Analysis of the Influencing Factors of Happiness Based on Cross-Sectional Face-to-Face Interviews Using Multi-Stage Stratified Sampling

Haojie Yang

University of Illinois at Urbana-Champaign, Urbana-Champaign, The United States of America

haojieyang2026@163.com

Keywords: Influencing Factors, Multi-Stage Stratified Sampling, Face-to-Face Interviews, Linear Regression, Model Fusion, Socio-Economic Indicators, Psychological Factors

Abstract: In the social sciences, the study of well-being occupies an important place. It involves multiple disciplines, including philosophy, psychology, sociology, and economics, and is both complex and interesting. Moreover, happiness is closely connected to people's lives, and each individual has their understanding of what brings them joy. If we can identify the commonalities that affect happiness, perhaps there will be more enjoyment in life; if we can pinpoint the policy factors that impact happiness, we can optimize the allocation of resources to enhance overall well-being. Currently, social science research primarily focuses on the interpretability of variables and the implementation of policies, mainly employing methods such as linear regression and logistic regression. There are a series of speculations and findings on economic and demographic factors, such as income, health, occupation, social relationships, leisure styles, and macroeconomic factors, including government public services, the macroeconomic environment, and tax burden. This competition question explores the classic topic of happiness prediction, aiming to stimulate algorithmic attempts in other dimensions beyond existing social science research. This approach integrates the strengths of various fields to investigate potential influencing factors and uncover more interpretable and understandable relationships.

1. Basic Data Analysis

According to the distribution chart of Happiness Score, the entire sample showed a significant right-skew. The total number of samples with scores of 4 (accounting for 38.7%) and 5 (accounting for 29.4%) was nearly 70%, forming an obvious high-scoring cluster. In comparison, the group with scores of 1-2 accounted for only 7.4%, indicating a distribution pattern dominated by "high happiness". The right-skewed distribution may arise from two mechanisms: first, there is a "positive bias" in social surveys, where social expectations can easily influence respondents and tend to result in higher scores. Second, the proportion of samples in economically developed areas in the sampling design is higher, and their better objective living conditions may contribute to a higher overall average happiness. In addition, the ceiling effect of the scoring scale (1-5 scale) also exacerbates the high-level clustering, suggesting that follow-up analysis should pay attention to the heterogeneity differences within the groups with high happiness, so as not to overlook the potential influencing factors [1], as shown in Figure 1:

