

# **Modeling and Text Analysis to Empower FAERS Adverse Event Assessment**

**Qais Hatim, PhD**

**Tom Sabo**

**Kendra Worthy, PharmD, MS**

Drugs in post-marketing period experience unexpected adverse drug events

Unnecessary, often fatal harm to patients

Pharmacovigilance involves early discovery and analysis of emerging adverse event signals

Much of the supporting information is textual data that requires review by safety experts to:

- Identify clinical entities
- Decide upon acquisition of additional information
- Consider whether any regulatory action is warranted

# Research Problems



# Research Problems, cont.

FDA uses FAERS to compute signal scores (statistical reporting associations).

Signals do not establish a causal adverse drug reactions relationship, but are rather considered initial warnings that require domain experts assessment.

Assessment normally consists of a complicated process

- drug safety evaluators look for supporting information:
  - temporal relationships,
  - published case reports,
  - biological and clinical plausibility,
  - clinical trials data,
  - etc.

A FAERS report contains both structured and unstructured data.

Some of the FAERS information is structured, a significant portion of it remains in narrative formats.

Plentiful of the information that is critical to risk assessments are imbedded in narrative text

### Challenge:

- Timely adverse event detection can save lives and further quality of life
- Determining conclusions from data can prove challenging for manual analysis alone

### Solutions:

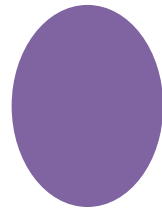
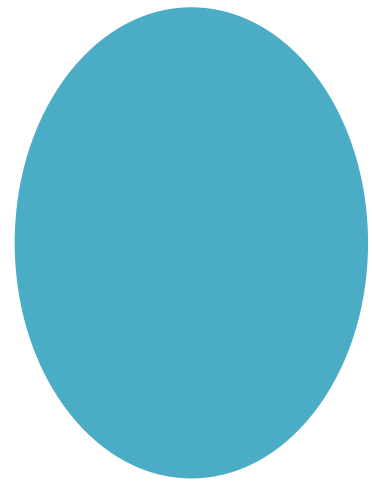
- Text analysis employs a wide range of statistical, machine learning, and linguistic techniques for automated/semi-automated processing of text data
- Generate structured data from unstructured, then apply modeling techniques

## Research Challenge and Solution

# Research Solution

- **Applying computational** approaches in the analysis of FAERS , **advance** the understanding of the feasibility of applying advance techniques to uncover relationships between drugs and hepatic failure.
- **Outputs** offer information on:
  - the drug combinations related to hepatic failure,
  - clarify factors that can predict a greater degree of the hepatic failure from serious (death) to minor events (treatable),
  - rate the serious impact of using these drugs on patients.





# METHODOLOGY

# DATA DESCRIPTION & PREPARATION

- Retrieving FAERS hepatic failure data (Nov.1997- March 2018).
- Events are customized using SMQ:
  - select drug related hepatic disorders-severe events only.
  - groupings of terms from one or more SOC's related to:
    1. defined medical condition
    2. area of interest
    3. terms related to signs, symptoms, diagnoses, syndromes, physical findings, laboratory test data related to DILI.



# DATA DESCRIPTION & PREPARATION

- 304,000 cases are retrieved and data was prepared for both the unsupervised and supervised learning.

| Age Group      | Total Cases   | Report Type  |               |               | Seriousness  |               | Reported Outcomes |               |              |            |           |            |               |
|----------------|---------------|--------------|---------------|---------------|--------------|---------------|-------------------|---------------|--------------|------------|-----------|------------|---------------|
|                |               | Direct       | Expedited     | Non-Expedited | Non Serious  | Serious       | DE                | HO            | LT           | DS         | CA        | RI         | OT            |
| <1 year        | 134           | 3            | 125           | 6             | 1            | 133           | 50                | 58            | 21           | 4          | 5         | 0          | 69            |
| 1 - <3 years   | 174           | 16           | 153           | 5             | 3            | 171           | 67                | 84            | 15           | 2          | 2         | 0          | 76            |
| 3 - <7 years   | 225           | 23           | 198           | 4             | 7            | 218           | 50                | 110           | 31           | 0          | 0         | 1          | 119           |
| 7 - <17 years  | 770           | 49           | 685           | 36            | 16           | 754           | 131               | 397           | 71           | 14         | 1         | 3          | 469           |
| 17 - <65 years | 19,456        | 901          | 17,385        | 1,170         | 753          | 18,703        | 3,849             | 10,106        | 1,530        | 482        | 8         | 70         | 12,100        |
| >=65 years     | 10,444        | 363          | 9,559         | 522           | 320          | 10,124        | 2,811             | 5,819         | 934          | 242        | 2         | 36         | 6,139         |
| NOT REPORTED   | 11,616        | 112          | 10,157        | 1,347         | 1,057        | 10,559        | 2,366             | 3,616         | 425          | 221        | 40        | 5          | 8,118         |
| <b>Total</b>   | <b>42,819</b> | <b>1,467</b> | <b>38,262</b> | <b>3,090</b>  | <b>2,157</b> | <b>40,662</b> | <b>9,324</b>      | <b>20,190</b> | <b>3,027</b> | <b>965</b> | <b>58</b> | <b>115</b> | <b>27,090</b> |

| Patient Sex  | Total Cases   | % of Cases    |
|--------------|---------------|---------------|
| Female       | 19,567        | 45.7%         |
| Male         | 18,250        | 42.6%         |
| Not Reported | 4,990         | 11.7%         |
| Unknown      | 12            | 0.0%          |
| <b>Total</b> | <b>42,819</b> | <b>100.0%</b> |

| Reported Outcomes             | Total Cases   |
|-------------------------------|---------------|
| Death                         | 9,324         |
| Hospitalized                  | 20,190        |
| Life Threatening              | 3,027         |
| Disabled                      | 965           |
| Congenital Anomaly            | 58            |
| Required Intervention         | 115           |
| Other Outcome                 | 27,090        |
| <b>Total (Distinct Cases)</b> | <b>42,819</b> |

# DATA DESCRIPTION & PREPARATION



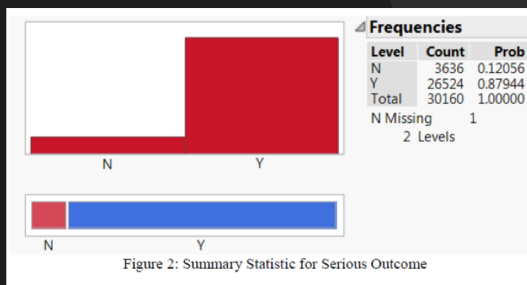
Data is dominated by cases with serious outcome value of Yes ( $Y=1$ ).

model with such dominate outcome will be biased.

To compensate for the rare proportion of No ( $No=0$ ) in the raw data, over-sampling is performed

produce a more balanced data set

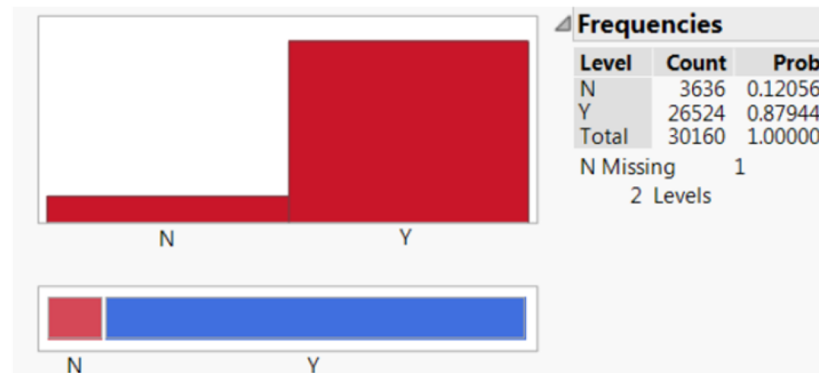
keep the patterns that appear in the data traceable in the sample.





