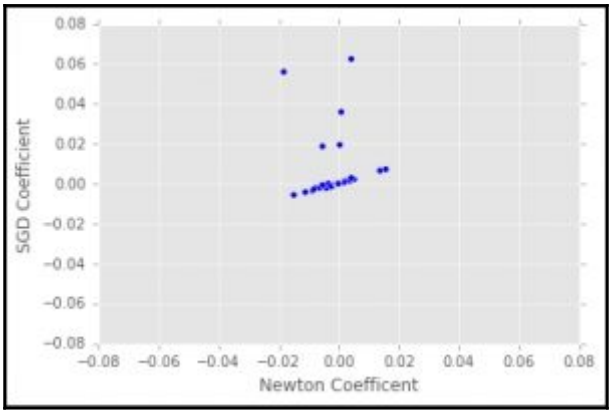
$$= (X)^T F(Z)\left(1 - F(Z)\right)(X) - I\lambda$$

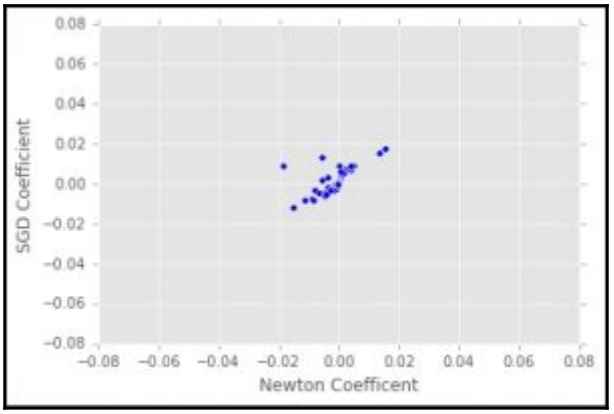
$$= X^T AX - I\lambda$$

$$\beta_t = \beta_{t-1} + \left(X^T AX - I\lambda\right)^{-1}\left(\sum_{i=1}^{n} x_i\left(y_i - F\left(z_i\right)\right) - \lambda\beta_t\right)$$

$$F\left(\alpha x_1 + \left(1-\alpha\right)x_2\right) \leq \alpha F\left(x_1\right) + \left(1-\alpha\right)F\left(x_2\right)$$

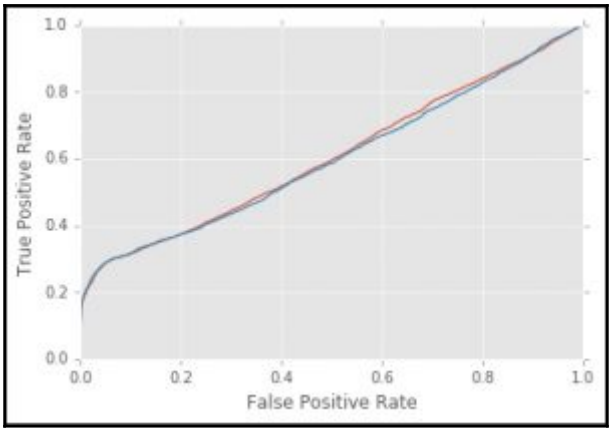


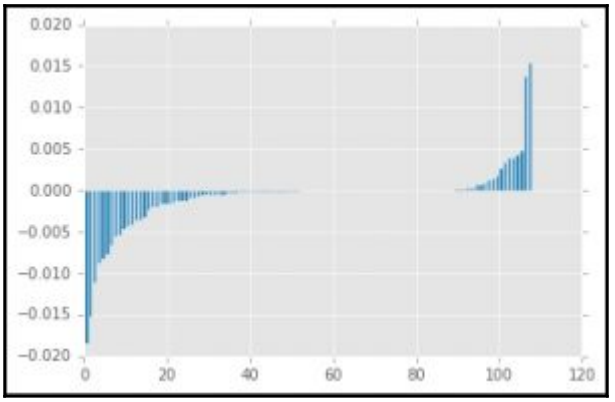




$$F_1 = 2\frac{(precision)(recall)}{(precision + recall)}$$

$$MCC = \frac{TPxTN - FPxFN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$





$$F(X) = X\beta^T + b$$

$$L(\beta, w, \alpha) = \frac{1}{2}\|B\|^2 - \sum_{i=1}^{n} \alpha_i \left(y_i \left(x_i \beta^T + b \right) - 1 \right)$$

$$\|\beta\| = \sqrt{\beta_1^2 + \beta_2^2 + \ldots + \beta_m^2}$$

$$\frac{\partial L}{\partial \beta} = \beta - \sum_{i=1}^{n} \alpha_i y_i x_i = 0$$

$$\beta = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$L(\beta, w, \alpha) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j - \sum_{i=1}^{n} \alpha_i \left(y_i x_i \left(\sum_{i=1}^{n} \alpha_i y_i x \right) - 1 \right)$$

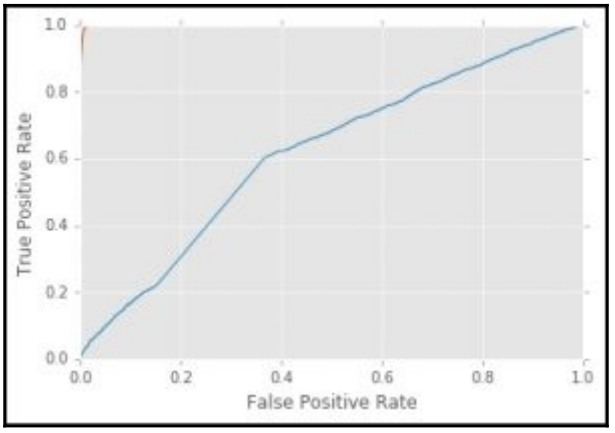
$$= \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j$$

$$\max\left(0,1-y_i\left(x_i\beta+b_i\right)\right)$$

$$<\varphi\left(x_i\right),\varphi\left(x_j\right)>$$

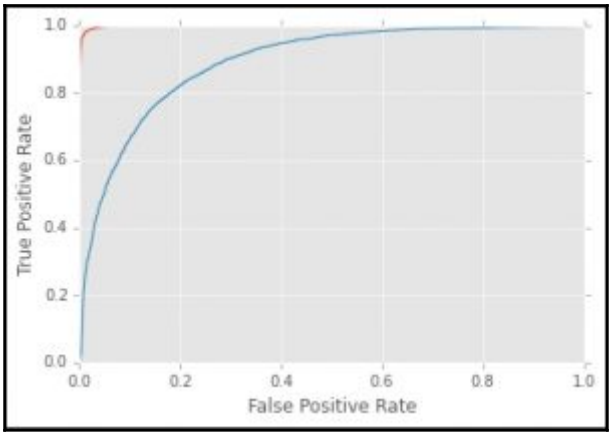
$$K\left(x_i,x_j\right)=e^{\frac{-\left\|x_i-x_j\right\|^2}{2\gamma^2}}$$

