Application of Machine-Learning in Identifying Mental Health Issues in the Tech Industry

Group members: Wanlin Zhou, Yifei Wang, Kunqi Zhao, Yuan Lin
Feb 17, 2024

1. Background, Motivation, & Business Question

In today's fast-paced tech industry, the mental well-being of employees is critical. We're here to provide solutions to address the issue. As an HR consultancy focusing on mental health, we serve tech companies across North America. We utilize comprehensive data from the "Mental Health in Tech Survey" to build tools to identify the signs when employees might be struggling with their mental health, particularly when it's affecting their work and well-being.

Our approach involves machine learning models. These models are designed to identify key predictors of mental health challenges, such as the extent of work interference caused by mental health issues, the availability of employer-provided mental health benefits, employee awareness of mental health care options, etc, to identify employees who may be silently enduring hardships. Our insights are invaluable for tech companies aiming to customize their Mental Wellness Programs. It's about reaching the employees who are adversely affected by mental health issues without seeking help.

By proactively initiating and continuously refining a Mental Wellness Program, a company does more than aid its employees; it lays the groundwork for a thriving business environment. This strategy leads to improved employee performance, heightened job satisfaction, better retention rates, and a deeper commitment to the company.

2. Data and Statistical Questions

As per the 2023 Mental Health in Tech Report, an increasing number of individuals within the technology sector are facing elevated levels of depression and anxiety regarding their career outlooks. A mentally distressed workplace poses various challenges, including a substantial employee turnover rate, diminished morale, and potential reputation issues within the public domain. To help our clients address the issue, we analyzed the dataset derived from a survey on mental health in the tech workplace.

2.1 The Statistical Questions and The Usefulness

As for the current data, we can develop a machine-learning model to predict if an employee would seek treatment based on the selected features from the survey. We've utilized various models and chosen one that could deliver the best prediction to help our client address their requirements.

It is a classification problem because our target is to see if the employee took the action to seek treatment or not. This model is useful in managing our clients' mental health conditions by providing key indicators that would drive the employees to take action. We've learned that some tech companies may have developed their wellness program but still lack informative methods to help their employees seek treatment. One of the primary goals of the model is to help our clients improve their wellness programs or raise awareness and utilize this tool to assess their effectiveness.

2.2 The Limitation of The Questions

The model may fall short on the following problems: First, it only considers the causes of mental health within the organizations, which means there might be other reasons (relationship problems, personal traumas) that cannot be explained by the model. Secondly, the classification model cannot assess the severity or types of mental issues; further information might be needed if the clients want to utilize the prediction results.

2.3 Feature Engineering Plan

For better model performance, we will apply a transformation on different features based on features' data types, such as imputation, scaling, and encoding. We will also remove some features based on the values of its information. The detailed processing of the feature will be illustrated in the EDA and Preprocessing sections.

2.4 Reflection on Selected Performance Metrics

Since we chose a classification model to solve the problem, we needed to perform a confusion matrix to derive accuracy, recall, precision, and f1 to select our model. As one of the main purposes of the model is to assess if the company's current wellness program is effective, we considered the precision matrix to be more informative in this case. In addition, we need to put more weight on the F-1 score since it considers both precision and recall, providing a balanced assessment of a model's accuracy.

3. Exploratory Data Analysis

3.1 The Dataset

This dataset originates from a 2014 study focused on assessing perceptions of mental health and the prevalence of mental health conditions within the technology industry.

3.2 Split Data

We split the data into a training set (train_df) and a test set (test_df).

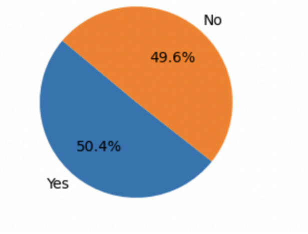
3.3 Describe Data

There are in total, 1,259 rows and 26 columns in the loaded datasets. For the training set, there are 1007 rows. Each row of the data records a survey response of an individual in the technical industry. The columns include information about the individual's personal attributes and survey responses. Age' is the only numeric feature that will be discussed in the later part; the others are categorical features.

