

A Study of FDA Medical Device Recalls

John T Giles

Mathematics and Computer Science Department

Lewis University

Romeoville, IL USA

johntgiles@lewisu.edu

Abstract—this article presents the results from the data analysis and mining of the FDA medical device recalls for the year 2015. The data is analyzed and mined to help predict if any of the Firms in the data will have medical device recalls in 2016.

Keywords—Medical Device Recall; Data Mining;

I. INTRODUCTION

The company who currently employs me manufactures a diverse line of products. Medical devices is one of the line of products' it makes. During my tenure at this company I've had the opportunity to work with a team whose task is to process all company manufactured medical devices removed from patients for any reason. All the information this team collects about these devices must be reported to the FDA. The FDA calls this process medical device tracking. This experience piqued the author's interest in finding out more information about medical devices reported to the FDA. The current presidential administration has mandated governmental agencies provide as much data as possible freely to the public. The FDA is one of the agencies which has released data available for analysis by anyone willing to retrieve it. The data is available on data.gov [1].

II. THE DATA

A. Dataset Creation

Data is available from the FDA for many different purposes including medical device tracking as well as medical device recalls. Medical device recalls were chosen for analysis. Although Data is available for medical device recalls starting from November 1, 2002 to the present, only data for the last full year was downloaded. FDA medical device recall data for the year 2015 was used to create the datasets analyzed for this project.

The main reason for the restriction of the dataset to one year was the download method provided by the FDA. The FDA only allows for the extraction of datasets up to 500 records at a time. The tool provided by the FDA was designed as a search tool to discover small sets of data based on a few parameters such as manufacturer name, product name, or date range. It did not appear to be designed for whole sale download and analysis of all the data available.

It was unknown how many medical device recalls occurred in 2015, but it appeared to be more than 500. Therefore, it was not possible to retrieve all the 2015 medical device recalls in one dataset. Through trial and error it was discovered there were 500 or less medical device recalls for each month in the year 2015. A dataset for each month was downloaded and then consolidated into a single dataset. The downloaded datasets were in a comma separated format (csv). The consolidated dataset was in a Microsoft Excel format. The final datasets used for analysis were extracted from the Excel consolidated dataset in csv and tab delimited formats.

B. Dataset Issues

Combining the monthly files into one file was accomplished in Excel. Excel was easily able to load and read the data downloaded from the FDA. A cursory review of the data in the Excel file showed no mismatched rows or columns.

The consolidated file failed to load into the Weka application and the Orange application. However, the information provided by the Weka application included very specific data issue details and the line number of the record containing the issue. On the other hand the Orange application failed to load providing little to no information. The Weka application became the main application used to discover data load issues.

Non printable control characters from the source datasets caused the initial data load issues. The non-printable control characters were not apparent in the Excel format. Excel was able to include the non-printable control characters in the appropriate columns. The subsequent consolidated csv and tab delimited files interpreted the non-printable control characters as intended and reformatted the records incorrectly.

The non-printable control characters appeared to be carriage returns and line feeds. The following function was created in Excel to help remove these characters:

```
=TRIM(CLEAN(SUBSTITUTE(Original!C2,CHAR(160),CHAR(32))))
```

Other issues resolved included: format issues, double quote issues, single quote issues, and several other hard to find data characters the Weka and Orange applications misinterpreted during the load process.

Data in the Product Description and Manufacturer Recall Reason columns continued to present load problems and were eliminated from the dataset. Two data rows continued to present data load issues and were also removed from the dataset.

Fig. 1. Weka Preprocess Screen Print

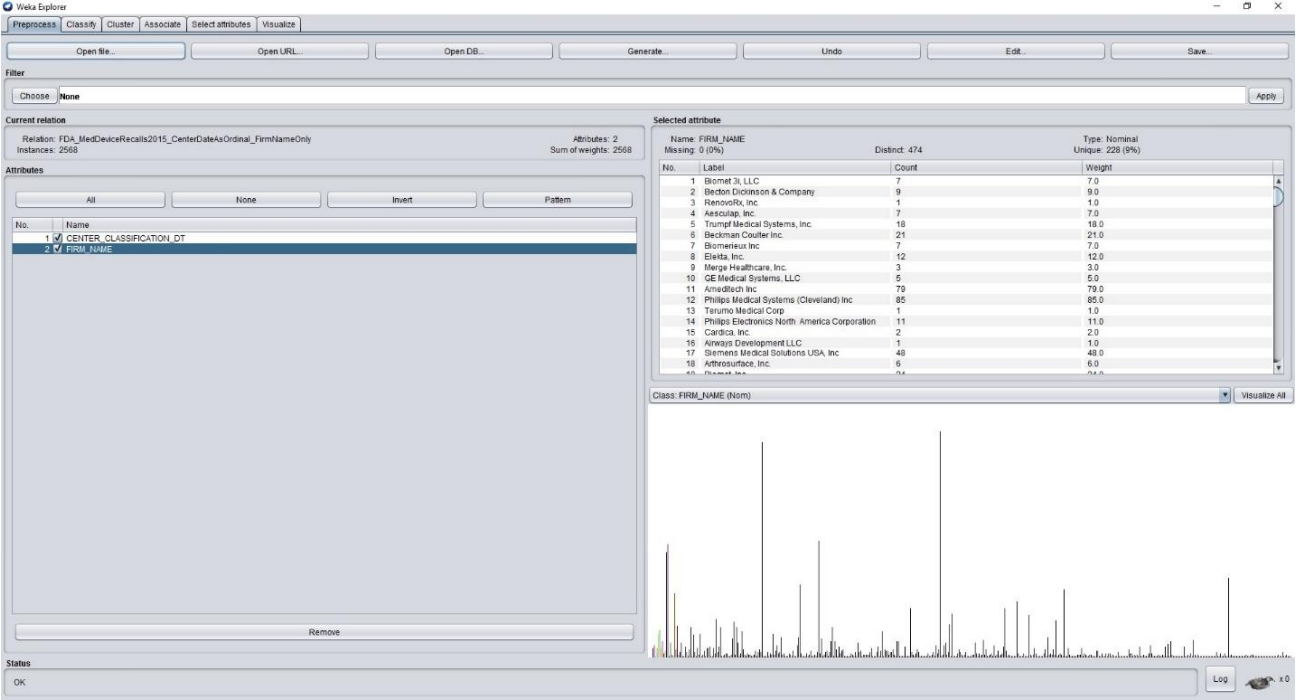


Table 1. Consolidated Excel dataset example excluding the WEB_ADDRESS, RECALL_NUMBER, PRODUCT_DESCRIPTION, and MANUFACTURER_RECALL_REASON attributes

