Students' attention is all you need to transform complexity to clarity

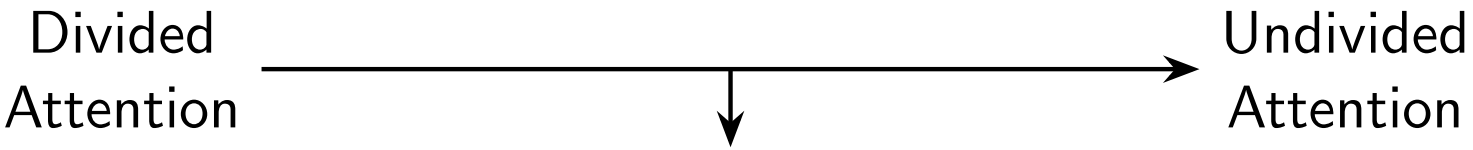
Prince Afriyie
School of Data Science
University of Virginia

July 19, 2025

Overview

Learning begins when we have our students' attention

From Divided to Undivided Attention



1. Motivation: Through Data And Useful Models
   Piqued Interest
   ↓
   Engagement
   ↓
   Enhanced Understanding
   ↓
   Captivated Attention
   ↓
   Deep Understanding and Inspiration

2. Intangibles: Being Human

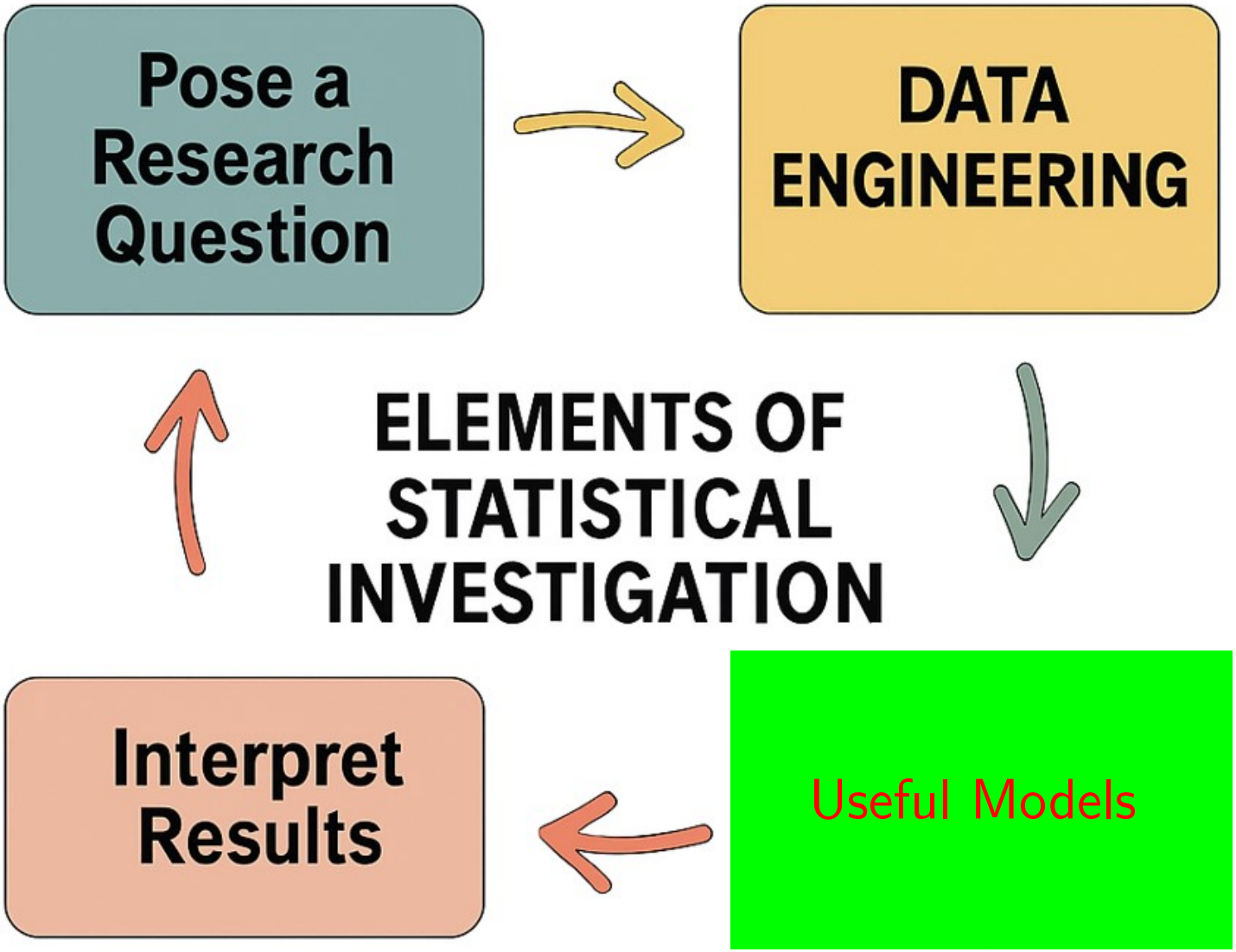
Keller's ARCS Model of Motivational Design

