COGNITELL
ENABLING DATA INTELLIGENCE

Retention Modeling

and Prescribed Interventions

CAIR 2020

Who We Are

Matt Swaffer

Technical architect with a background in Ed Psych and experience designing systems that extract information and knowledge from data enabling intelligent reasoning.

Mike Barber

Data scientist specializing in bringing machine learning models into production. He currently has models in production today at companies like StubHub, Paypal, eBay, and Cognitell.

What do we do?

Data Science

We provide data engineering and data science services specific to IR needs.

Data Wrangling

We provide ETL and data warehousing expertise to support advanced modeling.

Higher Education

We have a focus on Higher Education and understand the data and the environment.

Background of Wyoming Work

- Wyoming Community College Commission
- 7 Community Colleges
  - Many IR offices have 1-2 full time employees
  - Retention Modeling attempts in past failed
- Work performed
  - Initiated & Facilitate State-wide IR Research group
  - Provide Data Science Training specific to Institutional Researchers
  - Project Collaboration
  - Design, Develop, and Prototype Data Pipelines
  - LMS Data Value Analyses
  - Feature Engineering
  - Machine Learning Model Development

Preliminary Results

- LMS engineered features indicative of Student Engagement
  - Almost half of highest correlated combined features were LMS
  - Important LMS features
    - LMS Routine – frequency/how long/when
    - LMS Assignment and Quiz Activity
    - LMS Discussion Topic Participation

- Model Improvement of LMS data
  - Mean model accuracy increased between 10%-20%
  - False Positives (Type I error) reduced by as much as 15%

Self Report Survey Data

Gaining early insight into student behavior indicators

Challenges with existing intake assessments

- Expensive to administer
- Not customized
- Might not work for a small or rural student population
- Colleague SIS hard to integrate with captured data

1200 6th Avenue, Suite 800, Seattle, WA 98101 | (719) 314-3400 | www.cognitell.com

Background of CWC Work

- Worked with IE and Student Success to develop intake survey
- Iterated over scales and items to tailor to the population
- Incorporated items related to previously collected drop / withdrawal

Central Wyoming College

Psychosocial & Skills Factors (PSF)

Motivation and Skills	• Achievement motivation • Academic goals • Academic skills
Self-Regulation	• Academic self-efficacy • General self concept
Social Engagement	• Perceived social support • Social Involvement • Institutional Commitment

(Le et. al. 2005, Robbins et. al., 2004,2009)

Psychosocial & Skills Factors (PSF)

Relevant to Goals & Population

Motivation and Skills	• Achievement motivation • Academic goals • Academic skills
Self-Regulation	• Academic self-efficacy
Social Engagement	• Perceived social support

(Le et. al. 2005, Robbins et. al., 2004,2009)

Population Specific Constructs

Drop Withdrawal Reasons

- Gathered from CWC reports

COVID-19

- Specific issues for CWC students

Uses of Data

- Coaching of individual students
- Identifying trends
- Predictive analytics

Combining Disparate Data Sources

Using LMS, SIS, Survey and other data for modeling

Importance of Combining Disparate Data Sources

Importance of Combining Disparate Data Sources

- More Context is ALWAYS a good thing
- More Context allows for new Features, or different Features

