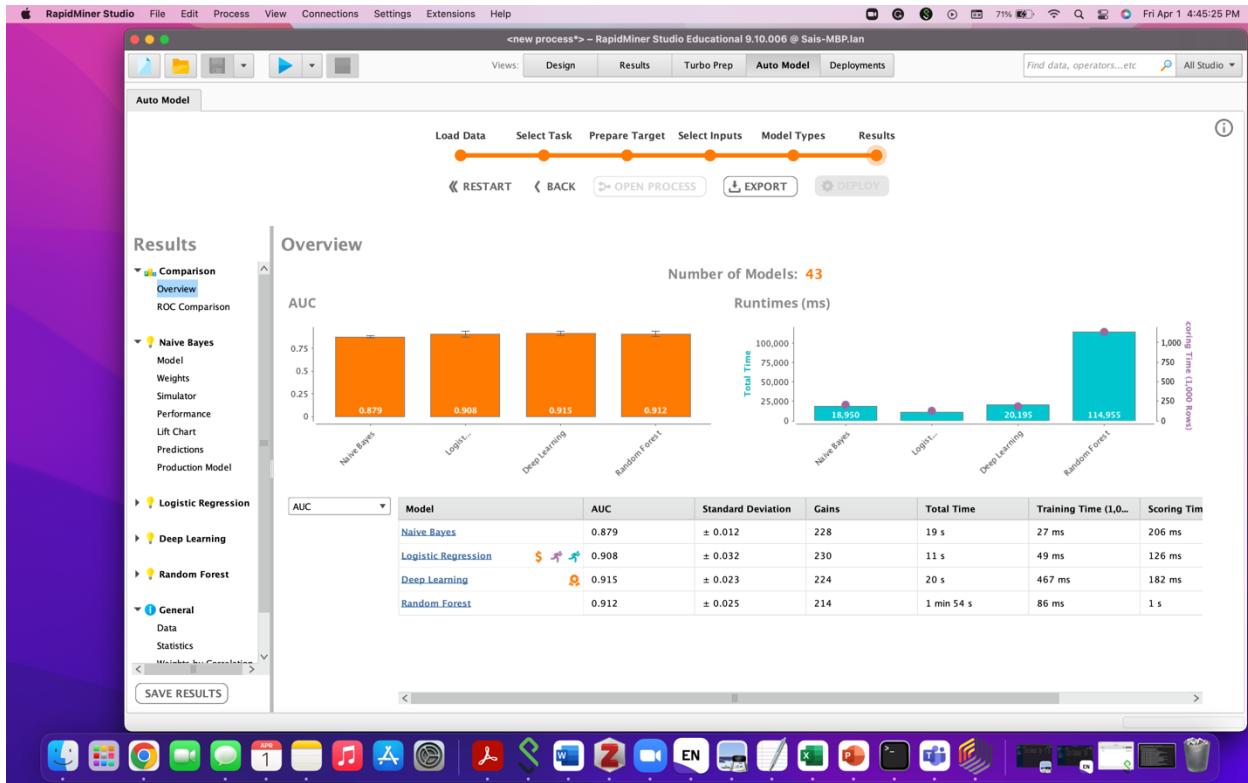


MACHINE LEARNING



AUC

Naive Bayes = 0.879

Logistic regression = 0.908

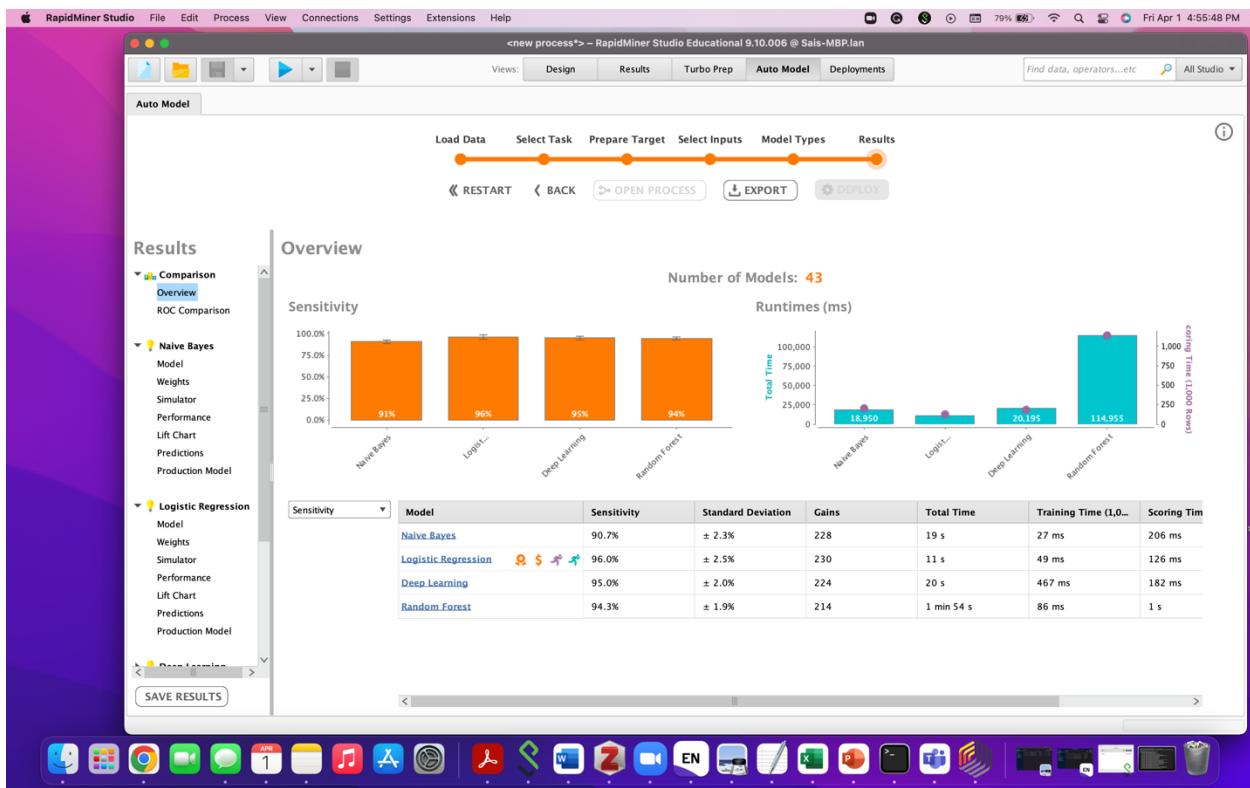
Deep learning = 0.915

Random forest = 0.912

Highest AUC – Deep learning algorithm (0.915)

Fastest results- Logistic regression (11s)

Slowest – Random Forest (1 min 54s)



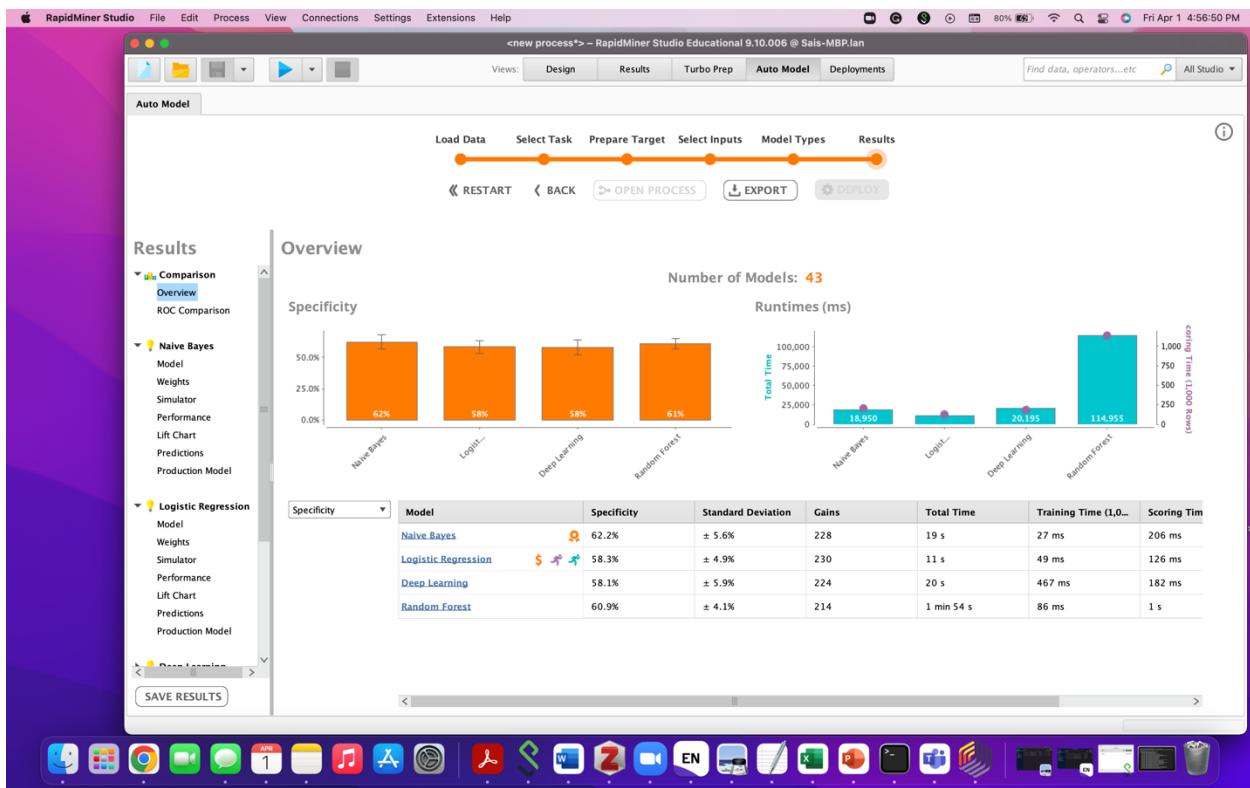
Sensitivity:

Highest sensitivity is for Logistic Regression (96%)

Lowest sensitivity is for Naive Bayes (90.7%).

