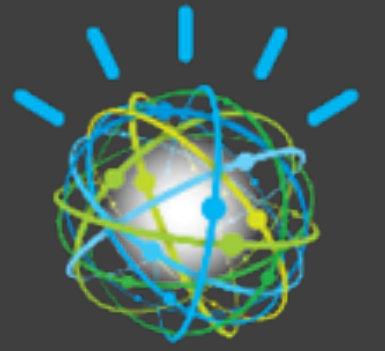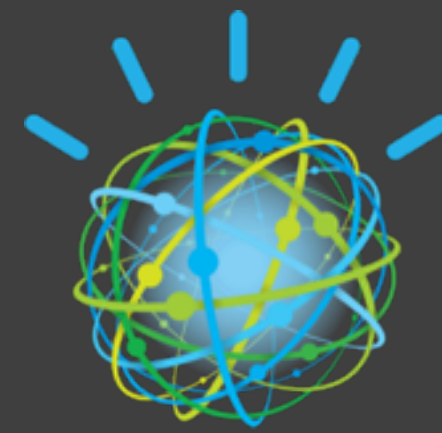# Keeping Students in School:
Harnessing Watson to Improve Dropout Predictions

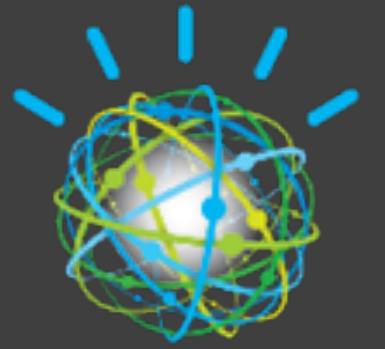Glenn J. Hansen, University of Oklahoma
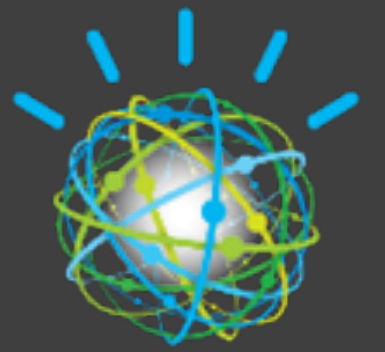Raghuram Srinivas, IBM

Summary

- Assisting each student to be successful is a primary focus for OU

- Success measurement: 1st year retention increased from 86.1% (2014 cohort) to 90.4% (2015 cohort).

- Goal: Use data analytics & cognitive technology to provide personalized support & services to the correct students.

Motivation

- Retention & graduation rates are the pulse of the modern university.

- We have an interest in investing in our algorithms; therefore, we are careful about using branded black retention boxes.

- The success of our analytic efforts can be measured in the economic value of our algorithms.

- Working with IBM allows us to better participate in the algorithm economy.

Our Work

- Make predictions <u>immediately after admission</u>:
  - Increase first-year retention
  - Predict likelihood of making next payment

- The only inputs available:
  - Admission application
  - Historical performance of HS
  - Home zip code info

- Important: the predictions will:
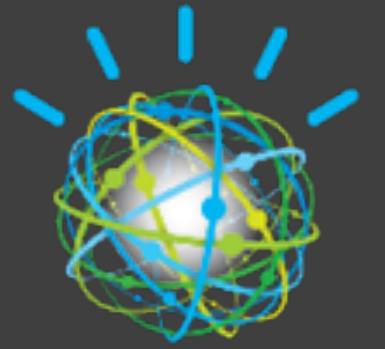 - Help get students appropriate assistance
  - NOT to make admission decisions
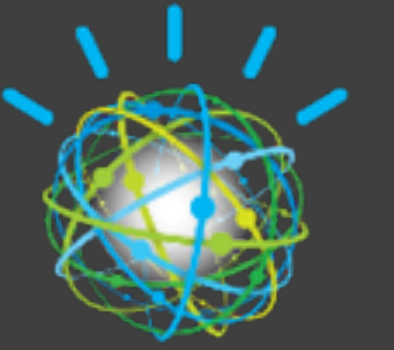
Using Watson to Improve Predictions

Exploration, Algorithms & Unstructured Data

- We work with different estimation algorithms and utilize the one that predicts most accurately out of sample.

- Utilized Watson services to glean new information from untapped, unstructured data: 34,000+ admission essays.

