

**THE
POWER
TO KNOW.**

Apport du Text Mining pour l'amélioration de la performance des modèles de Data Mining

Grégoire de Lassence



ESSEC
M.S.



- **Grégoire de Lassence**
Responsable Pédagogie et Recherche Département Académique

Tel : +33 1 60 62 12 19

gregoire.delassence@fra.sas.com

<http://www.sas.com/france/academic>



Présentation suite à :

Apport du Text Mining pour prédire le potentiel
d'accident d'un véhicule en assurance

BMBAF 2007 – EGC 2007 – Namur Belgique

Grégoire de Lassence, Sorin Moga, Philippe Lenca,

GET/ENST Bretagne – UMR 2872 TAMCIC

Avec la participation de

Manu Carricano

INSEEC

Plan

- Introduction au Text Mining
- Présentation du Cas NHTSA
- Le processus de Text Mining
- Application au cas NHTSA
- Conclusion

Le Text Mining ?

- “La quantité d’information non structurée double tous les deux mois dans les grandes entreprises” – IDC
- “Aujourd’hui, les entreprises ayant mis en place un système de gestion des données non structurées ont accru leur productivité de 15 % en moyenne”– Basex
- “L’employé qualifié moyen passe 2h30 / jour à rechercher des documents” - IDC – March 2002
- “80-90% de l’information figurant sur le net ou sur les réseaux d’entreprise n’est pas structurée” - Goldman Sachs

Qu'est ce que le Text Mining ?

- Processus automatique qui consiste à extraire l'information à partir de données textuelles, et à l'utiliser afin d'améliorer les décisions prises en entreprise.
- Nouvelle technologie qui :
 - Combine les analyses
 - Linguistiques
 - statistiques
 - détermine les idées clés contenues dans un corpus de documents (analyse descriptive)
 - affecte les documents traitant d'un même thème dans à une même classe (analyse prédictive)

