Model Card: Random Forest Regressor for Karting Lap Time Prediction

# 1. Model Details

Model Name: Tuned SVR model   
Framework: Scikit-learn (v1.4+)   
Creator: Karting Performance Prediction Project – Bindu Kummeta   
Date: 2025-06-09

# 2. Intended Use

This model is designed to predict lap times in competitive karting based on various features such as driver attributes, kart specifications, and environmental conditions. It is primarily intended for educational use and performance analysis.

# 3. Factors

The model takes into account the following features:   
- DriverAge   
- DriverWeight   
- KartPowerHP   
- TrackLength\_m   
- NumCorners   
- ElevationChange\_m   
- TirePressure\_psi   
- GearRatio   
- FuelLoad\_kg   
- ReactionTime\_ms   
- AggressivenessScore   
- TrainingHoursPerWeek   
- TrackTemperature\_C   
- Humidity\_%   
- WindSpeed\_kmh   
  
Performance may vary significantly based on environmental conditions and physical kart setup.

# 4. Metrics

The model was evaluated using the following metrics:   
- R² Score   
- Root Mean Squared Error (RMSE)   
  
Final Results on Test Data:   
- R² Score: 0.9566   
- RMSE: 4.90 seconds

# 5. Training Data

The model was trained on a simulated dataset containing 10,000 samples of karting lap performance data. The data includes a wide range of synthetic examples representing different racing conditions and driver profiles.

# 6. Evaluation Data

A separate 15% of the dataset was held out for final testing and evaluation. Cross-validation was used on the training set to ensure generalization.

# 7. Ethical Considerations

As this model is built using synthetic data based on one racer’s race data for educational purposes, it carries no direct ethical implications. However, real-world deployment should consider factors like fairness across age groups and physical capabilities. Since data is synthetically generated, all the models performed very well, overfitting to data.

# 8. Caveats and Recommendations

This model should not be used for real-time safety-critical decision-making in motorsport or commercial racing environments. It is an educational and analytical tool and requires real-world calibration and validation. Since this is synthetic data, it looks ideal.