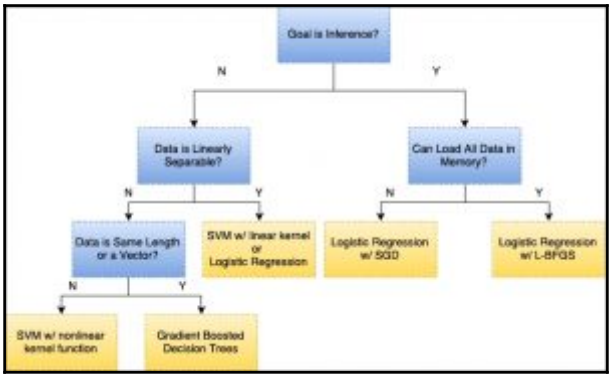
$$e^{\frac{-\left\|x_i-x_j\right\|^2}{2\gamma^2}}=e^{\frac{1}{2\gamma^2}(-x_i)^2(-x_j)^2}e^{2(x_ix_j)}=e^{\frac{1}{2\gamma^2}(-x_i)^2(-x_j)^2}\sum_{k=0}^{\infty}\frac{2^k x_i^k x_j^k}{k!}$$



$$\gamma_t=\frac{\arg\min}{\gamma}\sum_{i=1}^{n}L\left(y_i,F_{t-1}\left(x_i\right)+h_t\left(x_i\right)\right)$$

$$F_t=F_{t-1}\left(X\right)+\alpha\gamma_t h_t\left(X\right)$$





$$r_i = -\frac{\partial L\left(y_i, F\left(x_i\right)\right)}{\partial F\left(x_i\right)}$$

$$Y = X\beta^T$$

$$Y = logistic\left(X\beta^T\right)$$

$$logistic\left(X\beta^T\right) = logistic\left(z\right) = \frac{1}{1+e^{(-z)}}$$

Chapter 6: Words and Pixels – Working with Unstructured Data

$$W \leftarrow W \frac{X^T H}{W H^T H}$$

$$H \leftarrow H \frac{H X^T}{H H^T W^T}$$

$$\frac{1}{B(\alpha)} \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

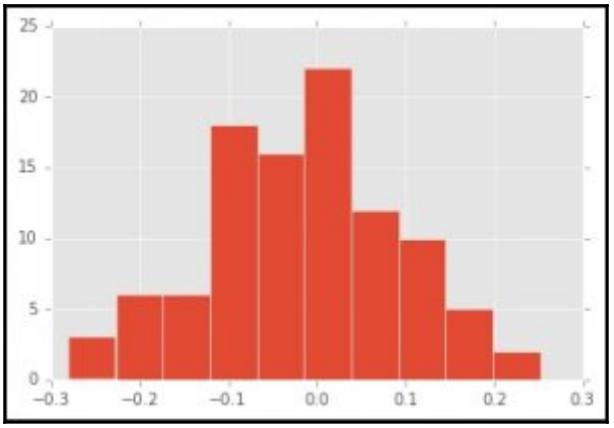
$$p = ui$$

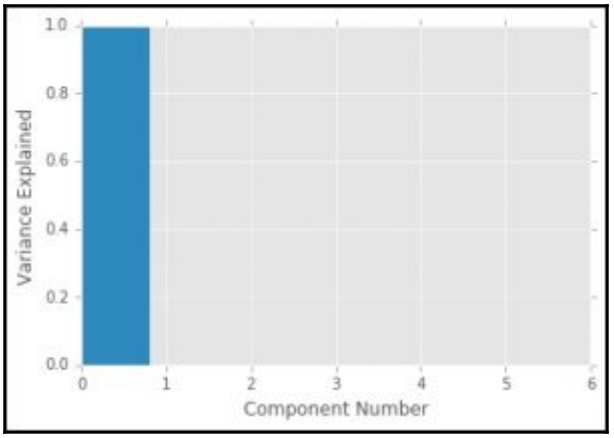
	label	text
0	ham	Go until jurong point, crazy.. Available only ...
1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...

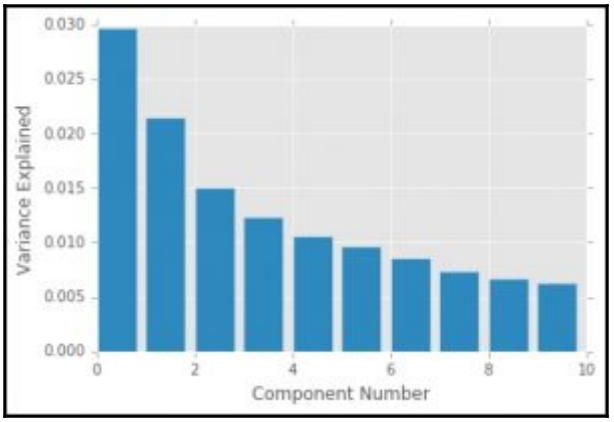
	label	text
0	ham	go until jurong point, crazy.. available only ...
1	ham	ok lar... joking wif u oni...
2	spam	free entry in 2 a wkly comp to win fa cup fina...
3	ham	u dun say so early hor... u c already then say...
4	ham	nah i don't think he goes to usf, he lives aro...

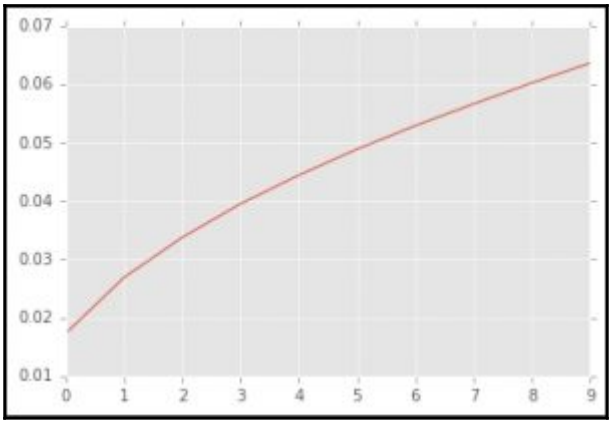
	label	text
0	ham	go jurong point , crazy.. avail bugi n grea...
1	ham	ok lar ... joke wif u oni ...
2	spam	free entri 2 wkli comp win fa cup final tkt...
3	ham	u dun say earli hor ... u c alreadi say ...
4	ham	nah n't think goe usf , live around though