	MODELTX	CRASH	FAIL_DATE	FIRE	SUMMARY	INJURIES	DEATHS	CO
1	TOWN AND COUNTRY	N	97-05-22	Y	TRANSMISSION COOLING UNIT AND LINES LEAK, CAUSING VEHICLE FIRE. *AK	0	0	POWER
2	BLAZER	N	97-07-02	N	BRAKES. THE BRAKES WERE INOPERATIVE. TOOK VEHICLE TO DEALER AND WAS TOLD THAT THE RETURN SPRINGS FOR THE S	0	0	BRAKES
3	S10	N		N	WINDSHIELD WIPERS ARE NOT WORKING AT ALL. *AK	0	0	VISUAL
4	BLAZER	N	97-07-02	N	BRAKES. THE BRAKES WERE INOPERATIVE. TOOK VEHICLE TO DEALER AND WAS TOLD THAT THE RETURN SPRINGS FOR THE S	0	0	BRAKES
5	VOYAGER	N	96-12-19	N	TRANSMISSION SHUDDERS WHEN STARTED UP FROM A STOPPED POSITION AFTER THE VEHICLE HAS BEEN WARMED UP. SOMETI	0	0	POWER
6	VOYAGER	N	96-12-19	N	TRANSMISSION SHUDDERS WHEN STARTED UP FROM A STOPPED POSITION AFTER THE CAR HAS WARMED UP. AT TIMES, IT ALSO	0	0	POWER
7	PICKUP	Y	95-09-24	N	FRONT SUSPENSION BROKE, LOST CONTROL OF VEHICLE THEN IT FLIPPED OVER. *AK CONSUMER WAS INJURED IN ACCIDENT.	1	0	SUSPE
8	CARAVAN	Y	91-09-25	N	BRAKE FAILURE CAUSED ACCIDENT. LW	1	0	BRAKES
9	K35	Y	94-04-15	N	BRAKES FAILED WITHOUT WARNING. CAUSED ACCIDENT. REPORTED BY TROOPER BONNELL; OWNER WILL PROVIDE ADDITIONAL I	1	0	BRAKES
10	S10	Y	91-02-19	N	LEFT FRONT UNIROVAL TIRE SEPARATED FROM RIM, VEHICLE SWERVED, HIT A UTILITY POLE, RESULTED IN NO APPARENT TIRE DA	0	0	TIRES
11	CARAVAN	Y	95-03-11	N	HR. PHUUPS WITNESSED INCIDENCE OF EJECTION OF A CHILD WHEN VEHICLE WAS HIT BROADSIDE, THEN REAR LIFTGATE FLEW O	1	1	STRUC
12	AEROSTAR	Y	94-09-30	N	ABS. APPLIED BRAKES, EXPERIENCED EXTENDED PEDAL FAGED, IMPACT 12:00 POSITION. SPEED UNKNOWN AIR BAG DID NOT DEPL.	1	0	BRAKES
13	S10	Y	94-04-15	N	PCV VALVE BLEW OFF HOSE, CAUSING HIGH VEHICLE IDLE, APPROX. 3000 RPM'S. TT	1	0	FUELF
14	VAN/WAGON	Y	91-07-08	N	HLE DRIVING VAN PULLED LEFT, ENDED UP IN CENTER DIVIDER, and FLIPPED OVER. LW	1	0	SUSPE
15	VAN/WAGON	Y	91-07-08	N	WHILE MAKING RIGHT TURN, REAR END OF VAN LOST TRACTION, WHEELS REACHED EDGE OF ROAD, and VAN ROLLED OVER. LW	1	0	POWER
16	LUMINA APV	Y	94-03-05	N	REAR SEAT BELTS DID NOT LOCK UP UPON IMPACT, EXTENDED OUT TO EXTREME, LEAVING OCCUPANTS UNRESTRAINED. TT	3	0	INTERI
17	CARAVAN	Y	94-11-04	N	UPON IMPACT AT 35MPH, 2ND ROW BUCKET SEAT BELT ANCHOR DETACHED FROM FLOOR, PASSENGER WAS THROWN OUT OF SEA	1	0	INTERI
18	DAKOTA	Y	89-09-08	N	TIE ROD BROKE RESULTING IN ACCIDENT. TRUCK ROLLED OVER TWICE. BI	0	0	STEERI
19	MPV	Y	91-01-31	N	WHILE PARKED IN DRIVEWAY, VAN DRIFTED BACKWARDS HITTING FENCE. SIPPED TRANSMISSION SYSTEM.	1	0	POWER
20	VOYAGER	Y	92-01-18	N	VEHICLE CROSSED CENTERLINE CAUSING ACCIDENT. *LDG	3	1	STEERI
21	VOYAGER	Y	93-09-21	N	FRONTAL COLLISION, 11:00 POSITION, VEHICLE WAS AT COMPLETE STOP, AIR BAG DID NOT DEPLOY. LAP / SHOULDER BELT FAILE	2	0	INTERI
22	F350	Y	93-11-30	N	REAR ABS BRAKES FAILED TO STOP VEHICLE EFFECTIVELY, CAUSING FRONT END COLLISION AT 15MPH. PLEASE DESCRIBE DETAI	1	0	BRAKES
23	PICKUP	Y	87-01-10	N	HIT POTHOLE IN DRIVEWAY AND STEERING BOX GAVE AND TRUCK HIT POLE. EC	1	0	STEERI
24	BLAZER	Y	95-03-28	N	CONSUMER IMPACTED ANOTHER VEHICLE AT 12:00 POSITION. SEAT BELTS DIDNT HOLD TO THEIR POTENTIAL. PASSENGERS BOTH	2	0	INTERI
25	PICKUP	Y	90-05-11	N	MASTER CYLINDER FAILURE, CAR WOULDN'T STOP NOR SLOW DOWN WHEN BRAKES APPLIED, FOOT ON BRAKE PEDAL ALMOST FL	5	0	BRAKES
