Selfe: A New Unit for Self Development

Marlon Wayne
Brainzones
marlon@brainzones.org

Abstract

The selfe model is designed to provide a holistic, multi-dimensional view of a student's social-emotional development. It generates three distinct, valuable metrics from the raw data of academic work: the Dynamic Score, the Consistency Score, and the Relational Network.

---

1. The Dynamic Selfe Score (S_i)

This is the primary measure of a student's current standing in a particular selfe. To ensure the score reflects recent growth, it is calculated as an **Exponentially Weighted Moving Average (EWMA)**, which gives more weight to recent data points.

The process begins by deriving a selfe contribution from a tagged academic task:

- $\mathbf{r_j}$: The normalized academic score on task *j* (on a scale of -1 to +1).
- $\mathbf{w_{ij}}$: The relevance weight of selfe *i* for task *j* (on a scale of 0 to 1).
- $\mathbf{s_{ij} = r_j \times w_{ij}}$: The selfe contribution from that single task.

When a new selfe contribution (s_{ij}) is recorded, the overall score S_i is updated using the EWMA formula:

$$S_{i,\text{new}} = \alpha \cdot s_{ij} + (1 - \alpha) \cdot S_{i,\text{old}}$$

Here, **α** is a smoothing factor (a constant between 0 and 1). A higher α makes the score more responsive to new information, while a lower α makes it more stable.

---

2. The Consistency Score (σ_i)

This metric captures the volatility of a student's performance in a given selfe. It is calculated as the **standard deviation** of all the individual selfe contributions (s_{ij}) for a given selfe *i* over a specific period.

A low $\boldsymbol{\sigma_i}$ indicates high consistency. A high $\boldsymbol{\sigma_i}$ indicates volatility, which is a powerful signal for an educator that a student may be struggling with inconsistent motivation or external factors.

---

3. The Relational Selfe Network: An Engine for Prediction

This is the most powerful analytic layer, revealing how a student's selfes interact. The network is modeled as a weighted, undirected graph, **G=(V,E)**, where the vertices **V** are the selfes, and the edges **E** represent the relationships between them. The weight of an edge between any two selfes, $\mathbf{e_{uv}}$, is their historical **correlation coefficient, $\boldsymbol{\rho(S_u,S_v)}$**.

This graph database drives two critical, actionable insights:

3.1 Identifying "Keystone Selfes"

To find the most influential skill for a student, we calculate the **Eigenvector Centrality** of each node in the graph. This metric identifies nodes that are connected to other highly influential nodes. The centrality score for a selfe *i*, $\mathbf{x_i}$, is given by:

$$x_i = \frac{1}{\lambda} \sum_{j \in N(i)} x_j$$

where N(i) are the neighbors of selfe *i* and λ is a constant. A high centrality score identifies a "keystone selfe"—the skill that, if improved, is predicted to cause the largest positive cascade across the entire network.

3.2 Simulating Interventions

To forecast the impact of an intervention, the network is treated as a **Bayesian Network**. The system learns a **Conditional Probability Table (CPT)** for each selfe, defining the probability of its state based on the state of its connected "parent" selfes: $\mathbf{P(S_i \mid Parents(S_i))}$.

This allows an educator to perform predictive modeling. By setting a target for a keystone selfe (e.g., simulating an intervention that raises the probability of a high Collaboration score), the model uses **belief propagation** to calculate the updated probabilities for all its connected selfes, such as Empathy and Communication. This provides a data-driven forecast of the most effective intervention, turning the relational network from a descriptive map into a predictive GPS for student development.

---

4. Methodological Considerations

The construction of the Relational Network is a sophisticated process. The initial graph structure is learned from historical data, modeling probabilistic dependencies, not claiming direct causation. The reliability of the network's predictive power is a direct function of the volume of data collected; its accuracy for any given student will naturally increase over the course of their academic career.

The computational models, while complex, are designed to be trained on aggregated data and run efficiently to provide insights without undue system load. This approach ensures the model is not only theoretically sound but practically feasible.

References

[1] Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988

[2] Bonacich, P. Power and centrality: A family of measures. *American Journal of Sociology*, 92(5):1170–1182, 1987

[3] Hunter, J. S. The exponentially weighted moving average. *Journal of Quality Technology*, 18(4):203–210, 1986

[4] Roberts, S. W. Control chart tests based on geometric moving averages. *Technometrics*, 1(3):239–250, 1959