

# DEMOGRAPHICS AND ACADEMIC PERFORMANCE IN NORTH CAROLINA'S SCHOOLS

UNC Data Analytics Bootcamp  
Final Presentation 2-12-2020



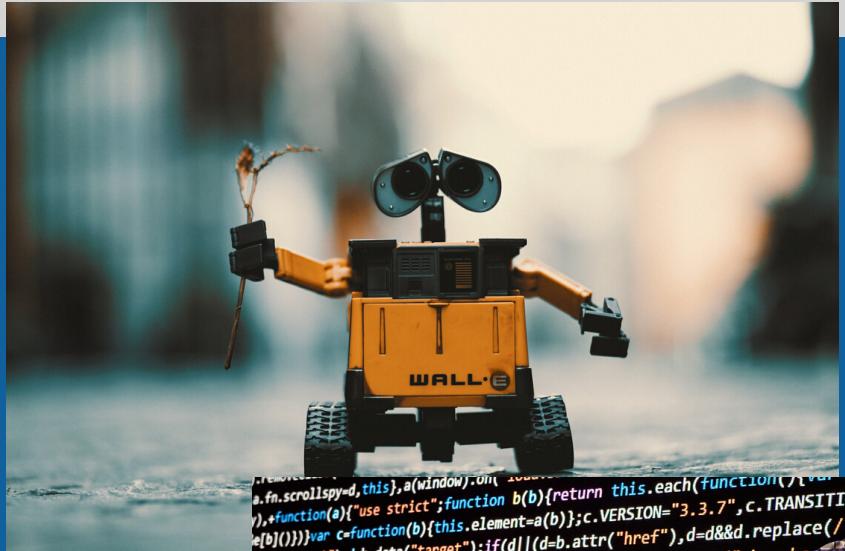


## What Factors did we look at?

- Economic status of students
- Race and poverty in county
- Education and employment status of county households
- Lead levels
- Child welfare stats in county