We will use the column "treatment" as our target value in the prediction model. As 'treatment' is categorical, We will use the classification method in analysis.

After checking the missing value, we find that 'comments' and 'state' have a large percentage of missing values (87% and 42%, respectively); we will exclude them from the model-building process and preprocess work_interfere and self_employed, which have 21% and 2% missing value respectively.

3.4 Discussion on Treatment



Graph 1: Distribution of Treatment in Training Set

The distribution of 'Yes' and 'No' is equal in the training dataset, showing we have a balanced dataset.

3.5 Discussion on Country and State

Table 1: Count on Interviewees from Different Countries

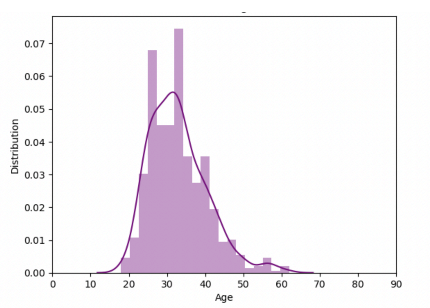
United States	593
United Kingdom	149
Canada	55
Germany	41
Netherlands	23
Ireland	22
Australia	18
France	11
New Zealand	8
Switzerland	6

The countries with the most representation are the United States, the United Kingdom, and Canada. Therefore, we have chosen to focus on the United States, the United Kingdom, and Canada for our discussion. This decision is based on the shared cultural similarities across English-speaking countries, allowing us to retain the broadest scope of our data. The state column is specific to the United States, and given that it only encompasses five states, its representativeness is limited. This further supports the decision to omit it.

3.6 Discussion Gender

Observing the variety of responses provided for gender, we cleaned our data by only applying four categories ('Trans', 'CIS female', 'CIS male', and 'Other') to ensure consistency in further analysis.
While incorporating gender into specific machine learning procedures can lead to bias, our findings indicate that, in our scenario, workplace gender differences contribute to certain mental health issues, as evidenced by research[1].

3.7 Age



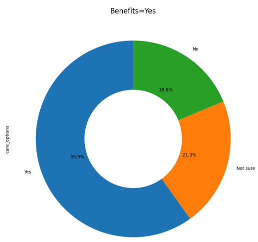
Graph 2: Distribution of Age in Training Set

[1] Harnois, C. E., & Bastos, J. L. (2018). Discrimination, Harassment, and Gendered Health Inequalities. *Journal of Health and Social Behavior, 59*(2), 283-299. American Sociological Association.

Upon examining the age data, we identified certain implausible values, including ages below 18 (the legal working age) and those above 70 (retirement age). Consequently, we removed these outliers from the dataset. After analyzing the age distribution, we discovered that the majority of individuals fall within the 25-40 age range.

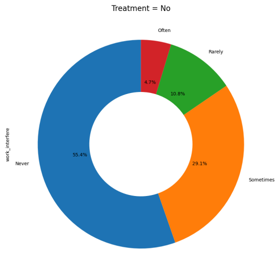
3.8 Discussion on Company-Related Features

In the previous step, we discovered that there is a feature tech_company, indicating that the data might include survey data from non-tech companies' employees. Thus, we applied the following filter to guarantee the accuracy of the data and would remove this feature before modeling.



Graph 3: Distribution of Care_options When Company Provides Benefits

Among companies offering mental health benefits, approximately 40% of employees are either unaware of or uncertain about the availability of these care options.