26	CARAVAN	Y	94-12-25	N	HIT FROM THE LEFT HAND SIDE LATCH OPENED AND 2 PASSENGERS WERE THROWN OUT OF THE CAR. TT	1	0	STRUC
27	VOYAGER	Y	95-04-01	N	AUTOMATIC TRANSMISSION FAILURE. TT	0	1	POWER
28	TRUCK AND VAN LINK	Y	85-11-01	N	BRAKES DO NOT WORK PROPERLY ABOVE 40 MPH, MUST PUMP BRAKES TO STOP COMPLETELY, ACCIDENT OCCURED WHILE REVER	0	0	BRAKES
29	SAMURAI	Y	88-07-27	N	TURNUED WHEEL SHARPLY TO GET ON RIGHT SIDE OF ROAD, VEHICLE WENT OUT OF CONTROL, RLYING OVER 4 TIMES. BG	1	0	SUSPE
30	CARAVAN	Y	87-11-03	N	SUDDEN BRAKING CAUSED VEHICLE TO SWERVE. STRAIGHT CONTROLLED STOP DIFFICULT TO ACHIEVE, 180 DEGREE TURN RESUL	1	0	BRAKES
31	F150	Y	92-08-16	N	TIRE JACK SLIPPED, CAUSING INJURIES. *LDG	1	0	EQUIP
32	ASTRO	Y	94-01-21	N	IN STOP POSITION WHEN REAR ENDED BOTH FRONT SEAT BACKS COLLASPED AND MIDDLE SEAT COLLASPED AND PULLED WHERE	6	0	INTERI
33	RAM	Y	95-01-29	N	WORKING UNDER THE HOOD THE SPRING ON DRIVERS SIDE POPPED OFF AND HOOD CAME DOWN. TT	0	0	STRUC
34	MPV	Y	93-02-10	N	VAN SWERVED OVER TWO LANES OF FREEWAY, LOST STEERING CONTROL. AK	0	0	STEERI
35	MPV	Y	93-02-10	N	BRAKES LOCKED UP ON DRY ROADS. AK	0	0	BRAKES
36	AEROSTAR	Y	91-12-09	N	EMERGENCY BRAKE LEVER LOOSEMED, CAUSING THE LOCK UP OF THE REAR BRAKES AND RESULTING IN AN ACCIDENT. AK	2	0	BRAKES
37	AEROSTAR	Y	87-05-05	N	POWER STEERING PUMP FAILURE, LEAKS FLUID.	0	0	STEERI
38	DAKOTA	Y	95-03-22	N	VEHICLE LOST CONTROL IN HEAVY SNOW, CAUSING ACCIDENT. *SKD	0	0	SUSPE
39	RANGER	Y	88-05-19	N	SHIFTING TO 5TH GEAR, STICK BROKE OFF, RETAINING COLLAR MADE OF PLASTIC, RAN INTO DITCH. EC. AK	0	0	POWER
40	E350	Y	86-07-25	N	FAULTY FUEL SYSTEM RESULTS IN FUEL PRESSURE BUILD UP IN HOT / HUMID WEATHER. INJURIES HAVE OCCURRED WHEN FUEL S	2	0	FUELF
41	E350	Y	86-07-25	N	FAULTY FUEL SYSTEM RESULTS IN FUEL PRESSURE BUILD UP IN HOT / HUMID WEATHER. INJURIES HAVE OCCURRED WHEN FUEL S	2	0	FUELF
42	PATHFINDER	N	97-05-29	N	BOLTS THAT ARE CONNECTED TO THE HUB OF THE WHEEL BROKE OFF, CAUSING THE DRIVER'S FRONT TIRE TO COME OFF. *AK	0	0	WHEEL
43	E350	Y	86-07-25	N	FAULTY FUEL SYSTEM RESULTS IN FUEL PRESSURE BUILD UP IN HOT / HUMID WEATHER. INJURIES HAVE OCCURRED WHEN FUEL S	2	0	FUELF
44	E350	Y	86-07-25	N	FAULTY FUEL SYSTEM RESULTS IN FUEL PRESSURE BUILD UP IN HOT / HUMID WEATHER. SEVERAL PARAMEDIC INJURIES HAVE DC	7	0	FUELF
45	SUBURBAN	Y	87-07-09	N	WHEN TRUCK WAS HIT, SEAT BELT DIDNT RESTRAIN THE DRIVER CAUSING INJURIES. LOOSE SEAT BELT RESTRAINT. DM	1	0	INTERI
46	C15	Y	90-05-19	N	TIE ROD BROKE / SUDDENLY DRIVER UNABLE TO CONTROL DIRECTION OF VEHICLE BY TURNING STEERING WHEEL, RESULTED IN A	0	0	STEERI
47	DAKOTA	Y	96-08-05	N	VEHICLE HIT A STONE WALL HEAD ON AND THE DRIVER'S SIDE AIR BAG DID NOT DEPLOY. *AK	0	0	INTERI
48	BLAZER	Y	92-02-14	N	ACCIDENT / INJURY CAUSED BY SUDDEN ACCELERATION. *LDG	2	0	FUELT
49	BLAZER	Y	92-02-14	N	SUDDEN ACCELERATION CAUSED ACCIDENT / INJURY. *AJ	2	0	FUELT
50	BLAZER	N		N	THE CONSUMER CALLED AND SET AN APPOINTMENT AND ON THE DAY OF HIS APPOINTMENT THE DEALER DIDNT KNOW ANYTHING A	0	0	ELECT
51	CHRYSLER	N	87-04-10	N	FRONT DEFLECTOR SEAT BELT IS DAMAGED, IT WILL NOT UNDO. MUST. THE CENTER FROM THE MIDDLE SEAT DOES NOT CLAY	0	0	INTERI