# Machine Learning!



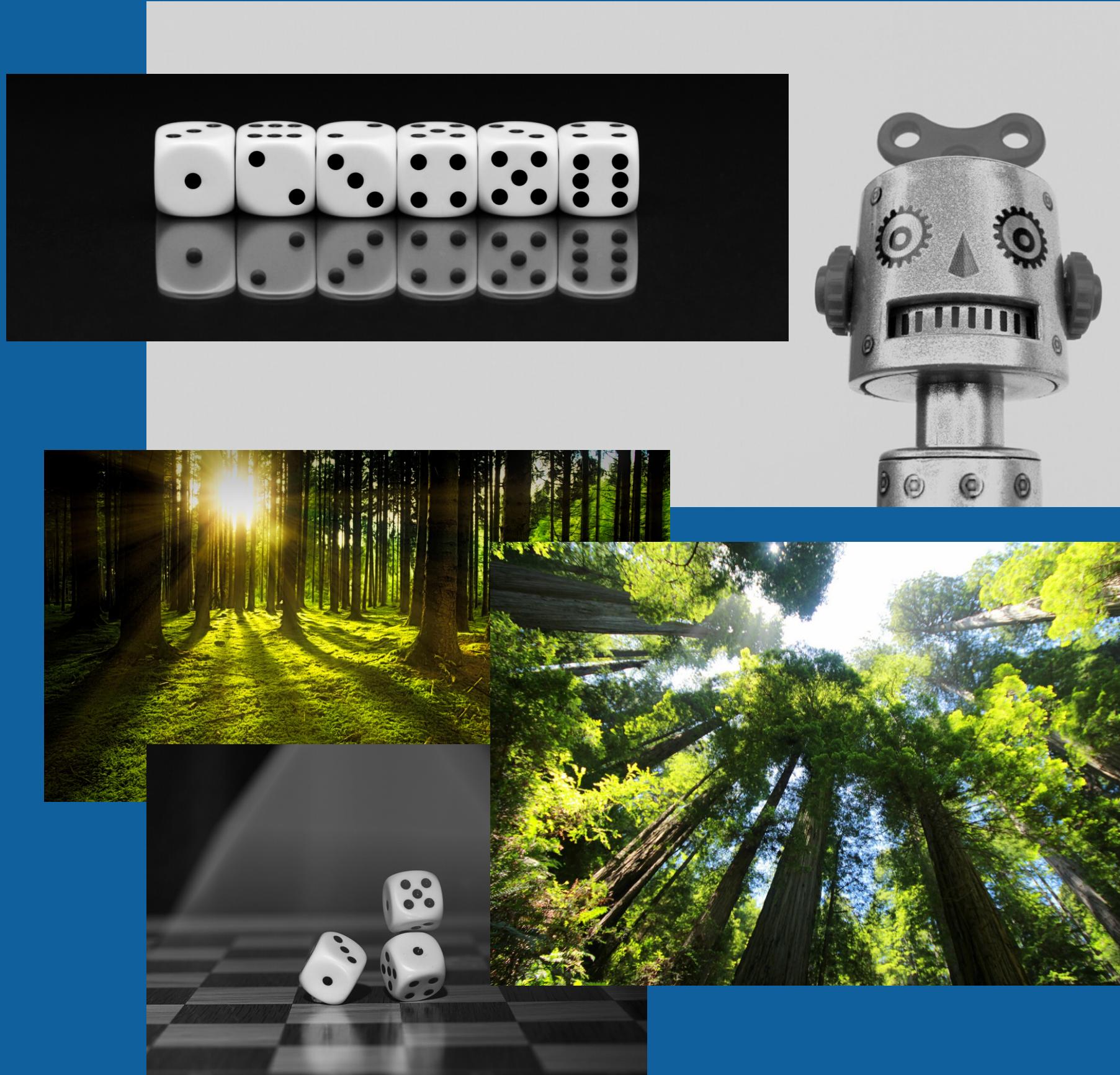
- What quintile did the school score within on end of grade (EOG) tests?
- Can the model guess based on demographic data?

