AOS C111, Introduction to Machine Learning
Dr. Lozinski, R. Matsuura
Jennifer Gipson
December 5, 2025

Machine Learning Modeling: Teacher Joy

Introduction

The intersection of academics and mental health for students is a lead concern for parents, school professionals, administrators, policymakers, healthcare professionals, and more. But an area that is frequently overlooked is the mental health and status of those that play a key role in student support systems: that is teachers. Teachers spend more face-to-face time with students in a typical K-12 school day than any other adult providing not only education, but physical and emotional management as well as support. It is crucial to consider the mental health of these professionals who, according to Agyapong (et al) in a 2022 paper, state that "Worldwide, stress and burnout continue to be a problem among teachers, leading to anxiety and depression. Burnout may adversely affect teachers' health and is a risk factor for poor physical and mental well-being," [4]. This report aims to address the several factors of teaching environments and personal teacher backgrounds to assess how well machine learning models can predict teacher joy.

Data

The data utilized for this report is provided by the Organisation for Economic Co-operation and Development (OECD) from the 2024 Teaching and Learning International Survey [1,2]. The survey collection (Table 1) covers teacher experiences from 50 countries and 7 territories. For the machine learning models employed, the target variable is 'teacher joy' with identification T4JOYTCH, measured on scale of extent. Full documentation of questions is provided in the metadata and codebook [3]. The input features were selected to capture a range of defining statistics. These variables are of high intention with hypothesized correlation based on decades of first-hand observations of teaching as a profession through hearing my parent's experiences.

The variables utilized to predict joy are as follows:

1. Years of teaching experience (T4TYEXPE)
2. Class size (T4TCSIZE)
3. Classroom management practices (T4CLASM)
4. Social utility motivations (T4MOSU)
5. Workload (T4WLOADT)
6. Student Behavior (T4STBEH)

Variable	Examples of Questions Surveyed
T4JOYTCH	*"I often feel happy while I teach.", "I generally teach with enthusiasm."*
T4CLASM	*"I tell students to follow classroom rules"; "I calm students who are disruptive"*
T4MOSU	*"Teaching allows me to influence the next generation"; "Teaching allows me to work against social disadvantage"*
T4WLOADT	Sources of stress in their work *"Having too much lesson preparation"; "Having too much marking"*
T4STBEH	Sources of stress in their work *"Being held responsible for students' achievement"; "Being intimidated or verbally abused by students".*

Table 1: *(Excluding the numeric variable inputs like T4TYEXPE and T4TCSIZE) Example questions asked in TALIS to contextualize metrics.*

Each variable, both input feature and target, is quantified by an associated scale based on how much they agree with surveyed questions, and values are standardized. The following table gives general statistics for the data (Table 2).

	T4TYEXPE	T4TCSIZE	T4CLASM	T4MOSU	T4WLOADT	T4STBEH	T4JOYTCH
count	16136.000000	16136.000000	16136.000000	16136.000000	16136.000000	16136.000000	16136.000000
mean	1.104425	2.392786	12.173607	13.287840	9.678535	9.216361	13.857346
std	0.430985	0.807868	1.757651	2.007848	2.068032	2.023821	2.074341
min	1.000000	1.000000	5.918700	5.974490	5.877050	5.722100	6.378620
25%	1.000000	2.000000	10.982310	11.295690	8.290420	8.077980	12.290320
50%	1.000000	2.000000	12.466030	13.538560	9.564900	9.092050	13.862500
75%	1.000000	3.000000	14.173850	15.199670	11.159610	10.666830	16.313870
max	4.000000	4.000000	14.173850	15.199670	13.893270	13.342500	16.313870

Table 2: *Baseline statistical values for all variables. Includes count of data points, mean, standard deviation (std), minimum value (min), lower quartile (25%), median (50%), upper quartile (75%), maximum value (max).*

Omitted or invalid answers or unadministered questions for variables T4TYEXPE and T4TCSIZE were stored as 9 and 8, respectively. Omitted or invalid answers or unadministered questions for all other variables were stored as 9999 and 9998, respectively. These values, for the purpose of this report, were parsed, stored as NaN, and removed from the data set. Likewise, any data point with a NaN value for any variable was entirely removed. After this preprocessing, 16,136 data points remained. The following figure gives the frequency distribution of three select predictor variables (Figure 1).

