Classification of Business Categories for Small and Medium Enterprises

December 8th, 2014  Amelia Hardjasa, Energy Research Scientist
Pulse Energy joins EnerNOC

Acquired by leading global energy intelligence software provider

Read more
Segmentation matters

- Social comparison can drive behavioral change:

You spend $1,740 less on energy than average laundromats.

Your electricity and natural gas costs over the last 12 months are compared to the average energy costs of laundromats with similar characteristics.

- If customers feel that they are being compared to the appropriate peer group.

"[The report should be] about artist's studios in [city] and not an anonymous building with 'employees'"

“We have printing presses running all day it can’t be compared to a warehouse facility used for storage”

Quoted from focus group participants
But the data needed to segment are poor

- Business type coding is usually provided by third parties like D&B
- Uses industry codes: NAICS, SIC, NACE
- But there are often problems:
  - Even when right, these are not designed to group by similar energy usage
Even when right, these are not designed to group by similar energy usage

NAICS 722213: Snack and Nonalcoholic Beverage Bars

Density

\[ \ln(\text{Average Daily Usage (kWh)}) \]
We developed a method for customer categorization

- Use account/organization name fields – not perfect, but self-identification
- Start with simple keyword searching, but what do you do with these?
Develop a method for customer categorization

- Neural networks are based on the structure of connections in the brain
- Each word is received by a separate “neuron”,
- combined with other neurons in a hidden layer,
- and then a prediction and confidence value are returned

- Two big advantages
  - Take in all the information as a whole
  - Learn from mistakes
Measuring Results - Sensitivity vs. Specificity

Green line = 95%

Accuracy

Percent with Prediction

Accuracy

Percent with Prediction

Green line = 95%
Results and Future work

- Used to add over 5000 customers with no known primary use to a program population
- Now using to revise NAICS codes in provided data

Future work:
- Combining usage and text
- Using word order instead of “bag of words”
- Using traits rather than categories
Amelia Hardjasa
Energy Research Scientist
778-331-0500
amelia.hardjasa@pulseenergy.com
www.pulseenergy.com

Thanks!
87% report that offering targeted solutions for their business needs is important or very important

* Source: Accenture New Energy Handbook 2013