MUKISA Vicent, 2025/HD05/26353U, 2500726353

Project Proposal: Predicting Mental Health Risk and Help-Seeking in the Tech Industry

1. Background

Mental health challenges in the technology sector are increasingly recognized as major concerns affecting productivity, workplace satisfaction, and individual well-being. The *Open Sourcing Mental Illness (OSMI)* survey collects demographic, workplace, and health-related responses from professionals in technology. While existing work often focuses on simple classification of "mental health condition: yes/no," this project aims to go further by exploring risk factors, workplace influence, and help-seeking behavior through machine learning models.

2. Objectives

1. **Predictive Modeling**: Build machine learning models to predict whether a respondent is likely to seek treatment for mental health issues based on survey features.
2. **Feature Importance**: Identify key demographic and workplace-related factors influencing help-seeking.
3. **Fairness Analysis**: Evaluate whether model predictions vary across sensitive groups such as gender or region.
4. **Model Comparison**: Compare the performance of baseline classifiers (Logistic Regression, Naïve Bayes, Decision Trees) with regularized models (Ridge, Lasso) using cross-validation.
5. **Interpretability**: Provide actionable insights into which workplace variables (e.g., supervisor support, openness about mental health) have the strongest predictive power.

3. Dataset

- **Source:** OSMI Mental Health in Tech Survey (Kaggle public dataset).
- **Format:** Survey responses with categorical (e.g., gender, family history, supervisor support) and numerical features (e.g., age, hours worked).
- **Target Variables:**
    - *Primary:* Whether the respondent has sought treatment.
    - *Secondary:* Perception of workplace support/stigma.

4. Methodology

1. **Data Preprocessing**: Clean inconsistent entries, handle missing values, encode categorical features, and scale numeric variables.
2. **Exploratory Data Analysis**: Visualize demographic and workplace trends in relation to mental health treatment.
3. **Model Development**:
    - Baseline models: Logistic Regression, Naïve Bayes.
    - Advanced models: RidgeCV, LassoCV, Decision Trees.
    - Pipelines for encoding + scaling + model training (per Lecture 4).
4. **Evaluation**: Use a stratified train/test split and cross-validation. Report Accuracy, Precision, Recall, F1, ROC-AUC.
5. **Fairness & Interpretability**: Evaluate model bias across gender/region and apply SHAP/LIME for feature importance.

5. Expected Outcomes

- A predictive model that identifies individuals at higher risk of mental health challenges or those less likely to seek help.
- A ranked list of workplace and demographic factors most strongly associated with treatment-seeking.
- A fairness analysis highlighting whether the model’s predictions are equitable across groups.
- Practical recommendations for improving workplace mental health support policies in the tech industry.

How to Ensure Good Model Performance

Clean the data

- Fix missing values, unify categories (e.g., gender), and remove outliers.

Preprocess features

- Scale numbers (StandardScaler/MinMaxScaler).
- Encode categoricals (OneHotEncoder).
- Use Pipelines to avoid leakage.

Cross-validation

- Always validate with k-fold CV, not just one split.

Tune hyperparameters

- Use GridSearchCV/RandomizedSearchCV for best alpha, C, max_depth, etc.

Handle imbalance

- Use class_weight='balanced', oversampling (SMOTE), or undersampling.

Interpret results

- Coefficients (Logistic/Lasso), Feature importance, SHAP values.