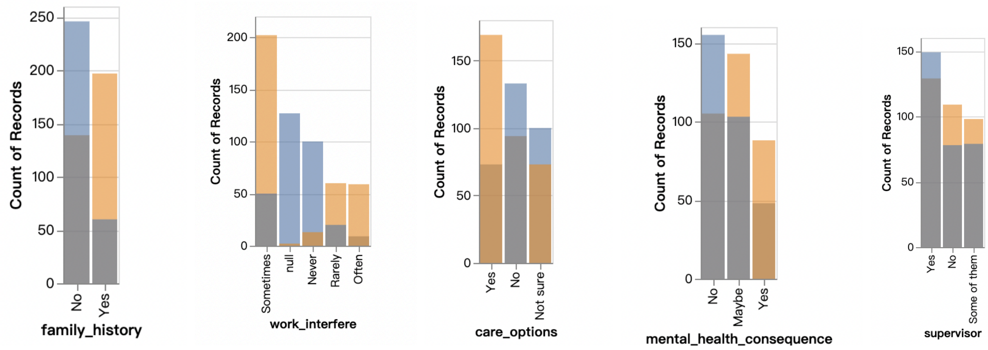


Graph 4: Distribution of Work_interfere When Interviewees Have Not Sought Treatment

Among individuals who have not sought treatment for a mental health condition, 35% still report that it interferes with their work.

3.9 Identify Potential Features

We used visualizations to see what features may contribute to the targets values. 'family_history','work_interfere','care_options','mental_health_consequence', and 'supervisor' are the features showing the most importance.



Graph 5: Distribution of Potential Features Based on Treatment

4. Preprocessing

4.1 Separating feature vectors and targets

After all the data cleansing, we could now split the training sets and test sets into target values ("treatment") and features. Meanwhile, we dropped the features that we think are not appropriate to include, as discussed in EDA, including "country", "comments", "stats", and "tech_company".

4.2 Building ColumnTransformer

We separated the features by their data types to go through different transformations. For the numeric features, we used SimpleImputer to fill the NaN values with the median of the feature, then used StandardScaler to implement scaling transformation; For the categorical features, we conducted imputation by filling in the missing values with "missing" and applied OneHotEncoder to encode categorical values. To prevent new categorical values from appearing in the test set and causing errors, we set the parameter "handle_unknown" to "ignore". Additionally, in EDA, we knew that some of the categorical features are binary, so we used the drop = "if_binary" parameter. Finally, we built a ColumnTransformer to combine different routes for transformations.

4.3 Model Pipelines

We first built a dummy model as a baseline. Then we explored various models that have been covered in our learning curriculum, applying them to the dataset. They are the Decision Trees model, k-Nearest Neighbours model, SVM model, and Logistic Regression model. To deepen our understanding and application of machine learning techniques, we integrated the random forest model into our project. We made a pipeline for each model we used, incorporating the preprocessor we built above to process the input data.

Given that our target variables are both categorical and binary, we employed classifier models in this process.

5. Model Evaluation and Hyperparameter Tuning

At this stage, we first used cross-validation to have a glance at the accuracy of the training set and test set. We then conducted hyperparameter tuning to improve accuracy performance.

Since the size of our datasets is relatively smaller, we used grid search to find the best hyperparameters and set all the cross-validation fold numbers to 5.

(For convenience, we only displayed the **accuracy after hyperparameter tuning** in this report.)

Table 2: Information of the Fitted Models after Hyperparameter Tuning

Model	Training set score	Test set score	Best Parameter
Dummy	0.52	-	-
Decision Tree	0.85	0.80	max_depth: 2 min_samples_leaf': 10
KNN	-	-	-
SVM	0.91	0.80	C: 1.0 gamma: 0.1
Logistic Regression	0.86	0.83	C: 20
Random Forest	1.00	0.81	n_estimators: 80

While other models can run normally, we cannot get results from KNN models. If we directly use the pipeline to fit and score the training set, an error message of the 'NoneType' object that has no attribute 'split' will be shown.

For the Random Forest model, the n_estimators are 80. It means the model has 80 decision trees in the forest. The higher this hyperparameter is, the more complex and accurate the model is. The highest accuracy score on the test set is 0.79. Yet the training score after fitting is 1, indicating the high possibility of overfitting.

6. Model Evaluation and Selection

For each model we would generate a report including its accuracy, recall, precision, and f1. By comparing these matrices with high priorities for precision and F-1 scores, we selected the model with the best performance.

We would choose from the models that could run successfully and have higher performance than the dummy model. They are the decision tree model, SVM model, logistic regression model, and random forest model.