Importance of Combining Disparate Data Sources

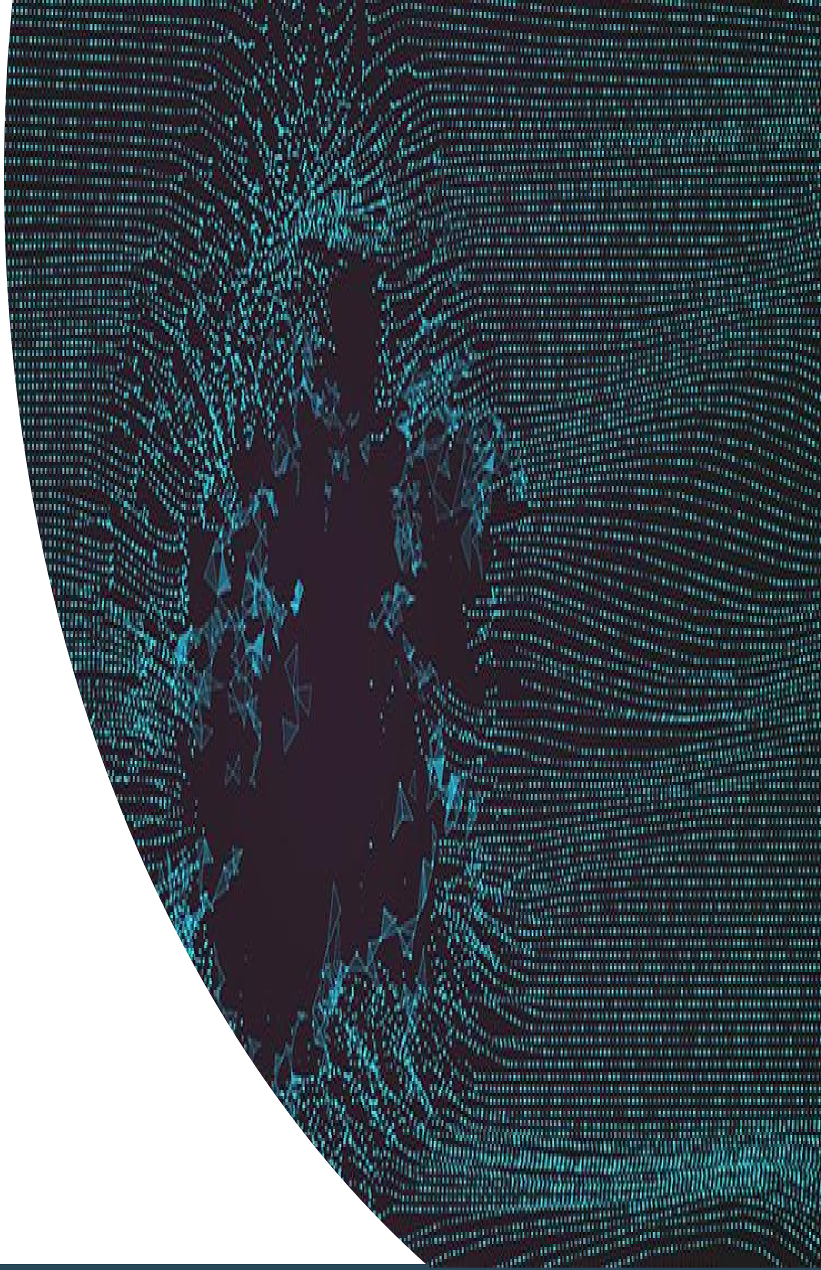
- More Context is ALWAYS a good thing
- More Context allows for new Features, or different Features
- More Data is not necessarily a good thing – unless it contains new context

Importance of Combining Disparate Data Sources

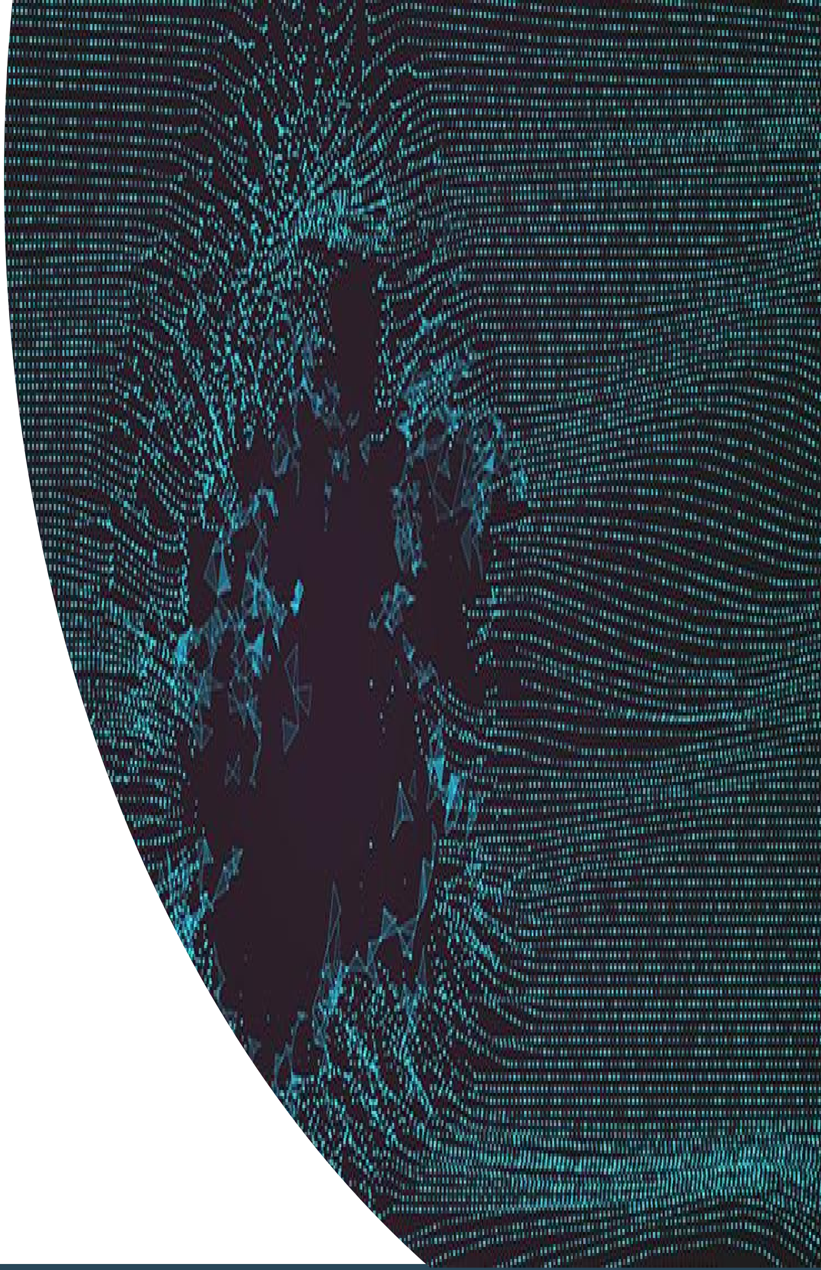
- Feature Shift is Real
- Our World is constantly changing around us -> our models should too!
- Education has schedules [semesters] that require model retraining & retuning