TRADE_NAME	RECALL_CLASS	CENTER	CENTER_CLASSIFICATION_DT	POSTED_INTERNET_DT	TERMINATION_DT	FIRM_NAME
BIOMET 3i ART 1036 Dental Implant Radiographic Transparency	2	CDRH	2015/12/31	2015/12/31 21:00:35		Biomet 3i, LLC
BD Vacutainer PLUS K2EDTA Blood Collection Tubes.	2	CDRH	2015/12/31	2015/12/31 21:00:32		Becton Dickinson & Company
BD Vacutainer Brand Plus EDTA Blood Collection Tubes	2	CDRH	2015/12/31	2015/12/31 21:00:32		Becton Dickinson & Company
BD Vacutainer PLUS K2EDTA Plus Blood Collection Tubes	2	CDRH	2015/12/31	2015/12/31 21:00:32		Becton Dickinson & Company

III. OBJECTIVES

A. Primary Objective

Use the Orange or Weka application to analyze and mine a dataset. Ensure to perform at least one of the following data mining activities to answer the corresponding questions:

- Clustering algorithms
 - What clusters were discovered
 - Could these clusters correspond to anything (i.e. some class of data)
- Association algorithms
 - What rules were found
 - How many rules were found
 - What is the support, confidence, or interest of these rules
 - Was any rule surprising
- Supervised machine learning algorithms
 - Is there anything interesting that could be predicted based on the data
 - What is the prediction performance (i.e. accuracy, precision, recall, etc.)

B. Secondary Objective

Predict if any of the firms found in the 2015 FDA Medical Device Recall dataset will have a device recall in 2016.

IV. DISCOVERY

A. Data Statistics

There were 2569 FDA medical device recall records in 2015. 471 firms were associated to all of the FDA medical device recalls in 2015. 1806 medical devices for these 471 firms were involved in one or more FDA medical device recalls.

B. Data Analysis

There are 11 attributes in this dataset. A description of these attributes was not found on the FDA website. The descriptions are derived from my past work experience with similar data, the data analysis process, and information gathered from the FDA website.

Many of the FIRM_NAME attribute values contain duplicate uncommon words or slight variations in punctuation. These slight differences appear to indicate duplicate firms or firms with multiple divisions being involved in the product recalls.

Table 2. Example of uncommon word repeating in firm name

Abbott Electrophysiology
Abbott Laboratories
Abbott Laboratories, Inc
Abbott Medical Optics, Inc.
Abbott Molecular
Abbott Point Of Care Inc.

Table 3. Example of slight variations in punctuation

Siemens Healthcare Diagnostics Inc
Siemens Healthcare Diagnostics Inc.
Siemens Healthcare Diagnostics, Inc
Siemens Healthcare Diagnostics, Inc.

These possible firm duplicates could skew the results of the objectives by reducing the count of the medical device recalls for the firms involved. No data alterations were made to adjust for this issue.

C. Data Mining

The entire dataset was initially mined using both the Weka and Orange applications. During the mining process it was determined only two of the eleven attributes were necessary to meet the secondary objective. These two attributes were also adequate to satisfy the primary objectives. The “Excluded Attributes” section below describes why attributes were ultimately eliminated from the data mining process. The exclusion of the attributes and reasons given are based on inferences made during the initial data mining activities.

Table 4. Dataset Attributes

Attribute	Type	Description
WEB_ADDRESS	String	Location on the FDA website for the document describing the details of the medical device recall represented by the current record in the dataset
RECALL_NUMBER	Discrete	Unique record identifier
PRODUCT_DESCRIPTION	String	Description of the product as described by the product manufacturer
TRADE_NAME	Discrete	Name of the product as defined by the product manufacturer
RECALL_CLASS	Discrete	Medical device recall classification as defined by the FDA
CENTER	Discrete	
CENTER_CLASSIFICATION_DT	Continuous	Approximate date recall reported to FDA
POSTED_INTERNET_DT	Continuous	Approximate date recall details posted to the searchable database on the FDA website
TERMINATION_DT	Continuous	Approximate date the medical device recall ended
FIRM_NAME	Discrete	Manufacturer name
MANUFACTURER_RECALL_REASON	String	Manufacturer description of why the medical device was recalled

1) *Excluded attributes*

a) *WEB_ADDRESS:*

The WEB_ADDRESS information may be useful for deeper analysis into answering the secondary objective, but would require more time than allowed for this project. It would also require data retrieval methods not provided by the FDA Medical Device Recall application.

b) *RECALL_NUMBER:*

The RECALL_NUMBER data element provides no useful information to answer either the primary or secondary objectives. It clutters up the results. The usefulness of this data is for communicating with the FDA data systems, which is not necessary to meet the primary or secondary objectives.

c) *PRODUCT_DESCRIPTION:*

The PRODUCT_DESCRIPTION attribute provides useful information for the firms associated with the medical device recall record. This data element may be useful for a deeper analysis of the dataset with regards to the firms themselves. This information could be used to communicate with the firms to get more information about the recalls, which may help verify the results of the secondary objective. Unfortunately, the data in this attribute caused load issues, and did not appear to add any value to the data mining process.

d) *TRADE_NAME:*

The TRADE_NAME attribute is similar to the PRODUCT_DESCRIPTION and provides useful information for the firms associated with the medical device recall record. The same explanation found in the PRODUCT_DESCRIPTION above applies to the TRADE_NAME as well.

e) *RECALL_CLASS:*

The RECALL_CLASS attribute is not useful for meeting the primary or secondary objectives of this project. It may be useful for a deeper analysis of the dataset and for communicating with the FDA to extract more information from them about the medical device recalls.

f) *CENTER:*

The CENTER is an unknown attribute. It appears to be a representation of the location, virtual or real, in which the medical device recall was first reported. Every record in the dataset contains the same value. This information is not necessary to for meeting the primary or secondary objectives.

g) *POSTED_INTERNET_DT:*

The POSTED_INTERNET_DT attribute appears to be the date in which some information regarding a medical device recall record was posted to the internet. It is unclear where on the internet this posting took place. In addition, only 2330 of the 2568 medical device records contain a value. This is a significant number for analyzing. However, it was excluded due to the fact there are 238 records without a POSTED_INTERNET_DT value. Instead the

CENTER_CLASSIFICATION_DT attribute contains a value for all medical device recall records in the dataset.

h) *TERMINATION_DT:*

The TERMINATION_DT attribute is useful for a deeper analysis of the dataset. However, it does not provide any useful information to help meet the primary or secondary objectives.

i) *MANUFACTURER_RECALL_REASON:*

The MANUFACTURER_RECALL_REASON attribute is similar to the PRODUCT_DESCRIPTION and TRADE_NAME attributes and provides useful information for the firms associated with the medical device recall record. The same explanation found in the PRODUCT_DESCRIPTION above applies to the MANUFACTURER_RECALL_REASON as well.

2) *Included Attributes*

There are only two attributes used to help meet the primary and secondary objectives. The author came to the conclusion these two attributes together may help provide enough information to infer with some level of certainty whether a firm will have a medical device recalled in 2016. These two attributes are the FIRM_NAME and the CENTER_CLASSIFICATION_DT.

The CENTER_CLASSIFICATION_DT is a continuous attribute type. This presented problems for clustering and in particular associate rule creation. It was converted to a discrete value for data mining purposes. The month was extracted from the date and defined as a discrete attribute type.

