

# Uncovering Behavioral Risk Signals: A Machine Learning Study of Alcohol Misuse

Andres Salcedo ([asalcedo@berkeley.edu](mailto:asalcedo@berkeley.edu)) Cassidy Bruner ([cassidybruner@berkeley.edu](mailto:cassidybruner@berkeley.edu))

Eyup Agtepe ([eyupagtepe@berkeley.edu](mailto:eyupagtepe@berkeley.edu)) Gabrielle Doran ([gabrielle\\_doran@berkeley.edu](mailto:gabrielle_doran@berkeley.edu))

**GitHub Repository:** <https://github.com/cassidy-bruner/mids-207-final-project-team-4.git>

**Abstract:** As public health datasets grow larger and more detailed, understanding how to extract meaningful patterns from them has become increasingly important. In this project, we explore whether machine learning models can help identify factors associated with risky drinking behavior and evaluate how well different approaches capture these patterns. We compare several models to see how they perform and where their strengths and limitations appear. We examine whether Adverse Childhood Experience (ACE) variables contribute useful signals and how model performance varies across groups such as sex, age, and race. The goal of the project is to use predictive modeling to understand how ACE experiences and other features relate to risky drinking and how these patterns differ across population groups.

**Introduction:** Alcohol misuse is a major public health concern that contributes to significant physical, psychological, and social harm. In the United States, excessive alcohol consumption remains one of the leading causes of preventable disease and death. Prior research, such as the study by Felitti et al. (1998) in the *American Journal of Preventive Medicine*, revealed a strong graded relationship between Adverse Childhood Experiences (ACEs) and the likelihood of developing alcoholism and other risky behaviors in adulthood. These findings highlight the long-term behavioral consequences of early-life trauma. Our team is motivated to explore whether machine learning can uncover and quantify similar patterns using recent survey data. Specifically, we aim to analyze which demographic and health factors predict risky drinking, and to what extent ACE factors contribute beyond those variables. The input to our algorithm is self-reported survey data containing demographic variables, health indicators, and ACE scores, and we use supervised machine learning models to predict whether an individual exhibits risky drinking behavior as the output. This framing allows us to evaluate which factors are most predictive of alcohol misuse and assess the extent to which ACEs contribute beyond demographic and health characteristics.

**Related Work:** Recent work has increasingly applied machine learning to understand how substance-use outcomes relate to traumatic childhood experiences, generally falling into two main approaches: predictive modeling and population-level epidemiological analysis. In the predictive-modeling category, Jing et al. (2020) used Random Forests to classify individuals at risk for substance misuse, reporting accuracies between 74–86%. Their work demonstrates the strength of flexible, nonlinear models that capture complex interactions between behavioral and demographic features. However, these models are often limited by interpretability and can overfit when class imbalance is severe. In contrast, Leza et al. (2021) represent a systematic-review approach, synthesizing evidence across dozens of studies and finding consistent, graded associations between childhood adversity and later substance use. While this approach is comprehensive and high-level, it does not provide individual-level predictive tools or evaluate model performance. A third line of work, exemplified by Song et al. (2023), applies large-scale survey analytics to datasets such as BRFSS to quantify health outcomes associated with adverse childhood experiences. These studies benefit from nationally representative samples but typically rely on traditional statistical methods that may miss nonlinear risk patterns. Compared to these prior approaches, our project aims to bridge the gap between population-level insight and predictive modeling by applying supervised machine learning methods to BRFSS data to estimate the risk of adult risky drinking based on ACE exposure. By evaluating multiple model families and addressing class imbalance directly, we aim to improve predictive performance and assess subgroup fairness, an area underexplored in existing literature.

**Dataset:** We used data from the CDC's [2024 Behavioral Risk Factor Surveillance System \(BRFSS\)](#), a nationally representative survey of U.S. adults. The survey is administered via telephone interviews and collects information on a wide range of health-related risk behaviors, chronic health conditions, and preventive service use.

The BRFSS 2024 dataset contains 457,670 respondents across 301 variables. We filtered to 39,283 respondents from 8 states where the ACE module was administered (for more information on this see the data limitations in the appendix(1)). We engineered a broader risky drinking target combining binge drinking (4+ drinks for women or 5+ drinks for men on one occasion), heavy drinking (8+ drinks for women or 15+ drinks for men per week) and high-frequency drinking (4+ drink occasions per day). This expanded prevalence from 12% to 37% of respondents, creating a more balanced classification problem. We tested this definition on a random 20% sample of all respondents and confirmed similar prevalence (37.4%).

