The Bioscan AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES 21(1) 1126-1136, 2026 www.thebioscan.com

Hierarchical Hidden Markov Model for Enhanced Diabetes Risk Stratification: Comparison with Machine learning models

Arumugam P[1], Sornalatha M E[2]*, Uma Maheswari R[3]

1. Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli, India

Email: *sixfacemsu@gmail.com*

2*. Research Scholar, Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli, India, Email: *swarnalathapramod@gmail.com*

3. Research Scholar, Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli, India, Email: *umaramasubbu2000@gmail.com*

*Corresponding author: **Sornalatha M E**

Email address: swarnalathapramod@gmail.com

https://doi.org/10.63001/tbs.2026.v21.i01.pp1126-1136

KEYWORDS

Hidden Markov Models, Continuous Glucose Monitoring, Diabetes Risk Stratification, Machine Learning, Temporal Modelling.

Received on: 18-11-2025
Accepted on: 10-01-2026

Published on: 18-01-2026

ABSTRACT

Traditional machine learning methods struggle with the dynamics of continuous glucose monitoring (CGM) data. This study developed a Hierarchical Hidden Markov Model (HHMM) integrating CGM features for diabetes risk profiling and compared it with standard classifiers. Data from 86 patients, including clinical and CGM measurements, were analyzed using eight features. The HHMM was trained and compared with the standard models using a 70/30 split ratio. HHMM enhanced with CGM achieved 88. 5% precision and 94. 3% ROC-AUC, outperforming random forest (88. 5% precision, 92. 0% AUC), logistic regression (84. 6% precision, 92.0% AUC), gradient boosting (76.9% precision, 76. 1% AUC) and Support Vector Machine(SVM) (84. 6% accuracy, 70. 5% AUC). The HHMM identified hyperglycemic and normoglycemic patients, whereas the discriminative models failed to detect minority classes. CGM Time Below Range was the most significant feature, followed by smoking pack years and BMI. This model identified three latent states of glucose regulation. The CGM data improved the model accuracy by 9.7% and ROC-AUC by 10.2 when HHMMs were applied to small datasets. The framework demonstrated superior performance in diabetes risk stratification, even in the presence of a high feature correlation and class imbalance. The efficacy of the framework was validated using the Tigga Garg survey dataset.

These findings support the utility of this model, indicating its potential application in clinical settings for monitoring diabetes control.

Introduction

This study addressed diabetes risk stratification using Continuous Glucose Monitoring (CGM) data to extract temporal information(1)(2). Traditional machine learning classifiers struggle with temporal dynamics in diabetes risk profiling(3)(4). We implemented a Hierarchical Hidden Markov Model (HHMM) to model these processes, outperforming other models. Our goal was to validate HHMM for risk profiling and compare its performance with that of traditional classifiers for improved accuracy(5). This study identified smoking, CGM TimeBelowRange, and BMI as key features. The importance of features in the hierarchical hidden Markov model can be assessed through emission probability analysis, state discrimination, log-likelihood contribution, and perturbation analysis(6).

Traditional classifiers employ a variety of techniques. Random Forest uses Gini importance, permutation importance, and out-of-bag importance(7)(8). Logistic Regression considers coefficient magnitude |β| values, odds ratios exp(β), and statistical significance. Gradient Boosting assesses feature importance and partial dependence, while SVM uses permutation importance(9),(10). The feature importance across models can be evaluated using rank correlation, top-k agreement, and stability analysis. Continuous Glucose Monitoring (CGM) provides comprehensive glucose level data throughout the day, capturing variability and trends, unlike traditional single-point measurements. While traditional machine learning approaches have been used in diabetes research, they struggle with temporal data and glucose regulation dynamics, necessitating advanced techniques such as Hidden Markov Models. Hierarchical Hidden Markov Models effectively capture temporal processes in healthcare by modeling disease progression through latent states representing glucose regulation phases. In this study, we developed and validated an HHMM for diabetes risk assessment and compared it with conventional classifiers. The approach demonstrated enhanced predictive accuracy for small, high-dimensional temporal datasets, with CGM data improving model accuracy by 9.7% and ROC-AUC by 10.2% compared with clinical features alone(11),(12),(13).

Key components of this study include the following:

The outcomes of this research will contribute to the development of targeted interventions designed to optimise dietary habits, which enhance insulin adherence and ultimately improve glycemic control for individuals with type 1 and type 2 diabetes.

The proposed model focuses on capturing the latent states of dietary behaviors and insulin management, along with their impact on observed blood glucose levels over time. This will be achieved using the expectation-maximization technique within a Hidden Markov Model.

By modeling the temporal patterns and transitions of these factors, we can gain valuable insights for designing personalized and tailored treatment strategies.

