

# Variable Importance Plots: An Introduction to the vip Package

by Brandon M. Greenwell and Bradley C. Boehmke

**Abstract** In the era of big data, it is becoming more of a challenge to not only build state-of-the-art predictive models, but also gain an understanding of what's really going on in the data. For example, it is often of interest to know which, if any, of the predictors in a fitted model are relatively influential on the predicted outcome. Some modern algorithms like random forests and gradient boosted decision trees have a natural way of quantifying the importance or relative influence of each feature. Other algorithms like naive Bayes classifiers and support vector machines are not capable of doing so and model-agnostic approaches are generally used to measure each predictor's importance. Enter **vip**, an R package for constructing variable importance scores/plots for many types of supervised learning algorithms using model-based and novel model-agnostic approaches.

## Introduction

Variable importance (VI).

## Model-specific VI

TBD.

## Model-agnostic VI

TBD.

## A TensorFlow example

TBD.

## Summary

TBD.

## Bibliography

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL <https://www.R-project.org/>. ISBN 3-900051-07-0. [p]

Brandon M. Greenwell  
Department of Mathematics and Statistics  
Wright State University  
3640 Colonel Glenn Hwy  
Dayton, OH 45435  
United States of America  
ORCID: [0000-0002-8120-0084](https://orcid.org/0000-0002-8120-0084)  
[greenwell.brandon@gmail.com](mailto:greenwell.brandon@gmail.com)

Bradley C. Boehmke  
University of Cincinnati  
2925 Campus Green Dr  
Cincinnati, OH 45221  
United States of America

ORCID: 0000-0002-3611-8516  
bradleyboehmke@gmail.com