Text Analytics

Updated: July 9, 2015

Azure Machine Learning provides specialized tools for helping you work with both structured and unstructured text:

- Feature hashing helps you efficiently extract features from text without requiring additional preprocessing or advanced linguistic analysis
- The Vowpal Wabbit libraries support very fast text analytics using advanced feature hashing techniques
- Named entity recognition is provided to help you easily extract the names of people, places, and organizations from unstructured text

Examples

For examples of text analytics using Azure Machine Learning, see these sample experiments in the Model Gallery (http://gallery.azureml.net/):

- The News Categorization (http://go.microsoft.com/fwlink/?LinkId=525167) sample uses feature hashing to classify articles into a predefined list of categories.
- The Find similar companies (http://go.microsoft.com/fwlink/?LinkId=525164) sample uses the text of Wikipedia articles to categorize companies.
- In the five-part Text Classification (http://go.microsoft.com/fwlink/?LinkId=525957) sample, text from Twitter messages is used to perform sentiment analysis.

List of Modules

The Modules References.Text Analytics category includes the following modules:

Module	**Description**
Feature Hashing (https://msdn.microsoft.com/en-us/library/azure/dn906018.aspx)	Converts text data to integer-encoded features using the Vowpal Wabbit library
Named Entity Recognition (https://msdn.microsoft.com/en-	Recognizes named entities in a text column

us/library/azure/dn905955.aspx)	
Vowpal Wabbit Score (https://msdn.microsoft.com/en-us/library/azure/dn905869.aspx)	Scores data using the Vowpal Wabbit machine learning system
Vowpal Wabbit Train (https://msdn.microsoft.com/en-us/library/azure/dn905861.aspx)	Trains a model from the Vowpal Wabbit machine learning system

See Also

Machine Learning / Initialize Model / Regression (https://msdn.microsoft.com/en-us/library/azure/dn905922.aspx)
Machine Learning / Initialize Model / Classification (https://msdn.microsoft.com/en-us/library/azure/dn905808.aspx)
Machine Learning / Initialize Model / Clustering (https://msdn.microsoft.com/en-us/library/azure/dn905908.aspx)

OpenCV Library Modules (https://msdn.microsoft.com/en-us/library/azure/dn905946.aspx)
Machine Learning Module Descriptions (https://msdn.microsoft.com/en-us/library/azure/dn906013.aspx)