Predicting Mood Using Self-Recorded Data Over Time

Megan Eddy

Bellevue University

Business Problem

This project will explore relationships between self-reported mood-tracking data and sleep-tracking data to predict mood.

Background

According to dictionary.com, mood is defined as “a prevailing emotion tone or general attitude”. Some lines of research “[assume] that mood conveys important information about oneself and our social context which influences our cognitions and actions” (Fernandez-Berrocal & Extremera, 2008). It seems, therefore, that being able to predict mood would aid people in their own understandings of themselves and how they view the world. For those with mood disorders, sleep patterns are considered a “diagnostic criterion” (Peterson et al., 2006). This suggests a correlation between sleep and mood already known to exist in those with mood disorders and possibly existing for those without as well. Predicting mood could help inform when to make life decisions and when to wait. With this information, people could also make daily informed decisions that may assist with a better overall mood.

Dataset

The datasets included in this project are sourced from an anonymous user of two health-related apps. Both are downloads directly from the user’s application histories. They were collected over a series of years. The user will not be referenced to protect their identity.

The first application is a mood-tracking app. Users enter an overall mood and record various activities they participated in during the recorded period. They have the capability of tracking these moods on their own cadence. This data was collected once a day with recordings entered before bed. The data includes the date of collection, day of the week, time of collection, mood and activities.

The second application tracks users’ sleep patterns via their cell phone. The application is turned on before bed and records a variety of data. This includes the start and end times of physically going to bed, the start and end times of sleep, sleep quality (as calculated by the application), and air pressure.

Methodology

The individual datasets were cleaned separately before being merged on the ‘mood date’. The ‘mood date’ is the date the mood was recorded. Sleep data fields were aggregated in instances where there were multiple sleep entries for a given night. This resulted in one entry from the date at bedtime to date at wake up. The date at wake up was considered the ‘mood date’.

The datasets were merged using a left-join on the sleep data. This was because there was more mood information than sleep, so this method reduced the number of missing values.

Encoding was used on categorical features. One-hot encoding was utilized on the activity features of the mood dataset and ordinal encoding was used on the mood data. Several features were also converted to Boolean fields.

Distributions of features were explored and transformations applied where applicable.

Correlation coefficients were calculated .

Decision Tree, Random Forest and Gradient Boosting models were all built and R-squared and MSE evaluated.

Analysis

Analysis of the target variable distribution revealed a large class imbalance between the five different moods.