- Example essay topic: Reflect on a time when you challenged a belief or idea. What prompted you to act? Would you make the same decision again?

- Tone Analyzer, Personality Insights, Alchemy Language

Watson Services

Tone Analyzer

Uses linguistic analysis to detect emotional tones, social propensities, and writing styles in written communication and provides user with three tones - Emotional, Social and Language Style
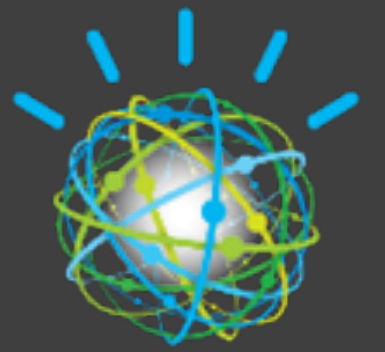
Personality Insights

Enables deeper understanding of people's personality characteristics, needs, and values to help engage users on their own terms

Alchemy Language

Offers text analysis through natural language processing to understand the content and context of text in webpages, news articles, and blogs
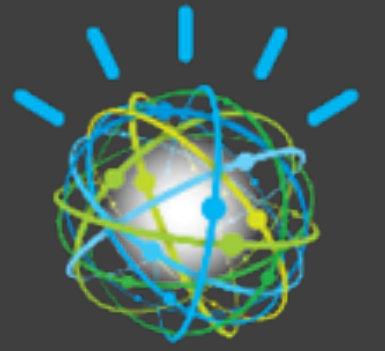
Watson on Sadness

- Tone analysis is less about analyzing how someone else feels, and more about analyzing how you are coming across to others

- Examples from student essays:
  - It still hurts sometimes knowing they won't be there to see me at special events in my life (2013 student essay).
  - Unfortunately, my team and I did not play well in our last postseason game (2014 student essay).
  - She passed away eventually, and I had barely spent any time with her even though I knew how much she wanted me to (2015 student essay).
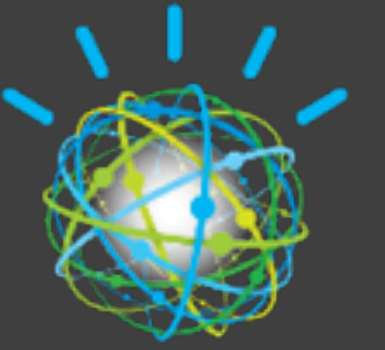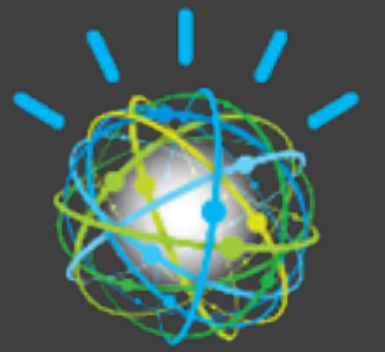
Results from work

Results

- For both outcomes structured features (e.g., HS GPA) are strong predictors.

- However, Watson derived features made a powerful difference.

- For example, where Tone Analyzer detected higher levels of "sadness," students were less likely to retain & make a payment on time.

- Sadness was the 6th (retention) and 7th (payment) best predictor out of 145.
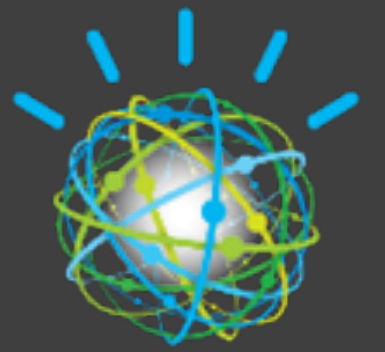
Results (cont.)

- Watson generated:
  - 13 of the top 20 predictors for payment model.
  - 14 of the top 20 predictors for retention model.

- Best features include:
  - Analytical (Personality Insights)
  - Disgust (Tone Analyzer)
  - Emotional Range (Tone Analyzer)
  - Conscientiousness (Personality Insights)

Tone Analyzer

Personality Insights

Next Steps

- Work with IBM to build production versions of both outcomes in the IBM Bluemix environment.

- Using Watson-based algorithm, create a "student financial health score" that:
  - Serves as an internal credit-like score
  - Identifies students for intervention (advisor, financial aid, etc.)
  - Focuses on likelihood of first year students to make tuition payments
  - Updates at key points along student journey, beginning at admission