Interpreting linear regression models

Advantages

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_n x_{ni} + \varepsilon_i$$

- Predict the target as a linear combination of the predictors (weighted sum of the predictors).
- We can use statistical tests to decide if / how much we can trust the model and its parameters (β_i).
- Intrinsically explainable by design.
- We can use regularization to reduce the feature space.
    - Optimize for interpretability.

Limitations

- The interpretation of the weight / coefficient is contrastive (depends on all other features).
    - Features with positive correlation coefficient show a negative weight.
- Make assumptions on the data ➔ when they are not met, we can't trust the model.
- Multicolinearity affects interpretability.
- Interactions between the variables won't be captured.

Train In Data

THANK YOU

www.trainindata.com