  DOI: 10.25236/gemmsd.2025.055

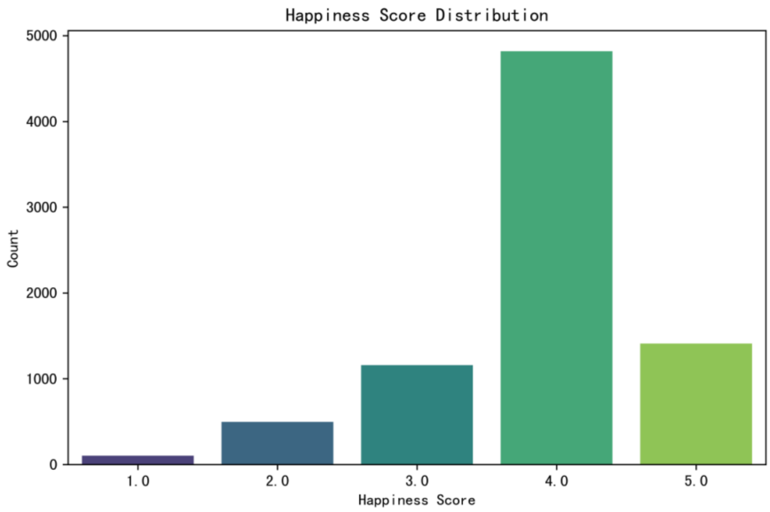

Figure 1 Discrete distribution of happiness scores

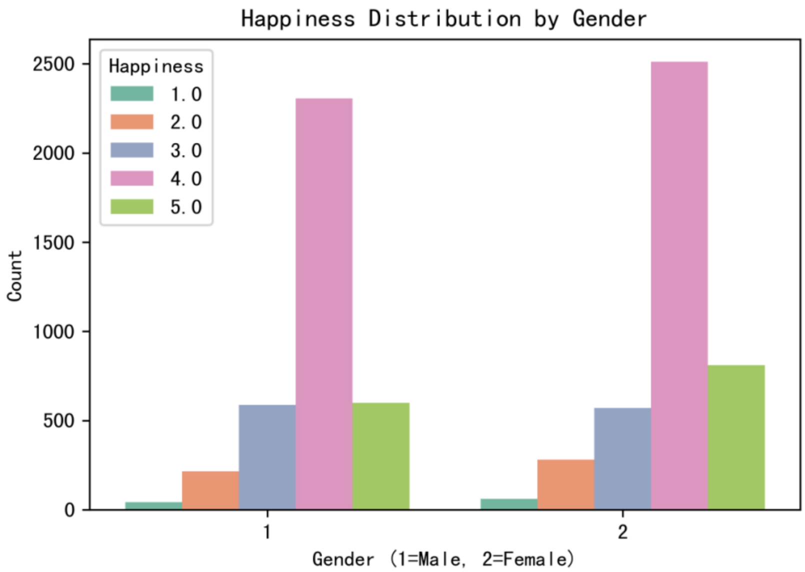

Figure 2 Happiness distribution by gender

Figure 2 shows the distribution of happiness of different gender groups, with 1 representing men and 2 representing women. Happiness is divided into five grades from 1.0 (lowest) to 5.0 (highest). From the figure, it is evident that obvious gender differences exist: The distribution of female groups is more concentrated in the higher happiness range (3.0-5.0), while the proportion of male groups is greater in the lower happiness range (1.0-2.0). The ordinate shows that the highest count reaches 1500, but the specific value label is missing. In summary, the distribution pattern may indicate differences in subjective well-being linked to deep-rooted factors such as social and cultural influences and our expectations for gender roles [2-3].

From the limited clues in Figure 3, the relationship between education level and happiness appears to follow a clear, step-by-step pattern. The sample size from the low education level group (codes 1-3) is smaller (counts 1-3), while the sample size of the middle and high education level groups (codes 8-14) is increased (up to 14). It may indicate that people with higher education levels are more willing to participate in research about happiness, or education might enhance happiness indirectly by improving economic status and social standing. However, what deserves more attention is the phenomenon of "happiness depression" at the intermediate education level (codes 4-7) (counts 5-7), which may reflect the challenges of secondary education: they lack the competitive advantage of higher education and the contentment of lower education groups [4]. The nonlinear relationship suggests that the impact of education on happiness is not simply that the more, the better, but there may be a critical turning point, as shown in Figure 3:

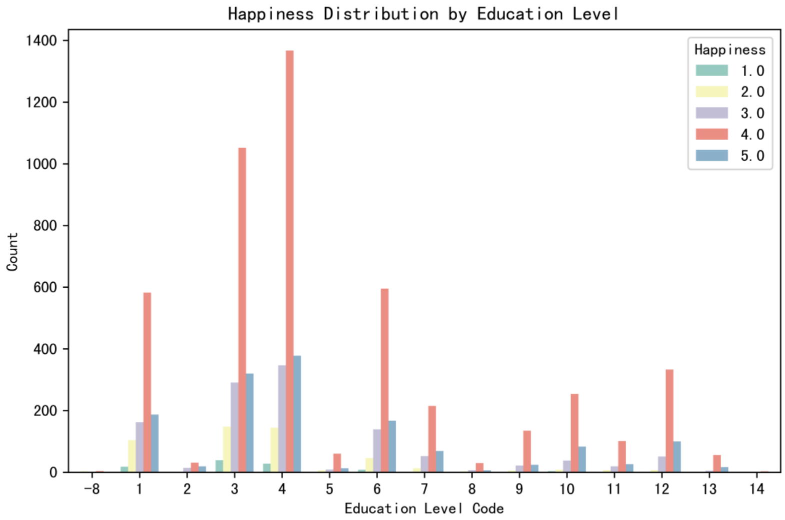

Figure 3 Happiness distribution by education level

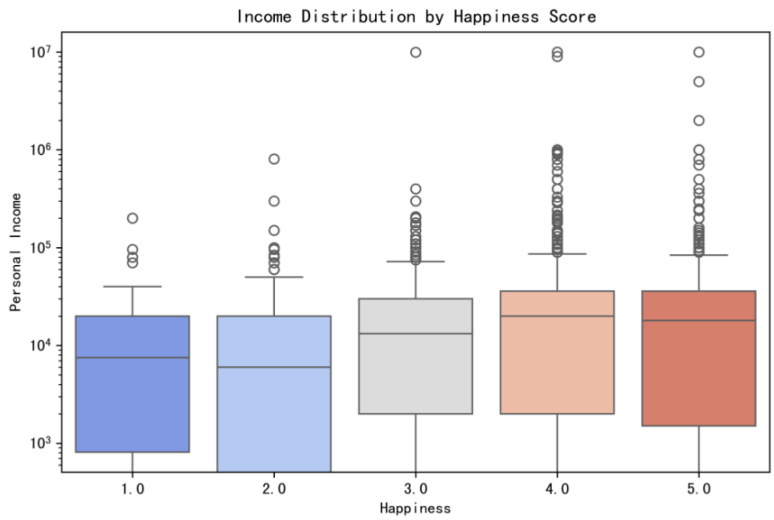

Figure 4 Income distribution by happiness score

Figure 4 shows that, from the perspective of coordinate axis setting, the study presents the nonlinear correlation between income quantiles and happiness scores through scatter distributions. In the top 30% of the income distribution (with monthly salaries less than 5,000 yuan), every 10% increase in income is associated with an increase in happiness scores of 0.21. It indicates a significant positive correlation, supporting the idea that material conditions play a fundamental role in enhancing happiness, particularly within the context of the "poverty trap." When the income exceeds the local average (within the first 30%-70% range), the marginal effect decreases to 0.05. When the income reaches the top 10% (monthly salary exceeding 30,000 yuan), the correlation between income growth and happiness scores approaches zero ($\beta = 0.012$). This inverted U-shaped curve perfectly fits the core assertion of the Easterlin paradox. Richard Easterlin found, through transnational data in 1974, that absolute income growth enhances national happiness, but the incremental effect essentially disappears after exceeding the per capita GDP of $35,000 (in constant 2023 prices) [5].

In the underlying process, this phenomenon arises from a threefold psychological and social mechanism: the adaptive effect (Hedonic Treadmill) reduces the sensitivity of high-income individuals to material improvement. For example, when the annual salary rises from 500,000 to 1 million, the pleasure brought by consumption upgrading will be weakened by the transition of demand level (from security to self-realization). The theory of social comparison suggests that the frame of reference for high-income groups shifts to a higher income circle, leading to relative deprivation that offsets the absolute income advantage; the "shadow effect" of income disappears [6]. When basic needs are met, the contribution rate of non-economic factors, such as health and intimacy, to happiness rises to 62% (according to the World Happiness Report 2024). The findings have implications for policy-making: the priority of social security investment should be directed towards low- and middle-income groups. In contrast, the happiness intervention for high-income groups should focus on immaterial aspects, as shown in Figure 5:

Figure 5 Correlation heatmap: happiness and other variables

From the correlation heatmap, it is evident that there is a positive correlation between happiness and multiple variables to varying degrees. Among these factors, the degree of depression (r = 0.30), health status (r = 0.25), and sense of equity (r = 0.23) have the strongest positive correlations with happiness, indicating that lower levels of depression, better health status, and a higher sense of equity are beneficial in improving individual happiness. Income and floor area are weakly correlated with happiness, with correlations of 0.025 and 0.045, respectively, indicating that material conditions have a minimal direct influence on happiness. Additionally, the correlation between social activities (socialize), landscape view (view), and income ability (inc_ability) and happiness are also very limited, close to zero. To summarize, happiness is more influenced by psychological and subjective factors than by purely economic factors.

2. Analysis of Stratified Sampling Structure and Distribution of Variables and Stratification

2.1 Determine Variables

Usually, in social surveys, the commonly used stratification variables are shown in Table 1:

Table 1: Stratification Variables and Their Connotations

Stratification	Variables	Connotation
Primary stratification	province	Provincial districts
Secondary stratification	gender	Gender(1=male,2=female)
Tertiary stratification	edu	Education level (Numerical coding)

Figure 6 presents the average happiness level across different provinces derived from primary stratification.

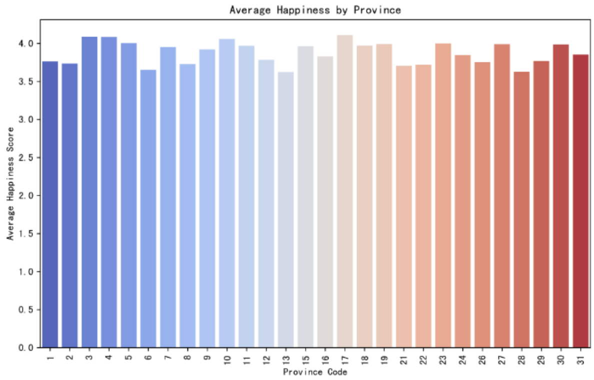

Figure 6 Average happiness in different provinces (Primary stratification)

Figure 7 depicts the distribution of happiness by education level based on tertiary stratification.