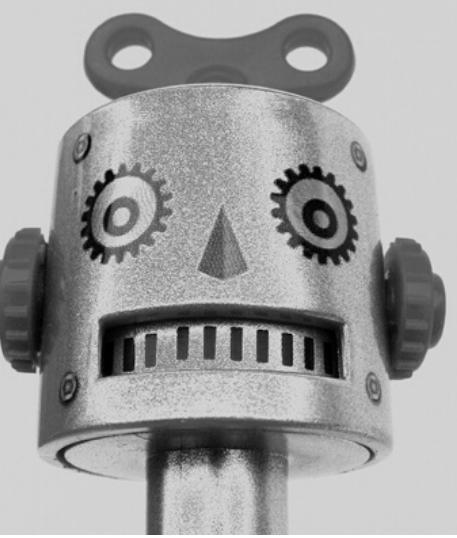
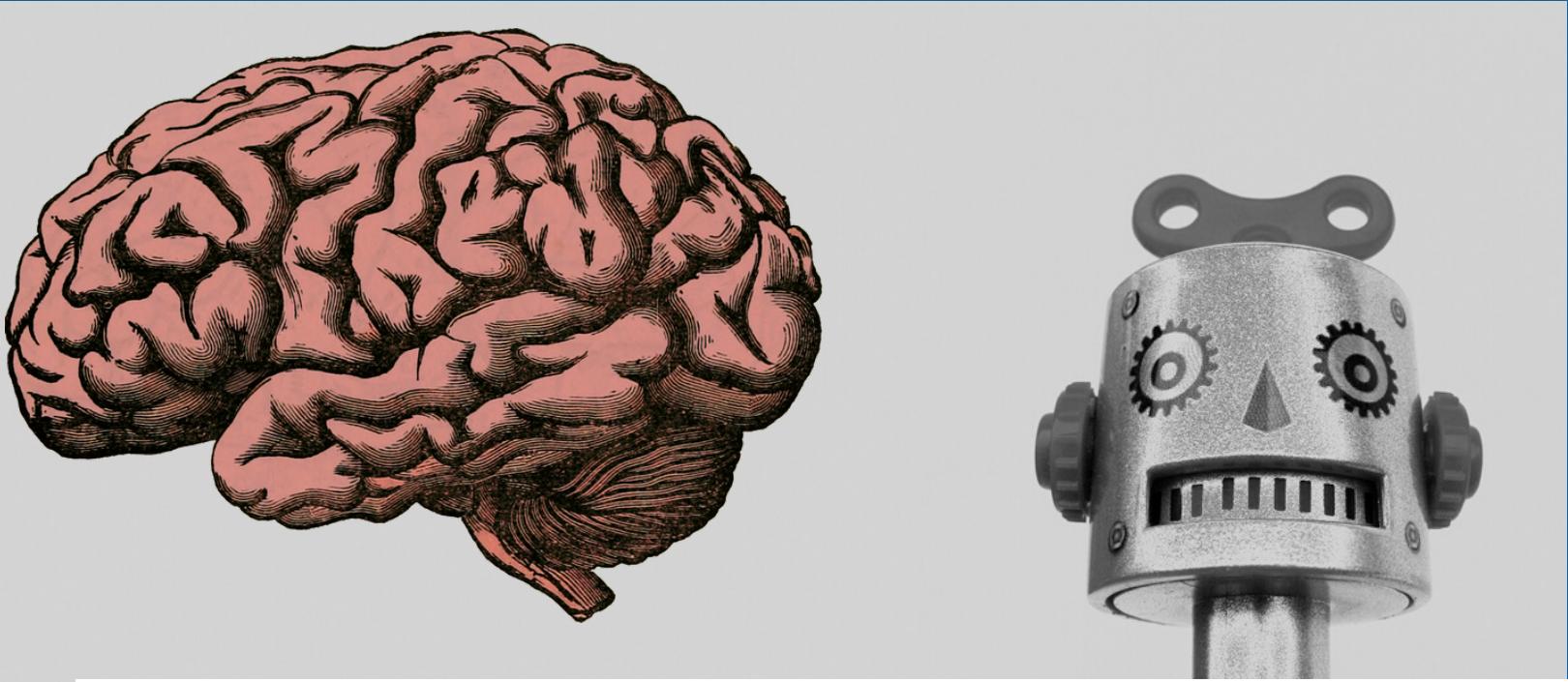
# Random Forest!

```
Out[13]: 0.48154093097913325
```

```
In [14]: sorted(zip(rf.feature_importances_, feature_na
```

```
Out[14]: [(0.7440865643535888, 'percent_EDS'),  
 (0.048522987433570135, 'county_poc'),  
 (0.027533033640828344, 'median_inc_county'),  
 (0.027302117085389502, 'parent_unemployed'),  
 (0.025978038750146027, 'poverty_county'),  
 (0.025539733650452, 'child_abuse_rate'),  
 (0.02530911585474701, 'children_conc_pov'),  
 (0.02520278537106873, 'juv_delinquent'),  
 (0.025013662221660224, 'No_HSdegree'),  
 (0.01950573394715777, 'elevated_lead'),  
 (0.006006227691391518, 'county_tier')]
```





# Neural Network!

## Quantify our Trained Model

```
In [16]: model_loss, model_accuracy = model.evaluate(  
    X_test_scaled, y_test_categorical, verbose=2)  
print(  
    f"Normal Neural Network - Loss: {model_loss}, Accuracy: {model_accuracy}")
```

623/623 - 0s - loss: 0.9350 - acc: 0.6067

Normal Neural Network - Loss: 0.9350250822583133, Accuracy: 0.6067415475845337