3) *Clustering Algorithms*

The Weka application was used to process the final consolidated dataset through the Canopy, Simple K Means, and Hierarchical algorithms. The complete results are found in the Appendix.

The Canopy algorithm found 83 clusters. These clusters could correspond to the FIRM_NAME attribute. The FIRM_NAME with the highest percentage of CENTER_CLASSIFICATION_DT attributes around it could be a good candidate for having medical device recalls in 2016.

The Simple K Means algorithm found 2 clusters. The clusters appear to correspond to CENTER_CLASSIFICATION_DT mean value and high occurring FIRM_NAME attributes. The FIRM_NAME attribute clusters could be good candidates for having medical device recalls in 2016.

The Hierarchical algorithm doesn't provide easily interpretable results. It would require more analysis to determine if the information provided is relevant to the secondary objective.

4) *Association Algorithms*

The Orange application was used to process the final consolidated dataset through an association algorithm. The complete results are found in the Appendix.

Table 5. Association Algorithm Rules Produced

Rules	Supp	Conf	Lift	Lev	Strg	Cov
CENTER_CLASSIFICATION_DT=12						
CENTER_CLASSIFICATION_DT=12 -> FIRM_NAME=Customed, Inc	0.012	0.142	2.148	0.006	0.78	0.085
CENTER_CLASSIFICATION_DT=10						
CENTER_CLASSIFICATION_DT=10						
CENTER_CLASSIFICATION_DT=10 -> FIRM_NAME=Stradis Healthcare, LLC.	0.012	0.138	6.961	0.01	0.235	0.085
CENTER_CLASSIFICATION_DT=11						
CENTER_CLASSIFICATION_DT=11 -> FIRM_NAME=Customed, Inc	0.011	0.133	2.014	0.005	0.81	0.082
CENTER_CLASSIFICATION_DT=05						
CENTER_CLASSIFICATION_DT=05 -> FIRM_NAME=Resource Optimization & Innovation Llc	0.01	0.12	3.496	0.007	0.406	0.085
CENTER_CLASSIFICATION_DT=04						
CENTER_CLASSIFICATION_DT=04 -> FIRM_NAME=Covidien LLC	0.011	0.133	2.114	0.006	0.771	0.082
CENTER_CLASSIFICATION_DT=04 -> FIRM_NAME=Navilyst Medical, Inc	0.011	0.138	5.91	0.009	0.286	0.082
CENTER_CLASSIFICATION_DT=06						
CENTER_CLASSIFICATION_DT=06 -> FIRM_NAME=Resource Optimization & Innovation Llc	0.011	0.129	3.752	0.008	0.419	0.082
CENTER_CLASSIFICATION_DT=01						
CENTER_CLASSIFICATION_DT=01 -> FIRM_NAME=Covidien LLC	0.011	0.123	1.955	0.005	0.714	0.088
CENTER_CLASSIFICATION_DT=01 -> FIRM_NAME=Customed, Inc	0.012	0.132	1.996	0.006	0.749	0.088
CENTER_CLASSIFICATION_DT=03						
CENTER_CLASSIFICATION_DT=03 -> FIRM_NAME=Covidien LLC	0.011	0.129	2.045	0.006	0.747	0.085
CENTER_CLASSIFICATION_DT=03 -> FIRM_NAME=Customed, Inc	0.012	0.138	2.088	0.006	0.783	0.085
CENTER_CLASSIFICATION_DT=02						
CENTER_CLASSIFICATION_DT=02 -> FIRM_NAME=Ameditech Inc	0.01	0.131	4.269	0.008	0.399	0.077
CENTER_CLASSIFICATION_DT=02 -> FIRM_NAME=Customed, Inc	0.011	0.136	2.06	0.005	0.859	0.077
CENTER_CLASSIFICATION_DT=02 -> FIRM_NAME=Covidien LLC	0.011	0.146	2.322	0.006	0.818	0.077
CENTER_CLASSIFICATION_DT=08						
CENTER_CLASSIFICATION_DT=08 -> FIRM_NAME=Philips Medical System (Cleveland) Inc	0.011	0.124	3.759	0.008	0.392	0.085

The Association Rules algorithm produced 15 rules. None of the rules were very surprising. The parameters were set wide open with minimal support and confidence allowing for a level as low as 1. The maximal number of rules was set 100000. The rules with the highest support and a high level of confidence may predict a FIRM_NAME having a recall in a particular month. Notice the FIRM_NAME Customed, Inc appears frequently in the rules. Also notice the FIRM_NAME **Customed Inc** has recalls throughout the year, including the 12th month. This might predict the FIRM_NAME having a recall in the same month of the next year.

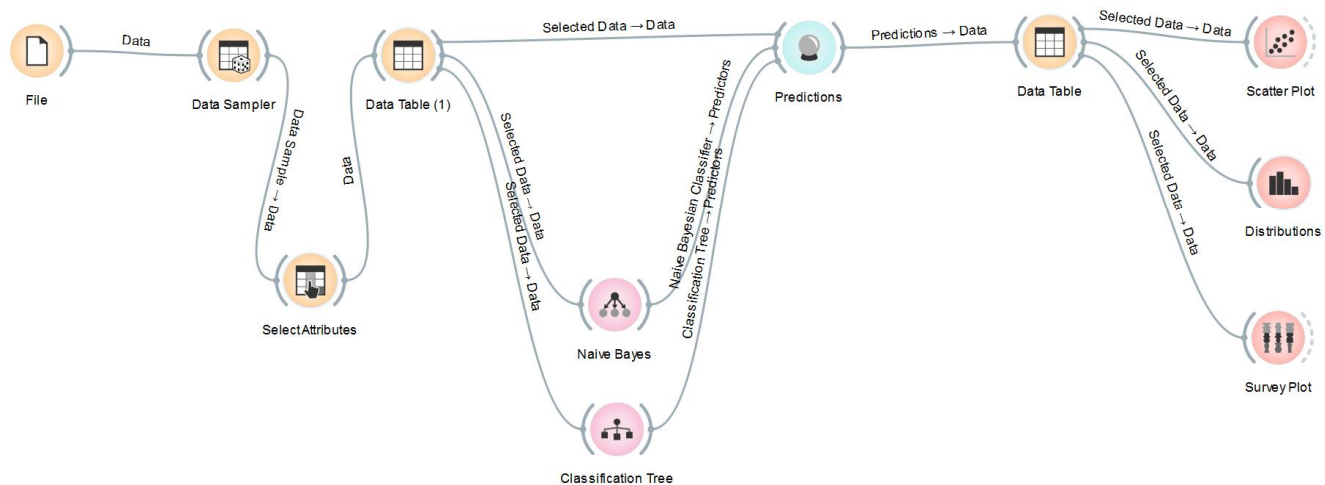
5) Prediction Algorithm

A prediction algorithm was built in the Orange application (Fig. 2) to assist in determining which companies might have a recall in 2016.

