The Art of Natural Language Interpretation

William Meisel

President, TMA Associates

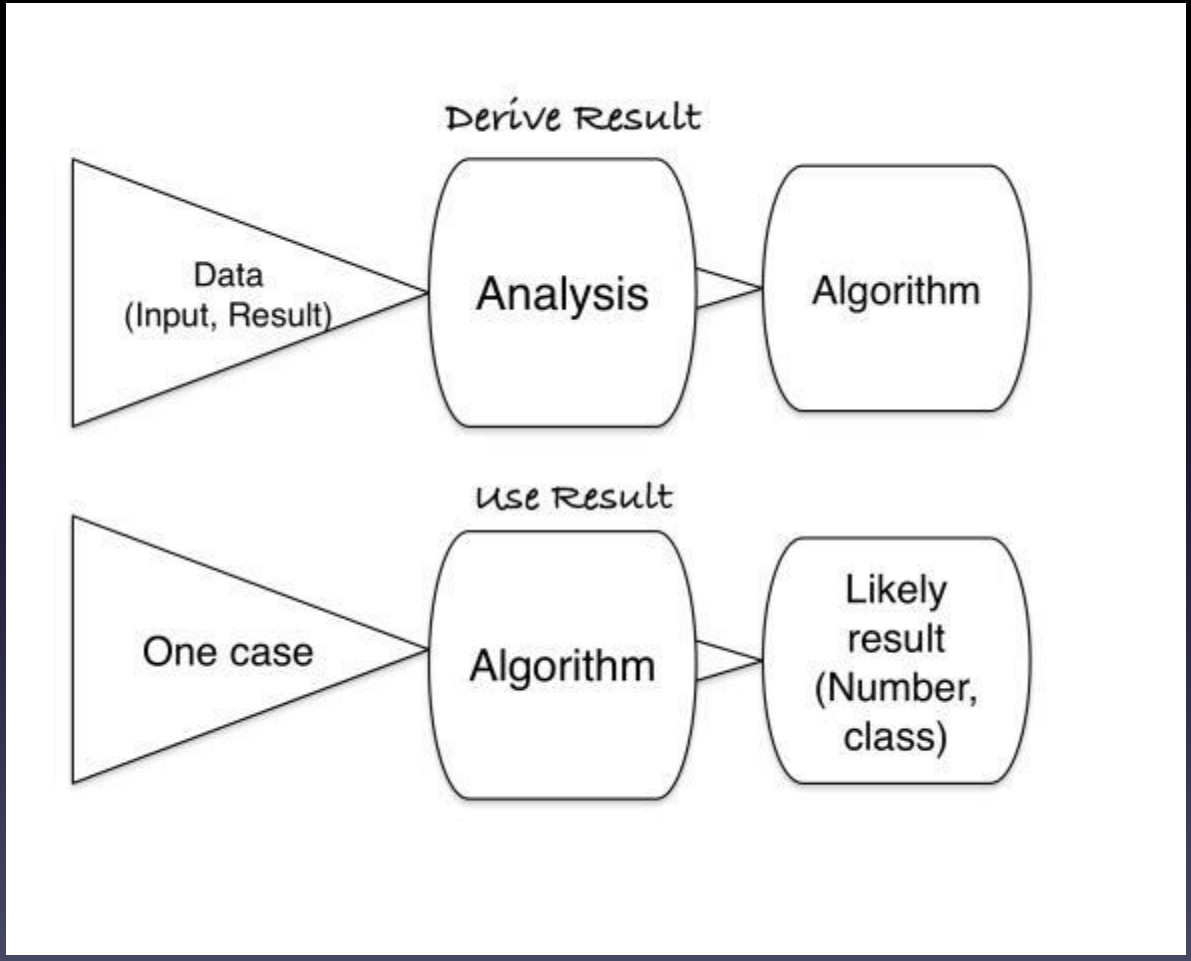
Natural Language Interpretation

- Keywords
  - Semantics
- Statistical (Machine Learning)
  - Data-driven

Keywords/Semantics

- Text from chat or speech-to-text
- Search for words or phrases that indicate nature of request
  - Semantic equivalents ["account balance", ("account" or "money") + "how much"]
  - Structures (e.g., dates, times)
  - Lists: E.g., travel destinations, songs
- **Lists and structures may be required for the speech-to-text processing in the language model**
  - **And the speech recognition language model can pass them on with a label**

Natural Language Processing (NLP) using machine learning



Machine learning for classification

- To predict “Class”
  - Class A: “order status”
  - Class B: “payment”
  - Class C: “requires agent”

Create machine learning function for each class separately

- For class A: F_A
  - Output is 1 if input values correspond to class A
  - 0 (zero) otherwise
- Output of F_A can be interpreted as the probability of Class A given specific input variables