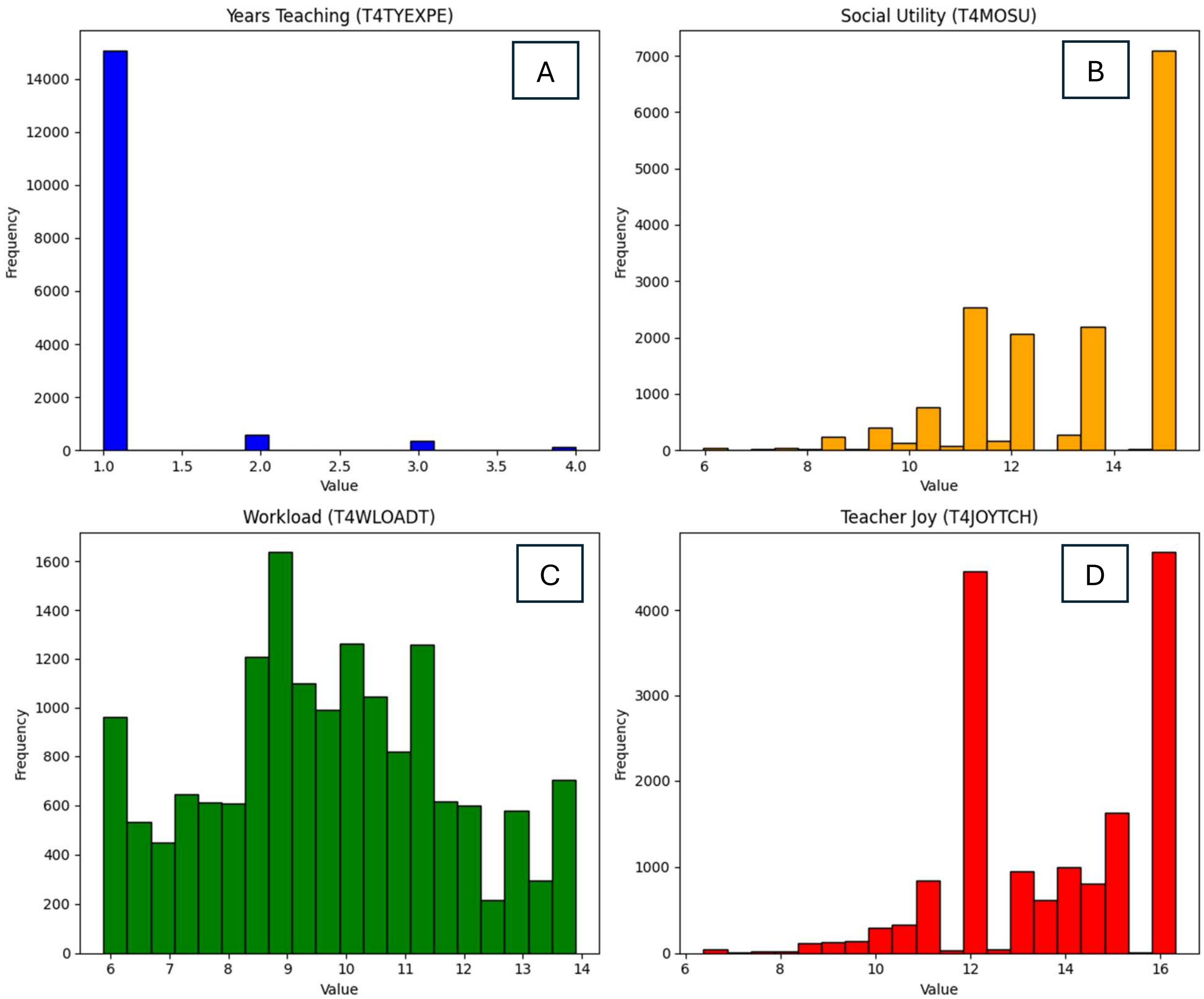

***Figure 1:** Histogram of standardized value frequency for four selected variables; **(A)** years teaching, **(B)** Social Utility, **(C)** Workload, **(D)** Teacher joy.*

Modeling

For the data set, several machine learning modeling types were implemented to assess which could best predict teacher joy. The initial motivation is rooted in unknown understanding of if the data would be best modeled linearly or would require more complex regression. For thorough coverage, the following models are applied: Linear Regression, Ridge Regression, Support Vector Regression (SVR), Multi-layer Perceptron Regressor (MLPRegressor), and Random Forest Regressor. It is important to note the application of StandardScaler as part of scikit-learn preprocessing to Ridge Regression, SVR, and MLPRegressor to standardize each feature that potentially see large range in associated value. All models are supervised regression types utilizing identical features. To truly assess how well each individual model performed, the root mean squared error (RMSE) and Regression Error Characteristic (REC) curve is applied. For all models, the data was split into training and test sets where 80% of the data was utilized for training and 20% for testing.