In sensitivity fastest results are obtained for Logistic Regression(11s)

Slowest are for Random Forest (1min 54s)



Specificity:

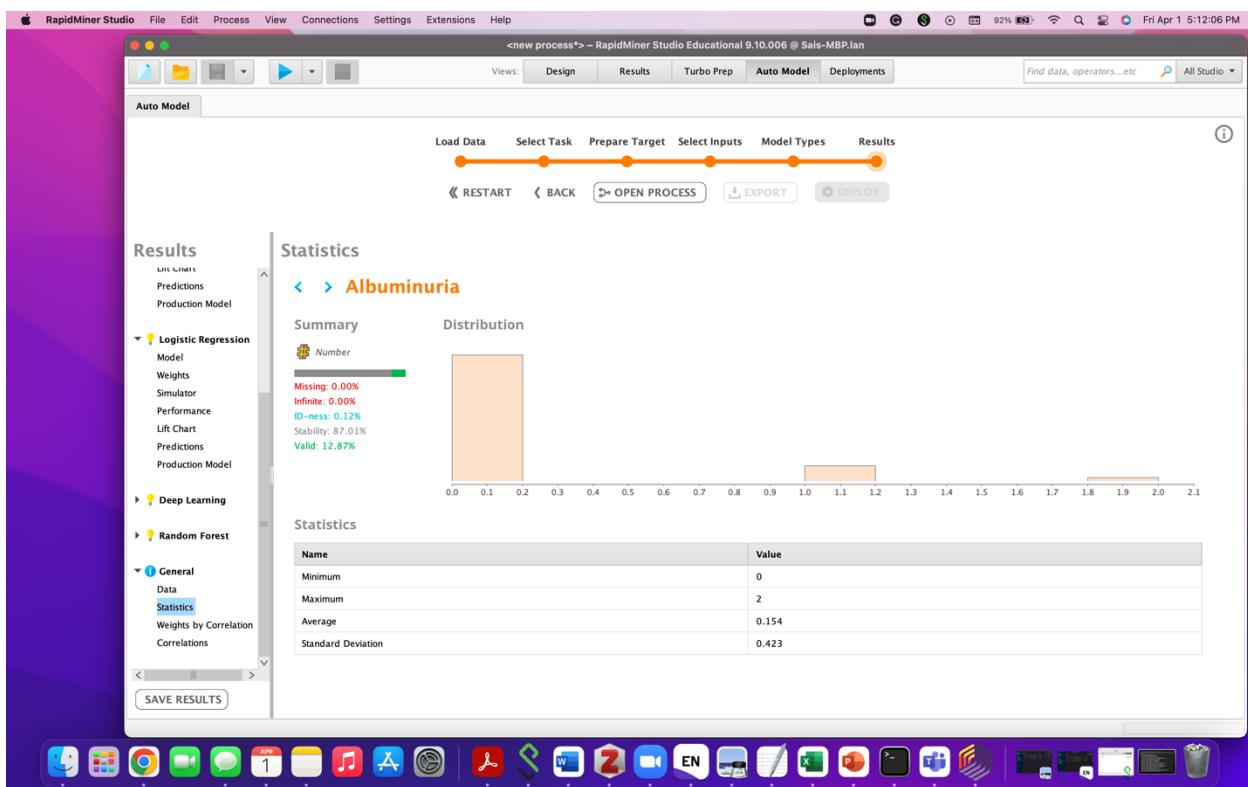
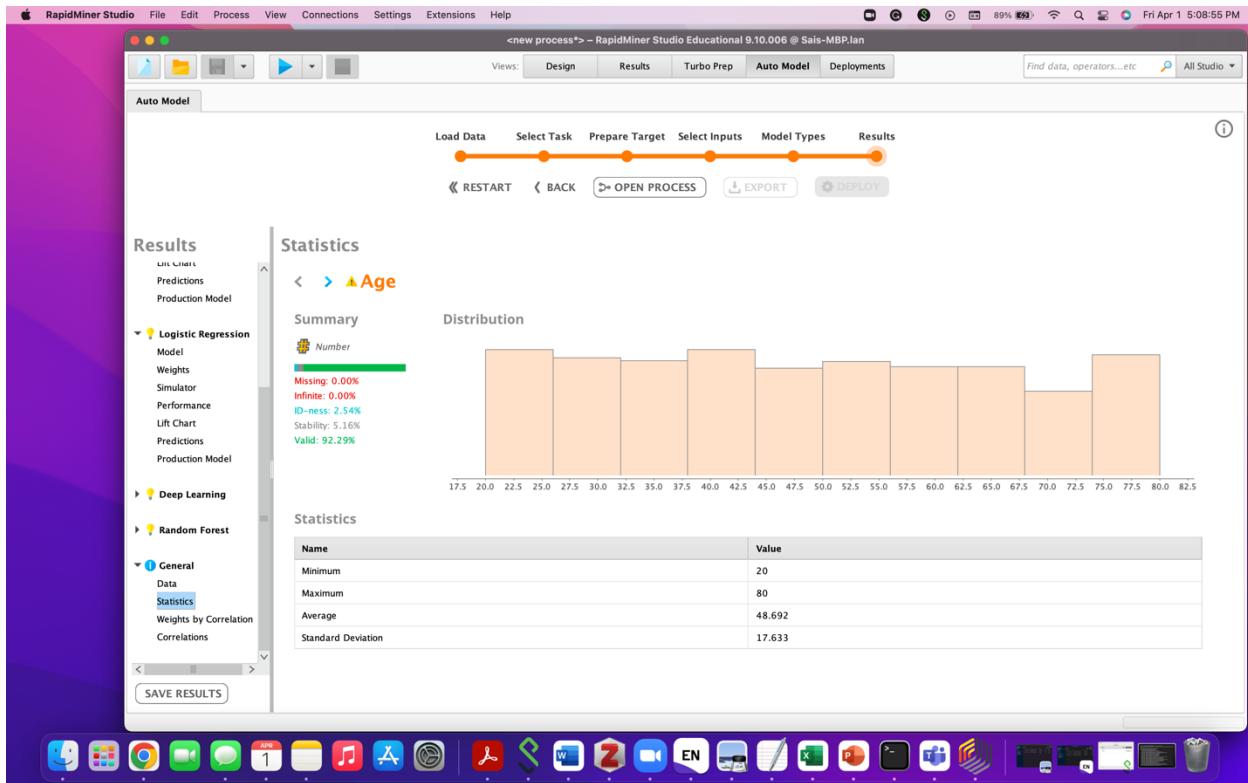
Highest specificity: Naïve Bayes (62.2%)

Lowest specificity: Deep Learning (58.1%)

Fastest results: Logistic regression (11s)

Slowest results: Random Forest (1min 54s)

STATISTICS



RapidMiner Studio 9.0.006 @ Sais-MBP.lan

Fri Apr 1 5:12:36 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics**
- Weights by Correlation
- Correlations

Statistics

BMI

Summary

Number

Missing: 1.08% Infinite: 0.00% ID-ness: 1.37% Stability: 1.01% Valid: 96.53%

Distribution

Statistics

Name	Value
Minimum	13.400
Maximum	68.700
Average	28.702
Standard Deviation	6.662

SAVE RESULTS

RapidMiner Studio 9.0.006 @ Sais-MBP.lan

Fri Apr 1 5:12:39 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics**
- Weights by Correlation
- Correlations

Statistics

BloodGlucose

Summary

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 7.79% Stability: 3.75% Valid: 88.46%

Distribution

Statistics

Name	Value
Minimum	39
Maximum	382
Average	108.248
Standard Deviation	34.821

RapidMiner Studio

<new process> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D+ OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics**
- Weights by Correlation
- Correlations

SAVE RESULTS

Statistics

HDL

Summary

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 3.87% Stability: 3.62% Valid: 92.50%

Distribution

Statistics

Name	Value
Minimum	14
Maximum	156
Average	53.369
Standard Deviation	15.186

Fri Apr 1 5:12:41 PM

RapidMiner Studio

<new process> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D+ OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics**
- Weights by Correlation
- Correlations

SAVE RESULTS

Statistics

Income

Summary

Number

Missing: 4.87% Infinite: 0.00% ID-ness: 0.58% Stability: 14.75% Valid: 79.79%

Distribution

Statistics

Name	Value
Minimum	300
Maximum	9000
Average	4005.254
Standard Deviation	2954.032

Fri Apr 1 5:12:43 PM

RapidMiner Studio 9.0.006 @ Sais-MBP.lan Fri Apr 1 5:12:48 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics**
- Weights by Correlation
- Correlations

SAVE RESULTS

Statistics

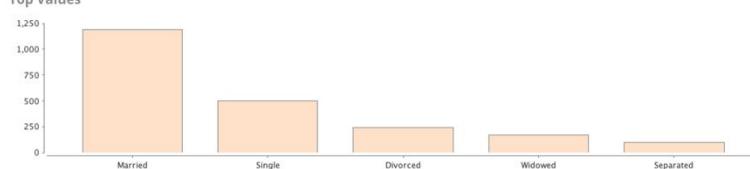
Marital

Summary

Category: Marital

Missing: 8.66% Infinite: 0.00% ID-ness: 0.21% Stability: 54.35% Valid: 36.77%

Top Values



Value	Count	Percentage
Married	1,192	54.35%
Single	498	22.71%
Divorced	242	11.04%
Widowed	166	7.57%
Separated	95	4.33%

5 Distinct Values:

Value	Count	Percentage
Married	1,192	54.35%
Single	498	22.71%
Divorced	242	11.04%
Widowed	166	7.57%
Separated	95	4.33%

SAVE RESULTS

RapidMiner Studio 9.0.006 @ Sais-MBP.lan Fri Apr 1 5:12:50 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics**
- Weights by Correlation
- Correlations

SAVE RESULTS

Statistics

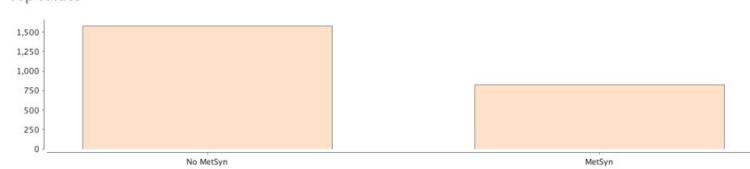
MetabolicSyndrome

Summary

Category: MetabolicSyndrome

Missing: 0.00% Infinite: 0.00% ID-ness: 0.08% Stability: 65.76% Valid: 34.15%

Top Values



Value	Count	Percentage
No MetSyn	1,579	65.76%
MetSyn	822	34.24%

2 Distinct Values:

Value	Count	Percentage
No MetSyn	1,579	65.76%
MetSyn	822	34.24%

SAVE RESULTS

RapidMiner Studio 9.10.006 @ Sais-MBP:lan Fri Apr 1 5:12:55 PM

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

RESTART BACK OPEN PROCESS EXPORT DEPLOY

Results

Logistic Regression Deep Learning Random Forest General Data Statistics Weights by Correlation Correlations

Statistics

Race

Summary

Category: White, Missing: 0.00%, Infinite: 0.00%, ID-ness: 0.25%, Stability: 38.86%, Valid: 60.89%

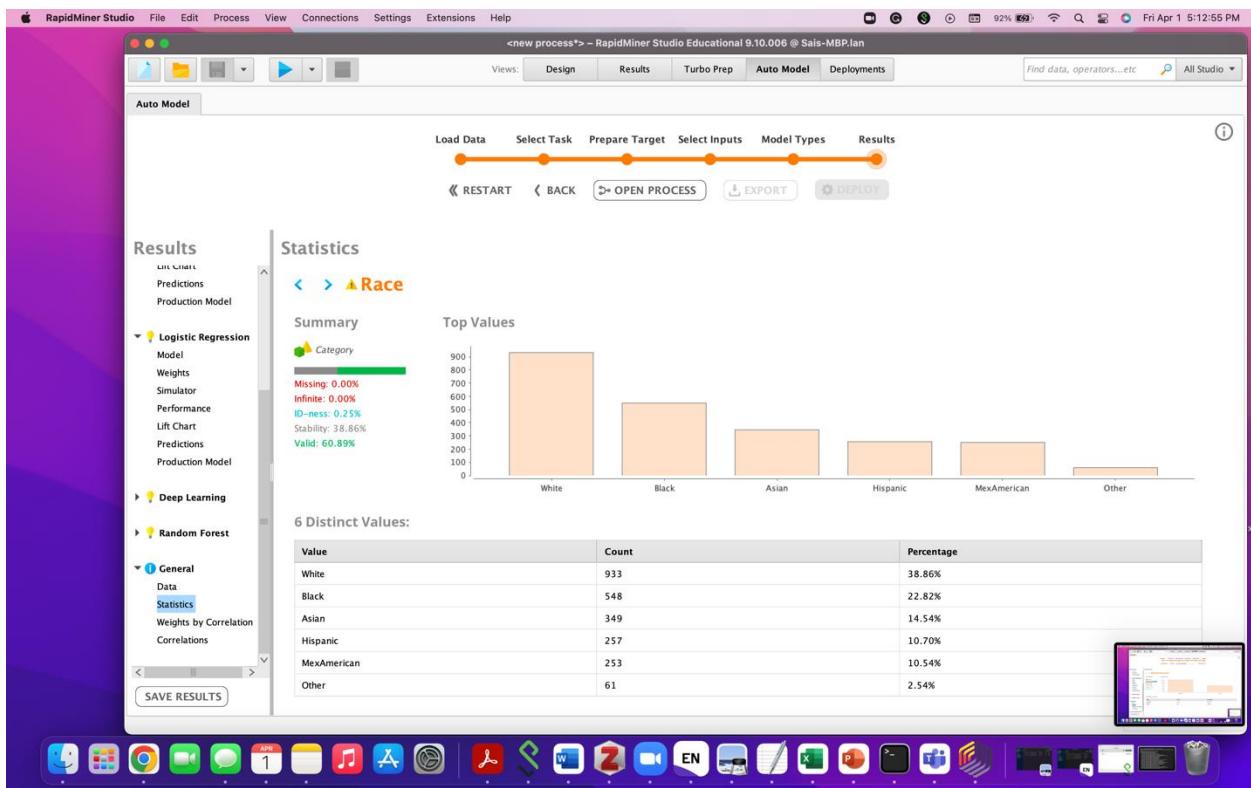
Top Values

Category	Count	Percentage
White	933	38.86%
Black	548	22.82%
Asian	349	14.54%
Hispanic	257	10.70%
MexAmerican	253	10.54%
Other	61	2.54%

