FACULTAD DE MEDICINA VETERINARIA Y ZOOTECNIA
DEPARTAMENTO DE MEDICINA Y ZOOTECNIA DE RUMIANTES

Application of Machine Learning Algorithms for Estimating Body Weight in Horses Using Morphometric Measurements

J. C. Angeles Hernandez[1], G. Mariano Hernandez[1], X. K. Aguilar Amaro[1], N. A. Cruz Gutierrez[2], E. Cardoso Gutierrez[3], R. Gonzalez Lopez[1]

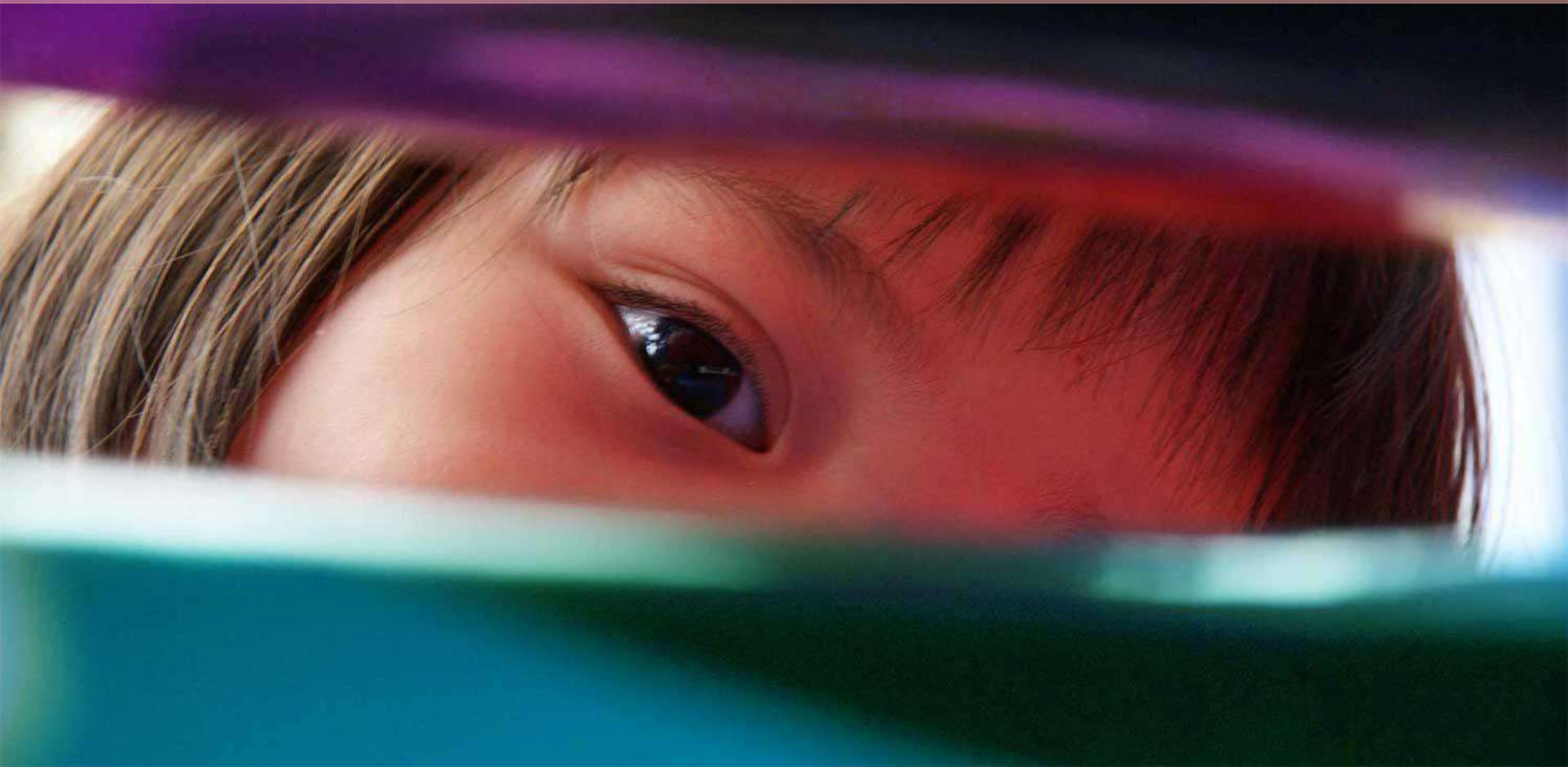
[1]Universidad Nacional Autónoma de México, Departamento de Medicina y Zootecnia de Rumiantes