Result

- For a given input, get F_A F_B F_C
- The highest output value (probability) is the most likely class given the specific input values
  - And you have a measure of the likelihood of one class versus another

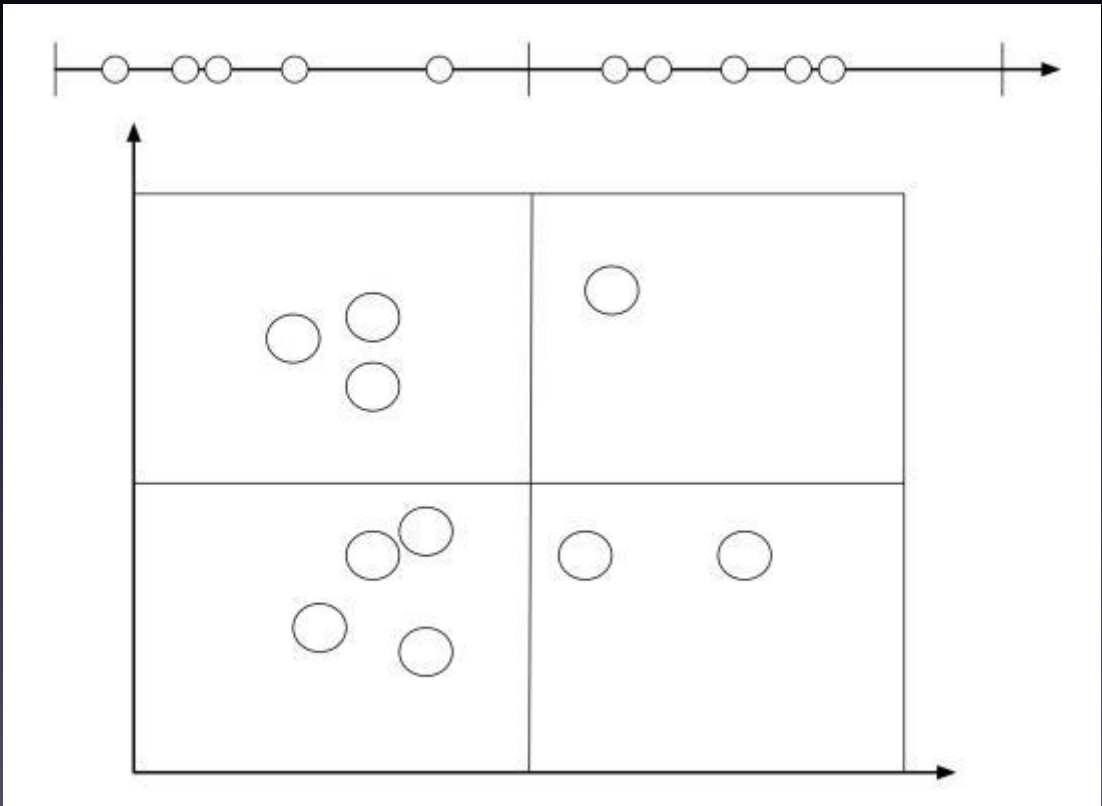
It's not automatic

- Cleaning the data
  - E.g., see Microsoft video tutorial "**Preprocessing Data in Azure Machine Learning Studio"**

But more fundamentally…

The "curse of dimensionality":

Describing the data with too many input variables



Defining input variables

- Be parsimonious
  - Choose features carefully
  - E.g., *presence* of words or phrases suggesting an "account balance" request may be one input variable
  - The *absence* of certain words can also be revealing

Defining input variables

- Be parsimonious
  - Choose features carefully
  - E.g., *presence* of words or phrases suggesting an "account balance" request may be one input variable
  - The *absence* of certain words can also be revealing
- Use your understanding of the problem to reduce the input space

Output classes

- What answers can you provide?
  - **A good place to start your analysis of the problem**
- What are you predicting (value or class)
  - Classes (categories predicted): "Account info," "equity line of credit prospect," "payment"…
  - Insight into feature definition

Output classes

- What answers can you provide?
  - **A good place to start your analysis of the problem**
- What are you predicting (value or class)
  - Classes (categories predicted): "Account info," "equity line of credit prospect," "payment"…
  - Insight into feature definition
- Number of classes also effects density of data
  - Combine classes if simple post-processing can separate similar classes
  - "Requires agent" might be a class

Model complexity

- More parameters in the model allow it to represent more complex relationships
  - Deep neural networks
- But also require more data to find the optimal parameters
  - And more power to run the final model in realtime

And can lead to suboptimal results

- Can lead to other than the best solution
  - Highest possible accuracy: The highest peak

Blind use of data-driven methods can lead to suboptimal results

- Human intelligence truly *understands* the application

You supply the art!