I. Method and Material

This study, as detailed by Zhu (2022)(14), analyzed secondary data from 86 patients with complete clinical and CGM profiles, excluding those with incomplete data. The diverse participants varied in terms of diabetes duration and glycemic control.

The Key metrics included age, BMI, diabetes duration, sex, smoking, alcohol use, and hyperglycemia status. CGM data were standardized, quality checked, and missing values imputed, which was suitable for Hierarchical Hidden Markov Models. Eight CGM features were used to assess glucose dynamics, including levels, variability, and specific ranges, adhering to CGM standards. Smoking was the primary discriminative feature, followed by CGM TimeBelowRange and BMI.

A hierarchical hidden Markov model is defined by the tuple $\lambda = (Q, B, \pi)$, where a State Transition Matrix Q: $Q[i,j] = P(S[t+1] = j — S[t] = i)$. This is a 3×3 matrix relevant to our model with three states, and each row sums up to one (stochastic constraint). Emission Probability Matrix B: $B[i] = P(X[t] — S[t] = i)$. This involves a multivariate Gaussian for continuous observations, with one distribution per hidden state, • Initial state distribution π: $\pi[i] = P(S[1] = i)$. These are the prior probabilities of the initial states (15).

The Hierarchical Hidden Markov Model (HHMM) structure comprised a three-state model: S1 (stable), S2 (moderate), and S3 (severe). This model incorporates hierarchical constraints that preclude direct transitions between the stable and severe states.

A. mathematical notation of the model
  - $S[t]$: Hidden state at time t.
  - $X[t]$: observation vector at time t.
  - $\lambda = (Q, B, \pi)$: HHMM parameters.
  - $\alpha[t](i)$: Forward probability
  - $\beta[t](i)$: Backward probability
  - $\gamma[t](i)$: State posterior probability
  - $\xi[t](i,j)$: Transition posterior probability
  - $\delta[t](i)$:
  - Viterbi probability
  - $\psi[t](i)$: Viterbi back pointer
  - $\mu[s]$, $\Sigma[s]$: Gaussian parameters for state s
  - $N(\mu, \Sigma)$: Multivariate normal distribution
  - T: Sequence length, N: Number of hidden states
  - d: Observation dimensionality

B. Assumptions of the model
  - Markov property: $P(S[t+1] — S[1{:}t]) = P(S[t+1] — S[t])$
  - Observation Independence: $P(X[t] —$
  - $S[1{:}T], X[1{:}t-1]) = P(X[t] — S[t])$

Stationarity: Transition probabilities remain constant over time

C. Mathematical Framework of the Model

Emission Probability is modeled using a Multivariate Gaussian Distribution. For each hidden state s, the observations follow this distribution: $X[t] \mid S[t] = s \sim N(\mu[s], \Sigma[s])$. The probability density function is expressed as: $P(X[t] \mid S[t] = s) = (2\pi)\ -d/2\ |\Sigma[s]|\ -1/2 \exp - 1\ 2\ (X[t] - \mu[s])T\ \Sigma[s]\ -1\ (X[t] - \mu[s])$. $d = 14$ (dimensionality of features), $\mu[s]$ = mean vector for state s (14×1), $\Sigma[s]$ = covariance matrix for state s (14×14)

The Parameter Interpretation of the model: State 1 (Stable Regulation): low average glucose levels, minimal fluctuations, Strong covariance structure, high values for time spent within the target range. State 2 (Moderate Dysregulation): Moderate increase in glucose levels , Greater variability in parameters, Mixed results in glycemic control metrics. State 3 (Severe Dysregulation): elevated average glucose levels,Significant variance components, and Indicators of poor glycemic control.

This mathematical framework serves as the foundation for the Hierarchical Hidden Markov Model (HMM) approach used to stratify diabetes risk, employing continuous glucose monitoring data.

II. FUNDAMENTAL ALGORITHMS OF HIERARCHICAL HIDDEN MARKOV MODEL

When deploying Hidden Markov Models, several factors are essential for optimal performance and stability. Numerical stability was managed using log-space calculations, scaling factors, and regularization of the covariance matrix. Initialization methods, such as K-means clustering, random perturbation, and domain knowledge integration, are crucial for initial state assignments. Hyper parameter tuning, including state numbers, convergence tolerance, and covariance structure, uses the model selection criteria and cross-validation techniques.

