TABLE I

EVALUATION METRICS SYSTOLIC

ALGORITHM	RMSE	R2	MAE	MAPE
LINIER REGRESSION	0.0	1.0	0.0	0.0
SUPPORT VECTOR REGRESSOR	0.06	0.99	0.04	0.04
KNN	2.2	0.9	1.6	1.04

Result and Discussion

This project employs the PPG signal blood pressure dataset, which comprises 54 rows of data per patient and 4200 ppg signal columns, along with systolic and diastolic values. The initial stage of this project involves the consolidation of 54 CSV files into a single comprehensive dataset. Following this, statistical analysis and data processing were conducted to address issues such as the elimination of empty data, duplicate data, and the identification and exclusion of outliers. Once pre-processing is complete and deemed suitable for use as training data, the data is then split into two distinct sets: 25% test data and 75% train data. The target variables used are systolic and diastolic in the data.

The algorithms utilized in this study include three distinct approaches: linear regression, support vector regression, and KNN. The resulting modeling training outcomes are presented in Tables I and II, respectively, for the systolic and diastolic evaluation metrics.

Table I presents the evaluation values for the three algorithms used. The linear regression algorithm is the most accurate, with an R2 value of 100% and RMSE, MAE, and MAPE values close to 0. Therefore, it can be concluded that the linear regression algorithm is able to provide highly accurate and stable predictions of systolic values. This indicates that the Linear Regression model can be used with confidence to predict systolic values on new data with a high degree of accuracy.

A comparison of Table II with the diastolic target reveals that the Linear Regression algorithm also demonstrates excellent performance in predicting diastolic values. With an R2 value of 100% and RMSE, MAE, and MAPE values close to zero, Linear Regression consistently produces highly accurate and stable predictions of diastolic values. This confirms that the Linear Regression model is not only suitable for predicting systolic values but also for diastolic values. Consequently, this model can be relied upon to provide accurate predictions of blood pressure values on new data, thereby conferring significant benefits in the effective and efficient management and monitoring of patient health.

It can be concluded that the Linear Regression algorithm consistently provided excellent prediction results for both systolic and diastolic values. The R2 or accuracy value was 100% in both evaluation tables, and the RMSE, MAE, and MAPE values were close to zero. These results demonstrate that Linear Regression has an excellent ability to model the relationship between input features and blood pressure values. This conclusion indicates that the Linear Regression model can be relied upon and trusted to provide accurate and stable predictions of blood pressure on new data, thus making an important contribution to health management and patient monitoring with a high degree of accuracy.

TABLE II

EVALUATION METRICS DIASTOLIC

ALGORITHM	RMSE	R2	MAE	MAPE
LINIER REGRESSION	0.0	1.0	0.0	0.0
SUPPORT VECTOR REGRESSOR	0.06	0.99	0.05	0.1
KNN	2.2	0.9	1.6	3.04

Nevertheless, it is possible that the three algorithms may still exhibit shortcomings, either in the preprocessing stage, due to the complexity of the data, or for other reasons. Consequently, as a writer, I would be grateful for input or evaluation.