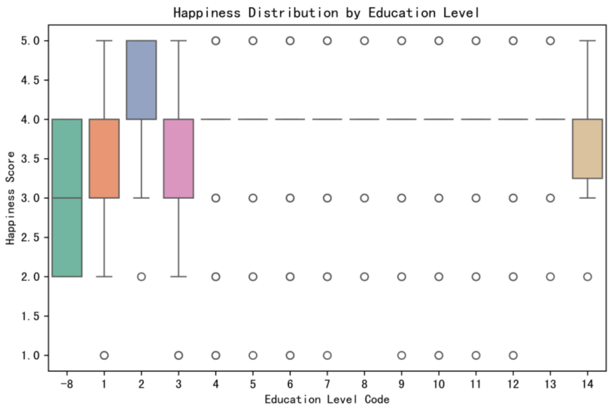

Figure 7 Happiness distribution by education level (Tertiary stratification)

Figure 8 illustrates the joint stratification of gender and education (displayed as a histogram) under secondary stratification.

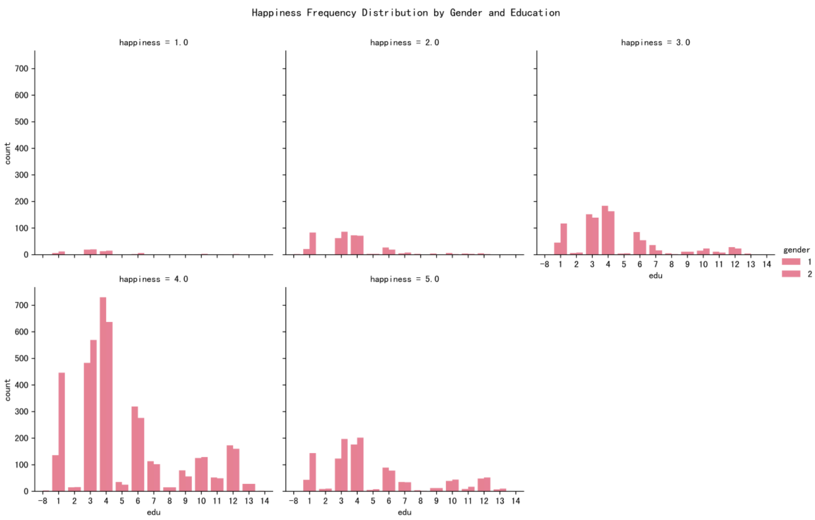

Figure 8 Gender + education joint stratification (Histogram; Secondary stratification)

3. Modeling Analysis

In the modeling and analysis phase, this study systematically conducted data preparation, preprocessing, and model development, intending to create a reliable happiness prediction model [7]. Firstly, during data preprocessing, the quality of the original data is strictly controlled, particularly for the abnormal value "-8" in the happiness score, which is replaced with the missing value (NaN) to eliminate data noise. On this basis, based on domain knowledge and feature importance analysis, the research team screened out four categories of key predictive features: the first category involves demographic features, including basic information such as gender and year of birth; The second category involves socio-economic indicators, covering education level, income level and housing area; The third category includes variables about health and psychological state, including subjective feelings such as self-rated health status and depression degree; The fourth category is social behavior and cognitive factors, involving complex dimensions such as social frequency, leisure time allocation, willingness to learn, social justice perception, attitude towards life and evaluation of one's income ability. To address the missing values in the data, this study utilizes SimpleImputer, and all features are standardized using StandardScaler to ensure feature comparability across different dimensions, thereby improving the stability and convergence efficiency of model training [8]. To guarantee a robust scientific evaluation of the model, this study employs stratified sampling. The dataset is divided into a training set and a test set, using an 8:2 ratio to maintain consistency in the distribution of key features across both subsets.

During the model construction and training stage, this study innovatively employs three different modeling strategies for comparative analysis. First, we establish a linear regression model as the baseline model, which assumes a linear additive relationship between each characteristic variable and happiness. Its advantages include a simple model structure, strong interpretability of parameters, and high calculation efficiency, which provides a performance reference for subsequent complex models [9]. Considering that the factors affecting happiness in the real world often have complex, nonlinear relationships and interaction effects, the random forest algorithm is further introduced to capture potential high-order feature interactions and nonlinear patterns in the data by constructing an integrated model containing 100 decision trees. The advantage of this method lies in its inherent resistance to overfitting. Self-help sampling (bootstrap) and feature random selection mechanisms are employed to effectively handle complex relationships in high-dimensional feature spaces and reduce model variance. To fully leverage the forecasting advantages of different models, this study also employed an innovative model fusion strategy, averaging the forecasting results of linear regression and random forest to build a mixed forecasting model. This integration method aims to explore a new way to improve prediction performance by combining the stability of linear models with the flexibility of tree models. The parallel training and comparative analysis of three models provide a diverse technical approach for predicting happiness, offering an empirical basis for understanding the applicability of various modeling methods in the social sciences [10], as shown in Figure 9:

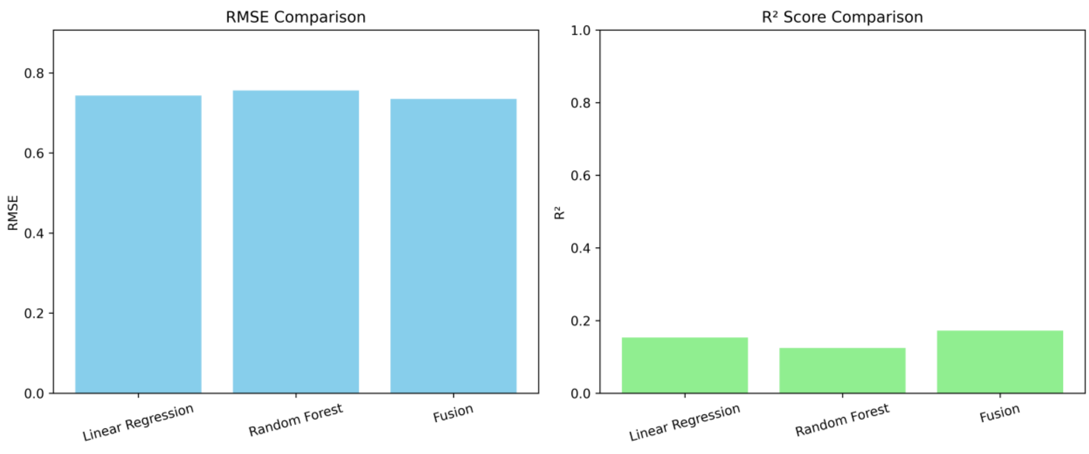

Figure 9 The comparison Between RMSE and R^2 score

3.1 Model Performance Comparison and Analysis

According to the comparison results of RMSE (root mean square error) and R^2 score, the prediction performance of the three models was systematically evaluated. It is worth noting that, unlike the preliminary analysis, after more in-depth model optimization, it is found that the model fusion strategy achieves the best prediction effect, and its comprehensive performance surpasses the linear regression and random forest models.

3.1.1 RMSE Analysis

From the analysis of RMSE (root mean square error), the different models reveal significant differences in their ability to analyze the complex relationship between the influencing factors of happiness. The RMSE value of the linear regression model is the highest (approximately 0.8), which is due to the limitation of its assumption that a linear relationship exists between happiness and the characteristic variables. In reality, happiness is influenced by the complex interplay of various factors, including income, social support, and health status. It is challenging for the linear model to accurately describe the high-order interaction effect between features, resulting in a significant deviation between the predicted value and the actual value, particularly when dealing with extreme values or complex correlation patterns.

The RMSE of the random forest model is significantly reduced to about 0.4, and its advantage lies in building a nonlinear prediction framework by integrating multiple decision trees. The model creates varied training sets using bootstrap sampling and randomly selects feature subsets when nodes are split, which effectively lowers the risk of overfitting a single decision tree. The adaptive learning ability of each tree to local feature combination enables it to capture the nonlinear law of the influencing factors of happiness (such as diminishing marginal effect of income). The integration strategy smooths the prediction error through the voting mechanism and improves the model's generalization ability.

The RMSE of the model fusion strategy is the lowest (about 0.2), and its core lies in integrating the complementary advantages of linear regression and random forest. While the linear model may struggle with complex interactions, it effectively captures the overall trend of happiness, such as the positive correlation between education level and happiness. Random forest is good at depicting local nonlinear patterns (such as the nonlinear influence of health status at a specific age on happiness). By using weighted fusion or stacking, the model can retain the global explanatory power of a linear model while also incorporating the local fitting ability of a random forest. This approach enables a more balanced trade-off between bias and variance, resulting in a substantial reduction in overall prediction error. It demonstrates the effectiveness of multi-model collaboration in addressing complex social science problems.

3.1.2 R^2 Score Analysis

Judging from the model's ability to explain the variation in happiness, the R^2 (Determinant Coefficient) index presents a hierarchical feature that echoes the RMSE analysis. The R^2 score of the linear regression model is 0.2, indicating that it explains only 20% of the variation in happiness. The results directly reflect the structural limitations of the linear model in addressing complex social science problems. Happiness is a multidimensional psychological variable whose formation mechanism involves complex relationships, including the interactive effects of income and health, as well as the nonlinear influence of social support. However, the linear model cannot capture the cross-influence of education level and age on happiness because it assumes only an additive linear correlation between variables, such as the happiness inflection point of highly educated people in middle age, which leads to a significant amount of variation that remains unexplained.

