OPERATION ANALYTICS

BIKES_HIRED

GROUP 2: OISIN BULTER, DANG DINH QUOC DUNG, QINGWEI YANG, CIARAN O'REILLY, CHIARA CENAMETI

LINEAR REGRESSION

- The bike_hires's Multiple R-square shows 54.53% of the variability in the dependent variable can be explained by the model. RMSE 8064.35.

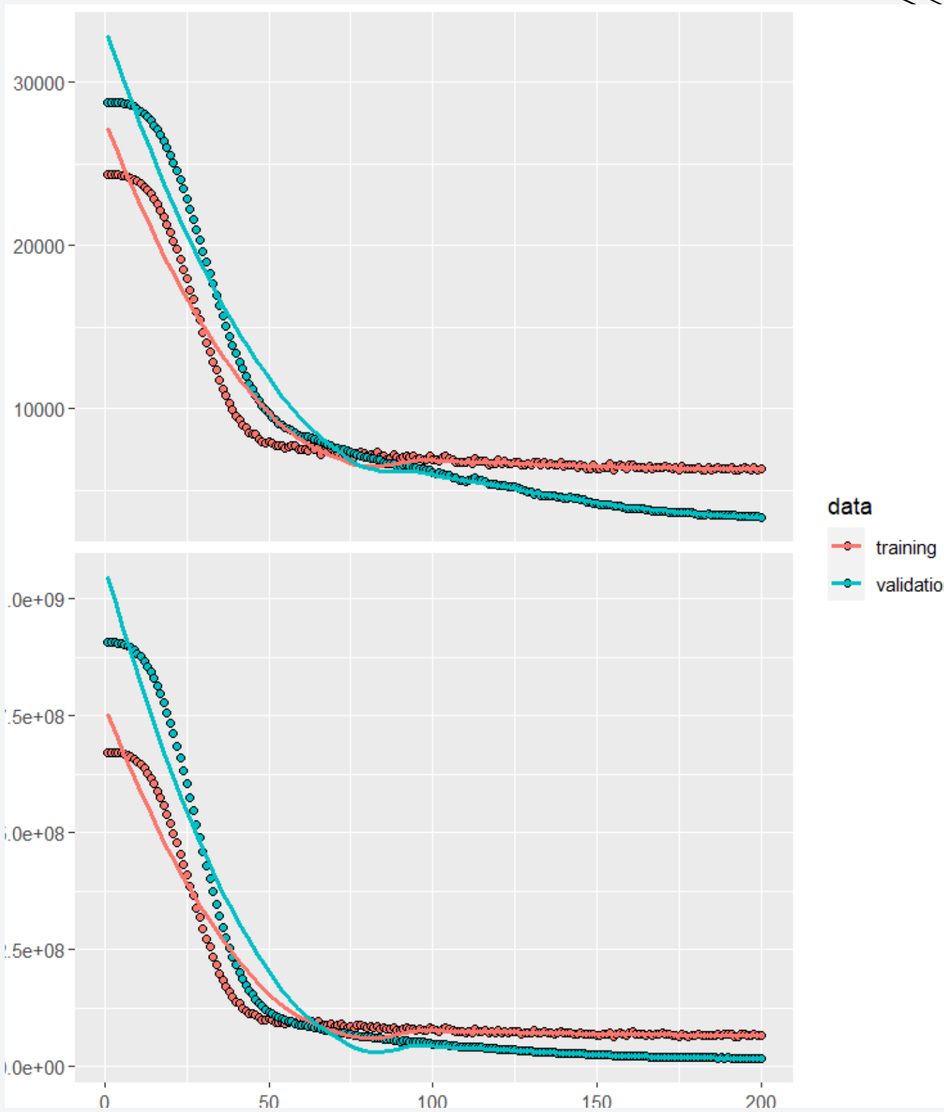
- The F-test is highly significant (p < 2.2e-16). Cloud_cover, humidity, pressure, radiation, precipitation, snow_depth, sunshine, mean_temp, and min_temp changed the bicycle hire by 177, 61, 8, 5, 34, 1380, 1380, 32, 1478, and 740 for each increase in one unit. Their p-values are significant, with snow depth and temperature having a large effect on bicycle hire.

```
> lr <- lm(bikes_hired ~ cloud_cover + humidity + pressure + radiat
ion + precipitation + snow_depth + sunshine + mean_temp +min_temp +
max_temp, data = bikes_hires)
> summary(lr)

Call:
lm(formula = bikes_hired ~ cloud_cover + humidity + pressure +
    radiation + precipitation + snow_depth + sunshine + mean_temp +
    min_temp + max_temp, data = bikes_hires)

Residuals:
   Min     1Q Median     3Q    Max
-29549  -4244    594   4631  34599

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   -66522.103  11091.607  -5.998 2.17e-09 ***
cloud_cover      177.460     83.654   2.121   0.0339 *
humidity         -61.232     15.691  -3.902 9.67e-05 ***
pressure           8.151      1.074   7.590 3.87e-14 ***
radiation          5.908      2.865   2.062   0.0393 *
precipitation    -34.199      2.974 -11.501  < 2e-16 ***
snow_depth     -1380.500    422.633  -3.266   0.0011 **
sunshine          32.803      7.424   4.418 1.02e-05 ***
mean_temp       1478.579    109.698  13.479  < 2e-16 ***
min_temp        -740.754     90.921  -8.147 4.81e-16 ***
max_temp          56.302     40.784   1.381   0.1675
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6650 on 4374 degrees of freedom
Multiple R-squared:  0.5453,    Adjusted R-squared:  0.5443
F-statistic: 524.5 on 10 and 4374 DF,  p-value: < 2.2e-16
```