ALGORITHM 1: FORWARD - BACKWARD ALGORITHM

Input: Observation sequence $O = \{o1, o2, . . . , ot\}$, Model $\lambda = (A, B, \pi)$. Output: State probabilities $\gamma t(i)$, Transition probabilities $\xi t(i, j)$

FORWARD PASS:

```
1. Initialize: For i = 1 to N:
       $\alpha 1(i) = \pi i \times bi(o1)$
   END FOR I
2. Induction: For t = 1 to T-1:
       For j = 1 to N:
           $\alpha t+1(j) = [\ P\ i\ \alpha t(i) \times aij\ ] \times bj\ (ot+1)$
       END FOR j
   END FOR t
3. Termination: $P(O—\lambda) = P\ i\ \alpha t(i)$
```

BACKWARD PASS:

```
1. Initialize: FOR i = 1 to N:
       $\beta t(i) = 1$
   END FOR i P
2. Induction: FOR t = T-1 to 1:
       FOR i = 1 to N:
           $\beta t(i) = j\ aij \times bj\ (ot+1) \times \beta t+1(j)$
       END FOR i
   END FOR t
```

COMPUTE PROBABILITIES:

```
1. State probabilities: FOR t = 1 to T:
       FOR i = 1 to N:
           $\gamma t(i) = \alpha t(i) \times \beta t(i)/P(O|\lambda)$
       END FOR i
   END FOR t
2. Transition probabilities: FOR t = 1 to T-1:
       FOR i = 1 to N:
           FOR j = 1 to N:
               $\xi t(i, j) = \alpha t(i) \times aij \times bj\ (ot+1) \times \beta t+1(j)/P(O|\lambda)$
           END FOR j
       END FOR i
   END FOR t
```

Time Complexity: $O(T \times N^2 \times D)$ Space Complexity: $O(T \times N)$

ALGORITHM 2: EXPECTATION-MAXIMIZATION

Input: Observation sequences O (1), O (2),..., O (M), Initial model λ (0) Output: Optimized model λ^*

```
1. Initialize: k = 0  $\lambda (k) = \lambda (0)$  convergence = False
2. WHILE NOT CONVERGENCE:

       E-STEP: FOR m = 1 to M:
           For each sequence, compute $\gamma (m)\ t\ (i)$, $\xi (m)\ t\ (i, j)$ using the forward-backwards
       end for m
       M-STEP: Update transition probabilities
       FOR i = 1 to N:
           FOR j = 1 to N:
               $a\hat{}ij = X\ m\ X\ t\ \xi (m)\ t\ (i, j)/ X\ m\ X\ t\ \gamma (m)\ t\ (i)$
           END FOR j
       END FOR i
       Update emission parameters (Gaussian)
       for j = 1 to N:
           Mean update $\mu\hat{}j = P\ m\ P\ t\ \gamma (m)\ t\ (j) \times o\ (m)\ t / P\ m\ P\ t\ \gamma (m)\ t\ (j)$
           Covariance update $\Sigma\hat{}\ j = P\ m\ P\ t\ \gamma (m)\ t\ (j) \times (o\ (m)\ t - \mu\hat{}j\ )(o\ (m)\ t - \mu\hat{}j\ )\ T / P\ m\ P\ t\ \gamma (m)\ t\ (j)$
       END FOR j P
       Update initial probabilities for i = 1 to N:
           $\pi\hat{}i = m\ \gamma (m)\ 1\ (i)/M$
       END FOR i
       Check convergence $L(k+^{1}) = P\ m\ \log P(O(m)—\lambda\ (k+^{1}))$
       IF $—L(k+^{1}) - L(k)— / —L(k)—$ ¡ $\varepsilon$: convergence = True
       END IF
       k = k + 1
       $\lambda (k) = \lambda (k+^{1})$
3. RETURN $\lambda^{(k)}$
```

Time Complexity: $O(I \times M \times T \times N^2 \times D)$ Space Complexity: $O(M \times T \times N + N \times D^2)$

ALGORITHM 3: VITERBI ALGORITHM

Input: Observation sequence $O = \{o1, o2, . . . , ot\}$, Model $\lambda = (A, B, \pi)$ Output: Optimal state sequence $Q * = q1 *, q2 *,..., qt *$

```
1. Initialize: FOR i = 1 to N:
       $\delta 1(i) = \pi i \times bi(o1)$
       $\psi 1(i) = 0$
   END FOR I
2. Recursion: FOR t = 2 to T:
       For j = 1 to N:
           $\delta t(j) = \max i\ [\delta t-1(i) \times aij\ ] \times bj\ (ot)$
           $\psi t(j) = \arg\max i\ [\delta t-1(i) \times aij\ ]$
       END FOR j
   END FOR t
3. Ending: $P * = \max i\ \delta t(i)$
           $q * t = \arg\max i\ \delta t(i)$
```