Component	Description
Attention	Capture learners' interest and curiosity.
Relevance	Link learning to their needs and goals.
Confidence	Build belief in their ability to succeed.
Satisfaction	Ensure they feel rewarded and fulfilled.

Reference: Keller, J. M. (1987). Development and use of the ARCS model of motivational design. Journal of Instructional Development, 10(3), 2–10.

Fundamentals

The Journey of Data



https://chatgpt.com/

Goals for a useful model

1. Usefulness in the future
2. Fast/Efficient
3. High Accuracy
4. Interpretable and explainable
5. Generalizable
6. Guaranteeable

`https://databookuw.com/`

Machine Learning Pipeline

Goal of Prediction

↓

Model Selection

↓

Loss Function Selection

↓

Model Training

Train/Test split, Resampling (Bootstrap/CV)

↓

Model Evaluation

Aggregate performance, Combine across splits

↓

Model Refinement

Tune hyperparameters

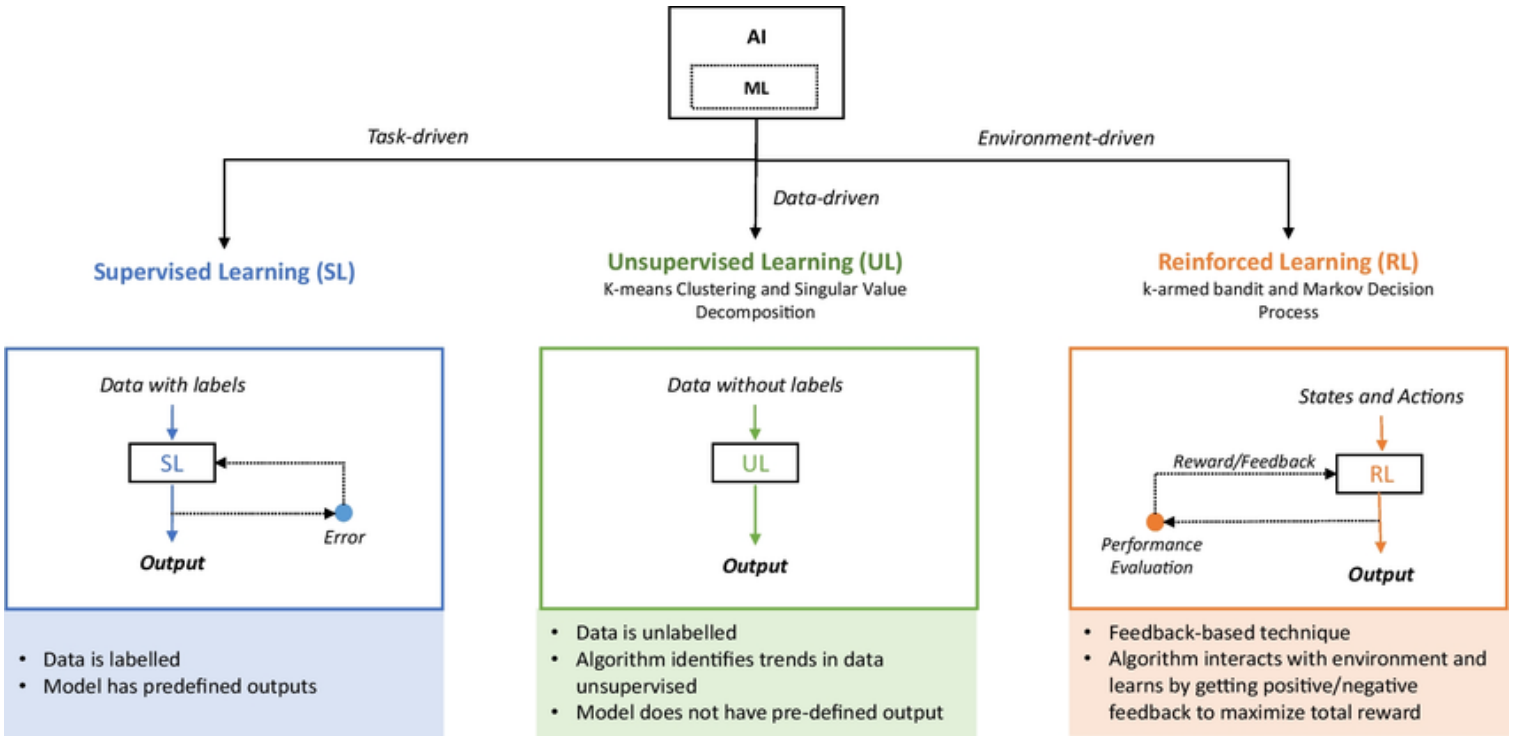
↓

Fit Final Model

How do you select a useful model?

Depends on your goals:

- Data-driven models: Unsupervised Learning
- Task-driven models: Supervised Learning
- Environment-driven models: Reinforcement Learning



[1] Hunter et al. World Journal of Emergency Surgery (2023) 18:16.

Supervised Learning Models

Model Type	**Target Variable Type**	**Examples**
Multiple Regression	Continuous numerical	Linear Regression, Ridge, Lasso, ElasticNet, SVR