Table 6. Association Algorithm Rule Statistics

Rule Statistics
Total number of rules: 15
Support: 1% - 1%
Confidence: 11% - 14%
Number of rules in the graph: 15
Support: 1% - 1%
Confidence: 11% - 14%
Selected rules: 15
Support: 1% - 1%
Confidence: 12% - 15%

Fig 2. Prediction Algorithm built in Orange



The prediction algorithm produced results depicting two FIRM_NAME attributes with a high number of medical device recalls corresponding to CENTER_CLASSIFICATION_DT attributes. The CENTER_CLASSIFICATION_DT attributes represents months in the later part of 2015. These FIRM_NAMES may be candidates for medical device recalls in 2016. Fig. 3 is a distribution graph of those results. Notice the large yellow areas in the 8 to 12 range of the CENTER_CLASSIFICATION_DT. The FIRM_NAMES Customed, Inc and Covidien LLC are represented by the large yellow areas in the distribution.

V. CONCLUSION

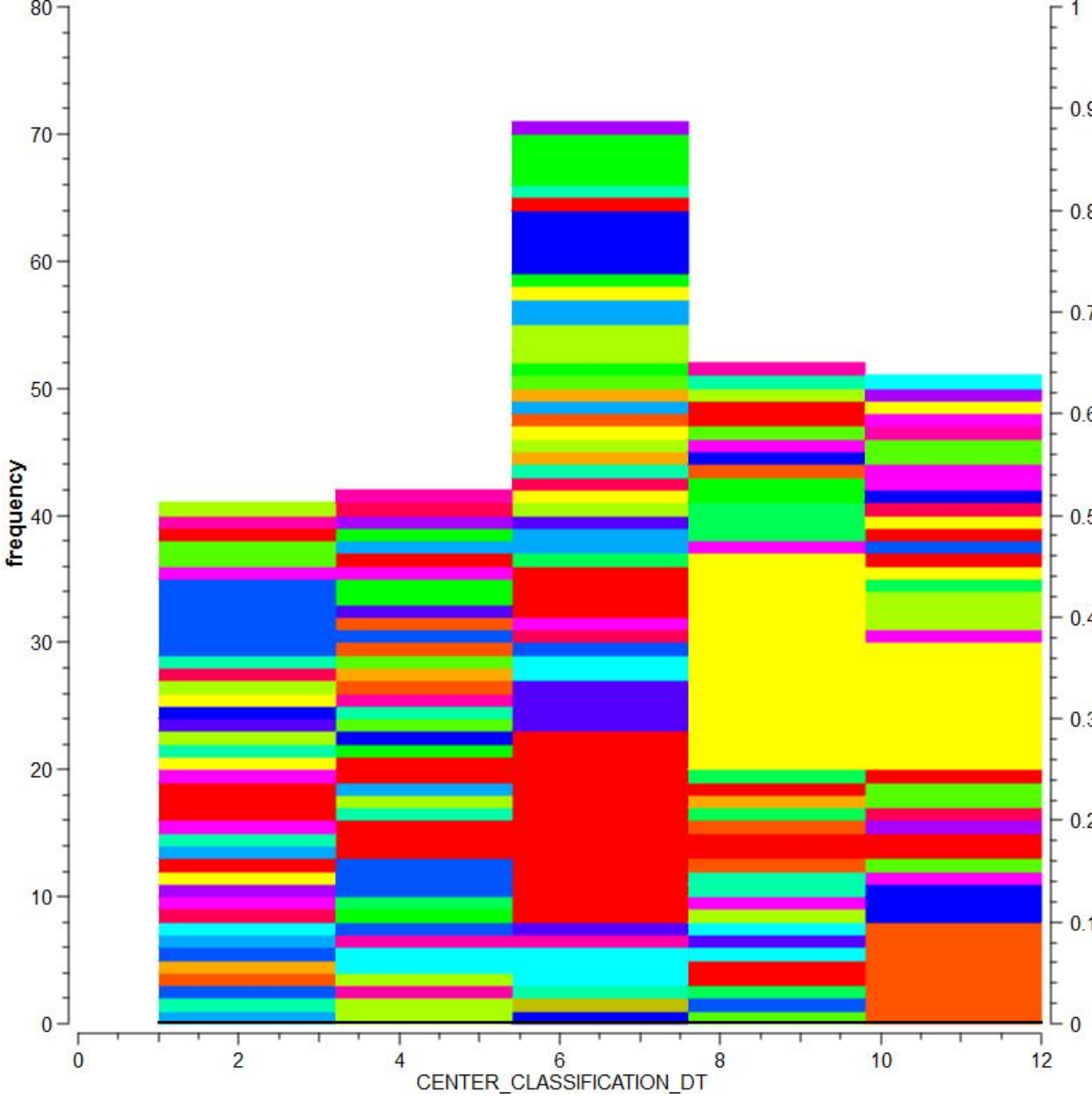
The data mining algorithms appear to produce two FIRM_NAME candidates which may have medical device recalls with the FDA in 2016. The firms are Customed, Inc and Covidien LLC. This conclusion is based on the number of recalls, and the clustering and prediction results. The best way to verify the prediction is to download another medical device recall dataset from the FDA for 2016 up to the current date, and analyze it. The predicted FIRM_NAMES having medical device recalls in 2016 may be an indicator the results of this data mining project are valid.

It was realized during the data mining activities a single year of medical device recall data may not be enough to predict a firm producing a medical device recall in the following year. It is unclear the optimum number of years for predicting whether or not a firm may produce a medical device recall in the following year.

REFERENCES

- [1] Recalls of Medical Devices. (2016). Retrieved October 19, 2016, from <https://catalog.data.gov/dataset/recalls-of-medical-devices-24040>
- [2] What is a Medical Device Recall? (n.d.). Retrieved October 19, 2016, from <http://www.fda.gov/MedicalDevices/Safety/ListofRecalls/ucm329946.htm>

Fig. 3. Prediction Distribution



VI. APPENDIX

A. Weka Canopy Cluster Algorithm Results

=== Run information ===

Scheme: weka.clusterers.Canopy -N -1 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t2 -1.0 -t1 -1.25 -S 1
Relation: FDA_MedDeviceRecalls2015_CenterDateAsOrdinal_FirmNameOnly
Instances: 2568
Attributes: 2
 CENTER_CLASSIFICATION_DT
 FIRM_NAME
Test mode: evaluate on training data