[2]Universitat de les Illes Bale, Departamento de Ciencias Matemáticas e Informática

[3]Universidad Autónoma del Estado de Hidalgo, Instituto darse Ciencias Agropecuarias

Orgullosamente UNAM

Introduction



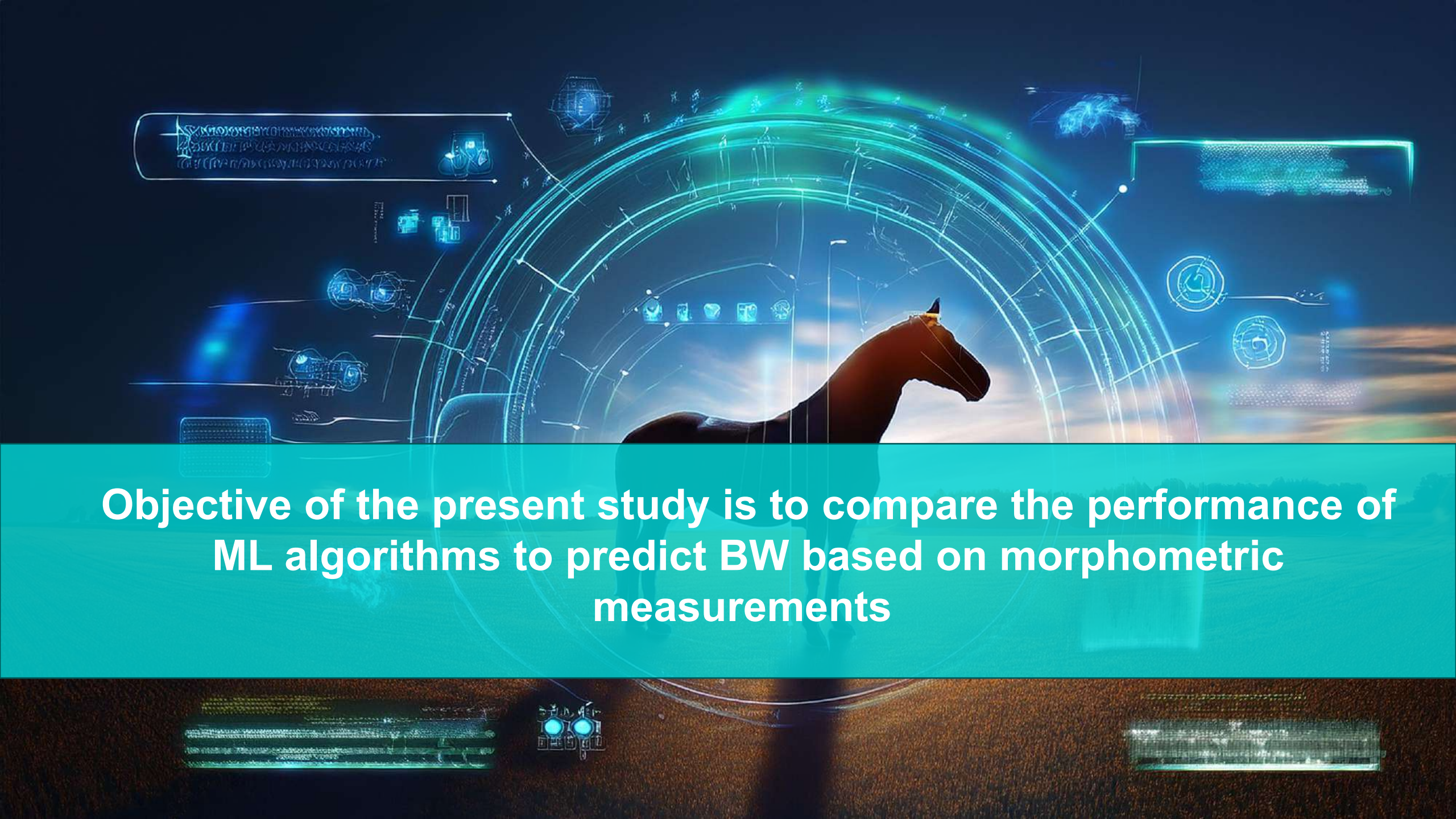
Estimation of body weight (BW) in horses is essential for effective health monitoring, dosing and nutritional planning.