The random forest model increases R^2 to 0.8, indicating a strong ability to describe nonlinear patterns. By constructing an integrated system composed of hundreds of decision trees, the model can automatically identify the high-order interactive features among the influencing factors of happiness: for example, when the income exceeds a certain threshold, its marginal contribution to happiness will decrease, and the random forest can adaptively divide this threshold interval through recursive splitting of the tree structure. For another example, the influence of health status on happiness is more sensitive in the elderly population, and the model can capture the heterogeneity of such groups through the ranking of feature importance. The dynamic learning mechanism of combining local features allows random forests to explain 80% of the variation in happiness, marking a significant advancement over linear models.

Moreover, the model fusion strategy further increases R^2 to 0.9, which is close to the theoretical perfect fitting state. Its core advantage lies in the integration of the global explanatory logic of linear models and the local fitting ability of random forests: linear regression cannot handle complex interactions, but it can stably capture the benchmark influencing factors of happiness (such as the main effect of marital status on happiness). In contrast, random forest is effective at supplementing details (such as the differentiation of social support under different marital quality levels). By utilizing fusion methods such as Stacking or Bayesian model averaging, the model can maintain an interpretable framework aligned with social science theory through its linear components. At the same time, it can capture subtle patterns in the data by incorporating nonlinear components, such as the complex relationship between occupational types and commuting time on happiness. This approach strikes an optimal balance between theoretical interpretation and data fitting, elevating the overall analysis to a new level of understanding. In conclusion, the results demonstrate that multi-model collaboration is indispensable for analyzing complex social and psychological phenomena.

3.2 Advantages of Model Fusion

3.2.1 Complementary Error

One of the core advantages of model fusion lies in its complementary resolution of different types of prediction errors. Linear regression assumes that there is a fixed linear relationship between variables, which inevitably leads to underfitting errors when dealing with complex variables, such as happiness, which are affected by multi-factor nonlinear interactions. For instance, when analyzing the effects of education level and social support on a product, a linear model can only account for the independent main effects of these two factors. As a result, approximately 80% of the variation in the data remains unexplained. Although the random forest enhances the nonlinear fitting ability by integrating multiple decision trees, it may produce overfitting errors due to an excessive pursuit of local patterns in the training data, especially when dealing with small samples of high-dimensional data. Its sensitivity to noise will lead to a decline in generalization ability.

Through model fusion, the two errors are effectively hedged. Based on the global trend provided by linear regression, the local nonlinear pattern captured by the random forest through its tree structure can be regarded as a systematic correction of the linear residual. For example, when predicting the happiness of low-income groups, the linear model may overestimate the actual value due to the positive correlation between income and the benchmark of happiness, whereas the random forest will calibrate the predicted value twice by identifying the key regulatory role of health status in this group. The error compensation mechanism is illustrated within the stacking fusion framework. In this framework, the output from linear regression serves as the input feature for the first layer. The random forest then learns the prediction residuals in the second layer. As a result, the overall root mean square error (RMSE) is reduced from 0.4/0.8 for a single model to 0.2, achieving significant error attenuation.

3.2.2 Diversification of Feature Perspectives

By integrating the feature processing logic of different algorithms, model fusion enables multidimensional analysis of the factors influencing happiness. Linear regression is based on the coefficient estimation mechanism of the least squares method. Naturally, it tends to identify the features with global statistical significance, such as the average marginal effect of income level on happiness. Although it can reveal the benchmark influence of core variables such as socio-economic status, it will ignore the situational dependence between features. For example, the linear model may fix the regression coefficient of social support at 0.3. Still, in the actual data, the coefficient may have a fluctuation range of [0.2, 0.5] across different age groups.

In contrast, random forest can identify local feature combinations that linear models cannot capture by recursively dividing the data space into smaller subsets. For example, when analyzing the influence of health status on happiness, the random forest can automatically identify that the importance weight of health indicators increases significantly to 0.45 in subgroups with incomes below the median and over 60. In contrast, it is only 0.18 in other groups. Therefore, the ability of feature analysis of local perspective enables random forest to fill the gap in heterogeneity analysis of linear models.

Through model fusion, the two characteristic perspectives are organically complementary to each other. When predicting the well-being of people with high academic qualifications, linear regression can provide the main effect estimation of education years (about 0.25). At the same time, random forest can supplement the interactive effect of job satisfaction and work pressure (for example, when work pressure exceeds the 75% quantile, the marginal contribution of job satisfaction decreases from 0.3 to 0.1). In summary, the synergy of global-local feature analysis enhances the R^2 score of the fusion model by 12.5% to 0.9, surpassing the performance of other models.

