ChatGPT

Advanced Methods for Modeling and Predicting Human Behaviour

Introduction

Markov chains provide a simple way to model sequences of events where the probability of the next state depends only on the current state (the *memoryless* property). However, human behaviour is rarely truly memoryless. More sophisticated models incorporate unobserved states, causal structure, nonlinear dynamics, network interactions, stochastic shocks and long–range dependencies. This report surveys several complementary modelling approaches—hidden Markov models, Bayesian networks, state-space models and Kalman filtering, stochastic differential equations, reinforcement learning, graph neural networks, agent-based simulations, chaos theory and information-theoretic methods—and describes how they can be combined to enhance AI-powered behavioural prediction. Each section cites original peer-reviewed research and includes a demonstration to provide an intuitive understanding of the method.

Hidden Markov Models (HMMs)

Concept and Applications

A basic Markov chain assumes that the current state is fully observable. In mental-health applications, true internal states (stress, mood or intent) are rarely measured directly. *Hidden Markov models* (HMMs) extend Markov chains by assuming that observations are generated by unobserved hidden states. An HMM is specified by a transition matrix for hidden states and an emission matrix for observations.

An experiment on mouse locomotion in a visual-cliff test used a three-state HMM to classify **resting**, **exploring** and **navigating** behaviour. The HMM analysis revealed that mice transition between these states rather than exhibiting simple avoidance, highlighting the complex structure of behaviour [1]. HMMs have also been used to detect speech patterns, stress levels and engagement in psychotherapy sessions.

Demonstration

The diagram above shows a simple three-state Markov chain (a reduced form of an HMM without hidden observations). Each arrow is labelled with a transition probability. In an HMM, observations (e.g., patient messages) would be emitted from hidden states and used by algorithms like the Viterbi algorithm to infer the most probable sequence of hidden states.

Bayesian Networks and Dynamic Bayesian Networks

Concept and Applications

Bayesian networks (BNs) are directed acyclic graphs in which nodes represent variables and edges encode conditional dependencies. They allow explicit representation of causal structure and can incorporate expert knowledge. When the graph is unrolled over time, the model becomes a *dynamic Bayesian network* (DBN). A DBN includes an initial network specifying the joint distribution at time 0 and a transition network encoding how variables at time t influence variables at time t+1 [2]. A study examining social determinants of mental health built a DBN where mental-health status at the next time point depended on general health, loneliness and community satisfaction [3].

DBNs are useful for **causal inference** and **early-warning systems** because they can simulate the effect of interventions on later outcomes. They naturally handle missing data through probabilistic reasoning and support online updating as new data arrive.

Demonstration

The figure shows a simple DBN with two variables (health and stress) across two time steps. Solid arrows represent causal relations. Inter-time arrows encode how health and stress at time t influence their values at time t+1; the diagonal arrow shows cross-temporal influence (stress $t \rightarrow health$). Real DBNs may include many variables, but the principle of modelling causal dependencies across time remains the same.

State-Space Models and Kalman Filtering

Concept and Applications

State-space models describe how a hidden continuous-valued state evolves over time. They consist of a state transition equation and an observation equation. A widely used linear–Gaussian special case is the **Kalman filter**, which provides optimal recursive estimation of the hidden state. Kalman filtering has been applied to estimate core temperature from heart-rate measurements in occupational settings [4]. However, many psychological processes are nonlinear, and linear models like vector autoregression (VAR) or Kalman filters may fail to capture dynamic feedback loops. Recent work suggests that **piecewise-linear recurrent neural network (PLRNN) state-space models** can better forecast mental-health trajectories because they allow nonlinear dynamics while remaining interpretable [5].

Demonstration

The plot shows a one-dimensional state-space model where the true position (blue) increases linearly with time, noisy observations (orange points) deviate randomly, and a Kalman filter (green) estimates the latent state. The filter combines prior predictions with new observations to produce smoothed estimates.

Continuous-Time and Stochastic Differential Equations

Many psychological phenomena evolve continuously. *Stochastic differential equations* (SDEs) model how states change under deterministic drift and stochastic diffusion. An SDE of the form dX_t = θ(μ − X_t) dt + σ dW_t (the **Ornstein–Uhlenbeck process**) captures mean-reverting dynamics. SDEs have been used to model affect dynamics with parameters representing attractor points and reactivity to inputs [6]. Continuous-time models handle irregular sampling and reveal subtle changes better than discrete-time models [7].

The time series above shows an Ornstein–Uhlenbeck process. The state wanders randomly but tends to revert toward a mean (0). This behaviour is analogous to mood fluctuations that drift around a baseline with random shocks.

Reinforcement Learning (RL)

Concept and Applications

Reinforcement learning models an agent interacting with an environment to maximize cumulative reward. RL has been adopted in mental-health research to learn *dynamic treatment regimes* that adapt interventions based on patient response. Reviews note a rapid increase in healthcare applications of RL since 2020 and highlight its potential to tailor care but caution about data quality and transparency [8]. RL has been proposed for preventing depression relapse by analyzing behavioural data, detecting risk and optimizing personalized interventions [9].