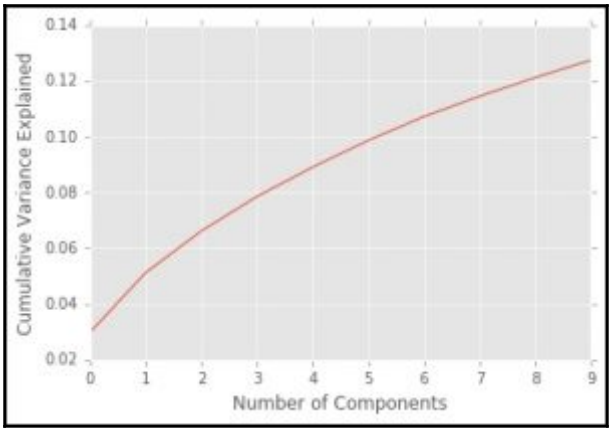
	label	text
0	ham	go jurong point , crazy.. avail bugi n grea...
1	ham	ok lar ... joke wif u oni ...
2	spam	free entri 2 wkli comp win fa cup final tkt...
3	ham	u dun say earli hor ... u c alreadi say ...
4	ham	nah n't think goe usf , live around though

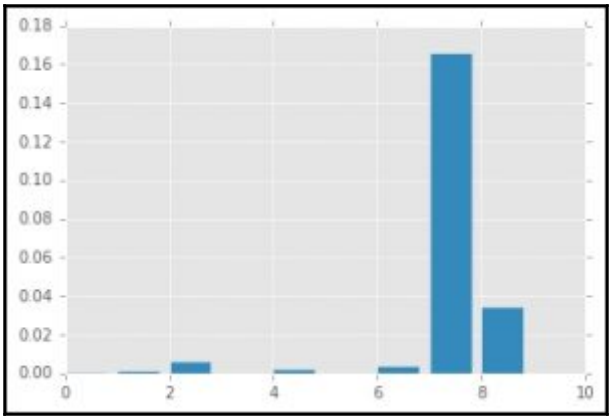

25
20
15
10
5
0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3

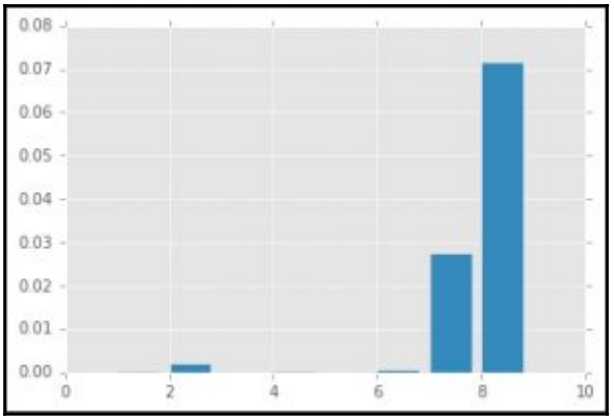

1.0
0.8
0.6
0.4
0.2
0.0
Variance Explained
0
1
2
3
4
5
6
Component Number

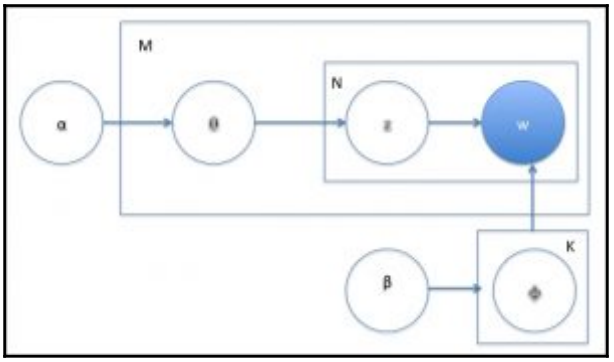

0.030
0.025
0.020
0.015
0.010
0.005
0.000
Variance Explained
0
2
4
6
8
10
Component Number

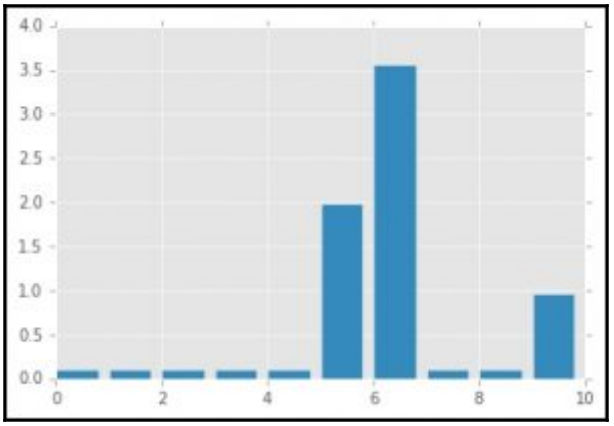

0.07
0.06
0.05
0.04
0.03
0.02
0.01
0 1 2 3 4 5 6 7 8 9

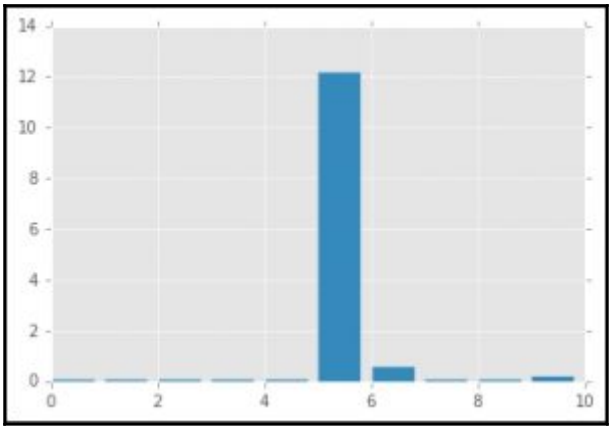

Cumulative Variance Explained
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0 1 2 3 4 5 6 7 8 9
Number of Components