Linear regression was first applied as a foundational modeling approach. This model was a best ‘first guess’ to assess whether the data could be represented linearly and minimize error. As a type of linear regression, ridge regression functions similarly with the addition of the L2 regularization to ensure correlated features do not lead to overfitting of the data. SVR is implemented as a regression type to model nonlinear relationships in the data applied with the default scikit-learn SVR parameters. It functions with a tolerance value of 0.1 such that errors (absolute value of the difference between actual teacher joy and predicted teacher joy) above this threshold are penalized for further parameter adjustment when optimizing. MLPRegressor is a neural network model that, like SVR, covers nonlinear relationships between the input features and teacher joy. The MLPRegressor function, provided as part of scikit-learn, functions by applying layers with respective neurons to receive input features, and with weighted numerical combinations, outputs patterns for the data allowing for nonlinear representation. For the purpose of this dataset, a baseline two hidden layers each with 30 neurons was run to capture potential nonlinearity. Finally, Random Forest Regressor was utilized to similarly capture nonlinear relations, but with decision trees that do not demand standardization. Trees, individually, split and sort a given random sampling of the data into optimal groups. Furthermore, predictions across these trees are averaged for a singular finalized prediction representing the entire dataset. The model used uses 350 decision trees to balance overfitting and stability.

Results

Model	RMSE
Linear Regression	1.918737
Ridge Regression	1.918736
Support Vector Regression	1.974472
MLP Regressor	1.921309
Random Forest Regressor	1.902117

