ITESM Campus Monterrey

# Proof of Concept: Improving Cognitive Modeling with Modern Estimation Techniques

Luis Alberto Portilla López
August 23rd, 2024

Research Stay - Going beyond Artificial Intelligence: Artificial Emotions

TC3073 | Group 573

# Introduction

In cognitive science, building accurate models of how humans think and make decisions is essential but challenging. Many traditional methods struggle with handling complex data, especially when the data doesn’t fit neatly into expected patterns. This Proof of Concept (PoC) will demonstrate how using modern estimation techniques can help build better cognitive models, making them more reliable and useful for understanding human thought processes.

## Business Problem

As AI and cognitive models become more integrated into our daily lives, there's a need for these models to be accurate and reliable. However, several issues make this difficult:

1. **Complex Relationships:** Human cognition often involves complex, non-linear patterns that are hard to capture with simple models.
2. **Data Issues:** Cognitive data often doesn't follow the expected patterns, making it hard to analyze.
3. **Small Sample Sizes:** Cognitive studies sometimes work with limited data, which can lead to inaccurate results.
4. **Computational Challenges:** As models become more complex, they require more computing power to process.

These challenges can lead to models that don't accurately represent human cognition, which in turn can result in AI systems that don't interact effectively with people.

# Proposed Solution

To address these challenges, this PoC proposes using modern estimation techniques that are better suited for handling complex data and small sample sizes. Specifically, the solution involves:

1. **Improved Estimation Techniques:** Applying more advanced statistical methods to better handle non-linear data, in this case, we aim to improve the parameter estimation via MLE with Stochastic Gradient Descent (SGD).

To enhance the computational efficiency and accuracy of the model, MLE via SGD is incorporated. MLE provides a principled way to estimate the parameters of the model by maximizing the likelihood function based on the observed data. However, when dealing with high-dimensional data or non-linear models, computing MLE directly can be computationally expensive. SGD addresses these challenges by optimizing the MLE iteratively over small batches of data, reducing computation time significantly and making it feasible to handle larger datasets. Additionally, SGD is particularly effective at escaping saddle points (a point where the slope of a function is zero, but it is not a minimum or maximum) in non-convex optimization problems, which are common in cognitive models. By using mini-batches, SGD updates the model parameters more frequently, leading to faster convergence compared to traditional gradient descent methods. Furthermore, it is flexible enough to handle latent variables, a common feature in cognitive models where not all variables are directly observed. This enables the model to learn from incomplete or missing data, making it more adaptable to real-world scenarios where data may be sparse or noisy.

2. **Bias Reduction:** Using techniques to reduce the errors that come from working with small datasets. In this case we will implement ridge regression regularization.

Regularization penalizes large parameter values, preventing overfitting and reducing the variance of the model. This results in more stable and generalizable estimates, particularly in situations where the available data is limited.

3. **Efficient Computation:** Implementing methods that allow for faster and more efficient processing of complex models. Distributed SGD allows parallel processing of large-scale cognitive datasets, which can significantly reduce computation time.

## Expected Outcomes

By implementing these techniques, the expected outcomes include:

- **Scalability**: The integration of MLE via SGD allows the model to scale efficiently, handling large datasets without a significant increase in computation time.
- **Improved Accuracy**: By iteratively refining parameter estimates, the model becomes more reflective of the underlying cognitive processes, leading to more accurate predictions.
- **Bias Reduction**: The incorporation of regularization techniques minimizes bias, resulting in more reliable parameter estimates, particularly in small sample sizes or noisy data.

## Conclusion

This PoC outlines a simple but effective approach to improving cognitive models using MLE via SGD, combined with regularization techniques, significantly enhancing the robustness of the cognitive model. This not only improves the scalability and accuracy of the model but also addresses key challenges such as overfitting and bias, ensuring that the model can be applied effectively in a wide range of data-driven applications.

## References


1. Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39(1), 1-22.

2. Firth, D. (1993). Bias reduction of maximum likelihood estimates. Biometrika, 80(1), 27-38.

3. Czepiel, S. A. (2002). Maximum likelihood estimation of logistic regression models: theory and implementation. Available at czep. net/stat/mlelr. pdf.

4. Tian Y, Zhang Y, Zhang H. Recent Advances in Stochastic Gradient Descent in Deep Learning. *Mathematics*. 2023; 11(3):682. https://doi.org/10.3390/math11030682.