



modelStudio: Interactive Studio with Explanations for ML Predictive Models

Hubert Baniecki¹, Przemysław Biecek¹

¹Faculty of Mathematics and Information Science, Warsaw University of Technology

Introduction

Automate explanation of machine learning predictive models.

The **modelStudio** R package generates advanced interactive and animated model explanations in the form of a serverless HTML site. It combines R with D3.js to produce plots and descriptions for various local and global explanations. Tools for model exploration unite with tools for EDA (Exploratory Data Analysis) to give a broad overview of the model behaviour. Using modelStudio is a fast and condensed way to get all the answers without much effort.

It requires only a few lines of code.

Customise your own dashboard to explore the model.

Interactive Model Studio

Lucas

Break Down [Local]
SHAP Values [Local]
Ceteris Paribus [Local]
Feature Importance [Global]
Partial Dependency [Global]
Accumulated Dependency [Global]
Feature Distribution [EDA]

+

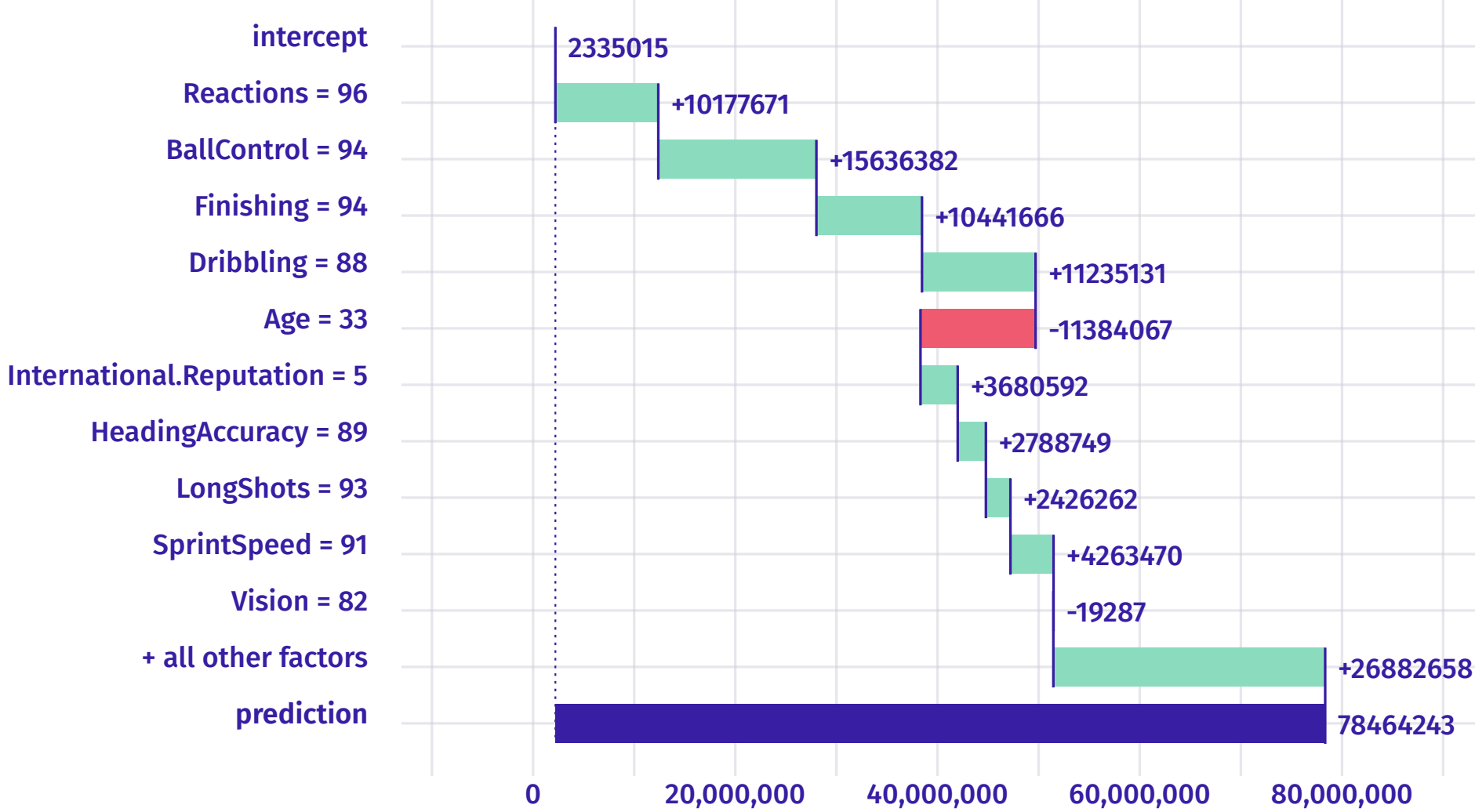
Site built with modelStudio v0.1.9 on 2019-11-20 15:14:25