How to Combine Disparate Data Sources

- Design, Develop, and Prototype Data Pipelines
- Perform A LOT of statistical experiments
- Apply Feature Engineering Continuously
- Employ Feature Selection & Evaluation Techniques



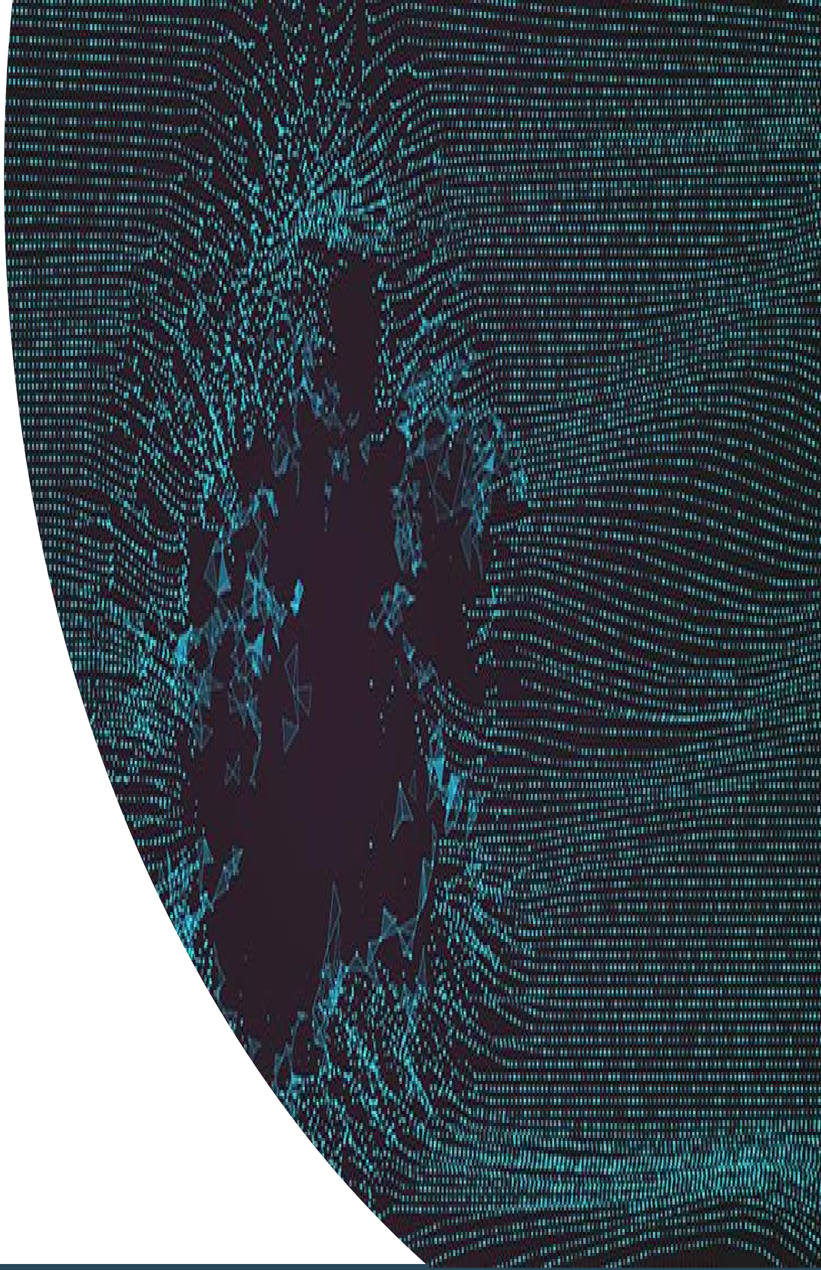
Feature Selection & Evaluation



Feature selection is primarily focused on removing non-informative or redundant predictors from the model.

