Evaluating regression algorithms

Evaluating regression algorithms:

- Evaluating machine learning models help us to answer the following questions:
    - How well is my model doing? Is it a useful model?
    - Will training my model on more data improve its performance?
    - Do I need to include more features?
- So, do we see how our model did for every observation on our data set?
    - No, we evaluate our model using a single number (metric).
- The most common metric used for regression is R^2

Importance of Evaluating Regression Models

- Evaluating regression models is crucial to ensure their accuracy and reliability.
- Proper evaluation helps in understanding how well the model fits the data and predicts future outcomes.

على أي أساس بختار مقياس التقييم ؟ الداتا labeled or not

كل ما كانت قيمتها اصغر ، أفضل دقة > MSE,MAE,RMES

(R^2) Coefficient of Determination

- Definition: Measures the proportion of the variance in the dependent variable that is predictable from the independent variables

Formula: $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

- R^2 ranges from 0 to 1
- **Example:** R^2= 0.85 means 85% of the variance is explained by the model

(R^2) Coefficient of Determination

- R^2 squared calculates how much regression line is better than a mean line.
- R-Squared is called coefficient of determination where it is commonly between 0 and 1 and as the value approaches 1 more, this gives intuition that the model representation
- is closer to the data.
- So, with help of R squared we have a baseline model to compare a model.

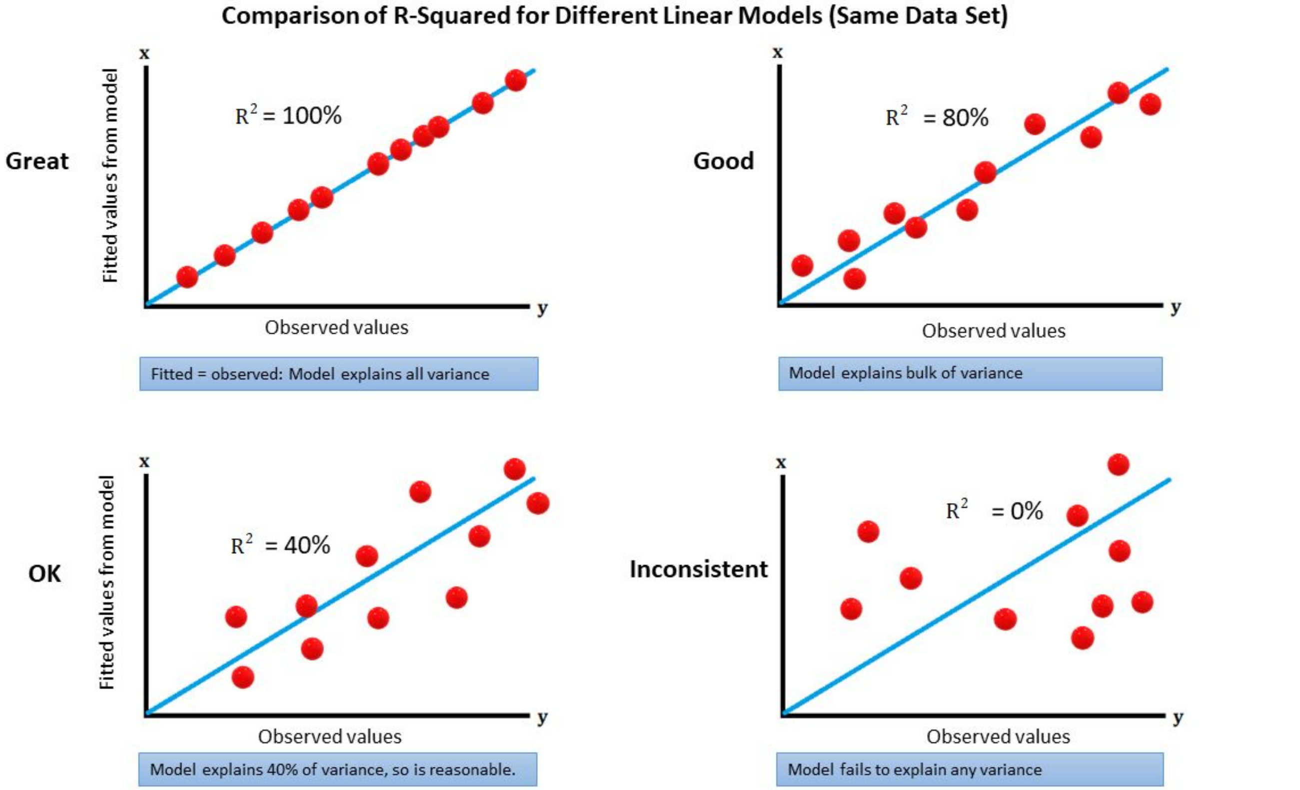
When to Use (R^2)