La base US National Highway Traffic Safety Administration

[ftp://ftp.nhtsa.dot.gov/Consumer Complaints/](ftp://ftp.nhtsa.dot.gov/Consumer%20Complaints/)

Enterprise Miner - Miner

Fichier Edition Affichage Actions Options Fenêtre Aide

Echantillon Exploration Modification Modélisation Evaluation Utilitaires Credit Scoring

Table des matières

Miner

- Sources de données
- Diagrammes
 - HCM
 - HMEQ
 - MM
 - NHTSA_Truck_C
- Package de modèles
- Utilisateurs

Propriétés Valeur

ID	EMVWS4
Nom	NHTSA_Truck_C
Etat	Ouvert
Notes	
Historique	

ID

Identifiant du diagramme, correspondant au

TRUCKS

Data Partition

Arbre sans données textes

Régression Robuste sans données textes

Raisonnement à base de cas

Régression Robuste

Réseau de neurones

Arbre

Arbre données textuelles seulement

Text Miner

Comparaison de modèles

Exécution terminée.

Connecté à SASMain - Logical Workspace Server

Processus de Text Mining

- Analyse grammaticale (Text Parsing)
 - Découpage des textes en sous éléments :
 - Mots, phrases, groupes de mots, expressions particulières, entités, marques de ponctuation ...
- Analyse des éléments du discours et étiquetage
 - Les termes sont affectés à une catégorie grammaticale ou à un type d'entité (adresse, date, lieu, compagnie...)

Processus de Text Mining

- Permet d'éliminer des termes en fonction de leur catégories ou d'en sur pondérer d'autres
- Permet de gérer les ambiguïtés induites par différents contextes.
 - Ex : aller au marché / faire un marché
- SAS TM permet de spécifier une stop-list ou in-list :
 - Stop list = Liste de mots à conserver (par ex taxinomie d'un thème particulier : pharma, industrie, économie ...)
 - in list = liste de mots à exclure (articles, prépositions, conjonctions)
- Des listes sont fournies dans SAS TM mais leur personnalisation enrichit énormément la richesse de l'analyse

Processus de Text Mining

- Réduction automatique des termes à leur racine
 - Réduit le nombre de termes à traiter
 - La contribution à l'analyse de la forme réduite sera plus forte que celle des multiples formes non réduites.
- Traitement des synonymes
 - SAS TM permet de créer une liste de synonymes dans laquelle seront stockés l'ensembles des groupes de termes déclarés comme équivalents
 - Ex : enseigner, instruire et éduquer

Processus de Text Mining

- Construction d'une représentation quantitative de l'ensemble des documents traités
 - A la fin du traitement, on obtient une table de fréquence qui sera la source des analyses ultérieures
 - En ligne les documents traités
 - En colonne les termes rencontrés

Processus de Text Mining

- Réduction des dimensions de la table de fréquence
 - Objectif = rendre l'information exploitable en déterminant les termes les plus significatifs
 - 2 méthodes :
 - Décomposition en valeur singulière (SVD) = projection des termes et des documents dans un même plan et analyse des proximités
 - Roll-up terms = conserver les N termes ayant le poids le plus élevée