We replaced BRFSS missing response codes 7, 77, and 777 (Don't know) and 9, 99, 999 (Refused) with NaN values. For MENTHLTH and PHYSHLTH we re-coded 88 (None) to 0 days. We examined continuous variables for outliers using the IQR method with 3 times IQR as the threshold and retained all values as flagged cases were valid responses of 30 poor health days. We re-coded all categorical variables to numeric formats appropriate for modeling. For binary ACE variables we re-coded (Yes/No) 1/0, ordinal ACE variables (Never/Once/More than once) to 0/1/ and protective ACE factors (1-5 scale) to 4-0 where higher values indicate less protection. We re-coded sex to 0/1 and applied one-hot encoding to the imputed race variable, creating 5 binary indicators with White as reference. We created two ACE features: `at_least_1_ace` (binary indicator) and `ACE_SCORE` (cumulative count, range 0-17, mean 1.83). After removing missing values we retained 34,513 complete observations. We selected 25 features: 12 individual ACE variables, 6 demographic and health variables (education, age, sex, mental health days, physical health days, smoking status), 5 race indicators, `at_least_1_ace` and `ACE_SCORE`. We excluded income and marijuana use due to high missingness rates. We split the data using stratified sampling with shuffling and `random_state` set to 42 for reproducibility. The training set contains 24,158 observations (70%), the validation set contains 5,178 observations (15%) and the test set contains 5,177 observations (15%). We applied `StandardScaler` to normalize all features, fitting the scaler on training data only and transforming all three sets to prevent data leakage.

**EDA:** We performed bivariate analyses exploring associations between demographic and behavioral features and target: `risky_drinking`. The classification plot (figure 1) shows the proportion of individuals classified as engaging in risky drinking (green) versus non-risky drinking (blue) by sex. Among men (0), 44% are classified as risky drinkers. Among women (1), 32% are risky drinkers. This suggests, men are more likely to engage in risky drinking than women in this sample. The classification plot (figure 2) shows the proportion of individuals classified as engaging in risky drinking (green) versus non-risky drinking (blue) by the response to "Live With Anyone Depressed, Mentally Ill, Or Suicidal?". Among those who said no (0), 37% are classified as risky drinkers. For those who said yes (1), 43% are risky drinkers. In the Age vs Risky Drinking boxplot (figure 3), the median age of risky drinkers appears noticeably younger, while non-risky drinkers are generally older. The distributions show that younger adults are more likely to engage in risky drinking behaviors, while older adults tend to report lower rates. Both groups have a few respondents in the lower age range, but there are no significant outliers. In the Education Level vs Risky Drinking bar chart (figure 4), individuals with higher education levels (categories 3 and 4, attended or graduated college) report more instances of risky drinking compared to those with lower education. However, the overall number of non-risky drinkers remains slightly higher in every education group. This pattern may indicate that people with higher education levels are more likely to engage in social drinking. We also examined correlations among the Adverse Childhood Experiences (ACE) variables to identify potential multicollinearity and shared patterns of adversity. As shown in the correlation heatmap (figure 5), most ACE variables exhibit moderate positive correlations ( $r = 0.2\text{--}0.4$ ), suggesting that individuals reporting one type of adverse experience often report others. These variations in risky drinking across subgroups is important to note in evaluating our models in different subgroups. In examining distributions of the variables (figure 6) you can see, both *MENTHLTH* and *PHYSHLTH* variables are right-skewed, with most respondents reporting 0 unhealthy days and smaller groups reporting more. This suggests that while most individuals in the sample report good overall health, a small subset experiences chronic mental or physical health difficulties, which could be relevant factors in understanding risky drinking behavior.

To assess feature importance we measured Pearson correlation coefficients between each feature and the target variable to measure linear relationships and a Random Forest classifier was trained using 100 trees and max depth 10 to extract feature importances capturing non-linear relationships and interactions. Demographic and health variables were the strongest predictors of risky drinking. Education level showed the highest correlation ( $r=0.150$ ). Age had the highest Random Forest importance (0.197). The top five predictors by both methods were education, age, sex, physical health days and race. Among ACE variables, *ACESWEAR* (verbal abuse) had the strongest correlation ( $r=0.060$ ). The cumulative `ACE_SCORE` ranked 6th overall in Random Forest Importance (0.052). The binary `at_least_1_ace` indicator showed weak predictive value with the second to lowest rank for Random Forest importance. Individual ACE experiences showed correlation ranging from  $r=0.002$  to  $r=0.060$ . Figure 7 gives us a better summary of the top 15 features ranked by importance.

**Methods:** We employed a progressive modeling strategy to predict risky drinking behavior using the features described above, following the machine-learning pipeline methodologies used throughout 207 by starting with a baseline model and introducing three increasingly complex improvements.

1. **Baseline Model - Majority Class Classifier:** Our baseline model predicts the majority class for all observations (no risky drinking). This provides a simple benchmark establishing the minimum level of performance our more

sophisticated models improve upon. A majority-class classifier is appropriate for this problem because the dataset is highly imbalanced, with non-risky drinkers far outnumbering risky drinkers. By evaluating this baseline, we can quantify how much predictive value the later models gain beyond simply exploiting class imbalance. Mechanically, the model assigns the most common label to every input, producing high apparent accuracy but zero recall and zero precision for detecting risky drinkers, further highlighting the need for more expressive models.