Logistic Regression	Binary or multiclass categorical	Binary Logistic, Multinomial Logistic, Regularized Logistic Regression
Classification	Categorical (binary or multiclass)	Decision Trees, Random Forest, KNN, SVM, Naive Bayes, XGBoost, Neural Networks

Unsupervised Learning Models

Model Type	**Target Variable Type**	**Examples**
Clustering	No predefined target variable	K-Means, Hierarchical Clustering, DBSCAN, Gaussian Mixture Models
Dimensionality Reduction	No predefined target variable	PCA, t-SNE, UMAP, Factor Analysis
Association Rule Learning	No predefined target variable	Apriori, Eclat

Reinforcement Learning Models

Model Type	**Target Variable Type**	**Examples**
Policy/Value Based Models	Reward signal (scalar feedback)	Q-Learning, SARSA, Deep Q-Networks (DQN)
Actor-Critic Models	Reward signal (scalar feedback)	A2C, DDPG, PPO

Case studies of essential toolkit of useful models

Example 1: Captivating Students' Attention through Modeling with Their Own Data

Leveraging student data to capture their attention on useful models

Data source: Data from my students

- Computer type
- Hours of sleep, hours on phone, and hours spent outside in a typical day
- Favorite sport to watch on TV
- Age in months
- Age difference between parents (father's age - mother's age)
- Favorite music genre
- In one word, summarize their summer (or winter) break

