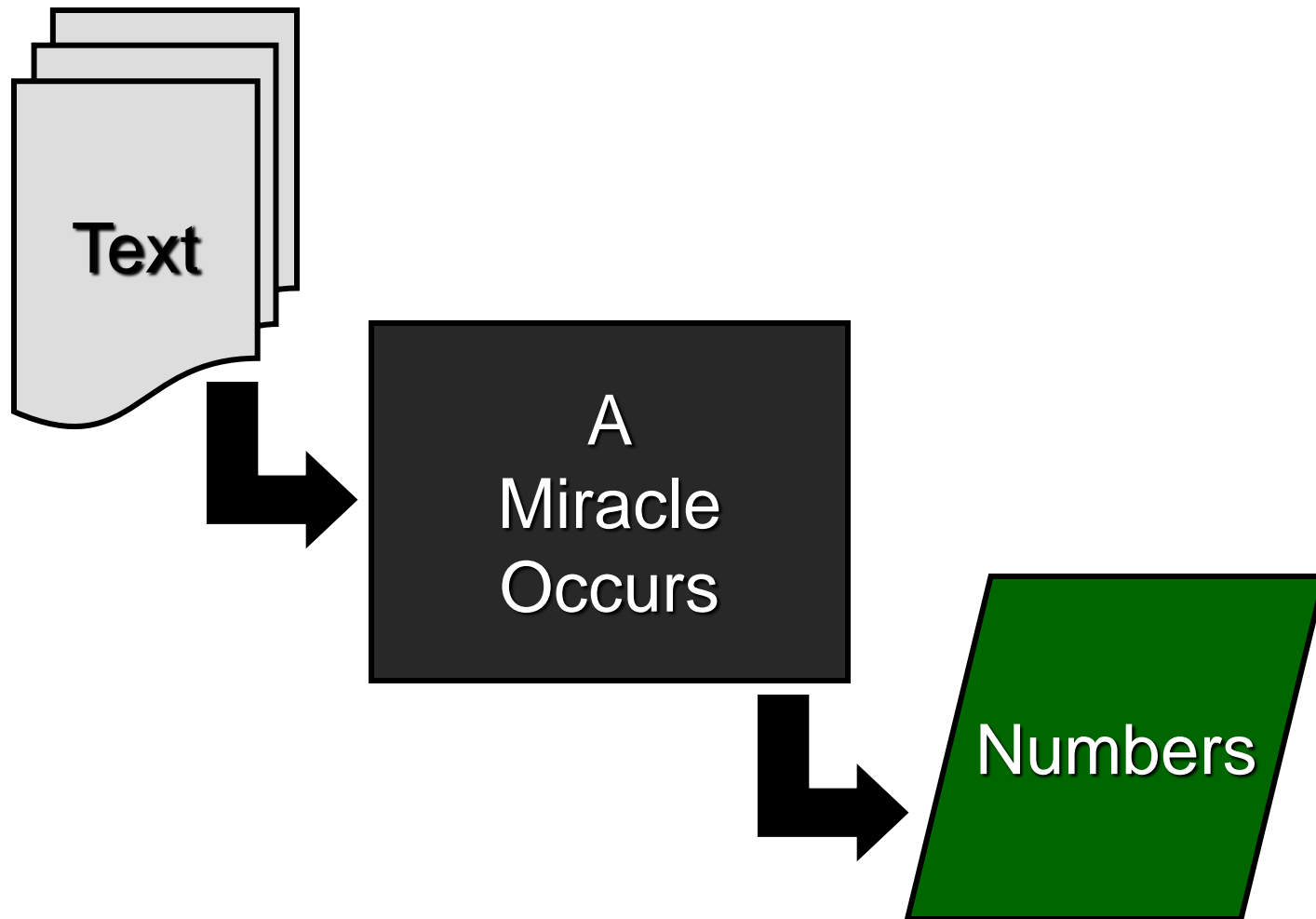


SAS Text Miner: From Soup to Nuts



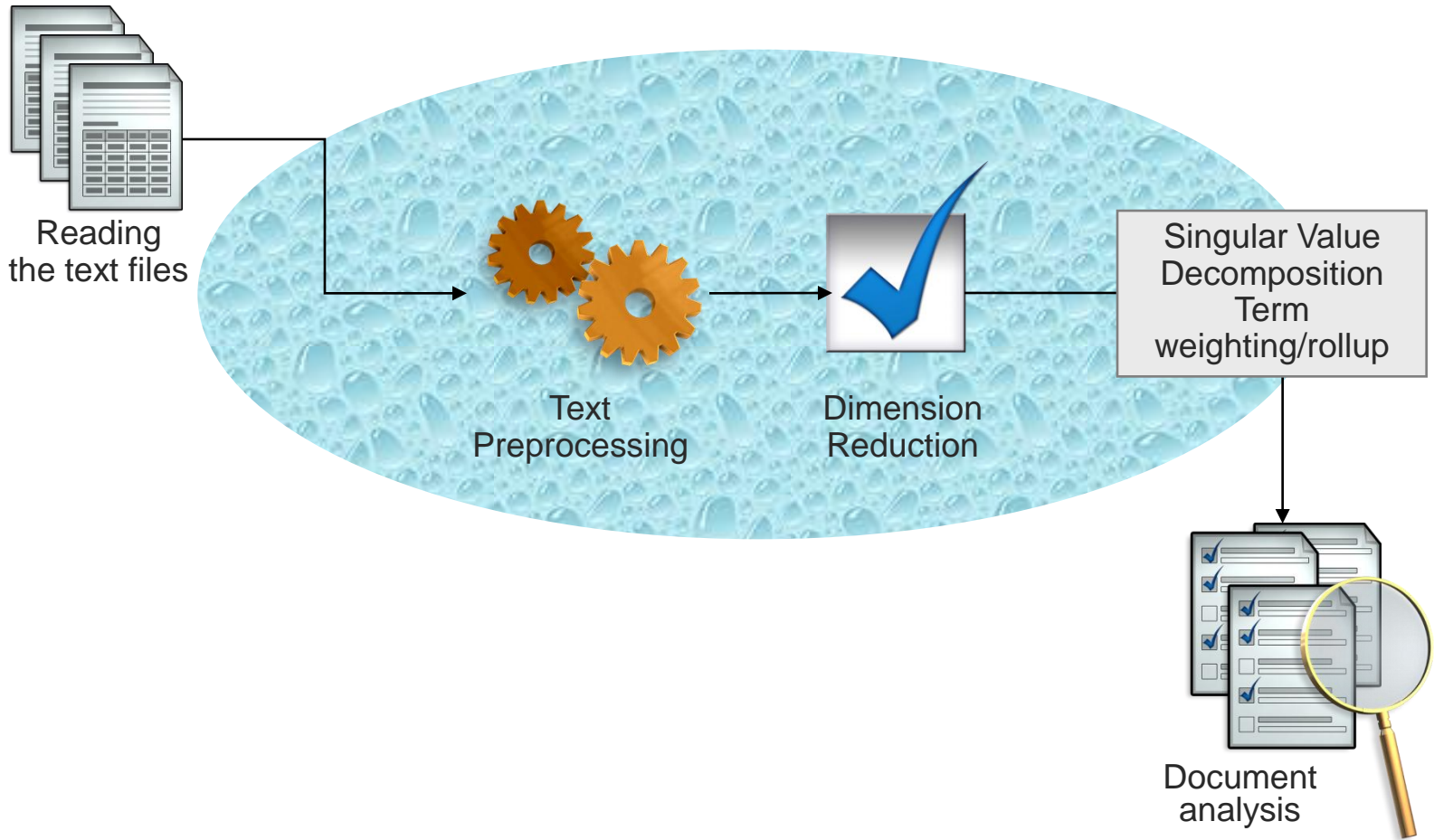
THE
POWER
TO KNOW®

Text Mining



Text Mining Process

S
E
M
M
A



Situation

- You are a member of the Special Investigative Unit (SIU) of an insurance company
- You are tasked with determining the Subrogation (recovery) potential of each claim that comes in.
- You have a target field (SubroFlag) that indicates whether the claim leads to any eventual recovery.
- You have the raw claim adjuster notes.
- You have indicator flag variables normally used to predict subrogation. (age, body parts affected, nature of the accident, occupation code, back injury flag, and vehicle flag).

Step 1

- Explore the information in the claims adjuster notes.
- Management is specifically interested in head and neck injuries, and wants to know what are the major causes of neck injuries.