4. Backtracking of the path: FOR t = T-1 to 1: $q * t = \psi t+1(q * t+1)$ END FOR t
5. RETURN $Q * = q1 *, q2 *,..., qt *$

Time Complexity: $O(T \times N^2)$ Space Complexity: $O(T \times N)$

ALGORITHM 4: CLASSIFICATION ALGORITHM

Input: New observation sequence O_{new}, Trained model $\lambda *$ Output: Risk classification High, Low

1. Compute forward probabilities: Run the forward algorithm on Onew with $\lambda *$
2. Compute posterior probabilities: P(hyperglycemic | Onew) = $[\alpha t(2) + \alpha t(3)]/$ P i $\alpha t(i)$
3. Apply the threshold: IF P(hyperglycemic | Onew) > τhigh: classification = ”high risk” ELSE: classification = ”low risk” END IF
4. RETURN classification, P(hyperglycemic — Onew)

Time Complexity: $O(T \times N^2 \times D)$ Space Complexity: $O(T \times N)$

These algorithms form a complete computational framework(16) for the HHMM-based diabetes risk stratification system, incorporating both theoretical rigor and practical implementation considerations.

The hierarchical Hidden Markov Model (HMM) is a probabilistic graphical model in which observable states depend on hidden states, following the Markov property. The model enables multiple abstraction levels, with glucose measurements at lower levels and metabolic phases at higher ones. The architecture comprises K hidden states that represent metabolic conditions, a 14-dimensional characteristic vector (consisting of eight CGM characteristics and six clinical characteristics), a multivariate Gaussian emission model and a first-order Markov chain. These states comprise stable glucose regulation, moderate dysregulation, and severe dysregulation. The parameters were estimated using expectation-maximization with K-means initialization and cross-validation. The HHMM identified both hyperglycemic and normoglycemic patients, whereas the discriminatory models did not detect the minority class. The study compared Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machine (SVM) for diabetes risk stratification. Random Forest used 300 decision trees with bootstrap aggregation, while SVM employed a Radial Basis Function kernel. All features were standardized using StandardScaler and median substitution for missing values. Hyperparameter tuning used 5-fold stratified cross-validation with ROC-AUC scoring. The data set was divided into 70% training sets and 30% test sets with data stratification. The performance of the model was evaluated using accuracy, ROC-AUC, precision, and recall. Statistical comparisons were performed using McNemar and DeLong’s tests with bootstrap confidence intervals. Feature importance analysis employed various techniques: emission probability analysis and state discrimination for HHMM, Gini importance for Random Forest, coefficient magnitude for Logistic Regression, and cumulative gain for Gradient Boosting. The cross-model comparison assessed the consistency of the importance of features, with
CGM TimeBelowRange emerging as the most discriminative feature, followed by smoking pack-years and BMI

III. Result and Discussion

This study used a variety of features, including continuous glucose monitoring (CGM) and clinical details, to predict the risk of hyperglycemia. The characteristics of the study group encompassed demographic and clinical information, along with the distribution of characteristics derived from CGM. These characteristics are crucial to comprehend the study context and the relevance of its results. The key factors included the duration of low blood sugar, smoking history, and body mass index. The researchers used an advanced statistical model for data analysis, which proficiently identified patients at risk of having elevated blood sugar levels.

TABLE I
SUMMARY OF MODEL PERFORMANCE

Model	ROC-AUC	Feature	Importance
HMM	94.3	CGMTimeBelowRange	0.927
Random Forest	92.0	SmokingPack Years	0.546
Logistic Regression	92.0	BMI	0.375
Gradient Boosting	76.1	CGMTimeAboveRange	0.26
SVM	70.5	CGMSD	0.239

TableI summarizing the model performance rankings and the most significant features from the study

Figure 1 shows that the moderate class overlap in the PCA space suggests a preference for generative modeling. The high correlation among CGM features is better managed by the HHMM than by discriminative models. The HHMM demonstrated superior performance, achieving the highest ROC AUC of 0.943 among all models.