2. **Logistic Regression:** Logistic regression was our first improvement over baseline. Logistic regression is appropriate because it provides interpretable coefficients that quantify the contribution of each feature including ACEs, demographics, and socioeconomic variables, to the likelihood of risky drinking. This interpretability is especially valuable in public-health contexts where understanding drivers of risk matters as much as prediction accuracy. The model works by applying a linear combination of input features and passing the result through a sigmoid function to produce a probability between 0 and 1. Training involves optimizing the weights via gradient descent to minimize binary cross-entropy loss. Because it is linear, logistic regression captures only additive relationships, but its transparency makes it a strong first model beyond baseline.
3. **XGBoost:** Our first more complex model is XGBoost (Extreme Gradient Boosting), a high-performance ensemble method that builds many shallow decision trees sequentially. Each new tree attempts to correct the errors made by the previous trees, allowing the model to capture non-linear interactions and subtle patterns in the data that simpler models may miss. XGBoost is well-suited for tabular data like ours because it can handle mixed feature types, missing values, and complex feature interactions with minimal preprocessing. Additionally, its built-in regularization (via L1 and L2 penalties) helps prevent overfitting—an important consideration given the imbalance between risky and non-risky drinkers. This model provides feature importance scores that help us identify which ACEs or demographic factors most strongly influence risky drinking predictions.
4. **Neural Network:** Our second advanced model is a feedforward neural network implemented using `keras.Sequential()`. Neural networks are appropriate for this task because they can capture high-dimensional nonlinear patterns that may emerge from combinations of ACEs, demographic factors, and behavioral indicators. The model architecture includes an input layer, multiple hidden layers with ReLU or tanh activation functions for nonlinear representation learning, and a final output layer with a sigmoid activation to generate class probabilities. During training, the network uses stochastic gradient descent with a tunable learning rate to minimize binary cross-entropy. By stacking multiple layers, the network can learn hierarchical representations of risk patterns which is something neither logistic regression nor tree-based models capture in exactly the same way. Although neural networks require more tuning and offer less interpretability, they can provide strong predictive performance when properly regularized.

## Experiments, Results and Discussion:

**Baseline Model:** For the baseline model, we trained a simple majority class classifier without any additional feature engineering, class-balancing techniques, or hyperparameter tuning beyond default settings. We evaluated performance using accuracy, precision, recall, and F1. Across all three splits, the baseline achieved nearly identical performance with approximately 62% accuracy and 0 precision/recall/F1 indicating that the model consistently predicted the majority class. This behavior is expected given the strong class imbalance in the outcome variable and the absence of strategies such as class weighting, resampling, or threshold tuning. Subgroup analyses revealed that accuracy varied somewhat across demographic and ACE-related subgroups (training accuracy ranged from 0.55 to 0.79 across education groups and 0.56 to 0.67 across sex), but precision and recall remained zero for all subgroups which makes sense given the nature of the model. Because the baseline model was too simplistic to capture the complexity of alcohol-misuse risk factors, its performance shows no evidence of overfitting, the near-identical results across training, validation, and test suggest actually a poor fit. These findings justify the need for more expressive models and hyperparameter-tuning strategies in subsequent experiments, as the baseline provides only a majority-class prediction and cannot meaningfully detect risky drinkers or support subgroup-fairness evaluation beyond accuracy reporting.

**Logistic Regression Model:** For the initial logistic regression, we used a Keras Sequential model with one dense layer with sigmoid activation. We used a kernel initializer of ones to make the model treat every feature equally, and a bias initializer of zero for the model to learn the bias so that our initial model doesn't assume anything about our data. The results seemed like the loss decreased from 2.60 to 2.30 during 15 epochs for the validation data and from 2.55 to 2.25 for the training data. However, both train accuracy and validation accuracy were reduced. This indicated that the initial logistic regression model did not learn anything and was just randomly guessing. The reason why training accuracy didn't drop much might be because we downsampled the majority class in the training data but not in the validation data. The second model uses zero bias initializers for the model to learn the bias by itself. For the kernel initializer, we used Glorot

Uniform to initialize the weights in a statistically balanced way. The results seemed like the logistic regression model still did not learn well from the training data and did not perform well on the validation data, since the loss remained basically the same (around 0.62 for train and around 0.6175 for validation). While accuracy stayed flaky around the 0.652 range for the validation data, it increased very little (0.001) for the training data. Our best learning rate was 0.01 for the logistic regression. Our bias term was  $-0.1135$ , which indicates that the model was seeing a 47% chance of risky drinking without learning any features. Our training loss was approximately 0.6199 and validation loss 0.6173. This indicated that there wasn't much difference between the losses and the model generalized fine. There wasn't a sign of overfitting. If we analyze the best logistic regression model's evaluation metrics on the output data, we can see that the model was good at identifying non-risky drinkers but did not perform well on identifying risky drinkers. Our logistic regression model's recall on risky drinking is 0.328, meaning the model misses a large portion of individuals who actually engage in risky drinking. Risky precision is moderate (0.578), meaning that when the model predicts risky, it's fairly reliable (a 57% chance of precision), but it predicts risky drinking very rarely. Model accuracy is 65%, which seems fine, but since the model is doing a pretty bad job at detecting risky drinkers, we cannot rely on this result. For the subgroup analysis, we saw that the model did better for the majority groups and younger people, but it didn't do well for women, older adults, or minority racial groups. The biggest differences showed up in recall, which means the model is much more likely to miss risky drinkers in some of these groups. This brings up some fairness concerns since the model isn't performing the same for everyone. From this analysis, we can clearly see that logistic regression may not be the best model with a lot of features and non-linear relationships.

