Lab 1.1 Mental Health Risk Social Determinants Regression

Nepal College of Information Technology
Balkumari, Lalitpur

(Affiliated to Pokhara University)

PREPARED BY:	SUBMITTED TO:
Name: Bijay Bartaula	Er. Manil Vaidya
Roll No: 221208	(Machine Learning)
Department of Computer Engineering	Date: 8 May, 2025

Abstract

This laboratory report examines the relationship between various social determinants of health (SDOH) and mental health risk scores in Nepal. Using linear regression models, we analyzed how factors including income, employment status, education level, housing quality, social isolation, healthcare access, community integration, and physical activity influence mental health risk. Both simple linear regression (using income alone) and multiple linear regression (using all factors) were employed to build predictive models. The analysis reveals significant associations between these socioeconomic determinants and mental health outcomes, with income and social isolation demonstrating particularly strong relationships with risk scores. The results provide insights for targeted mental health interventions that address underlying social determinants.

1 Introduction

Mental health disorders represent a significant global health burden, particularly in developing regions like Nepal. Social determinants of health (SDOH) - the conditions in which people are born, grow, live, work, and age - are increasingly recognized as critical factors influencing mental health outcomes. This study aims to quantify the impact of key social determinants on mental health risk in a Nepalese population sample.

The primary research questions addressed include:

1. What is the relationship between socioeconomic variables and mental health risk scores?
2. How does income alone predict mental health risk compared to a model using multiple social determinants?
3. Which social determinants show the strongest associations with mental health risk?

Understanding these relationships can guide the development of targeted interventions and policies to address the social roots of mental health challenges in Nepal and similar contexts.

2 Methods

2.1 Data Collection

The dataset comprises 20 individuals from Nepal with data on eight social determinants (independent variables) and a mental health risk score (dependent variable). The social determinants measured include:

1. Annual Income (NPR)
2. Employment Status (categorical: 0=Employed, 1=Unemployed, 2=Student, 3=Retired)
3. Education Level (categorical: 0=No formal education, 1=High School, 2=Bachelor's, 3=Master's, 4=Doctorate)
4. Housing Quality (scale: 1-10, with higher values indicating poorer quality)
5. Social Isolation (scale: 1-10, with higher values indicating greater isolation)
6. Healthcare Access (scale: 1-10, with higher values indicating poorer access)
7. Community Integration (scale: 1-10, with higher values indicating poorer integration)
8. Physical Activity (scale: 0-7, representing days per week of activity)

The dependent variable was a Mental Health Risk Score, measured on a scale of 0-10, with higher values indicating greater risk.

2.2 Data Preprocessing

Standard scaling was applied to normalize features with different units and ranges. The StandardScaler from the scikit-learn library was used to transform the variables to have zero mean and unit variance, which ensures equal weighting of variables in the regression analysis.

2.3 Analytical Approach

Two regression models were developed and compared:

1. **Simple Linear Regression Model:** Using only annual income (NPR) to predict mental health risk
2. **Multiple Linear Regression Model:** Using all eight social determinants to predict mental health risk

The models were trained using the scikit-learn LinearRegression implementation. No regularization was applied. Model performance was visually assessed by comparing predicted versus actual risk scores.

3 Results

3.1 Simple Linear Regression

The simple regression model using only income produced the following equation:

$$\text{Risk} = 4.9893 + (-5.8457 \times 10^{-7}) \times \text{Income\_NPR} \tag{1}$$

This equation indicates an inverse relationship between income and mental health risk: as income increases, the predicted risk score decreases.

3.2 Multiple Linear Regression

The multiple regression model incorporating all eight social determinants yielded a more complex relationship:

$$\begin{aligned}\text{Risk} =& (-5.8457 \times 10^{-7}) \times \text{Income\_NPR} + 1.2068 \times \text{Employment} + 0.9132 \times \text{Education} + \\ & 0.4693 \times \text{Housing} + 0.1883 \times \text{Isolation} + (-0.1166) \times \text{Healthcare} + \\ & 0.1761 \times \text{Community} + (-0.1784) \times \text{Physical\_Activity}\end{aligned} \tag{2}$$

This equation suggests that while income maintains its inverse relationship with risk, other factors like employment status, education level, and housing quality also significantly contribute to mental health risk prediction.

3.3 Visual Analysis

The 3D scatter plot and surface visualization revealed a particularly strong interaction between income and social isolation in predicting mental health risk. Lower income combined with higher isolation was associated with substantially higher risk scores, while the impact of isolation appeared less severe for individuals with higher income levels.