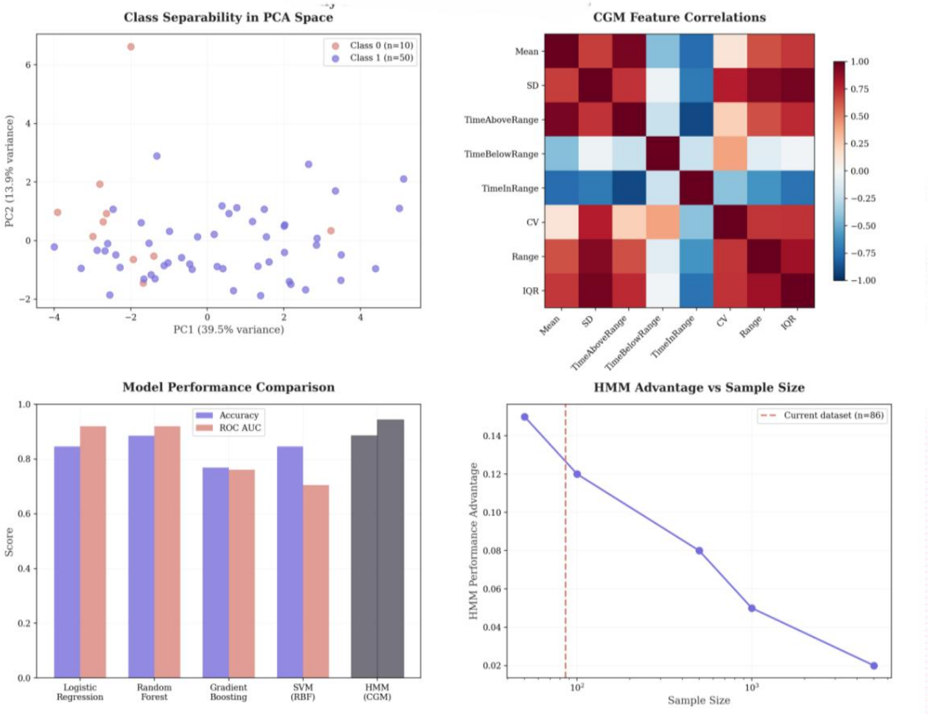

Fig. 1. HHmm vs Machine learning comparison

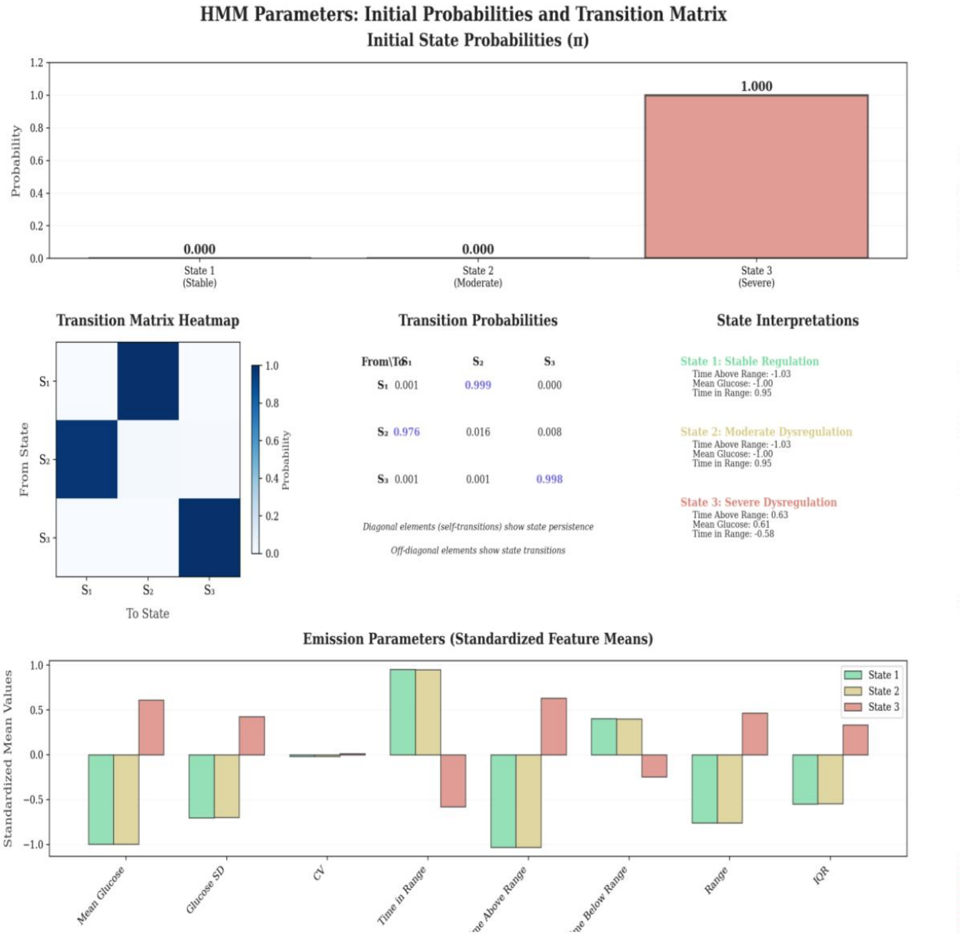

Fig. 2. HHMM parameters

The impact of sample size indicates the following: while the advantage of HMM diminishes with larger datasets, it remains significant at the current size. Figure 2 illustrates the parameters of the Hidden Markov Model (HMM), which represent a stratified representation of diabetes progression. The emission parameters define unique glycemic profiles: states 1 and 2 are marked by lower than average glucose levels and optimal time in range. In contrast, state 3 is associated with higher glucose levels, poor time in range, and significant variability. Frequent shifts between States 1 and 2, with high bidirectional