**XGBoost Model:** For the XGBoost model we trained a gradient boosting classifier using downsampling to address the class imbalance in the training data. We reduced the majority class (non-risky drinkers) to match the minority class, bringing the training set from 24,158 to 18,220 samples with a balanced 50/50 split. We used GridSearchCV with 5-fold cross-validation to tune hyperparameters, testing 1,296 combinations while optimizing for F1 score. The parameters tuned included `n_estimators` (10, 15, 20, 50, 75, 100), `max_depth` (2-7), `learning_rate` (0.001-0.2), and `min_child_weight` (1-6). The best configuration was `n_estimators` of 75, `max_depth` of 4, `learning_rate` of 0.1, and `min_child_weight` of 6 with a cross-validation F1 of 0.645. We evaluated the model using accuracy, precision, recall, F1, and ROC-AUC and compared results against a majority class baseline. The baseline achieved 62.3% accuracy but 0% recall and 0 F1, predicting all respondents as non-risky drinkers. On the test set XGBoost achieved 62.5% accuracy, 50.2% precision, 71.1% recall, 0.59 F1, and 0.70 ROC-AUC. While accuracy was similar to baseline, the model identified 71% of risky drinkers compared to 0% for baseline. This shows the model learned meaningful patterns despite the weak feature correlations observed in EDA. We examined overfitting by comparing train and test performance. The F1 gap was reasonable (0.67 train vs 0.59 test) and loss curves showed convergence with a small difference between training and validation loss of 0.01. We then ran subgroup analysis which showed performance disparities across demographics. Recall was higher for males (82%) than females (55%), higher for ages 30-44 (87%) than ages 60+ (58%), and higher for White respondents (77%) compared to Black (54%) or Other (56%) groups. This pattern suggests the model performs better for majority demographic groups, which is an important fairness consideration given our research question.

**FNN Model:** We developed a feedforward neural network with sigmoid activation for binary classification to predict Risky Drinking status, using cross entropy to measure loss. For the neural network optimization, we employed a hyperparameter tuning approach, testing various architectures with hidden layer configurations ranging from shallow (64, 32) to deeper networks (128, 64, 32, 16), along with different optimizers (SGD, Adam), learning rates (0.0001-0.01), and activation functions (ReLU). We validated model performance using a train-validation-test split, monitoring validation loss and recall across epochs to assess generalization. To address the class imbalance, we implemented multiple overfitting mitigation strategies including early stopping with a patience of 10 epochs, class weighting with a 1.2x penalty for false negatives, and consistent train-validation loss monitoring. We evaluated the model using recall, precision, and F1 score, prioritizing recall to maximize identification of risky drinkers for public health interventions. On the test set, the neural network achieved 71.7% recall, 49.6% precision, 0.586 F1, and 0.685 ROC-AUC. Subgroup analysis revealed similar fairness concerns as XGBoost but with more pronounced disparities; the neural network's recall was substantially higher for males (85%) than females (63%), higher for ages 18-29 (90%) than ages 60+ (66%), and higher for White respondents (81%) compared to Black (59%) or Other (63%) groups. While the model showed roughly equivalent performance in test recall to XGBoost, it came at the cost of lower precision, and with fairness concerns.

**Analysis:** Across the four modeling approaches we observed substantial differences in performance, overfitting behavior, and suitability for identifying individuals at risk of alcohol misuse across subgroups. The majority-class baseline performed as expected, achieving moderate accuracy (around 62%) but failing entirely to detect risky drinkers due to the severe class imbalance, underscoring the need for more expressive models. Logistic regression improved slightly beyond

baseline in accuracy (up to 65%) but remained limited by its linear form, producing low recall (around 33%) for risky drinkers and demonstrating clear difficulty capturing nonlinear relationships among ACEs, demographics, and socioeconomic factors. In contrast, both XGBoost and the neural network substantially improved recall by modeling higher-order interactions, but they differed in fairness and stability. XGBoost achieved strong recall (around 71%), with reasonable precision (around 50%), decent accuracy (63%) and showed only modest evidence of overfitting due to its built-in regularization and a well-tuned hyperparameter search. The neural network achieved similar recall (71%) but did so with reduced precision (down to 49%) and accuracy (down to 61%), a larger train-test gap, and more pronounced subgroup disparities. Taken together, these results indicate that while multiple complex models outperform simpler baselines, XGBoost provides the most reliable trade-off between detecting risky drinkers, maintaining generalizability, and offering interpretable insights for use in a more sensitive field of public-health.