— Page 488, Applied Predictive Modeling, 2013.

Feature Selection & Evaluation

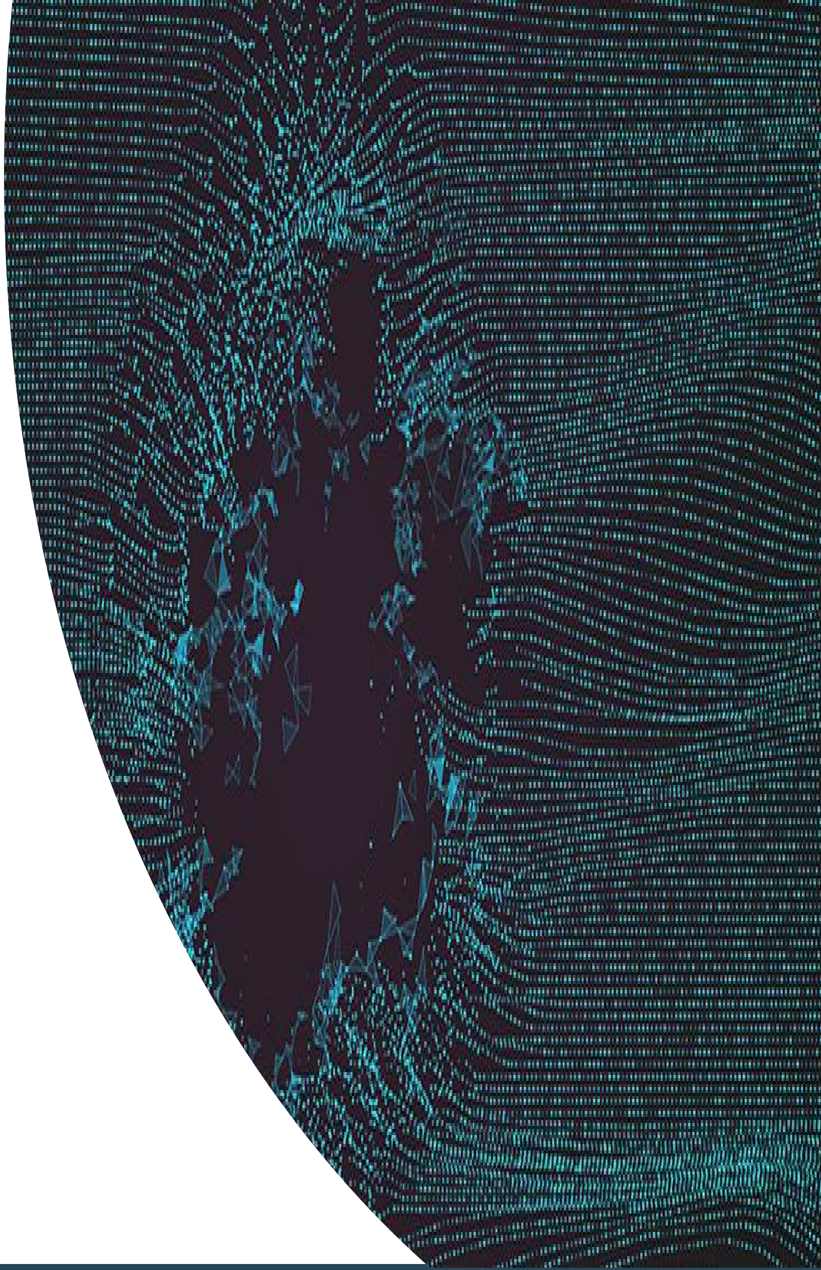


Many models, especially those based on regression slopes and intercepts, will estimate parameters for every term in the model. Because of this, the presence of non-informative variables can add uncertainty to the predictions and reduce the overall effectiveness of the model.