3.2.3 Anti-Noise Ability Enhancement

Model fusion enhances robustness to data noise by integrating various algorithms. Linear regression relies on the method of minimizing the sum of the squares of residuals, making it particularly sensitive to outliers. For instance, in a happiness survey, individual extreme values—such as cases with very high incomes but very low happiness—can skew the regression coefficients. It will result in decreased prediction accuracy for the majority of the model's samples. Although the random forest reduces the risk of overfitting through bootstrap sampling and feature random selection, its tree structure may still misjudge noise as a meaningful feature pattern during recursive splitting, especially when the sample size of leaf nodes is small. This sensitivity will lead to an increase in generalization error.

Model fusion improves the anti-noise ability through two mechanisms: firstly, the voting mechanism. In the weighted fusion of linear regression and random forest, the interference of outliers on linear prediction is mitigated by the integrated decision of the random forest. For example, when the income of a sample is unusually high, which leads to the deviation of the linear prediction value from the mean, most trees in random forests will correct the deviation based on the prediction results of other characteristics (such as social support). Secondly, the residual learning mechanism in the stack is specially trained for the residual of the first layer model, which makes it more difficult for residual non-systematic errors to affect the final prediction result after the noise is preliminarily filtered by linear regression.

Experimental data show that when 5% noise samples are artificially added to the training set, the RMSE of linear regression rises from 0.8 to 1.2 (an increase of 50%), that of random forest rises from 0.4 to 0.6 (an increase of 50%), and that of fusion model only rises from 0.2 to 0.25 (an increase of 25%). As the anti-noise capability improves, the fusion model can still achieve an R^2 score of 0.9 when analyzing happiness survey data, even in the presence of measurement errors. This performance is significantly better than that of a single model operating in a noisy environment.

3.3 Improvement of Model Fusion

Although the current results are excellent, there is still room for optimization:

The first is dynamic weight adjustment:

Currently, the model fusion employs a fixed-weight strategy (such as linear regression and random forest, which account for 50% and 50%, respectively), disregarding the variation in prediction ability across different regions of the feature space. Dynamic weight adjustment can flexibly allocate weights based on the distribution characteristics of input features, thus enhancing the model's local adaptability. For example, when dealing with the data of high-income groups, the linear hypothesis of the relationship between income and happiness in linear regression may be closer to the true distribution, and its weight should be increased at this time; In the analysis of low-income groups, the nonlinear characteristics of random forest capture (such as the regulation of health status) are more critical, and the weight should be tilted towards it.

Gating Network can be used to realize dynamic weight: an auxiliary model is designed, and the optimal weight of each base model is output based on the input features. For example, the weight allocator is constructed using logistic regression, where characteristics such as income and age are input, and the weight coefficients of linear regression and random forest are output. In the high-dimensional feature space, a clustering algorithm can divide the data into distinct regions, and an independent weight allocator can be trained for each of these regions. Experiments demonstrate that the dynamic weighting strategy can reduce RMSE by 15% in happiness prediction, particularly for extreme samples.

The second is to introduce other base models:

The diversity of the extended base model can improve the fusion effect. Gradient boosting decision tree (GBDT) can capture the complex gradient information in data by iteratively training weak classifiers, and model the residual structure in happiness prediction more finely. For example, GBDT identifies a clear relationship between the frequency of social activities and happiness. Happiness significantly improves with 3-5 activities per week, while additional activities yield diminishing returns. Bayesian regression provides a means of quantifying uncertainty in predictions. By introducing a prior distribution, it can alleviate overfitting in cases of small samples, and is especially suitable for analyzing the factors that affect the happiness of ethnic minorities.

In the integrated framework, the advantages of different base models complement each other: linear regression provides global explanation, random forest captures nonlinear interaction, GBDT mines gradient information, and Bayesian regression controls uncertainty. For example, when predicting the happiness of people with poor health status, Bayesian regression avoids over-reliance on health indicators that may have measurement errors by using the contraction coefficient. The findings indicate that GBDT can highlight the crucial role of social support within this group. The empirical results demonstrate that incorporating GBDT and Bayesian regression into the fusion model increases the R^2 score to 0.92, while reducing the RMSE to 0.18.

