Modelling Well-being: Comparing Statistical, ML, Hybrid, and Bayesian Network Models on Psychological Data

1. Introduction

Psychological outcomes (mental health, job satisfaction, life satisfaction) are affected by many interacting factors (personality, stress, sleep, social support, work environments). On a societal level, we capture patterns and correlations between living conditions and mental state academically, intuitively or empirically. While we do have a broad understanding of what needs must be met in order for people to achieve general wellbeing, a question would be how accurately machines capture and understand this intricate structure and predict human psychological health.

Scope of this thesis is a comparison between statistical, machine learning, hybrid and Bayesian Network models for predicting and explaining wellbeing under uncertainty, across individual and societal level data.

This study does not aim to perform clinical diagnosis or psychological assessment. The main focus is modelling self-reported mental health outcomes using different predictive approaches with emphasis on the causal structure of data.

Research Questions:

1. How do different modeling approaches perform when predicting well-being at individual vs societal levels?
2. How does interpretability differ between regression, ML, and Bayesian Networks?
3. Are Bayesian Networks better at capturing plausible causal structures in well-being data?
4. Can Bayesian Networks reveal causal structures that ML models cannot?
5. How robust are these models to missing data, noise and correlated predictors?

2. Sources & Related Work

1. Bayesian Networks & Causality → justify using BNs

Pearl, J. (2009). ***Causality: Models, Reasoning, and Inference***

Points: - framework for causal reasoning using directed graphical models

- establishes how BNs can represent cause-effect relationships rather than just statistical associations

- supports the use of BNs in this project to analyze well-being determinants beyond predictive correlation

Scutari & Denis (2014). ***Bayesian Networks: With Examples in R***

Points: - focus on structure learning, parameter estimation and probabilistic inference from data

- details score-based and constraint-based algorithms

- examples of learning network structures under uncertainty

2. ML vs Statistical Models in Psychology → justify comparison

Yarkoni & Westfall (2017). ***Choosing prediction over explanation in psychology***

Points: - argue that traditional statistical models in psychology prioritize explanation and hypothesis testing at the expense of predictive performance

- machine learning approaches often achieve better predictions for complex human behavior due to their ability to model nonlinear interactions

Breiman (2001). ***Statistical Modeling: The Two Cultures***.

Points: - distinguishes between the data modeling culture (statistical models) and the algorithmic modeling culture (predictive accuracy through machine learning methods)

- states that algorithmic models often outperform traditional statistical approaches when dealing with complex, real-world data

Shmueli (2010). ***To explain or to predict?***

Points: - distinguishes between explanatory modeling (testing theoretical relationships) and predictive modeling (accurate outcome prediction)

- models optimized for explanation often perform poorly in prediction tasks, and vice versa → evaluation strategy: compare models not only based on accuracy but also interpretability and explanatory capability

3. Mental Health / Well-Being Modeling

Diener et al. (1985). ***The Satisfaction With Life Scale***

Points: - introduce the Satisfaction With Life Scale as an instrument for measuring subjective well-being through self-reported life evaluation → strong reliability across diverse populations

- support for using life satisfaction and happiness-related outcomes as meaningful targets in well-being modeling

OECD (2020). ***Measuring Well-Being and Progress*** → validated outcome measures

https://www.oecd.org/en/topics/measuring-well-being-and-progress.html

https://www.oecd.org/en/publications/measuring-population-mental-health_5171eef8-en.html

https://www.oecd.org/en/publications/fitter-minds-fitter-jobs_a0815d0f-en.html

https://www.oecd.org/en/publications/how-to-make-societies-thrive-coordinating-approaches-to-promote-well-being-and-mental-health_fc6b9844-en.html

Points: - assess societal progress beyond traditional economic metrics (economic growth doesn't indicate an increase in human well-being)

- identifies key domains of well-being including income, employment, health, education, social support, work-life balance, trust, and subjective life satisfaction

- mental health as a core component of societal well-being, linking psychological outcomes to economic productivity, social cohesion, and quality of life

4. Interpretability & Explainability

Lundberg & Lee (2017). ***A Unified Approach to Interpreting Model Predictions (SHAP)***

Points: - propose SHAP, a unified framework for interpreting model predictions based on cooperative game theory

- attributes each feature's contribution to individual predictions while ensuring consistency and local accuracy

Molnar (2022). ***Interpretable Machine Learning***

Points: - overview of interpretability methods for machine learning models (feature importance, partial dependence plots, local explanation techniques)