# DATA DESCRIPTION and PREPARATION

- Data is dominated by cases with serious outcome value of Yes (Y=1).
  - building any model with such dominate outcome will be biased towards predicting serious adverse event mostly.
- To compensate for the rare proportion of No (No=0) in the raw data, over-sampling of the data was done to:
  - produce a more balanced data set
  - Keep the patterns that appear in the data traceable in the sample.



# PREDICTIVE ANALYSIS



## Text Analytics

“The term **text analytics** describes a set of linguistic, statistical, and machine learning techniques that model and structure the information content of textual sources for business intelligence, exploratory data analysis, research, or investigation.”\*

Or, put more simply...

*“Using technology to **scale** the human acts of **reading, organizing, and quantifying** freeform text in meaningful ways.”*



- Capture information embedded in text that is critical to risk assessments
  - Signs
  - Symptoms
  - Disease status/severity
  - Medical history

# PREDICTIVE ANALYSIS

## Text Analytics-Text Parsing and Text Filtering

- Stemming
- Misspellings
- Synonyms
- Noun groups
- Parts-of-Speech
- Term filtering
- Term Mapping
- Native Language Models

| Documents                                                                                         |                                              |                      |  |  |  |
|---------------------------------------------------------------------------------------------------|----------------------------------------------|----------------------|--|--|--|
| NARRATIVE                                                                                         | TEXTFILTER_SNIPPET                           | TEXTFILTER_RELEVANCE |  |  |  |
| Carol's disease [Dilatation intrahepatic duct congenital] ([Ascites], [Hepatosplenomegaly],       | ... /L, <b>alkaline phosphatase</b> 73       | 0.401                |  |  |  |
| This case was reported in a literature article and described the occurrence of liver failure in a | ... ), abnormal <b>alkaline</b>              | 0.401                |  |  |  |
| This case was reported in a literature article and described the occurrence of liver failure in a | ... ), abnormal <b>alkaline</b>              | 0.401                |  |  |  |
| This case was reported in a literature article and described the occurrence of liver failure in a | ... ), abnormal <b>alkaline</b>              | 0.401                |  |  |  |
| Serious, Related, Unlisted.                                                                       | ... pyruvic transaminase and <b>alkaline</b> | 0.401                |  |  |  |
| EVENTS: Hepatic cirrhosis, Anemia, Portal hypertension, Hepatobiliary insufficiency,              | ... hepatic function. <b>Alkaline</b>        | 0.429                |  |  |  |
| This 65-year-old female subject was enrolled in a blinded study for the treatment of diabetes     | ... laboratory results ( <b>alkaline</b>     | 0.401                |  |  |  |
| A healthcare professional reported on 27JAN98 that a 68 year old Caucasian female,                | ... ALP ( <b>alkaline</b>                    | 0.401                |  |  |  |
| A health care Professional reported on 20FEB98 that a 56 year old female patient, CAK,            | ... , and elevated <b>alkaline</b>           | 0.457                |  |  |  |
| A health care professional reported on 20MAY98 a 70 year old male patient who experienced         | ... albumin 6.7, <b>alkaline</b>             | 0.457                |  |  |  |
| Patient is a female with type II diabetes. Between March and June 1997, while taking Rezulin      | ... normal), <b>alkaline phosphatase</b>     | 0.429                |  |  |  |
| THE PREVIOUS NARRATIVE SENT ON 07FEB00 HAS BEEN SEPARATED TO REFLECT WHICH                        | ... RESULTS REVEALED AN <b>ALKALINE</b>      | 0.401                |  |  |  |
| A consumer's mother, who is a nurse, reported on 13APR00 that her son developed an                | ... ), and <b>alkaline phosphatase</b>       | 0.401                |  |  |  |
| Information was received initially on 13-OCT-00 from a healthcare professional describing a       | ... /L, <b>alkaline phosphatase</b> 301      | 0.485                |  |  |  |
| Additional information was received which included the recovery date for the event of             | ... aminotransferase 78, <b>alkaline</b>     | 0.401                |  |  |  |
| Additional information received: test results: Cytosol dosing, medical history, confirmation of   | ... ); blood <b>alkaline phosphatase</b>     | 0.401                |  |  |  |
| Additional information was received from a healthcare professional which added the events of      | ... except for his <b>alkaline</b>           | 0.541                |  |  |  |
| A 66-YEAR-OLD FEMALE PATIENT WAS HOSPITALISED DUE TO JAUNDICE, WEIGHT LOSS,                       | ... 1.65, AND <b>ALKALINE</b>                | 0.513                |  |  |  |
| Tasmar Regulatory Authority Report (Canada) (ID #120085)                                          | ... LDH INCREASED, <b>ALKALINE</b>           | 0.429                |  |  |  |
| Subject ID: G5001DAZ This 47-year-old male subject was enrolled in the Promacta REMS              | ... 2010, peak <b>alkaline</b>               | 0.457                |  |  |  |
| Subject ID: G5001DCE This approximately 77-year-old female subject was enrolled in the            | ... 1.33 and <b>alkaline phosphatase</b>     | 0.401                |  |  |  |

| Terms                 |       |        |                                     |        |      |           |
|-----------------------|-------|--------|-------------------------------------|--------|------|-----------|
| TERM                  | FREQ  | # DOCS | KEEP                                | WEIGHT | ROLE | ATTRIBUTE |
| alkaline              | 2327  | 1102   | <input checked="" type="checkbox"/> | 0.001  | Adj  | Alpha     |
| phosphatase           | 2240  | 1084   | <input checked="" type="checkbox"/> | 0.008  | Noun | Alpha     |
| phosphatases          | 13    | 10     | <input type="checkbox"/>            |        |      |           |
| phosphate             | 2227  | 1078   | <input type="checkbox"/>            |        |      |           |
| patient               | 10328 | 1046   | <input checked="" type="checkbox"/> |        |      |           |
| alkaline phosphatase  | 2012  | 1016   | <input checked="" type="checkbox"/> |        |      |           |
| alkaline phosphatases | 11    | 10     | <input type="checkbox"/>            |        |      |           |
| alkaline phosphatase  | 2001  | 1010   | <input type="checkbox"/>            |        |      |           |
| liver                 | 5183  | 975    | <input checked="" type="checkbox"/> |        |      |           |
| information           | 2758  | 881    | <input checked="" type="checkbox"/> |        |      |           |
| history               | 1676  | 873    | <input checked="" type="checkbox"/> |        |      |           |
| bilirubin             | 3094  | 854    | <input checked="" type="checkbox"/> |        |      |           |
| event                 | 2531  | 774    | <input checked="" type="checkbox"/> |        |      |           |
| medical               | 1433  | 758    | <input checked="" type="checkbox"/> |        |      |           |
| case                  | 1571  | 756    | <input checked="" type="checkbox"/> |        |      |           |
| discontinue           | 1252  | 742    | <input checked="" type="checkbox"/> |        |      |           |
| start                 | 1362  | 736    | <input checked="" type="checkbox"/> |        |      |           |
| medication            | 1327  | 706    | <input checked="" type="checkbox"/> |        |      |           |
| mg                    | 3150  | 700    | <input checked="" type="checkbox"/> |        |      |           |
| physician             | 2450  | 695    | <input checked="" type="checkbox"/> |        |      |           |
| hepatitis             | 2500  | 685    | <input checked="" type="checkbox"/> |        |      |           |
| year                  | 917   | 682    | <input checked="" type="checkbox"/> |        |      |           |
| include               | 1278  | 673    | <input checked="" type="checkbox"/> | 0.001  | Verb | Alpha     |
| concomitant           | 1004  | 672    | <input checked="" type="checkbox"/> | 0.0    | Adj  | Alpha     |
| reveal                | 1928  | 653    | <input checked="" type="checkbox"/> | 0.064  | Verb | Alpha     |