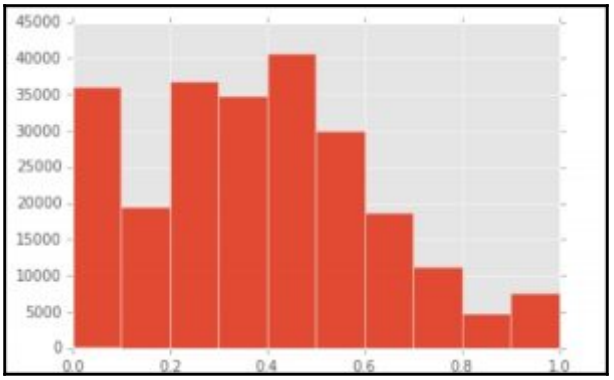

0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0 2 4 6 8 10

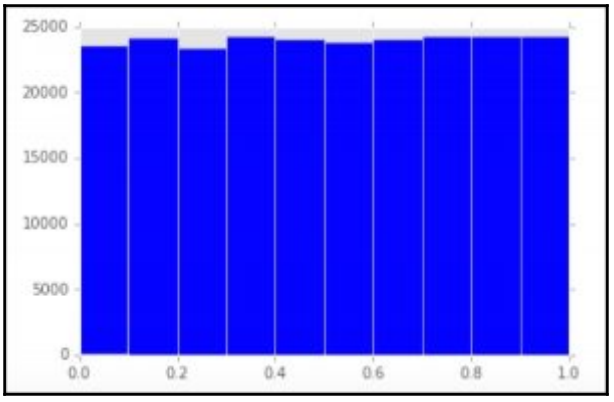

0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
2
4
6
8
10

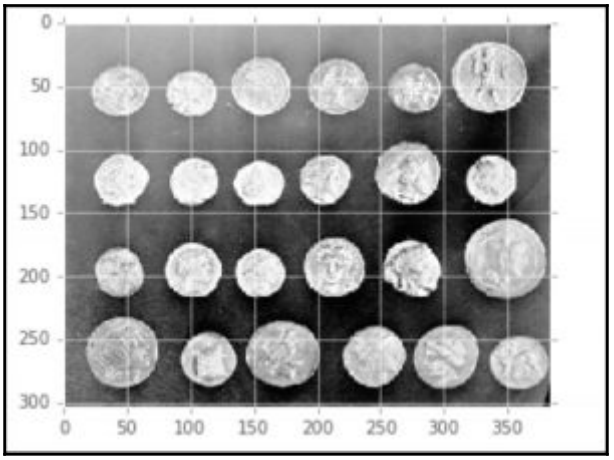

M
N
α
θ
z
w
K
β
φ

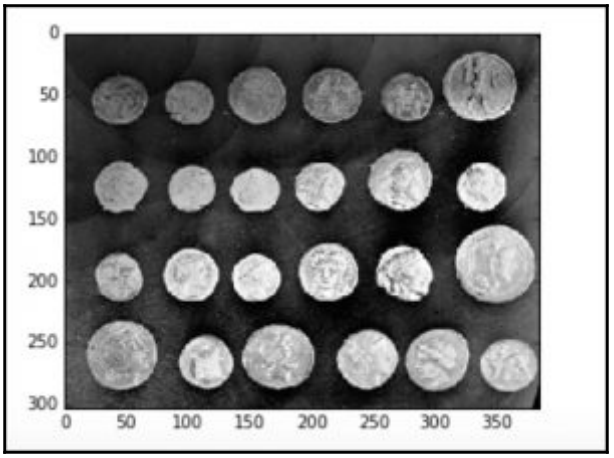

4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
2
4
6
8
10

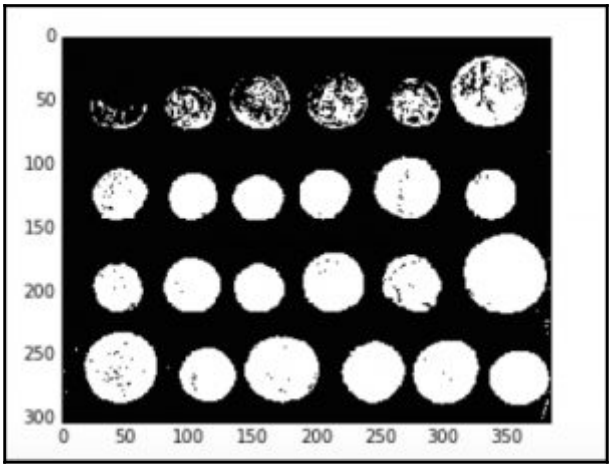

14
12
10
8
6
4
2
0
0
2
4
6
8
10

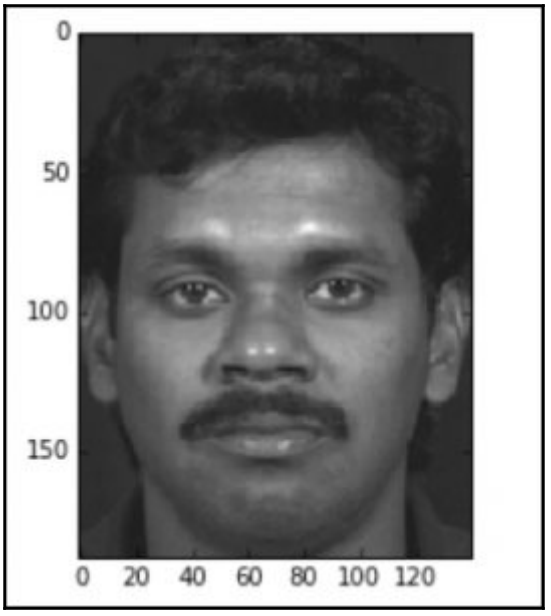

45000
40000
35000
30000
25000
20000
15000
10000
5000
0
0.0
0.2
0.4
0.6
0.8
1.0

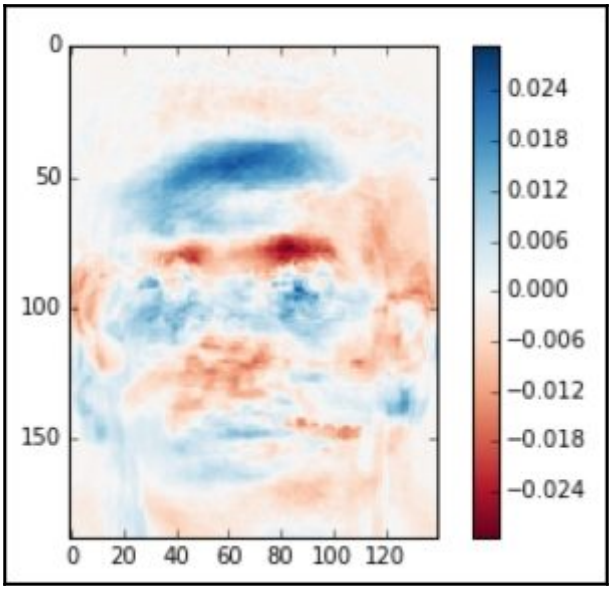

25000
20000
15000
10000
5000
0
0.0
0.2
0.4
0.6
0.8
1.0

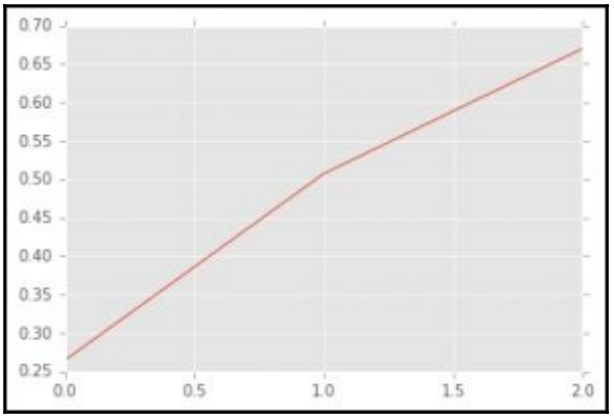

0
50
100
150
200
250
300
0
50
100
150
200
250
300
350

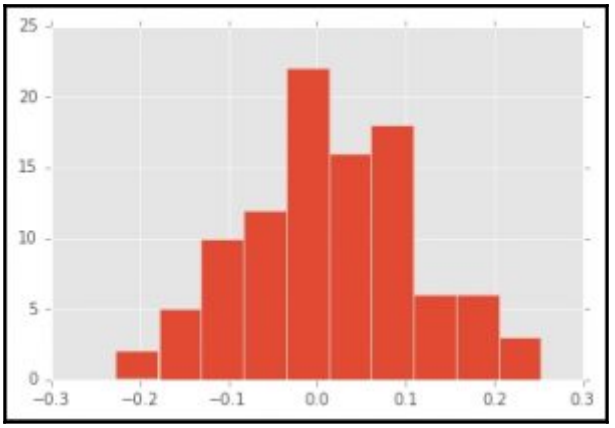

0
50
100
150
200
250
300
0
50
100
150
200
250
300
350

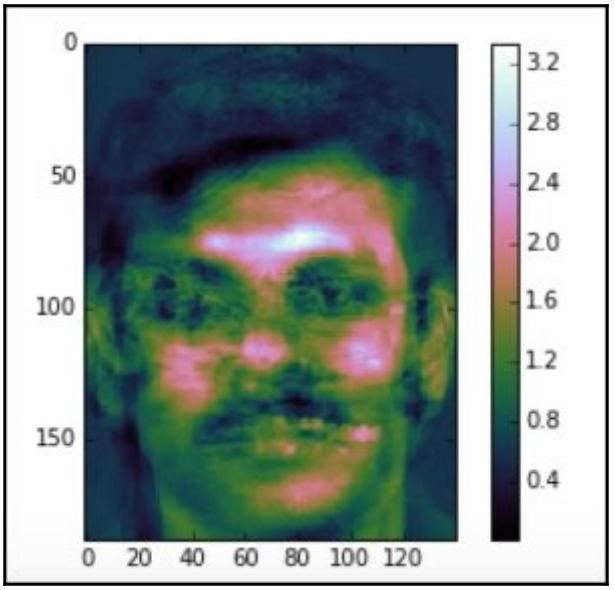

0
50
100
150
200
250
300
0
50
100
150
200
250
300
350


0
50
100
150
0
20
40
60
80
100
120


0
50
100
150
0 20 40 60 80 100 120
0.024
0.018
0.012
0.006
0.000
−0.006
−0.012
−0.018
−0.024