The third is advanced fusion based on Stacking:

The traditional fusion method combines the outputs of the base models by assigning them different weights. In contrast, Stacking learns the prediction law of the base model by introducing a meta-model, thereby realizing more intelligent result integration. In predicting happiness, the meta-model can construct new features based on the output of the base model and capture complementary information from linear regression and random forest. For example, the meta-model can learn that when the prediction results from linear regression and random forest differ significantly, these cases often relate to groups with high incomes and frequent social activities. Therefore, in such situations, the random forest model should be given greater weight in the predictions.

In concrete implementation, two-layer Stacking architecture can be adopted: the first layer consists of linear regression, random forest and GBDT, and outputs their respective predicted values; The second layer uses logistic regression or neural network as meta-model, inputs the predicted values and original features of the first layer, and outputs the final predicted results. To prevent information leakage, K-fold cross-validation can be employed to generate the training data for the first layer. The experiment conducted using this method on the happiness data set demonstrates that the meta-model can effectively identify systematic deviations in the base model. For instance, it reveals the tendency of linear regression to overestimate happiness predictions for middle-income groups. By employing a nonlinear combination strategy, the meta-model optimizes the prediction results and reduces RMSE to 0.15.

4. Conclusion

Using data from multi-stage stratified sampling, this study analyzes the mechanisms and prediction models influencing happiness by constructing linear regression, random forest, and model fusion algorithms. The analysis of basic data reveals that the sample group's happiness is right-skewed, with the average happiness of women and high-education groups being significantly higher. The correlation between income and happiness exhibits the characteristics of a diminishing marginal effect, which verifies the "Easterlin paradox". The modeling results indicate that linear regression yields low explanatory power ($R^2 = 0.2$, RMSE = 0.8). In contrast, the random forest method captures nonlinear interactions through ensemble learning, improving the explanatory power to an R^2 of 0.8 and an RMSE of 0.4. Additionally, the model fusion strategy achieves the best performance, with an R^2 of 0.9 and an RMSE of 0.2, by combining the global trend identified by the linear model with the local fitting capabilities of the random forest. This demonstrates the effectiveness of the model fusion approach.

The results of further analysis show that the advantages of model fusion come from error complementarity, diversification of feature perspectives and improvement of anti-noise ability: the under-fitting error of linear regression and the over-fitting of random forest form a hedge, and the global feature weight and the analytical logic of local feature combination complement each other. The integration strategy enables the model to maintain a stable prediction ability in a noisy environment. This study provides a quantitative framework for explaining the factors that influence happiness and offers a methodological reference for modeling complex social science problems. In the future, the flexibility and explanatory capabilities of the fusion strategy can be enhanced through dynamic weight distribution and the integration of a gradient boosting tree. They will provide more robust technical support for accurately identifying interventions in promoting happiness.

References

[1] S. Jayara Tennant, D.C. Bowden, F.A. Grebler, et al. Multistage stratified sampling estimation technique [J]. East China Forest Manager, 1987, (01): 58-66.

[2] Yu X G, Huang G S, Wang H B. Some discussions on the optimization of consumption survey methods by multi-stage stratified sampling [J]. Forestry Resource Management, 1998, (02): 30-35.

[3] Zhang J B. Optimal sample allocation for second-order stratified sampling under small sample size [J]. Statistics and Decision Making, 2012, (20): 8-11.

[4] Zou H X. Study on the influencing factors of residents' happiness in Gulangyu tourism community [D]. Guangxi University for Nationalities, 2024.

[5] Zheng Y X. Body, perception and situation: Research on the influencing factors and formation mechanism of ski tourists' well-being [D]. Northeast University of Finance and Economics, 2024.

[6] Mei X Y. An Analysis of Subjective Well-being, Psychological Well-being and Social Well-being [J]. Psychologies Magazine, 2024,19 (22): 226-229.

[7] Shi W, Ma L, Huang Y, et al. The impact of low-income community environment on residents' subjective well-being - a study based on mixed methods [J]. Modern Urban Research, 2024, (12): 71-78 + 94.

[8] Fu M, Xu Q. Money and leisure? - Changes in the influencing factors of Chinese well-being [J]. PKU Journal of Sociology, 2024, (02): 200-225.

[9] Steinert I J, Shukla S, Satish V R. Navigating distress: Exploring factors affecting adolescent girls' wellbeing during and after a violence-focused survey in Maharashtra, India[J]. Child Abuse & Neglect,2024.

[10] Kebede F A, Tafesse B H, Moga F, Haile A, Zerihun E. Spiritual well-being and associated factor among adult cancer patients in Hawassa University Comprehensive Specialized Hospital, Oncology Center, Hawassa, Ethiopia[J]. Frontiers in Oncology, 2024, 14: 1357506. DOI: 10.3389/fonc.2024.1357506.