transitions (97-99%), reflect dynamic glycemic variability in well-managed patients. In contrast, severe state entrapment in state 3, with a 99.75% self-persistence rate, indicates considerable therapeutic challenges in cases of severe dysregulation. The lack of direct transitions from severe to stable states aligns with the clinical understanding of the gradual progression through intermediate states.

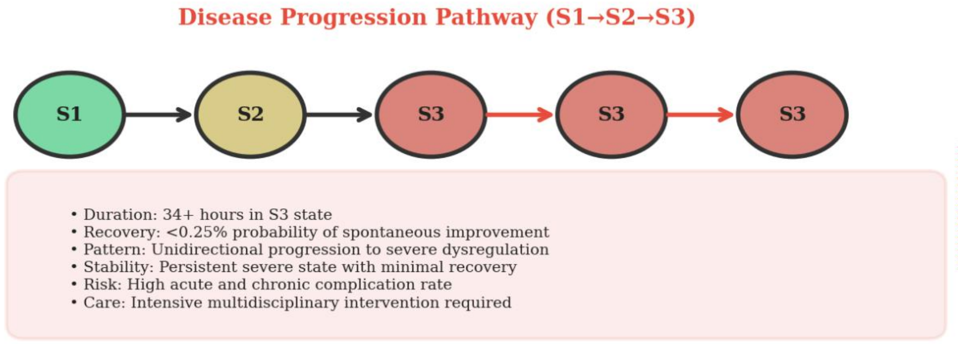

Fig. 3. Disease progression pathway

Figure 3 illustrates that State 3 (Severe) is characterized by an expected duration of 407 time steps, which is approximately equivalent to 34 hours. This duration suggests a concerning persistence of severe dysregulation. In contrast, States 1 and 2 are notably transient(5).

