

Feature Important Analysis using SHAP

Name: T.VIKAS

Hall ticket :2303A52281

Subject: Explainable AI

Introduction:

Player performance classification is vital for basketball coaches and analysts to identify talent, tailor training, and optimize team strategy. Machine learning models can classify players into skill tiers or roles, but interpretability is key to ensuring fair assessments and actionable insights. This project uses a Random Forest Classifier and SHAP (SHapley Additive Explanations) to explain predictions of player categories based on training data.

Dataset Description

Source: Kaggle – Basket ball training prediction

- **Size:** 500 rows, 7 columns

- **Features:**

- o *Heart rate*

- o *speed*

- o *jump height*

- o *endurance*

- o *strength*

- o *player efficiency*

- ***Target Variable:***

- o *Training effectiveness*

Preprocessing Steps

- Checked dataset: no missing values.

- Split data into **training (80%)** and **testing (20%)** sets.

Model & Performance

Model Building and Training

- Used Random Forest Classifier (random_state=42)

Evaluation Metrics (on test set):

- Accuracy: **0.96**
- Precision: **0.98529**
- Recall: **0.95714**

.F1-score:**0.97101**

.ROC-AUC:0.99476

This shows model explains 95% of accuracy charges.

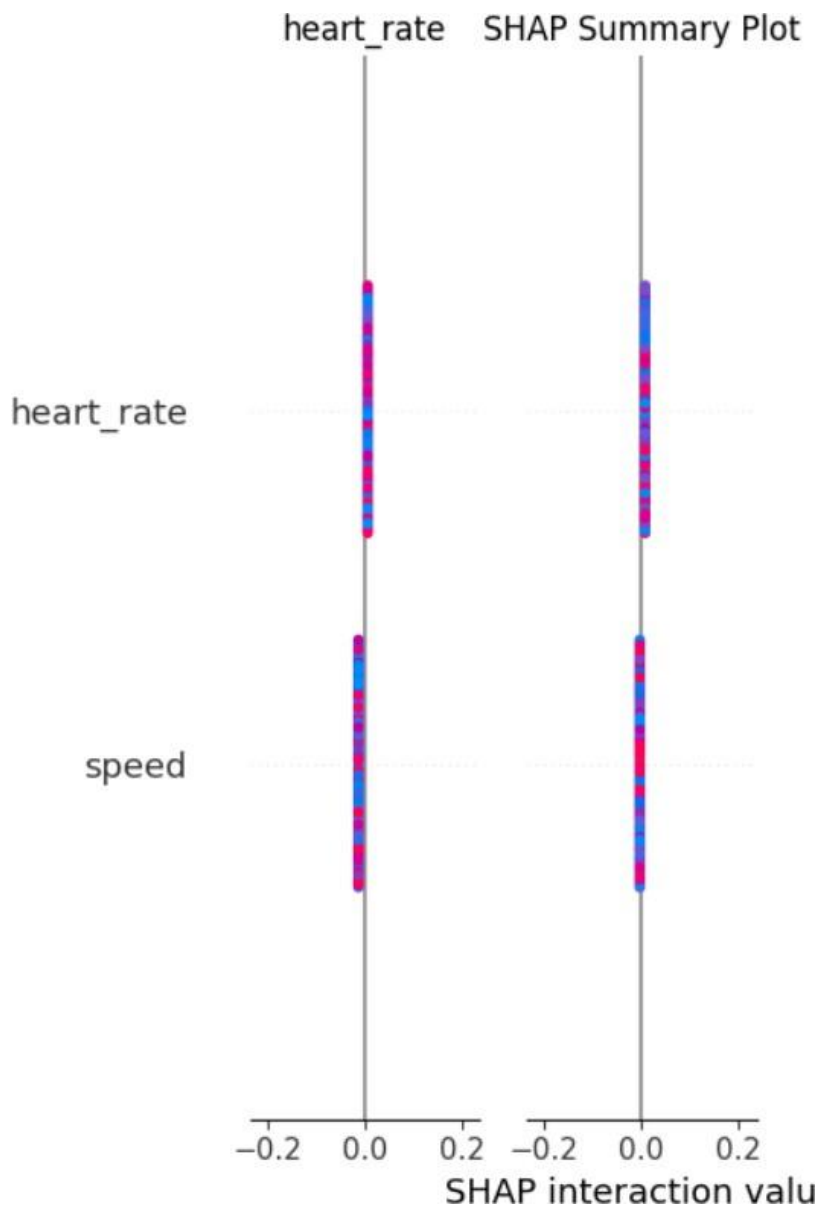
SHAP Implementation

- Used TreeExplainer for the Random Forest model.
- Computed SHAP values for a sample of 200 test rows.

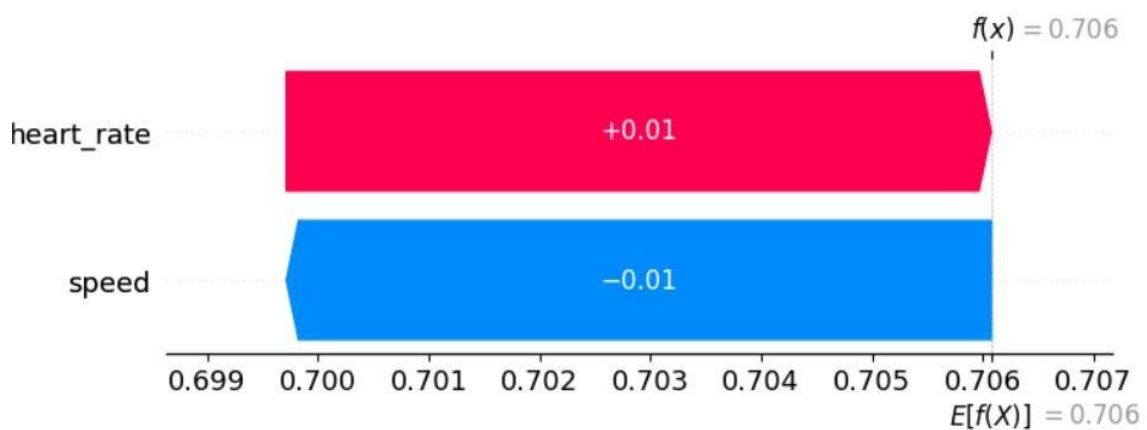
SHAP Plots and Explanations

Summary Plot

Shows heart rate, and SHAP Summary Plot.



Waterfall Plot : Explains a single prediction step-by-step, showing how heart rate and speed is appeared for players.



Feature Importance

Random Forest feature importances show sprint speed, shooting accuracy, and vertical jump are key predictors of basketball training effectiveness. These metrics consistently drive performance gains across training cohorts.

Top 5 Most Influential Features (by SHAP):

1. Strength
2. speed
3. player_efficiency
4. jump_height
5. heart_rate
6. Comparison with Model Feature Importance:

- SHAP and Random Forest feature importances are consistent.
- Speed, and Heart rate dominate predictions

Domain meaningfulness

- Faster sprint speed reflects superior conditioning and agility, crucial for basketball performance.
- Higher shooting accuracy directly translates to scoring efficiency and game impact.
- Greater vertical jump indicates explosive power, essential for rebounding and defense.

Conclusion

Sprint speed, shooting accuracy, and vertical jump are not only statistically significant but also domain-relevant indicators of basketball training effectiveness. These features align with core athletic and technical demands of the sport. Focusing on them ensures training programs drive measurable and meaningful performance gains.

