

# Analyzing Mental Health Attitudes in the Tech Industry: A Predictive Modeling Approach

Preethi Nigam Errala, Sai Kumary Chary Sripathi, Satya Valluri, Aneesh Reddy Pappireddy

## Background

- Mental-health concerns are widespread yet often under-reported among technology workers [1].
- High workloads, tight deadlines, and workplace stigma amplify psychological risks.
- Early identification is crucial—data-driven tools can flag employees who may need support.
- Research Question: Can machine-learning models predict whether a tech employee will seek mental-health treatment?

## Methods

**Dataset:** 2016 OSMI “Mental Health in Tech” survey; n = 1,386 after removing rows with > 30 % missing data (63 variables).

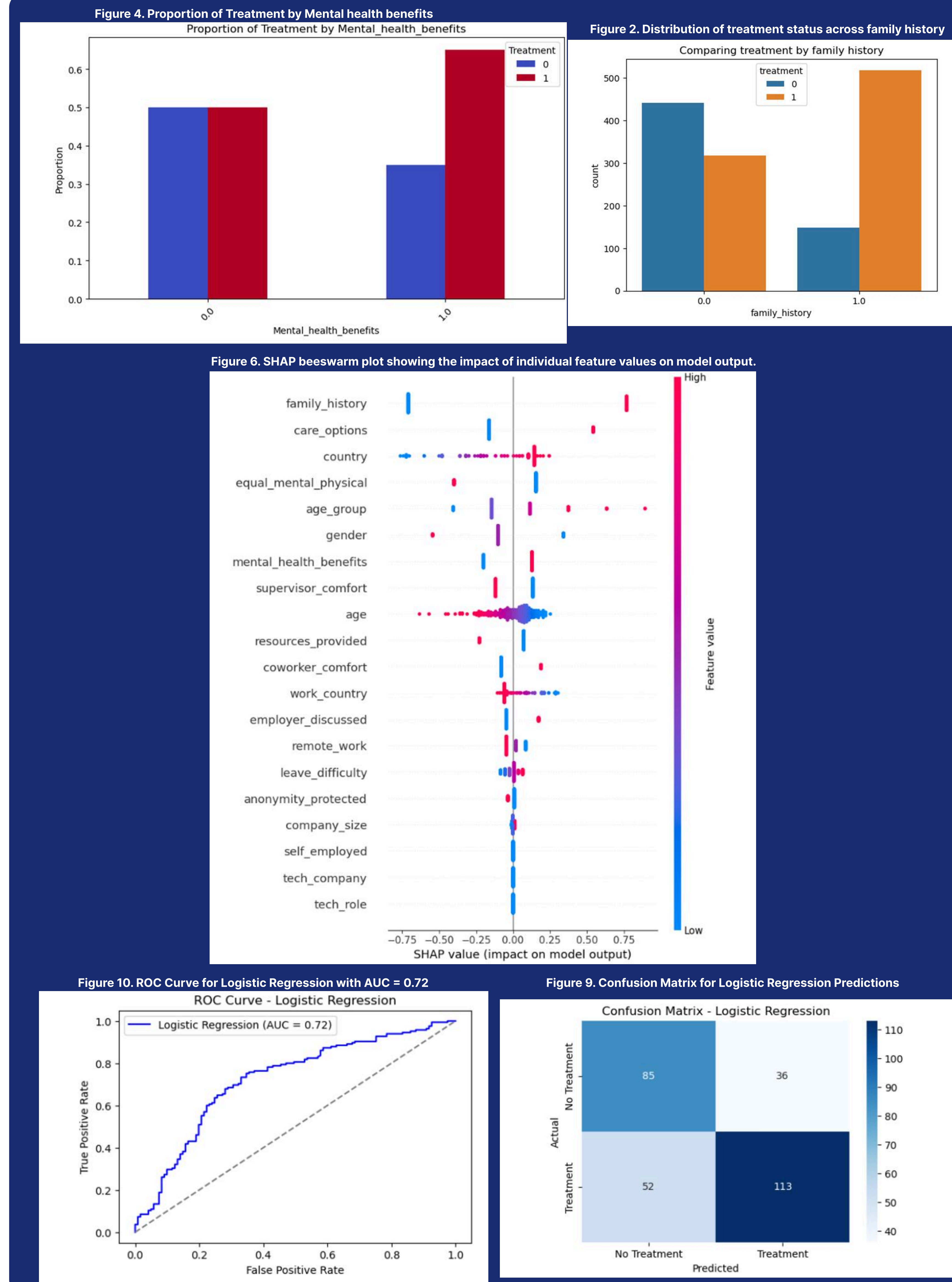
**Cleaning & Preparation:** Median / mode imputation, one-hot and ordinal encodings, z-score scaling. Class imbalance handled with `class_weight="balanced"` and SMOTE oversampling [3].

**Feature Engineering:** Region-level country grouping, Manager-Support Index (mean of three Likert items), Remote-Worker flag.  
**Models:** Logistic Regression, Random Forest, XGBoost, Support-Vector Machine (linear & RBF) plus majority-class baseline.

**Tuning & Evaluation:** GridSearchCV (5-fold stratified) on an 80 / 20 train–test split. Primary metrics: Accuracy & ROC-AUC; secondary: Precision, Recall, F1.

**Interpretability:** SHAP beeswarm for global feature importance [2]; force plots for representative cases.

**Stack:** Python 3.10 • scikit-learn 1.5 • XGBoost 2.0 • imbalanced-learn 0.13 • SHAP 0.44.



## Results

- Best Model: Logistic Regression (Accuracy = 69.2 %, AUC = 0.72); SVM-RBF close second (68.2 %).
- Top Predictors (via SHAP) [2]:
  - Family history of mental illness
  - Employer mental-health benefits
  - Perceived workplace support / care options
  - Country of employment
- Balanced precision–recall achieved despite class imbalance [3].

## Conclusion

- Machine-learning can moderately predict treatment-seeking behavior among tech workers.
- Employer benefits and stigma-reduction policies strongly influence outcomes [1].
- Transparent, interpretable models (e.g., SHAP) enhance stakeholder trust [2].
- Future Work: Collect larger longitudinal datasets, perform fairness / bias audits, and develop targeted interventions.

## References

- [1] World Health Organization. (2008). Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors. Geneva: WHO Press.
- [2] Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In Advances in Neural Information Processing Systems 30 (pp. 4765-4774).
- [3] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321-357.