Text Miner - Interactif

Fichier Edition Outils Affichage Fenêtre



Documents

SUMMARY	NHTSA T...	RECORD ...	MANUFA...	VEHICLE/...	VEHICLE/...	WAS VE...	DATE OF...	WAS VE...	NUMBER ...	NUMBER ...	SPECIFIC ...	CONSUM...
TRANSMISSION COOLING UNIT AND LINES LEAK, CAUSING VEHICLE FIRE. * AK	520896	1.0	DAIMLERC...	CHRYSLER...	TOWN AN...	N	1997-05-22	Y	0.0	0.0	POWER TR...	WA
BRAKES: THE BRAKES WERE INOPERATIVE. TOOK VEHICLE TOOK DEALER AND WAS TOL	813841	7.0	GENERAL ...	CHEVROLE...	BLAZER	N	1997-07-02	N	0.0	0.0	BRAKES:H...	WA
TRANSMISSION SHUDDERS WHEN STARTED UP FROM A STOPPED POSITION AFTER THE VEH	513883	9.0	DAIMLERC...	PLYMOUT...	VOYAGER	N	1996-12-19	N	0.0	0.0	POWER TR...	WA
FRONT SUSPENSION BROKE, LOST CONTROL OF VEHICLE THEN IT FLIPPED OVER. * AK C	804195	22.0	MAZDA (N...	MAZDA TR...	PICKUP	Y	1995-09-24	N	1.0	0.0	SUSPENSI...	NJ
ABS, APPLIED BRAKES, EXPERIENCED EXTENED PEDAL FADED, IMPACT 12:00 POSITION, S	952352	123.0	FORD MOT...	FORD TRU...	AEROSTAR	Y	1994-06-30	N	1.0	0.0	BRAKES:H...	TX
PCV VALVE BLEW OFF HOSE, CAUSING HIGH VEHICLE IDLE, APPROX. 3000 RPMs. TT	947918	158.0	GENERAL ...	CHEVROLE...	S10	Y	1994-04-15	N	1.0	0.0	FUEL:FUEL...	MS
HILE DRIVING VAN PULLED LEFT, ENDED UP IN CENTER DIVIDER, and FLIPPED OVER. LW	386832	259.0	MITSUBISHI...	MITSUBISHI...	VANWAG...	Y	1991-07-08	N	1.0	0.0	SUSPENSION	CA
UPON IMPACT AT 35MPH, 2ND ROW BUCKET SEAT BELT ANCHOR DETACHED FROM FLOOR,	956391	331.0	DAIMLERC...	DODGE TR...	CARAVAN	Y	1994-11-04	N	1.0	0.0	INTERIOR S...	CA
TIE ROD BROKE RESULTING IN ACCIDENT, TRUCK ROLLED OVER TWICE. BI	356119	360.0	DAIMLERC...	DODGE TR...	DAKOTA	Y	1989-08-08	N	0.0	0.0	STEERING:...	NC
WHILE PARKED IN DRIVEWAY, VAN DRIFTED BACKWARDS HITTING FENCE; SLIPPED TRANS	380196	367.0	MAZDA (N...	MAZDA TR...	MPV	Y	1991-01-31	N	1.0	0.0	POWER TR...	NY
VEHICLE CROSSED CENTERLINE CAUSING ACCIDENT. * LDG	410205	411.0	DAIMLERC...	PLYMOUT...	VOYAGER	Y	1992-01-18	N	3.0	1.0	STEERING:...	PA
FRONTAL COLLISION, 11:00 POSITION, VEHICLE WAS AT COMPLETE STOP, AIR BAG DID N	945207	442.0	DAIMLERC...	PLYMOUT...	VOYAGER	Y	1993-09-21	N	2.0	0.0	INTERIOR S...	WA
HIT POTHOLE IN DRIVEWAY AND STEERING BOX GAVE AND TRUCK HIT POLE. EC	H91035	523.0	TOYOTA M...	TOYOTA T...	PICKUP	Y	1987-01-10	N	1.0	0.0	STEERING:...	NM
HIT FROM THE LEFT HAND SIDE LATCH OPENED AND 3 PASSENGERS WERE THROWN OUT O	965420	550.0	DAIMLERC...	DODGE TR...	CARAVAN	Y	1994-12-25	N	1.0	0.0	STRUCTUR...	CA
AUTOMATIC TRANSMISSION FAILURE. TT	963782	558.0	DAIMLERC...	PLYMOUT...	VOYAGER	Y	1995-04-01	N	0.0	1.0	POWER TR...	CO
BRAKES DO NOT WORK PROPERLY ABOVE 40 MPH, MUST PUMP BRAKES TO STOP COMPLE	H90778	803.0	GENERAL ...	CHEVROLE...	TRUCK AN...	Y	1985-11-01	N	0.0	0.0	BRAKES:H...	NJ
SUDDEN BRAKING CAUSED VEHICLE TO SWERVE, STRAIGHT CONTROLLED STOP DIFFICULT	304646	870.0	DAIMLERC...	DODGE TR...	CARAVAN	Y	1987-11-03	N	1.0	0.0	BRAKES:H...	NY
IN STOP POSITION WHEN REAR ENDED BOTH FRONT SEAT BACKS COLLASPED AND MIDDLE	944860	959.0	GENERAL ...	CHEVROLE...	ASTRO	Y	1994-01-21	N	6.0	0.0	INTERIOR S...	MO
BRAKES LOCKED UP ON DRY ROADS. AK	936285	1113.0	MAZDA (N...	MAZDA TR...	MPV	Y	1993-02-10	N	0.0	0.0	BRAKES:H...	CA
EMERGENCY BRAKE LEVER LOOSEMED, CAUSING THE LOCK UP OF THE REAR BRAKES AN	912252	1131.0	FORD MOT...	FORD TRU...	AEROSTAR	Y	1991-12-09	N	2.0	0.0	BRAKES:H...	FL

