Iowa Initiative for Artificial Intelligence
Final Report

Project title:	Improving Data-Based Individualization through Artificial Intelligence: The IDEA Project
Principal Investigator:	Seth King, Associate Professor, Special Education
Prepared by (IIAI):	Avinash Reddy, Mudireddy
Other investigators:	Derek Rodgers, Clinical Professor of Special Education and Methodologist for Scanlan Center for School Mental Health; Shawn Datchuk, Associate Professor of Special Education. In addition to these faculty members, IDEA is receiving assistance from Tipton Elementary School (https://www.tipton.k12.ia.us/), which has agreed to furnish data for the project contingent on IRB approval, and the Iowa Reading Research Center (IRRC; https://iowareadingresearch.org/), which has agreed to assist in processing data and providing other services as needed.
Date:	July 03, 2025

Were specific aims fulfilled:	Y
Readiness for extramural proposal?	Y
If yes … Planned submission date	01/25; 06/25
Funding agency	Tool Competition; Call for Effective Knowledge Competition
Grant mechanism	
If no … Why not? What went wrong?	

Brief summary of accomplished results:

The project successfully demonstrated the feasibility of using machine learning to predict student literacy outcomes. A Transformer-style regression model was developed to predict student SystemScore based on a time series of intervention trials that included student attributes, intervention parameters, and teacher data. The model performed well on a held-out test set, achieving an **R-Squared (R^2) score of 0.74**, indicating it could explain over 74% of the variability in student performance. Other performance metrics included a **Mean Squared Error (MSE) of 2397.18** and a **Mean Absolute Error (MAE) of 37.26**. These results establish a strong proof-of-concept for using AI to generate actionable insights for personalizing literacy instruction.

Research report:
Aims (provided by PI):

The original specific aims of the project were:

1. To evaluate the acceptability of various decision-making rules for Data-Based Individualization (DBI).
2. To predict the efficacy of specific changes in instruction based on student data.
3. To develop a decision-making application compatible with the *Fastbridge* platform to increase adoption throughout Iowa.

However, the project's focus was modified from the original aims. Instead of using CBM data from Tipton Elementary and *Fastbridge* to evaluate DBI rules , the **project pivoted to a foundational goal: "predicting student performance scores"** using a different, readily available dataset from structured SQL databases. This new objective, under the name "Wordflight," focused on predicting metrics like SystemScore. This modification allowed the team to first establish a viable predictive model and a robust data processing pipeline, serving as a critical proof-of-concept before tackling the more complex, original aims of evaluating specific instructional interventions.

Data:

Structured SQL databases containing:

- Student demographics (grade, session duration, language proficiency)
- Intervention attributes (task identifiers)
- Educator profiles
- Literacy metrics (SystemScore)

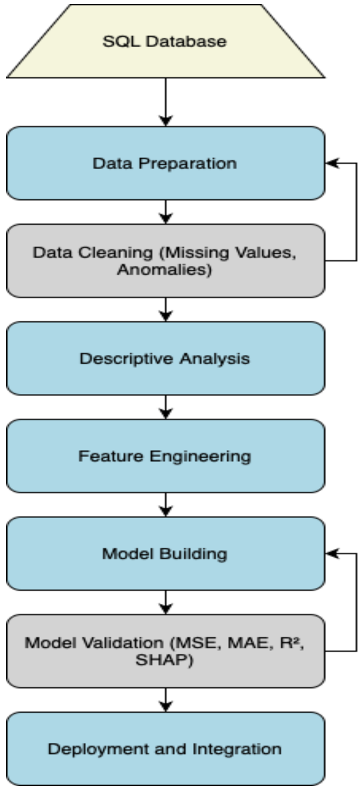
The final dataset for modeling was structured as a time series, represented as X ∈ ℝ^{N×T×F}, where N is the number of samples, T=540 is the number of time-steps, and F=14 is the number of features. The data pipeline involved merging tables, cleaning to handle missing values, and feature engineering to enhance predictive patterns. The full dataset was randomly split into training (with 20% reserved for validation) and held-out test sets.

AI/ML Approach:

The project employed a machine learning pipeline that started with data preparation and cleaning from SQL databases, followed by descriptive analysis, feature engineering, model building, and validation. The core of the analysis was a

Transformer-style regression network built in TensorFlow 2.7. The model architecture included:

- Layer normalization.
- Multi-head self-attention (8 heads).
- Dropout (rate = 0.1).
- A feed-forward projection layer.
- Global average pooling.
- A final linear output layer to produce the scalar prediction.