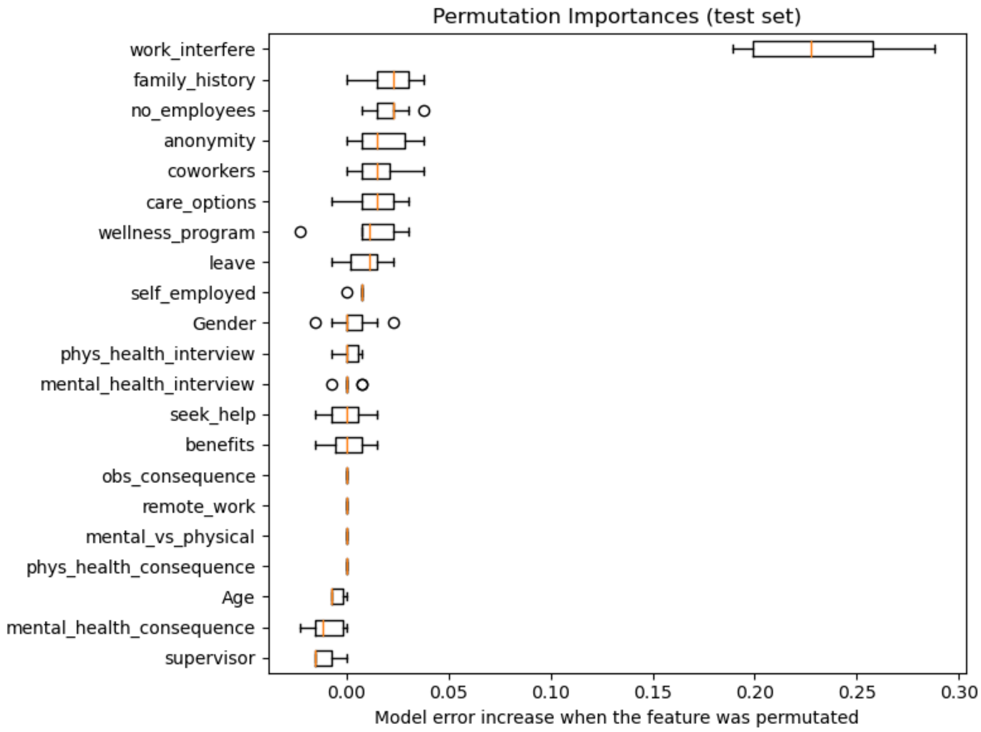
Table 3: Model Evaluation Metrics

Model	Accuracy	Precision	Recall	F-1 score
Logistic Regression	0.826	0.817	0.893	0.854
Random Forrest	0.811	0.784	0.920	0.847
Decision Tree	0.803	0.763	0.947	0.845
SVM	0.803	0.769	0.933	0.843

As we discussed in section 2.4, we will emphasize the F1 metric when selecting the model. Among the models we trained and fitted, the Logistic Regression model has the highest scores on both accuracy, precision, and F-1 Thus, we selected the Logistic Regression model with C = 20 to be our final prediction model.

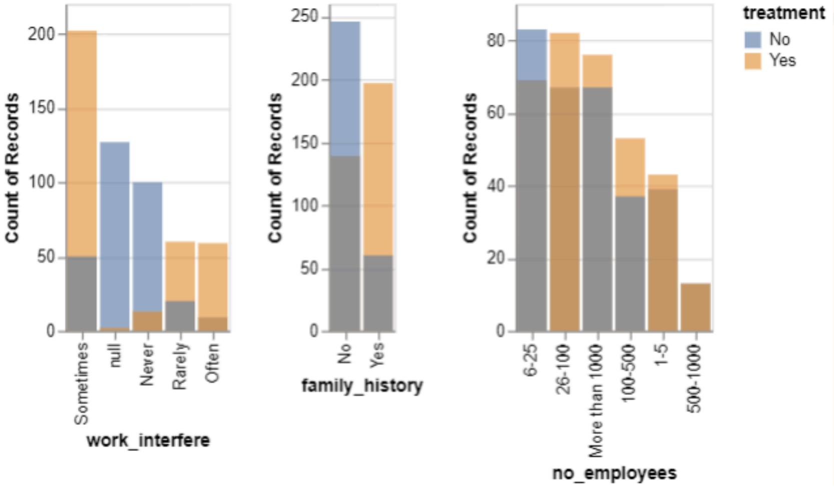
7. Feature Importances

Currently, we have developed a well-performing model that can predict whether employees require mental health treatment. Based on this, we can examine which features within the model have a significant impact on our target value. These features can provide some directions for companies in designing their Mental Health Well-being Programs.