— Page 488, Applied Predictive Modeling, 2013.

Feature Selection & Evaluation Techniques

- Feature Selection
  - Unsupervised
  - Supervised
    - Intrinsic Algo
    - Wrapper Methods
    - Filter Methods
- Dimension Redux



Learnings From Our Approach

- Feature Selection & Evaluations:
    - Reduce Model Overfitting
    - Improves Model Accuracy
    - Reduces Model Training Time

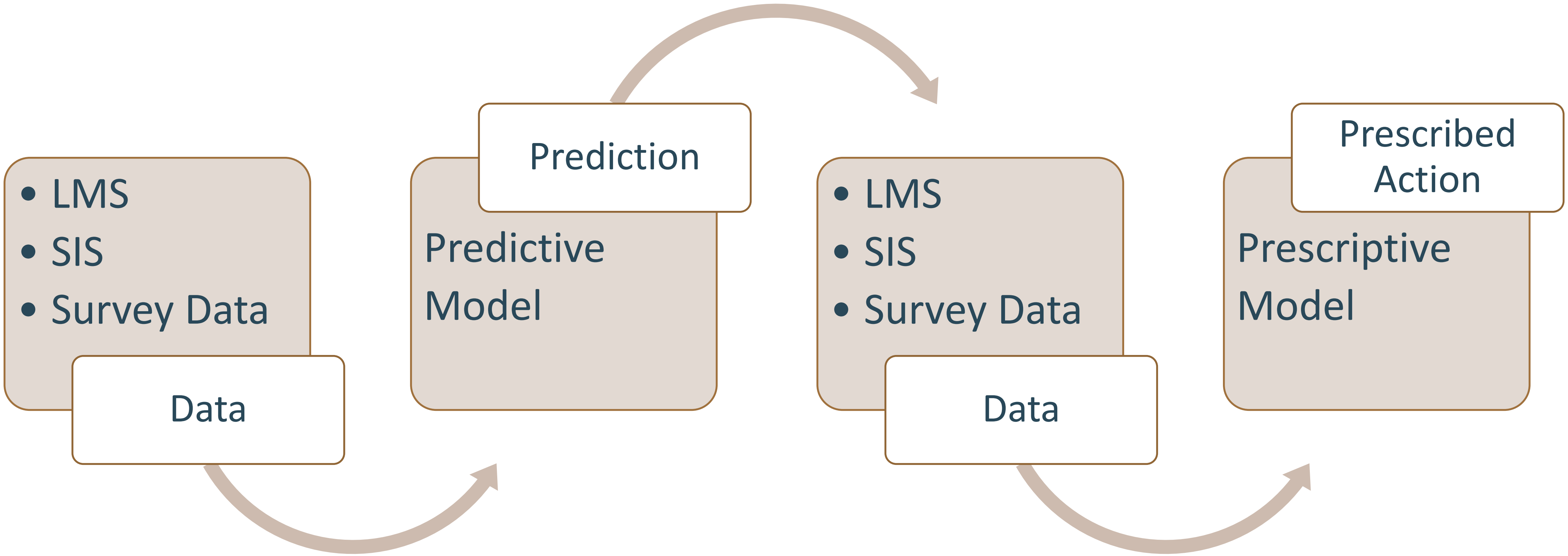
- There is no Best Feature Selection & Evaluation Technique.
- Make sure you apply the appropriate statistical test to the appropriate data type.
- Feature Selection must be an active part of the modeling process.
- Features should be routinely evaluated for model contribution.

COGNITELL
ENABLING DATA INTELLIGENCE
1200 6th Avenue, Suite 800, Seattle, WA 98101 | (719) 314-3400 | www.cognitell.com

