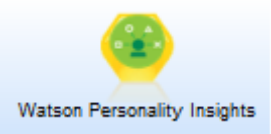
Watson Personality Insights

Analyze text using Watson Personality Insights resulting in scores on 52 attributes



Product: IBM® SPSS® Modeler

Extension type: Utility

Table of Contents

Description:

This SPSS Modeler node allows you to easily send text to the Watson Personality Insights Service and get scores on 52 different personality attributes. Watson Personality Insights allows users to derive insights from social media, enterprise data, or other digital communications. The service uses linguistic analytics to extract cognitive and social characteristics, including Big Five personality, values, and needs, from text. The results from Personality Insights can help businesses to understand their clients' preferences and improve customer satisfaction by anticipating customer needs and recommending the next best actions. IBM SPSS Modeler can use the results from User Modeling as additional predictors in machine learning models.

Requirements:

- SPSS Modeler v16.0 or later
- R: http://www.r-project.org/
- 'R Essentials for SPSS Modeler' plugin: https://developer.ibm.com/predictiveanalytics/downloads/
- A Bluemix account with Watson Personality Insights Service activated
- **Note that the Watson Personality Insights requires AT MINIMUM 100 words to give output.**
    - Results will be $null otherwise
    - Make sure there are no "\" in the data (including carriage return such as \r\n)

Installation:

Close SPSS Modeler. Save the .cfe file in the CDB folder of the IBM SPSS Modeler installation directory for Windows and Linux. The copy should reside in that same folder and not in a sub-folder.

For example, for Windows 7 the default location is "C:\ProgramData\IBM\SPSS\Modeler\16\CDB". If the ProgramData folder is hidden type the path manually.

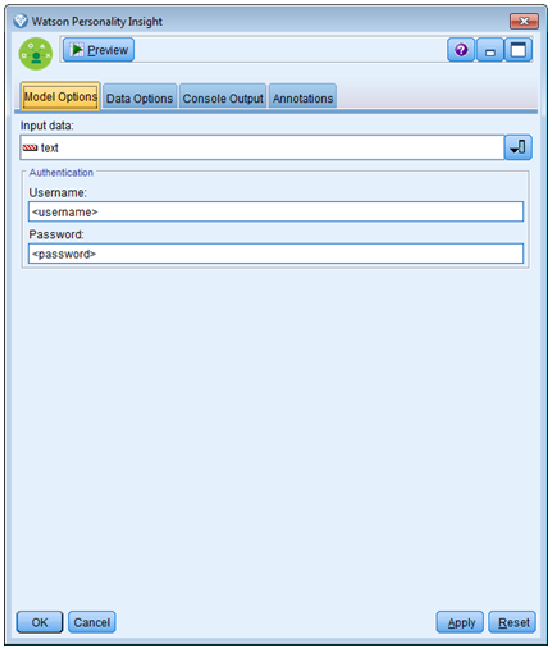
Restart SPSS Modeler: the node will now appear in the Field Ops palette.

R Packages used:

- 'httr' package created by Hadley Wickham Lang https://cran.r-project.org/web/packages/httr/
- 'jsonlite' package created by Jeroen Ooms, Duncan Temple Lang, Lloyd Hilaiel https://cran.r-project.org/web/packages/jsonlite/

User Interface

- Double click on the node to get to the options. There are the following fields:
  - Input data: click the drop down to select the field containing the text for analysis
  - User Name and Password: These are the credentials obtained in Bluemix when Watson Language Translation is added to your account (This is not your Bluemix user name and password).



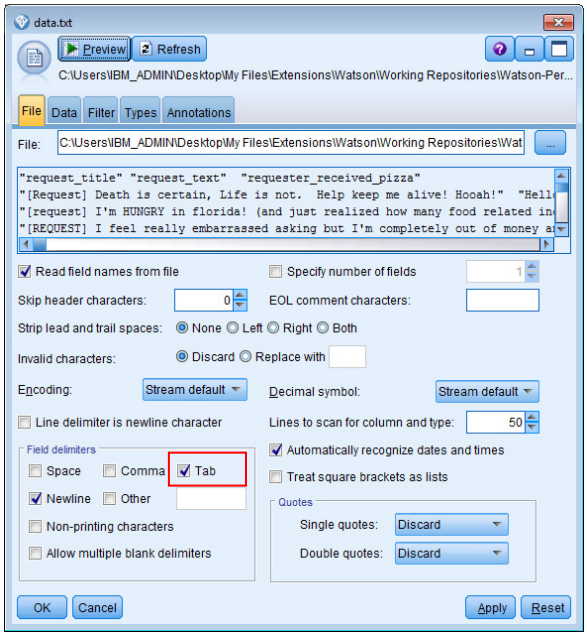
Example

This example will demonstrate how to use this extension on a sample dataset. The sample dataset contains 241 textual requests for pizza from the Reddit community Random Acts of Pizza (http://www.reddit.com/r/Random_Acts_Of_Pizza/) together with their outcome (successful/unsuccessful). Dr. Olav Laudy, Chief Data Scientist with IBM, built a model using this dataset and extension, which has been included in this repository. This example will only show a demonstration of using this extension.