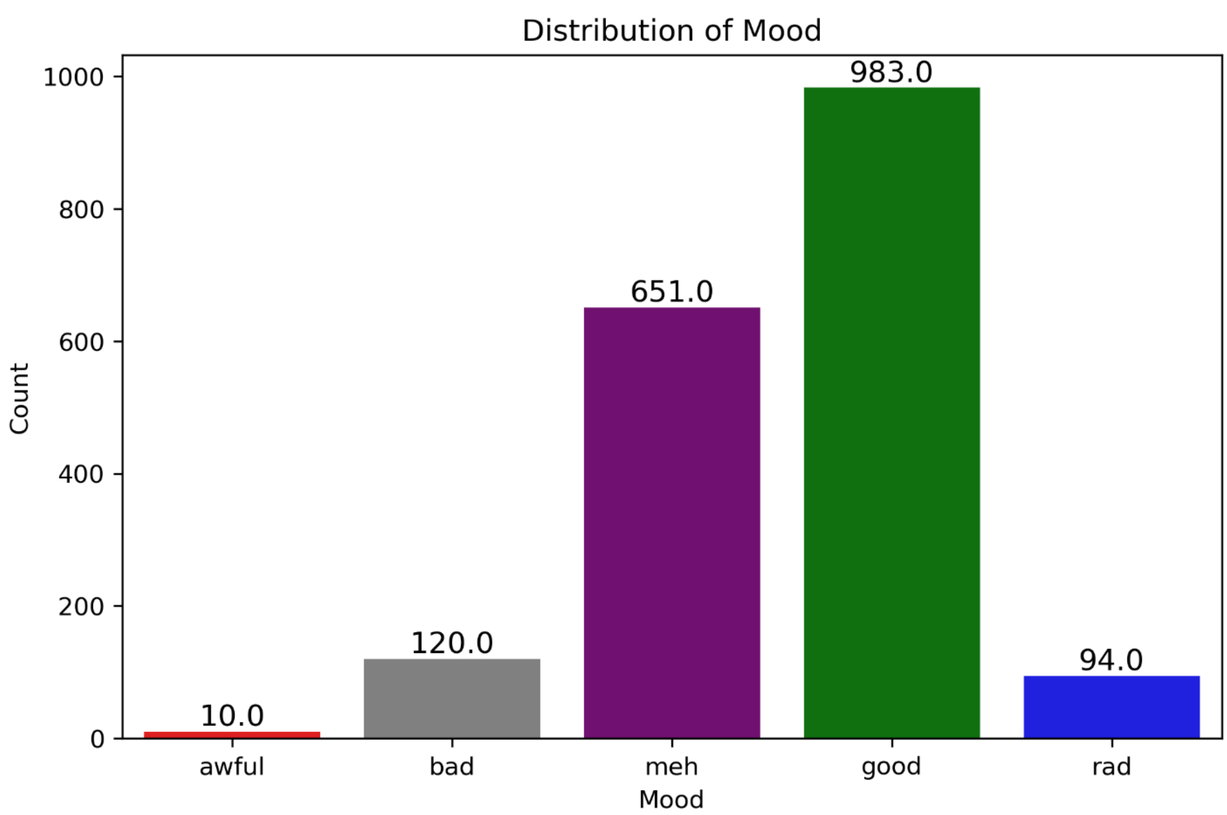

Figure 1: Mood Distribution

Exploration of the correlation between mood and other features did not reveal any obvious trends. Visualizations were used as well as a correlation matrix to evaluate correlation coefficients.

mood_numeric	1
activity relax	0.273827
Alarm	0.25145
weekday_encoded	0.215869
activity work	0.204496
activity Gym	0.20405
activity Racing Thoughts	0.196692
activity reading	0.183852
activity bad sleep	0.172165
activity SMART	0.16833
activity negative self talk	0.150371
activity Megan	0.13655
activity good meal	0.120349
activity travel	0.114315
activity Hike	0.113957
activity shopping	0.110002
activity friends	0.102366
activity date	0.100159
activity Don’t feel good	0.099997
Sleep Quality	0.098883
activity Pool/Water	0.082183
activity something fun	0.070569
Regularity	0.066748
activity really good sleep	0.06552
Movements per hour	0.059003
activity errands	0.05393
activity party	0.05342
activity nap	0.053121
Time before sleep (seconds)	0.047717
activity no megan	0.046853
activity house stuff	0.043258
activity Sports	0.042034
activity car trouble	0.04106
activity gaming	0.039352
activity Writing	0.038532
activity cleaning	0.037273
Time in bed (seconds)	0.035343
activity phone rabbit hole	0.035001
activity movies	0.031073
Time asleep (seconds)	0.024724
activity work from home	0.016055
activity Language	0.014949
activity Stocks	0.014112
activity Phone call	0.013588
activity therapy	0.012692
Snore time	0.007111
Disrupted Sleep	0.007019

activity family	0.006022
activity drinking	0.003032
activity urges	NaN

Figure 2: Absolute values of correlation coefficients of feature with mood

There were multiple features that showed significant skew and kurtosis. These were normalized using logarithmic transformation and Winsorization. The transformations did not significantly affect the correlation coefficients.