From Prediction to Prescription

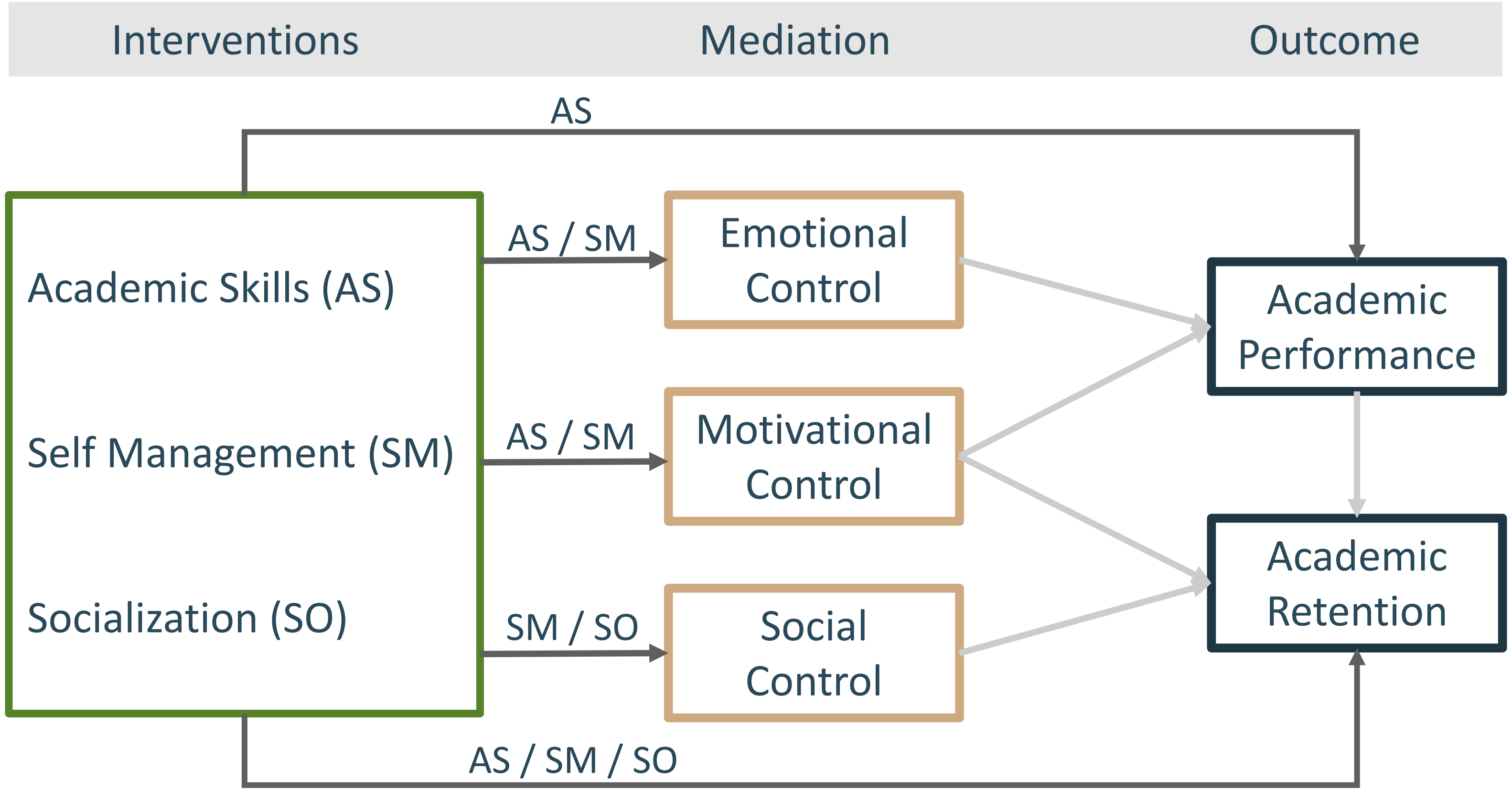
A theoretical model for prescribing interventions based on student behavior

From Prediction to Prescription



COGNITELL ENABLING DATA INTELLIGENCE
1200 6th Avenue, Suite 800, Seattle, WA 98101 | (719) 314-3400 | www.cognitell.com

The Theoretical Model



(Robbins et. al., 2009)

Intervention Categories

Academic Skills

- Task-related (i.e., underlining, note-taking etc.)
- Mnemonic training
- Minimize learning pathologies (i.e., anxiety, surface learning)
- Reading skills

Self-management

- Planning, implementing, and monitoring
- Time Management
- Motivation and self-concept
- General counseling

Social

- Learning communities
- Service Learning
- Social media
- Diversity / inclusivity programs

(Gray et al., 2013; Hatch, 2017; Hattie et al., 1996; Rigali-Oiler & Kurpius, 2013; Robbins et al., 2009; Ye, 2009)

Psychosocial & Skills Factors (PSF)