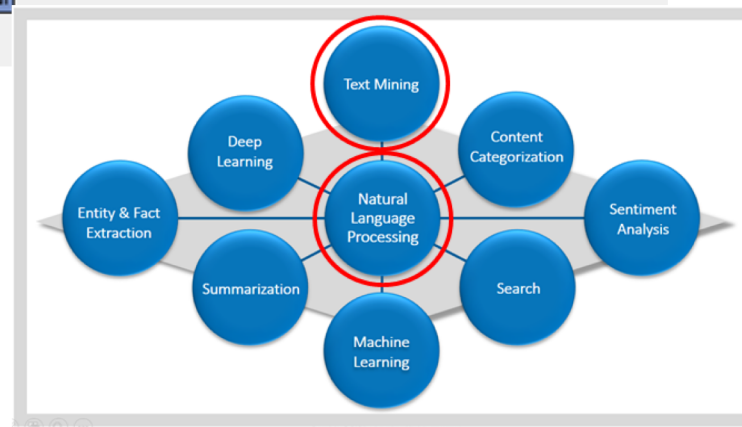
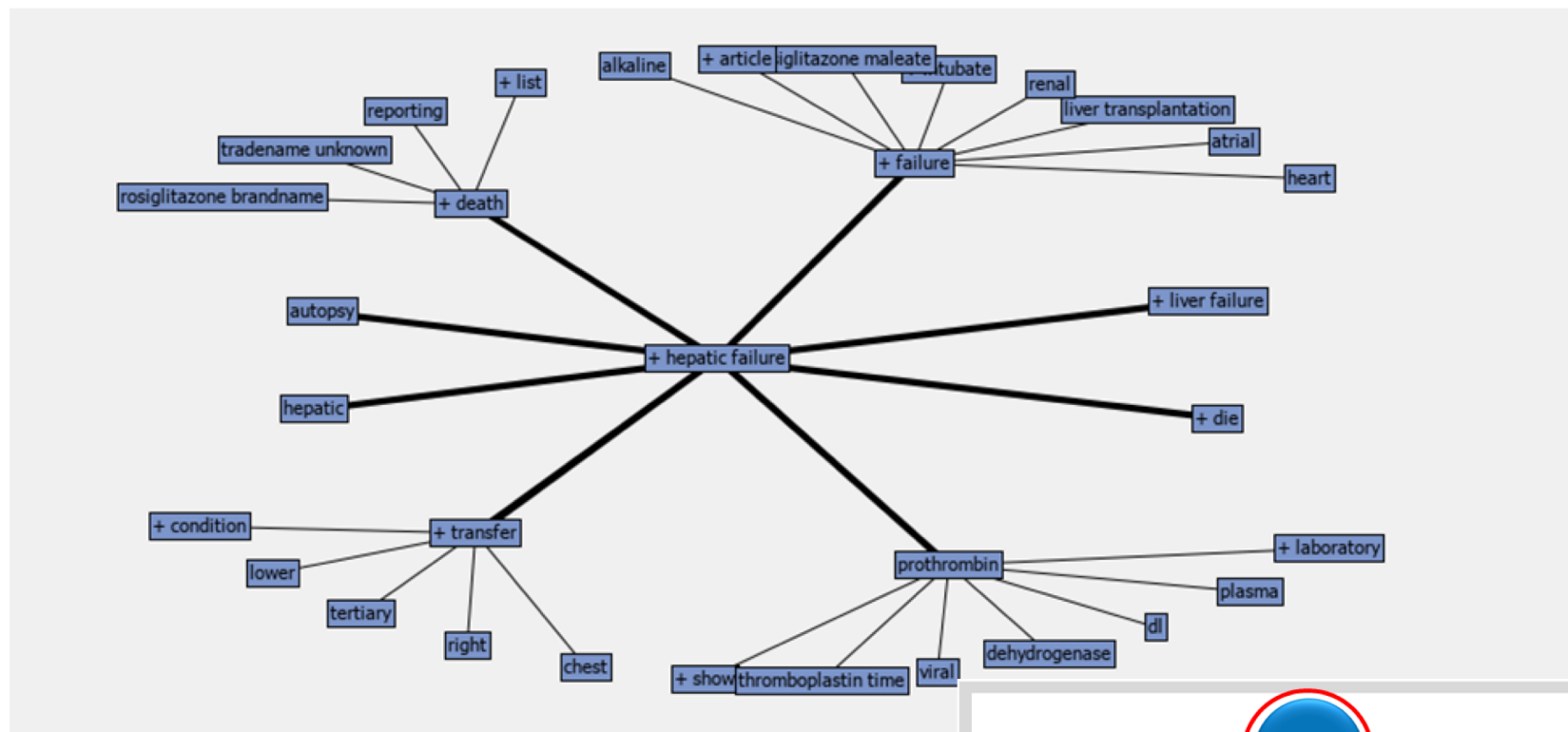
  

```

graph TD
    NLP((Natural Language Processing)) --- TextMining((Text Mining))
    NLP --- DeepLearning((Deep Learning))
    NLP --- ContentCategorization((Content Categorization))
    NLP --- SentimentAnalysis((Sentiment Analysis))
    NLP --- Search((Search))
    NLP --- MachineLearning((Machine Learning))
    NLP --- Summarization((Summarization))
    NLP --- EntityFactExtraction((Entity & Fact Extraction))
  
```

# PREDICTIVE ANALYSIS

## Text Analytics-Concept Linking



# PREDICTIVE ANALYSIS

## Supervised & Unsupervised Models

MBR

Decision  
Tree

Text Rule  
Builder

Text Topic

Neural  
Network

Text  
Cluster

Regression

# PREDICTIVE ANALYSIS

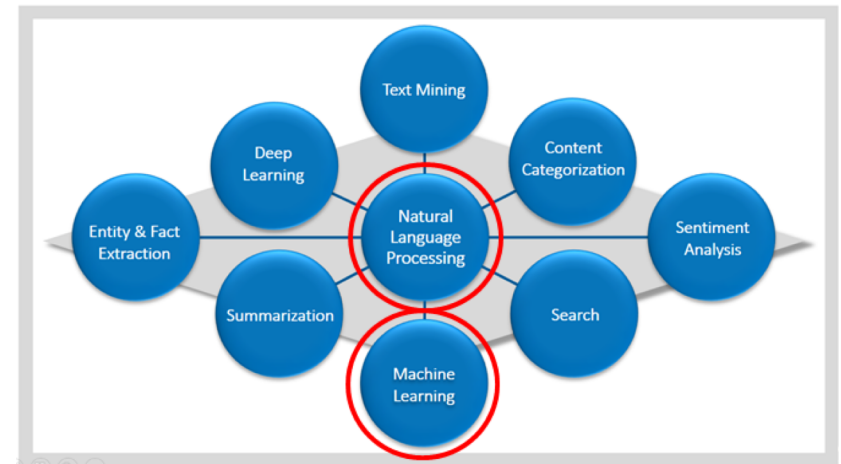
## Text Categorization-Rule Building

- **Boolean rules** are created from subsets of terms to predict a target variable.