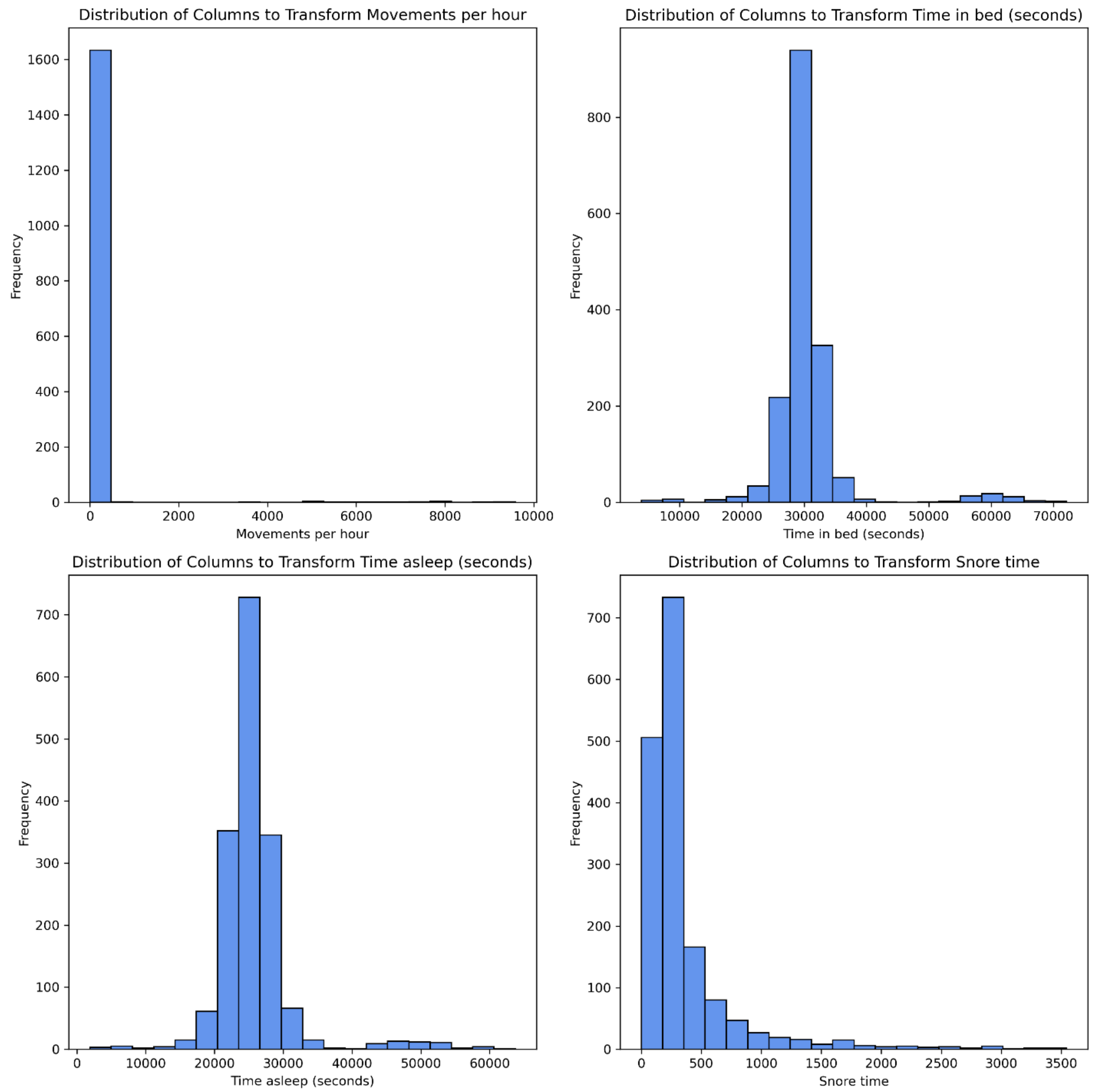

Figure 3: Features with non-normal distributions before transformation

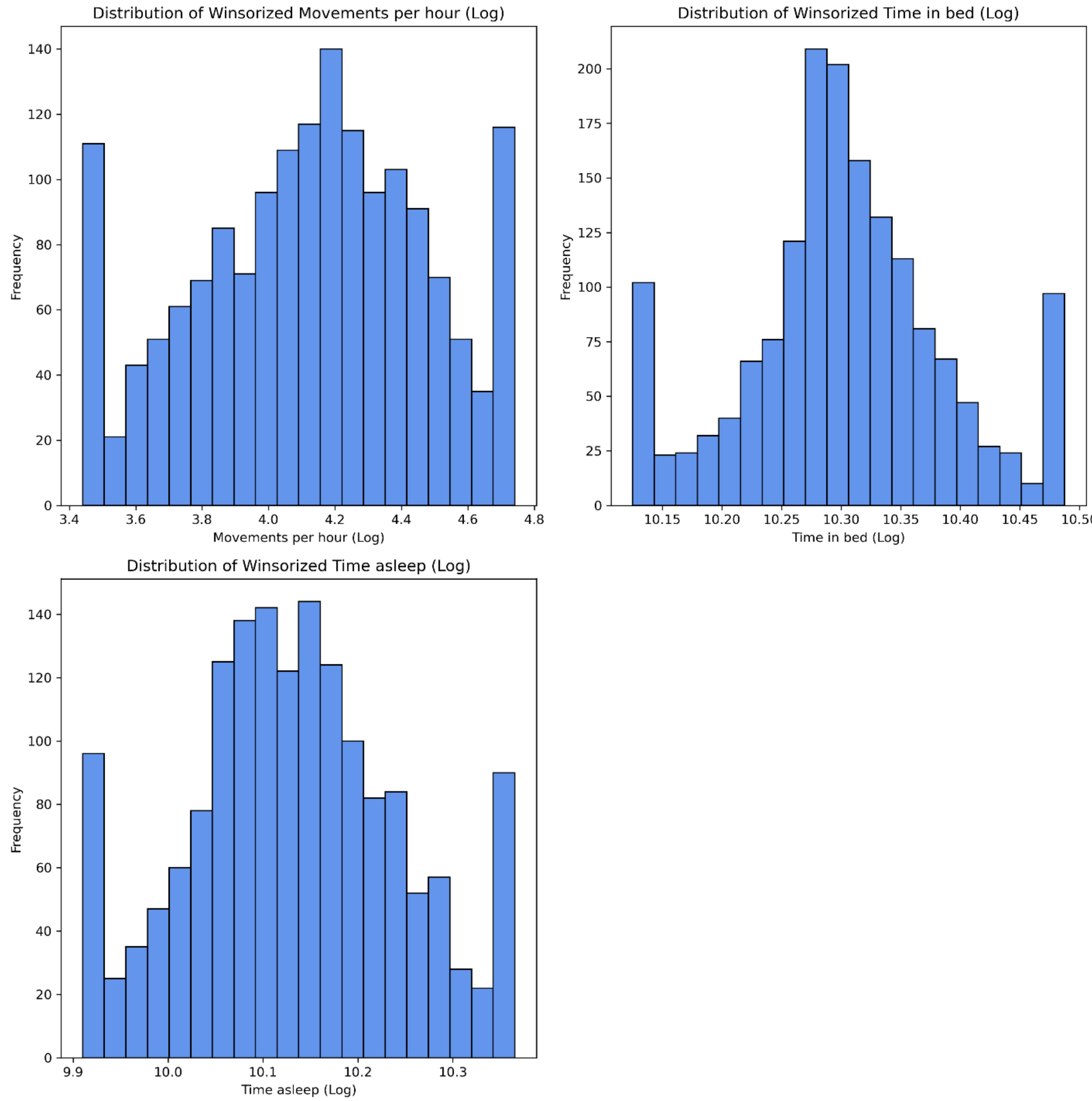

Figure 4: Features after Winsorization and logarithmic transformations

Decision Tree, Random Forest and Gradient Boosting models were tested due to their ability to handle potentially non-linear relationships. The Decision Tree model performed the worst while Gradient Boosting performed the best, but only with a weighted F1 score of 49.20%.