0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.0
0.5
1.0
1.5
2.0





$$tf - idf\left(t_i, d_j\right) = \frac{\sum_{k=1}^{V_j} 1_{t_i = v_k}}{\frac{\sum_{j=1}^{D} 1_{t_i \in V_j}}{D}}$$

$$Cov\left(x_i, x_j\right) = \frac{1}{n}\sum_{i,j=1}^{n}\left(x_i - \mu_i\right)\left(x_j - \mu_j\right)$$

$$X^T X = W^T \sigma W$$

$$X = CUR$$

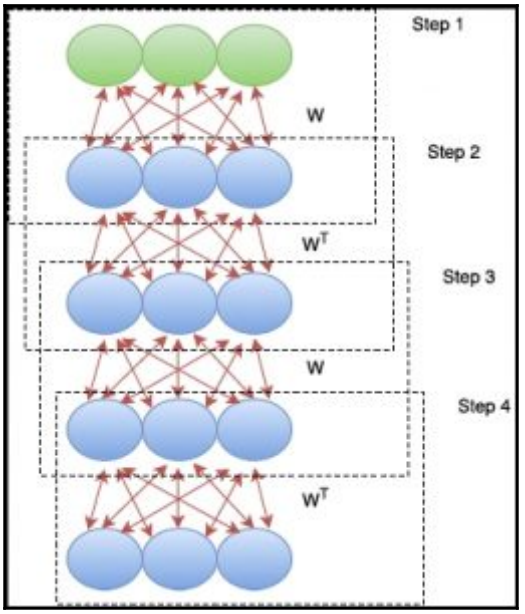
$$lv_j = \frac{1}{k}\sum_{i=1}^{k}\left(v_j^i\right)^2$$

$$X = WH$$

$$\left\|X - WH\right\|_F^2 = \left(X - WH\right)\left(X - WH\right)^T$$

$$\frac{\partial}{\partial W}\left(X - WH\right)\left(X - WH\right)^T \approx$$
$$-X^T H + WH^T H =$$
$$\frac{X^T H}{WH^T H}$$

Chapter 7: Learning from the Bottom Up – Deep Networks and Unsupervised Features



$$F(x)=\begin{cases}1 if\ wx>b\\0 otherwise\end{cases}$$

$$\Delta w=\alpha\left(y_i-F\left(x_i\right)\right)x_i$$

$$\frac{1}{n}\sum_{i=1}^{n}\left|y_i-F\left(x_i\right)\right|$$

$$F(z)=\frac{1}{1+e^{(-wx+b)}}=\frac{1}{1+e^{(-z)}}$$

$$E\left(x_i\right)=\frac{1}{2}\left(F\left(z_i\right)-y_i\right)^2$$

$$F(x) = e^{x^2}$$

$$\frac{\partial F}{\partial x} = \frac{\partial F \partial z}{\partial z \partial x}$$