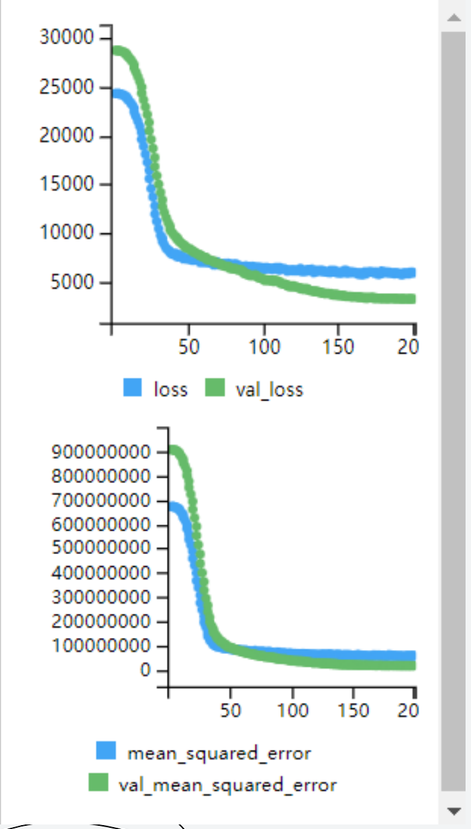
ARTIFICIAL NEURAL NETWORK

The close tracking of the validation loss with the training loss suggests that the model is generalizing well and not overfitting.

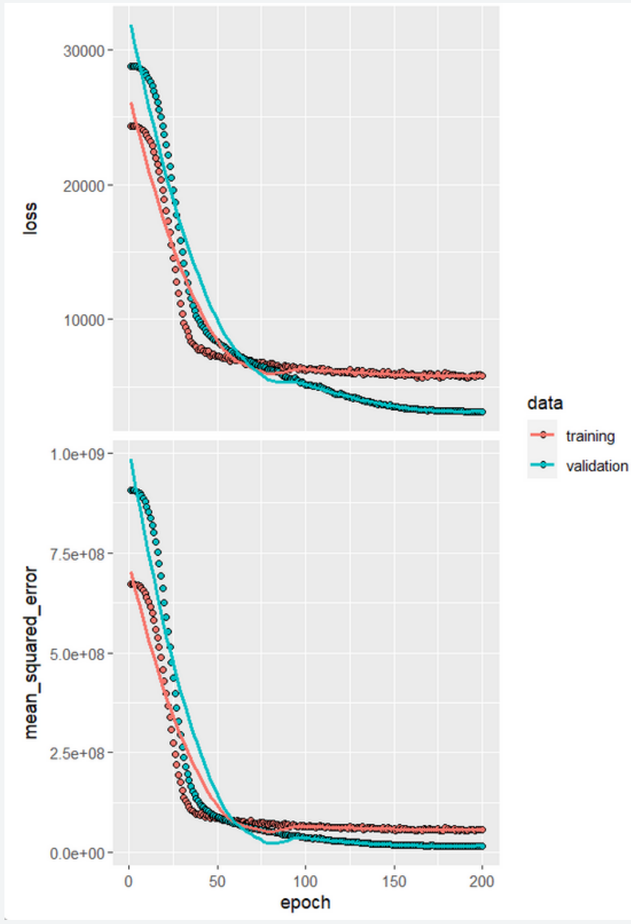
The bike_hires data has a big RMSE which is 8348.106. The data's predictions are off by a large margin on average.



LSTM



The plots represent the loss and MSE from training the LSTM model. There's a sharp decrease in loss and MSE at the start, which indicates initial rapid learning.



CONCLUSION

- *Adjusted R-squared: 54.43%. there is still about 45% of the variability in bike hires that is not explained by linear regression. Time seasonal data also should be considered.*
- *To fulfil the higher demand indicated by the models, bicycle stock should be raised on days with favourable weather, meaning there should be less cloud cover, moderate humidity, and warmer temperatures.*
- *Activities like maintenance and restocking could be planned for days with less favourable weather, when rental numbers are expected to be lower.*