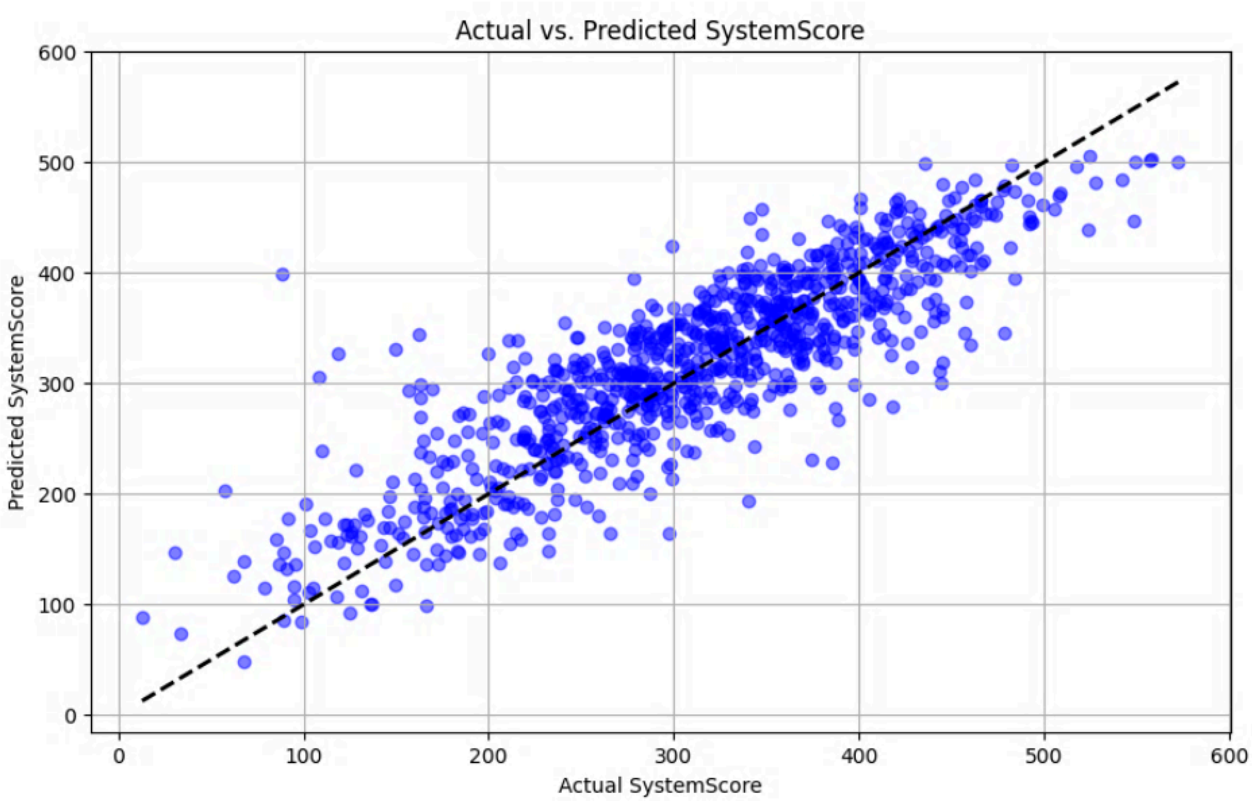
Experimental methods, validation approach: The model was trained for up to 100 epochs using a custom Adam optimizer with a learning rate of 5×10^{-4} and Mean Absolute Error (MAE) as the loss function. To prevent overfitting and improve performance, the training process incorporated EarlyStopping, which halted training when validation loss did not improve for 5 epochs, and ReduceLROnPlateau, which reduced the learning rate if validation loss stagnated. The final model was evaluated on the held-out test set, and performance was visualized using scatter plots of predicted versus actual values to assess calibration.

```
Model: "model"
__________________________________________________________________________________________________
 Layer (type)                   Output Shape         Param #     Connected to
==================================================================================================
 input_1 (InputLayer)           [(None, 540, 14)]    0           []

 layer_normalization (LayerNorm  (None, 540, 14)     28          ['input_1[0][0]']
 alization)

 multi_head_attention (MultiHea  (None, 540, 14)     30222       ['layer_normalization[0][0]',
 dAttention)                                                      'layer_normalization[0][0]']

 dropout (Dropout)              (None, 540, 14)      0           ['multi_head_attention[0][0]']

 layer_normalization_1 (LayerNo  (None, 540, 14)     28          ['dropout[0][0]']
 rmalization)

 dense (Dense)                  (None, 540, 64)      960         ['layer_normalization_1[0][0]']

 global_average_pooling1d (Glob  (None, 64)          0           ['dense[0][0]']
 alAveragePooling1D)

 dense_1 (Dense)                (None, 1)            65          ['global_average_pooling1d[0][0]'
                                                                 ]

==================================================================================================
Total params: 31,303
Trainable params: 31,303
Non-trainable params: 0
__________________________________________________________________________________________________
```