=== Clustering model (full training set) ===

Canopy clustering

=====

Number of canopies (cluster centers) found: 83

T2 radius: 0.638

T1 radius: 0.798

Cluster 0: 3.783333,'Navilyst Medical, Inc',{60} <0>
Cluster 1: 4.833333,'Bard Access Systems',{12} <1>
Cluster 2: 8.759259,'Philips Medical Systems (Cleveland) Inc',{54} <2>
Cluster 3: 1.45,'The Anspach Effort, Inc',{20} <3>
Cluster 4: 7,'Pega Medical Inc',{4} <4>
Cluster 5: 4.818182,'Resource Optimization & Innovation Llc',{88} <5>
Cluster 6: 8.75,'Cook Inc',{8} <6>
Cluster 7: 7.214286,'LeMaitre Vascular, Inc',{14} <7>
Cluster 8: 3,'Vital Signs Colorado Inc',{2} <8>
Cluster 9: 3.987342,'Ameditech Inc',{79} <9>
Cluster 10: 2,'Zipline Medical',{3} <10>
Cluster 11: 9.510204,'GE Healthcare',{49} <11>
Cluster 12: 6.583333,'Biomet, Inc',{24} <12>
Cluster 13: 3.312925,'Covidien LLC',{147} <13,33>
Cluster 14: 11.125,'Siemens Medical Solutions USA Inc',{8} <14>
Cluster 15: 8.923077,'CooperSurgical, Inc',{13} <15>
Cluster 16: 5.333333,'Zimmer, Inc',{12} <16>
Cluster 17: 7.733333,'Synthes, Inc',{15} <17>
Cluster 18: 6.5,'Arrow International Inc',{12} <18>
Cluster 19: 2.206186,'Customed, Inc',{97} <19>
Cluster 20: 7.344828,'Teleflex Medical',{29} <20>
Cluster 21: 8.142857,'Edwards Lifesciences, LLC',{7} <21>
Cluster 22: 9.615385,'Winco Mfg., LLC',{13} <22>
Cluster 23: 9.784314,'Stradis Healthcare, LLC',{51} <23>
Cluster 24: 7.333333,'DeRoyal Industries Inc',{6} <24>
Cluster 25: 9,'Ortho-Clinical Diagnostics',{12} <25,61>
Cluster 26: 9.666667,'Ethicon Endo-Surgery Inc',{3} <26>
Cluster 27: 6.142857,'HeartWare Inc',{7} <27>
Cluster 28: 5.833333,'Roche Diagnostics Operations, Inc',{18} <28>
Cluster 29: 11.705882,'Trumpf Medical Systems, Inc',{17} <29>
Cluster 30: 4,'Therasage, LLC',{3} <30>
Cluster 31: 4,'GE Medical Systems, LLC',{5} <31>
Cluster 32: 11.232877,'Customed, Inc',{73} <32>

Cluster 33: 12,'Covidien LLC',{15} <13,33>
 Cluster 34: 3.125,'Becton Dickinson & Company',{8} <34>
 Cluster 35: 7.066667,'Beckman Coulter Inc.',{15} <35>
 Cluster 36: 8,'Remel Inc',{6} <36>
 Cluster 37: 5.787879,'Halyard Health, Inc',{33} <37>
 Cluster 38: 7.866667,'Baxter Healthcare Corp.',{15} <38>
 Cluster 39: 8.756757,'Intuitive Surgical, Inc.',{37} <39>
 Cluster 40: 9.086957,'Siemens Medical Solutions USA, Inc.',{23} <40>
 Cluster 41: 3,'Steris Corporation',{2} <41>
 Cluster 42: 9,'Vilex In Tennessee Inc',{3} <42>
 Cluster 43: 4.714286,'Fujifilm Medical Systems U.S.A., Inc.',{7} <43>
 Cluster 44: 6.25,'Natus Medical Incorporated',{4} <44>
 Cluster 45: 10.666667,'GE Inspection Technologies, LP',{3} <45>
 Cluster 46: 7.666667,'Trilliant Surgical Ltd.',{6} <46>
 Cluster 47: 6.6,'Invacare Corporation',{5} <47>
 Cluster 48: 7.5,'SCC Soft Computer',{2} <48>
 Cluster 49: 7.066667,'DePuy Orthopaedics, Inc.',{15} <49,65>
 Cluster 50: 4.59375,'American Medical Systems Innovation Center - Silicon Valley',{32} <50>
 Cluster 51: 7,'Baylis Medical Corp *',{2} <51>
 Cluster 52: 11.941176,'The Anspach Effort, Inc.',{17} <52>
 Cluster 53: 8.7,'Siemens Healthcare Diagnostics, Inc.',{20} <53>
 Cluster 54: 7.535714,'Nobel Biocare Usa Llc',{28} <54>
 Cluster 55: 4.375,'Toshiba American Medical Systems Inc',{8} <55>
 Cluster 56: 7.1,'Stryker Howmedica Osteonics Corp.',{20} <56>
 Cluster 57: 9.238095,'Siemens Medical Solutions USA, Inc',{42} <57>
 Cluster 58: 11.5,'Hamilton Medical, Inc.',{4} <58>
 Cluster 59: 6.272727,'Philips Electronics North America Corporation',{11} <59>
 Cluster 60: 7,'Beaver-Visitec International Inc.',{5} <60>
 Cluster 61: 1,'Ortho-Clinical Diagnostics',{2} <25,61>
 Cluster 62: 10.384615,'Microtek Medical Inc',{13} <62>
 Cluster 63: 8.333333,'Medtronic Navigation, Inc.',{6} <63>
 Cluster 64: 7.75,'Terumo Cardiovascular Systems Corp',{8} <64>
 Cluster 65: 1,'DePuy Orthopaedics, Inc.',{5} <49,65>
 Cluster 66: 9.176471,'Boston Scientific Corporation',{17} <66>
 Cluster 67: 5.888889,'Volcano Corporation',{9} <67>
 Cluster 68: 1.5,'Baxter Corporation Englewood',{2} <68>
 Cluster 69: 5.5,'Exactech, Inc.',{6} <69>
 Cluster 70: 8,'Elite Biomedical Solutions LLC',{2} <70>
 Cluster 71: 7.5,'Biomet Spine, LLC',{12} <71>
 Cluster 72: 7.5,'Omnilife Science Inc.',{4} <72>
 Cluster 73: 4.5,'bioMerieux, Inc.',{8} <73>
 Cluster 74: 6,'Churchill Medical Systems, Inc.',{12} <74>
 Cluster 75: 9.75,'Maquet Cardiovascular Us Sales, Llc',{8} <75>
 Cluster 76: 6.222222,'Fresenius Medical Care Holdings, Inc.',{9} <76>
 Cluster 77: 4,'Beckman Coulter, Inc.',{9} <77>
 Cluster 78: 3.583333,'Siemens Healthcare Diagnostics Inc',{12} <78>
 Cluster 79: 3.25,'Hardy Diagnostics',{8} <79>
 Cluster 80: 5.952381,'Synthes (USA) Products LLC',{21} <80>
 Cluster 81: 9.285714,'Biomet 3i, LLC',{7} <81>
 Cluster 82: 6,'Stryker Craniomaxillofacial Division',{5} <82>