Table 3: *List of models and corresponding RMSE values (above).*

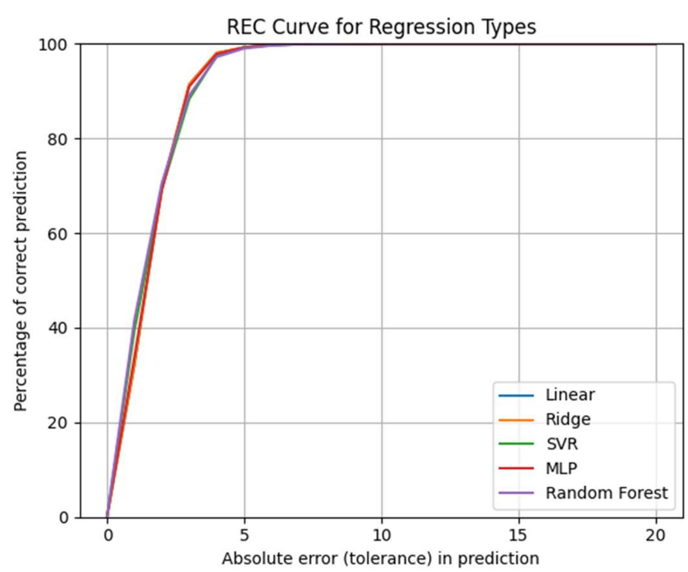

Figure 2: *REC curves for different regression types (right).*

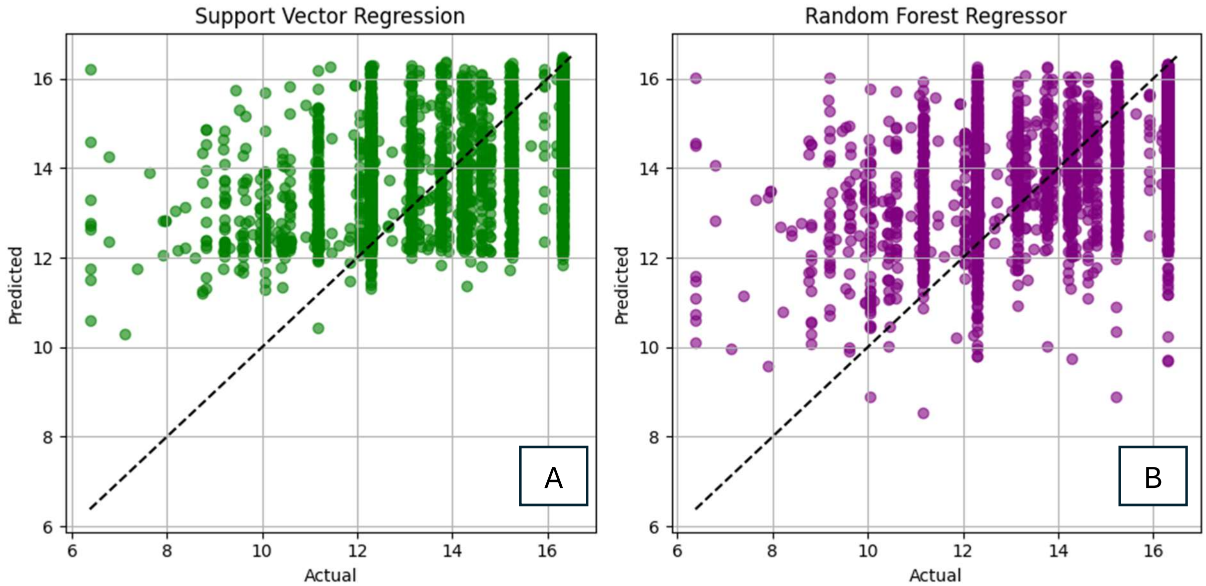

Figure 3: *Scatter plots of actual and predicted teacher joy for models* ***(A)*** *Support Vector Regression, and* ***(B)*** *Random Forest Regressor.*

The results variations in the RMSE (Table 3) are highest for SVR while lowest for Random Forest Regressor. The differences between the RMSE values across all models is extremely nominal with little significant variation. The REC curve (Figure 2) values across all models shows that with increasing absolute error in predictions, percentage of correct predictions also increases. The curves are nearly identical and almost reach 100 percent accuracy before a tolerance value of 5. The scatter plots (Figure 3) for both SVR and Random Forest Regressor show widespread values for the actual data and predicted outputs. Similarly, both plots indicate greater variance in the actual data as opposed to lower variance in the predicted values. The line of one-to-one relation running diagonally through each plot sees little to no consistency with the plotted points.

Discussion

The results of the modeling and analysis of different modeling types indicate strong resistance of the input features to the given machine learning models applied. The RMSE results show that there is little variation in the errors produced by the different models, but the data can potentially be better predicted when utilizing Random Forest Regressor. The resulting REC curve further supports the resulting RMSE values as the ability of the models to predict shows widespread agreement. Additionally, the REC curves' increase shows that absolute errors are relatively small, and predictions occur over a relatively narrow range. Lastly, the scatter plots for SVR and Random Forest Regressor indicate the predicted values deviate from the actual values of teacher joy. Given the strong similarities between these two plots, it can be further emphasized that there is little variation between the different model types.

Conclusion

The inability of the model types, linear regression, ridge regression, SVR, MLPRegressor, and Random Forest Regressor, to accurately predict teacher joy based on years of teaching experience, class size, classroom management practices, social utility motivations, workload, and student behavior demonstrates the lack of predictive relations. Error metrics across all models yield similar results as do REC curves and select scatter plots of SVR and Random Forest Regressor. The potential strongest model of the group favors Random Forest regressor, although the improvements seen when implementing are nominal compared to that of the linear regression types. The consistency across the models shows a key finding regarding the data set that none of the select input features can accurately point to the joy of teachers, even with correct modeling strategies.

Future Directions

The results of performing such machine learning modeling types to the survey data allow for strong reflection of the data applied. Predicting and quantifying behavioral health metrics has proven to be difficult for correlated and causal relations [5]. The challenge of predicting psychological constraints supports a wide field of research, and the application of machine learning models gives extensive potential for better quantification. Broadening the scope of this report would be best performed by observing all provided variables to compare RMSE and determine if linear or nonlinear regression models could capture the teacher experience data.

References

[1] TALIS: https://www.oecd.org/en/about/programmes/talis.html

[2] TALIS Data: https://www.oecd.org/en/about/programmes/talis.html#data

[3] Codebook and metadata https://app.powerbi.com/view?r=eyJrIjoiYTA0YzRhYmUtY2M0Ny00MGQxLWJhNTEtZTViNTU4NWRjYzM0IiwidCI6ImFjNDFjN2Q0LTFmNjEtNDYwZC1iMGY0LWZjOTI1YTJiNDcxYyIsImMiOjh9

[4] Agyapong B, Obuobi-Donkor G, Burback L, Wei Y. Stress, Burnout, Anxiety and Depression among Teachers: A Scoping Review. Int J Environ Res Public Health. 2022 Aug 27;19(17):10706. doi: 10.3390/ijerph191710706. PMID: 36078422; PMCID: PMC9518388

[5] Maslach, C., & Leiter, M. P. (2016). *Understanding the burnout experience: Recent research and its implications for psychiatry*. World Psychiatry, 15(2), 103–111. https://doi.org/10.1002/wps.20311