**Conclusion:** This project set out to evaluate whether machine learning methods can identify individuals at risk for alcohol misuse, with the broader goal of supporting early prevention and targeted public-health interventions. By analyzing demographic factors, health indicators, and ACE exposures, we examined both how well different models predict risky drinking and which features contribute most to those predictions. While interpretable models such as logistic regression offered useful directional insights, they were unable to capture the complex nonlinear relationships in the data and consequently missed many true cases of risky drinking. More flexible approaches like XGBoost and the feedforward neural network performed substantially better, with XGBoost providing the strongest balance of recall, precision, and generalizability; its especially high recall is critical for ensuring that most individuals who actually engage in risky drinking are detected early enough for intervention. However, both complex models displayed performance disparities across sex, age, and racial groups, highlighting fairness concerns and the need for additional behavioral or longitudinal features, alternative sampling strategies, and fairness-aware training to reduce subgroup gaps. Overall, while machine learning can meaningfully uncover and quantify risk factors for alcohol misuse, its use in real prevention efforts requires rigorous validation, attention to equity, and careful consideration of how predictive tools impact different population groups.

#### **Contributions:**

**Andres:** I conducted data preprocessing including missing value handling, feature recoding (binary, ordinal, reverse coding for ACE variables), categorical encoding (one-hot encoding for race), and feature engineering (ACE\_SCORE). I performed EDA focusing on feature distributions, class balance, and feature importance analysis using correlations and Random Forest methods. Based on my EDA findings, I helped reframe the research question and contributed expertise on our modeling approach. I developed the XGBoost gradient boosting model, implementing downsampling for class imbalance and GridSearchCV hyperparameter tuning (1,296 combinations, F1-optimized). I evaluated model performance across train/validation/test sets and conducted subgroup fairness analysis by sex, age, and race. My downsampling and subgroup analysis code was adopted by teammates for their models. I maintained the BRFSS\_data\_preprocessing, andres\_eda, and andres\_model notebooks.

**Cassidy:** I contributed to organizing group meetings, notes, and documents, developed the baseline model (documented in python\_notebook/baseline\_model) including evaluation, subgroup analysis, and associate visuals. I led the discussion/analysis of the various methods used and model results in the report, comparing and contrasting different approaches and findings. I also implemented a logistic regression and neural network (included in the baseline\_model notebook but not in the final report). I assisted in writing the data section including challenges and limitations and handled a portion of EDA, performing bivariate analyses to explore associations between features and the outcome.

**Eyup:** I created the logistic regression model, maintained the logistic\_reg.ipynb eyup\_eda.ipynb files, and wrote the abstract, introduction, motivation, logistic regression analysis for the reports and contributed to the literature review section. I also contributed to the related work, results, and linear regression sections of the presentation. Additionally, I communicated with the instructor about the team's progress and gathered feedback and suggestions.

**Gabby:** I performed a literature review on existing works, which grounded our analysis within the current research on ACEs and public health, particularly outcomes related to mental health and substance use disorders. I ideated and suggested the project topic and dataset. I participated in EDA so that our team may understand the demographics, survey questions, and class distributions of the 2024 CDC BRFSS dataset. As part of the feature selection process, I performed subgroup analyses and feature engineering to create our 'risky\_drinking' target class, which combines multiple features measuring respondents' drinking activity. I developed a feedforward neural network with weights to care for the target class imbalance and to train the model to penalize false negatives, and I optimized this model for recall to reflect our use case. I created a slide deck template for our presentation and completed the Context, Motivations, and Related Works slides.

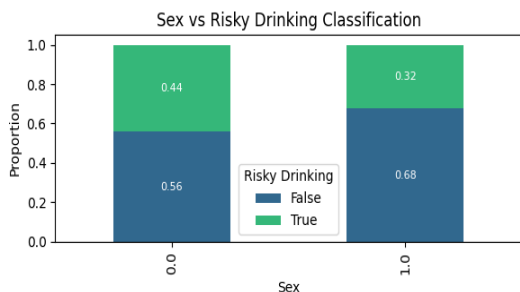
## References:

Jing, R., et al. (2020). Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6980708/>

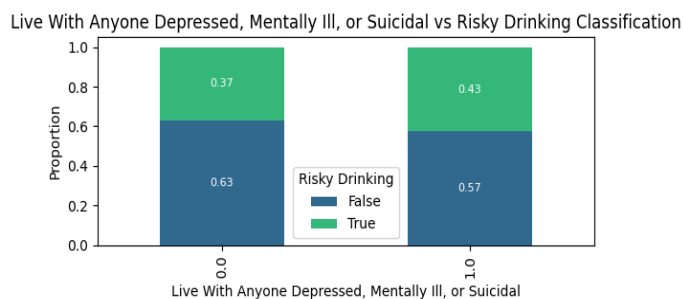
Leza, L., et al. (2021). Adverse childhood experiences (ACEs) and substance use disorder (SUD): A scoping review. <https://www.sciencedirect.com/science/article/abs/pii/S0376871621000582>