6 Distinct Values:

Value	Count	Percentage
White	933	38.86%
Black	548	22.82%
Asian	349	14.54%
Hispanic	257	10.70%
MexAmerican	253	10.54%
Other	61	2.54%

SAVE RESULTS



RapidMiner Studio 9.10.006 @ Sais-MBP:lan Fri Apr 1 5:12:57 PM

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

RESTART BACK OPEN PROCESS EXPORT DEPLOY

Results

Logistic Regression Deep Learning Random Forest General Data Statistics Weights by Correlation Correlations

Statistics

Sex

Summary

Category: Female, Missing: 0.00%, Infinite: 0.00%, ID-ness: 0.08%, Stability: 50.44%, Valid: 49.48%

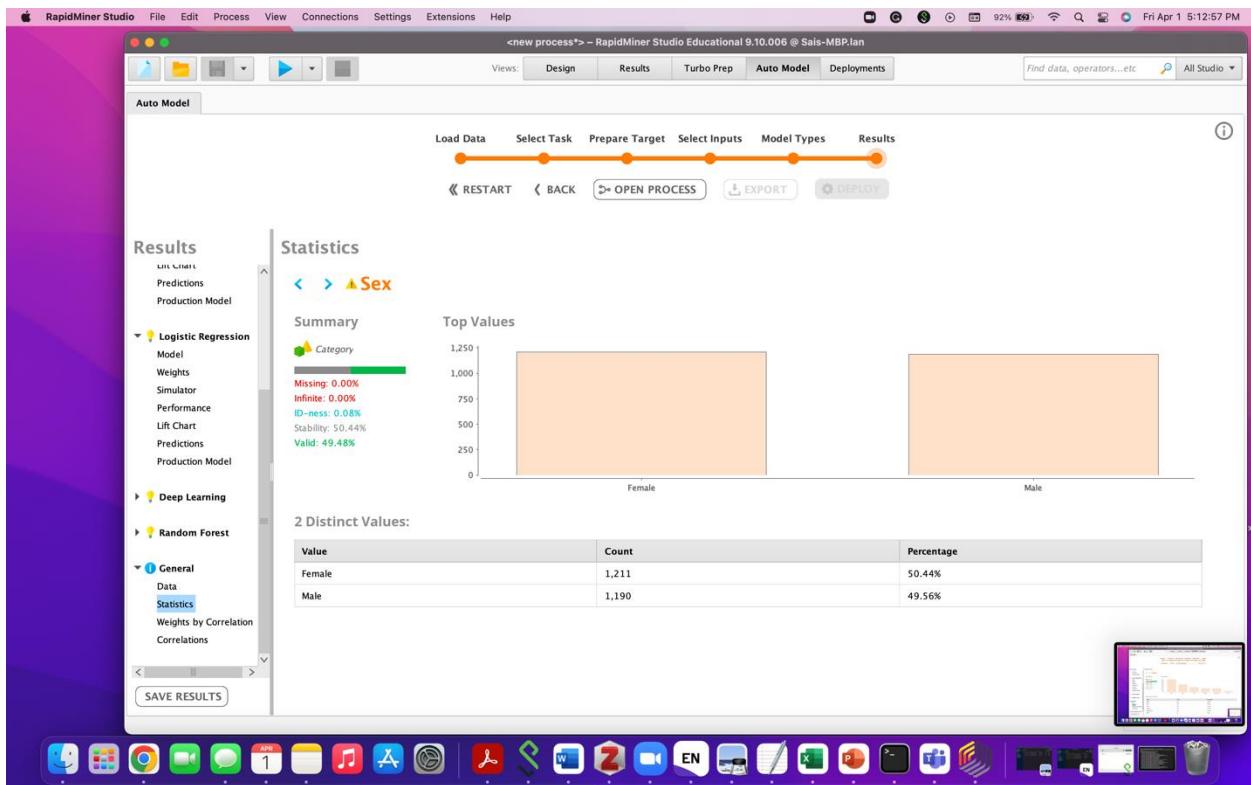
Top Values

Category	Count	Percentage
Female	1,211	50.44%
Male	1,190	49.56%

2 Distinct Values:

Value	Count	Percentage
Female	1,211	50.44%
Male	1,190	49.56%

SAVE RESULTS



RapidMiner Studio

File Edit Process View Connections Settings Extensions Help

<new process*> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

Logistic Regression Deep Learning Random Forest General

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 14.37% Stability: 1.42% Valid: 84.21%

Distribution

Statistics

Triglycerides

Summary

Name	Value
Minimum	26
Maximum	1562
Average	128.125
Standard Deviation	95.322

SAVE RESULTS

RapidMiner Studio

File Edit Process View Connections Settings Extensions Help

<new process*> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

Logistic Regression Deep Learning Random Forest General

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 1.33% Stability: 0.75% Valid: 97.92%

Distribution

Statistics

UrAlbCr

Summary

Name	Value
Minimum	1.400
Maximum	5928
Average	43.626
Standard Deviation	258.273

SAVE RESULTS

RapidMiner Studio

File Edit Process View Connections Settings Extensions Help

<new process*> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

Logistic Regression Deep Learning Random Forest General

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 1.33% Stability: 0.75% Valid: 97.92%

Distribution

Statistics

UrAlbCr

Summary

Name	Value
Minimum	1.400
Maximum	5928
Average	43.626
Standard Deviation	258.273

SAVE RESULTS

RapidMiner Studio File Edit Process View Connections Settings Extensions Help Fri Apr 1 5:13:04 PM

<new process*> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

UricAcid Predictions Production Model

Logistic Regression Model Weights Simulator Performance Lift Chart Predictions Production Model

Deep Learning

Random Forest

General Data Statistics Weights by Correlation Correlations

SAVE RESULTS

Statistics

UricAcid

Summary

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 0.37% Stability: 3.21% Valid: 96.42%

Distribution

1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10.0 10.5 11.0 11.5

Statistics

Name	Value
Minimum	1.800
Maximum	11.300
Average	5.489
Standard Deviation	1.439

EN

Fri Apr 1 5:13:04 PM

RapidMiner Studio File Edit Process View Connections Settings Extensions Help Fri Apr 1 5:13:06 PM

<new process*> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP.lan

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

WaistCirc Predictions Production Model

Logistic Regression Model Weights Simulator Performance Lift Chart Predictions Production Model

Deep Learning

Random Forest

General Data Statistics Weights by Correlation Correlations

SAVE RESULTS

Statistics

WaistCirc

Summary

Number

Missing: 3.54% Infinite: 0.00% ID-ness: 3.00% Stability: 0.99% Valid: 92.47%

Distribution

55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180

Statistics

Name	Value
Minimum	56.200
Maximum	176
Average	98.307
Standard Deviation	16.253

EN

Fri Apr 1 5:13:06 PM

RapidMiner Studio 9.0.006 @ Sais-MBP:lan Fri Apr 1 5:13:08 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

RESTART BACK OPEN PROCESS EXPORT DEPLOY

Results

- Logistic Regression
- Deep Learning
- Random Forest
- General
- Data
- Statistics
- Weights by Correlation
- Correlations

SAVE RESULTS

Statistics

seqn

Summary

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 100.00% Stability: 0.04% Valid: 0.00%

Distribution

62,000 62,500 63,000 63,500 64,000 64,500 65,000 65,500 66,000 66,500 67,000 67,500 68,000 68,500 69,000 69,500 70,000 70,500 71,000 71,500 72,000

Statistics

Name	Value
Minimum	62161
Maximum	71915
Average	67030.674
Standard Deviation	2823.565

CORRELATIONS

RapidMiner Studio - Fri Apr 1 5:26:01 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

RESTART BACK OPEN PROCESS EXPORT DEPLOY

Results

Predictions Production Model

Logistic Regression Model Weights Simulator Performance Lift Chart Predictions Production Model

Deep Learning

Random Forest

General Data Statistics Weights by Correlation Correlations

SAVE RESULTS

Correlations

Attribu...	Age	Albumi...	BloodG...	BMI	HDL	Income	Marital...	Marital...	Marital...	Marital...	Met...	Race = ...	Race = ...	Race = ...	
WaistCirc	0.221	0.118	0.250	0.883	-0.342	-0.061	0.072	0.029	-0.014	-0.106	0.056	-0.490	-0.274	0.073	0.001
BMI	0.079	0.076	0.209	1	-0.291	-0.079	0.050	-0.020	0.013	-0.046	0.032	-0.429	-0.268	0.147	0.016
Triglyc...	0.070	0.108	0.270	0.163	-0.388	-0.039	0.025	0.051	0.040	-0.104	0.011	-0.427	-0.004	-0.156	0.025
BloodGl...	0.237	0.199	1	0.209	-0.214	-0.092	0.044	0.016	-0.024	-0.075	0.072	-0.362	-0.042	0.046	0.008
Age	1	0.233	0.237	0.079	0.079	-0.012	0.117	0.198	-0.203	-0.400	0.330	-0.249	-0.082	-0.003	0.025
UricAcid	0.148	0.177	0.056	0.270	-0.253	0.008	0.012	0.057	-0.043	-0.061	0.037	-0.246	-0.016	0.040	-0.045
Albumin...	0.233	1	0.199	0.076	-0.024	-0.096	0.032	-0.027	-0.053	-0.070	0.138	-0.159	-0.036	0.048	0.004
UrAlbCr	0.118	0.536	0.106	0.038	0.002	-0.048	0.007	0.002	-0.033	-0.036	0.056	-0.085	-0.023	0.061	-0.015
Marital ...	0.330	0.138	0.072	0.032	0.073	-0.125	-0.091	-0.271	-0.084	-0.139	1	-0.049	-0.056	0.055	-0.036
Marital ...	0.198	-0.027	0.016	-0.020	-0.047	0.351	-0.332	1	-0.306	-0.508	-0.271	-0.047	0.118	-0.175	-0.015
Race = ...	-0.077	0.032	0.050	0.069	-0.077	-0.080	-0.029	-0.021	0.106	-0.048	-0.024	-0.047	-0.142	-0.187	-0.119
Race = ...	0.025	0.004	0.008	0.016	-0.067	-0.067	0.014	-0.015	0.056	-0.048	-0.036	-0.043	-0.143	-0.188	1
Marital ...	0.117	0.032	0.044	0.050	0.011	-0.115	1	-0.332	-0.103	-0.171	-0.091	-0.041	-0.083	0.019	0.014
Race = ...	0.109	-0.042	-0.050	0.015	-0.030	0.052	0.062	0.097	-0.024	-0.123	0.039	-0.032	-0.329	-0.434	-0.276
seqn	0.009	-0.023	-0.019	-0.003	0.056	0.019	0.035	0.015	-0.040	0.004	-0.041	-0.011	-0.025	0.008	0.009

