Predicting Depression Prevalence Using Deep Learning

DATA 5610 Final Project  
Spring 2025  
Student: Carter Grant

# Page 1: Introduction – What is the Problem and Why Does it Matter?

## Page 1: Introduction – What is the Problem and Why Does it Matter?  
  
Depression is one of the most pervasive public health challenges in the United States and globally. According to the CDC, the rate of adults who report having been diagnosed with depression has increased significantly over the last two decades. As policymakers, public health officials, and mental health professionals seek to better allocate resources and improve early intervention, understanding the demographic and geographic patterns of depression prevalence is crucial.  
  
This project addresses the question: \*\*Can we accurately predict depression prevalence using group-level demographic, geographic, and temporal data?\*\* Specifically, we aim to identify which population subgroups are most at risk and explore how those risks vary across states and over time.  
  
To explore this question, we use data from the Behavioral Risk Factor Surveillance System (BRFSS), which provides an extensive, annual, state-level survey of adult health behaviors and conditions, including mental health diagnoses. Our goal is to develop a deep learning model that can estimate depression prevalence across subgroups and serve as a predictive tool for public health analysis.  
  
The proposed approach leverages a feedforward neural network trained on one-hot encoded demographic, state, and year features, along with sample size normalization. Our key finding is that this model significantly outperforms traditional baselines (random forest, linear regression) in predicting prevalence values. Moreover, the model identifies interpretable subgroup trends that align with known disparities in mental health.  
  
This paper documents our process from data selection to model evaluation and concludes with next steps for enhancing prediction power and real-world application.

# Pages 2–3: Approach – How Did You Address the Question?

## Pages 2–3: Approach – How Did You Address the Question?  
  
### Data  
  
We use the publicly available BRFSS dataset (2011–present), focusing specifically on responses to the question: \*“Has a doctor or other healthcare provider ever told you that you have a form of depression?”\* We filter for respondents who answered "Yes" and then group the data by:  
  
- \*\*Year\*\* (2011 to most recent available)  
- \*\*State\*\* (Locationabbr code)  
- \*\*Demographic group\*\*, defined by combinations of gender, income, education, race, etc.  
  
The aggregated dataset includes:  
- `Data\_value`: Percentage of individuals in each group reporting a depression diagnosis  
- `Sample\_size`: Number of individuals in the group (used as a weighted feature)  
  
To make this data suitable for machine learning:  
- We combined Break\_Out\_Category and Break\_Out into a single "Demographic" field  
- Applied \*\*one-hot encoding\*\* to categorical features (`Year`, `Locationabbr`, `Demographic`)  
- Scaled sample size using `StandardScaler`  
- Split the dataset into training and test sets (80/20)  
  
We considered the sample size as a numeric input feature to control for statistical confidence across demographic groups. Groups with extremely low sample sizes were still retained to explore generalization across sparsely represented subpopulations.  
  
### Methodology  
  
#### Model Selection  
  
We implemented a fully connected \*\*feedforward neural network\*\* in PyTorch with the following architecture:  
- Input layer (number of encoded features + 1 for sample size)  
- Hidden layers: `[512 → 256 → 64]`  
- Activations: ReLU  
- Regularization: Batch Normalization + Dropout (tested at 0.1–0.3)  
- Output: A single continuous value (predicted depression prevalence)  
  
#### Training & Optimization  
  
- Optimizer: Adam  
- Loss Function: Mean Squared Error (MSE)  
- Learning Rate: 0.001  
- Scheduler: ReduceLROnPlateau (patience = 5, factor = 0.5)  
- Epochs: 150  
- Batch Size: 64  
  
We used both validation loss and final test performance for model selection. A \*\*grid search\*\* over dropout rates and hidden layer widths helped determine the optimal architecture. Additionally, we implemented \*\*checkpoint saving\*\*, storing the best-performing model by validation loss throughout training.  
  
Baseline comparisons included:  
- \*\*Linear Regression\*\* (with same encoded inputs)  
- \*\*Random Forest Regressor\*\*  
  
These benchmarks helped evaluate the value of nonlinear modeling via deep learning.