Emotional Control (Academic skills)	• Study skills • Problem solving skills • Communication skills • Emotional control skills (i.e., test anxiety)
Motivation Control	• Goal focus (academic goals) • Academic self-confidence (self-efficacy) • Conscientiousness
Social Control (Engagement)	• Sociability (e.g., perceived social support) • Social connection (involvement) • Teamwork

(Le et. al. 2005, Robbins et. al., 2009)

Behavior Related to PSFs

Study Skills

Time management				Preparation for Exams		Class Notes
Time between due date and submission	Time of day using LMS	Frequency and consistency of LMS use	Low long before a quiz to access study resources	Use of practice quizzes	Use of study resources	Download of PowerPoint

Behavior Related to PSFs

Motivation - Conscientious

Achievement Motivation		Goal Focus		Academic Self-confidence
Significant change for the worse in GPA	Failing a core class	Placement test performance	Performance in core courses	Has a major been declared?

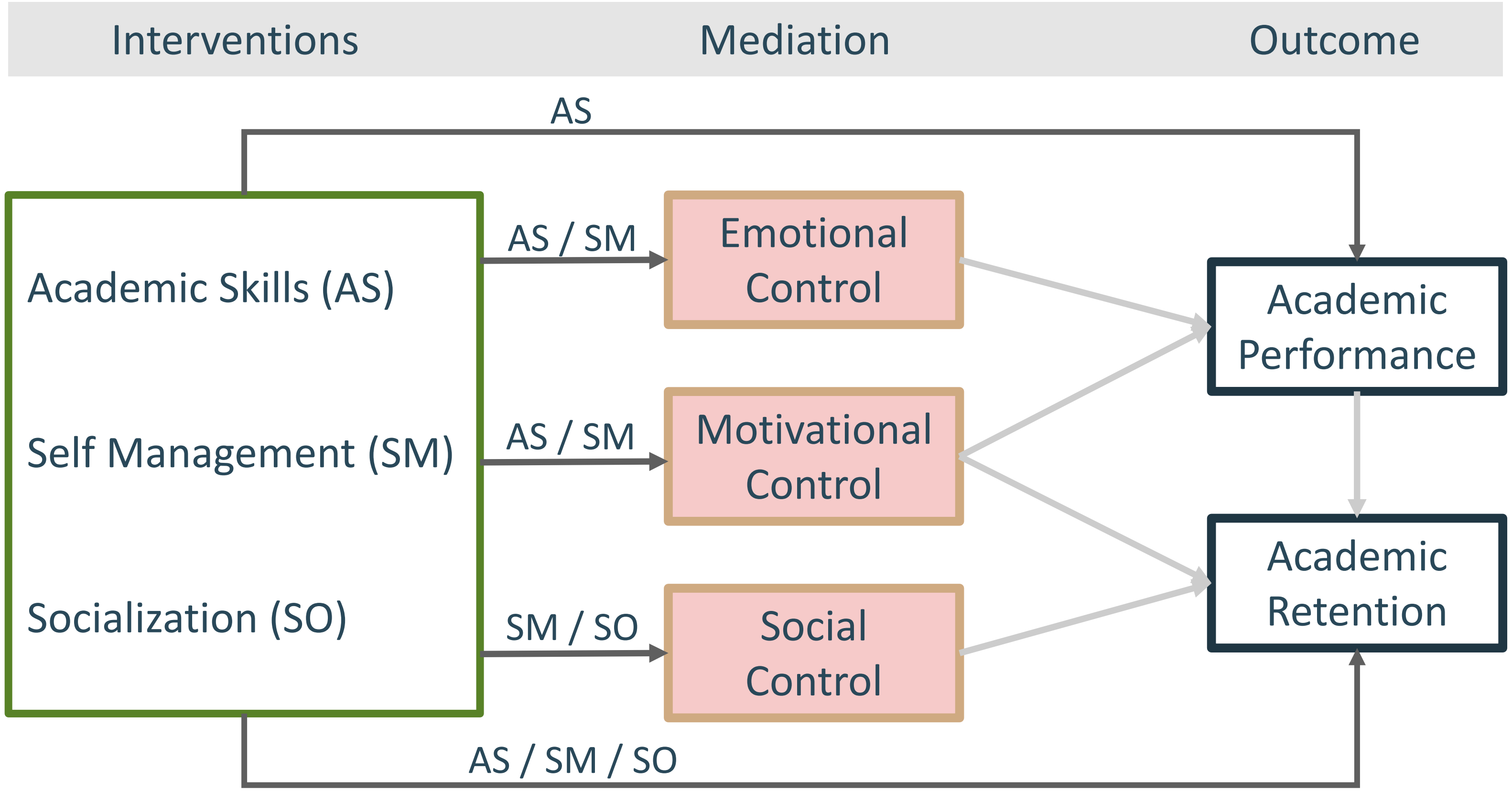
1200 6th Avenue, Suite 800, Seattle, WA 98101 | (719) 314-3400 | www.cognitell.com