Demonstration

In this two-armed bandit simulation, the agent chooses between two arms with unknown reward probabilities (0.2 and 0.8) using an epsilon-greedy strategy (ε=0.1). The red curve counts the number of times the optimal arm was selected; the blue curve accumulates rewards. Over time, the agent learns to favor the better arm, illustrating how RL can learn personalised intervention policies.

Graph Neural Networks (GNNs)

Concept and Applications

Human behaviour is embedded in social networks and multimodal brain-network data. *Graph neural networks* process data defined on graphs by iteratively aggregating information from neighboring nodes. A multimodal depression detection model called **EMO-GCN** used GNNs to integrate EEG and speech data, achieving 96.3% accuracy in diagnosing major depressive disorder [10]. A deep graph-learning model combining fMRI and EEG connectivity predicted antidepressant responses with modest but significant R^2 values and highlighted key brain regions [11].

Demonstration

The schematic network above represents individuals (nodes) connected by social ties (edges). GNNs propagate information along edges, enabling a node's hidden representation to reflect its context. In mental-health applications, nodes might represent brain regions or people, and edges represent functional connectivity or social interactions.

Agent-Based Modelling

Concept and Applications

Agent-based models (ABMs) simulate the decisions of individual agents and their interactions with each other and the environment. ABMs are particularly useful when complex system behaviour emerges from interactions rather than individual characteristics. A public-health article argues that agent-based modelling can represent the complex causal architecture of health behaviour by simulating entities whose actions influence and are influenced by their environment [12]. ABMs allow heterogeneity and interaction effects that are difficult to capture with aggregate models [13].

Demonstration

The image shows a Schelling segregation simulation. Two types of agents (blue and red) occupy a grid. Agents move to a vacant location if too few neighbours are of the same type. Even when agents only have a mild preference for similar neighbours (threshold 30%), the system evolves into segregated clusters, illustrating how complex group patterns can emerge from local rules.

Chaos Theory, Fractals and Complex Systems

Concept and Applications

Traditional health-behaviour models assume linear relationships—small inputs produce proportionally small outputs. Resnicow and Page argue that behaviour change often occurs through *quantum events* that resemble chaotic processes; change is sensitive to initial conditions, highly variable and difficult to predict [14]. A companion debate article describes four key principles: (1) chaotic systems can be mathematically modelled but are hard to predict; (2) they are sensitive to initial conditions; (3) complex systems involve many interacting parts; and (4) results are greater than the sum of parts [15]. Small changes in knowledge or attitude may dramatically alter outcomes, and interactions can produce nearly infinite patterns [16].

Chaos theory suggests that human behaviour may follow fractal patterns with strange attractors. Recognizing chaotic dynamics allows therapists to focus on tipping points rather than linear progress. For instance, sudden insights or epiphanies (quantum change) may trigger lasting transformations [17].

Demonstration

The logistic map x_{n+1}=a x_n (1−x_n) with a=3.9 exhibits chaos. Two trajectories with almost identical initial conditions (0.2 and 0.2001) diverge dramatically over time, illustrating *sensitivity to initial conditions*. Such unpredictability underscores why linear models may fail to anticipate behavioural shifts.

Information-Theoretic Approaches

Concept and Applications

Information theory quantifies uncertainty and relationships between variables. Shannon entropy measures the average information content of a random variable. Mutual information quantifies the dependence between two variables. In computational psychiatry, information theory provides a framework to describe information processing and its disorders. For example, Silverstein et al. argue that foundational concepts such as Shannon information, entropy, data compression and strategies to increase the signal-to-noise ratio can offer novel insights into cognitive impairments and guide computational models of schizophrenia [18]. Recent extensions such as *infomax* and *partial information decomposition* can model failures in sensory tuning, perceptual organization and cognitive control [19].

Information theory has been used to analyse cognitive control tasks: fMRI studies show that activation of the cognitive control network correlates with the entropy of a task [20]. Mutual information can decompose the information required for stimulus–response mapping into sensorimotor, contextual and episodic components [21]. Information-theoretic measures have also been applied to EEG to detect depression, to measure complexity of physiological signals and to evaluate treatment decisions.

Demonstration

The plot displays the entropy of a binary random variable with probability p of being 1. Entropy is maximal (1 bit) when p=0.5, reflecting maximum uncertainty, and decreases toward 0 as the outcome becomes certain. In behaviour analysis, rising entropy may indicate increased unpredictability or complexity in a person's actions, whereas decreasing entropy suggests more stereotyped behaviour.

Sequence Models: RNNs, LSTMs and Transformers

Concept and Applications

Recurrent neural networks (RNNs) and their gated variants (LSTMs, GRUs) process sequences by maintaining a hidden state that summarizes previous inputs. Unlike Markov chains, which only depend on the last state, RNNs can incorporate long histories. A 2022 study used LSTMs to forecast human decision-making in psychological tasks (Iterated Prisoner's Dilemma and Iowa Gambling Task). The researchers trained the networks on 168 386 decisions and showed that LSTMs outperformed logistic regression and vector autoregression in predicting individuals' action sequences [22]. The weights of the LSTM models for top performers had broader distributions than those for poor performers, suggesting differences in strategy [23].