Maybe for some stables it is common to have a scale, however, in many countries such facilities are not available, so it is useful to use proxis for live weight estimation


Traditional methods based on
morphometric measurements
such as height at withers, chest
girth and body length have been
widely used


Through the development of
multiple regression models


Machine learning (ML) algorithms have shown the potential to predict BW.



Objective of the present study is to compare the performance of ML algorithms to predict BW based on morphometric measurements

MATERIALS AND METHODS

Mexico: State of Mexico, Nuevo Leon and Mexico City

Location

Mexico: State of Mexico, Nuevo Leon and Mexico City

Database

142 horses (88 females and 54 males) over 2 years of age

Morphometric measurements

Height at withers (HW), chest girth (CG), neck circumference (NC) and body length (BL)

Additionaly: Age and BCS

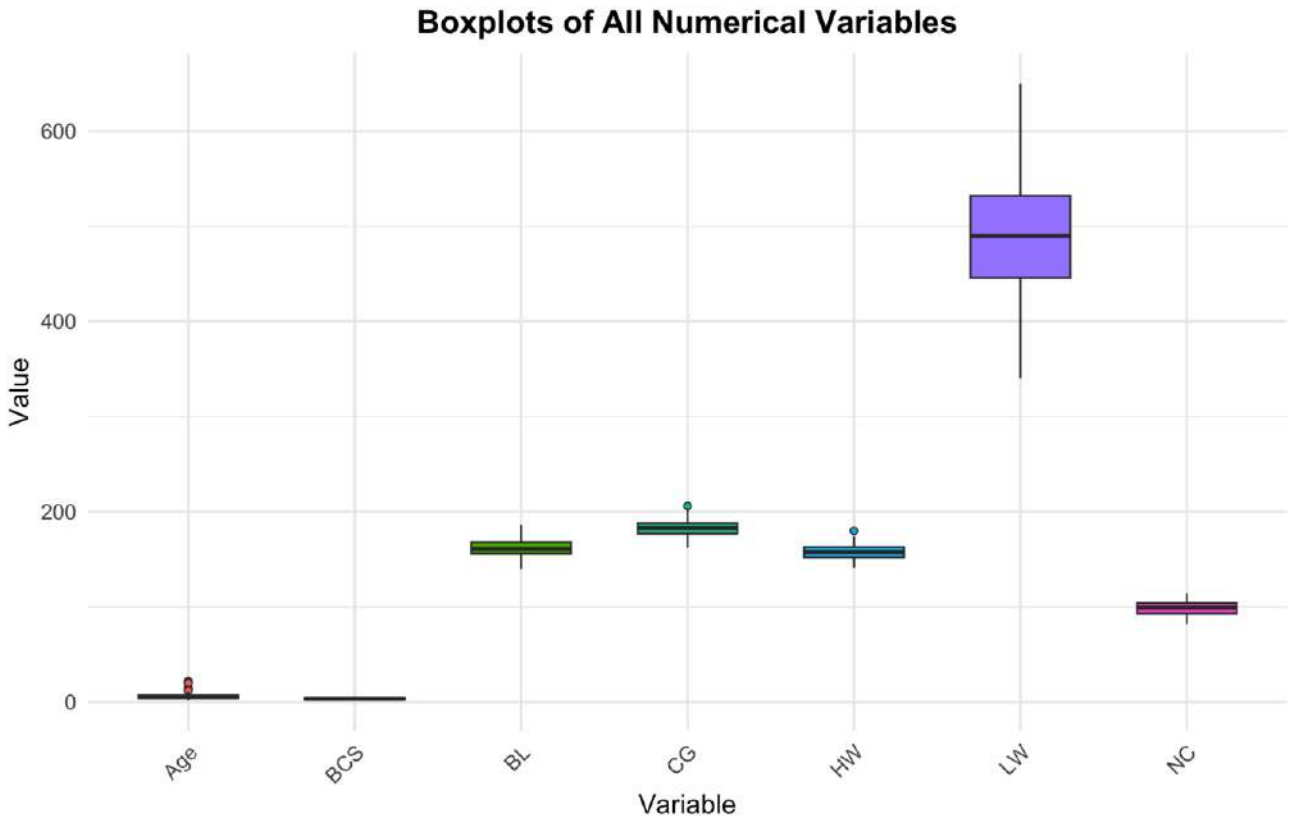
Partition

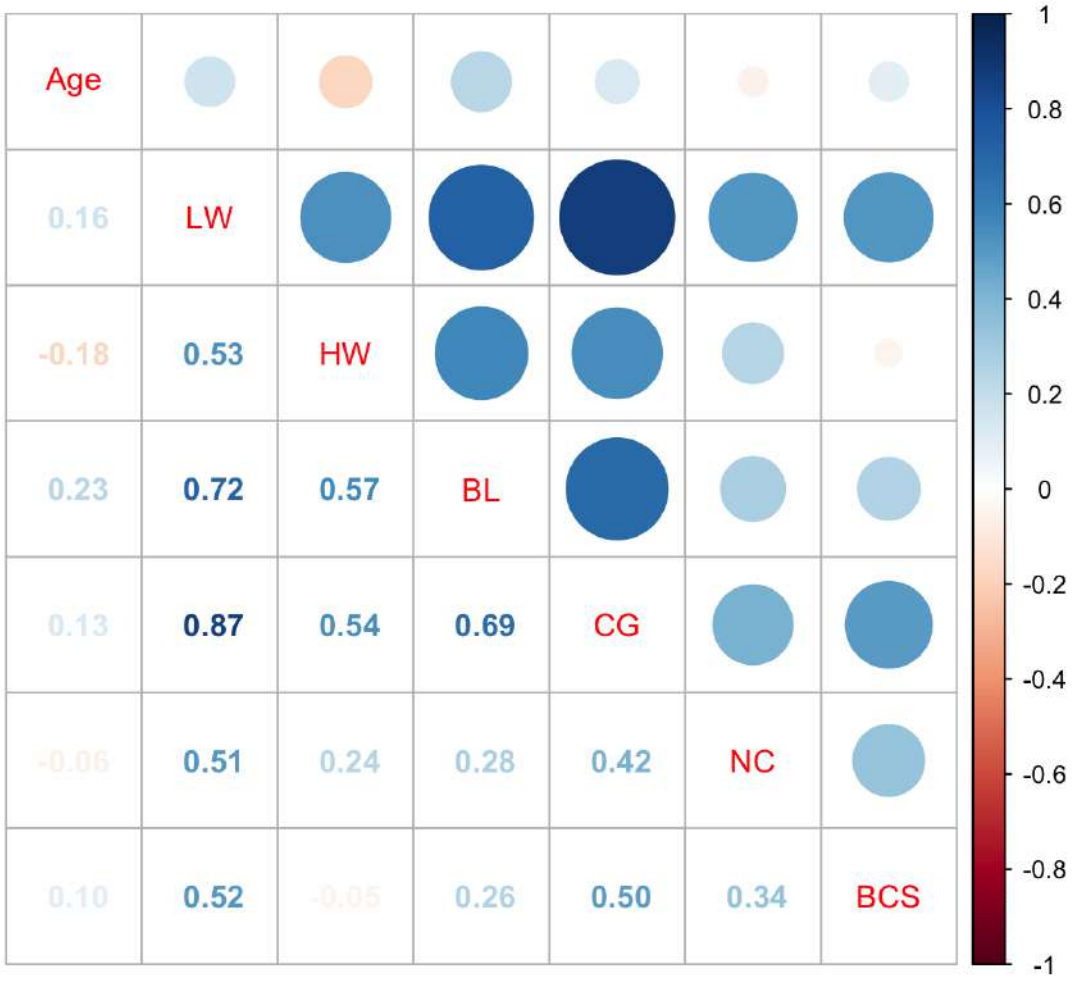
The data set was randomly divided into 70% for training and 30% for testing