- Suitable for simple linear regression
- Initial model evaluation
- Quick assessment of model performance

Limitations of

- Can be misleading with multiple predictors
- Does not account for model complexity or overfitting

Evaluating Model Performance (R^2)



adjusted R^2

- Definition: Adjusts the R^2 value based on the number of predictors in the model

$$\text{Formula: Adjusted } R^2 = 1 - \left(\frac{1-R^2}{n-p-1}\right)(n-1)$$

- **adjusted R^2** can decrease if unnecessary predictors are added

P: num of features
n: num of rows

When to Use Adjusted R^2

- Ideal for multiple regression models
- Comparing models with different numbers of predictors
- Feature selection and model refinement

Comparison of R^2 and Adjusted R^2

- R^2 does not penalize for extra predictors
- Adjusted R^2 penalizes for unnecessary predictors
- Adjusted R^2 is more reliable for multiple regression

Example

Data:

- y (house prices in \$1000s): [200, 250, 300, 350, 400]
- X (size in square feet): [1500, 1800, 2000, 2300, 2500]

- Predicted values ($\hat{y}$): [210, 260, 290, 340, 390]
- Mean of actual values ($\bar{y}$): 300
- R^2 Calculation: $R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$
- Result: $R^2 = 0.95$ (indicates a very good fit)

- **New Model**: Multiple linear regression with additional predictors:
    - Size (in square feet)
    - Number of bedrooms
    - Age of the house (years)
- Additional data:
    - Bedrooms: [3, 3, 4, 4, 5]
    - Age: [10, 15, 10, 20, 10]
- Predicted values ($\hat{y}$): [205, 255, 305, 355, 405]
- Adjusted R^2 Calculation: Suppose R^2 for the new model is 0.96.
- Adjusted $R^2 = 1 - \left(\frac{1-0.96}{5-3-1}\right)(5-1) = 0.94$

Adjusted $\mathbf{R^2}$ indicates **whether the new predictors** genuinely <u>improve the model or not.</u>

Evaluating Model Performance (adjusted R^2)

- R-Squared can lead sometimes to misleading results in case of overfitting where the model overfits on data so it gives a value of 1 (means that it kept 100% of the variance of the data) while it is not learning but memorizing.
- Adjusted R2 is better in showing if there is real improvement or not.
- If you add more features to the data, the R2 will either increase or remains the same even if they are useless, however in the adjusted R2, it is more intuitive as if the added feature is useless, it will decrease due to the k value.
- Notes:
  1. k is the number of features in the data while n is the number of samples in the data.
  2. Its value is commonly between 0 and 1 as well and less than or equals R2.

$$R_a^2 = 1 - [(\frac{n-1}{n-k-1}).(1-R^2)]$$

🔍 Comparison of Key Regression Metrics

Metric	Definition	Sensitive to Outliers?	Use Case
MSE (Mean Squared Error)	Average of squared differences between actual and predicted values	✅ Yes	Penalizes large errors more, useful for optimization
RMSE (Root Mean Squared Error)	Square root of MSE	✅ Yes	Interpretable like actual values, popular in evaluation
MAE (Mean Absolute Error)	Average of absolute differences	❌ Less	More robust to outliers, useful for noisy data
R^2 (R-squared)	Proportion of variance explained by the model	❌ No	Measures model fit, easy to interpret
Adjusted R^2	Penalized version of R^2 that accounts for number of features	❌ No	Used when comparing models with different numbers of features

📊 Numerical Example:

Let's say you have two prediction errors:

- A small error: `2`
- A large error (outlier): `20`

✅ MAE Calculation:

$$MAE = \frac{1}{2}(|2| + |20|) = \frac{22}{2} = 11$$

- The large error **contributes 20**, no more.
- Treated as a large error, but **not exaggerated**.

❌ MSE Calculation:

$$MSE = \frac{1}{2}(2^2 + 20^2) = \frac{4 + 400}{2} = \frac{404}{2} = 202$$

- The large error **contributes 400**!
- One outlier **dominates** the total error.

Squaring in MSE **magnifies large errors**, while MAE treats all errors **linearly**.