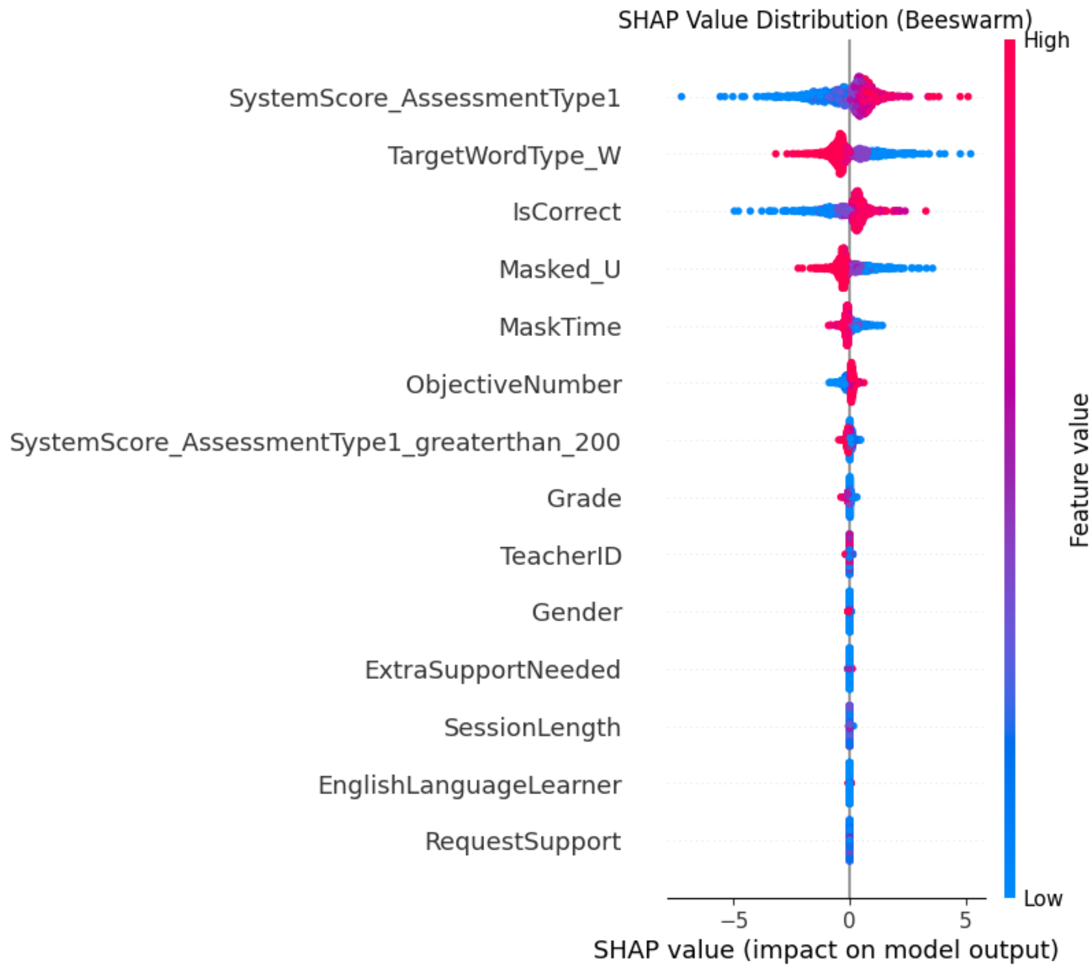
Results:

The model demonstrated strong predictive power for the `SystemScore` metric.

- **Quantitative Metrics**: The model achieved a **Mean Squared Error (MSE) of 2397.18**, a **Mean Absolute Error (MAE) of 37.26**, and an **R-Squared Score of 074**. The R^2 value indicates the model successfully captured over 74% of the variance in student performance data.
- **Qualitative Assessment**: A scatter plot of actual vs. predicted `SystemScore` values showed a strong positive correlation, with data points clustered tightly around the diagonal line, visually confirming the model's alignment with observed student performance.



- **Interpretability**: The methodology incorporates tools like SHAP values to help educators understand the influence of different variables on model predictions, enhancing trust and utility.



Ideas/aims for future extramural project: Collaborate with districts/education companies to couple AI-decision making with curricular adaptations.

Publications resulting from project: King, S.A., Muidreddy, A., & Rodgers, D. B. (Under review). Enhancing data-based individualization through artificial intelligence: Guidance from an ongoing project. *Learning Disabilities Quarterly*