Partition

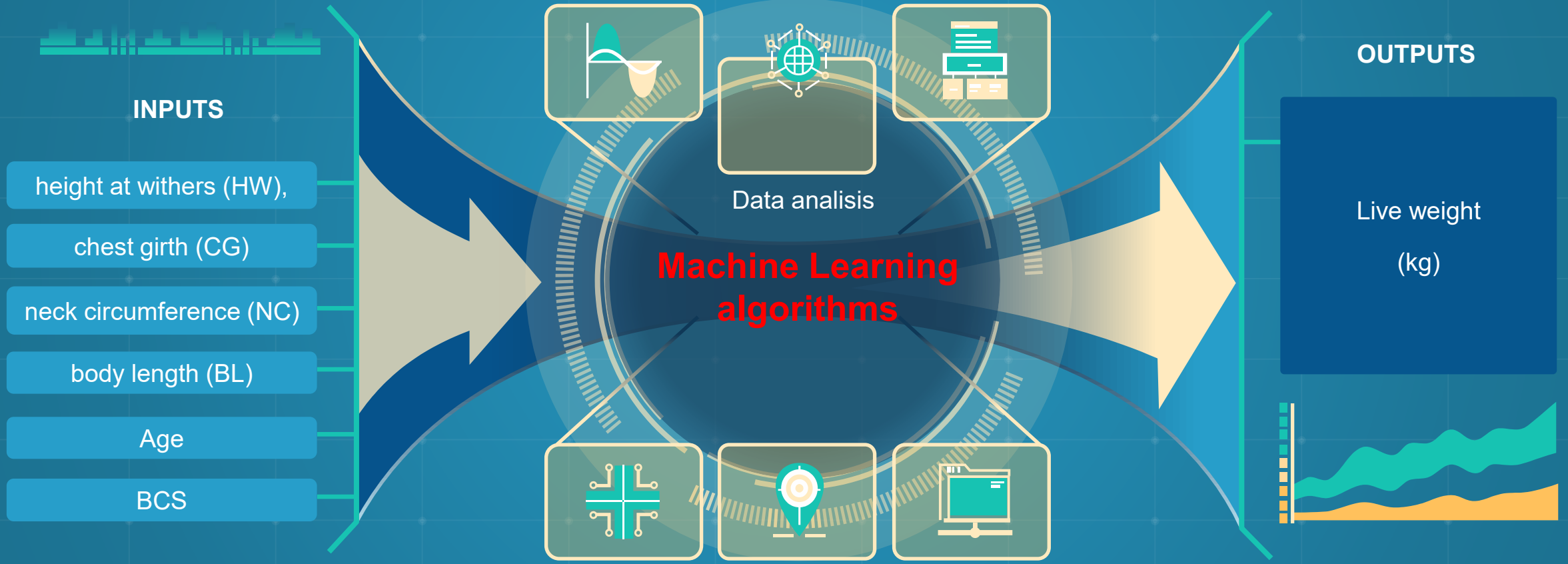
The data set was randomly divided into 70% for training and 30% for testing

Exploratory análisis





Data analysis



Machine Learning algorithms

- Multiple linear regression (MLR)
- MLR with interaction (MLRi)
- Stepwise MLR (MLRs)
- Artificial neural networks (ANN)
- Generalized additive model with penalized splines (GAMLSS)
- Support vector machines (SVM)

Fitting performance

- Coefficient of determination (r square)
- Root mean square prediction error (RMSE)
- Mean absolute percentage error (MAPE)