Example: Insurance 1

Interactive Topic Viewer

File Edit View

Search :

Concept

- +door,+walk,
- +vehicle accident
- +floor,+slip, ice
- struck, +strike
- neck,vehicle,head

Terms

driving

Topic Weight
1.540
1.320
1.314
1.288
1.000
0.500
0.302
0.297

vehicle

Documents

Topic Weight
0.97734163403338
0.97431881845973
0.81085996080732
0.69409968409640
0.69404885480187
0.32001930652275
0.24861686048410
0.22112556649293

Explore - EMWS1.TextTopic2_TRAIN

File Edit View Actions Window

Graph4

SubroFlag	struck, +strike (Sum)
0	8.5
1	8.0

Graph2

SubroFlag	+vehicle accident,+involve,+back,pain...
0	0
1	35

Graph3

SubroFlag	+floor,+slip, ice (Sum)
0	6.0
1	8.0

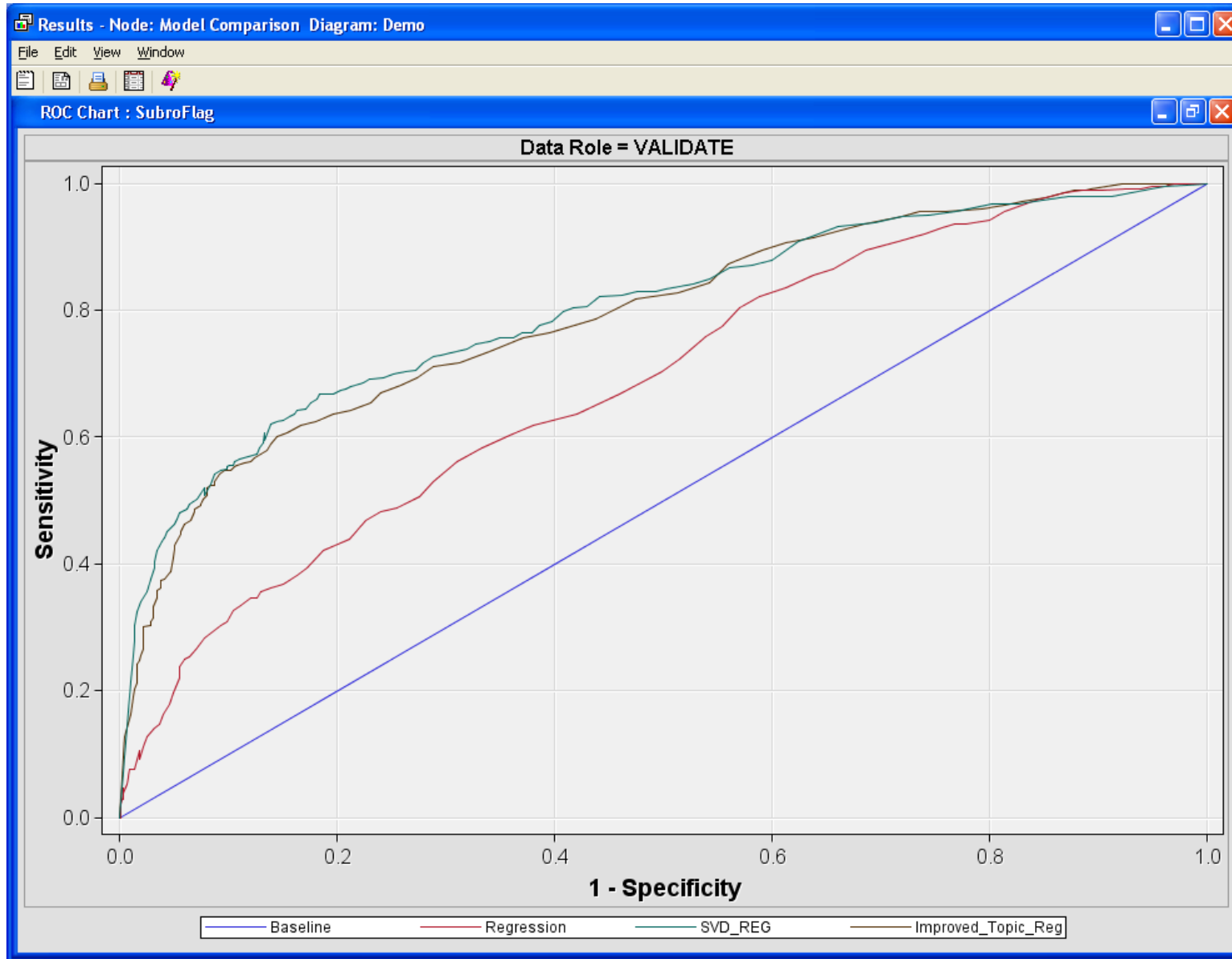
Graph1

SubroFlag	+door,+walk, (Sum)
0	7.0
1	3.0

Step 2

- Compare the use of adjustor notes in the prediction of successful subrogation to the existing regression model

Example: Insurance 2



Questions

tim.trussell@sas.com