Transformers use self-attention to model dependencies between all positions in a sequence, enabling parallel processing and capturing long-range interactions. A 2025 article introduced **MPHI-Trans**, a Transformer-based multimodal temporal model for adolescent mental-health prediction. By combining multimodal data (text, images, physiological signals) with temporal modelling and personalized features, the model provided accurate individual predictions and targeted interventions [24]. MPHI-Trans integrates LSTMs for temporal modelling with a Transformer self-attention mechanism for data fusion [25] and outperforms static multimodal models in capturing long-term emotional fluctuations [26].

Synergy and Hierarchical Modelling

Sequence models can serve as the perceptual front-end for probabilistic frameworks. For example, an LSTM or Transformer could recognise long-range patterns in a person's behaviour and produce a feature vector. This vector could then be used as a node in a Bayesian network that models causal relationships between high-level factors (e.g., stress → relapse). The Bayesian network can feed its predictions into an RL agent that schedules interventions. Such layering combines deep predictive power, causal interpretability and adaptive control.

Graphical Integration of Methods

To design a comprehensive AI system for behavioural prediction and intervention, the models above can be layered as follows:

1. **Perception layer:** Sequence models (LSTMs or Transformers) extract features from multimodal behavioural data (text, speech, physiological signals). Chaos theory and entropy measures assess variability and detect early warning signals. An HMM or clustering model segments behaviour into discrete states (e.g., mood states).
2. **Causal inference layer:** A dynamic Bayesian network models how high-level variables influence each other over time. State-space models and Kalman filters estimate latent trajectories of emotional states.
3. **Network layer:** Graph neural networks integrate information about social interactions or brain connectivity. Agent-based models simulate group dynamics and evaluate policy changes.
4. **Control layer:** Reinforcement learning optimizes personalised interventions, using predictions from the causal and network layers as state information. The RL agent can use entropy or chaos detection to adjust exploration versus exploitation.

This hierarchical architecture blends probabilistic reasoning with deep learning and complex systems theory. It affords interpretability (via BNs and state-space models), flexibility (via RNNs/Transformers), sensitivity to nonlinear dynamics (via SDEs and chaos), and adaptability (via RL).

Conclusion

Markov chains opened the door to modelling sequential behaviour, but modern AI offers a rich toolkit for understanding human dynamics. Hidden Markov models infer latent psychological states; Bayesian and state-space models capture causal and temporal dependencies; stochastic differential equations describe continuous fluctuations; reinforcement learning learns personalized intervention policies; graph neural networks and agent-based simulations incorporate social and neural networks; chaos theory explains sudden behaviour changes; and information-theoretic measures quantify complexity and predictability. Combining these approaches allows researchers to build robust, interpretable and adaptive systems for behavioural prediction and mental-health interventions.

---

[1] Hidden Markov models reveal behavioral state dynamics in depth-related locomotion in mice | PLOS One
https://journals.plos.org/plosone/article

[2] [3] Dynamic Bayesian network analysis of the social determinants of mental health - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC12281507/

[4] Frontiers | A real-time biphasic Kalman filter-based model for estimating human core temperature from heart rate measurements for application in the occupational field
https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2024.1219595/full

[5] Computational network models for forecasting and control of mental health trajectories in digital applications
https://www.medrxiv.org/content/10.1101/2025.07.03.25330825v1.full.pdf

[6] Modeling Intraindividual Dynamics Using Stochastic Differential Equations: Age Differences in Affect Regulation - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC6433426/

[7] Applying continuous-time models to ecological momentary assessments: A practical introduction to the method and demonstration with clinical data | NPP—Digital Psychiatry and Neuroscience
https://www.nature.com/articles/s44277-024-00004-x

[8] Reinforcement Learning and Its Clinical Applications Within Healthcare: A Systematic Review of Precision Medicine and Dynamic Treatment Regimes - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC12295150/

[9] Can reinforcement learning effectively prevent depression relapse? - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC12362656/

[10] An adaptive multi-graph neural network with multimodal feature fusion learning for MDD detection | Scientific Reports
https://www.nature.com/articles/s41598-024-79981-0

[11] Deep graph learning of multimodal brain networks defines treatment-predictive signatures in major depression | Molecular Psychiatry
https://www.nature.com/articles/s41380-025-02974-6

[12] [13] Developing agent-based models of complex health behaviour - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC6284360/

[14] [17] Embracing Chaos and Complexity: A Quantum Change for Public Health - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC2446457/

[15] [16] A chaotic view of behavior change: a quantum leap for health promotion | International Journal of Behavioral Nutrition and Physical Activity | Full Text
https://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-3-25

[18] [19] Implications of Information Theory for Computational Modeling of Schizophrenia - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC5774180/

[20] [21] Information Theory and Cognition: A Review - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC7513233/

[22] [23] Predicting human decision making in psychological tasks with recurrent neural networks - PMC
https://pmc.ncbi.nlm.nih.gov/articles/PMC9154096/

[24] [25] [26] Frontiers | Personalized prediction and intervention for adolescent mental health: multimodal temporal modeling using transformer
https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2025.1579543/full