Leveraging student data to capture their attention on useful models

Data source: Data from my students

APP URL: `https://prince-afriyie.shinyapps.io/USCOTS25_Student_Data_App/`

Example 2: Captivating Students' Attention through Modeling with Sports Data

Leveraging sports data to capture students' attention on useful models

Data source: NFL Play-by-Play Data: *nflfastR* R package

- Contains Play-by-play for all NFL games since 1999
- Rich real-world data: Combines numeric, categorical, and time-based features
- Predictive potential: Models can predict yards gained, play success, expected points added (EPA), or win probability
- Strategy analysis: Students can engage as coaches and analysts use it to optimize playcalling
- NFL play-by-play data records every individual play during a football game, capturing detailed information about:
  - Game context: which game, teams, quarter, time remaining
  - Play situation: down, distance, field position
  - Play description: run, pass, kick, penalty
  - Outcome: yards gained, touchdowns, turnovers, EPA

Leveraging sports data to capture students' attention on useful models

Data source: NFL Play-by-Play Data: *nflfastR* R package

APP URL:
`https://prince-afriyie.shinyapps.io/USCOTS25_NFL_App/`

Other available Packages

- *sportsdataverse* R ecosystem is a collection of data package which has data on WNBA, NCAA, MLB, etc
- NBA Player Stats: *nbastatR* R package
- Lahman Baseball Database: *Lahman* R package
- NHL team and player data: *hockeyR* R Package

Example 3: Captivating Students' Attention through Modeling with Financial Data

Leveraging financial data to capture students' attention on useful models

Data source: Freddie Mac Single-Family Loan-Level Dataset: `https://www.freddiemac.com/research/datasets`

- Data is public but requires registration
- Real-world mortgage data (millions of loans)
- Detailed variables similar to underwriting models used by banks
- Longitudinal performance to study defaults, prepayments, and credit risk
- Benchmarks for academic research, **mortgage default modeling**, prepayment analysis, housing market policy analysis
- Variables include credit score, original loan-to-value ratio, flag indicating first-time buyer, number of units in the property, delinquency status code, current unpaid principal balance, default status, etc.

Leveraging financial data to capture students' attention on useful models

Data source: https://www.freddiemac.com/research/datasets

APP URL: `https://prince-afriyie.shinyapps.io/USCOTS25_Freddie_Mac_App/`

Other available Packages

- Fannie Mae Single-Family Loan Performance Data: `https://www.fanniemae.com/`
- Lending Club Loan Data: `https://www.kaggle.com/datasets/wordsforthewise/lending-club`
- HMDA (Home Mortgage Disclosure Act) Data
- Fintech Lending Datasets (Prosper, SoFi)

Example 4: Captivating Students' Attention through Modeling with Data through APIs

Leveraging API data to capture students' attention on useful models

Data source: OpenWeather API: `https://openweathermap.org/api`

APP URL: `https://prince-afriyie.shinyapps.io/USCOTS25_Open_Weather_App/`

Other Free APIs for Instruction

- The Movie Database (TMDB) API: `https://developers.themoviedb.org/3`
- Data.gov APIs: `https://api.data.gov/`
- RestCountries API: `https://restcountries.com`
- Dog API (or Cat API): `https://dog.ceo/dog-api/`

Intangible Ways to Capture Students' Attention

Intangible Ways to Capture Students' Attention

Intangible: "low effort" but high impact

- Make an inviting syllabus
- Pre-course surveys
- Storey telling
- Being relatable

Wrap Up

How do you Gain Students' attention to transform complexity to clarity

- Motivation: Through Data And Useful Models
- Remove any unnecessary barriers to hamper students' learning

Be relatable to your students

Thank you so much for your ATTENTION!
Q&A or Discussion

References

1. Keller, J. M. (1987). Development and use of the ARCS model of motivational design. Journal of Instructional Development, 10(3), 2–10
2. https://databookuw.com/
3. Hunter et al. (2023) Science fiction or clinical reality: a review of the applications of artificial intelligence along the continuum of trauma care. World Journal of Emergency Surgery 18:16.
4. `https://chatgpt.com/`
5. `https://selectorgadget.com/`