Visualizing the Bias-Variance Tradeoff

Halldor Thorhallsson & Gursimran Singh
The Bias-Variance Tradeoff

- A fundamental concept in Machine Learning
  - Bias - \((E[\hat{f}(x)] - f(x))^2\)
  - Variance - \(E \left[ (\hat{f}(x) - E[\hat{f}(x)])^2 \right] \)
- Can be difficult to grasp
- Very few interactive visualisations out there that explain this!
TensorFlow Playground

Tinker With a Neural Network Right Here in Your Browser. Don't Worry, You Can't Break It. We Promise.
Demo!

Demo KNN Link

Demo Linear Models Link
VAD Analysis

What

- Multiple samples \((X^{(1)}, y^{(1)}), (X^{(2)}, y^{(2)}), \ldots (X^{(n)}, y^{(n)})\),
- Predictions \(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n\),
- Derived: Residuals \(\rightarrow y^{(i)} - \hat{y}_i\)

Why

- To understand the **bias-variance** tradeoff

How

- Bar plot shows bias and variance exactly
- Box plot shows distribution of residuals \((y^{(1)} - \hat{y}_1)\)
- Dart board shows bias intuitively
VAD Analysis - How
Challenges

- **Polynomial regression**
  - Hard to show bias and variance in the same plot.
  - Future work - show the distribution in the scatter plot.

- **KNN**
  - Changing parameters is a bit laggy for heatmap
  - What is high/low bias and variance?
Thanks!

Q&A
VAD Analysis ()

What

- True distribution f(x)
- Training sampled data from f(x) -> X(i), y(i)
- Testing/Generalization sample -> X_g(i), y_g(i)
- Predictions y_hat

Why

- To understand how close y_hat is to y(i) and y_g(i)

How

- Scatter plot
- Navigation to view generalization samples
Scatter plot

Degree 3 fit

Degree 4 fit
Challenge - Scaling of bias and variance

Before

After