Python technical report

Date: Dec 10, 2025

Problem

The model predicts whether the player will answer the next typing prompt correctly, whether that is capital letter small letter, left hand or right hand righting.

The goal is to adapt difficulty dynamically so that the game remains challenging.

Dataset

Each completed game generates a JSON record.

Features such as rolling accuracy, whether the previous prompt was correct, difficulty, typing speed, etc.

Over time the dataset grows and the model retrains or updates to learn the player's pattern.

Each record includes features such as:

- rolling_accuracy (recent accuracy trend)
- prev_correct (previous number of correct keystrokes)
- current_difficulty, prev_difficulty
- prev_time (time taken in prior session)
- current_length (length of text typed)
- avg_wpm (words per minute average)

method

We used **Logistic Regression.**

for reasons like Simple and lightweight (important for a game where you want fast updates).

- Works well for binary prediction (correct vs. incorrect).
- Easy to interpret feature weights, which helps you adjust difficulty logic.
- Performs well even with a relatively small dataset.

We thought about things we learned this semester which is simple linear Regression but it had less practicality and less results(only accuracy and error) so we thought Logistic Regression was the perfect fit.

Why not more complex models:

- Neural networks or random forests may overfit due to fewer samples.
- More complex models are slower to retrain inside a game loop.
- Logistic Regression provides transparency, which is useful when tuning difficulty rules.

Accuracy/Results

Accuracy	80.6%
Precision	80.2%
Recall	94.2%
F1-Score	86.6%
Baseline Accuracy	66.7%
Improvement	14.0%

Confusion Matrix

Correctly predicted	81
Training Samples	512

Conclusion

The logistic regression model successfully learned to adapt game difficulty based on player performance, achieving 80.6% accuracy—a 14% improvement over the baseline. Its design allows the game to progressively challenge players while maintaining interpretability and real-time responsiveness. By continuously updating from new session data, the system ensures personalized difficulty scaling, enhancing both engagement and skill development. Future improvements could explore ensemble methods or time-series models if data volume increases significantly.

Team

- **Student : Ahmed Haitham**
- **Student: Wed Ahmed**
- **Student: Ruqaya Alaa**