3. Datasets

Individual-Level Well-Being (Micro Scale)

1. Mental Health in Tech Survey (OSMI)

https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey

Characteristics:

- Mostly categorical variables → suitable for BN
- Realistic noise
- Causal intuition

2. Student Mental Health Dataset

https://www.kaggle.com/datasets/shariful07/student-mental-health

Characteristics:

- Simple structure
- Clear binary outcomes
- Educational domain (distinct from workplace)

Focus:

- Workplace factors
- Academic pressure
- Social environment
- Mental health outcomes

Societal-Level Well-Being (Macro Scale)

World Happiness Report (extended Kaggle version)

https://www.kaggle.com/datasets/khushikyad001/world-happiness-report

Level: Country / Year

Characteristics:

- Multi-domain predictors
- Predictions over multiple years
- Interpretable variables
- Suited for causal discussion (economy → stress → well-being)
- Suited for Regression, ML, BN Structure Learning (after discretization)

Focus:

- Economy (GDP, employment)
- Social trust & support
- Work–life balance

- Mental health index
- Institutional stability

Shared predictor categories

Category	Micro Level (Individual)	Macro Level (Societal)
Demographics	Age, gender	Population, urbanization
Work	Remote work, job policies	Employment rate, work-life balance
Stress	Work interference	Unemployment, crime
Social support	Coworkers, supervisors	Social support index
Well-being	Treatment, depression	Happiness, life satisfaction

4. Methodology

Individual-level survey data captures personal and workplace factors, while country-level indicators reflect structural and institutional influences. The same modeling frameworks are applied consistently to both levels to evaluate predictive performance, interpretability, and causal plausibility.

Model Strategy (For all datasets)

Statistical

- Logistic regression (binary outcomes: treatment)
- Linear regression (continuous outcomes: happiness score)
- Ordinal logistic regression (for work_interfere)
- Regularized regression (L1/L2 / Elastic Net)

ML

- Random Forest

- Gradient Boosting (XGBoost / LightGBM)

Hybrid

- Logistic regression + tree-based feature selection
- Statistical model (regression) + ML on residuals: training linear regression, then training XGBoost on residuals to capture nonlinear structure

Bayesian Networks

- Structure learning: defined DAGs, structure learned with PC (constraint-based) + hill-climbing + BIC (score-based)
- Inference: variable elimination to compute posteriors and interventions
- Missing data handling
- Parameters: MLE or Bayesian (Dirichlet priors), handling continuous with discretization

Outcomes to Predict

Micro-level

- work impact
- treatment
- depression / anxiety (student dataset)

Macro-level

- happiness score
- life satisfaction
- mental health index
- work life balance

Preprocessing & feature engineering

- **Missing data:** compare listwise deletion (baseline), simple imputation (+ missingness indicators) for regression/ML, BN handling missing evidence

- **Discretization** (for BN): binning for continuous predictors, 1–2 domain-informed bins as a sensitivity check
- **Feature engineering** (hybrids): encoding categorical values (one-hot for ML/regression, label/category for BN), train/test split strategy for all models, feature selection pipeline (from RF/XGBoost)

Evaluation

Predictive performance

- Regression: RMSE, MAE, R^2, calibration plots
- Classification: Accuracy, Precision/Recall, F1, ROC-AUC, Brier score for probabilities

Probabilistic quality

- Calibration (reliability curves), Brier score

Interpretability & explanation

- **Global interpretability**: regression coefficient, feature importance (RF/GBoost), SHAP values
- **Structural interpretability**: BN as visual DAG, presence of intuitive edges (relationships), edge frequency

Causal & what-if evaluation

- Check plausibility of discovered causal links (is BN structure plausible?)
- What if interventions for each dataset ("If work benefits increase → how does treatment probability increase")

Robustness tests

- Missingness stress test (increase number of missing values, compare models)
- Noise injection (add noise to numeric predictions)

5. Comparative Analysis

6. Implementation & tools (Python)

7. Visualization & UX

- **Model comparison dashboard**: side-by-side panels showing accuracy metrics, calibration plots, and global feature importance for each model type
- **BN interactive viewer**: clickable BN graph; clicking a node lets user set evidence and observe posterior updates in real time (bar charts for categorical or density plots for continuous)
- **Intervention simulator**: slider to change a variable and show predicted effect

8. Discussion & Limitations

9. Conclusion