Automating Market Trends Detection with Machine Learning

Hieu Nguyen
MS Statistics, Colorado State University

Problem Statement

For a company's marketing team to create attractive and exciting contents, it is important to identify topics that are highly trending. The question is:

Which topics (say, in the data infrastructure and analytics industry) are highly trending within the recent one year?

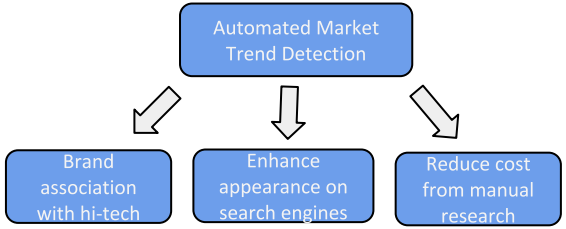


Figure 1. Values from detecting market trends.

Solution

Stage 1: Discover topics of interest relevant to the industry
Co-occurrences of words within same mention reveal their relationship.

Stage 2: Retrieve Google search topic popularity
Google search volumes, available at trends.google.com, show how many times people search for a topic over time.

Stage 3: Identify topics with consistent growth
Train a Convolutional Neural Network (CNN) to differentiate times-series with positive trend and others with stationary or declining trend.

Figure 2. Solution broken down into 3 stages.

Stage 1: Discover topics of interests

Each word is embedded as a numerical vector, based on how well it is predicted by the surrounding words.

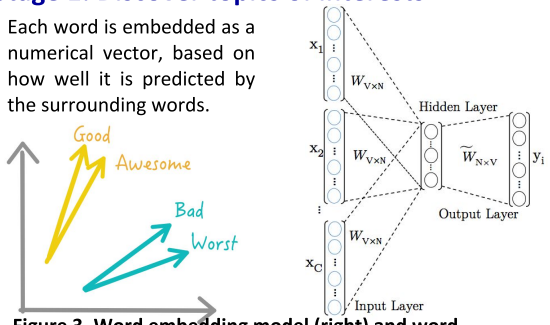


Figure 3. Word embedding model (right) and word associations based vector representation (left)).

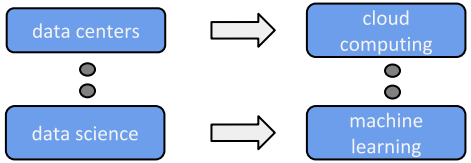


Figure 4. Words associated with data related topics.

Stage 2: Retrieve Google search volume for topics

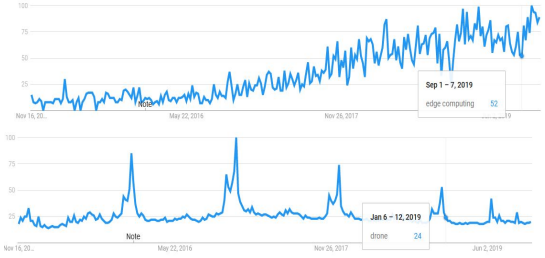


Figure 4. Google search volume time-series examples.

Stage 3: Identify topics with consistent growth

Convolutional neural network trained on synthetic data is able to discern whether Google search volume time-series has consistent growth pattern.

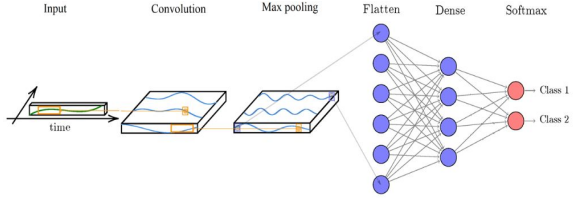
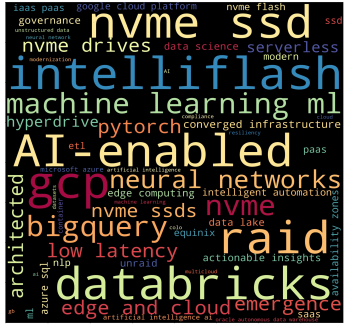


Figure 5. Convolutional neural network for time-series classification

Results

Figure 6. Discovered trending topics. Bigger topics grow faster in one-year period.



References

1. Mikolov, Tomas; Sutskever, Ilya; Chen, Kai; Corrado, Greg S.; Dean, Jeff (2013). *Distributed representations of words and phrases and their compositionality*. Advances in Neural Information Processing Systems.

2. LeCun, Yann; Léon Bottou; Yoshua Bengio; Patrick Haffner (1998). *Gradient-based learning applied to document recognition*. Proceedings of the IEEE. 86 (11): 2278–2324.