- Different setting for the

- **complexity,**
- **exhaustiveness,**

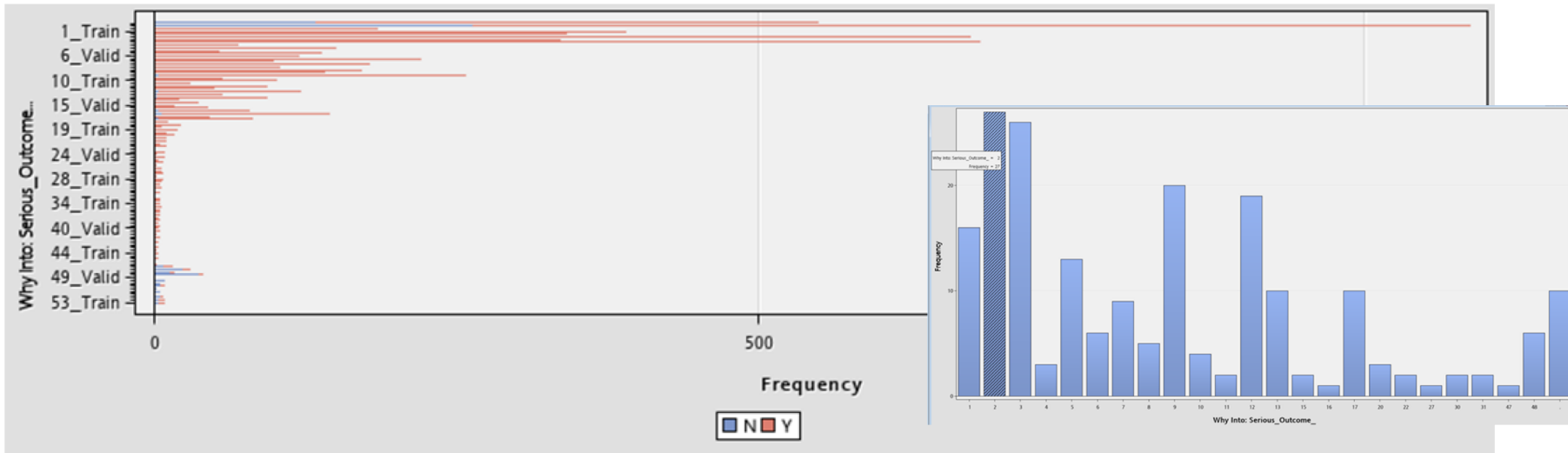
are applied to prevent **overtraining** and several analyses were compared to set the **optimal** text rule builder setting.



# PREDICTIVE ANALYSIS

## Text Categorization-Rule Building

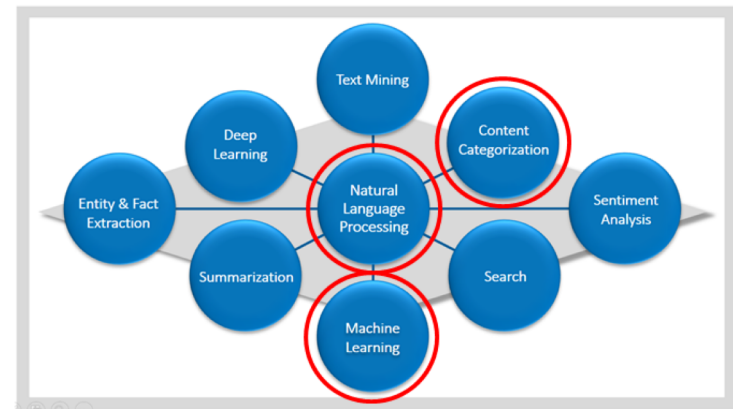
| Target Value | Rule # | Rule                                                                                                   | Precision | Recall | F1 score | Valid Precision | Valid Recall | Valid F1 score | True Positive/ Total | Valid True Positive/ Total |
|--------------|--------|--------------------------------------------------------------------------------------------------------|-----------|--------|----------|-----------------|--------------|----------------|----------------------|----------------------------|
| Y            | 1      | clinical outcome                                                                                       | 100.0%    | 8.50%  | 15.67%   | 100.0%          | 8.06%        | 14.92%         | 390/390              | 185/185                    |
| Y            | 2      | concomitant medication & drug                                                                          | 100.0%    | 23.19% | 37.64%   | 99.62%          | 22.83%       | 37.15%         | 722/722              | 370/372                    |
| Y            | 3      | medical & ~drug induced liver injury & ~auto-antibody & ~opioid dependence & ~epstein-barr & hepatitis | 100.0%    | 38.05% | 55.12%   | 99.77%          | 37.47%       | 54.48%         | 1,093/1,093          | 571/571                    |
| Y            | 4      | mg                                                                                                     | 100.0%    | 41.32% | 58.47%   | 99.79%          | 40.48%       | 57.59%         | 334/334              | 156/156                    |
| Y            | 5      | article                                                                                                | 100.0%    | 44.35% | 61.44%   | 99.80%          | 42.83%       | 59.94%         | 243/243              | 88/88                      |
| Y            | 6      | information & ~treatment information & ~auto-antibody & ~normal liver & ~fall & initial                | 100.0%    | 49.16% | 65.92%   | 99.73%          | 48.02%       | 64.82%         | 580/580              | 309/310                    |
| Y            | 7      | rezulin & ~lts                                                                                         | 100.0%    | 53.04% | 69.32%   | 99.75%          | 52.29%       | 68.61%         | 582/582              | 292/292                    |
| Y            | 8      | iu & ~healthy                                                                                          | 100.0%    | 56.77% | 72.42%   | 99.77%          | 56.82%       | 72.40%         | 697/697              | 356/356                    |
| Y            | 9      | failure & ~normal                                                                                      | 99.83%    | 62.28% | 76.70%   | 99.65%          | 62.88%       | 77.10%         | 698/703              | 375/377                    |
| Y            | 10     | suspect                                                                                                | 99.76%    | 64.44% | 78.30%   | 99.53%          | 65.23%       | 78.81%         | 623/625              | 317/319                    |
| Y            | 11     | induce & ~ana                                                                                          | 99.77%    | 66.49% | 79.80%   | 99.41%          | 66.45%       | 79.66%         | 460/462              | 213/215                    |
| Y            | 12     | significant & ~specific                                                                                | 99.65%    | 69.03% | 81.57%   | 99.37%          | 68.58%       | 81.15%         | 900/906              | 465/467                    |
| Y            | 13     | daily                                                                                                  | 99.57%    | 71.02% | 82.91%   | 99.21%          | 70.89%       | 82.69%         | 1,163/1,163          | 607/613                    |



# PREDICTIVE ANALYSIS

## Text Categorization-Clustering

- To utilize unstructured data, a text cluster is built.
- Latent Semantic Analysis (LSA) derives a numeric representation for each document.
  - terms and documents organized into a common semantic space based upon term co-occurrence.
- Derived numeric representation produces the clusters.
- An EM algorithm describes the contents clusters.