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial F(z)_i} \frac{\partial F(z)_i}{z} \frac{\partial z}{\partial w_{ij}}$$

$$\frac{\partial E}{\partial F(z)_i} = 2\frac{1}{2}\left(F(z)_i - y\right)(1) = F(z)_i - y$$

$$\frac{\partial F(z)_i}{z} = \frac{(0)\left(1+e^{-z}\right)-(1)\left(-e^{-z}\right)}{\left(1+e^{-z}\right)^2} =$$

$$\frac{1}{\left(1+e^{-z}\right)} \frac{1-1+e^{-z}}{\left(1+e^{-z}\right)} = F(z)_i \left(1-F(z)_i\right)$$

$$\frac{\partial z}{\partial w_{ij}} = F(z)_j$$

$$w_{ij-new} = w_{ij-old} - \alpha \frac{\partial E}{\partial w_{ij}}$$

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial F(z)_j} \frac{\partial F(z)_j}{z} \frac{\partial z}{\partial w_{jk}}$$

$$\frac{\partial E}{\partial F(z)_j} = \sum_{i=1}^{n}\left(\frac{\partial E}{\partial z}\frac{\partial z}{\partial F(z)_i}\right) = \sum_{i=1}^{n}\left(\frac{\partial E}{\partial F(z)_i}\frac{\partial F(z)_i}{\partial z} w_{ij}\right)$$

$$\frac{\partial z}{\partial w_{jk}} = x_k$$

$$\frac{\partial z}{\partial w_{jk}} = \frac{\partial E}{\partial F(z)_j}\frac{\partial F(z)_j}{z}\frac{\partial z}{\partial w_{jk}} = \left(\sum_{i=1}^{n}\frac{\partial E}{\partial F(z)_i}\frac{\partial F(z)_i}{\partial z} w_{ij}\right) F(z)_j\left(1 - F(z)_j\right) x_k$$

$$F(z) = max(z, 0)$$

$$\frac{\partial F(z)}{\partial z} = \begin{cases} 1 if\ z > 0 \\ else\ 0 \end{cases}$$

$$F(z) = max(z, \alpha z)$$

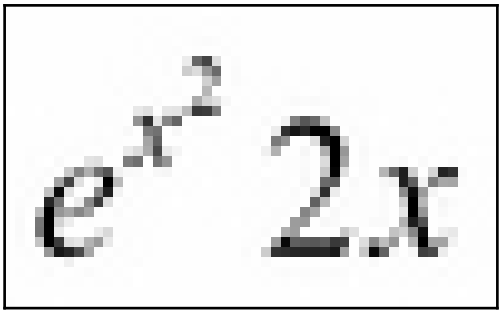
$$\frac{\partial F(z)}{\partial z} = \begin{cases} 1 if\ z > 0 \\ else\ \alpha \end{cases}$$

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{G_t + \varepsilon}} \frac{\partial E}{\partial w}$$

$$\frac{1}{T}\left(\frac{\partial E}{\partial w}\right)_t^2 = \gamma \frac{1}{T}\left(\frac{\partial E}{\partial w}\right)_{t-1}^2 + (1-\gamma)\left(\frac{\partial E}{\partial w}\right)_t^2$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{\frac{1}{T}\left(\frac{\partial E}{\partial w}\right)_t^2 \left(\begin{array}{c} 1 \\ 1-\beta_1 \end{array}\right) + \varepsilon}} \frac{\partial E}{\partial w}\left(\begin{array}{c} 1 \\ 1-\beta_2 \end{array}\right)$$



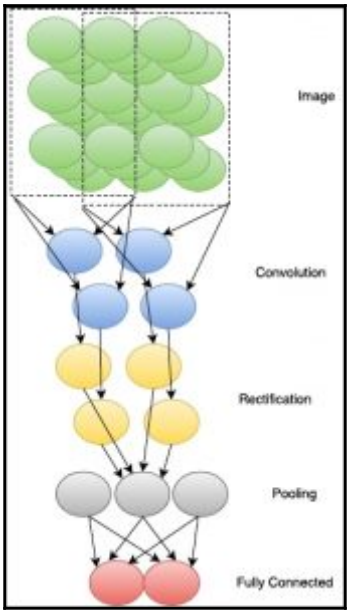
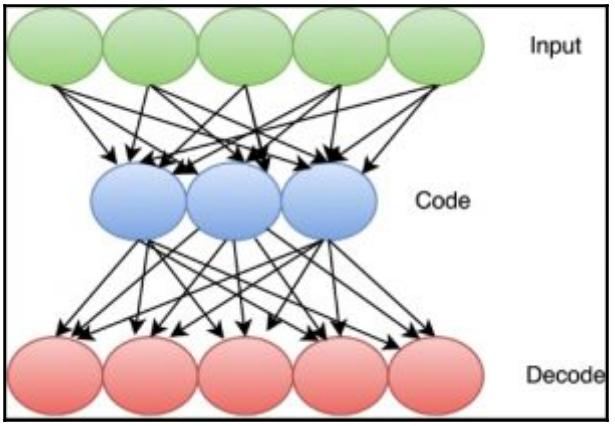
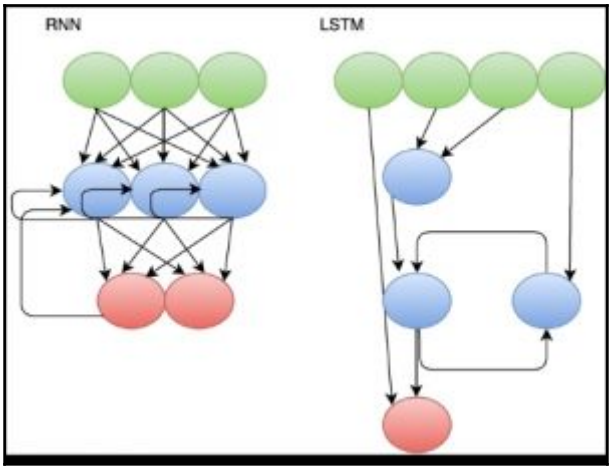
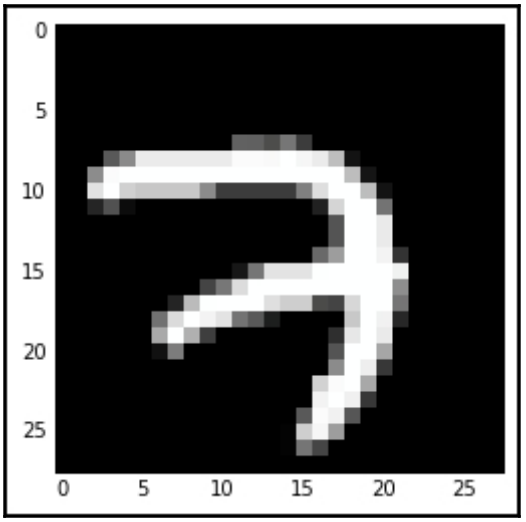
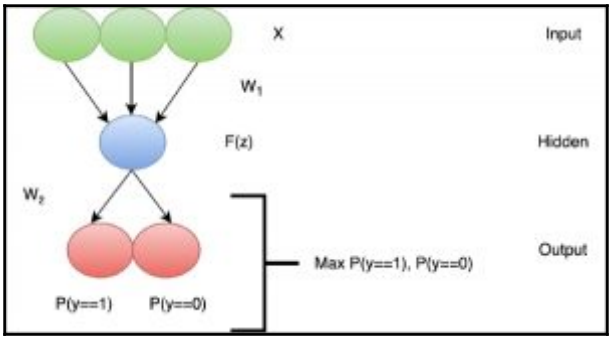