Song, E., et al. (2023). The association between adverse childhood experiences and mental health wellbeing during adulthood. <https://jphe.amegroups.org/article/view/9233/html>

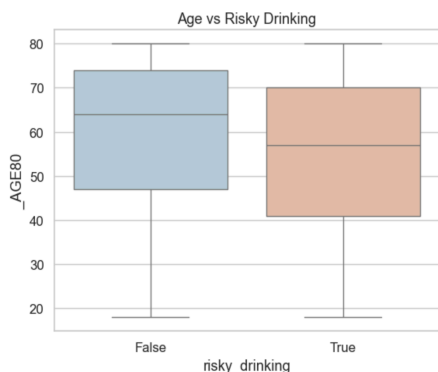
## Appendix: EDA Figures



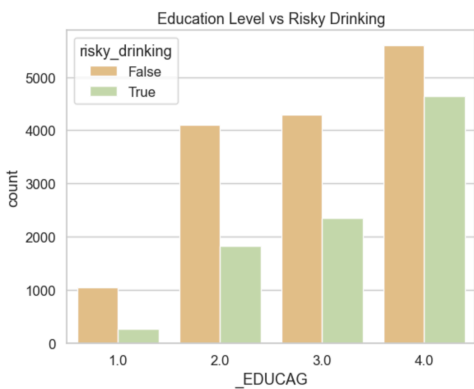
1. Figure 1:



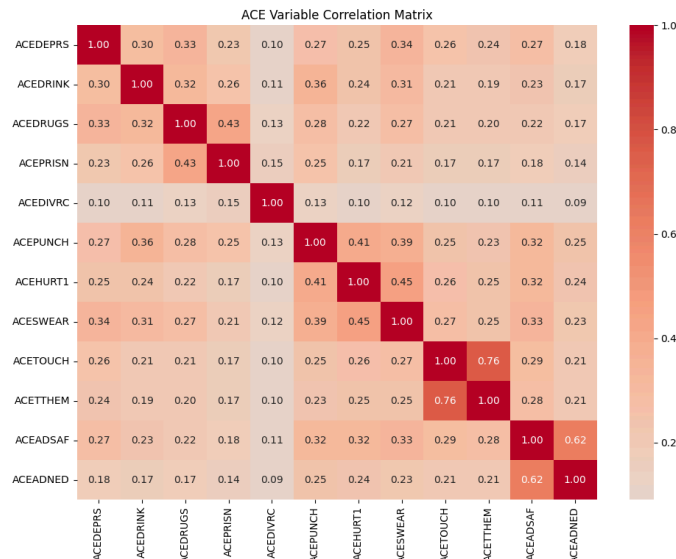
2. Figure 2:



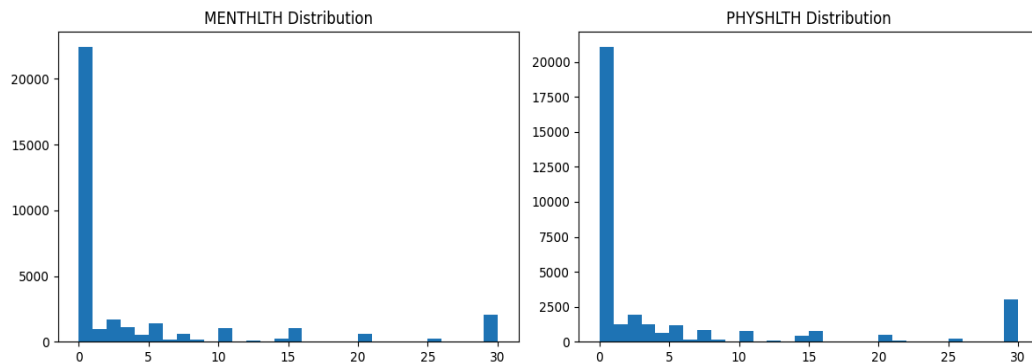
3. Figure 3:



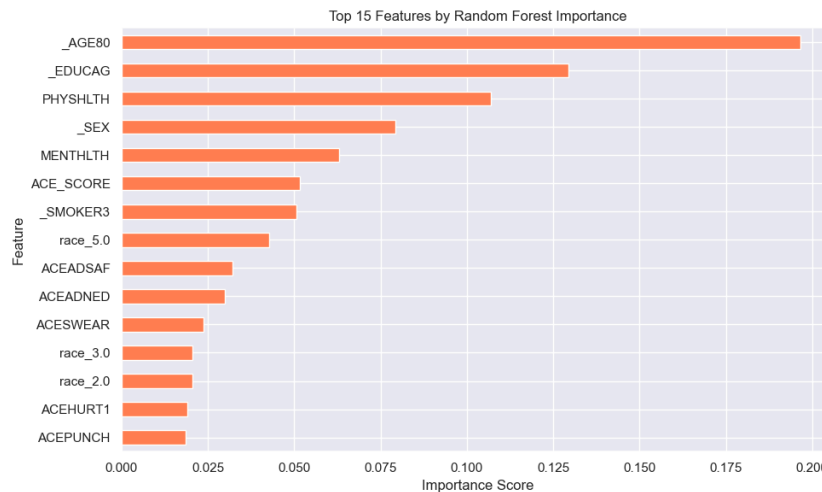
4. Figure 4:



5. Figure 5:



6. Figure 6:



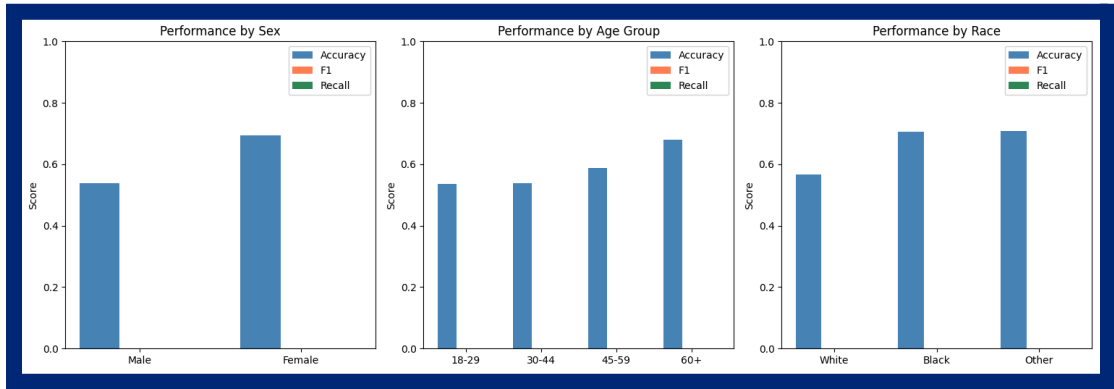
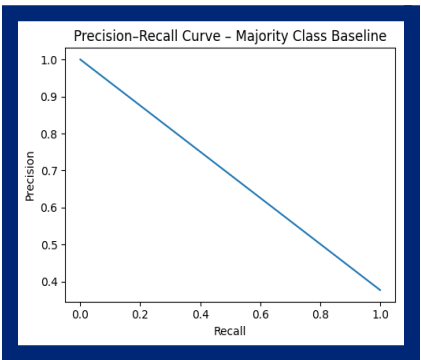
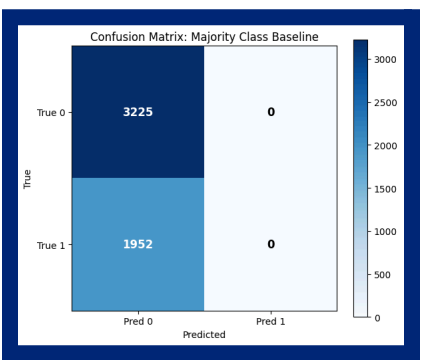
7. Figure 7:

**Data Challenges/Limitations:** Limited coverage of key variables: Only 8 states (Florida, Georgia, Hawaii, Virginia, North Dakota, Nevada, Puerto Rico, the Virgin Islands) included ACE data. These variables are important to our motivation, so we restricted the dataset to these states, reducing sample size, geographic diversity and the model's generalizability. Outcome variable construction: The target variable, risky\_drinking, was derived by combining three BRFSS questions on binge, frequent, and heavy drinking. Respondents were labeled as risky drinkers if they met our specified thresholds in any of the questions. This approach captures multiple risk types but may introduce bias from the chosen thresholds. Categorical and complex variable structure: Most BRFSS features are categorical and numerically coded, requiring cross-referencing documentation sources to interpret accurately. This process was time-consuming, and preparing the data for modeling demanded extensive encoding and cleaning to address missing or ambiguous responses.

