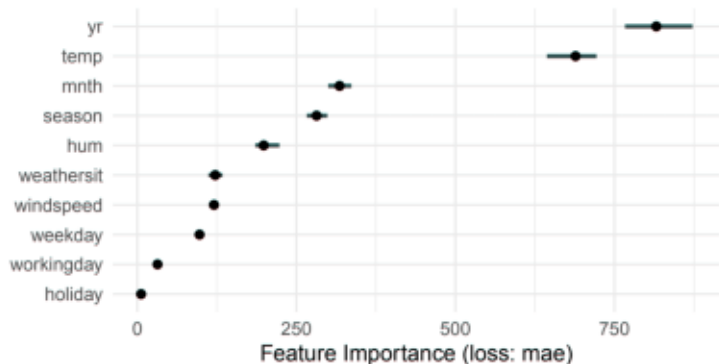


DISCOVER AND GAIN GLOBAL INSIGHTS

~> Gain insights about data, distribution and model

Example: Bike Sharing Dataset (predict number of bike rentals per day)

Exemplary question: Which feature influences the model performance and to what extent?



- Year (yr) and Temperature (temp) most important features
- Holiday (holiday) less important (Can we drop it?)

IMPROVE, DEBUG AND AUDIT MODELS

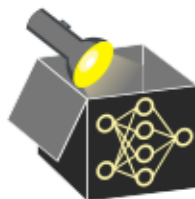
↪ Insights help to identify flaws (in data or model), which can be corrected

Example: Neural Net Tank [▶ gwern.net](https://www.gwern.net)



A cautionary tale (never actually happened):

- Train a neural network to detect tanks
- Good fit on training data
- Application outside training data: failure



IMPROVE, DEBUG AND AUDIT MODELS

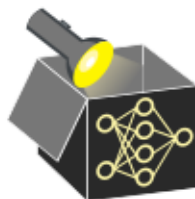
↪ Insights help to identify flaws (in data or model), which can be corrected

Example: Neural Net Tank [▶ gwern.net](http://gwern.net)



A cautionary tale (never actually happened):

- Train a neural network to detect tanks
- Good fit on training data
- Application outside training data: failure
- Reasons vary depending on input
 - ↪ NN based decision on irrelevant points



IMPROVE, DEBUG AND AUDIT MODELS

↪ Insights help to identify flaws (in data or model), which can be corrected

Example: Neural Net Tank [▶ gwern.net](https://www.gwern.net)



A cautionary tale (never actually happened):

- Train a neural network to detect tanks
- Good fit on training data
- Application outside training data: failure
- Reasons vary depending on input
 - ↪ NN based decision on irrelevant points
- E.g. model detects weather based on sky:
 - ↪ All photos with tanks show cloudy sky
 - ↪ Photos without tanks show sunny sky