# Pages 4–5: Analysis and Results – What Did You Find?

## Pages 4–5: Analysis and Results – What Did You Find?  
  
### Model Performance  
  
The final model achieved the following test metrics:  
- \*\*RMSE\*\*: [INSERT RMSE]  
- \*\*MAE\*\*: [INSERT MAE]  
- \*\*R²\*\*: [INSERT R2]  
- \*\*Adjusted R²\*\*: [INSERT ADJ\_R2]  
  
Compared to baselines:  
- \*\*Random Forest\*\*: R² = [INSERT], Adjusted R² = [INSERT]  
- \*\*Linear Regression\*\*: R² = [INSERT], Adjusted R² = [INSERT]  
  
These results demonstrate that the neural network outperformed both baselines in terms of overall predictive accuracy and generalization.  
  
### Predicted vs. Actual Plot  
  
Our predicted vs. actual scatterplot shows a strong linear alignment, with most points closely following the y=x line. The model slightly underpredicts for subgroups with very high depression prevalence (outliers) and overpredicts for low-prevalence groups with small sample sizes. However, on average, the model tracks observed trends well.  
  
### Subgroup Analysis  
  
We ranked the top 10 most at-risk groups based on predicted depression prevalence. These often included:  
- Female-identifying groups with low income  
- Individuals with less than a high school education  
- Hispanic and multiracial demographic groups in certain Southern and Midwest states  
  
The most at-risk state-level averages were found in [INSERT STATES], while the least affected included [INSERT STATES].  
  
### Year-by-Year Trends  
  
A time series plot comparing predicted vs. actual prevalence over the years showed high agreement, confirming that the model learns temporal dynamics effectively. Notably, the model captured the slight national rise in prevalence after 2015, aligning with known mental health trends during that period.

# Page 6: Discussion, Conclusions, and Next Steps

## Page 6: Discussion, Conclusions, and Next Steps  
  
### Interpretation of Findings  
  
Our model suggests that a well-tuned deep learning approach can learn complex patterns in mental health data, even when limited to categorical and group-level inputs. By using demographic, geographic, and temporal information, the model was able to estimate prevalence with high accuracy and generalize across underrepresented populations.  
  
This not only supports our hypothesis but provides a powerful tool for health departments and NGOs to monitor mental health risks more granularly than national averages allow.  
  
### Limitations  
  
Several limitations should be acknowledged:  
- BRFSS data is self-reported and subject to biases in recall and diagnosis  
- Groups with small sample sizes are overrepresented in model error  
- One-hot encoding is inherently sparse and doesn't capture relationships between groups (e.g., similarity between states or income brackets)  
- The model doesn't incorporate external context (economic, policy, pandemic-related changes)  
  
### Conclusion  
  
Deep learning models are capable of producing accurate, interpretable predictions of depression prevalence across demographic groups. Our work shows the viability of this approach, outperforming traditional regressors and providing actionable insights.  
  
### Future Work  
  
Opportunities to build on this project include:  
- Using \*\*embedding layers\*\* or \*\*transformers\*\* for categorical inputs  
- Adding external features like unemployment rate or insurance coverage  
- Deploying the model in a real-time dashboard for public health departments  
- Applying explainability tools (e.g., SHAP) to better understand feature influence  
  
We hope this work helps inform both methodological choices and strategic health planning moving forward.

## Appendix A: Visualizations

• Predicted vs. Actual Scatterplot

• Training vs. Validation Loss Curve

• Top 10 At-Risk Demographic Subgroups Table

• Yearly Depression Prevalence Trend Plot

## Appendix B: Code Excerpt – Final Model Definition

Refer to the Jupyter notebook file for the complete model architecture and grid search implementation.  
The final model structure is as follows:

class EnhancedDepressionNN(nn.Module):  
 def \_\_init\_\_(self, input\_dim):  
 super(EnhancedDepressionNN, self).\_\_init\_\_()  
 self.model = nn.Sequential(  
 nn.Linear(input\_dim, 512),  
 nn.ReLU(),  
 nn.BatchNorm1d(512),  
 nn.Dropout(0.3),  
 nn.Linear(512, 256),  
 nn.ReLU(),  
 nn.BatchNorm1d(256),  
 nn.Dropout(0.2),  
 nn.Linear(256, 64),  
 nn.ReLU(),  
 nn.Linear(64, 1)  
 )  
 def forward(self, x):  
 return self.model(x)