RapidMiner Studio - Fri Apr 1 5:26:10 PM

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

RESTART BACK OPEN PROCESS EXPORT DEPLOY

Results

Predictions Production Model

Logistic Regression Model Weights Simulator Performance Lift Chart Predictions Production Model

Deep Learning

Random Forest

General Data Statistics Weights by Correlation Correlations

SAVE RESULTS

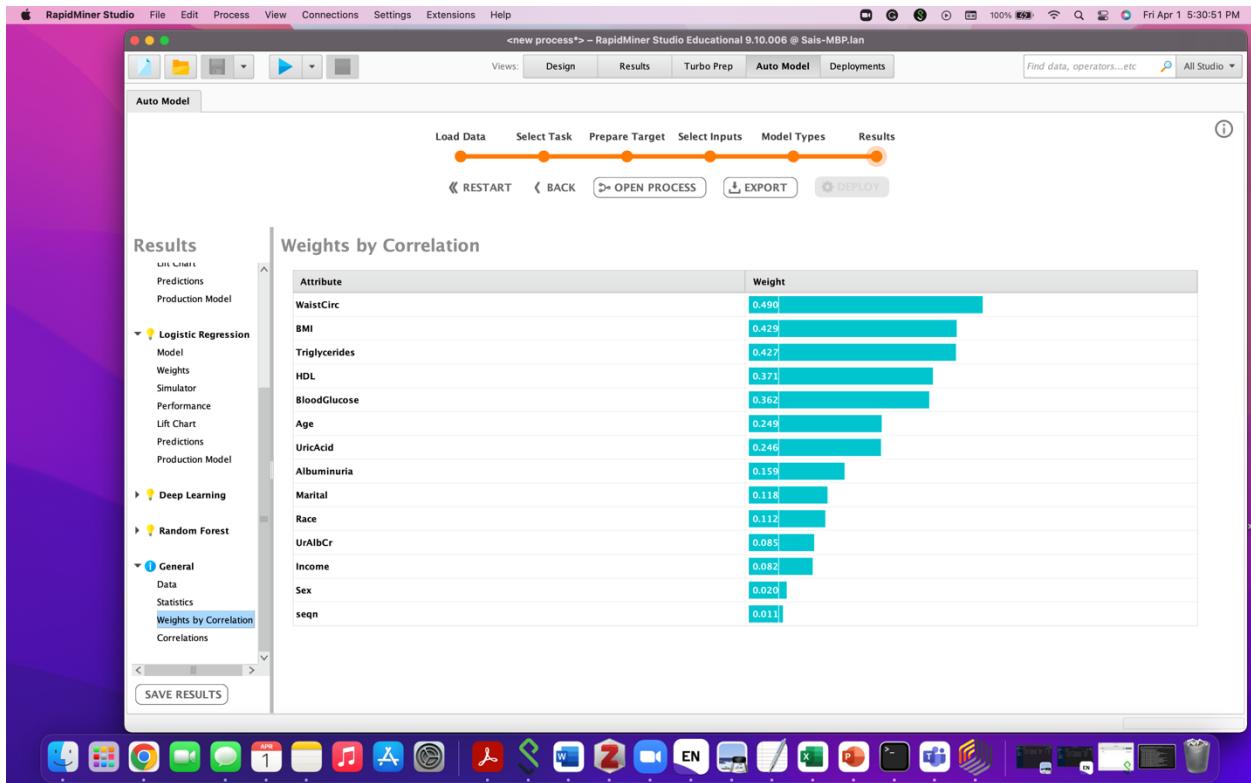
Correlations

Attribu...	Age	Albumi...	BloodG...	BMI	HDL	Income	Marital...	Marital...	Marital...	Marital...	Met...	Race = ...	Race = ...	Race = ...	
Metabol...	-0.249	-0.159	-0.362	-0.429	0.371	0.082	-0.041	-0.047	0.029	0.103	-0.049	1	0.096	0.014	-0.043
HDL	0.079	-0.024	-0.214	-0.291	1	0.076	0.011	-0.047	-0.057	0.032	0.073	0.371	0.028	0.115	-0.067
Marital ...	-0.400	-0.070	-0.075	-0.046	0.032	-0.136	-0.171	-0.508	-0.158	1	-0.139	0.103	0.040	0.153	-0.048
Race = ...	-0.082	-0.036	-0.042	-0.268	0.028	0.119	-0.083	0.118	-0.081	0.040	-0.056	0.096	1	-0.224	-0.143
Income	-0.012	-0.096	-0.092	-0.079	0.076	1	-0.115	0.351	-0.126	-0.136	-0.125	0.082	0.119	-0.040	-0.067
Marital ...	-0.203	-0.053	-0.024	0.013	-0.057	-0.126	-0.103	-0.306	1	-0.158	-0.084	0.029	-0.081	-0.026	0.056
Sex = ...	-0.012	-0.019	-0.090	0.051	0.318	-0.030	0.005	-0.089	-0.026	0.014	0.142	0.020	-0.014	0.055	0.004
Race = ...	-0.003	0.048	0.046	0.147	0.115	-0.040	0.019	-0.175	-0.026	0.153	0.055	0.014	-0.224	1	-0.188
seqn	0.009	-0.023	-0.039	-0.003	0.056	0.019	0.035	0.015	-0.040	0.004	-0.041	-0.011	-0.025	0.008	0.009
Race = ...	0.109	-0.042	-0.050	0.015	-0.030	0.052	0.062	0.097	-0.024	-0.123	0.039	-0.032	-0.329	-0.434	-0.276
Marital ...	0.117	0.032	0.044	0.050	0.011	-0.115	1	-0.332	-0.103	-0.171	-0.091	-0.041	-0.083	0.019	0.014
Race = ...	0.025	0.004	0.008	0.016	-0.067	-0.067	0.014	-0.015	0.056	-0.048	-0.036	-0.043	-0.143	-0.188	1
seqn	0.009	-0.023	-0.039	-0.003	0.056	0.019	0.035	0.015	-0.040	0.004	-0.041	-0.011	-0.025	0.008	0.009
Race = ...	-0.077	0.032	0.050	0.069	-0.077	-0.080	-0.029	-0.021	0.106	-0.048	-0.024	-0.047	-0.142	-0.187	-0.119
Marital ...	0.198	-0.027	0.016	-0.020	-0.047	0.351	-0.332	1	-0.306	-0.508	-0.271	-0.047	0.118	-0.175	-0.015
Marital ...	0.330	0.138	0.072	0.032	0.073	-0.125	-0.091	-0.271	-0.084	-0.139	1	-0.049	-0.056	0.055	-0.036

Highest positive correlation – HDL (0.371)

Highest negative correlation – WaistCirc(-0.490)

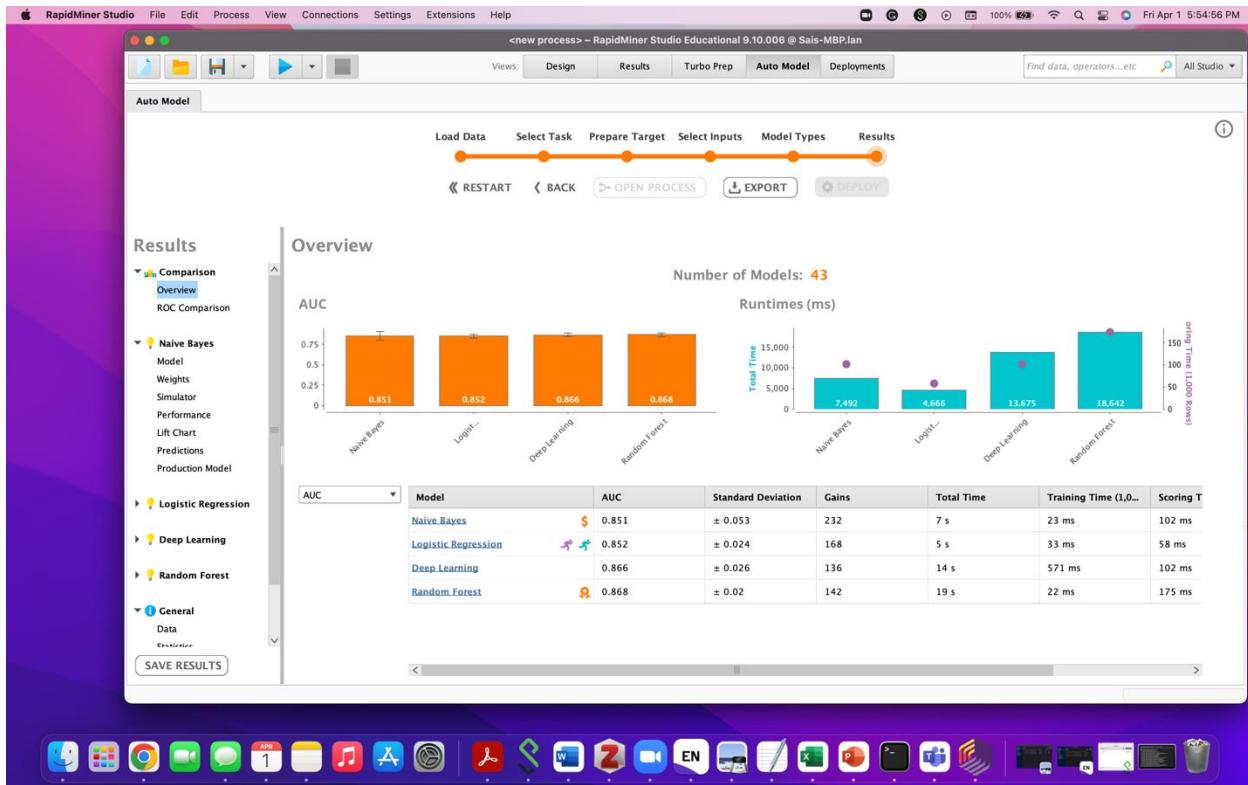
WEIGHTS CORRELATION



Highest 3 weights are:

1. WaistCirc (0.490)
2. BMI (0.429)
3. Triglycerides (0.427)

BLOOD GLUCOSE, HDL, TRIGLYCERIDES



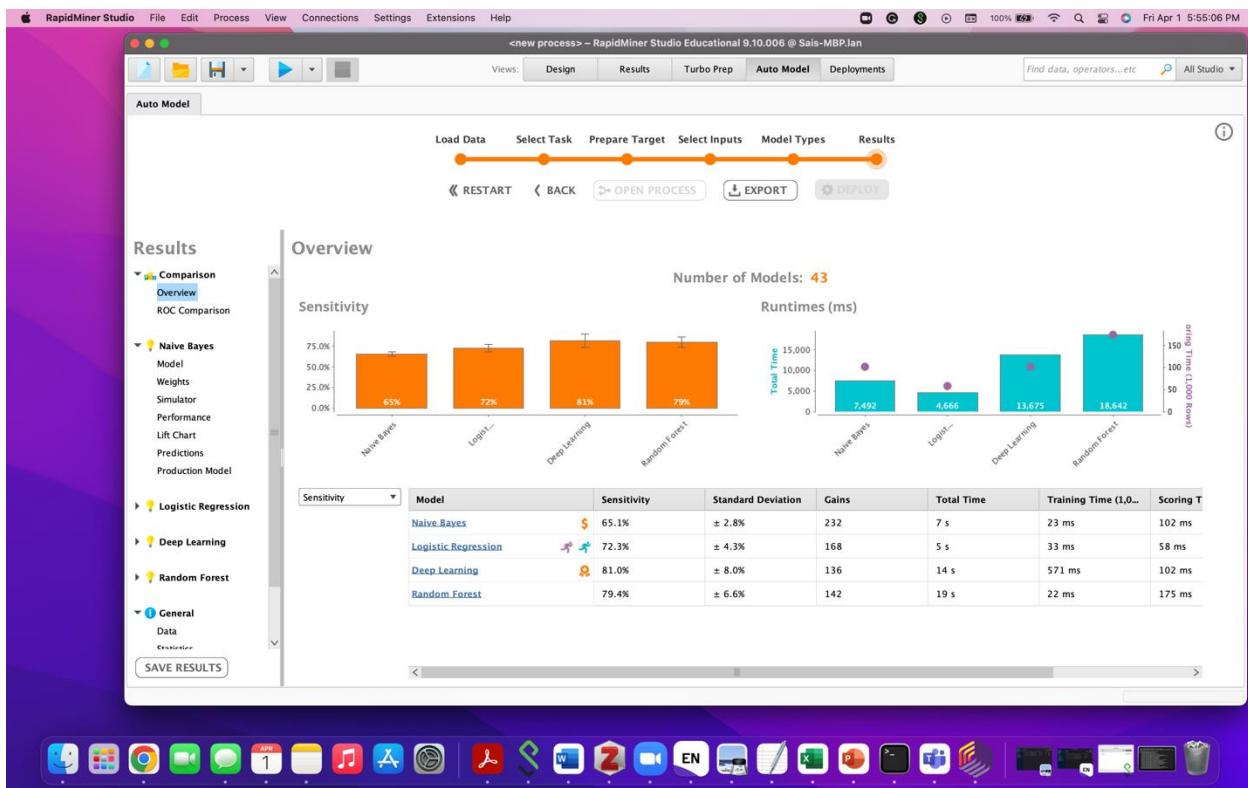
AUC:

Highest AUC: Random Forest (0.868)

Lowest AUC: Naïve Bayes (0.851)

Fastest results: Logistic Regression (5s)

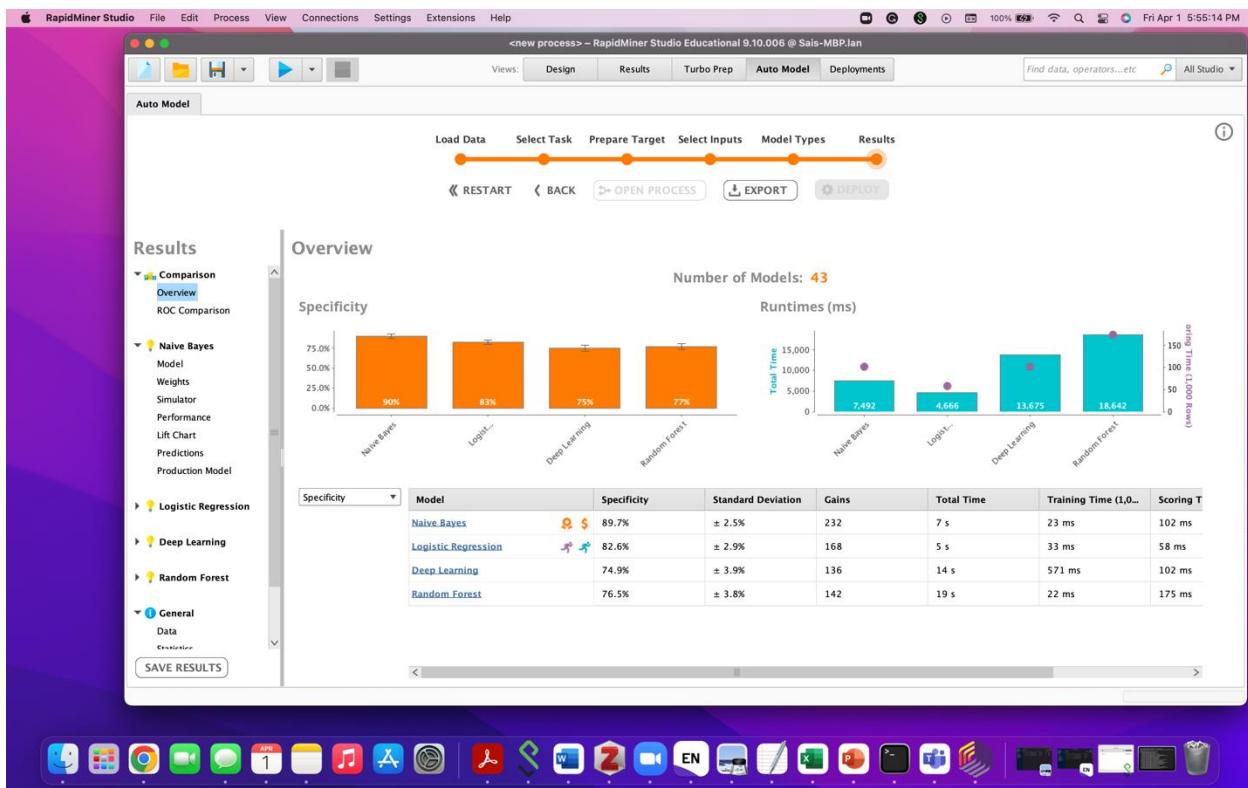
Slowest results: Random Forest (19s)



SENSITIVITY:

Highest sensitivity: Random Forest (79.4%)

Lowest sensitivity: (65.1%)

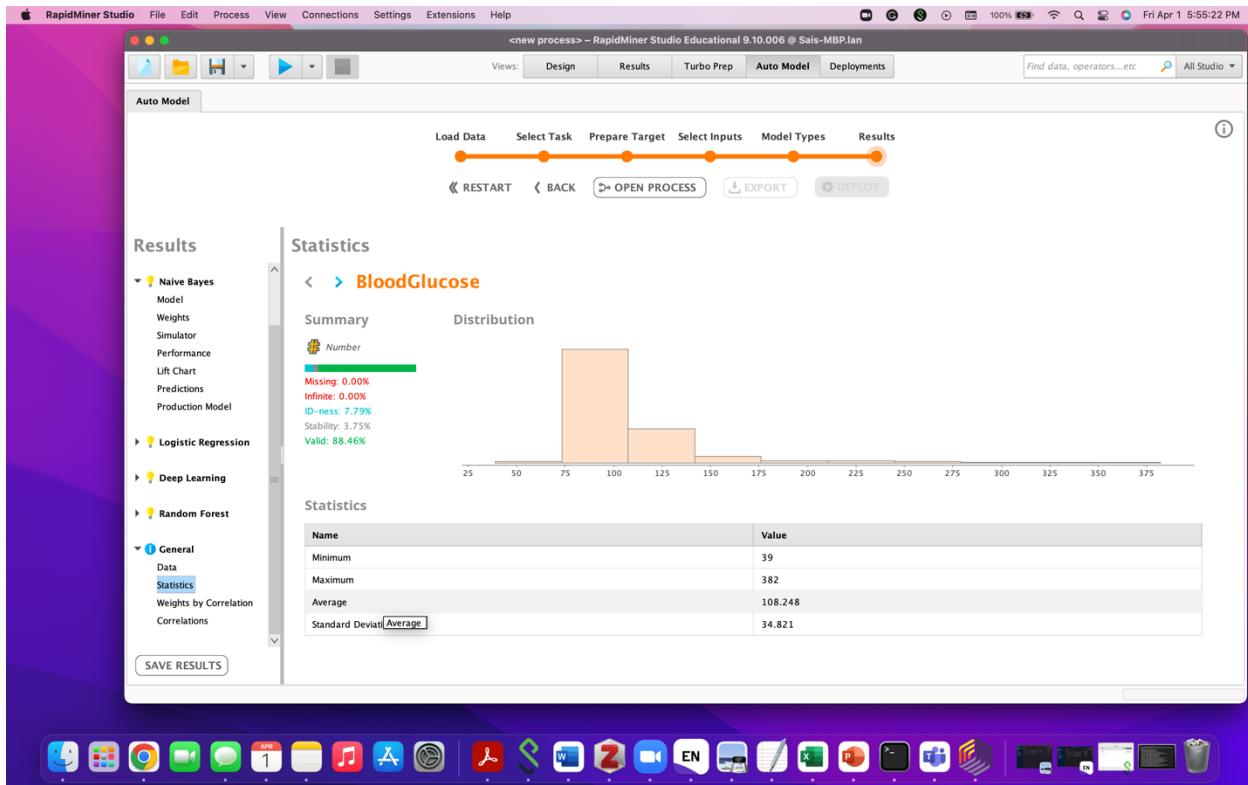


SPECIFICITY:

Highest: Naïve Bayes (89.7%)

Lowest: Deep Learning (74.9%)

STATISTICS



RapidMiner Studio File Edit Process View Connections Settings Extensions Help Fri Apr 1 5:55:29 PM

<new process> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP:lan

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

Naive Bayes Model Weights Simulator Performance Lift Chart Predictions Production Model

Logistic Regression

Deep Learning

Random Forest

General Data Statistics Weights by Correlation Correlations

SAVE RESULTS

Statistics

HDL

Summary

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 3.87% Stability: 3.62% Valid: 92.50%

Distribution

10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160

Statistics

Name	Value
Minimum	14
Maximum	156
Average	53.369
Standard Deviation	15.186

SAVE RESULTS

RapidMiner Studio File Edit Process View Connections Settings Extensions Help Fri Apr 1 5:55:36 PM

<new process> - RapidMiner Studio Educational 9.10.006 @ Sais-MBP:lan

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc All Studio

Auto Model

Load Data Select Task Prepare Target Select Inputs Model Types Results

« RESTART < BACK D OPEN PROCESS EXPORT DEPLOY

Results

Naive Bayes Model Weights Simulator Performance Lift Chart Predictions Production Model

Logistic Regression

Deep Learning

Random Forest

General Data Statistics Weights by Correlation Correlations

SAVE RESULTS

Statistics

Triglycerides

Summary

Number

Missing: 0.00% Infinite: 0.00% ID-ness: 14.37% Stability: 1.42% Valid: 84.21%

Distribution

0 100 200 300 400 500 600 700 800 900 1,000 1,100 1,200 1,300 1,400 1,500 1,600

Statistics

Name	Value
Minimum	26
Maximum	1562
Average	128.125
Standard Deviation	95.322

CORRELATIONS:

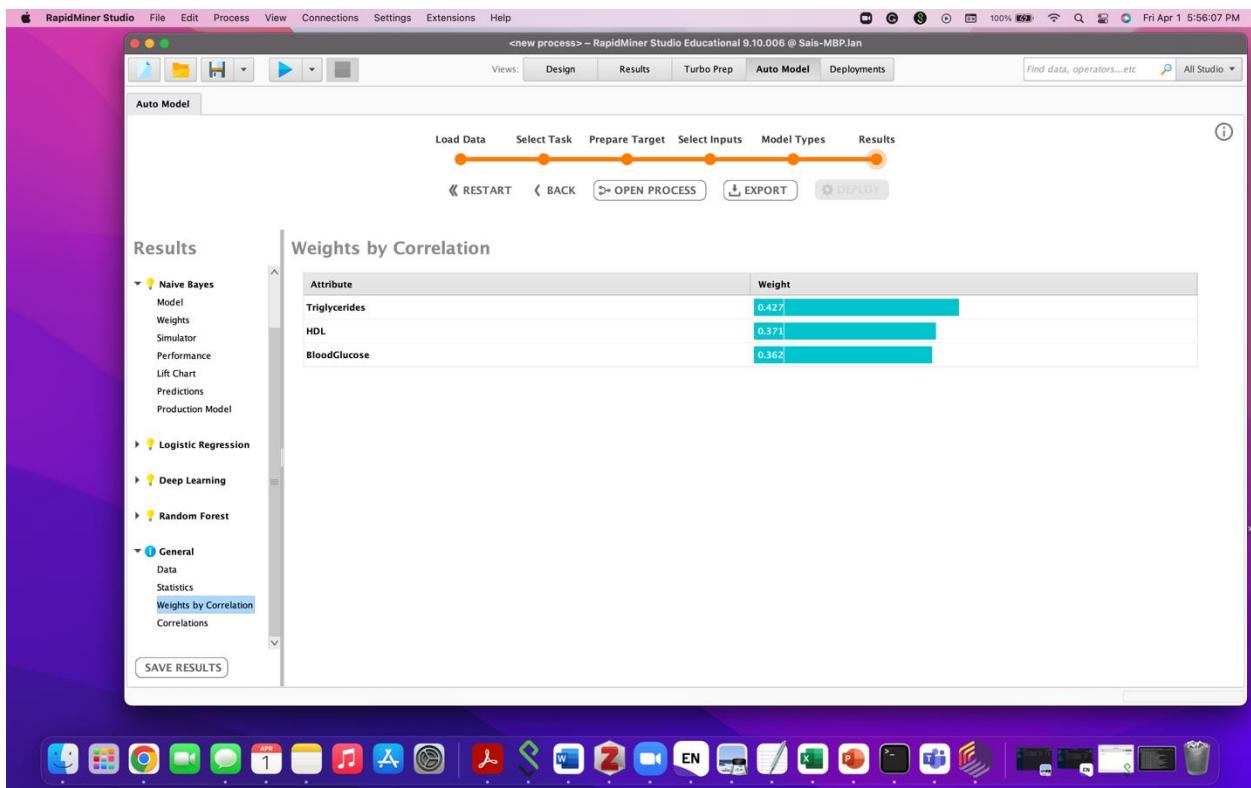
The screenshot shows the RapidMiner Studio interface with the 'Auto Model' tab selected. The main area displays a correlation matrix titled 'Correlations'. The matrix shows the following data:

Attribu...	BloodG...	HDL	Metabo...	Triglyc...
BloodGl...	1	-0.214	-0.362	0.270
HDL	-0.214	1	0.371	-0.388
Metabol...	-0.362	0.371	1	-0.427
Triglyce...	0.270	-0.388	-0.427	1

The sidebar on the left lists various models and general statistics, with 'Correlations' selected under 'General'. A 'SAVE RESULTS' button is located at the bottom left of the results panel.

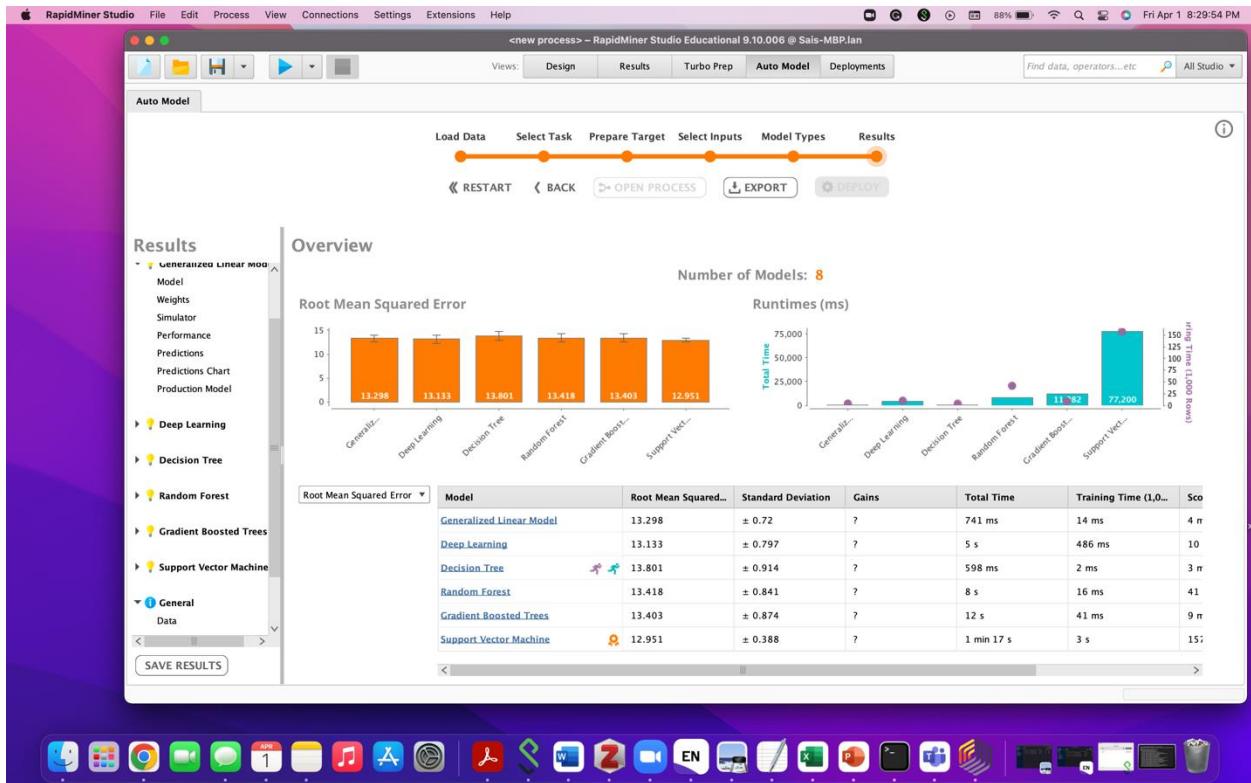
Highest positive correlation: HDL (0.371)

Highest negative correlation: Triglycerides (-0.427)

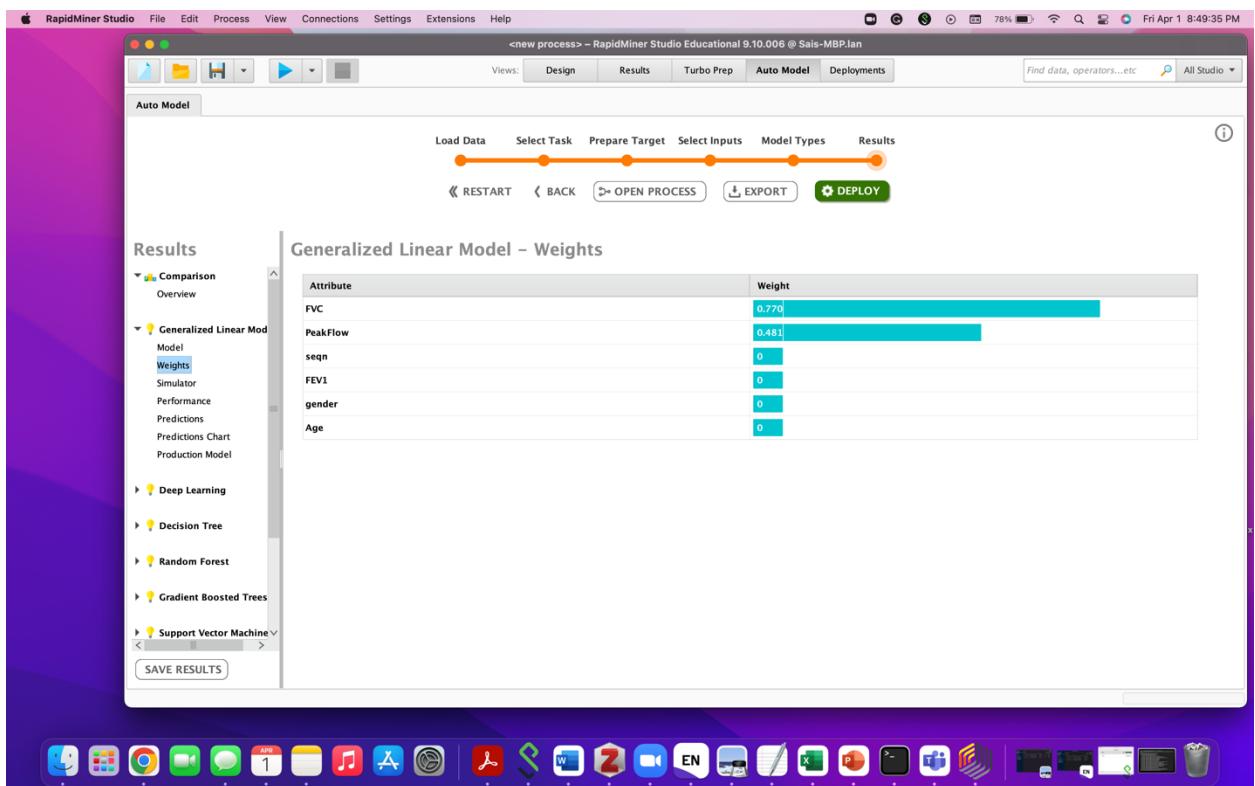


- Triglycerides has the highest weights by correlation and showed good performance than other variables.