Interactive Model Studio

Cristiano.Ronaldo

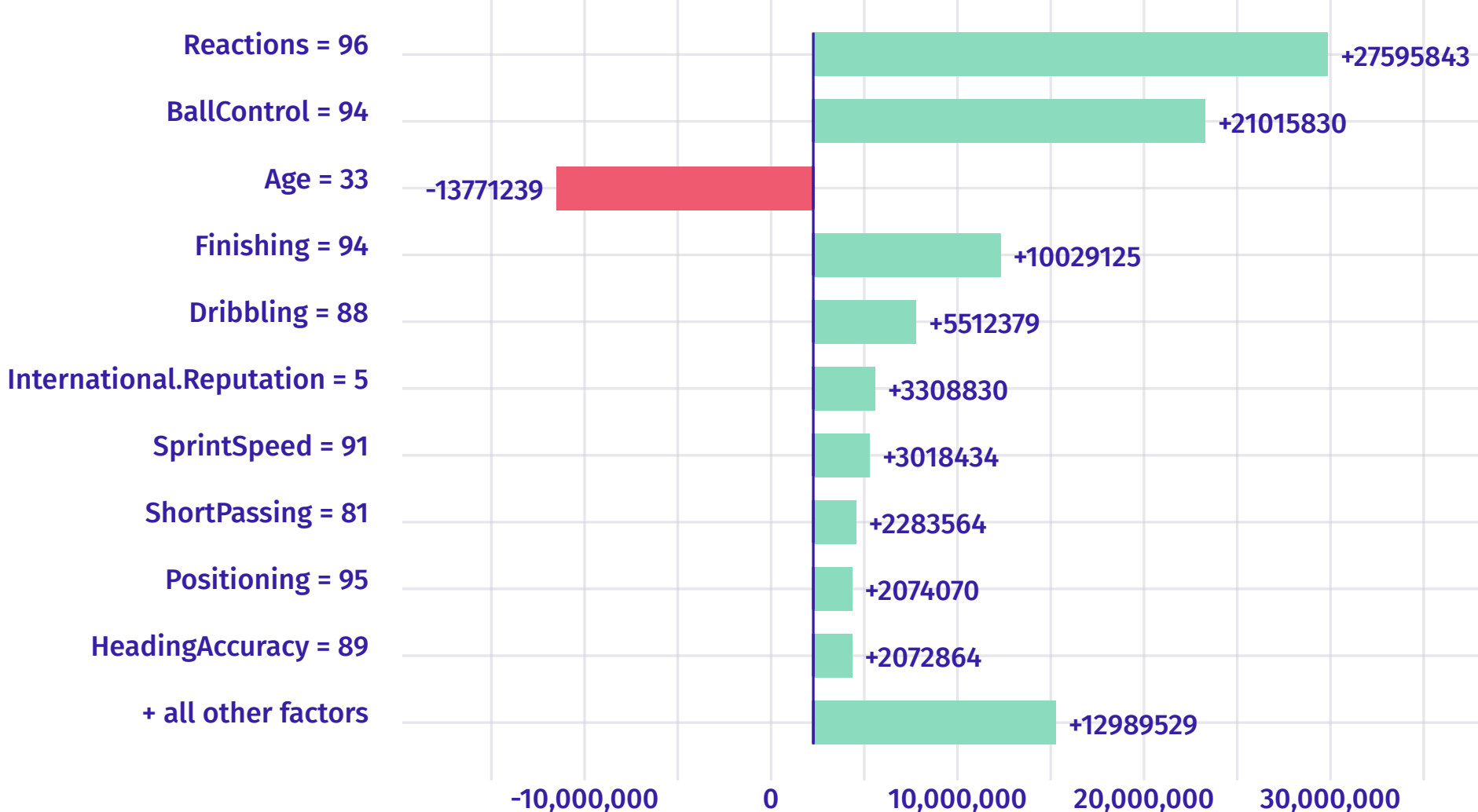
Break Down

D X



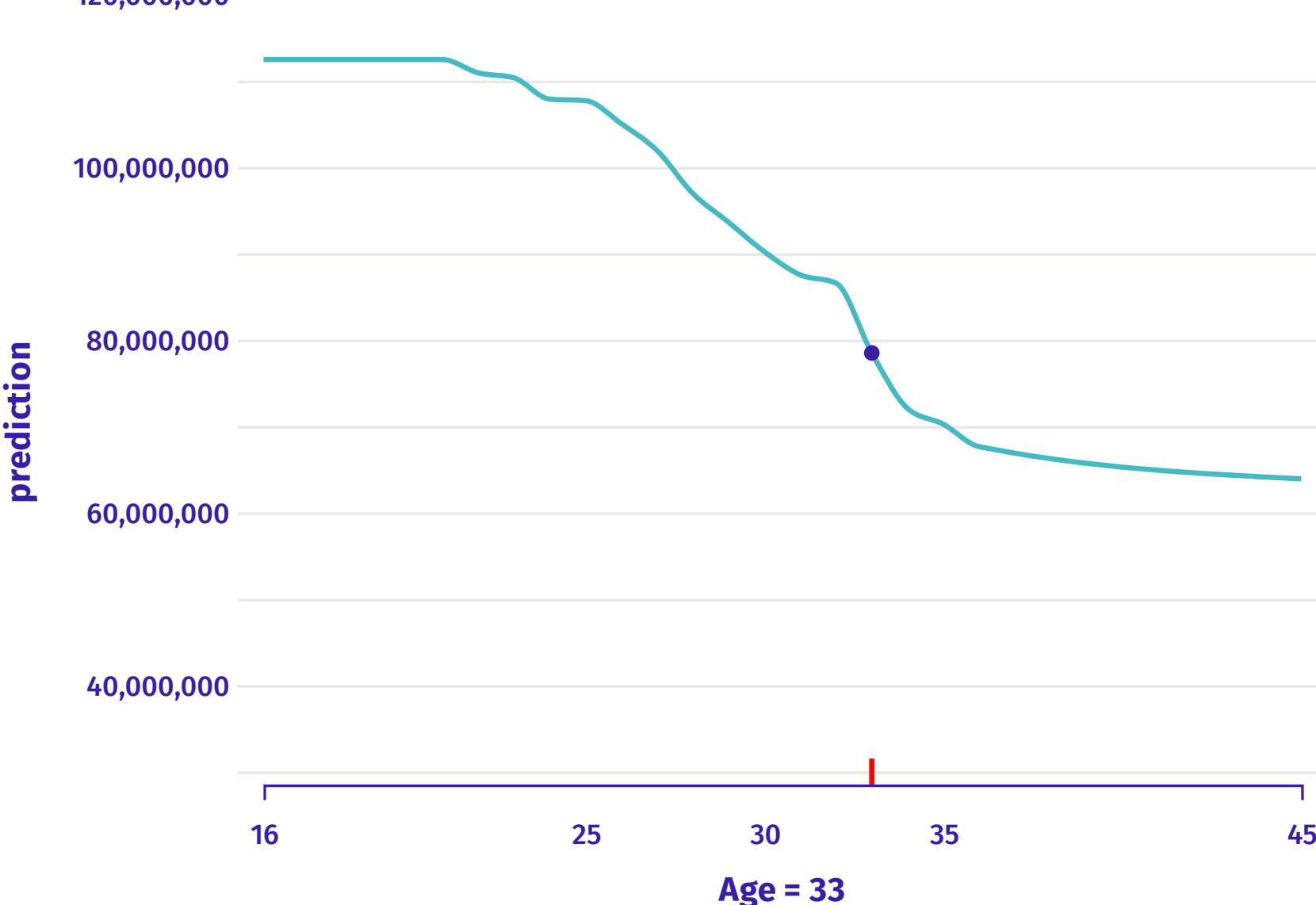
SHAP Values

D X



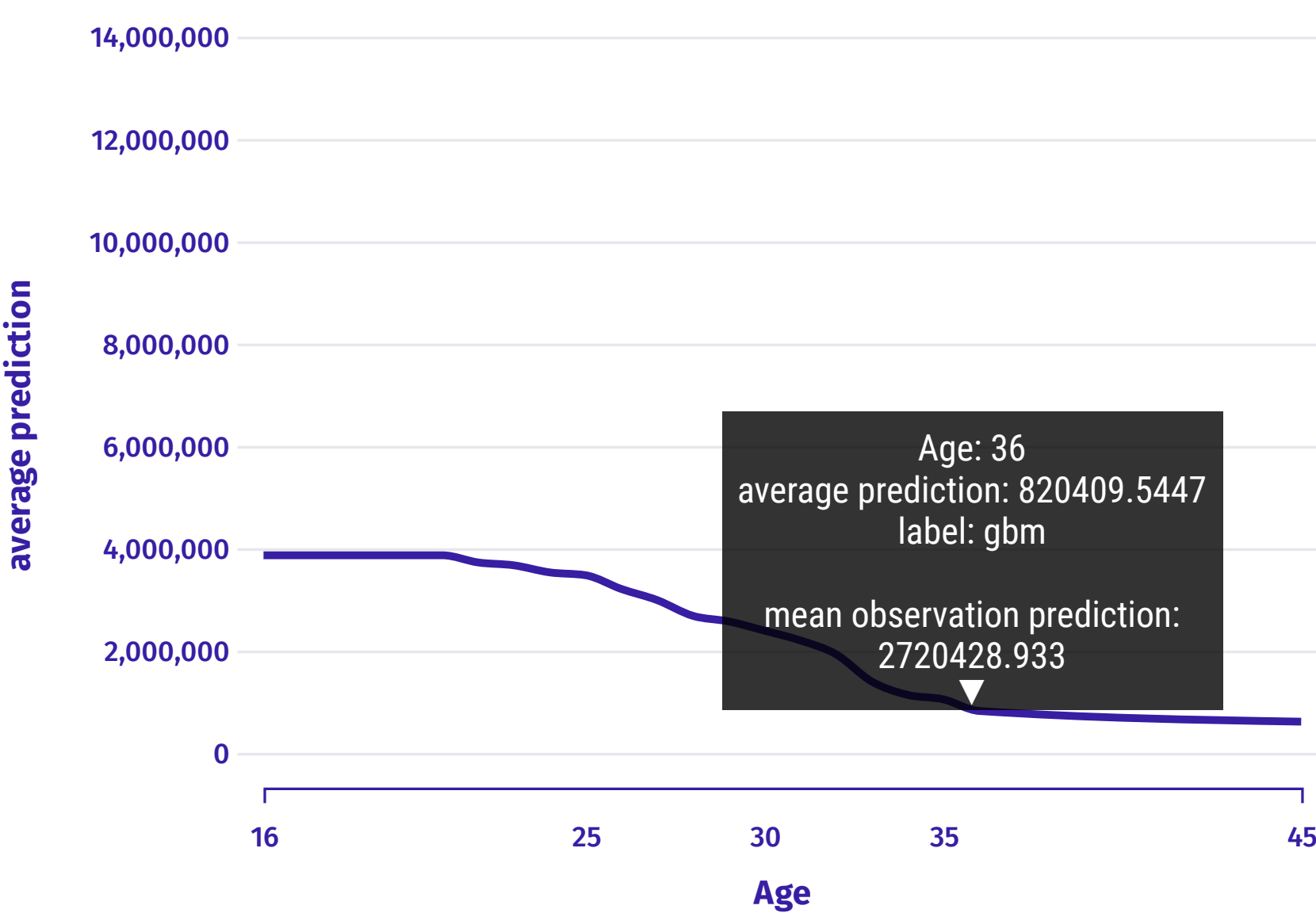
Ceteris Paribus

D X



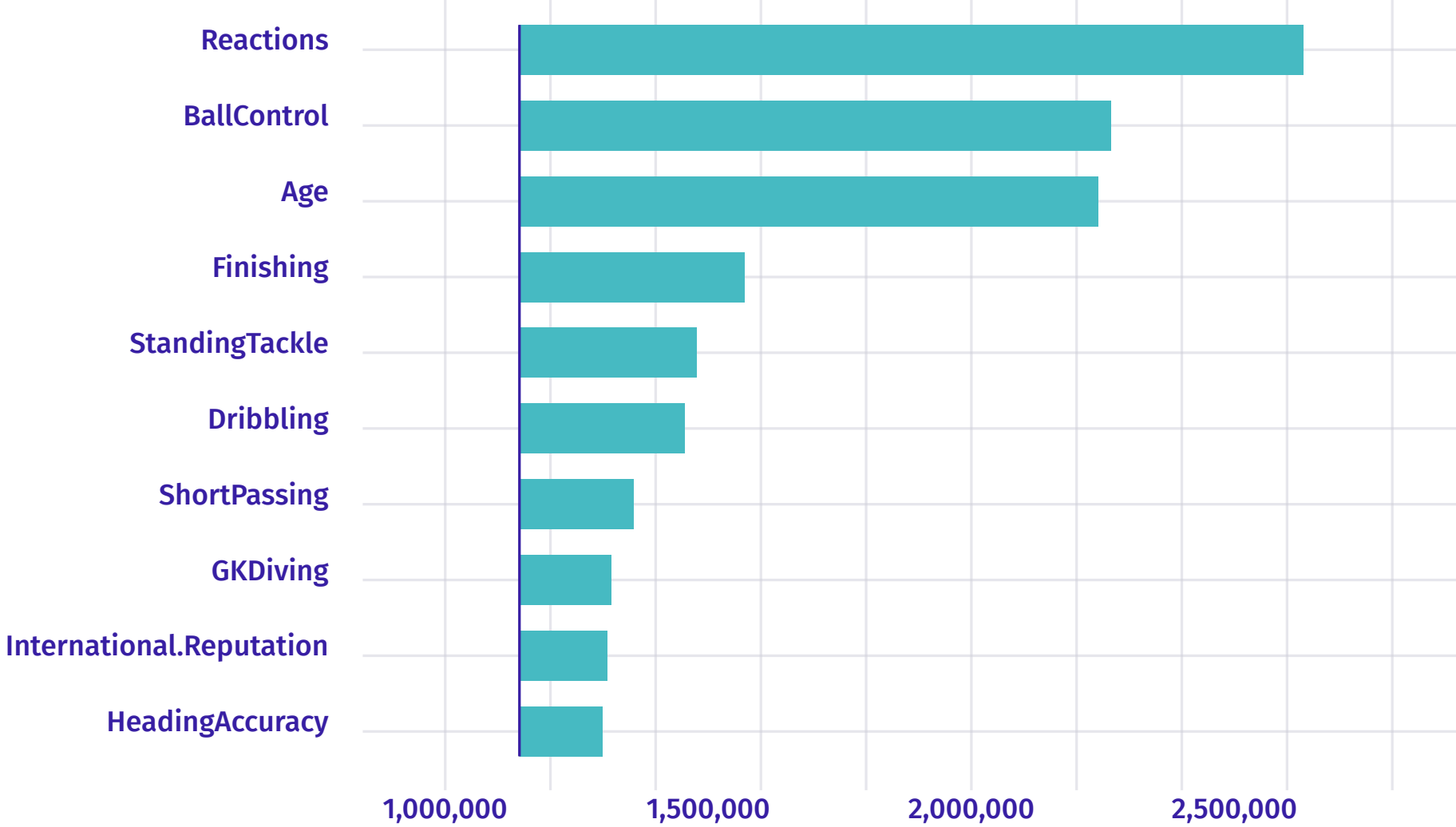
Partial Dependency

D X



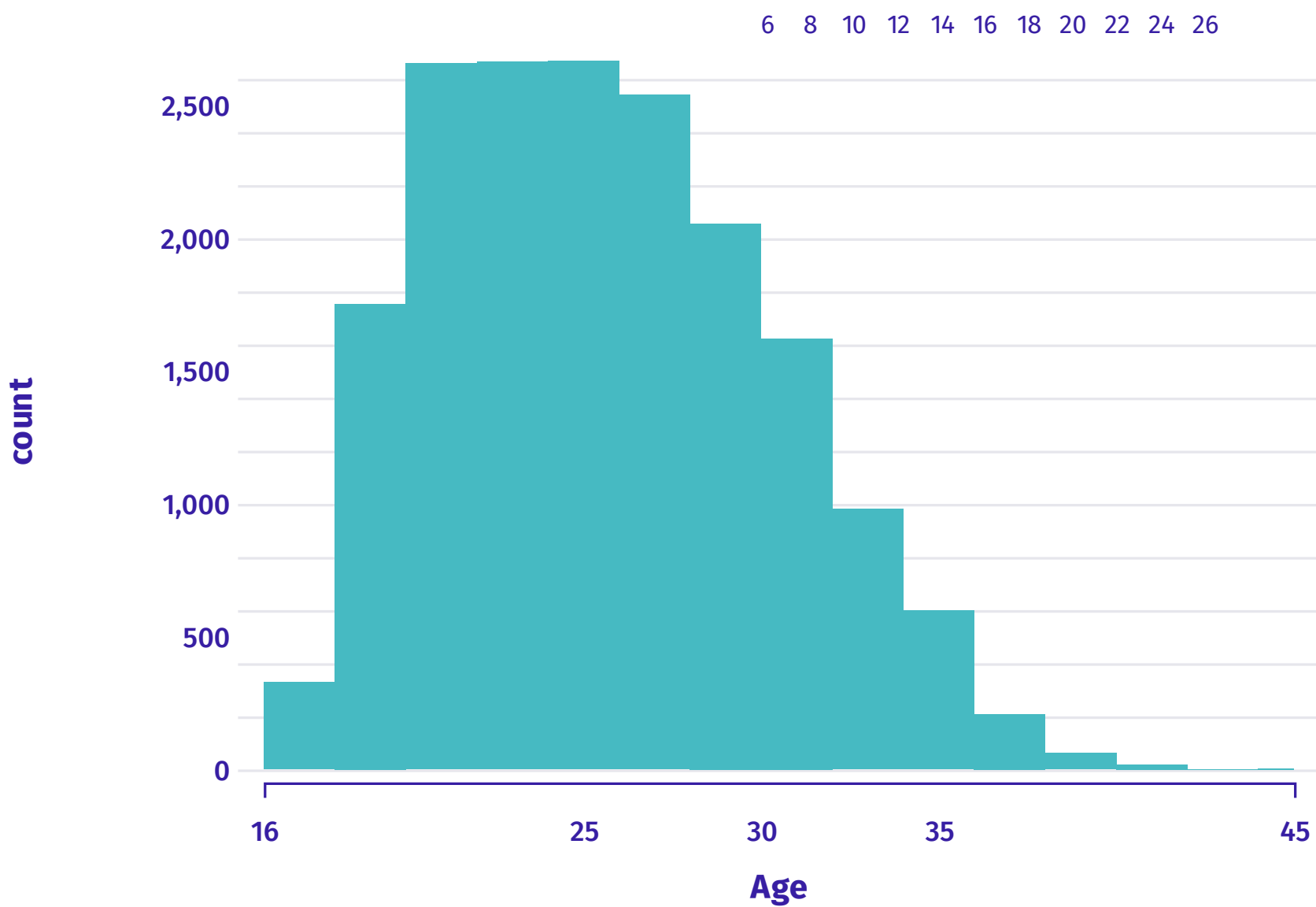
Feature Importance

D X



Feature Distribution

X



Site built with modelStudio v0.1.9 on 2019-11-21 22:59:49

Figure 1: Exemplary modelStudio HTML output[1] layout for the GBM model predicting player's value on the FIFA19 data.