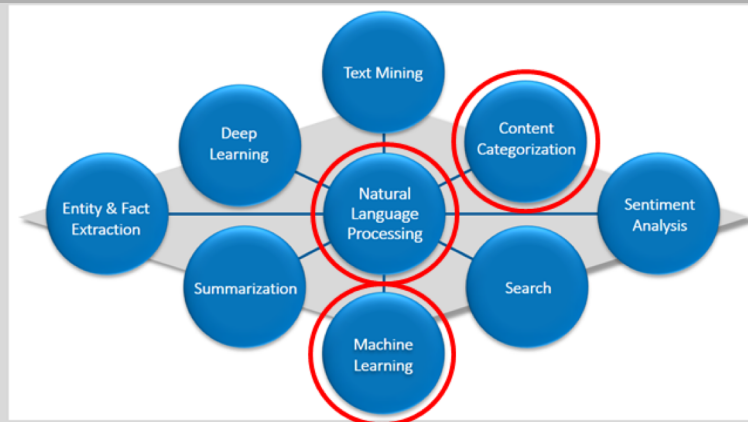




# PREDICTIVE ANALYSIS

## Text Categorization-Clustering

| Cluster ID | Descriptive Terms                                                                                                                                                                                                                                                                                                       |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1          | mi-qod betaseron-qod                                                                                                                                                                                                                                                                                                    |
| 2          | ap dronedarone ggt +value causal alat +relationship +increase further present +comment +recover written +bilirubin +phosphatase first alt +antibody +study alkaline +day substantial sclerosis +admit +purpose                                                                                                          |
| 3          | ast alt troglitazone alanine aminotransferase aspartate +additional information' aspartate aminotransferase' alanine aminotransferase' +total bilirubin' +request rezulin alkaline additional +alkaline phosphatase' total +type +clinical outcome' +phosphatase +laboratory follow-up daily +value diabetes +bilirubin |
| 4          | tsabri unknown iv 'monitored therapy' +onset 'tsabri therapy' +program monitored qm biogen therapy +enrol idex sclerosis 'biogen idex' +update 'multiple sclerosis' +causality +assess female 'onset unknown' +confirm +mar +nurse written                                                                              |
| 5          | disease +perform +biopsy +hospital +admit +alcohol +week first +condition +body +cell jaundice +initiate present acute +follow +severe +grade literature +injury +confirm +relationship +pain beta +failure                                                                                                             |
| 6          | +program betaseron sclerosis +interferon +refer 'multiple sclerosis' beta +consumer +state patient +reporter +dose unspecified +condition previously +causality +request +number +enzyme +drug +event +follow-up +confirm +time +outcome                                                                                |
| 7          | troglitazone 'cardiovascular system' +liver damage' +liver problem' +medical condition' +potential injury' +unspecified medical condition' +body +complaint +condition +damage +develop +experience +form +injury +pain +system +threaten attorney cardiovascular experienced life medical mental patient               |
| 8          | +antibody +disorder 'diabetes mellitus' mellitus lu dl sgot hepatic daily sgot +purpose +function +laboratory +value diabetes mg +body +test +reference +condition +total bilirubin' total +comment +follow +mar                                                                                                        |
| 9          | +tracking number' +assign +purpose +track naltrexone +no further detail' +reference +initiate +number +detail spontaneous male oral unknown serious +case +time further +unknown date' +resolve alanine +status registered +event aminotransferase                                                                      |
| 10         | tsabri unknown sclerosis iv 'multiple sclerosis' +assess +causality +onset 'tsabri therapy' +update qm +unknown date' +reference 'onset unknown' +mar apr written 'old female patient' female +confirm significant +antibody +nurse medically +diagnose                                                                 |
| 11         | +onset lu mellitus health 'diabetes mellitus' +transaminase sgot sgot diabetes +damage +physician concomitant daily +case +hospitalize +history mg +medical history' medical +increase causal total +concomitant medication' +number hepatic                                                                            |
| 12         | alk phos 'alk phos' dronedarone sgot sgot written abdominal normal +liver biopsy' +alkaline phosphatase' er +ultrasound alkaline +normal +phosphatase alt 'monitored therapy' +total bilirubin' total ast +biopsy monitored +study +threaten                                                                            |
| 13         | +nurse registered +request troglitazone +additional information' +type rezulin +total bilirubin' ggt 'diabetes mellitus' 'old female patient' diabetes total mellitus additional further +old aspartate 'aspartate aminotransferase' +clinical outcome' ast follow-up +discontinue +troglitazone therapy' +begin        |
| 14         | +study +enrol +grade +result literature dl +update acute +dose ap +medical condition' alanine aminotransferase' +disorder +assign spontaneous oral aspartate biogen causal concomitant +ultrasound +detail +drug +relate +include                                                                                       |
| 15         | iv +grade +cell +study +relationship +enrol +death acute er +condition +assess +die +dose +treat +initiate causal +male +system +failure +severe +blood +reference clinical +comment +perform                                                                                                                           |
| 16         | iv spontaneous +infection +assign +detail +comment +relationship +initiate serious +purpose +admit +male acute +day +condition sgot +reporter potential oral +antibody +hospital +treat causal sgot +causality                                                                                                          |
| 17         | beta betaseron +comment +female literature +severe +increase pt +cause 'alk phos' phos sclerosis +enzyme previously alk +cell apr +interferon +medication first +confirm +condition +begin +biopsy +follow-up                                                                                                           |
| 18         | er pt +begin +admit +pain +ultrasound jaundice +study +no further detail' +detail +lab +discharge ap +enzyme +system alk elevated +female +day +time normal +complaint phos first +week                                                                                                                                 |
| 19         | +ultrasound pt +disease +show +infection +status +female patient' hepatic negative +decrease +level betaseron alat +cause +normal +resolve acute first mental +test normal abdominal literature +reveal +day                                                                                                            |
| 20         | alat +male first acute hepatic +infection oral dronedarone +dose +comment causal +drug +resolve +decrease +relationship +event spontaneous daily 'aspartate aminotransferase' aminotransferase +treat +grade +female patient' alanine aminotransferase' +disorder                                                       |
| 21         | liver literature +alcohol +comment er +disorder ap +injury +biopsy written +liver biopsy' potential normal +assign +form +week +condition +liver problem' +enzyme medical serious +month +drug +begin +cell                                                                                                             |
| 22         | march alat +treatment +severe literature +male oral +injury +medical condition' +death +dose +die +time +unknown date' male +treat hepatic +follow dronedarone +day +condition +case +failure jaundice +cause                                                                                                           |
| 23         | sgot +state +consumer +admit +cause +death +hospital sgot +treat +blood +follow-up +discharge +infection +physician +reveal +cell +day +form +biopsy pt +die elevated follow-up +enzyme +perform                                                                                                                        |
| 24         | liver troglitazone +damage unspecified sgot diabetes sgot rezulin mellitus +injury 'diabetes mellitus' +liver damage' +troglitazone therapy' +clinical outcome' attorney anguish +patient +pain jaundice daily +physician medically summons +year +discontinue                                                          |
| 25         | phos alk 'alk phos' +liver biopsy' pt +biopsy limited literature +type +test rezulin +month +perform 'diabetes mellitus' +troglitazone therapy' sgot +improve hepatitis abdominal troglitazone +hospital +relationship mellitus +antibody negative                                                                      |

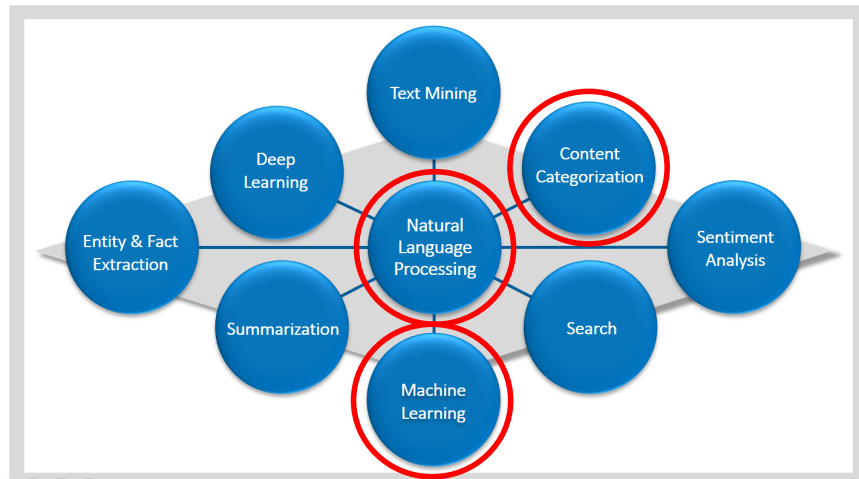


# PREDICTIVE ANALYSIS

## Supervised Model-Decision Tree

- Decision Tree is developed to perform:
  - Predict new cases
  - Select useful inputs
  - Optimize complexity.

| Predictive Modeling Task | General Principle                        | Decision Trees   |
|--------------------------|------------------------------------------|------------------|
| Predict new cases        | Decide, rank, or estimate                | Prediction Rules |
| Select useful inputs     | Eradicate redundancies and irrelevancies | Split Search     |
| Optimize complexity      | Tune models with validation data         | Pruning          |