Behavior Related to PSFs

Social Engagement

- **Perceived Social Support**
  - SES differences
  - Financial aid
  - 1st gen / FTFT
- **Social Connection**
  - Social media engagement
  - Extracurricular activities
- **Teamwork**
  - Group work engagement
  - Group work success

Revisit the Theoretical Model



(Robbins et. al., 2009)

COGNITELL
ENABLING DATA INTELLIGENCE

Ways we can help

Survey customization, data management, data analysis

Ways To Get In Touch

- Visit our CAIR 2020 conference page at https://www.cognitell.com/cair-2020
  - Here, you can download the slides, along with our brochures, and watch the presentation.
  - You may also use the calendar on the page to set up a time to meet with us to further discuss how we may help you.
- For any questions please feel free to email us at info@cognitell.com
- Follow us on LinkedIn here https://www.linkedin.com/company/cognitell

References

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312

Fernández, A. R., González, F. S., Merino, P. J. M., & Kloos, C. D. (n.d.). *A Data Collection Experience with Canvas LMS as a Learning Platform*. 15.

Gray, R., Vitak, J., Easton, E. W., & Ellison, N. B. (2013). Examining social adjustment to college in the age of social media: Factors influencing successful transitions and persistence. *Computers & Education*, *67*, 193–207. https://doi.org/10.1016/j.compedu.2013.02.021

Hatch, D. K. (2017). The Structure of Student Engagement in Community College Student Success Programs: A Quantitative Activity Systems Analysis: *AERA Open*. https://doi.org/10.1177/2332858417732744

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of Learning Skills Interventions on Student Learning: A Meta-Analysis. *Review of Educational Research*, *66*(2), 99–136. JSTOR. https://doi.org/10.2307/1170605

Krumm, A. E., Waddington, R. J., Teasley, S. D., & Lonn, S. (2014). A Learning Management System-Based Early Warning System for Academic Advising in Undergraduate Engineering. In J. A. Larusson & B. White (Eds.), *Learning Analytics: From Research to Practice* (pp. 103–119). Springer. https://doi.org/10.1007/978-1-4614-3305-7_6

Le, H., Casillas, A., Robbins, S. B., & Langley, R. (2005). Motivational and Skills, Social, and Self-Management Predictors of College Outcomes: Constructing the Student Readiness Inventory. *Educational and Psychological Measurement*, *65*(3), 482–508. https://doi.org/10.1177/0013164404272493

References

Morris, L., Finnegan, C., & Wu, S.-S. (2005). Tracking Student Behavior, Persistence, and Achievement in Online Courses. *The Internet and Higher Education*, *8*, 221–231. https://doi.org/10.1016/j.iheduc.2005.06.009

Robbins, S. B., Oh, I.-S., Le, H., & Button, C. (2009). Intervention effects on college performance and retention as mediated by motivational, emotional, and social control factors: Integrated meta-analytic path analyses. *Journal of Applied Psychology*, *94*(5), 1163–1184. https://doi.org/10.1037/a0015738

School of Humanities and Digital Sciences, Tilburg University, Netherlands, Shayan, P., & Zaanen, M. van. (2019). Predicting Student Performance from Their Behavior in Learning Management Systems. *International Journal of Information and Education Technology*, *9*(5), 337–341. https://doi.org/10.18178/ijiet.2019.9.5.1223

Smith, V. C., Lange, A., & Huston, D. R. (2012). *Predictive Modeling to Forecast Student Outcomes and Drive Effective Interventions in Online Community College Courses*. https://doi.org/10.24059/olj.v16i3.275

Tellakat, M., Boyd, R. L., & Pennebaker, J. W. (2019). How do online learners study? The psychometrics of students' clicking patterns in online courses. *PLOS ONE*, *14*(3), e0213863. https://doi.org/10.1371/journal.pone.0213863

Yeh, T. L. (2010). Service-Learning and Persistence of Low-Income, First-Generation College Students: An Exploratory Study. *Michigan Journal of Community Service Learning*, *16*(2), 50–65.

Yu, T., & Jo, I.-H. (2014). *Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education*. 269–270. https://doi.org/10.1145/2567574.2567594