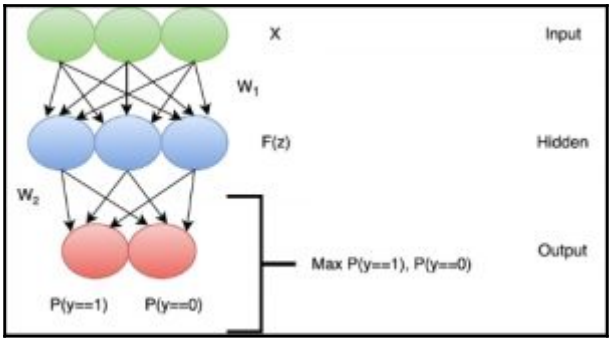
Image
Convolution
Rectification
Pooling
Fully Connected

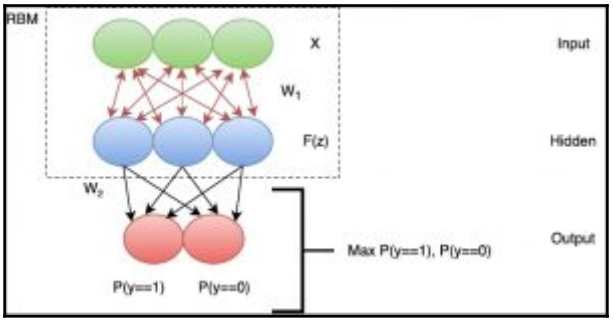

Input
Code
Decode

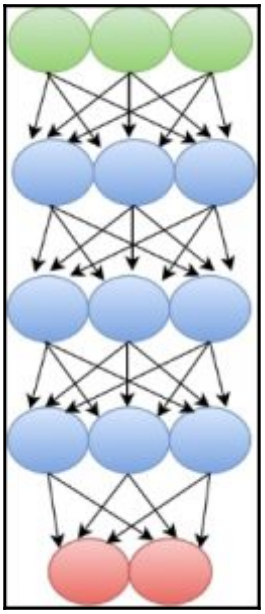

RNN
LSTM

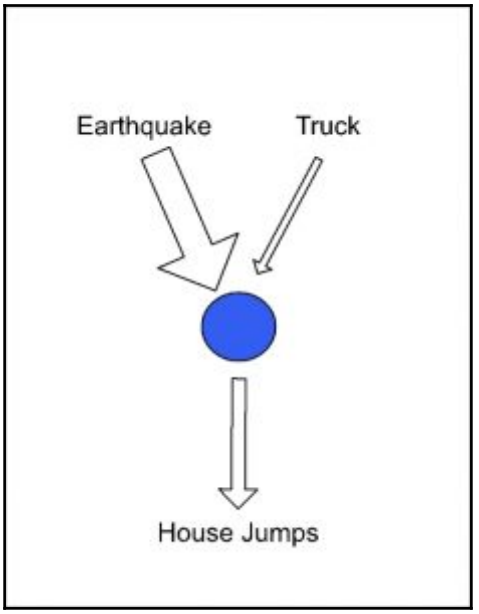

0
5
10
15
20
25
0
5
10
15
20
25


X
Input
W1
F(z)
Hidden
W2
Max P(y==1), P(y==0)
Output
P(y==1)
P(y==0)

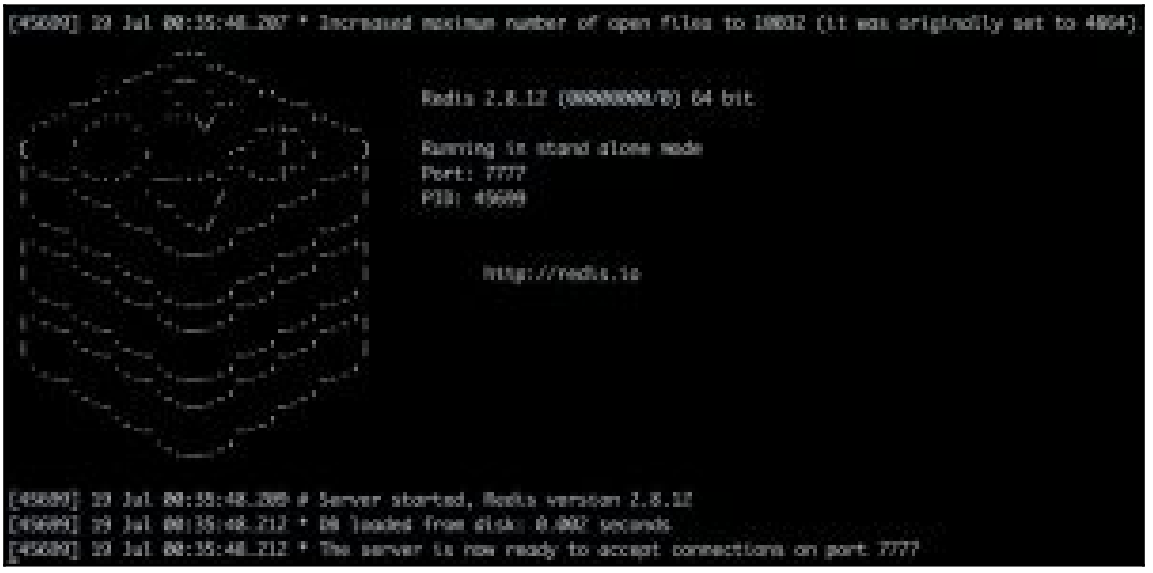

X
Input
W1
F(z)
Hidden
W2
Max P(y==1), P(y==0)
Output
P(y==1)
P(y==0)


RBM
X
Input
W1
F(z)
Hidden
W2
Max P(y==1), P(y==0)
Output
P(y==1)
P(y==0)