# PREDICTIVE ANALYSIS

## Supervised Model-Decision Tree



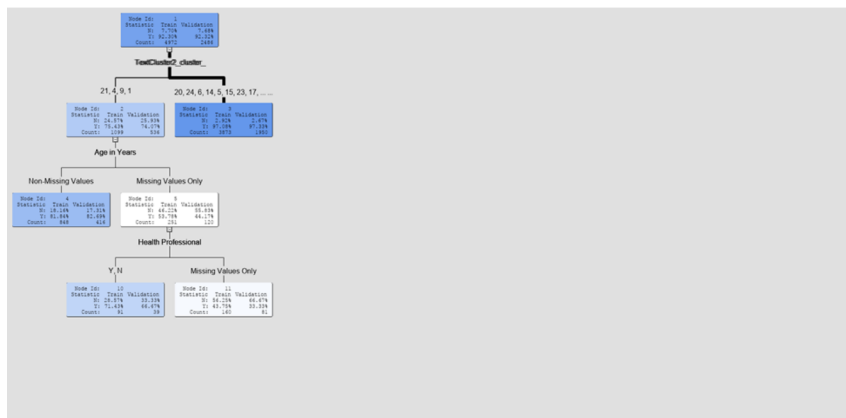
- The **output** from the cluster analysis is the **input** to the decision tree modeling.

| Cluster ID | Descriptive Terms                                                                                                                                                                                                                                                                                       |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1          | ...                                                                                                                                                                                                                                                                                                     |
| 2          | +grade clofarabine +leukemia cytarabine myeloid transplantation +regimen 'acute myeloid leukemia' aml +group median +study acute iv +cell +investigator +action kg +enrol +dose +age +die +platelet +total +count                                                                                       |
| 3          | +lesion +mri 'multiple sclerosis' sclerosis +age multiple +feature literature +author prior +neurologist left +suspect +refer +diagnose 'drug indication' +symptom +condition +level hepatic +sample +dose spontaneous +case +indication                                                                |
| 4          | cirrhosis troglitazone rezulin portal +esophageal varix +bleed fibrosis chronic +biopsy gastrointestinal +type elevated significant +liver biopsy' hypertension medically additional mellitus 'diabetes mellitus' +physician diabetes +request ascites +relate unknown                                  |
| 5          | liver jaundice +failure +pain ascites unspecified abdominal +injury heart +liver failure' +problem +abdominal pain' +unspecified injury' troglitazone +disease rezulin +request hepatitis +encephalopathy +die elevated diabetes autopsy +total bilirubin' +laboratory                                  |
| 6          | +patient troglitazone medical +increase +physician +discontinue +outcome daily +value experienced +recover serious unspecified +damage +transaminase sgpt sgot health diabetes iu mellitus 'diabetes mellitus' +concomitant medication' +condition follow-up                                            |
| 7          | +bleed gastrointestinal +failure +state +severe heart +die +liver damage' +cause +note +blood dilin +set rezulin autopsy +develop +encephalopathy jaundice +unit +death +sample march +period acute +continue                                                                                           |
| 8          | +death +die +cause +failure autopsy +liver failure' rezulin troglitazone heart +perform +hospital +late +develop +request +level unknown +encephalopathy hepatic diabetes +problem +physician +patient +condition additional +laboratory                                                                |
| 9          | portal +vein fibrosis +liver biopsy' +biopsy +cell cirrhosis chronic hypertension +lesion +show +note +mri hla +severe splenectomy +etiology +grade +score negative ascites gaucher +encephalopathy +disease +associate                                                                                 |
| 10         | gemtuzumab +result ozogamicin mylotarg acute myeloid 'gemtuzumab ozogamicin' aml ascites +concomitant therapy' +investigator +leukemia 'acute myeloid leukemia' +study cytarabine +assessment relatedness +injection +product +enrol 'blood bilirubin' +medical history' +cell 'alanine ...             |
| 11         | +count +cell +platelet white +blood dl +decrease iu +find abnormal +control +hospitalize 'diabetes mellitus' +injection mellitus follow-up negative binding 'alanine aminotransferase' alanine +total bilirubin' ap +old mg +year                                                                       |
| 12         | rosiglitazone max 'rosiglitazone maleate' maleate +liver injury' march 'unknown time' +period hepatotoxicity 'drug indication' 'severe hepatotoxicity' withrosiglitazone +accompany +article +author +define +determine +encephalopathy +subject transplantation +etiology +exclude +liver failure' ... |
| 13         | +pain left +abdominal pain' heart abdominal +present +etiology +type +note +effect +replacement +hospital +show +symptom opinion elevated +request review +site prior +accompany quality +late chronic +mri                                                                                             |
| 14         | cerezyme gaucher 'gaucher disease' transplantation +replacement 'enzyme replacement therapy' 'gaucher patient' +gaucher cell' splenectomy +genotype +esophageal varix' opinion unrelated kg +enzyme +bleed +begin fibrosis +regimen march +cell +type +age +severe +determine ...                       |

- Two decision tree models have been developed.
- 1<sup>st</sup> tree: the SVD numeric values have been **rejected** only the nominal values of cluster numbers will **input** the decision tree modeling with other FAERS input variables.
- 2<sup>nd</sup> tree: the SVDs is **utilized** as input to the decision tree with other FAERS variables and cluster number variable has been **rejected**.

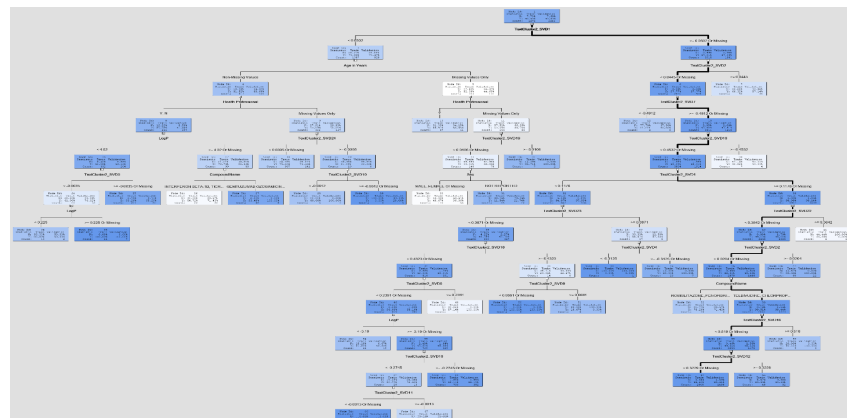
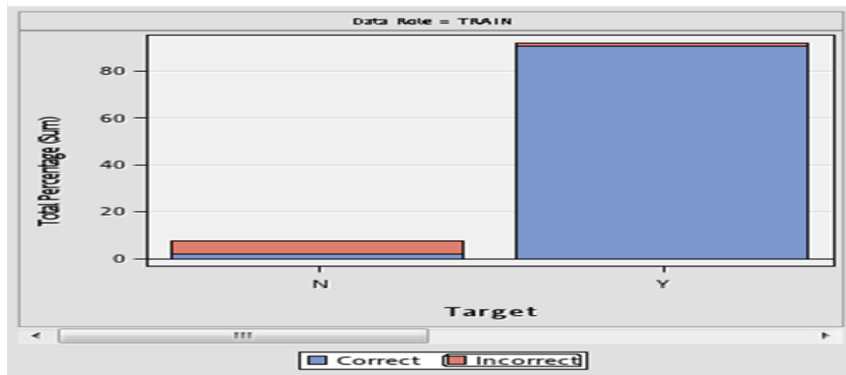
# Predictive Analysis

## Supervised Model-Decision Tree



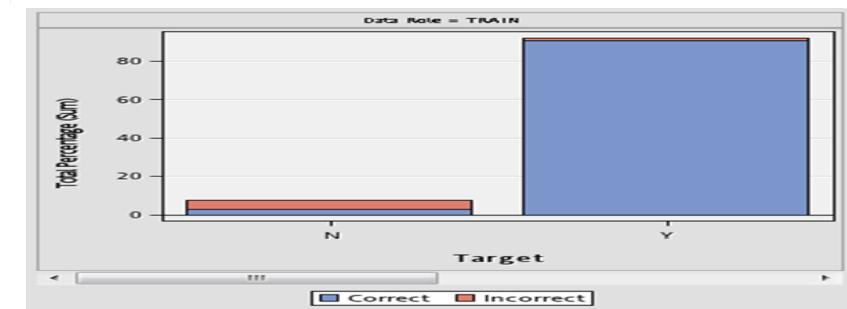
Variable Importance

| Variable Name         | Label               | Number of Splitting Rules | Importance | Validation Importance | Ratio of Validation to Training Importance |
|-----------------------|---------------------|---------------------------|------------|-----------------------|--------------------------------------------|
| TextCluster2_cluster_ |                     | 1                         | 1.0000     | 1.0000                | 1.0000                                     |
| Age_in_Years          | Age in Years        | 1                         | 0.6163     | 0.7513                | 1.2189                                     |
| Health_Professional   | Health Professional | 1                         | 0.3328     | 0.3681                | 1.1060                                     |