SMOTE was performed and the Gradient Boosting model tested again. Accuracy and weighted F1 scores improved but only slightly.

Several ensemble models were built using Random Forest, Gradient Booster and Logarithmic regression. A grid search was performed to determine the best parameters for each. The final ensemble model performed the best with an accuracy of 59.21%, an F1 score of 23.55% and a weighted F1 score of 53.51%.

Model	**Metric**		
	Accuracy	*F1*	*Weighted F1*
Decision Tree	43.50%	44.69%	44.44%
Random Forest	51.05%	20.42%	46.70%
Gradient Boosting	53.17%	24.13%	49.20%
Gradient Boosting + SMOTE	52.57%	27.04%	51.19%
Ensemble Model	53.47%	27.45%	51.48%
Ensemble Model w/ Logistic Regression	50.76%	28.98%	50.75%
Ensemble w/o SMOTE	57.10%	27.54%	53.38%
Ensemble / Best Parameters	59.21%	23.55%	53.51%

Figure 5: Models and evaluation metrics.

Conclusion

Multiple models were built to increase performance. The Ensemble model using the best parameters from Random Forest, Gradient Boosting and Logistic regression performed the best. Its accuracy is somewhat inflated due to the class imbalance in the target variable but its weighted F1 score showed the best performance.

Even this model is not ready for production. It performs slightly better than guessing. Due to lack of significant correlation between features and the target variable, it is unlikely that the participant's mood can be predicted reliably. The participant is most likely to rank each day as either 'meh' or 'good' regardless of sleep data or activities tracked that day.

Assumptions

The assumption was made that merging on the sleep tracker data would retain enough data with enough variation to create a working model.

Limitations

The major limitation in these datasets is the participant's consistency with mood measurements. This data is very subjective and likely influenced by factors not captured in either tracking application.

Challenges

The class imbalance in the target variable proved challenging as it may be affecting the outcome of the models.

Determining the best model to use with the low correlation coefficients also proved challenging.

Future Uses/Additional Applications

This model could form a basis for improvement in self-recorded tracking. The sleep-tracker app already includes a mood tracker, but it is less robust than the mood-tracking app. Combining the features of both could lead to further insights into how sleep, activities and daily life may affect perceived mood.

Recommendations

Merging the data on the mood dataset instead of the sleep dataset would provide more datapoints that might lead to a more reliable model. Missing sleep data could easily be imputed.

Gathering mood data from a larger group of users would likely help with variation in the data.

A feature of the sleep data could become the target variable instead of mood. This data is more objective.

Implementation

The class imbalance will be addressed using SMOTE and then the Gradient Boosting model retested. PCA or feature engineering may also be used to create features with more meaningful relationships to mood

Ethical Considerations

The data used in this model is private health data so any references to the user's identity have been removed.

It must be made clear that the output is not delivered by a medical professional. It also must be emphasized that correlation does not equal causation. This is a predictive model only and will not diagnose underlying health concerns. The output of the model is only as good as the data that was recorded.

The user is particularly interested in how their sleep affects their mood. They have been informed that if a potential correlation exists, it will be explored but it is not guaranteed that the two are related.

Perhaps the biggest ethical concern is that I am not unbiased as a researcher. I know the participant personally and already have my own thoughts and ideas concerning what affects their moods. Ideally, I would source the data anonymously so I would also be blind to identifying participant information.

References

Dictionary.com. (n.d.). Dictionary.com. https://www.dictionary.com/browse/mood

Fernandez-Berrocal, P., & Extremera, N. (2008). A review of trait meta-mood research. In A. M. Columbus (Ed.), *Advances in psychology research* (Vol. 55, pp. 17–). Nova Science Publishers. Retrieved from Google Books

Peterson, M. J., Benca, R. M., & Merlotti, L. A. (2006). Sleep in mood disorders. *The Psychiatric Clinics of North America, 29*(4), 1009–1032. https://www.psych.theclinics.com/article/S0193-953X(06)00080-3/abstract