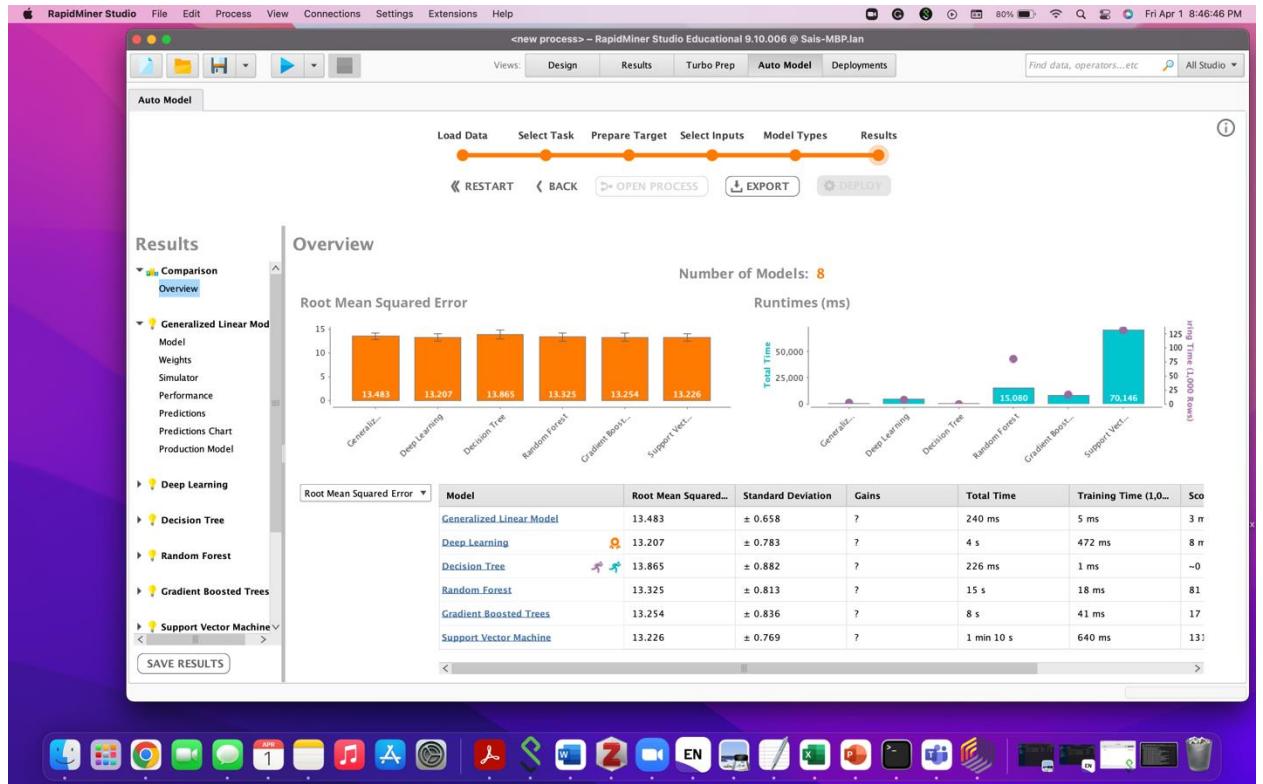
REGRESSION



- The algorithm that gave the lowest root mean squared errors (RMSEs) – Support Vector Machine (12.951) which indicates best performance.

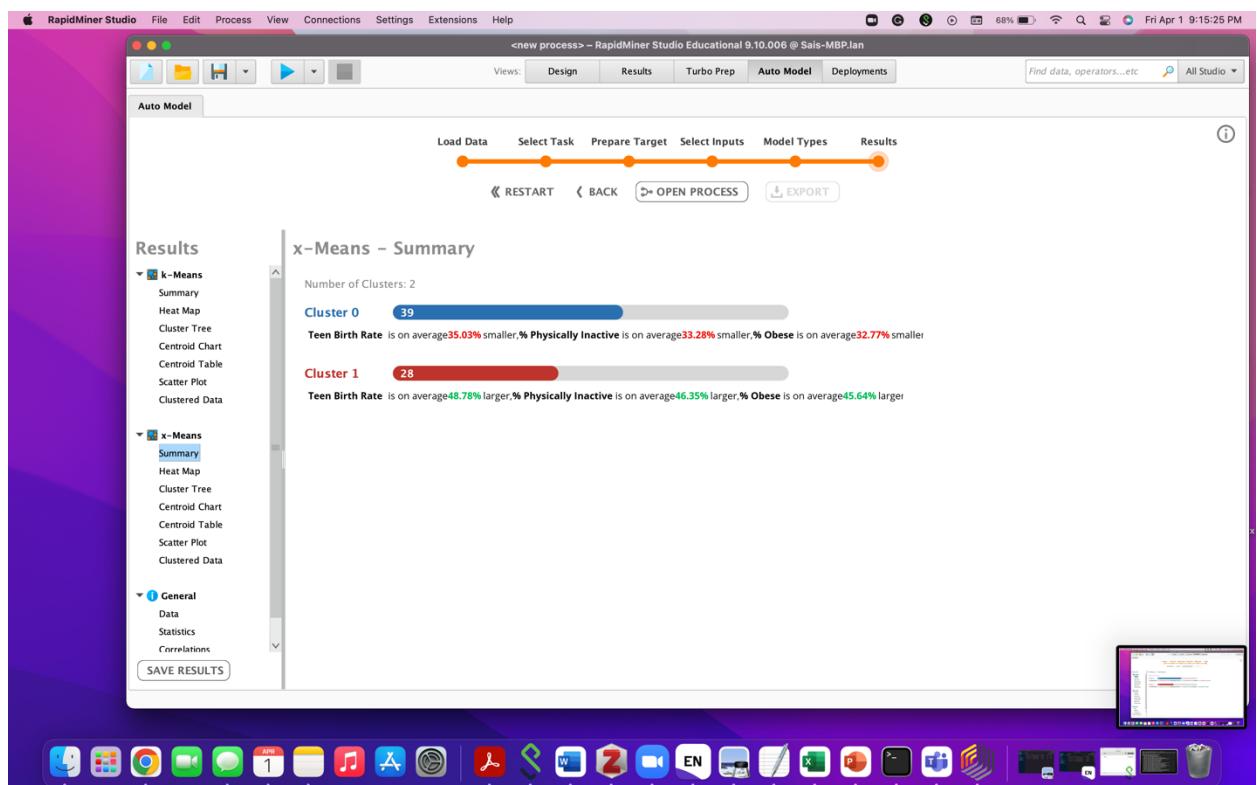
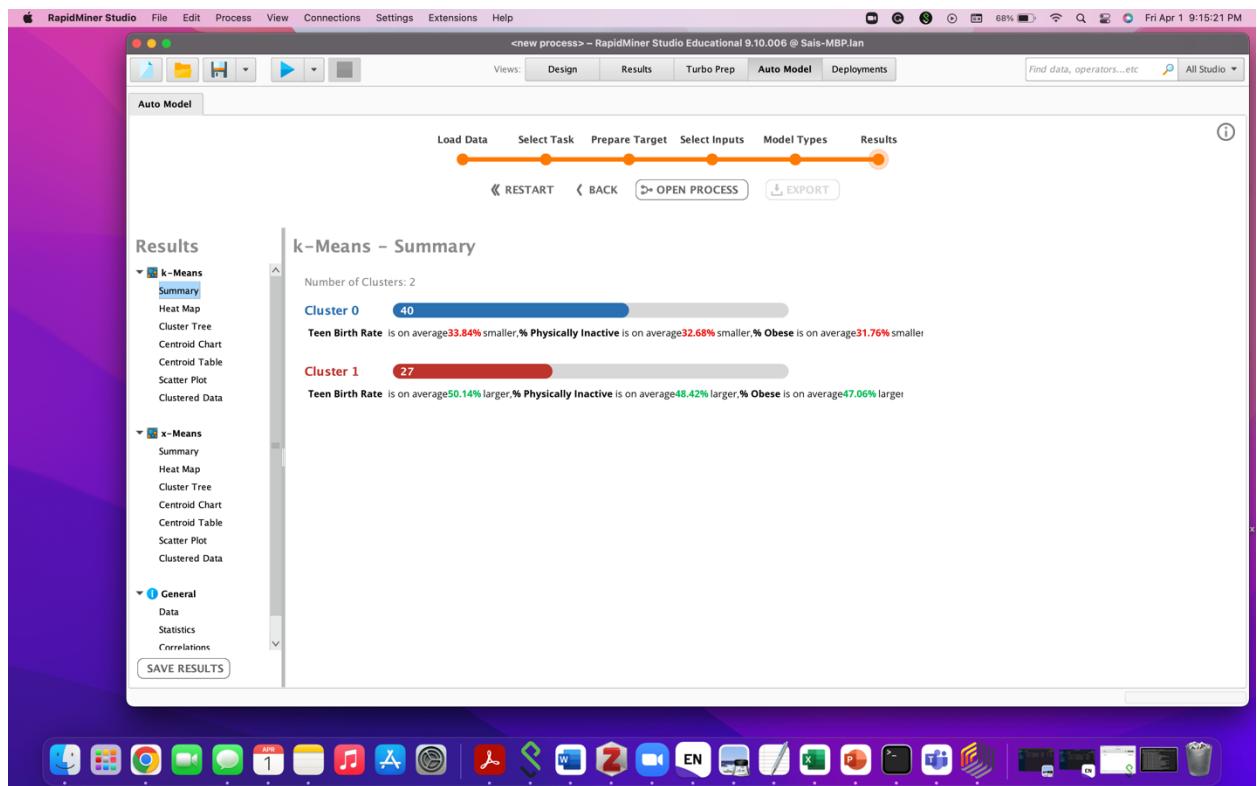


- Apart from FVC, good results can be obtained from Peak Flow as it shows faster results compared to other variables.



- Previously Support Vector Machine has the least RMSEs and best performance. When the variables have changed, Deep Learning algorithm indicated least RMSEs and best performance.

CLUSTERING:



- 2 clusters were found by both. No, the findings were not similar. Cluster coding was similar for k-Means and x-Means but within them, the cluster coding was different for Cluster 0 and Cluster 1.
- Variables that were predictive are: Teen Birth rate, Physically inactive, Obese.
- There isn't much numerical variation of the variables in the summary presented by k-means and x-means with clustering.

ARTIFICIAL INTELLIGENCE:

1.

The screenshot shows the Google Cloud Natural Language API demo page. At the top, there's a navigation bar with links for Google Cloud, Why Google, Solutions, Products (which is underlined), Pricing, and Getting Started. To the right are search, documentation, support, language selection, console access, and user account options. Below the navigation is a sub-navigation for Cloud Natural Language with Contact Us and Get started for free buttons.

The main content area has a DEMO header. On the left, a sidebar lists Natural Language AI features: Benefits, Demo (which is selected and highlighted in blue), Key features, Our customers, Documentation, Compare features, Pricing, and Take the next step. In the center, the title is "Natural Language API demo" with a "Try the API" button. A text input field contains the sentence "The patient was discharged from the hospital last week". Below it is a "See supported languages" link. To the right is a "RESET" button. Further down, there are tabs for Entities, Sentiment, Syntax, and Categories. The Entities tab is active, showing results for the sentence. It lists two entities: "1. patient" (PERSON) with Salience: 0.89 and "2. hospital" (LOCATION) with Salience: 0.11. At the bottom left, there's a "KEY FEATURES" section with the heading "Three natural language solutions that work".

cloud.google.com/natural-language/

Google Cloud Why Google Solutions Products Pricing Getting Started

Cloud Natural Language

DEMOS

Natural Language API demo

Try the API

The patient was discharged from the hospital last week

RESET

See supported languages

Entities Sentiment Syntax Categories

Dependency Parse label Part of speech Lemma Morphology

	det	nsubjpass	auxpass	root	prep	det	pobj	amod	tmod
The	DET	patient	was be VERB	discharged discharge VERB	from	the	hospital	last	week
	number=SINGULAR	mood=INDICATIVE number=SINGULAR person=THIRD tense=PAST		voice=PASSIVE tense=PAST	DET	number=SINGULAR		ADJ	number=SINGULAR

The -DET

Patient - NOUN

Was - VERB

Discharged - VERB

From - ADP

The - DET

Hospital - NOUN

Last - ADJ

Week – NOUN

2. She loves her new allergy medicine

The screenshot shows the Google Cloud Natural Language API interface. At the top, there's a navigation bar with links like 'Why Google', 'Solutions', 'Products' (which is underlined), 'Pricing', and 'Getting Started'. On the right side of the header are icons for 'Language', 'Console', and user profile.

The main area is titled 'Try the API' and contains a text input field with the text 'She loves her new allergy medicine'. Below this is a 'RESET' button. To the left of the input field is a sidebar titled 'Natural Language AI' with sections for 'Benefits', 'Demo' (which is currently selected), 'Key features', 'Our customers', 'Documentation', 'Compare features', 'Pricing', and 'Take the next step'.