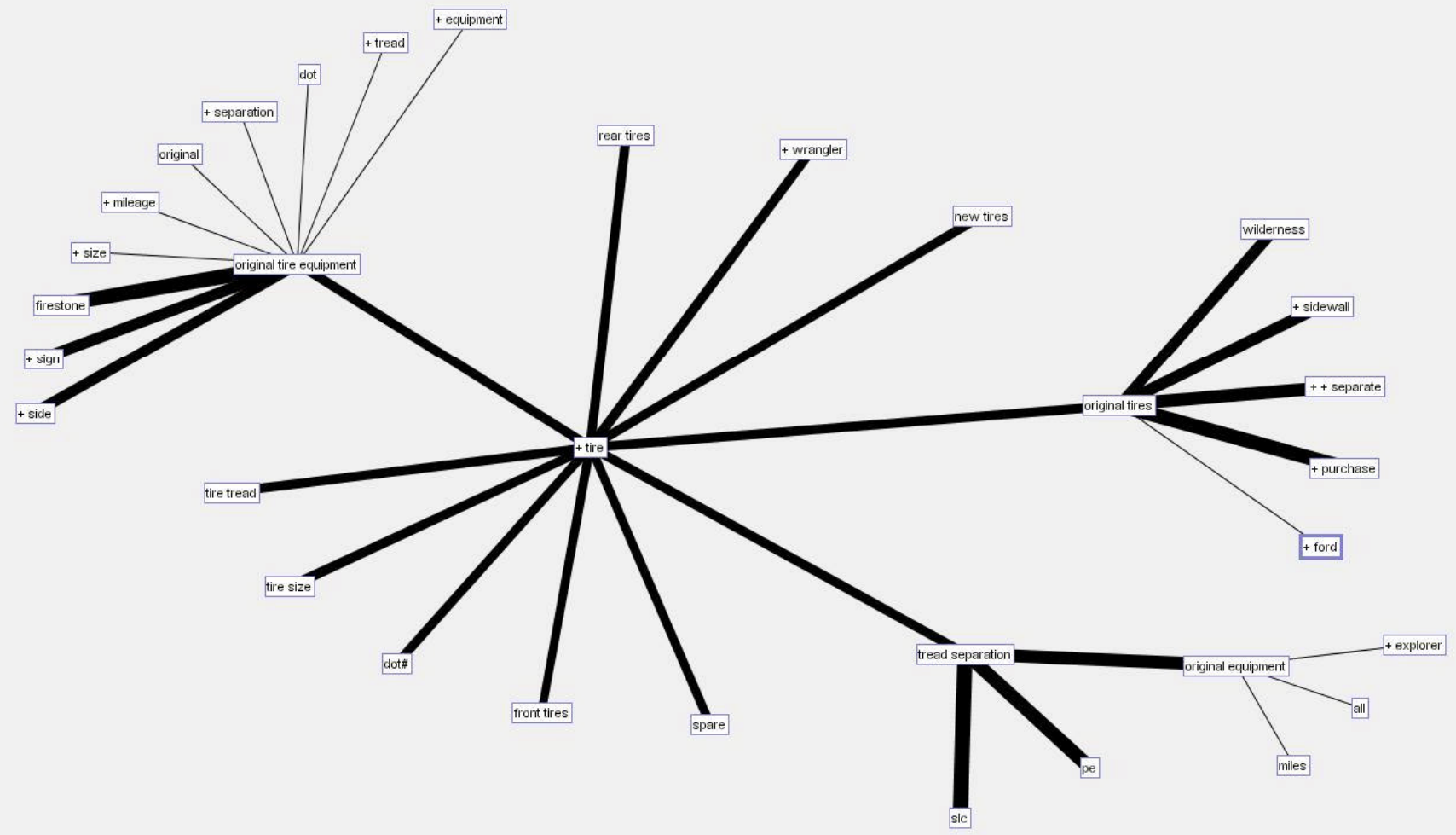
Termes

TERM	Freq	# Documents	Keep	WEIGHT	Role	Attribute
not	10320	7801	<input checked="" type="checkbox"/>	0.116	Part	Alpha
dealer	9854	7756	<input checked="" type="checkbox"/>	0.118	Noun	Alpha
in	10175	7666	<input checked="" type="checkbox"/>	0.119	Prep	Alpha
on	9473	7037	<input checked="" type="checkbox"/>	0.128	Prep	Alpha
consumer	8710	5925	<input checked="" type="checkbox"/>	0.146	Noun	Alpha
brake	9491	5709	<input checked="" type="checkbox"/>	0.153	Noun	Alpha
when	6031	5270	<input checked="" type="checkbox"/>	0.151	Conj	Alpha
cause	5548	5042	<input checked="" type="checkbox"/>	0.153	Verb	Alpha
problem	6465	4963	<input checked="" type="checkbox"/>	0.161	Noun	Alpha
drive	3735	3587	<input checked="" type="checkbox"/>	0.185	Verb	Alpha
do	3971	3409	<input checked="" type="checkbox"/>	0.195	Aux	Alpha
replace	4552	3325	<input checked="" type="checkbox"/>	0.207	Verb	Alpha
driver	3836	3176	<input checked="" type="checkbox"/>	0.203	Noun	Alpha
with	3116	2662	<input checked="" type="checkbox"/>	0.22	Prep	Alpha
mph	2678	2540	<input checked="" type="checkbox"/>	0.22	Prop	Alpha
time	2962	2482	<input checked="" type="checkbox"/>	0.227	Noun	Alpha
far	2291	2281	<input checked="" type="checkbox"/>	0.229	Adj	Alpha
accident	2445	2268	<input checked="" type="checkbox"/>	0.233	Noun	Alpha
rear	2652	2237	<input checked="" type="checkbox"/>	0.238	Adj	Alpha
will	2716	2235	<input checked="" type="checkbox"/>	0.238	Aux	Alpha
provide	2248	2235	<input checked="" type="checkbox"/>	0.231	Verb	Alpha

Classes

#	Descriptive Terms	Freq	Percentage	RMS Std.
1	+ accident, + park, + roll, + result, + gear	2053	0.090680212...	0.1039998...
2	+ far, + problem, + replace, + dealer, + provide	7101	0.313648409...	0.1110551...
3	mph, + not, + driver, + deploy, + impact	1796	0.079328621...	0.0886393...
4	+ apply, + brake, when, + pedal, + problem	3777	0.166828621...	0.1010373...
5	+ separate, + size, firestone, + tread, rear	1209	0.053401060...	0.0780535...
6	+ driver, + seat, + side, + break, + do	1341	0.059231448...	0.0878791...
7	+ wheel, + drive, + dealer, + problem, on	5363	0.236881625...	0.1108531...