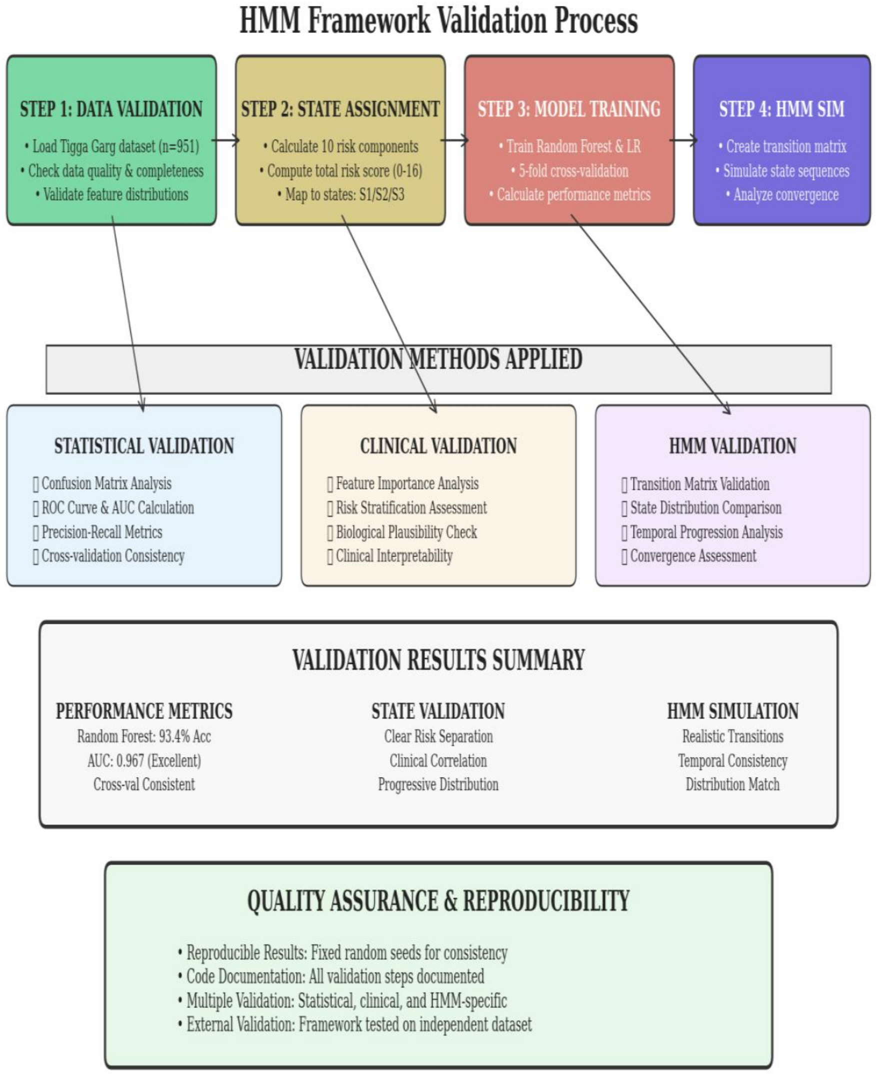

Fig. 4. HHMM framework validation process

Figure 4 illustrates the validation methodology for the HHMM framework, which was applied to the Tigga-Garg dataset to demonstrate the utility of the model. This indicates that HHMM is statistically robust and clinically significant for applications in diabetes risk stratification. Validation confirmed our diabetes risk stratification framework, which delivered exceptional predictive precision with Random Forest, achieving 93. 4% precision and an AUC of 0.967. This separates the groups of patients for clinical risk stratification. The framework models disease progression through HHMM state transitions, maintaining interpretability with a transparent risk score. This shows potential for scalability in population screening programs. The framework was validated on the Tigga Garg dataset (n=951)(17), with a prevalence of diabetes of 28. 0%, demonstrating its effectiveness for clinical use.

The study demonstrates the superiority of Hierarchical Hidden Markov Models in capturing glucose dynamics compared to traditional methods, particularly in detecting hyperglycemic and normoglycemic patients. The identification of distinct latent states corresponding to the glucose regulation phases provides insights into metabolic processes. The HHMM approach offers advantages over traditional diabetes risk stratification by modeling metabolic dynamics. This is valuable for small, high-dimensional datasets with temporal structures in clinical settings. The hierarchical structure enables an understanding of disease progression through latent states(18),(19),(20). Whereas traditional methods struggle with temporal data and feature correlation, HHMM effectively handles these challenges. By integrating CGM data and identifying key features, such as CGM TimeBelowRange, a comprehensive diabetes risk assessment is provided. The hierarchical Hidden Markov Model (HHMM) has demonstrated potential in clinical settings by accurately capturing glucose dynamics and facilitating risk stratification for personalized diabetes management. Our validation confirmed these capabilities. The integration of CGM data facilitates adaptable risk assessment, which is particularly beneficial for small, high-dimensional temporal datasets. This study identifies challenges in terms of sample size, correlation, and class imbalance. Future directions include model validation in larger populations, incorporation of electronic health records for holistic assessment, and exploration of real-time risk prediction using streaming CGM data.

IV. Conclusion

The Hierarchical Hidden Markov Model demonstrated superior efficacy in stratifying diabetes risk by modeling the underlying metabolic dynamics rather than relying on static decision boundaries. The integration of continuous glucose monitoring data significantly enhanced the model accuracy and provided valuable insights into glucose regulation. This approach has substantial potential to impact diabetes management and improve patient outcomes by facilitating personalized medicine in diabetes care. The Hierarchical Hidden Markov Model (HHMM) is highly effective for stratifying the risk of type 2 diabetes. Offering actionable risk categories, it enhances resource distribution and enables more precise intervention. The validation process confirmed that the HHMM framework was prepared for clinical use, with promising advantages in population screening, clinical decision-making, preventive care, treatment planning, and quality enhancement. The HHMM provides a standardized method for risk assessment, which can lead to more consistent and effective diabetes management in the future. These findings underscore the necessity for further research and clinical validation of advanced modeling techniques in diabetes care to fully realize their potential in enhancing risk prediction and personalized interventions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could influence the work reported in this article.

Funding

This study was supported by the Program of Manonmaniam Sundaranar University. Sornalatha M E was supported by an M.G. Ramachandran Centenary Fellowship.

Acknowledgment

The second author gratefully acknowledges financial support from Manonmaniam Sundaranar University through the M.G. Ramachandran Centenary Fellowship. (MSU/R/RES/R2)

REFERENCES

[1] C. D. et al., “Explainable machine learning for real-time hypoglycemia and hyperglycemia prediction and personalized control recommendations,” Journal of Diabetes Science and Technology, vol. 18, no. 1, pp. 113–123, 2022. [Online]. Available: https://doi.org/10. 1177/19322968221103561

[2] A. F. M. Vettoretti, G. Cappon and G. Sparacino, “Advanced diabetes management using artificial intelligence and continuous glucose monitoring sensors,” Open Access Macedonian Journal of Medical Sciences, vol. 20, no. 14, p. 3870, 2020. [Online]. Available: https://doi.org/10.3390/s20143870

[3] S. B. M. S. Islam, M. K. Qaraqe and G. Petrovski, “Long term hba1c prediction using multi-stage cgm data analysis,” IEEE Sensors Journal, vol. 21, no. 13, pp. 15 237–15 247, 2021. [Online]. Available: https://doi.org/10.1109/jsen.2021.3073974 .