## Model Results and Figures

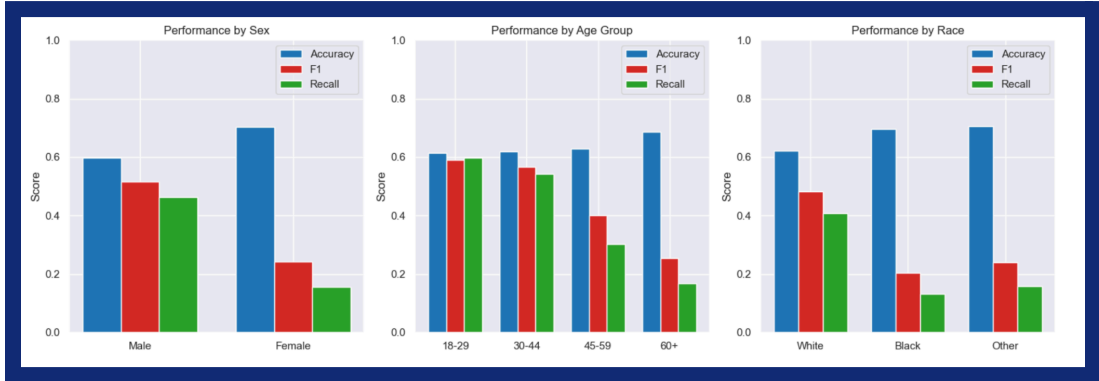
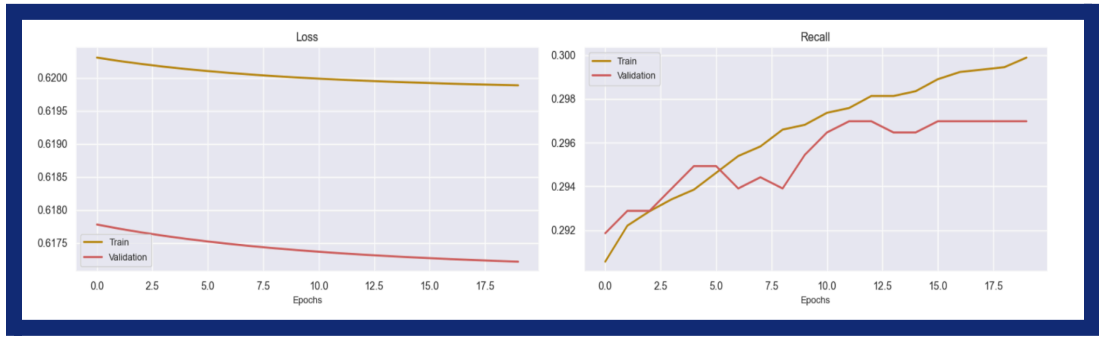
Baseline:

Data	Accuracy
Train	62.29%
Val	62.28%
Test	62.29%



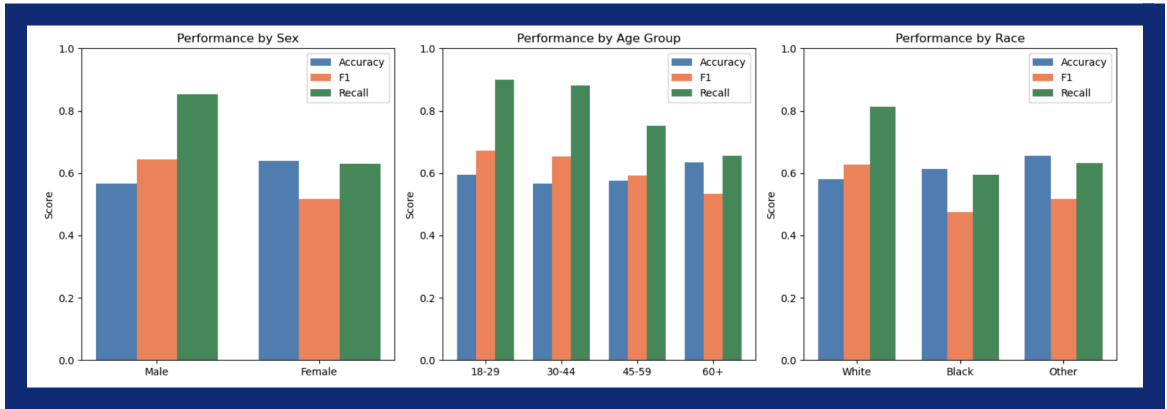
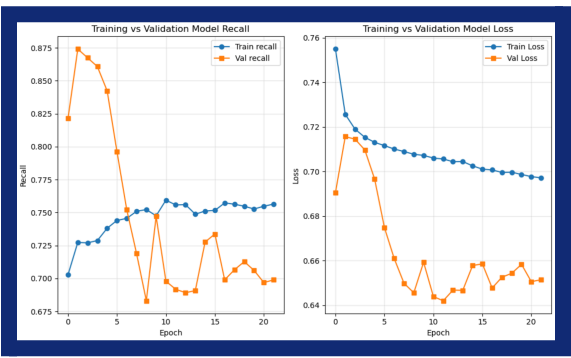
Logistic Regression:

Data	Accuracy	Recall
Train	64.60%	30.00%
Val	65.20%	29.85%
Test	65.64%	32.83%



FNN:

Data	Recall	Loss	AUC
Train	0.7564	0.6972	0.6819
Val	0.6989	0.6514	0.6714
Test	0.7172	0.6399	0.6845



XGBoost:

Data	Accuracy	Recall
Train	64.46%	70.77%
Val	61.39%	66.56%
Test	62.55%	71.06%

