LEARNING ANALYTICS PROCESS MODEL

START HERE

1. What **question** do we want the data to answer?
2. Do we have the **data** to answer this question? *No → Instrument to collect data.*
3. Translate into data questions.

Real-world question: Which students need intervention?
Data questions: How do I predict the probability? What's the cutoff for intervention?

	PREDICT	INFER	MINE
ANSWER THE QUESTION	How can I accurately predict new data points (e.g. learner outcomes)?	What meaning can be inferred from the data?	How can I isolate the most useful information from a large data set?
GOALS	A model to predict a single aspect of the data (value or category) with high accuracy and low error	An estimate of association between an outcome variable and predictor variables.	Discover features, patterns, correlations, or anomalies of a data set useful for decision making or further analysis
EXAMPLES	What's the probability a learner will successfully complete this course? Should we classify the learner as "likely" or "unlikely" to complete?	Do the data indicate that a learner is bored or frustrated? Do the data indicate that the learning activity results in learning?	What features of the data are most useful for creating a predictive model? What interactions between learners engaged in collaborative learning are most productive?

CHOOSE METHOD AND DATA SET

Predictive Modeling ← Inference Modeling ← FEATURE MINING *informs modeling* — Mining

CREATE A MODEL USING DATA

Regression Prediction	Classification	Knowledge Inference	Behavior Detection
Methods, e.g. • Linear regression • Support-vector machines	**Methods**, e.g. • Step regression • Logistic regression • Decision trees	**Methods**, e.g. • BKT +variants • PFA • IRT • DKT	**Methods**, e.g. • Carelessness detection algorithms

Preprocessing
- Analyze variables
- Data reduction

Data Mine
Rules from training discovery
- Anomalies
- Association rules
- Clustering
- Classifying
- Regression (model building)
- Summarizing

in a new data set

Test
with 3rd data set

CHECK MODEL GOODNESS — Use different data to **test the model** than to create or train it

CONFIDENCE

Regression Prediction	Classification	Knowledge Inference	Behavior Detection
examples: • Linear correlation • MAD • RMSE • AIC	examples: • Accuracy • Kappa • ROC • AUC ROC	Check model based on how well it predicts performance	Check model based on how well it predicts behavior

VALIDITY

Generalizability, Cross-learner validity, Ecological validity, Construct validity, Predictive validity, Substantive validity, Content validity, Conclusion validity, Cross validation methods (e.g. K-fold)

VISUALIZE (SOMETIMES)

	e.g., Learning Curves	e.g., In-Experience Visual Cues