The main content area has tabs for 'Entities', 'Sentiment' (which is selected), 'Syntax', and 'Categories'. Under the 'Sentiment' tab, there's a section for 'Document and Sentence Level Sentiment' with a table:

	Score	Magnitude
Entire Document	0.9	0.9
She loves her new allergy medicine	0.9	0.9

Below this is a 'Score Range' bar with three segments: green (0.25 - 1.0), yellow (-0.25 - 0.25), and red (-1.0 - -0.25).

Under the 'Entities' tab, there's a section for 'Entity Level Sentiment' showing one entity:

1. allergy medicine	OTHER
Sentiment: Score 0.8 Magnitude 0.8	

3. ENTITIES

cloud.google.com/natural-language/

Google Cloud Why Google Solutions Products Pricing Getting Started

Cloud Natural Language

Contact Us Get started for free

Natural Language AI

Benefits Demo Key features Our customers Documentation Compare features Pricing Take the next step

Try the API

63 year old type 2 diabetic man with diabetic retinopathy and nephropathy returns today for refills of his Lantus insulin and to get a pneumovax injection. He has a living will and a surrogate power of attorney.

RESET

See supported languages

Entities Sentiment Syntax Categories

63:12 year old type:2:13 diabetic:man:1 with:diabetic retinopathy:2 and:(nephropathy):6:returns:3 today:for:(refills):3:of:his:Lantus:7:(insulin):4 and:to: get:a:(pneumovax injection):8. He:has:a:(living will):9 and:a:surrogate:power:10:of:(attorney):11.

1. man	PERSON	OTHER
2. diabetic retinopathy		OTHER
3. refills	OTHER	
4. insulin		OTHER
5. returns	OTHER	
6. nephropathy		OTHER
7. Lantus	CONSUMER GOOD	
8. pneumovax injection		OTHER

Salience: 0.62
Salience: 0.08
Salience: 0.05
Salience: 0.05
Salience: 0.05
Salience: 0.05
Salience: 0.03
Salience: 0.02

cloud.google.com/natural-language/

Google Cloud Why Google Solutions Products Pricing Getting Started

Cloud Natural Language

Contact Us Get started for free

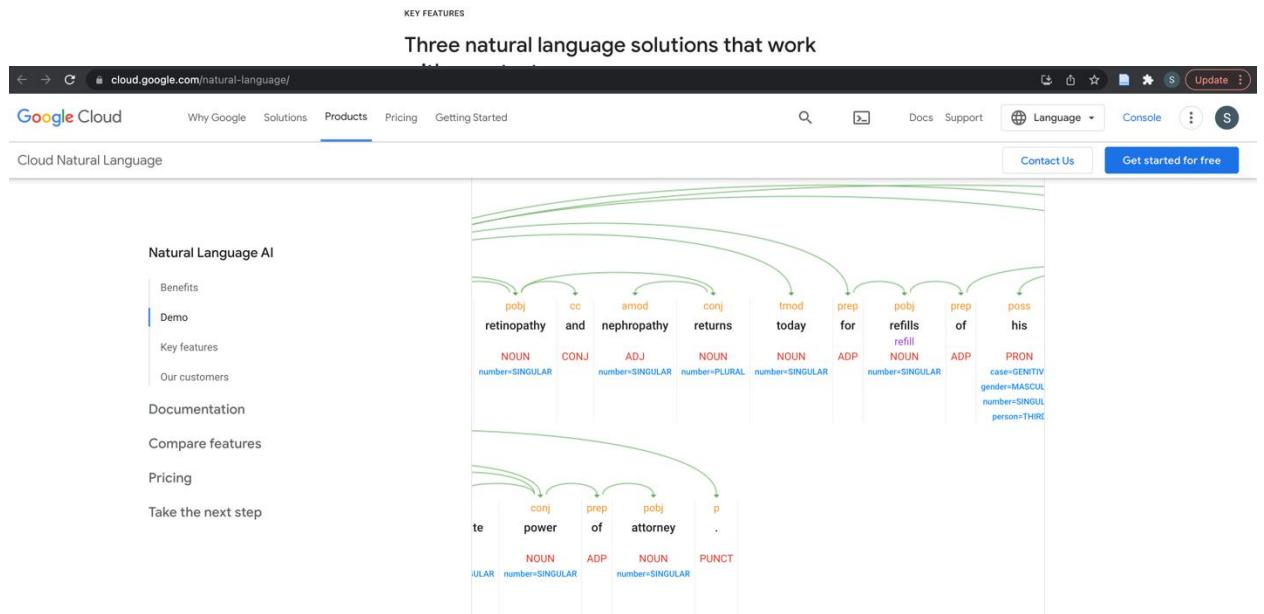
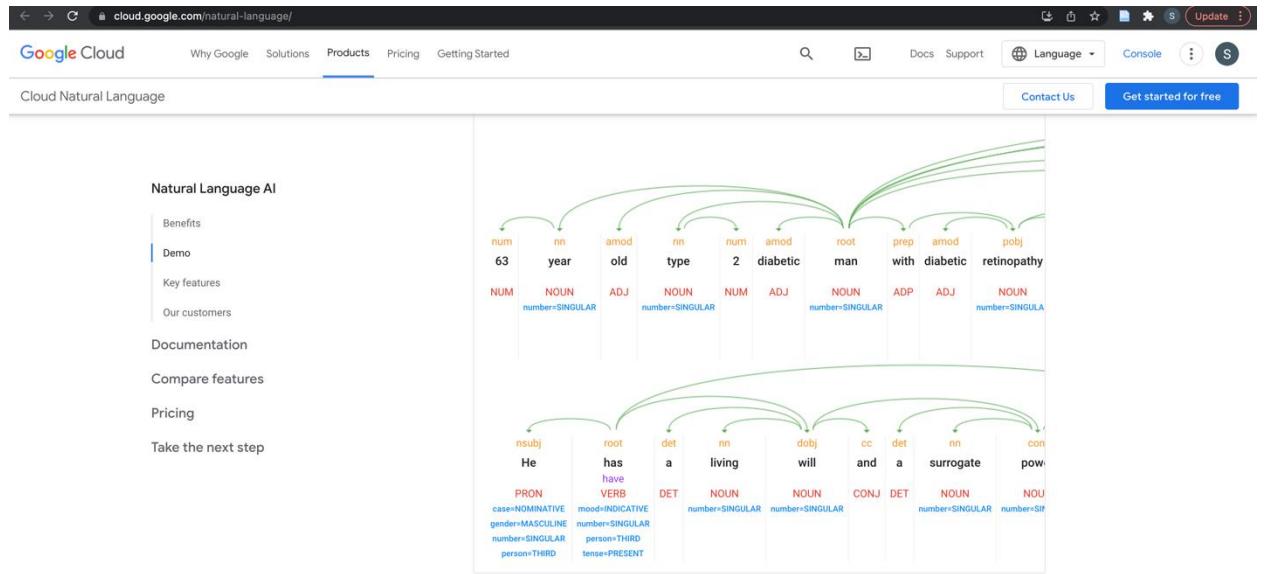
Natural Language AI

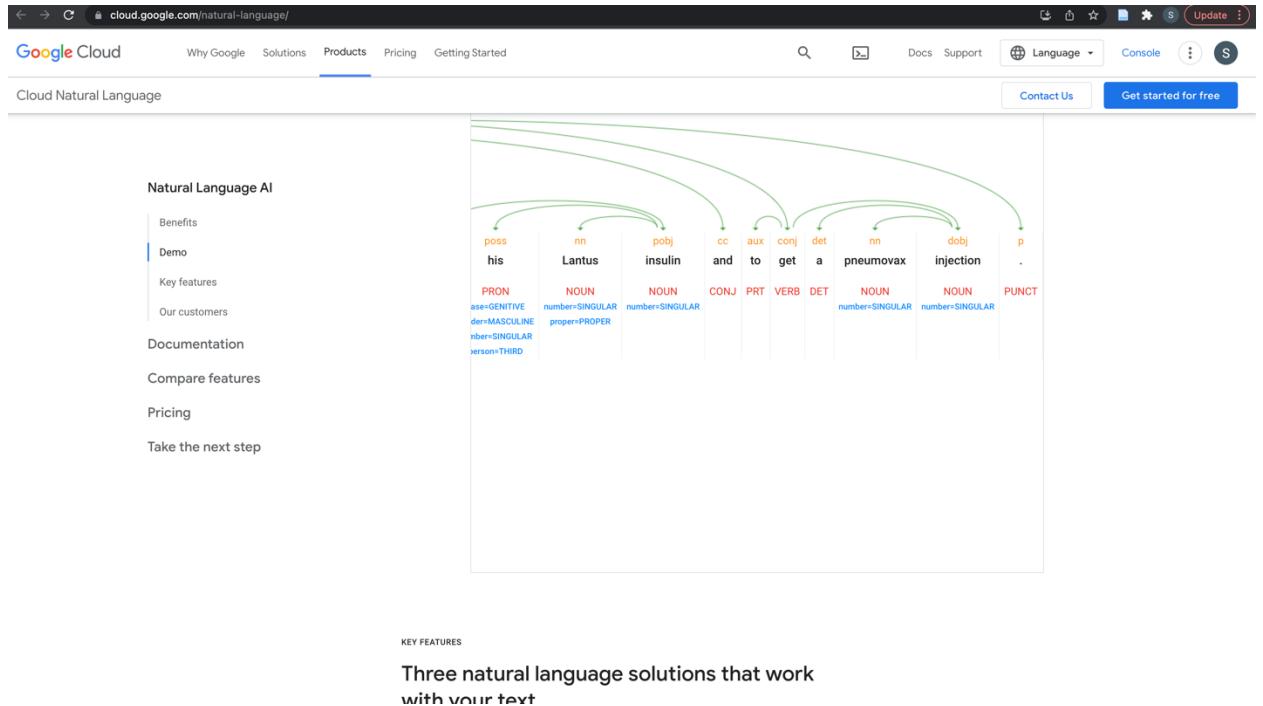
Benefits Demo Key features Our customers Documentation Compare features Pricing Take the next step

7. Lantus	CONSUMER GOOD	OTHER
8. pneumovax injection		OTHER
9. living will	OTHER	OTHER
10. power		OTHER
11. attorney	PERSON	
12. 63		NUMBER
13. 2		NUMBER

Salience: 0.03
Salience: 0.02
Salience: 0.02
Salience: 0.01
Salience: 0.01
Salience: 0.01
Salience: 0.01

SYNTAX





CATEGORIES:

The screenshot shows the Google Cloud Natural Language AI demo page. It features a sidebar with the same set of links as the previous page. In the main area, there's a section titled "Try the API" with a text input field containing a medical sentence: "63 year old type 2 diabetic man with diabetic retinopathy and nephropathy returns today for refills of his Lantus insulin and to get a pneumovax injection. He has a living will and a surrogate power of attorney." Below this, there's a "See supported languages" link. A "RESET" button is also present. The interface includes tabs for Entities, Sentiment, Syntax, and Categories (which is currently selected). Under the Categories tab, two categories are listed with their confidence scores: "/Health/Health Conditions/Diabetes" (Confidence: 0.97) and "/Health/Health Conditions/Endocrine Conditions" (Confidence: 0.97). A link "See a complete list of content categories." is provided at the bottom.

1. No, it didn't find all the products listed. Pnuemovax injection and insulin are listed under 'other' but they are to listed under products.
2. Pleural Refills is handled as 'NOUN' and morphology is described as number= singular.
3. Under health conditions, it identified 2 categories- 'Diabetes' and 'Endocrine Conditions'.