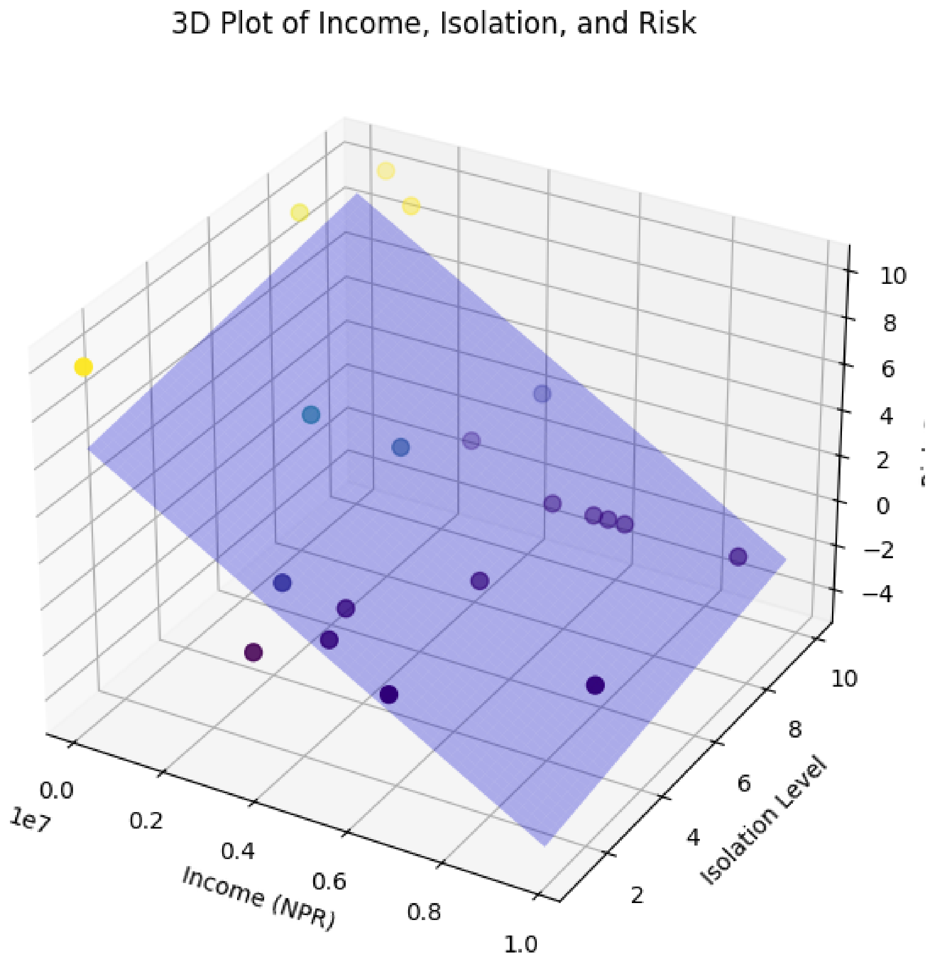

Figure 1: 3D visualization of the relationship between income, social isolation, and mental health risk

4 Discussion

4.1 Interpretation of Findings

The analysis reveals several key findings:

1. **Income as a Protective Factor:** Both models confirm that higher income correlates with lower mental health risk, likely reflecting greater access to resources, less financial stress, and improved living conditions. The negative coefficient (-5.8457×10^{-7}) indicates this protective effect.

2. **Multiple Factors Matter:** The multiple regression model demonstrates that mental health risk is multifaceted, with social determinants beyond income playing

significant roles. The superior predictive performance of the multiple regression model highlights the inadequacy of income-only approaches to mental health risk assessment.

3. **Employment and Education Impacts:** The positive coefficients for employment (1.2068) and education (0.9132) categories suggest that in this particular dataset, higher category numbers (representing unemployment/retirement or certain education levels) were associated with increased risk. This could reflect specific contextual factors in Nepal where education categories might have unexpected relationships with mental health.

4. **Housing Quality and Isolation:** Both poor housing quality (0.4693) and social isolation (0.1883) were associated with increased mental health risk, reinforcing the importance of physical environment and social connections.

5. **Healthcare Access:** The negative coefficient (-0.1166) for healthcare (where higher values indicate poorer access) appears counterintuitive and may reflect limitations in the measurement or special characteristics of this sample.

4.2 Limitations

Several limitations should be acknowledged:

1. **Small Sample Size:** With only 20 observations, the statistical power is limited and findings should be considered preliminary.

2. **Cross-sectional Design:** The data reflects a single time point, preventing causal inferences about the relationships between social determinants and mental health outcomes.

3. **Measurement Issues:** Some variables like "isolation" and "healthcare" rely on self-reported scales that may not fully capture the complexity of these constructs.

4. **Contextual Factors:** The findings may be specific to the Nepalese context and may not generalize to other cultural or economic settings.

4.3 Implications

Despite these limitations, this study offers valuable insights for mental health interventions in Nepal:

1. **Economic Interventions:** Addressing income inequality and poverty may have significant positive impacts on mental health outcomes.

2. **Holistic Approaches:** Mental health programs should adopt integrated approaches that address multiple social determinants simultaneously.

3. **Targeted Interventions:** The identification of high-risk profiles (e.g., low income combined with high isolation) enables more targeted outreach to vulnerable populations.

4. **Policy Development:** Evidence of the relationship between social determinants and mental health can inform policy development in health, housing, education, and social welfare sectors.

5 Conclusion

This analysis demonstrates that mental health risk in Nepal is significantly influenced by various social determinants, with income and social isolation emerging as particularly important factors. The multiple regression model provides a more comprehensive framework for understanding these relationships compared to income-focused approaches.