Graph 6: Permutation Feature Importances

Based on the figure above, we can identify three important features: "work_interfere", "family_history", and "no_employees". They respectively represent the assessment of "judgment of the extent to which mental illness may interfere with their capability with work", "the presence of a family history of mental illness", and "the number of employees in the company".



Graph 8: Distribution of Important Features Based on Permutation Importances

We dive deeper into these features to see how may they influence the possibility that an individual should seek for mental health treatment:

- The more the employees sense that mental health problems are interfering with their work, they are more likely to seek/need treatment.
- If an employee has a family history of mental health, they are more likely to seek/need mental health treatment.
- For companies with employee numbers from 26 to 500 and above 1000, their employees are more likely to seek/need treatment.

8. Communication of Results

As an HR consultancy committed to enhancing employee well-being and mental health, we advocate for the critical need of tech companies to embrace Mental Wellness Programs as a vital part of their long-term business strategy. Our developed model functions as a proactive early-warning system, analyzing key indicators that may signal an employee's mental health struggles. Essential factors such as work interference, family history, and company size are identified by the model as critical in determining the need for intervention. This insight is invaluable, guiding tech companies to effectively allocate resources and focus areas. It is imperative for companies to not only provide, but also to educate employees on recognizing mental health concerns and accessing available support resources. This proactive approach is essential in fostering a supportive workplace environment, demonstrating a company's commitment to its employees' mental health and overall well-being.

Here’s how a tech company applies our model to customizing its Mental Wellness Program:

- Prioritize Interventions: Focus on the factors with the highest importance in the model. For example, as ‘work_interfere’ shows high importance, if an employee frequently finds their work interfering with their mental health issues, this could be a sign they need support.
- Increase Awareness: Use the model to identify employees who might be unaware of mental health care options provided by the company and inform them how to access these resources.
- Continuously Improve Mental Wellness Programs: Tailor mental wellness programs based on the key predictors identified. For example, as 'anonymity' is a significant predictor, companies should be more aware of protecting privacy and creating a safe environment for employees to share their authentic feelings. It is worth mentioning that, even though family history is an important influencer, we should be careful not to force employees to provide such information.
- Proactive Support: Employ the model to proactively pinpoint employees who may be at risk, facilitating timely interventions. This approach is particularly beneficial for those who suffer in silence, enabling support even before they reach out for help.

As for the final test score and metrics, it tells us how accurately the model can identify employees who might need help. The higher the score, the better it is at prediction. We’ve ensured that our model reaches a balance between being sensitive to the signs of distress and being specific to the factors that truly matter.

To further improve our model, we might explore additional data sources or consider the impact of less obvious factors like workplace culture or personal life stresses. We could also continuously update the model with new survey data to keep it relevant, accurate, and reliable.

In conclusion, our model isn't just a technical tool; it's a means to a healthier, happier workplace. By identifying the signs early and offering targeted support, we're not just helping individual employees; we're cultivating a more productive and positive work environment. This proactive stance on mental well-being is not just good ethics; it's good business as human beings are always the most valuable resources for tech companies.

9. References

[1] Plusapn.com. (2023). 2023 Mental Health in Tech Report.
https://apn.com/wp-content/uploads/2023/06/APN-Tech-Execs-Mini-Report-Outline-3.pdf

[2] Harnois, C. E., & Bastos, J. L. (2018). Discrimination, Harassment, and Gendered Health Inequalities. *Journal of Health and Social Behavior, 59*(2), 283-299. American Sociological Association.