Dashboard Plots

Seven possible plots to choose from are implemented. There are three local explanation plots, three global explanation plots and a feature density plot. **Local explanations** are designed to better understand model behaviour around a single observation.

- **Break Down** plot and **SHAP Values** plot present variable contributions to a model prediction. Both of them come from the **iBreakDown** [3] R package.
- **Ceteris Paribus** plot presents model responses around a single point in the feature space.

Global explanations are designed to allow for better understanding of how the model works in general, for some population of interest.

- **Feature Importance** plot presents permutation based feature importance.
- **Partial Dependency** plot presents averages from N number of Ceteris Paribus Profiles
- **Accumulated Dependency** plot presents accumulated local changes in Ceteris Paribus Profiles.

The last explanations are implemented in the **ingredients** [4] R package.

Key Features

- **Serverless Architecture**
- **Easy to Extend and Deploy**
- **Fully Automated**
- **Multiple Observation Choice**
- **Great Customisability**
- **Supports Plot Descriptions**
- **Model and Language Agnostic**



SCAN ME

Conclusion

The **modelStudio** package is easy to use and its output is intuitive to explore. Automation is convenient and interactivity adds an another dimension to visualisations. All of this enhance explanation of machine learning predictive models. More examples can be found in the **vignette** and on **GitHub**.

- <https://qrco.de/modelStudioVignette>
- <https://github.com/ModelOriented/modelStudio>

References

- [1] modelStudio on the FIFA19 data. URL <https://pbiecek.github.io/explainFIFA19/>.
- [2] Hubert Baniecki and Przemyslaw Biecek. modelStudio: Interactive Studio with Explanations for ML Predictive Models. *The Journal of Open Source Software*, Nov 2019. doi: 10.21105/joss.01798. URL <https://doi.org/10.21105/joss.01798>.
- [3] Alicja Gosiewska and Przemyslaw Biecek. ibreakdown: Uncertainty of model explanations for non-additive predictive models, 2019. URL <https://arxiv.org/abs/1903.11420v1>.
- [4] Przemyslaw Biecek, Hubert Baniecki and Adam Izdebski. *ingredients: Effects and Importances of Model Ingredients*, 2019. URL <http://CRAN.R-project.org/package=ingredients>.

Acknowledgements

Hubert Baniecki was financially supported by the grant of Polish Centre for Research and Development *POIR.01.01.01-00-0328/17*.

Contact

- ✉ hbaniecki@gmail.com



SCAN ME