Time taken to build model (full training data) : 0.02 seconds

=== Model and evaluation on training set ===

Clustered Instances

0	60 (2%)
1	94 (4%)
2	85 (3%)
3	19 (1%)
4	96 (4%)
5	88 (3%)
6	8 (0%)
7	14 (1%)
8	84 (3%)
9	79 (3%)
10	86 (3%)
11	55 (2%)
12	24 (1%)
13	141 (5%)
14	91 (4%)
15	15 (1%)
16	17 (1%)
17	15 (1%)
18	12 (0%)
19	97 (4%)
20	29 (1%)
21	7 (0%)
22	13 (1%)
23	140 (5%)
24	6 (0%)
25	103 (4%)
26	3 (0%)
27	7 (0%)
28	18 (1%)
29	18 (1%)
30	62 (2%)
31	5 (0%)
32	73 (3%)
33	92 (4%)
34	9 (0%)
35	21 (1%)
36	101 (4%)
37	33 (1%)
38	15 (1%)
39	37 (1%)
40	23 (1%)
41	4 (0%)
42	3 (0%)
43	7 (0%)
44	4 (0%)
45	3 (0%)
46	6 (0%)
47	5 (0%)
48	2 (0%)
49	15 (1%)
50	32 (1%)
51	2 (0%)
52	18 (1%)
53	23 (1%)
54	28 (1%)

55	8 (0%)
56	20 (1%)
57	48 (2%)
58	5 (0%)
59	11 (0%)
60	5 (0%)
61	60 (2%)
62	13 (1%)
63	6 (0%)
64	8 (0%)
65	5 (0%)
66	18 (1%)
67	9 (0%)
68	2 (0%)
69	6 (0%)
70	2 (0%)
71	12 (0%)
72	4 (0%)
73	8 (0%)
74	86 (3%)
75	9 (0%)
76	9 (0%)
77	9 (0%)
78	12 (0%)
79	11 (0%)
80	23 (1%)
81	7 (0%)
82	5 (0%)

B. Weka Simple K Means Cluster Algorithm Results

=== Run information ===

Scheme: weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10
Relation: FDA_MedDeviceRecalls2015_CenterDateAsOrdinal_FirmNameOnly
Instances: 2568
Attributes: 2
 CENTER_CLASSIFICATION_DT
 FIRM_NAME
Test mode: evaluate on training data

=== Clustering model (full training set) ===

kMeans

=====

Number of iterations: 3
Within cluster sum of squared errors: 2345.277457806276

Initial starting points (random):

Cluster 0: 9,'Stradis Healthcare, LLC.'
Cluster 1: 4,'Beckman Coulter, Inc.'

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute	Cluster#	
	Full Data	0 1
	(2568.0)	(1361.0) (1207.0)
=====		
CENTER_CLASSIFICATION_DT		6.5031 8.9633 3.7291
FIRM_NAME		Customed, Inc Customed, Inc Covidien LLC

Time taken to build model (full training data) : 0.03 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 1361 (53%)
1 1207 (47%)

=== Run information ===

=== Clustering model (full training set) ===

$$(0.0:0.18182,(((0.0:0,0.0:0):0.09091,0.0:0.09091):0,((0.0:0,0.0:0):0,0.0:0):0.09091):0.09091)$$
[illegible]

0.1):(445.0:0,445.0:0):1):0,446.0:1):0,447.0:1):0,449.0:1):0,(440.0:0.09091,440.0:0.09091):0.90909):0,441.0:1):0,436.0:1):0,437.
0:1):0,438.0:1):0,439.0:1):0,442.0:1):0,443.0:1):0,432.0:1):0,(424.0:0.09091,(424.0:0,424.0:0):0.09091):0.90909):0,434.0:1):0,43
5.0:1):0,(429.0:0.27273,429.0:0.27273):0.72727):0,430.0:1):0,431.0:1):0,433.0:1):0,425.0:1):0,426.0:1):0,420.0:1):0,((421.0:0,42
1.0:0):0,421.0:0):1):0,422.0:1):0,423.0:1):0,((((((((((((((((((((427.0:0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.
.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,42
7.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0.09091,((((((((((((((((((((427.0:0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:
0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.
0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0.09
091):0,(((((((427.0:0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,427.0:0):0,4
28.0:1):0,416.0:1):0,417.0:1):0,(418.0:0.09091,(418.0:0,418.0:0):0.09091,418.0:0.09091):0):0.90909):0,419.0:1):0,412.0:1):0,41
3.0:1):0,414.0:1):0,415.0:1):0,(404.0:0.09091,404.0:0.09091):0.90909):0,409.0:1):0,410.0:1):0,411.0:1):0,405.0:1):0,406.0:1):0,4
07.0:1):0,408.0:1):0,400.0:1):0,(((((((401.0:0,401.0:0):0,401.0:0):0,401.0:0):0,401.0:0):0,401.0:0):0,401.0:0):0,401.0:0):0.09091,
(((401.0:0,401.0:0):0,401.0:0):0,401.0:0):0,401.0:0):0.09091):0.90909):0,(402.0:0,402.0:0):1):0,403.0:1):0,396.0:1):0,(397.0:0,3
97.0:0):1):0,392.0:1):0,(399.0:0,399.0:0):0.09091,399.0:0.09091):0.90909):0,395.0:1):0,390.0:1):0,388.0:1):0,398.0:1):0,(((((((
((384.0:0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):0,384.0:0):1
):0,391.0:1):0,393.0:1):0,(((((((394.0:0,394.0:0):0,394.0:0):0,394.0:0):0,394.0:0):0,394.0:0):0,394.0:0):0,394.0:0):0,394.0:0):0,394.0:0):1):0,(385.0:0.36
364,385.0:0.36364):0.63636):0,386.0:1):0,(((((((380.0:0,380.0:0):0,380.0:0):0,380.0:0):0,380.0:0):0,380.0:0):0,380.0:0):0,380.0:0):0.27273,
380.0:0.27273):0.72727):0,389.0:1):0,376.0:1):0,(((((((382.0:0,382.0:0):0,382.0:0):0,382.0:0):0,382.0:0):0,382.0:0):0,382.0:0):0.09091,(((((((382.0:0,382.0:0):0,382.0:
0):0,382.0:0):0,382.0:0):0,382.0:0):0.09091):0.90909):0,383.0:1):0,387.0:1):0,379.0:1):0,374.0:1):0,372.0:1):0,373.0:1):0,(375.0:
0,375.0:0):1):0,377.0:1):0,(378.0:0.63636,378.0:0.63636):0.63636):0,381.0:1):0,((368.0:0,368.0:0):0,368.0:0):1):0,(((((((369.0:0,
369.0:0):0,369.0:0):0,369.0:0):0,369.0:0):0,369.0:0):0,369.0:0):0,369.0:0):0,369.0:0):0,369.0:0):1):0,(364.0:0,364.0:0):1):0,371.0:1):0,360.0:
1):0,366.0:1):0,367.0:1):0,((370.0:0.27273,370.0:0.27273):0,370.0:0.27273):0.72727):0,(((((((361.0:0,361.0:0):0,361.0:0):0,361.0:
0):0,361.0:0):0,361.0:0):0,361.0:0):1):0,362.0:1):0,356.0:1):0,(357.0:0.09091,357.0:0.09091):0.90909):0,358.0:1):0,(3
59.0:0.90909,359.0:0.90909):0.09091):0,363.0:1):0,365.0:1):0,(((((((353.0:0,353.0:0):0,353.0:0):0,353.0:0):0,353.0:0):0,353.0:0):0,353.0:0):1
):0,((354.0:0,354.0:0):0,354.0:0):1):0,349.0:1):0,(355.0:0,355.0:0):1):0,350.0:1):0,351.0:1):0,352.0:1):0,355.0:1):0,((345.0:0.0909
1,345.0:0.09091):0.63636,(345.0:0.63636,345.0:0.63636):0.09091):0.54545):0,(341.0:0,341.0:0):1):0,342.0:1):0,348.0:1):0,337.0
:1):0,344.0:1):0,346.0:1):0,347.0:1):0,((338.0:0,338.0:0):0,338.0:0):1):0,(339.0:0.09091,(339.0:0,339.0:0):0.09091):0.90909):0,34
0.0:1):0,343.0:1):0,((333.0:0.09091,333.0:0.09091):0.09091,333.0:0.18182):0.81818):0,334.0:1):0,335.0:1):0,((336.0:0,336.0:0):0
09091,336.0:0.09091):0.90909):0,(325.0:0.09091,325.0:0.09091):0.90909):0,326.0:1):0,331.0:1):0,332.0:1):0,(330.0:0.63636,((
330.0:0,330.0:0):0.09091,330.0:0.09091):0,330.0:0.09091):0.27273):0.63636):0,328.0:1):0,327.0:1):0,(((((((329.0:0,329.0:0):0,329.
0:0):0,329.0:0):0,329.0:0):1):0,(321.0:0.45455,321.0:0.45455):0.54545):0,(324.0:0.27273,324.0:0.27273):0.72727):0,319.0:1):0,3
23.0:1):0,313.0:1):0,(((((((318.0:0.09091,318.0:0.09091):0.09091,(318.0:0.09091,(318.0:0,318.0:0):0.09091):0.09091):0,318.0:0.18
182):0.09091,318.0:0.27273):0.72727):0,((320.0:0,320.0:0):0.45455,320.0:0.45455):0.54545):0,322.0:1):0,(316.0:0,316.0:0):1):0,
(317.0:0.09091,317.0:0.09091):0.90909):0,(((((((305.0:0,305.0:0):0,305.0:0):0,305.0:0):0,305.0:0):0,305.0:0):0,305.0:0):0,305.0:0):0,305.0:
0):0,(305.0:0,