Variable Importance

| Variable Name       | Label               | Number of Splitting Rules | Importance | Validation Importance | Ratio of Validation to Training Importance |
|---------------------|---------------------|---------------------------|------------|-----------------------|--------------------------------------------|
| TextCluster2_SVD1   |                     | 2                         | 1.0000     | 1.0000                | 1.0000                                     |
| Age_in_Years        | Age in Years        | 1                         | 0.6182     | 0.8019                | 1.2972                                     |
| TextCluster2_SVD2   |                     | 2                         | 0.4848     | 0.5093                | 1.0506                                     |
| Health_Professional | Health Professional | 2                         | 0.4552     | 0.5756                | 1.2645                                     |
| LogP                | LogP                | 3                         | 0.3535     | 0.1742                | 0.4929                                     |
| TextCluster2_SVD24  |                     | 3                         | 0.3530     | 0.3004                | 0.8510                                     |
| TextCluster2_SVD10  |                     | 1                         | 0.3489     | 0.0000                | 0.0000                                     |
| CompoundName        | CompoundName        | 2                         | 0.3354     | 0.2472                | 0.7370                                     |
| TextCluster2_SVD4   |                     | 1                         | 0.3160     | 0.2357                | 0.7460                                     |
| TextCluster2_SVD5   |                     | 2                         | 0.2668     | 0.0906                | 0.3370                                     |
| Sex                 | Sex                 | 1                         | 0.2504     | 0.0576                | 0.2301                                     |
| TextCluster2_SVD11  |                     | 1                         | 0.2491     | 0.2753                | 1.1053                                     |
| TextCluster2_SVD9   |                     | 1                         | 0.2379     | 0.2486                | 1.0448                                     |
| TextCluster2_SVD23  |                     | 1                         | 0.2296     | 0.0000                | 0.0000                                     |
| TextCluster2_SVD22  |                     | 1                         | 0.2012     | 0.1759                | 0.8715                                     |
| TextCluster2_SVD16  |                     | 2                         | 0.1953     | 0.1696                | 0.8666                                     |
| TextCluster2_SVD19  |                     | 1                         | 0.1805     | 0.2112                | 1.1697                                     |
| TextCluster2_SVD12  |                     | 1                         | 0.0640     | 0.0989                | 1.5276                                     |

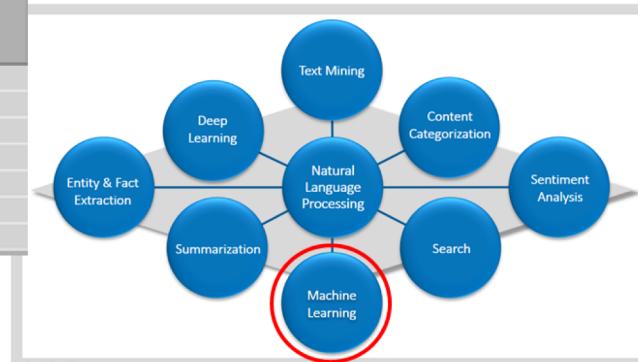


# Predictive Analysis

## Model Comparison Methodology

- Summary statistics used for model comparison.
- Summary statistics transform model performance to numerical scores.
- The model comparison tool collates statistics from different modeling nodes for easy comparison.

| Selected Model ▼ | Predecessor Node | Model Node | Model Description                 | Target Variable  |
|------------------|------------------|------------|-----------------------------------|------------------|
| Y                | Meta10           | Tree       | Decision Tree_Probability         | Serious_Outcome_ |
|                  | Meta13           | TextRule2  | Text Rule Builder_Aggressive      | Serious_Outcome_ |
|                  | Meta8            | Tree2      | Decision Tree_DefaultDecisionT... | Serious_Outcome_ |
|                  | Meta12           | TextRule7  | Text Rule Builder_VeryAggressive  | Serious_Outcome_ |
|                  | Meta11           | TextRule3  | Text Rule Builder_Default         | Serious_Outcome_ |
|                  | Meta9            | MBR        | MBR                               | Serious_Outcome_ |
|                  | Meta7            | Neural     | Neural Network                    | Serious_Outcome_ |

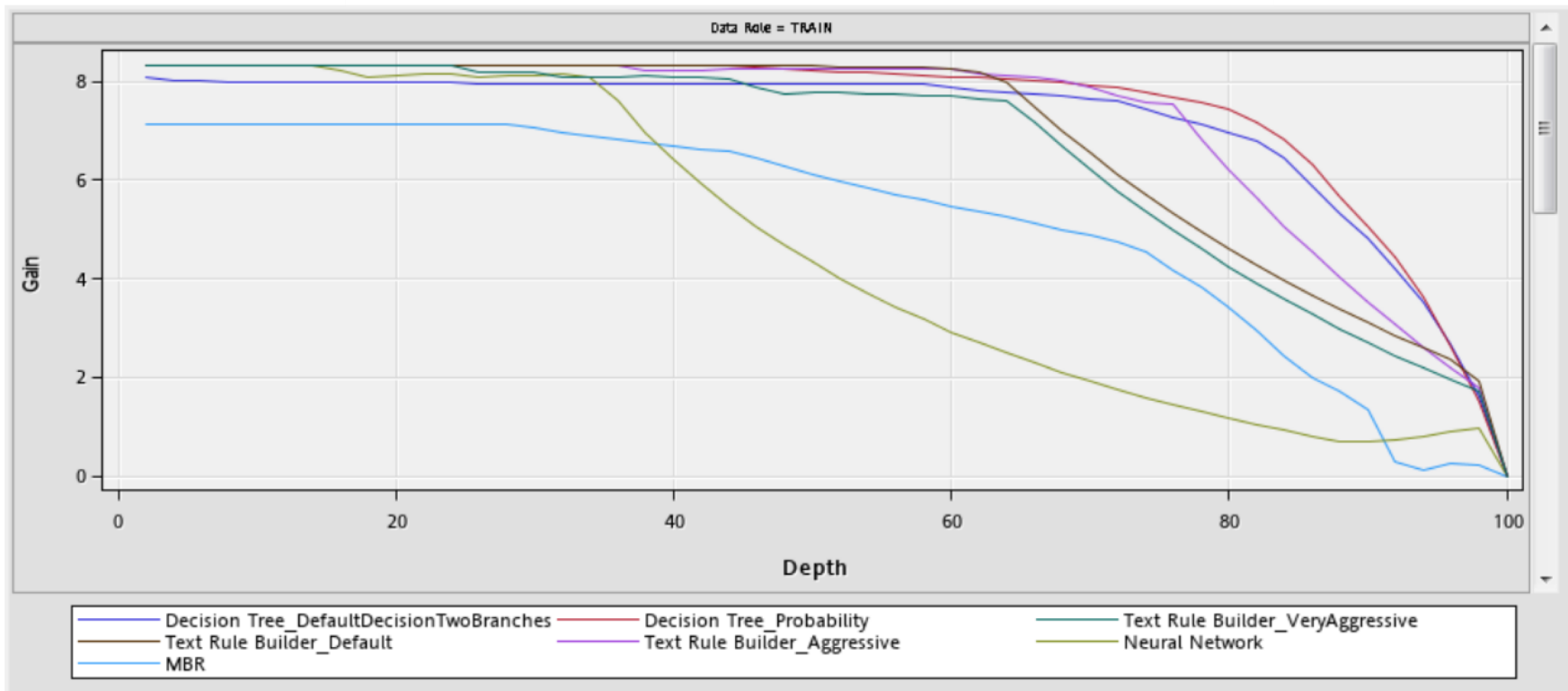


# Predictive Analysis

## Model Comparison Methodology

### Gain Chart

$$\text{Gain} = \left[ \frac{\% \text{ of events in decil}}{\text{random \% events in decile}} - 1 \right]$$



