Probabilistic Scoring

Update

Client: Alan Malik-Patient Tools

Lee Panter

Final Aims-Compare Accuracy

Problem

- Data has no measure of “truth”
- QDP does not measure classic depression risk (11 more specific traits)
- How do we define the accuracy of a method for which we have no way of determining truth in classification?

Approach #1

- Information from original data source:
  - Traditional classification is 88% accurate
  - 94% symmetrically sensitive and specific
- Use this information to reverse engineer the supervision set
- Introduce 12% independently sampled classification error into the traditional classification outcome
  - Higher probability of mis-classification in traditional scores that are closer to threshold areas
  - Allow for multiple mis-classification directions where appropriate (i.e. in the middle of the middle region)
- Perform accuracy analysis on new variable with error

Approach #2

- Information from original data source:
  - C1, C2, and C3 threshold values chosen to minimize traditional sum classification error compared when compared to THE ACTUAL TRUTH
- Probabilistic Scoring algorithm derives information on how subject answer sequences uniquely identify traditional sum classification
  - Higher correlations --> higher weight & (should) exhibit stability across training data variations
  - Lower correlations --> lower weight value, use the average across training sets as a representation of their value
- Use average weight value across training data sets to create new data set
- Perform accuracy analysis on new variable

Compare accuracy values generated by each approach

Completed and Outstanding-Just the Code

Generalization level	Data Processing	Weight Calc.	Prob. Score Sequence Calc.	Prob. Score Analysis	Accuracy Analysis vs Sup. Var. #1	Weight Convergence Analysis	Accuracy Analysis vs Sup. Var. #1
Subject	😀	😀	😀	• Did it converge? • How many Questions? • With what Probability? • What class did it converge to?	• Supervision variable already integrated • Accuracy calculations *should* be easy to create	• Average individual weight for each CV data set • Looking for convergence as training data size increased	Same
Arbitrary Single Train/Test data pairings	😀	😀	😀				
Arbitrary partitions of Full data into Test/Train data pairings	😀	😀	😀				

Generalization

INFERENCE

Coding is finished

Coding not finished but should be soon

Coding not finished, will need until end of Spring Break to finish

Completed & Outstanding-Other

- Completed
  - Replication of weight results from presentation
  - Been in communication with client regarding:
    - Goals
    - Desired outcomes
    - Deadlines and feasability
- Incomplete
  - Paper (all)
  - Analysis (most)
  - Methods Presentation (all)
  - Developing visualizations using methods presentation (hopefully)
  - …

What do I need to be Successful?

- I need to know if either/both/neither of these approaches are appropriate, with the understanding that:
  - Neither are sufficient for answering the clients proposed question
  - Both approaches are analyzing separate and essentially unrelated topics (if they are analyzing anything at all).

- No?
  - Is there time to re-examine the problem? Should I?
  - Do you have any suggestions?

- Yes?
  - Then the rest *should* be manageable:
    - Approximately 24-36 more hours programming (glitches & compiling time not included)
    - Methods paper on Self Organizing Maps, use SOM to develop visual tool(s) for Alan
    - Project paper
  - Possible Conflict(s) with MS presentation

Thank You

Project GitHub: https://github.com/leepanter/ProbabilisticScoring.git