[illegible]

[illegible]

[illegible]

0	7 (0%)
1	2561 (100%)

D. Orange Association Algorithm Results

File name: C:/Users/jgiles/Desktops/Training/Lewis University MSDS/70-510 Intro to Data Mining Analytics/Data sets/FDAME dDe vice R e ca lls2015/FDA_Me dDe vice R e ca lls2015_C e nte rDa te AsO rdina lFo rO ra nge .csv

Format: C o m m a -se pa rate d va lue s

Data

Examples: 2568

Attributes: 7 (TR ADE_NAME, R E C A L L _ C L A S S , C E N T E R , C E N T E R _ C L A S S I F I C A T I O N _ D T , P O S T E D _ I N T E R N E T _ D T , T E R M I N A T I O N _ D T , F I R M _ N A M E)

Meta attributes: 1 (R E C A L L _ N U M B E R)

Class: W E B _ A D D R E S S

Select A ttributes

Input data

Examples: 2568

A ttributes: 7 (TR ADE_NAME, R E C A L L _ C L A S S , C E N T E R , C E N T E R _ C L A S S I F I C A T I O N _ D T , P O S T E D _ I N T E R N E T _ D T , T E R M I N A T I O N _ D T , F I R M _ N A M E)

Meta attributes: 1 (R E C A L L _ N U M B E R)

Class: W E B _ A D D R E S S

Output data

Examples: 2568

A ttributes: 1 (C E N T E R _ C L A S S I F I C A T I O N _ D T)

Class: F I R M _ N A M E

Removed: 7 (W E B _ A D D R E S S , R E C A L L _ N U M B E R , P O S T E D _ I N T E R N E T _ D T , T R A D E _ N A M E , C E N T E R , R E C A L L _ C L A S S , T E R M I N A T I O N _ D T)

Data Table

Data

Examples: 2568

A ttributes: 1 (C E N T E R _ C L A S S I F I C A T I O N _ D T)

Class: F I R M _ N A M E

ENTIRE DATA TABLE NOT SHOWN

Itemsets

Settings

A lgorithm: s p a r s e d a t a

Minimal support: 1

Maximal number of rules: 100000

Itemsets Explorer

Itemset statistics

Number of itemsets: 3

Selected itemsets: 2

Covered examples: 2568

Itemsets

Itemsets Examples

C E N T E R _ C L A S S I F I C A T I O N _ D T 2568

F I R M _ N A M E 2568

F I R M _ N A M E 2568

C E N T E R _ C L A S S I F I C A T I O N _ D T 2568

A s s o c i a t i o n R u l e s

Settings

A lgorithm: a t t r i b u t e d a t a

Induction of classification rules: Y e s

Minimal support: 1

Minimal confidence: 1

Maximal number of rules: 100000

A s s o c i a t i o n R u l e s F i l t e r

Rules statistics

Total number of rules: 15

Support: 1% - 1%

Confidence: 11% - 14%

Number of rules in the graph: 15

Support: 1% - 1%

Confidence: 11% - 14%

Selected rules: 15

Support: 1% - 1%

Confidence: 12% - 15%

Rule matrix

No te : co lum ns co rre s po nd to s uppo rt, ro ws co rre s po nd to co nfide nce

Selected rules

Supp Conf Lift Lev Strg Cov A ntecedent -> Consequent

0.011 0.138 5.910 0.009 0.286 0.082 C ENTER _C LASSIFICATIO N_DT=04 FIRM_NAME=Na vilyst Me dica l, I nc

0.012 0.138 2.088 0.006 0.783 0.085 C ENTER _C LASSIFICATIO N_DT=03 FIRM_NAME=C usto m e d, Inc

0.011 0.129 3.752 0.008 0.419 0.082 C ENTER _C LASSIFICATIO N_DT=06 FIRM_NAME=R e so urce O ptim iza tio n & Inno va tio n Llc

0.011 0.136 2.060 0.005 0.859 0.077 C ENTER _C LASSIFICATIO N_DT=02 FIRM_NAME=C usto m e d, Inc

0.010 0.131 4.269 0.008 0.399 0.077 C ENTER _C LASSIFICATIO N_DT=02 FIRM_NAME=Am e dite ch Inc

0.010 0.120 3.496 0.007 0.406 0.085 C ENTER _C LASSIFICATIO N_DT=05 FIRM_NAME=R e so urce O ptim iza tio n & Inno va tio n Llc

0.011 0.123 1.955 0.005 0.714 0.088 C ENTER _C LASSIFICATIO N_DT=01 FIRM_NAME=C o vidie n LLC

0.011 0.124 3.759 0.008 0.392 0.085 C ENTER _C LASSIFICATIO N_DT=08 FIRM_NAME=P hilips Me dica l Syste m s (C le ve la nd) Inc

