Jackie Chan
301310345

Assignment 1: Regression

1 Probabilistic Modeling and Bayes' Rule

1. $P(M) = 0.01;\ P(T|M) = 0.95;\ P(T|\neg M) = 0.05$
   a. $P(T) = P(T|M) * P(M) + P(T|\neg M) * P(\neg M) = 0.95 * 0.01 + 0.05 * 0.99$
      $P(T) = 0.059$
   b. $P(M|T) = \frac{P(T|M)P(M)}{P(T)} = \frac{0.95(0.01)}{0.059} = \frac{19}{118} \approx 0.161$

2. Let P(rain today) be P(TO), let P(rain tomorrow) be P(TM)
   $P(TO) = 0.3;\ P(TM) = 0.6;\ P(TO, TM) = 0.25$
   $P(TM|TO) = \frac{P(TO,TM)}{P(TO)} = \frac{0.25}{0.3} = 0.8\overline{3}$

3. $P(biased\ odd) = P(1) + P(3) + P(5) = 0.1 + 0.2 + 0 = 0.3$
   $P(fair\ odd) = P(1) + P(3) + P(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 0.5$
   Probability of winning with a biased die is 0.3. This is worse than a fair die.

3 Training vs. Test Error

1. The validation error is not *always* higher than the training error. Validation error could be lower if the regression model generalizes well, and the validation data set happens to fit better with the trained regression model.
2. In unregularized regression, training error with a degree 10 polynomial is always lower than or equal to the training error with a degree 9 polynomial. This is the case because unregularized regression does not penalize larger coefficients (which would be the case with a degree 10 polynomial), and a higher degree polynomial will fit more specifically to the training data than a lower degree polynomial.
3. The testing error with a degree 20 using regularized regression is always lower than the testing error using unregularized regression. This is the case because the regularized regression model avoids overfitting by penalizing large coefficients, allowing the model to better fit similar but different data. The unregularized regression model will be overfit for the training data, and thus performs worse for any data that differs from the training data.
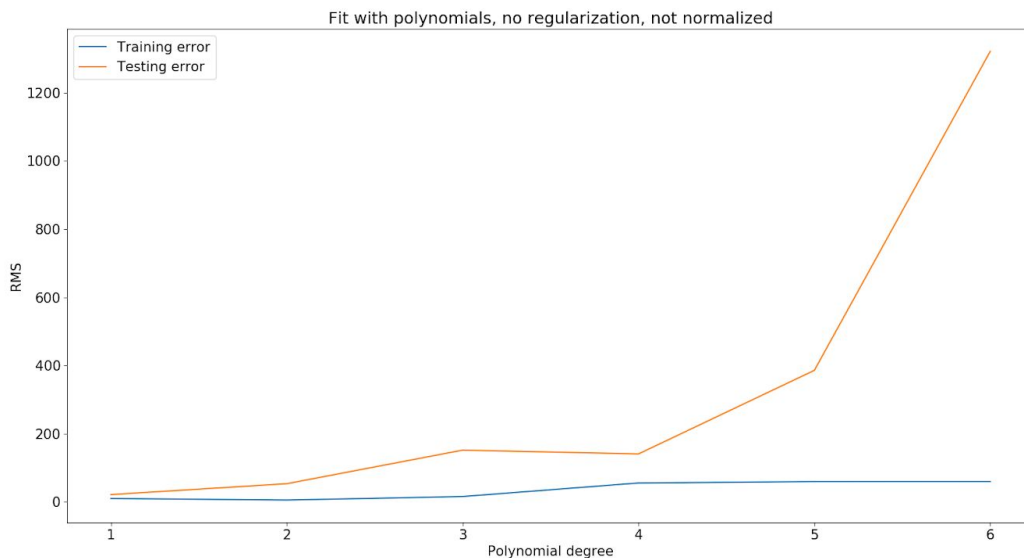
4 Regression