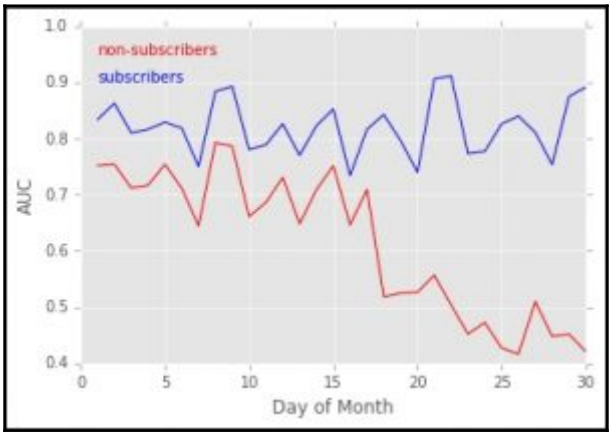
Chapter 8: Sharing Models with Prediction Services

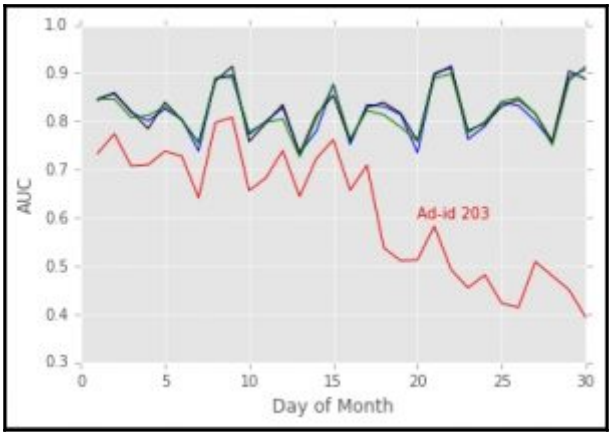


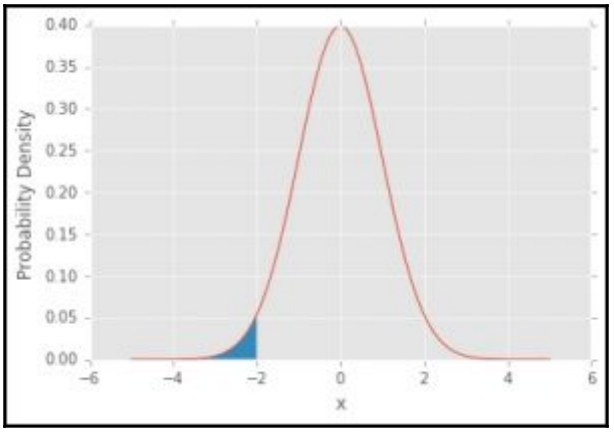
```
[19/Jul/2016:00:52:46] ENGINE Bus STARTING
[19/Jul/2016:00:52:46] ENGINE Started monitor thread 'Autoreloader'.
[19/Jul/2016:00:52:46] ENGINE Started monitor thread '_TimeoutMonitor'.
[19/Jul/2016:00:52:47] ENGINE Serving on http://0.0.0.0:5000
[19/Jul/2016:00:52:47] ENGINE Bus STARTED
```

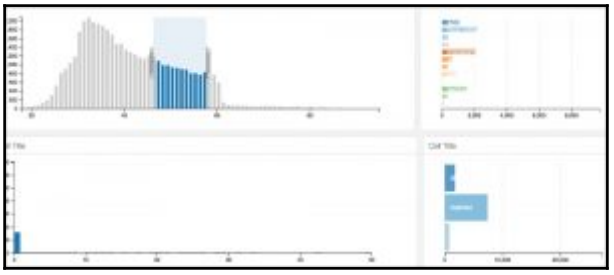
```
-------------- celery@lgml-jbabcock404 v3.1.22 (Cipater)
---- **** -----
--- * ***  * -- Darwin-15.5.0-x86_64-i386-64bit
-- * - **** ---
- ** ---------- [config]
- ** ---------- .> app:         modelservice:0x110cc87d0
- ** ---------- .> transport:   redis://localhost:6379//
- ** ---------- .> results:     redis://localhost:6379/
- *** --- * --- .> concurrency: 4 (prefork)
-- ******* ----
--- ***** ----- [queues]
 -------------- .> celery           exchange=celery(direct) key=celery


[2016-07-19 01:14:10,739: WARNING/MainProcess] celery@lgml-jbabcock404 ready.
```

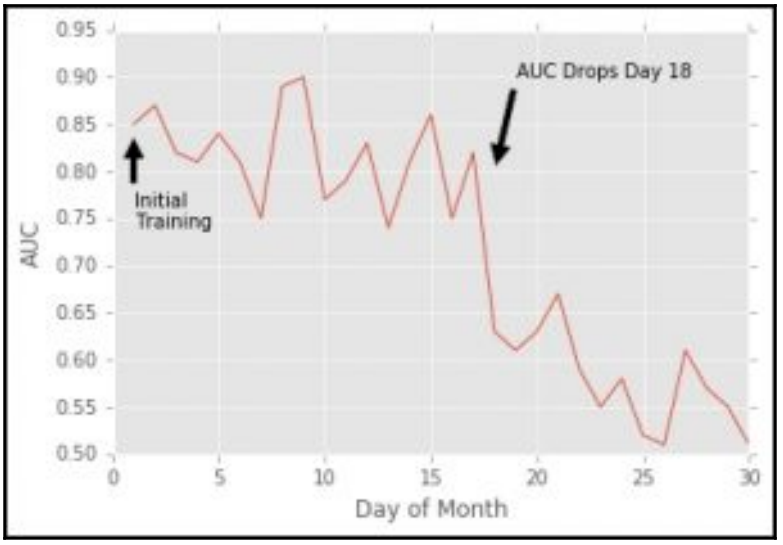
Chapter 9: Reporting and Testing – Iterating on Analytic Systems




AUC
Day of Month
Ad-id 203


Probability Density
x





$$T = \frac{Y_1 - Y_2}{S_{Y_1 - Y_2}}$$

$$S_{Y_1 - Y_2} = \sqrt{\frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}}$$

$$1.64 \leq \frac{10}{\sqrt{\frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}}}$$

$$\frac{100}{2.69} \geq \frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}$$

$$S_{Y_1 - Y_2} = \sqrt{\frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}}$$

$$1.64 \leq \frac{10}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$37.17 \geq \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

$$\frac{(predicted - observed)^2}{n}$$

$$s(i) = \frac{d'(i) - d(i)}{\max(d(i), d'(i))}$$

Table of Contents

Index