0.012 0.132 1.996 0.006 0.749 0.088 C ENTER _C LASSIFICATIO N_DT=01 FIRM_NAME=C usto m e d, Inc

0.011 0.133 2.114 0.006 0.771 0.082 C ENTER _C LASSIFICATIO N_DT=04 FIRM_NAME=C o vidie n LLC

0.011 0.133 2.014 0.005 0.810 0.082 C ENTER _C LASSIFICATIO N_DT=11 FIRM_NAME=C usto m e d, Inc

0.012 0.142 2.148 0.006 0.780 0.085 C ENTER _C LASSIFICATIO N_DT=12 FIRM_NAME=C usto m e d, Inc

0.011 0.129 2.045 0.006 0.747 0.085 C ENTER _C LASSIFICATIO N_DT=03 FIRM_NAME=C o vidie n LLC

0.012 0.138 6.961 0.010 0.235 0.085 C ENTER _C LASSIFICATIO N_DT=10 FIRM_NAME=Stradis He a lthca re , LLC .

0.011 0.146 2.322 0.006 0.818 0.077 C ENTER _C LASSIFICATIO N_DT=02 FIRM_NAME=C o vidie n LLC

Association Rules Explorer

Rule statistics

Number of rules: 15

Selected rules:

Examples covered by selected rules' precondition:

Matching examples:

Mismatching examples:

Rules

Rules Supp Conf Lift Lev Strg Cov

C ENTER _C LASSIFICATIO N_DT=12

C ENTER _C LASSIFICATIO N_DT=12 -> FIR M_NAME=C ustom e d, Inc 0.012 0.142 2.148 0.006 0.780 0.085

C ENTER _C LASSIFICATIO N_DT=10

C ENTER _C LASSIFICATIO N_DT=10

C ENTER _C LASSIFI C ATIO N_DT =10 -> FI R M_NAME=Stradis He a lthca re , LLC . 0.012 0.138 6.961 0.010 0.235 0.085

C ENTER _C LASSIFICATIO N_DT=11

C ENTER _C LASSIFICATIO N_DT=11 -> FIR M_NAME=C ustom e d, Inc 0.011 0.133 2.014 0.005 0.810 0.082

C ENTER _C LASSIFICATIO N_DT=05

C ENTER _C LASSIFICATIO N_DT =05 -> FIR M_NAME=R e source O ptim iza tio n & Inno va tio n Llc 0.010 0.120 3.496 0.007 0.406 0.085

C ENTER _C LASSIFICATIO N_DT=04

C ENTER _C LASSIFICATIO N_DT=04 -> FIR M_NAME=C o vidie n LLC 0.011 0.133 2.114 0.006 0.771 0.082

C ENTER _C LASSIFICATIO N_DT=04 -> FIR M_NAME=Na vilyst Me dica l, Inc 0.011 0.138 5.910 0.009 0.286 0.082

C ENTER _C LASSIFICATIO N_DT=06

C ENTER _C LASSIFICATIO N_DT =06 -> FIR M_NAME=R e source O ptim iza tio n & Inno va tio n Llc 0.011 0.129 3.752 0.008 0.419 0.082

C ENTER _C LASSIFICATIO N_DT=01
C ENTER _C LASSIFIC ATIO N_DT=01 -> FIR M_NAME=C o vidie n LLC 0.011 0.123 1.955 0.005 0.714 0.088
C ENTER _C LASSIFIC ATIO N_DT=01 -> FIR M_NAME=C ustom e d, Inc 0.012 0.132 1.996 0.006 0.749 0.088
C ENTER _C LASSIFICATIO N_DT=03
C ENTER _C LASSIFIC ATIO N_DT=03 -> FIR M_NAME=C o vidie n LLC 0.011 0.129 2.045 0.006 0.747 0.085
C ENTER _C LASSIFIC ATIO N_DT=03 -> FIR M_NAME=C ustom e d, Inc 0.012 0.138 2.088 0.006 0.783 0.085
C ENTER _C LASSIFICATIO N_DT=02
C ENTER _C LASSIFIC ATIO N_DT=02 -> FIR M_NAME=Am edite ch Inc 0.010 0.131 4.269 0.008 0.399 0.077
C ENTER _C LASSIFIC ATIO N_DT=02 -> FIR M_NAME=C ustom e d, Inc 0.011 0.136 2.060 0.005 0.859 0.077
C ENTER _C LASSIFIC ATIO N_DT=02 -> FIR M_NAME=C o vidie n LLC 0.011 0.146 2.322 0.006 0.818 0.077
C ENTER _C LASSIFICATIO N_DT=08
C ENTER _C LASSIFIC ATIO N_DT=08 -> FIR M_NAME=P hilips Me dica l Syste m s (C le ve la nd) Inc 0.011 0.124 3.759
0.008 0.392 0.085

E. What is a Medical Device Recall [2]?

A recall is an action taken to address a problem with a medical device that violates FDA law. Recalls occur when a medical device is defective, when it could be a risk to health, or when it is both defective and a risk to health.

A medical device recall does not always mean that you must stop using the product or return it to the company. A recall sometimes means that the medical device needs to be checked, adjusted, or fixed. If an implanted device (for example, a pacemaker or an artificial hip) is recalled, it does not always have to be removed. When an implanted device has the potential to fail unexpectedly, companies often tell doctors to contact their patients to discuss the risk of removing the device compared to the risk of leaving it in place.

Examples of the types of actions that may be considered recalls:

- Inspecting the device for problems
- Repairing the device
- Adjusting settings on the device
- Re-labeling the device
- Destroying device
- Notifying patients of a problem
- Monitoring patients for health issues

Sometimes a company may have concerns about a group of products, but it cannot predict which individual devices will be affected. To be on the safe side, the company may recall an entire lot, model, or product line.

A recall is either a correction or a removal depending on where the action takes place.

- Correction - Addresses a problem with a medical device in the place where it is used or sold.
- Removal - Addresses a problem with a medical device by removing it from where it is used or sold.

1) Who recalls medical devices?

In most cases, a company (manufacturer, distributor, or other responsible party) recalls a medical device on its own (voluntarily). When a company learns that it has a product that violates FDA law, it does two things:

- Recalls the device (through correction or removal)
- Notifies FDA.

Legally, FDA can require a company to recall a device. This could happen if a company refuses to recall a device that is associated with significant health problems or death. However, in practice, FDA has rarely needed to require a medical device recall.

More in Medical Device Recalls

(/MedicalDevices/Safety/ListofRecalls/default.htm)

2016 Medical Device Recalls (/MedicalDevices/Safety/ListofRecalls/ucm480134.htm)

2015 Medical Device Recalls (/MedicalDevices/Safety/ListofRecalls/ucm429489.htm)

2014 Medical Device Recalls (/MedicalDevices/Safety/ListofRecalls/ucm384921.htm)