Concept Linking



Enterprise Miner - Miner

Fichier Edition Affichage Actions Options Fenêtre Aide

Echantillon Exploration Modification Modélisation Evaluation Utilitaires Credit Scoring

Table des matières

Miner

- Sources de données
- Diagrammes
 - HCM
 - HMEQ
 - MM
 - NHTSA_Truck_C
- Package de modèles
- Utilisateurs

Propriétés Valeur

ID	EMVWS4
Nom	NHTSA_Truck_C
Etat	Ouvert
Notes	
Historique	

ID

Identifiant du diagramme, correspondant au

TRUCKS

Data Partition

Arbre sans données textes

Régression Robuste sans données textes

Raisonnement à base de cas

Régression Robuste

Réseau de neurones

Arbre

Arbre données textuelles seulement

Text Miner

Comparaison de modèles

Exécution terminée.

Connecté à SASMain - Logical Workspace Server

Results - Arbre

File Edit View Window



Tree

Statistic	Training	Validation
Y:	30.1%	30.1%
N:	69.9%	69.9%
N in Node:	22640	16979

_SVD_3

Statistic	Y:	N:	N in Node:

495

Statistic	Training	Validation
Y:	70.1%	70.1%
N:	29.9%	29.9%
N in Node:	2867	

16

Statistic	Training	Validation
Y:	82.0%	80.2%
N:	17.0%	19.7%
N in Node:	2478	1938

Statistic	Training	Validation
Y:	39.3%	40.1%
N:	60.7%	59.9%

Statistic	Training	Validation
Y:	58.0%	59.9%
N:	42.0%	40.1%
N in Node:	690	495

NUMBER OF OCCURRENCES

≥ 0.5

Statistic	Training	Validation
Y:	70.4%	74.2%
N:	21.6%	25.8%
N in Node:	458	332

< 0.5

Statistic	Training	Validation
Y:	17.7%	12.3%
N:	82.3%	87.7%
N in Node:	232	162

_SVD_11

< 0.04937

Statistic	Training	Validation
Y:	45.1%	44.0%
N:	54.9%	56.0%
N in Node:	113	91

_SVD_3

≥ 0.04937

Statistic	Training	Validation
Y:	89.2%	85.5%
N:	10.7%	14.5%
N in Node:	345	246

_SVD_4

≥ 0.03085

Statistic	Training	Validation
Y:	29.1%	34.4%
N:	70.9%	65.6%

< 0.03085

Statistic	Training	Validation
Y:	82.4%	66.7%
N:	17.6%	33.3%

< -0.0473

Statistic	Training	Validation
Y:	67.7%	62.7%
N:	32.3%	37.3%

≥ -0.0473

Statistic	Training	Validation
Y:	97.2%	94.3%
N:	2.8%	5.7%

Statistic	Y:	N:	N in Node:

Résultats - Comparaison de modèles

Fichier Edition Affichage Fenêtre

Tests d'ajustement

Selected Mo...	Predecess...	Valid: Misclassification Rate ▼
	DmineReg2	0.229048
	Tree	0.187172
	MBR	0.180812
	Tree3	0.12592
	Tree2	0.113847
	Neural	0.101243
Y	DmineReg	0.083397