Results

📊 **Table 1.** Goodness-of-Fit Metrics for Machine Learning Models

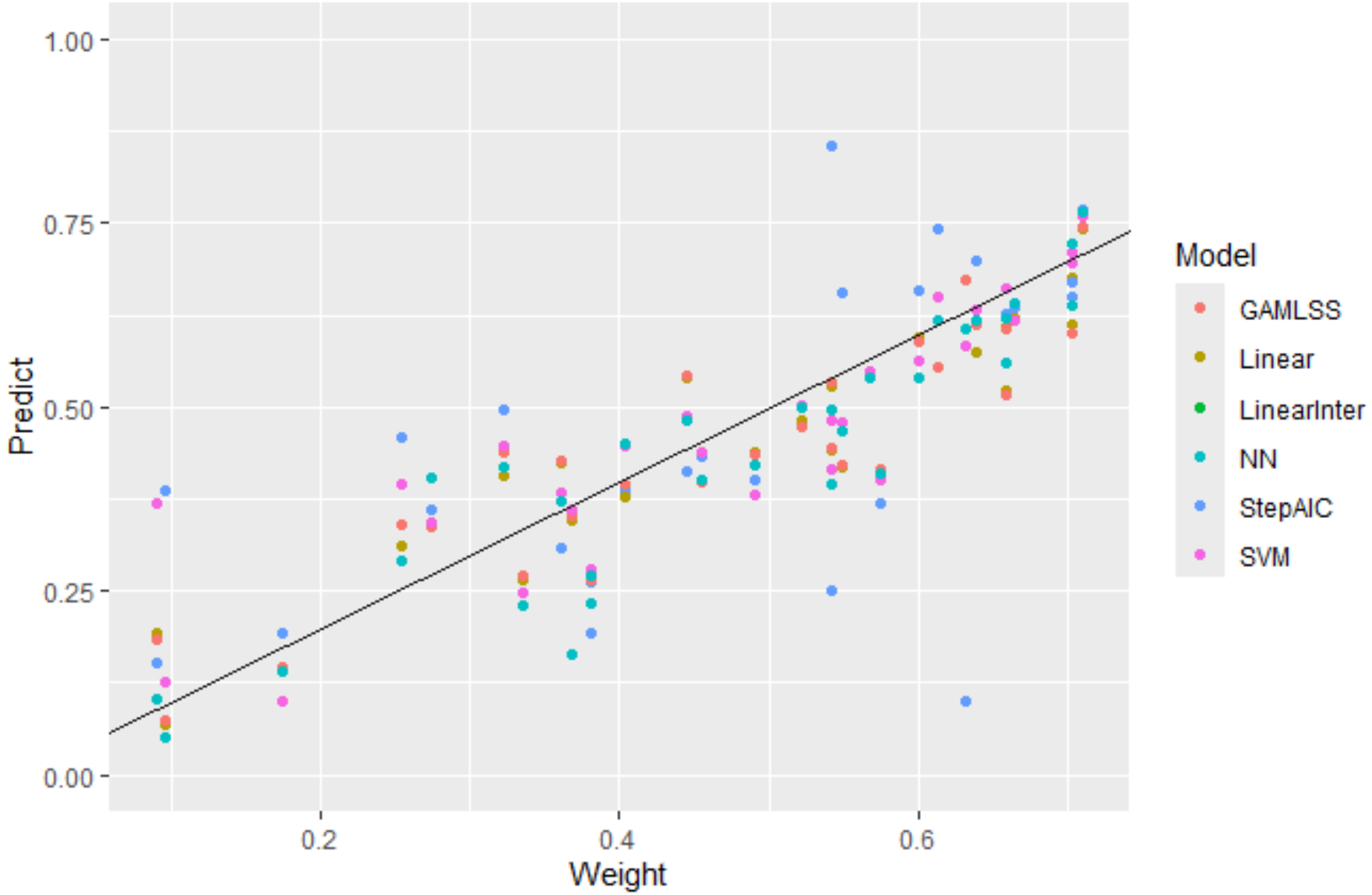
Model	R Square		MAPE		RMSE	
	Train	Test	Train	Test	Train	Test
MLR	**83.41%**	**81.70%**	2.08%	9.71%	0.05	0.05
MLR interactions	**94.94%**	**-53.07%**	2.95%	4.83%	0.06	0.06
MLR StepAIC	**94.94%**	**-53.07%**	3.57%	4.54%	0.07	0.07
GAMLSS	**84.12%**	**80.75%**	**2.08%**	**9.71%**	**0.05**	**0.05**
ANN	89.03%	77.71%	**4.51%**	**4.42%**	0.08	0.08
SVM	90.27%	74.19%	**4.48%**	**4.41%**	0.08	0.08

📊 **Table 1.** Goodness-of-Fit Metrics for Machine Learning Models

Model	R Square		MAPE		RMSE	
	Train	Test	Train	Test	Train	Test
MLR	**83.41%**	**81.70%**	2.08%	9.71%	0.05	0.05
MLR interactions	**94.94%**	**-53.07%**	2.95%	4.83%	0.06	0.06
MLR StepAIC	**94.94%**	**-53.07%**	3.57%	4.54%	0.07	0.07
GAMLSS	**84.12%**	**80.75%**	**2.08%**	**9.71%**	**0.05**	**0.05**
ANN	89.03%	77.71%	**4.51%**	**4.42%**	0.08	0.08
SVM	90.27%	74.19%	**4.48%**	**4.41%**	0.08	0.08

Overfitting

Figure 2. Scatterplot of observed (x-axis) vs. predicted (y-axis) values, illustrating model accuracy on test data.



Conclusions

01

Based in R square and RMSE, MLR and GAMLSS showed the best overall performance for BW estimation, high accuracy and generalisation

02

Accorgding with MAPE, SVM and ANN showed good predictive ability, however, their susceptibility to overfitting limits their reliability

Areas for Improvement / Next Steps

- **Expand sample size**: Increase the dataset to include **>500 horses** for better model robustness.
- **Include additional variables**: Consider **breed, activity level**, and other relevant factors.
- **Explore alternative ML algorithms**: Test models like **decision trees, rule-based systems, and random forest**.
- **Enhance algorithm usability**: Focus on **portability** and develop a **user-friendly app** for practical implementation.

tryadd®

Confiabilidad Global en Nutrición Animal

UAEH

Universidad Autónoma del Estado de Hidalgo

UIB Universitat de les Illes Balears

Universidad Nacional Autónoma de México — Facultad de Medicina Veterinaria y Zootecnia — 1853

Thank you

UNAM