WelcomeBike

House Price Estimate Report

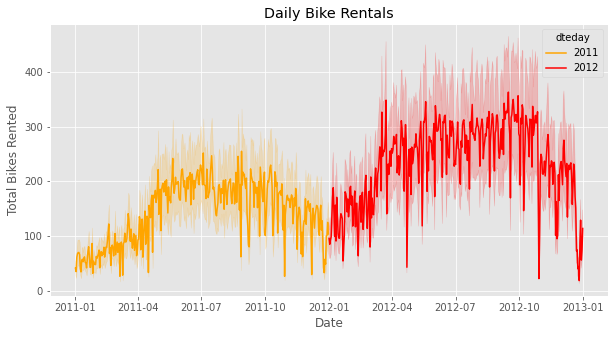
## presented by

Cody Overholt

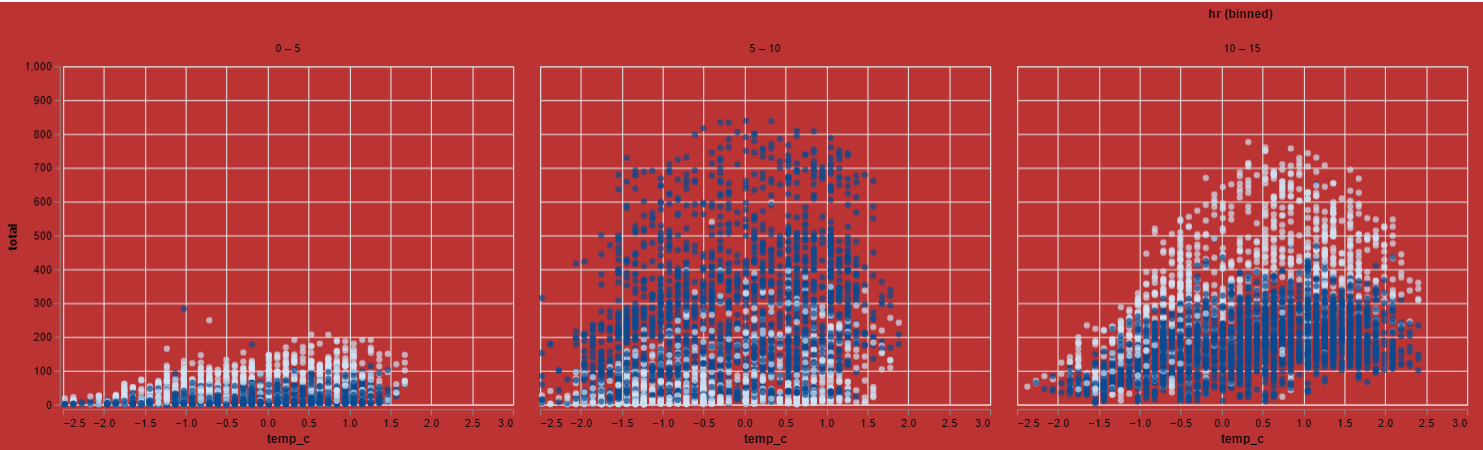
Kyle Mueller

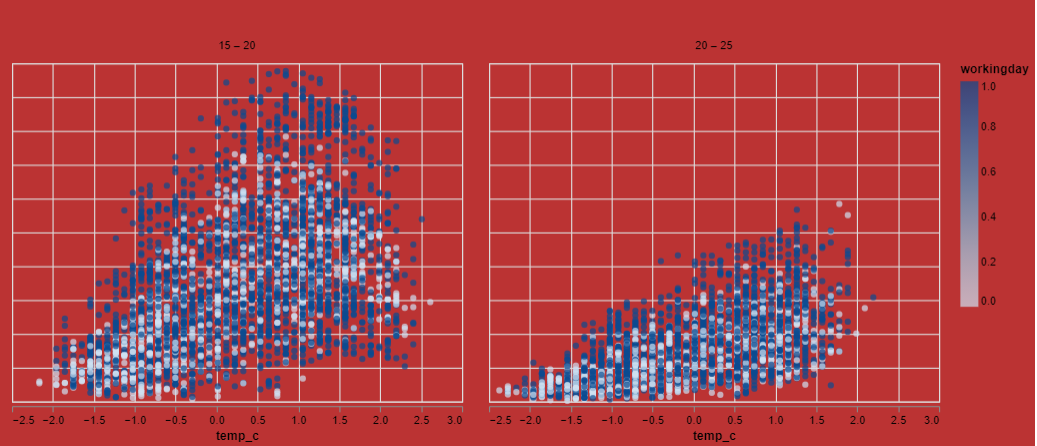
Jordan Carlson

I. Data Summary



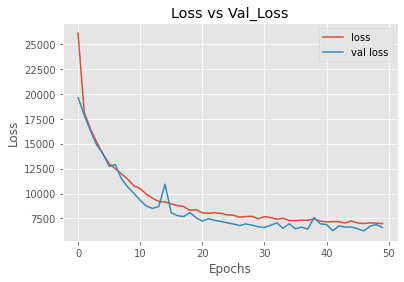
The number of bike rentals per day seems to be correlated to a number of different factors. When looking at the two years we have data from, there seems to be a consistent increase of rentals during the warm summer months and fewer rentals in the winter months. There are also massive drops in the total number of rentals surrounding certain holidays, such as christmas and thanksgiving.





The hour of day and the type of day (whether it is a working day or not) has an impact on how many bikes are rented. White dots indicate non-working days and blue dots indicate working days. Key points in the day include leaving for work and coming back home from work where the amount of bikes rented increases on working days. As the hours of the day increase the amount of rented bikes on non-working days become less than the amount of bikes rented on working days. As for indicating which days will have consistent sales, the ‘hr’ feature does not help in deciding the number of bikes rented on a given day since the hours of the day are the same for every day.

II. Model Summary

The model we trained used 7 of the features to predict the number of bikes rented out each hour of each day. It had a total of 5 layers, one input, one output, and 3 hidden layers. One hidden layer was a dropout layer that helped to prevent overfitting. In the end our model had a Mean Squared Error of 6596.69 with an Root Mean Squared Error of 81.22. To ensure that our model will perform well in the real world and not just in testing, we tested the loss of the function against the value loss. A model that isn't overfitted will have a value loss lower than the loss. The performance is shown by the graph below.

III. Conclusion

The use of a specially developed neural network along with normalized data enabled the production of a very accurate model. Data dropped from the learning algorithm included: ‘casual’, ‘registered’, and ‘dteday’, as well as a new feature we created, ‘total’.

To ensure that our model was sufficiently accurate, we added 2 hidden layers with 256 nodes. We also added a dropout layer between the hidden layers to prevent overfitting. Changing the activation functions of the layers also helped to increase accuracy and allowed the neural network to learn better.

This neural network model will accurately predict the number of bicycles to be rented for any given day of the year, and can help your business to be more efficient and profitable than ever before. Thank you.

IV. Python Notebooks

Below are Github Gist links to the notebooks we used during this case study:

<https://gist.github.com/JordanCarlson7/efa878e070a06c726b713b1aeb947f8c>

<https://gist.github.com/mueller14003/ff07364d3f20d909bd4f81eecfbd7c78>

<https://gist.github.com/codeholt/34cf8553eb7edb63a64f353561c04116>