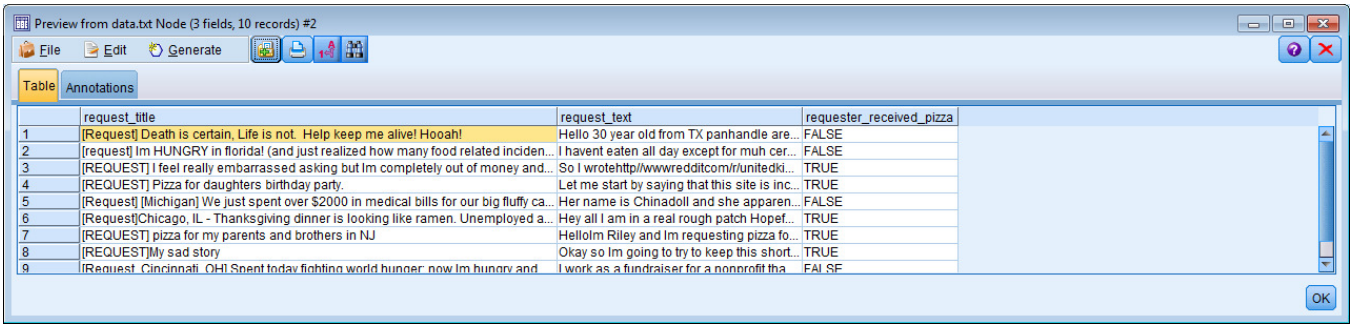
User Input

1. Add a Var. File node to the canvas from the Sources palette

2. Open the 'data.txt' file that can be downloaded from the Example directory of this repository. Make sure that the Field Delimiter is set to Tab

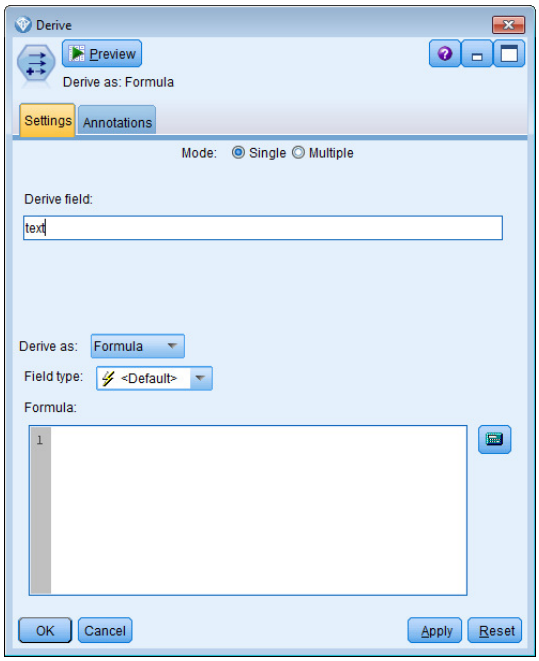


3. Click Preview to see that the data is in three columns



4. Go to the Field Ops palette and connect a Derive node to the data.txt node

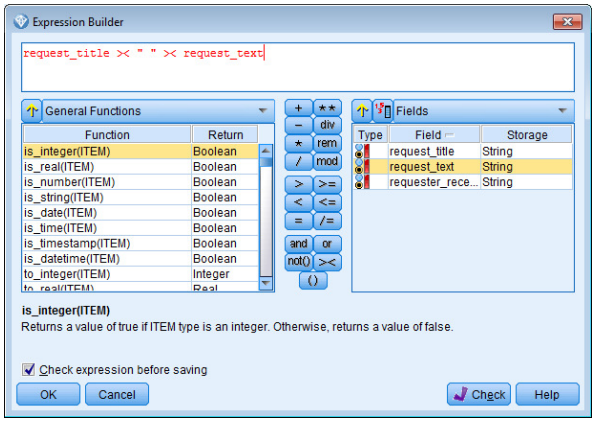
    a. Name the derived field “text”



    b. Click the calculator icon next to the Formula box to open the Expression Builder

        i. Double click “request_title”, then click the “><” button to concatenate text

        ii. Type in “ “ then click the “><” button again before clicking “request_text”



        iii. This will create a derived field that combines the title and the text of the post

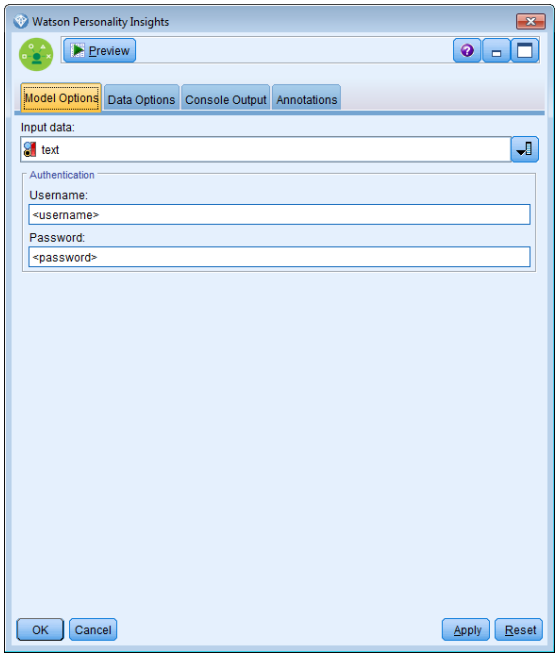
5. In the Record Ops palette, add the Sample node to the stream, this is to restrict the number of records analyzed for this demo