# Scoring

- The **champion model** classifies new data that **do not** contain a target value (i.e., no serious outcome)
- 4000 cases** have been selected randomly and assigned the role of score data set.

| Narrative                                                                                                | Probability for level Y of Serious_Outcome_ | Probability of Classification | Prediction for Serious_Outcome_ | Target Variable: Serious_Outcome_ ▲ |
|----------------------------------------------------------------------------------------------------------|---------------------------------------------|-------------------------------|---------------------------------|-------------------------------------|
| This spontaneous report was received from a MRI technician refers to a female patient of unknown ...     | 0.315789                                    | 0.684211N                     | N                               |                                     |
| Report of liver damage, splenomegaly, portal hypertension, ascites, listlessness, and withdrawal sy...   | 0.166667                                    | 0.833333N                     | N                               |                                     |
| A 48 year old female patient on TYSABRI (300 mg, IV, QM) since 01 Oct 2006 reported experiencing l...    | 0.076923                                    | 0.923077N                     | N                               |                                     |
| A 43 year old female patient on TYSABRI (300 mg, IV, QM) for Multiple Sclerosis since 17 Dec 2012 ...    | 0.076923                                    | 0.923077N                     | N                               |                                     |
| A currently 64 year old female patient on TYSABRI (300 mg, IV, QM) for Multiple Sclerosis since 14 M...  | 0.884615                                    | 0.884615Y                     | N                               |                                     |
| A 51 year old female patient on TYSABRI (300 mg, IV, QM), for Multiple Sclerosis, from 18 Jan 2012 t...  | 0.076923                                    | 0.923077N                     | N                               |                                     |
| Case reference number 2016BI00283601 is a Spontaneous post marketing case report by a nurse....          | 0.941176                                    | 0.941176Y                     | N                               |                                     |
| Case number# PHEH2017US025482, is an initial spontaneous report received from a consumer (p...           | 0.666667                                    | 0.666667Y                     | N                               |                                     |
| This is a spontaneous report from a contactable physician. A patient of unspecified age, ethnicity, a... | 0.614035                                    | 0.614035Y                     | N                               |                                     |
| ** Abstracted per FDA rep **Drug Induced Liver InjuryA 34 year old man with history of intravenous dr... | 0                                           | 1N                            | N                               |                                     |
| GI BLEEDESOPHAGEAL VARICESCIRRHOIS WITH MICRONADULAR CHANGESCHRONIC ACTIV...                             | 0.904762                                    | 0.904762Y                     | N                               |                                     |
| This is a spontaneous report received from a consumer (via medical information) which described ...      | 0.97619                                     | 0.97619Y                      | N                               |                                     |
| Serious; related; unlistedThis case report from BRAZIL was derived from medical literature on 24-J...    | 0.990991                                    | 0.990991Y                     | Y                               |                                     |
| Initial report received from a patient counselor executive on 29 Jan 2010: This patient was prescribe... | 1                                           | 1Y                            | Y                               |                                     |
| Case number PHHY2012KR060672 is an initial PMS (CLDT600AKR01) report received from an inv...             | 1                                           | 1Y                            | Y                               |                                     |
| This serious post-marketing case report from France received on 10-MAR-2014 by an investigator d...      | 0.991131                                    | 0.991131Y                     | Y                               |                                     |
| THIS IS A SECOND FOLLOW UP TO A REPORT INITIALLY SUBMITTED ON 08APR99. THE FIRST F...                    | 1                                           | 1Y                            | Y                               |                                     |
| THIS IS A FOLLOW-UP BASED ON INFORMATION REPORTED TO PFIZER ON 04MAY00. THE INITI...                     | 0.723684                                    | 0.723684Y                     | Y                               |                                     |
| This spontaneous case, reported by a cardiologist via a company sales representative, concerns a f...    | 0.941176                                    | 0.941176Y                     | Y                               |                                     |
| Initial consumer report was received from the patient's wife on 28 Sep 2009: This polymedicated pat...   | 1                                           | 1Y                            | Y                               |                                     |
| Septic shockDrug ineffectiveRenal failure acuteOesophageal varices haemorrhage                           | 0.972222                                    | 0.972222Y                     | Y                               |                                     |
| Literature report: Boyle-Vavra S, Jones M, Gourley BL, Holmes M, Ruf R, Balsam A, et al. Comparativ...   | 0.991131                                    | 0.991131Y                     | Y                               |                                     |
| Initial report received on 24 Jun 2009 : Patient 17 from centre 1486 was enrolled in study CSPP10...     | 1                                           | 1Y                            | Y                               |                                     |
| THIS IS A FOLLOW-UP REPORT BASED ON INFORMATION REPORTED TO PFIZER ON 02AUG99. ...                       | 1                                           | 1Y                            | Y                               |                                     |
| Information was received from a healthcare professional describing a 68 year-old female who recei...     | 1                                           | 1Y                            | Y                               |                                     |

## Scoring

- The champion model classifies new data that do not contain a target value (i.e., no serious outcome)
- 4000 cases have been selected randomly and assigned the role of score data set.

| Probability of Classification | Prediction for Serious_Outcome_ | Target Variable: Serious_Outcome_ ▲ |
|-------------------------------|---------------------------------|-------------------------------------|
| 0.684211N                     | N                               | N                                   |
| 0.833333N                     | N                               | N                                   |
| 0.923077N                     | N                               | N                                   |
| 0.923077N                     | N                               | N                                   |
| 0.884615Y                     | N                               | N                                   |
| 0.923077N                     | N                               | N                                   |
| 0.941176Y                     | N                               | N                                   |
| 0.666667Y                     | N                               | N                                   |
| 0.614035Y                     | N                               | N                                   |
| 1N                            | N                               | N                                   |
| 0.904762Y                     | N                               | N                                   |
| 0.97619Y                      | N                               | N                                   |
| 0.990991Y                     | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |
| 0.991131Y                     | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |
| 0.723684Y                     | Y                               | Y                                   |
| 0.941176Y                     | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |
| 0.972222Y                     | Y                               | Y                                   |
| 0.991131Y                     | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |
| 1Y                            | Y                               | Y                                   |



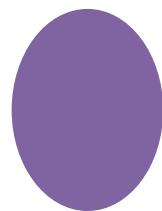
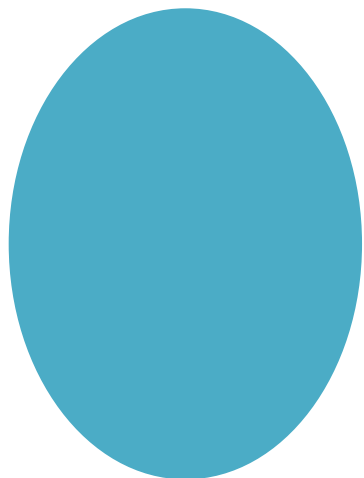
## Assessment Objectives - Revisited

- **Uncovered relationships** between these drugs and hepatic failure
- Offered information on **drug combinations** related to hepatic failure
- Clarified factors that can predict a greater degree of hepatic failure from **serious events** (death) versus **less serious** (treatable) events
- Assisted with **rating** the impact of using these drugs on patients

# Discussion and Conclusions

- Visualization of results in interactive reporting tool
- Model improvement
- Application to other adverse event scenarios





Q & A

Thank you !