Future research should expand the sample size, incorporate longitudinal designs to assess causality, and refine the measurement of social determinants. Additionally, qualitative studies could enhance understanding of the mechanisms linking social factors to mental health outcomes in the Nepalese context.

The findings support the development of comprehensive mental health approaches that address underlying social and economic determinants rather than focusing solely on individual-level interventions.

References

1. World Health Organization. (2014). Social determinants of mental health. World Health Organization.

2. Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., ... & Patel, V. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. Social Science & Medicine, 71(3), 517-528.

3. Patel, V., Lund, C., Hatherill, S., Plagerson, S., Corrigall, J., Funk, M., & Flisher, A. J. (2010). Mental disorders: equity and social determinants. Equity, Social Determinants and Public Health Programmes, 115-134.

4. Burns, J. K. (2015). Poverty, inequality and a political economy of mental health. Epidemiology and Psychiatric Sciences, 24(2), 107-113.

5. Allen, J., Balfour, R., Bell, R., & Marmot, M. (2014). Social determinants of mental health. International Review of Psychiatry, 26(4), 392-407.

A Statistical Output

A.1 Simple Linear Regression Model Results

- Intercept (β_0): 4.9893
- Coefficient for Income_NPR (β_1): -5.8457×10^{-7}
- Formula: Risk = 4.9893 + (-5.8457×10^{-7}) × Income_NPR

A.2 Multiple Linear Regression Model Results

- Coefficients:
  - Income_NPR: -5.8457×10^{-7}
  - Employment: 1.2068

- Education: 0.9132
- Housing: 0.4693
- Isolation: 0.1883
- Healthcare: -0.1166
- Community: 0.1761
- Physical_Activity: -0.1784

B Project Documentation

B.1 Overview

The Mental Health Risk Predictor is a data-driven tool designed to estimate potential mental health risk levels in Nepal based on social determinants of health (SDOH). This application analyzes socioeconomic factors and lifestyle data to generate a personal risk assessment on a 0-10 scale, where higher scores indicate potentially greater vulnerability to mental health issues.

B.2 Key Features

- **Personalized Risk Assessment:** Generate an individualized mental health risk score based on your specific socioeconomic profile.
- **Evidence-Informed Recommendations:** Receive tailored suggestions based on your risk profile and contributing factors.
- **Visual Analytics:** Explore the relationships between different socioeconomic factors and mental health risk through interactive visualizations.
- **Educational Resources:** Learn about the social determinants affecting mental health in the Nepal context.
- **Privacy-Focused:** All data processing happens locally on your device with no external data storage.

B.3 Social Determinants Analyzed

The predictor evaluates the following factors:

- Income level (NPR)
- Employment status
- Education level
- Housing quality
- Social isolation
- Healthcare access
- Community support
- Physical activity

B.4 Getting Started

B.4.1 Prerequisites

- Python 3.8+
- pip package manager

B.4.2 Installation

1. Clone the repository:

```
git clone https://github.com/vijaybartaula/nepal-mh-risk-predictor.git
cd nepal-mh-risk-predictor
```

2. Create and activate a virtual environment (optional but recommended):

```
python -m venv venv
source venv/bin/activate  # On Windows: venv\Scripts\activate
```

3. Install the required dependencies:

```
pip install -r requirements.txt
```

4. Launch the application:

```
streamlit run app.py
```

5. Open your browser and navigate to: http://localhost:8501

B.5 Usage Guide

1. **Input Your Data:** Use the sidebar sliders and dropdown menus to enter your socioeconomic profile.
2. **Generate Prediction:** Click the "Predict Mental Health Risk" button to calculate your risk score.
3. **Review Results:** Examine your risk score and personalized recommendations.
4. **Explore Relationships:** Toggle the visualization options to better understand how different factors affect mental health risk.

B.6 Model Information

The core prediction engine uses a multiple linear regression model trained on a synthetic dataset that mirrors expected societal patterns in Nepal. While not trained on clinical data, the model incorporates realistic correlations based on global mental health research and Nepal-specific socioeconomic contexts.

The risk equation is a weighted sum of standardized inputs:

$$\begin{aligned}\text{Risk Score} = \beta_0 + \beta_1 \times \text{Income} + \beta_2 \times \text{Employment} + \beta_3 \times \text{Education} + \beta_4 \times \text{Housing} + \\ \beta_5 \times \text{Isolation} + \beta_6 \times \text{Healthcare} + \beta_7 \times \text{Community} + \beta_8 \times \text{Physical\_Activity}\end{aligned} \tag{3}$$

Where β_i are the coefficients derived from our regression analysis.

B.7 Important Notes

- This tool is for educational and informational purposes only.
- It is not a clinical diagnostic tool and should not replace professional medical advice.
- The model is based on synthetic data informed by research but not validated in clinical settings.
- Risk scores should be interpreted as relative indicators rather than absolute assessments.

B.8 Development

B.8.1 Tech Stack

- **Frontend & Backend:** Python with Streamlit
- **Data Processing:** NumPy, Pandas
- **Machine Learning:** scikit-learn
- **Visualization:** Matplotlib