6. Next add the Watson Personality Insights node to the stream from the Field Ops palette,

7. Now fill in the fields:

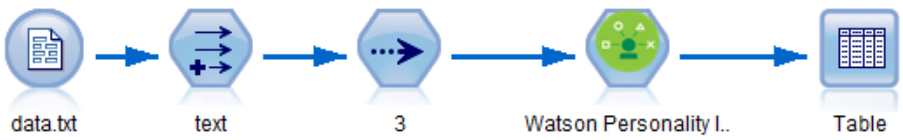
   Input data: Select text (the derived field)

   Username and Password: Enter credentials from Bluemix



8. Add a Table node from the Output palette to see the results from the Watson Service

9. Click on the table and run the stream



Results and Interpretation

If you used the first 3 cases from the sample dataset, your results should match the table below:

Table Annotations

	...	text	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Adventurousness	Artistic interests	Emotionality	Imagination	Intellect	Liberalism
1	...	[Request] Deat...	0.618	0.428	0.252	0.522	0.653	0.306	0.664	0.674	0.550	0.543	0.787
2	...	[request] Im HU...	0.561	0.071	0.170	0.275	0.980	0.217	0.009	0.819	0.364	0.375	0.886
3	...	[REQUEST] I fe...	0.280	0.254	0.249	0.598	0.975	0.203	0.017	0.729	0.013	0.225	0.808

	Achievement striving	Cautiousness	Dutifulness	Orderliness	Self-discipline	Self-efficacy	Activity level	Assertiveness
1	0.564	0.526	0.350	0.736	0.273	0.604	0.571	0.412
2	0.226	0.402	0.199	0.781	0.009	0.196	0.077	0.188
3	0.566	0.185	0.073	1.000	0.180	0.308	0.765	0.965

	Cheerfulness	Excitement-seeking	Friendline...	Gregariousness	Altruism	Cooperati...	Modesty	Morality	Sympathy	Trust	Anger	Anxiety	Depression	Immoderation	Self-consciousness
1	0.327	0.218	0.322	0.073	0.513	0.706	0.075	0.544	1.000	0.401	0.697	0.579	0.512	0.765	0.959
2	0.298	0.136	0.073	0.027	0.058	0.330	0.142	0.338	0.593	0.203	0.930	0.852	0.895	0.796	0.966
3	0.172	0.329	0.081	0.208	0.033	0.112	0.850	0.781	0.969	0.302	0.938	0.935	0.927	0.841	0.588

	Vulnera...	Challenge	Closeness	Curiosity	Excitement	Harmony	Ideal	Liberty	Love	Practicality	Self-expression	Stability	Structure	Conservation	Openness ...	Hedonis...	Self-enhan...	Self-transce...
1	0.697	0.128	0.089	0.143	0.241	0.557	0.057	0.309	0.584	0.617	0.103	0.104	0.093	0.019	0.959	0.940	0.757	0.339
2	0.947	0.095	0.586	0.846	0.448	0.956	0.558	0.731	0.691	0.925	0.764	0.626	0.699	0.053	0.673	0.484	0.159	0.944
3	1.000	0.201	0.392	0.127	0.119	0.986	0.096	0.250	0.775	0.989	0.036	0.049	0.030	0.011	0.676	0.074	0.899	0.420

This example demonstrates the power of combining Watson Services with IBM SPSS Modeler. A free text field was submitted and 52 scores were returned which can be used for creating a variety of models. IBM SPSS Modeler can use the results from Watson Personality Insights as additional predictors in machine learning models.

Important Links

Learn

- Learn more about SPSS software.
- To learn more about this service please visit http://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud/personality-insights.html
- Visit developerWorks Business analytics for more technical analytics resources for developers.
- The Comprehensive R Archive Network is the main site for the R project and each R package. The help pages and manuals that are associated with optimx, nlmrt, and Rcgmin are detailed. Numerous references are provided.
- Read "Do I need to learn R?" (Catherine Dalzell, developerWorks, September 2013) to learn why R is a valuable tool for data analytics that was expressly designed to reflect the way that statisticians think and work.
- "Calling R from SPSS" describes how to use R code inside IBM SPSS Modeler 16.
- Read "Create new nodes for IBM SPSS Modeler 16 using R" to learn how to create new extensions easily.

Discuss

- Visit the IBM SPSS Community to share tips and experiences with other IBM SPSS developers.
- Follow developerWorks on Twitter to be among the first to hear about new resources.