4.1 Getting Started

1. The country with the lowest child mortality rate in 1990 is Iceland, with a rate of 6.3.

2. The country with the lowest child mortality rate in 2011 is San Marino, with a rate of 1.8.
3. Missing features are filled in with the mean value for that feature from other countries.
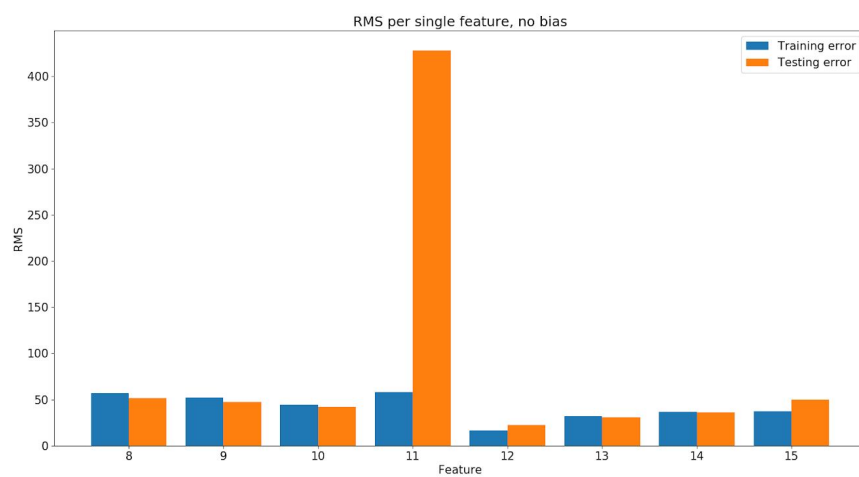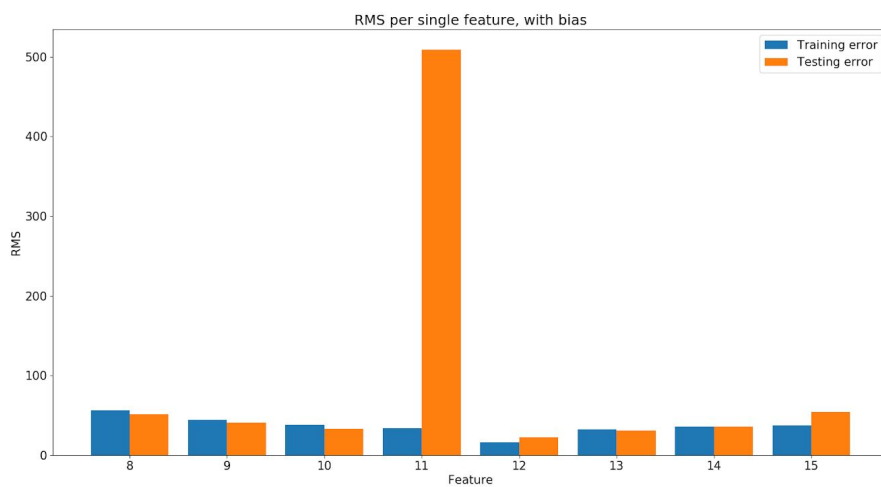
4.2 Polynomial Regression

1.

The issue with polynomial regression without regularization is the large increase in testing error as the polynomial degree increases, due to overfitting of the weights.

2.

RMS per single feature, with bias

RMS per single feature, no bias
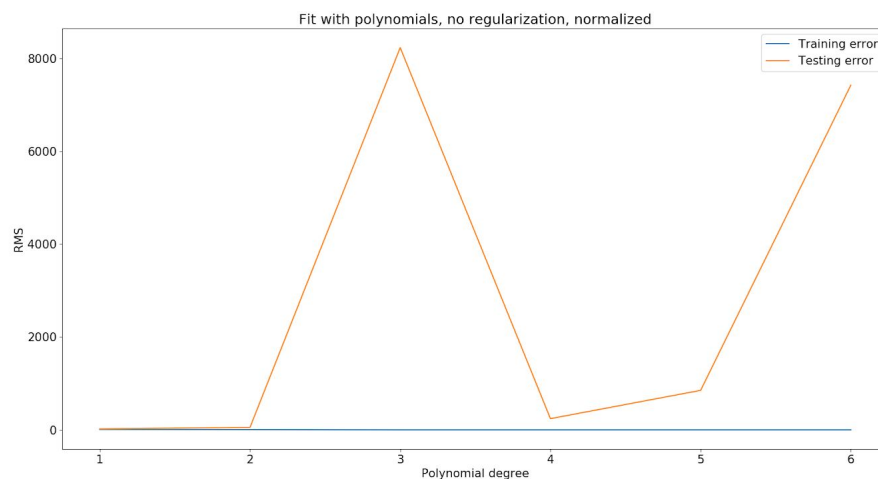
4.3 Sigmoid Basis Functions

	Training Error	Testing Error
μ = 100, *no bias*	28.555	34.041
μ = 10,000, *no bias*	38.409	39.187
μ = 100, *with bias*	28.573	34.089
μ = 10,000, *with bias*	38.422	39.250