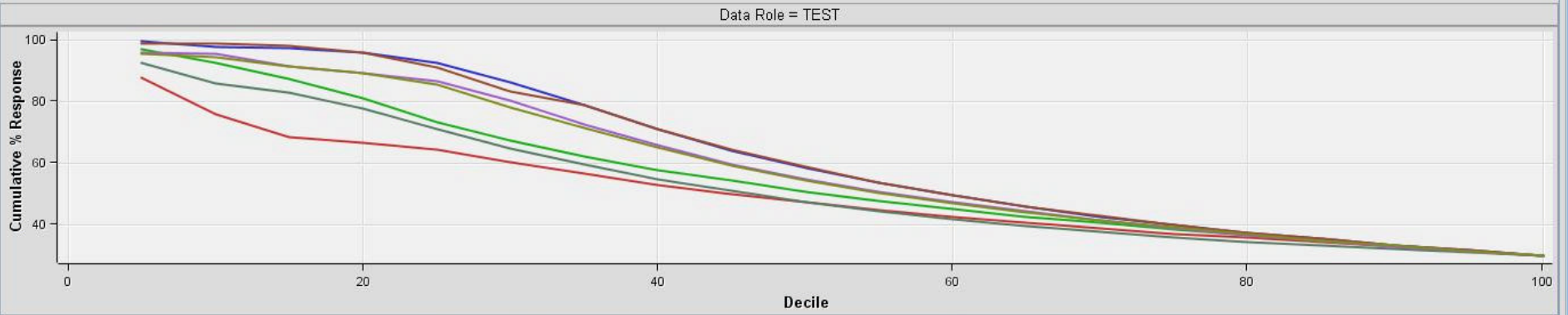
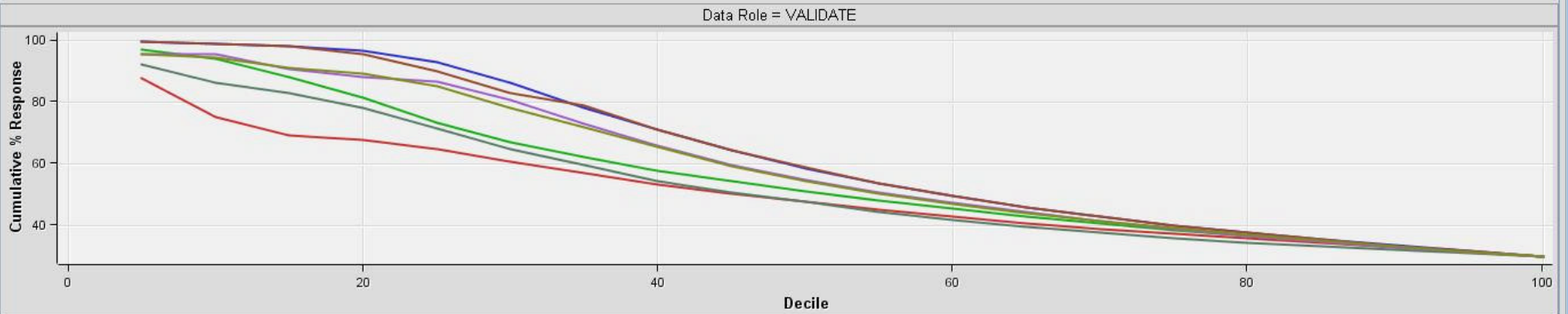
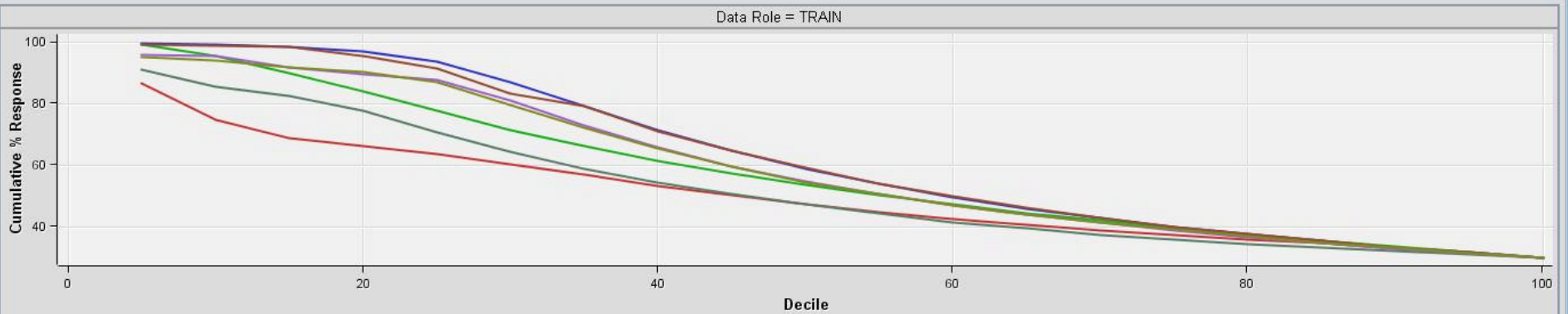
Résultats - Comparaison de modèles

Fichier Edition Affichage Fenêtre



Recouvrement des classements des scores : CRASH

Sélectionner un graphique : Cumulative % Response



— DmineReg — DmineReg2 — MBR — Neural — Tree — Tree2 — Tree3



THE
POWER
TO KNOW®



SAS France
Domaine de Grégy - BP 5
77166 Grégy-sur-Yerres
Tél : 01 60 62 11 11
Fax : 01 60 62 11 99

SAS Europe, Middle East & Africa
P.O. Box 10 53 40
Neuenheimer Landstr. 28-30
D-69043 Heidelberg, Germany
Tel: +49 6221 4160, Fax: +49 6221 474850

SAS, le Système SAS® sont les marques déposées de SAS Institute Inc., Cary, NC, USA.
Les autres noms de produits ou concepts sont des marques déposées des sociétés respectives.

www.sas.com/france