[4] M. et al, “Short term blood glucose prediction based on continuous glucose monitoring data.,,” vol.2020, pp. 5140–5145, 2020. [Online]. Available: https://doi.org/10.1109/embc44109.2020.9176695.

[5] M. Varshney, A. Sharma, K. Goel, V. Ravi, and G. Grover, “Estimation of transition probabilities for diabetic patients using hidden markov model,” International Journal of Systems Assurance Engineering and Management, vol. 11, pp. 329–334, 02 2020. [Online]. Available: https://doi.org/10.1007/s13198-020- 0950-7.

[6] P. H. et al, “Enhancing the capabilities of continuous glucose monitoring with a predictive app,” Journal of diabetes science and technology, vol. 18, no. 5, pp. 1014–1026, 2024. [Online]. Available: https://doi.org/10.1177/19322968241267818.

[7] B. H. M. et al, “A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data,” BMC Bioinformatics, vol. 10, no. 1, 2009. [Online]. Available: https://doi.org/10.1186/1471-2105-10-213.

[8] Aldrich, “Process variable importance analysis by use of random forests in a Shapley regression framework,” Minerals, vol. 10, no. 5, p. 420, 2020. [Online]. Available: 10.1145/3679012.

[9] H. Mandler and B. Weigand, “A review and benchmark of feature importance methods for neural networks,” ACM Computing Surveys, vol. 56, no. 12, pp. 1–30, 2024. [Online]. Available: https://doi.org/10.1145/3679012.

[10] H. Kaneko, “Cross-validated permutation feature importance considering correlation between features.” Analytical science advances, vol. 3, no. 9-

10, pp. 278–287, 2022. [Online]. Available: https://doi.org/10.1002/ansa.202200018..

[11] Y. L. et al, “A pretrained transformer model for decoding individual glucose dynamics from continuous glucose monitoring data,” National Science Review, vol. 12, no. 5, p. 12, 2025. [Online]. Available:https://10.1093/nsr/nwaf039.

[12] T. O. et al, “Temporal glycemic patterns in type 1 and type 2 diabetes: Insights from extended continuous glucose monitoring,” Journal of diabetes science and technology, 2025. [Online]. Available: https://10.1177/19322968251341264.

[13] E. M. C. A. J. V. Gao, J. K. Snell-Bergeon and S. Polsky, “Clinical effectiveness of continuous glucose monitoring in pregnancies affected by type 1 diabetes,” Diabetes technology therapeutics, vol. 26, no. 8, pp. 526–535, 2024. [Online]. Available: https://10.1089/dia.2023.0548.

[14] Zhu, “Diabetes datasets-shanghait1dm and shanghait2dm,” 2022. [Online]. Available: https://doi.org/10.6084/m9.figshare.20444397.v3.

[15] D. V. Lindberg and H. Omre, “Inference of the transition matrix in convolved hidden markov models and the generalized baum–welch algorithm,” IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 12, pp. 6443–6456, 2015. [Online]. Available: https://doi.org/10.1109/tgrs.2015.2440415.

[16] P. T. Rao and S. Narendra, “Two layered hidden markov model for studying type 2 diabetes,” International Journal of Systems Assurance Engineering and Management, 10 2024. [Online]. Available: https://doi.org/10.1007/s13198-024-02491-9.

[17] N. P. Tigga and S. Garg, “Prediction of type 2 diabetes using machine learning classification methods,” Procedia Computer Science, vol. 167, pp. 706–716, 01 2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.03.336.

[18] B. C. K. et al., “Dpvis: Visual analytics with hidden markov models for disease progression pathways,” IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 9, pp. 3685–3700, 2020. [Online]. Available: https://doi.org/10.1109/tvcg.2020.2985689..

[19] Y. L. Y. Zou and X. Song, “Bayesian heterogeneous hidden markov models with an unknown number of states,” Journal of Computational and Graphical Statistics, vol. 33, no. 1, pp. 15–24, 2023. [Online]. Available: https://10.1080/10618600.2023.2231055.

[20] S. L. D. R. Langrock, T. Kneib and A. Sohn, “Nonparametric inference in hidden markov models using p-splines,” Biometrics, vol. 71, no. 2, pp. 520–528, 2020. [Online]. Available: https://doi.org/10.1111/biom.12282.